

# STUDIES ON THREONINE SYNTHASE

Tracy Neal

A Thesis Submitted for the Degree of PhD  
at the  
University of St Andrews



1998

Full metadata for this item is available in  
St Andrews Research Repository  
at:

<http://research-repository.st-andrews.ac.uk/>

Please use this identifier to cite or link to this item:

<http://hdl.handle.net/10023/15427>

This item is protected by original copyright

L

# STUDIES ON THREONINE SYNTHASE

a thesis presented by  
Tracy Neal  
to the  
University of St. Andrews  
in application for

THE DEGREE OF DOCTOR OF PHILOSOPHY



St. Andrews

October 1997

ProQuest Number: 10170943

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 10170943

Published by ProQuest LLC (2017). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 – 1346

TH  
D15

## DECLARATIONS

I, Tracy Neal, hereby certify that this thesis, which is approximately 55, 000 words in length, has been written by me, that it is the record of work carried out by me and that it has not been submitted in any previous application for a higher degree.

Date: 20/11/98..... Signature of candidate.

I was admitted as a research student in October 1993 and as a candidate for the degree of Ph.D. in October 1997; the higher study for which this is a record was carried out in the university of St. Andrews between 1993 and 1997.

Date: 20/11/98..... Signature of candidate.

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of Ph.D. in the University of St. Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date: 20/11/98..... Signature of supervisor. (

## COPYRIGHT

In submitting this thesis to the University of St. Andrews I understand that I am giving permission for it to be made available for the use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. I also understand that the title and the abstract will be published, and that a copy of the work may be made and supplied to any *bona fide* library or research worker.

Date: ...2014.198.....

Signature of candidate

*To my mother  
and my grandparents*

*There are some things which cannot be learned quickly and time, which is all we have, must be paid heavily for their acquiring. They are the very simplest things, and because it takes a man's life to know them, the little new that each man gets from life is very costly and the only heritage he has to leave.*

*Ernest Hemingway*



## *Acknowledgements*

I wish to thank my supervisor, Professor David Gani for his support and encouragement throughout my PhD. I also thank Dr. Martin Ryan for being such a great surrogate supervisor, for helping me to understand and enjoy the world of molecular biology and for his hospitality.

I would like to thank Donald McNab and Graham Allan for all their help in lab 424 in Chemistry and for providing the entertainment! I would also thank all those in lab 41 at Biochemistry for their patience while I was learning so many new techniques, especially Sue Monaghan and Michelle Donnelly.

Thank you to the proof-readers of this thesis; to Martin Ryan, Mahmoud Akhtar, Elaine Brown and Saqib Furukkh. Thanks also to Melanja Smith for all the friendly NMR advice, to Colin Millar for providing the mass specs, sometimes at short notice, to Dr. John Wilkie for advice about sequence alignments and molecular modelling and to Alex Houston for all the sequencing. Thanks to my great friend Jane Smith for helping me to find last-minute references in Cardiff university library! Thanks also to Roger Pybus for all his help when I had computer troubles!

Thanks to my friends in St Andrews: Colin French, Arwel Lewis, Iain Greig, Neil Anderson, Sue Richardson, Karen Muirhead and Elaine for some most amusing and interesting nights both out and in which helped me to forget I was doing a PhD! Thanks also to Martin and Gillian Beaton, Nigel Pitt, Elaine and Fred Kroll for all the silly games of tennis!

I would especially like to thank some people without whose love and support I would never have got through this: thanks to Andrew Danby for helping me to survive the first year, to Panthea Hormozdiari for the free counselling service!, to Stacey for her encouragement and to Nigel who always listened.

Finally I thank the EPSRC for funding my PhD.

## Abstract

Threonine synthase (TS, EC 4.2.99.2), catalyses a  $\beta,\gamma$ -replacement reaction to convert (2*S*)-*O*-phosphohomoserine (**32**) into (2*S*,3*R*)-threonine (**33**). Although threonine as an essential amino acid is vital to all lifeforms, the enzyme is not expressed in mammals. This makes it an ideal target enzyme for herbicides, fungicides and bactericides, and several inhibitors of the enzyme have been produced with this in mind. In particular, the enzyme from *Escherichia coli* has previously been the subject of inhibition studies, and the reaction mechanism of this particular TS has also been partially elucidated. However, no product inhibition studies have been carried out on TS from *E. coli*. In order to carry out such studies, a [U-<sup>14</sup>C]-labelled version of the substrate (**32b**) was synthesised in this work, starting from [U-<sup>14</sup>C]-(2*S*)-aspartate (**28b**), *via* a route previously developed for producing the unlabelled substrate. Various analogues of the substrate have been synthesised and these were also to be tested with the enzyme, either as potential inhibitors, or to elucidate further the reaction mechanism and active-site structure of TS.

The mutant *E. coli* strain *K-12 Tir8*, which had been used previously as a source of TS, appeared to have reverted to wild-type, no longer over-expressing the enzymes of the *thr* operon. Therefore, the *thrC* gene from *E. coli*, coding for TS, was cloned into a pET-expression system. In a host cell containing such a construct TS could amount to 50% of total cell protein. *ThrC* was amplified *via* PCR and inserted into a cloning vector, pGEM-T, and then subcloned into pET-3a, pET-3b and pET-16b. pET-16b constructs produce a His-tagged version of the recombinant protein. Sequencing of the recombinant gene in two pGEM-T constructs revealed one deletion and two mutations in the *thrC* sequence, which probably occurred during the PCR-amplification of the gene. These alterations were confirmed in sequences obtained for the pET-constructs.

*N,N*-diisopropylidichlorophosphamidite (**78**), a precursor to a reagent used originally for the phosphorylation of  $\alpha$ -isopropyl-*N*-trifluoroacetyl-(2*S*)-homoserine (**75**), proved difficult to synthesise. Instead, dipentafluorophenyl phosphorochloridate (**81**) was used for the phosphorylation reaction, as it was easier to synthesise and gave a good yield of the phosphate ester. Deprotection of dipentafluorophenyl phosphates has, however, only been achieved previously on the solid-phase. The solid-phase synthesis of the substrate was therefore attempted using Wang, *p*-hydroxymethyl polystyrene and Merrifield resins. A new linker was attached to Merrifield, to produce the novel resin, polystyrene-4-oxymethyl-2-phenylethanol (**97**). Although selective opening of the *N*-trifluoroacetyl-(2*S*)-aspartic anhydride (**73**) was successfully accomplished to attach the  $\alpha$ -carboxyl group to these resins, subsequent reduction of the  $\beta$ -acid has not been achieved.  $\alpha$ -*p*-Benzyloxybenzylpolystyrene-*N*-trifluoroacetyl-(2*S*)-aspartate (**87**) proved unstable towards reducing agents and bases. It is hoped that compounds attached to **97** will prove more stable towards reducing agents.

## *List of Figures*

Figure 1.1: Dermatitis caused by pellagra.	4
Figure 1.2: Sequence homology between various eukaryotic and prokaryotic AspATs close to the active-site lysine residue.	9
Figure 1.3: A summary of the reactions carried out by enzymes of the $\beta$ -family of PLP-dependent enzymes.	12
Figure 1.4: Sequence homology of $\beta$ -family enzymes around the lysine motif and the glycine-rich turn.	13
Figure 1.5: Complex between cobalt (III) and PLP-glycine Schiff's base.	18
Figure 1.6: Active-site structures of the $\beta$ -subunit of wild-type and $\beta$ K87T TrS.	20
Figure 1.7: Casal's necklace - a common symptom of pellagra.	29
Figure 1.8: Route for the biosynthesis of threonine from aspartic acid postulated in the 1950s.	37
Figure 1.9: Sequence homology of amino acid residues around the active-site (underlined) of threonine synthase from various microorganisms.	41
Figure 1.10: Genetic map of the threonine operon of <i>E. coli</i> K-12.	45
Figure 1.11: Regulatory region of the <i>thr</i> operon.	46
Figure 1.12: Model for a typical repression mechanism of gene regulation.	46
Figure 1.13: The <i>thr</i> regulatory region from -71 to -27, which forms the termination structure.	47
Figure 1.14: Internal phosphate complex formed prior to elimination of phosphoric acid.	51
Figure 1.15: Inhibitors of PLP-enzymes.	57
Figure 1.16: Amino acid inhibitors of threonine synthase.	58
Figure 1.17: General structure of pET-vectors used in cloning experiments	68
Figure 2.1: Graphic map of the restriction sites on <i>thrC</i> used during cloning experiments.	80
Figure 2.2: DNA sequences of the primers used in the first PCR reaction.	80
Figure 2.3: 1% agarose gel electrophoresis of amplified <i>thrC</i> .	81
Figure 2.4: Primers complementary to the centre of <i>thrC</i> .	83
Figure 2.5: Results of overlap PCR experiments on <i>thrC</i> .	84
Figure 2.6: <i>Nde</i> I digest of pGEM-TTS <sub>A</sub> and pGEM-TTS <sub>B</sub> .	85
Figure 2.7: SP6 and T7 primer sequences.	86
Figure 2.8: Orientation of <i>thrC</i> in pGEM-TTS <sub>B</sub> .	87
Figure 2.9: Left: primers used for the third PCR experiment to amplify <i>thrC</i> . Right: result of PCR using primers TS5 and TS6.	89
Figure 2.10: Results of ligations of <i>thrC</i> into pGEM-T.	91
Figure 2.11: Sequences obtained for pGEM-TTS <sub>I</sub> and pGEM-TTS <sub>II</sub> .	93
Figure 2.12: Vector maps of pTN-TS constructs.	95
Figure 2.13: Gels of restriction digests of pTN-TS constructs.	96

Figure 2.14: Restriction analyses of pTN-TS constructs derived from pET-16b.	97
Figure 2.15: DNA sequence of the two termini of the $\beta$ -methylaspartase gene and the primers used in the PCR reaction.	98
Figure 2.16: Graphic map of the restriction sites on the gene for $\beta$ -methylaspartase used during cloning work.	99
Figure 2.17: Result of the first PCR experiment to amplify the gene for $\beta$ -methylaspartase.	99
Figure 2.18: Primers complementary to the centre of the gene for $\beta$ -methylaspartase.	100
Figure 2.19: Results of overlap PCR experiments.	100
Figure 2.20: <i>Nde</i> I restriction digest of pGEM-TMA <sub>I</sub> (1) and pGEM-TMA <sub>II</sub> (2).	101
Figure 2.21: Left: primers used in the third set of PCR reactions for $\beta$ -methylaspartase, right: the result of PCR using these primers.	102
Figure 2.22: <i>Nde</i> I restriction digest of several putative pGEM-TMA constructs.	102
Figure 2.23: Restriction digests on maxiprep DNA produced from samples shown in Fig. 2.22.	103
Figure 2.24: Alternative phosphorylating agents tried in the synthesis of (2 <i>S</i> )- <i>O</i> -phosphohomoserine.	106
Figure 2.25: NMR spectra of <i>N</i> -trifluoroacetyl-(2 <i>S</i> )- <i>O</i> -phosphohomoserine diphenyl ester ( <b>83</b> ).	109
Figure 2.26: Examples of NMR spectra comparing ratios of $\alpha$ - vs $\beta$ -substitution of <b>86</b> on Wang.	119
Figure 2.27: <sup>13</sup> C gel-phase NMR spectra of top: Wang resin, bottom: <b>87</b> , both in DMSO.	120
Figure 2.28: Hindered reducing agents.	130
Figure 2.29: <sup>1</sup> H-NMR spectrum of material cleaved from Wang resin after an attempted reduction of <b>87</b> using LiBH <sub>4</sub> .	131
Figure 2.30: <sup>13</sup> C gel-phase NMR spectra of top: Merrifield resin, middle: <b>97</b> and bottom: <b>99</b>	138
Figure 2.31: <sup>19</sup> F NMR spectra referred to in Table 2.14	139

## *List of Schemes*

Scheme 1.1: Some types of reaction catalysed by PLP-dependent enzymes.	2
Scheme 1.2: The metabolism of vitamin B <sub>6</sub> .	3
Scheme 1.3: Snell's "Shuttle Mechanism" for enzymic transamination.	5
Scheme 1.4: Conversion of bound cofactor to pyridoxamine phosphate - a process common to many PLP-enzymes.	6
Scheme 1.5: Condensation between $\epsilon$ -lysine residue and PLP at the enzyme active-site followed by transaldimination.	16
Scheme 1.6: Preferential cleavage of the bond orthogonal to the plane of the pyridinium ring as predicted by the Dunathan hypothesis.	17
Scheme 1.7: Formation of the $\alpha$ -aminoacrylate intermediate.	22
Scheme 1.8: The aspartate pathway.	36
Scheme 1.9: The reaction catalysed by threonine synthase.	37
Scheme 1.10: (2 <i>S</i> )- <i>O</i> -phosphohomoserine deaminase reaction.	38
Scheme 1.11: Alternative metabolic pathways for phosphohomoserine.	38
Scheme 1.12: Transaldimination reaction involving (2 <i>S</i> )- <i>O</i> -phosphohomoserine ( <b>32</b> )	49
Scheme 1.13: Labilisation of H $^{\alpha}$ by an enzyme-bound base.	49
Scheme 1.14: $\beta,\gamma$ -elimination of phosphoric acid.	51
Scheme 1.15: Labilisation of the 3- <i>pro</i> -S hydrogen atom.	52
Scheme 1.16: Comparison of proposed mechanisms for the elimination of phosphate.	53
Scheme 1.17: Resaturation of vinylglycine intermediate and release of threonine.	54
Scheme 1.18: Protocol for the purification of His-tagged TS	68
Scheme 2.1: Route A: Fickel and Gilvarg's synthesis of (2 <i>S</i> )- <i>O</i> -phosphohomoserine, Route B: Schnyder and Rottenberg's method.	70
Scheme 2.2: The biomimetic synthesis of (2 <i>S</i> )- <i>O</i> -phosphohomoserine ( <b>32</b> ) from (2 <i>S</i> )-aspartic acid ( <b>28</b> ).	72
Scheme 2.3: Synthesis of <i>N,N</i> -diisopropyl(bis)benzyl phosphoramidate ( <b>78</b> ).	73
Scheme 2.4: Final cloning procedure adopted for <i>thrC</i> .	78
Scheme 2.5: Plan for the ligation of <i>thrC</i> into pET-3a derived from a pTC construct.	90
Scheme 2.6: Synthesis of dipentafluorophenylphosphorochloridate ( <b>81</b> ).	107
Scheme 2.7: Attempted methods of phosphorylating <i>N</i> -trifluoroacetyl- $\alpha$ -isopropyl-(2 <i>S</i> )-homoserine ( <b>75</b> ).	108
Scheme 2.8: The proposed solid-phase synthesis of (2 <i>S</i> )- <i>O</i> -phosphohomoserine ( <b>32</b> ) on Wang resin.	113
Scheme 2.9: Selective reduction of a carboxylic group in the presence of <i>N</i> -trifluoroacetyl and methyl ester functionalities.	123

Scheme 2.10: Proposed scheme for the reaction of <b>87</b> with NMM and IBCF.	126
Scheme 2.11: Proposed mechanism for the diborane-induced cleavage of <b>86</b> from Wang resin.	127
Scheme 2.12: Coupling reactions with diethylamine.	128
Scheme 2.13: The KOH-mediated loading of <b>95</b> onto Merrifield.	134
Scheme 2.14: McKillop's ether synthesis.	135
Scheme 2.15: Mechanism for the reaction carried out by Davis and Muchowski.	136

## *List of Tables*

Table 1.1: PLP-dependent enzymes belonging to the $\alpha$ -family.	8
Table 1.2: Enzymes of the $\gamma$ -family.	14
Table 1.3: Minimal requirements of humans for some of the essential amino acids ( $\text{mg kg}^{-1} \text{ day}^{-1}$ ).	39
Table 1.4: Threonine synthase from microbial sources.	40
Table 1.5: Kinetic properties of threonine synthase from various higher plants.	43
Table 1.6: Inhibition of TS from different sources by various anions and salts.	59
Table 1.7: Inhibitors of threonine synthase.	61
Table 2.1: The results of ligations using <i>thrC</i> and pGEM-T.	86
Table 2.2: Results of ligations of new <i>thrC</i> fragment into pGEM-T.	91
Table 2.3: pTN-TS constructs produced from various pET-vectors and pGEM-TTS constructs.	94
Table 2.4: Description of primers used in sequencing experiments.	104
Table 2.5: Loading of Wang resin with <i>N</i> -TFA-asp under various conditions.	115
Table 2.6: The racemisation effects of varying amounts of DMAP.	117
Table 2.7: Effects of varying reaction conditions on the regioselectivity of Wang loading.	118
Table 2.8: Conditions for loading <i>p</i> -hydroxymethylpolystyrene resin.	121
Table 2.9: Conditions for attempted reductions of $\alpha$ - <i>p</i> -benzyloxybenzylpolystyrene- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-aspartate ( <b>87</b> ).	124
Table 2.10: Experiments to find reagent(s) responsible for cleavage of <b>86</b> from the resin.	125
Table 2.11: Reagents tried for the reduction of the activated $\beta$ -carboxyl group on <b>87</b> .	130
Table 2.12: Conditions used in experiments to load <b>95</b> onto Merrifield resin.	133
Table 2.13: Optimisation of the conditions for the loading of <b>95</b> onto Merrifield using potassium hydroxide.	137
Table 2.14: Comparison of the regioselectivity of substitution reactions of <b>73</b> onto <b>98</b> .	140
Table 3.1: Oligonucleotides used for amplification of gene fragments.	175
Table 3.2: Oligonucleotides used in sequencing experiments.	176

## *List of Compounds*

73	<i>N</i> -Trifluoroacetyl-(2 <i>S</i> )-aspartic anhydride	152
74	$\alpha$ -Isopropyl- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-aspartate	153
75	$\alpha$ -Isopropyl- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-homoserine	154
78	<i>N,N</i> -Diisopropylidichlorophosphamidite	155
79	<i>N,N</i> -Diisopropylbis(benzyl)phosphoramidate	156
76	$\alpha$ -Isopropyl- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-phosphohomoserinebis(benzyl) ester	157
32	(2 <i>S</i> )- <i>O</i> -phosphohomoserine	158
74b	[U- <sup>14</sup> C]- $\alpha$ -Isopropyl- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-aspartate	159
75b	[U- <sup>14</sup> C]- $\alpha$ -Isopropyl- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-homoserine	159
76b	[U- <sup>14</sup> C]- $\alpha$ -Isopropyl- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-phosphohomoserinebis(benzyl) ester	159
32b	[U- <sup>14</sup> C]-(2 <i>S</i> )- <i>O</i> -phosphohomoserine	159
81	Dipentafluorophenylphosphorochloridate	160
86	<i>N</i> -Trifluoroacetyl-(2 <i>S</i> )-aspartate	161
87	$\alpha$ - <i>p</i> -Benzyloxybenzylpolystyrene- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-aspartate	162
101	$\alpha$ - <i>p</i> -Benzyloxybenzylpolystyrene- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-aspartate- $\beta$ -methyl ester	163
94	$\alpha$ - <i>p</i> -Benzyloxybenzylpolystyrene- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-aspartate- $\beta$ -diethyl amide	164
102	$\alpha$ - <i>p</i> -Methylpolystyrene- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-aspartate	165
97	Polystyrene-4-oxymethyl-2-phenylethanol	166
103	$\alpha$ -[2-(4-Methoxyphenyl)-ethyl]- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-aspartate	166
99	$\alpha$ -(Polystyrene-4-oxymethyl-2-phenylethyl)- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-aspartate	167



## *Abbreviations*

AMP	adenosine monophosphate
Amp <sup>R</sup>	ampicillin resistance gene
AIDS	Acquired immune deficiency syndrome
AlaAT	alanine aminotransferase
AK	aspartokinase
APS	ammonium persulphate
AspAT	Aspartate aminotransferase
ATP	adenosine 5'-triphosphate
Boc	tert-butoxycarbonyl
bp	base pairs
<i>Br. flavum</i>	<i>Brevibacterium flavum</i>
<i>Br. lactofermentum</i>	<i>Brevibacterium lactofermentum</i>
BSA	bovine serum albumin
<i>B. stearothermophilus</i>	<i>Bacillus stearothermophilus</i>
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
cAspAT	cytosolic aspartate aminotransferase
Cbz (Z)	benzyloxycarbonyl
CD	circular dichroism
<i>C. glutamicum</i>	<i>Corynebacterium glutamicum</i>
CIAP	calf intestinal alkaline phosphatase
CNS	central nervous system
mCPBA	<i>m</i> -chloroperoxybenzoic acid
CS	cysteine synthase ( <i>O</i> -acetylserine sulfhydrylase)
Da	Dalton
DCM	dichloromethane
DDC	DOPA decarboxylase

ddNTP	chain terminating dideoxyribonucleotide triphosphates
DEA	diethylamine
DIPEA	diisopropylethylamine
DMAP	4-dimethylamino pyridine
DME	1,2-dimethoxyethane
DMF	<i>N,N</i> -dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphates
DOPA	3,4-dihydroxyphenylalanine
dsDNA	double-stranded DNA
DSDHT	D-serine dehydratase
DTT	dithiothreitol
EC	Enzyme Catalogue
<i>E. coli</i>	<i>Escherichia coli</i>
EDTA	ethylenediaminetetraacetic acid
eq	equivalents
Et <sub>3</sub> N	triethylamine
Et <sub>2</sub> O	diethyl ether
GABA	γ-aminobutyric acid
GABA-T	γ-aminobutyric acid transaminase
HDC	histidine decarboxylase
HEPES	( <i>N</i> -[2-hydroxyethyl]piperazine- <i>N'</i> -[2-ethanesulfonic acid])
<i>H. influenzae</i>	<i>Hermophilus influenzae</i>
HK	homoserine kinase
HRMS	high resolution mass spectrometry
HSD	homoserine dehydrogenase
IBCF	isobutylchloroformate
IPTG	isopropyl-β-D-thiogalactopyranoside

IR	infrared
LMP	low-melting point
kb	kilobases
$K_d$	dissociation constant
kDa	kilo Dalton
$K_i$	enzymic inhibition constant
$K_M$	Michaelis constant
KOH	potassium hydroxide
<i>lacZ</i>	gene for $\beta$ -galactosidase
LB	Luria-Bertani medium
LSDHT	L-serine dehydratase
<i>mAspAT</i>	mitochondrial aspartate aminotransferase
<i>M. glycogenes</i>	<i>Methylobacillus glycogenes</i>
<i>M. leprae</i>	<i>Mycobacterium leprae</i>
$M_r$	relative atomic mass
mRNA	messenger RNA
<i>M. tuberculosis</i>	<i>Mycobacterium tuberculosis</i>
NAD	nicotinamide adenine dinucleotide
NADP	nicotinamide adenine dinucleotide phosphate
<i>N. crassa</i>	<i>Neurospora crassa</i>
Ni-NTA	nickel-nitrilo-tri-acetic acid
NMDA	<i>N</i> -methyl- <i>D</i> -aspartate
NMM	<i>N</i> -methylmorpholine
NMR	nuclear magnetic resonance
OASS	<i>O</i> -acetylserine sulfhydrylase (cysteine synthase)
ODC	ornithine decarboxylase
<i>ompT</i>	outer membrane protease
O/N	overnight
ORF	open reading frame

ORI	replication origin
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
PBS	phosphate buffered saline
PCR	polymerase chain reaction
pET	plasmid for expression by T7 RNA polymerase
<sup>32</sup> P-γ-ATP	adenosine triphosphate, <sup>32</sup> P-labelled at the gamma phosphate
P <sub>i</sub>	inorganic orthophosphate
PLP	pyridoxal-5'-phosphate
PMP	pyridoxamine-5'-phosphate
PyBOP	benzotriazolylloxy-tris[pyrrolidino]-phosphonium hexafluorophosphate
RBS	ribosomal binding site
RNA	ribonucleic acid
SAM	S-adenosylmethionine
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i> (yeast)
SDS	sodium dodecylsulfate
SDS-PAGE	SDS-polyacrylamide gel electrophoresis
<i>S. marcescens</i>	<i>Serratia marcescens</i>
<i>S. pombe</i>	<i>Schizosaccharomyces pombe</i> (fission yeast)
ssDNA	single-stranded DNA
STET	sodium-tris-EDTA-triton buffer
TAE	Tris-acetate-EDTA buffer
<i>T. aquaticus</i>	<i>Thermus aquaticus</i>
TBE	Tris-boric-EDTA buffer
TBS	Tris buffered saline
TE	Tris-EDTA buffer
TEMED	<i>N,N,N',N'</i> -tetramethylethylenediamine
TES	triethylsilane
TFAA	trifluoroacetic anhydride

TFA	trifluoroacetic acid
THF	tetrahydrofuran
<i>thr</i>	threonine operon
<i>thrA</i>	aspartate kinase I- homoserine dehydrogenase I
<i>thrB</i>	gene for homoserine kinase
<i>thrC</i>	gene for threonine synthase
tlc	thin layer chromatography
T4 PNK	T4 polynucleotide kinase
tRNA	transfer ribonucleic acid
Tr S	tryptophan synthase
TS	threonine synthase
UV	ultraviolet
DV	$V_H/V_D$
DV/K	$(V_H/V_D)/(K_H/K_D)$
X-gal	5-bromo-4-chloro-3-indolyl- $\beta$ -D-galactopyranoside
Z-APPA	(2 <i>S</i> )-2-amino-5-phosphono-3- <i>cis</i> -pentenoic acid
3D	3-dimensional

## Contents

Acknowledgements	i
Abstract	ii
List of Figures	iii
List of Schemes	v
List of Tables	vii
List of Compounds	viii
Abbreviations	ix
1.0 Introduction	1
1.1 An Introduction to Pyridoxal-5'-Phosphate Dependent Enzymes	1
1.1.1 The Scope of Pyridoxal-5'-Phosphate Dependent Enzymes	1
1.1.2 The Discovery of Pyridoxal-5'-Phosphate	2
1.2 The Classification of PLP Enzymes	6
1.2.1 The Evolution of PLP-Dependent Enzymes	6
1.2.2 The PLP-Enzyme Families	7
1.2.2.1 The $\alpha$ -Family of PLP-Dependent Enzymes	8
1.2.2.2 The $\beta$ -Family of PLP-Dependent Enzymes	11
1.2.2.3 The $\gamma$ -Family of PLP-Dependent Enzymes	13
1.2.2.4 Other PLP Enzymes	14
1.2.2.5 Exceptions to the Family Rule	14
1.3 The Structure and Mechanism of PLP Enzymes	15
1.3.1 A General Mechanism of PLP-Enzyme Action	15
1.3.2 Recent Structural and Mechanistic Studies on $\beta$ -Family Enzymes	19
1.3.2.1 Recently Published Crystallographic Data on PLP Enzymes	19
1.3.2.2 Structure-Activity Studies on <i>O</i> -Acetylserine Sulphydrolase	21
1.3.2.3 Structure-Activity Studies on Tryptophan Synthase	24
1.4 The Metabolic Roles of Various Pyridoxal-5'-Phosphate Dependent Enzymes	27
1.4.1 Niacin (Nicotinic Acid) Synthesis From Tryptophan	27
1.4.1.1 Pellagra and Related Diseases	28
1.4.2 Decarboxylases	30
1.4.2.1 Glutamate Decarboxylase and $\gamma$ -Aminobutyric Acid Transaminase	30
1.4.2.2 Ornithine Decarboxylase (ODC)	31
1.4.2.3 Histidine Decarboxylase (HDC)	33
1.4.2.4 DOPA Decarboxylase (DDC)	34
1.4.3 Amino Acid Racemases	35

1.4.4	The Aspartate Pathway	36
1.5	Introduction to Threonine Synthase	37
1.5.1	The Discovery of Threonine Synthase	37
1.5.2	The Associated Deaminase Activity of Threonine Synthase	38
1.5.3	The Regulation of Threonine Synthase	39
1.6	The Structure and Character of Threonine Synthase from Various Organisms	40
1.6.1	Microbial Sources of Threonine Synthase	40
1.6.2	Threonine Synthase from Plant Sources	42
1.6.3	Threonine Synthase Expression in Mammalian Systems	44
1.6.4	Threonine Synthase as a Potential Target Enzyme for Herbicides	44
1.7	The Threonine Operon of <i>Escherichia coli</i>	45
1.7.1	The Structure of the Operon	45
1.7.2	The Regulation of <i>Thr</i> Expression	46
1.7.3	Cloning of <i>Thr</i>	47
1.8	The Mechanism of (2 <i>S</i> ,3 <i>R</i> )-Threonine Synthesis	48
1.8.1	Introduction	48
1.8.2	Reaction Mechanism	48
1.8.3	Unanswered Questions Concerning the Threonine Synthase Reaction	55
1.9	Inhibitors of Threonine Synthase	56
1.9.1	Functional Groups Which Inhibit PLP-Enzymes	56
1.9.2	Amino Acid Inhibitors of Threonine Synthase	57
1.9.3	Inhibition of Threonine Synthase by Various Anions	59
1.9.4	Phosphonic Acid Analogues of (2 <i>S</i> )- <i>O</i> -Phosphohomoserine	60
1.9.5	Phosphate Analogues	62
1.10	The Aims of This Project	63
1.10.1	Testing New Substrate Analogues as Inhibitors of Threonine Synthase	63
1.10.2	Product Inhibition Assays	64
1.10.3	Further Assessment of the Allosteric Activation of TS by SAM	65
1.10.4	Structure-Activity Studies on Threonine Synthase	65
1.11	pET Expression Vectors	66
1.11.1	Introduction	66
1.11.2	T7 RNA Polymerase	66
1.11.3	Expression of Target Genes	67
1.11.4	Vectors Selected For Expressing Wild-Type and Mutant <i>ThrC</i>	68

2.0 Results and Discussion	69
2.1 Substrate Synthesis	69
2.1.1 A Survey of the Literature Methods for Synthesising (2 <i>S</i> )- <i>O</i> -Phosphohomoserine.	69
2.1.2 The Biomimetic Synthesis of (2 <i>S</i> )- <i>O</i> -Phosphohomoserine.	71
2.1.3 Synthesis of [U- <sup>14</sup> C]-(2 <i>S</i> )- <i>O</i> -Phosphohomoserine.	74
2.2 The Growth of Threonine Synthase	76
2.3 Cloning of <i>ThrC</i> and the Gene for $\beta$ -Methylaspartase	77
2.3.1 The General Cloning Strategy	77
2.3.2 Cloning of <i>ThrC</i>	79
2.3.2.1 Amplification of <i>ThrC</i>	79
2.3.2.1.1 Isolation of Chromosomal DNA	79
2.3.2.1.2 Polymerase Chain Reaction on <i>ThrC</i>	79
2.3.2.2 Attempted Ligation of <i>ThrC</i> Into Cloning Vectors	81
2.3.2.3 Overlap Polymerase Chain Reaction Experiments.	83
2.3.2.4 Ligation of <i>ThrC</i> into pGEM-T.	84
2.3.2.5 Attempted Ligation of <i>ThrC</i> Into Expression Vectors.	87
2.3.2.5.1 pET-3a and pET-16b.	87
2.3.2.5.2 pKK223-3	87
2.3.2.6 New PCR Fragment.	88
2.3.2.7 Attempted Ligation of <i>ThrC</i> into pET-3a.	89
2.3.2.8 Ligation of New Gene Fragment Into pGEM-T.	90
2.3.2.8.1 Amplification of <i>ThrC</i>	90
2.3.2.9 Sequencing of <i>ThrC</i> in pGEM-TTS <sub>I</sub> and pGEM-TTS <sub>II</sub> .	92
2.3.2.10 Ligation of <i>ThrC</i> Into pET-Expression Vectors.	94
2.3.3 Cloning of the $\beta$ -Methylaspartase Gene	98
2.3.3.1 Cloning the Gene Into pGEM-T	98
2.3.3.1.1 Gene Amplification	98
2.3.3.1.2 Overlap PCR Experiments	99
2.3.3.1.3 pGEM-T	100
2.3.3.1.4 A New Gene Fragment	101
2.3.3.1.5 New pGEM-TMA Constructs	102
2.3.3.2 Further Work on the Gene for $\beta$ -Methylaspartase	103
2.4 Sequencing of the <i>ThrC</i> Gene Fragment in Recombinant Plasmids	104



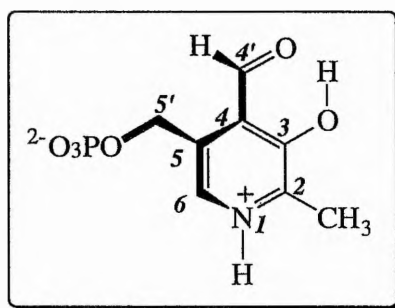
2.5 Attempted Solid-Phase Synthesis of (2 <i>S</i> )- <i>O</i> -Phosphohomoserine	106
2.5.1 Alternative Phosphorylating Agents	106
2.5.2 <sup>19</sup> F Gel-phase NMR Spectroscopy	111
2.5.3 <i>p</i> -Benzyloxybenzyl Alcohol (Wang) Resin	112
2.5.3.1 The Loading of Wang Resin With <i>N</i> -Trifluoroacetyl-(2 <i>S</i> )-aspartate	114
2.5.3.2 <i>p</i> -Hydroxymethylpolystyrene Resin	121
2.5.3.3 Attempted Reduction of $\alpha$ - <i>p</i> -Benzyloxybenzylpolystyrene- <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-aspartate	122
2.5.3.4 Experiments to Confirm the Activation of the Carboxylic Group on <b>87</b>	128
2.5.3.5 Reduction of $\alpha$ - <i>p</i> -Benzyloxybenzylpolystyrene- <i>N</i> -Trifluoroacetyl-(2 <i>S</i> )-aspartate	129
2.5.4 A New Solid Support	132
2.5.4.1 Attachment of 2-( <i>p</i> -Hydroxyphenyl)-ethanol to Merrifield Resin	132
2.5.4.2 Loading of the New Resin With <i>N</i> -trifluoroacetyl-(2 <i>S</i> )-aspartate	137
2.6 Conclusions and Future Work	142
3.0 Experimental	149
3.1 Synthesis	149
3.1.1 General Experimental Procedures	149
3.1.1.1 Compound Characterisation	149
3.1.1.2 Reagent and Solvent Preparation	150
3.1.1.3 Suppliers	151
3.1.2 Synthetic Procedures and Characterisations	152
3.2 Molecular Cloning	169
3.2.1 Cloning Techniques	169
3.2.2 Polymerase Chain Reaction	173
3.2.3 Oligonucleotides	175
3.2.4 Expression in <i>E. coli</i>	176
References	179
Appendices	191

**CHAPTER ONE**  
**INTRODUCTION**

## 1.1 An Introduction to Pyridoxal-5'-Phosphate Dependent Enzymes

### 1.1.1 The Scope of Pyridoxal-5'-Phosphate Dependent Enzymes

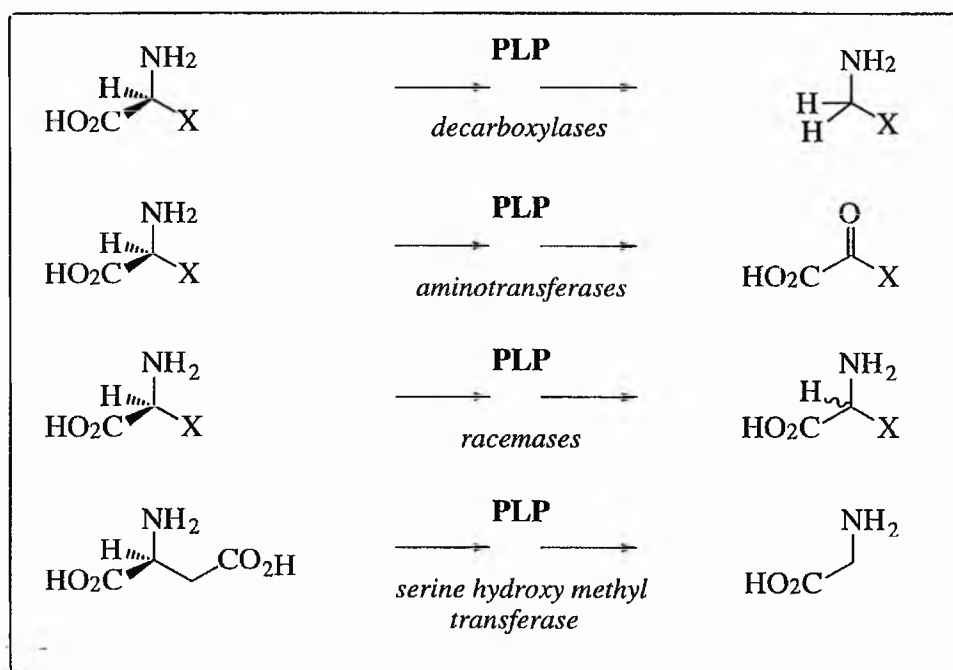
A wide variety of enzymes responsible for amino acid metabolism utilise pyridoxal 5'-phosphate (PLP, **1**) as their cofactor. The enzymes dependent on PLP include the transaminases<sup>1</sup> (or aminotransferases), through whose action the C<sup>α</sup>-N bonds of most amino acids are formed, the racemases,<sup>2</sup> many of which are involved in the biosynthesis of peptidoglycan in bacterial cell walls, and the decarboxylases which convert amino acids to amines (Scheme 1.1).<sup>3</sup> The latter function is one of the most important physiological roles of PLP.



*Pyridoxal-5'-phosphate (1).*

There are over 66 known transaminases altogether as well as 50 different decarboxylases. All the aforementioned enzymes alter a substituent at the  $\alpha$ -position on the substrate. Many other enzymes which are PLP-dependent, in addition to first removing the  $\alpha$ -proton of the amino acid, also eliminate or replace groups which are in a  $\beta$ - or a  $\gamma$ -position. The numerous examples of PLP-dependent enzymes indicate that PLP is one of nature's most versatile coenzymes; playing a crucial role in amino acid metabolism,<sup>4</sup> connecting the carbon and nitrogen cycles and providing entry into the 'one-carbon pool'. PLP also participates in the formation of biogenic amines as well as a number of other important processes.<sup>5</sup> The versatility of this coenzyme is reflected in the fact that PLP-dependent enzymes occur

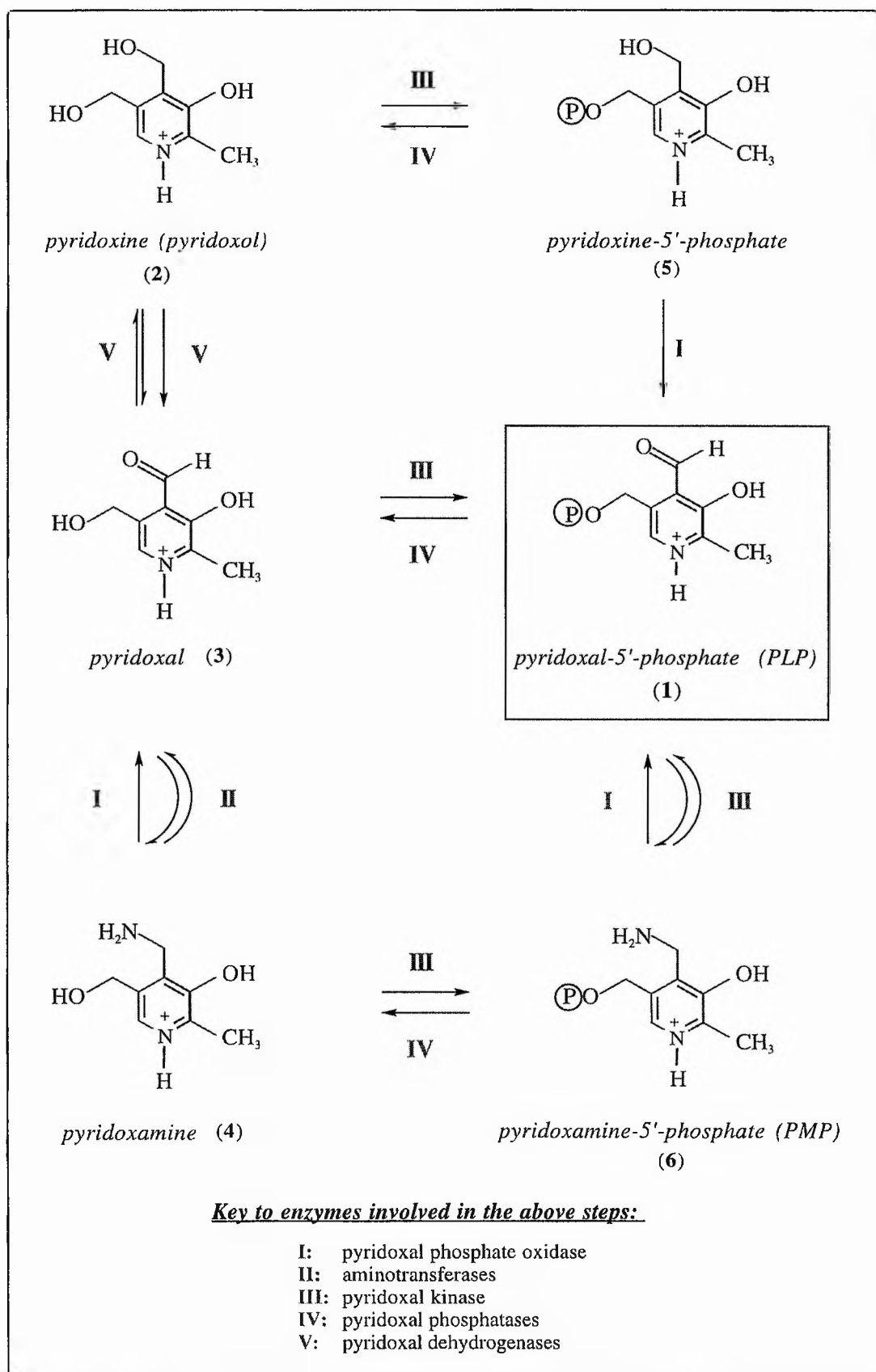
in four of the six EC classes of enzymes.<sup>6</sup>



**Scheme 1.1:** Some types of reaction catalysed by PLP-dependent enzymes.

### 1.1.2 The Discovery of PLP

PLP is a metabolite of vitamin B<sub>6</sub> (or pyridoxol (2), see Scheme 1.2), a nutritional factor which was first identified in 1934 by Paul György.<sup>7</sup> In the early 1930s, when purified preparations of thiamine (B<sub>1</sub>) and riboflavin (B<sub>2</sub>) first became available, it became apparent that a separate heat-stable vitamin, later found to be vitamin B<sub>6</sub>, could arrest the development of a pellagra-like dermatitis (similar to acrodermatitis) in rats (Fig. 1.1).<sup>4</sup> In fact, as early as 1926, Goldberger and Lille had noted that rats given a diet deficient in a so-called 'pellagra-preventative factor' developed this condition.<sup>8</sup> In 1932 Ohdake isolated a nitrogenous compound from rice polishings, but no vitamin significance was attached to it at the time.<sup>9</sup> It was only later that it was shown to be identical with vitamin B<sub>6</sub>. Birch and György studied the properties of the vitamin in 1936<sup>10</sup> and subsequent to this the compound was isolated independently by five

Scheme 1.2: The metabolism of vitamin B<sub>6</sub>.

laboratories in 1938 and its structure and synthesis were reported in 1939 when it was named pyridoxine.<sup>11-14</sup>

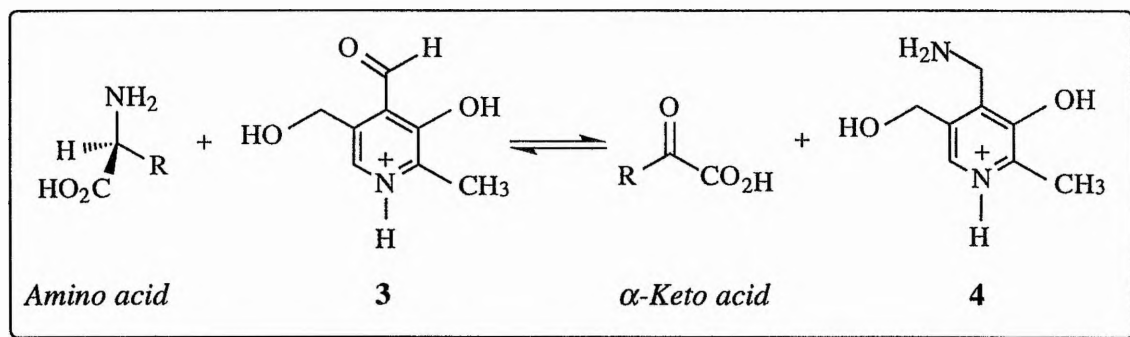


**Figure 1.1:** *Dermatitis caused by pellagra.*<sup>a</sup>

Originally the cofactor was named 'codecarboxylase' by many research workers as a result of its ability to reactivate the resolved inactive *apoenzyme* of (2*S*)-lysine decarboxylase.<sup>15</sup> It was not until Bellamy and Gunsalus were investigating the properties of tyrosine decarboxylase that the link between the nutritional factor B<sub>6</sub> and 'codecarboxylase' was made.<sup>16</sup> In their experiments media containing larger amounts of pyridoxine (2) and nicotinic acid than were necessary for growth of bacteria, were found to give high yields of the active enzyme. A couple of years later Snell realised that vitamin B<sub>6</sub> occurs in multiple free forms, and by studying the non-enzymic transaminations between pyridoxal (3) and amino acids he provided the first clues to the function of the vitamin.<sup>17</sup> The reaction shown in Scheme 1.3 was shown to occur in the case of (2*S*)-glutamic acid, and the data suggested that a similar reaction could

<sup>a</sup> reproduced from web-site: <http://www.mc.vanderbilt.edu/biolib/hc/pd4.html>

take place between pyridoxal (3) and certain other amino acids. It is now known that most amino acids are biosynthesised in this manner from (2*S*)-glutamic acid and the corresponding  $\alpha$ -keto acid *via* the action of transaminases.



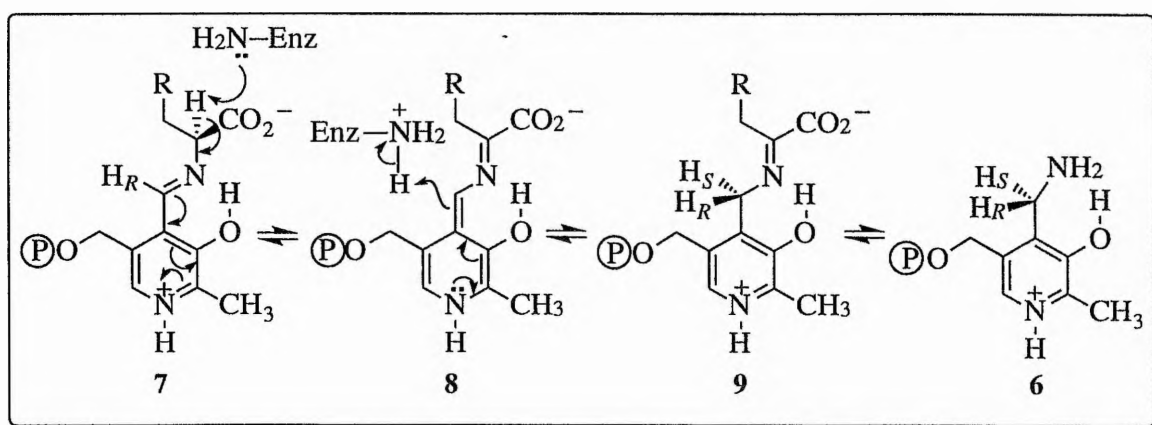
**Scheme 1.3:** Snell's "Shuttle Mechanism" for enzymic transamination.

The next decade witnessed the identification of tryptophanase, (2*R*)-serine dehydratase and cysteine desulfhydrase as PLP-enzymes.<sup>4</sup> The availability of further enzymes for study meant that much stereochemical, kinetic and primary structural information about PLP enzymes could be gained, providing insights into the general mechanisms of such enzymes and the role of PLP. 3D structures of most PLP enzymes and their enzyme-substrate complexes are, however, still scarce, except in the case of the transaminase group of enzymes.

## 1.2 The Classification of PLP Enzymes

### 1.2.1 The Evolution of PLP-Dependent Enzymes

All the PLP enzymes studied to date share certain similarities in the preliminary stages of their reaction mechanisms (see also Section 1.3). In 1974, Dunathan and Voet also noted that several PLP-dependent enzymes convert the bound cofactor to pyridoxamine phosphate (6) during normal or abortive transamination reactions and that this proceeds stereospecifically, with the same absolute stereochemistry in seven quite different PLP-enzymes.<sup>18</sup> The key step in this reaction is the protonation of the bound substrate-pyridoxal phosphate Schiff base (7) at the C-4' carbon of the cofactor, from the 4'-*si* face, after loss of one of the groups at the substrate C<sup>α</sup> (Scheme 1.4). They inferred that this demonstrated a remarkable regularity in the geometry of the binding of the cofactor to the apoenzyme in all PLP enzymes, and interpreted this as evidence for the evolution of this entire family of enzymes from a common progenitor. Jensen's hypothesis supports this theory, as it suggests that ancestral organisms produced a relatively small number of enzymes with flexible substrate specificity, which gradually mutated into more selective enzymes.<sup>19</sup>



**Scheme 1.4:** Conversion of bound cofactor to pyridoxamine phosphate- a process common to many PLP-enzymes.  $H^{\alpha}$ -elimination is illustrated here.



However, not only do PLP enzymes share many mechanistic similarities, they also demonstrate definite sequence and structural homologies. Mehta *et al.* found that most aminotransferases formed a group of homologous proteins, which could then be divided further into distinct evolutionary subgroups on the basis of their sequence similarities (see Section 1.2.2.1).<sup>20</sup> Subsequent profile analysis to detect distant relationships between amino acid sequences of various enzymes showed that the evolutionary relationships included other B<sub>6</sub> enzymes.<sup>21-23</sup> In 1994 Christen and co-workers carried out a comprehensive analysis of the amino acid sequences of PLP enzymes and of their 3D structures and, as a result, proposed that most of the enzymes could be assigned to one of three different families of homologous proteins; the  $\alpha$ -,  $\beta$ - or  $\gamma$ -families.<sup>24</sup> Although the sequence similarity is too low in some cases to allow meaningful alignments by standard programmes, hydropathy plots and secondary structure predictions do indicate evolutionary relationships between PLP enzymes.<sup>25</sup> Their affiliation with one of these structurally defined families also correlates in most cases with their regiospecificity.<sup>24</sup> The folding pattern of the polypeptide chain of the  $\gamma$ -family enzyme is similar to that of aspartate aminotransferase (AspAT), further suggesting an evolutionary relationship between the  $\alpha$ - and  $\gamma$ -families.<sup>24</sup> Christen *et. al* also cited various reports of the side reactions catalysed by B<sub>6</sub> enzymes as further evidence for a common ancestral enzyme with broad substrate specificity.<sup>24</sup> Decarboxylases catalyse transamination, aminotransferases cause racemisation, tryptophan synthase (TrS) produces racemisation and transamination and glycine hydroxymethyltransferase also initiates  $\alpha$ -decarboxylation, transamination and racemisation.

### 1.2.2 The PLP-Enzyme Families

The largest family of PLP-dependent enzymes, the  $\alpha$ -family, contains enzymes catalysing reaction at C <sup>$\alpha$</sup>  of the substrate. Similarly, the  $\beta$ -family contains those acting at C <sup>$\beta$</sup>  and the  $\gamma$ -family those acting at C <sup>$\gamma$</sup> . There are, however, some exceptions to this general rule, and some PLP-dependent enzymes do not seem to belong to any of the groups.

### 1.2.2.1 The $\alpha$ -Family of PLP-Dependent Enzymes<sup>24</sup>

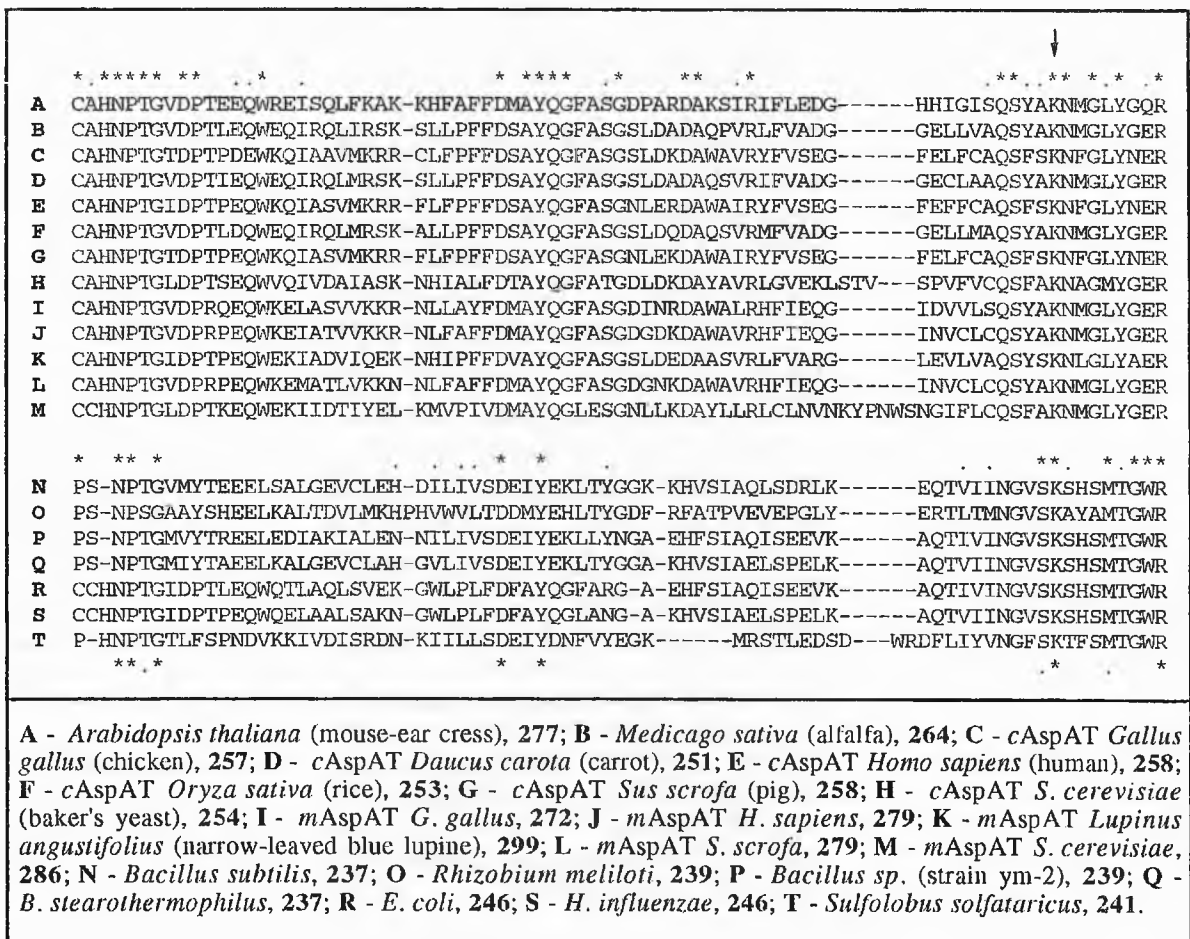
The  $\alpha$ -family is the largest of all the PLP-dependent groups and includes the aminotransferases (ATs) and the group II decarboxylases (DCs). In spite of the low level of sequence similarity between members of the  $\alpha$ -family, (only two residues are conserved throughout the  $\alpha$ -family; the lysine and aspartate residues which bind PLP),<sup>25</sup> many of these enzymes have 3D structures available for comparison and the conformational similarities between different types of enzyme can clearly be seen. John reviews these similarities in more detail, but they include eight or nine  $\alpha$ -helices packed around a seven-stranded  $\beta$ -sheet in the large domain.<sup>25</sup> In the  $\alpha$ -family the PLP-binding lysine residue occurs between residues 209 and 256 of the various members of this family, which are listed in Table 1.1.

**Table 1.1:** *PLP-dependent enzymes belonging to the  $\alpha$ -family.*

<i>Most aminotransferases (excluding branched chain and D-amino acid)</i> <sup>20, 26</sup>
<i>Group II amino acid decarboxylases</i>
<i>Tryptophanase</i>
<i>Tyrosine phenol-lyase</i> <sup>27</sup>
<i>Glycine hydroxymethyltransferase</i>
<i>1-aminocyclopropane-1-carboxylate synthase</i> <sup>23</sup>
<i>2-amino-6-caprolactam racemase</i>
<i>Glutamate-1-semialdehyde 2,1-aminomutase</i>
<i>Isopenicillin N-epimerase</i>
<i>2,2-dialkyl decarboxylase</i> <sup>28</sup>
<i>4-amino-4-deoxychorismate synthase</i> <sup>22, 29</sup>
<i>Glycine-C-acetyltransferase</i>
<i>5-Aminolevulinate synthase</i>
<i>8-Amino-7-oxononanoate synthase</i>
<i>The gene product of cobC (cobalamin synthesis)</i>
<i>The gene product of nifS (nitrogen fixation)</i>
<i>The gene product of malY (abolishes induction of the maltose system)</i>

The most intensely studied members of this family are the ATs. In their studies on the evolutionary relationship between ATs, Mehta and co-workers discovered 107 invariant

residues between the *cAspAT* and *mAspAT* sequences from vertebrates, which have an average length of around 400 amino acids.<sup>26</sup> Their average pairwise identity was 41%. The active-site sequences of AspATs from both prokaryotic and eukaryotic sources have been aligned here showing the conservation of various motifs, including the lysine residue for binding PLP (Fig. 1.2).



Sequences A - M are eukaryotic, sequences N - T are from prokaryotes. Residue number of the active-site lysine is shown in blue. Asterisks denote identical residues, periods denote conservative replacements (e.g. I-V-L-M-F, D-E, R-K, S-T). For the eukaryotic sequences, identical residues are highlighted in red. For the prokaryotic sequences, identical residues are highlighted in blue. The homology of each set of sequences is also indicated by asterisks and periods above the sequences. The overall homology between all the AspATs aligned here are indicated by the asterisks and periods underneath both sets of sequences. The lysine residue responsible for binding to PLP is indicated by an arrow. The alignment was performed using Clustalw (1.60).

**Figure 1.2:** Sequence homology between various eukaryotic and prokaryotic AspATs close to the active-site lysine residue.

When the various vertebrate ATs were aligned the percentage identity ranged from as low as 5% up to 75%, whereas the percentage similarity (which accounts for conservative replacements) ranged from 25% to as high as 87%.<sup>20</sup> Four residues were conserved throughout the sequences: G197 which participates in a  $\beta$ -turn at the domain interface, D/E222 which forms a salt-bridge or hydrogen bond to N1 of PLP, K258 which forms a Schiff base with PLP and R386 which forms a hydrogen bond with the  $\alpha$ -carboxylate of the substrate.<sup>20</sup> Mehta reviews various site-directed mutagenesis experiments to replace these residues in *E. coli* AspAT which resulted in significant reductions in the catalytic activity of the enzyme.<sup>20</sup>

On the basis of sequence similarities the ATs could be divided into four subgroups. 11 residues were invariant in subgroup I, 19 in subgroup II, 82 in subgroup III and 25 in subgroup IV.<sup>20</sup> Jensen and Gu have subsequently reviewed the evolutionary relationship between 47 group I ATs and proposed seven new subfamilies on the basis of sequence alignments.<sup>30</sup> They suggested that these originate from a single ancestral broad specificity AT. Another study has demonstrated structural homology between serine hydroxymethyltransferase (SHMT) and AspAT, with 66% of aligned residues around the central and C-terminal regions predicted to share the same conformation.<sup>31</sup> Bossa also cited a previous observation of weak sequence similarity between the active-site regions of decarboxylases and various PLP enzymes, including SHMT.<sup>31</sup> In addition to this, the NifS group of proteins (responsible for nitrogen fixation in bacteria) appears homologous to a new class of PLP-enzymes, which includes various aminotransferases and isopenicillin N epimerase,<sup>32</sup> although they all have very different functions. The sequence identity between distant members within this class is below 25%, but secondary structure predictions indicate that this class of enzymes might have an overall  $\alpha$ - $\beta$  topology. Sander and Ouzounis also reported sequence similarities of around 30% between mammalian peroxisomal serine-pyruvate aminotransferases and the small subunit of soluble hydrogenases.<sup>32</sup> Most recently Babbitt and Gerlt have observed sequence similarities between 5-keto-4-deoxyglucarate dehydratase/decarboxylase and various dehydratases and aldolases.<sup>33</sup> In a separate study on decarboxylases Kang and Joh have demonstrated significant homologies between porcine and bovine L-amino acid decarboxylases and feline glutamic acid

decarboxylase and tyrosine decarboxylase from *C. roseus* (31% identity between all sequences).<sup>34</sup> Christen *et al.* have more recently compared sequences from nine different DCs and divided them into four subgroups.<sup>35</sup> However, only group II demonstrated a distant relationship with the ATs. No evidence was obtained that the other DCs were in any way related to other PLP enzymes and profile analysis suggests that the decarboxylases have evolved along multiple lineages, even if they have the same substrate specificity.

#### 1.2.2.2 The $\beta$ -Family of PLP-Dependent Enzymes<sup>24</sup>

This family has seven members including *threonine synthase* (TS), which are shown in Fig. 1.3. Although few crystal structures are available for the comparison of the tertiary structures of these enzymes, their sequences do show significant homology. Parsot has demonstrated extensive homology between the TS enzymes of *E. coli* and *B. subtilis*, threonine dehydratase (TD) from *S. cerevisiae* and D-serine dehydratase (DSDHT) from *E. coli*.<sup>19</sup> There are 72 conserved amino acid residues (21%) and 40 conservative replacements between TS from *B. subtilis* and TD from yeast; an overall homology of 33%. Similar results were noted between the sequences of yeast TD and DSDHT of *E. coli*. As a result he postulated the evolutionary connection of these enzymes. Bork and Rohde also aligned various  $\beta$ -family enzymes from a number of microorganisms and identified three regions of significant homology in the sequences including a lysine motif and a glycine-rich turn, which are illustrated in Fig. 1.4.<sup>36</sup> They proposed a new alignment, which revealed a similar folding topology in long segments of the enzymes studied. The pairwise identities ranged from 15% to 19%. They also found that the sequences of TrS and *O*-acetylhomoserine thiolase, a  $\gamma$ -family enzyme, had an amino acid identity of 11% and suggested that there could exist an evolutionary relationship between these two families, although the findings of Christen *et al.* refute this.<sup>24</sup> Appendix 1 shows a more extensive alignment of  $\beta$ -family enzymes from *B. subtilis*, *E. coli* and yeast whose pairwise identities ranged from as low as 4% between DSDHT and TD of *B. subtilis* to 100% between TD and TS from this source. The average pairwise identity was 15.7%, which increased to 24.9% around the active-site.

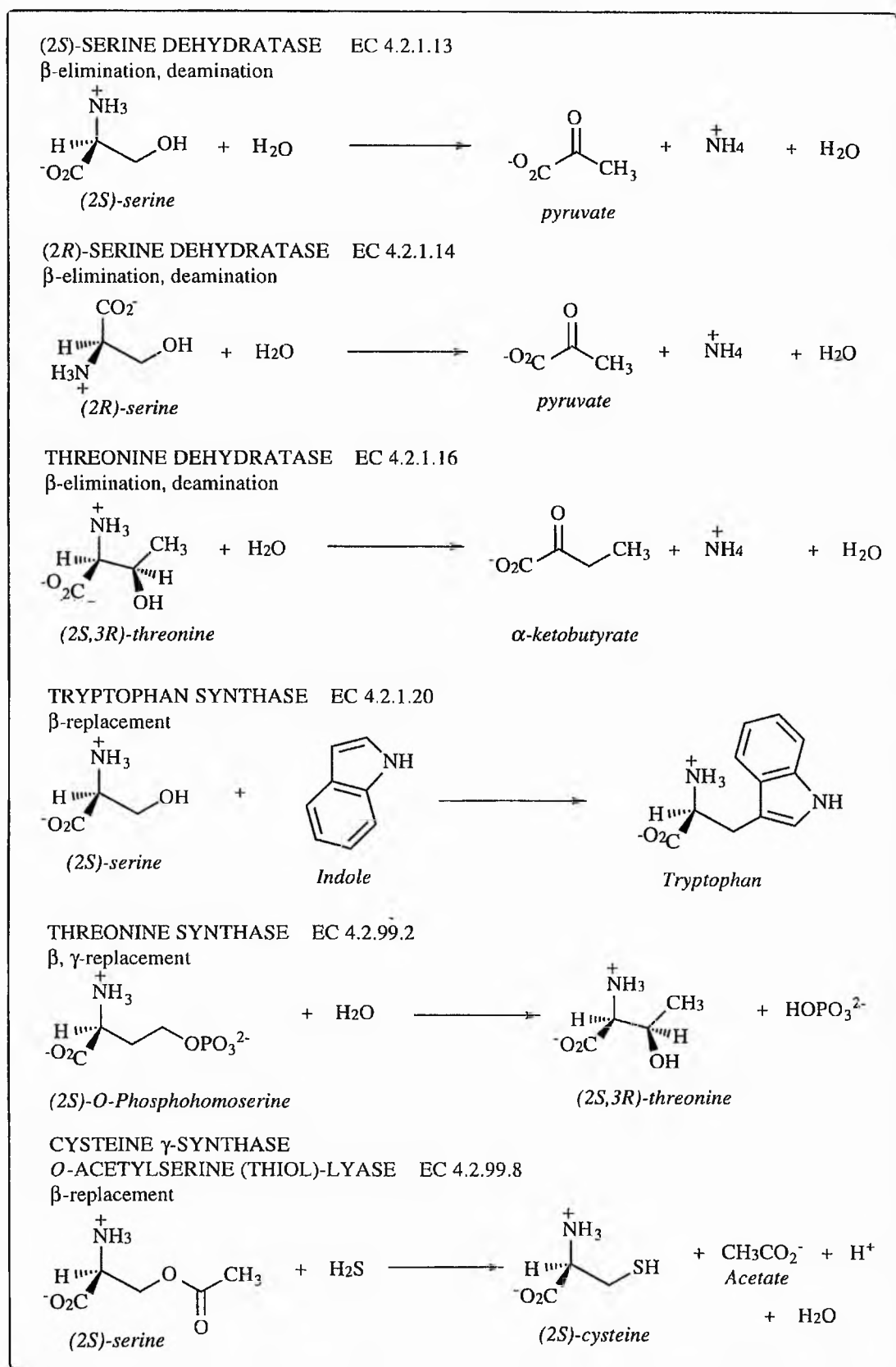


Figure 1.3: A summary of the reactions carried out by enzymes of the β-family of PLP-dependent enzymes

			↓			
cs_bacill	39	---	PG-SSVKDRI-GLAMIEAAE----	K	58	167 DD-QLDAFVAGIG-TGGTTTGAGEVLK---E 199
cs_e_coli	36	---	PS-FSVKCRI-GANMIWDAE----	K	55	165 DG-QVDVFIAGVG-TGGTLTGVSRYIKG-TK 183
cs_yeast	81	---	PA-GSAKDRV-ALNI IKTAE----	E	100	219 -KGNIDAFIAGCG-TGGTTTGVAKFLKERAK 248
dsdht_bac	111	HELPI	SGSIKARGGIYEVLYKAENLALQ		136	266 SLETPLFVYLPCG-VGGPGGVAFGLKLLYG 296
lsdht_yea	34	---	PG-GSFKSRG-IGHLIRKSN----	Q	53	171 SLPRVKALVCSVG-CGGLFSGIIRKGLDR-NH 200
td_bacill	54	---	PT-GSFKDRG-MVMAVAKAK----	E	73	174 -GEAPDVLAIPVG-NAGNITAYWKGFKEYHE 203
td_e_coli	57	---	PV-HSFKLRG-AYAMMAGLT----	E	76	178 --AHLDRVFVPVG-CGGLAAGVAVLIKQLMP 206
td_yeast	104	---	PV-FSFKLRG-AYNMIAKLD----	D	123	225 TANKIGAVFVPVG-CGGLIAGIGAYLKRVP 256
trs_bacil	85	---	HT-GSHKINN-ALGQALLAK----	K	104	223 EGTMPDKVVACVG-GGSNAMGMFQAFIN-ED 252
trs_e_col	81	---	HG-GAHKTNQ-VLGQALLAK----	R	100	219 EGRLPDVAIACVG-GGSNAIGMFADFIN-ET 248
ts_bacill	54	---	PT-GSFKDRG-MVMAVAKAK----	E	73	174 -GEAPDVLAIPVG-NAGNITAYWKGFKEYHE 203
ts_e_coli	102	FHGPT-	LAFKDFG-GRFMAQMLT----	H	121	236 ETRNQLVVVPSGNFGDLTAGLLAKSLG-LP 266
ts_yeast	119	HG-PT-	YAFKDVA-LQFVGNLFE----	Y	138	265 KDSKKVKFVVPVPSGNFGDILAGYFAKRMGLPI 296
			*			*
			---	P--GSFK-RG-----A-----		-----D----PVG--GG---G-----K----

The lysine residue responsible for binding to PLP is indicated by an arrow. The consensus sequence is given below the other sequences. Homologous residues are indicated by a period, identical residues by an asterisk.

**Figure 1.4:** Sequence homology of  $\beta$ -family enzymes around the lysine motif and the glycine-rich turn.

Further evidence of the evolutionary relationship between the  $\beta$ -family enzymes is found in their ability to catalyse similar reactions and produce similar intermediates, suggesting a similar active-site conformation. TS from *B. subtilis* and *E. coli* still exhibit some TD activity both *in vivo* and *in vitro* and both TD and TS produce  $\alpha$ -aminocrotonate as an intermediate.<sup>19</sup> TrS can catalyse the conversion of vinylglycine to  $\alpha$ -ketobutyrate, a reaction normally carried out by TS.<sup>37</sup> TS also catalyses various side-reactions (Appendix 2).

In contrast to members of the  $\alpha$ -family, in the  $\beta$ -enzymes the PLP-binding lysine is positioned in the NH<sub>2</sub>-terminal segment of the polypeptide chain between residues 41 and 118.<sup>24</sup> This difference may reflect the alternative PLP arrangement needed for the enzymes to carry out reactions at the C <sup>$\alpha$</sup> -centre rather than at the C <sup>$\beta$</sup> -atom.

### 1.2.2.3 The $\gamma$ -Family of PLP-Dependent Enzymes<sup>24</sup>

The  $\gamma$ -family contains four members and is therefore the smallest group of PLP enzymes (Table

1.2 ). Belfaiza *et al.* noted an overall homology of 36% (31% identity) between the two cystathionase enzymes.<sup>38</sup>

Alexander *et al.* found that the PLP-binding lysine residue is in the same sequence segment as in the  $\alpha$ -family suggesting a possible evolutionary relationship between these two families.<sup>24</sup> As stated earlier, significant sequence similarities between TrS (of the  $\beta$ -family) and *O*-acetylhomoserine sulfhydrylase of this family have also been shown.<sup>36</sup>

**Table 1.2:** *Enzymes of the  $\gamma$ -family .*

*O*-succinylhomoserine (thiol)-lyase (cystathionine- $\gamma$ -synthase)

*O*-acetylhomoserine (thiol)-lyase (methionine synthase)

Cystathionine- $\gamma$ -lyase ( $\gamma$ -cystathionase)

Cystathionine- $\beta$ -lyase ( $\beta$ -cystathionase)

#### 1.2.2.4 Other PLP Enzymes

A few of the PLP-dependent enzymes studied by Alexander *et al.* proved to be unrelated to any of the three family groups. These are alanine racemase, selenocysteine synthase and many amino acid decarboxylases which are not in group II.<sup>24</sup>

#### 1.2.2.5 Exceptions to the "Family Rule"

As already mentioned a few PLP-dependent enzymes do not belong to the family that would be suggested by their regioselectivity. Tryptophanase, tyrosine phenol lyase and 4-amino-4-deoxychorismate synthase are all members of the  $\alpha$ -family by virtue of their sequences, yet they catalyse  $\beta$ -elimination reactions.<sup>24</sup> 1-aminocyclopropane-1-carboxylate synthase, also an  $\alpha$ -family enzyme, catalyses an  $\alpha,\gamma$ -replacement reaction. Similarly, TS, which is a member of the  $\beta$ -family, catalyses a  $\beta,\gamma$ -replacement and cystathionine  $\beta$ -lyase, supposedly a  $\gamma$ -family enzyme, catalyses a  $\beta$ -elimination reaction.<sup>24</sup>



## 1.3 The Structure and Mechanism of PLP Enzymes

### 1.3.1 A General Mechanism of PLP-Enzyme Action

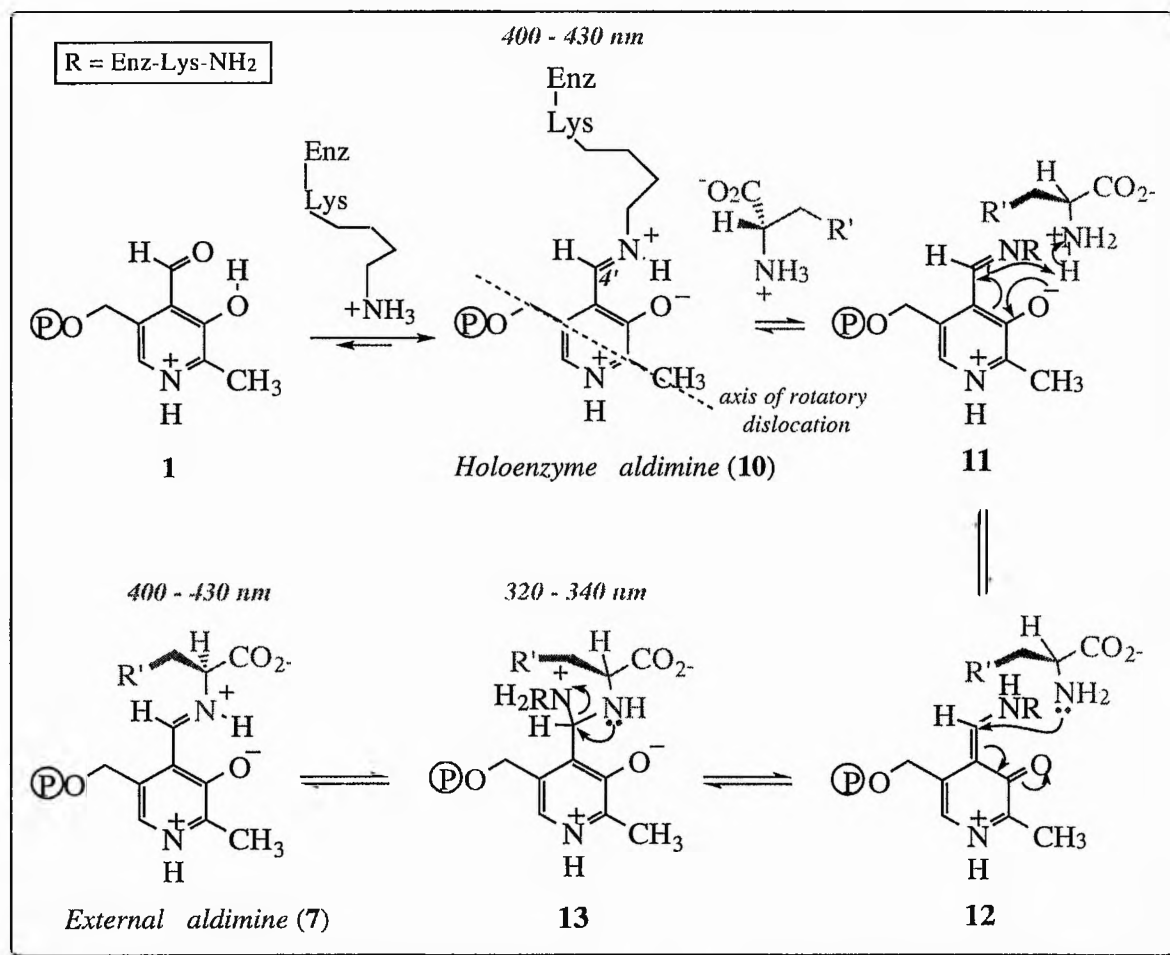
Much of the information concerning the general mechanism of PLP-mediated amino acid reactions and the structure of PLP enzymes has been gleaned from studies of the PLP-dependent transaminases. In particular the action of aspartate aminotransferase (AspAT) has received much attention. Between 1954 and 1955 two independent workers, Braunstein and Snell, developed a general theory for the mechanism of PLP enzymes.<sup>39, 40</sup> The following section gives a detailed discussion of this mechanism, and examines the experimental data that was used to prove it.

For all known PLP-enzymes the first reaction step involves the formation of the Michaelis complex (**10**) otherwise known as the holoenzyme aldimine or internal aldimine between the  $\epsilon$ -amino group of a specific enzyme-bound *lysine* residue and the C-4' carbonyl group of PLP (**1**, Scheme 1.5).<sup>4, 5</sup>

Upon formation of the holoenzyme the amino acid substrate can bind to give the Schiff's base substrate aldimine (or external aldimine, **7**).<sup>41</sup> This involves the condensation of the amino acid substrate with the internal aldimine (a process known as transaldimination<sup>4</sup> or transimination<sup>5</sup>) with the simultaneous release of the enzyme-bound lysine residue (Scheme 1.5).<sup>42</sup> This occurs *via* the formation of a tetrahedral transition complex. Formation of the external aldimine was originally observed for AspAT by a change to an absorption maximum at 430 nm in the UV-visible spectrum and a decrease in the positive CD value.<sup>43-45</sup>

The structure of this external aldimine at the active-site has not been rigorously established, but a *trans*-configuration at the imine double bond (C-4'-N) is most likely on steric grounds<sup>5</sup> as unfavourable interactions would occur if the 4'-N substituent were in the *cis*-coplanar conformation.<sup>46</sup> This structure would also allow hydrogen bonding between the imine

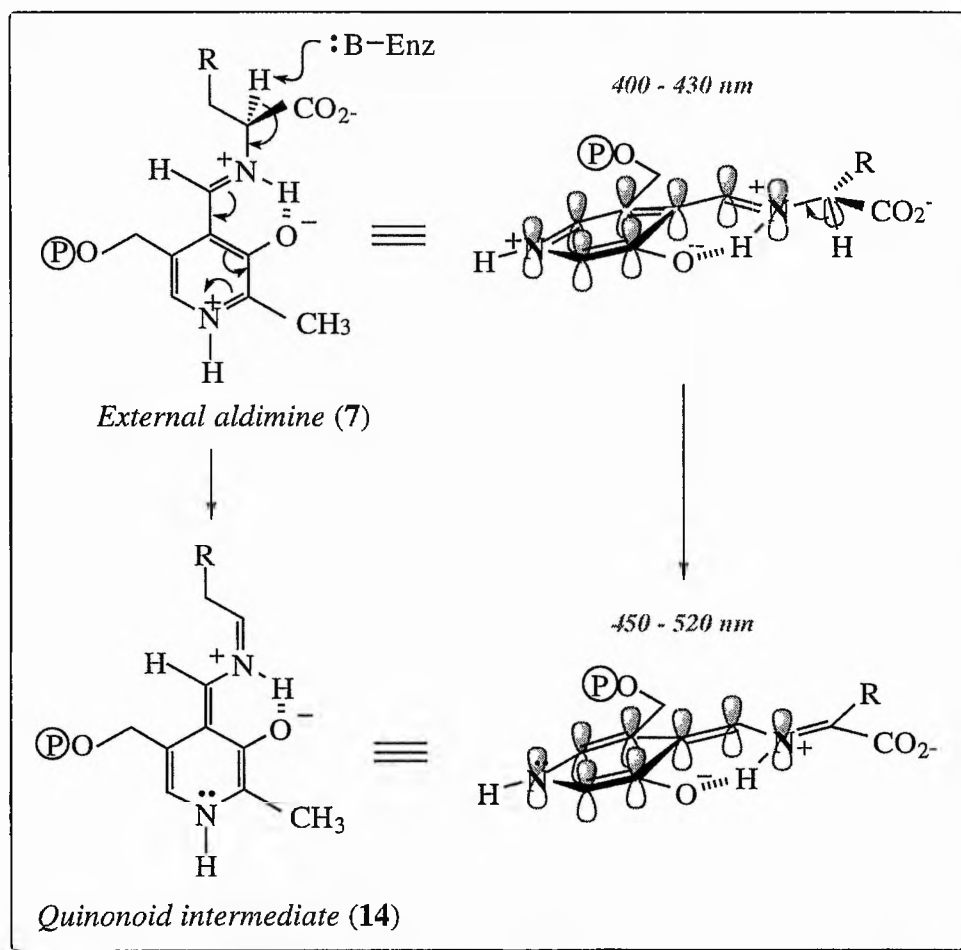
nitrogen and the hydroxyl group on the pyridine ring,<sup>5</sup> helping to hold the imine double bond in the plane of the pyridinium ring. Braunstein and Snell first postulated the formation of the external aldimine as a solution to the mechanistic problem of C<sup>α</sup>-bond cleavage leading to an intermediate carbanion which is too basic to be kinetically competent.<sup>39, 40</sup> They suggested that the developing negative charge is stabilised by the delocalisation of electrons mesomerically with the positively charged pyridinium nitrogen atom.<sup>47</sup>



**Scheme 1.5:** Condensation between lysine residue and PLP (1) at the enzyme active-site followed by transaldimination.  $\lambda_{max}$  absorbance values are given for various intermediates.

Dunathan's postulate states that, in the external aldimine, rotation about the C<sup>α</sup>-N bond is controlled by the enzyme, so that the bond to be cleaved at C<sup>α</sup> is held orthogonal to the plane of the pyridinium ring, thus allowing maximum orbital overlap between the developing

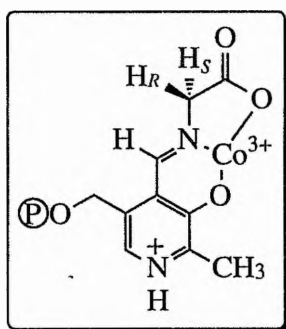
negative charge and the conjugated electron-deficient  $\pi$ -system (Scheme 1.6).<sup>4, 5</sup> The  $C^\alpha$ - $CO_2H$  bond is broken by the  $\alpha$ -decarboxylases whereas the amino acid transaminases and racemases cleave the  $C^\alpha$ -H bond.<sup>4</sup> These enzymes will be discussed further in Section 1.4. The formation of the conjugated enamine (or quinonoid intermediate, 14) in AspATs is accompanied by a characteristic absorption band at 502 nm and a sharp negative CD.<sup>43-45</sup>



**Scheme 1.6:** Preferential cleavage of the bond orthogonal to the plane of the pyridinium ring as predicted by the Dunathan hypothesis. This example shows  $C^\alpha$ -H cleavage but  $C^\alpha$ - $CO_2$  and  $C^\alpha$ -R bond cleavage can also occur. The diagram also shows the hydrogen bond between the PLP hydroxyl group and the nitrogen on the substrate, which helps to hold the imine bond in the plane of the pyridinium ring.

Quantum mechanical calculations have indeed implied that the strength of a  $\sigma$ -bond between  $C^\alpha$  of the Schiff's base and the electrofuge is at its weakest when the bond lies perpendicular to the pyridinium ring.<sup>48</sup> Floss and Vederas suggest that, if an enzyme were to bind the relatively rigid PLP-cofactor at the pyridine nitrogen and at the phosphate, attachment of a single distal group on the substrate, which in most cases would probably be the carboxyl group, would result in a three-point binding of the complex fixing a particular conformation of the  $C^\alpha$ -N bond.<sup>5</sup> There is little actual evidence to support this suggestion however plausible it may seem.

Abbot has provided experimental support for Dunathan's hypothesis by demonstrating that in a complex between cobalt (III) and PLP-glycine Schiff's base the two  $C^\alpha$ -hydrogens are held in distinct conformations relative to the  $\pi$ -system and apparently exchange at different rates, as measured by  $^1\text{H-NMR}$  spectroscopy (see Fig. 1.5).<sup>49</sup>



**Figure 1.5:** Complex between cobalt (III) and PLP-glycine Schiff's base

Karpeisky and Ivanov proposed a model for the mechanism of enzyme-catalysed transamination, suggesting that the transaldimination step requires reorientation of the coenzyme by rotatory dislocation round an axis passing through its 2-methyl and 5'-phosphate groups (indicated on structure **10**, Scheme 1.5).<sup>50, 51</sup> The reversal to the original orientation of the coenzyme ring would be linked with the transition from the PMP ketimine to the free PMP-enzyme and keto acid.

A large number of reaction pathways are available to a PLP-amino acid, and it is the enzyme present which determines which of the C $\alpha$  bonds are broken by orienting the C $\alpha$ -R bond and which determines the reaction specificity by the positioning of specific binding sites and catalytically active residues. Snell and Metzler postulated that the catalytic abilities of PLP-dependent systems are, to an exaggerated degree, those of the coenzyme and that enzymic and non-enzymic reactions proceed by similar mechanisms.<sup>40</sup> It has also been postulated that the reactions of PLP-enzymes all take place only on one face of the planar PLP-substrate complex (the 'exposed' or 'solvent' face) the other being covered by the protein. It has been noted that the 'exposed' face is always the *si*-face at C-4' of the cofactor (below the plane of the page in all structures shown).<sup>5</sup>

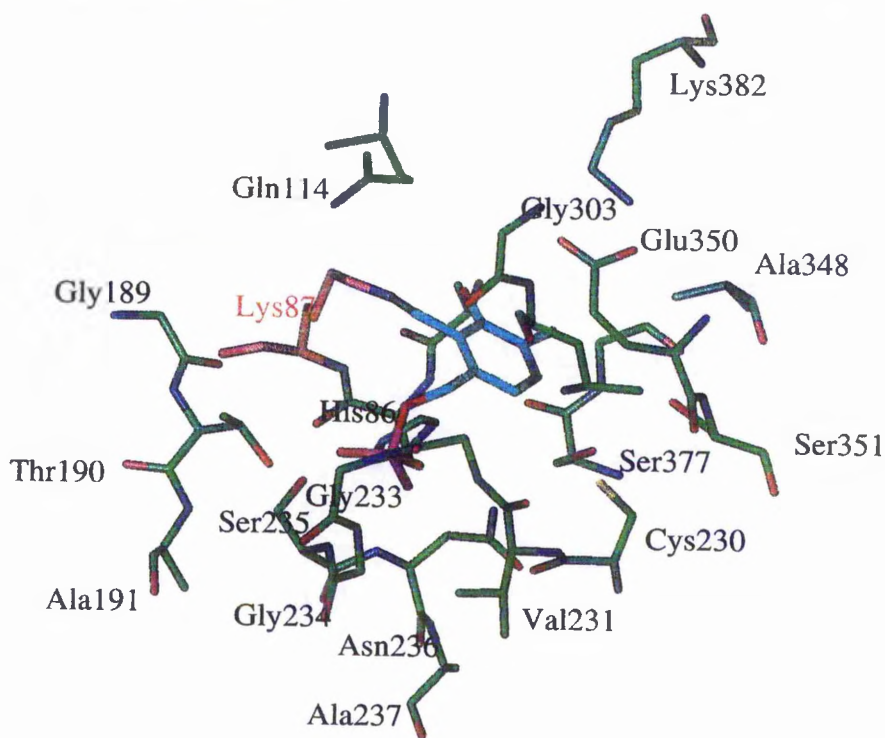
### 1.3.2 Recent Structural and Mechanistic Studies on $\beta$ -Family Enzymes

#### 1.3.2.1 Recently Published Crystallographic Data on PLP Enzymes

The first X-Ray structures of cytosolic chicken heart AspAT became available in 1975.<sup>52</sup> The most recent studies on AspATs include that on a R292D mutant complexed with PLP and sulfate at 2.8Å resolution in 1993, and a K259H mutant complexed with PLP at 2.1Å resolution in 1995.<sup>53, 54</sup> Arg292 is responsible for substrate charge specificity in the wild-type enzyme and Lys259 is the residue which forms a Schiff base with PLP in the active-site. The R292D mutant preferred arginine to aspartate as a substrate, but had a lower catalytic efficiency than the wild-type enzyme. This was attributed to small side-chain and main chain reorientations. The catalytic competence of the K259H mutant in the forward half-reaction was only 0.1% that of the wild-type enzyme. The most recently obtained structure was produced in 1997 of wild-type pig *c*AspAT refined to 1.6Å resolution.<sup>55</sup> However, for most of the  $\alpha$ - and  $\gamma$ -family enzymes there are few representative 3° structures available.

Of the  $\beta$ -family enzymes only crystal structures for TrS are available (Fig 1.6). Complexes of wild-type TrS with PLP in the presence of Cs<sup>+</sup> and K<sup>+</sup> ions have been resolved to 2.3 and

Wild-type TrS complexed with PLP



$\beta$ K87T TrS complexed with PLP and (2S)-serine

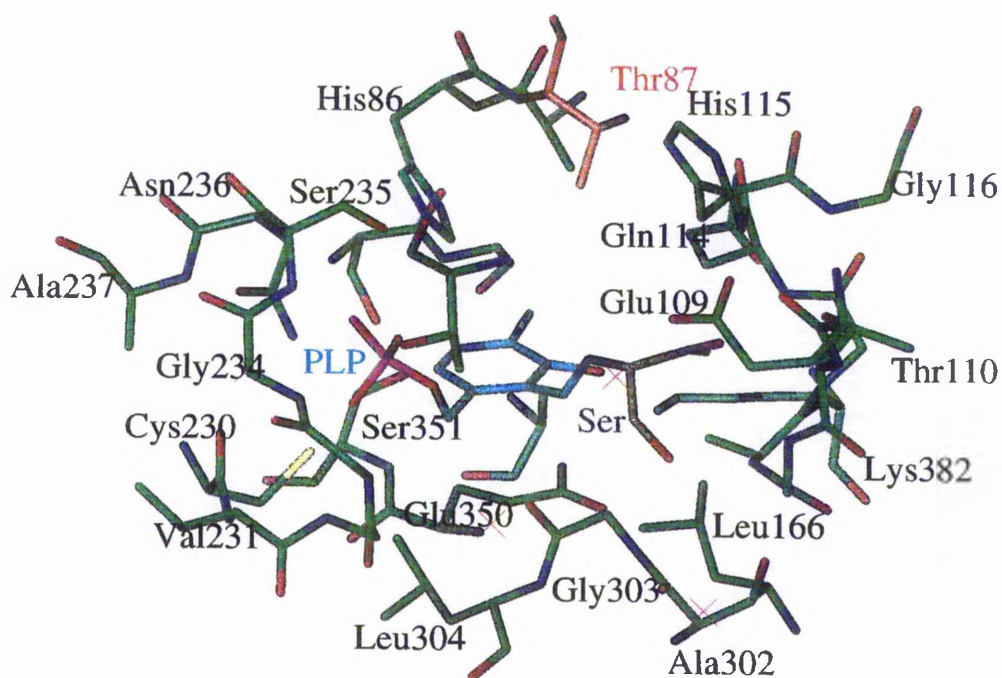


Figure 1.6: Active-site structures of the  $\beta$ -subunit of wild-type and  $\beta$ K87T TrS.<sup>56, 57, 58, 59</sup>

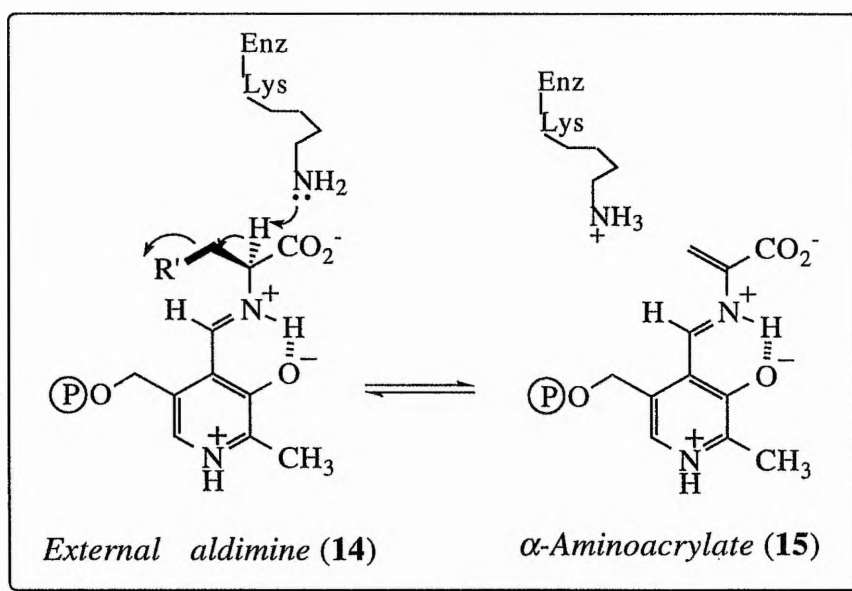
2.0Å respectively.<sup>60</sup> Substitution of these monovalent cations for Na<sup>+</sup> induce local and long-range changes in the 3D structure of the enzyme. A crystal structure of the wild-type holoenzyme is also available without cations at 2.5Å resolution (Fig. 1.6, top).<sup>56</sup> A mutant form of TrS ( $\beta$ K87T), lacking the PLP-binding lysine residue in the  $\beta$ -subunit, has been crystallised recently in two forms, complexed with PLP and the natural substrate, (2*S*)-serine (Fig. 1.6, bottom), or with PLP and the product, (2*S*)-tryptophan.<sup>57</sup> These structures were refined to 1.9Å resolution.

### 1.3.2.2 Structure-Activity Studies on *O*-Acetylserine Sulfhydrylase

In 1994 McClure and Cook carried out fluorescence studies on OASS to attempt to determine the effect of product and substrate binding on the structure of the enzyme.<sup>61</sup> Previously the emissions of both enzymic and synthetic PLP Schiff bases in both aqueous and non-aqueous buffers had been studied and these results were used for comparison. An emission maximum was observed at 497 nm with OASS, consistent with the presence of a PLP-Schiff base in an aqueous environment. In addition, emissions attributed to the presence of tryptophan residues in the enzyme were considerably quenched by the formation of the holoenzyme, consistent with a model in which energy transfer occurs between at least one Trp residue and the PLP Schiff base during transaldimination. A similar phenomenon had been observed with TrS and this was attributed either to singlet-singlet or triplet-singlet energy transfer from an excited Trp residue to the PLP Schiff base in the holoenzyme.<sup>62, 63</sup>

Other results with OASS indicated that product binding does not affect the equilibrium between tautomers of PLP that might be present in the enzyme. An enhanced fluorescence observed upon the binding of (2*S*)-cysteine to OASS was thought to indicate a conformational change that occurs upon transaldimination to increase the efficiency of energy transfer from a Trp residue. The presence of a Trp residue, namely Trp163, in the active-site was therefore postulated. More recent studies on Trp luminescence have produced further evidence that such a conformational change takes place.<sup>64</sup> Addition of *O*-acetylserine (OAS) to OASS led to

a splitting of the 0,0 vibronic band in the phosphorescence spectrum of Trp163 yielding poorly resolved peaks at 406 and 408.5 nm. This indicated that the polarity of the environment of the Trp residue had changed as the internal Schiff base was converted to  $\alpha$ -aminoacrylate (Scheme 1.7). The distance between Trp163 and PLP was estimated to be about 25 Å which is comparable to the distance between Trp177 and PLP in TrS.<sup>64</sup>



**Scheme 1.7:** Formation of the  $\alpha$ -aminoacrylate intermediate. R' = OAc for OASS or OH for TrS.

Studies by Cook *et al.* confirmed that transaldimination produces a conformational change in the enzyme, but their data indicated that this is reversed by the formation of the  $\alpha$ -aminoacrylate intermediate **15**.<sup>65</sup> They studied the characteristics of OASS with various substrate analogues. With (2*S*)-serine and (2*S*)-cysteine the enzyme was trapped in the external aldimine form, whereas with the natural substrate, OAS, the  $\alpha$ -aminoacrylate **15** was formed. UV-CD spectra recorded for OASS in the presence or absence of the three amino acids revealed a significant increase in the  $\alpha$ -helical content of the enzyme in the external Schiff base as compared to the internal aldimine. This was interpreted as the closing of the active-site into a catalytic conformation. No difference was observed, however, between the 2° structures of the internal aldimine and species **15**, evidence which was later refuted by investigations on



the phosphorescence of Trp163 (see earlier discussion).<sup>64</sup> The formation of the external aldimine from the holoenzyme aldimine was observed by a shift in the  $\lambda_{\max}$  value from 412 nm, for OASS and PLP, to 418 nm. The absorption band at 412 nm indicates that the Schiff base between the enzyme lysine and PLP is protonated. Cook *et al.* also suggest that the pyridinium nitrogen is protonated in the internal aldimine.<sup>66</sup> The external aldimine formed with (2*S*)-serine also produced a second absorbance band at 320 nm suggestive of the establishment of an equilibrium between different tautomeric forms of the Schiff's base. In the presence of OAS, absorbance at 481 nm increased, concomitant with a decrease in the absorbance at 412 nm, indicating *concerted* formation of **15**. No quinonoid intermediate was detected for OASS, whereas for TrS a quinonoidal species has been detected as a fleeting transient, before an equilibrium is established between the  $\alpha$ -aminoacrylate and the external Schiff base.<sup>67</sup>

The <sup>31</sup>P NMR spectra for the reactions of OASS with (2*S*)-serine and (2*S*)-cysteine were also evaluated.<sup>65</sup> The internal aldimine has a relatively high  $\delta_p$  value, indicative of a strong salt bridge formed between the 5'-phosphate group in the dianionic form and an enzyme residue, which is excluded from solvent. Similar findings for TrS and D-serine dehydratase (DSDHT) were also cited.<sup>68</sup> The line widths of the <sup>31</sup>P NMR signals for the external aldimines formed were considerably increased, indicating the presence of different tautomeric forms of the Schiff bases.

Initial velocity studies on the basis of pH indicated that in the internal aldimine an enzyme residue with a pK of 6.7 - 7 must be protonated for maximum activity.<sup>66</sup> This residue is thought to hydrogen bond to the *O*-acetyl group on the substrate to hold it in the correct conformation for  $\beta$ -elimination to occur. A similar mechanism may operate for TS to hold the  $\gamma$ -phosphate group of *O*-PHse orthogonal to the PLP ring prior to elimination. Additional data suggested that the enzyme lysine (Lys42) that forms the internal Schiff base with PLP also acts as a general base to accept a proton from the  $\alpha$ -carbon of the external Schiff base. A K42A mutant was also found to be incapable of forming **15**. The protonated lysine is thought to donate a proton back to the  $\alpha$ -carbon of **15** to form the product external Schiff base.

At the start of the second half-reaction, in which **15** undergoes nucleophilic attack by HS<sup>-</sup>, the lysine residue is protonated and the enzyme group with a pK of 6.7 - 7 is unprotonated.<sup>66</sup> The latter hydrogen-bonds to the incoming nucleophile. A similar mechanism may operate in other  $\beta$ -family enzymes. Nucleophilic attack at C-3 of **15** is accompanied by protonation of the  $\alpha$ -carbon. The unprotonated lysine residue then attacks C-4' of the product external aldimine to regenerate the internal aldimine.

More recent studies on OASS have involved stopped-flow and single wavelength absorbance stopped-flow experiments and kinetic isotope studies.<sup>67, 69</sup> These implied that the first half-reaction (formation of **15**) was limiting to the overall reaction. The external Schiff base formed rapidly, but its conversion to **15** was slower and quasi-irreversible.

### 1.3.2.3 : Structure-Activity Studies on Tryptophan Synthase

TrS exists as an  $\alpha_2\beta_2$  tetramer of which the  $\alpha$ -subunits catalyse the formation of indole and the  $\beta$ -subunits utilise PLP in catalysing the formation of (2*S*)-tryptophan. Various studies on TrS have attempted to elucidate the functions of a number of amino acid residues by using mutated enzymes. In one such study the role of  $\beta$ Glu350 was investigated.<sup>70</sup> Crystal structures had already revealed that this residue is located near the protonated pyridinium nitrogen of PLP and stacked parallel to the plane of the coenzyme in a geometry that might not favour strong interaction (Fig. 1.6 and Appendix 3). Comparisons of the wild-type,  $\beta$ E350A and  $\beta$ E350Q mutants indicated that although E350 is not essential for catalysis it may serve a role in stabilising an optimal conformation of the active-site. Mutation of this residue did not induce *gross* conformational change, however, as all the enzymes had many similarities in their spectroscopic properties and substrate specificities. A reduced intensity of the 412 nm peak for the internal Schiff base did suggest, that the orientation of PLP was altered in the mutants. The mutant enzymes also appeared incapable of  $\alpha$ -proton removal, or of catalysing  $\beta$ -elimination to form **15**. This may be attributed to an alteration of the positions of catalytic groups in the active-site relative to the  $\beta$ -hydroxyl group of (2*S*)-serine. Later studies on mutants  $\beta$ E109A,  $\beta$ E306A and  $\beta$ D305A showed similar alterations in their reaction specificities, indicating that

the active-site conformation of the wild-type enzyme is readily altered and that these changes affect the selectivity of the enzyme.<sup>70</sup>

A  $\beta$ K167T TrS mutant was observed to have a drastically reduced specific activity in its  $\beta$ -reaction (4% of wild-type) although the  $\alpha$ -reaction was unaffected.<sup>71</sup> The substrate and reaction specificities were also altered from the wild-type enzyme. The K167T  $\alpha_2\beta_2$  complex, like the wild-type  $\beta$ -subunit, was more active with  $\beta$ -chloro-(2*S*)-alanine than with (2*S*)-serine and had similar activities in the  $\beta$ -replacement and  $\beta$ -elimination reactions. This differed from the wild-type  $\alpha_2\beta_2$  complex. The spectroscopic properties of the K167T  $\alpha_2\beta_2$  complex also differed from that of the wild-type tetramer. In the presence of the substrate, (2*S*)-serine, the major intermediate for the mutant enzyme had a dominant absorbance at 420 nm, ascribed to the external aldimine. The *wild-type*  $\alpha_2\beta_2$  complex produced a major absorbance at 340 nm with the substrate, ascribed to **15**. Interestingly, upon binding of  $\alpha$ -subunit ligands, the properties of the mutant enzyme became closely similar to those of the wild-type enzyme. These results were seen to imply that the interaction of  $\beta$ K167 with  $\alpha$ D56 is important for promoting intersubunit communications in the tetramer.

Luminescence studies of TrS revealed a number of details concerning the allosteric activation of the  $\alpha_2\beta_2$  tetramer.<sup>63</sup> The binding of  $\alpha$ -subunit ligands caused a reduction in the thermal quenching of the fluorescence of Trp177 and slower triplet decay kinetics. (See earlier discussion on Trp fluorescence in OASS). From this it was inferred that there was a considerable tightening of the active-site region and N-domain. However, the constancy of Trp177 fluorescence and phosphorescence in the various complexes of the native enzyme suggested that the structural changes observed in phosphorescence do not involve a different distance or orientation of the PLP moiety relative to Trp177. This would indicate that the PLP ring does not tilt during transaldimination in TrS. Reorientation of the coenzyme has been observed, however, during the transaldimination reactions of  $\alpha$ -family enzymes.<sup>42</sup> In addition the internal Schiff base at K87 of the  $\beta$ -subunit was clearly demonstrated to exist in two tautomeric forms, that of a ketoenamine and an enolimine fluorescing at around 490-510 and 361-380 nm respectively.

Although studies have been carried out on the roles of the PLP-binding lysine residue of various  $\alpha$ -family enzymes, it was not until 1993 that the altered properties of a  $\beta$ K87T mutant TrS were investigated by E. W. Miles *et al.*<sup>72</sup> Spectroscopic and kinetic studies revealed that the mutant enzyme formed external aldimines much more slowly than the wild-type and only carried out two slow partial  $\beta$ -elimination reactions: the conversion of  $\beta$ -chloro-(2*S*)-alanine and (2*S*)-serine to **15**. The latter reaction was catalysed by ammonia, which partially replaced the deleted  $\epsilon$ -amino group. These results highlighted the role of K87 in facilitating the formation of external aldimines by transaldimination and its importance for substrate and product release, because (2*S*)-serine, (2*S*)-tryptophan and aminoacrylate dissociated very slowly from the mutant  $\alpha_2\beta_2$  complex. The residue does not play an essential role in cofactor binding, which is in agreement with results observed for other PLP enzymes.

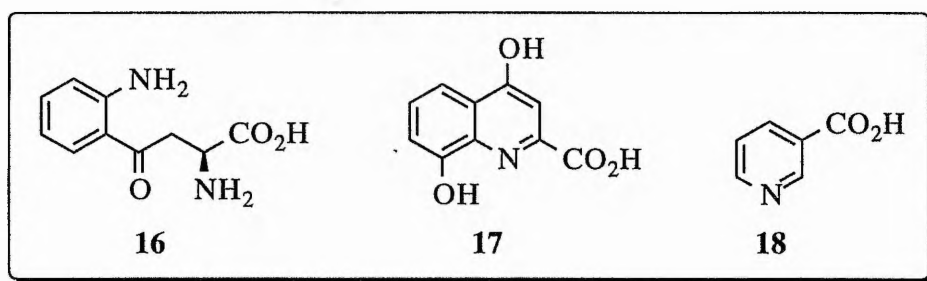
One of the most recent studies on the  $\beta$ -subunit of TrS involved the construction of 4 C-terminal deletion proteins and 11 point mutant proteins.<sup>73</sup> Multiple sequence alignments had revealed that residues 384 - 397 are variable, whereas residues 378 - 383 are conserved (see Appendix 3 for the positioning of these residues relative to PLP). Residues 383 - 393 form helix 13. Deletion of the variable region ( $\Delta$ 384 - 397) resulted in 80% inactivation of the enzyme, whereas the additional deletion of one or two conserved residues ( $\Delta$ 383 - 397 or  $\Delta$ 382 - 397) resulted in greater inactivation. Point mutants at the conserved residues R379, D381 or K382 exhibit significant activities (2 - 40%). The authors concluded that the C-terminal conserved and variable regions and helix 13 play no essential functional roles, but may serve structural roles.

## 1.4 The Metabolic Roles of Pyridoxal 5'-Phosphate Dependent Enzymes

In Section 1.1 introducing PLP-enzymes some of their metabolic functions were alluded to. This section describes the roles of various important PLP-dependent enzymes and some reference is also made to the biochemical effects of pyridoxine deficiency to further illustrate its physiological function.

### 1.4.1 Niacin (Nicotinic Acid) Synthesis From Tryptophan

The *tryptophan-kynurenine-nicotinic acid* metabolic pathway in most animal species is under the control of four B<sub>6</sub>-dependent enzymes. If a deficiency of pyridoxine (B<sub>6</sub>) occurs then kynurenine (**16**) is diverted away from its further conversion to niacin to become involved instead in the formation of xanthurenic acid. Excretion levels of this acid are therefore used as an indicator of the levels of B<sub>6</sub> deficiency.<sup>74</sup>



*Kynurenine (16), xanthurenic acid (17) and niacin (18).*

Niacin is an essential component of the two cofactors, nicotinamide adenine dinucleotide (NAD<sup>+</sup>) and nicotinamide adenine dinucleotide phosphate (NADP<sup>+</sup>), which are involved in a number of processes: tissue respiration, glycolysis, the hydrogen transport series, the formation of high energy phosphate bonds, pyruvate metabolism, pentose biosynthesis, lipid metabolism, amino acid and protein metabolism and photosynthesis. In fact more than forty

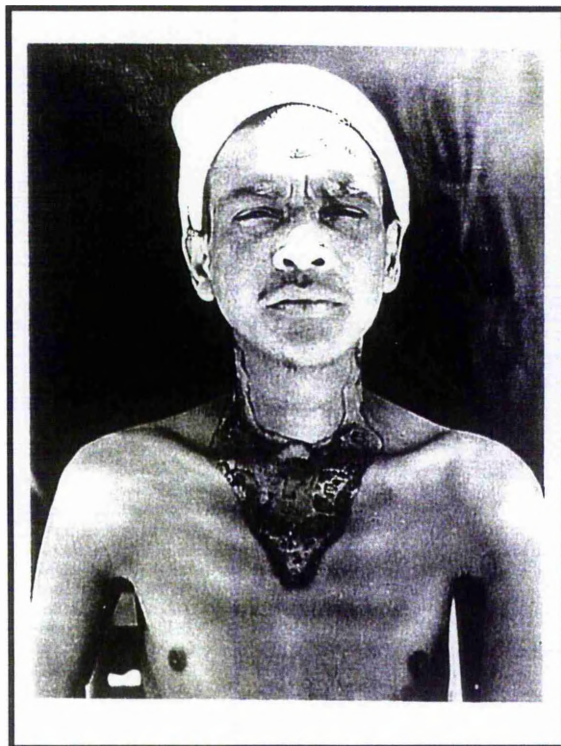
reactions are known to be dependent on these two cofactors. Niacin is therefore required by all living cells.

#### 1.4.1.1 Pellagra and Related Diseases

In man, if niacin is not obtained through the diet, this can result in a condition called *pellagra*. The name of this disease is derived from two Italian words: "*pelle*" meaning skin and "*agra*" meaning rough or unsightly.<sup>75</sup> Although pyridoxine deficiency rarely leads to the disease in adults, in infants the lack of pyridoxine is much more likely to be a factor in niacin deficiency.

In the 1940's a liquid baby formula in which a considerable portion of vitamin B<sub>6</sub> had been destroyed by heat sterilisation produced convulsions in infants.<sup>76</sup> Prompt administration of the vitamin to affected babies produced quick recovery. Similar symptoms were observed by Hawkins and Barsky when they attempted to produce experimental pyridoxine deficiency in human subjects, although one child developed hyperchromic anaemia instead.<sup>77</sup>

Pellagra has often been referred to as the disease of the four 'D's': dermatitis, diarrhoea, dementia and death.<sup>75</sup> Early signs of the disease include lassitude, anorexia, weakness, digestive disturbances (such as heartburn and nausea) and various psychic and emotional changes such as anxiety, depression and irritability.<sup>78</sup> One main indicator for the disease is an affection of the skin resembling severe sunburn, which is usually symmetrically disposed on those parts of the body which have been exposed to the sun (Fig. 1.7).<sup>79</sup> (In cases of chronic pellagra skin changes include thickening, scaling, hyperkeratinization and pigmentation). The mouth can also become sore, especially the tongue, and the gastrointestinal tract inflamed, leading to diarrhoea.<sup>79</sup> If the disease progresses to an advanced state the patient may experience delirium, hallucinations, disorientation, confusion and stupor.<sup>80</sup> Fortunately treatment of the disease with the administration of a B vitamin complex is swift to reverse the effects of the disease, with recovery sometimes produced within 24 hours.



**Figure 1.7:** *Casal's necklace - a common symptom of pellagra.*

As described in Section 1.1 rats fed a diet similar to that which caused pellagra in humans developed a dermatitis similar to a human disease known as acrodynia.<sup>75</sup> There was reddening and scaliness of the snout, the paws and the tail. It was because riboflavin (vitamin B<sub>2</sub>) and niacin (nicotinic acid) failed to cure the condition that B<sub>6</sub> was first identified as a new vitamin, which restored the rats to normal health.<sup>7</sup> The lack of pyridoxine also causes rats to lose hunger, thus aggravating any inherent problems of nutritional deficiency, with ensuing weight loss and a decrease in fat storage. There was also oedema, growth retardation, anaemia, proteinuria, nerve degeneration and abnormal reproductive performances in rats maintained on the pyridoxine-deficient diet.<sup>80</sup>

Various other animal species have also been tested for their response to pyridoxine deficiency in the diet. The vitamin is essential for the growth of chicks, dogs, pigs, mice and infant humans as well as for a number of microorganisms.<sup>81</sup> In dogs, as in several other animals, microcytic anaemia results from a lack of B<sub>6</sub>.<sup>82</sup> In monkeys Rinehart and Greenberg

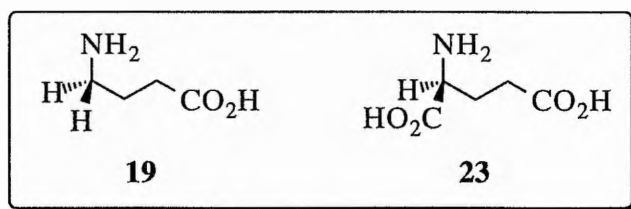
demonstrated that a lack of B<sub>6</sub> caused atherosclerosis and hypercholesterolemia.<sup>83</sup> This was attributed to the role of PLP in essential fatty acid synthesis. In chicks there was poor growth and feathering, inflammation of the mouth and diarrhoea as a result of B<sub>6</sub> deficiency, whereas rabbits suffered from subnormal weight gain, leukopenia and terminal diarrhoea.<sup>78</sup> Pyridoxine deficiency has also been implicated in kidney stone formation due to the accumulation of calcium oxalate caused by a lack of transaminase activity necessary for the conversion of oxalate to glycine. A loss of conditional reflex performance and some impairment of the immune response in various B<sub>6</sub>-deficient animals has also been found.<sup>84</sup>

#### 1.4.2 Decarboxylases.

PLP-dependent decarboxylases control the formation of a number of important amines. These include GABA (19), putrescine (20), histamine (21) and dopamine (22)

##### 1.4.2.1 $\gamma$ -Aminobutyric Acid Transaminase (GABA-T)

Glycine and  $\gamma$ -amino butyrate (GABA, 19) increase the permeability of postsynaptic membranes towards chloride ions.<sup>74</sup> This results in an improved chloride conductance causing membrane hyper polarisation. It is this effect which increases the threshold for the triggering of an action potential and thus effectively inhibits neurotransmission in the central nervous system (CNS). In a therapeutic capacity high cerebral concentrations of GABA have been shown to prevent convulsions by depressing the firing of neurons throughout the CNS.



*$\gamma$ -aminobutyrate (19) and glutamate (23).*

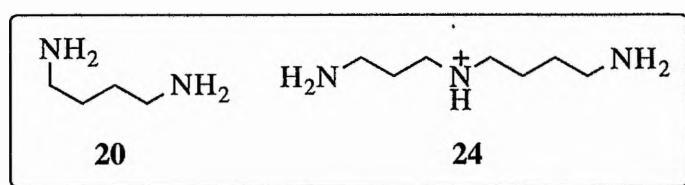


GABA is formed *via* the decarboxylation of glutamate (**23**), a reaction which is catalysed by glutamate decarboxylase in the mammalian brain. Both the decarboxylase and  $\gamma$ -amino butyric acid transaminase (GABA-T) are PLP-dependent enzymes involved in the regulation of the GABA-ergic system. The inhibition of GABA-T, the first enzyme on this catabolic pathway, may be important in anti-convulsive therapy. Several factors indicated the role of GABA in causing convulsions: its abundance in the brain, its ability to produce hyperpolarizing inhibition of almost all neurons, its association with benzodiazepines and the discovery that many convulsants inhibited its synthesis.<sup>85</sup> One anti-epileptic drug,  $\gamma$ -vinyl GABA (GVG), also known as vigabatrin, is a selective irreversible suicide inhibitor of GABA-T and increases brain and cerebrospinal fluid GABA levels, without a deleterious effect on cognitive function.<sup>86, 87</sup> This drug has been widely tested for side-effects and tolerance development, with encouraging results and is the only GABA-T inhibitor in clinical use.<sup>86-94</sup> Patients suffering from chronic epilepsy when treated with this drug maintained a 54.7% reduction of seizure frequency compared with an 18.6% *increase* for those on placebo.<sup>86</sup> GVG and ethanolamine *O*-sulfate have been shown to reduce rat brain GABA-T activity *in vitro* by 65 - 79% and increase GABA content by 40 - 99% compared with control values.<sup>95</sup> Other GABA-T inhibitors include gabaculine,  $\gamma$ -acetylenic GABA and aminooxyacetic acid (AOAA).<sup>91, 96</sup> AOAA has been shown to dose-dependently increase spinal GABA content and markedly suppress procaine-induced convulsions in rat spinal cord.<sup>97</sup> Pyridoxine has also been used as a treatment of epilepsy in infants.<sup>98, 99</sup> A treatment for migraine prophylaxis, valproic acid, is believed to act by inhibiting GABA-T and the antidepressant phenelzine also reduced GABA-T activity in rat brains.<sup>100, 101</sup>

#### 1.4.2.2 Ornithine Decarboxylase (ODC)

Ornithine decarboxylase (ODC) is the first, rate limiting enzyme in the biosynthetic pathway to putrescine (**20**) and higher polyamines.<sup>102</sup> Both spermine and spermidine (**24**) are synthesised from putrescine. These two positively-charged polyamines help to counteract the charge repulsion effects of negatively charged phosphate groups on DNA, thus ensuring that

chromosomal DNA is sufficiently compacted to fit inside a cell.



*Putrescine (20) and spermidine (24).*

The activity of ODC *in vivo* increases dramatically in response to cellular stimulation which promotes regeneration and replication. Cellular levels of the enzyme are therefore high during growth and protein synthesis and very low in growth arrested cells.<sup>103</sup> ODC activity has been shown to increase markedly in cells undergoing *apoptosis* followed by a paradoxical decrease of the intracellular content of polyamines.<sup>104</sup> Liver sepsis and traumatic brain injury in rats is also accompanied by increased ODC activity.<sup>105-107</sup>

The association of high polyamine levels with rapid cellular proliferation and protein biosynthesis led to the idea that polyamines may be required for RNA and DNA biosynthesis.<sup>108</sup> For this reason the enzyme has been identified as a target for cancer chemotherapy.<sup>4</sup> Many tumour promoters have been shown to increase ODC activity and a recent study confirms the potential role of ODC in causing tumor growth.<sup>109-111</sup> Constitutive expression of epidermal ODC in transgenic mice led to several abnormalities, including the development of dermal follicular cysts, excessive skin wrinkling, enhanced nail growth, alopecia and spontaneous tumor development.<sup>111</sup> The tumors themselves were analysed for ODC activity and the absolute values of ODC-specific activity were extremely high in all tumors. ODC activity has also been reported to be high in human neoplastic lesions from breast, stomach and colon.<sup>112</sup> Tamori and co-workers compared the nucleotide sequence of ODC cDNA with that of non-tumorous tissue in the same patients and discovered three point mutations in hepatoma ODC cDNA, which led to a change in two amino acids and the replacement of codon 415 (CAA) with TAA, a stop codon.<sup>112</sup> The truncated ODC from

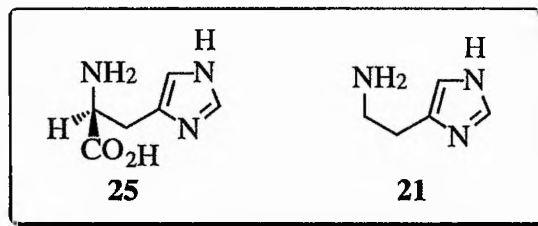
hepatoma tissue is stabilised, thus increasing its half-life compared to that of wild-type ODC. Studies have indicated that a high level of tumor ODC expression predicts a favourable prognosis in human colorectal carcinoma but not in breast cancer.<sup>113, 114</sup>

Many highly specific and potent inhibitors of ODC are based on the lead compound  $\alpha$ -difluoromethylornithine, an enzyme-activated irreversible inhibitor.<sup>109, 115, 116</sup>  $\beta$ -carotene and DL- $\alpha$ -tocopherylacetate also effect a reduction in ODC activity in human atrophic stomach mucosa and intestinal metaplasia, by 50% and 18% respectively.<sup>117</sup> Unfortunately  $\beta$ -carotene appears to increase the incidence of some types of cancer.<sup>118</sup> Some green tea polyphenols inhibit ODC induction dose-dependently and crude extracts of lowbush blueberry, cranberry and lingonberry are active inhibitors of ODC activity *in vitro*.<sup>119, 120</sup> Wilms' tumor suppressor has been shown to function by binding to multiple sites within the human ODC promoter.<sup>121</sup>

#### 1.4.2.3 Histidine Decarboxylase (HDC)

Histidine decarboxylase (HDC) catalyses the decarboxylation of histidine (**25**) to give histamine (**21**). In mammals histamine is important as a major receptor agonist and regulatory factor for peripheral blood circulation.<sup>122</sup> The overproduction of histamine is associated with many biological responses such as gastric secretion and various allergic and hypersensitivity reactions.<sup>123</sup> The development of antihistamines; antagonists for H<sub>1</sub> and H<sub>2</sub> receptors, is an important area in medicinal chemistry. An alternative, more recent approach to the problem, has been to evaluate methods for controlling the biosynthesis of histamine *in vivo* through the use of specific histidine decarboxylase inhibitors.<sup>4</sup> *In vivo* and *in vitro* experiments using the HDC irreversible inhibitor,  $\alpha$ -fluoromethylhistidine, produced virtually complete inactivation of HDC activity in the rat hypothalamus, similar to results obtained with mice.<sup>124</sup> 18-*O*- $\beta$ -glycyrrhetic acid has been shown to inhibit HDC activity *in vitro* in cultured mast cells.<sup>125</sup> Somastatin also inhibits HDC in cells isolated from rabbit gastric mucosa by a dual mechanism involving a decrease in the affinity of the enzyme for (2*S*)-histidine and a reduction in the number of functional HDC molecules.<sup>126</sup> HDC activity in rat oxyntic mucosa has been

shown to be suppressed by the administration of the prostaglandin I derivative, beraprost sodium.<sup>127</sup>



*Histidine (25) and histamine (21).*

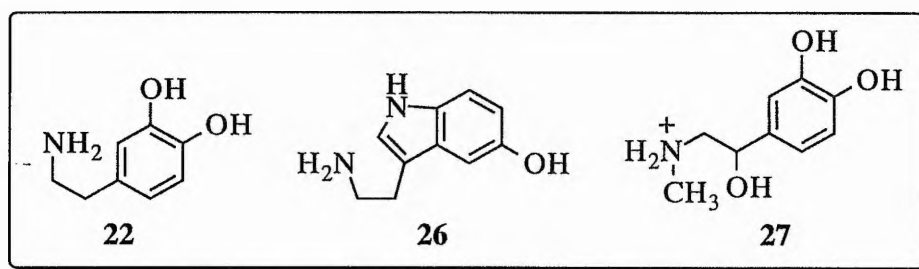
#### 1.4.2.4 DOPA Decarboxylase (DDC)

DOPA decarboxylase (DDC) is part of a family of aromatic acid decarboxylases which catalyse the formation of many pharmacologically important amines including catecholamines and indolamines.<sup>74</sup> These enzymes proliferate in bacteria, plants and animals. DOPA decarboxylase itself is found in mammals, and is broadly specific for a number of aromatic amino acid substrates.

Amines synthesised *via* the action of DDC include dopamine (22) and serotonin (26). Dopamine is a precursor in adrenaline biosynthesis. Both adrenaline (27) and serotonin are catecholamine neurotransmitters and act on smooth muscle tissue and the cardiovascular system. These two catecholamines are synthesised from tyrosine in sympathetic nerve terminals and in the medulla of the adrenal gland. Adrenaline causes an increase in blood pressure, strengthening of the heart-rate and widening of the passages of the lungs. Serotonin is also thought to be important in maintaining stable mental processes and schizophrenia has been attributed to abnormalities in the metabolism of this amine. Two non-dopaminergic psychotogenic drugs, lysergic acid diethylamide (LSD) and phencyclidine (PCP) significantly up-regulate DDC mRNA levels in rat brains, producing useful models of schizophrenia.<sup>128</sup> Amphetamine and vigabatrin, which act as dopaminergic system agonists, cause a down-regulation of DDC mRNA levels. The antipsychotic drugs, haloperidol and clozapine, relieved

the symptoms of PCP-treated animals. Evidence exists that the excess amounts of 2-phenylethylamine (2PE) and tryptamine produced by overactive DDC are associated with schizophrenia.<sup>128</sup> These compounds are modulators of all the neuroreceptors known to play a role in the disease.

The headache, chronic pain and depression associated with serotonin deficiency are often alleviated by administering B<sub>6</sub> supplements to the patient. DOPA and various inhibitors of the decarboxylase are used in the treatment of Parkinson's disease, a neurological disorder characterised by tremor, rigidity and disturbances of posture.<sup>74</sup>



3,4-dihydroxyphenylalanine (dopamine, 22), serotonin (26) and adrenaline (27).

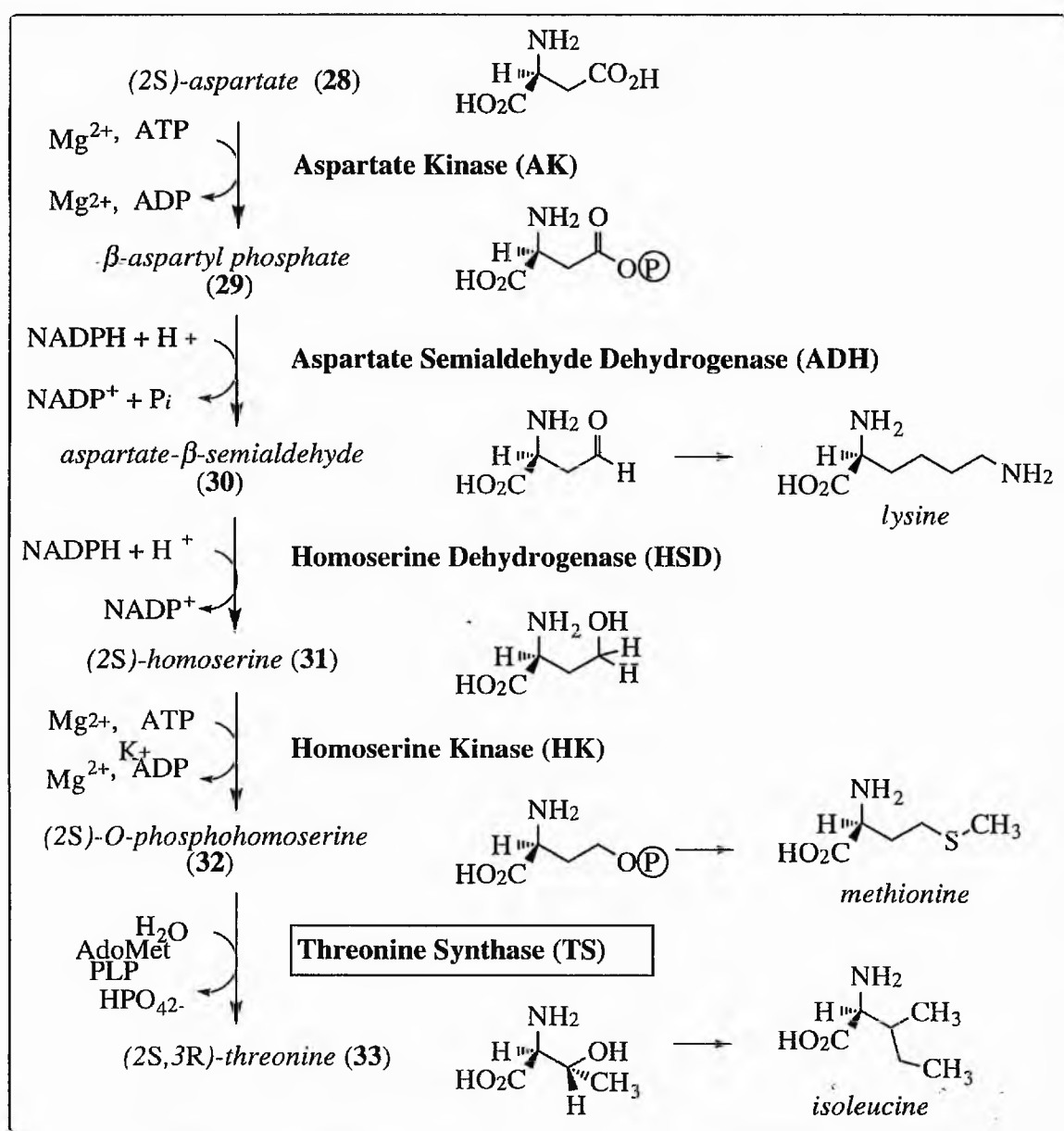
Serotonin has recently been shown to be a mechanism-based inactivator of DDC in both a time- and concentration-dependent manner.<sup>129</sup> This occurs *via* a decarboxylation-dependent transamination. Other substrates such as DOPA and *m*-tyrosine and their  $\alpha$ -methyl derivatives also undergo this minor reaction and it is thought that this could be the basis of a control mechanism.<sup>129</sup> Two quaternary benzophenanthridine alkaloids, sanguinarine and chelerythrine, have also exhibited strong inhibition of the enzyme from rat liver *in vitro*.<sup>130</sup>

### 1.4.3 Amino Acid Racemases

The racemases, which catalyse the interconversion of (2*S*)- and (2*R*)-amino acids, are extremely common in prokaryotes where they play a crucial role in the biosynthesis of peptidoglycan in bacterial cell walls.<sup>131</sup> These enzymes are therefore a possible target for chemotherapy.

### 1.4.4 The Aspartate Pathway

In both plants and bacteria, threonine synthase (TS) is the last of five enzymes involved in the biosynthesis of (2*S*,3*R*)-threonine (37) from (2*S*)-aspartic acid (32).<sup>54</sup> The first two steps of this biosynthetic pathway, known as the aspartate pathway, are common to lysine, threonine, isoleucine and methionine biosynthesis (see Scheme 1.8).<sup>132, 133</sup> (2*S*,3*R*)-threonine is one of the twenty common, nutritionally essential amino acids and as such is vital to all life forms.<sup>41</sup>



Scheme 1.8: The aspartate pathway.

## 1.5 Introduction to Threonine Synthase

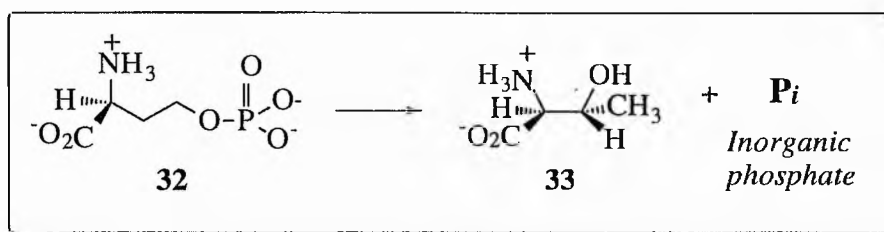
### 1.5.1 The Discovery of Threonine Synthase

The existence of the enzyme threonine synthase (EC 4.2.99.2)<sup>6</sup> was first suggested by H. J. Teas *et al.* in 1947.<sup>134</sup> In studies on mutant strains of both *Neurospora crassa* and *Bacillus subtilis*, which had a requirement of threonine (**33**) for growth, he demonstrated that homoserine (**31**) was a precursor to both threonine and methionine. During the early 1950s a number of other authors reported on work with various other species of bacteria which led them to assume that the pathway for threonine biosynthesis in all bacteria was as follows:<sup>135-139</sup>



**Figure 1.8:** Route for the biosynthesis of threonine from aspartic acid postulated in the 1950s.

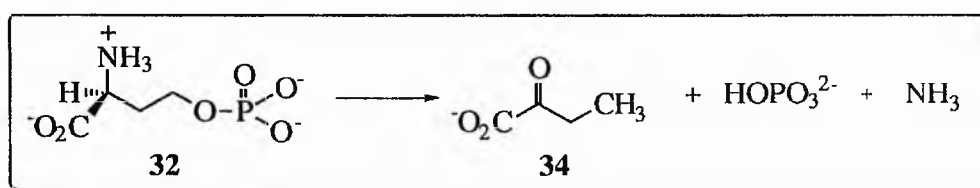
However, the enzyme was not properly identified in bacteria until 1960 when Flavin and Slaughter reported a pyridoxal phosphate-dependent enzyme, "which will be called threonine synthetase, [and which] catalyzed an elimination of orthophosphate coupled to isomerisation from  $\gamma$ - to  $\beta$ - hydroxy compound, to yield threonine from P-homoserine."<sup>140</sup> The conversion of (2*S*)-*O*-phosphohomoserine to threonine actually takes place stereospecifically *via* a  $\beta,\gamma$ -replacement, and a side-product of the reaction is inorganic phosphate (Scheme 1.9). It was not until 1975 that the enzyme was first isolated from plants.<sup>141</sup>



**Scheme 1.9:** The reaction catalysed by threonine synthase.

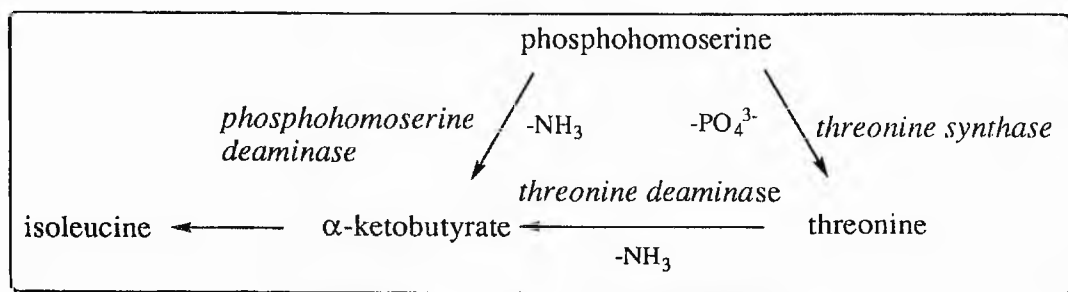
### 1.5.2 Associated Deaminase Activity of Threonine Synthase

TS is very selective for the substrate, and is only known to catalyse a few other reactions, often with substrates similar to (2*S*)-*O*-phosphohomoserine (see Appendix 2). In addition to catalysing the dephosphorylation of (2*S*)-*O*-phosphohomoserine, TS in *B. subtilis* and *E. coli* has also been shown to catalyse its subsequent deamination; a step usually carried out by the enzyme threonine deaminase (Scheme 1.10).<sup>142</sup> This reaction yields  $\alpha$ -ketobutyrate (34).



**Scheme 1.10:** (2*S*)-*O*-phosphohomoserine deaminase reaction.

Mutation events affecting TS activity also affected the deaminase activity suggesting that both activities were produced by the same enzyme.<sup>142, 143</sup> In contrast to the reaction involving threonine deaminase, the deaminase activity of TS has been shown not to pass through a threonine intermediate (Scheme 1.11).<sup>142</sup> This activity of TS was therefore named phosphohomoserine deaminase. This associated phosphohomoserine deaminase activity was shown to be a lesser activity of threonine synthase, as the two activities co-purified, were of the same molecular weight, co-inactivated upon heating and also displayed the same  $K_m$  value for the substrate, phosphohomoserine.<sup>144</sup> This associated deaminase activity appears unique to the TS enzymes of *B. subtilis* and *E. coli*.



**Scheme 1.11:** Alternative metabolic pathways for phosphohomoserine



### 1.5.3 Regulation of Threonine Synthase

As mentioned in Section 1.4.4 threonine synthase (TS) is the last of five enzymes involved in the biosynthesis of (2*S*,3*R*)-threonine (**33**) from (2*S*)-aspartic acid (**28**, Scheme 1.8).<sup>54</sup> Table 1.3 compares the human requirement for some of the common amino acids. Interestingly threonine is also one of only two essential amino acids containing two chiral centres, the other being isoleucine, itself synthesised from threonine. Threonine has also been found to play a role in the regulation of the aspartate pathway in a number of microorganisms by feedback inhibition of three of the enzymes involved.<sup>145</sup> Depending on the organism concerned it has been shown to inhibit aspartokinase,<sup>146-151</sup> homoserine dehydrogenase,<sup>152-155</sup> or homoserine kinase.<sup>140, 142, 156-160</sup> In plants TS does not generally appear to be subject to feedback inhibition by amino acids, but levels of TS activity are repressed by supplementation of the growth medium with methionine.<sup>161</sup> In wheat seedlings threonine almost completely inhibits its own synthesis without affecting methionine biosynthesis.<sup>161</sup> An inadequate supply of threonine results in a negative nitrogen balance by inhibiting protein synthesis due to stalling of the ribosome at codons for threonine (see Section 1.7). Other amino acids then accumulate and are transferred into degradative metabolic pathways causing a net loss in nitrogen.

**Table 1.3:** Minimal requirements of human beings for some of the essential amino acids (mg kg<sup>-1</sup> day<sup>-1</sup>)<sup>142</sup>

	Ile	Leu	Lys	Phe	Met	Cys	Thr	Trp	Val
<i>Child</i>	90.0	-	90.0	90.0	85.0	-	60.0	30.0	85.0
<i>Man</i>	10.4	9.9	8.8	4.3	1.3	11.6	6.5	2.9	8.8
<i>Woman</i>	5.2	7.1	3.3	3.1	4.7	0.5	3.5	2.1	9.2
<i>WHO*</i>	3.0	3.4	3.0	2.0	1.6	1.4	2.0	3.0	-

\*World Health Organisation norms

## 1.6 The Structure and Character of Threonine Synthase from Various Organisms

Until fairly recently little was known about the structure of threonine synthase (TS). Within the last ten years, however, the gene encoding the TS protein has been isolated and sequenced from a number of organisms, as has the enzyme itself. No 3D structures of the enzyme are available as yet. In all cases the enzyme has proved to be dependent on pyridoxal-5'-phosphate (PLP, 1).

### 1.6.1 Microbial Sources of Threonine Synthase

The gene for TS has been isolated and sequenced from *B. subtilis*,<sup>162</sup> *Brevibacterium lactofermentum*,<sup>143, 163</sup> *Corynebacterium glutamicum*,<sup>164</sup> *E. coli*,<sup>165</sup> *Pseudomonas aeruginosa*,<sup>166</sup> *Serratia marcescens*<sup>167</sup> and *Saccharomyces cerevisiae* (yeast).<sup>168, 169</sup> The enzyme has been isolated and characterised from these sources and also from *Neurospora crassa*.<sup>134</sup> Table 1.4 shows the various masses of the enzyme from different sources.

**Table 1.4.** *Threonine synthase from microbial sources*

<i>Species</i>	<i>Molecular Mass (M<sub>r</sub>) of TS (Da)</i>	<i>Number of Amino Acid Residues</i>	<i>Literature Reference</i>
<i>B. subtilis</i>	-	351	19
<i>Br. lactofermentum</i>	52, 807	481	170
<i>C. glutamicum</i>	54, 481	481	164
"	56, 000	-	133
<i>E. coli</i>	47, 060	428	165
( " )	36, 000	-	171
<i>M. glycogenes</i>	52, 167	475	172
<i>S. cerevisiae</i>	58, 000	514	169

Depending on the species in question, the TS protein has a molecular weight ranging from 36 to 58 kDa. As Fig. 1.9 shows, it is an enzyme which shows significant homology particularly in the amino acid residues in and around the active site (see also Appendix 6).<sup>173</sup> The average pairwise identity around the active-site is 41.8%. The lysine residue (indicated by an arrow) is thought by analogy with (2R)-serine dehydratase to be the residue responsible for binding pyridoxal phosphate to the enzyme by formation of the internal aldimine.<sup>19</sup> This idea is supported by the conservation of this residue and much of the surrounding sequence, such as the phenylalanine and aspartic acid residues to either side (underlined), in all of the microbes studied. As mentioned in Section 1.2.3, TS also shows significant sequence homology with and a similar active-site structure to other enzymes from the  $\beta$ -family of PLP-dependent enzymes, of which it is a member (see Figs. 1.3 and 1.4 and Appendix 1).<sup>24</sup>

	↓	
<i>B. subtilis</i>	-----TEGVN-PTGSFKDRG--MVM(A)VAKAKEEGN-----DTIMCASTGNTSAA	
<i>B. lactofermentum</i>	-----YEGAN-PTGSFKDRG--MVM(A)VAKAKEEGS-----EA(I)ICASTGNTSAS	
<i>C. glutamicum</i>	--IYLGH(L)SEGPTAAFKDMAMQLLGE(L)LFYELRRRN-----ETIN(I)LGATSGDT-GS	
<i>E. coli</i>	--VGCLE(L)FHGPTLAFKDFGGRFMAQMLTHIAGDKP-----VTIL(T)ATSGDT-GA	
<i>H. influenzae</i>	--IYALE(L)FHGPTLAFKDFGGRFMAQALAAVRGDGK-----ITIL(T)ATSGDT-GA	
<i>M. glycogenes</i>	--LYLLS(L)SNGPTLAFKDMAMQLLGN(L)LFYVLAQKG-----ETIN(I)LGATSGDT-GS	
<i>M. leprae</i>	-----VEGLN-PTGSFKDRG--MTMAVTDALARGQ-----RAV(L)CASTGNTSAS	
<i>M. tuberculosis</i>	-----VEGLN-PTGSFKDRG--MTMAVTDALAHGQ-----RAV(L)CASTGNTSAS	
<i>P. aeruginosa</i>	-----CVEL(F)FHGPTLAFKDFALQ(L)LLGRLLDHVLAARG-----ERVVIMGATSGDT-GS	
<i>S. pombe</i>	--LNVLE(L)FHGPTFAFKDVALQ(L)FLGN(L)FEFF(L)TRKNGNKPEDERDHLTVVGATSGDT-GS	
<i>S. marcescens</i>	--VSCLE(L)FHGPTLAFKDFGGRFMAQMLAEVAGEQP-----VTIL(T)ATSGDT-GA	
<i>S. cerevisiae</i>	ENLH(I)LE(L)FHGPTYAFK(D)VALQ(F)VG(N)LFY(F)LQ(R)INANLPEGEKKQ(I)TVVGATSGDT-GS	
	** .*** .	. . . * * *

Arrow indicates the lysine residue thought to be responsible for binding to PLP. Asterisks indicate conserved residues (highlighted in blue); homologous residues are shown by a period.<sup>41</sup> Note the conserved PT--FKD motif at the active-site in all enzymes. (Appendix 6 shows the alignment of full TS sequences and gives values for the similarity between each sequence). Multiple alignment was performed using Clustal W 1.60.

**Figure 1.9:** Sequence homology of amino acid residues around the active-site of threonine synthase from various microorganisms.

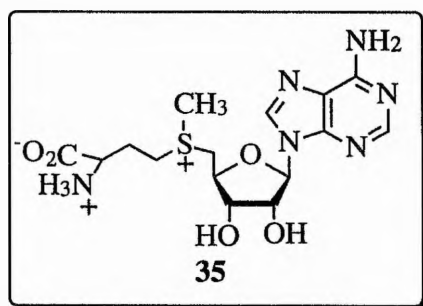
The first reported purification of threonine synthase was a 500-fold purification from *Neurospora* in 1960 through a series of acetone and ammonium sulfate fractionations followed by ion-exchange chromatography.<sup>137</sup> A crude preparation was obtained from *Brevibacterium flavum* by Miyajima *et al.* in 1968,<sup>145</sup> and a purified preparation from *C. glutamicum* by Han and co-workers in 1990.<sup>164</sup> However, there were no reports of the enzyme being purified to homogeneity until Parsot and his co-workers described a preparation of the enzyme from *E. coli* K-12 strain Tir 8 (strain Tir 8 is a constitutive mutant for the enzymes on the threonine operon).<sup>165, 174</sup> They carried out a 35-fold purification of the enzyme to 100% homogeneity in a protocol which has been adopted in subsequent experiments involving TS from this strain. Szczesiul and Wampler have since found TS from the same strain to have a pH optimum of 8.0 - 8.5,<sup>151</sup> and the same enzyme has an apparent  $K_m$  of 0.053 mM for phosphohomoserine.<sup>175</sup>

### 1.6.2 Threonine Synthase from Plant Sources

The study of the aspartate pathway in plants has aroused interest for two main reasons. Firstly, as a result of man's desire to improve the nutritional qualities of various crop plants and secondly, because an understanding of the aspartate pathway in plant tissues may aid in the design and synthesis of effective and selective herbicides. Cereals are deficient in the nutritionally essential amino acids lysine and methionine and legume seeds exhibit low quantities of methionine.<sup>176</sup> Increasing the efficiency of aspartate metabolism in plants would lead to improved levels of a number of amino acids in foodstocks.

Although there have been relatively few studies on TS in plants a number of groups have successfully isolated and purified threonine synthase from plant sources. The enzyme was first detected in 1975 by Schnyder *et al.* in the homogenates of pea seedlings and *Lemna minor*.<sup>141</sup> The following year Madison and Thompson identified the enzyme in sugar beet and radish leaves,<sup>177</sup> and it has since been partially purified from *Lemna pausicostata*,<sup>178</sup> barley seedlings,<sup>179</sup> pea seedlings,<sup>180</sup> soybean plants.<sup>181</sup> Wallsgrove and coworkers have also

shown that the enzymes of the aspartate pathway are confined to the chloroplasts of plants.<sup>182</sup> In all plants studied to date threonine synthase demonstrates an almost absolute requirement for S-adenosylmethionine (SAM, or AdoMet, **35**).<sup>132</sup> SAM acts as a potent allosteric activator of the enzyme, stimulating its activity as much as twenty-fold with a Hill coefficient of 2.0-2.5 (Table 1.5).<sup>177, 183</sup>



*S*-adenosylmethionine (**35**).

**Table 1.5:** Kinetic properties of threonine synthase from various higher plants.

<i>Plant Source</i>	<i>SAM</i>	$K_M$ (mM)	$V_{max}$	<i>Hill Coefficient</i>	<i>Literature Reference</i>
Sugar Beet	X	2.7	38.2 nmol h <sup>-1</sup>	-	165
"	✓	2.2	534.3 nmol h <sup>-1</sup>	-	165
Pea Seedlings	X	2.2	-	2.1	168
"	✓	0.67	4-fold increase	-	168
Barley Seedlings	✓	0.05	-	-	167
<i>Lemna pausicostata</i>	✓	2.2-6.9	70.8 pmol h <sup>-1</sup>	2.5	166

The Hill Coefficient is a measure of cooperativity, where the binding of a ligand (in this case SAM) to protein is at 50% saturation

Half-maximal activation of TS occurs at concentrations of between 40 and 200  $\mu$ M SAM.<sup>180</sup> Aarnes has found TS to have a pH optimum of around 8-9 in barley seedlings.<sup>179</sup> No isoenzymes of TS have been detected.<sup>132</sup>

### 1.6.3 Threonine Synthase Expression in Mammalian Systems

In order to improve the diet of non-ruminants it is necessary to supplement their feed with the essential amino acids, especially lysine and threonine (33). If the genes encoding the enzymes used in the biosynthesis of these amino acids could be expressed in animal cells, the need for supplementation of the diet would be eliminated.<sup>184</sup> Rees and Hay recently achieved the expression of aspartokinase and threonine synthase in mouse 3T3 cells.<sup>185</sup> These were then able to grow successfully on medium lacking threonine but containing (2*S*)-homoserine (31), the precursor to phosphohomoserine (32).

Threonine synthase and homoserine kinase have never been shown to occur naturally in mammals. It has, however, been noted that (2*S*)-*O*-phosphohomoserine is an antagonist of *N*-methyl-D-aspartate (NMDA) in the rat brain.<sup>186, 187</sup> (NMDA is an excitatory amino acid in the mammalian CNS). Phosphohomoserine has also been the subject of a patent for its use, together with the drug levemopamil, in the treatment of patients suffering from damage to neurons as a result of AIDS-related dementia, myelopathy, peripheral neuropathy and vision loss. This is caused by a reduction of the gp176 responsive rise in calcium ions.<sup>188</sup>

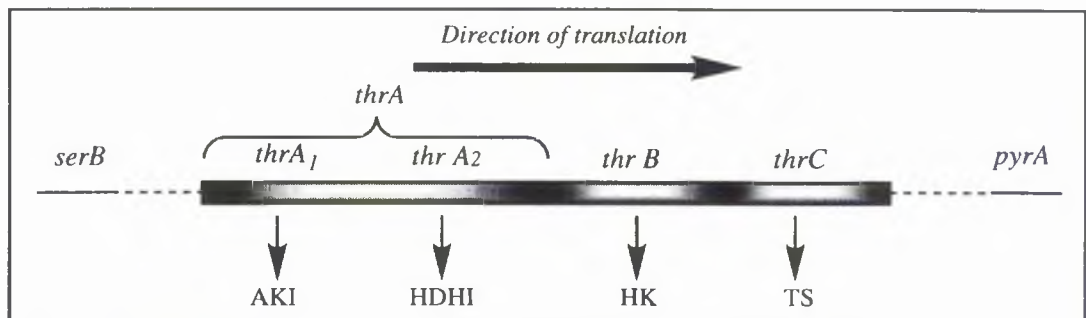
### 1.6.4 Threonine Synthase as a Potential Target Enzyme for Herbicides

Because of its high substrate specificity (see Section 1.5) and the range of organisms in which it is found, TS is an ideal target enzyme for herbicides, bactericides and fungicides. Selective inhibitors of the enzyme should not have adverse effects on animal species, which do not express TS. It is also known that blocking amino acid biosynthesis is a highly effective method of inhibiting growth in plants and bacteria.<sup>41</sup> Any structural and mechanistic information obtained about TS can be used in the rational design of potential inhibitors of the enzyme. Many inhibitors of TS have already been synthesised and tested and provide insight into the active-site structure of the enzyme. These are discussed fully in Section 1.9.

## 1.7 The Threonine Operon of *Escherichia coli*

### 1.7.1 The Structure of the Operon

The *thr* operon is at 0 min on the genetic map of *E. coli* and contains three structural genes, *thrA*, *thrB* and *thrC*.<sup>189, 190</sup> Transcription of the *thr* operon proceeds from *thrA* to *thrC* (Fig. 1.10) as established by analysis of phage Mu insertions and nonsense mutations in the *thr* cluster, and by analysis of a phage  $\lambda$  insertion into the *thr* regulatory region, which pleiotropically affected the expression of both *thrA* and *thrB*.<sup>191-193</sup> The bi-cistronic *thrA* gene encodes the threonine-sensitive bifunctional enzyme complex aspartate kinase I - homoserine dehydrogenase I (AKI-HDHI, Scheme 1.8).<sup>194, 195</sup> *E. coli* strain *K-12* contains two additional aspartate kinases (AKII and AKIII) and one additional homoserine dehydrogenase (HDHII), controlled by methionine and lysine respectively.<sup>196, 197</sup> The *thrB* gene encodes homoserine kinase (HK), which is inhibited by threonine.<sup>198</sup> *ThrC*, a gene of 1287 bp, encodes TS (Appendices 7 and 8).



**Figure 1.10:** Genetic map of the threonine operon of *E. coli* *K-12*.<sup>185, 190</sup>

The promoter for the operon has been localised approximately 230 nucleotides upstream of *thrA*, between a *BstE* II site at position -254 and a *Hha* I site at -184.<sup>199</sup> An operator region has been located between the promoter region and the structural genes.<sup>190</sup> Positions -147 to -82 upstream of *thrA*<sub>1</sub> on the operon encode a putative leader peptide of 21 amino acids (Fig. 1.11), containing eight threonine and four isoleucine residues, eleven of which occur consecutively.

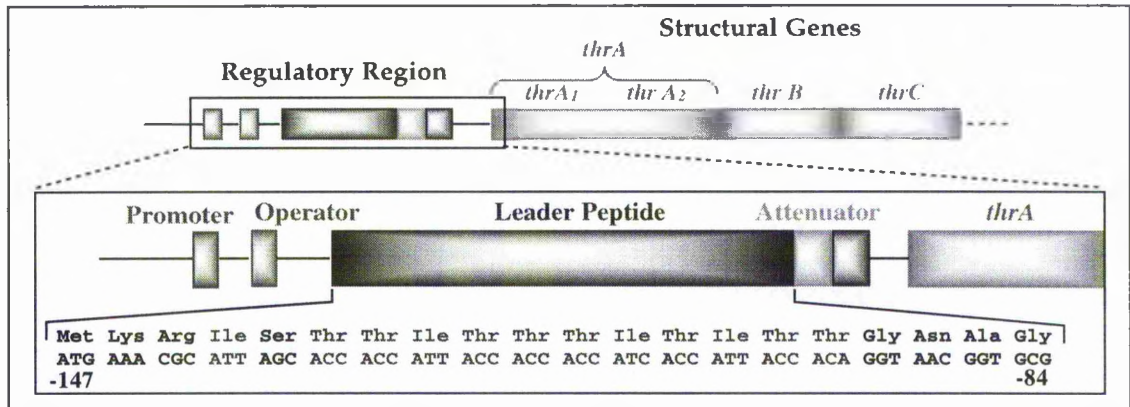


Figure 1.11: Regulatory region of the *thr* operon.

### 1.7.2 The Regulation of *Thr* Expression

At first the expression of the *thr* operon was thought to be under the control of a multivalent repression system dependent on the concentration of free threonine and isoleucine (Fig. 1.12).<sup>200</sup> However, Saint-Girons and Margarita were unable to detect the presence of any repressor bound to the operator and further research has shown that factors affecting threonyl tRNA synthase also affect *thr* expression.<sup>190, 172</sup> A model has therefore been proposed for the multivalent regulation of *thr* operon expression involving coupled transcription and translation of the leader region.<sup>172</sup> This is a common regulatory mechanism in phage and bacteria.

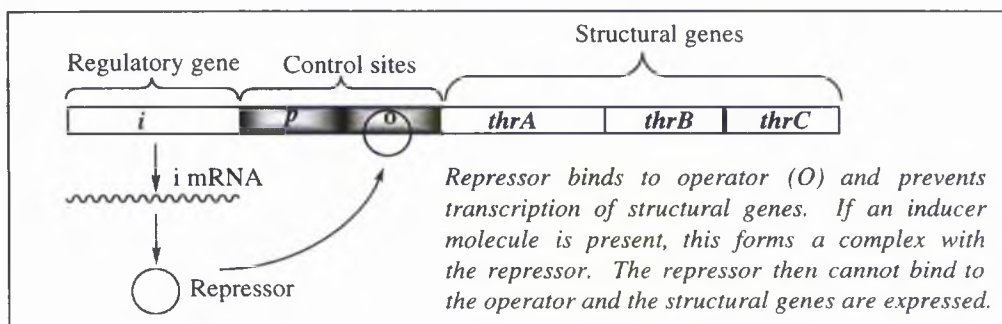
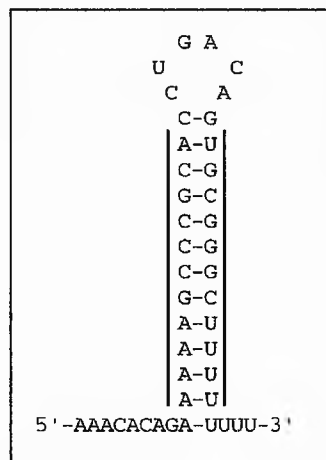


Figure 1.12: Model for a typical repression mechanism of gene regulation.<sup>74</sup>

If the intracellular levels of Ile and Thr are sufficient, RNA encoded by the *thr* regulatory region



from bases -71 to -27 folds to form a termination structure (Fig. 1.13).<sup>172, 201, 202</sup> This signals RNA polymerase to terminate transcription at a site (attenuator site) upstream from the structural genes. Under conditions of threonine and isoleucine *deficiency*, translation of the leader RNA sequence is stopped due to a shortage of threonyl- and isoleucyl-tRNA molecules. This pre-empts the termination structure, allowing transcription of the structural genes. A region of overlapping rotational symmetry from bases -93 to -49 overlaps the attenuator and can form of a structure which blocks the attenuator region.<sup>203</sup> A G-C rich region of dyad symmetry followed by a stretch of six uridine residues occurs 21 base pairs after the stop codon of *thrC*, which is thought to be a rho-dependent transcription termination site.<sup>204</sup>



**Figure 1.13:** The *thr* regulatory region from -71 to -27, which forms the termination structure.

### 1.7.3 Cloning of *Thr*

Amplification of the *thr* operon has been carried out by various groups<sup>205-207</sup> and has also been the subject of patents.<sup>208, 209</sup> The *thrC* gene has been cloned from *C. glutamicum* by Follettie and coworkers,<sup>210</sup> and from *Br. lactofermentum* by Malumbres *et al.*<sup>170</sup> A constitutive mutant *E. coli* K-12 Tir8 has also been produced, which over-expresses the enzymes of the *thr* operon.<sup>165</sup>

## 1.8 The Mechanism of (2S,3R)-Threonine Synthesis

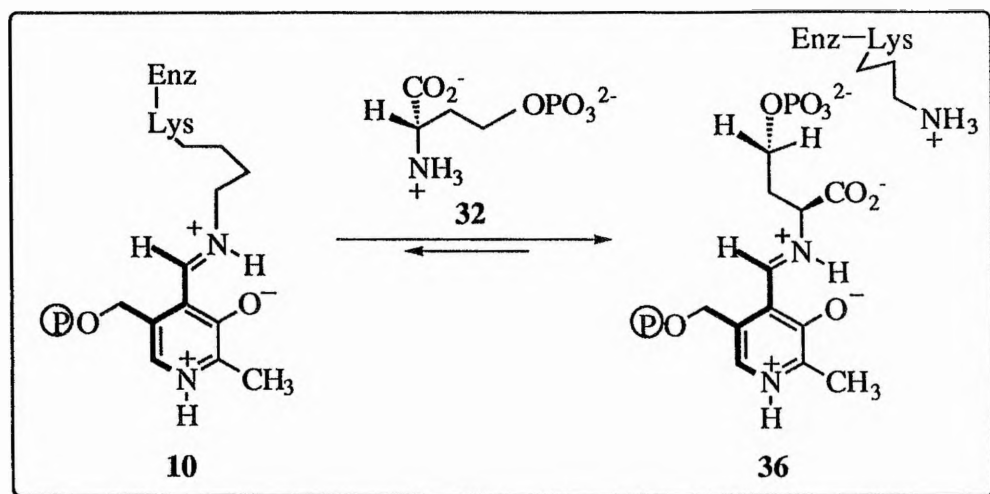
### 1.8.1 Introduction

Although there is a wide interest in PLP-dependent enzymes, relatively few studies have been carried out on the mechanism of the TS reaction.

The first few steps in the conversion of (2S)-O-phosphohomoserine (**32**) to (2S,3R)-threonine (**33**) are assumed to follow the same general mechanism as the reactions of other PLP-dependent enzymes acting at C-4 of the substrate. Most of the mechanistic information for such enzymes has been acquired from studies using cystathionine  $\gamma$ -synthase from *Salmonella*.<sup>4</sup> The reaction has been demonstrated to be essentially irreversible by Rognes.<sup>132</sup> He determined that no [<sup>14</sup>C]-O-phosphohomoserine was formed when *N. crassa* TS was incubated with [<sup>14</sup>C]-threonine and inorganic phosphate ( $P_i$ ), and that no enzymic exchange of <sup>32</sup>P<sub>i</sub> into phosphohomoserine occurred.

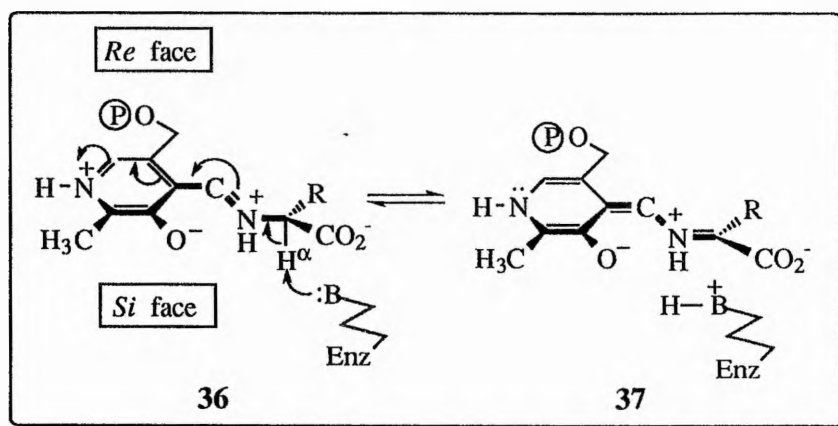
### 1.8.2 Reaction Mechanism

The first two steps in the TS reaction are described in more detail in Section 1.3.2. They involve the formation of a holoenzyme aldimine (**10**) between the  $\epsilon$ -amino group of a specific enzyme-bound lysine residue and the C-4' carbonyl group of PLP (Scheme 1.5).<sup>4, 5, 42</sup> The condensation of (2S)-O-phosphohomoserine with the internal aldimine (**10**) then leads to the formation of a Schiff's base substrate aldimine (**36**, Scheme 1.12), with the release of the enzyme-bound lysine residue.<sup>4, 42</sup> Farrington *et al.* found that substrate analogues, which were capable only of binding to PLP and undergoing C $^{\alpha}$  proton abstraction, exhibit rapidly reversible binding, suggesting that transaldimination is a *fast* step.<sup>175</sup>



**Scheme 1.12:** *Transaldimination reaction involving (2S)-O-phosphohomoserine (32) .*

The next stage in the reaction is labilisation of the  $C^\alpha$ -H bond to form the quinonoid intermediate (37, Scheme 1.13). As described in Section 1.3.2, Dunathan's postulate states that in the external aldimine the bond to be cleaved at  $C^\alpha$  is held orthogonal to the plane of the pyridinium ring, thus allowing maximum orbital overlap between the developing negative charge and the conjugated electron-deficient  $\pi$ -system.<sup>4, 5</sup>



**Scheme 1.13:** *Labilisation of  $H^\alpha$  by an enzyme-bound base.*

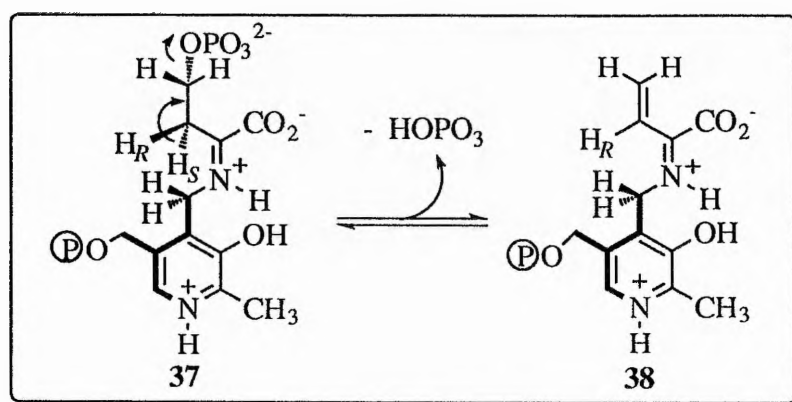
Fuganti suggests that not only the  $C^\alpha$ -H, but also the  $C^\beta$ -H and  $C^\gamma$ -OPO<sub>3</sub>H  $\sigma$  bonds in the Schiff's base, all of which are broken in the course of the TS reaction, must be held perpendicular to the extended  $\pi$ -system, in order to optimize orbital interactions in the bond

breaking and making processes.<sup>212</sup> It has been demonstrated that the known reactions of PLP-enzymes all take place only on the 'exposed', *i.e.* the *si*-face at C-4' of the cofactor (below the plane of the page in all structures shown here).<sup>5</sup> It is therefore assumed that TS follows this general pattern.

The cleavage of H $^{\alpha}$  is further facilitated by an enzyme-bound base, probably the unprotonated lysine residue originally involved in binding the enzyme to the coenzyme, PLP.<sup>4</sup> This has certainly been suggested by data produced in studies on OASS (Section 1.3.3.1).<sup>66</sup> The  $\epsilon$ -amino group of this lysine residue stabilizes the positive charge as it forms on the electrofuge.<sup>46</sup> The removal of H $^{\alpha}$  forms a carbanion at C $^{\alpha}$ , which is resonance stabilised as a pyridoxal-ketimine-*p*-quinonoid (37). In this intermediate the acidity of the protons at C-3 is increased through being adjacent to a conjugated system of double bonds, and they become more labile. OASS, another  $\beta$ -family enzyme, does not appear to form a quinonoid intermediate, whereas for TrS such a species has been fleetingly detected.<sup>67</sup> When F. Barclay incubated (2*S*)-[2-<sup>2</sup>H<sub>1</sub>]-*O*-phosphohomoserine with the enzyme from *E. coli*, isotope effects of 1.2 and 1.0 were obtained for <sup>D</sup>V and <sup>D</sup>(V/K) respectively. This indicates that the removal of C $^{\alpha}$ -H from the substrate is kinetically significant but not rate-limiting.<sup>41</sup> The fact that the value of <sup>D</sup>V is small, but larger than the value of <sup>D</sup>(V/K) for the C-2-deuteriated compound, strongly suggests that the high forward commitment to the C $^{\alpha}$ -H proton removal step occurs because transaldimination in the reverse reaction direction and/or dissociation of the substrate from the enzyme is very slow. Farrington *et al.* also proposed that immediately after C $^{\alpha}$ -elimination, the phenolic oxygen and C-4' of PLP are protonated (as shown in Scheme 1.14).<sup>175</sup> TS has been shown to produce PMP in half-transamination reactions with a number of amino acid substrates which do not possess a leaving group in the  $\beta$ -position.<sup>211</sup>

In 1979 Fuganti first demonstrated, using yeast TS, that it was the 3-*pro-S* rather than the 3-*pro-R* proton which was eliminated in the next stage of the reaction, using (3*R*)- and (3*S*)-[3-<sup>3</sup>H; U-<sup>14</sup>C]-(2*S*)-*O*-phosphohomoserine as substrates.<sup>212</sup> This is unusual, as most other PLP-dependent enzymes acting at C-4 of the substrate catalyse the removal of the 3-*pro-R* hydrogen.<sup>4</sup> This was also observed for the *E. coli* enzyme. The (2*S*,3*S*)-[3-<sup>2</sup>H<sub>1</sub>]-isotopomer of the substrate displayed isotope effects of 1.74 and 1.0 for <sup>D</sup>V and <sup>D</sup>(V/K) respectively

whereas the (2*S*,3*R*)-[3-<sup>2</sup>H<sub>1</sub>]-isotopomer produced no kinetic isotope effect.<sup>41</sup> The value of  $D(V/K)$  for both the (2*S*)-[2-<sup>2</sup>H<sub>1</sub>] and the (2*S*,3*S*)-[3-<sup>2</sup>H<sub>1</sub>]-isotopomers was unity, indicating that the first two reaction steps provide a large forward reaction commitment. The substrate can therefore not readily debind from the enzyme *via* reverse steps once the quinonoid intermediate (37) has formed. Laber *et al.* observed the accumulation of the quinonoid (observed as a 448 - 453nm chromophore) which implied that the elimination of the β-substituent is rate-limiting as for tryptophanase.<sup>211</sup> F. Barclay also noted that hydrogen elimination was slow compared to other steps in the reaction, suggesting that the steady-state concentration of 37 is quite high.<sup>41</sup>



Scheme 1.14:  $\beta,\gamma$ -Elimination of phosphoric acid.

Loss of the 3-*pro-S* hydrogen could occur in a concerted manner with elimination of the phosphate group, *via* the formation of an internal phosphate complex (Fig. 1.17 and Scheme 1.14). If this were the case, it is most likely that the  $\beta,\gamma$ -elimination of phosphoric acid would occur in a *syn*-fashion, although normally an *anti*-periplanar elimination would be assumed to be energetically more favourable, due to the enzyme covering the *re*-face of the coenzyme. This would produce an enzyme-bound vinylglycine intermediate.

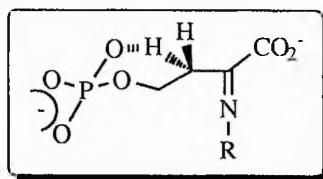
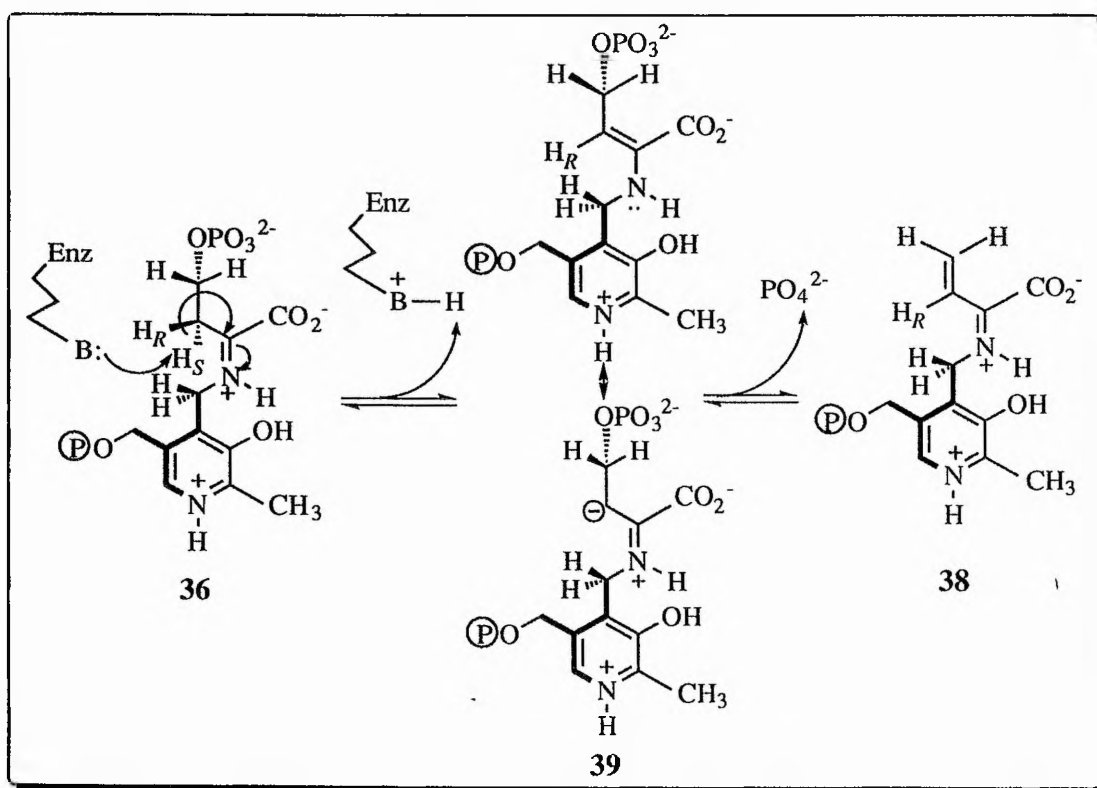


Figure 1.14: Internal phosphate complex formed prior to elimination of phosphoric acid.

Another postulate for the elimination of the phosphate group involves a sequential reaction, whereby the 3-*pro-S* hydrogen is removed first by a base approaching from the 4'-*si*-face to form a resonance stabilised C<sup>β</sup> carbanion (39, Scheme 1.15). Subsequent elimination of the phosphate group (the nucleofuge), also positioned on the 4'-*si*-face of the coenzyme, from C<sup>γ</sup> then generates the vinyglycine intermediate, as shown before (38, Scheme 1.15). Phosphorylation of homoserine to replace a hydroxyl group with the more electronegative phosphate group facilitates this  $\gamma$ -elimination.

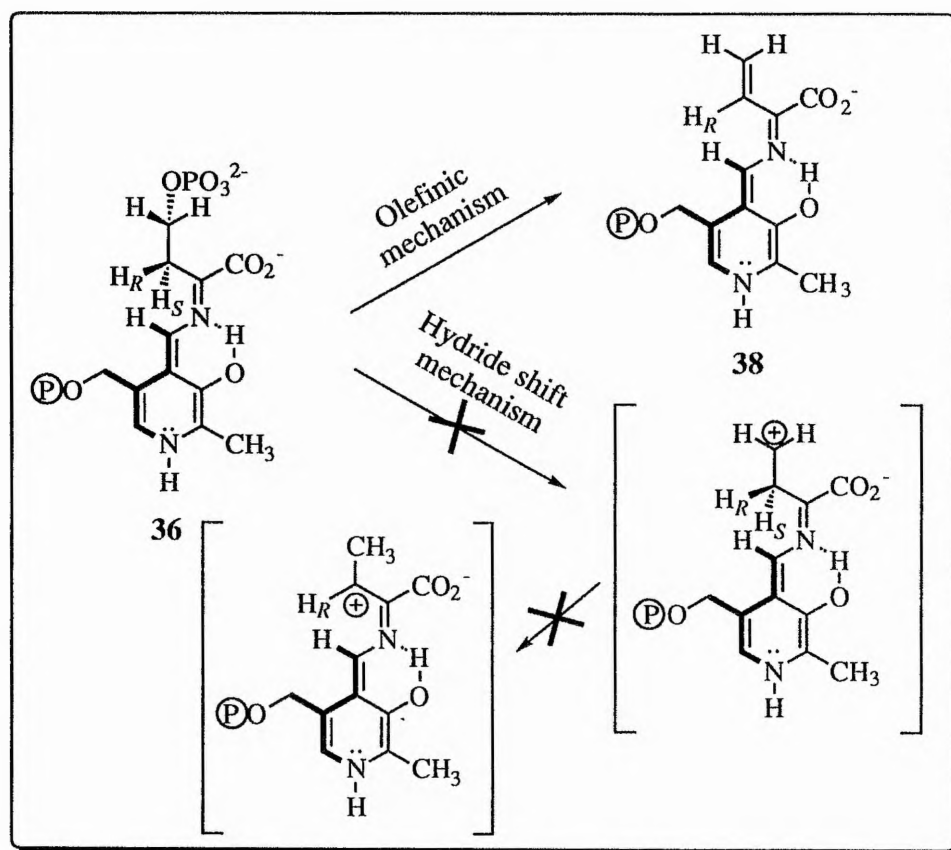


Scheme 1.15: Labilisation of the 3-*pro-S* hydrogen atom.

Laber *et al.* observed that vinyglycine was converted to threonine upon incubation with TS, confirming that the pyridoxal *p*-quinonoid of vinyglycine 38 is an intermediate in the reaction.<sup>211</sup> The initial absence of any long-wavelength chromophore was seen to argue against the transient accumulation of 38.

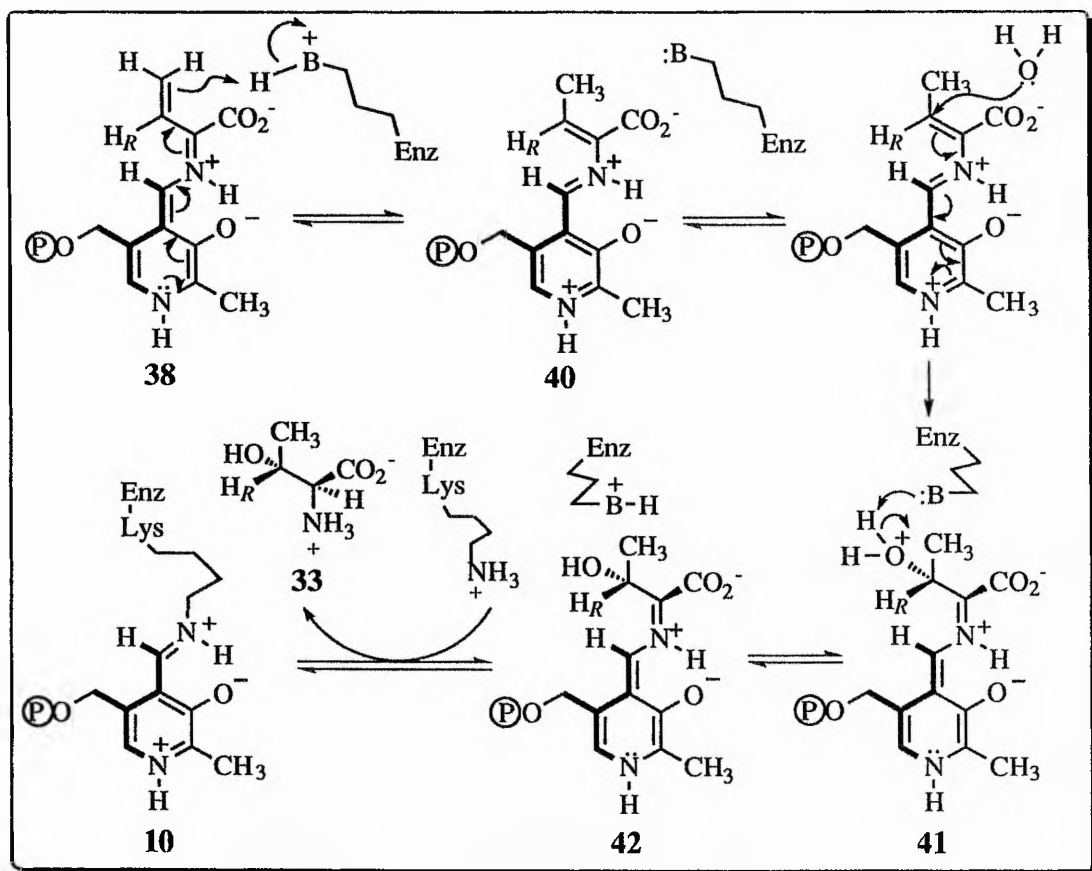
Although it is not yet known which of the aforementioned  $\gamma$ -elimination mechanisms are

preferred, it has been shown that inorganic phosphate is formed by *nonhydrolytic* elimination rather than by hydrolysis of the phosphate group at C $^{\gamma}$  (i.e. the C $^{\gamma}$ -O rather than the P-O bond is broken). This was discovered by Flavin and Slaughter in studies carried out in  $^{18}\text{O}$ -labelled water, where  $^{18}\text{O}$  was incorporated into (2*S*,3*R*)-threonine but not into phosphate.<sup>213</sup> They also proved, using  $^3\text{H}$ -labelling, that it was the vinylglycine that was formed, rather than the reaction proceeding *via* a hydride-shift mechanism (see Scheme 1.16).



**Scheme 1.16:** Comparison of proposed mechanisms for the elimination of phosphate.

The terminal alkene (or vinylglycine intermediate, 38) produced by the removal of the phosphate group is re-saturated by a hydrogen atom, which may come from the solvent, or from a protonated base on the enzyme, again probably from the 4'-*si*-face (40, Scheme 1.17). Farrington *et al.* propose that, immediately prior to this step, the phenolic oxygen and C-4' of PLP are deprotonated again.<sup>175</sup>



**Scheme 1.17:** Resaturation of vinylglycine intermediate and release of threonine.

It is known from incubations conducted in deuterium oxide that threonine synthase introduces two atoms of deuterium into the product, one at C-2 and one at C-4, leading to the assumption that no significant *internal* proton transfer takes place.<sup>4</sup>

The TS reaction is then completed by a hydroxyl group or water attacking at C-3 of the substrate.<sup>5</sup> As the absolute configuration of threonine is (2*S*,3*R*) this process must occur *via* a retentive mode.<sup>4</sup> (2*S*,3*R*)-threonine can now be released from the coenzyme *via* reverse transaldimination, which regenerates the holoenzyme aldimine.

The mechanism of this  $\beta,\gamma$ -replacement requires a completely different conformation at the active site compared with  $\beta$ -replacement systems. This difference in the stereochemical courses of the two types of reaction can be explained either by the transient quinonoid intermediate described earlier or by a two-base system.<sup>4</sup>



### 1.8.3 Unanswered Questions Concerning the Threonine Synthase Reaction

- The removal of the C-3 proton and the elimination of phosphate from the quinonoid intermediate (**37**) may be sequential or concerted (Section 1.8.2). The concerted elimination should occur in a *syn*-fashion, if the enzyme covers the *re*-face of PLP, but it is not yet certain from which face of the coenzyme the phosphate group leaves. It is also not clear whether the phosphate group or the  $\epsilon$ -amino group of the transaldiminated lysine residue would act as the base removing the C-3 proton (Scheme 1.15).
- It is not yet known whether the reprotonation of the enamine (**38**) to give the imino acid moiety (**40**) occurs with retention or inversion of configuration, whether the protonation occurs from the 4'-*si*-face as assumed and whether internal proton transfer from C-3 to C-4 occurs. Flavin and Slaughter saw no internal proton transfer when they assessed tritium transfer from the substrate to the product.<sup>213</sup> However, if the proton in the conjugate acid is readily exchangeable with solvent protons, then very few tritium atoms would be transferred back to the substrate in this type of assay. More careful analysis of the isotope content of the threonine methyl group from very short-time experiments is required. The reprotonation step may also be reversible.
- The attack of water at C-3 of **40** is defined by the stereochemistry at C-3 of threonine (**33**) (*i.e.* 3*R*) but the geometry of the attack is not known with respect to the coenzyme.

## 1.9 Inhibitors of Threonine Synthase

### 1.9.1 Functional Groups Which Inhibit PLP-Enzymes

In section 1.4 the metabolic role of various PLP-enzymes was discussed, demonstrating the wide scope of action of such enzymes. As a result of their varied functions the inhibition of PLP-enzymes has potential pharmacological and therapeutic utility. Generally PLP-enzymes are inhibited by carbonyl attacking reagents such as hydroxyl amines and hydrazines, and substrate analogues which bind at the active-site but cannot undergo further reaction.<sup>4</sup> In the absence of endogenous coenzyme many PLP-dependent enzymes are also inhibited by the production of PMP in abortive transaminase side reactions. The development of mechanism-based, suicide inhibitors of PLP-enzymes which are substrate analogues has been widely studied, both for their inhibitory properties and as a means of elucidating the mechanism of specific enzymes.<sup>214-216</sup> As the role of PLP is to stabilise carbanionic intermediates that form during the catalytic process (Section 1.3.1), inhibitors possessing functional groups which become activated by their close proximity to an enzyme-generated carbanion can break down to yield a reactive species. This may react with an active-site amino acid side-chain, or with tightly bound PLP-coenzyme. Either of these processes will lead to the inactivation of the enzyme, in the first instance blocking the active-site with an unreactive species, or in the second case causing the coenzyme to become unavailable to bind substrate.

Functional groups commonly used in the design of such inhibitors include acetylenic, olefinic,  $\beta$ -halo substituents and other leaving groups, nitriles, aryl sulfoxides, dihydroaromatics and phosphonamino acids.<sup>214, 217</sup> A number of olefinic substrate analogues such as vinylglycine,<sup>216, 218</sup> (*E*)-methoxy vinylglycine (**43**)<sup>219</sup> and  $\beta$ -methylene aspartate (**44**)<sup>220</sup> have been demonstrated to inhibit AspAT. Acetylenic analogues such as (*2S*)-propargyl glycine (**45**) inhibit a range of PLP-enzymes including the transaminases AspAT and  $\omega$ -ornithine aminotransferase and both  $\gamma$ -cystathionase and cystathionine- $\gamma$ -synthetase.<sup>221, 222</sup>

Some of the most common suicide substrates for transaminases are those which possess a good leaving-group at C $\beta$ , as they can undergo a facile elimination to generate an olefinic intermediate. (2S)-3-chloroalanine (**46**)<sup>223, 224</sup> inhibits AlaAT and various halogenated alanines are suicide substrates for tryptophan synthase (TrS).<sup>225</sup> (2S)-serine-O-sulfate (**47**) also acts as an inhibitor of various PLP-enzymes including AspAT, and glutamate decarboxylase.<sup>226</sup> GABA-T is inhibited by various acetylenic, olefinic and halogenated substrate analogues<sup>227, 228</sup> and a number of carbonyl attacking reagents such as hydroxylamine, aminooxyacetic acid, hydrazines and carbazides.<sup>4</sup>

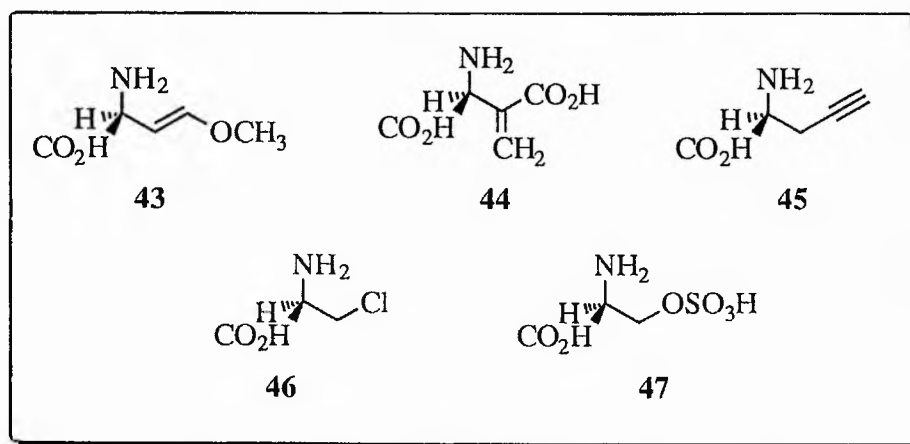


Figure 1.15: Inhibitors of PLP-enzymes.

### 1.9.2 Amino Acid Inhibitors of Threonine Synthase

Various studies have been carried out on the inhibitory properties of both biogenic and exogenous amino acids on enzymes of the aspartate pathway, some of which are proven inhibitors of threonine synthase. In 1967 Miyajima *et al.* found that TS from *Br. flavum* was scarcely inhibited by amino acids of the aspartate family suggesting that the enzyme may only be subject to weak feedback inhibition or repression *in vivo*.<sup>229</sup> Lysine, threonine and isoleucine only showed approximately 15 - 20% inhibition of the enzyme, and the presence of methionine negated the effects of lysine and isoleucine. However, cysteine (**48**) and

glutathione (**49**) showed 100% inhibition and alanine (**50**) showed 44% inhibition of TS. In *Lemna pausicostata* cysteine has also been shown to inhibit the stimulation of TS by SAM.<sup>180</sup> As cysteine and phosphohomoserine are converted to cystathionine by another enzyme it has been suggested that the conversion of (2S)-O-phosphohomoserine to either threonine or cystathionine is regulated by the relative amounts of SAM and cysteine present.<sup>180</sup> In 1984 Giovanelli and coworkers carried out a separate study on the activity of TS in *Lemna pausicostata* with mechanism-based inhibitors of PLP enzymes. Although DL-vinylglycine and cysteine both irreversibly inhibited TS to a small extent (22 and 17% respectively) both (2S)-propargyl glycine (PAG) and gabaculine failed to act as suicide substrates for the enzyme.<sup>178</sup> In a later study by Rognes *et al.* the TS activity in pea seedlings was reduced by up to one half in the presence of excess methionine (**51**).<sup>160</sup> Aarnes also tested a number of amino acids on TS from barley seedlings, but only (2S)-ornithine (**52**) and 2,4-diamino butyrate had any significant effect (75 and 77% of control activity respectively). The latter was also found to inhibit the enzyme in *Lathyrus sylvestris*.<sup>179</sup>

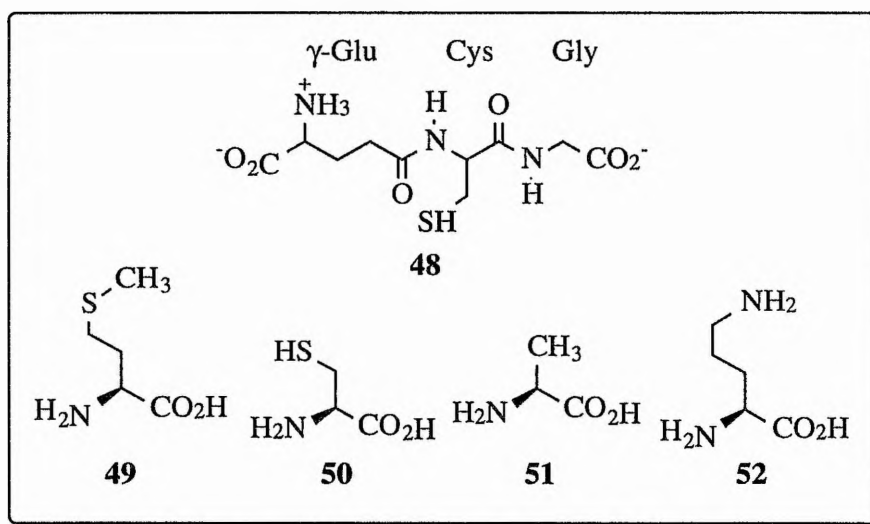


Figure 1.16: Amino acid inhibitors of threonine synthase.

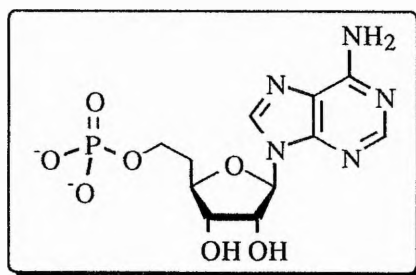
### 1.9.3 Inhibition of Threonine Synthase by Various Anions

Aarnes found that conditions of high ionic strength appeared to inhibit threonine synthase in barley seedlings.<sup>179</sup> The addition of 100 mmol of NaCl, KCl and  $(\text{NH}_4)\text{SO}_4$  reduced the activity of the enzyme by approximately 25%, although NaF had no effect. As shown previously (see Scheme 1.9) inorganic phosphate is produced as a side-product of the threonine synthase reaction and it has also been shown to strongly inhibit threonine synthase in *Lemna pausicostata* in a reversible manner.<sup>183</sup> Other anions tested for their effects on TS were less inhibitory (see Table 1.6). Adenosine monophosphate (AMP, 53) is also an extremely potent and structurally specific inhibitor of threonine synthase, each inhibition progressively decreasing with increasing concentrations of (2S)-O-phosphohomoserine (*i.e.* competitive inhibition).<sup>183</sup> Most analogues of AMP do not share its inhibitory properties, and only the N<sup>6</sup>-substituted derivatives resulted in inhibitions even approaching that of the parent compound.

**Table 1.6.** Inhibition of TS from different sources by various anions and salts.

Barley Seedlings			<i>L. pausicostata</i>		
Salt	Concentration	Inhibition	Anion <sup>a</sup>	Concentration	Inhibition
NaCl	100 mmol	25%	P <sub>i</sub>	1 mmol	50%
KCl	"	25%	PP <sub>i</sub>	"	15%
$(\text{NH}_4)\text{SO}_4$	"	25%	SO <sub>4</sub> <sup>2-</sup>	"	30%
NaF	"	none	Cl <sup>-</sup>	"	none

a - cations are not specified in the reference.



Adenosine monophosphate (AMP, 53)

#### 1.9.4 Phosphonic Acid Analogues of (2S)-O-Phosphohomoserine

Phosphonic acid structural analogues often exhibit high affinities (and are therefore potent inhibitors) for enzymes which catalyze phosphoryl transfer or phosphate hydrolysis reactions and threonine synthase is no exception. This is due to the fact that these analogues do not undergo complete catalysis due to the chemically inert nature of the C-P bond.<sup>175</sup> Ash *et al.* first explored the suitability of various sulfide analogues as inhibitors of *E. coli* TS, finding that 2-amino-3-[(phosphonomethyl)thio]propionic acid (**55**, see Table 1.7) was a slow-binding potent inhibitor of the enzyme.<sup>230</sup> The efficacy of this compound probably lies in the fact that it cannot proceed past the formation of the quinonoid intermediate (**37**, Scheme 1.14), but is already committed to the catalytic process. In 1993 Farrington *et al.* carried out further studies on phosphonic acid inhibitors of TS from *E. coli* K-12 Tir-8,<sup>165</sup> in particular analogues **55-58** and **61**.<sup>175</sup> The results of similar studies by Pohlenz and his group were published in 1994. They describe the synthesis of a number of analogues of the natural substrate (**55**, **59**, **60** and **64-68**) which they later tested on TS along with analogues **61-63**.<sup>231, 232</sup>

Although the sulfide analogue **55** is a good inhibitor, the  $\beta$ -sulfoxide and  $\beta$ -oxo-analogues of the substrate (**56** and **60**) have a significantly reduced potency (one order of magnitude). In the case of the sulfoxide analogue this is probably due to steric hindrance on the active-site preventing it from effectively binding to TS. The oxo-analogue **60** may be inert as an inhibitor due to the less nucleophilic nature of oxygen relative to sulfur. This may have some significance for the binding of substrates to TS. Thioethers also have the advantage that the sulfur atom, which is more readily polarised, will stabilise adjacent carbanions. This renders the  $\alpha$ -protons much more acidic than is the case for ethers. Thus, in the case for thioether the quinonoid intermediate is more readily formed (Scheme 1.15). The  $\beta$ -amino analogue **59** on the other hand demonstrates a  $K_i$  value (59  $\mu\text{M}$ ) comparable with the  $K_m$  of the substrate (53  $\mu\text{M}$ ).<sup>175</sup> Analogues with a chain length of three carbons are, generally speaking, more potent inhibitors, with the potency decreasing as the number of carbons along the chain increases, indicating that larger compounds have difficulty fitting into active-site of the enzyme. This also has implications for the positioning of the phosphate group in the active-site (see later

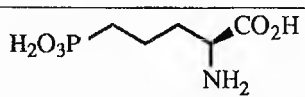
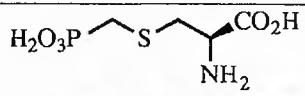
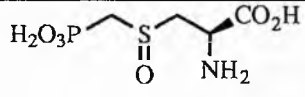
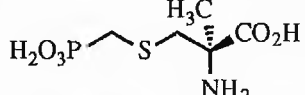
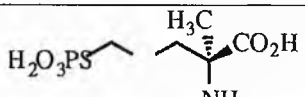
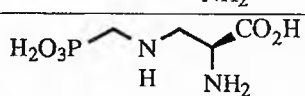
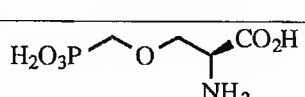
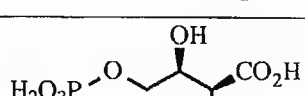
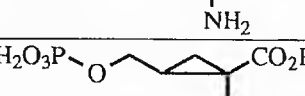
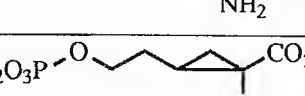

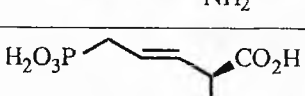
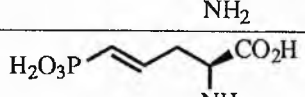
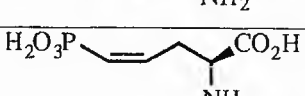
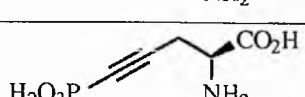
Structure	No.	$K_i$ (mM)	Type of Inhibition
	54	0.031	competitive
	55	0.011	slow-binding
	56	0.194	competitive
	57	0.225	competitive
	58	0.330	competitive
	59	0.059	enzyme-activated irreversible inhibition
	60	none	-----
	61	0.05	competitive
	62	0.01	competitive
	63	none	-----
	64	0.10	irreversible
	65	0.40	irreversible
	66	0.54	reversible competitive
	67	none	-----
	68	none	-----

Table 1.7: Inhibitors of threonine synthase.

discussion). Replacement of H $\alpha$  with a methyl group also hinders the binding of the substrate to the enzyme both sterically and because this precludes any reaction beyond the formation of the holoenzyme aldimine. These substrates are therefore less committed to the catalytic pathway.<sup>175</sup> Olefinic and acetylenic amino acids are well known inactivators of various PLP-dependent enzymes (Section 1.4.1). Indeed the olefinic analogues **64** and **65** do inhibit TS in *E. coli*, but the olefinic analogue **67** and the acetylenic analogue **68** both fail to inactivate TS. The most likely explanation for the lack of reactivity of these compounds is that the phosphate group is held too rigidly in a conformation which prevents it from interacting with specific residues on the enzyme. In a study by Giovanelli *et. al.* the importance of the  $\gamma$ -phosphate group on (2*S*)-*O*-phosphohomoserine for binding was clearly demonstrated.<sup>178</sup> A variety of *O*-acyl homoserine esters had a binding affinity for TS at least 2 orders of magnitude smaller than that of the natural substrate. Also the poor binding affinity of *O*-phosphoserine and *O*-phosphothreonine indicated that the optimal position for the phosphate group was at the  $\gamma$ -position. Analogue **64**, [(2*S*)-2-amino-5-phosphono-3-*cis*-pentenoic acid (Z-APPA)], has been shown to be the active component of plumbemycins<sup>233, 234</sup> and rhizocitins,<sup>235</sup> which are di- and tripeptides that possess bactericidal, fungicidal or acaricidal activity according to the nature of the peptide. The acid has been shown to decrease the activity of threonine synthase by acting as a suicide inhibitor and is also known to have herbicidal activity.

### 1.9.5 Phosphate Analogues

The use of naturally occurring antimetabolites as bacteriocidal agents has received much attention. In 1984 Shames *et al.* carried out studies on the action of aspartate-derived antimetabolites on the enzymes of the threonine biosynthetic pathway, and found that (2*S*)-*threo*-3-hydroxyhomoserine phosphate (**61**) was a potent inhibitor of threonine synthase,<sup>236</sup> a finding confirmed by Laber *et al.*<sup>232</sup> They also investigated the properties of cyclopropyl analogues **62** and **63** and showed that the first of these was a potent inhibitor of TS, but the latter showed no inhibition under standard assay conditions. This is probably due to the extended length of the compound, causing the phosphate group to be in the wrong position for binding.



## 1.10 The Aims of This Project

### 1.10.1 Testing New Substrate Analogues as Inhibitors of Threonine Synthase

Previous inhibition studies on TS indicate the following about potent inhibitors of the enzyme:

- The side-chain should be 3 carbons in length.
- A phosphate (or phosphonate) group should be in the  $\gamma$ -position for optimum binding.
- The inhibitor should undergo both C $^{\alpha}$ -H cleavage and 3-*pro*-S proton removal, exhibiting a high forward reaction commitment, but remain inert to further reaction.

Potential inhibitors of TS were therefore designed which should promote the removal of a proton in the C-3 position; (2*R*,3*S*)- and (2*R*,3*R*)-3-fluorophosphohomoserine.  $\beta$ -fluoro-amino acids have previously been shown to be good inhibitors of certain PLP-dependent enzymes. It was assumed that the small size of fluorine would allow these compounds to fit easily into the active site, but that the electronegativity of fluorine would cause perturbation of the enzyme mechanism. The analogues (2*S*,3*S*), (2*S*,3*R*)-, (2*S*,4*S*)- and (2*S*,4*R*)-methyl-*O*-phosphohomoserine were also designed to aid in the elucidation of the geometry of bound reaction intermediates at both C-3 and C-4. These substrate analogues have been synthesised by G. Allan (unpublished results). The aim of *this* work was to assay all these analogues with TS from *E. coli* K-12 Tir8 and determine their effectiveness as substrates or inhibitors of the enzyme. The rate of turnover of these analogues and the substrate was to be monitored using malachite green assays, which detect even nanomolar quantities of the side-product, inorganic phosphate.<sup>41, 238-241</sup>  $K_M$  values can be determined by measuring the initial reaction velocity ( $v$ ) for a series of different analogue concentrations [S]. This data can then be represented on a Lineweaver-Burk plot ( $1/v$  against  $1/[S]$ ). The slope of this plot is  $K_M/V_{max}$ , the y intercept is  $1/V_{max}$  and the x intercept is  $-1/K_M$ .<sup>242</sup> Alternatively the Eadie-Hofstee plot ( $v$  vs.  $v/[S]$ ) is more accurate and gives a slope of  $-K_M$ , a y intercept of  $V_{max}$  and an x intercept of  $V_{max}/K_M$ . Different *types* of inhibition can also be determined using these types of plot. Values of  $v$  are determined for a constant inhibitor concentration ([I]) over a range of [S] values (in this case S is the natural substrate). This data is determined for several values of [I] and the separate

lines for each  $[I]$  value are drawn on the same graph. *Competitive* inhibition is indicated by a family of lines with a common y intercept on both Eadie and Lineweaver plots. *Uncompetitive* inhibition (where I binds to the enzyme-substrate complex only) is indicated on a Lineweaver-Burk plot by parallel lines or on an Eadie plot by lines with a common x intercept. *Noncompetitive* inhibition (where I and S bind simultaneously) is shown on an Eadie plot by parallel lines and on a Lineweaver-Burk plot by a set of lines with a common, negative x intercept.  $K_i$  can be calculated once  $V_{max}$  and  $K_M$  and the type of inhibition are known.

The (2*R*,3*S*)-3-fluoro- and (2*S*,3*R*)-3-methyl- analogues are not expected to undergo reaction beyond formation of the quinonoid intermediate, as there is no 3-*pro*-S proton available. These substrates may therefore undergo abortive transamination reactions to produce PMP (**6**), which can be monitored by UV-Vis studies. The removal of the 3-*pro*-S proton in the (2*R*,3*R*)-3-fluoro- analogue may proceed more quickly than is the case for the natural substrate, due to the electron-withdrawing effects of the fluorine atom. This analogue may therefore prove to be an enzyme-activated irreversible inhibitor of TS. The analogues methylated at C-4 will indicate how constrained the  $\gamma$ -phosphate group is normally in the TS active-site, as they cannot bind effectively if there is too much steric hindrance.

### 1.10.2 Product Inhibition Assays

Inorganic phosphate ( $P_i$ ) is produced as a side-product of the TS reaction (Scheme 1.9) and has been shown to strongly inhibit the enzyme from *Lemna pausicostata* at 1 mmol concentration in a reversible manner. Threonine has also been shown to inhibit the enzyme from *Br. flavum* but only by 15 - 20%. It was considered of value to carry out product inhibition studies on the enzyme from *E. coli*, but malachite green cannot be used to monitor substrate turnover in this case. Therefore it was envisaged that part of *this* work would involve the synthesis of  $[U-^{14}C]$ -(2*S*)-*O*-phosphohomoserine (**32b**, Section 2.1.3). The substrate turnover in the presence of various starting concentrations of  $P_i$  or threonine was to be monitored by quenching the assays after a given time, separating the  $[U-^{14}C]$ -(2*S*,3*R*)-threonine (**33b**) produced from **32b**

via ion-exchange chromatography and determining the quantity of **33b** present in solution by scintillation counting.

### 1.10.3 Further Assessment of the Allosteric Activation of TS by SAM

The effect of SAM on TS from *E. coli* K-12 Tir8 has been investigated by F. Barclay for comparison with its effect on plant TS enzymes.<sup>41</sup> The value of  $K_M$  for TS was almost trebled in the presence of SAM (115.1  $\mu\text{M}$  compared to a value of 43.5  $\mu\text{M}$  without SAM), and the  $V_{\text{max}}$  value was increased almost two-fold (7.34  $\mu\text{M min}^{-1}$  compared to 4.18  $\mu\text{M min}^{-1}$  without SAM). More kinetic data is needed in order to determine the Hill coefficient,  $h$ , which at the upper limit is equal to the number of allosteric binding sites. The value of  $h$  can be determined from the slope of the following equation in the region of 50% saturation:

$$\log \frac{v}{V_{\text{max}} - v} = h \log[S] - \log K$$

It was also hoped that some of the other unanswered questions concerning the TS mechanism (see Section 1.8.3) could be addressed in this work.

### 1.10.4 Structure-Activity Studies on Threonine Synthase

No crystal structure of TS from any source has yet been produced. It was hoped that an X-ray structure of the enzyme complexed with PLP could be produced during this work, either using TS isolated from the Tir8 strain, or from *thrC* cloned into an expression system, such as the pET system (Section 1.11). This would elucidate the active-site geometry and provide information on how substrates bind to the enzyme. In Section 1.3.3 the altered properties of a number of mutant TrS enzymes were described. It was hoped that similar site-directed mutagenesis experiments could be carried out on TS to elucidate the exact role of key residues, such as Lys107 (responsible for binding PLP), which are conserved in all the known microbial TS sequences (Appendix 6). Mutant enzymes were to be expressed in pET vectors.

## 1.11 pET Expression Vectors

### 1.11.1 Introduction<sup>243, 244</sup>

The pET system is one of the most powerful systems to date for the cloning and expression of recombinant proteins in *E. coli*. Expression of target genes in pET plasmids is under the control of strong bacteriophage T7 transcription and translation signals and is therefore only induced by the presence of T7 RNA polymerase provided by the host cell. This polymerase is so selective and active that it can direct almost all the cell's resources towards expression of the target gene. Recombinant protein can amount to up to 50% of total cell protein.

### 1.11.2 T7 RNA Polymerase<sup>245</sup>

The RNA polymerase of bacteriophage T7 has a stringent specificity for its own promoters and active T7 promoters have never been found in DNAs unrelated to T7 DNA.<sup>246</sup> No T7 promoters exist on *E. coli* chromosomal DNA (a natural host for T7). Once an *E. coli* cell becomes infected with T7, the bacteriophage can use its highly selective RNA polymerase to direct gene transcription away from the host's DNA to its own. Transcription by T7 RNA polymerase is very active as well as selective, elongating RNA chains about five times faster than *E. coli* RNA polymerase.<sup>246-248</sup> T7 RNA polymerase also appears not be subject to the factors that cause termination of transcription by *E. coli* RNA polymerase.<sup>249</sup> The one terminator specifically for the polymerase in T7 DNA is also not entirely efficient.<sup>250</sup> Studier and co-workers therefore concluded that T7 RNA polymerase should be capable of efficiently producing complete transcripts from almost any DNA linked to a T7 promoter. Since T7 promoters are never utilized by *E. coli* RNA polymerase the presence of this promoter in the cell should have little effect on host gene expression. However, once T7 RNA polymerase is introduced into the host cell this induces active and selective transcription from the T7 promoter.

Rosenberg and co-workers first reported the construction of pET vectors (plasmid for Expression by T7 RNA polymerase) based on the plasmid pBR322.<sup>251-253</sup> T7 gene 1 is contained in the *expression* host; the plasmid containing both the T7 promoters and terminators either side of the cloning site.

### 1.11.3 Expression of Target Genes<sup>243</sup>

The target genes are initially cloned in hosts which do not contain the T7 RNA polymerase gene. Background expression of the gene product is minimal in the absence of T7 RNA polymerase helping to establish stable recombinants. This is particularly advantageous in cases where the gene product would be toxic to the cell.

For protein expression the recombinant plasmid is transferred to a host containing a chromosomal copy of T7 gene 1. These are lysogens of the bacteriophage  $\lambda$ DE3. Of the two lysogenic strains suitable for gene expression, BL21(DE3) and HMS174(DE3), the former lacks the *ompT* outer membrane protease which can degrade some proteins during purification. Both strains contain the polymerase gene under the control of the inducible *lacUV5* promoter. Expression of target genes can therefore be halted until the growing culture has reached a suitable stage. Upon the addition of the inducer isopropyl- $\beta$ -D-thiogalactoside (IPTG) T7 RNA polymerase is expressed and the gene contained on the plasmid is transcribed. Some pET vectors also contain a T7lac promoter operator system. This means that they contain a *lac* operator sequence just downstream of the T7 promoter, and also the natural promoter and coding sequence for the lac repressor (*lacI*). The latter is oriented so that the T7lac and *lacI* promoters diverge. In DE3 lysogens the *lac* repressor acts at the *lacUV5* promoter to repress background transcription of the T7 RNA polymerase gene by the hosts RNA polymerase. The *lac* repressor also acts at the T7lac promoter to block transcription of the target gene by any T7 RNA polymerase that is made.

### 1.11.4 Vectors Selected For Expressing Wild-Type and Mutant *ThrC*

The three vectors chosen for expressing *thrC*, pET-3a, pET-3b and pET-16b (Fig. 1.17), are translation vectors containing efficient translation initiation signals.<sup>243</sup> All these vectors carry the strong ribosome binding site of the T7 gene 10 protein (the major capsid protein of phage T7). pET-16b also contains the T7lac promoter operator system and produces recombinant protein with a His•Tag™ affinity handle at the N-terminus.<sup>243</sup> This permits one-step purification of expressed proteins by metal chelation chromatography on Ni-NTA resin (see Scheme. 1.18).<sup>254</sup>

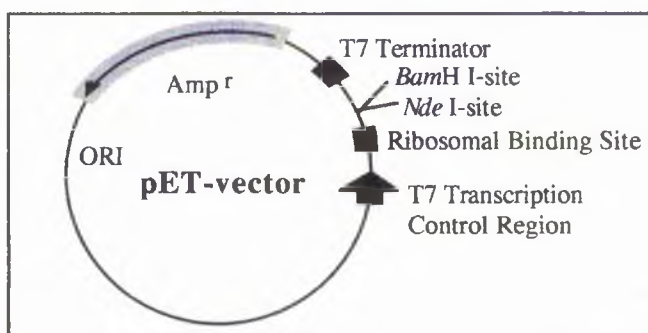
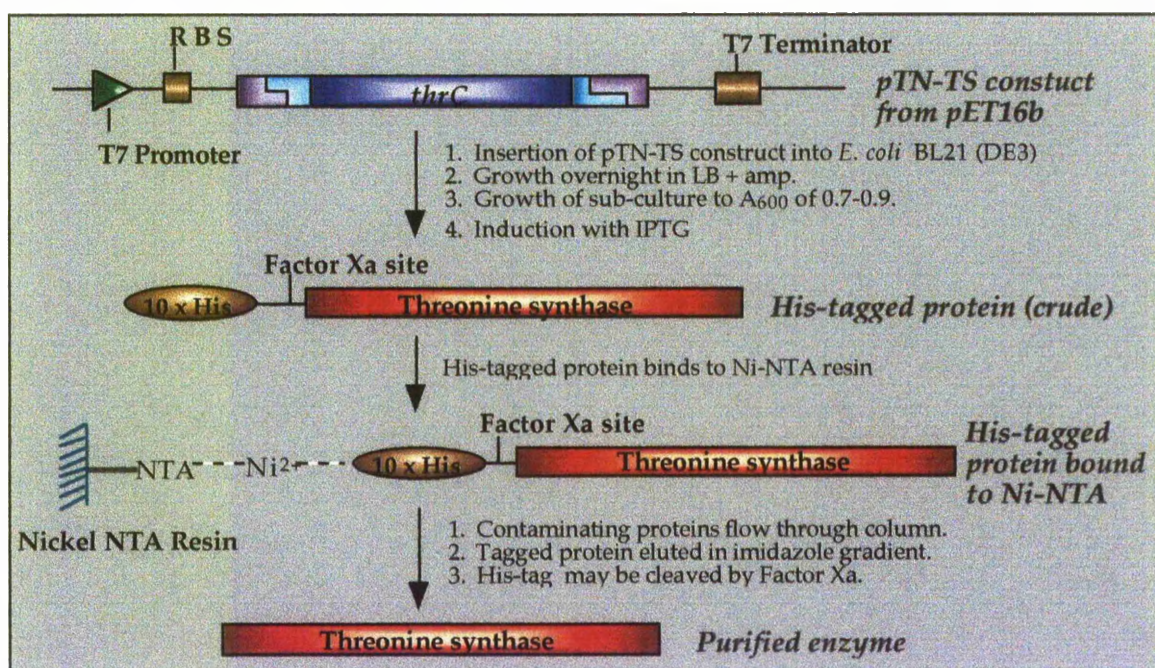


Figure 1.17: General structure of pET-vectors to be used in cloning experiments.



Scheme 1.18: Protocol for the purification of His-tagged TS.

**CHAPTER TWO**  
**RESULTS AND DISCUSSION**

## 2.1 Substrate Synthesis

A protocol for the synthesis of (2*S*)-*O*-phosphohomoserine (**32**) had previously been developed by F. Barclay.<sup>41, 255</sup> This approach was adapted for the solution-phase synthesis of both the unlabelled and [U-<sup>14</sup>C]-labelled substrate in these studies, with a view to continuing the elucidation of the TS reaction mechanism. Several other approaches to the compound had been reported in the literature, but none of these were very efficient, and many of them relied on enzymic synthesis.

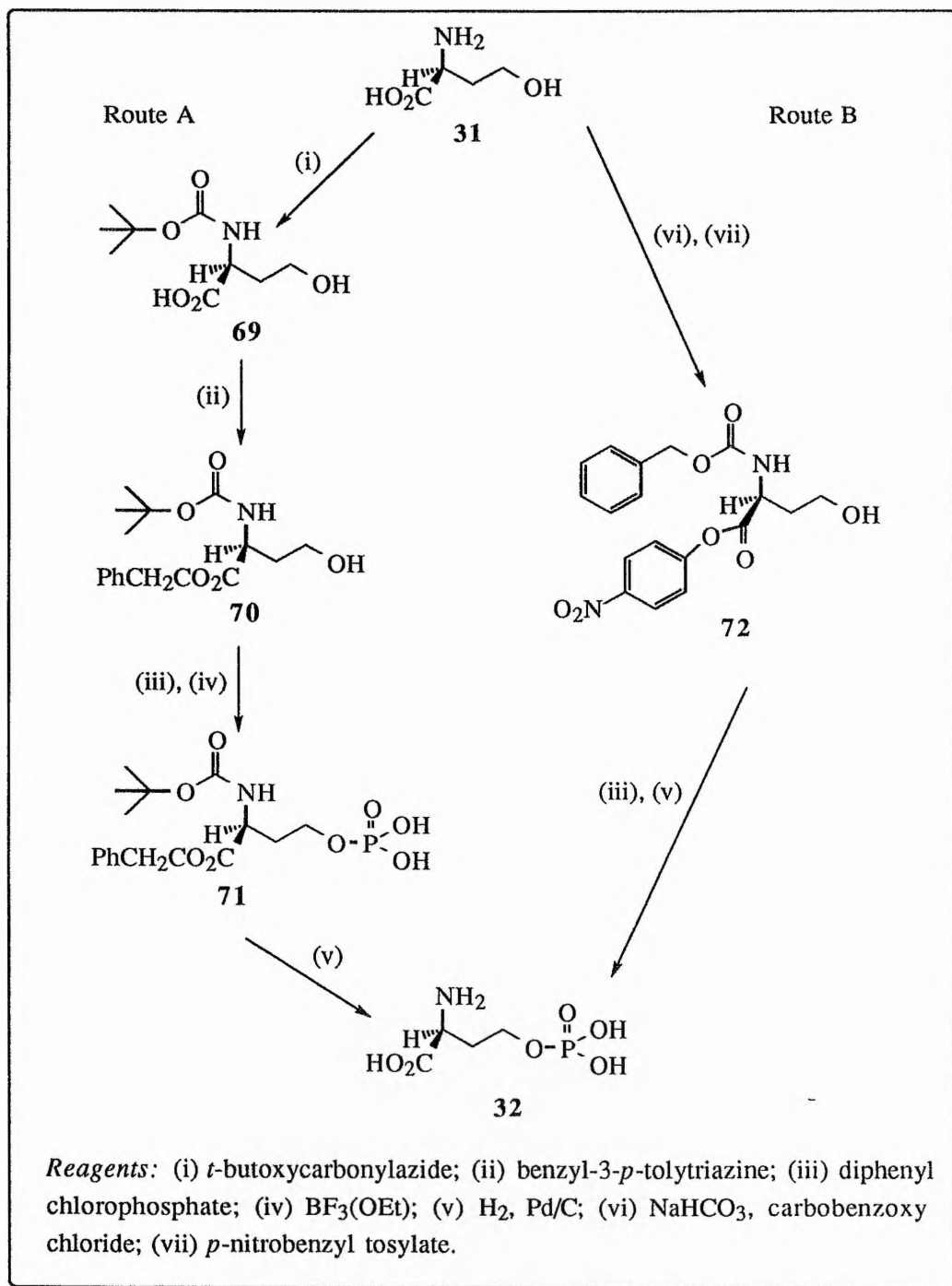
### 2.1.1 A Survey of the Methods for Synthesising (2*S*)-*O*-Phosphohomoserine

Ågren reported the isolation of phosphohomoserine from trichloroacetic acid extracts of *Lactobacillus casei* in 1962, obtaining 55 mg of substrate from 650 g (dry weight) of cells.<sup>256</sup> A few years later Watanabe *et al.* reported the production of phosphohomoserine by the treatment of homoserine with a yeast-derived preparation of homoserine kinase.<sup>257</sup> Skarstedt and Greer also phosphorylated (2*S*)-homoserine (**31**) enzymically, using aspartokinase purified from a threonine-deficient strain of *B. subtilis*.<sup>143</sup> Contaminating phosphate-containing side-products were completely removed by ion-exchange chromatography.

The first *chemical* synthesis of the compound was in 1973 by Fickel and Gilvarg, starting from (2*S*)-homoserine (Scheme 2.1, Route A).<sup>258</sup> The starting material was *N*-protected with *t*-butoxycarbonylazide, esterified with benzyl 3-*p*-tolyltriazine and phosphorylated with diphenylchlorophosphate. The protecting groups were then removed using boron trifluoride and hydrogenolysis. This synthetic route unfortunately led to 20% racemisation of the end product. Schnyder and Rottenberg developed this route further, synthesising the substrate *via*  $\alpha$ -*p*-nitrobenzyl-*N*-benzyloxycarbonyl-(2*S*)-homoserine (**72**) instead, (Scheme 2.1, Route B).<sup>162</sup> This was again phosphorylated with diphenylchlorophosphate, and deprotected by



hydrogenolysis, producing the substrate (**32**) in 49% overall yield. F. Barclay developed the first *biomimetic* synthesis of the substrate, starting from (2*S*)-aspartic acid (**32**).<sup>41, 255</sup>



**Scheme 2.1:** Route A: Fickel and Gilvarg's synthesis of (2*S*)-O-phosphohomoserine, Route B: Schnyder and Rottenberg's method.

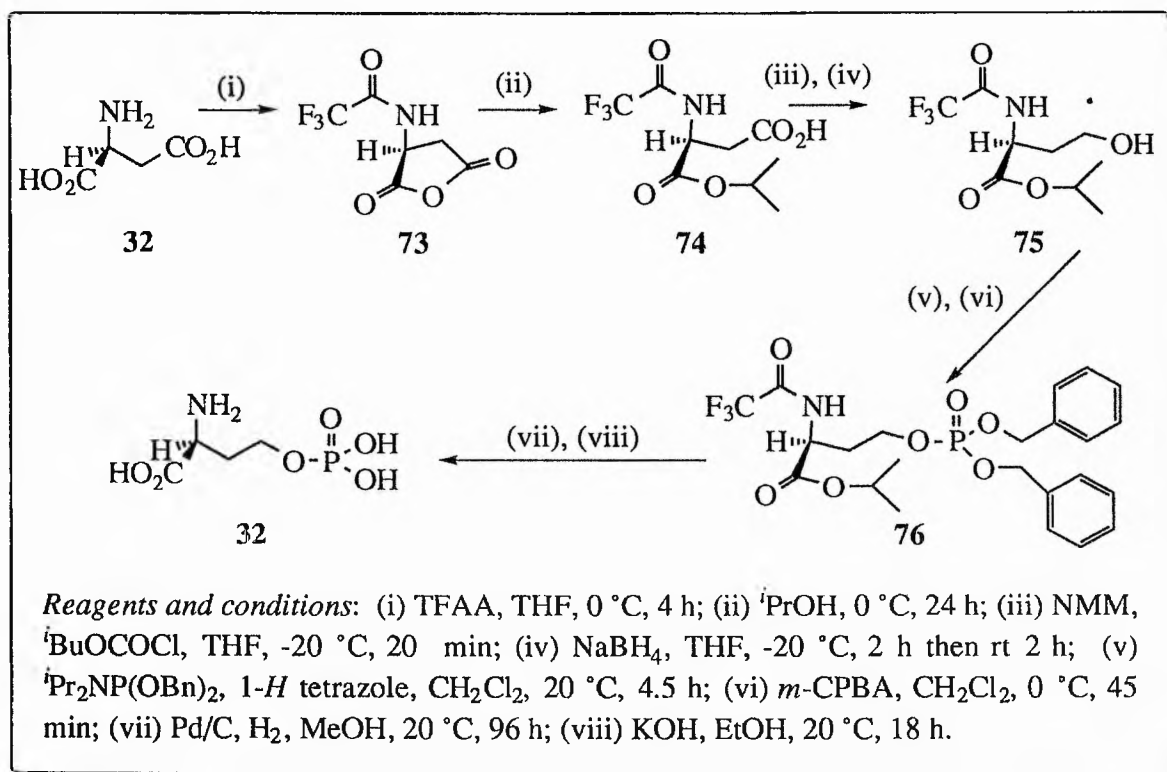
### 2.1.2 The Biomimetic Synthesis of (2*S*)-*O*-phosphohomoserine

The biomimetic synthesis of the substrate developed by F. Barclay<sup>41, 255</sup> was repeated in this study of TS as follows:

(2*S*)-aspartic acid (**28**) was *N*-protected using trifluoroacetic anhydride in tetrahydrofuran, giving the *N*-trifluoroacetyl cyclic anhydride (**73**) in excellent yield (99%, Scheme 2.3). Other *N*-protecting groups tried previously in this laboratory by F. Barclay such as *t*-butoxycarbonyl and benzyloxycarbonyl gave much poorer yields.<sup>41</sup> The IR spectrum of **73** confirmed the presence of anhydride carbonyl groups {1886 m (C=O, anhydride) and 1722 s (C=O, anhydride)} and of a secondary amide carbonyl group {1556 m (C=O, 2° amide)}, as well as various peaks corresponding to C-F bonds. The <sup>13</sup>C-NMR spectrum confirmed the presence of two new anhydride carbonyl groups {169.28 (αC=O) and 167.61 (βC=O)} and two quartets corresponding to the *N*-trifluoroacetyl protecting group {157.43 (q,  $J_{F_3C,C}$  39, F<sub>3</sub>CCO) and 115.34 (q,  $J_{F,C}$  285, F<sub>3</sub>C)}.

The anhydride was regioselectively opened using isopropanol to give the α-isopropyl ester β-acid (**74**) in quantitative yield after one recrystallisation from ether and light petroleum. Other alcohols used by F. Barclay were less selective in opening the anhydride to gain the α-ester due to their higher nucleophilicity. The <sup>1</sup>H-NMR spectrum of **74** produced new signals corresponding to the α-isopropyl ester {5.11 (1 H, sept,  $J_{CH_3,H}$  6.25, <sup>*i*</sup>PrCH) and 1.26 and 1.23 (6 H, d,  $J_{CH_3,H}$  6.25, 2 x <sup>*i*</sup>PrCH<sub>3</sub>)} and IR spectroscopy confirmed the presence of ester and acid carbonyl groups {1738 s (C=O, ester), 1709 s (C=O, carboxylic)}. In the <sup>13</sup>C-NMR spectrum the signals for the carbonyl groups had shifted downfield {175.54 (βCO<sub>2</sub>H) and 168.56 (αCO<sub>2</sub><sup>*i*</sup>Pr)}.

Reduction of the free β-acid was performed using mixed anhydride methodology. The mixed β-aspartic isobutylcarbonic anhydride was formed from the acid by treatment with 1.1 equivalents each of *N*-methyl morpholine and isobutylchloroformate at -20 °C. The

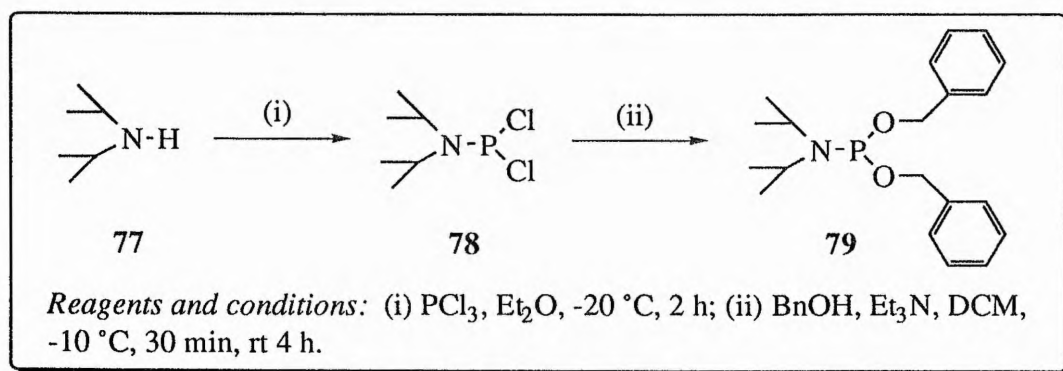


**Scheme 2.2:** The biomimetic synthesis of (2*S*)-*O*-phosphohomoserine (**32**) from (2*S*)-aspartic acid (**28**).

precipitated salts were then removed, and the mixed anhydride treated with sodium borohydride in THF. This is a problematic reaction, and yields varied widely depending on the equivalents of NMM and IBCF and of sodium borohydride used. *N*-selectride was also tested as a reducing agent for this reaction by G. Allan (unpublished results), but although this produced a much cleaner product, the yield was significantly lower (20% crude). The results using this reducing agent were also not easily reproducible. The best yield of  $\alpha$ -isopropyl-*N*-trifluoroacetyl-(2*S*)-homoserine (**75**) obtained after purification by flash silica chromatography, eluting with 3% ethanol in DCM gave **75** as a yellow oil in a yield of 35%. The <sup>13</sup>C-NMR spectrum revealed the loss of the signal at 175.54 ppm corresponding to the  $\beta$ CO<sub>2</sub>H group of **74**, and the appearance of a new signal at 58.58 ppm, corresponding to the  $\gamma$ CH<sub>2</sub>OH of **75**. The carboxylic C=O signal (1709) had disappeared from the IR spectrum. F. Barclay reported that the deprotection of this compound to give (2*S*)-homoserine (**31**) showed

that no racemisation had occurred even during saponification, the  $[\alpha]_D$  being the same as that obtained for an authentic sample of **31**.

Phosphorylation of **75** was achieved using a reagent which was first developed for the synthesis of *myo*-inositol phosphates; *N,N*-diisopropyl dibenzyl phosphoramidite (**79**, Scheme 2.3).<sup>259</sup> This compound is produced from *N,N*-diisopropyl dichlorophosphoramidite (**78**, Scheme 2.3). **78** was synthesised by the addition of two equivalents of diisopropylamine (**77**) to phosphorous trichloride in diethyl ether under a nitrogen atmosphere (HRMS: Found:  $M^+$  201.0235, Calc. for  $C_6H_{14}NPCl_2$  201.0241). Although this reaction worked well in some cases (up to 81% yield was achieved) it was very moisture sensitive, and the product had a tendency to hydrolyse whilst the contaminating hydrochloride salts were being removed by filtration, prior to solvent removal and distillation of the product. Although this problem was originally overcome by the use of cannula filtration under nitrogen to syphon off the product in solution, later experiments to reproduce this method were unsuccessful, and a new phosphorylating reagent had to be selected (see Section 2.6). G. Allan has since successfully performed this reaction under argon, suggesting that the nitrogen source used here contained moisture which caused the product to hydrolyse.



**Scheme 2.3:** Synthesis of *N,N*-diisopropyl(bis)benzyl phosphoramidate (**79**).

Esterification of **78** via nucleophilic displacement of chloride with excess benzyl alcohol in the presence of triethylamine yielded the phosphoramidate (**79**), which was purified by silica

column chromatography using 25% ethyl acetate-light petroleum as the eluant, in 84% yield. (HRMS: Found:  $[M + H]^+$  346.1600, Calc. for  $C_{20}H_{29}NO_2P$  346.1936). The  $^{13}C$ -NMR spectrum showed the presence of aromatic carbons, some of which were coupled to phosphorous: 138.11 and 138.01 (2 C, d,  $J$  4.89, Ar-C quaternary), 128.31, 127.41 and 126.87 (10 C, Ar-CH) and 64.41 and 64.19 (2 C, d,  $J_{CH_2,OP}$  10.88, 2 x  $OCH_2Ph$ ).

The alcohol, **75** was treated with **79** and the intermediate was oxidised with *m*-CPBA as described by Yu and coworkers.<sup>259</sup> Only one equivalent of phosphoramidate was necessary to produce the phosphate triester (**76**) in 63% yield. The  $^{13}C$ -NMR spectrum revealed two new signals corresponding to the benzyl groups {127.93 ppm (10 x Ar-C) and 69.46 ppm ( $-CH_2Ph$ )} and  $^{31}P$ -NMR spectroscopy produced a single signal at -1.07 ppm. Deprotection of this to give the substrate (**32**) was achieved by catalytic hydrogenolysis of the benzyl protecting groups and base-catalysed hydrolysis of the *N*-trifluoroacetyl and  $\alpha$ -isopropyl groups. This proceeded in excellent yield (91%) after purification by cation-exchange chromatography (Dowex 50-W,  $H^+$ ), eluting with water. mp 167-169 °C (lit.,<sup>41</sup> 170 °C);  $\delta_H$ (200 MHz;  $^2H_2O$ ) 3.97 (H, dq,  $J_{AX}$  7.6,  $J_{BX}$  4.8,  $H^\alpha$ ), 3.93 (H, dq,  $J_{CX}$  11.6,  $J_{DX}$  5.4,  $H^\gamma$ ) and 2.09 (2 H, m,  $H^\beta$ );  $\delta_C$ (50.31 MHz;  $^2H_2O$ ) 175.09 ( $\alpha CO_2H$ ), 64.67 ( $\alpha C$ ), 54.19 and 54.03 (d,  $J_{C,OP}$  8.14,  $\gamma C$ ) and 33.40 and 33.25 (d,  $J_{C,COP}$  7.63,  $\beta C$ );  $\delta_P$ (121.42 MHz;  $^2H_2O$ ) 0.59;  $m/z$  (EI) 200 (5%,  $[M + H]^+$ ), 199 (5,  $M^+$ ), 181 (9,  $M - H_2O]^+$ ), 119 (17,  $[M - H_3PO_3]^+$ ) and 104 (23,  $[M - PO_4]^+$ ). The overall yield of **32** achieved here by this method was 20%, although F. Barclay reported a yield of 66% from (2*S*)-aspartic acid (**28**),<sup>41</sup> an improvement on the previous best yield of 49% reported by Schnyder and Rottenberg.<sup>162</sup>

### 2.1.3 Synthesis of $[U-^{14}C]$ -(2*S*)-*O*-Phosphohomoserine

Previous methods for studying the kinetics of the reactions of TS from *E. coli* in this group involved the use of malachite green assays to detect inorganic phosphate, which is produced as a side-product.<sup>238-241</sup> Lanzetta *et al.* first reported the detection of nanomolar quantities of

phosphate by measuring the absorption at 660 nm of a sample treated with malachite-molybdate complex.<sup>241</sup> Although this has proved a very accurate method for monitoring the progress of the TS reaction, it precludes product inhibition studies.

In order to avoid using malachite green as an indicator of product concentration, an assay was designed using [U-<sup>14</sup>C]-labelled substrate, **32b**. In the case of malachite green assays, the colorimetric dye was also used to quench the reaction. For these assays it was envisaged that the reactions would be quenched using trichloroacetic acid as reported by Rognes.<sup>132</sup> The product, [U-<sup>14</sup>C]-(2*S*,3*R*)-threonine (**33b**), could then be isolated by passing the reaction mixture down an anion-exchange column (*e.g.* Dowex 50-W, Cl<sup>-</sup>), which would bind unreacted labelled phosphohomoserine. The amount of radioactivity still present in the eluted solution could then be determined by scintillation counting to monitor turnover of the substrate. To attempt such an assay **32b** was synthesised from [U-<sup>14</sup>C]-(2*S*)-aspartic acid (**28b**), using the protocol described for the synthesis of the unlabelled substrate, in 32% overall yield.

## 2.2 The Growth of Threonine Synthase

In her work on threonine synthase, F. Barclay used an *E. coli* constitutive mutant for the enzymes encoded by the *thr* operon.<sup>165</sup> The use of this mutant, in which the enzymes of the *thr* operon are continually expressed, was necessary because threonine synthase is normally only expressed under conditions where threonine and isoleucine become limiting for growth, as described in Section 1.7.

Attempts to isolate the enzyme from the *E. coli* strain K-12 Tir8 in this work, using the protocol developed by F. Barclay, were unsuccessful. No TS activity could be detected after just a few purification steps and malachite green assays on crude cell-free extracts were not possible due to the high background levels of phosphate. The radiolabelled substrate was prohibitively expensive for use during these enzyme purification experiments. The reason for the lack of enzyme activity was not fully investigated at the time, but it was thought possible that the mutant had reverted to wild-type. It has been reported that such constitutive mutants do revert readily to wild-type and that wild-type cells will rapidly overtake a growing culture.<sup>174</sup> It is also possible that excessive cell lysis or lengthy purification procedures may have inactivated the enzyme. After a few abortive attempts to purify the enzyme from Tir8 cells grown from fresh frozen glycerol stocks, attempts were made to clone the gene for threonine synthase (*thrC*) into a pET expression vector.

## 2.3 Cloning of *ThrC* and the Gene for $\beta$ -Methylaspartase

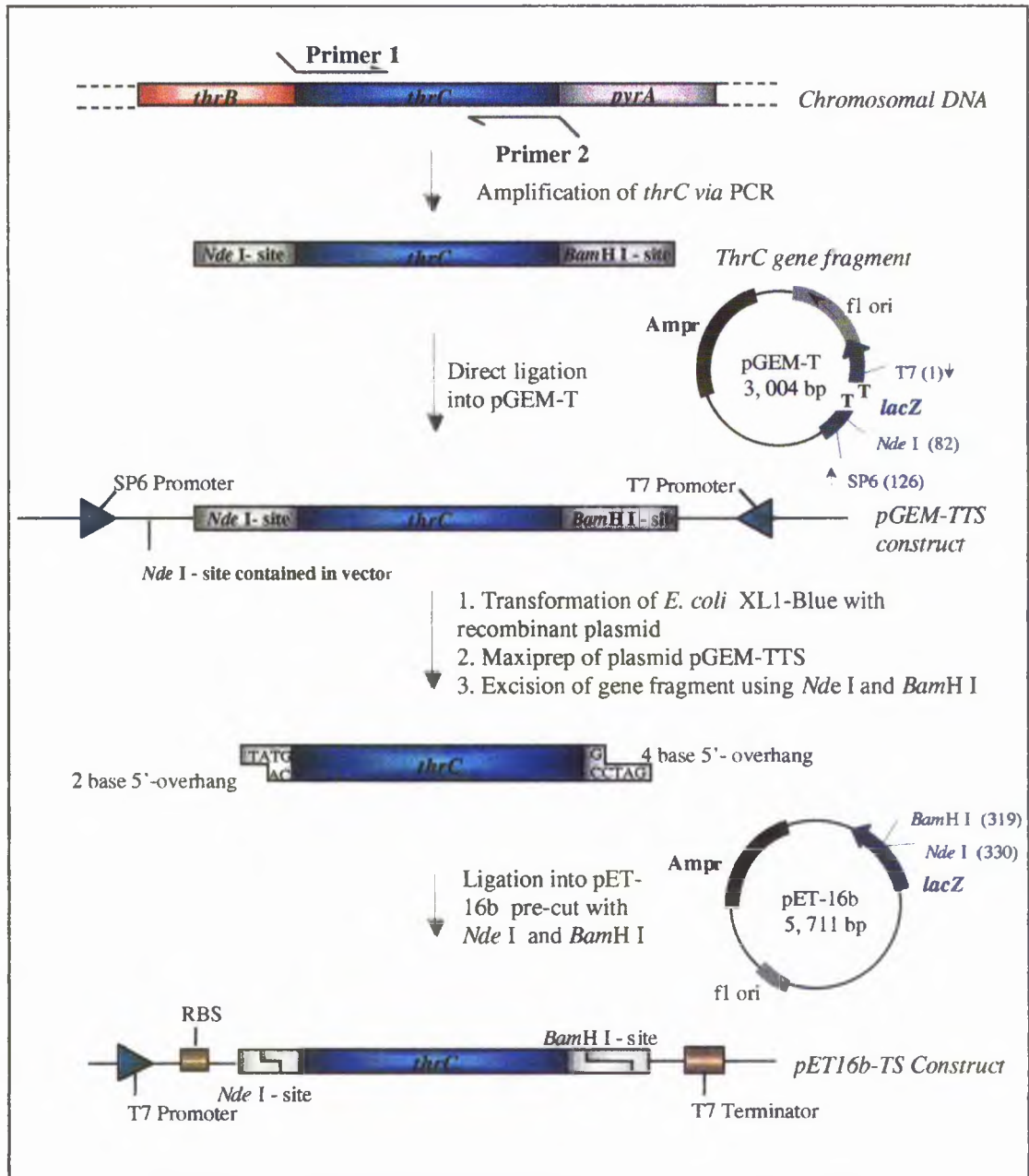
### 2.3.1 The General Cloning Strategy

Experiments were undertaken to attempt the cloning of both *thrC* and the gene for  $\beta$ -methylaspartase into expression vectors. The latter had already been cloned into a different expression system, which was no longer producing sufficient enzyme.

Amplified genes were not cloned directly into pET vectors, because previous research in the group had indicated that the ligation of gene fragments with these vectors was problematic. The intention was that each gene be cloned into a pGEM cloning vector so that large quantities of DNA could be produced relatively easily, removing the need for repeated PCR experiments. The strategy for cloning *thrC* and the gene for  $\beta$ -methylaspartase into a pET™ expression vector was therefore as follows (see Scheme 2.4):

- Amplification of *thrC* (from *E. coli* chromosomal DNA) and of the gene for  $\beta$ -methylaspartase [from plasmid pSG4 contained in *E. coli* TG1 (*K-12*  $\Delta$  [*lac-pro*] *supE thi hsdR5 / F'-traD36 proA<sup>+</sup> B<sup>+</sup> lac*]  $\phi$  *lacZ*  $\Delta$  *M15* )],<sup>260</sup> adding specified restriction sites at each terminus. These were chosen such that the restricted gene fragment could be ligated to a similarly restricted vector.
- Insertion of the gene fragment into a cloning vector; in this case the pGEM™ series of vectors was chosen.
- Maxiprep purification of the recombinant plasmid produced in the above step.
- Excision of the gene fragment using the restriction sites at each terminus of the gene.
- Subcloning of the gene fragment into a pET expression vector cut with the same enzymes.





**Scheme 2.4:** Final cloning procedure adopted for *thrC*. A similar strategy was applied to the gene for  $\beta$ -methylaspartase and other expression vectors were also used.

## 2.3.2 Cloning of *ThrC*

### 2.3.2.1 Amplification of *ThrC*

#### 2.3.2.1.1 Isolation of Chromosomal DNA.

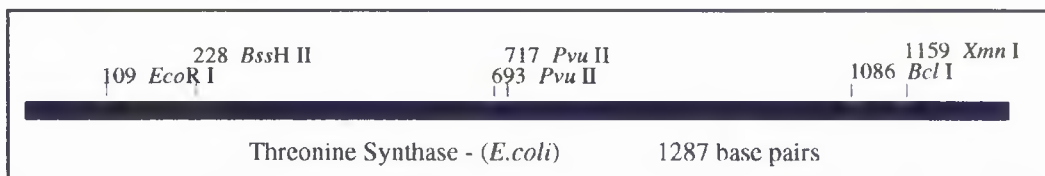
Preliminary attempts to isolate chromosomal DNA from *E. coli* JM109 using freeze-thaw and sonication methods for lysing cells were unsuccessful. The freeze-thaw method appeared not to be harsh enough to disrupt the cell walls and sonication led to fragmentation of the DNA making it unsuitable for use in PCR experiments. However, in an experiment for the minipreparation of *pSG4* for the cloning of  $\beta$ -methylaspartase a small amount of chromosomal DNA also remained, which was a suitable template for amplifying *thrC*.

DNA was isolated using a variation of the alkaline lysis method of miniprep plasmid purification.<sup>261, 262</sup> All minipreps were carried out in this manner. Cells were harvested from an overnight culture in LB and ampicillin and lysed with lysozyme in STET buffer. Cells were then incubated on ice with SDS to denature bacterial proteins and potassium acetate to precipitate insoluble cell debris, such as genomic DNA and protein, leaving the plasmid in solution. The supernatant was treated with RNase. The aqueous DNA solution was further purified by repeated extraction with phenol/chloroform. DNA was precipitated with cold absolute ethanol, allowed to dry briefly and stored in water at -20 °C.

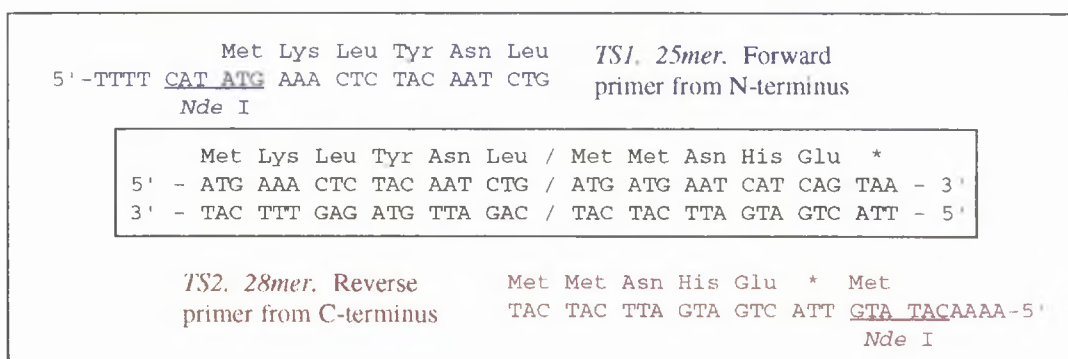
#### 2.3.2.1.2 Polymerase Chain Reaction on *ThrC*

The ligation of cohesive termini is known to be more efficient than blunt-end ligation, as indicated by the shorter incubation time and much smaller amount of ligase (50 times less to achieve the same extent of ligation) necessary for the former.<sup>263, 264</sup> The pET vectors chosen for the expression of *thrC* each contain a unique *Nde* I-site (see Appendix 9 and Fig. 1.17) in the multi-cloning region. As *thrC* does not contain this site (Fig. 2.1 and Appendix 8), the first primers for PCR experiments were designed to produce *Nde* I-sites on either side of the

termini of *thrC* (Fig. 2.2). The gene was amplified via 20 cycles of PCR (annealing temp. 42 °C, polymerisation step 2 min) using the AmpliTaq™ form of *Taq* DNA polymerase. This thermophilic polymerase lacks a 3' - 5' exonuclease proofreading activity, and therefore has a high error rate per nucleotide polymerized ( $110 \times 10^{-4}$  at 10mM MgCl<sub>2</sub> and 1mM dNTPs).<sup>265</sup> This results in a base substitution rate of one in 9 000 and a frameshift rate of one in 41 000 nucleotides.<sup>265</sup> *Taq* was, however, readily available and had previously provided satisfactory results in the group, so it was used again during these experiments. The enzyme *does* possess a 5' - 3' structure specific nuclease activity, which means that it will degrade any DNA annealed to the template strand while adding on bases to the 3' end of another annealed primer.<sup>264</sup> In addition the presence of a 3' - 5' exonuclease activity can sometimes cause some degradation of primers from the 3' end, although this has only been observed to be a problem with primers of a length greater than 30 bases.<sup>264</sup>

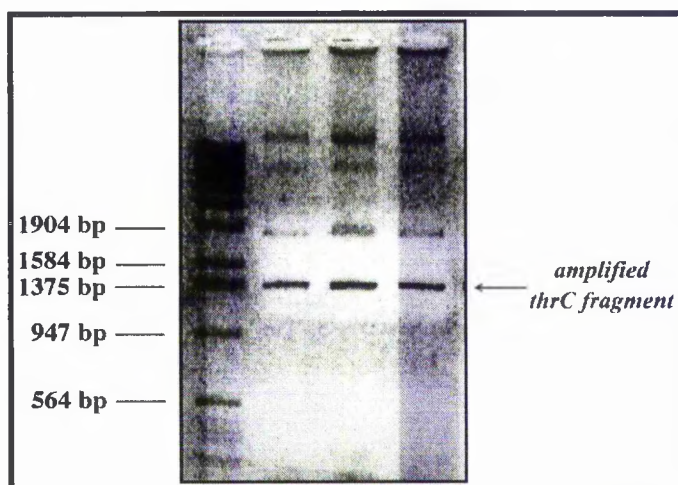


**Figure 2.1:** Graphic map of the restriction sites on *thrC* used during cloning experiments.



**Figure 2.2:** DNA sequences of the primers used in the first PCR reaction. *Nde I*-sites are underlined. The middle sequences show the two termini of *thrC*.

PCR produced a DNA fragment of approximately 1300 bp, consistent with an amplified gene fragment (1287 bp) with 17 additional bases added at the termini by primers. This was determined by agarose gel electrophoresis, staining with the fluorescent dye ethidium bromide and using a 1 kb DNA ladder molecular weight marker for comparison (Fig. 2.3). The amplified gene was gel-isolated by running the band of DNA into a section of low-melting agarose, which was cut out and incubated at 70°C in phenol/acetate. The DNA was then purified and extracted from the aqueous layer in the usual manner.



**Figure 2.3:** 1% agarose gel electrophoresis of amplified *thrC* (left lane contains a  $\lambda$  DNA ladder).

### 2.3.2.2 Attempted Ligation of *ThrC* Into Cloning Vectors.

For all cloning work, competent *E. coli* cells were prepared using the  $\text{CaCl}_2$  method,<sup>266, 267</sup> and blue-white screening for  $\alpha$ -complementation was used to test for the presence of putative recombinant plasmids in individual colonies.<sup>266, 268, 269</sup> All maxipreps of recombinant plasmids were performed using the QIAGEN™ maxiprep plasmid purification protocol, using strain XL1-Blue. This slow-growing strain of *E. coli* is known to produce a very high quality of DNA, which works well for sequencing. Other strains (TG1 and the JM100 series for example) release large amounts of carbohydrate during lysis and have high endonuclease activities.

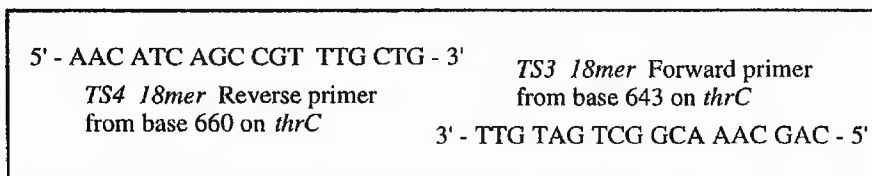
The restriction endonuclease, *Nde* I, has a very short half-life (15 minutes) and can often be inefficient at digesting DNA.<sup>264</sup> *Nde* I is also particularly sensitive to impurities present in some DNA preparations; for example, DNA purified by standard miniprep procedures is reportedly cleaved at lower rates.<sup>264</sup> Some ligations carried out prior to the results reported here, using *Nde* I-restricted DNA, have produced a high vector background, which may in part be attributed to this inefficiency. Remnants of undigested plasmid DNA are frequently undetected by routine agarose gel electrophoresis and this can lead to a significant proportion of cells transfected with the undigested vector as opposed to recombinants.<sup>263</sup> In addition to having a short half-life, *Nde* I has been reported to require a segment of at least 8 DNA bases before the *Nde* I-site in order to function correctly.<sup>264</sup> As the *thrC* PCR fragment had been designed with just 4 bases before the restriction site at each terminus (Fig. 2.2) it was thought that the enzyme would not, in any case, cleave the termini of the PCR product. Some apparent disruption of the *lacZ* region of cloning vectors after *Nde* I digestion has also been noted, causing even self-ligated plasmids to produce white colonies. For these reasons the use of *Nde* I was avoided in the preliminary cloning work in favour of blunt-end ligations, in spite of their reduced efficiency.

pGEM-3Zf(+) (3199 bp, Appendix 9) contains a unique site for *Sma* I in its multi-cloning region, an enzyme which produces blunt-ended DNA fragments. The first experiments involving *Sma* I-digested vector and the *thrC* PCR fragment (uncut) produced few white colonies and control experiments indicated a significant background of self-ligated vectors with disrupted *lacZ* regions. In order to decrease the vector background in ligation experiments, the 5'-phosphate groups on the linearised vector could have been removed using calf intestinal alkaline phosphatase (CIAP). However, the use of CIAP in previous ligation experiments with other vectors had failed to produce successful recombinant plasmids. CIAP is very difficult to remove from DNA and the continued presence of the enzyme will impede subsequent ligation reactions. Even if great care is taken to remove CIAP, organic contaminants or particles of agarose from DNA preparations, any of these may contribute to the failure of ligation reactions. The use of CIAP in this case would also have necessitated

use of T4 polynucleotide kinase to add 5'-phosphate groups onto the PCR product and this was considered too inefficient. Increasing the ratio of PCR product to vector did produce several putative recombinants, of which three were selected for maxiprep plasmid purification. These plasmid preparations were subjected to restriction analyses using the enzymes *Bam*H I, *Eco*R I, *Hpa* I, *Hind* III, and double restrictions using *Bam*H I and *Kpn* I, *Eco*R I and *Hind* III, *Eco*R I and *Bam*H I and *Hind* III and *Bam*H I. Only the restriction using *Bam*H I produced the predicted size of DNA fragment (corresponding to the linearised recombinant plasmid). This suggested a possible mutation of the gene insert or that the wrong gene or segment of chromosomal DNA had been amplified during the original PCR experiment.

### 2.3.2.3 Overlap Polymerase Chain Reaction Experiments

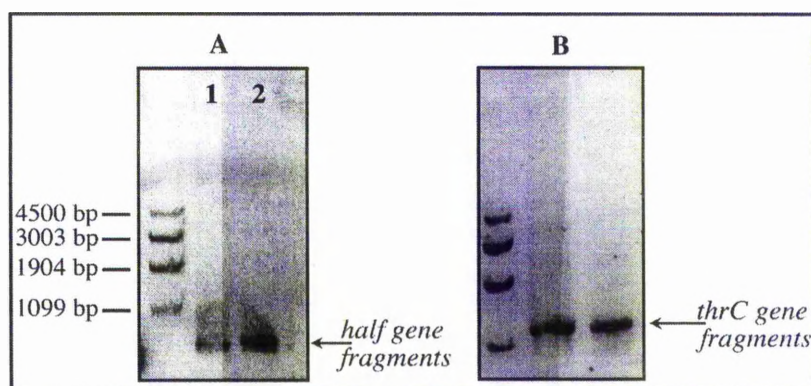
To quickly confirm that the original PCR experiments had produced the desired *thrC* fragment, overlap PCR experiments were carried out using the original chromosomal DNA preparation as a template. Two experiments were performed to produce half-gene fragments originating from either terminus of the gene. This was achieved using the original external primers (Fig. 2.2) and two new primers complementary to the centre of the gene (Fig. 2.4). The two half-gene fragments were then subjected to overlap PCR in a reaction containing the half-gene fragments as templates and the outside primers TS1 and TS2.



**Figure 2.4:** Primers complementary to the centre of *thrC*.

Both sets of experiments successfully produced DNA fragments of the correct apparent sizes as visualised on agarose gels (Fig. 2.5). The new PCR gene fragment was purified more stringently on this occasion, using Wizard Prep™ DNA purification columns, to rule out the

possibility of contamination by agarose or organic solvents, which could have been a contributing factor to the previous failure of ligation experiments.

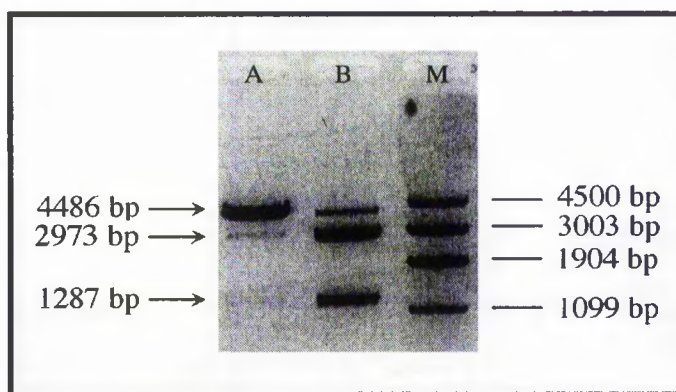


**Figure 2.5:** Results of overlap PCR experiments on *thrC* (1287 bp). Gel A shows the two different half-gene fragments produced from TS1 and TS4 (lane 1) or TS2 and TS3 (lane 2). Gel B shows the results of PCR reactions with the half-gene fragments and forward and reverse oligonucleotides TS1 and TS2.

#### 2.3.2.4 Ligation of *ThrC* into pGEM-T

Once it was confirmed that *thrC* had been amplified in PCR experiments, a new vector, pGEM-T (see Scheme 2.4) was chosen for ligation experiments. pGEM-T is formed by linearising pGEM-5Zf(+) with *EcoR* V and has one thymidine overhang added at each end. These single thymidine overhangs greatly improve the efficiency of ligation of PCR products because of the nontemplate dependent addition of a single deoxyadenosine to the 3' end of PCR product by thermostable polymerases. *Taq* is known to produce predominantly two kinds of ends on DNA fragments, some of which are blunt ends, some of which are single base 3' overhangs.<sup>264</sup> Proofreading polymerases, such as *Vent*, on the other hand, generate 95% blunt-ended fragments.<sup>264</sup> If the proportion of blunt-ended fragments generated by *Taq* was too low, this may account for the apparent failure of blunt-end ligations attempted between *thrC* and pGEM-3Zf(+). pGEM-T is also supplied with a control sample of DNA of known concentration for inserting into the vector, giving an accurate guide to the

overall efficiency of ligations and transformations using the pGEM-T system. Ligations were carried out using various ratios of insert: vector and incubated at room temperature overnight, as summarised in Table 2.1. Competent JM83 were used for transformations. Colonies taken from experiment 4 produced a number of putative recombinants which were subjected to restriction analyses using *Apa* I, *Bss*H II, *Eco*R I, *Sca* I, *Ssp* I and *Xmn* I. Two of these plasmids produced the predicted results with all enzymes. Maxiprep plasmid preparations were made of these two recombinants (denoted pGEM-TTS<sub>A</sub> and pGEM-TTS<sub>B</sub>) and restricted with *Nde* I (expected fragments 31, 1287 and 2973 bp). This indicated that pGEM-TTS<sub>A</sub> had lost one of its *Nde* I-sites as no *thrC* PCR fragment was excised (Fig. 2.6).



**Figure 2.6:** *Nde* I-digest of pGEM-TTS<sub>A</sub> and pGEM-TTS<sub>B</sub>.

In previous ligation experiments the digested vectors had been purified by phenol/chloroform extraction and ethanol precipitation, rather than gel-isolation. Although this should have removed any enzyme and buffer components from the DNA solution, undigested vectors may have remained, which are smaller and more likely to be taken up by competent cells. As pGEM-T is provided in a linear form, such vector background is eliminated. Linear DNA is taken up 50 times more slowly than supercoiled circular DNA molecules by *E. coli* cells.<sup>263</sup>

An attempt was made to sequence the gene from pGEM-TTS<sub>B</sub>, the plasmid which had retained both *Nde* I-sites, and from which the gene could therefore be subcloned, using the fmol™ cycle sequencing protocol. SP6 and T7 primers (Fig. 2.7) were used for this purpose as these promoter sites are found on all pGEM™ plasmids (Appendix 9). A successful



sequence was obtained for just the first 317 bases at the C-terminus of the gene (slightly downstream of the SP6 promoter). This confirmed that the C-terminal *Nde* I-site was still intact.

**Table 2.1:** The results of ligations using *thrC* and *pGEM-T*.

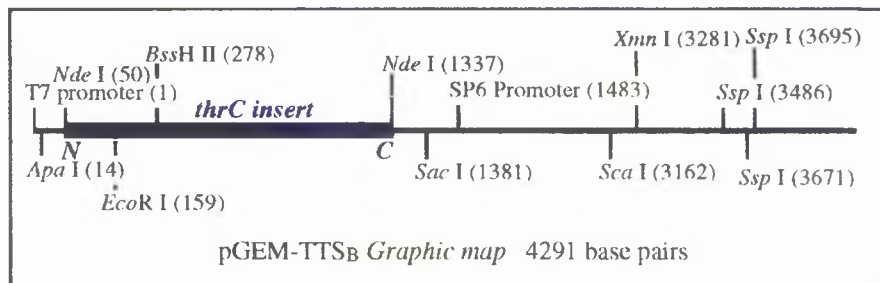
No.	Ratio of Insert : Vector	Amount of Transformed Cells per Plate	Number of Colonies	% Whites
A	Vector + control insert	-	-	45%
B	Vector only control	100µl	10	10%
1	1:1	50µl	20	40%
2	1:1	100µl	50	20%
3	2:1	50µl	20	0%
<b>4</b>	<b>2:1</b>	<b>100µl</b>	<b>50</b>	<b>20%</b>
5	4:1	50µl	20	15%
6	4:1	100µl	50	60%
7	20:1	50µl	15	80%
8	20:1	100µl	25	60%
10	200:1	100µl	10	70%
9	200:1	50µl	2	100%

Highlighted row indicates ligation experiment from which recombinant plasmids were isolated.

**SP6 Promoter Primer:** 5' - d(GATTTAGGTGACACTATAG) - 3'

**T7 Promoter Primer:** 5' - d(TAATACGACTCACTATAGGG) - 3'

**Figure 2.7:** SP6 and T7 primer sequences



**Figure 2.8:** Orientation of *thrC* in pGEM-TTS<sub>B</sub>. Restriction sites used in analysis of the construct are also shown.

### 2.3.2.5 Attempted Ligation of *ThrC* Into Expression Vectors

#### 2.3.2.5.1 pET-3a and pET-16b

The recombinant *thrC* gene was excised from maxipreps of pGEM-TTS<sub>B</sub> using *Nde* I and purified using the phenol/acetate method of gel-isolation described earlier, or the Wizard™ Prep DNA Purification System, according to the manufacturer's instructions. The pET-vectors named above were restricted with *Nde* I, gel-purified and treated with CIAP prior to ligations to try to eliminate the self-ligation of plasmids. Although numerous ligations were attempted with varying ratios of insert:vector, and a number of white colonies were obtained when selecting for the *lac*<sup>+</sup> phenotype, no recombinants were isolated. Unphosphatased linear plasmids also failed to produce successful recombinants, indicating that CIAP was not the sole causative factor in the failure of ligation experiments. Various putative recombinants failed to produce the expected fragments on digestion with *Nde* I and *EcoR* I. As many of these produced fragments corresponding to the linear vector, but not an excised *thrC* fragment, this suggested that the plasmids were self-ligating, as had been observed earlier in ligations experiments involving cloning vectors.

#### 2.3.2.5.2 pKK223-3

Once again, in an attempt to avoid using *Nde* I where possible, in case the enzyme was contributing to the failure of ligations, blunt-end ligation of the excised *thrC* fragment into an

expression vector was attempted. The vector chosen for these experiments, pKK223-3, contains a *Sma I*-site in the multiple-cloning region (Appendix 9).

The pKK223-3 vector contains the strong *tac* promoter, which in an appropriate host is regulated by the *lac* repressor. Expression of gene inserts can be induced by the use of IPTG as is the case for pET-vectors. Immediately downstream of the *tac* promoter is the *pUC8* multiple-cloning site (MCS) and the strong *rrn* ribosomal terminator which stabilises this vector-host system by inhibiting transcription initiated from the *tac* promoter in the parent plasmid. The linearised vector was dephosphatased and purified in the usual manner to remove CIAP. The *thrC* gene fragment was excised from pGEM-TTS<sub>B</sub> using *Nde I* and incubated with T4 DNA polymerase to flush the 5'-protruding ends with dNTPs. Ligations (at room temperature) and transformations were carried out in-gel using SeaPlaque™ GTG agarose gel in low EDTA TAE buffer, as recommended by the manufacturer. This was to prevent any loss of the *thrC* fragment during purification procedures.. After two attempts at using this protocol without success a new strategy for cloning was proposed.

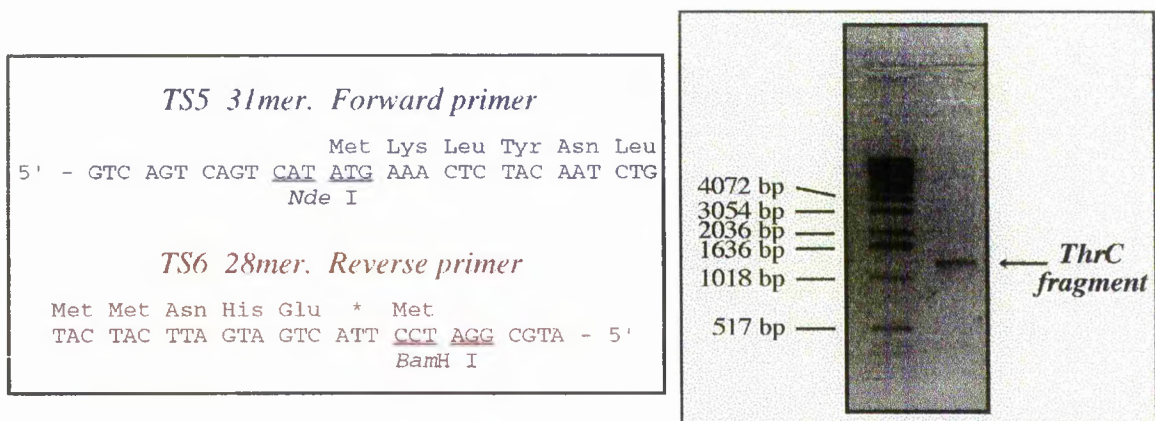
#### 2.3.2.6 New PCR Fragment

For some reason the ligation of gene fragments with *Nde I*-sites at each terminus into pET-vectors is very problematic. This may be due to *Nde I* non-specifically digesting the ends of restricted DNA, as observed by the disruption of the *lacZ* region of cloning vectors digested with the enzyme. It may also be due, in part, to the inefficiency of the enzyme in restricting DNA. In addition, the orientation of such *thrC* fragments in pET-vectors cannot be controlled, further decreasing the efficiency of ligations. The use of CIAP also presents a problem. It is beneficial to hinder the self-ligation of the expression vector in order to increase the efficiency of ligation reactions. However, because CIAP is difficult to remove, the enzyme may also inhibit ligations between DNA fragments and dephosphatased vectors.

Some success had been achieved previously in the group ligating gene fragments with different restriction sites at each terminus to pET-vectors. For this reason two new primers

were designed for PCR experiments (Fig. 2.9). The forward primer (TS5) contains an *Nde* I-site directly before the gene start codon. As a minimum of 8 bases are needed before the site for the enzyme to function efficiently, a ten base sequence was also added before this restriction site. The reverse primer (TS6) contains a *Bam*H I-site, which is found on all the pET-vectors that had been used previously (see Fig. 1.17).

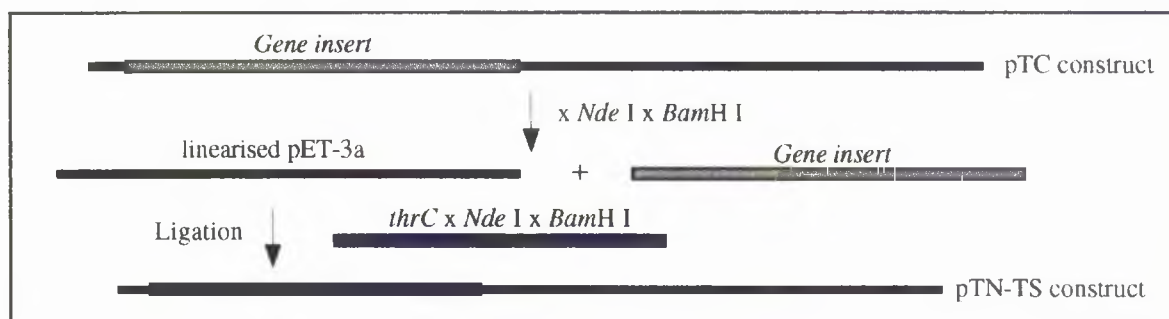
*ThrC* was amplified via PCR as before, from the original chromosomal DNA preparation, gel-isolated, restricted with *Nde* I and *Bam*H I and gel-purified, using the Wizard™ protocol, before carrying out further ligation reactions.



**Figure 2.9:** Left: primers used for the third PCR experiment to amplify *thrC*. Right: result of PCR using primers TS5 and TS6.

### 2.3.2.7 Attempted Ligation of *ThrC* into pET-3a

Attempts were made to ligate the doubly-restricted PCR product directly into pET-3a. In order to monitor the progress of the restriction of pET-3a, the original vector was not used in digestion experiments. Instead, constructs made previously in the group, from pET-3a and various gene inserts, were restricted with *Nde* I and *Bam*H I. These had been produced using the same protocol now being applied to *thrC*, the gene insert having an *Nde* I-site at one end and a *Bam*H I-site at the other. The progress of the restrictions could be monitored by the observation of a gene insert and linearised vector on agarose gels (Scheme 2.5).



**Scheme 2.5:** Plan for the ligation of *thrC* into pET-3a derived from a pTC construct.

A number of different pTC constructs were linearised using *Nde* I. Subsequent *Bam*H I-digests produced an observable gene fragment and linearised pET-3a vector (4640 bp). The linearised vector was then isolated from Nusieve™ agarose gel. Two sets of ligations were attempted, using CIAP-treated vector only in the second case, but no recombinants were identified.

### 2.3.2.8 Ligation of New Gene Fragment Into pGEM-T

#### 2.3.2.8.1 Amplification of *ThrC*

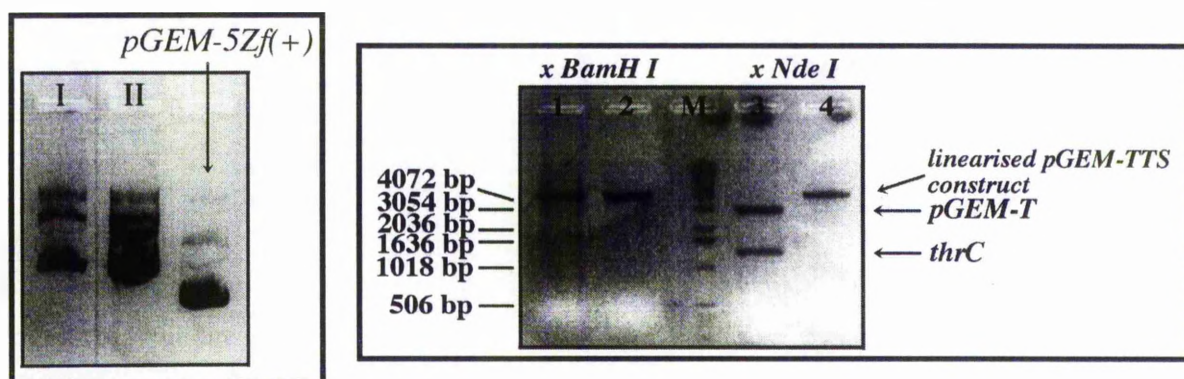
It appeared that, even if the termini of the *thrC* gene fragment contained 2 different restriction sites, ligation of the gene into a pET-vector would prove difficult. In order to prevent the need to perform repeated PCR experiments the gene was cloned as before into pGEM-T. At this stage the original chromosomal DNA preparation (also containing pSG4) from TG1 became ineffective as a template for PCR. Chromosomal DNA was therefore isolated from XL1-Blue using the SDS-Proteinase K protocol.<sup>270</sup> PCR experiments were carried out using either this chromosomal DNA preparation or pGEM-TTS<sub>B</sub> as a template. The first set of these experiments produced fragments of DNA too small to be *thrC*, but the use of freshly synthesised primers and a lower annealing temperature (40 °C) achieved the desired result.

*ThrC* fragments from PCR were gel-isolated and purified using the Wizard™ PCR Preps DNA Purification System. Direct ligation experiments produced the recombinant plasmids

pGEM-TTS<sub>I</sub> and pGEM-TTS<sub>II</sub> (see Table 2.2). QIAGEN maxiprep plasmid purifications of these vectors were subjected to *Nde* I- and *Bam*H I-digestion, with the predicted results (Fig 2.10).

**Table 2.2:** Results of ligations of the new *thrC* fragment into pGEM-T.

No.	Vector: PCR product ratio	Number of Colonies	% Whites
1	1:0.8	49	80%
2	1:0.8	37	81%
3	1:1	Lawn	95%
4	1:1	Lawn	95%
5	1:1.5	Lawn	95%
6	1:1.5	Lawn	95%
7	1:2	11	55%
8	1:2	8	63%
9	1:3	15	67%
10	1:3	11	55%



**Figure 2.10:** Results of ligations of *thrC* into pGEM-T. Left: plasmid miniprep DNA preparations pGEM-TTS<sub>I</sub> and II, taken from ligation experiments 9 and 10 respectively, showing they are putative recombinants. Right: restriction digests of these recombinants using *Nde* I and *Bam*H I. Lanes 2 and 4 contain restricted pGEM-TTS<sub>I</sub>, 1 and 3 pGEM-TTS<sub>II</sub>. The *Nde* I-restriction indicates that the gene is oriented differently in the two vectors.

### 2.3.2.9 Sequencing of *ThrC* in pGEM-TTS<sub>I</sub> and pGEM-TTS<sub>II</sub>.

Sequences were obtained of the termini of the recombinant gene in both vectors. This was achieved using the fmol™ protocol for cycle sequencing with SP6 and T7 primers (Fig. 2.7). Both primers were <sup>32</sup>P-labelled at the 5'-terminus by incubation with <sup>32</sup>P-γ-ATP and T4 PNK. PCR experiments were performed with each reaction containing one of the primers and one of four chain-terminating ddNTPs. The radiolabelled DNA fragments obtained in this manner were separated by high-resolution denaturing polyacrylamide gel electrophoresis using different lanes for each reaction. The sequences obtained are shown in Fig. 2.11, confirming that the genes are oriented in different ways in either vector. Attempts were also made to sequence the middle of the recombinant genes at this time, using primers TS3 and TS4 (Fig. 2.5), but without success. This could be due to the plasmid DNA being too dilute, the primer solutions having deteriorated during storage, or their concentration having altered since last being calculated. There could also have been some residual contaminants in the plasmid DNA preparation, which made the polymerase reaction too inefficient. Whilst further attempts were made to obtain fuller sequence information on the two recombinant genes, experiments were carried out to subclone the genes into pET vectors.

**pGEM-TTS<sub>I</sub> x SP6** 1st 120 bases of gene from N-terminus

```

TT GGA GCT CTC CAT ATG GTC CGG CTC TCT GCA GGC GGC CGC ACT CAG TGA TTG
TCA GTC AGT CAT ATG AAA CTC TAC AAT CTG AAA GAT CAC AAC GAG CAG GTC AGC
      Nde I
TTT GCG CAA GCC GTA ACC CAG GGG TTG GGC AAA AAT CAG GGG CTG TTT TTT CCG
CAC GAC CTG CCG GAA TTC AGC CTG

```

**pGEM-TTS<sub>I</sub> x T7** 1st 170 bases of gene from C-terminus

```

CC CGG CCG CCA TGG CGG GAT TTA TGC GGA TCC TTA CTG ATG ATT CAT CAT CAA
      BamH I
TTT ACG CAA CGC AGC AAA ATC GGC GGG CAC ATT ATG TGA ACT TTT GGC AGA TCC
AAC GTT TCA CCG AGA ATC GCT TCC ACG CTC TCT TTA AAT TTC GCC GGA TGC GCG
GTG CCG A

```

**pGEM-TTS<sub>II</sub> x SP6** 1st 80 bases of gene from C-terminus

```

GGA GCT CTC CAT ATT GGC GAC ATC TGC AGG CGG CCG CAC TAG TGA TTA TGC GGA
      Bam
TCC TTA CTG ATG ATT CAT CAT CAA TTT ACG CAA CGC AGC AAA ATC GGC GGG CAG
H I
ATT ATG TGA AAG CAG CAC GTT CTG CCA GCT CA

```

**pGEM-TTS<sub>II</sub> x T7** 1st 200 bases of gene from N-terminus

```

AT TGT ACA GTC AGT CAT ATG AAA CTC TAC AAT CTG AAA GAT CAC AAC GAG CAG
      Nde I
GTC AGC TTT GCG CAA GCC GTA ACC CAG GGG TTG GGC AAA AAT CAG GGG CTG TTT
TTT CCG CAC GAC CTG CCG GAA TTC AGC CTG ACT GAA ATT GAT GAG ATG CTG AAG
CTG GAT TTT GTC ACC CGC AGT GCG AAG ATC CTC TCG CGC TTT ATT GGT GAT GAA
AT

```

**Figure 2.11:** Sequences obtained for pGEM-TTS<sub>I</sub> and pGEM-TTS<sub>II</sub>. Primer sequences are shown in red except for the Nde I and BamH I-restriction sites which are shown in blue and underlined. The first base of the thrC gene is indicated by the arrow.



### 2.3.2.10: Ligation of *ThrC* Into pET-Expression Vectors.

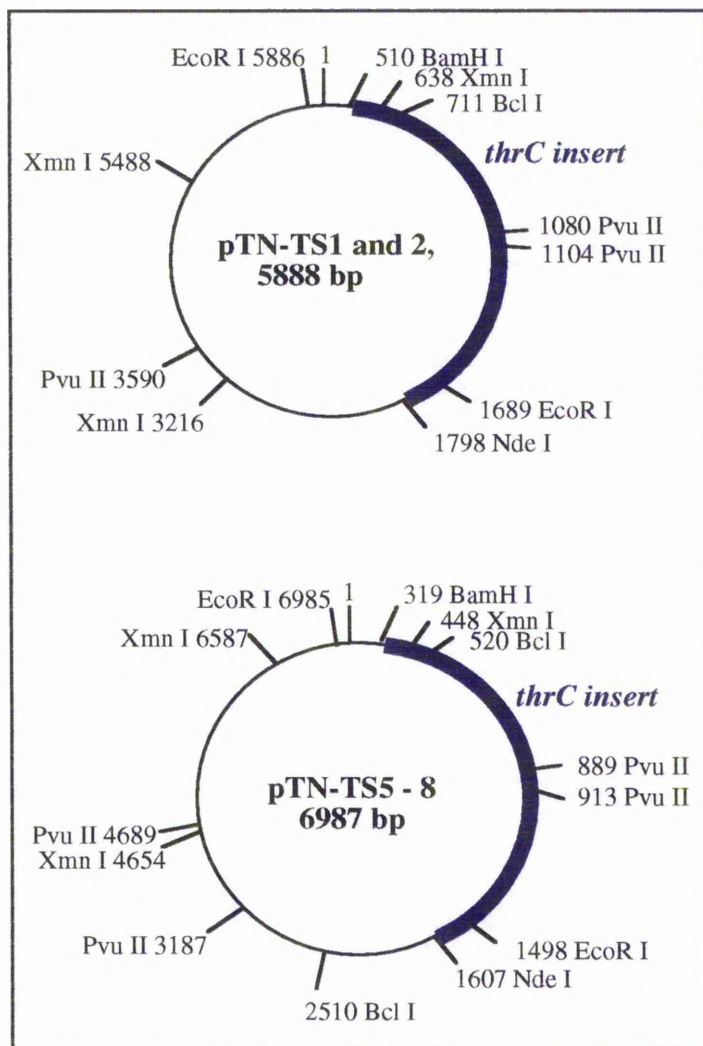
Once the *thrC* gene had been cloned into pGEM-T it was a simple procedure to excise the gene fragment using *Nde* I and *Bam*H I and gel-purify it for use in ligations. The pET-vectors pET-3a, -3b and -16b were also restricted with these enzymes, but not treated with CIAP as the enzyme was thought to be inhibiting ligations. Various ligation experiments were carried out using different ratios of insert : vector and a number produced successful recombinants. This could not be determined from the plasmid minipreps, as the DNA was smeared when run on a gel. Therefore maxiprep plasmid DNA was produced, as previously described, from 8 putative recombinants (see Table 2.3).

**Table 2.3:** *pTN-TS constructs produced from various pET-vectors and pGEM-TTS constructs.*

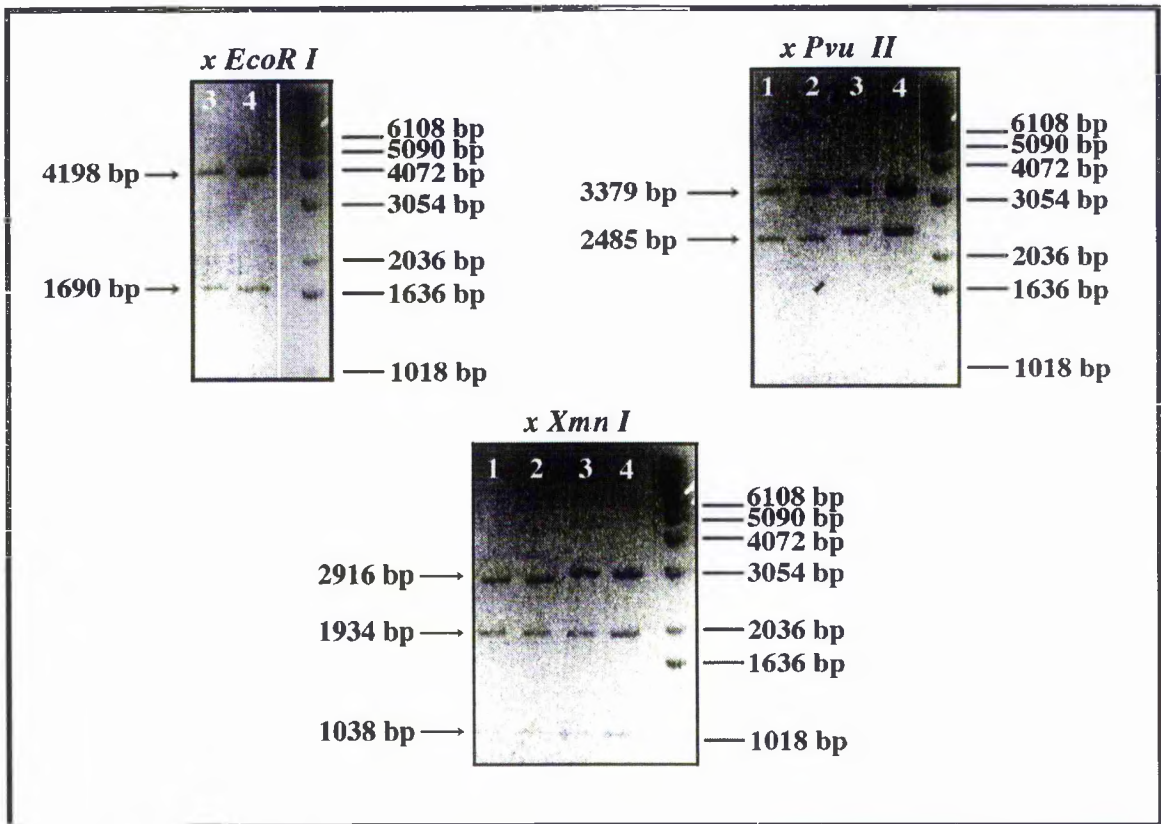
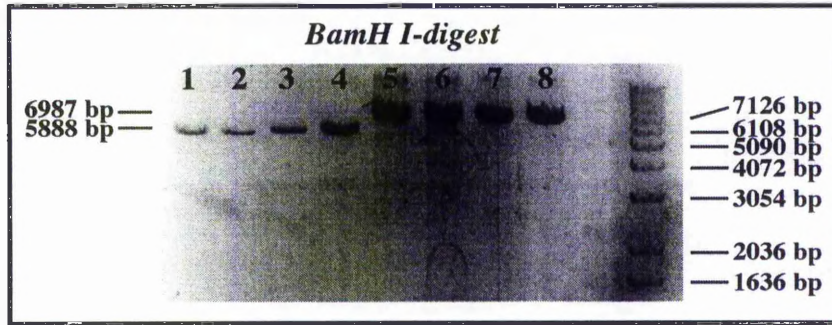
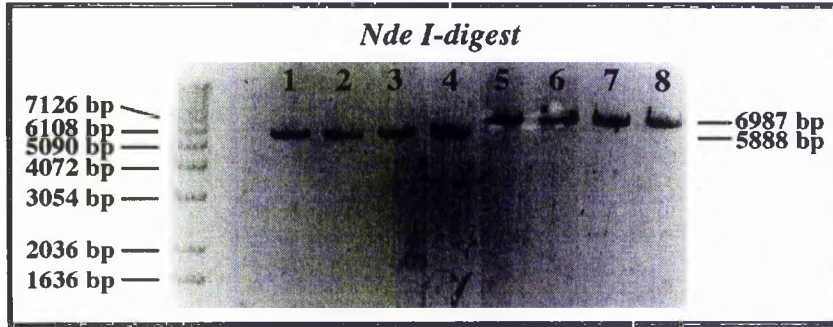
No.	pET-vector	Source of <i>ThrC</i>	Number of Colonies <sup>a</sup>	pTN-TS construct(s) obtained
1	pET-3a	pGEM-TTS <sub>I</sub>	8	-
2	"	pGEM-TTS <sub>II</sub>	8	pTN-TS1 and pTN-TS2
3	pET-3b	pGEM-TTS <sub>I</sub>	11	pTN-TS3
4	"	pGEM-TTS <sub>II</sub>	18	pTN-TS4
5	pET-16b	pGEM-TTS <sub>I</sub>	30	pTN-TS5 and pTN-TS6
6	"	pGEM-TTS <sub>II</sub>	45	pTN-TS7 and pTN-TS8

*a* - only white colonies were produced.

The recombinant vectors derived from pET-3a and -3b were subjected to restriction analyses using *Nde* I, *Bam*H I, *Eco*R I, *Pvu* II and *Xmn* I and produced the predicted fragments in all cases. The vectors derived from pET-16b were also restricted using these enzymes and *Bcl* I, producing the predicted fragment sizes in all cases. Fig. 2.12 shows the graphic maps of the pTN-TS constructs. The gels for each restriction digest are shown in Figs. 2.13 and 2.14.



**Figure 2.12:** Vector map of *pTN-TS* constructs. Restriction sites for *pTN-TS3* and *4* are similar to those in *pTN-TS1* and *2* but orientated in the reverse direction.



**Figure 2.13:** Gels of restriction digests of pTN-TS constructs. Lanes are marked 1 - 8 denoting the pTN-TS constructs 1 - 8. The *EcoR* I-digest of pTN-TS1 and 2 produced the same fragment sizes as shown for pTN-TS3 and 4.

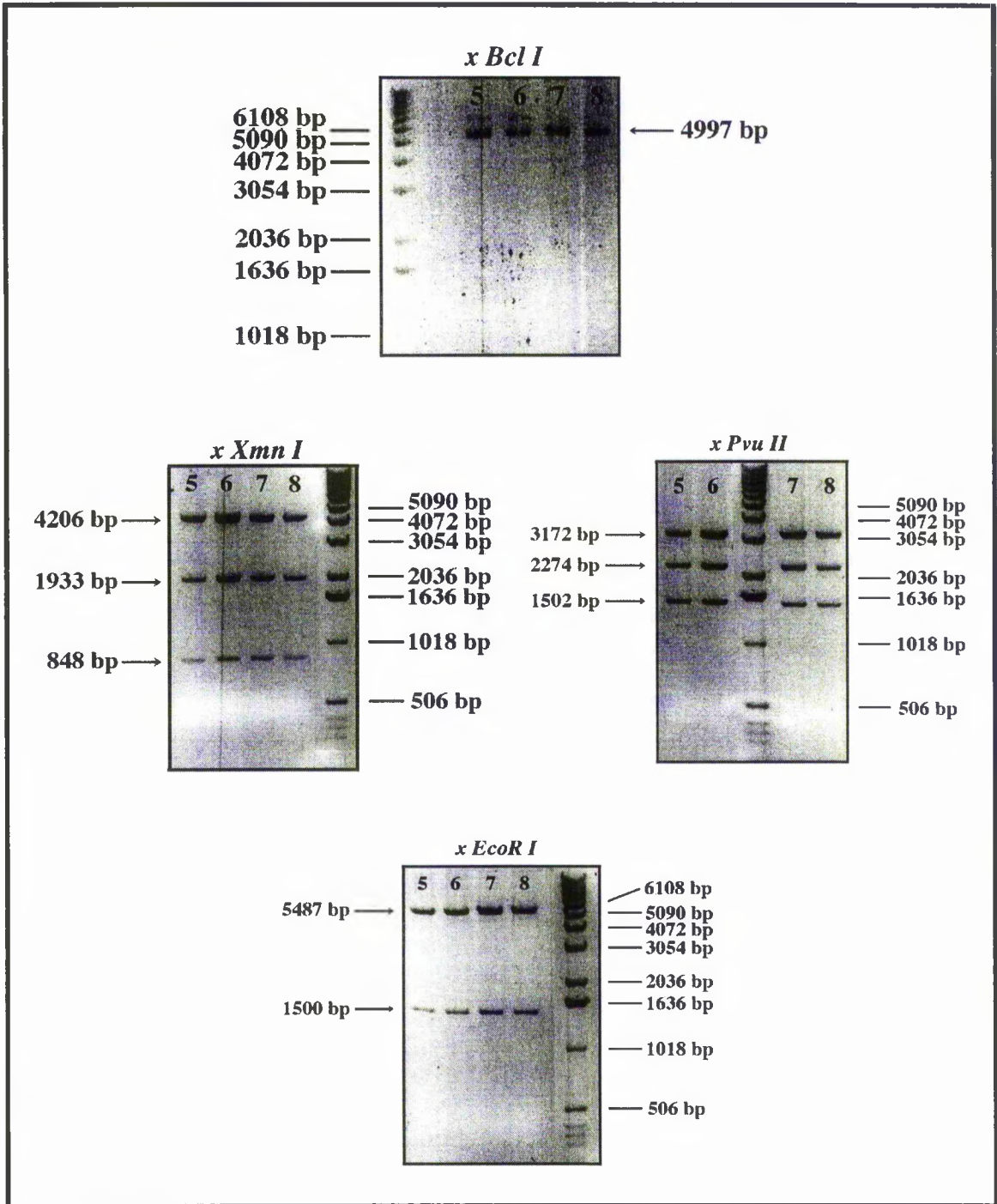


Figure 2.14: Restriction analyses of pTN-TS constructs derived from pET-16b.

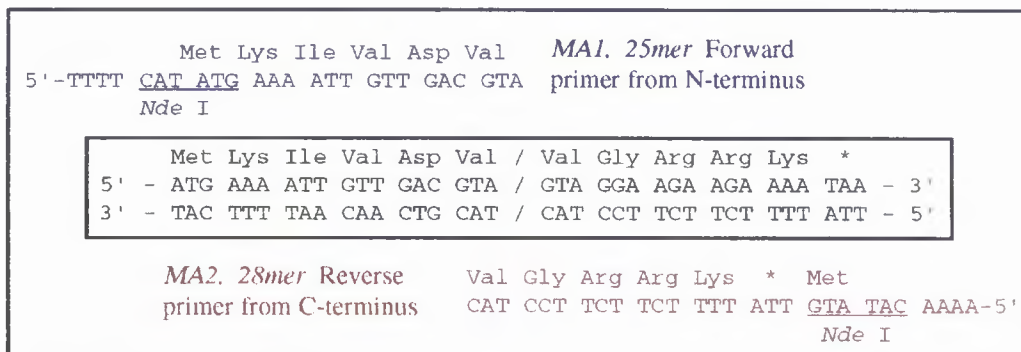
### 2.3.3 Cloning of the $\beta$ -Methylaspartase Gene

Much of the work carried out during attempts to clone the gene for  $\beta$ -methylaspartase into pET-vectors was similar to that already described for *thrC*, as the two sets of experiments were performed simultaneously. Cloning experiments will therefore not be discussed in as much detail in this section.

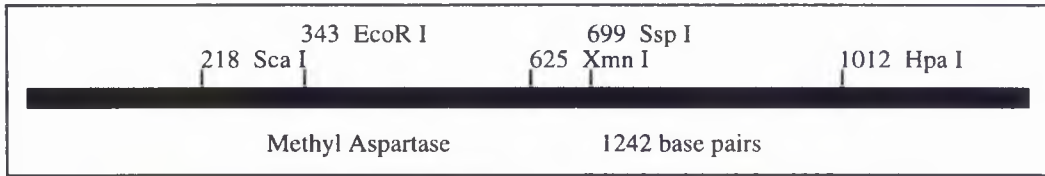
#### 2.3.3.1 Cloning the Gene into pGEM-T

##### 2.3.3.1.1 Gene Amplification

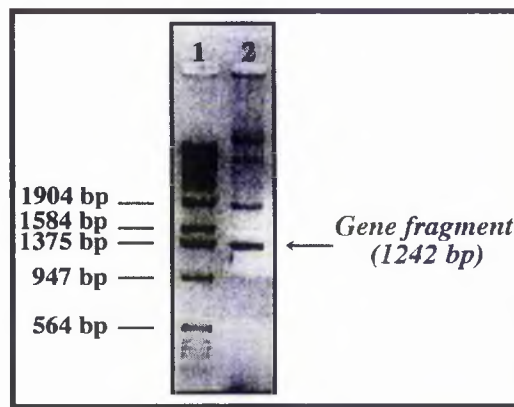
The plasmid pSG4 was isolated from *E. coli* strain TG1 (*K-12 [lac-pro] supE thi hsdR5 / F'-traD36 proA+ B+ lacI Q lacZ M15*)<sup>260</sup> using the alkaline lysis method (Section 2.3.2.1.1).<sup>261, 262</sup> The  $\beta$ -methylaspartase gene was originally amplified using primers which incorporated *Nde* I-sites into each terminus of the gene fragment (Fig. 2.15) as for *thrC*. These sites are not contained on the gene (Fig. 2.16 and Appendix 10). The PCR product was then gel-isolated.



**Figure 2.15:** DNA sequence of the two termini of the  $\beta$ -methylaspartase gene (centre) and the primers (top and bottom) used in the PCR reaction. *Nde* I-sites on the primers are underlined.



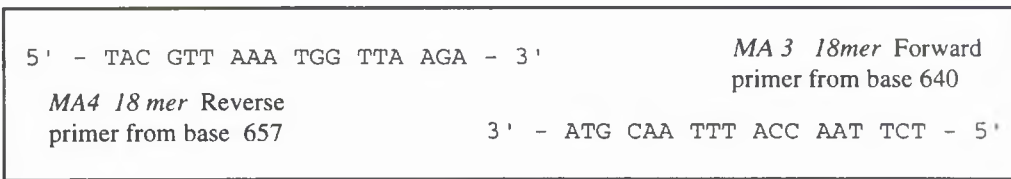
**Figure 2.16:** Graphic map of the restriction sites on the gene for  $\beta$ -methyl-aspartase used during cloning work..



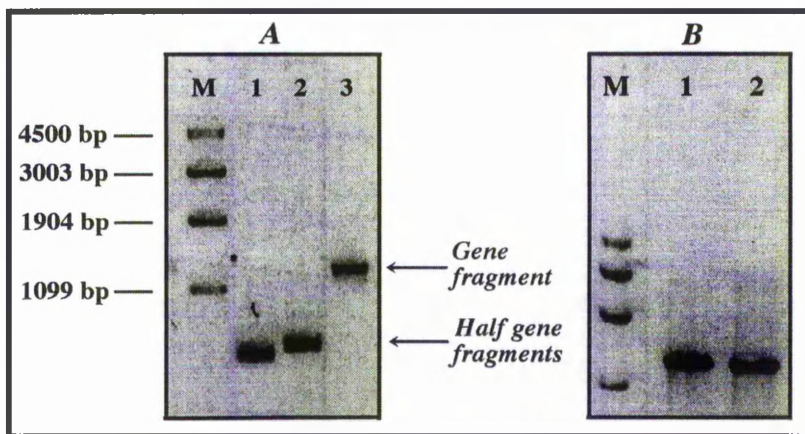
**Figure 2.17:** Result of the first PCR experiment to amplify the gene for  $\beta$ -methylaspartase (1242 bp). Lane 1 contains a  $\lambda$  DNA ladder, lane 2 a portion of the PCR reaction mixture after 20 cycles.

#### 2.3.3.1.2 Overlap PCR Experiments

As was the case with *thrC*, preliminary attempts to ligate the amplified gene into cloning vectors were unsuccessful and overlap PCR was performed to confirm that the correct gene had been produced, as for *thrC* (Section 2.3.2.3). Two DNA fragments were produced using the two external primers MA1 and MA2, and two new primers complementary to the centre of the gene (Fig. 2.18), which were of the correct size to be half-gene fragments. These were used as a template in overlap PCR experiments, producing cloned DNA of approximately 1300 bp, consistent with an amplified gene fragment of 1259 bp (1242 bp for the gene plus 17 additional bases produced by the primers, Fig. 2.19).



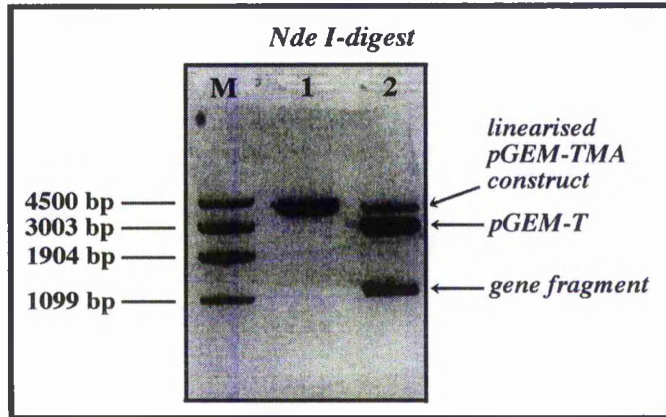
**Figure 2.18:** Primers complementary to the centre of the gene for  $\beta$ -methylaspartase.



**Figure 2.19:** Results of overlap PCR experiments. A: Lanes 1 and 2: half gene fragments produced using MA2 & MA3 or MA1 & MA4 respectively, lane 3 whole gene from PCR using MA1 and MA2. B: amplified gene (1242 bp) produced from either half-gene fragment (see text for details).

### 2.3.3.1.3 pGEM-T

The gene fragment produced from overlap PCR experiments was directly ligated to pGEM-T using similar vector:insert ratios to those described for *thrC* (Table 2.1). Competent JM83 cells were used for the subsequent transformations. Seven putative recombinants were identified and subjected to restriction analyses using *Apa* I, *Bss*H II, *Eco*R I, *Sca* I, *Ssp* I and *Xmn* I, which suggested that five of them contained the desired DNA insert. Two of these (denoted pGEM-TMA<sub>I</sub> and pGEM-TMA<sub>II</sub>) were used to prepare maxiprep plasmid DNA. Upon restriction of these vectors with *Nde* I it became apparent that pGEM-TMA<sub>I</sub> had lost one *Nde* I-site, as no gene fragment was excised from this construct (Fig 2.20).



**Figure 2.20:** *Nde I*-restriction digest of pGEM-TMA<sub>I</sub> (1) and pGEM-TMA<sub>II</sub> (2).

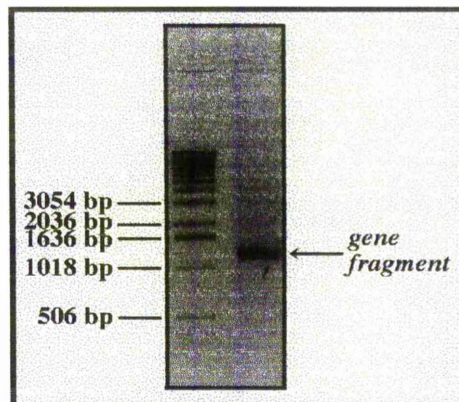
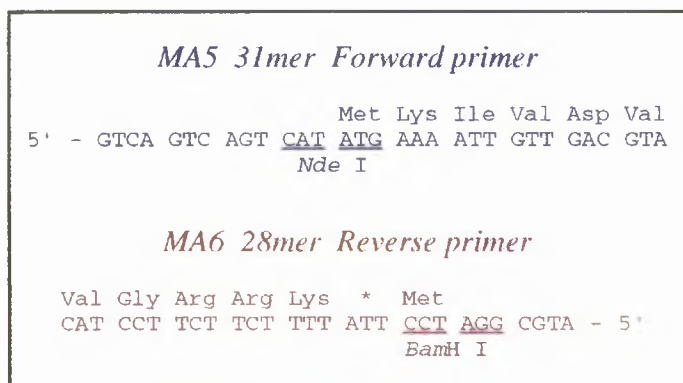
It was confirmed that the gene insert in pGEM-TMA<sub>I</sub> had only retained one *Nde I*-site by sequencing of the gene insert in each recombinant plasmid, using the fmol™ protocol, with the SP6 and T7 promoter primers (Fig. 2.7). The correct sequence for the first 120 bases of the gene from the SP6 promoter (the C-terminus of the gene) and the first 120 bases from the T7 promoter (the N-terminus) were obtained from pGEM-TMA<sub>II</sub>.

The β-methylaspartase gene fragment was excised from maxipreps of pGEM-TMA<sub>II</sub> using *Nde I* and purified by agarose gel electrophoresis. As with *thrC*, attempts to ligate this into various expression vectors were unsuccessful.

#### 2.3.3.1.4 A New Gene Fragment

After several failed attempts at cloning the gene fragment into various expression vectors, a new gene fragment was produced with an *Nde I*-site at the N-terminus and a *BamH I*-site at the C-terminus (Fig. 2.21). The PCR product was gel-isolated in the usual manner and attempts were made to clone this directly into pGEM-T.

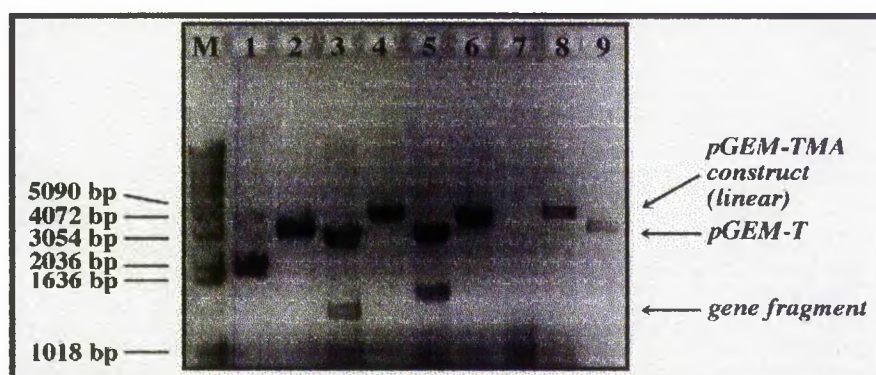




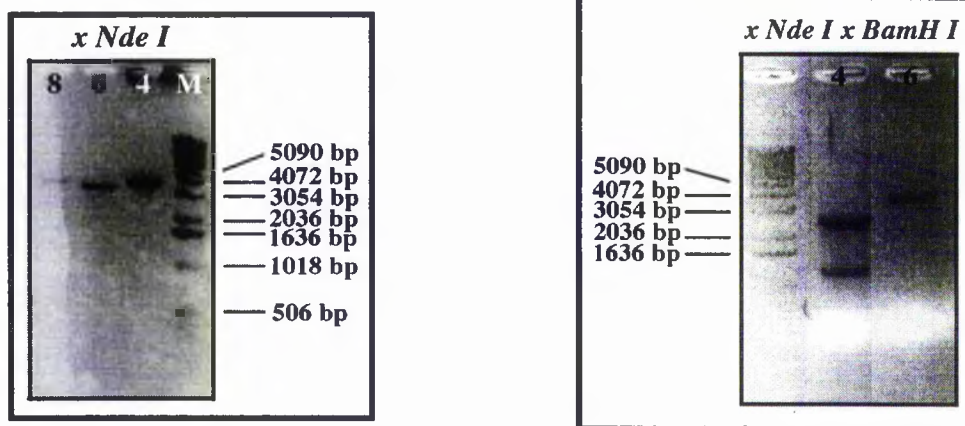
**Figure 2.21:** Left: primers used in the third set of PCR reactions for  $\beta$ -methylaspartase, right: result of PCR using these primers.

### 2.3.3.1.5 New pGEM-TMA Constructs

Ligation of the new gene fragment into pGEM-T produced several putative recombinants. A number of these were digested with *Nde* I to produce linearised constructs (samples 4, 6 and 8, 4245 bp, Fig. 2.22). DNA inserts were also excised from recombinants 3 and 5 (Fig. 2.22). However, some of the putative recombinants produced fragments which did not correspond to any of the expected *Nde* I-restriction products. The putative recombinants shown in lanes 4, 6 and 8 of Fig. 2.22 were used to produce maxiprep plasmid DNA which was subjected to restriction with *Nde* I and *Bam*H I (Fig. 2.23). Only sample 4 appeared of the correct size for a recombinant and subsequent restriction with *Bam*H I apparently excised a cloned DNA insert. However the remaining DNA fragment is too small to be pGEM-T (3003 bp), indicating that more than just the desired gene insert was excised from the plasmid, when digested with *Nde* I.



**Figure 2.22:** *Nde* I-restriction digest of several putative pGEM-TMA constructs.



**Figure 2.23:** Restriction digests on maxiprep DNA produced from samples shown in Fig. 2.22.

### 2.3.2.3 Further Work on the Gene for $\beta$ -Methylaspartase

Further ligation experiments were attempted to produce successful pGEM-TMA constructs, and several more putative recombinants were identified. However, at this stage efforts were focused on the work to clone the *thrC* gene into pET-vectors and further work to clone the gene for  $\beta$ -methylaspartase into pGEM-T and then into expression vectors was continued by other researchers. Successful pGEM-TMA constructs were indeed obtained and the gene fragment excised from these was inserted into pET-5a. Unfortunately, although some success was achieved in purifying the recombinant enzyme from this source, efforts made by S. McMahon of J. Naismith's group to produce crystals in order to obtain a 3D-structure of the enzyme were without success. This was thought to be due to the presence of the His • Tag, which could not be cleaved with Factor Xa as expected. The gene was therefore excised and cloned into a new pET-vector to prevent the formation of a His • Tag. However crystals could not be obtained of the enzyme from this source either. Sequencing experiments later revealed that the stop codon at the end of the gene had mutated and further amino acid residues were being added to the end of the protein. The gene then had to be cloned directly from *C. tetanomorphum*.

## 2.4 Sequencing of The *ThrC* Gene Fragment in Recombinant Plasmids

Sequencing of each vector was carried out by the University of St. Andrews DNA Sequencing Unit. They produce sequence data using the ABI PRISM™ DNA Sequencer, which uses a different fluorescent dye for each DNA extension reaction: A, C, G and T. They also use AmpliTaq FS, a mutant form of *Taq* DNA polymerase, which is reported to read easily sequences which are otherwise difficult. The fluorescent dye is incorporated on each dideoxy terminator, although it is also possible to use dye primers which are fluorescently labelled instead. All four extension reactions are then loaded onto a single lane on a polyacrylamide gel. Perkin-Elmer software is used to read the sequence directly from the gel by scanning for the four fluorescent dyes.

**Table 2.4:** Description of primers used in sequencing experiments.

<b>Primer</b>	<b>Description</b>	<b>Sequence</b>
<b>SP6</b>	<i>SP6 Promoter Primer: External primer complementary to the SP6 promoter on pGEM vectors.</i>	5' - d(GATTTAGGTGACACTATAG) - 3'
<b>TS3</b>	<i>Internal primer producing sense strand of <u>thrC</u> (starting base 643).</i>	5' - d(AAC ATC AGC CGT TTG CTG) - 3'
<b>TS4</b>	<i>Internal primer producing anti-sense strand of <u>thrC</u> (starting base 660).</i>	5' - d(CAG CAA ACG GCT GAT GTT) - 3'
<b>TS6</b>	<i>External primer producing sense strand of <u>thrC</u> from N-terminus of gene.</i>	5' - d(ATGC GGA TCC TTA CTG ATG ATT CAT CAT) - 3'
<b>T7</b>	<i>T7 Promoter Primer: External primer complementary to the T7 promoter on pET and pGEM vectors.</i>	5' - d(TAATACGACTCACTATAGGG) - 3'

The constructs pGEM-TTS<sub>I</sub> and pGEM-TTS<sub>II</sub> were submitted for sequencing using the vector promoter primers SP6 and T7 and the gene internal primers TS3 and TS4 (Table 2.4). Full sequences were obtained of the gene inserts (Appendix 11). The following changes from the wild-type *thrC* DNA sequence were observed:

- In pGEM-TTS<sub>I</sub> base base 320 (adenosine) of *thrC* had been deleted, disrupting the reading frame of most of the gene. Base 944 was also mutated from thymidine to cytidine. Bases 340 and 716 were also mutated (A to G in both cases).
- In pGEM-TTS<sub>II</sub> the mutations described for pGEM-TTS<sub>I</sub> were found, except that base 340 (A) remained unaltered.

Each pTN-TS vector was also sequenced using the T7 promoter primer, TS6, which is complementary to the sense strand at the N-terminus of the gene and the two internal primers, TS3 and TS4. The data obtained from these constructs confirmed the findings for the pGEM-TTS constructs (Appendix 12). It should be noted that the gene inserts in pTN-TS3, pTN-TS5 and pTN-TS6 were subcloned from pGEM-TTS<sub>I</sub> and all others were subcloned from pGEM-TTS<sub>II</sub> (Section 2.3.2.10). Some additional mutations were observed in sequences of the pTN-TS constructs as follows :

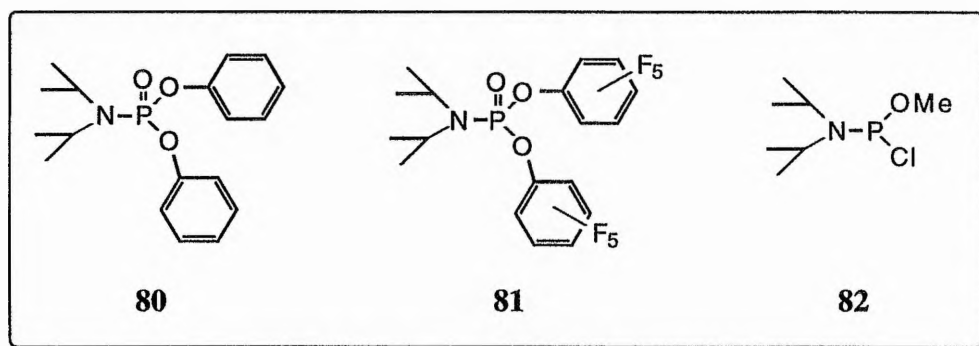
- In pTN-TS2 base 1051 was mutated (G to A).
- In the construct pTN-TS5 base 278 was mutated (T to A).

The deletion of base 320 of *thrC* will produce a truncated protein, as this leads to a stop codon (TGA) after a further sequence of 19 amino acids. In fact, the frameshift caused by this deletion leads to a total of 18 stop codons along the DNA sequence.

## 2.5 Solid-Phase Synthesis of (2S)-O-Phosphohomoserine

### 2.5.1 Alternative Phosphorylating Agents

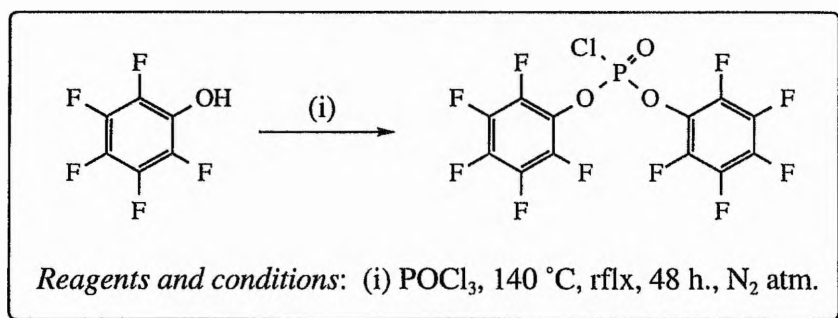
Due to the difficulties encountered in the synthesis of *N,N*-diisopropyldichlorophosphamidite (**78**), a key reagent for the phosphorylation of *N*-trifluoroacetyl- $\alpha$ -isopropyl-(2S)-homoserine (**75**) in the solution-phase protocol for substrate synthesis (Section 2.2.2), it was necessary to find some new means of attaching the phosphate group to **75**. To this end a few readily available phosphorylating agents were tried: diphenylphosphorochloridate (**80**), dipentafluorophenylphosphorochloridate (**81**) and *N,N*-diisopropyl methyl phosphonamidic chloride (**82**, Fig. 2.24).



**Figure 2.24:** Alternative phosphorylating agents tried in the synthesis of (2S)-O-phosphohomoserine.

Both the phosphonamidic chloride (**82**) and diphenylphosphorochloridate (**80**) are commercially available. There are many reported uses of **80**, with different methods for the subsequent deprotection of the phenyl protecting groups.<sup>271-283</sup> The main method for the removal of the phenyl groups involves hydrogenolysis using platinum dioxide (Adams catalyst) or platinum/activated carbon.

**81** was first patented in 1967, as a reagent in the synthesis of base stocks for hydraulic fluids, as heat transfer media and lubricants in aircraft systems.<sup>284, 285</sup> The compound was developed in this group as a phosphorylating agent for peptide synthesis by P. Hormozdiari.<sup>286</sup> **81** is synthesised by refluxing phosphorous oxychloride and 1.8 equivalents of pentafluorophenol at 140 °C under nitrogen for at least 20 h.<sup>287, 288</sup> The method was repeated here, continuing the reaction for 48 h (Scheme 2.6) The undesired side-products of this reaction were readily removed by distillation under reduced pressure, giving a pale brown oil, which solidified on cooling, in quantitative yield. The <sup>13</sup>C-NMR spectrum showed signals between 136.77 - 142.91 ppm corresponding to aromatic carbons; <sup>31</sup>P-NMR spectroscopy showed one main signal at -15.14 ppm and the <sup>19</sup>F-NMR spectrum showed signals at -154.89 and -154.96 ppm (4 F, d *J* 19.75, *ortho* Ar-CF), -157.96 ppm (2 F, dt, *J* 3.95, *J* 21.73, *para* Ar-CF) and -162.38 ppm (4 F, t, *meta* Ar-CF).

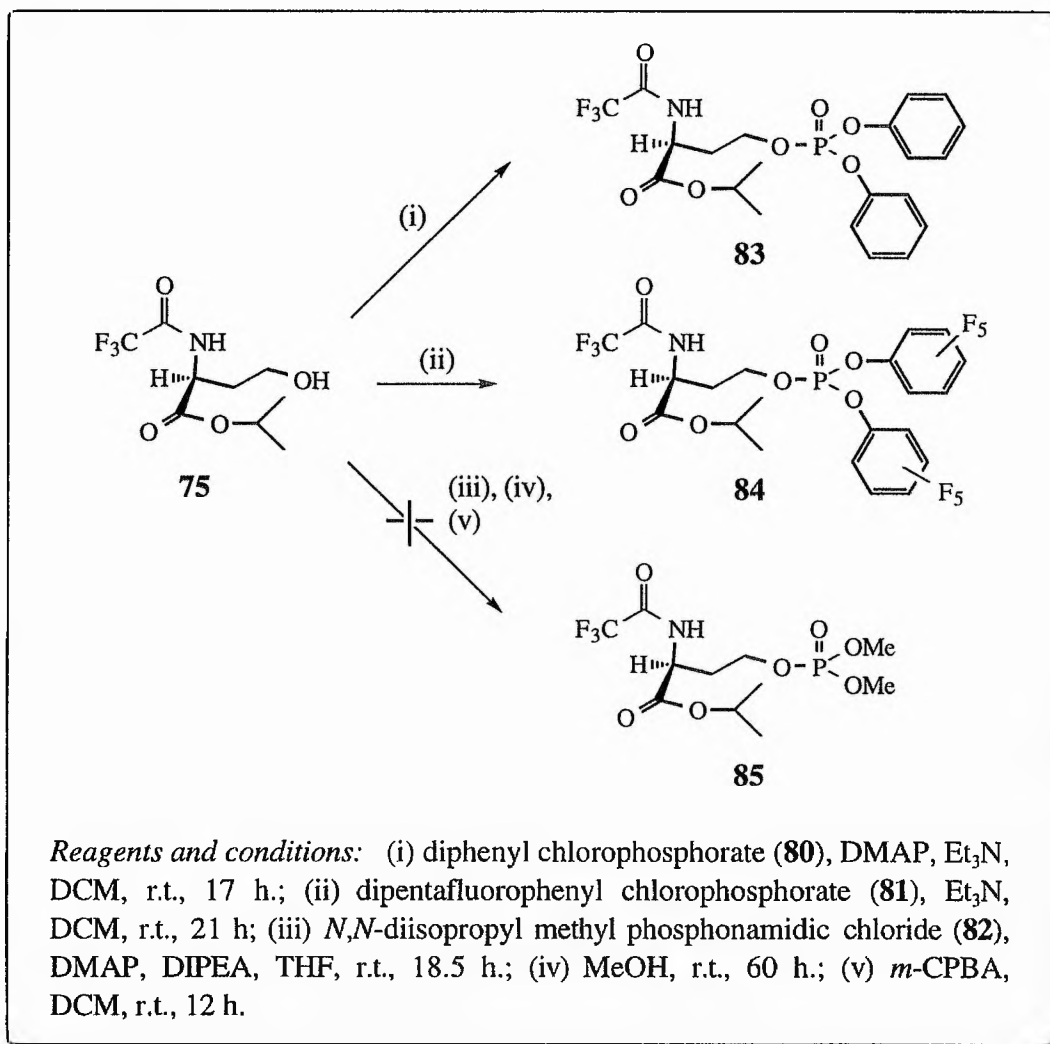


**Scheme 2.6:** Synthesis of dipentafluorophenylphosphorochloridate (**81**).

Phosphorylation reactions carried out using **80** and **81** were successful, but **82** failed to produce the desired phosphate methyl ester (**84**, Scheme 2.7).

The phosphorylation reaction using **80** produced the diphenylphosphate ester (**83**) in 82% yield. This was purified by silica column chromatography, eluting with ethyl acetate-light petroleum (1:1) (*R<sub>f</sub>* 0.42). LRMS gave a peak corresponding to the mass ion [*m/z* (EI) 489 (*M*<sup>+</sup>)]. The <sup>1</sup>H-NMR spectrum showed new aromatic signals between 7.35 and 7.17 ppm, and the <sup>13</sup>C-NMR spectrum indicated phosphorous coupling to βC and γC and one of the

aromatic carbons (Fig. 2.25). The  $\gamma\text{C}$  signal was also shifted downfield by 6.6 ppm:  $\delta_{\text{C}}$ (50.31 MHz;  $\text{C}^2\text{HCl}_3$ ) 169.49 ( $\text{CO}_2^i\text{Pr}$ ), 150.56 and 150.42 (d,  $J$  7.04, Ar-C), 130.11 (Ar-C), 124.80 (Ar-C), 120.21 and 120.13 (d,  $J$  17.21, Ar-C), 70.88 ( $^i\text{PrCH}$ ), 65.17 and 65.06 (d,  $J_{\text{P,C}}$  5.49,  $\gamma\text{C}$ ), 50.40 ( $\alpha\text{C}$ ), 31.82 and 31.68 (d.,  $J$  6.71,  $\beta\text{C}$ ), 21.79 ( $^i\text{PrCH}_3$ ).  $^{31}\text{P}$ -NMR spectroscopy produced just one signal at -11.57 ppm.



**Scheme 2.7:** Attempted methods of phosphorylating *N*-trifluoroacetyl- $\alpha$ -isopropyl-(2*S*)-homoserine (**75**).

The phosphorylation using **81** also proceeded in good yield. The  $^{13}\text{C}$ -NMR spectrum of **84** showed a downfield shift of the  $\gamma\text{C}$  signal and phosphorous coupling to  $\beta\text{C}$  and

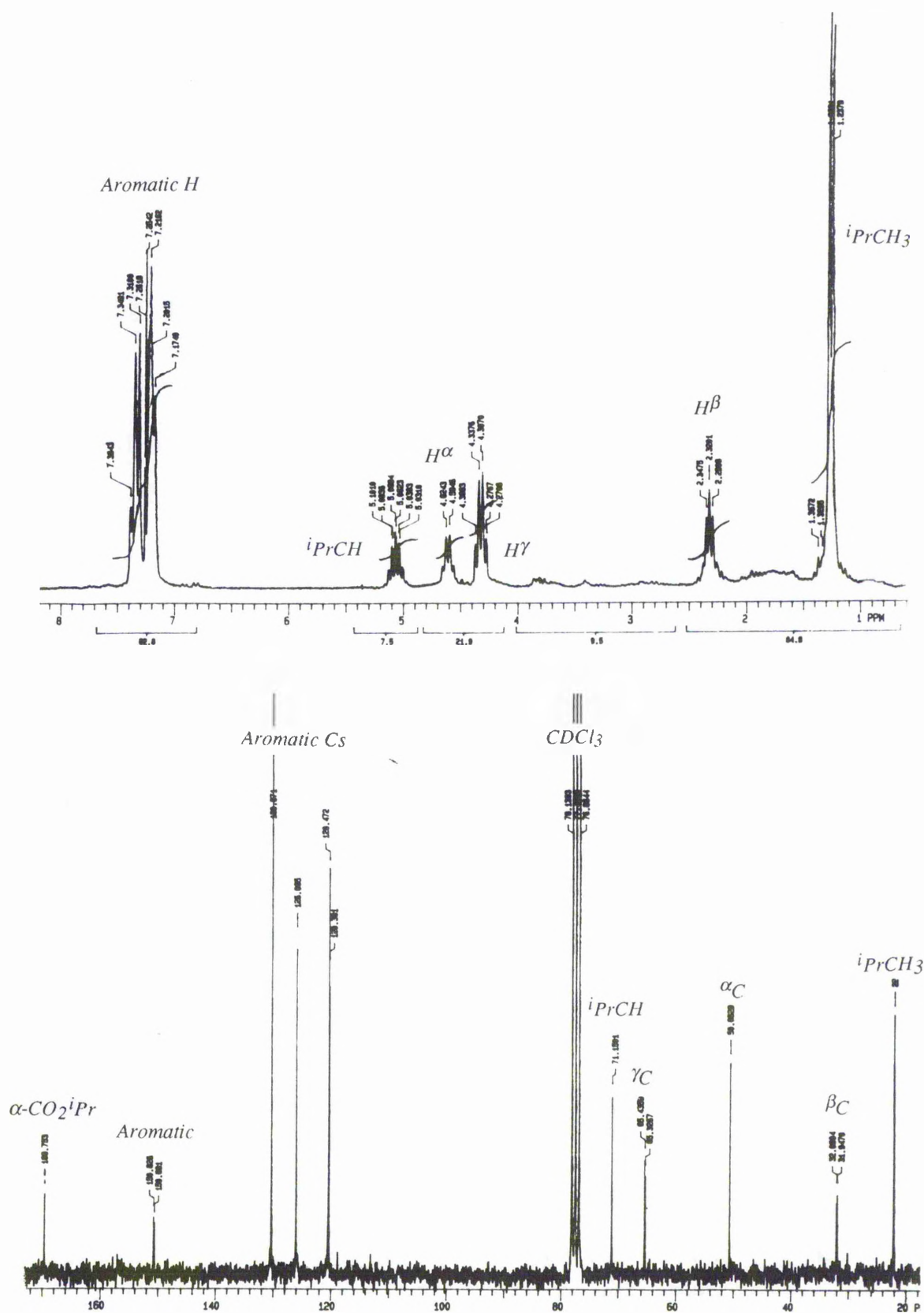


Figure 2.25: NMR spectra of N-trifluoroacetyl-(2S)-phosphohomoserine diphenyl ester (83). Top:  $^1\text{H-NMR}$  spectrum, bottom:  $^{13}\text{C-NMR}$  spectrum.



$\gamma\text{C}$ .  $\{\delta_{\text{C}}(50.31 \text{ MHz}; \text{C}^2\text{HCl}_3) 169.46 (\text{CO}_2^i\text{Pr}), 157.82 (\text{F}_3\text{CCO}), 138.46 (\text{Ar-C}), 71.19 ({}^i\text{PrCH}), 67.44 \text{ and } 67.32 (\text{d.}, J_{\text{P,C}} 5.90, \gamma\text{C}), 50.23 (\alpha\text{C}), 32.03 \text{ and } 31.88 (\text{d.}, J_{\text{P,C-P}} 7.32, \beta\text{C}), 21.69 ({}^i\text{PrCH}_3)\}$ . The  ${}^{31}\text{P}$ -NMR spectrum gave a new signal at  $-10.77 \text{ ppm}$ , and  ${}^{19}\text{F}$ -NMR spectroscopy of the crude material showed new multiplet signals at  $189.99, 188.27, 187.25 \text{ and } 185.45 \text{ ppm}$ .  ${}^{19}\text{F}$  signals for **81** in this spectral window appeared at  $198.10, 195.06 \text{ and } 190.60 \text{ ppm}$ .

Having established two new methods for the phosphorylation of **75**, attempts were made to deprotect the phosphate esters **83** and **84** using base-catalysed hydrolysis. If this were successful then the removal of all protecting groups could be performed in one step to give the substrate (**32**), a major advantage over the previous route (Scheme 2.3).

Although **83** was subjected to hydrolysis in  $1\text{M}$  sodium hydroxide, for several days, the deprotection appeared to be incomplete. The  ${}^1\text{H}$ -NMR spectrum still showed aromatic signals at between  $7.28 \text{ and } 7.00 \text{ ppm}$ , and although  ${}^{31}\text{P}$ -NMR spectroscopy showed one signal at  $0.67 \text{ ppm}$ , which possibly indicated the presence of the desired product  $\{\delta_{\text{P}}(121.42 \text{ MHz}; {}^2\text{HOH}) 0.59 \text{ for } (32)\}$ , a number of other  ${}^{31}\text{P}$  signals were also present. P. Hormozdiari found, whilst using this phosphorylating agent, that hydrogenolysis of diphenyl phosphates at atmospheric pressure, using either Adams catalyst or platinum/ activated carbon, failed to produce complete removal of the phenyl protecting groups.<sup>286</sup> Reports in the literature, however, state that complete deprotection can be effected using this method within  $30 \text{ min to } 3 \text{ h}$ .<sup>271-283</sup>

**84** was subjected to overnight hydrolysis in  $1\text{M}$  sodium hydroxide. A white precipitate was formed, and tlc on cellulose [eluant  ${}^i\text{PrOH}/\text{H}_2\text{O}/\text{NH}_3 (20:5:6)$ , developing with ninhydrin] showed a band with a similar  $R_f$  value to that of the target compound. The  ${}^1\text{H}$ -NMR spectrum was, however, inconclusive, and the  ${}^i\text{PrCH}_3$  signal of the  $\alpha$ -ester was still present, indicating that hydrolysis was incomplete. The  ${}^{13}\text{C}$ -NMR spectrum contained many aromatic signals at between  $142.64 \text{ and } 128.50 \text{ ppm}$ , indicating that the pentafluorophenyl groups were still present. P. Hormozdiari had attempted a number of methods of hydrolysing off the

pentafluorophenyl groups, using water, 5% NaHCO<sub>3</sub>, or an excess of base, but the reaction would not go to completion, even when an ethanolic solution of KOH and an organic base, pyridine/ H<sub>2</sub>O mixture was used.<sup>286</sup>

These two routes to the substrate appeared then to be unviable and the phosphate esters were not characterised further. Fortunately, the acid-lability of pentafluorophenol esters has been reported previously in the literature,<sup>288, 289</sup> and some success had been achieved in this group prior to these studies in the deprotection of such phosphates on the solid-phase, using acidic conditions.<sup>286</sup> With this in mind the solution-phase protocol for the synthesis of (2*S*)-*O*-phosphohomoserine (**32**) was modified in order to apply it to the solid-phase.

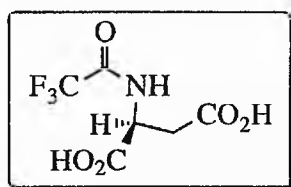
### 2.5.2 <sup>19</sup>F Gel-phase NMR Spectroscopy

One drawback to solid-phase synthesis is that monitoring of reaction progress is problematic, unless one of the reagents in solution is a strong UV-chromophore. Often the compound must also be cleaved from the resin to allow proper characterisation. Conventional <sup>1</sup>H- and <sup>13</sup>C-NMR spectroscopy of compounds attached to solid supports (gel-phase NMR) usually requires prolonged spectral acquisition, gives broad resonances and is dominated by signals from the solid support.<sup>290</sup> Various approaches to overcome these problems include magic angle spinning and the use of <sup>13</sup>C-enriched building blocks, but these methods are expensive. Recently, however, it has been shown that <sup>19</sup>F gel-phase NMR is a very sensitive technique for measuring transformations on solid-phase supports involving fluorine-containing compounds. <sup>19</sup>F chemical shifts are spread over a wide frequency range and structural changes which are even quite distant from the fluorine atom will affect the position of the signal. Svensson and coworkers found that the correlation between the <sup>19</sup>F chemical shifts for the support-linked compounds and the corresponding soluble references was excellent ( $\leq 0.1$  ppm difference) and that there was a significant change in the <sup>19</sup>F chemical shift ( $\geq 1.0$  ppm) for each transformation they carried out.<sup>290</sup> Shapiro *et al.* also reported the successful use of

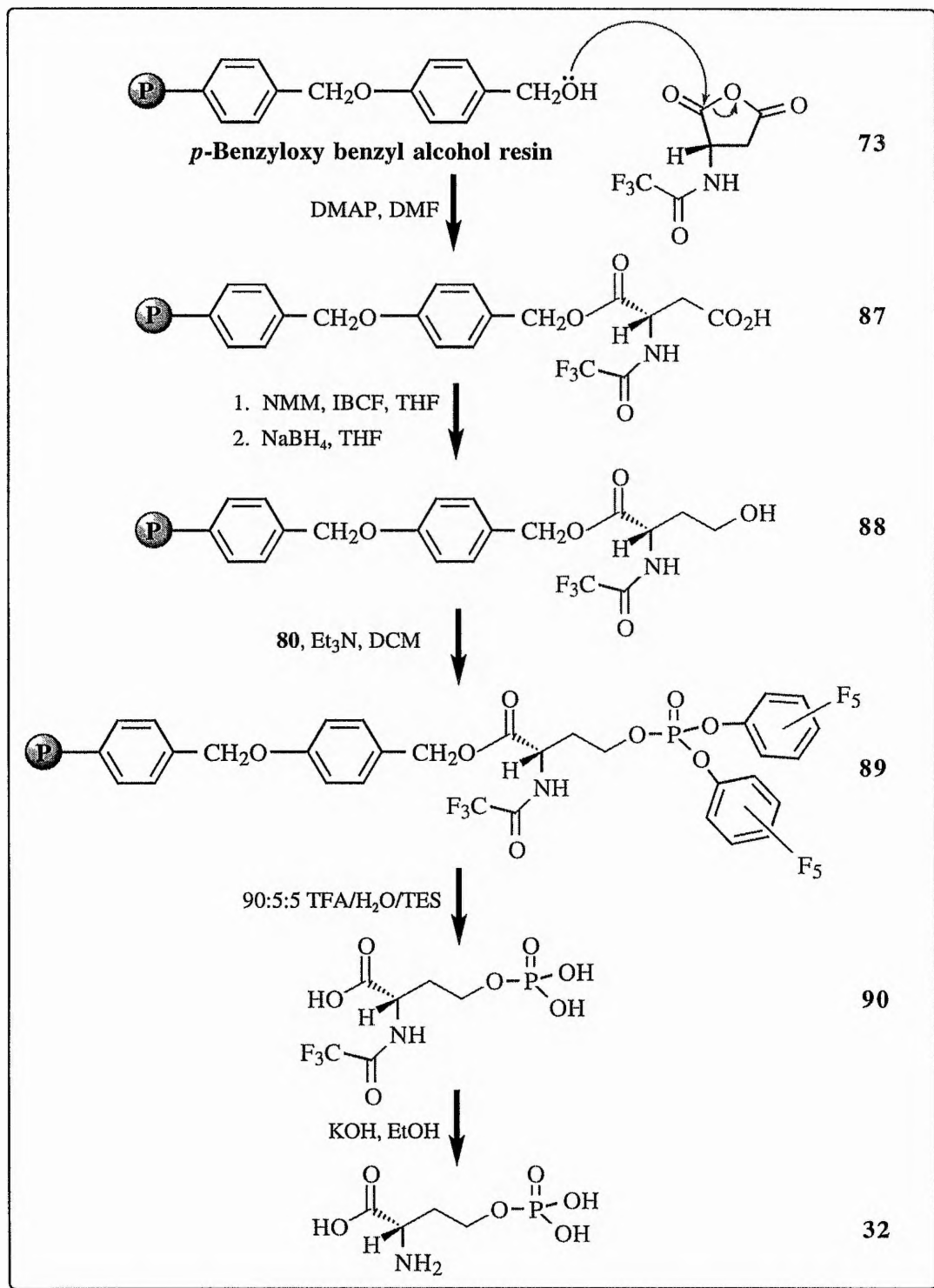
this technique for monitoring the time-course of nucleophilic displacement of fluorine from a solid-support.<sup>291</sup> In our experiments to perform the solid-phase synthesis of (2*S*)-*O*-phosphohomoserine, the presence of the *N*-trifluoroacetyl group lent itself readily to the use of this technique, as well as FT-IR and <sup>13</sup>C gel-phase NMR spectroscopy, for the characterisation of compounds on the solid support.

### 2.5.3 *p*-Benzyloxybenzyl Alcohol (Wang) Resin

Wang resin had previously been widely used in the group and was therefore an obvious first choice as a solid support. The strategy developed for the solid-phase synthesis of (2*S*)-*O*-phosphohomoserine (**32**, Scheme 2.8) was similar to the solution-phase procedure (see Scheme 2.3). The only major modification was that Wang resin was to be loaded with *N*-trifluoroacetyl-(2*S*)-aspartate (**86**) by using the alcohol group on the resin, in the place of propan-2-ol, for the ring opening of the symmetric anhydride (**73**). It was hoped that, due to the bulky nature of the resin, the opening of the anhydride would proceed regioselectively as with isopropanol. The β-carboxyl group was to be reduced in the usual manner by mixed anhydride methodology and the resulting alcohol group phosphorylated using dipentafluorophenyl chlorophosphate (**81**). The phosphate group would then be deprotected using TFA/H<sub>2</sub>O/TES (90:5:5) which would also cleave *N*-trifluoroacetyl-(2*S*)-*O*-phosphohomoserine (**90**) from the resin. This would leave only the hydrolytic removal of the *N*-TFA-group using base to achieve the target compound. The solid-phase methodology also removes the need for time-consuming purification steps and hydrogenolysis of the phosphate benzyl protecting groups of **76**, which often appeared to be inefficient.



*N*-trifluoroacetyl-(2*S*)-aspartate (**86**)



**Scheme 2.8:** The proposed solid-phase synthesis of (2S)-O-phosphohomoserine (32) on Wang resin.

### 2.5.3.1 The Loading of Wang Resin With *N*-Trifluoroacetyl-(2*S*)-aspartate

When using propan-2-ol to selectively protect the  $\alpha$ -carboxylic group of **73** the reaction was carried out at room temperature, after the addition of the alcohol at 4 °C (Scheme 2.3). Experiments to load **86** onto Wang were carried out as indicated in Scheme 2.8 and using the conditions listed in Table 2.5. A loading temperature of 4 °C was tested with Wang resin first of all (experiments 1-3, Table 2.5), as this low temperature might be expected to enhance the regioselectivity of the substitution. The reaction proceeded too slowly, however, to be useful at this temperature. Higher temperatures and catalytic amounts of DMAP were necessary to produce satisfactory loading within a reasonable time. Experiments 4-9 (Table 2.5) established that the optimum temperature for loading was probably between room temperature (23 °C) and 45 °C. In most cases a reaction temperature of 30-35 °C was used, as the regioselectivity of the substitution was better at this temperature than at higher temperatures. IR spectroscopy indicated the formation of an ester linkage  $\{\nu_{\max}(\text{KBr})/\text{cm}^{-1} 1735 \text{ s (C=O, ester)}\}$  and both the  $^{13}\text{C}$  and  $^{19}\text{F}$  gel-phase NMR spectra confirmed the substitution of **86** onto the resin:  $\delta_{\text{C}}$ (74.76 MHz; DMSO) 171.33 ( $\alpha\text{CO}_2\text{R}$ ), 162.48 ( $\beta\text{CO}_2\text{H}$ ), 130.02 ( $\text{F}_3\text{CCO}$ ), 117.84 ( $\text{F}_3\text{C}$ ), 49.40 ( $\alpha\text{C}$ ), 34.81 ( $\beta\text{C}$ );  $\delta_{\text{F}}$ (282.2 MHz; DMSO) -74.91 ( $\text{F}_3\text{C}$  on  $\alpha$ -ester). A portion of *N*-TFA-aspartate (**86**) was also cleaved from the resin using TFA/H<sub>2</sub>O/TES (90:5:5) and characterised. The  $^{13}\text{C}$ -NMR spectrum showed two signals for carboxylic groups  $\{172.54 (\alpha\text{CO}_2\text{H}) \text{ and } 171.06 (\beta\text{CO}_2\text{H})\}$ , shifted downfield from the carbonyl signals for **73**  $\{169.28 (\alpha\text{C=O}) \text{ and } 167.61 (\beta\text{C=O})\}$ . There were also two quartets which corresponded to the *N*-trifluoroacetyl group  $\{156.70 (\text{q}, J_{\text{F}_3\text{C},\text{C}} 37.18, \text{F}_3\text{CCO}) \text{ and } 115.43 (\text{q}, J_{\text{F},\text{C}} 285.26, \text{F}_3\text{C})\}$ . The optical rotation of the cleaved material was in good agreement with that obtained for the solution-phase product  $\{-25.09 (c. 0.87, \text{EtOH}) \text{ compared to } -24.35 \text{ for the solution-phase product}\}$ , indicating that the use of DMAP as an acylation catalyst does not cause significant racemisation, although P. Sieber found it did when attaching Fmoc-amino acids to Wang resin.<sup>292</sup>

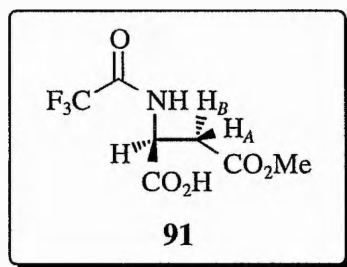
The regioselectivity of substitution reactions under various conditions were determined by

Table 2.5: Loading of Wang resin with N-TFA-asp under various conditions

Experiment	Temperature	Time (hours)	Equivalents of DMAP	Equivalents of Anhydride	Stirring	% Loading <sup>a</sup>
1	4°C	15	none	10	X	minimal
2	4°C	15	0.5	"	X	none
3	4°C	15	1.5	"	X	none
4	0°C / rt	15 / 6	none	"	X	none
5	0°C / rt	15 / 6	1	"	X	none
6	rt	19.5	none	"	X	none
7	rt	19.5	1	"	X	none
8	rt / 45°C	28.5 / 16	none	"	X	12
9	rt / 45°C	30 / 10.5	1	"	X	62
10	rt / 60°C	55.5 / 8	1	"	X	65 <sup>b</sup>
11	rt	65	2	"	✓	31
12	rt	160	2	"	✓	82
13	30°C	118.5	2	"	✓	100
14	25°C	36.5	1	"	✓	none
15	15°C / 35°C	48 / 116	1	"	✓	75
16	rt / 30°C	14 / 49	1	"	✓	58
17	30°C	142	1	"	✓	46
18	30°C	"	1	5	✓	56
19	30°C	"	1	2.5	✓	12
20	30°C	"	2	10	✓	79
21	30°C	"	4	"	✓	81

*a* - loading was determined by comparing the mass of the resin before and after loading experiments (after careful washing and drying of loaded resin), values given are only approximate; *b* - a reaction was carried out using 1 eq. pyridine under similar conditions, giving a loading of 60%

methylating the free acid group of **87** with diazomethane. The methyl-ester **91** was then cleaved from the resin using TFA/H<sub>2</sub>O/TES (90:5:5), and the ratio of  $\alpha$ - vs  $\beta$ -methyl ester peaks on <sup>1</sup>H-NMR spectra was determined {3.70 (3 H, s,  $\alpha$ CO<sub>2</sub>CH<sub>3</sub>, 16%), 3.63 (3 H, s,  $\beta$ CO<sub>2</sub>CH<sub>3</sub>, 84%)}. <sup>19</sup>F solution-phase NMR spectroscopy of the methyl ester **91** { $\delta_F$ (282.2 MHz; C<sup>2</sup>HCl<sub>3</sub>) -76.66 (F<sub>3</sub>C on  $\beta$ -ester)} and <sup>19</sup>F gel-phase NMR spectroscopy on *N*-TFA-asp-Wang (**87**) { $\delta_F$ (282.2 MHz; (C<sup>2</sup>H<sub>3</sub>)<sub>2</sub>SO) -74.27 (12%, F<sub>3</sub>C on  $\beta$ -ester), -74.91 (88%, F<sub>3</sub>C on  $\alpha$ -ester)} also confirmed the findings of these NMR experiments (Fig. 2.26 and Table 2.8). This indicated that the lower temperatures produced better selectivity for  $\alpha$ -substitution. An experiment left for a week at room temperature (12, Table 2.5) gave a high loading (82%) and one of the the highest regioselectivities achieved { $\alpha/\beta$  5.2:1 as judged from the <sup>1</sup>H-NMR spectrum of the methyl ester (**91**)}.



*The  $\beta$ -methyl ester (**91**) produced by diazomethane treatment of **87**.*

Some experiments were necessary in order to optimise the equivalents of anhydride and DMAP used in the reaction. It was also of interest to ascertain the effect of varying the number of equivalents of DMAP, with respect to the racemisation of the loaded amino acid and the regioselectivity of substitution. Experiments 17-20 (Table 2.6) show the effect of different quantities of anhydride and DMAP on the  $[\alpha]_D$  values of **86** cleaved from the resin. Loading experiments 8, 9, 10, 12 and 22 - 27 (Table 2.7) show the effect of different temperatures and equivalents of DMAP on the *regioselectivity* of the nucleophilic substitution reaction. Although 2.0 equivalents or more of DMAP gives a markedly faster loading of the resin compared to 1.0 equivalents, this has an adverse effect on the regioselectivity and causes greater racemisation of **86**.

**Table 2.6:** The racemisation effects of varying amounts of DMAP.

<i>Experiment<sup>a</sup></i>	<i>Ratio of Equivalents of DMAP vs. Anhydride</i>	<i>[<math>\alpha</math>]<sub>D</sub></i>
17	1:10	-25.09
18	1:5	-20.25
19	2:5	- b
20	1:5	-18.91
21	2:5	-20.89
<b>Solution phase product<sup>c</sup></b>	-	-24.35

*a* - numbers refer to experiments listed in Table 2.5; *b* - insufficient 86 was loaded on the resin to obtain an [ $\alpha$ ]<sub>D</sub> reading; *c* - synthesised by hydrolysis of 78.

All reactions were carried out at 30 °C. *N*-TFA-asp (86) for these measurements was cleaved using 90:5:5 TFA/H<sub>2</sub>O/TES and recrystallised from ether.

The use of sonication in the attachment of compounds to solid-supports has been reported in the literature.<sup>293</sup> To pursue this method a loading experiment was carried out overnight at room temperature using 1 equivalent of DMAP and sonicating the reaction for three bursts of 15 minutes. This produced a loading of 26%, which appeared quite efficient for a relatively short reaction time. However, an attempt to repeat this method, sonicating the resin and anhydride for longer failed to produce any loading. There was insufficient time in these studies to optimise the amount of sonication that can be tolerated by the resin, before cleavage of the loaded amino acid occurs. No studies were made at this time of the regioselectivity and racemisation effects of this technique, but sonication remains a promising method for efficiently loading amino acids onto Wang resin.



Table 2.7: Effects of varying reaction conditions on the regioselectivity of Wang loading.

Experiment <sup>a</sup>	Temperature	Equivalents of DMAP	Ratio of $\beta$ - vs $\alpha$ -Me ester		Ratio of $\alpha$ - vs $\beta$ -Wang
			<sup>1</sup> H solution-phase NMR <sup>b</sup>	<sup>19</sup> F solution-phase NMR <sup>c</sup>	<sup>19</sup> F gel-phase NMR <sup>d</sup>
8	45 °C	none	3.1:1	-	-
9	45 °C	1	3.6:1	3.4:1	-
10	60 °C	1	4:1	-	-
12	rt	2	5.2:1	-	-
22	rt	1	4.1:1	3.5:1	4.6:1
23	rt	2	2.4:1	3.2:1	11.2:1
24	30 °C	1	3.8:1	3.6:1	4.1:1
25	30 °C	2	3.0:1	3.9:1	2.1:1
26	60 °C	1	4.3:1	3.5:1	8.6:1
27	60 °C	2	2.2:1	1.8:1	1.1:1

*a* - experiment number refers to the number in Table 2.5; *b* - calculated from the ratio of the two methyl peaks at 3.63 (3 H, s,  $\beta$ CO<sub>2</sub>CH<sub>3</sub>) and 3.70 (3 H, s,  $\alpha$ CO<sub>2</sub>CH<sub>3</sub>); *c* - calculated from the relative integral values of <sup>19</sup>F signals for **91** at -76.61 (F<sub>3</sub>C on  $\alpha$ -ester) and -76.66 (F<sub>3</sub>C on  $\beta$ -ester); *d* - calculated from the ratio of signals for **87** at -74.27 (F<sub>3</sub>C on  $\beta$ -ester), -74.91 (F<sub>3</sub>C on  $\alpha$ -ester)

Fig. 2.26 gives some examples of NMR spectra comparing the ratios of  $\alpha$ - vs  $\beta$ -substitution of **86** on Wang resin at different temperatures and equivalents of DMAP. Solution-phase spectra (in C<sup>2</sup>HCl<sub>3</sub>) are of the methylation-product of *N*-TFA-asp-Wang after cleavage from the resin. Gel-phase spectra (in C<sub>6</sub><sup>2</sup>H<sub>6</sub>) are of *N*-TFA-asp-Wang. The spectra shown are of top: experiment 22, middle: 24 and bottom: 25.

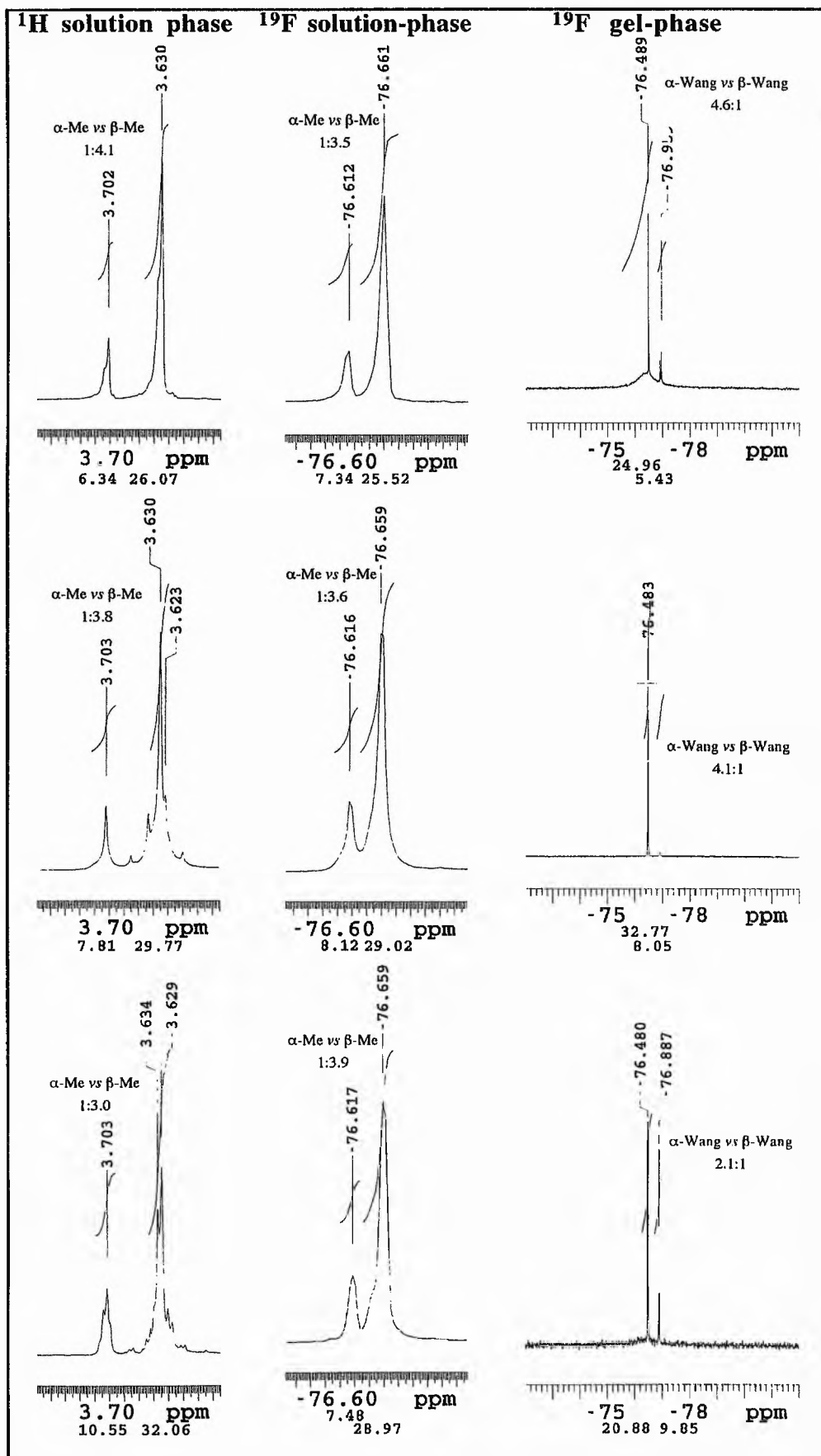


Figure 2.26: Examples of NMR spectra comparing ratios of  $\alpha$ -vs  $\beta$ -substitution of 86 on Wang.

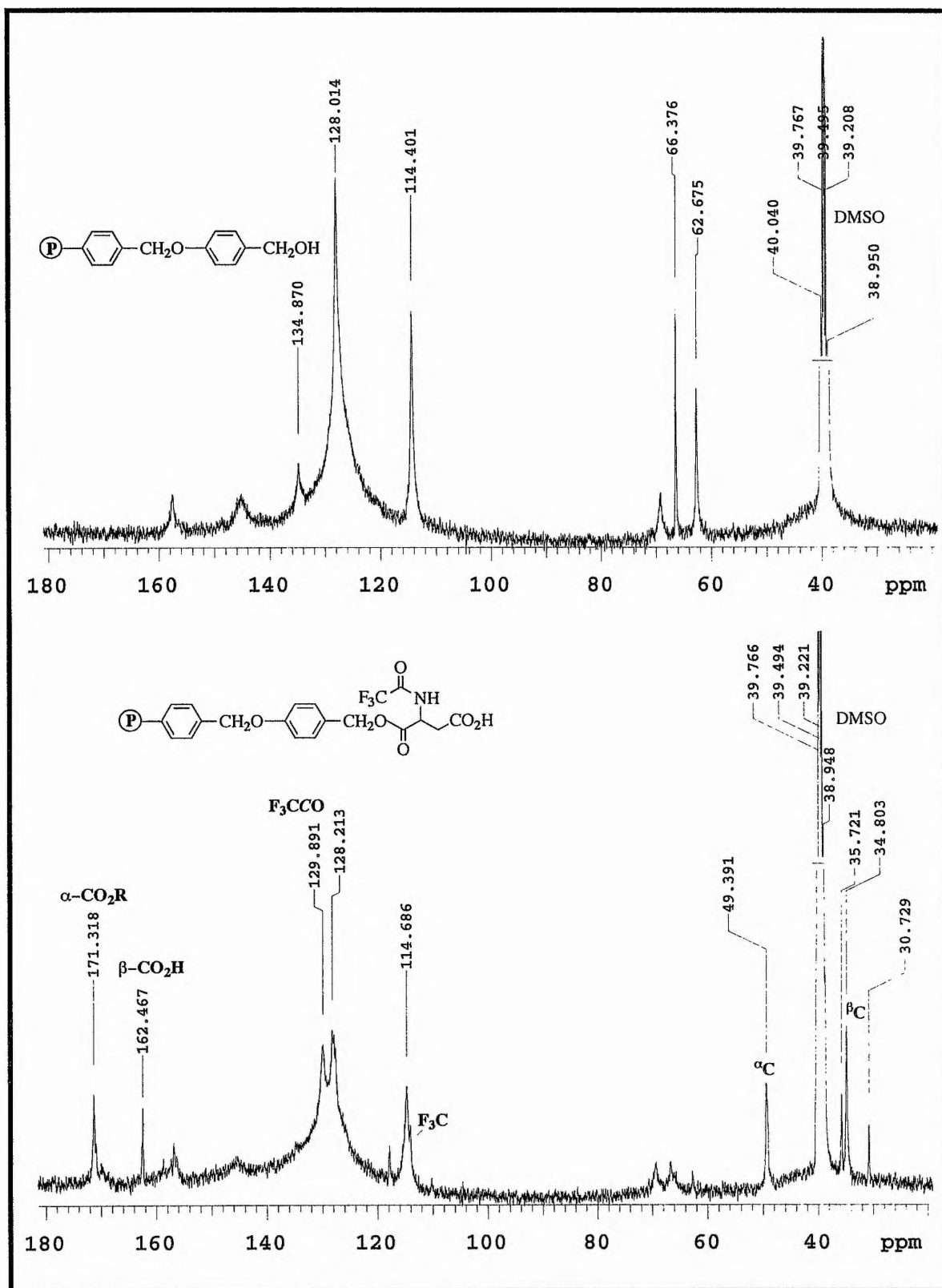
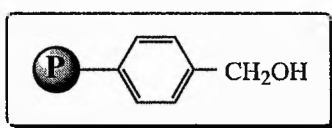


Figure 2.27:  $^{13}\text{C}$ -gel-phase NMR spectra of top: Wang resin, bottom: 87, both in DMSO.

### 2.5.3.2 *p*-Hydroxymethylpolystyrene Resin

*N*-trifluoroacetyl-(2*S*)-aspartate (**86**) was also successfully loaded onto *p*-hydroxymethylpolystyrene resin (**92**) using the protocol developed for Wang resin. Preliminary small-scale experiments to optimise the loading temperature established that the reaction must be heated to 35 °C or above (Table 2.8). Unfortunately the nucleophilic substitution reaction appeared to proceed more slowly with this resin, so this route to the solid-phase synthesis of the substrate (**32**) was not pursued.



*p*-Hydroxymethylpolystyrene resin (**92**).

**Table 2.8:** Conditions for loading *p*-hydroxymethylpolystyrene resin.

Experiment	Temperature	Time (hours)	Equivalents of DMAP	Equivalents of Anhydride	Stirring	% Loading <sup>a</sup>
1	0 °C	170	none	5	X	none
2	0 °C	"	1	"	X	none
3	rt	"	none	"	✓	none
4	rt	"	1	"	✓	none
5	35 °C	"	none	"	✓	none
6	35 °C	"	1	"	✓	36

*a* - calculated by comparing mass of resin before and after loading experiment (values are approximate)

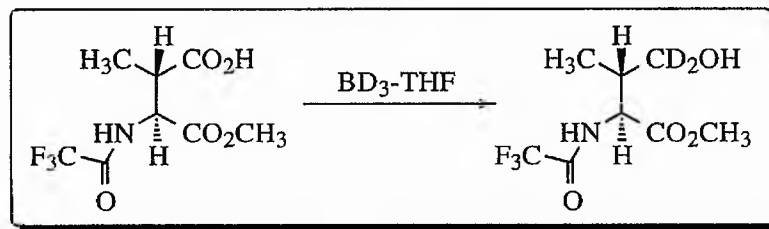
The loading of **86** onto *p*-hydroxymethylpolystyrene resin was indicated by several new peaks on the IR spectrum:  $\nu_{\max}(\text{KBr})/\text{cm}^{-1}$  1735 s (C=O, aryl ester), 1541 (2° amide, solid state) and 1170 s (C-F or C-O ester). <sup>19</sup>F and <sup>13</sup>C gel-phase NMR spectra showed several new signals compared to the resin alone, further indicating the loading of **86** onto the resin:

{ $\delta_{\text{C}}$ (74.76 MHz;  $(\text{C}^2\text{H}_3)_2\text{SO}$ ) 171.37 ( $\alpha\text{CO}_2\text{R}$ ), 162.58 ( $\beta\text{CO}_2\text{H}$ ), 157.00 ( $\text{F}_3\text{CCO}$ ), 117.87 ( $\text{F}_3\text{C}$ ), 49.45 ( $\alpha\text{C}$ ), 34.79 ( $\beta\text{C}$ );  $\delta_{\text{F}}$ (282.2 MHz;  $\text{C}^2\text{HCl}_3$ ) -74.87 ( $\text{F}_3\text{C}$ )}

### 2.5.3.3 Attempted Reduction of $\alpha$ -*p*-Benzyloxybenzylpolystyrene-*N*-trifluoroacetyl-(2*S*)-aspartate

Once the procedure for loading **86** onto Wang resin had been established, attempts were made to reduce the free  $\beta$ -carboxylic group of **87** to an alcohol using the mixed anhydride methodology developed previously for the solution-phase synthesis of **75**. The first attempt using this method led to cleavage of most of the *N*-trifluoroacetyl-(2*S*)-aspartate (**86**) from the resin. The failure of this method, and a desire to avoid the awkward step of removing excess NMM and IBCF under nitrogen *via* cannulae, led us to investigate a number of other methods for the selective reduction of carboxylic acid groups in the presence of both amide and ester functionalities.

The use of boranes in organic synthesis has been widely reported in the literature.<sup>294</sup> Some borane complexes can selectively reduce carboxyl groups to primary alcohols in the presence of other functionalities, without the need for prior activation of the acid group. A popular reagent is diborane in THF, which, according to the literature, reacts extremely rapidly and selectively with acids.<sup>295-299</sup> H. C. Brown *et al.* observed that aliphatic acids were reduced under mild conditions within 15-30 minutes and aromatic acids within 24 hours.<sup>295</sup> In fact, the relative reactivity of this reagent with various functionalities is almost the reverse of that of alkali metal borohydrides.<sup>294</sup> This is a result of the Lewis acid nature of diborane, which functions best as a reducing agent in attacking groups at positions of high electron density.<sup>295</sup> Of particular interest was a carboxylic acid reduction carried out by Kluender and coworkers in the presence of *N*-trifluoroacetyl and methyl ester groups (Scheme 2.9).<sup>300</sup>



**Scheme 2.9:** *Selective reduction of carboxylic group in the presence of N-trifluoroacetyl and methyl ester functionalities.*

Other reagents for the selective reduction of carboxylic acids to alcohols in the presence of esters and amides at low temperatures include sodium borohydride/ iodine, sodium borohydride/ catechol and sodium borohydride/ trifluoroacetic acid.<sup>301-303</sup> The latter could obviously not be used with Wang resin as TFA promotes the cleavage of compounds from the solid support. The action of diborane has also been found to be enhanced by the presence of catalytic amounts of sodium borohydride.<sup>294</sup>

These reducing reagents were tested for their efficacy under a number of different reaction conditions (Table 2.9). Unfortunately, in many cases, the amino acid was merely cleaved from the resin. After each reaction, any material remaining bound to the resin was cleaved using TFA/H<sub>2</sub>O/TES (90:5:5) and analysed by <sup>1</sup>H-NMR spectroscopy. Neither the acid group or the N-trifluoroacetyl amide group were reduced in any reaction as no H<sup>γ</sup> signals were observed and the NH doublet of **86** was still visible at 7.83 and 7.81 ppm in the <sup>1</sup>H-NMR spectrum.

It seemed possible that the ester linkage to the resin was much more labile than the α-isopropyl ester linkage of **74**, and that reducing agents were preferentially attacking this function.

**Table 2.9:** Conditions for attempted reductions of  $\alpha$ -p-benzyloxybenzylpolystyrene-N-trifluoroacetyl-(2S)-aspartate.

<i>Exp.</i>	<i>Solvent</i>	<i>Reducing Reagent</i>	<i>Temp.</i>	<i>Time</i>	<i>% cleavage</i>
1	DMF-THF	2.5 eq ea NMM and IBCF 2.5 eq NaBH <sub>4</sub>	-5 °C -5 °C/ rt	30 min 2 h/ O/N	65%
2	DMF/ THF	2.5 eq ea NMM and IBCF 2.5 eq NaBH <sub>4</sub>	0 °C 0 °C	2 h 24 h	-
3	THF	1.1 eq ea NMM and IBCF 0.6 eq NaBH <sub>4</sub>	0 °C 0 °C/ rt	30 min 3 h/ 14 h	35%
4	DMF-THF	5 eq ea NMM and IBCF 2.5 eq NaBH <sub>4</sub>	0 °C 0 °C	30 min 6 h/ 1.5 h	-
5	DME	5 eq ea NMM and IBCF 0.6 eq NaBH <sub>4</sub>	0 °C 0 °C	O/N 5 h	-
6	DME	5 eq ea NMM and IBCF 0.55 eq NaBH <sub>4</sub>	0 °C 0 °C/ rt	O/N 1.5 h/ 1.5 h	13%
7	DMF-THF	1.1 eq BH <sub>3</sub> -THF	0 °C	O/N	72%
8	DMF-THF	1.1 eq BH <sub>3</sub> -THF	0 °C/ rt	7 h/ O/N	-
9	THF	1 eq BH <sub>3</sub> -THF, 0.1 eq NaBH <sub>4</sub>	0 °C	5 h	-
10	DMF-THF	1.2 eq NaBH <sub>4</sub> , 0.5 eq I <sub>2</sub> (dark)	0 °C	14 h	7%
11	DMF	1.2 eq NaBH <sub>4</sub> , 0.6 eq I <sub>2</sub>	0 °C/ rt	5 h/ O/N	-
12	THF/ EtOH	1.2 eq NaBH <sub>4</sub> , 0.5 eq I <sub>2</sub> (dark)	0 °C	24 h	-
13	DMF	2.5 eq ea NMM and IBCF 2.1 eq N-selectride	0 °C 0 °C	30 min 2 h/ 30 min	39%

DMF was used in preliminary reactions due to its resin-swelling properties. When several reactions failed in this solvent, the solvent was changed to THF, as recommended in the solution-phase procedures. It was thought that the reducing reagents might be interacting with DMF, rendering them less effective. After reactions in THF also failed DME was selected as a possible alternative. Sodium borohydride is soluble in few organic solvents excepting alcohols and DME.<sup>294</sup> Experiments to reduce the protected aspartate **74** using this reagent were successful in THF probably because it was less essential to solvate the sodium borohydride as **74** was already in solution. In this case, however, the reaction was hindered by both the reducing agent and the substrate being in the solid-phase. DME also has the advantage of being slightly more dipolar than THF, and therefore has better resin swelling properties.

Reactions had been carried out successfully in the group previously using NMM and IBCF with compounds on Wang resin. The results obtained here were therefore difficult to explain.

In order to better understand the mechanism by which **86** was being cleaved, a number of control experiments were carried out using just these two reagents (Table 2.10). Some of these reactions resulted in significant proportions of **86** being cleaved from the resin.

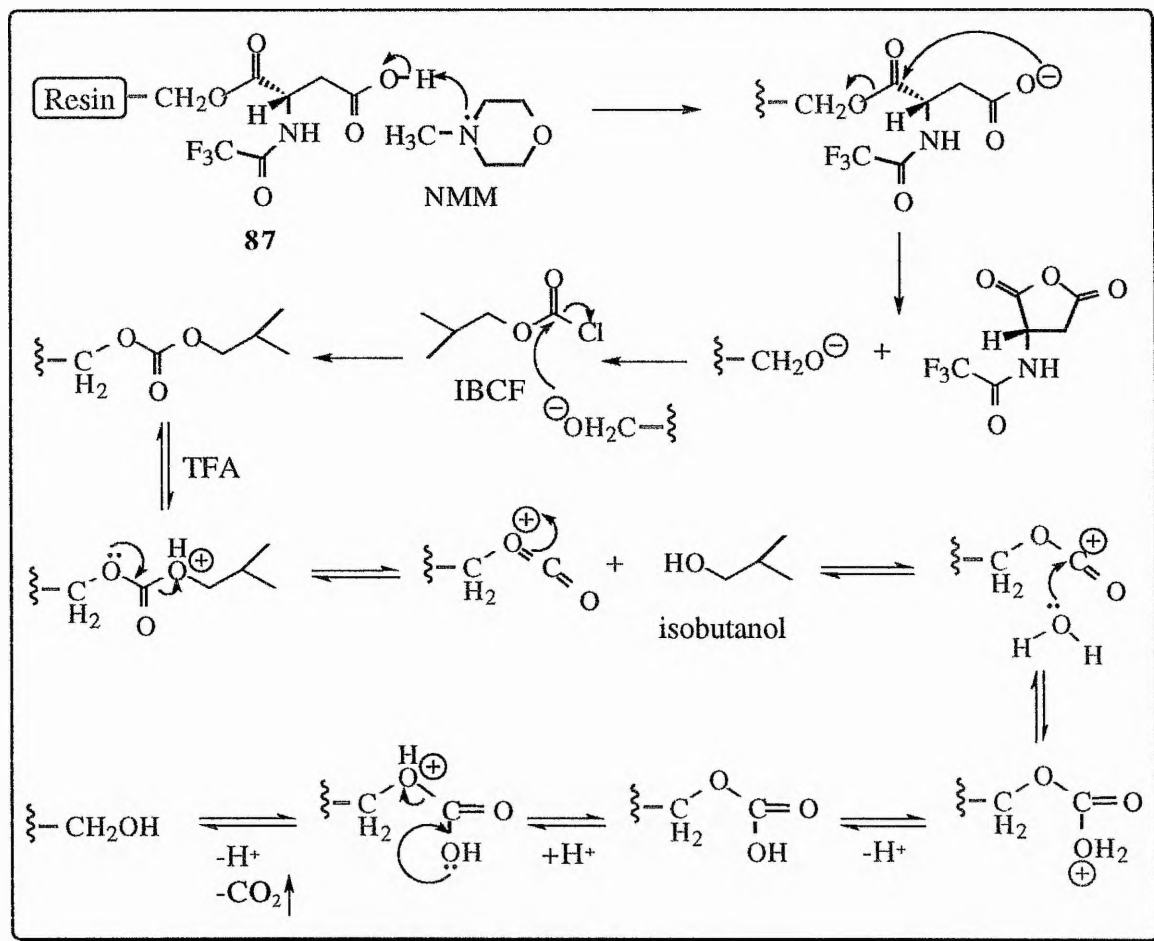
The washings from experiments 3-6 were analysed by  $^1\text{H}$ - and  $^{19}\text{F}$ -NMR spectroscopy and all were found to contain at least some traces of **86**. These findings, and the elimination of **86** caused by NMM alone (experiments 2, 5 and 6, Table 2.10), indicated that the cleavage of **86** from the resin could be induced by the presence of base. It was also noted that, in some early experiments using NMM and IBCF, the  $^1\text{H}$ -NMR spectrum of material cleaved from the resin displayed multiplet signals at around 2 and 4 ppm, indicative of an isobutanol moiety derived from IBCF. On the basis of these findings the elimination mechanism shown in Scheme 2.10 was proposed.

**Table 2.10:** Experiments to find the reagent(s) responsible for cleavage of **86** from the resin.

<i>Exp.</i>	<i>Reagent</i>	<i>Equivalents<sup>a</sup></i>	<i>Time</i>	<i>% cleavage</i>
1	NMM	5	3 h	none
2	NMM	5	4	84.5
3	(i) NMM	5	2 h	40
	(ii) IBCF	5	2.5 h	
4	(i) NMM and IBCF	5 of each	30 min	54.4
	(ii) DIEA	2	3.5 h	
5	NMM	2.5	30 min	49.7
6	NMM	2.5	12 h	32

All reactions listed in this table were carried out at 0 °C in DME. a - relative to loading of **86** on the resin.



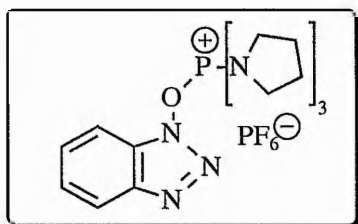


**Scheme 2.10:** Proposed scheme for the reaction of **87** with NMM and IBCF.

The postulated elimination mechanism shown in Scheme 2.10 was confirmed when experiments were carried out to provide evidence for the activation of the carboxylic acid group (Section 2.5.3.4). It was found that if the bases used for these reactions were added gradually in dilute solution, this resulted in less compound being cleaved from the resin. IBCF or the coupling reagent PyBOP (**93**) were also added before any base. This meant that as small amounts of carboxylate ion were formed, it reacted preferentially with the relative excess of activating agent, leaving little carboxylate available to cause elimination/ cleavage.

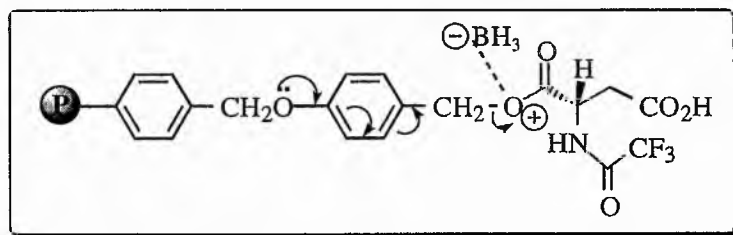
It was assumed that, because the formation of a carboxylate ion in **74** did not lead to cleavage of the  $\alpha$ -isopropyl group, the carboxylate ion formed by **87** would also not attack an ester function. There are two possible explanations for the difference in reactivity between these two

esters. Firstly it may be that the ester linkage is weaker in **87** than in **74**. Another possible explanation is based on the different phases of **87** and **74**. The carboxylate ion formed by **74** is fully solvated. The carboxylate ion of **87**, on the other hand, cannot interact to such an extent with the solvent, and is therefore much more nucleophilic towards the ester linkage.



*Benzotriazolyl-oxy-tris[pyrrolidino]-phosphoniumhexafluorophosphate (PyBOP, 93)*

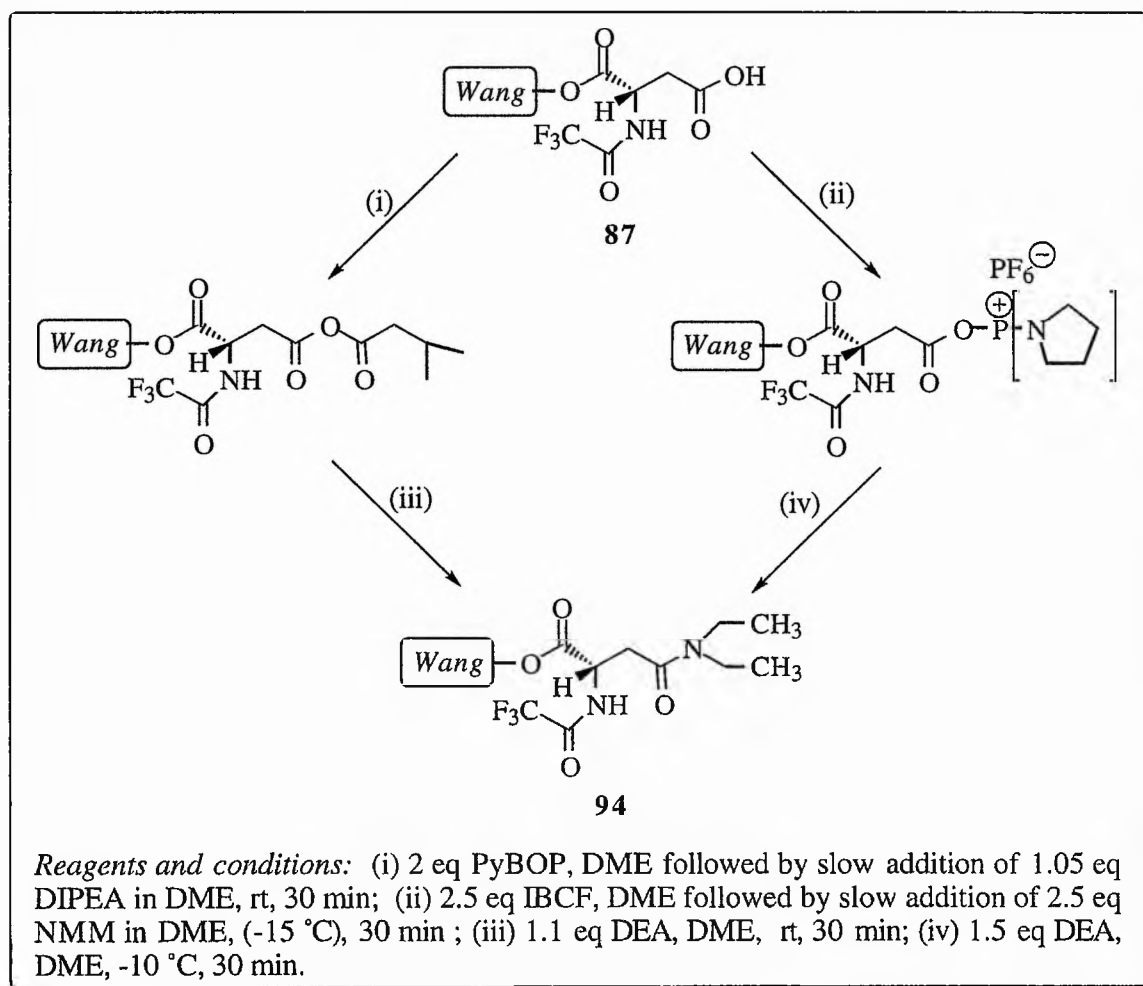
Although the mechanism given in Scheme 2.10 could explain the cleavage of **86** in the presence of NMM, there was also significant cleavage of the compound in the presence of diborane. This cleavage must therefore occur *via* a different mechanism. One possibility is that, due to its Lewis acidic nature, diborane readily coordinates with the oxygen atom linking **86** to the resin, which further weakens the ester bond (Scheme 2.11) One advantage of using *p*-hydroxymethylpolystyrene resin (**92**) instead of Wang resin, is that this elimination mechanism would not be possible.



**Scheme 2.11:** Proposed mechanism for the diborane-induced cleavage of **86** from Wang resin.

### 2.5.3.4 Experiments to Confirm the Activation of the Carboxylic Group

Although the mixed anhydride methodology resulted in some cleavage of **86** from the resin, if fewer equivalents of IBCF and NMM were used, and the reaction time for the activation step was minimised to 15 minutes, then at least half of the *N*-TFA-asp would remain on the resin. In order to establish the reason why no reduction of the  $\beta$ -carboxylic acid group had occurred, it was necessary to determine whether the mixed anhydride was formed at all, or whether some other means were necessary to activate the  $\beta$ -CO<sub>2</sub>H group. To do this simple coupling experiments were devised, using diethylamine to attack any activated acid. Reactions were conducted in DME at 0 °C, and the bases used were diluted in DME and added slowly over 15 minutes. These reactions are summarised in Scheme 2.12.



**Scheme 2.12:** Coupling reactions with diethylamine.

Both methods successfully produced the amide **94**, although a consistent yield could not be obtained, and some cleavage still occurred in nearly all cases. A quantitative yield was obtained in one case using PyBOP. The PyBOP activation also produced the cleaner product: For characterisation the amide was cleaved from the resin using TFA/H<sub>2</sub>O/TES (90:5:5) (*m/z* (EI) 284 (42%, M<sup>+</sup>)). The <sup>1</sup>H-NMR spectrum showed signals for a carboxylic group {172.29 (αCO<sub>2</sub>H)} and an amide {169.91 [βC(O)N]} and two sets of signals corresponding to the ethyl groups {42.52 and 40.89 (CH<sub>3</sub>CH<sub>2</sub>N, 2 conformations) and 13.50 and 12.46 (CH<sub>3</sub>CH<sub>2</sub>N, 2 conformations)}.

#### 2.5.3.5 Reduction of α-*p*-Benzyloxybenzylpolystyrene-*N*-trifluoroacetyl-(2*S*)-aspartate

Having established two methods of activating the acid group and the way to minimise cleavage of **86** from the resin, a number of different reducing agents were tested (Fig. 2.28). Some more hindered reducing agents were tried, as these might have greater selectivity toward the activated acid rather than the ester linkage. PyBOP was used in most cases to activate the β-carboxyl group, as indicated in Scheme 2.12. The results of these experiments are summarised in Table 2.11. The success or failure of the reduction reactions was monitored by cleaving any remaining material from the resin using TFA/H<sub>2</sub>O/TES (90:5:5), and examining the product by <sup>1</sup>H-NMR spectroscopy. A successful reduction was indicated by an upfield shift in the signals for H<sup>β</sup> to approximately 1.9 ppm and the appearance of new signals for H<sup>γ</sup> at around 3.3 ppm {*cf.* <sup>1</sup>H-NMR spectrum of **86**, Fig. 2.29}. Although a number of reagents managed to produce some level of reduction, the amount of **86** eliminated from the resin was still too high. Any excess DIPEA and PyBOP or NMM and IBCF was removed prior to the addition of the reducing agent to reduce the opportunity for base-catalysed cleavage/ elimination of **86**. It seemed, therefore, that there was another mechanism *via* which elimination could occur. Experiments were carried out in which Wang resin, or **87** were stirred overnight with 0.6 equivalents sodium or lithium borohydride to check their stability towards these reagents.

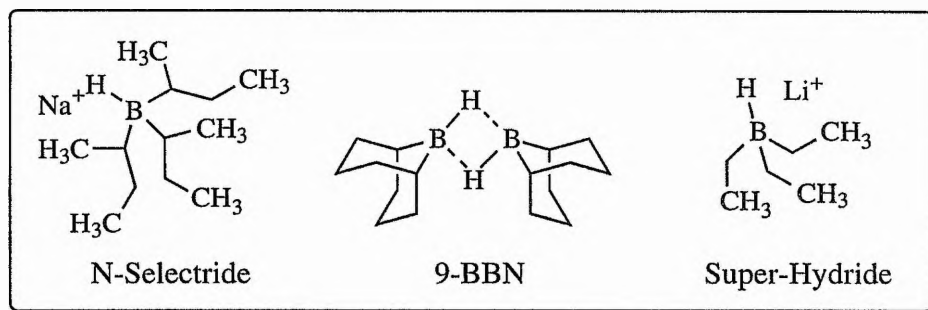


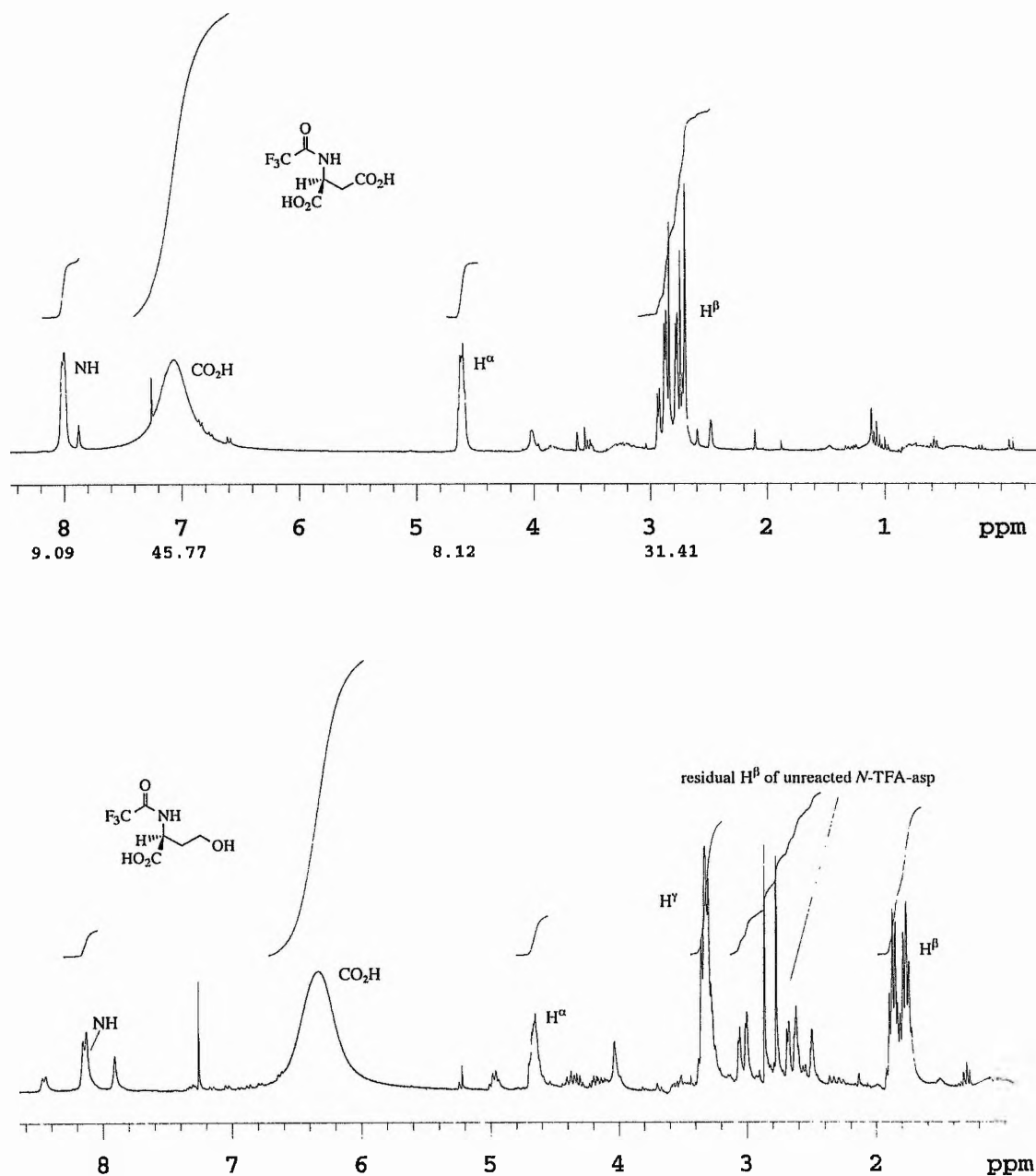
Figure 2.28: Hindered reducing agents.

Table 2.11: Reagents tried for the reduction of the activated  $\beta$ -carboxyl group.

Reducing Agent	Equivalents	Temperature	Time	Reduction	% Cleavage
N-selectride	3	-15 °C/ rt	1 h/16 h	x	95
9-BBN	3	-15 °C/ rt	20 min/ 25 h	x	95
Super-Hydride	3	-15 °C/ rt	30 min/ 11 h	✓	-
Super-Hydride	2.5	0 °C/ rt	2.5 h/ 1 h	-	100
BH <sub>3</sub> -THF <sup>a</sup>	1.1	- 10 °C/ rt	50 min/ 21 h	x	47
NaBH <sub>4</sub> <sup>b</sup>	2.5	0 °C	12 h	minor	24
NaBH <sub>4</sub>	2.5	0 °C/ rt	4 h/ 10.5	-	92
NaBH <sub>4</sub>	0.6	rt/ 30 °C	11 h/ 10 h	✓	95
LiBH <sub>4</sub> <sup>b</sup>	1.1	0 °C	1.5 h	✓	27
LiBH <sub>4</sub>	0.6	rt	18 h	✓	13
LiBH <sub>4</sub>	0.6	rt/ 30 °C	10.5 h/ 10 h	✓	95

*a* - non-activated  $\beta$ CO<sub>2</sub>H, *b* - NMM/ IBCF activation

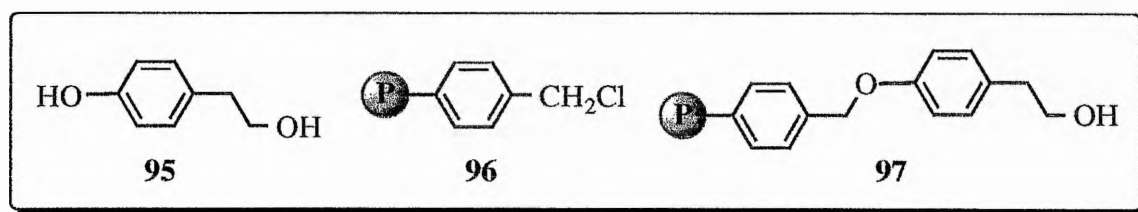
The results of the control experiments with sodium and lithium borohydride were not entirely conclusive. There was some loss of mass from both the Wang resin and **87**, although some of this may be due to losses incurred during the washing procedure. At this stage attempts to synthesise **32** using Wang resin were abandoned in favour of searching for a new solid-support system which would prove more stable to reducing agents.



**Figure 2.29:** Top:  $^1\text{H-NMR}$  spectrum of material cleaved from 87 (i.e. 86), bottom:  $^1\text{H-NMR}$  spectrum of material cleaved from the resin after treatment of 87 with PyBOP/DIPEA then LiBH<sub>4</sub>.

### 2.5.4 A New Solid Support

The problems encountered when trying to reduce the free acid group of **86** were mainly due to the ease with which the aspartate moiety underwent base-catalysed elimination from the resin. A new linker, **95** was obtained for attachment to chloromethylpolystyrene-divinylbenzene (Merrifield) resin (**96**) which would produce a resin similar to Wang, but with an ethyl rather than a methyl alcohol attachment (**97**). This would ensure that the ester linkage from the resin to **85** would not be in conjugation with the aromatic ring system on the resin, and would leave the  $\beta$ -carboxyl group less hindered by the bulky resin. It was envisaged that the synthesis of **32** using the new resin would proceed as originally intended with Wang resin (Scheme 2.8).



*The new linker, 2-(p-hydroxyphenyl)-ethanol (**95**), Merrifield resin (**96**) and the new resin (**97**).*

#### 2.5.4.1 Attachment of 2-(p-Hydroxyphenyl)-ethanol to Merrifield Resin

A number of methods were tried for the attachment of the alcohol named above to Merrifield. These have been summarised in Table 2.12. It was preferable to keep the reaction temperature low to prevent damage to the resin. The best method was found to be the use of 2.5 equivalents of the alcohol and 1 equivalent each of  $K_2CO_3$  and potassium iodide in acetonitrile at room temperature (97% loading after approximately 40 hours). However the method most often used, for convenience, involved 2.5 equivalents each of the alcohol and powdered potassium hydroxide in acetonitrile at 35 °C (97% after 4 days, see Scheme 2.13). A catalytic amount of 18-crown-6 for solvating cations was essential for either method to succeed. Attempts were made to monitor the progress of reactions by UV, but it was found

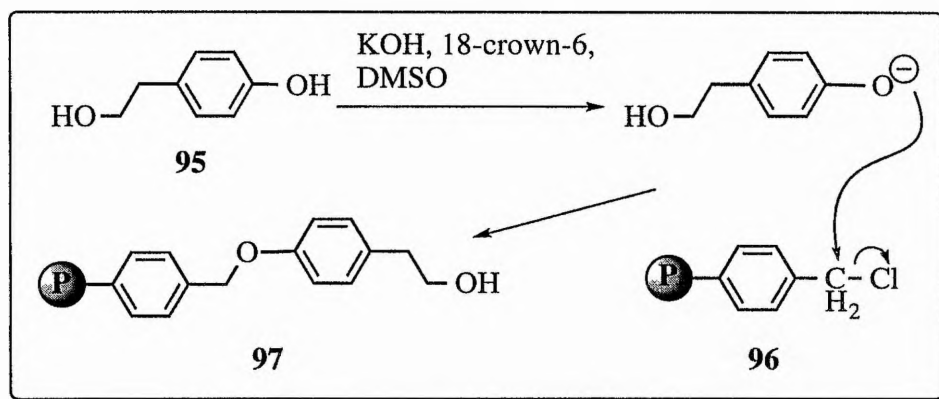
that the absorbance values of the various phenoxide species were too different from that of the phenol to produce coherent readings.

**Table 2.12:** Conditions used in experiments to load 95 onto Merrifield resin.

	Equivalents of (96)	Other Reagents	Solvent	Temp (°C)	Time (hours)	% Loading <sup>a</sup>
1	10	10 eq KOH (solid)	DMSO	23	290	none
2	10	10 eq KOH	DMSO	35	192	none
3	5	5 eq KOH	DMSO	23/ 100	30/ 17.5	99
4	2.5	2.5 eq KOH, 18-crown-6 (cat)	CH <sub>3</sub> CN	40	85.5	97
5	5	2 eq Et <sub>4</sub> NF	DMF	40	46	none
6	5	2.5 eq CsF	DMF	35	24	none
7	5	2.5 eq CsF	DME	35	24	none
8	10	10 eq NaH	DMF	23	290	none
9	10	2 eq NaH	DMF	35	192	none
10	2.5	1 eq NaH, 18-crown-6 (cat)	CH <sub>3</sub> CN	40	85.5	15
11	5	1 eq NaH, 18-crown-6 (cat)	THF	40	160	21
12	5	1 eq NaH, 18-crown-6 (cat)	THF	23/ 67	46/ 17.5	47
13	5	1.5 eq NaOEt/ EtOH	THF	40	42	none
14	5	1.5 eq NaOEt/ EtOH	THF	23/ 67	30/ 10.5	none
15	2.5	1.5 eq NaOEt/ EtOH, 18-crown-6 (cat)	CH <sub>3</sub> CN	40	85.5	none
16	5	1 eq K <sub>2</sub> CO <sub>3</sub>	DMF	40	46.5	69
17	2.5	1 eq K <sub>2</sub> CO <sub>3</sub> , 1 eq KI, 18-crown-6 (cat)	CH <sub>3</sub> CN	23	41.5	97
18	5	1 eq K <sub>2</sub> CO <sub>3</sub> , 1 eq KI, 18-crown-6 (cat)	CH <sub>3</sub> CN	23	41.5	70
19	2.5	1 eq K <sub>2</sub> CO <sub>3</sub> , 1 eq KI, 18-crown-6 (cat)	CH <sub>3</sub> CN	30	41.5	53
20	5	1 eq K <sub>2</sub> CO <sub>3</sub> , 1 eq KI, 18-crown-6 (cat)	CH <sub>3</sub> CN	30	41.5	50
21	5	1 eq K <sub>2</sub> CO <sub>3</sub> , 1 eq KI, 18-crown-6 (cat)	CH <sub>3</sub> CN	40	46.5	82

*a* - calculated by comparing the mass of the resin before and after loading experiments



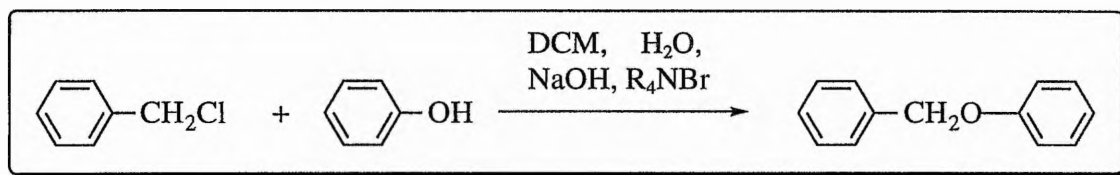


**Scheme 2.13:** The KOH-mediated loading of 95 onto Merrifield.

According to the literature, the best general method for the preparation of unsymmetrical ethers is the *Williamson reaction*, discovered in 1850. The normal method involves treatment of the halide with an alkoxide or aroxide ion derived from an alcohol or phenol. A variation on this procedure was reported in 1979 by Johnstone and Rose, involving the combination of the alkyl halide and alcohol with powdered potassium hydroxide in DMSO at room temperature.<sup>309</sup> Although they only describe a method involving alkyl bromides and iodides, it seemed plausible that Merrifield resin (an alkyl chloride) might also undergo reaction under these conditions. In our studies using Merrifield resin, however, no detectable reaction occurred at room temperature or 35 °C. It was necessary either to raise the reaction temperature to 100 °C or use a crown ether (18-crown-6) to solvate cations. A number of factors were producing this slow reaction. The chloride ion is a poorer leaving group than bromide or iodide ions. KOH is also only slightly soluble in DMSO, hence the need for the crown ether. This problem was alleviated by using acetonitrile as it is slightly more polar than DMSO. Another factor hindering the reaction is the bulky nature of the resin. It was noted in the paper that more hindered alkyl groups, not surprisingly, reacted more slowly and that this necessitated a higher reaction temperature.

Phase transfer catalysis has been shown to promote efficient reaction of alcohols and alkyl halides. Freedman and Dubois reported an improved Williamson ether synthesis using

tetrabutylammonium bisulphate (TBAB).<sup>305</sup> McKillop *et al.* also used a quaternary ammonium salt as a phase transfer catalyst in the synthesis of phenol ethers.<sup>306</sup> Of particular interest was the reaction shown in Scheme 2.14, which is similar to what we were trying to achieve, which produced a yield of 86% of the desired product.



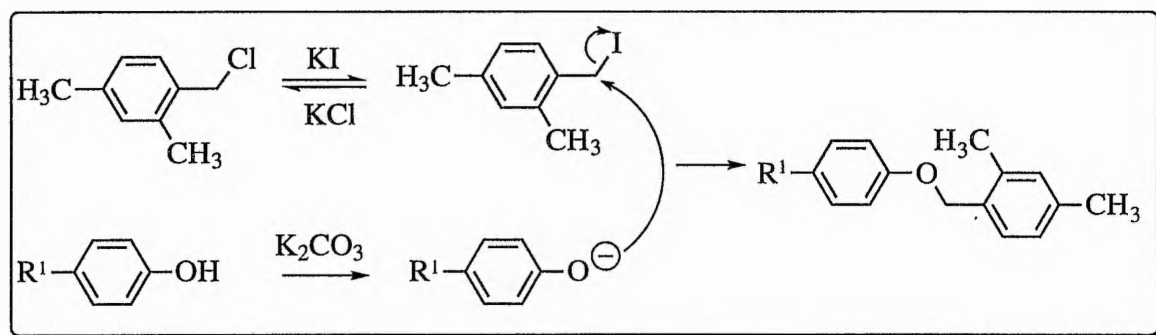
**Scheme 2.14:** McKillop's ether synthesis.

J. M. Miller *et al.* also reported on the use of a quaternary ammonium halide for the promotion of the synthesis of alkyl phenyl ethers, including the product shown in Scheme 2.14.<sup>307</sup> He used tetraethylammonium fluoride in DMF as a source of fluoride ions. The fluoride ion can act as a powerful hydrogen bond donor. In an article published in 1977 it was shown that phenol and fluoride can form a strong H-bond together, resulting in a large shift of the hydroxyl stretching band in the IR spectrum  $\{\Delta\nu_s(\text{OH}) \sim 1000 \text{ cm}^{-1}\}$ .<sup>308</sup> An attempt was made to use this reagent for promoting the alkylation (experiment 5, Table 2.12), but to no avail. Again steric hindrance may be the reason why this reaction failed and a higher reaction temperature may prove necessary. In the same paper Miller reports that the condensation of phenol with halogenoalkanes is promoted by the presence of  $\text{KF}$ .<sup>308</sup> The reaction is accelerated by an order of magnitude if  $\text{CsF}$  is used instead. Although, in our hands, this reagent failed to produce a successful reaction, this may be attributed to the low solubility of the alkali metal fluoride in either of the solvents used. Again the use of 18-crown-6 and a higher reaction temperature would probably have solved this problem. The use of fluoride-containing reagents was not pursued further, however, because other, more cost-effective methods had produced satisfactory results.

The use of sodium hydride to form an alkoxide ion which could react with benzyl bromide was reported by D. J. Hart *et al.* in their work on the total synthesis of lythrancepines.<sup>309</sup> 3.6

equivalents of NaH were added to the alcohols in DMF at room temperature. After 1 hour benzyl bromide was added and the reaction left for 36 hours. Harsher reaction conditions were necessary here to produce a reaction between **95** and Merrifield resin (**96**). The best loading (47%) was obtained by refluxing the reagents in THF for over 60 hours in the presence of a catalytic amount of 18-crown-6. F. Kroll and D. Stones (unpublished results) have achieved some success with the loading of various phenols onto Merrifield resin using NaH in DMF at 60 °C.

By far the best method found for loading **95** onto Merrifield was a protocol developed originally for the protection of phenols with a functionality that would be stable to hydrogenolysis at atmospheric pressure (Scheme 2.15).<sup>310</sup> This involved the use of potassium carbonate and potassium iodide in acetonitrile with a phenol and the alkyl chloride. The iodide replaces the chloride group on a small proportion of the alkyl chloride. This acts as a better leaving group when attacked by a phenoxide ion, thus accelerating the reaction. A reaction carried out using just K<sub>2</sub>CO<sub>3</sub> produced a good loading of **95** on the resin (69%), but was not as efficient as the method using KI in conjunction with K<sub>2</sub>CO<sub>3</sub>.



**Scheme 2.15:** Mechanism for the reaction carried out by Davis and Muchowski.

The loading of **95** onto Merrifield was confirmed by the presence of several new peaks in the IR spectrum compared that for the resin alone:  $\nu_{\max}(\text{KBr})/\text{cm}^{-1}$  3453 br m (O-H, alcohol), 1517 m (aromatic C=C), 1500 (aromatic C=C), 1390 (O-H, alcohol). The <sup>13</sup>C-NMR spectrum also showed the presence of several new signals:  $\delta_{\text{C}}$ (74.76 MHz; C<sub>6</sub><sup>2</sup>H<sub>6</sub>) 158.55

(Ar-CO, quaternary), 130.68 (Ar-CH), 115.52 (Ar-CH), 70.33 (PhCH<sub>2</sub>), 64.12 (CH<sub>2</sub>OH) (Fig. 2.30). The conditions for the loading of **95** onto Merrifield using potassium hydroxide were optimised, as summarised in Table 2.13.

**Table 2.13:** *Optimisation of the conditions for the loading of 95 onto Merrifield using potassium hydroxide.*

	<i>Equivalents of (96)</i>	<i>Equivalents of KOH</i>	<i>Temp (°C)</i>	<i>Time (hours)</i>	<i>% Loading</i>
1	1	1	23	41	none
2	1	1	35	41	22
3	1	1	60	41	none
4	2.5	2.5	23	41	50
5	2.5	2.5	35	41	22
6	5	5	23	41	none
7	5	5	35	41	24
8	5	5	60	41	none

All reactions were conducted in CH<sub>3</sub>CN with a catalytic amount of 18-crown-6. Equivalents given are with respect to the original loading of the resin (1.0 mmol/g).

#### 2.5.4.2 Loading of the New Resin With *N*-trifluoroacetyl-(2*S*)-aspartate

Before attempting to load **85** onto the new resin, **97**, model reactions were carried out using 2-(4-methoxyphenyl)-ethanol (**98**) as a structural analogue of **97**. Reactions were carried out as for the loading of Wang and **91** with the anhydride **73** using varying equivalents of DMAP at different temperatures to determine how this affected the regioselectivity of the substitution. The ratio of signals on the <sup>19</sup>F-NMR spectra were used to determine the ratio of  $\alpha$ - vs  $\beta$ -ester (Fig. 2.31 and Table 2.14).

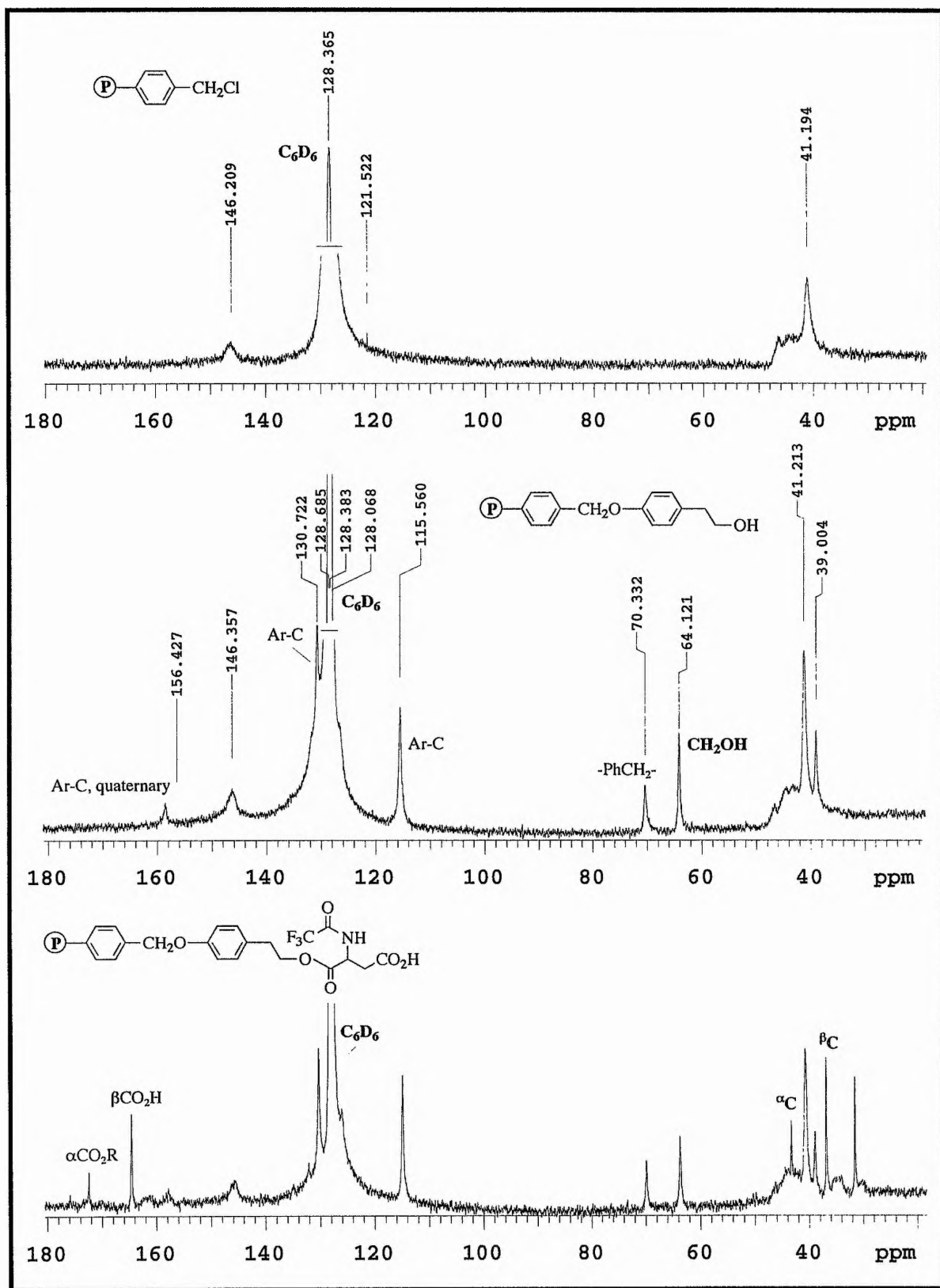
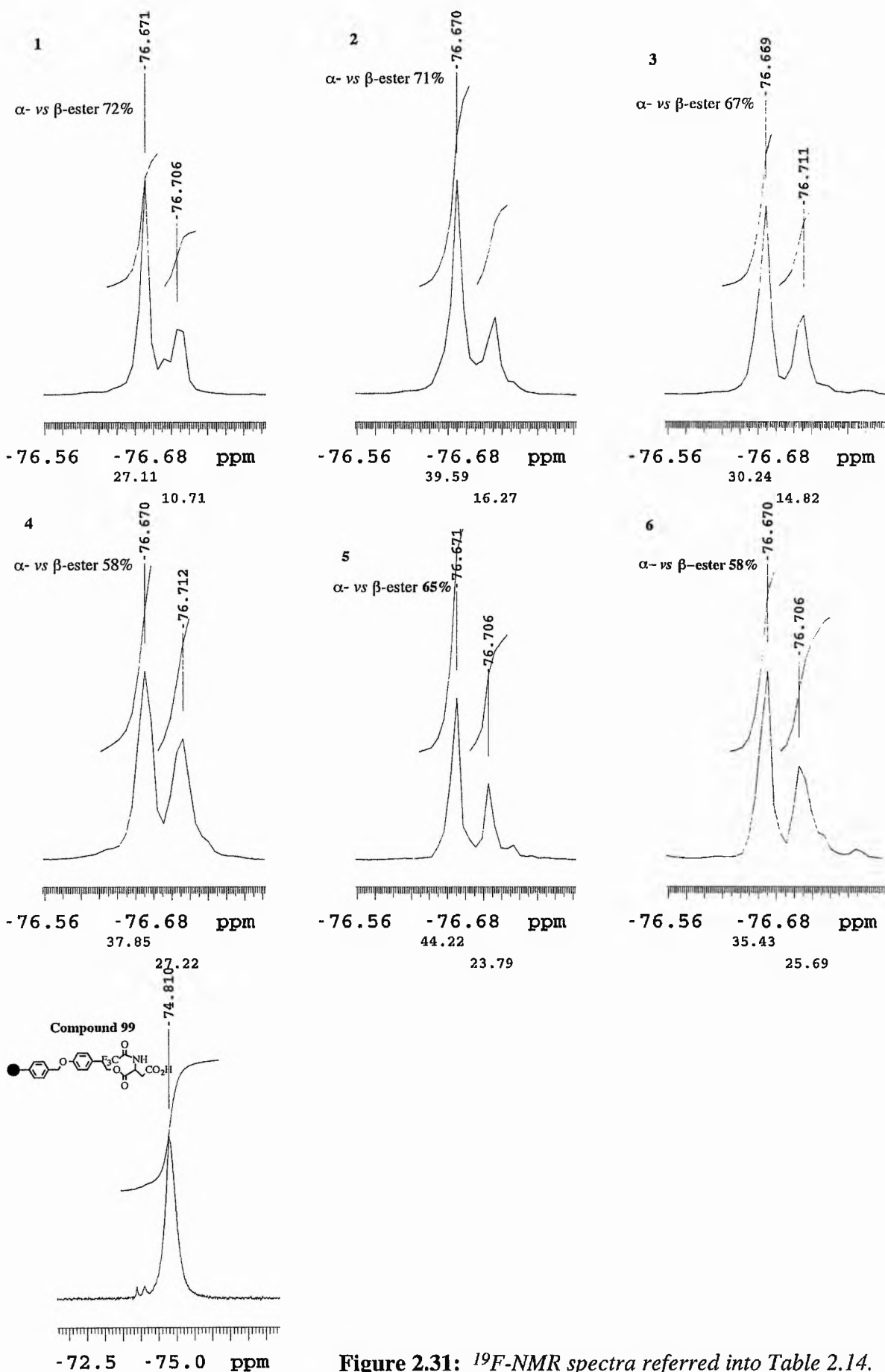
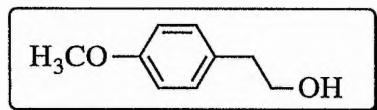


Figure 2.30:  $^{13}\text{C}$  gel-phase NMR spectra of top: Merrifield resin, middle: 97 and bottom: 99.

Figure 2.31:  $^{19}\text{F}$ -NMR spectra referred into Table 2.14.

2-(4-methoxyphenyl)-ethanol (**98**)**Table 2.14:** Comparison of the regioselectivity of substitution reactions of **73** onto **98**.

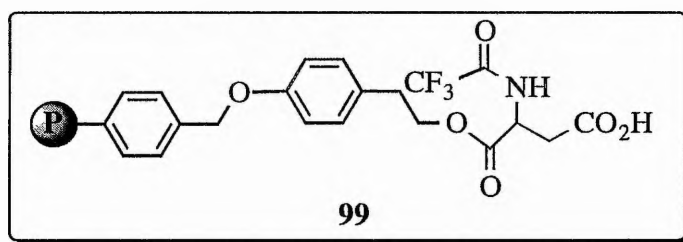
<i>Experiment</i>	<i>Equivalents of DMAP</i>	<i>Temperature</i>	<i>α- vs β-ester<sup>a</sup></i>
<b>1</b>	0	0 °C	72%
<b>2</b>	0.5	0 °C	71%
<b>3</b>	1.0	0 °C	67%
<b>4</b>	0	rt	58%
<b>5</b>	0.5	rt	65%
<b>6</b>	1.0	rt	58%

Reactions were carried out in DMF using 25 mg/ml **98** and 175 mg/ml **73**.

Reactions 3, 5 and 6 (Table 2.14) appeared to be complete within 17 hours as shown by tlc [band on baseline; eluant ethyl acetate-light petroleum (60%)], but reaction 2 took a few days to go to completion, and reactions 1 and 4 (no catalyst) were not completed even after 6 days. It was therefore obvious that the reaction to load **73** onto the new resin (**97**) would have to be performed at room temperature or higher using a minimum of 0.5 equivalents of DMAP.

The first two loading experiments using the new resin were carried out at 0 °C and room temperature respectively using 5 equivalents of the anhydride **73** and 0.5 equivalents of DMAP. Both were left for three days, but no protected aspartate **85** had loaded onto the resin as determined by comparison of the mass of the resin before and after the loading experiments. Both <sup>19</sup>F gel-phase NMR and IR spectroscopy confirmed that no *N*-trifluoroacetyl-(2*S*)-aspartate was present. A later experiment using the same equivalents of DMAP and **73** but increasing the temperature to 35 °C did not produce a significant loading after 4 days, by

comparison of the mass of the resin at the start and end of the experiment. However, both  $^{19}\text{F}$  and  $^{13}\text{C}$  gel-phase NMR spectroscopy indicated that some of the protected aspartate **85** had attached to the resin:  $\delta_{\text{C}}$ (74.76 MHz;  $\text{C}_6^2\text{H}_6$ ) 34.09 ( $\beta\text{C}$  of aspartate), 48.42 ( $\alpha\text{C}$  of aspartate), 117.44 ( $\text{F}_3\text{C}$ ), 169.11 ( $\text{F}_3\text{CCO}$ ), 173.77 ( $\beta\text{CO}_2\text{H}$  of aspartate);  $\delta_{\text{F}}$ (282.2 MHz;  $(\text{C}^2\text{H}_3)_2\text{SO}$ ) -74.81 ( $\text{F}_3\text{C}$ ). The IR spectrum showed the presence of an aryl ester linkage {1727 s (C=O)} and the *N*-trifluoroacetyl amide group {1653 m ( $2^\circ$  amide, solid state)}. The most recent experiment was carried out using 1 equivalent of DMAP and 5 equivalents of **73** at 35 °C. After 6 days the loading appeared to be quantitative, and the synthesis of **99** was confirmed by  $^{19}\text{F}$  and  $^{13}\text{C}$  gel-phase NMR and IR spectroscopy (Fig. 2.30 and 2.31). A method for cleaving **85** from the new resin has yet to be found. A number of methods for the selective cleavage of esters have been reported. Boron tribromide has been used for the deprotection of a number of ester protecting groups to the corresponding acids, without affecting amide bonds.<sup>311</sup> Iodotrimethylsilane has also been shown to be a mild and effective reagent for the non-saponificative hydrolysis of esters under mild conditions.<sup>312-314</sup> This can be synthesised *in situ* from chlorotrimethylsilane and sodium iodide.<sup>314</sup> This reagent does, however, also cleave ethers.





## 2.6 Conclusions and Future Work

### 2.6.1 Substrate Synthesis

#### 2.6.1.1 Solid-Phase Synthesis

The synthesis of both the non-labelled and [U-<sup>14</sup>C]-labelled substrate using previously established protocols has been reported here. *N,N*-diisopropyl dibenzyl phosphoramidite (**79**) has always been used in the past as a phosphorylating agent in this synthesis. However, problems can be encountered in the synthesis of this reagent, due to the instability of its precursor *N,N*-diisopropyl dichlorophosphoramidite (**78**). Two *new* reagents were therefore identified for the phosphorylation of *N*-trifluoroacetyl- $\alpha$ -isopropyl-(2*S*)-homoserine (**75**), one of which was available commercially (**80**), the other, **81**, being synthesised easily from POCl<sub>3</sub> and pentafluorophenol. These gave good clean yields of the phosphorylated compounds **83** and **84**. However, these phosphate groups could not be subsequently deprotected.

*N*-trifluoroacetyl-(2*S*)-aspartate (**86**) has been successfully loaded onto Wang, *p*-hydroxymethyl polystyrene (**92**) and polystyrene-4-oxymethyl-2-phenylethanol (**97**) resins during attempts to synthesise the substrate on the solid-phase. The latter resin is a *novel* resin, which was synthesised by the base catalysed addition of 2-(*p*-hydroxyphenyl)-ethanol (**95**) to Merrifield resin (**96**) in 97% yield. This was promoted by the use of KI to exchange chloride for the better leaving group, iodide, and by 18-crown-6 to solvate the ionic species. Both K<sub>2</sub>CO<sub>3</sub> and powdered KOH could be used as bases for this reaction. Acetonitrile was found to be the best of the solvents tried in this reaction.

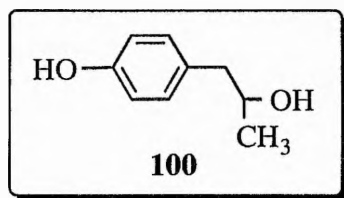
The nucleophilic addition reactions of each resin with **85** proceeded with good regioselectivity to favour protection of the  $\alpha$ -carboxylic group of **86**. Wang resin was the most selective for the  $\alpha$ -ester (84%  $\alpha$ - vs.  $\beta$ -ester). This is due to the more hindered nature of the hydroxyl group on Wang compared to **92** and **97**. This regioselectivity compares reasonably well with the ring opening of **85** by isopropanol, which produces a ratio of  $\alpha$ - vs.  $\beta$ -ester of 95%. The ring opening of **85** using resins does, however, proceed much more slowly, and requires

heating of the reaction to at least 30-35 °C, which in turn lowers the regioselectivity.

The compound produced from the reaction of **85** with Wang resin, **87**, was found to be unstable to reducing agents. It was thought that the compound produced using *p*-hydroxymethyl polystyrene (**92**) would also be unstable to reducing agents, as it resembles the structure of **87**. Resin **92** was, in any case, less selective for the  $\alpha$ -ester than Wang resin. The novel resin, **97**, was also less selective than Wang resin for the  $\alpha$ -ester (72%  $\alpha$ - vs.  $\beta$ -ester), but it is hoped that **98** will prove more stable to reducing agents.

### 2.6.1.2 Suggestions for Future Synthetic Work

The conditions for the loading of **86** onto **97** have yet to be optimised, and a mild and selective method for the cleavage of the ester linkage must be developed. It would be preferable to find a method which would simultaneously deprotect the dipentafluorodiphenylphosphate triester to minimise the number of deprotection steps necessary in the solid-phase synthesis of **32**. The possibility of reducing the  $\beta$ -carboxyl group on **99** using either the mixed anhydride or the PyBOP method of activation has yet to be explored, and it remains to be seen whether the new linker prevents the cleavage/elimination of **86** from the resin. It may be possible to use diborane with the new resin, as the ester linkage is no longer adjacent to a system of conjugated double bonds. If cleavage still occurs, more sterically hindered linkers such as **100** may be necessary, which resemble more closely the structure of the isopropyl group of **74**, which is stable to reduction. A more hindered linker may also improve the regioselectivity of the nucleophilic addition reaction in favour of the  $\alpha$ -ester.



If a resin is developed with enhanced stability towards reducing agents, this may have a wider application for general use in the solid-phase synthesis of peptides and amino acid derivatives.

## 2.6.2 A Viable Source of Threonine Synthase

### 2.6.2.1 The Tir8 Mutant

The mutant form of *E. coli* K-12 Tir8 is was first characterised by its resistance to thiaisoleucine (2-amino-3-methylthiobutyric acid).<sup>174</sup> The growth of wild-type K-12 cells is prevented by the presence of this compound, due to the inhibition of threonine deaminase activity. The growth inhibition is reversed by isoleucine at a concentration one-tenth that of the inhibitor. Tir mutants have altered isoleucyl tRNA synthetases and derepressed levels of three of the five enzymes involved in the biosynthesis of isoleucine and valine. These mutants are also slower growing than wild-type cells.

In Section 2.2 it was reported that the Tir8 mutant might have reverted to wild-type during experiments to isolate TS and so that it longer over-expressed the enzymes coded for by the *thr* operon. One of the possible reasons for this is that wild-type cells are faster growing than Tir8 cells and can rapidly overtake a growing culture. As Tir8 exhibits enhanced resistance to thiaisoleucine one method for detecting any remaining mutant cells is to plate out a fresh culture, grown from frozen glycerol stocks of Tir8, onto L-agar plates containing the compound and incubate these overnight. This would provide growth conditions selective for the mutant colonies, which could then be isolated and grown up in LB-medium for further attempts to isolate TS from this source. If this were still unsuccessful or if no mutants were isolated, then it would be necessary to immediately carry out more cloning work to insert *thrC* into a pET vector. Cloning the gene into such vectors does also have some advantages over the use of Tir8. Only *thrC* would be over-expressed in cells transformed with these constructs, facilitating the purification of the enzyme. A His-tagged version of the enzyme should also make the purification of TS much quicker and easier (see Section 1.11).

### 2.6.2.2 Cloning Amplified *thrC* Into pET Vectors

A protocol for the cloning of amplified gene fragments into pET vectors has been established and reported here. This involves cloning the gene first into pGEM-T from which it is subcloned into pET vectors. This first cloning step allows for the production of a large amount of the amplified gene DNA without the need for repeated PCR experiments, during which mutations can arise. Cloning of gene fragments with *Nde* I-sites at either terminus into pET vectors *via* ligation of cohesive termini does not appear to be possible, but fragments with a *Nde* I-site at one terminus and a *Bam*H I site at the other can be ligated into pET vectors restricted with these two enzymes.

It has been found that ligations using pET-vectors have a specific temperature requirement of 16 °C, whereas ligations into pGEM-T can tolerate warmer conditions (*i.e.* room temperature). The DNA insert must be extremely pure for ligation into pET expression systems. Successful results were obtained using PCR products and excised gene fragments which had been purified on Wizard Prep™ DNA purification columns and using recombinant plasmids which had been isolated using the QIAGEN maxiprep protocol and purified where necessary using Wizard™ DNA Clean-up Purification Resin.

Both the *thrC* gene and the gene coding for  $\beta$ -methylaspartase, containing *Nde* I-sites at each terminus, were amplified and cloned into the pGEM-T transcription vector. The identity of the gene inserts was identified by a number of means:

- Diagnostic PCR - whereby half as well as whole amplified gene fragments have been obtained using internal and external primers .
- Restriction analyses - carried out on the recombinant plasmids which gave distinct fragments of the predicted sizes.
- Subsequent di-deoxynucleotide sequencing of the recombinants which gave the sequence of the first 127 bases of the C-terminus of the *thrC* gene, including the *Nde* I-site, and the first

120 bases of both the C- and N-termini of the gene for  $\beta$ -methylaspartase. These gene fragments could not be subcloned into pET vectors and the constructs were therefore discarded before full sequencing could be carried out.

The two genes were amplified again to produce gene fragments with an *Nde* I-site at the N-terminus and a *Bam*H I-site at the C-terminus. This also improved the efficiency of ligations into pET vectors, as the orientation of gene inserts in the vector could now be controlled. These PCR products were cloned into pGEM-T. Two recombinants for *thrC* (pGEM-TTS<sub>I</sub> and pGEM-TTS<sub>II</sub>) were identified by a number of restriction analyses. Preliminary sequencing experiments produced correct sequences for the two termini of the gene in both constructs. Subsequent sequencing of the whole *thrC* gene insert in both constructs has revealed one deletion and two mutations of the recombinant gene (Section 2.4).

The *thrC* fragment was excised from pGEM-TTS<sub>I</sub> and pGEM-TTS<sub>II</sub> for subcloning into the three pET-vectors pET-3a, -3b and -16b. Recombinants were identified by restriction analysis. Sequencing of these vectors was also carried out (Section 2.4) confirming the gene mutations observed in the pGEM-TTS constructs.

The asymmetric  $\beta$ -methylaspartase PCR product was also inserted into pGEM-T, and further cloning work has been carried out by other researchers to insert the gene fragment into pET-vectors (Section 2.3.2.3).

#### 2.6.2.3 Future Work on *ThrC*

Since it has been determined that the recombinant *thrC* genes do not have the same sequence as the wild-type gene, if the Tir8 mutant is confirmed to have completely reverted to wild-type it will be necessary to repeat the cloning procedure which has been optimised during these studies.

*Taq* DNA polymerase lacks a 3' - 5' exonuclease proofreading activity, and therefore has a relatively high error rate. The mutation frequency using this enzyme is  $110 \times 10^{-4}$  at 10mM  $MgCl_2$  and 1mM dNTPs.<sup>265</sup> A mutation is therefore highly likely to occur after relatively few cycles of PCR using this enzyme.<sup>265</sup> *Vent*<sup>TM</sup> DNA polymerase has a fidelity 5 - 15 fold higher than that of *Taq*, having a mutation frequency of  $190 \times 10^{-6}$ , and is the only thermophilic polymerase with a 3' - 5' exonuclease proofreading activity.<sup>264, 265</sup> Therefore, in experiments to repeat the PCR amplification of *thrC*, it is obvious that *Vent* should be used instead of *Taq*. A new amplified *thrC* fragment should be produced with an *Nde* I-site at the N-terminus and a *Bam* H I-site at the C-terminus using this enzyme. As a precaution the PCR product(s) should then be sequenced before undertaking any further work, to determine whether the PCR reaction has been successful. If a mutation occurs at an early stage in the PCR reaction, then that error will be amplified in subsequent cycles of PCR. If this is the case then in subsequent ligation experiments most of the recombinant plasmids will contain the mutant gene, and this is probably what happened when *Taq* was used in PCR reactions.

Many factors can affect the fidelity of *Taq* and these have been detailed by Eckert and Kunkel.<sup>265</sup> A discussion of these may offer an explanation for the PCR errors encountered in this work. These factors may also affect the fidelity of other thermophilic polymerases, which have been less extensively studied.

- The concentration of dNTPs can have a detrimental effect on *Taq* fidelity if too high. A 1000  $\mu$ M dNTP concentration causes the error rate to be doubled in comparison to using 1  $\mu$ M dNTPs. In the work reported here the dNTP concentration was 125  $\mu$ M, which gives a result very similar to that obtained for 1000  $\mu$ M dNTPs. If a polymerase lacks an exonuclease to remove terminal mismatches, then unextended errors will be lost because they do not yield full-length DNA products for further amplification. At low dNTP concentrations the level of mispairing is reduced and the error discrimination is enhanced. The concentration of dNTPs should therefore be lowered in future PCR reactions.

- The fidelity of *Taq* can be improved by decreasing the magnesium chloride concentration relative to the concentration of dNTPs present in the reaction. The concentration of  $MgCl_2$  should be equimolar with the dNTP concentration for optimum fidelity. The reactions reported here were carried out using 2.5mM  $MgCl_2$ , as recommended by the manufacturer. In future reactions the concentration of  $MgCl_2$  relative to the dNTP concentration could be lowered to reduce the level of mutation.
- The error rate of *Taq* is reduced by decreasing the reaction pH to 5 - 6 and the frequency of mutations continues to increase as the pH is raised through 8.2. At pH 5.1, the error rate is reported to be 1/180 000 using 1000  $\mu M$  dNTPs and 10 mM  $MgCl_2$  whereas at pH 8.2 the error rate is increased to 1/3 200 (both reactions at 70 °C). However, the incubation of DNA at high temperatures and low pH causes DNA damage, increasing the potential for mutations, so this must also be taken into consideration if lowering the reaction pH of PCR reactions in order to improve polymerase fidelity.

If the *thrC* gene is successfully cloned into pET vectors then it may prove possible to purify TS in a more efficient way than previously possible using the Tir8 mutant. If this were the case, it would then be useful to attempt the first X-ray crystallographic studies on the enzyme, particularly in order to elucidate the active-site structure of the enzyme. TS from either source could also be used, as originally intended, for product inhibition assays. The substrate analogues, which have been synthesised by G. Allan (these are discussed in Section 1.10), could also be tested as inhibitors of the enzyme or to further elucidate the TS reaction mechanism, as appropriate.

Site-directed mutagenesis experiments could later be carried out to monitor the effects of altering residues, which are conserved in all microbial TS sequences, on the structure and catalytic activity of the enzyme from *E. coli*. Mutant *thrC* genes produced *via* PCR could again be cloned into pET constructs using the established protocol.

**CHAPTER THREE**  
**EXPERIMENTAL**



## 3.1 Synthesis

### 3.1.1 General Experimental Procedures

#### 3.1.1.1 Compound Characterisation

Elemental microanalyses were performed in the departmental microanalytical laboratory.

NMR spectra were recorded on Varian Gemini 300 ( $^1\text{H}$ , 300 MHz;  $^{13}\text{C}$ , 74.76 MHz;  $^{31}\text{P}$ , 121.42 MHz;  $^{19}\text{F}$ , 282.2 MHz), or Varian Gemini 200 ( $^1\text{H}$ , 200 MHz;  $^{13}\text{C}$ , 50.31 MHz) spectrometers.  $^1\text{H}$ - and  $^{13}\text{C}$ -NMR spectra are described in parts per million downfield shift from  $\text{SiMe}_4$  and are reported consecutively as position ( $\delta_{\text{H}}$  or  $\delta_{\text{C}}$ ), relative integral, multiplicity (s = singlet, d = doublet, t = triplet, q = quartet, dd = double of doublets, dt = doublet of triplets, dq = doublet of quartets, sept = septet, m = multiplet, and br = broad), coupling constant ( $J/\text{Hz}$ ) and assignment (numbering according to the IUPAC nomenclature for the compound).  $^1\text{H}$ -NMR spectra were referenced internally on  $^2\text{HOH}$  ( $\delta$  4.80 ppm),  $\text{C}^2\text{HCl}_3$  ( $\delta$  7.27 ppm) or  $(\text{C}^2\text{H}_3)_2\text{SO}$  ( $\delta$  2.50 ppm).  $^{13}\text{C}$ -NMR spectra were referenced on  $\text{C}^2\text{HCl}_3$  ( $\delta$  77.23 ppm),  $\text{C}_6^2\text{H}_6$  ( $\delta$  128.39 ppm) or  $(\text{C}^2\text{H}_3)_2\text{SO}$  ( $\delta$  39.51 ppm).  $^{19}\text{F}$ -NMR spectra were referenced internally on  $\text{CFCl}_3$  and described in parts per million downfield shift from this compound.  $^{31}\text{P}$ -NMR spectra were referenced on external  $\text{H}_3\text{PO}_4$ .

IR spectra were recorded on a Perkin-Elmer 1710 FT-IR spectrometer. The samples were prepared as Nujol mulls, thin films between sodium chloride discs or as dispersions in potassium bromide discs. The frequencies ( $\nu$ ) as absorption maxima are given in wavenumbers ( $\text{cm}^{-1}$ ) relative to a polystyrene standard at  $1603\text{ cm}^{-1}$ . Signal strengths are reported as s = strong, m = medium, w = weak and br = broad.

Mass spectra and accurate mass measurements were recorded on VG 70-250 SE, Kratos MS-50, VG Platform E/S. Major fragments were given as percentages of the base peak intensity (100%).

UV data was obtained on Pye-Unicam SP8-500 or SP8-100 spectrophotometers.

Flash chromatography was performed according to the method of Still *et al.*<sup>315</sup> using Sorbsil C60 (40-60  $\mu\text{m}$  mesh) silica gel. Analytical thin layer chromatography (tlc) was carried out on 0.25 mm precoated silica gel plates (Macherey-Nagel SIL g/UV254) or 0.25 mm precoated cellulose plates and compounds were visualised using UV fluorescence, potassium permanganate solution, vanillin, or ninhydrin. Preparative tlc was performed on glass-backed cellulose plates.

Melting points were taken on an Electrothermal or Gallenkamp melting point apparatus and are uncorrected.

Optical rotations were measured at 23 °C on an Optical Activity AA-1000 polarimeter using 10 cm path length cells. Readings are given in  $10^{-1}$  deg  $\text{cm}^2 \text{g}^{-1}$  and the solution concentration is given in  $\text{g cm}^{-3}$ .

### 3.1.1.2 Reagent and Solvent Preparation

The solvents used were either distilled or of Analar quality and light petrol ether refers to that portion boiling between 40 and 60 °C. Solvents and bases were dried according to literature procedures, distilling under nitrogen. Ethanol and methanol were dried using magnesium turnings and iodine and stored over freshly activated Linde 4Å molecular sieves. Isopropanol and acetonitrile were distilled over  $\text{CaH}_2$  and stored over activated Linde 4Å molecular sieves. Toluene,  $\text{CH}_2\text{Cl}_2$  and diisopropylamine were distilled over  $\text{CaH}_2$  directly before use. Diisopropylethylamine was distilled from KOH. Triethylamine and N-methylmorpholine were distilled from ninhydrin, then KOH. THF and  $\text{Et}_2\text{O}$  were distilled from the sodium ketal of benzophenone directly before use, under nitrogen. DME was also distilled under nitrogen from the sodium ketal of benzophenone onto 4Å molecular sieves. DMF and benzyl alcohol were distilled under reduced pressure onto activated Linde 4Å molecular sieves.

**Diazomethane preparation:** N-nitroso-N-methyl-4-toluene-sulphonamide (Diazald) (21 g,

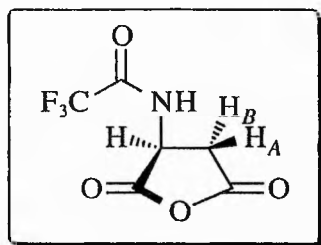
98 mmol) in ether (150 cm<sup>3</sup>) was added slowly to a solution of potassium hydroxide (4.8 g, 86 mmol) in water (8 cm<sup>3</sup>) and isopropanol (12 cm<sup>3</sup>). The solution was heated in a clear jointed distillation apparatus and the ethereal diazomethane distilled into ether (10 cm<sup>3</sup>) at 0 °C. The yellow solution obtained was used in the synthesis of methyl ester derivatives.

### 3.1.1.3 Suppliers

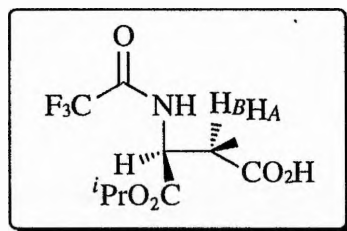
2-(*p*-Hydroxyphenyl)-ethanol and 2-(4-methoxyphenyl)-ethanol were purchased from Aldrich. *p*-Benzyloxybenzyl alcohol (Wang), hydroxymethylpolystyrene and chloromethylpolystyrene-divinylbenzene (Merrifield) resins were purchased from Novabiochem.

### 3.1.2 Synthetic Procedures and Characterisations

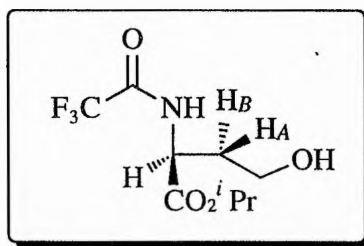
#### *N*-Trifluoroacetyl-(2*S*)-aspartic anhydride **73**.



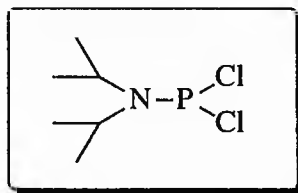
To a stirred suspension of (2*S*)-aspartate (500 mg, 3.76 mmol) in dry THF at  $-10 \rightarrow 0$  °C was added dropwise trifluoroacetic anhydride (2.65 cm<sup>3</sup>, 18.8 mmol). The reaction was stirred under nitrogen at room temperature. The progress of the reaction was monitored by tlc [ $R_f$  0.38: ethyl acetate-light petroleum (3:2)]. Upon completion (48 h), the solvent was removed under reduced pressure to give **73** as a white solid (785 mg, 99%); mp 110-115 °C;  $\nu_{\max}$ (Nujol)/cm<sup>-1</sup> 3350 m (N-H stretching, 2° amide), 1886 m (C=O, anhydride), 1722 s (C=O, anhydride), 1556 m (C=O, 2° amide), 1371 s (C-F), 1342 s (C-F), 1225 s (C-F, C-O), 1176 s (C-F, C-O) and 1083 m (C-F, C-O);  $\delta_{\text{H}}$ (200 MHz; C<sup>2</sup>HCl<sub>3</sub>) 3.01 (2 H, dq, ABX,  $J_{\text{AB}}$  18.5,  $J_{\text{AX}}$  10.1,  $J_{\text{BX}}$  6.8, H <sup>$\beta$</sup> ), 4.85 (1 H, dt,  $J_{\text{NH,H}}$  7.0,  $J_{\text{AX}}$  10.1,  $J_{\text{BX}}$  6.8, H <sup>$\alpha$</sup> ) and 9.57 (1 H, d,  $J_{\text{NH,H}}$  7.0, amide NH);  $\delta_{\text{C}}$ (74.76 MHz; C<sup>2</sup>HCl<sub>3</sub>) 34.09 (C <sup>$\beta$</sup> ), 48.67 (C <sup>$\alpha$</sup> ), 115.34 (q,  $J_{\text{F,C}}$  285.28, F<sub>3</sub>C), 157.43 (q,  $J_{\text{F}_3\text{C,C}}$  38.58, F<sub>3</sub>CCO), 167.61 ( $\beta$ C=O) and 169.28 ( $\alpha$ C=O);  $\delta_{\text{F}}$ (282.2 MHz; C<sup>2</sup>HCl<sub>3</sub>) -76.37 (F<sub>3</sub>C);  $m/z$  (CI) 230 (100%, [M + H + H<sub>2</sub>O]<sup>+</sup>) and 212 (14, [M + H]<sup>+</sup>).

**$\alpha$ -Isopropyl-*N*-trifluoroacetyl-(2*S*)-aspartate 74.**

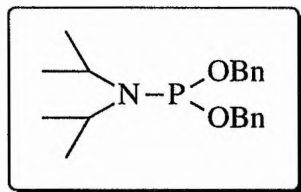
To a stirred suspension of (2*S*)-aspartic acid (1.15 g, 8.61 mmol) in dry THF (30 cm<sup>3</sup>) at -5 °C was added dropwise (8 cm<sup>3</sup>, 56.7 mmol) of trifluoroacetic anhydride. After 4 h the solvent was removed under reduced pressure, and the anhydride **73** was dissolved in dry isopropanol at room temperature. The progress of the alcoholysis was monitored by tlc [*R<sub>f</sub>* 0.46: ethyl acetate-light petroleum (3:2)] and after 1 day the isopropanol was removed under reduced pressure, giving a clear yellow oil, which on standing formed a white solid. Recrystallisation from ether gave a white crystalline solid (2.31g, 99%), mp 92 °C (lit.,<sup>35</sup> 91 °C); [ $\alpha$ ]<sub>D</sub> -49.95 (*c* 1.0 in MeOH), [lit.,<sup>35</sup> -40.7 (*c* 1.0 in MeOH)];  $\nu_{\max}$ (Nujol)/cm<sup>-1</sup> 3310 s (N-H, 2° amide), 3115 s (OH, acid), 1738 s (C=O, ester), 1709 s (C=O, carboxylic), 1566 s (N-H, amide II), 1467 s (CH<sub>2</sub>, CH<sub>3</sub>, C-H deformations), 1390 and 1378 s (C-F), 1313 s (C-F), 1280, 1236, 1208 and 1186 s (C-O and C-F stretching) and 1106 s (C-F);  $\delta_{\text{H}}$ (200 MHz; C<sup>2</sup>HCl<sub>3</sub>) 1.25 (6 H, d,  $J_{\text{CH}_3,\text{H}}$  6.25, 2 x <sup>*i*</sup>PrCH<sub>3</sub>), 3.08 (2 H, dq, ABX,  $J_{\text{AB}}$  17.83,  $J_{\text{AX}}$  4.42,  $J_{\text{BX}}$  4.25, H <sup>$\beta$</sup> ), 4.21 (1 H, br s,  $\beta$ CO<sub>2</sub>H), 4.78 (1 H, dt,  $J_{\text{NH,H}}$  6.60, ABX,  $J_{\text{AX}}$  4.42,  $J_{\text{BX}}$  4.25, H <sup>$\alpha$</sup> ), 5.11 (1 H, sep,  $J_{\text{CH}_3,\text{H}}$  6.25, <sup>*i*</sup>PrCH) and 7.38 (1 H, d,  $J_{\text{NH,H}}$  6.60, amide NH);  $\delta_{\text{C}}$ (74.76 MHz; C<sup>2</sup>HCl<sub>3</sub>) 21.30 (<sup>*i*</sup>PrCH<sub>3</sub>), 35.06 (C <sup>$\beta$</sup> ), 48.80 (C <sup>$\alpha$</sup> ), 70.96 (<sup>*i*</sup>PrCH), 115.65 (q,  $J_{\text{F,C}}$  285.28, F<sub>3</sub>C), 157.19 (q,  $J_{\text{F}_3\text{C,C}}$  37.53, F<sub>3</sub>CCO), 168.56 ( $\alpha$ CO<sub>2</sub><sup>*i*</sup>Pr) and 175.54 ( $\beta$ CO<sub>2</sub>H);  $\delta_{\text{F}}$ (282.2 MHz; C<sup>2</sup>HCl<sub>3</sub>) -76.77 (F<sub>3</sub>C); *m/z* (EI) 226 (4%, [M - CO<sub>2</sub>H]<sup>+</sup>), 212 (7, [M - CH<sub>2</sub>CO<sub>2</sub>H]<sup>+</sup>), 184 (36, [M - CO<sub>2</sub><sup>*i*</sup>Pr]<sup>+</sup>), 139 (35, [M - CO<sub>2</sub><sup>*i*</sup>Pr and CO<sub>2</sub>H]<sup>+</sup>), 96 (22, F<sub>3</sub>CCO), 88 (8, [<sup>*i*</sup>PrO<sub>2</sub>CH]<sup>+</sup>), 69 (36, F<sub>3</sub>C<sup>+</sup>), 59 (8, CH<sub>2</sub>CO<sub>2</sub>H<sup>+</sup>), 43 {100, (CH<sub>3</sub>)<sub>2</sub>CH<sup>+</sup>} and 28 (35, CO<sup>+</sup>).

**$\alpha$ -Isopropyl-*N*-trifluoroacetyl-(2*S*)-homoserine 75.**

To a stirred solution of the isopropyl ester **74** (2.02g, 7.46 mmol) in dry THF (30 cm<sup>3</sup>) at -20 °C was added *N*-methyl morpholine (0.82 cm<sup>3</sup>, 7.46 mmol) and isobutylchloroformate (1.02 cm<sup>3</sup>, 7.46 mmol). After 20 min the solution was filtered into a suspension of sodium borohydride (0.169 g, 4.48 mmol) in dry THF at -20 °C. The solution was stirred for 2 h before being allowed to warm to room temperature and then stirred for a further 2 h. Excess borohydride was destroyed by the addition of water (5 cm<sup>3</sup>), and stirring continued for a further 30 min. The THF was removed under reduced pressure and the product extracted from the aqueous layer using ether (10 x 10 cm<sup>3</sup>). The combined organic extracts were washed with saturated NaHCO<sub>3</sub> solution (12 cm<sup>3</sup>) and saturated brine (12 cm<sup>3</sup>) and dried (MgSO<sub>4</sub>). The ether was removed under reduced pressure giving the crude alcohol as a clear orange oil. Purification of the crude oil by flash silica column chromatography [eluant DCM-EtOH (97:3) containing a few drops of acetic acid] gave the pure protected homoserine **75** as a clear yellow oil (0.672 g, 35%);  $\nu_{\max}$ (Nujol)/cm<sup>-1</sup> 3318 br m (OH and NH superimposed), 2923 m (OH, acid/ alcohol), 1723 s (C=O ester), 1560 w (NH), 1458 w (CH<sub>2</sub>, CH<sub>3</sub>, C-H deformations), 1377 m (CH<sub>3</sub>), 1184 bs (C-O, C-F) and 1106 m (C-O, C-F);  $\delta_{\text{H}}$ (200 MHz; C<sup>2</sup>HCl<sub>3</sub>) 1.29 (6 H, d,  $J_{\text{CH}_3,\text{H}}$  6.25, 2 x *i*PrCH<sub>3</sub>), 2.10 (2 H, m, H <sup>$\beta$</sup> ), 3.73 (2 H, m, H <sup>$\gamma$</sup> ), 4.69 (1 H, dt,  $J_{\text{NH,H}}$  4.35, ABX,  $J_{\text{AX}}$  12.00,  $J_{\text{BX}}$  3.20, H <sup>$\alpha$</sup> ), 5.09 (1 H, sept,  $J_{\text{CH}_3,\text{H}}$  6.25, *i*PrCH) and 7.82 (1 H, d,  $J_{\text{NH,H}}$  4.35, NH);  $\delta_{\text{C}}$ (50.31 MHz; C<sup>2</sup>HCl<sub>3</sub>), 21.46 (*i*PrCH<sub>3</sub>), 32.90 (C <sup>$\beta$</sup> ), 51.32 (C <sup>$\alpha$</sup> ), 58.58 (C <sup>$\gamma$</sup> ), 70.17 (*i*PrCH), 115.74 (q,  $J_{\text{F,C}}$  287.37, F<sub>3</sub>C), 157.09 (q,  $J_{\text{F}_3\text{C,C}}$  37.78, F<sub>3</sub>CCO) and 170.47 ( $\alpha$ CO<sub>2</sub>*i*Pr);  $\delta_{\text{F}}$ (282.2 MHz; C<sup>2</sup>HCl<sub>3</sub>) -76.74 (F<sub>3</sub>C);  $m/z$  (EI) 258 (100%, [M + H]<sup>+</sup>), (17, [M - OH]<sup>+</sup>), 216 (33, [M - *i*Pr + 2H]<sup>+</sup>), 198 (85, [M - *i*PrO]<sup>+</sup>), 170 (80, [M - CO<sub>2</sub>*i*Pr]<sup>+</sup>), 140 (86, [M + H - CH<sub>2</sub>OH and CO<sub>2</sub>*i*Pr]<sup>+</sup>), 69 (28, F<sub>3</sub>C<sup>+</sup>) and 57 (20, F<sub>3</sub><sup>+</sup>).

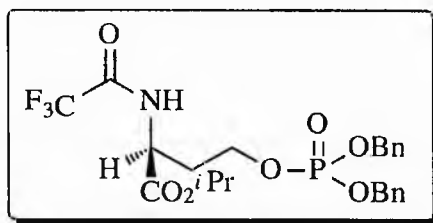
***N,N*-Diisopropyldichlorophosphamidite 78.**

To a vigorously stirred solution of  $\text{PCl}_3$  (11.62 g, 7.38  $\text{cm}^3$ , 85 mmol) in dry ether (75  $\text{cm}^3$ ) at  $-20^\circ\text{C}$  under nitrogen was added dropwise *via* cannulae dry diisopropylamine (17.11 g, 23.7  $\text{cm}^3$ , 169 mmol) dissolved in 75  $\text{cm}^3$  dry ether. After 2 h the reaction was allowed to warm to room temperature and stirred for a further 1.5 h. The precipitated salts were removed by filtration under a nitrogen atmosphere, washing with dry ether (250  $\text{cm}^3$ ). The solvent was then removed under reduced pressure to give a pale yellow liquid which was fractionally distilled. The product, **78**, which was obtained as a colourless liquid at room temperature, solidified on cooling (13.78 g, 81%), bp  $74\text{--}76^\circ\text{C}$ ; 7 mm Hg (lit.,<sup>35</sup>  $72\text{--}74^\circ\text{C}$ ; 7 mm Hg) (HRMS: found:  $M^+$ , 201.0235 Calc. for  $\text{C}_6\text{H}_{14}\text{NPCl}_2$ : 201.0241);  $\nu_{\text{max}}$ (thin film)/ $\text{cm}^{-1}$  2374 s (C-H), 1664 w, 1462 m ( $\text{CH}_3$ , C-H deformations), 1398 m and 1370 m ( $-\text{CH}(\text{CH}_3)_2$ ), 1202 m, 1171 m, 1155 m, 1123 s, 1027 m, 976 s and 882 w;  $\delta_{\text{H}}$ (200 MHz;  $\text{C}^2\text{HCl}_3$ ) 1.27 (6 H, d,  $J_{\text{CH}_3,\text{H}}$  6.83,  $^i\text{PrCH}_3$ ) and 3.92 (2 H, sept,  $J_{\text{CH}_3,\text{H}}$  6.83,  $^i\text{PrCH}$ );  $\delta_{\text{C}}$ (50.31 MHz;  $\text{C}^2\text{HCl}_3$ ) 23.61 (d,  $J_{\text{C,P}}$  8.54, 2 x  $^i\text{PrCH}_3$ ) and 48.35 (d,  $J_{\text{C,P}}$  13.94, 2 x  $^i\text{PrCH}$ );  $\delta_{\text{P}}$ (121.42 MHz;  $\text{C}^2\text{HCl}_3$ ) 169.90 [lit.,<sup>35</sup>  $\delta_{\text{P}}$  169.00];  $m/z$  (EI) 203 {5%,  $M^+$  ( $^{35}\text{Cl}$ ,  $^{37}\text{Cl}$ )}, 201 {8,  $M^+$  ( $^{35}\text{Cl}$ )}, 186 {100,  $[\text{M} - \text{CH}_3]^+$  ( $^{35}\text{Cl}$ )}, 168 {11,  $[\text{M} - \text{Cl}]^+$  ( $^{37}\text{Cl}$ )}, 166 {37,  $[\text{M} - \text{Cl}]^+$  ( $^{35}\text{Cl}$ )}, 148 (7,  $[\text{PrP}^{37}\text{Cl}_2]^+$ ), 146 (24,  $[\text{PrP}^{35}\text{Cl}^{37}\text{Cl}]^+$ ), 144 (42,  $[\text{PrP}^{35}\text{Cl}_2]^+$ ), 88 (28,  $[\text{PN}^i\text{Pr}]^+$ ), 43 (87,  $(\text{CH}_3)_2\text{CH}^+$ ) and 32 (100,  $\text{C}_2\text{H}_8^+$ ).

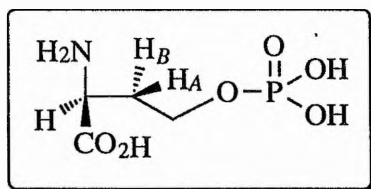
***N,N*-Diisopropylbis(benzyl)phosphoramidate 79.**

To a stirred solution of the dichlorophosphamidite **78** (1.52 g, 7.51 mmol) in dry DCM (5 cm<sup>3</sup>) at -10 °C under a nitrogen atmosphere was added *via* cannulae a solution of dry benzyl alcohol (2.13 g, 19.6 mmol) and dry Et<sub>3</sub>N (2.8 cm<sup>3</sup>, 20.1 mmol) in dry DCM (10 cm<sup>3</sup>). After 30 min the reaction was allowed to warm to room temperature and stirred for a further 4 h. The solution was then diluted with DCM (15 cm<sup>3</sup>), washed with NaHCO<sub>3</sub> solution (10 cm<sup>3</sup> of 5g *w/w*) and saturated brine solution (10 cm<sup>3</sup>), and dried (MgSO<sub>4</sub>). The solvent was then removed under reduced pressure to yield the crude product as a pale yellow clear liquid. Purification of the crude oil by silica column chromatography on triethylamine-basified silica [eluant petroleum-ether/ethyl acetate 4:1] gave the phosphoramidate **79** as a colourless oil (2.19 g, 84%), (HRMS: found: [M + H]<sup>+</sup>, 346.1600, Calc. for C<sub>20</sub>H<sub>29</sub>NO<sub>2</sub>P: 346.1935);  $\nu_{\max}$ (thin film)/ cm<sup>-1</sup> 3070 m, 3060 s, 3010 s (triplet, C-H), 2900 br s (CH<sub>2</sub>, CH<sub>3</sub>, C-H), 1930 w, 1860 w, 1780 w, 1600 m and 1580 m (doublet, Ar ring), 1490 s and 1450 s (CH<sub>2</sub>, CH<sub>3</sub>, C-H deformations), 1370 s (CH<sub>3</sub> symmetrical deformation), 1010 br s (P-O), 910 s and 870 m (Ar C-H), and 770 br s (CH-Ar out of plane);  $\delta_{\text{H}}$ (200 MHz; C<sup>2</sup>HCl<sub>3</sub>) 1.23 (6 H, d,  $J_{\text{CH}_3,\text{H}}$  6.84, 2 x <sup>*i*</sup>PrCH<sub>3</sub>), 3.72 (2 H, m, 2 x <sup>*i*</sup>PrCH), 4.75 (4 H, q,  $J_{\text{C,P}}$  8.05, PhCH<sub>2</sub>) and 7.35 (10 H, m, Ar-H);  $\delta_{\text{C}}$ (50.31 MHz; C<sup>2</sup>HCl<sub>3</sub>) 24.56 (d,  $J_{\text{CH}_3,\text{H}}$  7.22, 2 x <sup>*i*</sup>PrCH<sub>3</sub>), 42.95 (d,  $J_{\text{CH}_3,\text{H}}$  7.22, <sup>*i*</sup>PrCH), 64.30 (d,  $J_{\text{CH}_2,\text{OP}}$  10.88, 2 x OCH<sub>2</sub>Ph), 128.31, 127.41 and 126.87 (Ar-C), 138.06 (d,  $J_{\text{P,C}}$  4.89, Ar-C quaternary);  $\delta_{\text{P}}$ (121.42 MHz; C<sup>2</sup>HCl<sub>3</sub>) 140.01 [lit.,<sup>35</sup>  $\delta_{\text{P}}$ (36.2 MHz; C<sup>2</sup>HCl<sub>3</sub>) 148.10];  $m/z$  (EI) 346 (4%, [M + H]<sup>+</sup>), 261 (3, [NH<sub>2</sub>P(OBn)<sub>2</sub>]<sup>+</sup>), 91 (34, [<sup>*i*</sup>PrNHPH<sub>2</sub>]<sup>+</sup>), 86 (39, <sup>*i*</sup>PrNCH<sub>2</sub>]<sup>+</sup>), 44 {100, (CH<sub>3</sub>)<sub>2</sub>CH<sub>2</sub><sup>+</sup>}, 43 (27, <sup>*i*</sup>Pr<sup>+</sup>), 42 {17, (CH<sub>3</sub>)C<sup>+</sup>} and 28 (16, C<sub>2</sub>H<sub>4</sub><sup>+</sup>).



**$\alpha$ -Isopropyl-*N*-trifluoroacetyl-(2*S*)-phosphohomoserinebis(benzyl) ester 76.**

To a stirred mixture of the alcohol **75** (0.192 g, 0.75 mmol) and 1-*H* tetrazole (157 mg, 2.24 mmol) in dry DCM (12 cm<sup>3</sup>) was added *N,N*-diisopropylbis(benzyl)phosphoramidite **79** (388 mg, 1.13 mmol). The mixture was stirred at room temperature for 4.5 h then cooled to -40 °C. *m*-CPBA (258 mg, 1.50 mmol) in dry DCM (10 cm<sup>3</sup>) was added dropwise to the reaction mixture and the resulting solution was stirred at 0 °C for 45 min before being diluted further with DCM (30 cm<sup>3</sup>), washed with Na<sub>2</sub>SO<sub>3</sub> (7 x 10 cm<sup>3</sup> of 10% w/w), NaHCO<sub>3</sub> (6 x 10 cm<sup>3</sup> of 5% w/w), water (30 cm<sup>3</sup>) and saturated brine solution (15 cm<sup>3</sup>), and then dried (Na<sub>2</sub>SO<sub>4</sub>). The solvent was removed under reduced pressure, giving the product as a clear yellow oil (0.244 g, 63%);  $\nu_{\max}$ (thin film)/cm<sup>-1</sup> 1786 w, 1725 m (C=O, ester), 1560 w (2° amide), 1457 m (-CH<sub>2</sub>-, -CH<sub>3</sub>, CH deformations), 1377 w (-CH<sub>3</sub>, symmetrical deformations), 1262 m (C-F, C-O or P=O), 1215 m (C-O, C-F), 1182 m (C-F, C-O), 1106 w (C-O, C-F) and 1018 m (C-F);  $\delta_{\text{H}}$ (200 MHz; C<sup>2</sup>HCl<sub>3</sub>) 1.26 (6 H, m, 2 x <sup>*i*</sup>PrCH<sub>3</sub>), 2.08 (2 H, m, H <sup>$\beta$</sup> ), 3.75 (2 H, m, H <sup>$\gamma$</sup> ), 4.71 (1H, m, H <sup>$\alpha$</sup> ), 5.07 (5 H, m, 2 x PhCH<sub>2</sub> and <sup>*i*</sup>PrCH superimposed), 7.39 (10 H, m, Ar-H) and 7.62 (1 H, m, NH);  $\delta_{\text{C}}$ (50.31 MHz; C<sup>2</sup>HCl<sub>3</sub>) 18.99 (<sup>*i*</sup>PrCH<sub>3</sub>), 21.60 ( <sup>$\beta$</sup> C), 47.28 ( <sup>$\alpha$</sup> C), 68.39 (d,  $J_{\text{C,OP}}$  5.49,  <sup>$\gamma$</sup> C), 69.46 (CH<sub>2</sub>Ph), 70.31 (<sup>*i*</sup>PrCH), 127.93 (10 x Ar-C), 135.33 (q,  $J_{\text{F,C}}$  217.69, F<sub>3</sub>C) and 169.67 ( <sup>$\beta$</sup> CO<sub>2</sub><sup>*i*</sup>Pr);  $\delta_{\text{P}}$ (121.42 MHz; C<sup>2</sup>HCl<sub>3</sub>) -1.07 [lit.,<sup>35</sup>  $\delta_{\text{P}}$  0.55].

**(2S)-O -Phosphohomoserine 32.**

To a stirred solution of phosphate triester **76** (0.201 g (0.39 mmol) in methanol (10 cm<sup>3</sup>) was added 5% palladium charcoal (45 mg). The mixture was purged with hydrogen for 30 min and left to stir under a hydrogen atmosphere. After 4 days (when the reaction was judged to be complete by tlc [eluant MeOH-DCM 19:1]), the mixture was filtered through a celite pad and concentrated under reduced pressure. The residue was dissolved in 1M KOH (4 cm<sup>3</sup>) in ethanol (5 cm<sup>3</sup>) which resulted in the solution turning yellow. The mixture was left stirring for 18 h and then concentrated under reduced pressure. The resulting pale brown solid was subjected to ion-exchange chromatography (Dowex 50-W H<sup>+</sup>, 200-400 mesh, 8% cross-linked, 1cm x 6 cm) eluting with water. The ninhydrin positive fractions were combined and lyophilized to yield **32** as a white crystalline solid (700 mg, 91%), mp 167-169 °C (lit.,<sup>35</sup>, 170 °C);  $\nu_{\max}$ (Nujol)/cm<sup>-1</sup> 1733 w (C=O, saturated acid), 1462 m (C-H deformations), 1249 w (P=O), 1020 w (C-O);  $\delta_{\text{H}}$ (200 MHz; <sup>2</sup>H<sub>2</sub>O) 2.09 (2 H, m, H<sup>β</sup>), 3.93 (2 H, dq,  $J_{\text{CX}}$  11.6,  $J_{\text{DX}}$  5.4, H<sup>γ</sup>) and 3.97 (1 H, dq,  $J_{\text{AX}}$  7.6,  $J_{\text{BX}}$  4.8, H<sup>α</sup>);  $\delta_{\text{C}}$  (50.31 MHz; <sup>2</sup>H<sub>2</sub>O) 33.40 and 33.25 (d,  $J_{\text{C,OP}}$  7.63, βC), 54.19 and 54.03 (d,  $J_{\text{C,OP}}$  8.14, γC), 64.67 (αC) and 175.09 (αCO<sub>2</sub>H);  $\delta_{\text{P}}$ (121.42 MHz; <sup>2</sup>H<sub>2</sub>O) 0.59 [lit.,<sup>35</sup>  $\delta_{\text{P}}$  0.245];  $m/z$  (EI) 200 (5%, [M + H]<sup>+</sup>), 199 (5, M<sup>+</sup>), 181 (9, [M - H<sub>2</sub>O]<sup>+</sup>), 119 (17, [M - H<sub>3</sub>PO<sub>3</sub>]<sup>+</sup>), 104 (23, [M - PO<sub>4</sub>]<sup>+</sup>), 97 (15, [H<sub>2</sub>PO<sub>4</sub>]<sup>+</sup>), 91 (73, [M - COPO<sub>3</sub>H]<sup>+</sup>), 57 (23, [NH<sub>2</sub>CHCH<sub>2</sub>CH<sub>2</sub>]<sup>+</sup>), 55 (37, [C<sub>3</sub>H<sub>5</sub>N]<sup>+</sup>) and 44 (100, CO<sub>2</sub><sup>+</sup>).

**[U-<sup>14</sup>C]- $\alpha$ -Isopropyl-*N*-trifluoroacetyl-(2*S*)-aspartate 74b**

This compound was prepared in a manner identical to the unlabelled ester **74** starting from [U-<sup>14</sup>C]-(*2S*)-aspartic acid (700 mg, 5.26 mmol). The reaction progress was monitored by tlc [ $R_f$  0.46; ethyl acetate-light petroleum (3:2)]. Recrystallisation from ether-light petroleum gave the product as a white solid (1.0644g, 75%). All spectroscopic and analytical data was identical to that for **74**.

**[U-<sup>14</sup>C]- $\alpha$ -Isopropyl-*N*-trifluoroacetyl-(2*S*)-homoserine 75b**

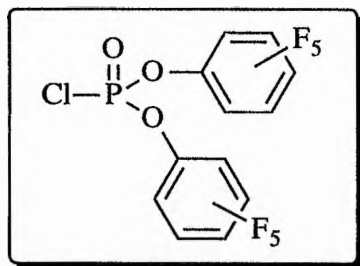
This compound was prepared in a manner identical to the alcohol **75** starting from **74b** (1.0644 g, 3.93 mmol). Purification by silica column chromatography [eluant DCM-EtOH (97%)] gave the product as a yellow oil (0.518g, 51%). All spectroscopic and analytical data was identical to that for **75**.

**[U-<sup>14</sup>C]- $\alpha$ -Isopropyl-*N*-trifluoroacetyl-(2*S*)-phosphohomoserinebis(benzyl) ester 76b**

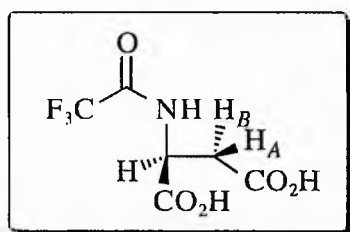
This was prepared in a similar manner to the phosphate triester **76**, starting from **75b** (518 mg, 2.02 mmol). The reaction was monitored by tlc [ethyl acetate-petroleum ether (1:4)]. The crude product was obtained as a yellow oil (1.085g, >100% due to benzyl alcohol impurities). All spectroscopic and analytical data was identical to that for **76**.

**[U-<sup>14</sup>C]-(*2S*)-*O* -phosphohomoserine 32b**

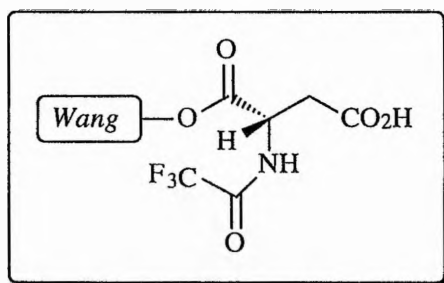
This compound was prepared in a similar manner to **32**, starting from **76b** (1.085g, 2.1 mmol). Any organic impurities were then removed by extraction with ether (3 x 10 cm<sup>3</sup>), and the product recrystallised from water-ethanol (346 mg, 83%). All spectroscopic and analytical data was identical to that for **32**.

**Dipentafluorophenylphosphoro chloridate 81.**

Using a modification of the procedure of Baer<sup>287</sup> freshly distilled  $\text{POCl}_3$  (7.64 g, 50 mmol) and pentafluorophenol (17.41 g, 94.6 mmol) were refluxed at 140 °C under a nitrogen atmosphere for 48 h. The contaminants were distilled under reduced pressure to leave the product as a pale brown oil which solidified upon cooling (28.99 g, >100%), bp 181-183 °C, 20 mm Hg;  $\nu_{\text{max}}$ (thin film)/ $\text{cm}^{-1}$  1346 m (C-F), 1318 m (C-F), 1239 m (C-O, C-F, P-O-aryl), 1156 s (C-F, C-O), 1034 (C-F, C-O) and 1001 (C-F, C-O);  $\delta_{\text{C}}$ (74.76 MHz;  $\text{C}^2\text{HCl}_3$ ) 136.75 (*meta* Ar-CF), 139.70 (*para* Ar-CF), 140.13 (*ortho* Ar-CF), 143.01 (Ar-CO);  $\delta_{\text{P}}$ (121.42 MHz;  $\text{C}^2\text{HCl}_3$ ) main signal -15.14;  $\delta_{\text{F}}$ (282.2 MHz;  $\text{C}^2\text{HCl}_3$ ) -154.89 and -154.96 (4 F, d  $J$  19.75, *ortho* Ar-CF), -157.96 (2 F, dt,  $J$  3.95,  $J$  21.73, *para* Ar-CF), -162.38 (4 F, t, *meta* Ar-CF);  $m/z$  (EI) 449 (40%,  $\text{M}^+$ ), 413 (30,  $[\text{M} - \text{Cl}]^+$ ), 265 (14,  $[\text{M} - \text{OC}_6\text{F}_5]^+$ ) and 249 (17  $[\text{M} - \text{O}_2\text{C}_6\text{F}_5]^+$ ).

**N-Trifluoroacetyl-(2S)-aspartate 86.**

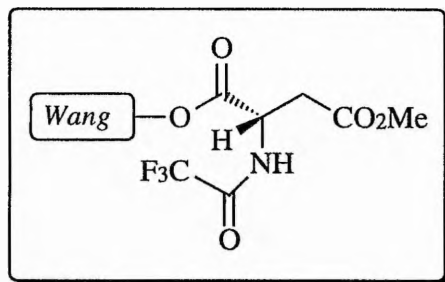
A solution of the anhydride **73** (0.555 g, 2.629 mmol) in distilled water (20 cm<sup>3</sup>) was stirred overnight at rt until a yellow oil crashed out of solution. The mixture was extracted with ether (25 cm<sup>3</sup>), washed with water (2 x 10 cm<sup>3</sup>), 5% citric acid (10 cm<sup>3</sup>) and saturated brine solution (10 cm<sup>3</sup>) and dried (MgSO<sub>4</sub>). The ether was removed under reduced pressure to give a white solid, which was recrystallised from ether-hexane to give the diacid **86** (0.147 g, 25%); mp 132-134 °C; [ $\alpha$ ]<sub>D</sub> -24.35 (*c* 0.87 in EtOH);  $\nu_{\max}$ (thin film)/cm<sup>-1</sup> 2974 s (acid OH), 2877 s (acid OH), 2604 w (acid OH), 1717 (C=O, saturated acid), 1566 m (NH, 2° amide), 1371 m (C-F), 1303 m (C-F, C-O) and 1195 m (C-O, C-F);  $\delta_{\text{H}}$ (200 MHz; C<sup>2</sup>HCl<sub>3</sub>) 2.91 (2 H, dq, ABX,  $J_{\text{AB}}$  17.47,  $J_{\text{AX}}$  4.76,  $J_{\text{BX}}$  4.16, H <sup>$\beta$</sup> ), 4.71 (1 H, dt,  $J_{\text{NH,H}}$  7.65,  $J_{\text{AX}}$  4.76,  $J_{\text{BX}}$  4.16, H <sup>$\alpha$</sup> ), 7.84 (1 H, d,  $J_{\text{NH,H}}$  7.65, amide NH) and 8.37 (2 H, br s, CO<sub>2</sub>H);  $\delta_{\text{C}}$ (74.76 MHz; C<sup>2</sup>HCl<sub>3</sub>) 34.90 ( $\beta$ C), 48.63 ( $\alpha$ C), 115.43 (q,  $J_{\text{F,C}}$  285.26, F<sub>3</sub>C), 156.70 (q,  $J_{\text{F}_3\text{C,C}}$  37.18, F<sub>3</sub>CCO), 171.06 ( $\beta$ CO<sub>2</sub>H) and 172.54 ( $\alpha$ CO<sub>2</sub>H);  $\delta_{\text{F}}$ (282.2 MHz; C<sup>2</sup>HCl<sub>3</sub>) -76.62 (F<sub>3</sub>C);  $m/z$  (CI) 230 (28%, [M - H<sub>2</sub>O]<sup>+</sup>), 184 (100, [M - HCO<sub>2</sub>H]<sup>+</sup>), 114 (22, [F<sub>3</sub>CC(O)NHCH<sub>2</sub>]<sup>+</sup>), 69 (35, F<sub>3</sub>C<sup>+</sup>) and 45 (32, CO<sub>2</sub>H<sup>+</sup>).

**$\alpha$ -*p*-Benzyloxybenzylpolystyrene-*N*-trifluoroacetyl-(2*S*)-aspartate **87**.**

To a stirred suspension of Wang resin (200 mg, loading: 1.16 mmol/g) in dry DMF (5 cm<sup>3</sup>) under a nitrogen atmosphere was added the anhydride **73** (0.4897 g, 2.32 mmol) and DMAP (283 mg, 0.232 mmol). The reaction was warmed to 30-35 °C. After 1 week the resin was filtered, washed copiously with DCM, DMF, 5% citric acid-DMF and methanol, then dried under reduced pressure (248.9 mg, 99.9% loading, 0.931 mmol/g of *N*-TFA-asp);  $\nu_{\max}(\text{KBr})/\text{cm}^{-1}$  1735 s (C=O, ester linkage);  $\delta_{\text{C}}(74.76 \text{ MHz}; (\text{C}^2\text{H}_5)_2\text{SO})$  34.81 ( $\beta\text{C}$ ), 49.40 ( $\alpha\text{C}$ ), 117.84 (F<sub>3</sub>C), 130.02 (F<sub>3</sub>CCO), 162.48 ( $\beta\text{CO}_2\text{H}$ ) and 171.33 ( $\alpha\text{CO}_2\text{R}$ );  $\delta_{\text{F}}(282.2 \text{ MHz}; (\text{C}^2\text{H}_5)_2\text{SO})$  -74.27 (12%, F<sub>3</sub>C on  $\beta$ -ester) and -74.91 (88%, F<sub>3</sub>C on  $\alpha$ -ester).

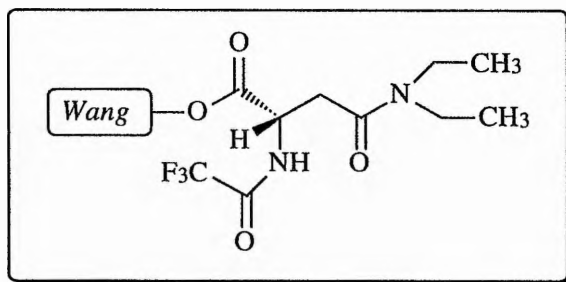
A portion of **87** was treated with TFA/H<sub>2</sub>O/TES (90:5:5) for 2 h at room temperature to produce the free acid **86**. All spectroscopic and analytical data was identical to that for **86**,  $\{[\alpha]_{\text{D}} -25.09 (c 0.87 \text{ in EtOH})\}$ .

**$\alpha$ -*p*-Benzyloxybenzylpolystyrene-*N*-trifluoroacetyl-(2*S*)-aspartate- $\beta$ -methyl ester 101.**



To **87** (500 mg, loading: 0.913 mmol/g of *N*-TFA-asp) was added an excess of ethereal diazomethane (10 cm<sup>3</sup>). The mixture was left at room temperature until all bubbling had ceased, and after the addition of another aliquot of diazomethane (5 cm<sup>3</sup>) the mixture was left for a further 30 min. Excess diazomethane was removed by bubbling nitrogen through the solution and the resin was filtered and washed copiously with ether, DMF, 5% citric acid-DMF and methanol. The resin was dried under reduced pressure (524 mg, 99%). The ester was cleaved from the resin using TFA/H<sub>2</sub>O/TES (90:5:5) at room temperature for 2 h, filtered off and dried under reduced pressure, giving the monoester **91** as a yellow oil;  $\nu_{\max}$ (thin film)/cm<sup>-1</sup> 3111 s and 2974 s (acid O-H), 2604 m (acid O-H), 1727 s (C=O, saturated acid), 1566 s (C=O, 2° amide), 1454 s (saturated C-H), 1298 s (C-O, C-F), 1240 (C-O, C-F), 1186 s (C-O ester, C-F);  $\delta_{\text{H}}$ (200 MHz; C<sup>2</sup>HCl<sub>3</sub>) 2.93 (2 H, dq, ABX,  $J_{\text{AB}}$  17.3,  $J_{\text{AX}}$  5.02,  $J_{\text{BX}}$  4.74, H <sup>$\beta$</sup> ), 3.63 (3 H, s,  $\beta$ CO<sub>2</sub>CH<sub>3</sub>, 84%), 3.70 (3 H, s,  $\alpha$ CO<sub>2</sub>CH<sub>3</sub>, 16%), 4.68 (1 H, dt,  $J_{\text{NH,H}}$  7.2,  $J_{\text{AX}}$  5.02,  $J_{\text{BX}}$  4.74, H <sup>$\alpha$</sup> ), 6.56 (1 H, br s,  $\alpha$ CO<sub>2</sub>H) and 7.86 and 7.82 (1 H, d,  $J_{\text{NH,H}}$  7.2, amide NH);  $\delta_{\text{C}}$ (74.76 MHz; C<sup>2</sup>HCl<sub>3</sub>) 34.85 ( $\beta$ C), 48.54 ( $\alpha$ C), 51.77 ( $\beta$ CO<sub>2</sub>CH<sub>3</sub>, major product), 52.71 ( $\alpha$ CO<sub>2</sub>CH<sub>3</sub>, minor product), 115.43 (q,  $J_{\text{F,C}}$  285.28, F<sub>3</sub>C), 156.72 (q,  $J_{\text{F}_3\text{C,C}}$  36.48, F<sub>3</sub>CCO) and 170.81 ( $\alpha$ CO<sub>2</sub>H);  $\delta_{\text{F}}$ (282.2 MHz; C<sup>2</sup>HCl<sub>3</sub>) -76.66 (F<sub>3</sub>C);  $m/z$  (EI) 244 (17%, [M + H]<sup>+</sup>), 212 (27, [M - MeOH]<sup>+</sup>), 198 (100, [M - CO<sub>2</sub>H]<sup>+</sup>), 184 (55, [M - CO<sub>2</sub>Me]<sup>+</sup>), 139 (38, [M - CO<sub>2</sub>H and CO<sub>2</sub>Me]<sup>+</sup>), 114 (25, [F<sub>3</sub>CC(O)NH<sub>3</sub>]<sup>+</sup>), 99 (43, [F<sub>3</sub>CC(O)H<sub>2</sub>]<sup>+</sup>), 69 (47, F<sub>3</sub>C<sup>+</sup>) and 59 (41, [CO<sub>2</sub>Me]<sup>+</sup>).

**$\alpha$ -*p*-Benzyloxybenzylpolystyrene-*N*-trifluoroacetyl-(2*S*)-aspartate- $\beta$ -diethyl amide **94**.**



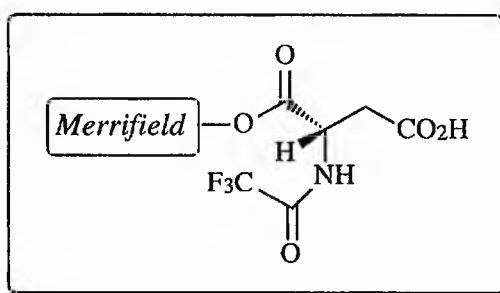
To a stirred suspension of **87** (250 mg, loading: 0.169 mmol of *N*-TFA-asp) in dry DME (5 cm<sup>3</sup>) at room temperature under a nitrogen atmosphere was added *via* a cannula a solution of PyBOP (0.1756 g, 0.338 mmol) in DME (5 cm<sup>3</sup>) followed by the dropwise addition of DIPEA (30.9 mm<sup>3</sup>, 0.177 mmol) in DME (5 cm<sup>3</sup>). The resin was then stirred for 15 min and washed with portions of dry DME, DMF and DME again *via* cannulae to remove excess base and activating agent. A solution of diethylamine (19.3 cm<sup>3</sup>, 0.186 mmol) in DME (5 cm<sup>3</sup>) was then added slowly over 30 min and then the resin was filtered and washed excessively with DCM, DMF, 5% citric acid-DMF and methanol. The resin was dried under reduced pressure (254.9 mg, 53%);  $\nu_{\max}(\text{KBr})/\text{cm}^{-1}$  1637 m (2° amide, solid state);  $\delta_{\text{C}}(74.76 \text{ MHz}; (\text{C}^2\text{H}_3)_2\text{SO})$  19.53 (CH<sub>3</sub>CH<sub>2</sub>N), 35.08 ( $\beta\text{C}$ ), 49.55 ( $\alpha\text{C}$ ), 52.03 (CH<sub>3</sub>CH<sub>2</sub>N), 130.48 (F<sub>3</sub>CCO), 162.97 ( $\beta\text{CONEt}_2$ ) and 170.56 ( $\alpha\text{CO}_2\text{R}$ ).

The amide was cleaved using TFA/H<sub>2</sub>O/TES (90:5:5) at room temperature for 2 h to give the free  $\alpha$ -acid as a pale yellow oil which was dissolved in acetonitrile and purified by reverse-phase HPLC [eluant 0 - 30% acetonitrile in water, eluting at 15% acetonitrile];  $\delta_{\text{H}}(200 \text{ MHz}; \text{C}^2\text{HCl}_3)$  1.19 (6 H, m., 2 x CH<sub>3</sub>CH<sub>2</sub>N), 2.83 (2 H, m, H $\beta$ ), 3.36 (4 H, m, 2 x CH<sub>2</sub>N), 4.44 (1 H, m, H $\alpha$ ), 7.99 and 7.95 (1 H, d,  $J_{\text{NH,H}}$  7.0, NH) and 8.96 (1 H, br s,  $\alpha\text{CO}_2\text{H}$ );  $\delta_{\text{C}}(74.76 \text{ MHz}; \text{C}^2\text{HCl}_3)$  13.50 and 12.46 (CH<sub>3</sub>CH<sub>2</sub>N, 2 conformations), 34.01 ( $\beta\text{C}$ ), 42.52 and 40.89 (CH<sub>3</sub>CH<sub>2</sub>N, 2 conformations), 49.01 ( $\alpha\text{C}$ ), 117.60 (q, F<sub>3</sub>C), 157.19 (q,  $J_{\text{F}_3\text{C,C}}$  37.88, F<sub>3</sub>CCO), 169.91 (C(O)N) and 172.29 ( $\alpha\text{CO}_2\text{H}$ );  $\delta_{\text{F}}(282.2 \text{ MHz}; \text{C}^2\text{HCl}_3)$  -76.72

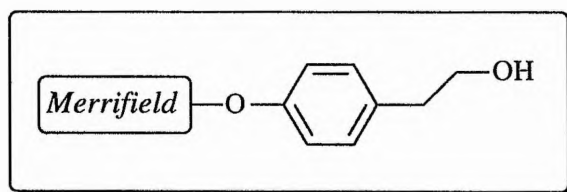


(F<sub>3</sub>C); *m/z* (EI) 284 (42%, M<sup>+</sup>), 239 (15, [M - CO<sub>2</sub>H]<sup>+</sup>), 184 (14, [M - CON(CH<sub>2</sub>CH<sub>3</sub>)<sub>2</sub>]<sup>+</sup>), 114 (29, [CH<sub>2</sub>C(O)N(CH<sub>2</sub>CH<sub>3</sub>)<sub>2</sub>]<sup>+</sup>), 100 (59, [CON(CH<sub>2</sub>CH<sub>3</sub>)]<sup>+</sup>), 74 (46, [(CH<sub>2</sub>CH<sub>3</sub>)<sub>2</sub>NH<sub>2</sub>]<sup>+</sup>), 72 (83, [(CH<sub>2</sub>CH<sub>3</sub>)<sub>2</sub>N]<sup>+</sup>), 58 (100, [CH<sub>3</sub>NCH<sub>2</sub>CH<sub>3</sub>]<sup>+</sup>) and 44 (36, CO<sub>2</sub><sup>+</sup>).

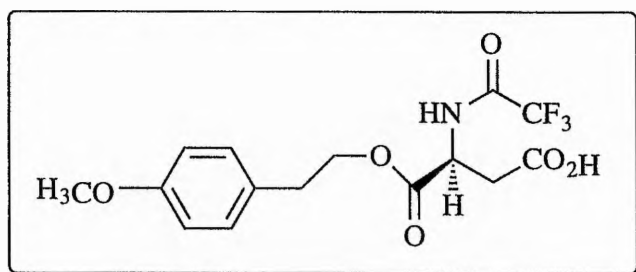
***α-p*-Methylpolystyrene-*N*-trifluoroacetyl-(2*S*)-aspartate 102.**



To a stirred suspension of hydroxymethylpolystyrene resin (1.00 g, loading: 1.16 mmol/g) in dry DMF (10 cm<sup>3</sup>) under a nitrogen atmosphere was added the anhydride **73** (2.4487 g, 11.60 mmol) and DMAP (0.1417 g, 1.16 mmol). The reaction mixture was left at 30 °C for 10 days. The resin was then filtered, washed as copiously in DCM, DMF, 5% citric acid-DMF and methanol and dried under reduced pressure (1.2450 g, 100% loading, 0.863 mmol/g of *N*-TFA-asp);  $\nu_{\max}$ (KBr)/cm<sup>-1</sup> 1735 s (C=O, aryl ester), 1541 (2° amide, solid state), 1170 s (C-F or C-O ester);  $\delta_{\text{C}}$ (74.76 MHz; (C<sup>2</sup>H<sub>3</sub>)<sub>2</sub>SO) 34.79 ( $\beta_{\text{C}}$ ), 49.45 ( $\alpha_{\text{C}}$ ), 117.87 (F<sub>3</sub>C), 157.00 (F<sub>3</sub>CCO), 162.58 ( $\beta_{\text{CO}_2\text{H}}$ ) and 171.37 ( $\alpha_{\text{CO}_2\text{R}}$ );  $\delta_{\text{F}}$ (282.2 MHz; C<sup>2</sup>HCl<sub>3</sub>) -74.87 (F<sub>3</sub>C on  $\alpha$ -ester).

**Polystyrene-4-oxymethyl-2-phenylethanol 97.**

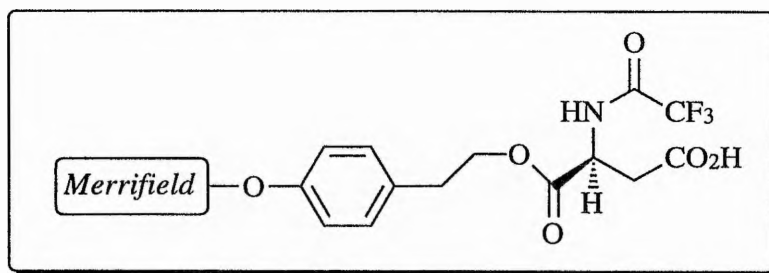
To a stirred suspension of chloromethylpolystyrene-divinylbenzene (Merrifield) resin (200 mg, loading: 0.2 mmol) and 2-(*p*-hydroxyphenyl)-ethanol (0.0691 g, 0.5 mmol) in dry acetonitrile (5 cm<sup>3</sup>) was added powdered potassium hydroxide (0.0281 g, 0.5 mmol) and 18-crown-6 (10 mg, 0.038 mmol). The reaction was left for 3.5 days at 40 °C under a nitrogen atmosphere. The loaded resin was filtered and washed copiously with DCM, 5% citric acid-DMF, DMF and methanol and dried under reduced pressure (226.7 mg, 97% yield, 0.859 mmol/g loading);  $\nu_{\max}(\text{KBr})/\text{cm}^{-1}$  3453 br m (O-H, alcohol), 1517 m (aromatic C=C), 1500 (aromatic C=C) and 1390 (O-H, alcohol);  $\delta_{\text{C}}(74.76 \text{ MHz}; \text{C}_6^2\text{H}_6)$  64.12 ( $\text{CH}_2\text{OH}$ ), 70.33 ( $\text{PhCH}_2$ ), 115.52 (Ar-CH), 130.68 (Ar-CH) and 158.55 (Ar-CO quaternary).

 **$\alpha$ -[2-(4-Methoxyphenyl)-ethyl]-*N*-trifluoroacetyl-(2*S*)-aspartate 103.**

To a stirred solution of 2-(4-methoxyphenyl)-ethanol (0.50 g, 3.29 mmol) in dry DMF (15 cm<sup>3</sup>) under a nitrogen atmosphere was added the anhydride **73** (3.4676 g, 16.43 mmol) and DMAP (0.2007 g, 1.643 mmol). After 2 days the reaction was judged to be complete by tlc [ $R_f$  0.0: ethyl acetate-light petroleum (3:2)]. The solvent was removed under reduced pressure

and the crude product redissolved in ether (20 cm<sup>3</sup>), washed with water (2 x 5 cm<sup>3</sup>), aqueous citric acid (25% w/w 5 cm<sup>3</sup>) and saturated brine (5 cm<sup>3</sup>) and dried (MgSO<sub>4</sub>) (1.98 g, >100% due to impurities). The crude product was purified by silica column chromatography [ethyl acetate-petroleum ether (3:1)] (1.09 g, 91%), mp 89-90 °C; (HRMS: found: 363.0930 C<sub>15</sub>H<sub>16</sub>NO<sub>6</sub>F<sub>3</sub> requires 363.2897);  $\nu_{\max}$ (thin film)/cm<sup>-1</sup> 1736 s (C=O, aryl ester), 1522 m (C=O, 2° amide), 1259 s (C-O, C-F), 1229 s (C-O, C-F) and 1181 s (C-O, C-F),  $\delta_{\text{H}}$ (200 MHz; C<sup>2</sup>HCl<sub>3</sub>) 2.85 - 3.14 (4 H, m, H <sup>$\beta$</sup>  and CH<sub>2</sub>O- overlapping), 3.77 (3 H, s, ArOCH<sub>3</sub>), 4.33 (2 H, m, PhCH<sub>2</sub>), 4.82 (1 H, dt,  $J_{\text{NH,H}}$  7.78, H <sup>$\alpha$</sup> ), 6.85 and 6.80 (2 H, d,  $J$  8.45, Ar-H), 7.11 and 7.07 (2 H, d,  $J$  8.45, Ar-H), 7.54 and 7.50 (1 H, d,  $J_{\text{NH,H}}$  7.78, NH), 7.69 (1 H, br s, CO<sub>2</sub>H);  $\delta_{\text{C}}$ (74.76 MHz; C<sup>2</sup>HCl<sub>3</sub>) 33.66 ( $\beta$ C), 34.88 (CH<sub>2</sub>O-), 48.50 ( $\alpha$ C), 55.09 (PhOCH<sub>3</sub>), 67.04 (PhCH<sub>2</sub>), 114.04 (Ar-CH), 115.46 (q,  $J_{\text{F,C}}$  287.65, F<sub>3</sub>C), 128.98 (Ar-C quaternary), 129.69 (Ar-CH), 157.05 (q,  $J_{\text{F}_3\text{C,C}}$  38.91, F<sub>3</sub>CCO), 158.42 (ArC quaternary) and 168.87 (CO<sub>2</sub>H);  $\delta_{\text{F}}$ (282.2 MHz; C<sup>2</sup>HCl<sub>3</sub>) -76.67 (F<sub>3</sub>C);  $m/z$  (EI) 363 (11%, M<sup>+</sup>), 134 (100, [NH<sub>3</sub>CH(CO<sub>2</sub>H)CH<sub>2</sub>CO<sub>2</sub>H]<sup>+</sup>).

**$\alpha$ -(Polystyrene-4-oxymethyl-2-phenylethyl)-*N*-trifluoroacetyl-(2*S*)-aspartate  
99.**



To a stirred suspension of **97** (1.50 g, loading: 1.431 mmol/g) in dry DMF was added the anhydride **73** (1.007 g, 4.77 mmol) and DMAP (0.0583 g, 0.477 mmol). The reaction was stirred at room temperature for 10 days. The resin was filtered, washed copiously with DCM, 5% citric acid-DMF, DMF and methanol and dried under reduced pressure (1.711 g, 97%,

0.541 mmol/g in *N*-TFA-asp);  $\nu_{\max}(\text{KBr})/\text{cm}^{-1}$  1727 s (C=O, aryl ester), 1653 m (2° amide, solid state) and 1078 w (C-O, C-F);  $\delta_{\text{C}}(74.76 \text{ MHz; C}_6\text{H}_6)$  34.09 ( $\beta\text{C}$  of aspartate), 40.04 (CH<sub>2</sub>O-), 48.42 ( $\alpha\text{C}$  of aspartate), 114.31 (Ar-CH), 117.44 (F<sub>3</sub>C), 129.35 (Ar-CH), 145.15 (Ar-C quaternary), 157.08 (Ar-CO), 169.11 (F<sub>3</sub>CCO), 173.77 (CO<sub>2</sub>H of aspartate);  $\delta_{\text{F}}(282.2 \text{ MHz; (C}^2\text{H}_3)_2\text{SO})$  -74.81 (F<sub>3</sub>C).

## 3.2 Molecular Cloning

### 3.2.1 Cloning Techniques

**Bacterial Strains and Plasmids.** Chromosomal DNA for the amplification of *thrC* was prepared from *E. coli* strain TG1 or XL1-Blue. Minipreparation of the plasmid pSG4 containing the gene for  $\beta$ -methylaspartase was from *E. coli* strain TG1 (K-12  $\Delta$  [*lac-pro*] *supE thi hsdR5 / F'-traD36 proA<sup>+</sup> B<sup>+</sup> lac*]  $\mathcal{Q}$  *lacZ*  $\Delta$  *M15*).<sup>260</sup> *E. coli* strains JM83 and XL1-Blue were used as hosts for cloning experiments and BL21(DE3) was used as the host strain for analytical expression of recombinant proteins. Transcription vectors pGEM<sup>TM</sup>-3Zf(+), pGEM<sup>TM</sup>-5Zf(+) and pGEM-T<sup>TM</sup> and the expression vector pKK223-3 were purchased from Promega. The expression vectors pET-3a, pET-3b and pET-16b were purchased from Novagen, AMS Biotechnology. All vectors contain the ampicillin resistance gene (Amp<sup>R</sup>) and a multi-cloning site contained in the N-terminal region of the *lacZ* operon.

**Isolation of chromosomal DNA from *E. coli*.** This was performed either as described for the minipreparation of plasmid DNA from *E. coli* strain TG1 (K-12  $\Delta$  [*lac-pro*] *supE thi hsdR5 / F'-traD36 proA<sup>+</sup> B<sup>+</sup> lac*]  $\mathcal{Q}$  *lacZ*  $\Delta$  *M15*) or using the SDS-Proteinase K method of large-scale chromosomal DNA preparation as follows:<sup>270</sup>

A starter culture of XL1-Blue cells in LB (10 cm<sup>3</sup>) was incubated at 37 °C for 3 h with shaking. Luria broth (200 cm<sup>3</sup>) was inoculated with 200  $\mu$ l of this starter culture and incubated overnight at 37 °C with shaking. Cells were pelleted by centrifugation (4 °C for 15 min at 6,000 xg) and stored at -20 °C for 2 h before resuspension in 40 cm<sup>3</sup> TE (10 mM Tris. HCl (pH 8.0), 1 mM EDTA). The cell suspension was stored on ice for 10 min then centrifuged (4 °C for 45 min at 3,000 rpm). The pellet was then resuspended in 3.2 cm<sup>3</sup> 50 mM Tris.HCl (pH 8.0) and 0.7 M sucrose. To this was added 0.6 cm<sup>3</sup> lysozyme (20 mg/cm<sup>3</sup>) and the mixture was stored on ice for 5 min. 0.6 cm<sup>3</sup> EDTA (pH 8.0) and 0.5 cm<sup>3</sup> 10%

(w/v) SDS were added and the mixture stored on ice for a further 5 min. The suspension was adjusted to 1% (w/v) SDS and 0.5 mg/cm<sup>3</sup> Proteinase K (Promega) and incubated at 55 °C for 7.5 h. The reaction was halted by the addition of 5 cm<sup>3</sup> phenol (pre-equilibrated with 1 M Tris.HCl (pH 8.0)) and mixed for 3 h. The mixture was transferred to a Falcon tube and centrifuged (25 °C, 10 min at 4, 000 rpm), discarding the lower phenol layer. 1.0 vol phenol/ chloroform was added to the aqueous layer, mixed gently and the mixture centrifuged at room temperature (10 min at 3, 000 rpm). The upper aqueous layer was transferred to a fresh tube and stored on ice for 5 min before the addition of 2.5 cm<sup>3</sup> 7.5 M ammonium acetate and 10 cm<sup>3</sup> ethanol to precipitate the DNA. The mixture was stored at -20 °C for 30 min after which the DNA was pelleted by centrifugation at room temperature (10 min at 3, 000 rpm), rinsed with 70% (v/v) ethanol and air-dried for 2 min. The DNA was dissolved in H<sub>2</sub>O and stored at -20°C before further use.

**Restriction enzyme digestions.** Restriction enzymes and buffers were purchased from Promega or New England Biolabs. All plasmid DNA or gene fragment digestions with restriction enzymes were carried out under the conditions specified by the supplier. In general 1µg of DNA was digested with 1 unit of restriction enzyme in the appropriate buffer in a total volume of 20 µl. DNA was purified by phenol/chloroform extraction.

**CIAP-digestions.** In general 1µg of DNA was digested with 1 unit of calf intestinal alkaline phosphatase (CIAP, Promega) in the appropriate buffer in a total volume of 20 µl incubating at 37 °C for 1 h.. The plasmid DNA was purified by extraction 3 times with phenol/chloroform, precipitation using 2.5 volumes ethanol and 1/20 volumes of 4M sodium acetate and the mixture was stored at -20 °C for a minimum of 20 min. The DNA was pelleted by centrifugation (13, 000 rpm for 15-20 min and redissolved in H<sub>2</sub>O to a final concentration of 1 µg/µl.

**Agarose gel electrophoresis.** Flat bed 1% agarose gels were prepared using 1 x TAE (0.04M Tris-acetate, 0.001M EDTA) containing 0.2 µg/cm<sup>3</sup> ethidium bromide. A 1 kb DNA

ladder (Gibco) was run against DNA fragments to determine their approximate size.

**Gel purification of DNA fragments.** To isolate DNA fragments from low melting point (LMP) agarose the following methods were used:

(i) *Phenol/acetate method.* The slice of LMP agarose containing the DNA fragment to be isolated was placed in 500  $\mu$ l of phenol/acetate (phenol equilibrated with 0.3 M sodium acetate) and heated at 70 °C until the agarose was melted. After vortexing the tube was incubated on ice for 10 min and then spun in a microfuge (10 min at 13,000 rpm). The upper, aqueous layer was removed to a fresh tube and subjected to a further extraction with phenol/acetate. DNA was precipitated from the resulting aqueous fraction by ethanol precipitation, dried in a vacuum dessicator and then redissolved in TE (10 mM Tris. HCl (pH 8.0), 1 mM EDTA) or H<sub>2</sub>O as appropriate to give an approximate concentration of 1 mg/cm<sup>3</sup> DNA.

(ii) *WIZARD™ Prep DNA Purification System.* DNA was isolated from a section of LMP agarose using the WIZARD™ Prep DNA purification system (Promega) according to the manufacturer's instructions.

(iii) *Seaplaque™ agarose.* The desired DNA fragment was run into a section of Seaplaque agarose (FMC Bioproducts) which was then excised, melted by incubation at 70 °C and used directly for in-gel ligations according to the manufacturer's instructions.

**Blunt-end repair of overhangs.** 5'-Overhangs were filled in to produce blunt ends using T4 DNA polymerase. 1  $\mu$ g of plasmid DNA was incubated for 5 min with 5 units of T4 DNA polymerase (Promega), 100  $\mu$ M of each dNTP in T4 polymerase buffer (33 mM Tris.acetate (pH 7.9), 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM DTT, 0.1 mg/cm<sup>3</sup> BSA). The reaction was halted by heating to 75 °C for 10 min.

**Ligations.** Ligation reactions were routinely carried out with 0.5 - 1 µg of vector DNA, 1.5 µg of insert DNA, 4 Weiss units of T4 DNA ligase (Promega) in a total of 25 µl of T4 DNA ligase buffer [30 mM Tris.HCl (pH 7.8), 10 mM MgCl<sub>2</sub>, 10 mM DTT, 0.5 mM ATP]. The reactions were incubated at room temperature or 16 °C overnight. Ligation reactions for the insertion of gene fragments into pGEM-T were carried out using 50 ng pGEM-T vector (Promega), with vector:insert ratios ranging from 1:1 and 1:10 using 1 unit T4 DNA ligase in a total volume of 10 µl. The concentration of cloned DNA and plasmid DNA solutions were estimated from their luminescence on agarose gels. The following formula was used to calculate the amount of insert needed for a ratio of 1:1 insert:vector:

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{insert : vector molar ratio} = \text{ng of insert}$$

In-gel ligations were performed using Seaplaque™ agarose according to the manufacturer's instructions.

**Transformation of *E. coli*.** *E. coli* was transformed with plasmid DNA after making competent cells using the calcium chloride method (adapted from Cohen *et al.*<sup>267</sup>). 10 cm<sup>3</sup> cultures of cells were grown to exponential phase in Luria Bertani broth [LB broth: 1% (w/v) bacto-tryptone, 0.5% (w/v) bacto-yeast extract, 1% (w/v) NaCl, pH 7.5] cooled to 4 °C for 10 min and pelleted by centrifugation (4 °C for 10 min at 4,000 rpm). Each pellet was resuspended in 10 cm<sup>3</sup> of ice-cold 0.1 M CaCl<sub>2</sub>, stored on ice for 10 min and pelleted again as before. Cells were then suspended in 200 µl 0.1 M CaCl<sub>2</sub> per pellet and stored at 4 °C for at least 30 min to attain competency. Cells could be stored in CaCl<sub>2</sub> solution at 4 °C for 24 - 48 h before losing competency.<sup>266</sup> 200 µl of competent cells were gently mixed with plasmid DNA (up to 50 ng in 10 cm<sup>3</sup> or less of solution) in sterile eppendorfs and the mixture stored on ice for 30 min. The tubes were transferred to a circulating water bath preheated to 42 °C for exactly 90 sec and then rapidly transferred to an ice-bath for 1-2 min. 800 µl of LB was added to each tube which were then incubated at 3°C for 45 min. Up to 200 µl of each culture was



then transferred to L-agar plates coated with 40  $\mu$ l of X-gal (20 mg/ml in DMF) and 4  $\mu$ l IPTG (200 mg/ml in DMF) for overnight incubation at 37 °C.

**Minipreparation of plasmid DNA.** 10 cm<sup>3</sup> of LB containing 10  $\mu$ g/cm<sup>3</sup> ampicillin was inoculated with a single colony of transformed *E. coli* JM83 or XL1-Blue and incubated overnight at 37 °C with shaking. The cells were pelleted (4 °C, 4,000 rpm, 10 min) and the supernatant removed. Pellets were resuspended in 700  $\mu$ l of STET buffer (0.1 M NaCl, 10 mM Tris.HCl (pH 8.0), 1 mM EDTA (pH 8.0), 5% (v/v) Triton X-100), containing 1 mg/cm<sup>3</sup> lysozyme, and incubated at room temperature for 10 min. 30  $\mu$ l of 10% (w/v) SDS was added and mixed in then 75 cm<sup>3</sup> of 4 M potassium acetate was added and mixed. After incubation on ice for 15 min the mixture was spun in a microfuge for 15 min at 13,000 rpm. The supernatant was transferred to a fresh tube, 1  $\mu$ l of 1 mg/cm<sup>3</sup> RNase A (Sigma) was added and the mixture incubated at room temperature for 15 min before repeated extraction with phenol/chloroform. The aqueous phase was removed to a fresh tube, and the DNA ethanol precipitated twice. The DNA precipitate was washed with 70% (v/v) ethanol and resuspended in 100 cm<sup>3</sup> H<sub>2</sub>O.

**Maxipreparation of plasmid DNA.** This was performed according to the QIAGEN maxiprep plasmid purification protocol. LB was inoculated with a single colony of *E. coli* XL1-Blue, or with a starter culture of XL1-Blue cells which had been transformed with miniprep plasmid DNA isolated as above, and grown overnight at 37 °C with shaking. For high copy number plasmids (pGEM vectors) 150 cm<sup>3</sup> cultures were used, for low copy number plasmids such as pET vectors 500 cm<sup>3</sup> cultures were grown. 10 cm<sup>3</sup> starter cultures in LB were incubated at 37 °C for 3 or 4 h.

### 3.2.2 Polymerase Chain Reaction

**Standard PCR Protocol.** Sequences encoding *thrC* or the gene for  $\beta$ -methylaspartase were amplified from chromosomal DNA *via* PCR. Reactions were carried out in TAQ DNA

polymerase buffer [50 mM KCl, 10 mM Tris.HCl (pH 9.0), 0.1% (v/v) Triton X-100, 2.5 mM MgCl<sub>2</sub>] with 1 mM DTT, 0.125 mM of each dNTP, 200 pmol of each primer and up to 1 µg of template DNA in a total 100 µl reaction volume. Reactions were overlaid with 100 cm<sup>3</sup> mineral oil (Sigma), heated to 94 °C for 5 min and then held at 85 °C whilst 2.5 units of TAQ DNA polymerase (Boehringer Mannheim or Promega) were added to the aqueous phase of each reaction.

PCR was carried out using the following programs:

For gene fragments with *Nde* I-sites at each terminus: 1: 94 °C, 2 min; 2: 42 °C, 1 min; 3: 72 °C, 2 min; 4: 94 °C, 1 min; 5: Cycle of steps 2-4, x 20; 6: 72 °C, 5 min; 7: Hold at 4 °C indefinitely.

For gene fragments with *Nde* I-sites at the N-terminus, *Bam*H I-sites at the C-terminus: as above except step 2: 40 °C, 1 min.

Products of PCR reactions were extracted against phenol/chloroform and ethanol precipitated before further use.

**Overlap PCR.** The reaction conditions were as stated above except that half-gene fragments were produced using either the external forward primer and the internal reverse primer, or the external reverse primer and the internal forward primer. These were gel-purified and used as the template for a PCR reaction using outside primers only to produce a whole gene fragment.

**Cycle Sequencing.** Plasmid DNA sequencing was carried out using the fmol™ kit (Promega) according to the supplier's instructions. Sequencing gels were made using 31.5 g urea, 15 cm<sup>3</sup> Sequagel™ Concentrate (National Diagnostics), 15 cm<sup>3</sup> x5 TBE [x5 composition: 54 g Tris base, 27.5 g boric acid, 0.5 M EDTA (pH 8.0) and dH<sub>2</sub>O to 1l], making up to a total volume of 75 cm<sup>3</sup>. To 68 cm<sup>3</sup> of this was added 300 µl 10% APS solution and 100 µl TEMED. Gels were preheated to 44 °C before applying samples for separation, run at 200 Vand dried under vacuum on a slab-gel drier for 2 h at 80 °C. Dried gels were exposed on X-Ray photographic film overnight.

### 3.2.3 Oligonucleotides

Oligonucleotides were synthesised using phosphoramidite chemistry on an oligonucleotide synthesizer in the Division of Cell and Molecular Biology at the University of St. Andrews or obtained from Oswel. These oligonucleotides are listed in Table 3.1 and 3.2.

**Table 3.1:** *Oligonucleotides used for amplification of gene fragments.*

<b>TS1.</b> Forward primer at N-terminus of <i>thrC</i> . <i>Nde</i> I-site	Met Lys Leu Tyr Asn Leu 5'-TTTT <u>CAT ATG</u> AAA CTC TAC AAT CTG- 3' <i>Nde</i> I
<b>TS2.</b> Reverse primer at C-terminus of <i>thrC</i> . Stop codon, <i>Nde</i> I-site.	Met * Glu His Asn Met Met 5' -AAAA <u>CAT ATG</u> TTA CTG ATG ATT CAT CAT- 3' <i>Nde</i> I
<b>TS3.</b> Forward primer from base 643 on <i>thrC</i>	His Gln Pro Phe Ala 5' - AA <u>CAT CAG</u> CCG TTT GCT G - 3'
<b>TS4.</b> Reverse primer from base 660 on <i>thrC</i>	5' - CAG CAA ACG GCT GAT GTT - 3'
<b>TS5.</b> Forward primer at N-terminus of <i>thrC</i> . <i>Nde</i> I-site.	Met Lys Leu Tyr Asn Leu 5' - GTCA GTC AGT <u>CAT ATG</u> AAA CTC TAC AAT CTG - 3' <i>Nde</i> I
<b>TS6.</b> Reverse primer at C-terminus of <i>thrC</i> . Stop codon, <i>Bam</i> H I-site.	Met * Glu His Asn Met Met 5' - ATGC <u>GGA TCC</u> TTA CTG ATG ATT CAT CAT - 3' <i>Bam</i> H I
<b>MA1.</b> Forward primer at N-terminus of the gene coding for $\beta$ -methylaspartase. <i>Nde</i> I-site.	Met Lys Ile Val Asp Val 5'-TTTT <u>CAT ATG</u> AAA ATT GTT GAC GTA- 3' <i>Nde</i> I
<b>MA2.</b> Reverse primer at C-terminus of the gene coding for $\beta$ -methylaspartase. Stop codon, <i>Nde</i> I-site.	Met * Lys Arg Arg Gly Val 5' -AAAA <u>CAT ATG</u> TTA TTT TCT TCT TCC TAC- 3' <i>Nde</i> I
<b>MA 3.</b> Forward primer from base 640 of the gene coding for $\beta$ -methylaspartase	Arg * Met Val Lys 5' - TA CGT TAA ATG GTT AAG A - 3'
<b>MA4.</b> Reverse primer from base 657 of the gene coding for $\beta$ -methylaspartase	5' - TCT TAA CCA TTT AAC GTA - 3'
<b>MA5.</b> Forward primer at N-terminus of the gene coding for $\beta$ -methylaspartase. <i>Nde</i> I-site	Met Lys Ile Val Asp Val 5' - GTCA GTC AGT <u>CAT ATG</u> AAA ATT GTT GAC GTA - 3' <i>Nde</i> I
<b>MA6.</b> Reverse primer at C-terminus of the gene coding for $\beta$ -methylaspartase. Stop codon, <i>Bam</i> H I-site.	Met * Lys Arg Arg Gly Val 5' - ATGC <u>GGA TCC</u> TTA TTT TCT TCT TCC TAC - 3' <i>Bam</i> H I

Oligonucleotide primers were designed to contain restriction sites to enable PCR products to be easily cloned. For forward coding sense primers the encoded amino acid sequence is shown above the nucleic acid sequence. Stop codons are indicated on the complementary strand by a \*.

**Table 3.2:** Oligonucleotides used in sequencing experiments.

<p><b>SP6 promoter primer.</b></p> <p>5' - d(GAT TTA GGT GAC ACT ATA G) - 3'</p>
<p><b>T7 promoter primer.</b></p> <p>5' - d(TAA TAC GAC TCA CTA TAG GG) - 3'</p>

Primer concentrations were calculated from their  $A_{260}$  values using the following formula:

$$\text{Concentration (pmol } \mu\text{l}^{-1}) = \left[ \frac{A_{260}}{(nA \times 1.52) + (nC \times 0.75) + (nG \times 1.17) + (nT \times 0.92)} \right] \times \text{dilution factor}$$

### 3.2.4 Expression in *E. coli*

**SDS-polyacrylamide gel electrophoresis (SDS-PAGE).** The discontinuous buffer gel system based on that of Laemmli was used.<sup>317</sup> All gels were constructed with a 4% stacking gel and a 10% polyacrylamide resolving gel. Gels were run at 40 mA upper current limit and 250 V upper voltage limit. Generally proteins were run through the stacking gel at 120 V and run through the resolving gel at 225 V. Gels were stained overnight in Coomassie blue according to the method of Sambrook *et al.*,<sup>266</sup> destained in destaining buffer (5% methanol, 10% acetic acid in aqueous solution) and dried under vacuum on a slab-gel drier at 60 °C for 2 h.

**Analytical expression of proteins.** 10 cm<sup>3</sup> cultures of LB-ampicillin (LB with 100 µg/cm<sup>3</sup> ampicillin) were inoculated with single colonies of *E. coli* BL21(DE3) transformed

with pTN-TS expression vectors and grown overnight at 37°C with shaking. 2 x 1 cm<sup>3</sup> of each overnight culture was transferred to 2 x 9 cm<sup>3</sup> of fresh LB-ampicillin (one culture for induction and one as a control) and incubated at 37 °C for 3 h with shaking. To the cultures to be induced was added 100 µl of 100 mM IPTG and all cultures were incubated for a further 2 h at 37 °C with shaking. The cultures were then kept at 4 °C for a short time before the cells were pelleted (4°C, 4, 000 rpm, 10 min), resuspended in 500 µl chilled PBS buffer containing 12 mg/cm<sup>3</sup> lysozyme and transferred to sterile eppendorfs. The cells were then lysed using a probe sonicator in several 10 sec bursts resting on ice in between. Insoluble material was removed by centrifugation in a microcentrifuge (4 °C, 10, 000 rpm, 30 min) and 20-40 µl of the lysate was subjected to SDS-PAGE analysis using molecular weight markers of 66 and 45 kDa. The total protein concentration of the lysates was measured using the BCA protein assay (Pharmacia) as recommended by the manufacturer, using a BSA calibration curve for comparison. Varying amounts of two of the lysates were applied to a gel first of all in order to determine what protein concentration would produce the clearest results. Approximately 50 µg of protein from each sample was loaded onto gels.

**Analytical purification of His-tagged proteins.** Ni-NTA Resin synthesised by C. Botting at the Division of Cell and Molecular Biology was prepared for use by equilibrating with 10 volumes of the sonication buffer described above for 1 h. 250 µl of the cell lysate from cells transformed with vectors pTN-TS5 - 8 were stirred with between 100 and 140 µl of a 50% Ni<sup>2+</sup> resin slurry at room temperature for 1 h. The slurry was pelleted by centrifugation on a microfuge (10, 000 rpm, 10 min) and the supernatant retained for analysis. The resin was then washed with 150 µl TBS (20 mM Tris, 0.3 M NaCl, pH 8.0), centrifuged as before and the supernatant retained for analysis. The protein was then eluted from the resin by washing with 150 cm<sup>3</sup> of elution buffer (250 mM imidazole, 20 mM Tris, 0.3 M NaCl, pH 7.5). This was centrifuged and the supernatant containing the His-tagged proteins removed. The supernatant from each of the Ni<sup>2+</sup> treatment steps was analysed by SDS-PAGE.

**Plasmid Stability Test.** *E. coli* BL21(DE3) freshly transformed with pTN-TS vectors 5-8

were grown in 5 cm<sup>3</sup> portions of LB-medium containing 100mg/cm<sup>3</sup> ampicillin overnight at 37 °C with shaking. A small subculture was transferred into 500 cm<sup>3</sup> of fresh LB medium containing 100 mg/cm<sup>3</sup> ampicillin and allowed to grow at 37 °C to an OD<sub>600</sub> reading of 0.6. Three aliquots of 1,000 mm<sup>3</sup>, 147 mm<sup>3</sup> and 135 mm<sup>3</sup> of LB were transferred to three sterile eppendorfs. A primary dilution of bacteria was prepared by transferring 1 mm<sup>3</sup> of log phase culture into 1 cm<sup>3</sup> of LB. Two secondary dilutions were prepared by transferring 3 mm<sup>3</sup> and 15 mm<sup>3</sup> of the primary dilution into tubes containing 147 mm<sup>3</sup> (dilution A) and 135 mm<sup>3</sup> (dilution B) of LB, respectively. Bacteria were plated onto L-agar plates containing no additives, 100µg/ml ampicillin only, 0.4 mM IPTG only or both 100µg/ml ampicillin and 0.4mM IPTG., as indicated in Table 2.6. The plates were incubated overnight at 37 °C.

## **REFERENCES**

---

## References

---

1. P. P. Cohen, 'The Enzymes', ed. J. B. Sumner and K. Myrbäck, Academic Press Inc., New York, 1951, vol. 2, p. 1040.
2. W. A. Wood and I. C. Gunsalus, *J. Biol. Chem.*, 1951, **190**, 403 - 416.
3. E. F. Gale, *Adv. Enzymol.*, 1946, **6**, 1 - 32.
4. D. Gani, 'Pyridoxal Dependent Systems', Ch. 6.5, in 'Comprehensive Medicinal Chemistry - The Rational Design, Mechanistic Study and Therapeutic Application of Chemical Compounds', ed. C. Hansch, P. G. Sammes, and J. B. Taylor, Pergamon Press, Oxford, 1990, vol. 2, pp. 213 - 254.
5. H. G. Floss and J. C. Vederas, 'Stereochemistry of pyridoxal phosphate-catalyzed reactions', Ch. 4, in 'Stereochemistry', ed. C. Tamm, Elsevier Biomedical Press, 1982, pp. 161 - 165 and 188 - 195.
6. 'Enzyme Nomenclature', in 'Recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology', Academic Press, 1992.
7. P. György, *Nature*, 1934, **133**, 498 - 501.
8. J. Goldberger and R. D. Lille, *Public Health Reports (US)*, 1926, **41**, 1025 - 1027.
9. S. Ohdake, *Bull. Chem. Soc. Jpn.*, 1932, **8**, 11 - 15.
10. T. W. Birch and P. György, *Biochem. J.*, 1936, **30**, 304 - 315.
11. J. C. Keresztesy and J. R. Stevens, *Proc. Soc. Exp. Biol. Med.*, 1938, **38**, 64 - 69.
12. E. T. Stiller, J. C. Keresztesy and J. R. Stevens, *J. Am. Chem. Soc.*, 1938, **61**, 1237 - 1242.
13. S. A. Harris and K. Folkers, *Science*, 1939, **89**, 347 - 350.
14. R. Kuhn and G. Wendt, *Ber. Deutsch. Chem. Ges.*, 1938, **71**, 780 - 783.
15. E. F. Gale and H. M. R. Epps, *Biochem. J.*, 1944, **38**, 232 - 256.
16. W. D. Bellamy and I. C. Gunsalus, *J. Bacteriol.*, 1943, **46**, 573 - 578.
17. E. F. Snell, *J. Am. Chem. Soc.*, 1945, **67**, 194 - 197.
18. H. C. Dunathan and J. G. Voet, *Proc. Natl. Acad. Sci. USA*, 1974, **71**, 3888 - 3891.
19. C. Parsot, *EMBO J.*, 1986, **5**, 3013 - 3019.
20. P. K. Mehta, T. I. Hale and P. Christen, *Eur. J. Biochem.*, 1993, **214**, 549 - 561.
21. M. Gribskov, R. Lüthy and D. Eisenberg, *Methods Enzymol.*, 1990, **183**, 146 - 159.
22. P. K. Mehta and P. Christen, *Eur. J. Biochem.*, 1993, **211**, 373 - 376.
23. P. K. Mehta and P. Christen, *Biochem. Biophys. Res. Commun.*, 1994, **198**, 138 - 143.
24. F. W. Alexander, E. Sandmeier, P. K. Mehta and P. Christen, *Eur. J. Biochem.*, 1994, **219**, 953 - 960.
25. R. A. John, *Biochim. Biophys. Acta*, 1995, **1248**, 81 - 96.
26. P. K. Mehta, T. I. Hale and P. Christen, *Eur. J. Biochem.*, 1989, **186**, 249 - 253.
27. A. A. Antson, T. V. Demidkinar, P. Gollnick, Z. Dauter, R. L. Von Teresh, J. Long, S. N. Berzshnoy, R. S. Philips, E. H. Harutyunyan and K. S. Wilson, *Biochemistry*, 1993, **32**, 4195 - 4206.
28. M. D. Toney, E. Hohenester, S. W. Cowan and J. N. Jansonius, *Science*, 1993, **261**, 756 - 759.
29. J. M. Green, W. K. Merkel and B. P. Nichols, *J. Bacteriol.*, 1992, **174**, 5317 - 5323.



30. R. A. Jensen and W. Gu, *J. Bacteriol.*, 1996, **178**, 2161 - 2171.
31. S. Pascarella, V. Schirch and F. Bossa, *FEBS Lett.*, 1993, **331**, 145 - 149.
32. C. Ouzounis and C. Sander, *FEBS Lett.*, 1993, **332**, 159 - 164.
33. C. Babbitt and J. A. Gerlt, *FASEB J.*, 1995, **9**, A1266.
34. U. J. Kang and T. H. Joh, *Mol. Brain Res.*, 1990, **8**, 83 - 87.
35. E. Sandmeier, T. I. Hale and P. Christen, *Eur. J. Biochem.*, 1994, **221**, 997 - 1002.
36. P. Bork and K. Rohde, *Biochem. Biophys. Res. Commun.*, 1990, **171**, 1319 - 1325.
37. C. Parsot, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 5207 - 5210.
38. J. Belfaiza, C. Parsot, A. Martel, C. Bouthier de la Tour, D. Margarita, G. N. Cohen and I. Saint-Girons, *Proc. Natl. Acad. Sci. USA*, 1986, **83**, 867 - 871.
39. A. E. Braunstein and M. M. Shemyakin, *Biokhimiya*, 1953, **18**, 393 - 411.
40. D. E. Metzler, M. Ikawa and E. E. Snell, *J. Am. Chem. Soc.*, 1954, **76**, 648 - 652.
41. F. Barclay, PhD Thesis, 1994, University of St Andrews.
42. D. M. Smith, N. R. Thomas and D. Gani, *Experientia*, 1991, **47**, 1104 - 1118.
43. H. Lis, *Biochim. Biophys. Acta*, 1958, **28**, 191 - 197.
44. W. T. Jenkins, D. A. Yphantis and I. W. Sizer, *J. Biol. Chem.*, 1959, **234**, 51 - 52.
45. H. Lis, P. Fasella, C. Turano and P. Vecchini, *Biochim. Biophys. Acta*, 1960, **45**, 529 - 536.
46. D. Gani and J. A. Robinson, 'Stereochemical and Mechanistic Features of Pyridoxal-5'-phosphate-Dependent Enzymes', Ch. 1, in 'Natural Product Reports', 1986, pp. 1 - 50.
47. V. C. Emery and M. Akhtar, 'Pyridoxal Phosphate-Dependent Enzymes', Ch. 18, in 'Enzyme Mechanisms', ed. M. I. Page and A. Williams, Royal Society of Chemistry, London, pp. 345 - 389.
48. V. P. Almazov, Y. V. Morozov, F. A. Savin and B. S. Sukhareva, *Int. J. Quantum Chem.*, 1979, **16**, 769 - 775.
49. J. R. Fischer and E. H. Abbott, *J. Am. Chem. Soc.*, 1979, **101**, 2781 - 2782.
50. M. Y. Karpeisky and V. I. Ivanov, *Nature*, 1966, **210**, 493 - 496.
51. V. I. Ivanov and M. Y. Karpeisky, *Adv. Enzymol.*, 1969, **32**, 21 - 54.
52. M. Yu Torchinsky and A. E. Braunstein, *FEBS Symp.*, 1979, **52**, 293 - 303.
53. S. C. Almo, D. L. Smith, A. T. Danishefsky and D. Ringe, *Protein Eng.*, 1994, **7**, 405 - 412.
54. V. N. Malashkevich, J. Jager, M. Ziak, U. Sauder, H. Gehring, P. Christen and J. N. Jansonius, *Biochemistry*, 1995, **34**, 405 - 414.
55. S. Rhee, M. M. Silva, C. C. Hyde, P. H. Rogers, C. M. Metzler, D. E. Metzler and A. Arnone, *J. Biol. Chem.*, 1997, **272**, 17293 - 17302.
56. C. C. Hyde and E. W. Miles, *Biotech.*, 1990, **8**, 27 - 32.
57. S. Rhee, K. D. Parris, C. C. Hyde, S. A. Ahmed, E. W. Miles and D. R. Davies, *Biochemistry*, 1997, **36**, 7664 - 7680.

58. F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer Jr., M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi and M. Tasumi, *J. Mol. Biol.*, 1977, **112**, 535 - 542.
59. E. E. Abola, F. C. Bernstein, S. H. Bryant, T. F. Koetzle and J. Weng, 'Protein Data Bank' in 'Crystallographic Databases - Information Content, Software Systems, Scientific Applications', ed. F. H. Allen, G. Bergerhoff and R. Sievers, Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester, 1987, pp. 107 - 132.
60. S. Rhee, K. D. Parris, S. A. Ahmed, E. W. Miles and D. R. Davies, *Biochemistry*, 1996, **35**, 4211 - 4221.
61. G.D. McClure and P. F. Cook, *Biochemistry*, 1994, **33**, 1674 - 1683.
62. G. B. Strambini, P. Cioni, A. Peracchi and A. Mozzarelli, *Biochemisry*, 1992, **31**, 7527 - 7534.
63. G. B. Strambini, P. Cioni, A. Peracchi and A. Mozzarelli, *Biochemistry*, 1992, **31**, 7535 - 7542.
64. G. B. Strambini, P. Cioni and P. F. Cook, *Biochemistry*, 1996, **35**, 8392 - 8400.
65. K. D. Schnackerz, C-H. Tai, J. W. Simmons, T. M. Jacobson, G. S. Jagannatha Rao and P. F. Cook, *Biochemistry*, 1995, **34**, 12152 - 12160.
66. C-H. Tai, S. R. Nalabolu, J. W Simmons, T. M. Jacobson and P. F. Cook, *Biochemistry*, 1995, **34**, 12311 - 12322.
67. E. U. Woehl, C-H. Tai, M. F. Dunn and P. F. Cook, *Biochemistry*, 1996, **35**, 4776 - 4783.
68. P. F. Cook, S. Hara, S. Nalabolu and K. D. Schnackerz, *Biochemistry*, 1992, **31**, 2298 - 2303.
69. C-C. Hwang, E. U. Woehl, D. E.Minter, M. F. Dunn and P. F. Cook, *Biochemistry*, 1996, **35**, 6358 - 6365.
70. A. M. Kayastha, Y. Sawa, S. Nagata and E. W. Miles, *J. Biol. Chem.*, 1991, **266**, 7618 - 7625.
71. X-J. Yang and E. W. Miles, *J. Biol. Chem.*, 1993, **268**, 22269 - 22272.
72. Z. Lu, S. Nagata, P. McPhie and E. W. Miles, *J. Biol. Chem.*, 1993, **268**, 8727 - 8734.
73. A. Ahmed and E. W. Miles, *FASEB J.*, 1996, **10**, 1374.
74. 'Biochemistry', 3rd Ed., ed. Lubert Stryer, W. H. Freeman and Co., New York, 1988.
75. 'The Vitamins', Ch. 8 in 'Basic Nutrition', ed. E. W. McHenry and G. H. Beaton, Pitman Medical Publishing Co. Ltd., London, 1963, p. 196.
76. A. D. Hunt, J. Stokes, W. W. McCroby and H. H. Stroud, *Pediatrics*, 1954, **13**, 140 - 142.
77. W. W. Hawkins and J. Barsky, *Science*, 1948, **108**, 284 - 287.
78. G. A. Goldsmith, 'The B Vitamins: Thiamine, Riboflavin and Niacin', Ch. 2 in 'Nutrition: A Comprehensive Treatise', ed. G. H. Beaton and E. W. McHenry, Academic Press, New York and London, 1964, vol. 2, pp. 109 - 206.
79. M. Pyke, 'Deficiency Diseases', Ch. 7 in 'Man and Nutrition', World University Library, 1970, pp. 117 - 121.
80. B. F. Chow, 'The B Vitamins: B<sub>6</sub>, B<sub>12</sub>, Folic Acid, Panthoic Acid and Biotin', Ch. 3 in 'Nutrition: A Comprehensive Treatise', ed. G. H. Beaton and E. W. McHenry, Academic Press, New York and London, 1964, vol. 2, pp. 207 - 264.
81. E. E. Snell, *Phys. Rev.*, 1953, **33**, 509 - 524.

82. M. M. Wintrobe, R. H. Follis Jr., M. H. Miller and H. J. Stein, *Bull. John Hopkins Hosp.*, 1943, **72**, 1 - 9.
83. J. F. Rinehart and L. D. Greenberg, *Am. J. Path.*, 1949, **25**, 481 - 485.
84. S. Gerhoff, F. Faragalla, D. A. Nelson and S. B. Andrus, *Am. J. Med.*, 1959, **27**, 72 - 77.
85. S. R. Snodgrass, *J. Child Neurol.*, 1992, **7**, 77 - 86.
86. H. A. Ring, A. J. Heller, I. N. Farr, E. H. Reynolds, *J. Neurol. Neurosurgery Psych.*, 1990, **53**, 1051 - 1055.
87. P. R. M. Debittencourt, S. Mazer, T. Marcourakis, M. M. Bigarella, Z. S. Ferreira and J. P. Mumford, *Epilepsia*, 1994, **35**, 373 - 380.
88. A. Sabers and L. Gram, *Pharmacol. Toxicol.* 1992, **70**, 237 - 243.
89. E. Benmenachem, A. Hamberger and J. Mumford, *Epilepsy Res.*, 1993, **16**, 241 - 243.
90. W. Loscher and D. Horstermann, *Naunyn-Schmiedebergs Arch. Pharmacol.*, 1994, **349**, 270 - 278.
91. I. E. Leprik, *Epilepsia*, 1994, **35**, S29 - S40.
92. R. H. Mattson, O. Petroff, D. Rothman and K. Behar, *Epilepsia*, 1994, **35**, S29 - S32.
93. M. C. T. F. M. Dekrom, N. Verduin, E. Visser, M. Kleijer, F. Scholtes and J. H. M. Degroen, *Seizure*, 1995, **4**, 159 - 162.
94. R. Pacific, P. Zuccaro, P. Ianetti, U. Raucci and C. Imperato, *Epilepsia*, 1995, **36**, 423 - 426.
95. M. Qume and L. J. Fowler, *Brit. J. Pharmacol.*, 1997, **122**, 539 - 545.
96. H. Golan and Y. Grossman, *J. Neurophysiol.*, 1994, **71**, 48 - 58.
97. K. Sawaki, K. Ouchi, T. Sato and M. Kawaguchi, *Jpn. J. Pharmacol.*, 1991, **56**, 327 - 335.
98. J. Pietz, C. Benninger, H. Schäfer, D. Sontheimer, G. Mittermaier and D. Rating, *Epilepsia*, 1993, **34**, 757 - 763.
99. S. B. Coker, *Pediatrics*, 1992, **90**, 221 - 223.
100. K. Taylor and J. Goldstein, *Headache*, 1996, **36**, 514 - 515.
101. K. F. McKenna, D. J. McManus, G. B. Baker and R. T. Coutts, *J. Neurol. Transmission - Supplementum*, 1994, 115 - 122.
102. O. Heby, *Adv. Enz. Regul.*, 1985, **24**, 103 - 124.
103. C. W. Taylor and H. Tabor, *Annu. Rev. Biochem.*, 1976, **45**, 285 - 301.
104. E. Grassilli, M. Alfonsina Desiderio, E. Bellesia, P. Salomoni, F. Benatti and C. Franceschi, *Biochem. Biophys. Res. Commun.*, 1995, **216**, 708 - 714.
105. G. Tiao, Y. Noguchi, M. A. Lieberman, J. E. Fischer and P. O. Hasselgren, *Shock*, 1995, **4**, 403 - 410.
106. M. F. K. Baskaya, A. M. Rao, M. R. Prasad and R. J. Dempsey, *Neurosurgery*, 1996, **38**, 140 - 145.
107. C. M. Henley, C. Muszynski, L. Cherian and C. S. Robertson, *J. Neurotrauma*, 1996, **13**, 487 - 496.
108. D. H. Russell, in '*Polyamines in Normal and Neoplastic Growth*', Raven, New York, 1973.
109. A. E. Pegg, L. M. Shantz and C. S. Coleman, *J. Cell. Biochem.*, 1995, 132 - 138.

110. G. Chen, E. M. Perchellet, X. M. Gao, S. W. Newell, R. W. Hemingway, V. Bottari and J. P. Perchellet, *Anticancer Res.*, 1995, **15**, 1183 - 1189.
111. L. Megosh, S. K. Gilmour, D. Rosson, A. P. Soler, M. Blessing, J. A. Sawicki and T. G. O'Brien, *Cancer Res.*, 1995, **55**, 4205 - 4209.
112. A. Tamori, S. Nishiguchi, T. Kuroki, N. Koh, K. Kobayashi, Y. Yano and S. Otani, *Cancer Res.*, 1995, **55**, 3500 - 3503.
113. N. Matsubara, O. A. Hietala, S. K. Gilmour, K. Y. Yum, S. Litwin, P. Watts, E. J. Brennan and T. G. O'Brien, *Clin. Cancer Res.*, 1995, **1**, 665 - 671.
114. A. Manni, D. Mauger, P. Gimotty and B. Badger, *Clin. Cancer Res.*, 1996, **2**, 1901 - 1906.
115. F. L. Meyskens and E. W. Gerner, *J. Cell. Biochem.*, 1995, 126 - 131.
116. S. Merali and A. B. Clarkson, *Antimicrobial Agents and Chemotherapy*, 1996, **40**, 973 - 978.
117. Y. V. Bukin, V. A. Draudinkylenko, E. N. Orlov, Y. P. Kuvshinov, B. K. Poddubny, O. V. Vorobyeva and M. A. Shabanov, *Cancer Epidemiology Biomarkers and Prevention*, 1995, **4**, 865 - 870.
118. Y. Okai, K. Higashiokai, Y. Yano and S. Otani, *Cancer Lett.*, 1996, **99**, 15 - 21.
119. M. Wadatani, K. Ohtani, K. Kageyama, Y. Yano, T. Hasuma, S. Otani, S. Sakanaka, H. Kim and I. Matsuiyuasa, *Cancer J.*, 1996, **9**, 161 - 167.
120. J. Bomser, D. L. Madhavi, K. Singletary and M. A. L. Smith, *Planta Medica*, 1996, **62**, 212 - 216.
121. J. A. Moshier, M. Skunca, W. Wu, S. M. Boppana, F. J. Rauscher and J. Dodescu, *Nucleic Acids Res.*, 1996, **2**, 1149 - 1157.
122. 'International Encyclopaedia of Pharmacology and Therapeutics', ed. M. Schlachter, Pergamon, Oxford, 1973, Section 74.
123. M. J. Jung, *Bioorg. Chem.*, 1986, **14**, 429 - 443.
124. J. J. Skratt, L. B. Hough, J. W. Nalwalk and K. E. Barke, *Biochem. Pharmacol.*, 1994, **47**, 397 - 402.
125. Y. M. Lee, S. Hirota, T. Jippokanemoto, H. R. Kim, T. Y. Shin, Y. Yeom, K. K. Lee, Y. Kitamura, S. Nomura and H. M. Kim, *Int. Arch. Allergy Immunol.*, 1996, **110**, 272 - 277.
126. F. Hollande, A. Choquet, J. P. Ball and R. Magous, *Endocrinol.*, 1997, **138**, 955 - 962.
127. T. Kaneko, Y. Nagamachi and S. Matsuzaki, *Prostaglandins*, 1997, **53**, 37 - 47.
128. P. R. Buckland, R. Marshall, P. Watkins and P. McGuffin, *Mol. Brain Res.*, 1997, **49**, 266 - 270.
129. M. Bertoldi, P. S. Moore, B. Maras, P. Dominici and C. B. Voltattorni, *J. Biol. Chem.*, 1996, **271**, 23954 - 23959.
130. J. Drsata, J. Ulrichova and D. Walterova, *J. Enz. Inhib.*, 1996, **10**, 231 - 237.
131. J. T. Park, *Symp. Soc. Gen. Microbiol.*, 1958, **8**, 49 - 61.
132. S. E. Rognes, 'Threonine Biosynthesis', Ch. 19, in 'Methods in Plant Biochemistry', Academic Press Ltd., 1990, vol. 3, pp. 315 - 324.
133. A. J. Sinskey, International Patent Application, Int. App. Number: **PCT/US88/02029**, 1988.
134. H. J. Teas, N. H. Horowitz and M. Fling, *J. Biol. Chem.*, 1947, **172**, 651 - 658.
135. H. E. Umbarger and E. A. Adelberg, *J. Biol. Chem.*, 1951, **192**, 883 - 889.

136. G. Ehrensward, L. Reio, E. Saluste and R. Stjernholm, *J. Biol. Chem.*, 1951, **189**, 93 - 108.
137. P. H. Abelson, E. T. Bolton and E. Aldous, *J. Biol. Chem.*, 1952, **198**, 173 - 178.
138. K. McQuillen and R. B. Roberts, *J. Biol. Chem.*, 1954, **207**, 81 - 95.
139. P. H. Abelson, *J. Biol. Chem.*, 1954, **206**, 335 - 343.
140. M. Flavin and C. Slaughter, *J. Biol. Chem.*, 1960, **235**, 1103 - 1108.
141. J. Schnyder, M. Rottenberg and K. H. Erismann, *Biochem. Phys. Pflanzen*, 1975, **167**, 605 - 608.
142. I. Schildkraut and S. B. Greer, *J. Bacteriol.*, 1973, **115**, 777 - 785.
143. M. T. Skarstedt and S. B. Greer, *J. Biol. Chem.*, 1973, **248**, 1032 - 1044.
144. I. Schildkraut, PhD Thesis, University of Miami, 1974.
145. R. Miyajama, S. Otsuka and I. Shio, *J. Biochem.*, 1968, **6**, 139 - 148.
146. E. R. Stadtman, G. N. Cohen, G. LeBras and H. Robichon-Szulmajster, *J. Biol. Chem.*, 1961, **236**, 2033 - 2038.
147. H. Paulus and E. Gray, *J. Biol. Chem.*, 1964, **239**, PC4008 - 4009.
148. P. Datta and H. Gest, *Proc. Natl. Acad. Sci. USA*, 1964, **52**, 1004 - 1009.
149. K. Nakayama, K. Tanaka, H. Ogino and S. Kinoshita, *Agr. Biol. Chem.*, 1966, **30**, 611.
150. J. Giovanelli, S. H. Mudd and A. H. Datko, *Plant Phys.*, 1989, **90**, 1577 - 1583.
151. M. Szczesiul and D. E. Wampler, *Biochemistry*, 1976, **15**, 2236 - 2244.
152. J-C. Patte, G. LeBras, T. Loviny and G. N. Cohen, *Biochim. Biophys. Acta*, 1963, **67**, 16 - 30.
153. K. D. Gibson, A. Neuberger and G. H. Tait, *Biochem. J.*, 1962, **84**, 483 - 490.
154. T. Nara, H. Samejima, C. Fujita, M. Ito, K. Nakayama and S. Kinoshita, *Agr. Biol. Chem.*, 1961, **25**, 532.
155. R. Miyajima and I. Shio, *J. Biochem.*, 1970, **68**, 311 - 319.
156. E. H. Wormser and A. B. Pardee, *Arch. Biochem. Biophys.*, 1958, **78**, 416 - 432.
157. R. Miyajima and I. Shio, *J. Biochem.*, 1972, **71**, 219 - 226.
158. H. Yen and H. Gest, *Arch. Microbiol.*, 1974, **101**, 187 - 210.
159. B. Burr, J. Walker, P. Truffa-Bachi and G. N. Cohen, *Eur. J. Biochem.*, 1976, **62**, 519 - 526.
160. S. E. Rognes, R. M. Wallsgrove, J. S. H. Kueh and S. W. J. Bright, *Plant Sci.*, 1986, **43**, 45 - 50.
161. S. W. J. Bright, P. J. Shewry and B. J. Mifflin, *Planta*, 1978, **139**, 119 - 125.
162. J. Schnyder and M. Rottenberg, *Helv. Chim. Acta*, 1975, **58**, 518 - 521.
163. M. Malumbres, L. M. Mateos, C. Guerrero and J. F. Martin, *Nucleic Acids Res.*, 1988, **16**, 9859.
164. K. S. Han, J. A. C. Archer and A. J. Sinskey, *Mol. Microbiol.*, 1990, **4**, 1693 - 1702.
165. C. Parsot, P. Cossart, I. Saint-Girons and G. N. Cohen, *Nucleic Acids Res.*, 1983, **11**, 7331 - 7345.

166. C. Clepet, F. Borne, V. Krishnapillai, C. Baird, J. C. Patie and B. Cami, *Mol. Microbiol.*, 1992, **6**, 3109 - 3119.
167. K. Omon, Y. Imai, S-I. Suzuki and S. Komatsubara, *J. Bacteriol.*, 1993, **175**, 785 - 794.
168. S. F. Aas and S. E. Rognes, *Nucleic Acids Res.*, 1990, **18**, 665.
169. G. Mannhaupt, G. Van der Linden, I. Vetter, K. Maurer, U. Pilz, R. Planta and H. Feldmann, *Yeast*, 1990, **6**, 353 - 361.
170. M. Malumbres, L. M. Mateos, M. A. Lumbreras, C. Guerrero and J. F. Martin, *Appl. Environ. Microbiol.*, 1994, **60**, 2209 - 2219.
171. S. P. Lynn and J. F. Gardner, 'The Threonine Operon of *Escherichia coli*', 'Ch. 10 in *The Amino Acids: Biosynthesis and Genetic Regulation*', ed. K. Hermann and R. Somerville, Addison-Wesley Publishing Co., pp. 173 - 189.
172. H. Motoyama, K. Maki, H. Anazawa, S. Ishino and S. Teshiba, *Appl. Environ. Microbiol.*, 1994, **60**, 111 - 119.
173. B. Cami, C. Clepet and J-C. Patte, *Biochimie*, 1993, **75**, 487 - 495.
174. A. Szentirmai, M. Szentirmai and H. E. Umbarger, *J. Bacteriol.*, 1968, **95**, 1672 - 1679.
175. G. K. Farrington, A. Kumar, S. L. Shames, J. I. Ewaskewicz, D. E. Ash and F. C. Wedler, *Arch. Biochem. Biophys.*, 1993, **307**, 165 - 174.
176. P. R. Shewry, B. J. Mifflin, B. G. Forde and S. W. J. Bright, *Sci. Prog. Oxf.*, 1981, 575 - 600.
177. J. T. Madison and J. F. Thompson, *Biochem. Biophys. Res. Commun.*, 1976, **71**, 684 - 691.
178. J. Giovanelli, K. Veluthambi, G. A. Thompson, S. H. Mudd and A. H. Datko, *Plant Physiol.*, 1984, **76**, 285 - 292.
179. H. Aarnes, *Planta*, 1978, **140**, 185 - 192.
180. Å. Thoen, S. E. Rognes and H. Aarnes, *Plant Sci. Lett.*, 1978, **13**, 113 - 119.
181. J. M. Greenberg, J. F. Thompson and J. T. Madison, *Plant Cell Rep.*, 1988, **7**, 477 - 480.
182. R. M. Wallsgrove, P. J. Lea and B. J. Mifflin, *Plant Phys.*, 1983, **71**, 780 - 784.
183. J. Giovanelli, S. H. Mudd, A. H. Datko and G. A. Thompson, *Plant Phys.*, 1986, **81**, 577 - 583.
184. W. D. Rees, M. F. Fuller and H. J. Flint, *Biotech.*, 1990, **8**, 629 - 633.
185. W. D. Rees and S. M. Hay, *J. Bacteriol.*, 1993, **291**, 315 - 322.
186. J. H. Connick, G. C. Heywood, D. A. S. Smith and T. W. Stone, *Neurosci. Lett.*, 1986, **68**, 249 - 251.
187. T. W. Stone, J. H. Connick, G. C. Heywood and D. A. S. Smith, *Neurol. Neurobiol.*, 1987, **24**, 177 - 180.
188. S. A. Lipton, International Patent Application, Int. App. Number: **PCT WO92/0313**, 1992.
189. J. Thèze, D. Margarita, G. N. Cohen, F. Borne and J. C. Patte, *J. Bacteriol.*, 1974, **117**, 133 - 143.
190. I. Saint-Girons and D. Margarita, *Mol. Gen. Genet.*, 1978, **162**, 101 - 107.
191. J. Thèze and I. Saint-Girons, *J. Bacteriol.*, 1974, **118**, 990 - 998.
192. J. F. Gardner, O. H. Smith, W. W. Fredricks and M. A. McKinney, 1974, *J. Mol. Biol.*, **90**, 613 - 631.

193. J. F. Gardner and O. H. Smith, *J. Bacteriol.*, 1975, **124**, 161 - 166.
194. P. Truffa-Bachi, M. Veron and G. N. Cohen, *CRC Crit. Rev. Biochem.*, 1974, **2**, 379 - 415.
195. G. N. Cohen and A. Dautry-Varsat, in 'Multifunctional Proteins', ed. H. Bisswanger and E. Schminke-Ott, John Wiley, New York, 1980, pp. 49 - 121.
196. J. C. Patte, G. le Bras, T. Loviny and G. N. Cohen, *Biochim. Biophys. Acta*, 1963, 16 - 30.
197. E. R. Stadtman, G. N. Cohen, G. le Bras and H. de Robichon-Szulmajster, *J. Biol. Chem.*, 1961, **236**, 2033 - 2038.
198. J. Thèze, L. Kleidman and I. Saint-Girons, *J. Bacteriol.*, 1974, **118**, 577 - 581.
199. S. P. Lynn, J. F. Gardner and W. S. Reznikoff, *J. Bacteriol.*, 1982, **152**, 363 - 371.
200. M. Freundlich, *Biochem. Biophys. Res. Commun.*, 1963, **10**, 277 -282.
201. C. Yanofsky, *Nature*, 1981, **289**, 751 - 758.
202. S. P. Lynn, W. S. Burton, J. H. Donohue, R. M. Gould, R. I. Gumport and J. F. Gardner, *J. Mol. Biol.*, 1987, **194**, 59 - 69.
203. J. F. Gardner, *Proc. Natl. Acad. Sci. USA*, 1979, **76**, 1706 - 1710.
204. M. Malumbres, L. M. Mateos, C. Guerrero and J. F. Martin, *Nucleic Acids Res.*, 1988, **16**, 9859.
205. R. Katsumata, T. Mizukami, Y. Kikuchi and K. Kino, *5th Int. Symp. Genet. Ind. Microorgs.*, 1986, 217 - 226.
206. M. Ishida, E. Yoshino, R. Makihara, K. Sato, H. Enei and S. Nakamori, *Agric. Biol. Chem.*, 1989, **53**, 2269 - 2271.
207. B. J. Eikmanns, M. Metzger, D. Reinscheid, M. Kircher and H. Sahm, *Appl. Microbiol. Biotechnol.*, 1991, **34**, 617 - 622.
208. R. Katsumata, T. Mizukami, M. Yokoi, Y. Kikuchi and T. Oka, International Patent Application, Int. App. Number: **PCT 86 107595**, 1986.
209. A. J. Sinskey, M. T. Follettie, W. Liebl and O. P. Peoples, International Patent Application, Int. App. Number: **PCT WO88/09819**, 1988.
210. M. T. Follettie, H. K. Shin and A. J. Sinskey, *Mol. Microbiol.*, 1988, **2**, 63 - 75
211. C. Harde, K. H. Neff, E. Nordhoff, K-P. Gerbling, B. Laber and H-D. Pohlenz, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 273 - 278.
212. C. Fuganti, *J. Chem. Soc. Chem. Commun.*, 1979, 337 - 339.
213. M. Flavin and C. Slaughter, *J. Biol. Chem.*, 1960, **235**, 1112 - 1118.
214. C. T. Walsh, in 'Vitamin B<sub>6</sub>: Pyridoxal Phosphate, Chemical, Biochemical and Medical Aspects', ed. D. Dolphin, R. Poulson and O. Avramovic, Wiley, New York, 1986, p. 43.
215. C. Danzin and M. J. Jung, in 'Chemical and Biological Aspects of Vitamin B<sub>6</sub> Catalysis', ed. A. E. Evangelopolous, Liss, New York, 1984, parts A and B, p. 377.
216. R. Rando, *Biochemistry*, 1974, **13**, 3859 - 3863.
217. M. Johnston and C. Walsh, in 'Molecular Basis for Drug Action', ed. T. Singer and R. Ondarza, Elsevier, North Holland, 1981, pp. 167 - 184.
218. H. Gehring, R. Rando and P. Christen, *Biochemistry*, 1977, **16**, 4832 - 4836.

219. R. Rando, N. Pelyea and L. Cheng, *J. Biol. Chem.*, 1976, **251**, 3306 - 3312.
220. A. Cooper, S. Fitzpatrick, C. Kaufman and P. Dowd, *J. Am. Chem. Soc.*, 1982, **104**, 332.
221. J. P. Sconnel, D. L. Preuss, T. C. Denny, F. Weiss, T. Williams and A. Stempel, *J. Antibiot.*, 1971, **4**, 239 - 334.
222. Y. Morino and M. Okamoto, *Biochem. Biophys. Res. Commun.*, 1973, **50**, 1061 - 1067.
223. Y. Morino, A. M. Osman and M. Okamoto, *J. Biol. Chem.*, 1974, **249**, 6684 - 6692.
224. R. B. Silverman and R. Abeles, *Biochemistry*, 1976, **15**, 4718 - 4723.
225. B. Lippert, B. W. Metcalf, M. J. Jung and P. Casara, *Eur. J. Biochem.*, 1977, **74**, 441 - 445.
226. P. J. Schechter and Y. Tranier, in '*Enzyme Activated Irreversible Inhibitors*', ed. N. Seiler, M. H. Jung and J. Knoch-Weser, Elsevier, Amsterdam, 1978, p. 14.
227. L. J. Fowler and R. A. John, *Biochem. J.*, 1972, **130**, 569 - 573.
228. R. Rando and F. W. Bangertter, *J. Am. Chem. Soc.*, 1976, **98**, 6762 - 6764.
229. R. Miyajama, S. Otsuka and I. Shio, *J. Biochem.*, 1968, **63**, 139 - 148.
230. G. K. Farrington, A. Kumar, J. E. Ewaskiewicz, D. E. Ash, S. L. Shames and F. C. Wedler, *Arch. Biochem. Biophys.*, 1993, **307**, 165 - 174.
231. C. Harde, K. H. Neff, E. Nordhoff, K-P. Gerbling, B. Laber and H-D. Pohlenz, *Bioorg. Med. Chem. Lett.*, 1994, **4**, 273 - 278.
232. B. Laber, K-P. Gerbling, C. Harde, K-H. Neff, E. Nordhoff and H-D. Pohlenz, *Biochemistry*, 1994, **33**, 3413 - 3423.
233. W. Loeffler, W. Katzer, S. Kremer, M. Kugler, F. Petersen, G. Jung, C. Rapp and J. S-M. Tschen, *Naturwissenschaftliche Mikrobiologie*, 1990, **3**, 156 - 163.
234. C. Rapp, G. Jung, M. Kugler and W. Loeffler, *Leibigs Ann. Chem.*, 1988, 655 - 661.
235. B. K. Park, A. Hirota and H. Sakai, *Agric. Biol. Chem.*, 1976, **40**, 1905 - 1906.
236. S. L. Shames, D. E. Ash, F. C. Wedler and J. J. Villafranca, *J. Biol. Chem.*, 1984, **259**, 15331 - 15339.
237. A. Fersht, '*Active-site-directed and Enzyme-activated Irreversible Inhibitors: "Affinity Labels" and "Suicide Inhibitors"*', Ch. 9 in '*Enzyme Structure and Mechanism*', 2nd Ed., W. H. Freeman and Co., New York, 1985, pp. 248 - 262.
238. K. Itaya and M. Ui, *Clin. Chim. Acta*, 1966, **14**, 361 - 366.
239. P. Drueckes, A. Schinzel and D. Palm, *Anal. Biochem.*, 1995, **230**, 173 - 177.
240. P. S. Chen, T. Y. Toriban and H. Warmer, *Anal. Chem.*, 1956, **28**, 1756 - 1758.
241. P. A. Lanzetta, L. J. Alvarez, P. S. Reinach and O. A. Candia, *Anal. Biochem.*, 1979, **100**, 95 - 97.
242. A. Fersht, '*The Basic Equations of Enzyme Kinetics.*', Ch. 3 in '*Enzyme Structure and Mechanism*', 2nd Ed., W. H. Freeman and Co., New York, 1985, pp. 98 - 120.
243. AMS Biotechnology (UK) Ltd., Catalogue, 1996.
244. F. W. Studier, *Meth. Enzymol.*, 1990, **185**, 60 - 89.
245. F. W. Studier, B. A. Moffatt, *J. Biol. Chem.*, 1986, **189**, 113 - 130.



246. M. Chamberlin, J. McGrath, L. Waskell, *Nature*, 1970, **228**, 227 - 231.
247. M. Chamberlin, J. Ring, *J. Biol. Chem.*, 1973, **248**, 2235 - 2244.
248. M. Golomb, M. Chamberlin, *J. Biol. Chem.*, 1974, **249**, 2858 - 2862.
249. M. W. Holmes, T. Platt, M. Rosenberg, *Cell*, 1983, **32**, 1029 - 1032.
250. A. D. Carter, C. E. Morris, W. T. McAllister, *J. Virol.*, 1981, **37**, 636 - 642.
251. A. H. Rosenberg, B. N. Lade, D-S Chui, S-W Lin, J. J. Dunn, F. W. Studier, *Gene*, 1987, **56**, 125 - 135.
252. F. Bolivar, R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, H. W. Boyer, J. H. Crosa, S. Falkow, *Gene*, 1977, **2**, 95 - 113.
253. J. G. Sutcliffe, *Cold Spring Harbor Symp. Quant. Biol.*, 1979, **43**, 77 - 90.
254. QIAGEN™, *'The QIAexpressionist' Booklet*, 1992.
255. F. Barclay, D. Gani and E. Chrystal, *J. Chem. Soc. Chem. Commun.*, 1994, **7**, 815 - 816.
256. G. Ågren, *Acta Chem. Scand.*, 1962, **16**, 1607 - 1615.
257. Y. Watanabe and K. Shimura, *J. Biochem.*, 1965, **43**, 283 - 285.
258. T. E. Fickel and C. Gilvarg, *J. Org. Chem.*, 1973, **38**, 1421 - 1423.
259. K-L Yu and B. Fraser-Reid, *Tetrahedron Lett.*, 1988, **29**, 979 - 982.
260. S. K. Goda, N. P. Minton, N. P. Botting and D. Gani, *Biochemistry*, 1992, **31**, 10747 - 10756.
261. H. C. Birnboim and J. Doly, *Nucleic Acids Res.*, 1979, **7**, 1513 - 1516.
262. D. Ish-Horowicz and J. F. Burke, *Nucleic Acids Res.*, 1981, **9**, 2989 - 2998.
263. F. Gannon and R. Powell, *'Construction of Recombinant DNA Molecules'*, Ch. 7 in *'Essential Molecular Biology - A Practical Approach'*, ed. T. A. Brown, IRL Press, Oxford University Press, Oxford, vol. 1, 1994, pp. 143 - 160.
264. a) New England Biolabs Catalog, 1998; b) [http://www.neb.com/neb/frame\\_tech.html](http://www.neb.com/neb/frame_tech.html).
265. K. A. Eckert and T. A. Kunkel, *'The Fidelity of DNA Polymerases Used in the Polymerase Chain Reactions'*, Ch. 14 in *'PCR - A Practical Approach'*, ed. M. J. McPherson, P. Quirke and G. R. Taylor, Information Press Ltd., Oxford, vol. 1, 1994, pp. 225 - 244.
266. J. Sambrook, E. F. Fritsch and T. Maniatis, *'Molecular Cloning: A Laboratory Manual'*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989.
267. S. N. Cohen, A. C. Y. Chang and L. Hsu, *Proc. Natl. Acad. Sci. USA*, 1972, **69**, 2110 - 2114.
268. A. Ullmann, F. Jacob and J. Monod, *J. Mol. Biol.*, 1967, **24**, 339 - 343.
269. J. P. Horowitz, J. Chua, R. J. Curby, A. J. Tomson, M. A. Darooge, B. E. Fisher, J. Maurico and I. Kundt, *J. Med. Chem.*, 1964, **7**, 574 - 578.
270. P. Towner, *'Purification of DNA'*, Ch. 3 in *'Essential Molecular Biology - A Practical Approach'*, ed. T. A. Brown, IRL Press, Oxford University Press, Oxford, vol. 1, 1994, pp. 52 - 53.
271. J. W. Perich, R. M. Valerio and R. B. Johns, *Tetrahedron Lett.*, 1986, **27**, 1377 - 1380.
272. J. H. van der Neut, J. H. Uhlenbroek and P. E. Verkade, *Recl. Trav. Chim. Pays-Bas*, 1953, **72**, 365 - 376.

273. J. H. Uhlenbroek and P. E. Verkade, *Recl. Trav. Chim. Pays-Bas*, 1953, **72**, 395 - 410.
274. J. W. Perich and R. B. Johns, *J. Chem. Soc. Chem. Commun.*, 1988, 664 - 666.
275. J. W. Perich and R. B. Johns, *J. Org. Chem.*, 1988, **53**, 4103 - 4105.
276. A. Paquet and M. Johns, *Int. J. Peptide Protein Res.*, 1990, **36**, 97 - 103.
277. M. Jones, K. K. Rana, J. G. Ward and R. C. Young, *Tetrahedron Lett.*, 1989, **30**, 5353 - 5356.
278. J. G. Ward and R. C. Young, *Tetrahedron Lett.*, 1988, **29**, 6013 - 6016.
279. S. J. Hecker, M. L. Minich and K. Lackey, *J. Org. Chem.*, 1990, **55**, 4909 - 4911.
280. N. Mora, J. M. Lacombe and A. A. Pavia, *Tetrahedron Lett.*, 1993, **34**, 2461 - 2464.
281. J. W. Perich, P. F. Alewood and R. B. Johns, *Tetrahedron Lett.*, 1986, **27**, 1373 - 1376.
282. T. B. Johnson and J. K. Coward, *J. Org. Chem.*, 1987, **52**, 1771 - 1779.
283. T. Baasov, S. Sheffer-Dee-Noor, A. Kohen, A. Jakob and V. Belakhov, *Eur. J. Biochem.*, 1993, **217**, 991 - 999.
284. R. H. Boschan and J. P. Holder, *Chem. Abs.*, 1968, **68**, P77961a.
285. R. H. Boschan and J. P. Holder, U. S. Patent Office 3,341,630, 1967.
286. P. Hormozdiari, PhD Thesis, University of St Andrews, 1996.
287. E. Baer, *Biochem. Prep.*, 1951, **1**, 50.
288. A. V. Podolskii and M. A. Bulatov, *Chem. Abs.*, 1976, **85**, 32557h.
289. M. I. Kabachnik, L. S. Zakharov, I. Yu. Kudryavtsev, A. P. Kharchenko, G. P. Tataurov and L. I. Korobeinikova, *Izv. Akad. Nauk. SSSR, Ser. Khim.*, 1979, **4**, 896 - 900.
290. A. Svensson, T. Fex and J. Kihlberg, *Tetrahedron Lett.*, 1996, **37**, 7649 - 7652.
291. M. J. Shapiro, G. Kumaravel, R. C. Petter and R. Beveridge, *Tetrahedron Lett.*, 1996, **37**, 4671 - 4674.
292. P. Sieber, *Tetrahedron Lett.*, 1987, **28**, 6147 - 6150.
293. M. V. Anuradha and B. Ravindranath, *Tetrahedron*, 1995, **51**, 5671 - 5674.
294. H. C. Brown, '*Boranes in Organic Chemistry*', Cornell University Press, 1972.
295. H. C. Brown and B. C. Subba Rao, *J. Am. Chem. Soc.*, 1960, **82**, 681 - 686.
296. H. C. Brown, P. Heim and N. M. Yoon, *J. Am. Chem. Soc.*, 1976, **92**, 1637 - 1646.
297. N. M. Yoon and C. S. Pak, *J. Org. Chem.*, 1973, **38**, 2786 - 2792.
298. C. F. Lane, *Chem. Rev.*, 1976, **76**, 773 - 799.
299. A. Pelter, *Chem. and Ind.*, 1976, 888 - 896.
300. H. Kluender, F.-C. Huang, A. Fritzberg, H. Schnoes, C. J. Sih, P. Fawcett and E. P. Abraham, *J. Am. Chem. Soc.*, 1974, **96**, 4054 - 4055.
301. J. V. Bhaskar Kanth and M. Periasamy, *J. Org. Chem.*, 1991, **56**, 5964 - 5965.
302. A. S. Bhanu Prasad, J. V. Bhaskar Kanth and M. Periasamy, *Tetrahedron*, 1992, **48**, 4623 - 4628.

303. Y. Suseela and M. Periasamy, *Tetrahedron*, 1992, **48**, 371 - 376.
304. R. A. W. Johnstone and M. E. Rose, *Tetrahedron*, 1979, **35**, 2169 - 2173.
305. H. H. Freedman and R. A. Dubois, *Tetrahedron Lett.*, 1975, **38**, 3251 - 3254.
306. A. McKillop, J.-C. Fiaud and R. P. Hug, *Tetrahedron*, 1974, **30**, 1379 - 1382.
307. J. M. Miller, K. H. So and J. H. Clark, *Can. J. Chem.*, 1979, **57**, 1887 - 1889.
308. J. H. Clark and J. M. Miller, *J. Am. Chem. Soc.*, 1977, **99**, 498 - 502.
309. D. J. Hart, W-P Hong and L-Y Hsu, *J. Org. Chem.*, 1987, **52**, 4665 - 4673.
310. R. Davis and J. M. Muchowski, *Synthesis*, 1982, 987 - 988.
311. A. M. Felix, *J. Org. Chem.*, 1974, **39**, 1427 - 1429.
312. T. L. Ho and G. A. Olah, *Angew. Chem. Int. Ed. Engl.*, 1976, **15**, 774 - 775.
313. M. E. Jung and M. A. Lyster, *J. Am. Chem. Soc.*, 1977, **99**, 968 - 969.
314. G. A. Olah, S. C. Narang, B. G. Balaram Gupta and R. Malhotra, *J. Org. Chem.*, 1979, **44**, 1247 - 1251.
315. W. C. Still, M. Kahn and A. Mira, *J. Org. Chem.*, 1978, **43**, 2923 - 2925.
316. M. Dagert and S. D. Ehrlich, *Gene*, 1979, **6**, 23 - 25.
317. U. K. Laemmli, *Nature*, 1970, **227**, 680 - 685.

# APPENDICES



Appendices

<i>CS_B. subtilis</i>	18	KLNRLADENS	----	ADVYLK	LE	YMN	----	PG	-	SS	VK	DR	46
<i>CS_E. coli</i>	18	RLNRIGN	----	GRILAK	VE	SRN	----	PS	-	FS	VK	CR	43
<i>CS_Yeast</i>	60	KIRSLTKATG	----	VNIYAK	LE	LCN	----	PA	-	GS	AK	DR	88
<i>DSDHT_B. subtilis</i>	82	EVQHMKGKLE	----	AAYQPP	FP	GRWLLKCDHEL		PI	S	GS	IK	AR	119
<i>LSDHT_Yeast</i>	13	RQVFNNGKTN	----	SWFYVK	HE	ILQ	----	PG	-	GS	FK	SR	41
<i>TD_B. subtilis</i>	33	HLPKLSEQLG	----	IELHVK	TE	GVN	----	PT	-	GS	FK	DR	61
<i>TD_E. coli</i>	36	KMEKLSRRLD	----	NVILVK	RE	DRQ	----	PV	-	HS	FK	LR	64
<i>TD_Yeast</i>	83	QGVGLSSRLN	----	TNVIK	RE	DLL	----	PV	-	FS	FK	LR	111
<i>TrS_B. subtilis</i>	63	YADRVTEYLG	----	GAKIYLK	RE	DLN	----	HT	-	GS	HK	IN	92
<i>TrS_E. coli</i>	60	KCQNITAGTN	----	TTLYLK	RE	DLL	----	HG	-	GA	HK	TN	88
<i>TS_B. subtilis</i>	33	HLPKLSEQLG	----	IELHVK	TE	GVN	----	PT	-	GS	FK	DR	61
<i>TS_E. coli</i>	72	LEERVRAAFA	----	FPAVAN	VE	SDVGCLELPHHG		PT	-	LA	FK	DF	109
<i>TS_Yeast</i>	85	KRSYSTFRSDEVTPLVQNV		TDGK	EN	LHILELPHHG		PT	-	YA	FK	DV	126

<i>CS_B. subtilis</i>	47	I-GLAMIEAAE	----	KEGK	LK	AGNT	----	----	----	----	----	----	II	68
<i>CS_E. coli</i>	44	I-GANMIWDAE	----	KRGV	LK	PGVE	----	----	----	----	----	----	LV	65
<i>CS_Yeast</i>	89	V-ALNIIKTAE	----	ELGE	LVR	GEP	----	----	----	----	----	----	GWVF	112
<i>DSDHT_B. subtilis</i>	120	GGIYEVLKYAENLALQEGMLQETDDYRILQEERFTGFFSRYSTIA												163
<i>LSDHT_Yeast</i>	42	G-IGHLIRKSN	----	QQPL	SE	SGK	----	----	----	----	----	----	LAVF	65
<i>TD_B. subtilis</i>	62	G-MMVAVAKAK	----	EEGN	DI		----	----	----	----	----	----	MC	80
<i>TD_E. coli</i>	65	G-AYAMMAGLT	----	EEQK	AHG		----	----	----	----	----	----	VI	83
<i>TD_Yeast</i>	112	G-AYNMIAKLD	----	DSQR	NQG		----	----	----	----	----	----	VI	130
<i>TrS_B. subtilis</i>	93	N-ALGQALLAK	----	KMGK	TKI		----	----	----	----	----	----	IA	111
<i>TrS_E. coli</i>	89	Q-VLGQALLAK	----	RMGK	TEI		----	----	----	----	----	----	IA	107
<i>TS_B. subtilis</i>	62	G-MMVAVAKAK	----	EEGN	DI		----	----	----	----	----	----	MC	80
<i>TS_E. coli</i>	110	G-GRFMAQMLT	----	HIAG	DKP		----	----	----	----	----	----	VTIL	130
<i>TS_Yeast</i>	127	A-LQFVGNLFE	----	YFLQ	RTNANLPEGEK		----	----	----	----	----	----	QITVV	157

<i>CS_B. subtilis</i>	69	EPTS	GNT	GIGL	AMVAA	AKG-LKAILV	MP	DT	-	MS	MERR	-	NLLR	107																									
<i>CS_E. coli</i>	66	EPTS	GNT	GIALAY	VAAARG-YKLT	LTMP	PET	-	MS	IERR	-	KLLK	104																										
<i>CS_Yeast</i>	113	EGTS	GST	GISIA	VVCNALG-YRAHIS	LP	D	-	DT	SLEKL	LALLE		151																										
<i>DSDHT_B. subtilis</i>	164	VGST	GNL	GLS	IGI	GAA	LG-FRVT	VHMSA	-	DAKQ	WK	DLLR	202																										
<i>LSDHT_Yeast</i>	66	SSSG	GNA	GLAAA	TACR	SMA-LNCS	VV	PK	-	TT	KPR	-	MVKK	102																									
<i>TD_B. subtilis</i>	81	ASTG	-	NTSAAA	AYA	ARAN-MKC	IV	IPN	-	GKIA	FG	-	KLAQ	117																									
<i>TD_E. coli</i>	84	TASAG	NH	AQGV	A	FSSAR	LG-VKAL	IV	MP	TA-TAD	IK	VD	-	AVRG	123																								
<i>TD_Yeast</i>	131	ACSAG	NH	AQGV	A	FAAK	HLK-IPAT	IV	MP	VC-T	PS	IK	YQ	-	NVSR	170																							
<i>TrS_B. subtilis</i>	112	ETGAG	QH	GV	A	PATV	AAK	FG-FS	CT	V	FM	GE	-	ED	VAR	QSL	NVFR	151																					
<i>TrS_E. coli</i>	108	ETGA	QH	GV	A	SAL	A	SALLG-LK	CRI	Y	M	GA	-	KD	VER	QSP	NVFR	147																					
<i>TS_B. subtilis</i>	81	ASTG	-	NTSAAA	AYA	ARAN-MKC	IV	IPN	-	GKIA	FG	-	KLAQ	117																									
<i>TS_E. coli</i>	131	TATS	GDT	GAA	V	A	HAFY	GLP	NV	K	V	ILY	P	R	-	GK	IS	PL	QEK	LFC	171																		
<i>TS_Yeast</i>	158	GATS	G	D	T	G	SAA	I	Y	G	L	R	G	K	D	V	S	V	F	I	L	Y	P	T	G	R	I	S	P	I	Q	E	E	-	Q	M	T	T	199

<i>CS_B. subtilis</i>	108	A-	YG	AE	LV	-LTP	GA	EG	-	-	MKG	A	IK	KA	EE	LA	E	K	H	G	-	-	Y	F	V	P	141																	
<i>CS_E. coli</i>	105	A-	LG	AN	LV	-LTE	GA	KG	-	-	MKG	A	IQ	KA	EE	I	V	AS	NP	E	-	-	K	Y	L	L	140																	
<i>CS_Yeast</i>	152	S-	LG	AT	V	NK	V	K	P	AS	I	V	D	P	N	Q	Y	V	NA	AK	KA	C	N	E	L	K	KS	G	-	-	NG	I	R	190										
<i>DSDHT_B. subtilis</i>	203	Q-	KG	V	T	V	M	-EY	E	T	D	-	-	-	Y	SE	A	V	NE	G	R	R	Q	A	E	Q	D	P	-	FC	Y	F	I	D	236									
<i>LSDHT_Yeast</i>	103	I	Q	S	A	G	A	K	V	I	-	I	H	G	D	H	W	G	-	-	E	A	D	E	Y	L	R	H	K	L	M	A	Q	E	S	Q	-	-	H	G	S	K	138	
<i>TD_B. subtilis</i>	118	A	V	M	Y	G	A	E	I	I	-A	I	D	G	N	-	-	-	F	D	D	A	L	K	I	V	R	S	I	C	E	K	S	P	-	-	-	-	-	-	147			
<i>TD_E. coli</i>	124	F	-	-	G	E	V	L	-	L	H	G	A	N	-	-	-	F	D	E	A	K	A	K	A	I	E	L	S	Q	Q	G	-	-	F	-	-	-	-	151				
<i>TD_Yeast</i>	171	L	-	-	G	S	Q	V	-	-	L	Y	G	N	D	-	-	F	D	E	A	K	A	E	C	A	K	L	A	E	E	R	G	-	-	L	-	-	-	198				
<i>TrS_B. subtilis</i>	152	M	K	L	G	A	E	V	V	P	V	T	S	G	N	G	T	-	-	L	K	D	A	T	N	E	A	I	R	Y	W	V	Q	H	C	D	H	F	Y	M	I	191		
<i>TrS_E. coli</i>	148	M	R	L	M	G	A	E	V	I	P	V	H	S	G	S	A	T	-	-	L	K	D	A	C	N	E	A	L	R	D	W	S	G	S	Y	E	T	A	H	Y	M	L	187
<i>TS_B. subtilis</i>	118	A	V	M	Y	G	A	E	I	I	-A	I	D	G	N	-	-	-	F	D	D	A	L	K	I	V	R	S	I	C	E	K	S	P	-	-	-	-	-	-	-	147		
<i>TS_E. coli</i>	172	T	L	G	G	N	I	E	T	V	-A	I	D	G	D	-	-	-	F	D	A	C	Q	A	L	V	K	Q	A	F	D	D	E	E	-	-	-	L	K	V	A	205		
<i>TS_Yeast</i>	200	V	P	D	E	N	V	Q	T	L	S	-V	T	G	T	-	-	-	F	D	N	C	Q	D	I	V	K	A	I	F	G	D	K	E	-	-	-	F	N	S	K	233		

<i>CS_B. subtilis</i>	142	Q	Q	F	N	N	P	S	-	N	P	E	I	H	-	-	R	Q	T	T	G	K	E	I	V	E	Q	F	G	-	-	-	DD	-	QL	D	A	F	V	A	G	176				
<i>CS_E. coli</i>	141	Q	Q	F	S	N	P	A	-	N	P	E	I	H	-	-	E	K	T	T	G	P	E	I	W	E	D	T	-	-	-	DG	-	Q	V	D	V	F	I	A	G	174				
<i>CS_Yeast</i>	191	A	V	F	A	D	Q	F	E	N	E	A	N	W	K	V	H	Y	Q	T	T	G	P	E	I	A	H	Q	T	-	-	-	K	G	N	I	D	A	F	I	A	G	228			
<i>DSDHT_B. subtilis</i>	237	D	E	H	S	R	Q	L	F	L	G	Y	A	V	A	A	S	R	L	K	T	Q	L	D	C	M	N	I	K	P	-	-	-	S	L	E	T	P	L	F	V	L	P	276		
<i>LSDHT_Yeast</i>	139	T	L	Y	V	H	P	F	D	N	E	T	I	W	E	G	-	H	S	T	I	V	D	E	I	I	E	Q	L	K	E	N	D	I	S	L	P	R	V	K	A	L	V	C	S	181
<i>TD_B. subtilis</i>	148	I	A	L	V	N	S	V	-	N	P	Y	R	I	E	G	-	Q	R	T	A	A	F	E	V	C	E	Q	L	-	-	-	G	E	A	P	D	V	L	A	I	P	183			
<i>TD_E. coli</i>	152	-	T	W	V	P	P	F	D	H	P	M	V	I	A	G	-	Q	G	T	L	A	L	E	L	L	Q	D	-	-	-	A	H	L	D	R	V	F	V	P	186					
<i>TD_Yeast</i>	199	-	T	N	I	P	P	F	D	H	P	Y	V	I	A	G	-	Q	G	T	V	A	M	E	I	L	R	Q	V	R	-	-	-	T	A	N	K	I	G	A	V	F	V	P	236	
<i>TrS_B. subtilis</i>	192	G	S	V	G	P	H	P	Y	P	Q	V	V	R	E	F	Q	M	I	G	E	E	A	K	D	Q	L	K	R	-	-	I	E	G	T	M	P	D	K	V	A	C	233			
<i>TrS_E. coli</i>	188	G	T	A	A	G	P	H	P	Y	P	T	I	V	R	E	F	Q	M	I	G	E	E	T	K	A	Q	I	L	E	-	-	R	E	G	L	P	D	A	V	I	A	C	229		
<i>TS_B. subtilis</i>	148	I	A	L	V	N	S	V	-	N	P	Y	R	I	E	G	-	Q	R	T	A	A	F	E	V	C	E	Q	L	-	-	-	G	E	A	P	D	V	L	A	I	P	183			
<i>TS_E. coli</i>	206	L	G	L	N	S	A	N	S	I	N	I	S	R	L	L	A	Q	I	C	Y	F	E	A	V	A	Q	L	P	-	-	Q	E	T	R	N	Q	L	V	S	V	P	246			
<i>TS_Yeast</i>	234	H	N	V	G	A	V	N	S	I	N	W	A	R	I	L	A	Q	M	T	Y																									

Appendices

<i>CS_B. subtilis</i>	177	IG-TGGGTITGAGEVLK--EAYPSIKIYA-----	201
<i>CS_E. coli</i>	175	VG-TGGTTLTGVSRYIKG-TKGKTDLISVA-----	201
<i>CS_Yeast</i>	229	CG-TGGTITGVAKFLKERAKIPCHVVLAD-----	256
<i>DSDHT_B. subtilis</i>	277	CG-VGGGPGGVAFGLKLLYGDDVHVFFAE-----	304
<i>LSDHT_Yeast</i>	182	VG-GGGLFSGI IKGLDR-NHLAEKIPVVA-----	208
<i>TD_B. subtilis</i>	184	VG-NAGNITAYWKGFKKEYHE-KNGTGLPK-----	210
<i>TD_E. coli</i>	187	VG-GGGLAAGVAVLIKQLMPQIKVIAVEAE-----	223
<i>TD_Yeast</i>	237	VG-GGGLIAGIGAYLKRVPAPHIKIIGVETY-----	273
<i>TrS_B. subtilis</i>	234	VG-GGSNAMGMFQAFLN-ED-VELIGAEAA-----	263
<i>TrS_E. coli</i>	230	VG-GGSNAIGMFADFIN-ETNVGLIGVEPG-----	260
<i>TS_B. subtilis</i>	184	VG-NAGNITAYWKGFKKEYHE-KNGTGLPK-----	210
<i>TS_E. coli</i>	247	SGNFGDLTAGLLAKSLG-LPVKRFIAATN-----	274
<i>TS_Yeast</i>	276	SGNFGDILAGYFAKKMGLPIEKLAIAATNENDILDRFLKSGLYER	319
<i>CS_B. subtilis</i>	202	VE---PSDSP---V-LSGG-----KPGPHKIQQ-----	222
<i>CS_E. coli</i>	202	VE---PTDSP---V-IAQALAGEEI-KPGPHKIQQ-----	228
<i>CS_Yeast</i>	257	-----PQGSGFYNR-VNYGVMYDYVEKEGTRRRHQ-----	285
<i>DSDHT_B. subtilis</i>	305	-----PTHSPCMLLGLYSLGHEKISVQDIGLDNQ-----	334
<i>LSDHT_Yeast</i>	209	---VETAGCD---V-LNKSLKKG---SPVTLEKLT-----	233
<i>TD_B. subtilis</i>	211	MR---GFEAE---GAAAIVRNEV-IENPETIAT-----	236
<i>TD_E. coli</i>	224	LDAGHPVDLPRVGL-FAEGVAVKRI-GDETFRLCQEYLDIITV	265
<i>TD_Yeast</i>	274	LQRNQRTPLPVVGT-FADGTSVRMI-GEETFRAVQVVDEVVLV	315
<i>TrS_B. subtilis</i>	264	IDT--PLHAA--T-ISKGTGVVIH-GSLTYLIQDEFQ---QII	297
<i>TrS_E. coli</i>	261	IET--GEHGA--P-LKHGRVGIYF-GMKAPMMQTEDG---QIE	294
<i>TS_B. subtilis</i>	211	MR---GFEAE---GAAAIVRNEV-IENPETIAT-----	236
<i>TS_E. coli</i>	275	VN---DTVPR--F-LHDGQWSPKA-TQATLSNAM-----	301
<i>TS_Yeast</i>	320	SDKVAATLSPAMDI-LISSNFERLLWYLAREYLANGDD-----L	357
<i>CS_B. subtilis</i>	223	---IGAGF-VPDILNTEVYDEIFP-----VKNEEAF	249
<i>CS_E. coli</i>	229	---IGAGF-IPANLDDLKLVKVI-----ITNEEAI	255
<i>CS_Yeast</i>	286	-VDTIVEGIGLNRITHNFHMGEKFI-----ESIRVNDNQAI	321
<i>DSDHT_B. subtilis</i>	335	AADGLAVGR-PSGFVGLKLI EPLLSG-----CYTVEDNTLY	368
<i>LSDHT_Yeast</i>	234	---SVATSLASPYIASFAFESFNKYGC-----KSVVLSDDVL	268
<i>TD_B. subtilis</i>	237	---AIRIGNPASWDKAVKAAEESNGK-----IDEVTDDEIL	269
<i>TD_E. coli</i>	266	DSDAICAAM-KDLFEDVRAAEPSG-----ALALAGMKKY	299
<i>TD_Yeast</i>	316	NTDEICA AV-KDIFEDTRSIVEPSG-----ALSVAGMKKY	349
<i>TrS_B. subtilis</i>	298	EPYSISAGLDYPGIGPEHAYLHKSGRV-----TYDSITDEEAV	335
<i>TrS_E. coli</i>	295	ESYSISAGLDFPSVGPQHAYLNSTGRA-----DYVSIITDDEAL	332
<i>TS_B. subtilis</i>	237	---AIRIGNPASWDKAVKAAEESNGK-----IDEVTDDEIL	269
<i>TS_E. coli</i>	302	---DVSQPNWPRVEELFRRKIVQLKE-----LGYAAVDETT	336
<i>TS_Yeast</i>	358	KAGEIVNNWFQELKTNKGKQVDKSIEGASKDFTSERVSNEETS	401
<i>CS_B. subtilis</i>	250	EYARRAAREEG---ILGGISSGAAIYAALQVAKKL-----G	282
<i>CS_E. coli</i>	256	STARRLMEEEG---ILAGISSGAAVAALKLQEDL-----S	288
<i>CS_Yeast</i>	322	RMAKYLSVNDG---LFGSSTAINAVAIIQVAKTL-----P	354
<i>DSDHT_B. subtilis</i>	369	TLHLMLAVSED---KYLEPSALAGMFGPVQLFSTEEGR-----R	404
<i>LSDHT_Yeast</i>	269	ATCLR YADDYN---FIVEPACGASLHLCYHPEILED-----I	302
<i>TD_B. subtilis</i>	270	HAYQLIARVEG---VFAEPGSCASIAGVLKQVKSG-----E	302
<i>TD_E. coli</i>	300	I ALHNIRGER---LAHILSGANVNFHGLRYVSR-----C	331
<i>TD_Yeast</i>	350	I STVHPEIDHTK-NTYVPILSGANMNFDRLRFSR-----A	385
<i>TrS_B. subtilis</i>	336	DALKLLSEKEG---ILPAIESAHALAKAFKLAKE-----M	367
<i>TrS_E. coli</i>	333	EAFKTLCLHEG---IIPALESSHALAHALKMMREN-----P	365
<i>TS_B. subtilis</i>	270	HAYQLIARVEG---VFAEPGSCASIAGVLKQVKSG-----E	302
<i>TS_E. coli</i>	337	QQTMR ELKELG---YTSEPHAAVAYRALRDQLNPGEY-----L	372
<i>TS_Yeast</i>	402	ETIKKIYESSVNPKHYILDPHTAVGVCATERLIAKDNDKSIQYI	445
<i>CS_B. subtilis</i>	283	KGKK---VLAIIPS---NGERYLSTPLYQFD-----	307
<i>CS_E. coli</i>	289	FTNKN---IVVILPS---SGERYLSTALFADLFTEKELQQ---	322
<i>CS_Yeast</i>	355	HGSN---IVIIACD---SGSRHLSKFWKEAKEIDHDVSL EEV	390
<i>DSDHT_B. subtilis</i>	405	YAQKYKMEHAVHVWGT---GGS MVPKDEMAAYNRIGADLLKRN	446
<i>LSDHT_Yeast</i>	303	LEQKIYEDDII IAC---GSGCMTYEDLVKASSTLNVS-----	338
<i>TD_B. subtilis</i>	303	IPKGSK---VVAVLTG---NGLKDPNTAVDISEIKPVTLP TDED	340
<i>TD_E. coli</i>	332	ELGEQREALLAVTIPEE---KGSFLKFCQLLGGRSVTEFN YRFAD	373
<i>TD_Yeast</i>	386	VLGEGKEVFMVLVTL PDV---PGA FKKMOKIIHPRSVTEFSYRYNE	427
<i>TrS_B. subtilis</i>	368	DRGQL---ILVCLSG---RGDKDVNTLMNVLEEEVKRHV----	400
<i>TrS_E. coli</i>	366	DKEQL---LVVNL SG---RGDKDIFTVHDLKARGEI-----	396
<i>TS_B. subtilis</i>	303	IPKGSK---VVAVLTG---NGLKDPNTAVDISEIKPVTLP TDED	340
<i>TS_E. coli</i>	373	FLGTAHPAKFKESVEAI---LGETL DLPKELAE RADLPLLSHNLP	414
<i>TS_Yeast</i>	446	SLSTAHPAKFA DAVNNALSGFSNYSFEKDVLP EELKKLSTLKKK	489

## Appendices

<i>CS_B. subtilis</i>	0	-----	307
<i>CS_E. coli</i>	0	-----	322
<i>CS_Yeast</i>	391	INI-----	393
<i>DSDHT_B. subtilis</i>	447	GK-----	448
<i>LSDHT_Yeast</i>	0	-----	338
<i>TD_B. subtilis</i>	341	SILEYVKGAARV-----	352
<i>TD_E. coli</i>	374	AKN-----ACIFVGVRLSRGLEERKEILQMLNDGGYSVVDL	409
<i>TD_Yeast</i>	428	HRHESSEVPKAYIYTSFSVVDREKEIKQVMQQLNALGFEAVDI	471
<i>TrS_B. subtilis</i>	0	-----	400
<i>TrS_E. coli</i>	0	-----	396
<i>TS_B. subtilis</i>	341	SILEYVKGAARV-----	352
<i>TS_E. coli</i>	415	ADFAALRKLMMNHQ-----	428
<i>TS_Yeast</i>	490	LKFIERADVELVKNAIEEELAKMKL-----	514
<i>CS_B. subtilis</i>	0	-----	307
<i>CS_E. coli</i>	0	-----	322
<i>CS_Yeast</i>	0	-----	393
<i>DSDHT_B. subtilis</i>	0	-----	448
<i>LSDHT_Yeast</i>	0	-----	338
<i>TD_B. subtilis</i>	0	-----	352
<i>TD_E. coli</i>	410	SDDEMAKLVHRYMVGGRPSHPLQERLYSFEFPESPGALLRFLNT	453
<i>TD_Yeast</i>	472	SDNELAKSHGRYLVGG-ASKVPNERIISFEFPERPGALTRFLGG	514
<i>TrS_B. subtilis</i>	0	-----	400
<i>TrS_E. coli</i>	0	-----	396
<i>TS_B. subtilis</i>	0	-----	352
<i>TS_E. coli</i>	0	-----	428
<i>TS_Yeast</i>	0	-----	514
<i>CS_B. subtilis</i>	0	-----	307
<i>CS_E. coli</i>	0	-----	322
<i>CS_Yeast</i>	0	-----	393
<i>DSDHT_B. subtilis</i>	0	-----	448
<i>LSDHT_Yeast</i>	0	-----	338
<i>TD_B. subtilis</i>	0	-----	352
<i>TD_E. coli</i>	454	LGTYWNIISLFHYRSHGTDYGRVLAAPFELG-DHEPDFETRLNELG	496
<i>TD_Yeast</i>	515	LSDSWNLTLFHYRNHGADIGKVLAGISVPPRENLTFOKFLLEDLG	558
<i>TrS_B. subtilis</i>	0	-----	400
<i>TrS_E. coli</i>	0	-----	396
<i>TS_B. subtilis</i>	0	-----	352
<i>TS_E. coli</i>	0	-----	428
<i>TS_Yeast</i>	0	-----	514

### Abbreviations given are as follows:

cs - cysteine synthase (*O*-acetylserine sulfhydrylase); dsdht - D-serine dehydratase; lsdht - L-serine dehydratase; td - threonine dehydratase (threonine deaminase); trs - tryptophan synthase ( $\beta$ -subunit); ts - threonine synthase.



## Appendix 2

*Table of Reactions Catalysed by Threonine Synthase.*

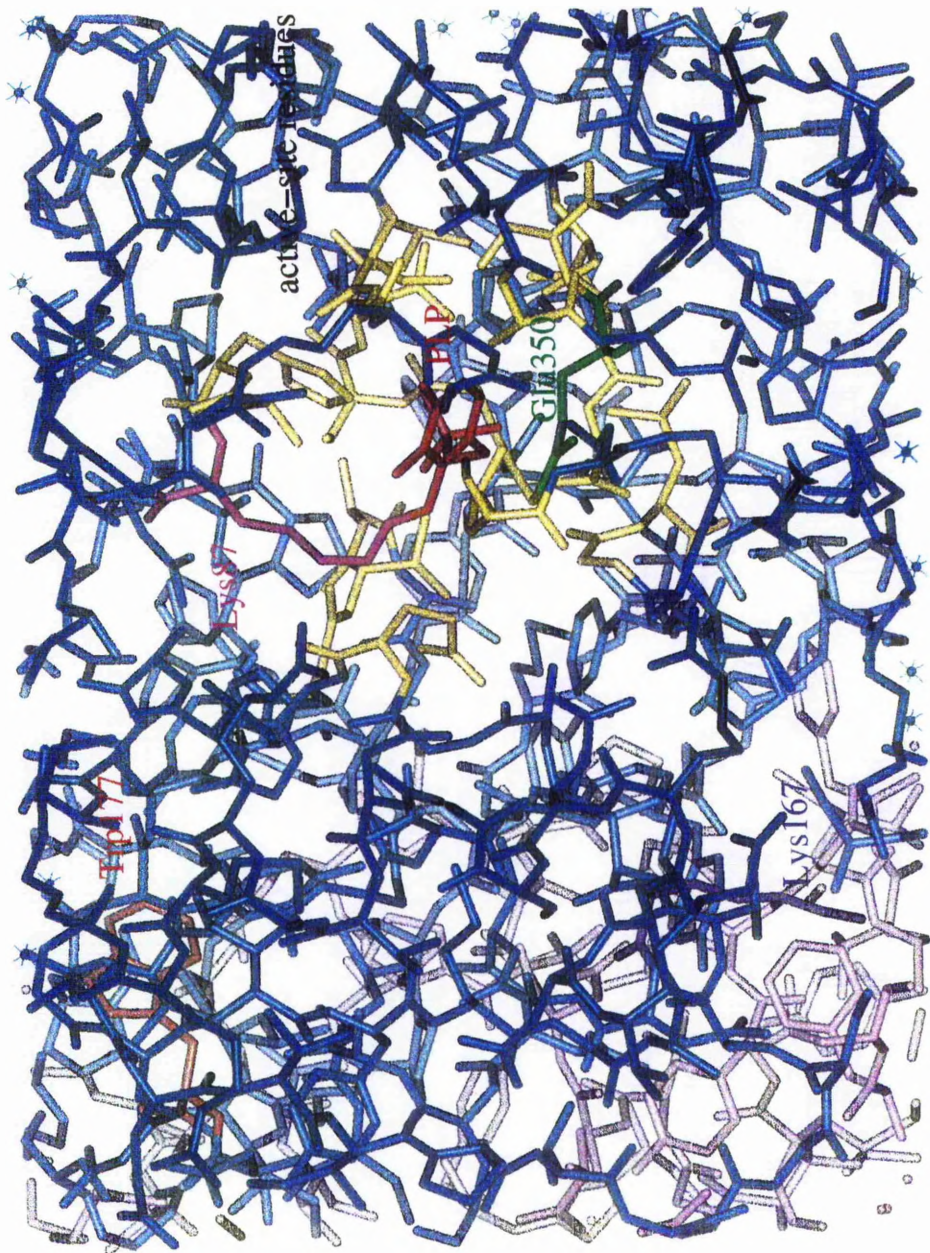
SUBSTRATE	TYPE OF REACTION	PRODUCT
(2 <i>S</i> )-homoserine	half-transamination	PMP (from PLP) <sup>a</sup>
(2 <i>S</i> )-serine	deamination ( $\beta$ -elimination)	pyruvate
(2 <i>S</i> )- <i>O</i> -phosphohomoserine	$\beta$ , $\gamma$ -replacement	(2 <i>S</i> ,3 <i>R</i> )-threonine
DL- <i>O</i> -methylserine	half-transamination	PMP (from PLP) <sup>a</sup>
(2 <i>S</i> ,3 <i>R</i> )-threonine	deamination ( $\beta$ -elimination)	$\alpha$ -ketobutyrate
(2 <i>S</i> )- <i>allo</i> -threonine	deamination ( $\beta$ -elimination)	$\alpha$ -ketobutyrate
DL-vinylglycine <sup>b</sup>	-	(2 <i>S</i> ,3 <i>R</i> )-threonine
DL-allylglycine	half-transamination	PMP (from PLP) <sup>a</sup>
(2 <i>S</i> )-alanine	half-transamination	PMP (from PLP) <sup>a</sup>
DL-3-chloroalanine	deamination ( $\beta$ -elimination)	pyruvate
DL-3,3,3-trifluoroalanine	half-transamination	PMP (from PLP) <sup>a</sup>
(2 <i>S</i> )-norvaline	half-transamination	PMP (from PLP) <sup>a</sup>
DL-3-hydroxynorvaline	deamination ( $\beta$ -elimination)	$\alpha$ -ketovalerate
(2 <i>S</i> )-2-aminobutanoic acid	half-transamination	PMP (from PLP) <sup>a</sup>
(2 <i>S</i> )- <i>threo</i> -2-amino-3-chlorobutanoic acid	deamination ( $\beta$ -elimination)	$\alpha$ -ketobutyrate
DL-2-amino-5-phosphonopentanoic acid	half-transamination	PMP (from PLP) <sup>a</sup>

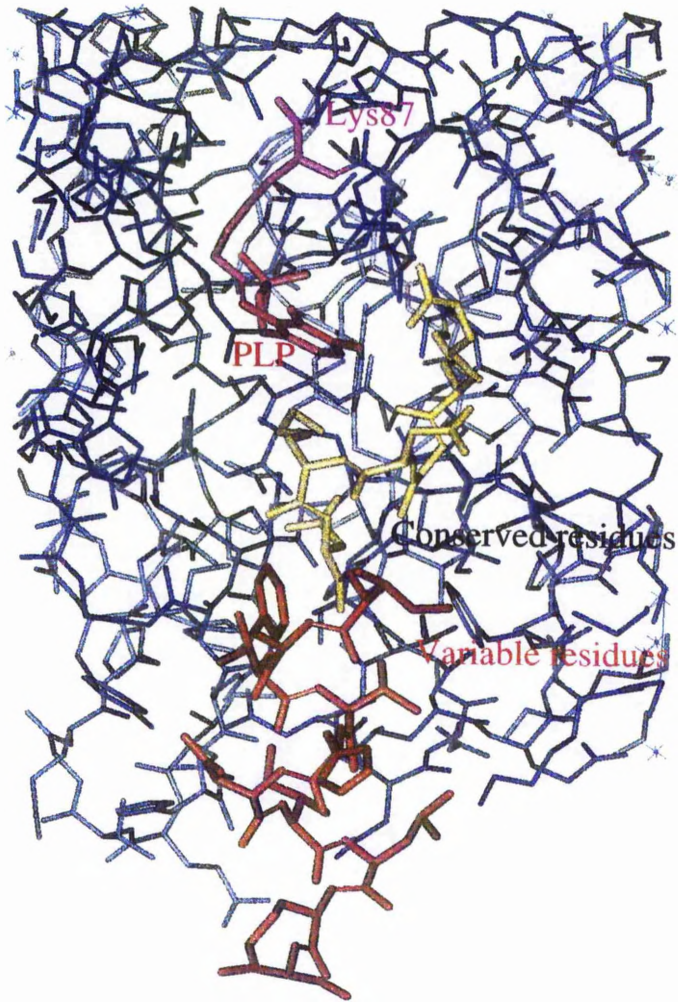
a - for the structure of PMP (pyridoxamine-5'-phosphate) see Scheme 1.2

b - (2*S*)-vinylglycine is an intermediate in the synthesis of (2*S*,3*R*)-threonine from (2*S*)-*O* -phosphohomoserine (see Schemes 1.16 - 1.18)

### Appendix 3

*Wild-type tryptophan synthase structures, highlighting residues involved in mutation experiments discussed in Section 1.3.*





Appendix 4

Table of Amino Acid Codes

Amino Acid	Code	Side-Chain Structure
Alanine	Ala, A	-CH <sub>3</sub>
Arginine	Arg, R	
Asparagine	Asn, N	
Aspartic acid	Asp, D	HO <sub>2</sub> C-
Cysteine	Cys, C	HSCH <sub>2</sub> -
Glutamic acid	Glu, E	HO <sub>2</sub> C-
Glutamine	Glu, Q	
Glycine	Gly, G	-H
Histidine	His, H	
Isoleucine	Ile, I	Me(Et)CH-
Leucine	Leu, L	(Me) <sub>2</sub> CHCH <sub>2</sub> -
Lysine	Lys, K	H <sub>3</sub> N <sup>+</sup> -
Methionine	Met, M	MeS-
Phenylalanine	Phe, F	PhCH <sub>2</sub> -
Proline	Pro, P	
Serine	Ser, S	HOCH <sub>2</sub> -
Threonine	Thr, T	
Tryptophan	Trp, W	
Tyrosine	Tyr, Y	HO-
Valine	Val, V	(Me) <sub>2</sub> CH-

Appendix 5

The Genetic Code

5'-OH Terminal Base	Middle Base				3'-OH Terminal Base
	U	C	A	G	
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	STOP <sup>1</sup>	STOP <sup>2</sup>	A
	Leu	Ser	STOP <sup>3</sup>	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met <sup>4</sup>	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

1 - Ochre; 2 - Opal; 3 - Amber; 4 - chain-initiating N-formyl-methionine

Appendix 6

Sequence alignments for threonine synthase from various microorganisms.

● A representation of the Clustal W alignment of TS enzymes from various microorganisms generated using SeqVu. Homologous residues are shaded and identical residues are boxed.

<i>S. cerevisiae</i>	1	MPNASQVYRSTRSSSPKTI	SFEEAIIQGLATDGGGLFIPPTIPQV	44			
<i>S. pombe</i>	1	-MSSQVSYLSTRGGSS-NFSFEEAVLKGLANDGGGLFIPSEIPQL		42			
<i>C. glutamicum</i>	1	-----MDYISTRDASRTPARFSDILLGGLAPDGGGLYLPATYPQL		39			
<i>M. glycogenes</i>	1	-----MKYISTRGQS-PALSFSEILLGGLAPDGGGLYLPPEQYPQF		38			
<i>E. coli</i>	1	-----MKLYNLKDHNE-QVSFAQAVTQGLGKNQGLFFPHDLPEF		38			
<i>P. aeruginosa</i>	1	-----MRYISTRGQAP-ALNFEDVLLAGLASDGGGLYVPENLPRF		38			
<i>S. marcescens</i>	1	-----MKLYNLKDHNE-QVSFAQAIKQGLGKQQGLFFPLDLPEF		38			
<i>H. influenzae</i>	1	-----MNLYNIKHEEE-QVTEFSQAVRQGLGRDQGLFFPEEVIPE		36			
<i>M. leprae</i>	1	-----MSG-QQTTTHQPWP	GVIAA-----YRDRLP	24			
<i>M. tuberculosis</i>	1	-----MTV-PTATHQPWP	GVIAA-----YRDRLP	24			
<i>B. subtilis</i>	1	-----MVKGLIHQ	-----YKEFLP	14			
<i>Br. lactofermentum</i>	1	-----MYKGLLKKQ	-----YASYLP	14			
<i>S. cerevisiae</i>	45	DQATLFDNDWSKLS	----FQDLAFAIMRLYIAQEEIPDADLKDLI	84			
<i>S. pombe</i>	43	PSG-WIEAWKDKS	----FPRIKFEVMSLYIPRSEISADELKKLV	81			
<i>C. glutamicum</i>	40	DDA-QLSKWREVLANEGYAALAREVISL	LVFV--DDIPVEDIKAIT	80			
<i>M. glycogenes</i>	39	SAD-ALSAMRGN	----YNDLAEFTILSRLI--DDIPADDLRIIV	75			
<i>E. coli</i>	39	SLT-EIDEMLRD	----FVTRSAKILSAFIG-DEIPQEI	76			
<i>P. aeruginosa</i>	39	TLE-ETASWVGLP	----YHELAFRVMRPFVAG-SIADADFKKIL	76			
<i>S. marcescens</i>	39	ELT-EIDHLLLEOD	----FVTRSSRILSAFIG-EEVPETALKRV	76			
<i>H. influenzae</i>	37	QLN-NINELLELP	----LVERSQKILGALID-GELPQATLDAMV	74			
<i>M. leprae</i>	25	----VGDDWT	PVT----LLEGGTPLIAAPR--LSEQTGCTIHL	57			
<i>M. tuberculosis</i>	25	----VGDDWT	PVT----LLEGGTPLIAATN--LSKQTGCTIHL	57			
<i>B. subtilis</i>	15	----VTDQTEALT	----LHEGNTPLIHLPK--LSEQLGIELHV	47			
<i>Br. lactofermentum</i>	15	----VNEKTPDVS	----LMEGNTPLIPLLN--ISKQLGVQLYG	47			
<i>S. cerevisiae</i>	85	KR-----SYSTFRSDEVT	PLVQNV	TGDKENLHILEL	FHGP	119	
<i>S. pombe</i>	82	DR-----SYSTRHRHPET	TPLKSLKNG	----LNVLEL	FHGP	112	
<i>C. glutamicum</i>	81	AR-----AYTIFKNS	EDIVPVT	LEDN	----IYLGHL	SEGP	113
<i>M. glycogenes</i>	76	DKTYRADVYAYAREGQDA	EDITPT	TYKLEDD	----LYLLSLS	SNGP	115
<i>E. coli</i>	77	RA-----AFAP	-APVANVESD	----VGCLEL	FHGP	102	
<i>P. aeruginosa</i>	77	ERTYG-----VFAHDA	SGAAPPV	ERTNG	----CVLEL	FHGP	108
<i>S. marcescens</i>	77	QA-----AFAP	-APVAKV	TTDD	----VSCLEL	FHGP	102
<i>H. influenzae</i>	75	KN-----AFTFP	-APLEK	VEEN	----IYALEL	FHGP	100
<i>M. leprae</i>	58	K-----			----VEGLN	-P	64
<i>M. tuberculosis</i>	58	K-----			----VEGLN	-P	64
<i>B. subtilis</i>	48	K-----			----TEGVN	-P	54
<i>Br. lactofermentum</i>	48	K-----			----YEGAN	-P	54
<i>S. cerevisiae</i>	120	TYAFKDV	ALQFVGNLF	FEYFLQRTNANL	PEGEKKQITV	VGATSGD	163
<i>S. pombe</i>	113	TFAFKDV	ALQFLGNLF	FEFLTRRKN	GNKPEDERDHL	TVVGATSGD	156
<i>C. glutamicum</i>	114	TAAFKDM	AMQLL	GFELRRKN	-----ETINIL	LGATSGD	149
<i>M. glycogenes</i>	116	TLAFKDM	AMQLL	GFVYVLAQKG	-----ETTNI	ILGATSGD	151
<i>E. coli</i>	103	TLAFKDF	GGRFMAQML	THIAGDKP	-----VTIL	TATSGD	136
<i>P. aeruginosa</i>	109	TLAFKDF	ALQLL	GRLLDHVLA	KRG-----ERVVI	MGATSGD	144
<i>S. marcescens</i>	103	TLAFKDF	GGRFMAQML	AEVAGEQP	-----VTIL	TATSGD	136
<i>H. influenzae</i>	101	TLAFKDF	GGRFMAQAL	AAVRGDGK	-----ITIL	TATSGD	134
<i>M. leprae</i>	65	TGSFKDRG	---MTMAV	TDALARGQ	-----RAVLC	ASTGN	95
<i>M. tuberculosis</i>	65	TGSFKDRG	---MTMAV	TDALAHGO	-----RAVLC	ASTGN	95
<i>B. subtilis</i>	55	TGSFKDRG	---MVM	AVAKAKEEGN	-----DTIM	CASTGN	85
<i>Br. lactofermentum</i>	55	TGSFKDRG	---MVM	AVAKAKEEGS	-----EAIIC	ASTGN	85



Appendices

<i>S. cerevisiae</i>	372	TNGK	FQVDKSIIEGAS	KDFTSERV	SNEETSETIKKIYES	SVNPK	415
<i>S. pombe</i>	376	RDG	TVTVRPEVLEAAR	RDFVSSERV	NDETIDAIIKKIYES	----	415
<i>C. glutamicum</i>	358	DDAN	-----	FEKAAMEYGF	ASGRSTHADR	VATIADVHSR	----L
<i>M. glycogenes</i>	360	DGG	-----	WPAKVA	DYGEVSGSSN	HANRMO	TIKATHER
<i>E. coli</i>	325	ELG	-----	YAAVDDE	-----	TTQOTMRELKEL	----
<i>P. aeruginosa</i>	341	ASGG	KLSVEDQRWTE	RKLFDSL	SLAVSDEQTCETIA	EVYRS	----S
<i>S. marcescens</i>	325	ELG	-----	HAAVSD	-----	TTKDTMRELAEL	----
<i>H. influenzae</i>	323	DLG	-----	SGMLSD	-----	ETEDTLKAMQSK	----
<i>M. leprae</i>	268	SKG	-----	RFAATDE	-----	EILAAYHLAKA	----E
<i>M. tuberculosis</i>	268	SKG	-----	RFLAASDE	-----	EILAAYHLAVR	----E
<i>B. subtilis</i>	256	SNG	-----	KIDVETD	-----	EILHAYQLIARV	----E
<i>Br. lactofermentum</i>	256	SHG	-----	EIDMVSDE	-----	EILHAYRLLAKT	----E
<i>S. cerevisiae</i>	416	HYILD	DPHTAVGV	CATERLIAK	DNDKSIQYIS	LSTAHPAKFADAV	459
<i>S. pombe</i>	416	HYILD	DPHTAVGV	ETGLRCLEK	TKDQDITYICLS	STAHPAKFDKAV	459
<i>C. glutamicum</i>	393	DVLID	DPHTADGV	HVARQRD	---EVNTP	PIIVLETALPVK	FADTI
<i>M. glycogenes</i>	394	GVTID	DPHTADGL	KVALEHR	---EAGT	PMLVLETALPA	KFEDA
<i>E. coli</i>	347	GYTSE	PHAAAVAY	RALRDQLN	---PGEY	GLFLGTAHPAK	FKESV
<i>P. aeruginosa</i>	381	GELL	DPHTAIGV	RAARECRR	---SLSV	PMTLGLTAHP	PKFPEAV
<i>S. marcescens</i>	347	GYISE	PHAAAIAY	RALRDQLQ	---EGE	FGLFLGTAHP	PKFESV
<i>H. influenzae</i>	345	GYLCE	PHGAIA	YQVVKDQK	---ASET	GIPLCTAHP	PKFESV
<i>M. leprae</i>	292	GVFVE	PASAA	SIAGLLK	AID	---GGW	VARGSTVV
<i>M. tuberculosis</i>	292	GVFVE	PASAA	SIAGLLK	AID	---GGW	VARGSTVV
<i>B. subtilis</i>	280	GVFAE	PGSCA	SIAGV	LKQVK	---SGE	IPKGSKVV
<i>Br. lactofermentum</i>	280	GVFAE	PGS	NASLAG	VIKHVE	---SGK	IKKGETVV
<i>S. cerevisiae</i>	460	NNAL	SGESNY	SFEK	DVLP	PEELK	KLSTLKKKLKFI
<i>S. pombe</i>	460	NLAL	SSYS	DYNF	NTOVLP	PIE	FDGLLDEERTC
<i>C. glutamicum</i>	434	VEAL	GEAP	QOT	-----	PER	FAAIMDAPFKV
<i>M. glycogenes</i>	434	VEAL	GHKPER	-----	-----	PHS	LEGLS
<i>E. coli</i>	387	EAIL	GETLDL	-----	-----	PKEL	AERADLPLLS
<i>P. aeruginosa</i>	422	EKAG	IGQAPA	-----	-----	LPAH	LADLFE
<i>S. marcescens</i>	387	EAIL	GQELPL	-----	-----	PKAL	ALRAELPLLS
<i>H. influenzae</i>	385	ERIL	GIQPL	-----	-----	PETL	DKHNQLPLLS
<i>M. leprae</i>	323	CTIT	GNG	LKD	-----	PDTAL	K-DM-PSV
<i>M. tuberculosis</i>	323	CTVT	GNG	LKD	-----	PDTAL	K-DM-PSV
<i>B. subtilis</i>	311	AVT	GNG	LKD	-----	PNTA	VDISE--IKP
<i>Br. lactofermentum</i>	311	AVLT	GNG	LKD	-----	PDI	AISSNQ--LD
<i>S. cerevisiae</i>	504	AIEE	ELAKM	KL	---	---	514
<i>S. pombe</i>	504	II	EVTLS	REMA	---	---	514
<i>C. glutamicum</i>	470	YIV	DAI	ANTS	VK	---	481
<i>M. glycogenes</i>	470	FIVE	RI	---	---	---	475
<i>E. coli</i>	423	LMM	RQ	---	---	---	428
<i>P. aeruginosa</i>	459	FV	SCH	GNP	GP	PL	470
<i>S. marcescens</i>	423	FL	MGL	PA	---	---	429
<i>H. influenzae</i>	421	YLL	KS	---	---	---	425
<i>M. leprae</i>	356	QL	GLV	---	---	---	360
<i>M. tuberculosis</i>	356	KE	GLA	---	---	---	360
<i>B. subtilis</i>	345	YVR	GAA	RV	---	---	352
<i>Br. lactofermentum</i>	345	HI	KGV	IMS	---	---	352



*Appendices*

Protein sequences for TS were obtained from the **SwissProt** database (web-site: <http://expasy.hcuge.ch/cgi-bin/sprot-search-de>) under the following accession numbers:

No.	Organism	Accession Number
1	<i>B. subtilis</i>	P04990
2	<i>B. lactofermentum</i>	P09123
3	<i>C. glutamicum</i>	P23669
4	<i>E. coli</i>	P00934
5	<i>H. influenzae</i>	P44503
6	<i>M. glycogenes</i>	P37145
7	<i>M. leprae</i>	P45837
8	<i>M. tuberculosis</i>	Q10610
9	<i>P. aeruginosa</i>	P29363
10	<i>S. pombe</i>	Q42598
11	<i>S. marcescens</i>	P27735
12	<i>S. cerevisiae</i> (yeast)	P16120

*Percentage identities between the various threonine synthase enzymes (organisms are numbered as in the table above).*

	1	2	3	4	5	6	7	8	9	10	11
2	68										
3	17	15									
4	17	12	28								
5	18	16	28	66							
6	10	11	57	29	28						
7	55	53	18	16	16	17					
8	55	55	18	14	13	16	93				
9	18	11	36	31	27	40	15	15			
10	17	13	33	28	30	37	13	16	37		
11	17	13	25	83	66	27	14	15	33	28	
12	13	12	35	31	30	33	13	12	35	55	30

Appendices

*Comparison between the threonine synthase enzymes quoted in reference 170.*

	<i>P. aeruginosa</i>	<i>B. subtilis</i>	<i>C. glutamicum</i>	<i>Br. lactofermentum</i>	<i>S. cerevisiae</i>
<i>E. coli</i>	34.9	24.2	30.2	21.9	37
	43	47.5	48.8	40.8	56
	261.8	137.7	229.5	131.4	252.9
<i>P. aeruginosa</i>		25.9	40.4	16.8	40.8
		48.7	56.8	30.2	59.7
		145.1	311.6	129.7	332.5
<i>B. subtilis</i>			22.5	67.8	24.5
			49.5	79.8	47
			139	397.4	143.3
<i>C. glutamicum</i>				17.4	39.4
				36.3	58.4
				134.3	324.1
<i>Br. lactofermentum</i>					16.1
					33.2
					143.5

First line, percentage of identity; second line, percentage of similarity; third line, quality scores.

Appendix 7.

Gene Sequence and Translation of *ThrC* from *E. coli*.

1/1  
 ATG AAA CTC TAC AAT CTG AAA GAT CAC AAC GAG CAG GTC AGC TTT GCG CAA GCC GTA ACC  
 M K L Y N L K D H N E Q V S F A Q A V T  
 61/21  
 CAG GGG TTG GGC AAA AAT CAG GGG CTG TTT TTT CCG CAC GAC CTG CCG GAA TTC AGC CTG  
 Q G L G K N Q G L F F P H D L P E F S L  
 121/41  
 ACT GAA ATT GAT GAG ATG CTG AAG CTG GAT TTT GTC ACC CGC AGT GCG AAG ATC CTC TCG  
 T E I D E M L K L D F V T R S A K I L S  
 181/61  
 GCG TTT ATT GGT GAT GAA ATC CCA CAG GAA ATC CTG GAA GAG CGC GTG CGC GCG GCG TTT  
 A F I G D E I P Q E I L E E R V R A A F  
 241/81  
 GCC TTC CCG GCT CCG GTC GCC AAT GTT GAA AGC GAT GTC GGT TGT CTG GAA TTG TTC CAC  
 A F P A P V A N V E S D V G C L E L F H  
 301/101  
 GGG CCA ACG CTG GCA TTT AAA GAT TTC GGC GGT CGC TTT ATG GCA CAA ATG CTG ACC CAT  
 G P T L A F K D F G G R F M A Q M L T H  
 361/121  
 ATT GCG GGT GAT AAG CCA GTG ACC ATT CTG ACC GCG ACC TCC GGT GAT ACC GGA GCG GCA  
 I A G D K P V T I L T A T S G D T G A A  
 421/141  
 GTG GCT CAT GCT TTC TAC GGT TTA CCG AAT GTG AAA GTG GTT ATC CTC TAT CCA CGA GGC  
 V A H A F Y G L P N V K V V I L Y P R G  
 481/161  
 AAA ATC AGT CCA CTG CAA GAA AAA CTG TTC TGT ACA TTG GGC GGC AAT ATC GAA ACT GTT  
 K I S P L Q E K L F C T L G G N I E T V  
 541/181  
 GCC ATC GAC GGC GAT TTC GAT GCC TGT CAG GCG CTG GTG AAG CAG GCG TTT GAT GAT GAA  
 A I D G D F D A C Q A L V K Q A F D D E  
 601/201  
 GAA CTG AAA GTG GCG CTA GGG TTA AAC TCG GCT AAC TCG ATT AAC ATC AGC CGT TTG CTG  
 E L K V A L G L N S A N S I N I S R L L  
 661/221  
 GCG CAG ATT TGC TAC TAC TTT GAA GCT GTT GCG CAG CTG CCG CAG GAG ACG CGC AAC CAG  
 A Q I C Y Y F E A V A Q L P Q E T R N Q  
 721/241  
 CTG GTT GTC TCG GTG CCA AGC GGA AAC TTC GGC GAT TTG ACG GCG GGT CTG CTG GCG AAG  
 L V V S V P S G N F G D L T A G L L A K  
 781/261  
 TCA CTC GGT CTG CCG GTG AAA CGT TTT ATT GCT GCG ACC AAC GTG AAC GAT ACC GTG CCA  
 S L G I P V K R F I A A T N V N D T V P  
 841/281  
 CGT TTC CTG CAC GAC GGT CAG TGG TCA CCC AAA GCG ACT CAG GCG ACG TTA TCC AAC GCG  
 R F L H D G Q W S P K A T Q A T L S N A  
 901/301  
 ATG GAC GTG AGT CAG CCG AAC AAC TGG CCG CGT GTG GAA GAG TTG TTC CGC CGC AAA ATC  
 M D V S Q P N N W P R V E E L F R R K I  
 961/321  
 TGG CAA CTG AAA GAG CTG GGT TAT GCA GCC GTG GAT GAT GAA ACC ACG CAA CAG ACA ATG  
 W Q L K E L G Y A A V D D E T T Q Q T M  
 1021/341  
 CGT GAG TTA AAA GAA CTG GGC TAC ACT TCG GAG CCG CAC GCT GCC GTA GCT TAT CGT GCG  
 R E L K E L G Y T S E P H A A V A Y R A  
 1081/361  
 CTG CGT GAT CAG TTG AAT CCA GGC GAA TAT GGC TTG TTC CTC GGC ACC GCG CAT CCG GCG  
 L R D Q L N P G E Y G L F L G T A H P A  
 1141/381  
 AAA TTT AAA GAG AGC GTG GAA GCG ATT CTC GGT GAA ACG TTG GAT CTG CCA AAA GAG CTG  
 K F K E S V E A I L G E T L D L P K E L

---

Appendices

---

1201/401

GCA GAA CGT GCT GAT TTA CCC TTG CTT TCA CAT AAT CTG CCC GCC GAT TTT GCT GCG TTG  
A E R A D L P L L S H N L P A D F A A L

1261/421

CGT AAA TTG ATG ATG AAT CAT CAG TAA  
R K L M M N H Q \*

1231/411



Appendices

ACGTTTATGCTGCGACCAACGTGAACGATACCGTGCCACGTTTCCTGCACGACGGTCAGTGGTCACCCAAAGCGACTC 880  
TGCAAATAACGACGCTGGTTGCACTTGCTATGGCACGGTGCAAAGGACGTGCTGCCAGTCACCAGTGGGTTTCGCTGAG  
BstE II  
863

AGGCGACGTTATCCAACGCGATGGACGTGAGTCAGCCGAACAAC TGCCCGGTGTTGGAAGAGTTGTTCCGCCGAAAATC 960  
TCCGCTGCAATAGGTTGCGCTACCTGCACTCAGTCGGCTTGTGACC GGCGCACACCTTCTCAACAAGCGCGCTTTTAG  
BstX I PshA I  
893 904

TGGCAACTGAAAGAGCTGGGTTATGCAGCCGTGGATGATGAAACCACGCAACAGACAATGCCGTGAGTTAAAAGAACTGGG 1040  
ACCGTTGACTTCTCGACCCAATACGTGGGCACCTACTACTTTGGTGCCTGCTGTACGCACTCAATTTCTTGACCC  
Bcl I  
1086

CTACACTTCGGAGCCGCACGCTGCCGTAGCTTATCGTGCCTGCGT GATCAGTTGAATCCAGGCGAATATGGCTTGTTC 1120  
GATGTGAAGCCTCGGCGTGCGACGGCATCGAATAGCACGCGACGC ACTAGTCAACTTAGGTCCGCTTATACCGAACAAGG  
Dra I Xmn I Psp1406 I  
1144 1159 1176

TCGGCACCGCGCATCCGGCGAAATTTAAAGAGAGCGTGGAAGCGATTCTCGGTGAAACGTTGGATCTGCCAAAAGAGCTG 1200  
AGCCGTGGCGCGTAGGCCGCTTAAATTTCTCTCGCACCTTCGCTAAGAGCCACTTTGCAACCTAGACGGTTTCTCGAC  
GCAGAACGTGCTGATTTACCCTTGCTTTCACATAATCTGCCCGCCGATTTTGCTGCGTTGCGTAAATGATGATGAATCA 1280  
CGTCTTGACAGACTAAATGGGAACGAAAGTGTAATTAGACGGCGGCTAAAACGACGCAACGCATTTAACTACTACTTAGT  
TCAGTAA 1287  
AGTCATT

## Appendix 9.

### *Accession numbers for vector DNA sequences.*

Full sequences for the pET-expression vectors can be obtained via e-mail on the Internet at the following address: novatech@novagen.com. The sequences for pBR322 and pKK223-2 are available from the EMBL database under the following accession numbers:

pBR322	VB0004 <sup>a</sup>
pKK223-2	M77749

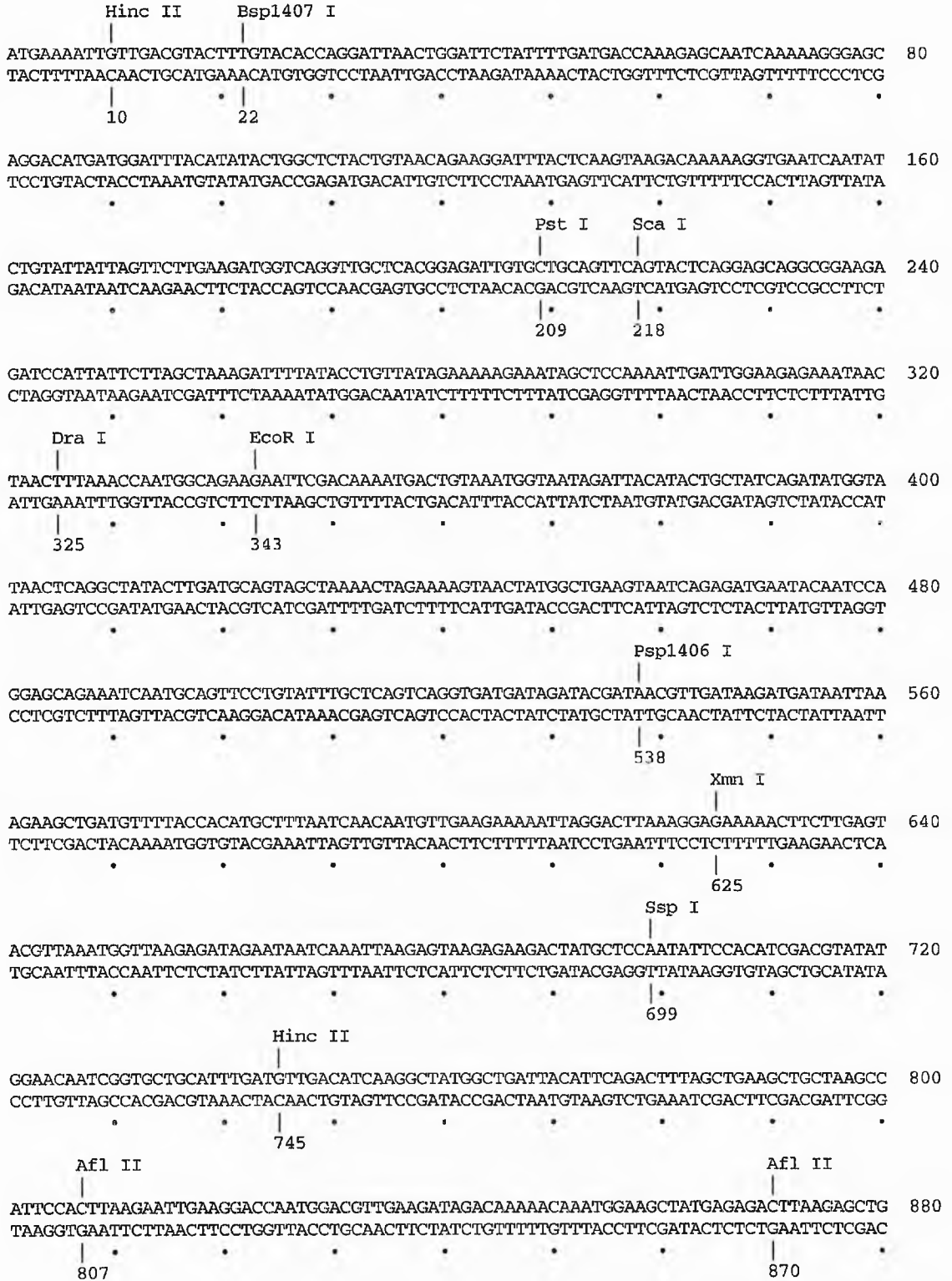
a - also in the following journal: *Gene*, 1982, **9**, 287 - 305.

Full sequences for the cloning vectors pGEM-3Zf(+), pGEM-5Zf(+) and pGEM-T are available from the EMBL database under the following accession numbers:

pGEM-3Zf(+)	X65304
pGEM-5Zf(+)	X65308
pGEM-T	X65308

Appendix 10.

Positions of Common Restriction Endonuclease Sites on the Gene for  $\beta$ -Methylaspartase.







## Appendix 11.

*The Sequence Data Obtained for the ThrC Recombinant Gene in pGEM-TTS<sub>I</sub> and pGEM-TTS<sub>II</sub>*

### 1. pGEM-TTS<sub>I</sub> x SP6:

Sense strand sequence from **N-terminus** of *thrC* (first codon underlined in red) to base **568**. Primer sequence is given in red at start of sequence. Mutations (underlined in the main sequence): **319-321** (deletion of A); **340** (A to G).

GTCAGTCAGT CATATGAAAC TCTACAATCT GAAAGATCAC AACGAGCAGG TCAGCTTTGC GCAAGCCGTA ACCCAGGGGT  
 TGGGCAAAAA TCAGGGGCTG TTTTTCCTCCG ACGACCTGCC GGAATTCAGC CTGACTGAAA TTGATGAGAT GCTGAAGCTG  
 GATTTTGTCA CCCGCAGTGC GAANATCCTC TCGGCGTTTA TTGGTGATGA AATCCCACAG GAAATCCTGG AAGAGCGCGT  
 GCGCGCGGCG TTTGCCTTCC CGGCTCCGGT CGCCAATGTT GAAAGCGATG TCGGTGTCTT GGAATTTGTT CACGGGCCAA  
 CGCTGGCATT TAAGATTTTCG GCGGTGCTTT TGTGGCACAA ATGCTGACCC ATATTGCGGG TGATAAGCCA GTGACCATTG  
 TGACCGCGAC CTCGGTGAT ACCGGAGCGG CAGTGGCTCA TGCTTTCTAC GGTTTACCGA ATGTGAAAGT GGTTATCCTC  
 TATCCACGAN GCAAAATCAG TCCACTGCAA GAAAACTGT TCTGTACATT GGGCGGCAAN ATCGAAACTG TTGCCATCNA  
 CGGCGATTTG NATGCCTGTC

### 2. pGEM-TTS<sub>I</sub> x T7:

Anti-sense strand sequence from **C-terminus** of *thrC* (stop codon shown in blue) to base **899**. Primer sequence is given in red at start of sequence.

ATGCGGATCC TTA CTGATGA NTCATCATCA ATTTACGCAA CGCAGCANAA TCGGCGGGCA GATTATGTGA AAGCAAGGGT  
 AAATCAGCAC GTTCTGCCAG CTCTTTTGGC AGATCCAACG TTTACCCGAG AATCGCTTCC ACGCTCTCTT TAAATTTGCG  
 CGGATGCNCG GTGCCGAGGA ACAAGCCATA TTTCGCCTGGA TTCAACTGAT CACGCANCGC ACNATAAGCT ACGGCAGCGT  
 GCGGCTCCGN AGTGTAGCCC AGTTCTTTTA NCTCAGCAT TGTCTGTTCG GTGGTTTCAT CATCCACGGC TGATAAACC

### 3. pGEM-TTS<sub>1</sub> x TS3:

Sense strand sequence from base **683** to base **1191**. Mutations (underlined): **716** (A to G); **944** (T to C).

AAGCTGTTGC GCAGCTGCCG CANGAGACCC GCAQCCAGCT GGTTGTCTCG GTGCCAAGCG GAAACTTCGG CGATTTGACG  
 GCGGGTCTGC TGGCGAAGTC ACTCGGTCTG CCGGTGAAAC GTTTTATTGC TCGGACCAAC GTGAACGATA CCGTGCCACG  
 TTTCTGACAC GACGGTCAGT GGTCAACCAA AGCGACTCAG GCGACGTTAT CCAACGCGAT GGACGTGAGT CAGCCGAACA  
 ACTGGCCGCG TGTGGAAGAG TCGTTCCGCC GCAAATCTG GCAACTGAAA GAGCTGGGTT ATGCAGCCGT GGATGATGAA  
 ACCACGCAAC AGACAATCGG TGAGTTAAAA GAACTGGGCT ACACTTCGGA GCCGCACGCT GCCGTAGCTT ATCGTGCGCT  
 GCGTGATCAN GTTGAATCCA GGCGAATATG GCTTGTTCCT CGGCACCGCG CATCCGGCGA AATTTTAAAGA NAGCGTGGA  
 GCGATTCTCG GTGAAACGTT GGATCTGCCA

### 4. pGEM-TTS<sub>1</sub> x TS4:

Anti-sense strand sequence from base **605** to base **25**. Mutations (underlined): **340** (T to C); **319-321** (deletion of T).

AGTTCTTCAT CATCAAACGC NTGCTTCACC ANGCCTTGA CAGGCATCGA AATCGCCGTC GATGGCAACA GTTTGATAT  
 TGCCGCCCAA TGTACAGAAC AGTTTTTCTT GCAGTGGACT GATTTTGCCT CGTGATAGA GGATAACCAC TTTCACATTC  
 GGTAAACCGT ANAAAGCATG AGCCACTGCC GCTCCGGTAT CACCGGAGGT CGCGGTCAGA ATGGTCACTG GCTTATCACC  
CGCAATATGG GTCAGCATTT GTGCCAAAA GCGACCGCCG AAATCTTAAA TGCCAGCGTT GGCCCGTGGA ACAATTCCAG  
 ACAACCGACA TCGCTTTCAA CATTTGGCGAC CGGAGCCGGG AAGGCAAACN CCGCGCGCAC GCGCTCTTCC AGGATTTTCT  
 GTGGGATTTT ATCACCATA AACGCCGAGA GGATCTTCGC ACTGCGGGTG ACAAAATCCA GCTTCAGCAT CTCATCAATT  
 TCAGTCAGGC TGAATTCGG CAGGTCGTGC GGAAAAACA GCCCTGATTT TTTGCCANC CCCTGGGTTA CGGCTTGGCC  
 AAAGCTGACC TGCTCGTGT

### 5. pGEM-TTS<sub>II</sub> x SP6:

Anti-sense strand sequence from **C-terminus** of *thrC* (stop codon is shown in blue) to base **786**. Primer sequence is shown in red. Mutations (underlined in main sequence): **944** (A to G), **785** (A to G, not confirmed in other sequencing experiments).

ATGCGGATCC TTTACTGATGA TTTCATCATCA ATTTTACGCAA CGCAGCAAAA TCGGCGGGCA GATTATGTGA AAGCAAGGGT  
 AAATCAGCAC GTTCTGCCAG CTCTTTTGGC AGATCCAACG TTTTACCAGAG AATCGCTTCC ACGCTCTCTT TAAATTTCCG  
 CGGATGCGCG GTGCCGAGGA ACAAGCCATA TTCGCCCTGGA TTCAACTGAT CACGCAGCGC ACGATAAGCT ACGGCAGCGT  
 GCGGCTCCGA AGTGTANCCC AGTTTCTTTTA ACTCACGCAT TGTTCTGTTC GTGGTTTCAT CATCCACGGC TGCATAACCC  
 ANCTCTTTCA GTTGCCAGAT TTTGCGGCGG AACACTCTT CCACACGCGG CCAGTTGTTC GGCTGACTCA CGTCCATCGC  
 GTTGGATAAC GTCNCCTGAG TCGCTTTGGG TGACCACTGA CCGTCGTGCA GGAAACGTGG CACGGTATCG TTTACGTTGG  
 NCGCAGCAAT AAAACGTTTC ACCGGCAGAC

### 6. pGEM-TTS<sub>II</sub> x T7:

Sense strand sequence from **N-terminus** of *thrC* (first codon is underlined in red) to base **348**. Primer sequence is shown in red. Mutation (underlined in main sequence): **319-321** (deletion of A).

GTCAGTCAGT CATATGAAAC TCTACAATCT GAAAGATCAC AACGAGCAGG TCAGCTTTTGC GCAAGCCGTA ACCCAGGGGT  
 TGGGCAAAAA TCAGGGGCTG TTTTFTCCGC ACGACCTGCC GGAATTCANC CTGACTGAAA TTGATGAGAT GCTGAAGCTG  
 GATTTTGTCA CCCGAGTGC GAAGATCCTC TCGGCGTTTA TTGGTGATGA AATCCCACAG GAAATCCTGG AAGAGCGCGT  
 GCGCGGGCG TTTGCCTTCC CGGCTCCGGT CGCCAATGTT GAAAGCGATG TCGGTGTCTT GGAANTGTTC CACGGGCCAA  
 CGCTGGCATT TAAGATTTG GCGGTCGCTT TATGGCAGAN

**7. pGEM-TTS<sub>II</sub> x TS3:**

Sense strand sequence from base **671** on *thrC* to base **1220** . Mutation (underlined): **944** (C to T). Note a mutation at **785** was not found in this sequencing experiment (*cf.* pGEM-TTS<sub>II</sub> x SP6).

GCTACTACTT TGAAGCTGTT GCNCAGCTGC CGCANGAGAC GCGCAACCAG CTGGTTGTCT CGGTGCCAAG CGGAAACTTC  
 GCGGATTTGA CGGCGGGTCT GCTGGCGAAG TCACTCGGTC TGCCGGTGAA ACGTTTTTATT GCTGCGACCA ACGTGAACGA  
 TACCGTGCCA CGTTTCCTGC ACGACGGTCA GTGGTCACCC AAAGCGACTC AGGCGACGTT ATCCAACCGG ATGGACGTGA  
 GTCAGCCGAA CAACTGGCCG CGTGTGGAAG AGTCGTTCGG CCGCAAAATC TGGCAACTGA AAGAGCTGGG TTATGCAGCC  
 GTGGATGATG AAACCACGCA ACAGACAATG CGTGAGTTAA AAGAACTGGG CTACACTTCG GAGCCGCACG CTGCCGTANC  
 TTATCGTGCG CTGCGTGATC AGTTGAATCC AGGCGAATAT GGCTTGTTCC TCGGCACCGC NCATCCGGCG AAATTTAAAG  
 ANAGCGTGGA AGCGATTCTC GGTGAAACGT TGGATCTGCC AAAAGANCTG GCANAACGTG CTGATTTTACC

**8. pGEM-TTS<sub>II</sub> x TS4:**

Anti-sense strand sequence from base **611** on *thrC* to base **252**. Mutation (underlined): **319-321** (deletion of T).

ACTTTCAGTT CTTCATCATC AAACGCCTGC TTCACCACNC GCCTGACAGG CATCGAAATC GCCGTCGATG GCAACAGTTT  
 CGATATTGCC GCCCAATGTA CAGAACAGTT TTTCTTGCAG TGGACTGATT TTGCCTCGTG GATAGAGGAT AACCACTTTC  
 ACATTCGGTA AACCGTAGAA AGCATGAGCC ACTGCCGCTC CGGTATCACC GGAGGTCGCG GTCAGAAATGG TCACTGGCIT  
 ATCACCCGCA ATATGGGTCA GCATTTGTGC CATAAAGCGA CCGCCGAAAT CTTAAATGCC AGCGTTGGCC CGTGGAACAA  
 TTCCAGACAA CCGACATCGC TTCAACATT GCGGACCGGA

## Appendix 12.

### *The Sequence Data Obtained for the ThrC Recombinant Gene in pTN-TS Constructs.*

The results for pTN-TS3, pTN-TS5 and pTN-TS6 should be compared with those for pGEM-TTS<sub>I</sub>; all others should be compared with those for pGEM-TTS<sub>II</sub>.

#### 1. pTN-TS1 x TS3:

Sense strand from base 712 on *thrC* to base 1287. Mutation (underlined): 944 (C to T).

```
CGCAACCAGC TGGTTGTCTC GGTGCCAAGC GGAAACTTCG GCGATTTGAC GGCGGGTCTG CTGGCGAACT CACTCGGTCT
GCCGGTGAAA CGTTTTATTC CTGCGACCAA CGTGAACGAT ACCGTGCCAC GTTTCCCTGCA CGACGGTCAG TGGTCACCCA
AAGCGACTCA NGCGACGTTA TCCAACGCGA TGGACGTGAG TCAGCCGAAC AACTGGCCGC GTGTGGAAGA GTCTTTCCGC
CGCAAAATCT GGCAACTGAA AGAGCTGGGT TATGCAGCCG TGGATGATGA AACCACGCAA CAGACAATGC GTGAGTTAAA
AGAACTGGGC TACACTTCGG AGCCGCACGC TGCCGTAGCT TATCGTGCGC TGCGTGATCA GTTGAATCCA GGCGAATATG
GCTTGTTCTT CGGCACCGCG CATCCGGCGA AATTTAAAGA GAGCGTGGAA GCGATTCTCG GTGAAACGTT GGATCTGCCA
AAAGAGCTGG CAGAACGTGC TGATTACCCT TGCTTTTACA TAATCTGCCC GCGATTTTCG CTGCGTTGCC GTAAATTGAT
GATGAATCAT CAGTNA
```

#### 2. pTN-TS1 x TS4:

Anti-sense strand sequence from base 588 of *thrC* to base 28. Mutation (underlined): 319-321 (deletion of T).

```
GCCTGCTTCA CCAGCGCCTG ACAGGCATCG AAATCGCCGT CGATGGCAAC AGTTTCGATA TTGCCGCCCA ATGTACAGAA
CAGTTTTTCT TGCACTGGAC TGATTTTGCC TCGTGGATAG AGGATAACCA CTTTCACATT CGGTAAACCG TAGAAAGCAT
GAGCCACTGC CGCTCCGGTA TCACCGGAGG TCGCGTTCAG AATGGTCACT GGCTTATCAC CCGCAATATG GGTCACTATT
TGTGCCATAA AGCGACCGCC GAAATCTTAA ATGCCAGCGT TGGCCCGTGG AACAAATCCA GACAACCGAC ATCGCTTTCA
ACATTTGGCGA CCGGAGCCGG GAAGGCAAAC GCCCGCGCA CGCGTCTTTC CAGGATTTCC TGTGGGATTT CATCACCAAT
AAACGCCGAG AGGATCTTCG CACTGCGGGT GACAAAATCC AGCTTCAGCA TCTCATCAAT TTCAGTCAGG CTGAATTCCG
GCAGGTCGTG CGGAAAAAAC AGCCCTTGAT TTTTGGCCAA CCCCTGGGTT ACGGCTTGCG CAAAGCTGAC CTGCTCGTT
```

### 3. pTN-TS1 x T7:

Failed to produce a readable sequence.

### 4. pTN-TS1 x T7T:

Anti-sense strand sequence from base **1284** of *thrC* to base **770**. Mutation (underlined): **944** (A to G).

```

CTGATGATTTC ATCATCAATT TACGCAACGC AGCAAAATCG GCGGGCAGAT TATGTGAAAG CAAGGGTAAA TCAGCACGTT
CTGCCAGCTC TTTTGGCAGA TCCAACGTTT CACCGAGAAT CGCTTCCACG CTCTCTTTAA ATTTCGCCGG ATGCGCGGTG
CCGAGGAACA AGCCATATTTC GCCTGGATTTC AACTGATCAC GCAGCGCACG ATAAGCTACG GCAGCGTTCG GCTCCGAAGT
GTAGCCCAGT TCTTTTAACT CACGCATTGT CTGTTGCGTG GTTTCATCAT CCACGGCTGC ATAACCCAGC TCTTTCAGTT
GCCAGATTTT GCGGCGGAAC GACTCTTCAC ACGCGGCCAG TTGTTCCGGCT GACTCACGTC CATCGCGTTG GATAACGTCG
CCTGAGTTCG TTTGGGTGAC CACTGACCGT CGTGCAGGAA ACGTGGCACG GTATCGTTCA CGTTGGTTCG AGCAATAAAA
CGTTTCACCG GCAGACOGAG TGACTTCGNC AGCA
    
```

### 5. pTN-TS2 x TS3:

Sense strand from base **717** on *thrC* to base **1110**. Mutation (underlined): **944** (T to C).

```

CGCACCAGCT GGTGTCTCG GTGCCAAGCG GAAACTTCGG CGATTGACG GCGGGTCTGC TGGCGAAGTC ACTCGGTCTG
CCGGIGAAAC GTTTTATTCG TCGACCAAC GTGAACGATA CCGTGCCACG TTCTCTGCAC GACGGTCAGT GGTACCCAA
AGCGACTNAG GCGACGTTAT CCAACGCGAT GGACGTGAGT CAGCCGAACA ACTGGCCGCG TGTGGAAAAG TGTTCCGCC
GCAAAATCTG GCAACTGAAA GAGCTGGGTT ATGCAGCCGT GGATGATGAA ACCACGCAAC AGACAATGCG TGAGTTAAAA
GAACTGGGCT ACACTINGGA GCCGNACGCT GCGTAGCTT ATCGNGCGCT GCGTGANAG TTGAATCCAG GCGAATAT
    
```

## 6. pTN-TS2 x TS4:

Anti-sense strand sequence from base **637** of *thrC* to base **30**. Mutation (underlined): **319-321** (deletion of T).

```

TTTTAACCTA GCGCCACTTT CAGTTCTTCA TCATCAAACG CCTGCTTCAC CAGCGCCTGA CAGGCATCGA AATCGCCGTC
GATGGCAACA GTTTTCGATAT TGCCGCCCAA TGTACAGAAC AGTTTTTCTT GCAGTGGACT GATTTTGCCT CGTGGATAGA
GGATAACCAC TTTTCACATTC GGTA AACCGT AGAAAGCATG AGCCACTGCC GCTCCGGTAT CACCGGAGGT CGCGGTCAGA
ATGGTCACTG GCTTATCACC CGCAATATGG GTCAGCATTT GTGCCATAAA GCGACCGCCG AAATCTTAAA TGCCAGCGTT
GGCCCGTGGA ACAATTCAG ACAACCGACA TCGCTTTCAA CATTGGCGAC CGGAGCCGGG AAGGCAAACG CCGCGCGCAC
GCGCTCTTCC AGGATTTCTT GTGGGATTTT ATCACCAATA AACGCCGAGA GGATCTTTCG ACTGCGGGTG ACAAATCCA
GCTTCAGCAT CTCATCAAIT TCAGTCAGGC TGAATTCGG CAGGTCGTGC GGAAAAACA GCCCCTGATT TTGCCCAACC
CCTGGGTTAC CGCTTGCACA AAGCTGCCTG CTCG
    
```

## 7. pTN-TS2 x T7:

Failed to produce a readable sequence.

## 8. pTN-TS2 x T7T:

Anti-sense strand sequence from base **1283** of *thrC* to base **1044**. Mutation (underlined): **1051** (C to T).

```

TGATGATTCA TCATCAATTT ACGCAACGCA GCAAAATCGG CGGGCAGATT ATGTGAAAGC AAGGGTAAAT CAGCACGTTT
TGCCAGCTCT TTTGGCAGAT CCAACGTTTC ACCGAGAATC GCTTCCACGC TCTCTTTAAA TTTCCGCGGA TGCGCGGTGC
CGAGGAACAA GCCATATTCG CCTGGATTNA ACTGATNACG CAGCGCACGA TAAGCTACGG CAGCGTCCGG NTTCGAAGTG
    
```



**9. pTN-TS3 x TS3:**

Sense strand sequence from base **686** on *thrC* to base **1045**. Mutations (underlined): **716** (A to G); **944** (T to C).

CTGTTGCGCA GCTGCCGCAT GAGACGCGCA GCCAGCTGGT TGTCTCGGTG CCAAGCGGAA ACTTCGGCGA TTGACGGCG  
GGTCTGCTGG CGAAGTCACT CGGTCTGCCG GTGAAACGTT TTAATTGCTGC GACCAACGTG AACGATACCG TGCCACGTTT  
CCTGCACGAC GGTCACTGGT CACCCAAAGC GACTNAGGCG ACGTTATCCA ACGCGATGGA CGTGAGTCAG CCGAACAACT  
GGCCGCGTGT GGAANAGTG TTCCGCCCGCA AAATCTGGCA ACTGAAAGAG CTGGGTTATG CAGCCGTGGA TGATGAAACC  
ACGCAACAGA CAATGCCGTGA GTTAAAAGAA CTGGGCTACA

**10. pTN-TS3 x TS4:**

Anti-sense strand sequence from base **589** of *thrC* to base **80**. Mutations (underlined): **319-321** (deletion of T); **340** (T to C)

ACGCCTGCTT CACCAGCGCC TGACAGGCAT CGAAATCGCC GTCGATGGCA ACAGTTTTGA TATTGCCGCC CAATGTACAG  
AACAGTTTTT CTGCACTGG ACTGATTTTG CCTCGTGGAT AGAGGATAAC CACTTTCACA TTCGGTAAAC CGTANAAAGC  
ATGAGCCACT GCCGCTCCG TATCACCGGA GGTCCGGTTC AGAATGGTCA CTGGCTTATC ACCCGCAATA TGGGTCAGCA  
TTTGTGCCAC AAAGCGACCG CCGAAATCTT AAATGCCAGC GTTGGCCCGT GGAACAATTT CAGACAACCG ACATCGCTTT  
CAACATTTGG GACCGGAGCC GGGAAAGCAA ACGCCGCGG CACGCGCTCT TCCAGGATTT CCTGTGGGAT TTCATNACCA  
ATAAACGCCG AGAGGATCTT CGCACTGCGG GTGACAAAAT CCAGCTTCAG CATCTCATCA ATTTCACTCA GGCTGAATTC  
CGGCAGGTG TCGGAAAAA ACAGCCCCTG

**11. pTN-TS3 x T7:**

Sense strand sequence from N-terminus of *thrC* (*Nde* I site shown in red, start codon underlined) to base **443**.

Mutations (underlined in the main sequence): **319-321** (deletion of A); **340** (A to G).

CATATGAAAC TCTACAATCT GAAAGATCAC AACGAGCAGG TCAGCTTTTC GCAAGCCGTA ACCCAGGGGT TGGGCAAAAA  
 TCAGGGGCTG TTTTTTCCGC ACGACCTGCC GGAATTNAGC CTGACTGAAA TTGATGAGAT GCTGAAGCTG GATTTTGTAC  
 CCGCAGTGC AAGATCCTCT CGGCGTTTAT TGGTGATGAA ATCCACAGG AAATNCTGNA AGAGCGCGTG CGCGGGCGT  
 TGCCTTCCCG GCTCCGGTGC CCAATGTGA AAGCGATGTC GGTGTCTGG AATTGTTCCA CGGGCCAACG CTGGCATTAA  
 GATTTCCGGC GTCGCTTTGT GGCACAAATG CTGACCCATA TTGCGGGTGA TAAGCCAGTG ACCATCTGA CCGCGACCTCC  
 GGTGATACCG GAGCGGCAGT GCCTCATGCT TTCTACGGTT TA

**12. pTN-TS3 x T7T:**

Anti-sense strand sequence from C-terminus of *thrC* (*Bam*H I site shown in red, gene stop codon in blue) to base

**813**. Mutation (underlined in the main sequence): **944** (A to G).

GGATCCTTAC TGATGATTCA TCATCAATTT ACGCAACGCA GCAAAATCGG CGGGCAGATT ATGTGAAAGC AAGGGTAAAT  
 CAGCACGTTT TGCCAGCTCT TTTGGCAGAT CCAACGTTTC ACCGAGAATC GCTTCCACGC TCTCTTTAAA TTTCCCGGA  
 TGCGCGGTGC CGAGGAACAA GCCATATTCC CCTGGATTCA ACTGATCAGC CAGCGCACGA TAAGCTACGG CAGCGTCCG  
 CTCCGAAGTG TAGCCAGTT CTTTTAACTC ACGCATTTGC TGTTCGGTGG TTTTCATCATC CACGGCTGCA TAACCCAGCT  
 CTTTCAGTTG CCAGATTTTG CGGCGGAACG ACTCTTCCAC ACGCGGCCAG TTGTTCCGCT GACTCACGTC CATCGCGTTG  
 GATAACGTCG CCTGAGTCCG TTTGGGTGAC CACTGACCGT CGTCCAGGAA ACGTGGCACG GTATCGTCA CGTTGGTCC

### 13. pTN-TS4 x TS3:

Sense strand sequence from base **687** on *thrC* to base **1239**. Mutation (underlined): **944** (T to C).

TGTTGCGCAG CTGCCGCAGG AGACGCGCAA CCAGCTGGTT GTCTCGGTGC CAAGCGGAAA CTTCGGCGAT TTGACGGCGG  
 GTCTGCTGGC GAAGTCACTC GGTCTGCCGG TGAAACGTTT TATTGCTGCG ACCAACGTGA ACGATACCGT GCCACGTTTC  
 CTGCACGACG GTCAGTGGTC ACCCAAAGCG ACTCAGGCGA CGTTATCCAA CGCGATGGAC GTGAGTCAGC CGAACAACTG  
 GCCCGGTGTG GAAGAGTCGT TCCGCCGCAA AATCTGGCAA CTGAAAGAGC TGGGTTATGC AGCCGTGGAT GATGAAACCA  
 CGCAACAGAC AATGCGTGAG TTA AAAAGAAC TGGGCTACAC TTCGGAGCCG GACGCTGCCG TAGCTTATCG TCGCTGCGT  
 GATCAGTTGA ATCCAGGCGA ATATGGCTTG TTCCTGGCA CCGCGCATCC GCGGAAATTT AAAGAGAGCG TGGAAAGCGAT  
 TCTGGGTGAA ACGTTGGATC TGNCAAAGA GCTGGCAGAA CGTGCTGATT TACCCTTGCT TTCACATAAT CTG

### 14. pTN-TS4 x TS4:

Anti-sense strand sequence from base **657** of *thrC* to base **85**. Mutation (underlined): **319-321** (deletion of T).

GCCTGCTTCA CCAGCGCCTG ACAGGCATCG AAATCGCCGT CGATGGCAAC AGTTTCGATA TTGCCGCCCA ATGTACAGAA  
 CAGTTTTTCT TGCAGTGGAC TGATTTTGCC TCGTGGATAG AGGATAACCA CTTTCACATT CGGTAAACCG TAGAAAGCAT  
 GAGCCACTGC CGCTCCGGTA TNACCGGAGG TCGCGGTCAG AATGGTCACT GGCTTAINAC CCGCAATATG GGTGAGCAAT  
 TGTGCCATAA AGCGACCGCC GAAATCTTAA ATGCCAGCGT TGGCCCCTGG AACAAATINCA GACAACCGAC ATCGCTTTCA  
 ACATTGGCGA CCGGAGCCGG GAAGGCAAAC GCCGNGCGCA CGCGCTCTTT CAGGANITTC TGNGGGAAIT CATCACCAAT  
 AAACGCCGAG AGGATCTTCG CACTGCGGGT GACAAAATNC AGCTTCAGCA TCTCATCAAT TTCAGTCAGG CTGAATTCCG  
 GCAGGTCGTG CGGAAAAAAC AG

**15. pTN-TS4 x T7:**

Sense strand sequence from bases **70** to **408** on *thrC*. Mutations (underlined): **88** (T to G); **106** (C to T); **282-284** (deletion of T); **395-396** (CG to TN). The sequence above (pTN-TS4 x TS4) does not confirm any of these alterations, and shows a mutation at **319-321** not found in this sequencing experiment.

```
GGCAAAAATC ANGGGCTGGT TTTTCCGCAC GACCTGTCGG AATTCAGCCT GACTGAAATT GATGAGATGC TGAAGCTGGA
TTTTGTACCC CGCAGTGCGA AGATCCTCTC GCGTTTTATT GGTGATGAAA TCCCACAGGA AATNCTGGAA GAGCGCGTGC
GGCGGGCGTT TGCCTTCCCG GCTCCGGTCC CCAATGTTGA AAGCGATGTC CGGTCTCTGG AATTGNCCCA CGGGCCAACC
CTGGCATTTA AGATTTCCGC GGTCCGCTTA TGGCACAAAT GCTGACCCAT ATTGCGGGTG ATAAGCCANT GACCATTCTG
ACCGTNACCT NCGGTGAT
```

**16. pTN-TS4 x T7T:**

Anti-sense strand sequence from base **1283** on *thrC* to base **786**. Mutation (underlined in the main sequence): **944** (A to G).

```
TGATGATTC A TCATCAATTT ACGCAACGCA GCAAAATCGG CGGGCAGATT ATGTGAAAGC AAGGGTAAAT NAGCACGTTT
TGCCAGCTCT TTTGGCAGAT CCAACGTTTN ACCGAGAAIN GCTTCCACGC TCTCTTTAAA TTTCCGCGGA TGCGCGGTGC
CGAGGAACAA GCCATATTCG CCTGGATTNA ACTGATNACG CAGCGCACGA TAAGCTACGG CAGCGTGCGG CTNCGAAGTG
TAGCCCAGTT CTTTAACTN ACGCATTGNC TGNITCGTGG NITCATNAIN CACGGCTGCA TAACCCAGCT CTTTNAGTTG
CCAGANPTTG CGGCGAACC ACTCTTCNAC ACGCGNCAG TIGNTCGGCT GACTNACGIN CATCGGTTG GATAACGTCG
CCTGAGTCGC TTTGGGTGAC CACTGACCGT CGTGCANGAA ACGTGGCACG GNATCGTTCA CGTTGGTCCG AGCAATAAAA
CGTTTNACCG GCAGACCGAG TGACTTCG
```

**17. pTN-TS5 x TS3:**

Sense strand from base **680** on *thrC* to base **1182**. Mutation (underlined): **944** (T to C).

CTGTTGCGCA GCTGCCGCGAG GAGACGCGCA ACCAGCTGGT TGTCTCGGTG CCAAGCGGAA ACTTCGGCGA TTTGACGGCG  
 GGTCTGCTGG CGAAGTCACT CGGTCTGCCG GTGAAACGTT TTATTGCTGC GACCAACGTG AACGATACCG TGCCACGTTT  
 CCTGCACGAC GGTCACTGGT CACCCAAAGC GACTCAGGCG ACGTTATCCA ACGCGATGGA CGTGAGTCAG CCGAACAACT  
 GGCCGCGTGT GGAAGAGTG TTCCGCGCGA AAATCTGGCA ACTGAAAGAG CTGGGTTATG CAGCCGTGGA TGATGAAACC  
 ACGCAACAGA CAATGCGTGA GTTAAAAGAA CTGGGCTACA CTTCCGGAGCC GNACGCTGCC GTAGCTTATC GTGCGCTGCC  
 TGATCAGTTG AATCCAGGCG AATATGGCTT GTTCCTCGGC ACCGCGCATC CCGCGAAATT TAAAGAGAGC GTGGAAGCGA  
 TTCINNGTGA AACGTTGGAC TTT

**18. pTN-TS5 x TS4:**

Anti-sense strand sequence from base **609** of *thrC* to base **100**. Mutation (underlined): **319-321** (deletion of T).

TTTCAGTTCT TCATCATCAA ACGCCTGCTT CACCAGCGCC TGACAGGCAT CGAAATCGCC GTCGATGGCA ACAGTTTCGA  
 TATTGCCGCC CAATGTACAG AACAGTTTTT CTTCAGTGG ACTGATTTIG CCTCGTGGAT AGAGGATAAC CACTTTCACA  
 TTCGGTAAAC CGTAGAAAGC ATGAGCCACT GCCGCTCCGG TATCACCGBA GGTCCGGTIC AGAATGGTCA CTGGCTTATC  
 ACCCGCAATA TGGGTCAGCA TTTGTGCCAT AAAGCGACCG NCGAAATCTT AAATGCCAGC GTTGGCCCGT GGAACAATTC  
 CAGACAACCG ACATCGCTTT CAACATIGGC GACCGGAGCC GGAAGGCAA ACGCCCGNGC GCACCGCTC TTCAGGATT  
 TCCTGTGGGA TTTTCATCACC AATAAACGCC GAGAGGATCT TCGCACTGCG GGTGACAAAA TNCAGCTTNA GCATCTCATN  
 AATTTCACTC AGGCTGAATT CCGGCAGGTC

**19. pTN-TS5 x T7:**

Sense strand sequence from **N-terminus** of *thrC* (*Nde* I site shown in red, start codon underlined) to base **440**.  
Mutations (underlined in the main sequence): **278** (T to A); **319-321** (deletion of A).

CATATGAAAC TCTACAATCT GAAAGATCAC AACGAGCAGG TCAGCTTTGC GCAAGCCGTA ACCCAGGGGT TGGGCAAAAA  
TCAGGGGGCTG NTTTTTCCGC ACGACCTGCC GGAATTCAGC CTGACTGAAA TTGATGAGAT GCTGAAGCTG GATTTTTGTA  
CCCGCAGTGC GAAGATCCTC TCGGCGNTTA TTGGTGATGA AATCCCACAG GAAATCCTGG AANAGCGCGT GCGCGCGCG  
TTTGCCTTCC CGGCTCCGGT CGCAATGTTG AAAGCGATG CGTTGTCTG GAATTTTCC ACGGGCCAAC GCTGGCATTT  
AGGANNNCNG CGGNCGCTTT ATGGCACAAA TGCTGACCCA TATTGCGGGT GATAAGCCAG TGACCATTCT GACCCGNACC  
TCGGTGATA ~~CCGAGGCGG~~ CAGTGGCTCA TGCTTTCTAC ~~GAG~~

**20. pTN-TS5 x T7T:**

Anti-sense strand sequence from base **1283** on *thrC* to base **1134**. No mutations noted.

TGATGATTCA TCATCAATTT ACGCAACGCA GCAAAATCGG CGGGCAGATT ATGTGAAAGC AAGGGTAAAT NAGCACGTTT  
TGNCAGCTCT TTTGGCAGAN NCAACGTTTIN ACCGAGAATN GTTTNCACGC TNINITTTAAA TTTCCCGGA

**21. pTN-TS6 x TS3:**

Sense strand sequence from base **716** on *thrC* to base **1056**. Mutation (underlined): **944** (T to C)

AACCAGCTGG TTGTCTCGGT GCCAAGCGGA AACTTCGGCG ATTTGACGGC GGGTCTGCTG GCGAAGTCAC TCGGTCTGCC  
GGTGAAACGT TTTATTGCTG CGACCAACGT GAACGATACC GTGCCACGTT TCCTGCACGA CGGTCAGTGG TCACCCAAAG  
CGACTCAGGC GACGTTATCC AACCGATGG ACGTGAGTCA GCGAACAAC TGGCCCGCTG TGGAAGAGT GTTCCGCCGC  
AAAATCTGGC AACTGAAAGA GCTGGGTTAT GCAGCCGTGG ATGATGAAAC CAGCAACAG ACAATGCGTG AGTTAAAAGA  
ACTGGGCTAC ACTTCGGAGC CG.

## 22. pTN-TS6 x TS4:

Anti-sense strand from base **583** to base **103**. Mutation (underlined): **321-319** (deletion of T).

GCTTCACCAG CGCCTGACAG GCATCGAAAT CGCCGTCGAT GGCAACAGTT TCGATAATTGC CGCCCAATGT ACAGAACAGT  
 TTTTCTTGCA GTGGACTGAT TTTGCCTCGT GGATAGAGGA TAACCACATT CACATTCCGT AAACCGTANA AAGCATGAGC  
 CACTGCCGCT CCGGTATCAC CGGAGGTCGC GGTCAGAATG GTCACATGGT TATCACCCGC AATATGGGTC AGCATTTGTG  
 CCATAAAGCG ACCGCCGAAA TCTTAAATGC CAGCGTTGGC CCGTGGAAACA ATTCCAGACA ACCGACATCG CTTTCAACAT  
 TGGCGACCGG AGCCGGGAAG GCAAACGCCG CGCGCACGCG CTCTTCCAGG ATTTCCCTGTG GGATTTTCATC ACCAATAAAC  
 GCGGAGAGGA TTTTGCCTACT GCGGGTGACA AAATCCAGCT TCAGCATCTC ATCAATTTCA GTCAGGCTGA ATTCCGGCAG

## 23. pTN-TS6 x T7:

Sense strand sequence from **N-terminus** of *thrC* (*Nde* I site shown in red, start codon underlined) to base **460**.

Mutations (underlined in the main sequence): **319-321** (deletion of A); **340** (A to G).

CATATGAAAC TCTACAATCT GAAAGATNAC AACGAGCAGG TCAGCTTTGC GCAAGCCGTA ACCCANGGGT TGGGCAAAAA  
 TNANGGGCTG NTTTTTCCGC ACGANCTGCC GGAATTCANC CTGACTGAAA TTGATGAGAT GCTGAAGCTG GATTTTGTCA  
 CCGCAGTGC GAAGATCCTC TCGGCGTTTA TTGGTGATGA AATCCCACAG GAAATNCTGG AAGAGCGCGT GCGCGCGGCG  
 TTTGNCITCC CGGCTCCGGT CGNCAATGTT GAAAGCGATG TCGGTTGTCT GGAATTTGTTT CACGGGCCAA CGCTGGCATT  
TAAGANTTCG GCGGNCGCTT TGTGGCACAA ATGCTGACCC ATATTCGGG TGATAAGCCA GTGACCATTG TGACCGCGAC  
 CTGGTGATA CCGGAGCGG AGTGGCTAT GCTTCTACN GNTTACGAA TGTGAAGTG G

**24. pTN-TS6 x T7T:**

Anti-sense strand sequence from **C-terminus** of *thrC* (*Bam*H I site shown in red, gene stop codon in blue) to base **770**. Mutation (underlined): **944** (A to G).

GGATCCTTAC TGATGATTCA TCATCAATTT ACGCAACGCA GCAAAATCGG CGGCAGATT ATGTGAAAGC AAGGGTAAAT  
 CAGCACGTTT TGCCAGCTCT TTTGGCAGAT CCAACGTTTC ACCGAGAATC GCTTCCACGC TCCTCTTTAAA TTTTCGCCGA  
 TGGCGGTGC CGAGGAACAA GCCATATTTC CCTGGATTCA ACTGATCAG CAGCGCACGA TAAGCTACGG CAGCGTGGG  
 CTCGGAAGTG TAGCCAGTT CTTTTAACTC ACGCATTTGC TGTTCGGTGG TTTTCATCATC CACGGCTGCA TAACCCAGCT  
 CTTTCAGTTG CCAGATTTTG CGGCGGAACG ACTCTTCCAC ACGCGGNCAG TTGTTCCGGT GACTCACGTC CATCGCGTTG  
 GATAACGTCG CCTGAGTCCG TTTGGGTGAC CACTGACCGT CGTGCAGGAA ACGTGGCACG GTATCGTTCA CGTTGGTCCG  
 AGCAATAAAA CGTTTCACCG GNAGACCGAG TGACTTCGCC AGCA

**25. pTN-TS7 x TS3:**

Sense strand sequence from base **712** on *thrC* to base **1287** (gene stop codon shown in blue). Mutation (underlined): **944** (T to C)

CGCAACCAGC TGGTTGTC TC GGTGCCAAGC GGAAACTTCG GCGATTTGAC GCGGGTCTG CTGGCGAAGT CACTCGGTCT  
 GCCGGTGAAA CGTTTTATTG CTGCGACCAA CGTGAACGAT ACCGTGCCAC GTTTCCTGCA CGACGGTCAG TGGTCACCCA  
 AAGCGACTNA GCGGACGTTA TCCAACCGCA TGGACGTGAG TCAGCCGAAC AACTGGCCGC GTGTGGAAGA GTGTTCCGC  
 CGCAAAATCT GGCAACTGAA AGAGCTGGGT TATGCAGCCG TGGATGATGA AACCACGCAA CAGACAATGC GTGAGTTAAA  
 AGAACTGGGC TACACTTCGG AGCCGNACGC TGCCGTAGCT TATCGTGGC TGGTGTATCA GTTGAATCCA GGCGAATATG  
 GCTTGNTCCT CGGCACCGG CATCCGGCGA AATTTAAAGA GAGCGTGGAA GCGATTCTCG GTGAAACGTT GGATCTGACA  
 AAAGAGCTGG CAGAACGTGC TGATTTACCC TTGCTTTTAC ATAATCTGCC GCCGANTTTG CTGCGTTGCG TAAATTTGATG  
 ATGAATCATC AGTAA



**26. pTN-TS7 x TS4:**

Anti-sense strand from base **602** to base **169**. Mutation (underlined): **321-319** (deletion of A).

TCTTCATCAT CAAACGCCTG CTTCAACCAGC GCCTGACAGG CATCGAAATC GCCGTCGATG GCAACAGTTT CGATATTGCC  
 GCCCAATGTA CAGAACAGTT TTTCTTGCAG TGGACTGATT TTGCCCTCGTG GATAGAGGAT AACCACTTTC ACATTCCGTA  
 AACCGTAGAA AGCATGAGCC ACTGCCGCTC CGGTATNACC GGAGGTCGCG GTCAGAATGG TCACTGGCTT ATNACCCGCA  
 ATATGGGTCA GCATTTGTGC CATAAAGCGA CCGCCGAAAT NTTAAATGCC AGCGTTGGCC CGTGGAACAA TTCCAGACAA  
 CCGACATNGC TTINAACATT GCGACCGGA GCCGGGAAGG CAAACGCCG CCGCACGCGC TCTTNCAGGA TTTCTGTGG  
 GATTTTATTA CCAATAAACG CCGAGAGGAT CTT

**27. pTN-TS7 x T7:**

Sense strand sequence from **N-terminus** of *thrC* (*Nde* I site shown in red, start codon underlined) to base **273**.

CATATGAAAC TCTACANTCT GAAAGATCAC AACGAGCAGG TCAGCTTTGC GCAAGCCGTA ACCCANGGGT NGGGCAAAAA  
 TCAGGGGCTG TTTTTCCGC ACGACCTGCC GGAAATTCANC CTGACTGAAA TTGATGAGAT GCTGAAGCTG GATTTTGTCA  
 CCCGCGNGC GAAGATCCTC TCGGCGTTTA TTGGTGTATGA AATCCACAG GAAATCCTGG AANAGCGCGT GCGCGCGCG  
 NTTGCCTTCC CGGCTCCGGT CGCCAATGTT GAAAGCGATG TCCGNT

**28. pTN-TS7 x T7T:**

Anti-sense strand sequence from **C-terminus** of *thrC* (gene stop codon in blue) to base **910**. Mutation (underlined): **944** (A to G)

TTACTGGATG ATTCATCATC AATTTACGCA ACGCAGCANA ATCGGCGGGC AGATTATGTG AAAGCAAGGG TAAATNAGCA  
 CGTTCGCCA GCTCTTTTGG CAGATCCAAC GTTINACCGA GAATCGCTTC CACGCTCTCT TTAATTTTCG CCGGATGCGC  
 GGTGCCGAGG AACAGCCAT ATTCGCCTGG ATTCAACTGA TCACGCAGCG CACGATAAGC TACGGCAGCG TGCGGCTCCG  
 AAGTGTAGCC CAGTCTTTT AACINACGCA TIGNCTGTG CGTGGTTTCA TNATCCACGG CTGCATAACC CAGCTCTTTC  
 AGTTGCCAGA TTTTGGGCG GAACGACTCT TNCACACGCG GNCAGTTGTT CCGCTGACT

**29. pTN-TS8 x TS3:**

Sense strand sequence from base 712 on *thrC* to base 1184 . Mutation (underlined): 944 (T to C).

GCAACCAGCT GGTGTGCTCG GTGCCAAGCG GAAACTTCGG CGATTTGACG GCGGGTCTGC TGGCGAAGTC ACTCGGTCTG  
 CCGGTGAAAC GTTTTATTGC TCGGACCAAC GTGAACGATA CCGTGCCACG TTTCTCTGCAC GACGGTCAGT GGTACCCAA  
 AGCGACTNAG GCGACGTTAT CCAACGCGAT GGACGTGAGT CAGCCGAACA ACTGGCCGCG TGTGGAAGAG TCGTCCGCC  
 GCAAAATCTG GCAACTGAAA GAGCTGGGTT ATGCAGCCGT GGATGATGAA ACCACGCAAC AGACAATGCG TGAGTTAAAA  
 GAACTGGGCT ACACPTCGGA GCCGNACGCT GCCGTAGCTT ATCGTGCCT CCGTGATCAG TTGAATNCAG GCGAATATGG  
 CTTGNTCCTC GGCACCGCGC ATNCGCGAA ATTTAAAGAG AGCGTGAAG CGAATCTINGG TGAAACGTTG GA

**30. pTN-TS8 x TS4:**

Anti-sense strand from base 587 to base 88. Mutation (underlined): 321-319 (deletion of T).

GCCTGCTTCA CCAGCGCCTG ACAGGCATCG AAATCGCCGT CGATGGCAAC AGTTTCGATA TTGCCGCCCA ATGTACAGAA  
 CAGTTTTTCT TGCAGTGGAC TGATTTTGCC TCGTGGATAG AGGATAACCA CTITCACATF CCGTAAACCG TAGAAAGCAT  
 GAGCCACTGC CGCTCCGTA TNACCGGAGG TCGCGGTCAG AATGGTCACT GGC~~TTT~~ATTAC CCGCAATATG GGTGAGCATT  
 TGTGCCATAA AGCGACCGCC GAAATC~~TT~~TAA ATGCCAGCGT TGGCCCGTGG AACAA~~TT~~INCA GACAACCGAC ATNGCTTTNA  
 ACATTTGGCGA CCGGAGCCCG GAAGGCAAAC GCCGNCGCA CGCGCTCTTC AGGAATTCCT GTGGGATTTIC ATTACCAATA  
 AACGCCGAGA GGATCTTTNGC ACTGCGGGTG ACAAATCCA GCTINAGCAT CTNATNAATT TCAGTCAGGC TGAATTCGG  
 CAGGTCGTGC GGAAAAA

**31. pTN-TS8 x T7:**

Failed to produce a readable sequence.

**32. pTN-TS8 x T7T:**

Anti-sense strand sequence from **C-terminus** of *thrC* (gene stop codon in blue) to base **694**. Mutation (underlined): **944** (A to G)

TTACTGATGA TTCATCATCA ATTTACGCAA CGCAGCAAAA TCGGCGGGCA GATTATGTGA AAGCAAGGGT AAATCAGCAC  
 GTTCTGCCAG CTCTTTTGGC AGATCCAACG TTTCACCGAG AATCGCTTCC ACGCTCTCTT TAAATTTCCG CGGATGCGCG  
 GTGCCGAGGA ACAAGCCATA TTCGCCCTGGA TTCAACTGAT CACGCAGCGC ACGATAAGCT ACGGCAGCGT GCGGCTCCGA  
 AGTGTAGCCC AGTTCTTTTA ACTCACGCAT TGTCTGTTCG GTGGTTTCAT CATCCACGGC TGCATAACCC AGCTCTTTCA  
 GTTGCCAGAT TTTGCGGCGG AACACTCTT NCACACGCGG NCAGTTGTTC GGCTGACTCA CGTCCATCGC GTTGGATAAC  
 GTCGCCCTGAG TCGCTTTGGG TGACCACTGA CCGTCGTGCA GGAAACGTGG CACGGTATCG TTCACGTTGG TCGCAGCAAT  
 AAAACGTTIN ACCGGCAGAC CGAGTGACTT CGCCAGCAGA CCGNCGTNA AATCGCCGAA GTTTNCGCTT TGGCACCGAG  
 ACAACCAGCT GGTGCGCGT CTNCTGCGGN AGCTG