Protein Synthesis

by

Nicola Robertson

A thesis submitted for the
degree of Doctor of Philosophy

University of Edinburgh
September 1996

To my family

This thesis is submitted in part fulfilment of the requirements of the degree of Doctor of Philosophy at the University of Edinburgh. Unless otherwise stated the work described is original and has not been previously submitted, in whole or in part, for any degree at this or any other university.

Nicola Robertson
University of Edinburgh
September 1996

Acknowledgements

First I would like to express my sincere thanks to Professor R. Ramage for giving me the opportunity to study here in Edinburgh and for his constant advice and encouragement throughout the course of my PhD and beyond.

Special thanks are due to Kevin Shaw and Brian Whigham for assistance with the chemical syntheses of polypeptides and for MALDI TOF MS/amino acid analyses respectively. I also thank those involved in the departmental technical services for their efficient work.

I am grateful to Karen Sage and Dr. Sabine Flitsch for their useful advice during the erythropoietin project and also to Karen for carrying out Western Blot analyses on the erythropoietin products.

Many thanks are due to Drs Lu Jiang and Angus Brown for their continued help and advice throughout the course of my PhD and especially to Lu for her support during the HOCt work.

I am extremely grateful to Dr. Kirstie Urquart for her time and patience in proof reading this thesis and also for being a good colleague.

I would also like to thank the BBSRC for funding the project and last but not least a big thank you to all my friends and colleagues in Edinburgh for making the past three years so enjoyable and memorable.

Abstract

A novel coupling reagent for solid phase peptide synthesis has been synthesised under optimum conditions and tested for racemisation with all naturally occurring chiral amino acids. The coupling reagent has been applied to the total chemical synthesis of deglycosylated human erythropoietin (166 residues). The affinity for carbon of the N^{α}–protecting group tetrabenzo[a,c,g,i]fluorenyl-17-methoxycarbonyl (Tbfmoc) has been exploited to simplify the purification of dhEPO and a purification protocol has been developed.

A versatile linker which allows derivatisation at the C-terminus of peptides to either the hydrazide or the amide has been synthesised and attachment of the linker to a suitable solid support has been investigated. The versatile linker has been used in all azide fragment condensation syntheses described herein.

Investigations have been carried out into enzyme cleavable protecting groups for the N^{ε} function of lysine. The protecting groups are designed to prevent unambiguous coupling during fragment condensation of peptide segments. This methodology has been used in the azide condensation of two fragments of Salmon Calcitonin I. In addition to this a chemically cleaved protecting group for lysine, developed for use in the fragment coupling strategy, has been applied to the azide and active ester (using HOCt) coupling of two small fragments of hEPO.

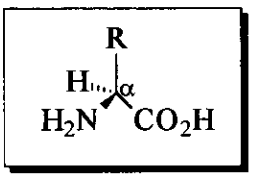
Abbreviations

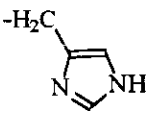
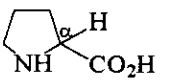
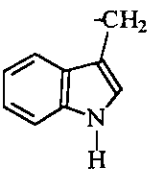
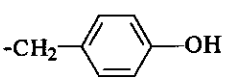
a.a	amino acid
A.A.A	amino acid analysis
ABI	applied biosystems
Acm	acetamidomethyl
AcOH	acetic acid
b	broad
Bepa	4-benzyloxyphenylacetomidomethyl
Boc	tert-butoxycarbonyl
BOP	benzotriazolyl-1-yl-oxy-tris-(dimethylamino)-phosphonium hexafluorophosphate
Bum	N(π)-*t*-butoxymethyl
Bupa	4-*t*-butoxyphenylacetomidomethyl
CD	circular dichroism
cDNA	complimentary deoxyribonucleic acid
CHO	chinese hamster ovary
CMP	chloromethylpolystyrene
d	doublet
Da	dalton
DCC	N,N'-dicyclohexylcarbodiimide
DCM	dichloromethane
dhEPO	deglycosylated human erythropoietin
DIC	N,N'-diisopropylcarbodiimide
DIEA	N,N-diisopropyl ethyl amine
DIU	diisopropylurea
DMAP	4-dimethylaminopyridine
DMF	N,N-dimethylformamide
DMSO	dimethylsulphoxide
DNA	deoxyribonucleaic acid
DTT	dithiolthreitol
EDT	ethane-1,2-dithiol
EDTA	ethylenediaminetetracetic acid
EI	electron ionisation
EPO	erythropoietin
FAB	fast atom bombardment
FC	fragment coupling
Fmoc	9-fluorenylmethyloxycarbonyl
FPLC	fast purification liquid chromatography
Gdm.Cl	guanidinium hydrochloride
h	hour(s)
hEPO	human erythropoietin
HMPA	hexamethylphosphoric triamide

HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole
HOCt	ethyl 1-hydroxy-1H-1,2,3-triazole-4-carboxylate
Hopa	4-hydroxyphenylacetomidomethyl
HPLC	high performance liquid chromatography
HR	high resolution
IL	interleukin
IR	infrared
k	kilo
m	meta
μ	micro
m	multiplet
MALDI	matrix assisted laser desorption ionisation
Mbh	4,4'-dimethoxybenzhydryl
Mepa	4-metoxyphenylacetomidomethyl
min	minutes
Mpt	melting point
MS	mass spectrometry
MWCO	molecular weight cut off
nm	nano metres
NMR	nuclear magnetic resonance
OPfp	pentafluorophenyl
p	para
p.a.	per annum
PAGE	polyacrylamide gel electrophoresis
PAM	*p*-aminomethylated
PGC	porous graphitised carbon
phenac	phenylacetomidomethyl
Pmc	2,2,5,7,8-pentamethylchroman-6-sulphonyl
PTH	phenylisothiocyanate
PVDF	polyvinyldifluoro
PyBOP	benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate
q	quartet
rhEPO	recombinant human erythropoietin
RP	reverse phase
R_t	retention time
s	singlet
SCT	salmon calcitonin
SDS	sodium dodecyl sulphate
SPPS	Solid Phase Peptide Synthesis
SPS	Solid Phase Synthesis
t	tertiary
t	triplet
Tbfmoc	17-tetrabenzo[*a,c,g,i*]fluorenylmethoxycarbonyl

TBTU	2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluroniumtetrafluoroborate
TEA	triethylamine
TFA	trifluoroacetic acid
TFMSA	trifluoromethanesulphonic acid
THF	tetrahydrofuran
TMSBr	trimethylbromosilane
Tnm	N^{ε}-1,5-dioxaspiro-[5,5]-undecane-3-nitro-3-methoxycarbonyl
TOF	time of flight
Tris	tris(hydroxymethyl)aminomethane
Trt	triphenylmethyl
uhEPO	urinary human erythropoietin
UV	ultraviolet
Z	benzyloxycarbonyl

The Naturally Occurring Amino Acids



Amino Acid	3 letter code	1 letter code	R Group
Alanine	Ala	A	$-CH_3$
Arginine	Arg	R	$-(CH_2)_3NHC(NH)NH_2$
Asparagine	Asn	N	$-CH_2CONH_2$
Aspartic acid	Asp	D	$-CH_2CO_2H$
Cysteine	Cys	C	$-CH_2SH$
Glutamic acid	Glu	E	$-(CH_2)_2CO_2H$
Glutamine	Gln	Q	$-(CH_2)_2CONH_2$
Glycine	Gly	G	$-H$
Histidine	His	H	
Isoleucine	Ile	I	$-CH(CH_3)CH_2CH_3$
Leucine	Leu	L	$-CH_2CH(CH_3)_2$
Lysine	Lys	K	$-(CH_2)_4NH_2$
Methionine	Met	M	$-CH_2CH_2SCH_3$
Phenylalanine	Phe	F	$-CH_2(C_6H_6)$
Proline	Pro	P	
Serine	Ser	S	$-CH_2OH$
Threonine	Thr	T	$-CH(CH_3)OH$
Tryptophan	Trp	W	
Tyrosine	Tyr	Y	
Valine	Val	V	$-CH(CH_3)_2$

Contents

Page Nº

Chapter one: Solid Phase Peptide Synthesis 1

1.1 General introduction 1

1.2 Nature of the Polymer Support 3

1.3 N^{α} Protection 4

1.3.1 Boc strategy 4
1.3.2 Fmoc strategy 5

1.4 The Resin Linkage 7

1.4.1 Synthesis of C-terminal acids 7
1.4.2 Synthesis of C-terminal amides and hydrazides 9

1.5 Activation and Coupling 11

1.5.1 Carbodiimides 12
1.5.2 Symmetrical anhydrides 14
1.5.3 Active esters 15

1.5.3.1 HOBt esters 15
1.5.3.2 Pentafluorophenyl esters 16

1.6 References 17

Chapter two: A Novel Coupling Reagent for Solid Phase Peptide Synthesis 19

2.1 Introduction 19

2.2 Synthesis of Ethyl 1-hydroxy-1H-1,2,3-triazole-4-carboxylate 20

2.3 Racemisation Studies with HOCt 22

2.3.1 Racemisation studies with Fmoc-His(Trt) and HOCt 23

2.4 Conclusions 27

2.5 References 28

Chapter three: Stepwise Solid Phase Synthesis of Deglycosylated Human Erythropoietin 29

3.1 Introduction 29

3.2 The Aims 32

3.3 Chemical Synthesis of Deglycosylated Human Erythropoietin 33

3.4 Purification of Deglycosylated Human Erythropoietin 35

3.4.1 Affinity purification of dhEPO using Tbfmoc on charcoal 35

3.4.2 FPLC purification of dhEPO by gel filtration 42

3.4.2.1 Molecular weight determination of dhEPO from FPLC 43

3.5 Characterisation of Purified dhEPO 44

3.5.1 Tryptic digest and mass spectrometry data 45

3.5.2 Sequencing analysis 46

3.5.3 SDS PAGE and western blot analysis 47

3.6 Oxidation and Refolding 48

3.7 Future Research 50

3.8 References 53

Chapter four: Convergent Synthesis 56

4.1 Introduction 56

4.2 Coupling of Minimally Protected Peptide Fragments 57

4.2.1 Chemical Ligation 57

4.2.2 Enzyme catalysed amide bond formation 59

4.2.3 Amide bond formation utilising the azide method 59

4.3 References 64

Chapter five: N^{ε}-Lysine Protection for use in Fragment Coupling 66

5.1 Introduction 66

5.2 Penicillin Acylase Cleaved N^{ε} Protecting Groups 67

5.2.1 N^{ε}-Phenylacetyl-lysine and other derivatives 68

5.2.1.1 Synthesis of N^{ε}-Fmoc-4-*t*-butoxyphenylacetyl-lysine: initial observations 70

5.2.1.2 Investigation into the synthesis of N^{ε}-Fmoc-4-*t*-butoxyphenylacetyl-lysine 72

5.2.1.3 Synthesis and studies on N^{ε}-Fmoc-4-*t*-methoxyphenylacetyl-lysine 73

5.3 Recent Developments 75

5.4 References 76

Chapter six: Investigation into Fragment Coupling Utilising Novel Lysine Protection 78

6.1 Introduction 78

6.2 Synthesis of the Dibenzosuberone Linker 79

6.3 Synthesis of Protected Fragments of SCT for Azide Fragment Coupling 83

6.4 Azide Coupling of Protected Fragments of SCT 85

6.4.1 Enzyme deprotection of SCT 1-32 $(Acm)_2(Hopa)_2$ 87

6.5 Summary 87

6.6 Recent Developments in Lysine Protection 88

6.7 Preliminary Investigations into the Stability of the Tnm Group for Fragment Coupling 89

6.7.1 Azide fragment coupling 89

6.7.2 Ester fragment coupling: preliminary investigation 92

6.8 References 94

Chapter seven: Outlook and Conclusions 95

7.1 References 96

Chapter eight: Experimental 97

8.1 Notes 97

8.2 Solid Phase Peptide Synthesis 98

8.2.1 Side-chain protecting groups used in Fmoc SPPS 98
8.2.2 C-terminal acids 99
8.2.3 C-terminal amides 99
8.2.4 C-terminal hydrazides 100
8.2.5 The Fmoc loading test 100
8.2.6 Automated SPPS 100

8.2.6.1 Capping 100
8.2.6.2 Deprotection 101
8.2.6.3 Coupling 101

8.2.7 The Kaiser test 102
8.2.8 Loading Tbfmoc onto the peptide resin 103
8.2.9 The Tbfmoc loading test 103

8.3 Experimental Details 104

8.4 References 135

1. Solid Phase Peptide Synthesis

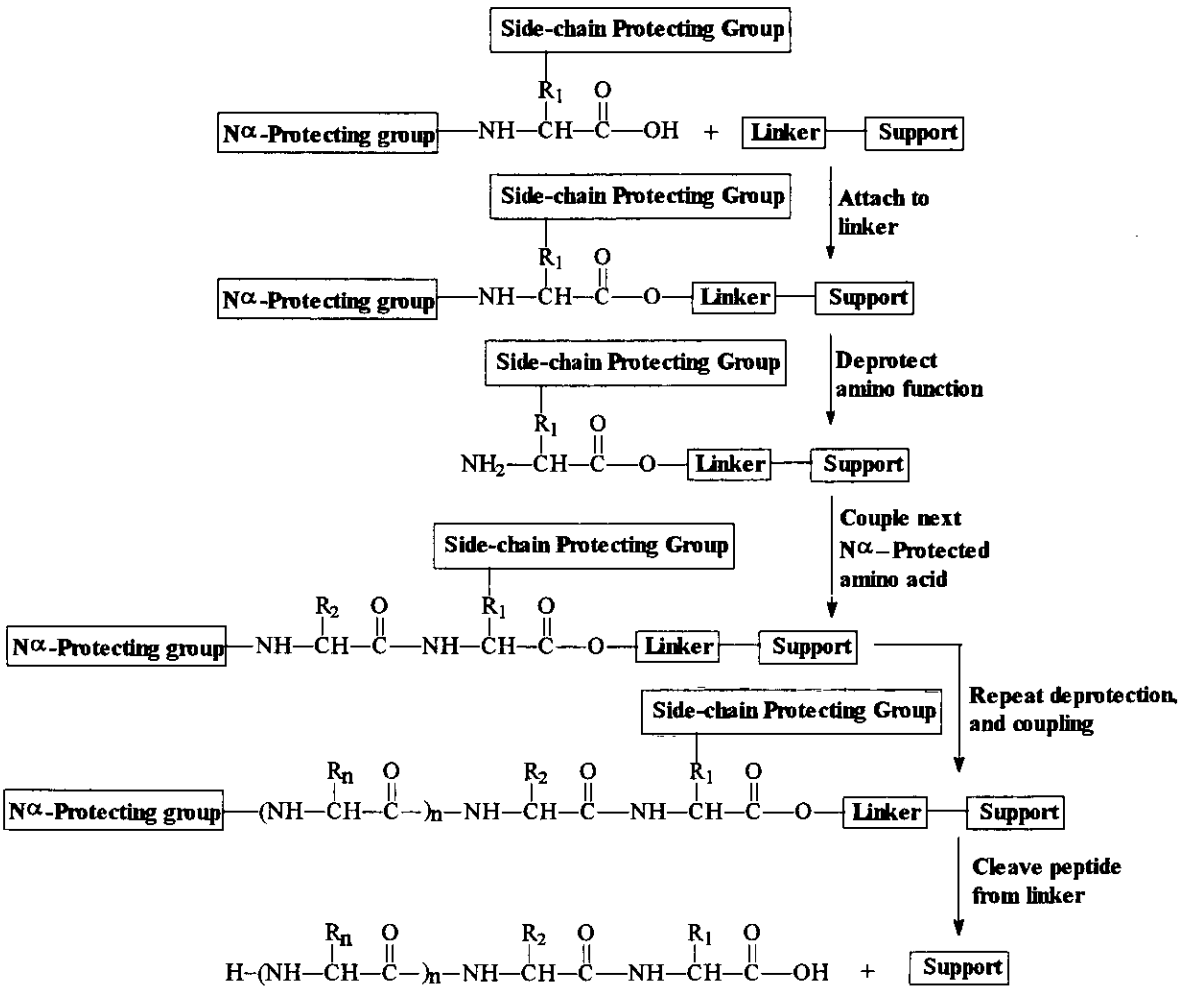
1.1 General Introduction

Advances in the methodology of peptide synthesis have always been stimulated by the existence of interesting target molecules. The discovery of naturally occurring, biologically active peptides created a surge in synthetic activity throughout the 1950's, laying the foundation for the modern era of peptide and protein synthesis. The improved synthetic capabilities of current peptide and protein research originate from a revolutionary development, namely the introduction of the solid phase method by Bruce Merrifield[1], the magnitude of which was recognised by the award of the Nobel Prize for Chemistry to Merrifield in 1984.[2]

Prior to the solid phase method, the synthesis of peptides in solution was a labourious and lengthy process which involved purification and isolation of each intermediate compound. Thus, even the assembly of a relatively short peptide sequence required highly skilled and experienced chemists. Adopting Merrifield's method of chain assembly from an insoluble resin, meant that many of the impurities produced during peptide synthesis could simply be washed away leaving the peptide bound to the resin.

Solid phase methods now dominate synthetic peptide research and the principle has been extended to other fields such as oligonucleotide synthesis, combinatorial libraries and related technologies. For peptide synthesis, Merrifield and others have developed his original chemistry to a fine art and today numerous peptides can be synthesised by automated solid phase synthesis, owing to the simplicity of its strategy and convenience.

The basic concept of Solid Phase Peptide Synthesis (SPPS) has been outlined in Scheme 1.1 below.



Scheme 1.1 Generalised approach to SPPS

An N^{α} derivatised amino acid is attached to an insoluble support *via* a linker. The N^{α}-blocking group is then removed and the next N^{α}-protected amino acid is coupled to the free amino function of the first. The deprotection and coupling cycles are repeated, washing the resin thoroughly between each cycle, until the desired sequence of amino acids is generated. Finally the completed peptide is cleaved from the linker and its side-chain protecting groups simultaneously.

1.2 Nature of the Polymer Support

The general requirements for a suitable solid support for SPPS were outlined by Merrifield and Erickson in 1976[3] and are set out below.

The polymer support must:

(a) have reactive sites at which the peptide can be attached, synthesised and obtained in good yield.
(b) allow good contact between peptide and reagents.
(c) have properties which allow easy separation from excess reagents and by-products.
(d) be stable to the reaction conditions.
(e) minimise interactions between the growing peptide chains.

In his original publication[1] Merrifield used a beaded polystyrene resin containing reactive chloromethyl groups. This was prepared by polymerisation of styrene plus 1% divinylbenzene to produce a rigid cross-linked resin, followed by chloromethylation. Effectively this resin contained a percentage of CH_2Cl groups which could be used for substitution of the Cl by the carboxyl group of an α- amino acid. Hydrolysis of the ester bond on completion of the peptide sequence could then occur *via* alkyl oxygen fission.

The small spherical beads of the resin are about 50μm in diameter but swell to five or six times their original volume in organic solvents such as DCM[4], allowing the reagents access to the peptide; hence coupling and deprotection can occur within the resin and soluble unwanted products can filter through.

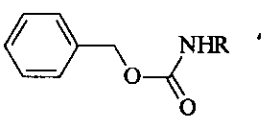
Subsequently, the British scientist Robert Sheppard argued that a resin more compatible with the polarity of the growing peptide chain would be more

advantageous to solid phase methodology. Accordingly he developed a beaded polyamide resin[5,6] with swelling characteristics which are complementary to Merrifield's polystyrene resin. The polyamide resins have come into widespread use and when supported on kieselguhr are mechanically robust proving suitable for continuous flow SPPS.

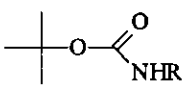
1.3 N^{α} Protection

1.3.1 Boc strategy

A year after his preliminary publication in which Merrifield[1] established the principles of SPS he described a synthesis of bradykinin.[7] The original benzyloxycarbonyl (Z) group (1), for N-terminal protection, had been replaced by the more acid labile *t*-butoxycarbonyl (Boc) group (2) and Z was now used as a permanent side-chain protecting group.



(1)



(2)

In this synthesis the C-terminus of the first amino acid residue in the target sequence was anchored to the chloromethylpolystyrene resin *via* its caesium salt displacing chloride ion and forming an ester bond. Removal of Boc, from the N^{α} positon, with mild acid followed by neutralisation gave the free amino group, to which the second amino acid residue was coupled using N,N'-dicyclohexylcarbodiimide (DCC) for activation. The peptide chain was lengthened in this way until the desired sequence was fully assembled. After cleavage of the peptide from the resin with HBr/TFA the side chains, protected with Z, were deprotected by catalytic hydrogenation to give a

quantitative yield of the nonapeptide. At each stage the completely insoluble polymer was washed copiously to remove excess reagents and co-products.

1.3.2 Fmoc strategy

Over the years, N^{α} protecting group strategy and the linkage of the growing peptide chain was studied in detail and as the synthetic targets increased not only in size but in sensitivity, the limitations of Boc SPPS became more apparent. Both the repetitive cleavage of Boc groups and especially the final cleavage of the peptide from the resin involved treatment with strong acids which were clearly destructive to some peptide sequences. Thus a fully orthogonal approach to SPPS was developed. The 9-fluorenylmethyloxycarbonyl (Fmoc), N^{α} protecting group[8,9] (3), is very stable towards acidic reagents but is cleaved swiftly under certain basic conditions. Piperidine (4) is the reagent routinely used and deprotection takes only seconds at room temperature *via* a β-elimination process (Fig. 1.1). Thus, Fmoc temporary protection can be combined with *t*-butyl derived acid labile side chain protection.

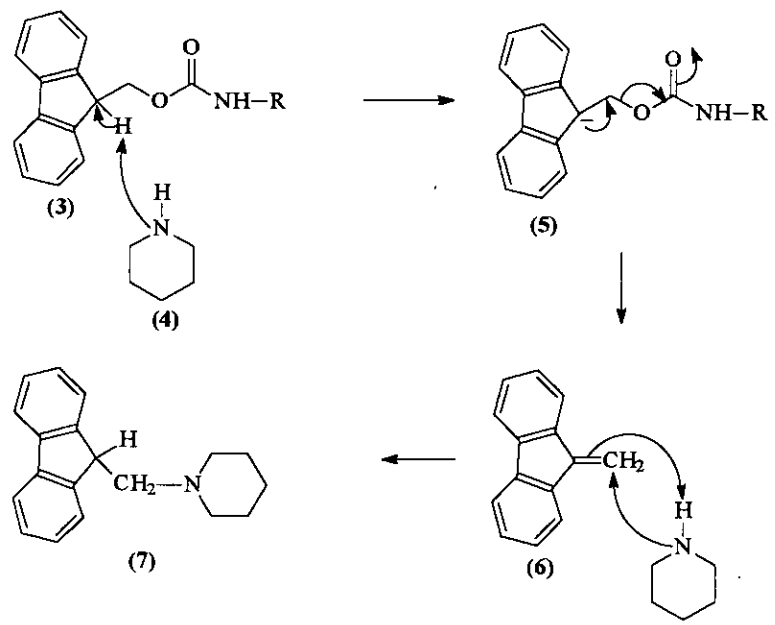
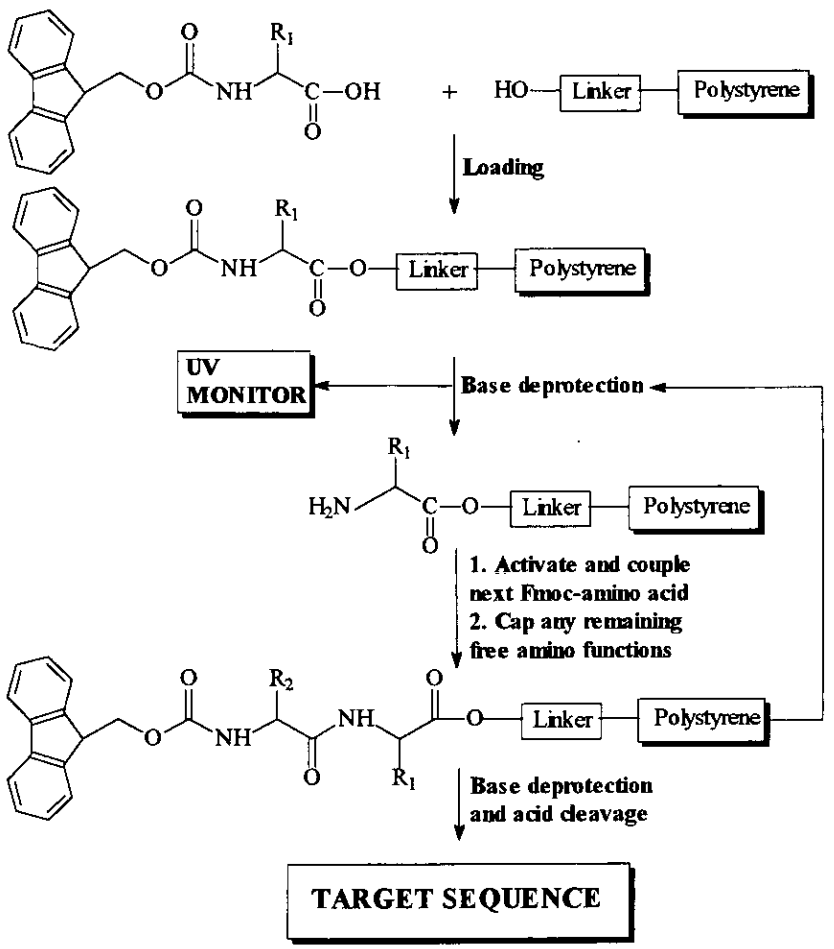


Fig 1.1 Mechanism of Fmoc deprotection

The dibenzofulvene (6) produced from the stabilised dibenzocyclopentadienide anion (5) forms the adduct (7) on reaction with excess piperidine (4) and in this way is prevented from reacting with the free amine of the peptide. The fulvene piperidine adduct (7) has a strong UV absorbance at 302nm and this property has been used in developing a system for monitoring the efficiency of each coupling step in automated Fmoc SPPS,[10] the general scheme for which is outlined in Scheme 1.2.



Scheme 1.2 SPPS utilising Fmoc amino acids

The initial protected Fmoc-amino acid is first coupled *via* an ester linkage to the exposed hydroxy function on the linker in the presence of a catalytic amount of DMAP. The resin bound Fmoc-amino acid is base deprotected by treatment with

20% piperidine/DMF and, after thorough washing, the second protected amino acid is coupled, as either a preactivated species (symmetrical anhydride), or without preactivation (*in situ*) where several different activators may be utilised. A capping cycle is then introduced to block any remaining free amino function to prevent its participation in subsequent couplings. This sequence of events is repeated until the target peptide is obtained. The completed peptide is then released from the acid labile linker and the base stable side-chain blocking groups by treatment with acid (usually TFA). Each coupling step is monitored by UV as mentioned above.

Fmoc SPPS and the reagents used will be discussed in detail in the chapters to follow along with its applications to peptide synthesis in practice.

1.4 The Resin Linkage

1.4.1 Synthesis of C-terminal acids

The introduction of linkers between the support and the peptide has allowed fine tuning of the strength of the peptide-resin bond, so that it can be made more, or less labile, to suit the synthetic strategy.

With the advent of Fmoc methodology and the use of *t*-butyl derived side-chain protecting groups came the need for linkers which were more acid labile than the conventional polystyrene or Merrifield PAM[11] resin (8) used in the original Boc chemistry. The Wang linker[11] (9), first introduced for the synthesis of protected peptide fragments, is cleaved in mild acid (50% TFA in aprotic solvents) and has remained popular in Fmoc SPPS (Fig 1.2).

In both Boc and Fmoc methodologies the cleavage mechanism is an ester hydrolysis by alkyl-oxygen fission. In the case of the *p*-benzyloxyalcohol (Wang) linker the additional *p*-oxygen donates electrons into the benzyl alcohol ring, stabilising the

carbocation formed during the fission process, thus increasing the lability of the peptide.

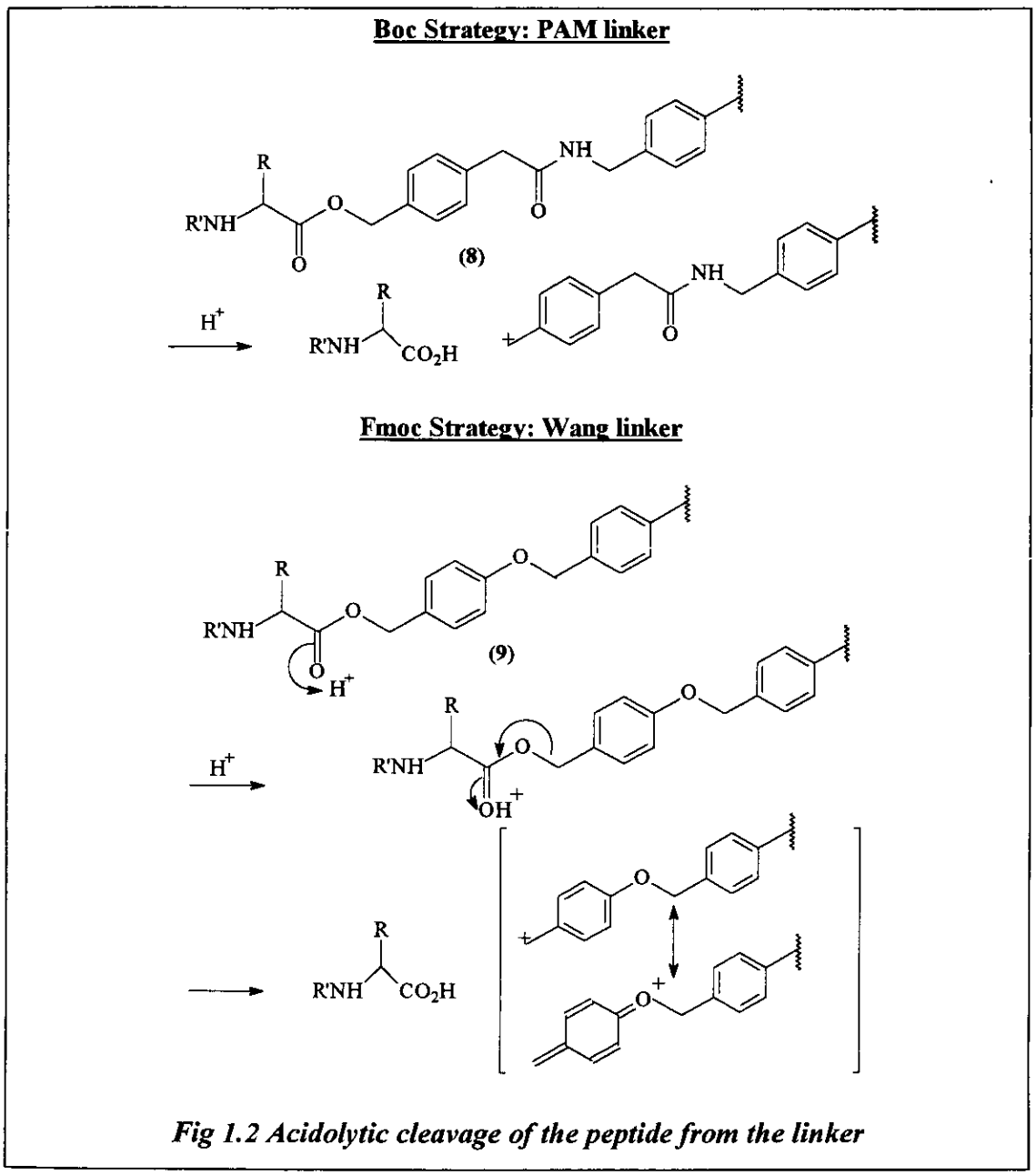
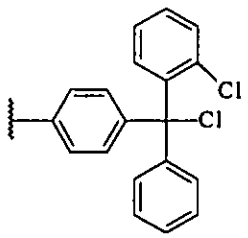


Fig 1.2 Acidolytic cleavage of the peptide from the linker

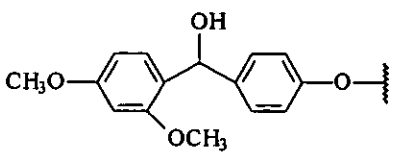
One disadvantage in using the Wang linker is 2,5-diketopiperizine[12] formation with C-terminal proline. Utilising the highly acid labile 2-chlorotrityl chloride linker[13,14] (10)

(cleaved in 10% acetic acid), this problem can be overcome, diketopiperizine formation being suppressed by steric bulk at the C-terminal ester.

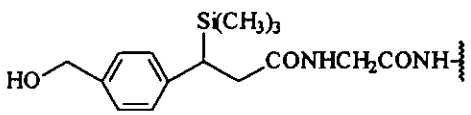


(10)

More recently linkers, including the chlorotrityl linker, which cleave under extremely mild conditions such as dilute TFA[15] (11) and fluoride ion[16,17] (12) have been developed to produce protected peptide fragments for convergent synthesis (Chapter 4).



(11)



(12)

1.4.2 Synthesis of C-terminal amides and hydrazides

The use of linkers also allows the functionality at the C-terminus of the peptide to be adapted to specific requirements e.g linkers are available for the generation of C-terminal amides[15,18-21] and hydrazides[18,22] under standard cleavage conditions (Fig 1.3).

Synthesis of C-terminal peptide amides is highly desirable as many biologically important peptides have a C-terminal amide functionality (Chapter 6) e.g. hormones and neurotransmitters and the synthesis of peptide hydrazides is important for

convergent synthesis using acid azides (Chapter 4). Utilising linkers to derivatise the C-terminus allows these C-terminal functionalities to be generated using mild procedures which are not destructive to the peptide chain.

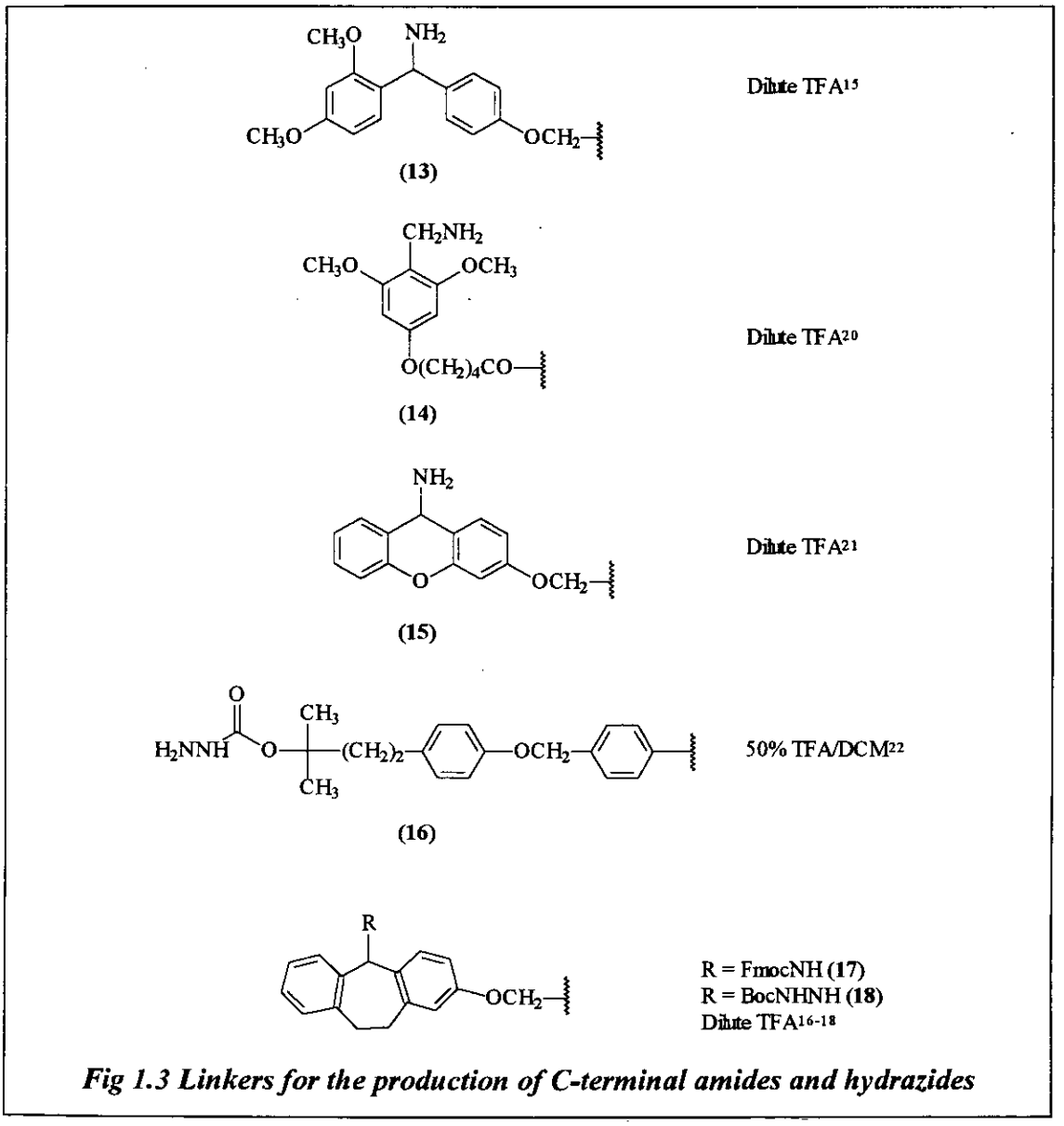


Fig 1.3 Linkers for the production of C-terminal amides and hydrazides

1.5 Activation and Coupling

The principal reaction in the synthesis of a peptide is acylation of the amino group of an amino acid by the carboxyl group of another amino acid to form an amide bond; the process is known as coupling. In order to convert carboxylic acids into acylating agents their hydroxyl group must be replaced by an electron withdrawing group, X, in order to enhance the electrophilicity of the carboxyl group (Fig 1.4).

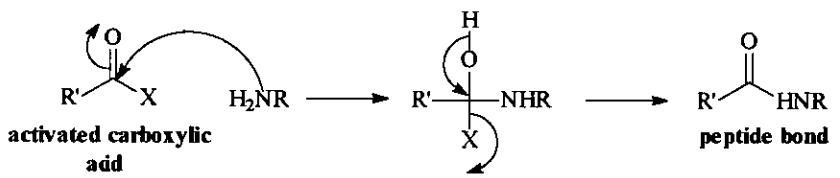


Fig 1.4 Acylation of an amino acid: coupling

Obviously, the success of a peptide synthesis relies mainly upon the ability to couple amino acids together efficiently to form the peptide bonds. Accordingly this area has attracted much attention and many coupling reagents have been proposed.[23] Reagent suitability is often ruled by the need to retain the chiral integrity of each amino acid on coupling, avoiding racemisation.

Racemisation of amino acids can occur upon activation of the carboxy function with an electron withdrawing substituent, X (19). The mechanism[24] (Fig 1.5) involves the formation and enolisation of an oxazolone, (20) and (21), on treatment with base. The oxazolones formed are themselves activated carboxylic acid derivatives and reaction with an amine can lead to peptide formation. However racemisation *via* the stabilised anion (22) occurs more rapidly than aminolysis and both possible epimers (23) and (24) are formed.

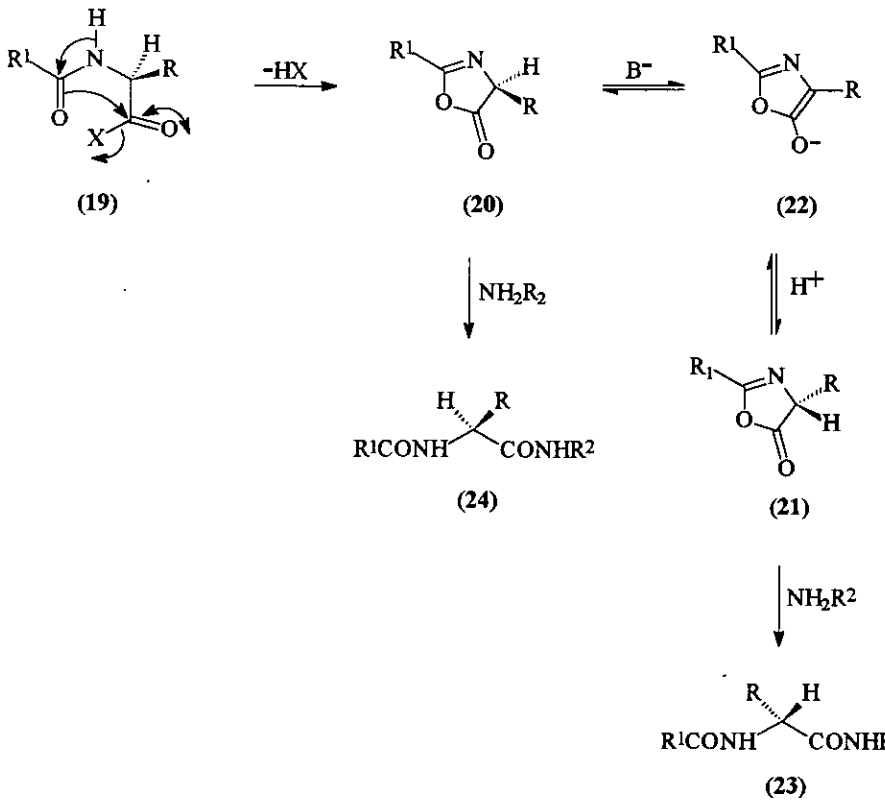


Fig 1.5 Base catalysed racemisation of activated amino acids via the oxazolone intermediate

It has been estimated that over 140 methods for peptide bond formation exist[25] but only those coupling techniques which have generally been applied to Fmoc SPPS will be discussed here.

1.5.1 Carbodiimides

Carbodiimides are among some of the most popular activating reagents employed in peptide synthesis. N,N'-Dicyclohexylcarbodiimide (DCC)[26] and N,N'-diispropylcarbodiimide (DIC)[27] (24) are used extensively in SPPS although DIC is preferred since the by-product, N,N'-diisopropyl urea (26), formed from the amidation reaction (Fig 1.5) is more soluble than its counterpart N,N-dicyclohexylurea.

The carboxy function of the α-amino acid reacts with the carbodiimide (25) to form an O-acylurea (26), which is a potent acylating reagent. This can then react with the free amino function of the growing peptide to form a peptide bond, or more usually react with another oxygen nucleophile to form either the symmetrical anhydride (28) (Section 1.5.2) or the active ester (29) (Section 1.5.3).

Although activation *via* carbodiimides generates an acylating reagent with higher reactivity than (28) or (29), this method of activation can cause serious side-reactions such as racemisation (Fig 1.4), carboxamide dehydration of asparagine (30) and glutamine and also the formation of N-acyl ureas (31) (Fig 1.7).

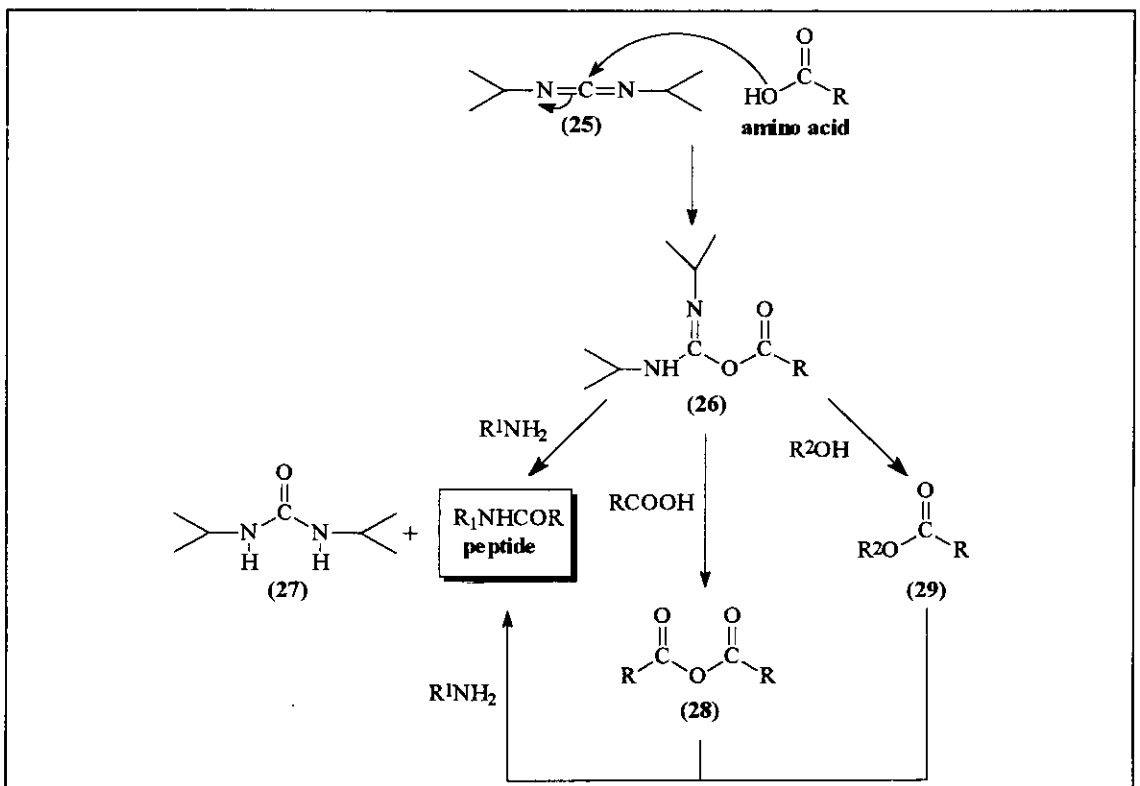


Fig 1.6 Activation and coupling using the carbodiimide procedure

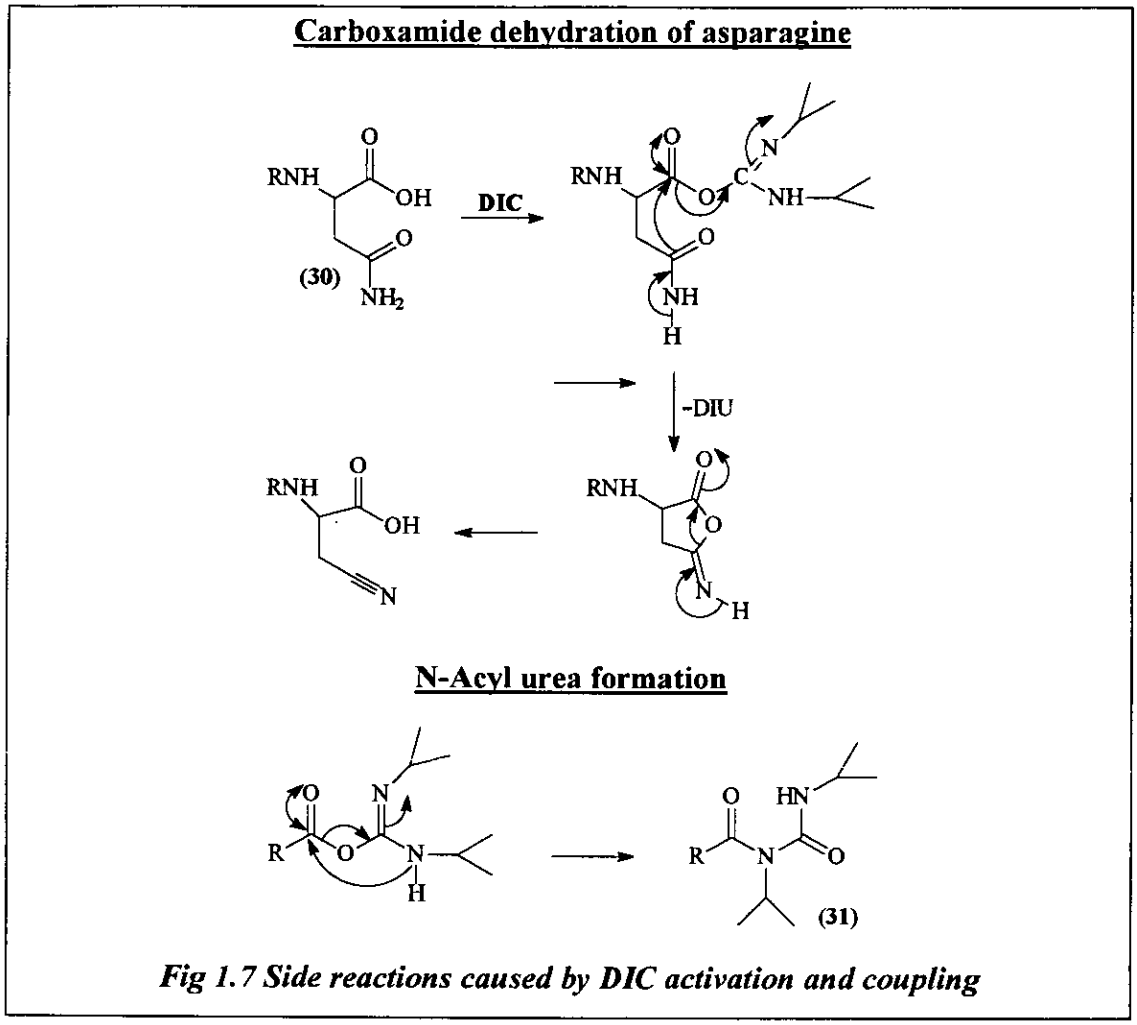


Fig 1.7 Side reactions caused by DIC activation and coupling

1.5.2 Symmetrical anhydrides

Symmetrical anhydrides (28) are prepared from reaction of two equivalents of amino acid derivative with one equivalent of carbodiimide (25). Although they are normally crystalline solids, they are routinely pre-formed automatically during synthesis and are used without isolation.

Aminolysis of symmetrical anhydrides is unambiguous and it has thus found widespread use in Fmoc SPPS[28], but the price for this is that only half of the expensive Fmoc-amino acid used is incorporated into the product. Another major

drawback to it's use are the side reactions apparent with asparagine, glutamine (Fig 1.7) and histidine (racemisation).[29]

1.5.3 Active esters

Although the coupling rates of activated esters are considerably slower than those of the symmetrical anhydrides, they have considerable advantages over symmetrical anhydrides in that none of the amino acid is 'wasted' and side reactions are minimal. Extensive studies have been carried out on the use of Fmoc amino acid active esters and these are now routinely used in SPPS.[30] Some important ones are described below (Fig 1.8).

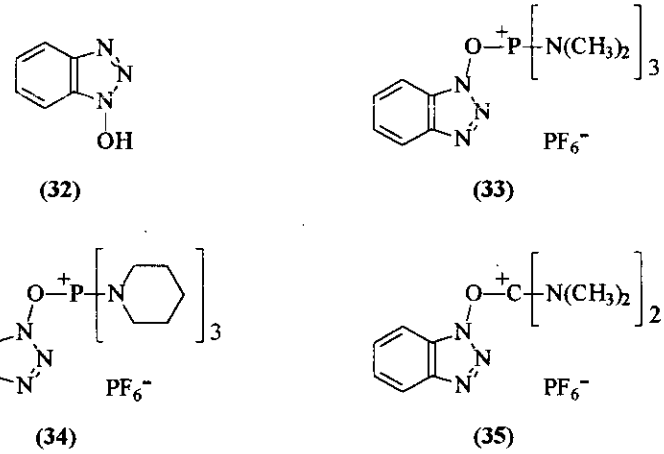
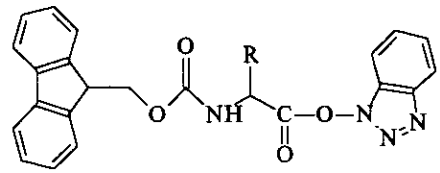


Fig 1.8 Coupling reagents used in Fmoc SPPS

1.5.3.1 HOBt esters

1-Hydroxybenzotriazole (HOBt) (32), first introduced into peptide synthesis as an additive to the carbodiimide method to prevent racemisation,[31] forms an active ester which has found great success in Fmoc SPPS. HOBt[32] (36) esters of protected amino acids are easily formed for example from DIC/HOBt *in situ* (Fig 1.6).



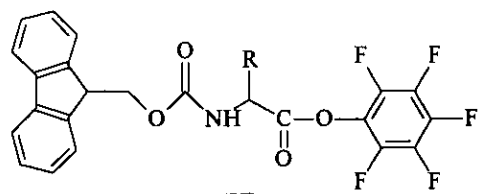
(36)

HOBt esters can also be formed from either its phosphonium or uronium derivatives e.g BOP[33] (33), PyBOP[34] (34) or TBTU[35] (35) (Fig 1.8).

The phosphonium reagent BOP (33) is an excellent peptide coupling reagent but the formation of the carcinogen hexamethyl phosphoric triamide (HMPA) on coupling has led to its replacement by PyBOP (34) and more recently, by the uronium salt, TBTU (35). These coupling reagents rate as highly as, if not better, than BOP with no carcinogenic by-products.

1.5.3.2 Pentafluorophenyl esters

Pentafluorophenyl (OPfp) esters[36] (37) are also efficient acylating agents and due to the bulk of their chemical structures, virtually no side-reactions are observed. They react considerably slower than symmetrical anhydrides but the addition of HOBt (32) significantly increases the rate of reaction, making these esters very useful activated species for SPPS.



(37)

Overall *in situ* activating reagents have become widely accepted in SPPS because of their ease of use, fast reactions and their general lack of side-reactions.

1.6 References

(1) R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149-2154.

(2) R. B. Merrifield, *Science*, 1986, **232**, 341-347.

(3) B. W. Erickson and R. B. Merrifield, *The Proteins*, eds. H. Neurath and R. L. Hill, 1976, vol. 2 3rd edn., Academic Press, New York, 255-257.

(4) B. Gutte and R. B. Merrifield, *J. Biol. Chem.*, 1971, **246**, 1922.

(5) E. Atherton, D. L. J. Clive and R. C. Sheppard, *J. Am. Chem. Soc.*, 1975, **97**, 6585

(6) R. Ashardy, E. Atherton, D. L. J. Clive and R. C. Sheppard, *J. C. S., Perkin Trans. I*, 1981, 1151.

(7) R. B. Merrifield, *Biochemistry*, 1964, **3**, 1385-.

(8) L. A. Carpino and G. Y. Han, *J. Org. Chem.*, 1972, **37**, 3404-3409.

(9) E. Atherton, C. J. Logan and R. C. Sheppard, *J. C. S. Perkin I*, 1981, 538-545.

(10) K. M. Otteson, R. L. Nobel, P. D. Hoeprich, Jr., K. T. Shaw and R. Ramage *Applied Biosystems Res. News*, June 1993.

(11) S.S. Wang, *J. Am. Chem. Soc.*, 1973, **95**, 1328-1333.

(12) B. F. Gissin and R. B. Merrifield, *J. Am. Chem. Soc.*, 1972, **94**, 3102-3103.

(13) K. Barlos, D. Gatos, S. Kapolos, G. Paphotu, W. Schafer and Y. Wenqing, *Tet. Lett.*, 1989, **30**, 3947-3950.

(14) K. Barlos, D. Gatos, S. Kapolos, G. Paphotu, W. Schafer and Y. Wenqing, *Tet. Lett.*, 1989, **30**, 3943-3946.

(15) H. Rink, *Tet. Lett.*, 1987, **28**, 3787-3790.

(16) R. Ramage, C. A. Barron, S. Beilecki and D. W. Thomas, *Tet. Lett.*, 1987, **28**, 4105-4108.

(17) R. Ramage, C. A. Barron, S. Beilecki and D. W. Thomas, *Tet. Lett.*, 1992, **48**, 499-514.

(18) R. Ramage, S. L. Irving and C. M^{c}Innes, *Tet. Lett.*, 1993, **34**, 6599-6602.

(19) J. P. Tam, *J. Org. Chem.*, 1985, **50**, 5291-5298.

(20) F. Albericio, N. Kneib-Cordonier, S. Biancalana, L. Gera, R. I. Masada, D. Hudson and G. Barany, *J. Org. Chem.*, 1990, **55**, 3730-3743.

(21) P. Seiber, *Tet. Lett.*, 1987, **28**, 2107-2110.

(22) S. S. Wang, *J. Org. Chem.*, 1975, **40**, 1235-1239.

(23) E. Gross and J. Meienhofer, *The Peptides: Analysis, Synthesis, Biology, eds.*, 1979, **11**, Academic Press, New York.

(24) M. Bodanszky, in *Peptide Chemistry*, 1988, Springer-Verlag, Berlin pp.115-118.

(25) H. D. Jakubke, *The Peptides*, 1987, Vol 9, Academic Press Inc., pp.103

(26) J. C. Sheehan and G. P. Hess, *J. Am. Chem. Soc.*, 1955, **77**, 1067-1068.

(27) D. Sarantakis, J. Teichman, E. L. Lien and R. L. Fenichel, *Biochem., Biophys. Res. Commun.*, 1976, **73**, 336.

(28) E. Brown et al., *J. Chem. Soc., Perkin Trans. I*, 1983, 75

(29) H. Rink and B. Riniker, *Helv. Chim. Acta*, 1974, **57**, 831

(30) G. B. Fields and R. L. Noble, *Int. J. Pept. Prot. Res.*, 1990, **35**, 161-214.

(31) W. Konig and R. Geiger, *Ber. Dtsch. Chem. Ges.*, 1970, **103**, 788

(32) W. Konig and R. Geiger, *Ber. Dtsch. Chem. Ges.*, 1970, **103**, 2024, 2034.

(33) B. Castro, J. R. Dormoy, G. Evin and C. Selve, *Tet. Lett.*, 1975, 1219-1222.

(34) J. Coste, D. Le-Nguyen and B. Castro, *Tet. Lett.*, 1990, **31**, 205-208.

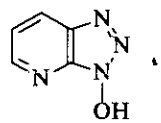
(35) B. Castro, J. R. Dormoy, B. Dourtoglon, G. Evin and C. Selve and J. C. Zeigler, *Synthesis.*, 1976, 751.

(36) L. Kisfaludy and I. Schon, *Synthesis*, 1983, 325-327.

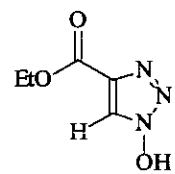
2. A Novel Coupling Reagent for Solid Phase Peptide Synthesis

2.1 Introduction

A recent development in the field of HOBt based coupling reagents has been the design of new, improved analogues, 1-hydroxy-7-azabenzotriazole (HOAt)[1] (38) and ethyl 1-hydroxy-1H-1,2,3-triazole-4-carboxylate (HOCt)[2-4] (39).

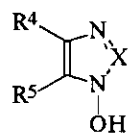


(38)



(39)

The HOCt analogue which was designed and synythesised, in this laboratory, by Amanda Davison[2], originated from a series of N-hydroxy compounds formulated from the dissection of HOBt to the basic structure (40).



(40)

Davison studied a series of imidazoles (X = CR) and triazoles (X = N), of these HOCt (39) showed the most promising properties, giving increased acidity compared to HOBt, rendering it a better leaving group. Another attractive feature of HOCt was it did not absorb at 302nm. This meant that a monitoring system, based upon the

Fmoc chromophore could be utilised whereby the efficiency of the acylation reaction could be determined prior to Fmoc removal. Thus, if the coupling of an amino acid was poor it could easily be recoupled at this stage. This is not possible using the monitoring system previously described for Fmoc SPPS (Section 1.3.2) since this is based solely on the deprotection of Fmoc after coupling.

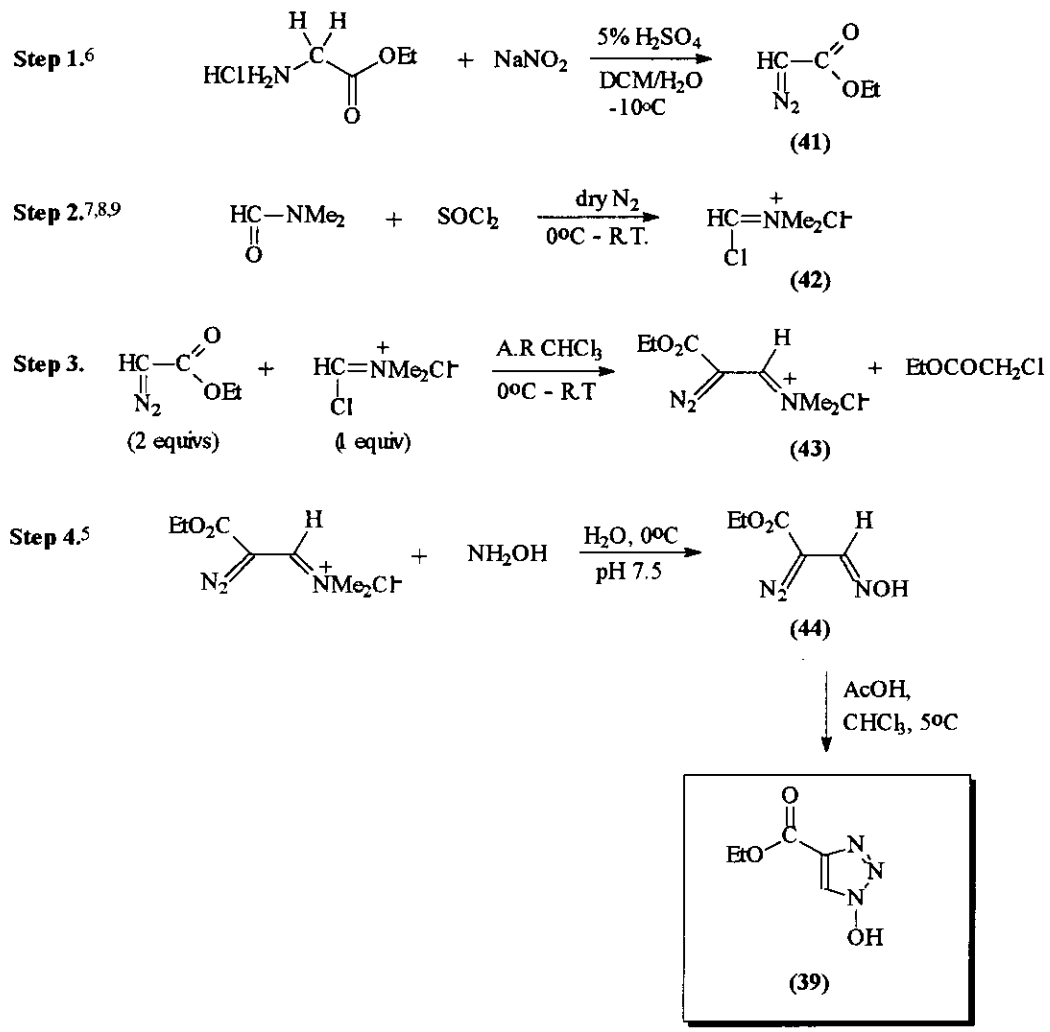
These promising features encouraged the development of a synthetic route towards HOCt. The initial route developed by Davison was based on the literature method of Stojonavic.[5] However, there were a number of problems inherent in this synthesis, mainly its irreproducibility.

Jiang[3] investigated the synthesis of HOCt further, optimising the individual steps and found that the final step could in fact be carried out in water which was a more favourable solvent than the absolute ethanol previously used as no rotary evaporation was required since the product precipitated from water. In addition the previously insoluble Na_2CO_3 was now completely solubilised. However, it was found that replacement of ethanol with water meant a reduction in the yield of HOCt. Thus, further work was carried out to try and improve the yield for this step as it was preferrable to use water as opposed to ethanol.

2.2 Synthesis of Ethyl 1-hydroxy-1H-1,2,3-triazole-4-carboxylate

It was noted that cyclisation of the diazo ester (44), formed in step 4, to HOCt (39) (Scheme 2.1) was carried out in acidic conditions. Thus, for the formation of the diazo ester, if the pH of the aqueous reaction was acidic, then some HOCt could be formed at this stage which, due to its solubility in water, could be lost on separation of the precipitated diazo ester from the water soluble inorganics. This would result in a lower overall yield of HOCt. In order to investigate this problem several experiments were carried out. The first synthesis was a repeat of Jiang's work in order to determine the pH of the aqueous solution in step 4. It was measured as

acidic (pH 4.5). Thus, in order to test the above idea step 4 was carried out adjusting the pH with 10% Na_2CO_3 to obtain less acidic conditions.



Scheme 2.1 Synthesis of HOCt

The results from the experiments are summarised below in Table 1.

Table 1: pH studies in HOCt synthesis

pH	Overall Yield (%)
4.5	15.3
6.5	19.4
7.5	24.1

The results suggest that the above explanation is true since improved yields of HOCt are observed when the diazo ester is formed in less acidic conditions thus preventing premature cyclisation at this stage. Scheme 2.1 is now the standard protocol for the synthesis of HOCt. It is convenient, short (two working days as opposed to one week by Davison's method), inexpensive and an average overall yield of 30-35% can be achieved reproducibly.

2.3 Racemisation Studies with HOCt

Racemisation still remains the most serious side-reaction in peptide synthesis and a considerable research effort has been put into this area.[10] Thus, any new coupling reagent designed for SPPS should be investigated thoroughly.

In order to study whether HOCt would cause racemisation on amino acid activation, a series of tripeptides with the general structure Ala-X-Gly were synthesised (where X = amino acid). Histidine was treated as a separate case (Section 2.3.1) since it is known to be problematic.

Ala-X-Gly was chosen as the model peptide since Gly is not chiral and Ala has a characteristic doublet between 1 and 2 ppm in the ^{1}H NMR corresponding to its side-chain methyl. Thus, if any of the amino acids should racemise with HOCt a second doublet would then be observed for Ala due to the formation of the D isomer (Fig 2.1) thus providing an easy method of analysis.

Each tripeptide was synthesised manually in a sonic bath and the amino acids with their various side-chain protecting groups were coupled, under standard conditions (Chapter 8), as their active HOCt esters. The peptides were analysed using 360 MHz ^{1}H NMR and the results are summarised below in Table 2.

The NMR results obtained for the series of trimers studied showed no evidence of racemisation with any of the amino acids listed in Table 2 on activation with HOCt, in other words only one doublet was observed between 1 and 2 ppm in the ^{1}H NMR spectrum. These results strongly indicate that HOCt does not promote racemisation on activation with any of these naturally occurring chiral amino acids.

Table 2: Racemisation studies with HOCt: using Ala-X-Gly as model

X	Side chain protecting group	Racemisation (%)
Ala	N/A	0
Arg	Pmc	0
Asp	*t*-Bu	0
Asn	Mbh, Trt	0
Cys	Acm, *t*-Bu, Trt	0
Gln	Trt	0
Glu	*t*-Bu	0
Ile	N/A	0
Leu	N/A	0
Lys	Boc	0
Met	N/A	0
Phe	N/A	0
Pro	N/A	0
Ser	*t*-Bu	0
Thr	*t*-Bu	0
Trp	indole NH not protected	0
Tyr	*t*-Bu	0
Val	N/A	0

2.3.1 Racemisation studies with Fmoc-His(Trt) and HOCt

Of all the twenty naturally occuring amino acids, histidine (45) is especially prone to racemisation. In addition to the possibility of racemisation *via* the formation of an oxazolone[11] (Fig 1.5, Section1.5), the side-chain[12] imidazole ring of histidine can also play a part (Fig 2.1).

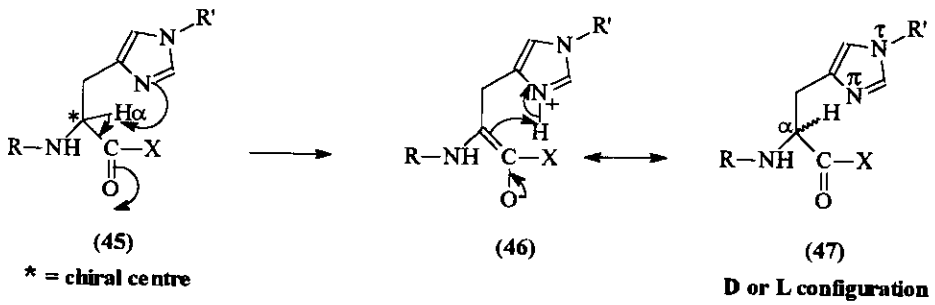
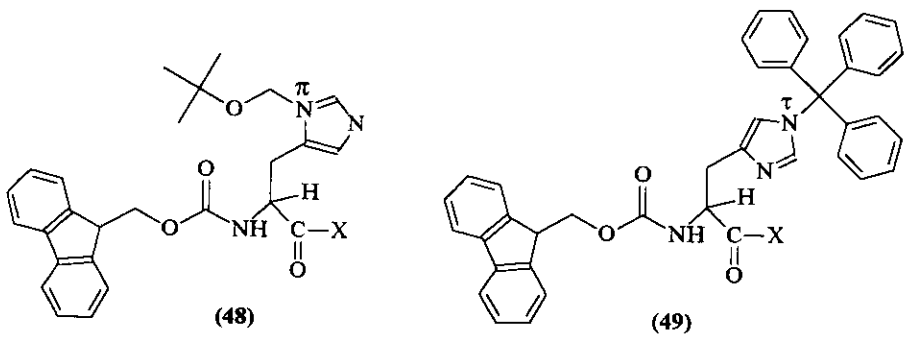


Fig 2.1 Effect of imidazole ring on the racemisation of histidine

The imidazole ring is a weak base but is strong enough to cause intramolecular proton abstraction to form (46) which can then form either the D or L amino acid (47).

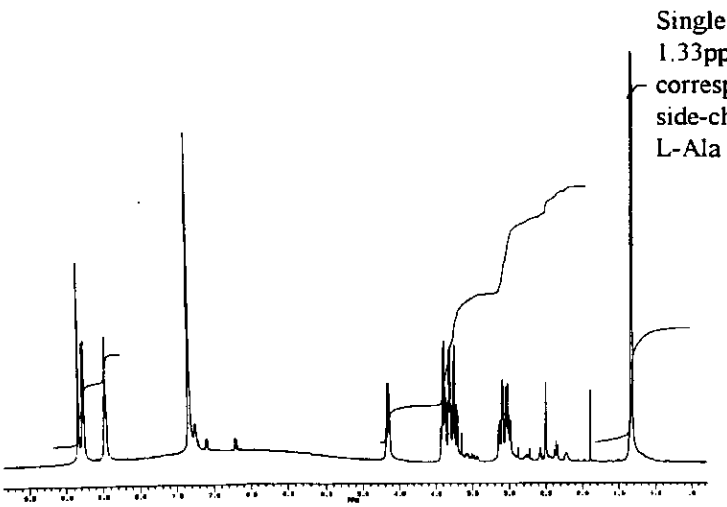
Masking the π-nitrogen in the imidazole function would seem to be a worthwhile objective, however, the synthesis of π-protected histidine is very difficult and the only commercially available analogue is Fmoc-His(Bum)[13] (48), which is impure and very expensive due to its poor yield on synthesis. Fortunately, by introducing a bulky protecting group on the τ-nitrogen e.g trityl (Trt) (49), the racemisation of histidine can be reduced.



The racemisation of histidine with HOCt was studied using the model peptide Ala-His-Gly and the results and reaction conditions are tabulated below (Table 3).

0.1%. Fig 2.2 shows two examples of the 360MHz ^{1}H NMR obtained in the racemisation studies with histidine illustrating the additional doublet obtained for Ala when racemisation occurs.

A: example of no racemisation with His(Trt)



B: example of racemisation with His(Trt)

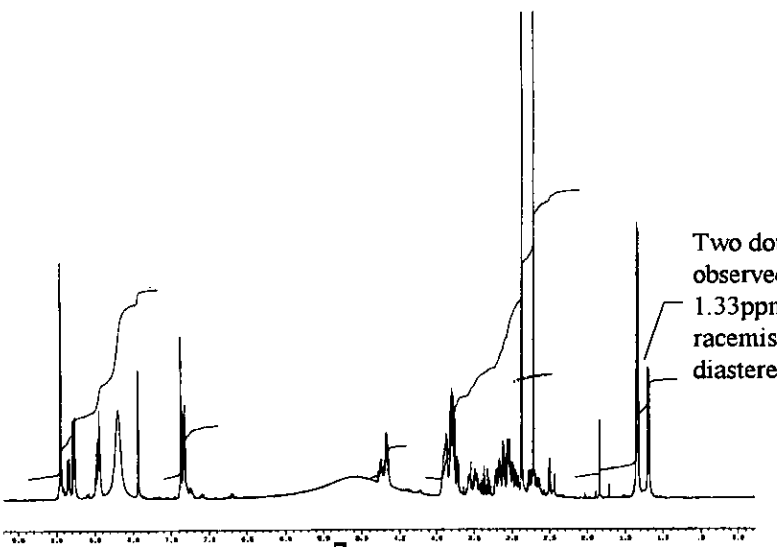


Fig 2.2 Examples of ^{1}H NMR obtained for racemisation studies with His.

2.4 Conclusions

The results suggest that not only is racemisation of histidine, when activated with HOCt, affected by time but also by temperature. The fact that the amount of racemisation is reduced on addition of an excess of HOCt indicates that HOCt, due to its acidic nature ($pK_a = 2.1$), is protonating the π-nitrogen in the imidazole ring and, in doing so, inhibits intramolecular proton abstraction (Fig 2.2).

The reason for the observed increase in racemisation of histidine, with HOCt compared to HOBt, is probably due to $^-$OCt being a better leaving group than $^-$OBt due to its stability. This means the $^{\alpha}$H is more acidic when histidine is activated with HOCt, rendering it extremely labile, thus facilitating racemisation by proton abstraction *via* the imidazole ring (Fig 2.1).

Prior to these results Fmoc-His(Trt) could not be coupled as the HOCt active ester, due to extensive racemisation under standard conditions, and was incorporated onto the synthesiser as its HOBt active ester. Now all amino acids can be successfully incorporated as their active HOCt esters.

HOCt has been proven to be a superior coupling reagent to HOBt/DIC even when double coupling cycles were used for the latter. Thus the synthesis of larger polypeptides using HOCt would seem feasible since coupling of amino acids should be more efficient, decreasing the possibility of truncated peptides. The following chapter discusses the application of HOCt in the synthesis of a 166 amino acid protein with reference to other proteins synthesised with this novel coupling reagent.

2.5 References

(1) L. A. Carpino, *J. Am. Chem. Soc.*, 1993, **115**, 4397-4398.

(2) A. Davison, *PhD Thesis*, 1992, Edinburgh University.

(3) L. Jiang, *PhD Thesis*, 1996, Edinurgh University.

(4) R. Ramage, G. W. Biggin, A. R. Brown, A. J. Comer, A. Davison, L. C. Draffan, L. Jiang, G. H. Morton, N. Robertson, K. T. Shaw, G. Tennant, K. Urqhuart and J. Wilken, *Proc. 4th International Symposium in Solid Phase Synthesis & Combinatorial Chemical Libraries*, Edinburgh, 1996.

(5) F. M. Stojanovic and Z. Arnold, *Collect. Czech. Commun.*, **32**, 2155

(6) N. E. Searle, *Org. Synth.*, 1955, **36**, 25-28.

(7) H. H. Bossard, R. Mory, M. Schmid and Hch. Zollinger, *Helv. Chim. Acta.*,1959, **42**, 1653-1658.

(8) J.-P. Dulcere, *Tet. Lett.*, 1981, **22**, 1599-1600.

(9) C. Jutz, *Adv. Org. Chem.*, 1976, **9**, 225-342.

(10) D. S. Kemp, in *The Peptides*, 1979, Vol 1 (E. Gross, J. Meienhofer, Eds.), pp 315, New York, Academic Press.

(11) M. Bergmann and L. Zervas, *Biochemistry*, 1928, **207**, 280.

(12) J. H. Jones, W. I. Ramage and M. J. Witty, *Int. J. Pept. Prot. Res.*, 1980, **15**, 301.

(13) R. Colombo, F. Colombo and J. H. Jones, *J. C. S., Chem. Commun.*, 1984, 292-293.

3. Stepwise Solid Phase Synthesis of Deglycosylated Human Erythropoietin

3.1 Introduction

Erythropoietin (EPO) is the principal hormone involved in the regulation and maintenance of red blood cell (erythrocyte) production in the bone marrow.[1,2] The hormone is produced in the kidney of the adult and in the liver during foetal life,[3] production being stimulated under conditions of hypoxia.[4] EPO operates by interaction with its receptors on the surface of erythroid precursor cells, in the bone marrow, to promote their differentiation into mature erythrocytes.[5,6] However, when there is progressive destruction of kidney mass, such as in chronic renal failure, an anaemia results due to a decrease in the production of EPO.[7]

In 1977 Miyake *et al*[8] purified EPO to homogeneity from the urine of patients with severe aplastic anaemia but, due to the scarcity of material, only limited amino acid sequence data was available. This led to the isolation of both cDNA and genomic clones of the human EPO (hEPO) gene in 1985.[9,10] Following the cloning and expression of the cDNA, it has been possible to isolate recombinant hEPO (rhEPO) from chinese hamster ovary (CHO) cells in much larger quantities than was previously possible for the naturally occurring hormone. As a result rhEPO has been characterised, and is thus valuable for research and is licensed for clinical use. The primary sequence of rhEPO consists of 166 amino acids (Fig 3.1) and has a molecular mass of 30.4 kDa. It is heavily glycosylated with the oligosaccharide chains comprising 40% of the molecular mass, the deglycosylated protein has a molecular mass of 18.5 kDa.[11-13] The protein has N-linked glycosylation sites at asparagines 24, 38 and 83 and one O-linked glycosylation site at serine 126[12] but it is widely accepted

that only the N-linked oligosaccharides are essential for full biological activity *in vivo*[14-21] with the O-linked oligosaccharide having little influence on either *in vivo* or *in vitro* bioactivity.[16,17,20] The carbohydrate structures of human and rhEPO have been determined.[22,23] There are two disulphide bridges in hEPO at cysteines 29-33 and 7-161 but only the latter has been shown to be essential for its bioactivity.[12,13,24]

```
1                                          10
Ala Pro Pro Arg Leu Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu
         20                  *                      30
Glu Ala Lys Glu Ala Glu Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser
               *       40                                      50
Leu Asn Glu Asn Ile Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp
                                   60
Lys Arg Met Glu Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala
      70                                       80            *
Leu Leu Ser Glu Ala Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser
                  90                                      100
Gln Pro Trp Glu Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu
                                  110
Arg Ser Leu Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile
120                        *                  130
Ser Pro Pro Asp Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp
               140                                     150
Thr Phe Arg Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu
                                  160                      166
         Lys Leu Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg
```

*** Glycosylation sites: Asn N-linked, Ser O-linked**

Fig 3.1 Amino acid sequence of hEPO showing glycosylation sites

Imai[25] and co-workers have found that CHO-derived hEPO is indistinguishable both physiochemically and biologically from urinary hEPO (uhEPO), thus over the last decade, the availability of the recombinant hormone has had a dramatic effect on renal medicine. As a consequence rhEPO is used routinely in the care of patients with chronic kidney failure[26] who are unable to produce sufficient endogenous hEPO. The first clinical trials using the recombinant therapeutic agent were initiated in Seattle in 1985[27] and in 1986 in London.[28] Subsequently in 1988,[29] in the US, only 6 out of 247 patients failed to respond to rhEPO and, in Seattle, patients who had entered maintenance therapy continued to respond to the hormone without developing an

immune response.[30] It would therefore seem that rhEPO is an effective and well tolerated drug.

The major market for rhEPO based therapies is in Japan. This is because brain death is not recognised, thus organ donation and hence transplants are very rare. The result is a rapidly growing population (10,000 p.a) of patients receiving regular and expensive rhEPO dialysis treatment to survive. This situation contrasts with that of the UK where kidney transplants are available and the rhEPO market is small in comparison.[31]

In addition to its clinical use rhEPO has been the subject of much research concerning its structure-function relationships. Both its 3-D structure and those features which contribute to its bioactivity through a receptor binding domain remain to be elucidated. Various structural models have been proposed[24,32] (Fig 3.2) which postulate a substantial α-helical content with a structural motif shared by related molecules such as growth hormone and prolactin. It has been suggested that these molecules constitute a helical cytokine superfamily with common ancestral origin.[24]

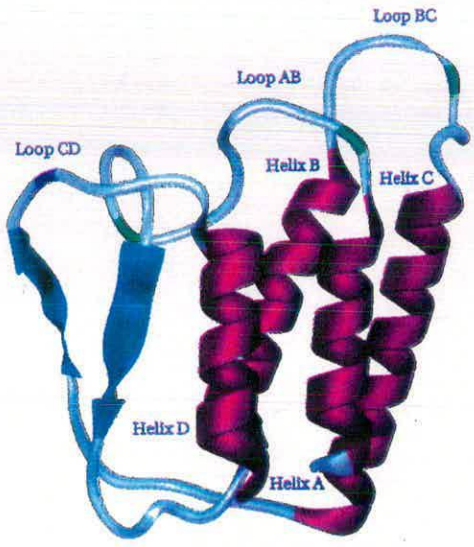


Fig 3.2 Postulated structural model for hEPO; a ribbon diagram of the predicted EPO tertiary structure. The four α-helices are labelled A-D

3.2 The Aims

Due to the interest in larger synthetic targets it was the aim of the research to synthesise a protein >150 amino acids in order to test HOCt as a coupling reagent and other methodology developed in this laboratory for the synthesis of large polypeptides. Proteins in this size range have previously been unattainable by other methodologies with the only total chemical synthesis of a polypeptide of this size (140 a.a) reported by Kent in 1986[33] who synthesised Interleukin-3 utilising Boc SPPS and double coupling each amino acid *via* the symmetrical anhydride (Section 1.1.5.2). Standard coupling with HOCt requires only single couple cycles and no amino acid is effectively 'wasted'. hEPO was chosen as a model as it is a biologically important protein with therapeutic value and was within the limits of the target size.

It was hoped that, if the synthesis and purification of deglycosylated hEPO (dhEPO) were successful this would lead to its crystallisation, NMR structural data and ultimately the synthesis of a glycosylated form of hEPO or similar with biological activity.

The advantages of producing such a protein chemically are outlined below:

(1) Guaranteed homogeneous product with the correct sequence of amino acids (such guarantees cannot always be met by recombinant means due to DNA mismatching).
(2) Protein structure can be modified easily for structure function studies using chemical means (recombinant techniques using site directed mutagenesis are essentially limited to the twenty natural amino acids. In addition these methods are not always adequate due to their time consuming nature and the fact that expression and purification of the altered product is not always straightforward).
(3) NMR probe nuclei can be incorporated at any predetermined single atom site in the protein to aid 3-D structure determination.

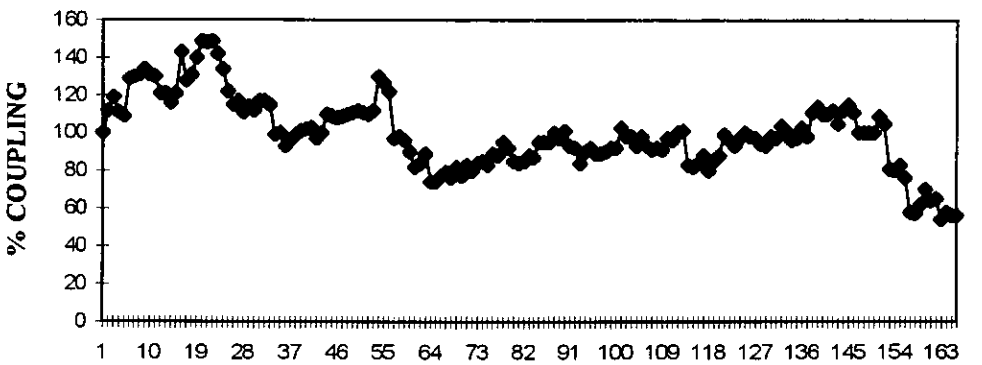
(4) Novel improved forms of the hormone could be synthesised *via* chemical glycosylation with suitable oligosaccharides to give enhanced therapeutic value.

(5) N-linked glycosylation plays a key role in *in vivo* protein activity, thus it is possible to mimic the naturally occurring glycosylation pattern to discover whether any of the glycoforms are more active than others.

3.3 Chemical Synthesis of Deglycosylated Human Erythropoietin

To synthesise and purify a protein of this size would be a challenge for any chemist and initially there was no guarantee that the methodology could stretch to such limits. Therefore it was not surprising that the synthesis had to be repeated several times to obtain more protein. This was due to loss of material during the purification and not because the synthesis had been unsuccessful. In fact the synthesis proceeded extremely well, considering the number of chemical steps involved.

The graph below shows the deprotection profile of the optimised synthesis for synthetic deglycosylated hEPO (dhEPO) (50).



Graph 1: Deprotection profile for dhEPO

Assembly of the dhEPO sequence (Fig 3.1) was carried out on a 0.1mmol scale using Fmoc-Arg(Pmc) functionalised Wang resin and an ABI 430A peptide synthesiser. The amino acids were single coupled as their active HOCt esters with the exception of the last 20 residues which were double coupled. Preliminary results indicated that double coupling from Lys 20 to Ala 1 may improve on yield since the majority of the impurities had a molecular mass of approximately 16 kDa as estimated by SDS PAGE.

The deprotection profile shows coupling percentages of greater than 100% which is possibly due to difficulties with resin swelling. The profile was therefore treated only as an indication that the synthesis had not 'crashed' and that there were no significant 'drops' in the coupling efficiency. Overall the synthesis proceeded extremely well and a final coupling of 56% is very promising for a molecule of this size when compared to other syntheses of smaller polypeptides in the laboratory. An amino acid analysis was carried out at this stage on a few milligrams of the resin bound protein (Table 4) to confirm the success of the synthesis.

Table 4 A.A.A of resin bound dhEPO (50)

Amino acid	Found	Expected
Asx	12.7	12
Thr	10.4	11
Ser	8.07	10
Glx	20.1	19
Pro	9.4	8
Gly	11.5	9
Ala	20.3	19
Cys	3.6	4
Val	10.4	11
Met	0.97	1
Ile	3.7	5
Leu	21.7	23
Tyr	4.4	4
Phe	3.83	4
His	1.8	2
Lys	8.71	8
Arg	12.92	13
Trp	n.d	3

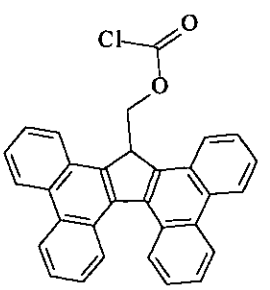
The values obtained from the amino acid analysis were in good agreement with the expected composition, thus with the knowlege that the synthesis had been successful a purification protocol was developed.

The following discussion will mainly concentrate on protein obtained from the optimised synthesis and the final purification protocol adopted, although difficulties encountered in the development of this from preliminary syntheses will be mentioned.

3.4 Purification of Deglycosylated Human Erythropoietin

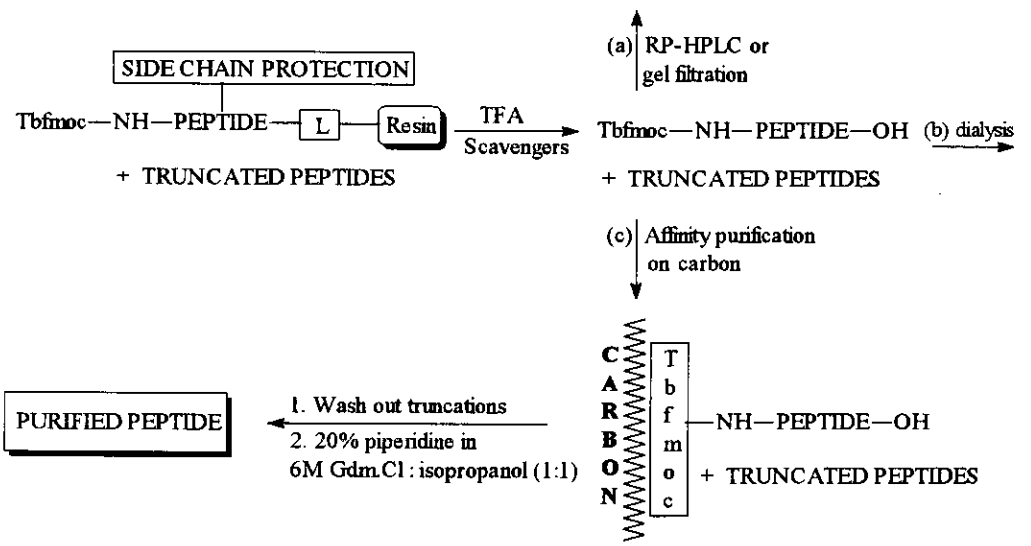
Although the synthesis of dhEPO was initially considered the most difficult stage due to its uncertainty, the main obstacle to the success of the synthesis lay in the subsequent purification. In fact it was expected that the crude product may largely be composed of impurities due to either side-reactions occurring during assembly and cleavage[34], or due to incomplete coupling steps forming truncated peptides (N-terminal acetylated peptides formed on capping). Separation of the target protein from the truncated sequences can be labour intensive, sometimes even impossible due to the physical and chemical similarities between the desired sequence and the shorter truncated impurities. This can often require protracted purification in order to isolate the target protein in pure form,[35] and this in turn can result in a low yield of the isolated product.

3.4.1 Affinity purification of dhEPO using Tbfmoc on charcoal



(51)

Recently, several methods have been reported which allow differentiation between the target sequence and any acetylated truncations. The free N-terminus of the resin bound polypeptide is derivatised with a group, which then allows separation of the desired sequence from the acetylated truncations either by affinity type binding[36,37] or covalent attachment[38] to a solid support. Ramage *et al* have developed a base labile N^{α}-protecting group, tetrabenzo[*a,c,g,i*]fluorenyl-17-methoxycarbonyl (Tbfmoc) (51), for both affinity purification of polypeptides on porous graphitised carbon (PGC)[39,40] and also as a hydrophobic chromatographic probe to simplify the purification of peptides by RP-HPLC or gel filtration.[40] Tbfmoc methodology has been applied to the purification of the chemically synthesised dhEPO (50) using denaturing conditions. The principles involved in this method are outlined below in Scheme 3.1.



Scheme 3.1 Tbfmoc purification

On completion of the synthesis the resin bound product (50) was sonicated in the presence of acetic anhydride/HOBt/DIEA in DCM in order to cap any remaining free amino groups. The N-terminal Fmoc group was then removed and after thoroughly washing the resin, Tbfmoc was introduced by treatment with the chloroformate (51) and DIEA in DCM. The resin bound Tbfmoc labelled dhEPO (52) was thoroughly

washed with DCM and dried to afford 2g of resin bound Tbfmoc-dhEPO. A trial cleavage was carried out on a portion of the resin bound material in order to optimise the time required for cleavage. This was achieved by acidolytic cleavage using a scavenger cocktail containing EDT, thioanisole, water and phenol. Samples of the cleavage mixture were analysed after 3h and every hour afterwards until a total of 6h had passed. HPLC analysis indicated that the optimum time for cleavage was approximately 4.5h since no significant change was observed in the HPLC profile after this time. More importantly it was considered that subjecting the protein to TFA for longer than this may cause shearing of the chain although this was not proven at this stage.

A large scale cleavage was carried out using the remaining 2g of resin. The crude Tbfmoc-labelled protein was isolated, after removal of the resin by filtration followed by rapid concentration of the filtrate *in vacuo*, ether precipitation and lyophilisation affording 914mg of crude Tbfmoc-dhEPO (53). Amino acid analysis was performed on the crude material over a range of hydrolysis times in order to optimise the time required for breakdown of the protein into its consecutive amino acids. The results are summarised in Table 5 below. The results indicate that the optimum time required for complete hydrolysis was approximately 48h (highlighted in bold) as these values gave the best comparison with the expected composition.

Table 5: A.A.A of crude Tbfmoc-dhEPO (53) at various hydrolysis times						
	Found; Hydrolysis time (hours)					
Amino acid	**24**	**36**	**48**	**60**	**72**	**Expected**
Asx	11.6	11.3	**12.7**	10.3	8.3	12
Thr	12.5	10.6	**11.2**	11.1	10.3	11
Ser	11	10.8	**10.7**	6.0	9.8	10
Glx	18.0	18.2	**18.8**	17.0	15.0	19
Pro	6.3	5.3	**7.3**	5.1	6.0	8
Gly	11.0	11.3	**13.5**	9.0	10.3	9
Ala	18.4	17.6	**20.1**	16.3	15.0	19
Cys	5.3	2.26	**3.9**	2.9	3.2	4
Val	8.2	9.5	**10.6**	10.0	7.12	11
Met	0.55	0.57	**1.0**	1.0	1.96	1
Ile	5.1	5.0	**5.0**	5.0	5.2	5
Leu	24.7	21.7	**22.5**	19.7	15.7	23
Tyr	3.9	4.2	**4.8**	2.4	4.0	4
Phe	3.25	3.53	**4.1**	1.7	3.2	4
His	1.04	1.2	**1.5**	1.13	2.7	2
Lys	8.8	8	**8.8**	11.8	?	8
Arg	11.97	13.0	**12.9**	10.2	9.1	13
Trp	n.d	n.d	**n.d**	n.d	n.d	3

As outlined in Scheme 3.1 three options were now available for the first stages of the purification of dhEPO with Tbfmoc and are as follows:

(a) Purification by HPLC or gel filtration using Tbfmoc as a hydrophobic chromatographic probe (UV absorbance at 364nm).

(b) Dialysis

(c) Affinity purification using carbon

In the early stages of the research method (c) was not employed due to the fact that removal of the protein from the carbon bound Tbfmoc group involved strongly basic conditions (10% piperidine) which had previously been shown to result in the formation of polymeric aggregates in the presence of free thiol groups[41,42] (dhEPO has four). Deprotection using either route (a) or (b) can be accomplished at pH 8 where any thiols can be maintained in the reduced state by the addition of Clellands reagent (DTT)[41]. This method of deprotection cannot be applied when utilising affinity

purification as it takes several hours and this would be too long to leave the protein bound to carbon.

Method (a) was chosen at first but attempts at purification by both HPLC and gel filtration failed.

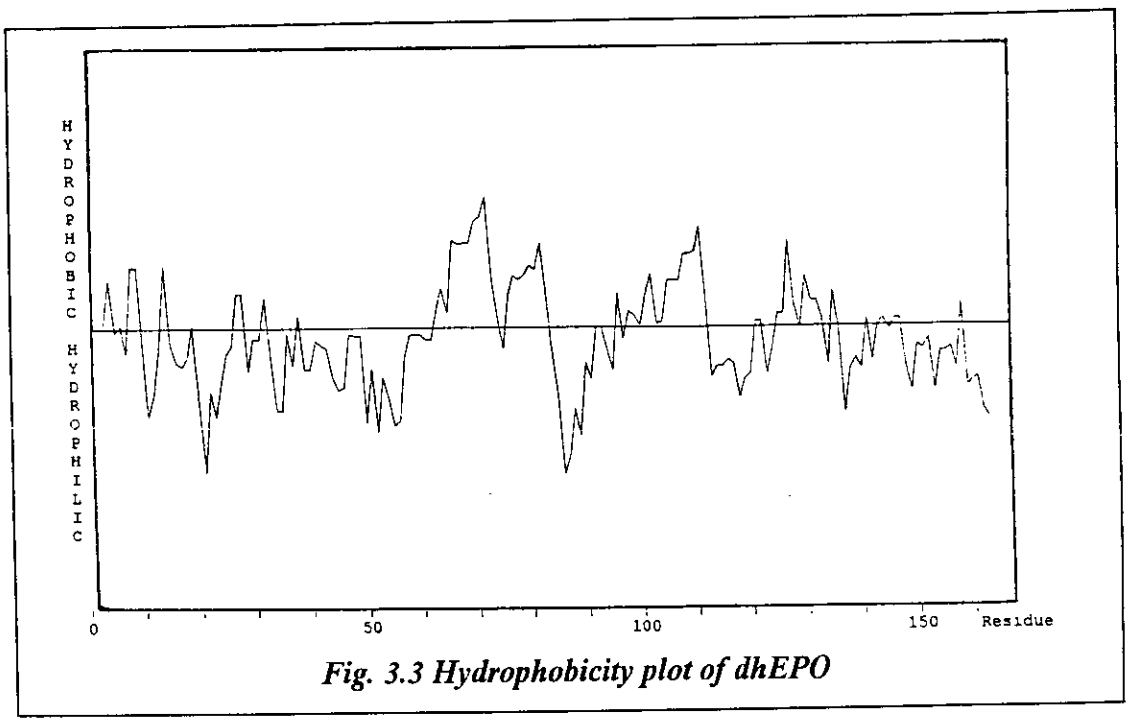


Fig. 3.3 Hydrophobicity plot of dhEPO

The combination of a very hydrophobic protein (Fig 3.3) with a hydrophobic probe resulted in the majority of the tagged protein sticking to the HPLC column; thus virtually no protein was recovered. HPLC was therefore used for analysis purposes only.

Gel filration using Sephadex G-50 monitoring at 226 and 365nm seemed promising at first as some separation was observed but all fractions showed an absorbance at 365nm, indicating the presence of Tbfmoc which, in theory, should only have bound to the dhEPO. Analysis by HPLC of the fractions collected gave identical traces for all, thus there was no resolution and the protein was merely eluting at different stages depending on how it had bound to the gel.

Dialysis of the tagged protein using a molecular weight cut off (MWCO) of 10,000 cleaned the protein to a certain extent but the procedure took several weeks and consumed litres of solvent due to the number of times the solution had to be changed.

Following these results it was proposed that an attempt to purify the material by method (c) should be made. It was suggested that by acidifying the piperidine mixture to pH 8 with addition of DTT immediately after fast release of the protein from Tbfmoc on carbon would prevent polymerisation *via* the cysteine side-chains. If this was successful not only would it be less time consuming than dialysis but most importantly it would show that the methodology was appropriate for large hydrophobic polypeptides containing cysteine residues.

The method was carried out first of all using material from the preliminary syntheses and PGC was used as the adsorption medium. The purification step went extremely well and no polymerisation was observed. Owing to the success of this, the method was applied to material from the optimised synthesis but PGC was replaced by much cheaper charcoal as the adsorption medium.

Prior to adsorption, the charcoal had to be washed thoroughly with the solvents used for the method (10% piperidine/6MGdm.Cl:isopropanol (1:1) (A) and 6MGdm.Cl:isopropanol (1:1) (B)). It is advisable to wash with (A) first then wash copiously with (B) as any trace of piperidine will cause cleavage of Tbfmoc from the protein on adsorption.

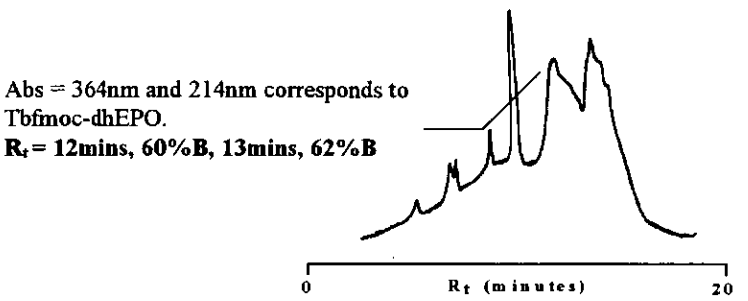


Fig. 3.4 HPLC trace showing crude Tbfmoc-dhEPO (53)

The crude Tbfmoc tagged protein (53) (Fig 3.4) was dissolved in (B) and added to the freshly washed charcoal. Adsorption of Tbfmoc-dhEPO onto charcoal was monitored by HPLC (UV set at 364nm) and was assumed to be complete after 30 minutes as a flat baseline was observed. The charcoal was then washed with (B) and washing was continued until no absorbance was observed by HPLC (UV set at 214nm). Experiments have shown that vortexing the mixture, centrifuging the carbon to a pellet then discarding the supernatant is much more effective than using a carbon column. Once the truncated peptides had been removed the protein was released from Tbfmoc by vortexing the charcoal mixture in (A) for 15 minutes then repeating and finally washing with (B). (B) was combined with (A) and the solution was reduced *in vacuo* to remove isopropanol. The pH of the protein solution was adjusted to 8 and incubated at 37°C after addition of DTT to reduce the cysteine side-chains. Again the purification proceeded well with no polymerisation and a single broad peak with only a few shoulders on its side was observed on HPLC analysis (Fig 3.5) of the Tbfmoc purified protein (54). The shoulders apparent on the side of the main peak suggest the presence of some remaining impurities but this was to be expected considering the size of the protein.

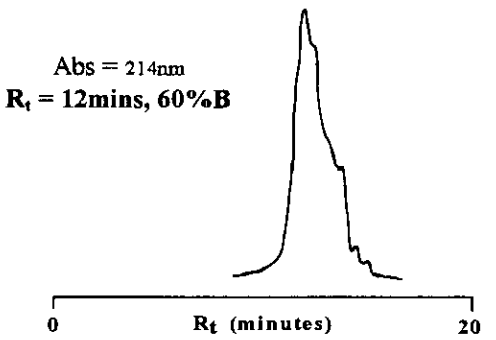


Fig. 3.5 HPLC trace showing Tbfmoc purified dhEPO (54)

The impurities could be either high molecular weight truncated material which has not been removed on washing, perhaps due to binding of the aromatic side-chains to the charcoal or the protein has sheared during cleavage with TFA creating non-acetylated truncations which can then couple to Tbfmoc during loading and thus be carried through in the purification or, most likely, the impurities are deletion peptides .

Deletions are similar to truncated peptides but, unlike truncations, are not acetylated during the capping step. This can happen if the growing peptide chain in some way inhibits the resin from swelling to its full potential, thus preventing all free amino functions from being exposed to the acetic anhydride used for capping. However, these sites can become available for coupling later in the synthesis, again due to a change in swelling properties of the peptide-resin. Thus, as well as incorporating Tbfmoc onto the target sequence, the deletion sequences can also become labelled.

3.4.2 FPLC purification of dhEPO by gel filtration (55)

In order to separate the remaining impurities from the desired protein FPLC gel filtration was applied, adopting the principle of size exclusion (largest molecular weight elutes first). The column was calibrated with known molecular weight standards (Ribonuclease A, MWt 13,700 and Chymotripsinogen A, MWt 25,000) prior to purification to ensure the protein synthesised was of the correct molecular

weight since techniques such as mass spectrometry (MS) and SDS PAGE had proven unsuccessful so far. The protein was injected onto the size exclusion column (Superdex^{TM} 75) and eluted with 6M Gdm.Cl. The procedure followed for MWt determination is described below.

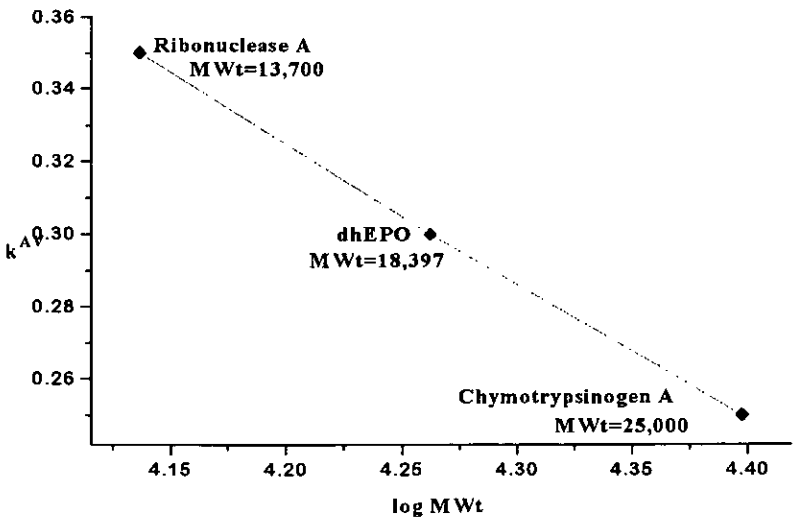
3.4.2.1 Molecular weight determination of dhEPO from FPLC

The technique assumes the same relationship between molecular size and molecular weight for all proteins. A graph can be constructed by plotting the elution volume parameter (k_{AV}) of the standards *versus* the logarithm of their molecular weight.

k_{AV} is calculated using the following equation;

$$k_{AV} = \frac{Ve - Vo}{Vt - Vo}$$

where Ve = elution volume, Vo = column void volume and Vt = column bed volume. Thus, the MWt of the protein synthesised can be determined by calculating its k_{AV} value and, using this, the logarithm of its MWt can be found from the graph (Graph 2).



Graph 2 Plot of logMWt vs k_{AV} to determine MWt of dhEPO fractions eluting from FPLC using two known MWt standards

From the graph the log MWt calculated for dhEPO is 4.2625 which compares well with the expected value of 4.2647. Thus, this is a good indication that the protein synthesised is dhEPO as it is eluting from FPLC at the correct molecular weight with respect to the standards.

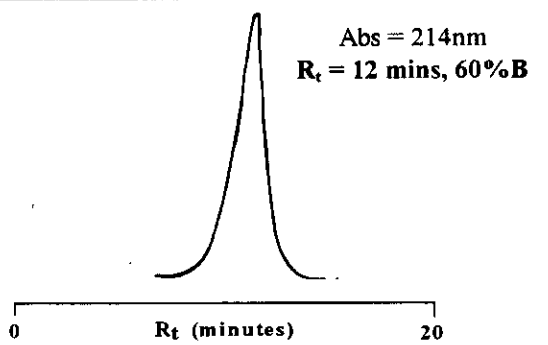


Fig. 3.6 HPLC trace showing FPLC gel filtered dhEPO (55)

The fractions collected were analysed by HPLC and this showed a single broad peak (Fig 3.6) (55) with no shoulders for the resolved protein. Fractions producing identical traces were pooled and concentrated to a small volume. Any unresolved material was recombined, concentrated and reapplied to the column in order to maximise recovery. Protein content was determined by UV analysis[43] to give an approximate yield of 12.6mg of purified protein.

3.5 Characterisation of Purified dhEPO (55)

Amino acid analysis data (Table 6) was obtained for the purified protein and, as before, the found composition was in good agreement with the expected composition.

Table 6 A.A.A of FPLC purified dhEPO (55)		
Amino acid	**Found**	**Expected**
Asx	12.3	12
Thr	10.2	11
Ser	8.5	10
Glx	20.5	19
Pro	8.3	8
Gly	10.4	9
Ala	19.2	19
Cys	0.1	4
Val	12.2	11
Met	1.2	1
Ile	4.4	5
Leu	22.0	23
Tyr	4.6	4
Phe	4.1	4
His	2.7	2
Lys	8.7	8
Arg	12.9	13
Trp	n.d	3

Although the amino acid analysis data and the molecular weight determination from FPLC are very encouraging and a good indication that the correct material has been synthesised, more substantial evidence was required.

3.5.1 Tryptic digest and mass spectrometry data

It has been mentioned previously that MS analysis using MALDI TOF MS failed for the intact protein. However, it was thought that on hydrolysing the protein into smaller fragments using a proteolytic enzyme, MS data could be obtained for these smaller peptides.

Trypsin was the enzyme of choice as it has only two cleavage sites (cuts at the C-terminus of Lys and Arg) which would minimise the number of fragments obtained and, due to its robust nature, should be able to withstand the 3M Gdm.Cl solution used (it was found that dialysis of the protein into low salt, unless the protein solution was very dilute, resulted in its precipitation).

Protein digestion followed by MS analysis of the mixture showed evidence for five main sites of cleavage thus the 3M Gdm.Cl solution had caused some enzyme inhibition as not all possible cuts were observed. However, this was considered an advantage as the molecular mass of some of the predicted fragments would have been too small to detect by MALDI TOF MS. The results are tabulated below.

Table 7 MS data obtained from tryptic digest

	Mass	
Fragment	**Found**	**Expected**
5-54	5754.49 (TFA salt)	5655.80
46-97	5967.87 (K^+ salt)	5931.12
59-103	5478.36 (Na^+ salt)	5453.30
98-140	4435.25	4435.48
111-150	3810.34 (H_2O)	3795.03

From the results obtained there was no evidence for the N-terminal amino acids (1-4) and it was important to know that the synthesis had proceeded to the end in order to confirm that this was not simply a deletion sequence which had been purified, thus other methods had to be employed to analyse this portion.

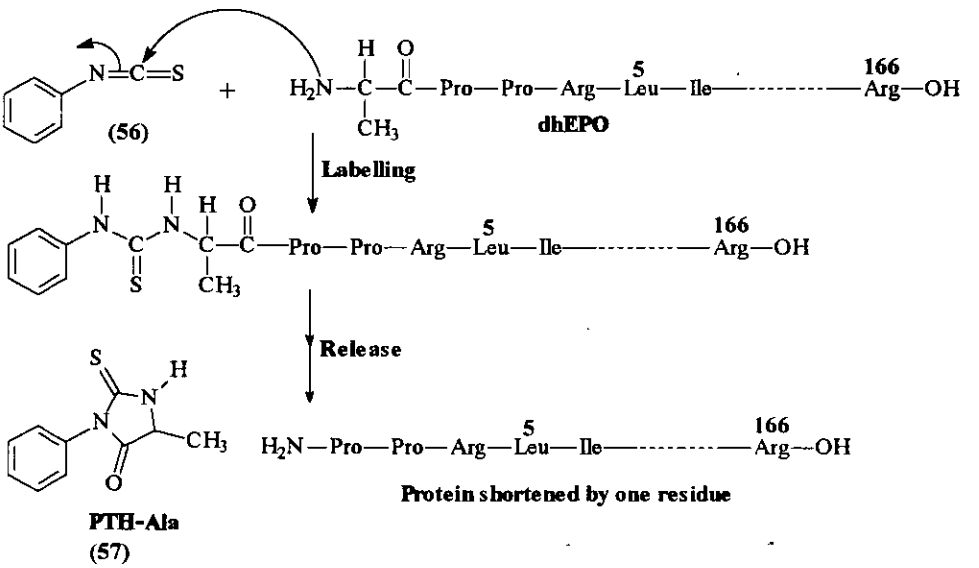
3.5.2 Sequencing analysis[44]

The intact protein was then sequenced (Wellmet) from the N-terminus[44] in an attempt to characterise this part of dhEPO. This was performed by Edman degradation (Scheme 3.2), labelling the N-terminus with phenyl isothiocyanate (56).[45]

Phenyl isothiocyanate (PTH) (56) is attached to the free N-terminus of the purified protein. The labelled amino terminal residue (PTH-Ala) (57) can then be released without hydrolysing the rest of the protein. Hence the amino-terminal residue of the shortened protein (dhEPO 2-166) can be determined in the second round.

A total of 19 rounds of Edman degradation were performed satisfactorily on the purified dhEPO confirming the N-terminal region. It was not possible to sequence

further due to the size of the protein. However combining these results with MS analysis from the tryptic digest and amino acid analysis stongly indicates that the correct material had been synthesised.



Scheme 3.2 Process of Edman degradation using phenyl isothiocyanate

3.5.3 SDS PAGE[46] and western blot analysis[47,48]

It was stated previously that SDS PAGE had been very unsuccessful for dhEPO due to either no band observed on staining or if a band was visible it was badly streaked. Mainly through trial and error gels have been obtained for the purified material and are shown in Fig 3.7 below.

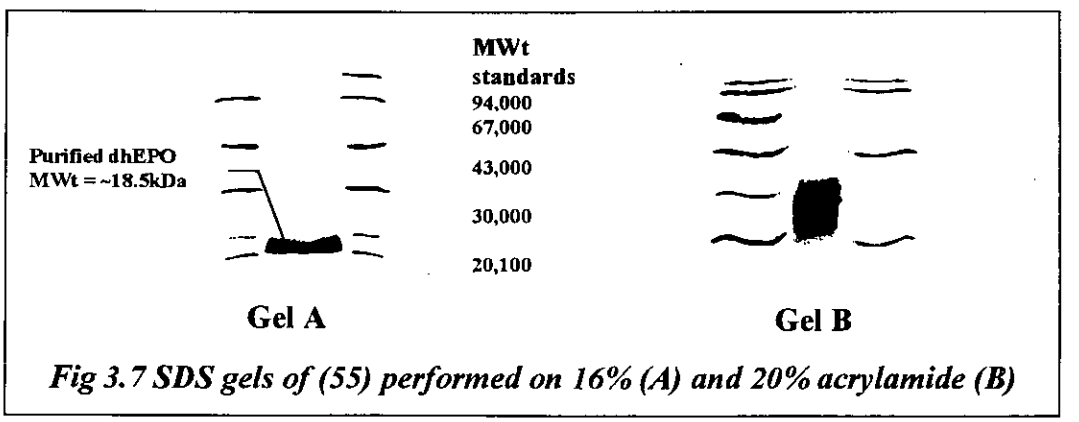


Fig 3.7 SDS gels of (55) performed on 16% (A) and 20% acrylamide (B)

Gel A was performed on 20% acrylamide and B was performed on 16% acrylamide using the method of Laemmli.[46] As can be seen there is still significant smearing but the band is within the correct molecular weight and this has also been confirmed by Western blotting analysis.

A monoclonal mouse anti-hEPO antibody (AE7A5, Genzyme diagnostics, West Mailing, UK) exists and is specific for the first 26 residues of hEPO. Western blotting[47,48] using this antibody was carried out on the purified dhEPO and also on samples from each stage of the purification protocol (Fig.3.8). The bands obtained were compared with a set of molecular weight markers (New England Biolabs) and a sample of His_{10}-rhEPO[47] (20kDa).

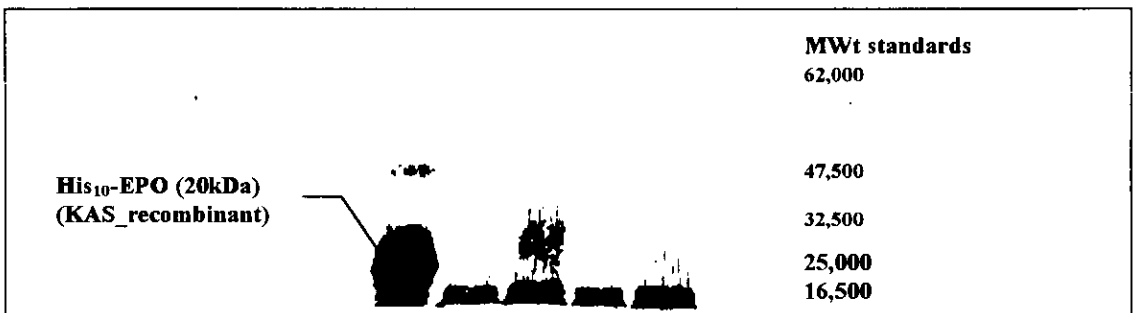


Fig 3.8 Western blot analysis of (53), (54), (55), (58) on PVDF membrane; all samples contained protein of mass ~18kDa to which the anti-EPO antibody binds

The gel shows that the antibody was active for all samples and the bands for each sample are within the correct molecular weight for dhEPO (~18.5kDa)

3.6 Oxidation and Refolding

The normal biologically active form of a protein is the folded state which is the thermodynamically preferred conformation for the protein. This thermodynamically stable three dimensional structure is maintained by relatively weak interatomic forces such as hydrogen bonding, hydrophobic interaction and ionic interaction.[49] Covalent bonds between the sulphur atoms on the cysteine side-chains form intramolecular

disulphide bridges in the polypeptide chain (or in some cases intermolecular disulphide bridges are formed between separate polypeptide chains). These disulphide connections must be formed in the correct way in order to give rise to the desired biological activity.

Generation of disulphide bonds can often be hindered by practical difficulties such as instability or insolubility of the reduced form, the latter was evident during the refolding of dhEPO which has four cysteine residues. It was found that very dilute solutions (1mg in 1l) were required to avoid precipitation of the protein when dialysing from the high salt concentrations (6M Gdm.Cl) required for denaturing into the very low salt concentrations (50mM Tris.HCl) required for folding.

Protein folding was accomplished by using a literature method developed for deglycosylated rhEPO.[24] The fully reduced denatured dhEPO was slowly dialysed against copper sulphate (40μM) in Tris.HCl buffer at pH 8, 5°C. Folding was monitored by HPLC and was considered complete when a sharp peak was observed in place of the broad hump for the reduced, denatured material (Fig. 3.9). This is usually a good indication that folding has occurred but in order to prove that this was folded protein, a sample was removed from dialysis, denatured and as expected the beautifully sharp peak collapsed to a broad hump corresponding to the denatured protein.

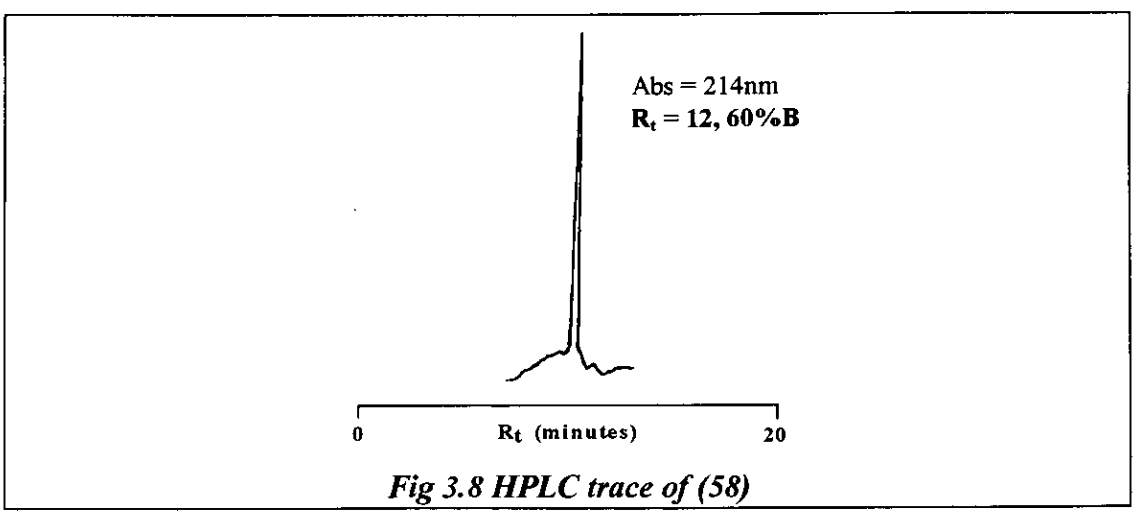


Fig 3.8 HPLC trace of (58)

An attempt has been made to obtain structural analysis on the folded dhEPO (58) but problems were encountered when trying to concentrate the protein into a small enough volume for this. It was found that the protein stuck to the membranes used during concentration and material was lost due to this. dhEPO is known to be a very sticky protein and extremely difficult to handle. Despite this a sample was eventually obtained by using a vaccuum centrifuge to concentrate the solution. However, it was very difficult to acquire an accurate estimate, by UV, of the quantity of protein present and CD analysis does require accurate protein concentrations to determine helical and β-sheet content, thus it was not surprising that the analysis was unsuccessful. Other explanations are possible.

It could be that the protein is unstable unless maintained in very dilute solution, as it has been noted that the commercially available protein is not supplied as an isolated product, but as a dilute solution (2.5μg/ml). Furthermore, in the plasma hEPO has a concentration of 0.1μg/L (range 0.5-2.5μg/L). In addition to this, the synthetic material lacks the oligosaccharide side-chains which have been recognised to be important for the stability of glycoproteins.[50] Moreover hEPO is known to have a very short half life of 7.5h (range 4-12h).

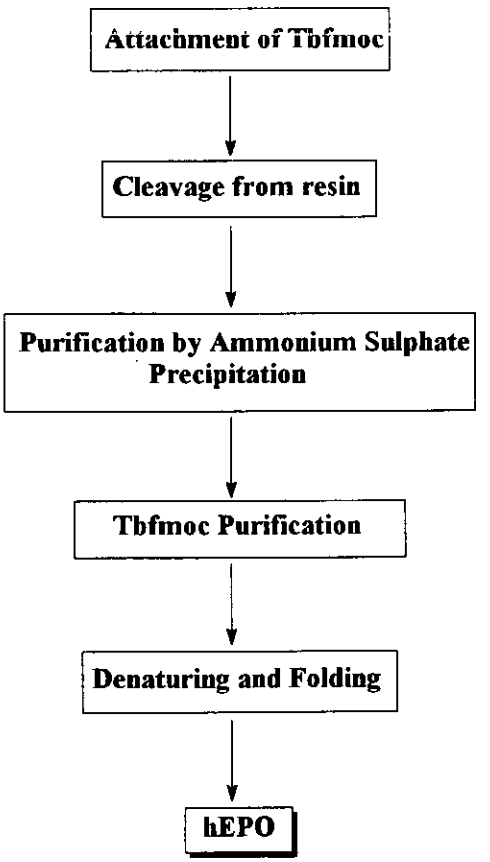
As well as structural analysis it is also important that biological testing should be carried out; even though it is known that the deglycosylated protein does not have *in vivo* activity it does have comparible *in vitro* activity. Unfortunately the success rate in finding a group to set up the required assay has been poor. Thus the research has been stalled for the time being.

3.7 Future Research

Obviously much more work is required in this area as the main objective is to obtain structural analysis (crystallography, NMR etc) on dhEPO. To do this the protein

must be synthesised again, and a new purification protocol from that adopted here may have to be developed in order to obtain a good sample of the protein.

For instance it has been shown that the biological activity of hEPO can be severely reduced if the protein is stored in 6M Gdm.Cl for long periods of time.[51] If the biological activity is affected, then it would be reasonable to assume that this is harmful to the protein structure in general. It would therefore be advantageous not to use FPLC as this would minimise the length of time the protein is in 6M Gdm.Cl (a few hours as opposed to several weeks). Also chromatography can result in loss of valuable material due to irreversible binding of the protein to the functional groups in the column medium. The modified protocol is outlined below.



Scheme 3.3 New purification protocol proposed for synthetic dhEPO

This new protocol (Scheme 3.3) still employs Tbfmoc purification as it proved to be an excellent method of purification and, although it is essentially a chromatographic technique, recovery from charcoal was almost quantitative. Tbfmoc would now be the final step in the purification and would follow precipitation. Precipitation is the purification technique commonly applied by biologists when purifying recombinant proteins. It works on the principal that different polypeptides will precipitate from solution at different salt concentrations (ammonium sulphate is normally used).

Utilising Tbfmoc would also provide a means for easy identification of the target protein by UV during precipitation, and its hydrophobic and bulky nature, should also facilitate this procedure. Immediately reducing the protein after Tbfmoc purification then folding, should produce pure dhEPO ready for crystallisation. If successful in this it would then be of interest to produce a glycoform of the protein.

3.8 References

(1) E. Goldwasser, *Fed. Proc. Fed. Am. Soc. Exp. Biol.*, 1975, **34**, 2285-2292.

(2) S. E. Graber and S. B. Krantz, *Annu. Rev. Med.*, 1978, **29**, 51-66.

(3) W. Fried, *Blood*, 1972, **40**, 671-677.

(4) A. J. Erslev, *Blood*, 1955, **10**, 954-996.

(5) J. L. Spivak, *Annu. Rev. Med.*, 1993, **44**, 243-253.

(6) K. W. Harris, R. A. Mitchell and J. C. Winkelmann, *J. Biol. Chem.*, 1992, **267**, 15205-15209.

(7) A. J. Erslev, J. Wilson and J. Caro, *J. Lab. Clin. Med.*, 1987, **109**, 429-433.

(8) T. Miyake, C. K.-H. Kung and E. Goldwasser, *J. Biol. Chem.*, 1977, **252**, 5558-5564.

(9) F.-K. Lin, S. Suggs, C.-H. Lin, J. K. Browne, R. Smalling, J. C. Egrie, K. K. Chen, G. M. Fox, F. Martin, Z. Stabinsky, S. M. Badrawi, P.-H. Lai and E. Goldwasser, *Proc. Natl. Acad. Sci.*, 1985 **82**, 7580-7584.

(10) K. Jacobs, C. Shoemaker, R. Rudersdorf, S. D. Neill, R. J. Kaufman, A. Mufson, J. Seehra, S. S. Jones, R. Hewick, E. F. Fritsch, M. Kawakita, T. Shmizu and T. Miyake, *Nature*, 1985, **313**, 806-810.

(11) J. M. Davis, T. Arakawa, T. W. Strickland and D. A. Yphantis, *Biochemistry*, 1987, **26**, 2633-2638.

(12) P-H. Lai, R. Everett, F.-F. Wang, T. Arakawa and E. Goldwasser, *J. Biol. Chem.*, 1986, **261**, 3116-3121.

(13) L. O. Narhi, T. Arakawa, K. H. Aoki, R. Elmore, M. F. Rohde, T. Boone and T. W. Strickland, *J. Biol. Chem.*, 1991, **266**, 23022-23026.

(14) K. Yamaguchi, K. Akai, G. Kawanishi, M. Ueda, S. Masuda and R. Sasaski, *J. Biol. Chem.*, 1991, **266**, 20434-20439.

(15) S. Dube, J. W. Fisher, J. S. Powell, *J. Biol. Chem.*, 1988, **263**, 17516-17521.

(16) L. C. Wasley, G. Timony, P. Murtha, J. Stoudemire, A. J. Dorner, J. Caro, M. Kreiger and R. J. Kaufman, *Blood*, 1991, **77**, 2624-2632.

(17) E. Delorme, T. Lorenzini, J. Giffin, F. Martin, F. Jacobsen, T. Boone and S. Elliot, *Biochemistry*, 1992, **31**, 9871-9876.

(18) T. Takeuchi and A. Kobata, *Glycobiology*, 1991, **1**, 337-346.

(19) E. Tsuda, G. Kawanishi, M. Ueda, S. Masuda and R. Sasaki, *Eur. J. Biochem.*, 1990, **188**, 405-411.

(20) M. Higuchi, M. Oh-eda, H. Kuboniwa, K. Tomonoh, Y. Shimonaka and N. Ochi, *J. Biol. Chem.*, 1992, **267**, 7703-7709.

(21) M. N. Fukuda, H. Sasaki, H. Lopez and M. Fukuda, *Blood*, 1989, **73**, 84-89.

(22) H. Sasaki, B. Bothner, A. Dell and M. Fukuda, *J. Biol. Chem.*, 1987, **262**, 12059-12076.

(23) M. Tekeuchi, S. Takasaki, H. Miyazaki, T. Kato, S. Hoshi, N. Kochibe and A. Kobata, *J. Biol. Chem.*, 1988, **263**, 3657-3663.

(24) J-P. Boissel, W. R. Lee, S. R. Presnell, F. E. Cohen and H. F. Bunn, *J. Biol. Chem.*, 1993, **268**, 15983-15993.

(25) N. Imai, A. Kawamura, M. Higuchi, M. Oh-eda, T. Orita, T. Kawaguchi and N. Ochi, *J. Biochem.*, 1990, **107**, 352-359.

(26) G. D. Holt and G. W. Hart, *J. Biol. Chem.*, 1986, **261**, 8049-8057.

(27) J. W. Eschbach, J. C. Egrie, M. R. Downing, J. K. Browne and M. D. Adamson, *New Engl. J. Med.*, 1987, **316**, 73-78.

(28) C. J. Winearls, D. O. Oliver, M. J. Pippard, C. Reid, M. R. Downing and P. M. Cotes, *Lancet*, 1986, **2**, 1175-1177.

(29) J. W. Eschbach and J. W. Adamson, *Kidney Int.*, 1988, **33**, 189A.

(30) D. Swinbanks, *Nature*, 1991, **354**, 4.

(31) S. Narasimhan, J. C. Fried and H. Schacter, *Biochemistry*, 1985, **24**, 1694-1700.

(32) J. F. Bazan, *Immunology Today*, 1990, **11**, 350-354.

(33) I. Clark-Lewis, R. Aebersold, H. Ziltener, J. W. Schrader, L. E. Hood and S. B. H. Kent, *Science*, 1986, **231**, 134-139.

(34) G. B. Fields and R. L. Noble, *Int. J. Pept. Prot. Res.*, 1990, **35**, 161-214.

(35) R. Ramage, J. Green, T. W. Muir, O.M. Ogunjobi, S. Love and K. Shaw, *Biochem. J.*, 1994, **299**, 151-158.

(36) T. J. Lobl, M. R. Deibel Jr and A. W. Yem, *Anal. Biochem.*, 1988, **170**, 502-511.

(37) M. Wilchek, *Adv. Exp. Med. Biol.*, 1974, **42**, 15-31.

(38) S. Funakoshi, H. Fukuda and N. Fujii, *Proc. Natl. Acad. Sci. U.S.A.*, 1991, **88**, 6981-6985.

(39) R. Ramage and G. Raphy, *Tet. Lett.*, 1992, **33**, 385-388.

(40) A. R. Brown, S. L. Irving and R. Ramage, *Tet. Lett.*, 1993, **34**, 7129-7132.

(41) A. R. Brown, *PhD Thesis*, 1995, Edinburgh University.

(42) S. L. Irving, *PhD Thesis*, 1994, Edinburgh University.

(43) H. Edeloch, *Biochemistry*, 1967, **6**, 1948-1954.

(44) A. D Cronshaw, *Wellmet Sequencing*, Edinbiurgh University, 1996.

(45) J. D. Hayes, L. A. Kerr and A. D. Cronshaw, *Biochem. J.*, 1989, **264**, 437-445.

(46) U. K. Laemmel, *Nature*, 1970, **227**, 680-685.

(47) K. Sage, Oxford University

(48) H. Towbin, T. Staehelin and J. Gordon, *Proc. Natl. Acad. Sci. USA*, 1979, **76**, 4350-4354.

(49) T. E. Creighton, WO 86/05809

(50) N. J. Davis and S. L. Flitsch, *Tet. Lett.*, 1991, **32**, 6793-6796.

(51) J. A. Caley, *Part II Thesis*, Oxford University.

4. Convergent Synthesis

4.1 Introduction

The stepwise synthesis of hEPO (chapter 3) and other biologically important proteins[1-10] has demonstrated the scope and utility of stepwise SPPS. These synthetic protocols, however, cannot always be extended to the assembly of large polypeptides due to problems with aggregation and interaction of the growing peptide chain with the resin which is sequence dependant.[11] This is demonstrated by the surprisingly small number of successful protein syntheses reported in the literature, and by the large number of ambiguous syntheses typically yielding poorly characterised products.[12]

An alternative approach to the stepwise construction of long peptide chains is to adopt a convergent strategy in which smaller peptide segments of the protein are individually synthesised and linked together to produce the desired sequence. In principle, convergent synthesis has the advantage that small peptide fragments are straightforward to synthesise and purify, thus giving increased yields. More importantly, the difference in size between product and reactants should allow easy purification of the target molecule.

The main problem with this strategy is ensuring selectivity of the coupling reaction. Thus, the original convergent strategy involved coupling of maximally protected fragments *via* the active ester. However, the low solubility of large protected peptide fragments in organic solvents has meant that few laboratories have used this approach successfully.

4.2 Coupling of Minimally Protected Peptide Fragments

In order to overcome the solubility problems encountered with maximally protected peptide fragments, chemical coupling of minimally protected peptide segments has been explored by some groups.[13-17]

4.2.1 Chemical Ligation

In 1992 Scholzner and Kent[13] reported the synthesis of a fully biologically active analogue of HIV-1 protease containing a thioester pseudo-peptide bond at the ligation site (^{51}Gly-^{52}Gly). The N-terminal fragment contained a thioacid function whilst the C-terminus incorporated a bromo acetyl function. The two segments were ligated with no side-chain protection, even on cysteine residues.

A further variation of this work was reported in 1994 for the synthesis of IL-8[14] (Fig 4.1) (62). The ligation reaction this time involved nucleophilic attack by the thiol group of an N-terminal cysteine residue on one fragment (60), with the C-terminus of the other fragment which was functionalised with a thioester (59). Displacement of the thioester formed the intermediate (61) which then spontaneously rearranged to give the natural amide bond of the product protein (62) and regenerated the thiol of the cysteine residue; again no side-chain protection was required.

More recently the chemically robust thioether linkage has been introduced into the field as a more generally applicable approach to chemical ligation. This again was applied to the synthesis of HIV-1 protease.[15]

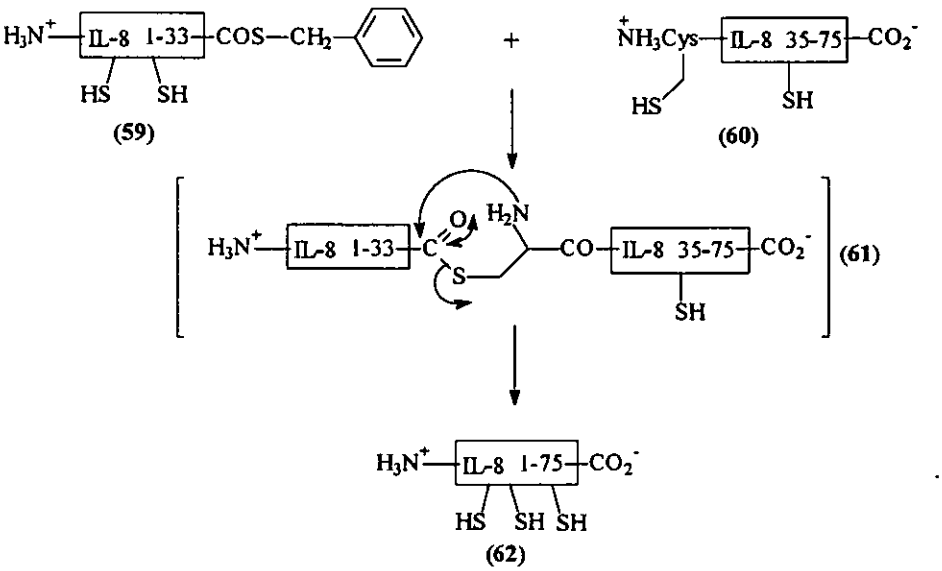


Fig 4.1 Synthesis of IL-8: chemical ligation via Cys side-chain

Chemical ligation of peptide fragments would therefore seem an attractive approach to the synthesis of proteins. However, the methods have certain limitations. The thioester surrogate amide bond is unstable at elevated pH and while the thioether is an attractive alternative both result in the introduction of an altered backbone structure into the protein. Whilst these factors were unimportant in the case of HIV-1 protease, the general application of these methods remains to be demonstrated in other systems.

Union of fragments *via* the formation of the native amide bond still remains the most obvious choice and, although this has been achieved using chemical ligation, as in the case of IL-8, the method is restricted to union at cysteine residues only. Thus, a more general approach to fragment coupling would be favourable.

4.2.2 Enzyme catalysed amide bond formation

Normally, enzymes catalyse the degradation of proteins and peptides by hydrolytic cleavage of the peptide bond. However, developments in peptide synthesis have shown that proteolytic enzymes can be used to catalyse the formation of amide bonds by reversing the hydrolysis reaction.[16-18]

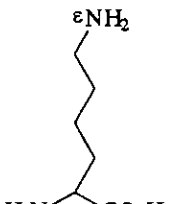
One of the major enzymatically catalysed systems involves displacement of the equilibrium in favour of amide bond formation. To achieve this the product peptide from the bond forming reaction must be removed from the system either by precipitation[19] or by use of a compound that forms a specific complex with the product removing it from the equilibrium,[20] e.g. monoclonal antibody. Biphasic systems are also successful in removing the target peptide from solution. In this case the peptide passes out of the aqueous phase into an organic solvent as it is formed which, once again, perturbs the system, shifting the balance of the equilibrium in a favourable direction.[21]

The advantages of enzyme catalysed coupling are that it is fast and racemisation free and requires only minimally protected fragments.

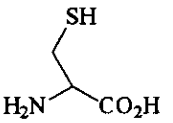
4.2.3 Amide bond formation utilising the azide method

The azide method for the formation of the peptide bond was introduced by Curtius over seventy years ago[22] and, to date, is one of the most successful amide bond segment condensation methods.

Azide couplings are advantageous in that the starting materials are easily prepared, it leads to the least racemisation during coupling and once the azide has been selectively introduced at the C-terminus, it is only necessary to protect the side-chain amino function of lysine (63) and the sulphydryl function of cysteine (64).[23]



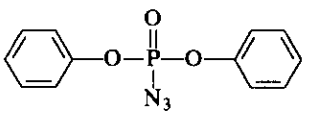
(63)



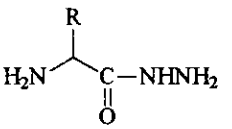
(64)

The azide method has been applied with success in the synthesis of ribonuclease A,[24] secretin III[25] and derivatives of ribonuclease T[26] to name but a few.

Generally azides can be generated in two ways, either directly using diphenylphosphorazidate[27] (65) and a carboxylic acid or *via* the hydrazide (66). Formation *via* the hydrazide is the preferred route as the former method requires that all side-chain carboxylic functions be protected.



(65)



(66)

Hydrazides can be incorporated into the C-terminus of peptides by treating a C-terminal alkyl ester with hydrazine. This lends itself well to SPPS since the peptide is held to the resin through an alkyl ester bond (67). Hydrazinolysis of the peptide from the resin followed by acidolysis to remove the side-chain protection yields the peptide hydrazide (68) (Fig 4.2).

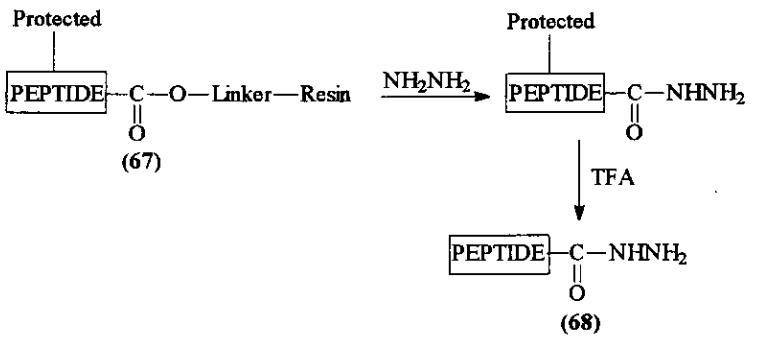


Fig 4.2 Formation of a peptide hydrazide via hydrazinolysis of an ester

Although some hydrazinolysis can occur at the *t*-butyl protected Asp and Glu[28,29] residues this can be minimised by using only rapid contact with dilute hydrazine (typically 2% aqueous hydrazine for 2-3 minutes).[30] However, there are some significant disadvantages with the method as hydrazine can deprotect Fmoc from the N-terminus[31] and there is also evidence that side reactions with the guanidine function of Arg can occur to give ornithine or aminoguanidine.[32]

The development of hydrazine resin linkers such as (18) (Section 1.4.2) has eliminated the need for hydrazinolysis.

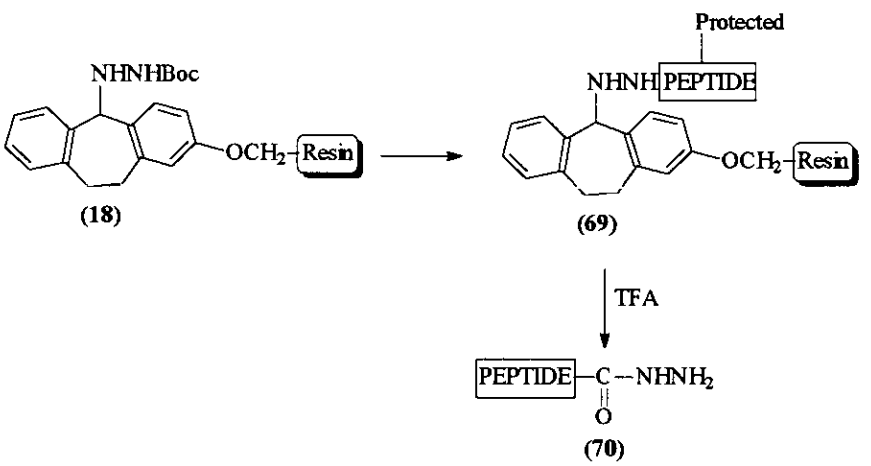


Fig 4.3 Formation of a peptide hydrazide via a linker

This technique selectively incorporates the hydrazide to the C-terminal carboxyl of the protected peptide on the solid support (69) and acidolysis on completion of the synthesis then yields the peptide hydrazide (70).

The subsequent conversion of the hydrazide to the azide has been traditionally carried out using the Curtius procedure[33] which utilises aqueous acid and sodium nitrite. However, the Honzl-Rudinger method[34] (Fig 4.4) is far superior.

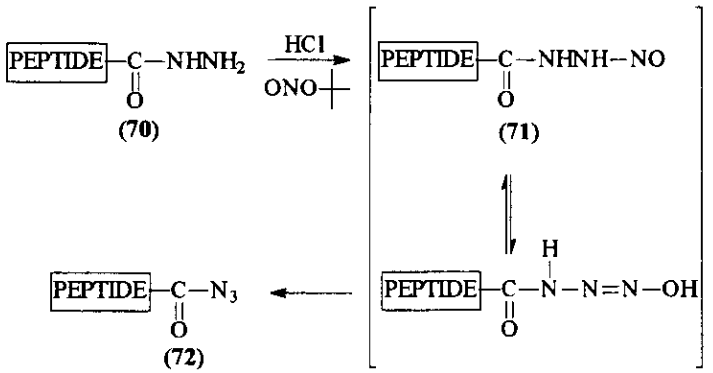


Fig 4.4 Azide formation via the Honzl-Rudinger method

Reacting the peptide hydrazide (70) with an alkyl nitrite such as *t*-butyl nitrite generates an N-nitroso intermediate (71) which, on loss of water, forms the corresponding peptide azide (72). However, the azide (72) itself is prone to Curtius rearrangement to give the isocyanate (73), which can react with the amine component to form a urea derivative (74) (Fig 4.5). This can be suppressed by maintaining the temperature at -10°C and by immediate addition of the amine component once azide formation is complete.

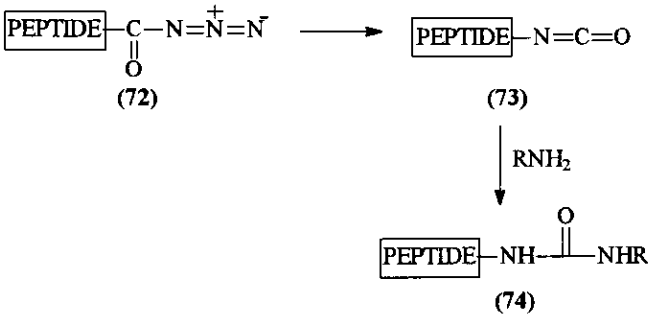
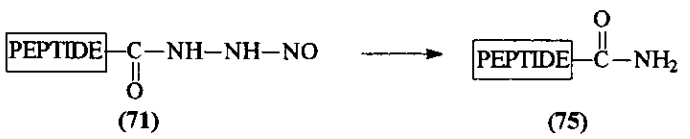


Fig 4.5 Urea formation via Curtius rearrangement of the azide

Another notable side-reaction is amide formation, occurring when the N-nitroso intermediate (71) loses N_2O to form the unreactive amide (75) (Fig 4.6).



(71) (75)

Fig 4.6 Decomposition of the N-nitroso intermediate to an amide

Despite these side reactions, the azide procedure is a very powerful method for fragment coupling using minimally protected fragments. One major problem exists, however, namely combining Fmoc SPPS with azide condensation. The combination requires protection for the strongly nucleophilic N^{ε} function of lysine and the sulphydryl function of cysteine, both throughout the synthesis of the fragment itself and during the condensation of the fragments to form the desired product. This is a difficult problem as the protecting groups must not only be orthogonal to the deprotection conditions for Fmoc but must also be stable to acid. In addition to these stringent requirements the group should preferrably be of a nature that would aid solubility of the fragment and be cleavable under mild conditions.

4.3 References

(1) B. Gutte and R. B. Merrifield, *J. Biol. Chem.*, 1971, **246**, 1922-1941.

(2) I. Clark-Lewis, R. Abersold, H. Ziltener, J. W. Schrader, L. E. Hood and S. B. H. Kent, *Science*, 1986, **231**, 134-139.

(3) C. H. Li, D. Yamashiro, D. Gospodarowicz, S. L. Kaplan and G. van Vliet, *Proc. Natl. Acad. Sci. USA*, 1983, **80**, 2216-2220.

(4) W. F. Heath and R. B. Merrifield, *Proc. Natl. Acad. Sci. USA*, 1986, **83**, 6367-6371.

(5) I. Clark-Lewis, B. Moser, A. Walz, M. Baggiolini, G. J. Scott and R. Abersold, *Biochemistry*, 1991, **30**, 3128-3184.

(6) J. Schneider and S. B. H. Kent, *Cell*, 1988, **54**, 363-368.

(7) R. F. Nutt, S. F. Brady, P. L. Darke, T. M. Ciccarone, D. Colton, E. M. Nutt, J. A. Rodkey, C. D. Bennett, L. H. Waxman, I. S. Sigal, P. S. Anderson and A. F. Veber, *Proc. Natl. Acad. Sci. USA*, 1988, **85**, 7129-7133.

(8) R. Ramage, J. Green and O. M. Ojunjobi, *Tet. Lett.*, 1987, **30**, 2149-2153.

(9) M. Kohmura, N. Nio and Y. Ariyoshi, *J. Biol. Chem.*, 1991, **55**, 539-545.

(10) G. H. Morton, *Personal Communication*, Edinburgh University.

(11) L. Jiang, *Personal Communication*, Edinburgh University.

(12) T. W. Muir and S. B. H. Kent, *Biotechnology*, 1993, **4**, 420-427.

(13) M. Scholzner and S. B. H. Kent, *Science*, 1992, **256**, 221-225.

(14) P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. H. Kent, *Science*, 1994, **266**, 776-779.

(15) D. R. Englebretsen, B. G. Garnham, D. A. Bergman and P. F. Alewood, *Tet. Lett.*, 1995, **36**, 8871-8874.

(16) G. A. Homandberg and M. Laskowski Jr., *Biochemistry*, 1979, **18**, 586-592.

(17) L. Graf and C. H. Li, *Biochemistry*, 1981, **78**, 6135-6138.

(18) P. Clapes, J-L. Torres and P. Adlercreutz, *Bioorg. and Med. Chem.*, **3**, 245-255.

(19) H. D. Jakubke, P. Kuhl and A. Konnecke, *Angew. Chem. Int. Ed. Engl.*, 1985, **24**, 85-93.

(20) G. A. Homandberg, A. Komiriya and I. M. Chaiken, *Biochemistry*, 1982, **21**, 3385-3389.9

(21) G. A. Homandberg, J. A. Mattis and M. Jaskowski Jr., *Biochemistry.*, 1978, **17**, 5220-5227.

(22) Y. S. Klausner and M. Bodanzky, *Synthesis*, 1974, **8**, 549-559.

(23) R. Hirschmann *et al*, *J. Am. Chem. Soc.*, 1966, **91**, 502-507.

(24) N. Fujii and H. Yajima, *J. C. S. Perkin I*, 1981, 789-841.

(25) M. A. Ondetti, V. L. Narayanan, M. Saltza, J. T. Sheehan, E. F. Sabo and M. Bodanzky, *J. Am. Chem. Soc.*, 1968, **90**, 4711-4715.

(26) J. Beacham, G. Dupuis, F. M. Finn, H. T. Storey, C. Yanaihara, N. Yanaihara and K. Hofmann, *J. Am. Chem. Soc.*, 1971, **93**, 5526-5539.

(27) T. Shioiri, K. Ninomiya and S. Yamada, *J. Am. Chem. Soc.*, 1972, **94**, 6203-6205.

(28) E. Scoffne, F. Marchiori, A. M. Tamburro and R. Rocchi, *Gazz. Chim. Ital.*, 1964, **94**, 695-709.

(29) B. Weinstein, R. M. Cook and J. T. Hefferman, *Int. J. Pept. Prot. Res.*, 1976, **8**, 21-25.

(30) J. Meienhofer, E. Schnabel and D. Gillessen, *Liebigs Ann. Chem.*, 1963, **667**, 164.

(31) L. A. Carpino and G. Y. Han, *J. Org. Chem.*, 1972, **37**, 3404-3409.

(32) W. Danho and C. H. Li, *Int. J. Pept. Prot. Res.*, 1971, **3**, 81-92.

(33) R. A. Boissonas, S. Guttmann, P. A. Jaquenord and J. P. Waller, 1955, *Helv. Chim. Acta.*, **38**, 1491-1501.

(34) P. Seiber *et al*, *Helv. Chim. Acta.*, 1970, **53**, 2135

5. N^{ε}-Lysine Protection for use in Fragment Coupling

5.1 Introduction

As mentioned in Chapter 4 when azide fragment coupling (FC) is carried out on components containing lysine and cysteine residues, it is necessary to temporarily mask the chemical reactivity of their side-chains to prevent chemical coupling reactions at these sites. In addition to this it is also necessary to protect the N^{α}-terminal amino group of the azide component preventing it from reacting with itself.

The Fmoc N^{α} protecting group used in SPPS is stable to the acidic conditions required for resin peptide cleavage, therefore if it is retained on the N-terminus of the hydrazide component after synthesis it can be used as an effective blocking group for the N^{α} function during FC.

The sulphhydryl function of cysteine can be protected using the acetamidomethyl (Acm) group as it is stable to the acidic and basic conditions required for Fmoc methodology and is also retained during azide FC.

Lysine must be protected either throughout synthesis with a group stable to both acid and base, or be reprotected following cleavage from the resin. The latter is achieved by treatment with a capping agent which temporarily blocks all exposed amino functions during FC. Lysine, usually protected with the acid labile Boc group, can be reprotected with Boc after cleavage *via* its azide[1] or symmetrical anhydride,[2] however, Boc has been found to cleave under the acidic conditions of the Honzl Rudinger procedure (Chapter 4).

Other groups, which are introduced onto the lysine side-chain post cleavage, have been investigated[1,3-8] but all have major drawbacks in that incomplete capping of the N^{ε} function takes place or capping of functionalities other than amines occurs. In addition, removal of these protecting groups can be incomplete or damaging to the final peptide product.

Thus, it would be highly desirable to develop an N^{ε} protecting group which could be introduced, with lysine, as the peptide chain is being assembled. This group would have to be stable to all the conditions mentioned above and easily removed so as not to cause damage to the final product. Several such groups have been proposed using a variety of orthogonal deprotection strategies[9-16] but, despite this, no ideal derivative has been discovered.

5.2 Penicillin Acylase Cleaved N^{ε} Protecting Groups

Penicillin acylase (E.C. 3.5.1.11) which normally catalyses the hydrolysis of penicillin G (76) into 6-amino penicillinic acid (77) and phenyl acetic acid (78)[17] (Fig 5.1) has also been shown to cleave phenylacetyl (Phenac) derivatives.

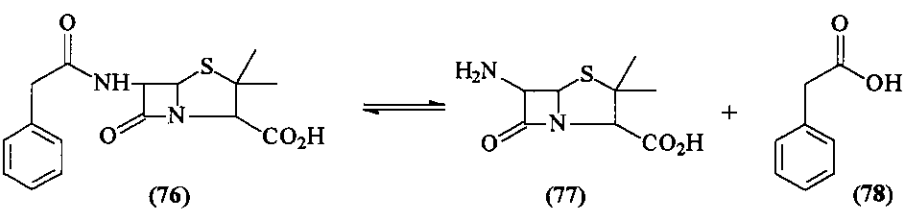
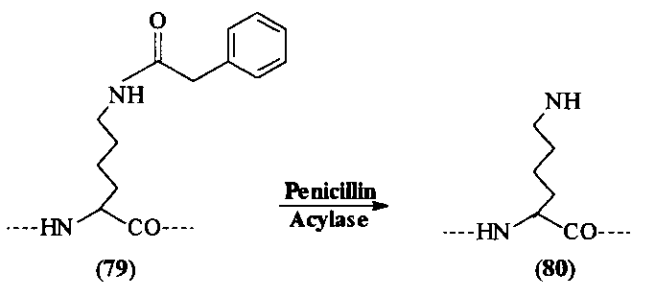


Fig 5.1 The hydrolysis of penicillin G with penicillin acylase

Its role as a selective hydrolysis reagent for N^{ε}-Phenac lysine was first demonstrated by Brtnik *et al* in the solution synthesis of vasopressin.[18] On completion of the synthesis, the Phenac lysyl peptide (79) was simply deprotected by incubating with the enzyme in solution to give the lysyl derivative (80) (Fig 5.2).



(79) (80)

Fig 5.2 Hydrolysis of a phenylacetyl lysine residue with penicillin acylase

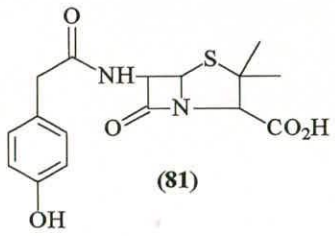
The potential advantages of such a protecting group are obvious. The amide bond formed should be as stable as the rest of the peptide bonds in the molecule and its resistance to acid and base is almost guaranteed.

5.2.1 N^{ε}-Phenylacetyl-lysine and other derivatives

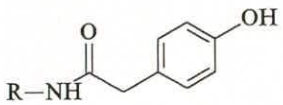
The Phenac protecting group has been found successful in FC reactions performed in this laboratory.[20] However, it was found that a typical deprotection required several days utilising conditions (40-50°C, pH 7.1-8.5) that are not ideal for many peptides, especially those which are readily hydrolised or oxidised.

One of the major problems associated with N^{ε} lysyl capping is the loss of protonation sites on the peptide and hence decreased solubility. A more hydrophilic protecting group would therefore be highly desirable, as would a group whose electronic properties made it more susceptable to cleavage.

Several workers have carried out research attempting to discover different substrates for penicillin acylase[21] and although this work was mainly directed towards penicillin chemistry, an improved substrate could be a potential protecting group. It was discovered that 4-hydroxybenzylpenicillin (81) was the only substrate of those tried which cleaved faster than penicillin G. The corresponding N^{ε} protecting group would be the 4-hydroxyphenylacetyl (Hopa) group (82).



(81)



(82)

This increased rate of hydrolysis can be rationally explained by modelling substrates within the active site of penicillin acylase, the crystal structure of which has recently been reported.[22]

The benzyl moiety is held in the active site by a hydrophobic cleft formed by three phenylalanine residues. At the front of this cleft is a β1 serine residue which is responsible for the hydrolysis of the amide bond. However at the back of this cleft is a second serine which could hydrogen bond to a function at the 4 position, drawing it quickly into the active site. The substrate is then held there until hydrolysis is complete. Fig 5.3 shows a molecular model of the phenylacetyl derivative in the active site. As can be seen, modifications other than those at the 4 position would interfere with the hydrophobic phenylalanine residues causing a worse fit into the cleft and a slower rate of hydrolysis.

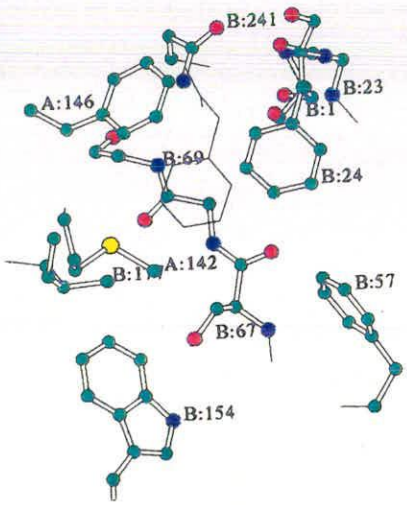
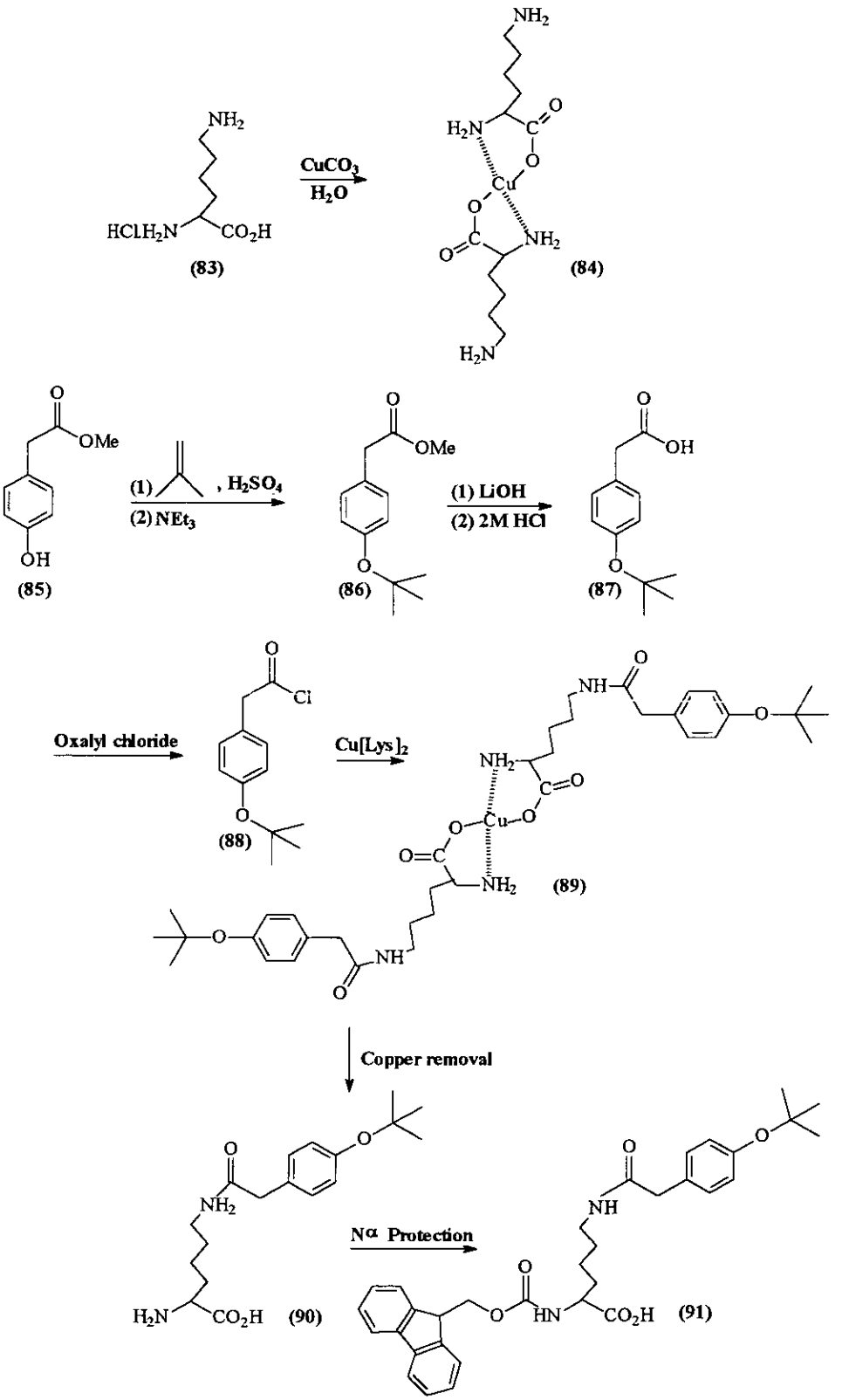


Fig 5.3 Molecular model of the active site of penicillin acylase hydrolysing the phenylacetyl substrate

5.2.1.1 Synthesis of N^{α}-Fmoc-4-*t*-butoxyphenylacetyl-lysine: initial observations

In order to incorporate the Hopa lysine residue into a peptide the hydroxy function of the protecting group must itself be protected during synthesis. Normally such functions are protected as the acid labile *t*-butyl ether therefore Hopa was protected in this way. The amino acid derivative introduced onto the synthesiser would then be N^{α}-Fmoc-4-*t*-butoxyphenylacetyl-lysine (Fmoc-Lys-Bupa) (91), cleavage of the peptide from the resin would regenerate the Hopa derivative which could then be cleaved after FC using the enzyme acylase.

Pallin[20] attempted to synthesise (91) (Scheme 5.1) using the same route as that devised for the Phenac derivative but the synthesis proved to be very difficult and problems arose on removal of copper from the copper-Lys(Bupa) complex (89) with EDTA. However, it was found that the complex could be broken up using sulphur containing compounds and that, of the reagents tried, ammonium sulphate was the most successful. Fmoc was attached without any problems but the product could not be crystallised or obtained in purity. Despite this a small pentapeptide, H-Phe-Gly-Lys-Ala-Gly-OH, with lysine Hopa protected was prepared and tested against the same sequence with lysine Phenac protected. Two exciting results were found: not only did the Hopa group give increased solubility over Phenac, but it cleaved almost four times as fast as the Phenac group. As this appeared to be such a promising protecting group for FC it was desirable to optimise the synthesis to improve on yield and purity.



Scheme 5.1 Synthesis of Fmoc-Lys-Bupa

5.2.1.2 Investigation into the synthesis of N^{α}-Fmoc-4-*t*-butoxyphenylacetyl-lysine

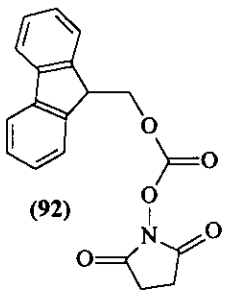
Pallin's synthesis was repeated first of all, but, it was found after a number of attempts that removal of copper from (89) using ammonium sulphide was very difficult and the synthesis was not reproducible by this route. Therefore a different approach was tried. Attempts were made to couple Bupa directly to the N^{ε} amino function of N^{α}-Fmoc-lysine using either the acid chloride or various activated esters, but all proved to be impossible. Attempts to couple to the Z-lysine derivative were also fruitless (Table 8).

Table 8 Attempted syntheses of Fmoc-Lys-Bupa

	Y = Fmoc or Z		
Bupa Activation	**Y-Lys**	**Y-Lys-Tos**	**Y-Lys-TMS**
Acid chloride	x	x	x
HOCt ester	x	x	x
HOBt ester	x	x	x
Succinamide ester	x	x	x

At this point the copper route was attempted again. The acid chloride (88) was used as the active species and used immediately, adding it dropwise to a solution of the lysine copper complex (84) in water while maintaining the pH of the reaction at 10. A blue precipitate formed which corresponded to the copper-lysine(Bupa) complex (89).

As sulphur compounds had proven to be unproductive at copper removal in previous experiments the disodium salt of EDTA was tried instead. It was found that using high dilution and stoichiometric amounts of EDTA the desired product (90) could be obtained in 48% yield. The N^{α} protecting group (Fmoc) was then attached to the N-terminus using Fmoc-ONSu (92) to give the pure amino acid derivative (91), but only a very poor yield (2%) was obtained.



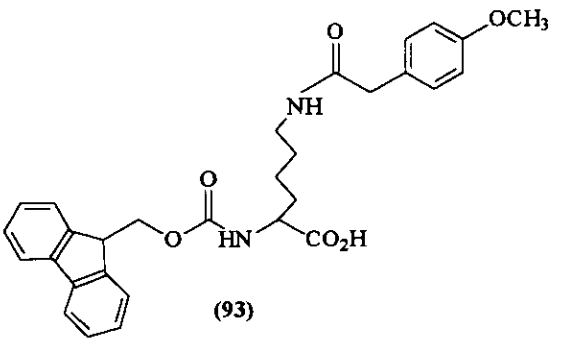
(92)

Several attempts were made to improve on this but each failed and problems started to arise during incorporation of Bupa into the copper-lysine complex (89). Due to the random variation in results found in this research, in Pallin's work and the irreproducibility of the syntheses, it is thought that Bupa becomes unstable when activated due to a combination of the inductive effect and the mesomeric effect inherent in the compound.

When Bupa is activated as the acid chloride there is a strong pull of electrons away from the *t*-butyl group which can leave forming an extremely stable carbocation. The remainder of the molecule is then stabilised due to the mesomeric effect and as a result of this resonance stabilisation can survive.

5.2.1.3 Synthesis and studies on N^{α}-Fmoc-4-methoxyphenylacetyl-lysine

It was evident from the number of problems encountered with Bupa that a different protecting group was required. The *p*-methoxyphenylacetyl (Mepa) derivative (93) was chosen as it is stable when activated and molecular modelling studies have shown that it is structurally compatible with the active site. Thus Mepa should hydrolyse as fast as Hopa since the oxygen at the para position should still be available for hydrogen bonding with the serine residue in the active site.



(93)

Fmoc-Lys(Mepa) (93) was synthesised using the procedure developed for Fmoc-Lys(Bupa) (Scheme 5.1). The synthesis proceeded without any difficulty and the final product was obtained in approximately 60% yield. The group's solubility when incorporated into a peptide and its relative rate of hydrolysis were investigated and compared with results obtained by Pallin for the Phenac and Hopa derivatives. The same model peptide (H-Phe-Gly-Lys-Ala-Gly-OH) as used by Pallin was synthesised with lysine Mepa protected. Unfortunately the results were very disappointing. Not only did the Mepa protecting group render the peptide even more hydrophobic than the Phenac derivative (Fig 5.3) but it had a slower rate of hydrolysis (191h as opposed to 51h). This would suggest that the largest group tolerated at the para position of phenylacetyl in the active site is the hydroxyl as even the methyl group seems to hinder access of the the serine residue to hydrogen bond to the oxygen.

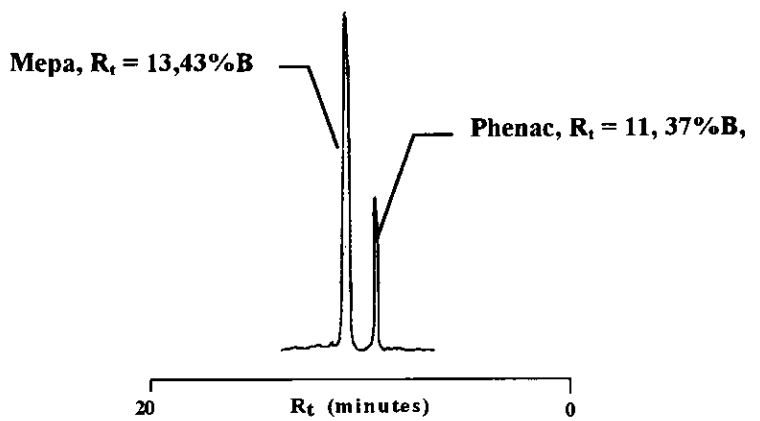
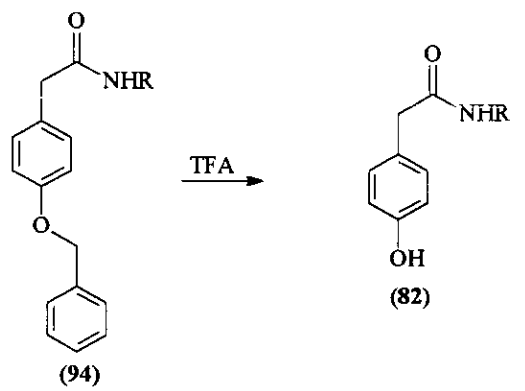


Fig 5.3 HPLC trace illustrating differnce in R_t between Mepa and Phenac peptides

Following these disappointing results a method was found in the literature which reported cleavage of the methyl ether of tyrosine to the hydroxy function using a thioanisole-TFMSA acid system.[23] It was thought this method could be applied to the Mepa protected pentapeptide in order to generate the Hopa derivative which as explained above is more desirable. The peptide was stirred in the same cocktail but only partial cleavage was observed even after 2.5h. The Mepa protecting group was abandoned at this point due to the long time required to cleave the methyl ether as this could be destructive to large polypeptide chains.

5.3 Recent Developments



(94) (82)

Fig 5.4 TFA cleavage of the benzyl ether to generate Hopa

Further investigation into this area by Draffan[24] has led to the development of a suitable protecting group for lysine. It was found that the Hopa derivative could be incorporated successfully onto the side-chain of lysine, using Scheme 5.1, if the hydroxy function was protected *via* a benzyl ether (94), (Fig 5.4). Cleavage of the benzyl ether to regenerate Hopa is compatible to the conditions used in cleaving the peptide from the resin (TFA). Chapter 7 discusses the utility of the group in azide FC.

5.4 References

(1) A. R. Rees and R. T. Offart, *Biochem. J.*, 1976, **159**, 467-486.

(2) L. Moroder, A. Hallet, E. Wunsch, O. Keller and G. Wersin, *Physiol. Chem.*, 1976, **357**, 1651-1653.

(3) P. G. J. Butler J. Harris, B. S. Hartley and R. Lebermann, *Biochem. J*, 1969, **112**, 679-689.

(4) H. B. F. Dixon and R. Perham, *Biochem. J.*, 1968, **109**, 312-314.

(5) L. Gibbons and R. Perham, *Biochem. J.*, 1968, **116**, 843-849.

(6) R. Geiger and D. Langner, "The Peptides 1974" (Wolman Y., Ed.), 1974, pp 159-163.

(7) G. I. Tesser and I. C. Balvert, *Int. J. Pept. Prot. Res.*, 1975, **7**, 295-305.

(8) M. J. Hunter and M. Ludwig, *J. Am. Chem. Soc.*, 1962, **84**, 3491-3504.

(9) R. Hirschmann, R. Nutt, D. F. Veber, R. A. Vitali, S. L. Varga, T. A. Jacob, F. W. Holly and R. G. Denkewalter, *J. Am. Chem. Soc.*, 1969, **91**, 503-507.

(10) D. Stevenson and G. T. Young, *J. Chem. Soc. Section C.*, 1969, 2389-2398.

(11) H. Kunz and U. Unverzagt, *Angew. Chem. Int. Ed. Engl.*, 1984, **23**, 436-437.

(12) H. Kunz and M. Joachim, *Angew. Chem. Int. Ed. Engl.*, 1988, **27**, 1375-1377.

(13) H. Kunz and K. Bruch, *Angew. Chem. Int. Ed. Engl.*, 1990, **29**, 1457-1460.

(14) H. Kunz and R. Barthals, *Angew. Chem. Int. Ed. Engl.*, 1983, **95**, 799.

(15) D. S. Kemp and G. Hanson, *J. Org. Chem.*, 1081, **46**, 4971-4975.

(16) B. W. Bycroft, W. C. Chan, S. R. Chhabra and N. D. Hone, *J. Chem. Soc. Chem. Commun.*, 1993, 778-779.

(17) M. Cole and T. Savidge, *Methods Enzymol.*, 1975, **43**, 705-728.

(18) F. Brtnik, T. Barth and K. Jost, *Coll. Czech. Chem. Commun.*, 1981, **46**, 1983-1989.

(19) F. Brtnik, T. Barth and K. Jost, *Coll. Czech. Commun.*, 1981, **46**, 1983-1989.

(20) T. D. Pallin, *PhD Thesis*, 1994, Edinburgh University.

(21) M. Cole, *Process Biochem.*, 1967, **2**, 35.

(22) H. J. Duggleby, S.P. Tolley, C. P. Hill. E. J. Dodson, G. Dodson and P. C. E. Moody, *Nature* (London), 1995, **373**, 264.

(23) Y. Kiso, S. Nakamura, K. Ito, K. Ukawa, K. Kitagawa, T. Akita and H. Moritoki, *J. C. S. Chem. Commun.*, 1979, 971-972.

(24) L. C. Draffan, *Personal Communication*, Edinburgh University.

6. Investigations into Fragment Coupling Utilising Novel Lysine Protection

6.1 Introduction

With the new 4-benzyloxyphenacetyl (Bepa) (94) protecting group in hand, (Chapter 5) it was desirable to test its ability to withstand the conditions necessary for azide fragment coupling (FC). Draffan[1] had already tested its durability during SPPS by incorporating it into the model pentapeptide system (Phe-Gly-Lys-Ala-Gly) and had shown that cleavage of the peptide from the resin with TFA also cleaved the benzyl ether to give the Hopa derivative (82) as expected. Deprotection of this single Hopa protected lysine residue with penicillin acylase was fast (~24h) and the results were comparible to those obtained by Pallin[2] for this system.

Salmon Calcitonin I (SCT) was chosen as a model system as this had already been synthesised by Pallin using the azide FC technique when testing Phenac as a suitable protecting group for lysine.

SCT[3] is a 32 residue peptide hormone with an amino terminal 1-7 disuphide bridge and a prolinamide at the carboxyl terminus[4,5] (95) (Fig 6.1). A calcitonin is a peptide hormone produced in the thyroid gland of mammals or the ultimobranchial gland of birds and fish. It is secreted in response to high calcium levels in the blood and has a regulatory function in the calcium phosphorous metabolism.[6] The calcitonins are widely used as drugs combating various bone metabolism disorders such as Pagets disease,[7] hypercalcemia[8] and osteoporosis.[9] Fish calcitonins are generally more potent than mammalian calcitonins, possibly due to reduced degradation,[10] and are currently used therapeutically.

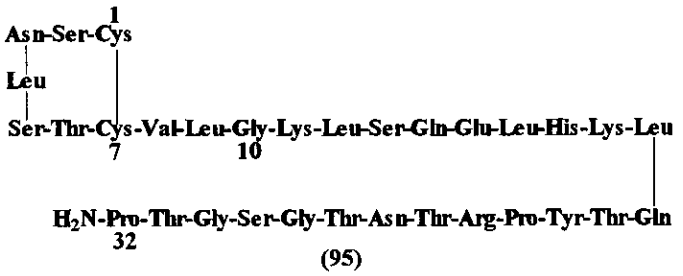


Fig 6.1 Amino acid sequence of SCT

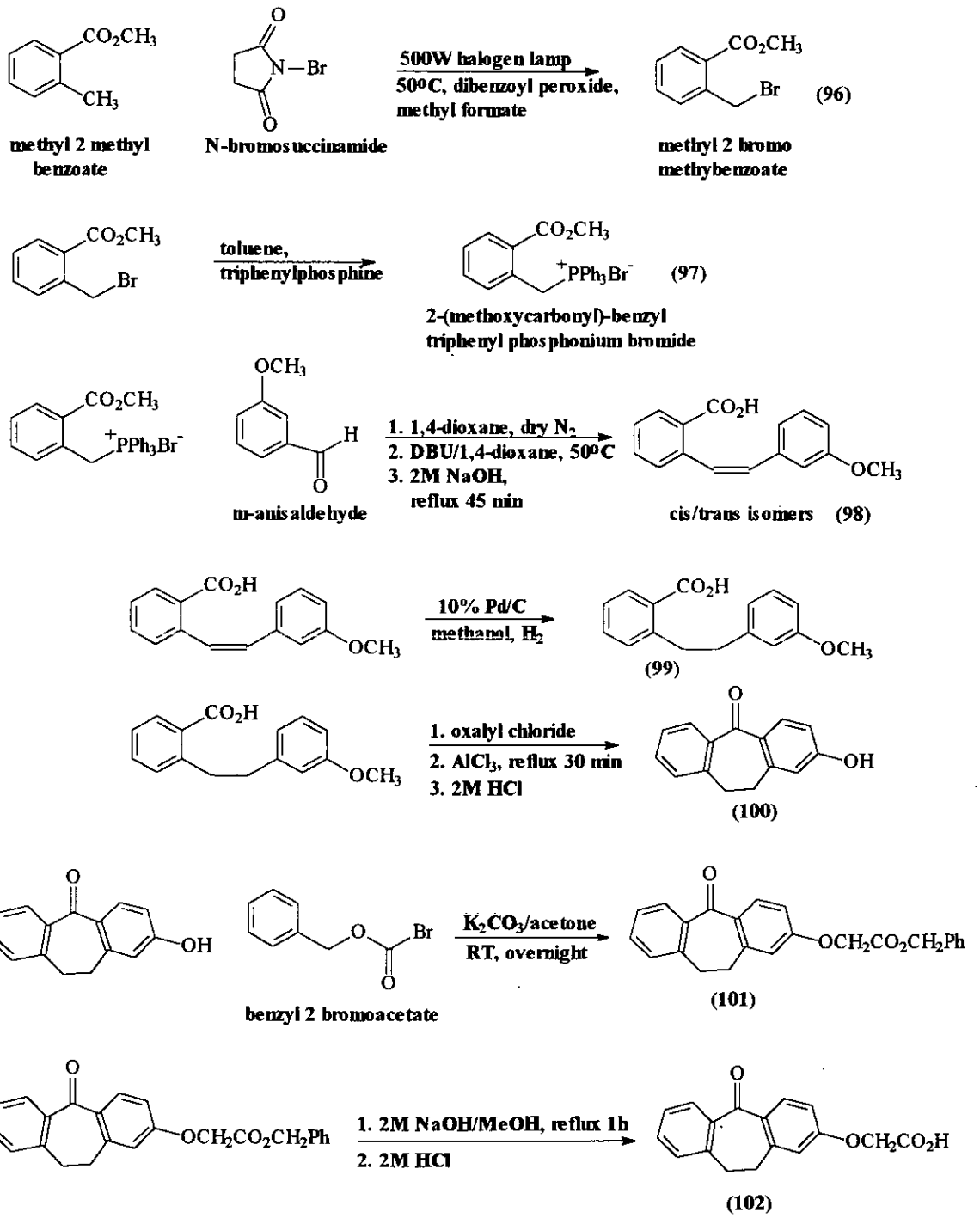
There have been several syntheses of the calcitonins, some utilising the fully protected fragment approach[11] and others utilising the enzymatic coupling approach;[12] thus no problems were envisaged for this synthesis.

Pallin chose to couple fragments 1-10 and 11-32 of the calcitonin (Fig 6.1) using the azide FC technique. This meant the amide bond to be formed lay between ^{10}Gly and ^{11}Lys. Conversion of glycine to the azide avoids both racemisation of the azide and minimises steric interaction during coupling, though conveniently, the cut happens to lie between the "head" and the "tail" of the peptide. Thus the same fragments were chosen again for the synthesis with lysine Bepa protected.

Both components to be coupled required derivatisation of their C-terminal residues and this could be achieved simply by utilising the versatile dibenzosuberone derived linker (17) for the C-terminal hydrazide (fragment 1-10) and (18) for the C-terminal amide (fragment 11-32). Thus, before embarking on a synthesis of the fragments to be coupled, the linker had to be synthesised.

6.2 Synthesis of the Dibenzosuberone Linker

The linker (102) was synthesised with no major difficulties using the synthetic route (Scheme 6.1) devised by M^{c}Innes[13] and Irving.[14]

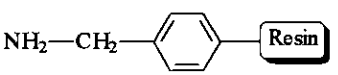


Scheme 6.1 Synthesis of dibenzosuberone linker

Once the linker is synthesised the next step is attachment to a solid support and subsequent reactions are then carried out while the linker is immobilised.

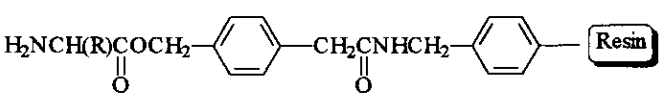
The first step is reduction of the ketone function, which is then reacted with the appropriate reagent in order to derivatise the linker to suit the functionality required at the C-terminus, be it the hydrazide or the amide.

Irving utilised aminomethylated polystyrene resin (103) as a support for the linker (102). This allowed the resin functionality to be controlled by monitoring the amount of remaining free amine (Kaiser test) which could then be capped to prevent further reaction during synthesis of a peptide e.g. low resin loadings are required for large polypeptides. This support also provided a strong amide bond between the linker and the resin.



(103)

Problems were encountered, however, while reproducing Irving's work. It was found that the support contained fines which prevented the resin from swelling sufficiently. This prohibited access of reagents to the linker, as well as blocking the filters on the peptide synthsiser. New batches of resin were purchased but these were equally unsuitable. Therefore a different support had to be used. The *p*-aminomethylated (PAM) resin (104), which is functionalised by an amino acid, still has all the advantages of the aminomethylated resin and was shown to be free of fines.

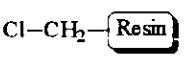


(104)

R = CH_3 (Ala)
$CH_2CH(CH_3)_2$ (Leu)
$CH(CH_3)_2$ (Val)

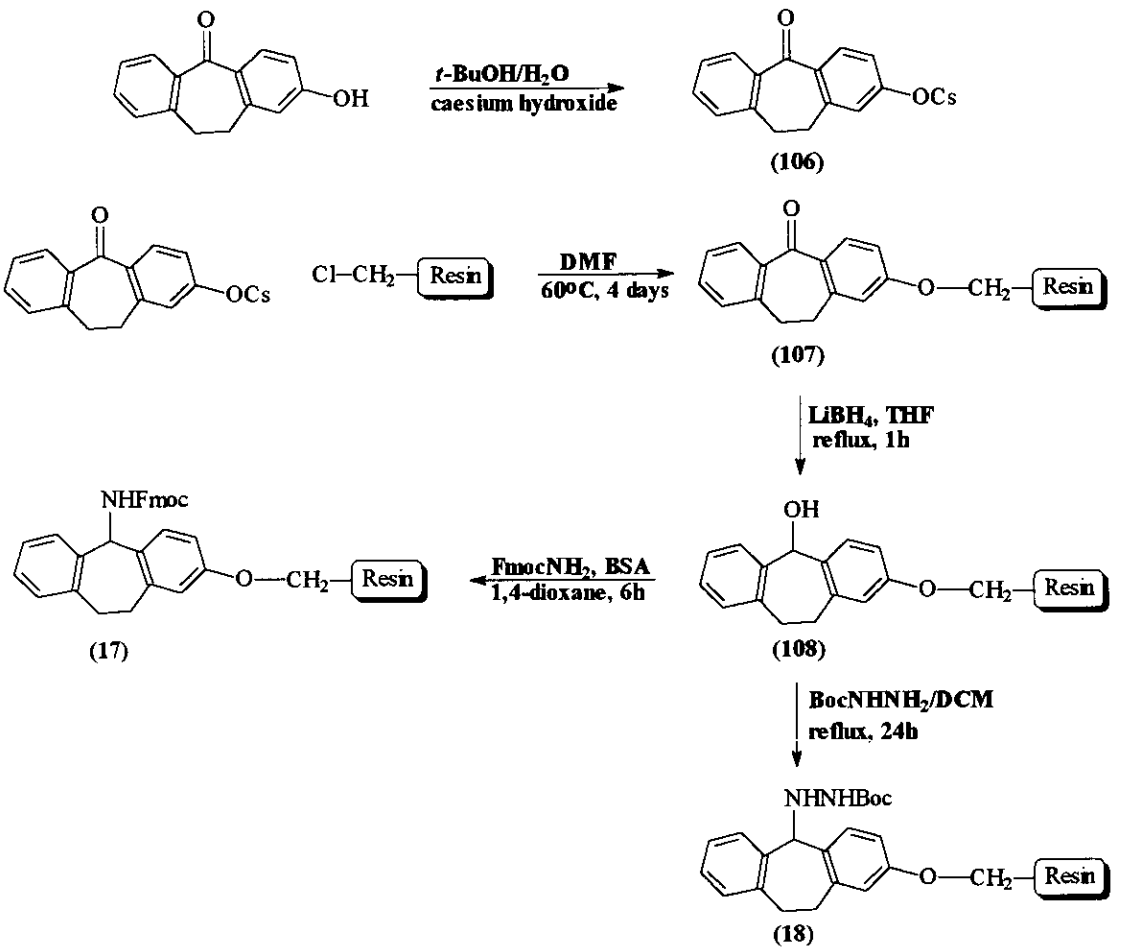
Unfortunately difficulties still ocurred after reduction of the ketone as the resulting system, unless used immediately, appeared to be unstable. Despite this, material was obtained that was suitable for use on the synthesiser but a more stable system would obviously be preferrable.

The chloromethylated polystyrene (Merrifield) resin (105) had previously been used by M^{c}Innes. This involves attachment of the linker (100) to the support *via* its caesium salt (106) forming a strong and stable ether linkage (107). The only disadvantage in this system is controlling the resin functionality and it was for this reason that the aminomethylated resin (105) was investigated by Irving. M^{c}Innes's synthesis was repeated (Scheme 6.2) and when the ketone was reduced the resulting system was found to be extremely stable.



(105)

The linker was then functionalised forming either the C-terminal hydrazide (18) or amide (17) by reaction with the Boc hydrazine or Fmoc carbamate respectively (Scheme 6.2).



Scheme 6.2 Synthesis of hydrazide and amide linker via attachment to chloromethylated polystyrene resin.

6.3 Synthesis of Protected Fragments of SCT for Azide Fragment Coupling

The first amino acid, in the hydrazide fragment (109), Fmoc-Gly-OH, was coupled manually to linker (18) *via* its acid chloride. The rest of the chain was then assembled on the synthesiser employing standard HOBt coupling conditions. Several syntheses of this fragment were attempted before a suitable combination of protecting groups for the N-terminus and the side-chain of cysteine were found.

Pallin, in his synthesis, protected cysteine with the Acm sulphydryl protecting group and used Fmoc for protection at the N^{α} position. However, the combination of the Acm groups with Fmoc formed an extremely hydrophobic peptide which was very difficult to purify. Thus, for this investigation it was desirable to produce the peptide in a more soluble form to avoid difficulties in purification.

It was proposed that the hydrazide fragment could be coupled without cysteine protection to improve the solubility of the peptide if the cysteines were oxidised beforehand, forming the disulphide bridge and thus preventing them from taking part during azide FC. The cysteine residues were therefore protected during synthesis with the trityl group which is acid labile and is cleaved when the peptide is released from the resin with TFA. Whilst in theory this should have been possible, attempts to synthesise the trityl protected peptide failed. The coupling efficiency dropped dramatically (~30%) after coupling the last residue in the fragment which is ^{1}Cys(Trt). Attempts to improve this employing extended double coupling cycles and manual coupling in a sonic bath were equally fruitless. The dramatic drop in coupling efficiency is probably caused by steric hindrance from the ^{t}Bu side-chain protecting group on ^{2}Ser to which the bulky ^{1}Cys(Trt) residue is to be coupled. Thus, this route was abandoned and Acm protection had to be used.

Further research within the group[15] showed that Hopa could be used successfully as an N^{α}-protecting group in SPPS and that it could aid solubility of peptide fragments not normally soluble with the Fmoc group. Thus, it was hoped that by switching to Hopa for N^{α}-protection in the SCT 1-10 synthesis that this would aid solubility of the Acm protected fragment for purification. This worked very well and solubility of the fragment improved dramatically compared to Pallin's synthesis. The fragment (109) (Fig 6.2), now mostly soluble in aqueous media, was purified by RP HPLC in a 28% yield.

The amine component (111) (Fig 6.2) is a peptide amide and was synthesised using the linker (17). ^{18}Lys and ^{11}Lys were coupled as their Bepa derivatives without any problems and the crude peptide was purified using the Tbfmoc/carbon procedure (Chapter 3). The yield of this peptide was lower than expected (~19%) but it was discovered later that this was due to an inaccurate initial resin functionality. The cleavage time, initially carried out for 5h, had to be repeated for a further 3h on the crude material as MS analysis showed evidence of incomplete cleavage of the benzyl ether on the Bepa protecting group to the Hopa derivative.

6.4 Azide Coupling of Protected Fragments of SCT

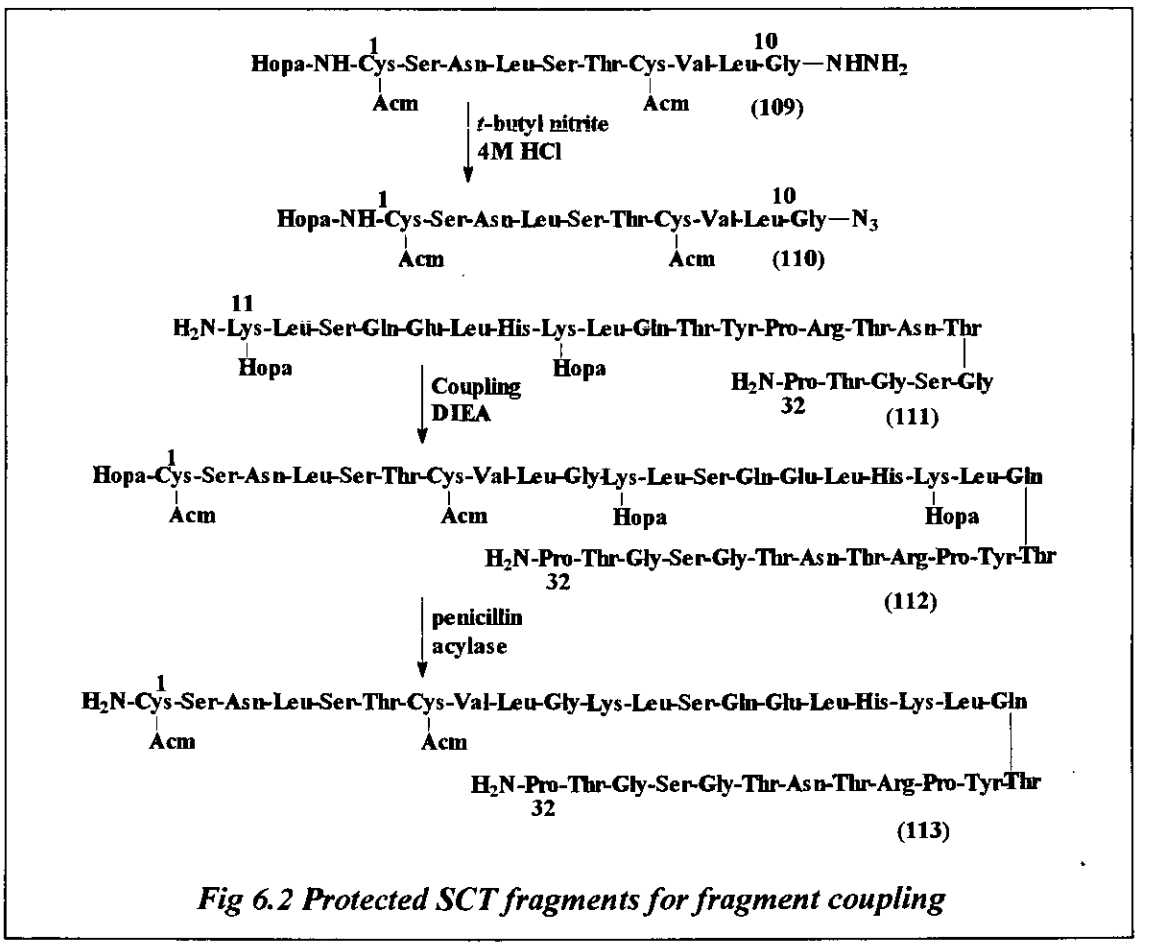


Fig 6.2 Protected SCT fragments for fragment coupling

The peptide hydrazide (109) was converted to the peptide azide (110) smoothly by the addition of *t*-butyl nitrite at low temperature. On addition of the amine component (111) the reaction was stored at 0°C for 24h. More azide was added and after a further 24h a third aliquot of azide was added as some amine component still remained unreacted. The coupling was monitored by HPLC (Fig 6.3) and was stopped after 4 days as no further reaction was observed. The product (112) was isolated by RP HPLC in 20% yield.

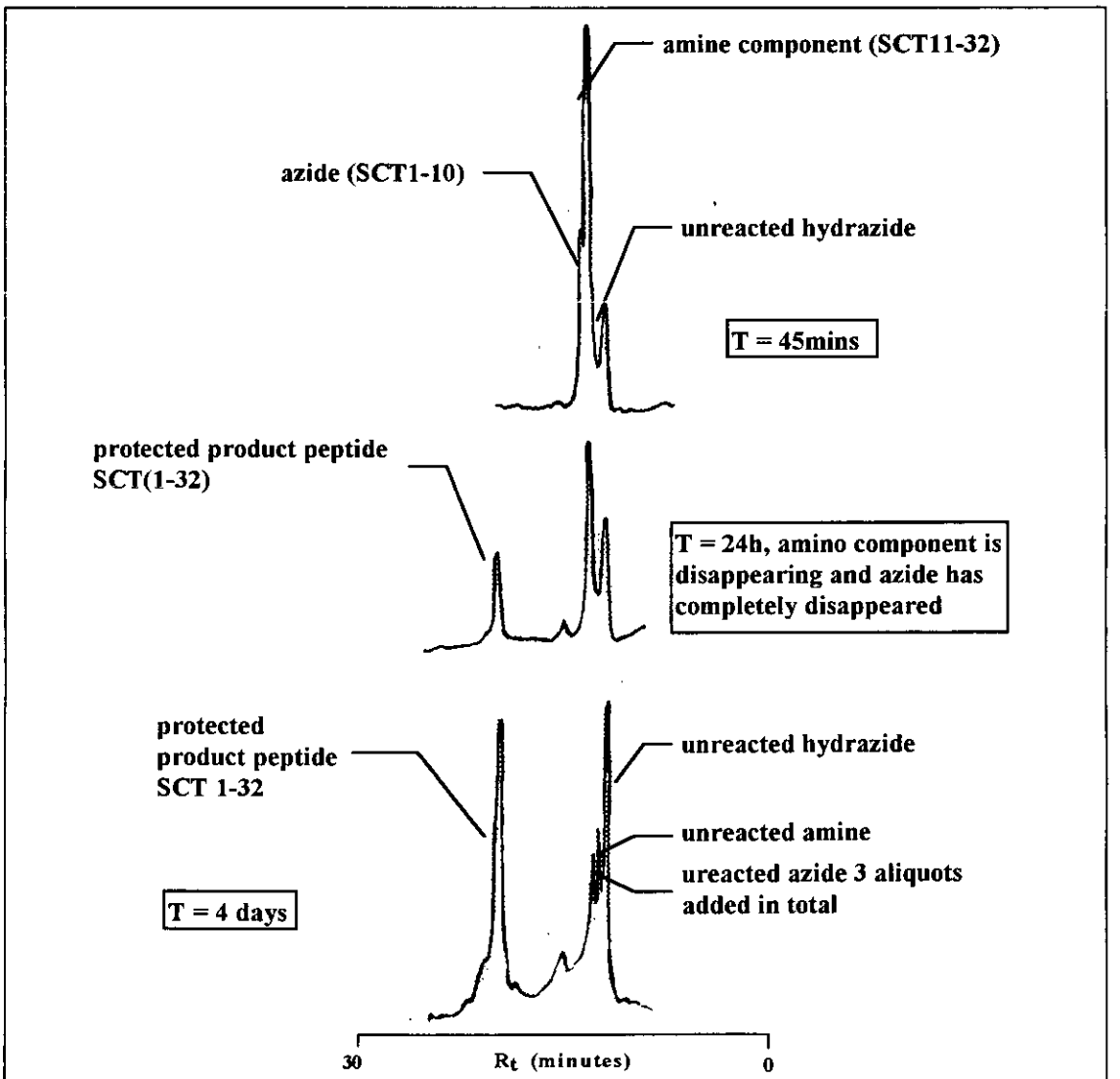


Fig 6.3 HPLC traces showing the progress of the azide coupling of SCT(1-10)(110) and SCT (11-32) (111)

6.4.1 Enzyme deprotection of SCT 1-32 $(Acm)_2(Hopa)_3$

Once the protected product peptide (112) had been isolated the N-terminal Hopa group and the Hopa lysine protecting groups were removed using immobilised penicillin acylase, incubating at 37°C and pH 8.1. Deprotection took only 48h and the final product (113) was isolated by filtration to remove the resin bound enzyme, followed by purification by HPLC which resulted in a very poor yield (2%). Due to the low yield there was insufficient material to carry out an Acm deprotection and obtain analysis therefore the Acm groups were not removed from cysteine. It was thought that the low yield was due to the peptide chain binding to the resin on which penicillin acylase is immobilised. The synthesis was not optimised due to other developments within the research group concerning Bepa. These are mentioned below.

6.5 Summary

Studies using SCT have shown that Bepa is a suitable protecting group for use in the azide fragment condensation technique but a more suitable support for the enzyme is required. However, recent investigations[1] have shown that this protecting group strategy is only applicable to small peptide systems since no deprotection of the lysine residues was observed when Bepa was incorporated into a larger polypeptide system, namely ubiquitin (76 residues of which 7 are lysine).

Fig 6.4 shows the 3-D crystal structure of the enzyme which might provide some explanation to the differences observed in the deprotection of small systems as opposed to larger peptide systems. The B chain N-terminal serine, responsible for hydrolysis, is just discernable at the apex of the β-strand buried deep in the centre of the molecule. Clearly the relative inaccessability of the active site means that, whilst small unstructured molecules are allowed access to the binding pocket and are rapidly

deacetylated, larger protein molecules (which can adopt some degree of structure under the reaction conditions) are less likely to be deprotected, if at all.

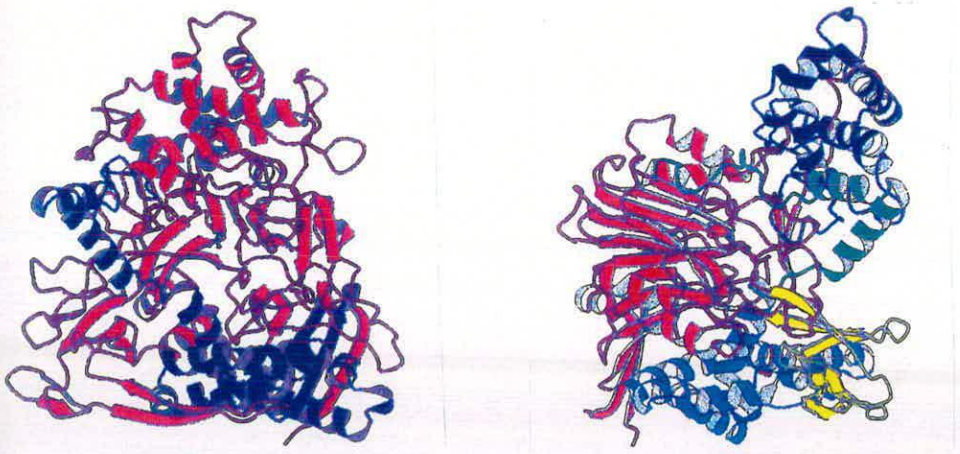
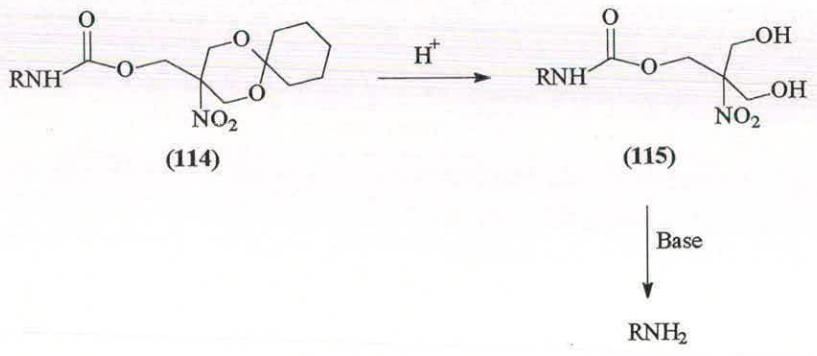


Fig 6.4 MOLSCRIPT representation of the 3-D crystal structure of penicillin acylase (left) Blue = A chain, Red = B chain showing ball and stick active site serine side chain and (right) an orthogonal view, coloured as described previously.[14]

6.6 Recent Developments in Lysine Protection



(114) (115)

Fig 6.5 Tnm protecting group for N^{ε} function of lysine: cleavage conditions

Owing to these results, a completely new approach to lysine protection had to be adopted. Comer[16] recently developed a group, which has found application in lysine protection.[1] The protecting group, referred to as Tnm (114), is stable to all the conditions required for Fmoc SPPS, but cleaves at the acetal during acid cleavage of the peptide from the resin to yield a protecting group (115) which is more solubilising and which now may be cleaved under mildly basic conditions (Fig 6.5). The group has been incorporated into ubiquitin[1] and has been shown to deprotect easily from the lysine residues.

6.7 Preliminary Investigations into the Stability of the Tnm Group for Fragment Coupling

6.7.1 Azide fragment coupling

It was mentioned in Chapter 3 that a convergent synthesis of dhEPO may be better yielding than the stepwise synthesis of the protein. Thus, in order to test the stability of the Tnm group to FC, two small fragments of hEPO (Fig 6.6) were chosen as a model for an azide FC.

Fmoc-NH-^{103}Arg-Ser-Leu-Thr-Thr-Leu-Arg-Ala-Leu-^{112}Gly-$NHNH_2$

(116)

H_2N-^{113}Ala-Gln-Lys-Glu-Ala-Ile-Ser-Pro-Pro-Asp-Ala-^{125}Ala-OH

Tnm (117)

Fig 6.6 Proposed protected fragments of hEPO for azide fragment coupling

Coupling between ^{112}Gly and ^{113}Ala seemed to be the most appropriate site in the protein sequence (Fig 3.1 Chapter 3) since glycine cannot racemise and coupling to Ala should minimise steric interactions. An attempt was made to synthesise the hydrazide fragment (116) but problems arose after ^{107}Thr was coupled. The next

amino acid in the sequence was ^{106}Thr but coupling this residue proved to be virtually impossible, even after extended double coupling cycles were employed followed by additional coupling in a sonic bath. Obviously this difficulty was not envisaged when choosing appropriate fragments since the chain assembly for the entire protein had proceeded smoothly. One possible explanation for this is that the isolated fragment is too hydrophobic in the absence of the rest of the protein chain. This means that the growing chain of the small peptide can interact with the resin causing the chain to wrap itself tightly around its surface. This would render the N-terminus inaccessable to the next activated amino acid which would explain why ^{106}Thr would not couple to ^{107}Thr. Despite this, more suitable fragments were not chosen as it was more important to demonstrate the utility of the Tnm protecting group in FC. Thus the hydrazide fragment was synthesised once more but only as far as ^{107}Thr (118) and the amino component (117) was assembled with no apparent difficulties (Fig 6.7).

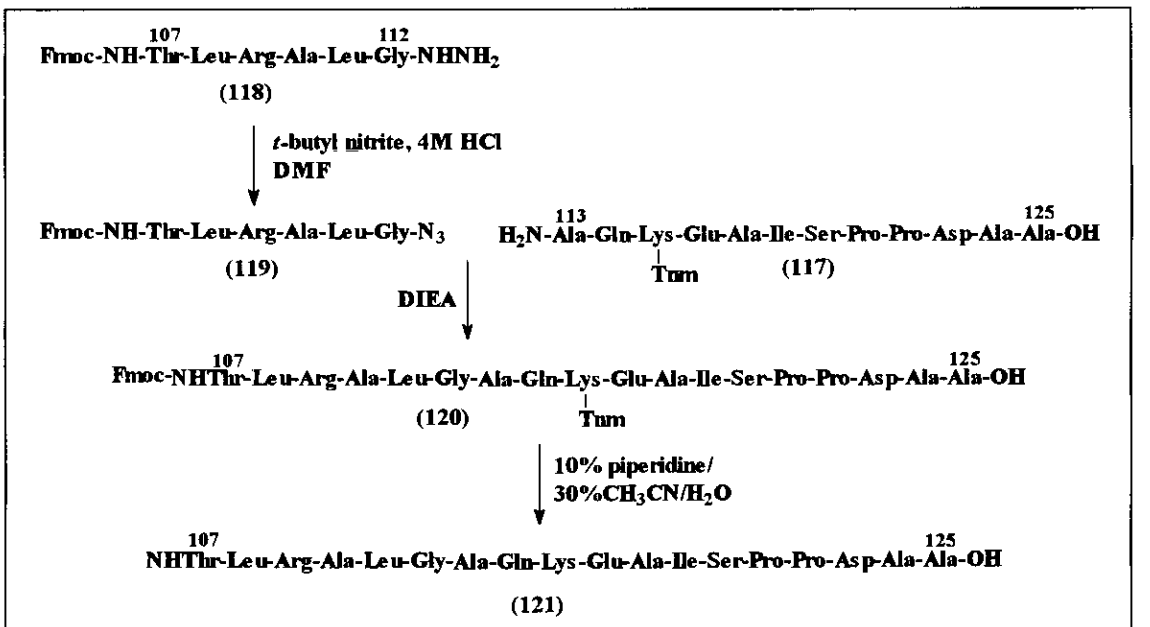


Fig 6.7 Azide fragment coupling of two fragments of hEPO utilising the Tnm protecting group for lysine

The hydrazide (118) was converted to the azide (119) and the segments were coupled in exactly the same way as described for SCT (Fig 6.7). Coupling was monitored by HPLC (Fig 6.8) and proceeded smoothly. MS analysis of the final product showed

that the Tnm group had not cleaved in the azide coupling mixture and thus is suitable for azide FC. The protected product peptide (120) was isolated by HPLC and lysine protection along with Fmoc N^{α} protection were removed simultaneously with a 10% piperidine solution to give the fully deprotected product peptide (121).

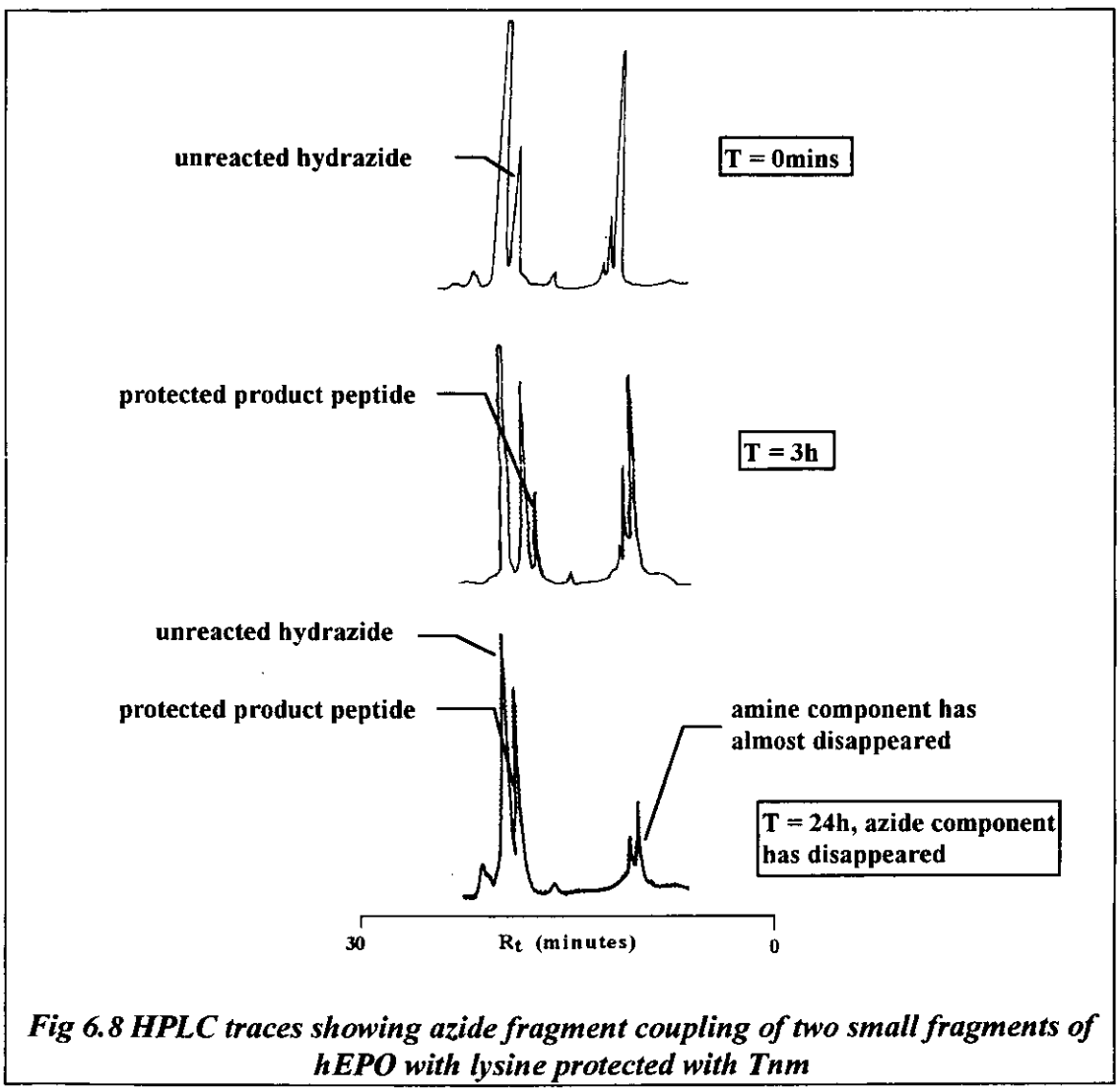


Fig 6.8 HPLC traces showing azide fragment coupling of two small fragments of hEPO with lysine protected with Tnm

6.7.2 Ester fragment coupling: preliminary investigation

Coupling peptide fragments *via* the active ester has been carried out successfully by some groups.[17] However, this technique has not gained as much interest as the azide method due to the requirement for protection of all hydroxyl and carboxylic acid side-chains in addition to those mentioned for the azide. Despite this it has certain advantages over the azide route. Coupling fragments *via* the active ester requires less sensitive conditions than the azide procedure (higher temperatures can be employed) and an excess of the activated species is not required for coupling to proceed effectively (this is an important consideration for large, expensive polypeptides). It would, therefore, be desirable to develop a route whereby the activated ester of minimally protected fragments could be formed i.e only lysine and cysteine need be protected as in the azide method. It has been proposed that this could be achieved if the minimally protected peptide azide was formed but instead of coupling the amino component the azide is first converted to the active ester by reacting it with HOCt (Fig 6.9). Forming the active ester in this way would not require DIC thus hydroxyl and carboxylic side-chains need not be protected.

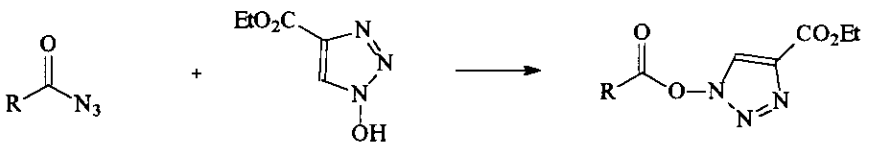


Fig 6.9 Formation of the active HOCt ester via the azide for use in ester fragment coupling

The method was applied to the two small fragments of hEPO used in the azide coupling described above (Fig 6.7). Three equivalents of HOCt were used for conversion of the azide to the active ester and additional base was added (molar equivalent to HOCt) to compensate for the acidity of HOCt. The coupling was monitored by HPLC, as before, and gave an identical profile to that obtained for the azide coupling (Fig 6.7). The results however, remain inconclusive at this time, as no

MS data could be obtained to confirm formation of the active ester. Thus, there is no evidence to confirm coupling actually proceeded *via* this active species.

Obviously further investigation is required into this area as this could provide a very promising method for producing minimally protected fragments for active ester fragment coupling. Unfortunately due to limited time this could not be taken further.

6.8 References

(1) L. C. Draffan, *Personal Comunication*, Edinburgh University.

(2) T. D. Pallin, *PhD Thesis*, Edinburgh University.

(3) J. T. Potts, *Bone and Mineral*, 1992, **16**, 169-173.

(4) I. MacIntyre, in *'Endocrinology'* (L. J.Degroot, Ed.), 1989, **2**, 892-901.

(5) M. Azira, in *'The Calcitonins: Physiology and Pharmacology'*, pp 4-10, A. G. Karger, Basel, Switzerland.

(6) L. A. Austin and H. Heath, *N. Engl. J. Med.*, 1981, **304**, 296-297.

(7) A. Aurimedes, R. K. Baker and S. Wallach, *Metabolism*, 1974, **23**, 1037-1046.

(8) D. J. Hosking and S. R. Heller, *Eur. J. Clin. Pharmacol.*, 1986, **31**, 27-31.

(9) S. Wallach *et al.*, *Curr. Ther. Res.*, 1977, **22**, 556-572.

(10) D. M. Findley *et al.*, *Endocrinology*, 1983, **112**, 1288-1291.

(11) P. Seiber, B. Riniker, M. Brugger, B. Kamber and W. Rittel, *Helv. Chim. Acta.*, 1970, **53**, 2135-2150.

(12) M. Gondo, H. Yamashita, K. Sakakibara and Y. Isowa, *Peptide Chemistry* (T. Shiori Ed.), 1981, Protein Research Foundation, Osaka, Japan.

(13) C. M^{c}Innes, *PhD Thesis*, Edinburgh University.

(14) S. L. Irving, *PhD Thesis*, Edinburgh University.

(15) A. Kelly, *Personal Comunication*, Edinburgh University.

(16) A. J. R Comer, *Personal Comunication*, Edinburgh University.

(17) H. Benz, *Synthesis*, 1994, **4**, 337-358.

7. Outlook and Conclusions

The preceeding chapters have described the methodology required for both stepwise and convergent syntheses of proteins. The pioneering stepwise syntheses of dhEPO and other large proteins within the group (stromelysin[1], obese gene[2,3] and interferon-γ[4]) have demonstrated that larger polypeptide systems previously unattainable by other methods are now possible to synthesise linearly, purifiy and characterise using our methodology. It would therefore be reasonable to assume that a convergent synthesis *via* the union of two large units of a polypeptide is within our reach as a suitable protecting group strategy now exists. Utilising convergent synthesis for the production of proteins should ease purification and give improved yields compared to those obtained from a linear synthesis.

However, it is evident from the experimental results that the combination of the two techniques has to be given careful consideration and their application could in fact be dictated according to protein sequence. This was apparent in the fragment coupling of SCT and also in the union of the two small segments of hEPO where choice of peptide fragments, though apparently straightforward, played an essential role in the FC technique. In fact, choice of peptide fragments could decide whether a FC synthesis is indeed a possibility for some protein sequences at all.

Obviously much investigation is still required in this field as synthesis and purification are only the stepping stones to obtaining biologically important proteins. The most difficult challenge will be to produce these proteins in a biologically active form which usually requires that the protein be folded. Indeed, the precisely folded three-dimensional structure of a protein is perhaps the ultimate synthetic goal.

7.1 References

(1) G. H. Morton, *Personal Communication*, Edinburgh University.

(2) Y. Zangh, *Personal Communication*, Edinburgh University.

(3) L. Jiang, *Personal Communication*, Edinburgh University.

(4) L. C. Draffan, *Personal Communication*, Edinburgh University.

8. Experimental

8.1 Notes

Sonication was carried out in a Decon FS300b sonic bath. Melting points were recorded in open capillaries using a Buchi 510 oil immersion melting point apparatus and are uncorrected. Nuclear magnetic resonance (NMR) spectra were recorded on either a Jeol FX-60 (60MHz), a Brucker WP-200 (200MHz), a Brucker AC-250 (250MHz) or a Brucker WH-360 (360MHz). UV spectra were recorded on either a Varian Carey 210 double beam spectrophotometer, a Perkin Elmer UV/vis single beam spectrophotometer or a Unicam UV/vis double beam spectrophotometer in the solvents described in the text. High and low resolution fast atom bombardment mass spectra (FAB MS) were measured on a Kratos MS50TC mass spectrometer, electron ionisation mass spectra (EI MS) were measured on a Kratos 902 MS and matrix-assisted laser desorption/ionisation (MALDI) time of flight mass spectra (TOF MS) were recorded on a PerSeptive Biosystems Voyager™ Biospectrometry™ Workstation from Vestec mass spectrometry products. Infrared spectra were recorded on a BIO-RAD SPC3200 instrument. Elemental analyses were performed on a Perkin Elmer 2400 CHN elemental analyser. All amino acids were purchased from Bachem with the exception of Fmoc-Lys(Phenac), Fmoc-Lys(Mepa), Fmoc-Lys(bepa), Fmoc-Lys(Tnm). Fmoc-Cys(Phenac) and Fmoc-Cys(Mepa) which were prepared in this laboratory by the methods described in this text, or in the texts of A.R. Brown. A.J.R Comer and L.C. Draffan. All amino acids were of the L configuration. Gel filtrations were carried out on a Pharmacia LKB p500 FPLC system using Superdex™ 75 HR 10/30 or 26/60. Amino acid analyses were performed on an LKB 4150 alpha amino acid analyser on the hydrosylate obtained from heating samples in 6N HCl at 110°C (number of hours stated in text) in a sealed

Carius tube, followed by evaporation to dryness. Protein sequencing was performed on an ABI 477A protein sequencer at the Welmet sequencing facility, University of Edinburgh. High performance liquid chromatography (HPLC) was carried out using either an ABI system comprising 2 x 1406A solvent delivery systems, a 1480 injector/mixer and a 1783A detector/controller, or a Gilson system comprising 2 x 306 solvent delivery systems, an 811C dynamic mixer, an 805 manometric module, a 119 UV/VIS detector and a Gilson software-driven gradient controller. Components were eluted from various columns, as described in the text, with a linear gradient of acetonitrile (%B), far UV grade (Rathburn Chemicals) in Milli-Q grade water (%A), where both solvents contained 0.1% v/v of HPLC grade TFA (Fisons). Peptide synthesis grade trifluoroacetic acid (TFA) was purchased from Applied Biosystems (ABI). Peptide synthesis grade dimethylformamide (DMF), 1,4-dioxane and piperidine were obtained from Rathburn Chemicals, Walkerburn, Scotland. Throughout the text ether refers to diethyl ether.

8.2 Solid Phase Peptide Synthesis

The polypeptides described were synthesised on an ABI 430A automated peptide synthesiser with on line UV monitoring using an ABI 758A detector. All polypeptides were synthesised using the 9-fluorenylmethoxycarbonyl (Fmoc) strategy of N^{α} protection. This involves the complementary use of orthogonal acid labile side chain protection and an acid labile peptide-resin linker.

8.2.1 Side-chain protecting groups used in Fmoc SPPS

The side-chain protecting groups were as follows: *t*-butyl (^{t}Bu) ethers for Ser, Thr and Tyr; ^{t}Bu esters for Asp and Glu; τ–triphenylmethyl (Trt) for His; 2,2,5,7,8-pentamethylchroman-6-sulphonyl (Pmc) for Arg. The carboxamide side chains of Asn

8.2.4 C-terminal hydrazides

The first amino acid is loaded manually *via* its acid chloride and is always Fmoc-Gly-Cl to avoid racemisation. The preparation of the hydrazide resin is given in the following text.

8.2.5 The Fmoc loading test

Dry Fmoc-amino acid resin (2-3mg), was placed in a 10ml volumetric flask and 20% piperidine/DMF was added to the volumetric level. The mixture was sonicated at room temperature for 15 minutes and the UV spectrum of the supernatant was recorded between 280 and 320nm. The resin functionality (mmol/g) and percentage coupling were calculated using the Beer-Lambert law ($\varepsilon 302 = 15400$ for the fulvene piperidine adduct).

8.2.6 Automated SPPS

Synthetic procedures were preprogrammed into the ABI 430A synthesiser prior to the the synthesis. Each synthetic cycle, resulting in the coupling of a single amino acid, involved: **Capping**, to block any unreacted amino groups; **Deprotection**, to remove the Fmoc group from the Nα position; **Coupling** of the next N^{α}-Fmoc-amino acid. Each step was followed by thorough washing of the resin and each cycle was repeated with the chosen amino acid in order to build the desired polypeptide. The preprogrammed synthetic cycles are summarised below.

8.2.6.1 Capping

The resin was vortexed with a solution of acetic anhydride (0.5M), DIEA (0.125M) and HOBt (0.2% w/v) in DMF:1,4-dioxane (1:1, 10ml) for 10 minutes, before the capping solution was drained from the reaction vessel and the resin washed with 6 portions of DMF:1,4-dioxane (1:1).

8.2.6.2 Deprotection

The resin was vortexed with a solution of 20% piperidine in DMF:1,4-dioxane (1:1, 10ml) for 6 minutes before being drained and washed 4 times with DMF. The deprotection cycle was repeated twice but vortexing was only for 1.5 minutes. Finally the resin was washed with 6 portions of DMF:1,4-dioxane (1:1).

The percentage of amino acid successfully coupled in the previous cycle was determined by the absorbance of the fulvene-piperidine adduct at 302nm from the first deprotection. An aliquot of the deprotection filtrate was passed through an UV detector, with on-line integration. Comparison of the area of each succesive deprotection peak with that of the first amino acid allowed an estimate of the percentage coupling of each residue.

8.2.6.3 Coupling

Residues were incorporated in one of two ways; either using the HOBt method, or the HOCt method described below.

The HOBt method: uses double couple cycles in which the first coupling cycle utilises a preformed symmetrical anhydride, followed by a second coupling using a preformed HOBt ester. The exceptions to this were asparagine, glutamine and histidine, which were coupled twice as their HOBt esters and glycine, which was coupled once as the symmetrical anhydride.

Both the symmetrical anhydride and the activated ester were preformed in an 'activation vessel' before being transferred to the reaction vessel. For symmetrical anhydride formation, the Fmoc-amino acid (1mmol) was reacted with DIC (0.5mmol) in DMF:1,4-dioxane (1:1, 8ml) for 10 minutes after which time it was transferred to the reaction vessel containing the resin. The mixture was vortexed for 30 minutes, the resin was drained and washed with 4 portions of DMF:1,4-dioxane (1:1). The resin was then vortexed for a second 30 minute period with the Fmoc-amino acid HOBt

active ester, preformed from the Fmoc-amino acid (0.5mmol), HOBt (0.5mmol) and DIC (0.5mmol). The resin was drained and washed with 4 portions of DMF:1,4-dioxane (1:1). For certain difficult couplings the coupling time was extended or, alternatively, a third coupling, using the HOBt ester, was included as required.

The HOCt method: uses a single coupling cycle in which the Fmoc-amino acid HOCt active ester is used. The exception to this was histidine which was coupled as the HOBt active ester (HOBt (1mmol) was placed in the cartridge with histidine (1mmol) prior to synthesis). The HOCt active ester was preformed as for the HOBt active ester but a 1mmol scale was used.

8.2.7 The Kaiser test

A sample of dried resin (between 2-3mg) was placed in a test tube and ninhydrin reagents according to the following ratio were added:

Monitor 1: phenol/EtOH (75μl)
Monitor 2: KCN/pyridine (100μl)
Monitor 3: ninhydrin/EtOH (75μl)

The resin was incubated in a heating block at 110°C for 7 minutes before 60% EtOH/H_2O (4.8ml) was added. The mixture was vortexed to mix the solution thoroughly and the resin was allowed to settle to the bottom again. The absorbance

(Abs) of the supernatant at 570nm was recorded using the 60% EtOH/H_2O solution as the standard. The percentage coupling was calculated as follows:

$$\mu mol/g = \frac{[Abs(sample) \times dilution(ml)] \times 10^6}{\text{Extinction coefficient} \times \text{sample weight (mg)}}$$

Extinction coefficient = 15,000 $M^{-1}cm^{-1}$
Dilution = 5ml

$$\% \text{ coupled} = \left[1 - \frac{\text{amine}(\mu\text{mol/g})}{10^3 \times \text{substitution}(\text{mmol/g})} \right] \times 10^3$$

8.2.8 Loading Tbfmoc onto the peptide-resin

Using the Fmoc loading test the resin functionality (mmol/g) of the peptide-resin was calculated. Tbfmoc chloroformate (3 equivalents) was sonicated in DCM until all had dissolved. The solution was added to the resin followed by DIEA (1 equivalent). Additional DCM was added if required to swell the resin fully. The mixture was sonicated in the dark for 3h after which time the resin was filtered, washed with copious DCM then ether, dried *in vacuo* and a Tbfmoc loading test was carried out. The resin was swollen in 1,4-dioxane and stored in the freezer until required.

8.2.9 The Tbfmoc loading test

Dry Tbfmoc-peptide-resin (2-3mg) was placed in a 10ml volumetric flask and 20% piperidine/1,4-dioxane was added to the volumetric level. The mixture was sonicated for 15 minutes. The UV spectrum of the supernatant was recorded between 340 and 400nm. The resin functionality was calculated from the equation:

Resin functionality (mmol/g) = [0.613 x Abs at 364nm]/weight of resin (mg)thus, % loading of Tbfmoc = mmol/g (Tbfmoc-peptide-resin)/mmol/g (peptide-resin)

8.3 Experimental Details

Ethyl diazoacetate (41)[4,5]

A solution of ethyl glycinate hydrochloride (210g, 1.5mol) in water (500ml) was mixed with DCM (500ml) in a 3-neck 3l flask fitted with a mechanical stirrer and a nitrogen inlet tube and cooled to -15°C (acetone/dry ice). A cold solution of sodium nitrite (124.5g, 1.8mol) in water (200ml) was added in one portion with stirring. The temperature was lowered to -20°C and an ice cold solution of 5% (w/w) H_2SO_4 (142.5g) was added dropwise over 30 minutes while keeping the temperature below -10°C. When all the acid solution had been added, the temperature was allowed to rise to 0°C(**N.B.** aqeous and DCM layers must be mixed thoroughly throughout). The reaction mixture was transferred to an ice cold 2l separating funnel. The yellow organic layer was run into an ice cold solution of 5% sodium carbonate (1l) and the water layer was extracted once with DCM. The combined DCM and Na_2CO_3 portions were stirred until no more gas evolved. The two layers were separated and the water layer was extracted once with DCM. The organic layer was dried ($MgSO_4$) and evaporated *in vacuo* behind a screen. The yellow liquid, which was pure enough for the next step, was stored below 4°C overnight in a well wrapped container.
Yield 95-100%.

Diazoacetic ester salt (43)[4,5]

DMF (52.8g, 0.72mol) was cooled by ice-salt bath under argon. Freshly distilled thionyl chloride (53ml, 0.72 mol) was added dropwise with stirring. When all the thionyl chloride had been added, the mixture was stirred at room temperature for 30 minutes. The colourless gel was then evaporated *in vacuo* (*via* oil pump) for 2h (water bath ~40°C), until a white solid was afforded (42). Care should be taken when open to the air. Assumed yield 100%. The white solid (42) was dissolved in A.R. chloroform (400ml) and cooled *via* ice salt bath. Ethyl diazoacetate (41), prepared above, was added dropwise under nitrogen keeping the temperature below 5°C.

When the addition was complete, the reaction mixture was stirred at room temperature for 30 minutes then evaporated *in vacuo*. Fresh ether was added to the residue and the yellow solid afforded was filtered under nitrogen and washed with ether. The yellow solid was stored in a dessicator overnight and used in the next step. Yield 130g, 88%.

1-Hydroxyl-4-ethoxy carbonyl-1,2,3-triazole (HOCt) (39)[4,5]

$NH_2OH.HCl$ (44g, 0.63mol) was dissolved in water (200ml), cooled *via* ice salt bath and the pH adjusted to ~7-7.5 with solid Na_2CO_3. The diazonium salt (43) (130g 0.63mol) was added as the solid while keeping the pH at 7-7.5 with 10% Na_2CO_3. The yellow product (44), which precipitated, was stirred for 5 minutes, filtered, washed with a little ice cold water and dissolved in A.R chloroform. Any remaining water was removed by separation in a separating funnel followed by drying with $MgSO_4$. After filtration, acetic acid (0.5ml) was added to the chloroform solution to catalyse the cyclisation and the solution was left at 4°C until cyclisation was complete. (Yellow colour fades but cyclisation was confirmed by ^1H NMR). The chloroform solution was reduced *in vacuo* to a white solid which was filtered and washed with a little cold EtOAc. The filtrate was reduced to a small volume and stored at 4°C until no more crystals formed.

N.B. If an oil formed after evaporation of the chloroform solution, it was taken up into ethyl acetate then reduced *in vacuo* to remove any residual chloroform. More EtOAc was added and the concentrated solution was left at 4°C to crystallise.

Overall yield 22g, 50%.

Mpt. 103-104°C. CHN: (Found, C 38.43%, H 4.48%, N 26.77%, $C_5H_8N_3O_3$ Requires, C 38.2%, H 4.5%, N 26.8%. UV (DCM) λ_{max} = .234nm (ε = 224637$dm^3mol^{-1}cm^{-1}$). MS (FAB) m/z = 158 (MH^+). HRMS (FAB) m/z = 158.05779 (MH^+, $C_5H_8N_3O_3$ Requires, 158.05657). IR (DCM) υ_{max} = 2000-3200 (OH), 1730cm^{-1} (C=O). δH (200MHz, $CDCl_3$) 1.32 (3H, t, CH_3), 4.31-4.4 (2H, q, CH_2), 8.08 (1H, s, olefinic CH) and 10.43ppm (1H, bs, OH). δC ($CDCl_3$, 50MHz) 14.0 (CH_3), 61.7 (CH_2), 121.7 (olefinic C) and 159.2 ppm (C=O).

H_2N-Ala-X-Gly-OH for racemisation studies on HOCt

The synthesis of each tripeptide was carried out on a 0.25mmol scale using Fmoc-Gly loaded Wang resin (330mg, 0.25mmol, 0.87mmol/g). The peptides were synthesised manually using a sonic bath for agitation and each of the amino acids were single coupled as their active HOCt esters. The active esters were made by dissolving the amino acid (1mmol) and HOCt (157mg, 1mmol) in DMF (10ml) and sonicating this with DIC (150µl, 1mmol) for 20 minutes at room temperature. The active ester was then added to the above resin, preswollen in DMF (2ml), and the final mixture was sonicated for 2h. The coupling of each amino acid was confirmed by the UV monitoring method. The resin was capped using the normal acetic anhydride method and Fmoc deprotected using 20% piperidine/DMF before the next amino acid was coupled. All peptides were cleaved from the resin immediately using 95% TFA/H_2O (5ml) and EDT (1ml). The cleavage mixture was reduced *in vacuo* via oil pump to an oil and the peptide was precipitated from ether. The ether was decanted off and the peptide was lyophilised and stored at 0°C.

X = Ala, Arg (Pmc), Asn, Asn (Mbh), Asn (Trt), Asp (^{t}Bu), Cys (Acm), Cys (^{t}Bu), Cys (Trt), Gln, Gln(Trt), Glu (^{t}Bu), Ile, Leu, Lys (Boc), Met, Phe, Pro, Ser (^{t}Bu), Thr (^{t}Bu), Trp, Tyr (^{t}Bu), and Val. NMR studies, δH (360MHz, d^6DMSO), showed no evidence of racemisation for any of the above peptides.

For X = His (Trt): The peptides containing His (Trt) were synthesised by automated SPPS and using the following combination of temperatures, activation times and amounts of HOCt

Table 9 Racemisation studies on His with HOCt

Temperature (°C)	Activation Time (min)	HOCt (mmol)
ambient	15	1
ambient	10	1
ambient	5	1
zero	15	1
ambient	1	2
ambient	1	3
ambient	15	3

Pk_a determination of HOCt

HOCt (0.1M) was titrated against NaOH (0.1M, 50ml). 34.3 ml of HOCt (5mmol) was required to reach the equivalence point which occurred at pH 8.15 as calculated from the titration curve. The K_a value was calculated as 0.00727 and taking the negative of $\log_{10}$ of this value gave a Pk_a value for HOCt of 2.1.

Chemical synthesis of dhEPO (50)

The initial resin was Fmoc-Arg(Pmc)-Wang resin (1g, 0.1mmol), using the method for a low loading. The protecting groups were as stated in the notes for Ser, Thr, Tyr Asp, His and Arg. Trt protection was used for Asn, Cys and Gln and Boc protection on Lys. The amino acids were coupled using the HOCt method apart from histidine which was single coupled as the active HOBt ester. The last 20 residues, Lys(20)-Ala(1), were double coupled as their HOCt active esters. Approximately half the resin was removed after Ala(114) and the remaining resin was used to continue the synthesis until Val(46) when the remaining resin was removed. The sequence was continued to the end, simultaneously, from these two portions. Fmoc was not removed from the last amino acid, but the final resin functionality could not be determined due to the size of the protein and the initial low loading. The resins from the two syntheses were combined, filtered, washed with DMF, DCM then ether and dried *in vacuo* to an off white solid.

Yield 2.2g. Amino acid analysis (48h hydrolysis), Asx_{12}12.7, Thr_{11}10.4, Ser_{10}8.07, Glx_{19}20.1, $Pro_8$9.4, $Gly_9$11.5, Ala_{19}20.3, $Cys_4$3.6, Val_{11}10.4, $Met_1$1.0, $Ile_5$3.7, Leu_{23}21.7, $Tyr_4$4.4, $Phe_4$3.8, $His_2$1.8, $Lys_8$8.7, Arg_{13}12.9.

Tbfmoc loading (51)

The resin bound protein (50) (2g) was given an additional capping (acetic anhydride/HOBt/DIEA), using DCM as the solvent and a sonicator for agitation. The resin was filtered, washed with DMF, 1,4-dioxane, DCM then ether and dried. The dried resin was sonicated in 20% piperidine/DMF for 15 minutes to remove the N-terminal Fmoc group. The resin was filtered, washed with DMF, DCM, then ether

and dried. Tbfmoc was loaded onto the peptide-resin, as described in the notes, and the resin was filtered, washed with copious DCM then ether and dried.

Cleavage of Tbfmoc-dhEPO-resin (52)

The resin bound Tbfmoc-protein (51) (2g) was stirred in EDT (4ml), thioanisole (1ml), H_2O (1ml) and phenol (1.5g) for 10 minutes. TFA (20ml) was added and the mixture was stitrred for 4.5h under dry nitrogen in the absence of light. The resin was filtered, washed with TFA (4ml) and reduced *in vacuo* to a yellow oil. The peptide was precipitated from ether, filtered and washed with ether. The protein was partially solubilised in 20% AcOH/H_2O (it is difficult to solubilise the protein in the aqueous systems required for freeze drying) and lyophilised in the absence of light to a white fluffy solid.

Yield 914mg. HPLC (ABI, Aquapore RP300 C_4, 4.6 x 100, 7μ, A = H_2O, B = CH_3CN, 0.1% TFA, 1ml/min. 30-80%B over 20 minutes. λ_1 = 214nm, λ_2 = 364nm). R_t = 12 minutes, 60%B and 13 minutes, 62%B. Amino acid analysis (48h hydrolysis), Found, Asx_{12}12.7, Thr_{11}11.2, Ser_{10}10.7, Glx_{19}18.8, $Pro_7$7.3, $Gly_9$13.5, Ala_{19}20.1, $Cys_4$3.9, Val_{11}10.6, $Met_1$1.0, $Ile_5$5.0, Leu_{23}22.5, $Tyr_4$4.8, $Phe_4$4.1, $His_2$1.5, $Lys_8$8.8, Arg_{13}12.9.

Affinity purification on charcoal (53)

Charcoal (5g) was washed with 10% piperidine/6MGdm.Cl:isopropanol, 1:1, (100ml) followed by 6MGdm.Cl:isopropanol, 1:1, (9 x 100ml). The crude Tbfmoc-protein (52) (900mg, 50μmol) was solubilised in 6MGdm.Cl:isopropanol, 1:1, (60ml) and added to the freshly washed charcoal (5g). The mixture was vortexed for 10 minutes and centrifuged to a pellet. The supernatant was decanted and analysed by HPLC (ABI, Aquapore RP300 C_4, 100 x 4.6mm, 7μ, A = H_2O, B = CH_3CN, 0.1% TFA, 1ml/min. 30-80%B over 20 minutes. λ_1 = 214nm, λ_2 = 364nm). No absorbance was visible at 364nm; therefore it was assumed that all the Tbfmoc material had been adsorbed onto the charcoal. The charcoal was washed with 6MGdm.Cl:isopropanol 1:1 until no absorbance at 214nm was visible by HPLC. The protein was cleaved

from Tbfmoc by vortexing the charcoal with 10% piperidine/6M Gdm.Cl:isopropanol 1:1 (2 x 15ml) for 10 minutes and then the mixture was centrifuged to a pellet. The supernatants from this were combined and reduced *in vacuo* to remove the isopropanol. The mixture was then acidified to pH 4 with AcOH and DTT (0.1M) added. The solution was left at room temperature overnight. The solution (200ml) was dialysed (Spectra/Por 6 membrane, MWCO 10,000, supplied by Pierce and Warriner) against 6M Gdm.Cl for 2 days, changing the dialysis solution twice daily. The solution was removed from dialysis and concentrated to a volume of 40ml (Amicon ultrafiltration, YM10 membrane, MWCO 10,000).

HPLC (ABI, Aquapore RP300 C_4, 100 x 4.6mm, 7μ, A = H_2O, B = CH_3CN, 0.1% TFA, 1ml/min. 30-80%B over 20 minutes. λ_1 = 214nm, λ_2 = 280nm). R_t = 12 minutes, 60%B. Amino acid analysis (48h hydrolysis), Found, Asx_{12}12.9, Thr_{11}9.1, Ser_{10}8.5, Glx_{19}25.2, $Pro_8$6.2, $Gly_9$12.8, Ala_{19}19.4, Cys_4/Val_{11}14.8 (not resolved), $Met_1$1.5, $Ile_5$4.9, Leu_{23}21.6, $Tyr_4$3.1, $Phe_4$4.2, $His_2$1.5, $Lys_8$8.7, Arg_{13}9.6.

FPLC gel filtration (54)

Gel filtration was carried out on (53) using Superdex^{TM} 75 HR26/60 column and eluting with 6M Gdm.Cl over a column volume of ml. Fractions absorbing at 280nm were analysed by HPLC (ABI, Aquapore RP300 C_4, 100 x 4.6mm, 7μ, A = H_2O, B = CH_3CN, 0.1% TFA, 1ml/min. 30-80%B over 20 minutes. λ_1 = 214nm, λ_2 = 280nm). Fractions 20-25 eluted with the correct retention (R_t = 11 minutes) time and were combined. Fractions 26-35 eluted with the correct retention time but were still slightly impure and therefore were concentrated, combined and reapplied to the column. All fractions with the correct retention time were combined and concentrated (MWCO 10,000) to a volume of 40ml.

HPLC (ABI, Aquapore RP300 C_4, 100 x 4.6mm, 7μ, A = H_2O, B = CH_3CN, 0.1% TFA, 1ml/min. 30-80%B over 20 minutes. λ_1 = 214nm, λ_2 = 280nm). R_t = 11 minutes, 60%B. Amino acid analysis (48h hydrolysis), Found, Asx_{12}12.3, Thr_{11}10.2, Ser_{10}8.5, Glx_{19}20.5, $Pro_8$8.3, $Gly_9$10.4, Ala_{19}19.2, $Cys_4$0.1, Val_{11}12.2, $Met_1$1.2, $Ile_5$4.4, Leu_{23}22.0, $Tyr_4$4.6, $Phe_4$4.1, $His_2$2.7, $Lys_8$8.7, Arg_{13}12.9.

Denaturing and folding

The pH of the above solution (40ml) was adjusted to 8 by addition of Tris.HCl (50mM, pH, 8 37°C) and DTT (0.1M) added. The mixture was incubated at 37°C for 2h. UV analysis[7] gave a protein content of approximately 12.6mg. The fully denatured dhEPO (12mg) was diluted to 500ml with Tris.HCl (50mM, pH8, 5°C), 40μM $CuSO_4$, 0.1% N-lauroyl sarkosine (folding buffer)[6] in 3M Gdm.Cl, and the mixture was stirred overnight. A precipitate formed and was removed by centrifugation, denatured (as above) and stored in the freezer till required. The supernatant was dialysed against folding buffer in 1M Gdm.Cl overnight. Again a precipitate formed and this was removed and treated as before. The supernatant was then dialysed (MWCO 10,000) against folding buffer only to give the folded protein (1mg/100ml).

HPLC (ABI, Aquapore RP300 C_4, 100 x 4.6mm, 7μ, A = H_2O, B = CH_3CN, 0.1% TFA, 1ml/min. 30-80%B over 20 minutes. λ_1 = 214nm, λ_2 = 280nm). R_t = 11 minutes, 60%B.

Tryptic digest

3ml of denatured protein solution (approximately 0.2mg by UV analysis[7]) was dialysed (MWCO 10,000) against Tris.HCl (50mM, pH8, 37°C) in 3M Gdm.Cl. Trypsin (5% w/v) was added and the mixture was incubated at 37°C for 8h before a second aliquot of trypsin (5% w/v) was added. The resulting mixture was incubated at 37°C for 16h.

HPLC (ABI Aquapore RP300 C_{18}, 220 x 4.6mm, 7μ, A = H_2O, B = CH_3CN, 0.1% TFA; 1ml/min. 10-90 %B over 30 minutes. λ = 214nm). R_t = 10-22 minutes, 24-60%B. MS (MALDI TOF) m/z = 2959 (MH^+, fragment 117-143, K^+ salt), 3807 (MH^+, fragment 140-166, K^+ salt), 4435 (MH^+, fragment 15-53, K^+ salt) and 5961(MH^+, fragment 46-97, Na^+ salt).

Protein sequencing[8]

3ml of denatured protein solution (approximately 0.2mg by UV analysis) was dialysed (MWCO 10,000) against H_2O for 2 weeks changing dialysis solution twice daily. Some protein precipitated from solution and was solubilised in 100% CH_3CN, diluted with H_2O, combined with soluble material from the dialysis and lyophilised to an off white solid.

Yield (~0.2mg).

Sequencing analysis from N-terminus Found, residue$(R)_1$ = Ala, R_2 = Pro, R_3 = Pro, R_4 = Arg, R_5= Leu, R_6 = Ile, R_7 = n.d, R_8 = Asp, R_9 = ?, R_{10} = ?, R_{11} = Val, R_{12} = Leu, R_{13} = Glu, R_{14} = ?, R_{15} = Tyr, R_{16} = Leu, R_{17} = Leu, R_{18} = Glu, R_{19} = Ala $C_{815}H_{1327}N_{344}O_{218}S_4$, Requires, R_1 = Ala, R_2 = Pro, R_3 = Pro, R_4 = Arg, R_5= Leu, R_6 = Ile, R_7 = Cys, R_8 = Asp, R_9 = Ser, R_{10} = Arg, R_{11} = Val, R_{12} = Leu, R_{13} = Glu, R_{14} = Arg, R_{15} = Tyr, R_{16} = Leu, R_{17} = Leu, R_{18} = Glu, R_{19} = Ala.

Copper[lysine]$_2$ (84)[1]

Lys.HCl (83) (36.8g, 0.2mol) was dissolved in water (500ml). The solution was heated to 50°C before basic copper carbonate (66g, 0.3mol) was added. The resulting suspension was refluxed for 3h. The suspension was filtered while hot, to remove the excess copper carbonate, and washed with cold water (30ml). The filtrate was reduced *in vacuo* to yield the *title compound* as a blue solid.

Yield 50g, 83%.

4-^{t}Butoxyphenylmethylacetate (86)[1]

Methyl-4-hydroxyphenylacetate (85) (25g, 0.15mol) was dissolved in DCM (120ml) in a 1l Buchi flask. The solution was cooled to -45°C (acetone/dry ice bath) while stirring and *iso*butylene (~120ml) was condensed into the solution. Concentrated H_2SO_4 (2ml) was added and thoroughly mixed in. The flask was sealed and the solution was stirred at room temperature overnight behind a safety screen. The flask was cooled to -45°C and carefully opened behind the screen. TEA (8ml) was added immediately (confirm alkaline pH) and the solution was allowed to warm to room

temperature while stirring. The solution was then warmed to 35°C to release the excess isobutylene as a gas. The final solution was reduced *in vacuo* to an orange oil which was taken into DCM (100ml). The organic solution was washed twice with 10% sodium carbonate and once with dilute HCl (pH2). The organic layer was dried ($MgSO_4$), filtered and reduced *in vacuo* to yield a pale yellow oil which was used without further purification.

Yield 28g, 84%.

R_f (hexane:EtOAc, 9:1) = 0.6. MS (FAB) m/z = 223 (MH^+). IR (DCM) υ_{max} = 2978-3034 (CH), 1746 (C=O), 1508 and 1609cm^{-1} (C=C). δH ($CDCl_3$, 200MHz), 6.88-7.39 (q, 4H, aromatic CH), 3.79 (s, 3H, OCH_3), 3.64 (s, 2H, CH_2), 1.32ppm (s, 9H, ^{t}Bu). δC ($CDCl_3$, 50MHz), 172.18 (C=O), 154.24 (C-O), 129.55 and 124.09 (CH x 4), 128.62 (quaternary aromatic C-CH_2), 53.29 (OCH_3), 40.31 (CH_2), 29.29ppm (^{t}Bu).

4-^{t}Butoxyphenyl acetic acid (87)[1]

4-^{t}Butoxyphenylmethylacetate (86) (28g, 0.13mol) was taken into acetone:H_2O (500ml, 4:1). The solution was cooled to 0°C and lithium hydroxide monohydrate (10.53g, 0.25mmol) in H_2O (168ml) was added. The solution was stirred for 4h before the acetone was removed *in vacuo*. The resulting solution was washed with ethyl acetate, to remove any unreacted starting material, and the pH was adjusted to 3.5 by slow addition of 2M HCl. The product precipitated from solution and was extracted from the aqueous layer with ethyl acetate, dried ($MgSO_4$) and reduced *in vacuo* to yield the *title compound* as shiny white needles which were recrystallised from ether, filtered and washed with a little cold hexane.

Yield 20g, 80%.

Mpt. 77-79°C. R_f (hexane:EtOAc, 9:1) = 0.14. CH: Found, C 68.94%, H 7.78% ($C_{12}H_{16}O_3$ Requires, C 69.23%, H 7.69%). MS (EI) m/z = 207.9. IR (nujol mull), υ_{max} = 2670-2724 (CH), 1703 (C=O), 1605 and 1508cm^{-1} (C=C). δH ($CDCl_3$, 200MHz), 9.09 (brs, 1H, OH), 6.95-7.19 (q, 4H, aromatic CH), 3.59 (s, 2H, CH_2),

1.33ppm (s, 9H, ^{t}Bu). δC ($CDCl_3$, 50MHz), 177.86 (C=O), 154.33 (C-O), 129.71 and 124.14 (CH x 4), 127.99 (quaternary aromatic C), 40.26 (CH_2), 29.33ppm (^{t}Bu).

Copper[N^{ε}-4-^{t}butoxyphenylacetyl lysine]$_2$ (89)[1]

4-^{t}Butoxyphenylacetic acid (87) (10g, 0.05mol) was dissolved in 1,4-dioxane (80ml) and phenylphthalene (1 drop) was added. The solution was titrated with NaOH (2M), 40ml being required to turn the indicator pink. The solution was reduced *in vacuo* to a white solid which was dried in a dessicator (10.3g). The dry sodium salt (10g, 0.04mol) was stirred in sodium dried benzene (40ml) and freshly distilled oxalyl chloride (3.8ml, 0.04mol) followed by DMF (0.1ml) were added. The mixture was stirred for 3h and the solution was reduced *in vacuo* to yield the acid chloride (88). The acid chloride (8g, 0.03mol) was dissolved in 1,4-dioxane (20ml) and added slowly, over a period of 2h, to a solution of copper[lysine]$_2$ (6g, 0.02mol) in H_2O (50ml) while keeping the pH at 10 with 10% sodium carbonate. The product precipitated from solution as the reaction progressed. The product was filtered, washed with H_2O and dried to yield the *title compound* as a pale blue solid (89).

Yield 8.48g, 27%.

N^{ε}-4-^{t}Butoxyphenylacetyl lysine (90)

Dry, finely powdered copper[N^{ε}-4-^{t}butoxyphenylacetyl-lysine]$_2$ (89) (1.49g, 2.03mmol) was suspended in a stirred solution of EDTA disodium salt (0.77g, 2.03mmol) in H_2O (105ml). The suspension was stirred at room temperature overnight. The suspension was filtered and the white solid was washed with H_2O and left to dry.

Yield 0.33g, 48%

Mpt. 215°C. CHN: Found, C 64.36%, H 8.72%, N 8.26%, ($C_{18}H_{28}N_2O_4$ Requires, C 64.3%, H 8.3%, N 8.3%). MS (FAB), m/z = 337 (MH^+), HRMS (FAB), m/z = 337.21273 (MH^+ $C_{18}H_{28}N_2O_4$ Requires, 337.21272). IR (nujol mull), υ_{max} = 3316 (NH). 1739 and 1646 (C=O), 1589cm^{-1} (C=C). δH (d^6DMSO, 200MHz), 6.77-7.02

(q, 4H, aromatic CH), 2.91-2.95 (d, 2H, CH_2CH), 2.61-2.64 (t, 1H, $^{\alpha}CH$), 1.2-1.34 (m, 8H, CH_2 x 4), 1.14ppm (s, 9H, ^{t}Bu). δC (d^6DMSO, 50MHz), 183.42 and 174.12 (C=O), 152.54 (C-O), 131.13 (quaternary aromatic C), 124.67 and 129.57 (aromatic CH), 55.71 ($^{\alpha}CH$), 39.12, 34.39, 28.06 and 22.22 (CH_2), 27.73ppm (^{t}Bu).

N^{α}-9-Fluorenylmethoxycarbonyl-N^{ε}-4-^{t}butoxyphenylacetyl lysine (91)

N^{ε}-4-^{t}butoxyphenylacetyl-lysine (90) (1.76g, 5.2mmol) was suspended in H_2O (30ml). A solution of TEA (1.1ml, 0.79mmol) in 1,4-dioxane (30ml) followed by solid Fmoc-ONSu (1.68g, 5mmol) were added. The mixture was stirred overnight then diluted with H_2O and acidified to pH2 with 1M HCl. The product which precipitated was extracted with EtOAc, washed with H_2O, dried ($MgSO_4$) and reduced *in vacuo* to an oil. The oil was crystallised from EtOAc/hexane to yield the *title compound* (91) as a white solid which was filtered and dried.

Yield 0.2g, 7.2%.

Mpt. 105°C. CHN: Found, C 70.72%, H 6.86%, N 4.47% ($C_{33}H_{38}N_2O_6$ Requires, C 71%, H 6.8%, N 5%). UV λ_{max} (MeOH) 253 (ε = 17232), 271 (ε = 5108), 290nm (ε = 6043$dm^3mol^{-1}cm^{-1}$). MS (FAB), m/z = 559 (MH^+). HRMS (FAB), m/z = 559.28083, ($C_{33}H_{38}N_2O_6$ Requires, 559.28079). IR (bromoform mull), υ_{max} = 3388-3413 (NH), 2823-3040 (CH), 1739, 1718 and 1655 (C=O), 1509cm^{-1} (C=C). δH (d^6DMSO, 200MHz), 6.85-8.07 (m, 12H, aromatic CH), 4.2-4.31 (m, 3H, $^{\alpha}CH$, $^{\alpha}NH$, $^{\varepsilon}NH$), 3.96-3.98 (m, 1H, fluorenyl CH), 3.01-3.1 (d, 2H, fluorenyl CH_2), 1.36-1.67 (m, 8H, CH_2 lysine), 1.21 (s, 9H, ^{t}Bu). δC (d^6DMSO, 50MHz), 174.14, 170.22 and 169.07 (C=O), 156.3 (quaternary aromatic C-O), 143.94 and 156.3 (quaternary fluorenyl C), 130.22 (quaternary aromatic C), 120.25, 121.51, 123.64, 125.42, 127.2, 127.78, 129.56 and 130.08 (aromatic CH), 65.72 (fluorenyl CH_2), 53.89 ($^{\alpha}CH$), 46.77 (fluorenyl CH), 23.23, 30.54, 38.35 and 41.79 (CH_2, lysine), 28.65ppm (^{t}Bu).

4-Copper[N^{ε}-4-methoxyphenylacetyl lysine]$_2$

4-Methoxyphenylacetic acid (8g, 0.05mol) was suspended in sodium dried benzene (40ml) and freshly distilled oxalyl chloride (3.8ml, 0.04mol) followed by DMF (0.1ml) were added. The mixture was stirred for 3h and the solution was reduced *in vacuo.* The acid chloride (8g, 0.03mol) was dissolved in 1,4-dioxane (20ml) and added slowly, over a period of 2h, to a solution of copper[lysine]$_2$ (6g, 17mmol) in H_2O (50ml) while keeping the pH at 10 with 10% sodium carbonate. The product precipitated from solution as the reaction progressed. The product was filtered, washed with H_2O and dried to yield the *title compound* as a pale blue solid.

Yield 6.9g, 25%.

N^{ε}-4-Methoxyphenylacetyl lysine

Dry, finely powdered copper[N^{ε}-4-methoxyphenylacetyl lysine]$_2$ (6g, 9.24mmol) was suspended in a stirred solution of EDTA disodium salt (3.4g, 9.24mmol) in H_2O (600ml). The suspension was sonicated for two days. The suspension was filtered and the white solid was washed with H_2O and left to dry.

Yield 3.6g, 66.5%

Mpt. 241-242°C. CHN: (Found, C 61.2%, H 7.53%, N 9.52%, ($C_{15}H_{23}N_2O_4$, Requires, C 61.16%, H 7.63%, N 9.39%). MS (FAB), m/z = 295 (MH^+), HRMS (FAB), m/z = 295.16562 (MH^+ $C_{15}H_{23}N_2O_4$, Requires, 295.16578). IR (bromoform mull), υ_{max} = 3721 (NH). 1668 and 1635 (C=O), 1540 and 1513cm^{-1} (C=C). δH (d^6DMSO, 200MHz), 6.69-7.09 (q, 4H, aromatic CH), 3.58 (s, 3H, CH_3), 2.85-2.9 (m, 3H, $^{\alpha}$CH and CH_2), 1.31-1.69 (m, 8H, CH_2 lysine). δC (d^6DMSO, 50MHz), 183.42 and 174.12 (C=O), 156.91 (quaternary aromatic C-O), 129.66 (quaternary aromatic C), 113.86 and 130.01 (aromatic CH), 55.33 ($^{\alpha}$CH), 40.06, 34.23, 31.45 and 22.04 (CH_2, lysine), 26.33ppm (OCH_3).

N^{α}-9-Fluorenylmethoxycarbonyl-N^{ε}-4-methoxyphenylacetyl lysine (93)

N^{ε}-4-Methoxyphenylacetyl-lysine (3g, 10.2mmol) was suspended in H_2O (56ml). A solution of TEA (2.14ml, 15.6mmol) in 1,4-dioxane (56ml) followed by solid 9-fluorenylmethylsuccinimidyl carbonate (3.2g, 9.7mmol) were added. The mixture was stirred overnight at room temperature. The mixture was diluted with H_2O (100ml) and acidified to pH2 with 2M HCl. The product which precipitated was extracted with EtOAc, washed with H_2O, dried ($MgSO_4$) and reduced *in vacuo* to a clear oil. The product was crystallised from EtOAc/ether to yield the *title compound* as a white solid which was filtered and dried.

Yield 3.2g, 60.7%.

Mpt. 143-144°C. CHN: Found, C 69.5%, H 6.35%, N 5.53% ($C_{30}H_{31}N_2O_6$ Requires, C 69.77%, H 6.01%, N 5.43%). UV λ_{max} (MeOH) 255 (ε = 16466), 279 (ε = 4345), 291nm (ε = 5088$dm^3mol^{-1}cm^{-1}$). MS (FAB), m/z = 517 (MH^+). HRMS (FAB), m/z = 517.23388, (MH^+, $C_{30}H_{31}N_2O_6$, Requires, 517.23384). IR (bromoform mull), υ_{max} = 3336-3418 (NH), 2823-3040 (CH), 1743, 1711 and 1693 (C=O), 1509cm^{-1} (C=C). δH (d^6DMSO, 200MHz), 6.76-7.77 (m, 12H, aromatic CH), 5.7-5.75 (m, NH x 2), 4.15-4.43 (m, 3H, $^{\alpha}$CH and CH_2), 3.73 (s, 3H, CH_3), 3.44-3.52 (t, 2H, CH_2), 1.11-1.78 (m, 8H, CH_2 lysine). δC (d^6DMSO, 50MHz), 174.59, 172.88 and 158.78 (C=O), 156.13 (quaternary aromatic C-O), 143.75 and 143.57 (quaternary fluorenyl C), 125.05 (quaternary aromatic C), 119.82, 126.05, 126.95, 127.78 and 127.56 (aromatic CH), 66.9 (fluorenyl CH_2), 55.14 ($^{\alpha}$CH, lysine), 53.41 (fluorenyl CH), 42.39 (CH_2), 46.96 (CH_3), 21.85, 28.71, 31.57 and 39.07ppm (CH_2, lysine).

H-Phe-Gly-Lys(Mepa)-Ala-Gly-OH

The initial resin was Fmoc-Gly-Wang resin (0.4g, 0.25mmol, 0.63mmol/g). The amino acids were coupled using the HOBt method and lysine was incorporated as the 4-methoxyphenylacetamidomethyl (Mepa) derivative. The peptide-resin was cleaved with a solution of TFA (9.5ml) and H_2O (0.5ml) for 1h and the resin was removed by filtration. The resin was washed with TFA (2ml) and the filtrate was reduced *in*

vacuo to an oil. The peptide was precipitated from ether, filtered and washed with ether. The peptide was dissolved in 20% acetic acid/H_2O and lyophilised to yield the *title compound* as a white fluffy solid.

Yield 135mg, 86%.

HPLC (Vydac C_{18}, 250 x 2.6mm, 5μ, A = H_2O, B = CH_3CN, 0.1% TFA; 1ml/min. 10-90 %B over 30 minutes. λ = 214nm). R_t = 13 minutes, 43%B. MS (FAB) m/z = 627 (MH^+).

H-Phe-Gly-Lys(Phenac)-Ala-Gly-OH

The initial resin was Fmoc-Gly loaded-resin (0.45g, 0.25mmol, 0.55mmol/g). The amino acids were coupled using the HOBt method and lysine was protected as its phenylacetamidomethyl (Phenac) derivative. The peptide-resin was cleaved with a solution of TFA (9.5ml) and H_2O (0.5ml) for 1h and the resin was removed by filtration. The resin was washed with TFA (2ml) and the filtrate was reduced *in vacuo* to an oil. The peptide was precipitated from ether, filtered and washed with ether. The peptide was taken into 20% acetic acid/H_2O and lyophilised to yield the *title compound* as a white fluffy solid.

Yield 130mg, 87%.

HPLC (Vydac C_{18}, 250 x 2.6mm, 5μ, A = H_2O, B = CH_3CN, 0.1% TFA; 1ml/min. 10-90 %B over 30 minutes. λ = 214nm). R_t = 11 minutes, 37%B. MS (FAB) m/z = 597 (MH^+).

H-Phe-Gly-Lys-Ala-Gly-OH

Route A: from H-Phe-Gly-Lys(Mepa)-Ala-Gly-OH.

Route B: from H-Phe-Gly-Lys(Phenac)-Ala-Gly-OH.

As a general method, a 12.7mg (0.2mmol) portion of each peptide was added to Tris.HCl buffer (50 mM, 6ml, pH 8.1). Resin bound penicillin acylase (60mg, E.C.3.5.1.11, 1325M.units/Kg, supplied by SmithKline Beecham Pharmaceuticals Ltd) was added and the mixture was allowed to mix gently whilst incubating at 37°C. The progress of each reaction was monitored by HPLC (Vydac C_{18}, 250 x 2.6mm,

5μ, A = H_2O, B = CH_3CN, 0.1% TFA; 1ml/min. 10-40 %B over 30 minutes. λ = 214nm). Route A took 191h for full deprotection while route B took only 51h. On completion of the reaction the resin bound enzyme was removed by filtration and washed with water. The deprotected peptides were isolated by analytical HPLC and lyophilised to yield the *title compound* as a white fluffy solid.

Yield 5mg, 52%.

HPLC (Vydac C_{18}, 250 x 2.6mm, 5μ, A = H_2O, B = CH_3CN, 0.1% TFA; 1ml/min. 10-40 %B over 30 minutes. λ = 214nm). R_t = 5 minutes, 21%B. MS (FAB), m/z = 479 (MH^+). HRMS (FAB), m/z = 479.26184 (MH^+, $C_{22}H_{35}N_6O_6$ Requires, 479.26179). Amino acid analysis (12h hydrolysis): $Gly_2$2.1, $Ala_1$1.1, $Phe_1$1.0, $Lys_1$1.2.

H-Phe-Gly-Lys(Hopa)-Ala-Gly-OH

The initial resin was Fmoc-Gly-Wang resin (0.53g, 0.25mmol, 0.48mmol/g). The amino acids were coupled using the HOBt method. Lysine was protected as the 4-methoxyphenylacetamidomethyl derivative. Tbfmoc was loaded onto the completed resin bound peptide (0.58g), as described in the notes at the beginning of this chapter and a Tbfmoc loading test gave a final resin functionality of 0.354mmol/g. The Tbfmoc loaded peptide-resin (0.3g, 0.12mmol) was stirred in thioanisole (0.5ml), TFMSA (0.9ml) and TFA (2ml), in the absence of light, for 20 minutes at 0°C then for 2h at room temperature. The cleaved resin was removed by filtration and washed with 70% CH_3CN/H_2O. The filtrate was applied to a column of PGC and the eluate was reapplied 3 times to ensure 100% adsorption of the Tbfmoc-peptide. The column was washed copiously with 70% CH_3CN/H_2O. The peptide was cleaved from Tbfmoc by applying 10% piperidine in 70%CH_3CN/H_2O (50ml) to the column. The filtrate was neutralised by addition of AcOH and reduced *in vacuo*. The *title compound* was isolated by HPLC (ABI Aquapore RP300 C_{18}, 250 x 10mm, 20μ, A = H_2O, B = CH_3CN, 0.1% TFA; 5ml/min. 10-40%B over 40 minutes. λ = 214nm. R_t = 9 minutes, 30%B) and lyophilised to yield a white fluffy solid.

Yield 32mg, 21%.

HPLC (Vydac C_{18}, 250 x 2.6mm, 5μ, A = H_2O, B = CH_3CN, 0.1% TFA; 1ml/min. 0-30 %B over 20 minutes. λ = 214nm). R_t = 9 minutes, 22%B. MS (FAB), m/z = 615 (MH^+). HRMS (FAB), m/z = 613.29858 (MH^+, $C_{30}H_{41}N_6O_8$ Requires, 613.29856). Amino acid analysis (12h hydrolysis): $Gly_2$2.0, $Ala_1$1.2, $Phe_1$1.0, $Lys_1$1.0.

Methyl 2-bromobenzoate (96)[2,3]

N-Bromosuccinimide (65.0g, 0.37mol) and a catalytic amount of dibenzoylperoxide were added to a solution of freshly distilled methyl 2-methyl-benzoate (50.0g, 0.33mol) in methyl formate (500ml). The mixture was heated at 50°C under reflux for 2h while irradiating with a halogen lamp (500W), switching the lamp off periodically to control the vigour of the reaction. Hexane (1l) was added and the precipitate was removed by filtration. The filtrate was reduced *in vacuo* to yield the *title compound* as a yellow oil which was used without further purification.

Yield 72.8g, 87.5%.

2-(Methoxycarbonyl)benzyltriphenylphosphonium bromide (97)[2,3]

Methyl 2-bromobenzoate (96) (72.8g, 0.32mol) was dissolved in toluene (500ml) together with triphenylphosphine (66.1g, 0.29mol) and the solution was stirred over the weekend. A white precipitate formed which was filtered, washed with toluene and dried in a dessicator to yield the *title compound* as a white solid.

Yield 100.75g, 64.3%.

Mpt. 234-235°C (lit[3] 230-235°C).

2-(3'-Methoxyphenethenyl) benzoic acid (98)[2,3]

2-(Methoxycarbonyl)benzyltriphenylphosphonium bromide (97) (100.75g, 0.21mol) and *m*-anisaldehyde (29.95g, 0.22mol) were stirred in 1,4-dioxane (350ml) under an atmosphere of dry nitrogen. A solution of DBU.HBr (32.75g, 0.215mol) in 1,4-dioxane (50ml) was added and the mixture was stirred at 50°C for 3h. After cooling, the DBU.HBr which precipitated was filtered and washed with 1,4-dioxane (200ml). Water (160ml) followed by 2M NaOH (250ml) were added to the filtrate and the

mixture was heated under reflux for 45 mins. The solution was concentrated *in vacuo*, H_2O (1l) added and the yellow precipitate which formed was filtered and washed with H_2O (200ml). The combined filtrate was washed with ethyl acetate (2 x 500ml) and ice added to the aqueous layer. The aqueous layer was acidified to pH1 with 2M HCl and a precipitate formed. The precipitate was filtered, taken up into ether, dried ($MgSO_4$) and the solvent was removed *in vacuo* to give the *title compound* as a cream solid.

Yield 41.82g, 78.4%.

Mpt. 110°C. R_f = 0.33 (6:4 EtOAc/hexane). CH: (Found, C 74.56%, H 5.6%. $C_{16}H_{14}O_3$ Requires, C 75.6%, H 5.5%). UV λ_{max} (MeOH) = 292.6 (ε = 26746.2), 209.8 (ε = 25069.8), 257.8nm (ε = 8915.4$dm^3mol^{-1}cm^{-1}$). MS (FAB), m/z = 255 (MH^+). HRMS (FAB), m/z = 255.10212 (MH^+, $C_{16}H_{14}O_3$ Requires, 255.10211). IR (bromoform mull), υ_{max} = 2750-3030 (OH), 1690 (C=O), 1501, 1580, 1597cm^{-1} (C=C). δH ($CDCl_3$, 200MHz) 9.3 (1H, br s, OH), 6.57-8.13 (10H, m, 8 aromatic CH and 2 olefinic CH), 3.82 and 3.55ppm (3H, s, OCH_3 cis/trans isomers). δC ($CDCl_3$, 50MHz) 172.4 and 171.9 (COOH, cis/trans isomers), 159.6 and158.9 (quaternary aromatic $COCH_3$, cis/trans isomers), 112-140 (aromatic CH and olefinic CH), 55.0 and 54.7ppm (CH_3 cis/trans isomers).

2-(3'-Methoxyphenethyl)benzoic acid (99)[2,3]

2-(3'-Methoxyphenethenyl)benzoic acid (98) (41.75g, 0.16mol) was dissolved in methanol (500ml) and added to 10% Pd/C (2.5g). The resulting mixture was stirred under an atmosphere of dry nitrogen until all acid had dissolved. The mixture was hydrogenated overnight and ethyl acetate was added to dissolve the precipitated product. The mixture was filtered through celite to remove the catalyst and this was washed with ethyl acetate to recover any undissolved product. Ethyl acetate was removed *in vacuo* and the residue was recrystallised from methanol to give the *title compound* as a white solid.

Yield 25.46g, 60.64%.

Mpt. 119-121°C (lit.[3] 120-120.5°C), CH: (Found, C 75.04%, H 6.22%. $C_{16}H_{16}O_3$ Requires, C 75%, H 6.25%. UV λ_{max} (MeOH) 280 (ε = 5180), 273 (ε = 4970), 209nm (ε = 35200$dm^3mol^{-1}cm^{-1}$). MS (EI), m/z = 256.9, 238.9 and 121. IR (bromoform mull), υ_{max} = 2750-3022 (OH), 1693 (C=O), 1495, 1581, 1595cm^{-1} (C=C). δH (200MHz, d^6DMSO) 11.8 (1H, br s, OH), 6.74-8.14 (8H, m, aromatic CH), 3.78 and 3.84 (3H, s, OCH_3, two conformations), 2.9-3.59ppm (4H, m, CH_2 x 2). δC (50MHz, d^6DMSO) 173.31 and 173.17 (C=O, two conformations), 159.7 and 159.44 (quaternary aromatic $COCH_3$, two conformations), 144.7, 143.4, 140.0 and 138.64 (quarternary aromatic C), 111.42-133.04 (aromatic CH), 55.09 and 54.95 (CH_3, two conformations), 38.05 (CH_2) and 36.91ppm (CH_2).

2-Hydroxydibenzo[*a,d*]cycloheptadien-5-one (100)[2,3]

2-(3'-Methoxyphenethyl)benzoic acid (99) (25.0g, 0.1mol) was taken in dry benzene (60ml) and oxalyl chloride (17.1ml, 0.2mol) followed by DMF (0.1ml) were added while the mixture was cooled on an ice/salt bath. The mixture was stirred at room temperature under an atmosphere of dry nitrogen for 2h, after which time the acid had dissolved and the gas evolution had ceased. Excess oxalyl chloride and benzene were removed *in vacuo* and the residue taken up in sodium dried benzene (100ml). A suspension of fresh aluminium chloride (37.02g, 0.28mol) was slowly added to this whilst cooling and stirring on an ice/salt bath. The mixture was heated under reflux for 30min then cooled via ice/salt bath before slowly adding 2M HCl (250ml) with mixing. Ether was added to dissolve any precipitated product and a black insoluble solid was removed by filtration. The etherial extract was washed with water, dried ($MgSO_4$) and the solvent removed *in vacuo*. The residue was recrystallised from ether/n-hexane to yield the *title compound* as pale needles.

Yield 18.2g, 91%.

Mpt. 141-142°C (lit.[3] 141-141.5°C). CH: (Found, C 80.4%, H 5.38%, $C_{15}H_{12}O_2$ Requires, C 80.4%, H 5.4%). UV λ_{max} (MeOH) = 302 (ε = 16700), 240 (ε = 12300), 208nm (ε = 28800$dm^3mol^{-1}cm^{-1}$). MS (EI), m/z = 224, 196, 181 and 165. IR (CH_2Cl_2), υ_{max} = 3570 (OH), 1632 (C=O), 1495, 1567, 1603cm^{-1} (C=C). δH

(200MHz, $CDCl_3$), 8.14-6.82 (7H, m, aromatic CH), 7.65 (1H, s, OH), 3.12ppm (4H, s, CH_2 x 2). δC (50MHz, $CDCl_3$) 192.2 (quaternary aromatic CO), 162.0 (quaternary aromatic, C=O), 145.8, 141.9 and 139.2 (quaternary aromatic C), 134.0, 132.2, 130.5, 129.0, 126.7, 116.1 and 114.4 (aromatic CH), 35.6 (CH_2), 34.4ppm (CH_2).

Dibenzo[*a,d*]cycloheptadien-5-one-(2-oxyphenylacetate) (101)[2,3]

2-Hydroxydibenzo[*a,d*]cycloheptadien-5-one (100) (3.0g, 13.4mmol) was taken in acetone (100ml) along with anhydrous potassium carbonate (18.5g, 134.0mmol) and benzyl-2-bromoacetate (3.09g, 13.5mmol). The resulting mixture was stirred at room temperature overnight. The excess potassium carbonate was removed via filtration and the filtrate concentrated *in vacuo*. The residue was taken into ethyl acetate (100ml), washed with a fresh solution of 10% sodium carbonate (3 x 25ml) then H_2O (25ml). The organic layer was dried ($MgSO_4$) and the solvent was removed *in vacuo* to yield the *title compound* as a white solid which was washed with a little ether.

Yield 3.36g, 67.4%.

Mpt. 79-80°C (lit.[3] 81.5-82°C). CH: (Found, C 77.76%, H 5.39%, $C_{22}H_{19}O_4$ Requires, C 77.4%, H 5.4%). MS (FAB), m/z = 373 (MH^+), HRMS (FAB), m/z = 373.14399 (MH^+ $C_{22}H_{19}O_4$ Requires, 373.14397). IR (bromoform mull), υ_{max} = 2625-3020 (aromatic CH and CH_2), 1767 (C=O, ester), 1633 (C=O, ketone), 1495, 1595, 1598cm^{-1} (C=C). δH ($CDCl_3$, 200MHz), 6.68-8.13 (12H, m, aromatic CH), 5.23 (2H, s, OCH_2CO_2), 4.71 (2H, s, CH_2Ph), 3.12ppm (4H, s, CH_2 x 2). δC ($CDCl_3$, 50MHz), 193.13 (quaternary aromatic CO), 168.05 (C=O, ketone), 160.05 (C=0, ester), 144.93, 141.45, 138.74 and 131.64 (quaternary aromatic C), 134.83, 133.64, 132.02, 130.50, 128.36, 126.50, 114.94 and 112.36 (aromatic CH), 66.99, 64.90, 35.55 and 34.50ppm (4 x CH_2).

Dibenzo[*a,d*]cycloheptadien-5-one-(2-Oxyacetic acid) (102)[2,3]

Dibenzo[*a,d*]cycloheptadien-5-one-(2-oxyphenylacetate) (101) (3.35g, 9.01mmol) was dissolved in methanol (45ml) along with 2M NaOH (9.2ml). The mixture was heated under reflux for 1h before removing the methanol *in vacuo*. Water (50ml) was

added to the residue followed by ice and the mixture was acidified to pH1 with 2M HCl. The aqueous layer was extracted with ethyl acetate (3 x 25ml) and the organic layer dried ($MgSO_4$). Ethyl acetate was removed *in vacuo* to a white solid which was washed with a little cold methanol to yield the *title compound* as a white solid.

Yield 2.31g, 91%.

Mpt. 163-164°C (lit.[3] 164-165°C). CH: (Found, C 72.42%, H 4.96%, $C_{17}H_{14}O_4$ Requires, C 72.3%, H 4.96%). MS (EI), m/z = 282, 195, 73. IR (bromoform mull), υ_{max} = 2750-3032 (aromatic CH and CH_2), 1729 (C=O, acid), 1646 (C=O, ketone), 1502, 1604cm^{-1} (C=C). δH ($CDCl_3$, 200MHz), 6.31-7.65 (7H, m, aromatic CH), 4.23 (2H, s, CH_2), 2.72ppm (4H, s, CH_2 x 2). δC ($CDCl_3$, 50MHz), 191.79 (C=O, ketone), 168.97 (C=O, acid), 159.9 (quaternary aromatic CO), 144.0, 140.48, 137.66 and 130.08 (quaternary aromatic C), 132.37, 131.07, 129.25, 127.78, 125.42, 113.81 and 111.55 (aromatic CH), 63.61 (CH_2), 34.53 and 33.4ppm (2 x CH_2).

Dibenzo[*a,d*]cycloheptadien-5-one-(2-oxyacetate-X-PAM) resin

(X = Ala or Leu or Val)

Boc-X-PAM-resin (104) (0.82g, 0.71mmol/g/NH_2, 0.58mmol) was stirred mechanically in 95% TFA/H_2O (10ml) under an atmosphere of dry nitrogen for 30 minutes to remove the Boc group. The resin was filtered, washed with TFA (2ml), DCM, DIEA (2ml), copious DCM then ether and left to dry *in vacuo* (0.73g). Dibenzo[*a,d*]cycloheptan-5-one-(2-oxyacetic acid) (102) (1.35g, 4.8mmol) and HOBt (9.6ml, 4.8mmol) were dissolved in DMF and to this DIC (460.2µl, 2.95mmol) was added. The mixture was sonicated for 30 minutes at room temperature before adding it to the above resin (1.4g, 0.56mmol/g/NH_2, 0.78mmol) preswollen in DMF along with 2,6-lutidine (1.12ml, 9.6mmol). The final mixture was stirred (mechanical stirrer) at room temperature overnight under an atmosphere of dry nitrogen after which time the resin showed only pale blue to the Kaiser test. The resin was filtered, washed with DMF, DCM then ether and dried *in vacuo* to yield 1.42g of product. Quantitative Kaiser test showed only 0.002mmol/g of unreacted amine.

IR (KBr disc) υ_{max} = 3315-3426 (NH), 2080-3840 (aromatic CH and CH_2), 1741 (C=O, ester), 1685 (C=O, amide I), 1655 (C=O, amide II), 1639 (C=O, ketone), 1598cm^{-1}(C=C).

Dibenzo[*a,d*]cycloheptadien-5-ol-(2-oxyacetate-X-PAM) resin

Dibenzo[*a,d*]cycloheptadien-5-one-(2-oxyacetate-X-PAM) resin (1.0g, 0.56mmol) was preswollen in sodium dried THF and $LiBH_4$ (0.01g, 0.56mmol) was added. The resulting mixture was stirred mechanically for 1h at room temperature. The reaction was monitored by IR (KBr disc) and was considered to be complete when the carbonyl stretch at 1645cm^{-1} corresponding to the ketone functionality had disappeared. The resin was filtered, washed with THF, THF/methanol (1:1), THF/acetone (1:1), 1mM HCl in THF then ether and was allowed to dry *in vacuo* (0.8g).

IR (KBr disc) υ_{max} = 3422-3480 (NH), 2840-3080 (aromatic CH and CH_2), 1742 (C=O, ester), 1686 (C=O, amide I), 1654 (C=O, amide I), 1600cm^{-1} (C=C).

Dibenzo[*a,d*]cycloheptadien-5-Fmoc-amino-(2-oxyacetate-X-PAM) resin (amide PAM resin)

Dibenzo[*a,d*]cycloheptadien-5-ol-(2-oxyacetate-X-PAM) resin (0.4g, 0.22mmol) was preswollen in a minimum amount of 1,4-dioxane and stirred mechanically under an atmosphere of dry nitrogen. 9-Fluorenylmethylcarbamate (prepared by the method of Carpino) (0.32g, 1.34mmol) was dissolved in a minimum amount of 1,4-dioxane and added to the resin through a cotton wool filter to remove any insoluble material. Benzene sulphonic acid was added to the mixture while stirring until the pH had been adjusted to 3 and the mixture was stirred for 6h. The resin was filtered, washed with 1,4-dioxane followed by ether and dried *in vacuo* (0.44g, functionality of resin 0.21mmol/g by Fmoc loading test (Section 9.1))

IR (KBr disc), υ_{max} = 3325-3426 (NH), 2840-3080 (aromatic CH and CH_2), 1737 (C=O, ester), 1720 (C=O, amide III), 1685 (C=O, amide II), 1655 (C=O, amide I), 1600cm^{-1} (C=C). UV (20% piperidine/DMF) λ_{max} = 300 and 290nm.

Dibenzo[*a,d*]cycloheptadien-5-Boc-hydrazine-(2-oxyacetate-X-PAM) resin (Boc hydrazide PAM resin)

Dibenzo[*a,d*]cycloheptadien-5-ol-(2-oxyacetate-X-PAM) resin (2g, 1.17mmol) was preswollen in DCM (80ml) and to this *t*-butyloxycarbonyl carbazate (1.08g, 8.17mmol) along with benzene sulphonic acid (0.1g, 0.58mmol) were added. The resulting mixture was refluxed under dry nitrogen for 6h. The mixture was allowed to cool and the resin was filtered, washed with DCM, DMF then ether and dried *in vacuo* (1.86g).

IR (KBr disc) υ_{max} = 3317-3423 (NH), 2840-3080 (aromatic CH and CH_2), 1741 (C=O, ester), 1716 (C=O, carbamate), 1655 (C=O, amide I), 1685 (C=O, amide II), 1600cm^{-1} (C=C).

Fmoc Gly hydrazide PAM resin

To Fmoc-Gly-OH (1.49g, 5mmol) stirred in DCM (30ml) at 0°C was added oxalyl chloride (0.86ml, 10mmol) and DMF (10μl). The mixture was stirred at room temperature for 30 minutes at which time gas evolution had ceased and reactants were completely in solution. Excess oxalyl chloride was removed by rotary evaporation and the residue was redissolved in DCM (10ml) and added to Boc hydrazide PAM resin (1g, 0.527mmol) preswollen in DCM (20ml) and pyridine (5ml). The mixture was refluxed under nitrogen for 4h and the resin isolated by filtration. The resin was washed with DCM, DMF, DCM then ether and dried *in vacuo* to give the title compound as an orange solid (1.2g, functionality 0.23mmol/g by Fmoc loading test (Section 9.1))

IR (KBr disc) υ_{max} = 3329-3430 (NH), 2840-3080 (aromatic CH and CH_2), 1736, 1720, 1701 (C=O), 1655 (C=O, amide I), 1685 (C=O, amide II), 1600cm^{-1} (C=C).

Dibenzo[*a,d*]cycloheptadien-5-one-(2-oxy-methylpolystyrene) resin (107)[2,3]

2-Hydroxydibenzo[*a,d*]cycloheptadien-5-one (100) (5.0g, 22.3mmol) was dissolved in ^{t}BuOH/H_2O (1:1, 50ml) and cesium hydroxide (3.75g, 22mmol) was added. The solution was stirred for 10 minutes before the *t*-BuOH was removed *in vacuo*. The

cesium salt (106) was dried by azeotropic distillation with pyridine (2 x 100ml) and DMF (3 x 100ml). The salt was dissolved in DMF (50ml) and added to chloromethylpolystyrene resin (105) (5.38g, 6mmol, 1.06mmol/g) preswollen in DMF (25ml). The reaction was stirred mechanically under reflux for 4 days at 60°C. The resin was filtered washed with copious DMF, *i*-PrOH, H_2O, DMF and finally *i*-PrOH. The resin was dried *in vacuo* to give the *title compound* as an off white solid (6.48g, 1.0mmol/g).

IR (KBr disc), υ_{max} = 3000-3100 (aromatic CH), 2840-3000 (CH_2), 1638 (C=O, ketone), 1490, 1560 and 1596cm^{-1} (C=C). Cl analysis, Found, <0.3%. Expected, 4.2% for CMP resin.

Dibenzo[*a,d*]cycloheptadien-5-ol-(2-oxy-methylpolystyrene) resin (108)[2,3]

THF (30ml) was added to dibenzo[*a,d*]cycloheptadien-5-one-(2-oxy-methylpolystyrene) resin (107) (0.25g, 0.62mmol) along with $LiBH_4$ (0.11g, 4.9mmol) and the mixture was heated under reflux for 1h, under nitrogen, while stirring mechanically. The reaction mixture was filtered, washed with THF, MeOH then ether and dried *in vacuo* to give the *title compound* as a white solid (0.28g).

IR (KBr disc), υ_{max} = 3428 (OH), 3000-3100 (aromatic CH), 2840-2980 (CH_2), 1490, 1585, 1605cm^{-1} (C=C).

Dibenzo[*a,d*]cycloheptadien-5-Boc-hydrazide-(2-oxy-methylpolystyrene) resin (MP resin) (18)[2,3]

N-*t*-Butoxycarbonylhydrazine (0.18g, 1.33mmol) was dissolved in DCM (20ml) and added to resin (108) (0.18g, 0.22mmol) followed by benzene sulphonic acid (0.04g, 0.22mmol). The mixture was heated under reflux for 24h before it was filtered. The resin was washed with DCM, DMF then ether and dried *in vacuo* (0.23g).

IR (KBr disc), υ_{max} = 3410 (NH), 2980-3100 (aromatic CH), 2840-2980 (CH_2), 1719 (C=O), 1494, 1585, 1600cm^{-1} (C=C).

Fmoc-Gly hydrazide MP resin[2,3]

To FmocGly-OH (0.5g, 1.68mmol) was added DCM (20ml), thionyl chloride (1.2ml, 16.8mmol) and DMF (0.1ml). The mixture was stirred for 30 minutes at which time all reactants had gone into solution and gas evolution had ceased. Excess thionyl chloride was removed *in vacuo* and the residue was redissolved in DCM and added to Boc-hydrazide-MP resin (18) preswollen in DCM (10ml) and pyridine (0.6ml). The mixture was heated under reflux for 6h. The resin was filtered, washed with DCM, DMF then ether and dried *in vacuo* (0.3g). Functionality was shown to be 0.53mmol/g by Fmoc loading test (Section 9.1)

IR (KBr disc), υ_{max} = 3421 (NH), 2966-3100 (aromatic CH), 2840-2928 (CH_2), 1674-1736 (C=O), 1493,1585, 1605cm^{-1} (C=C). UV (MeOH), λ_{max} = 289 and 300nm.

Dibenzo[*a,d*]cycloheptadien-5-Fmoc-amino-(2-oxy-methylpolystyrene) resin (amide-MP-resin) (17)[2,3]

9-Fluorenylmethylcarbamate (prepared by the method of Carpino) (0.45g, 1.86mmol) was dissolved in 1,4-dioxane (30ml) and added to the resin (108) (0.25g, 0.31mmol) through a cotton wool filter to remove any insoluble material. Benzene sulphonic acid (0.05g, 0.31mmol) was added and the mixture was stirred (*via* mechanical stirrer) under dry nitrogen for 6h. The resin was filtered, washed with 1,4-dioxane, DMF then ether and dried *in vacuo* (0.28g, functionality of resin 0.59mmol/g by Fmoc loading test (Section 9.1))

IR (KBr disc), υ_{max} = 3325-3426 (NH), 2840-3080 (aromatic CH and CH_2), 1686-1750 (C=O, urethane), 1485, 1585 and 1605cm^{-1} (C=C). UV (20% piperidine/DMF) λ_{max} = 300 and 290nm.

N^{α}-Hopa-Cys(Acm)-Ser-Asn-Leu-Ser-Thr-Cys(Acm)-Val-Leu-Gly-NHNH2 (SCT 1-10) (109)

The initial resin was Fmoc-Gly-hydrazide CMP resin (0.47g, 0.25mmol, 0.56mmol/g). All amino acids were coupled using the HOBt method and cysteine side chains were

protected with the acetomidomethyl (Acm) group. Asn was protected with Trt and Thr and Ser were protected as stated in the notes. The completed peptide-resin was filtered, washed with DMF, DCM then ether and an additional capping step was performed manually in a sonic bath. Fmoc was removed from the resin (600mg) by sonication in 20% piperidine/DMF for 10 minutes. The resin was filtered and washed with DMF, DCM then ether. The deprotected peptide-resin (0.48g) was swollen in DMF (10ml) and a solution of 4-'butoxyphenylacetic acid (0.21g, 1mmol) with DIC (68µl, 0.5mmol) in DMF (2ml) was added. The mixture was sonicated for 2h before the resin was filtered, washed with DMF, DCM then ether. Quantitative Kaiser test showed only 0.01% unreacted amine. The protected peptide-resin was stirred in EDT (0.5ml), thioanisole (1.17ml), *m*-cresol (0.1ml) and TFA (7.5ml) for 10 minutes at 0°C under nitrogen before TMSBr (1.32ml) was added. The resulting mixture was stirred for 45 minutes. The resin was removed by filtration and the filtrate was reduced *in vacuo*. The peptide was precipitated from ether, filtered and washed with ether. The crude peptide (184mg) was solubilised in 6M Gdm.HCl and purified by HPLC (ABI Aquapore RP300 C_8, 250 x 10mm, 20µ, A = H_2O, B = CH_3CN, 0.1% TFA; 5ml/min. 10-40 %B over 20 minutes. λ = 214nm). R_t = 15 minutes, 36%B and lyophilised to yield the *title compound* as a white fluffy solid.

Yield 90mg, 28%.

HPLC (ABI Aquapore RP300 C_8, 220 x 4.6mm, 7µ, A = H_2O, B = CH_3CN, 0.1% TFA; 1ml/min. 10-90 %B over 30 minutes. λ = 214nm). R_t = 17 minutes, 50%B. MS (MALDI TOF) m/z = 1309.9 (MH^+, Na^+ salt). Amino acid analysis (24h hydrolysis): $Asx_1$1.0, $Thr_1$0.89, $Ser_2$2.2 $Gly_1$1.1, $Cys_2$0.4, $Val_1$1.1, $Leu_2$2.0.

H-Lys(Hopa)-Leu-Ser-Gln-Glu-Leu-His-Lys(Hopa)-Leu-Gln-Thr-Tyr-Pro-Arg-Thr-Asn-Thr-Gly-Ser-Gly-Thr-Pro-NH_2 (SCT 11-32) (111)

The initial resin was N-9-fluorenylmethoxycarbonyl-N'-2-copoly (styrene-1%-divinylbenzene) methoxydibenzocycloheptadien-5-yl amide resin (provided by Bachem, 0.38g, 0.25mmol, 0.65mmol/g). The amino acids were coupled using the HOBt method and lysine was coupled as its 4-'benzyloxyphenylacetamidomethyl

(Bepa) derivative which was given an extended coupling time. Gln and Asn were protected with Trt and the remaining amino acid side-chains were protected as stated in the notes. The completed peptide-resin (1.3g, 0.14mmol, 0.11mmol/g) was sonicated in 20% piperidine /DMF to remove the Fmoc protecting group. The resin was filtered, washed with DMF, DCM then ether. Tbfmoc was loaded onto the dried peptide-resin (1.15g, 0.14mmol, 0.12mmol/g) as described previously. A Tbfmoc loading test gave a final resin functionality of 0.08mmol/g. The peptide was cleaved from the resin by stirring in EDT (2ml), thioanisole (0.5ml), H_2O (0.5ml) and phenol (0.75g) for 10 minutes. TFA (10ml) was then added and the resulting mixture was stirred in the absence of light for 5h. The cleaved resin was filtered off and washed with TFA (2ml). The filtrate was reduced *in vacuo* and the peptide was precipitated from ether, filtered and lyophilised to a gum. The gum was taken into 30% CH_3CN/H_2O and adsorbed onto PGC (0.8g). The PGC was washed copiously with 30% CH_3CN/H_2O to remove impurities and the peptide was cleaved from Tbfmoc by washing the PGC with 10% piperidine in 30%CH_3CN/H_2O. The mixture was separated by preparative HPLC (ABI Aquapore RP300 C_{18}, 250 x 10mm, 20μ, A = H_2O, B = CH_3CN, 0.1% TFA; 5ml/min. 10-50 %B over 20 minutes. λ = 214nm). Three fractions were collected and lyophilised R_t = 11, 14 and 17 minutes, 33, 41 and 48%B. MS (MALDI TOF) of the three fractions showed that there was incomplete cleavage of the lysine protecting group from the 4-benzyloxy to the 4-hydroxyphenylacetamide, thus fractions 2 and 3 were combined and stirred in cleavage mixture, as above, for 3h. The mixture was reduced *in vacuo*, lyophilised and combined with fraction 1 to give the *title compound* as a white fluffy solid.

Yield 21.8mg, 8.7%.

HPLC (ABI Aquapore RP300 C_{18}, 220 x 4.6mm, 7μ, A = H_2O, B = CH_3CN, 0.1% TFA; 1ml/min. 10-90 %B over 30 minutes. λ = 214nm). R_t = 14 minutes, 42%B. MS (MALDI TOF) m/z = 2722.8 (MH^+). MS (FAB), m/z = 2726 (MH^+). HRMS (FAB), m/z = 2725.40374 (MH^+, $C_{122}H_{189}N_{33}O_{38}$ Requires, 2725.39135). Amino acid analysis (24h hydrolysis), $Asx_1$1.1, $Thr_4$3.8, $Ser_2$1.7, $Glx_3$3.3, $Pro_2$2.4, $Tyr_1$0.98, $His_1$1.0, $Gly_2$2.1, $Leu_3$2.8, $Lys_2$2.1, $Arg_1$0.8.

N^{α}-Hopa-Cys(Acm)-Ser-Asn-Leu-Ser-Thr-Cys(Acm)-Val-Leu-Gly-Lys(Hopa)-Leu-Ser-Gln-Glu-Leu-His-Lys(Hopa)-Leu-Gln-Thr-Tyr-Pro-Arg-Thr-Asn-Thr-Gly-Ser-Gly-Thr-Pro-NH_2 (Hopa-SCT1-32) (112)

The peptide hydrazide (109) (14.1mg, 0.01mmol) was taken into DMF (3ml) and cooled to -20°C in an acetone/dry ice bath. 4M HCl in 1,4-dioxane (1.03µl) was added and the mixture was allowed to warm to -15°C. *t*-Butyl nitrite (1.63µl) was added and the temperature was allowed to rise to -10°C and remained at -10°C for 10 minutes to allow formation of the azide (110). A solution of the amino component (111) (20.8mg, 0.008mmol) in precooled DMF (1ml) was added dropwise. A stock solution of DIEA (100µl) in DMF (10ml) was made and an aliquot (532µl) was added. The mixture was incubated at 0°C for 24h and for the first 6h base (22µl) from the stock solution was added every hour. A second aliquot of azide (110) was added after 24h and a third after a further 24h with addition of base as above. The mixture remained at 0°C for 2 days. The coupling was monitored by HPLC (ABI Aquapore RP300 C_8, 220 x 4.6mm, 7µ, A = H_2O, B = CH_3CN, 0.1% TFA; 1ml/min. 10-90 %B over 30 minutes. λ = 214nm) and was stopped after a total of 4 days incubation by adding AcOH (1% v/v, 1ml). The mixture was reduced *in vacuo* to an oil which was taken into 6M Gdm.HCl and purified by HPLC (ABI Aquapore RP300 C_8, 250 x 10mm, 20µ, A = H_2O, B = CH_3CN, 0.1% TFA; 5ml/min. 10-90 %B over 60 minutes. λ = 214nm). Three fractions were isolated R_t = 12, 15 and 18 minutes; 38, 44 and 53%B. Fraction 1 corresponded to unreacted amine, fraction 2 corresponded to unreacted azide and fraction 3 corresponded to the desired product which was lyophilised to yield the *title compound* as a white fluffy solid.

Yield 6mg, 20%.

HPLC (ABI Aquapore RP300 C_8, 220 x 4.6mm, 7µ, A = H_2O, B = CH_3CN, 0.1% TFA; 1ml/min. 10-90 %B over 30 minutes. λ = 214nm). R_t = 21 minutes 61%B. MS (MALDI TOF), m/z = 3975.5. Amino acid analysis (48h hydrolysis): $Asx_2$2.1, $Thr_5$5.5, $Ser_4$4.4, $Glx_3$2.7, $Pro_2$1.7, $Gly_3$3.3, $Cys_2$0.15, $Val_2$2.2, $Leu_5$5.4, $Tyr_1$1.0, $His_1$1.1, $Lys_2$2.4, $Arg_1$1.3.

H-Cys(Acm)-Ser-Asn-Leu-Ser-Thr-Cys(Acm)-Val-Leu-Gly-Lys-Leu-Ser-Gln-Glu-Leu-His-Lys-Leu-Gln-Thr-Tyr-Pro-Arg-Thr-Asn-Thr-Gly-Ser-Gly-Thr-Pro-NH_2 (113)

(112) (6mg, 1.5μmol) was added to Tris.HCl (6ml, 50mM, pH8.1) and resin bound penicillin acylase (138mg, E.C.3.5.1.11, 1325M.units/Kg, supplied by SmithKline Beecham Pharmaceuticals Ltd) was added. The mixture was allowed to mix gently whilst incubating at 37°C. The progress of the reaction was monitored by HPLC (ABI C_8 Aqupore RP300. 100 x 4.6, 7μ, A = H_2O, B = CH3CN, 0.1% TFA, 1ml/min. 10-90%B over 30 minutes. λ = 214nm). Fresh enzyme (50mg) was added after 47h and the reaction was stopped after 48h by addition of AcOH, purified by analytical HPLC (column as above). R_t = 18 minutes, 55%B, 19minutes, 57%B. Fraction 2 was partially deprotected peptide (one Hopa on) and fraction 1 corresponded to the *title compound* and was lyophilised to a white fluffy solid.

Yield 1.5mg, 28%.

HPLC (ABI Aquapore RP300 C_8. 220 x 4.6mm, 7μ, A = H_2O, B = CH_3CN, 0.1% TFA, 1ml/min. 10-90%B over 30 minutes. λ = 214nm). R_t = 18 minutes, 55%B. MS (MALDI TOF), m/z = 3578.5 (MH^+). Amino acid analysis (36h hydrolysis): $Asx_2$1.9, $Thr_4$4.1, $Ser_4$3.7, $Glx_3$3.2, $Pro_2$2.0, $Gly_3$3.2, $Val_1$1.1, $Leu_5$5.0, $Tyr_1$1.2, $His_1$1.0, $Lys_2$2.2, $Arg_1$1.1, $Cys_2$1.2.

H-Ala-Gln-Lys(Tnm)-Glu-Ala-Ile-Ser-Pro-Pro-Asp-Ala-Ala-OH (hEPO 114-125) (117)

The initial resin was Fmoc-Ala loaded Wang resin (0.5g, 0.25mmol, 0.5mmol/g). All amino acids were coupled as their HOCt active esters, lysine was incorporated as its 1,5-dioxaspiro-[5,5]-undecane-3-nitro, 3-methoxycarbonyl (Tnm) derivative and Fmoc was removed from the last amino acid. Gln was protected with Trt and Asp, Ser and Glu were protected as stated in the notes. After synthesis the resin was filtered, washed with DMF, DCM then ether and dried. The peptide was cleaved from the resin by first stirring in EDT (2ml), anisole (0.5ml) and H_2O (0.5ml) for 10 minutes and then after addition of TFA the mixture was stirred for 2h. The resin was

filtered, washed with TFA (2ml) and the filtrate was reduced *in vacuo* to an oil. The peptide was precipitated from ether, filtered and washed with ether. The peptide was solubilised in 10%AcOH/H_2O and lyophilised to yield the *title compound* as a white fluffy solid.

Yield 260mg, 76%.

HPLC (Vydac C_8, 250 x 2.6mm, 5μ, A = H_2O, B = CH_3CN, 0.1% TFA; 1ml/min. 10-90 %B over 30 minutes. λ = 214nm). R_t = 18 minutes, 56%B. MS (MALDITOF), m/z = 1374.66. Amino acid analysis (24h hydrolysis), Found $Asx_1$1.0, $Ser_1$0.8, $Glx_2$1.97, $Pro_2$2.4, $Ala_4$3.9, $Ile_1$1.0, $Lys_1$0.9.

Fmoc-NH-Thr-Leu-Leu-Arg-Ala-Leu-Gly-$NHNH_2$ (Fmoc hEPO 107-113-$NHNH_2$) (118)

The initial resin was Fmoc-Gly hydrazide MP resin (0.55g, 0.25mmol, 0.43mmol/g). The amino acid side-chains were protected as stated in the notes. All amino acids were coupled using the HOCt method and threonine was double coupled. Fmoc was not removed from the last amino acid. The resin (0.63g) was cleaved by stirring in EDT (2ml), thioanisole (0.5ml), H_2O (0.5ml) and phenol (0.75g) for 10 minutes and then after the addition of TFA (10ml) for 3h. The resin was filtered, washed with TFA (2ml) and the filtrate was reduced *in vacuo* to an oil. The peptide was precipitated from ether, filtered and taken up into 70%CH_3CN/H_2O. The peptide was lyophilised and isolated by preparative HPLC (ABI Aquapore RP300 C_8. 250 x 10mm, 20μ, A = H_2O, B = CH_3CN, 0.1% TFA, 5ml/min. 10-90%B over 30 minutes. λ = 214nm). R_t = 17 minutes, 58%B. The purified peptide was lyophilised to yield the title compound as a white fluffy solid.

Yield 98mg, 40%.

HPLC (ABI Aquapore RP300 C_8. 220 x 4.6mm, 7μ, A = H_2O, B = CH_3CN, 0.1% TFA, 1ml/min. 10-90%B over 30 minutes. λ = 214nm). R_t = 26 minutes, 74%B. MS (MALDI TOF), m/z = 979.95. Amino acid analysis (24h hydrolysis), $Thr_1$1.1, $Gly_1$1.2, $Ala_1$1.0, $Leu_3$2.8, $Arg_1$0.95.

H_2N-Thr-Leu-Leu-Arg-Ala-Leu-Gly-Ala-Gln-Lys-Glu-Ala-Ile-Ser-Pro-Pro-Asp-Ala-Ala-OH (hEPO 107-125) (121)

Route A: Coupling Fmoc-hEPO 107-113$NHNH_2$ (118) to hEPO 114-125 (117) *via* convertion to the azide.

Route B: Coupling Fmoc-hEPO 107-113$NHNH_2$ (118) to hEPO 114-125 (117) *via* conversion to the HOCt ester.

Route A and Route B: Fmoc-hEPO 107-113-$NHNH_2$ (118) (49mg, 0.05mmol) was dissolved in DMF (1ml) and the solution was cooled to -20°C. A solution of HCl in 1,4-dioxane (4M, 37.2µl, 0.15mmol) was added and the temperature was allowed to rise to -15°C, *t*-Butyl nitrite (7.3µl, 0.06mmol) was added and the mixture was left for 10 minutes. This allows conversion to the azide (119).

Route A: A solution of hEPO 114-125 (117) (47.5mg, 0.04mmol) in precooled DMF (1ml) was added to the azide (119) followed by DIEA (4µl, 0.03mmol). The mixture was allowed to warm to 0°C and more base (18µl, 0.13mmol) was added at a rate of 4µl every 0.5h.

Route B: HOCt (0.02g, 0.45mmol) was added to the azide (119) and after 10 minutes the temperature was allowed to rise to -10°C. A solution of hEPO 114-125 (117) (47.5mg, 0.04mmol) in precooled DMF (1ml) was added followed by DIEA (29µl, 0.45mmol) and the mixture was allowed to warm to 0°C. More base (23µl, 0.13mmol) was added at a rate of 4µl every 0.5h.

Route A and Route B: The mixture was stored at 0°C overnight. The product (120) started to precipitate from solution as the reaction proceeded. The coupling was monitored by HPLC (Vydac C_8. 250 x 2.6mm, 5µ, A = H_2O, B = CH_3CN, 0.1% TFA, 1ml/min. 10-90%B over 30 minutes. λ = 214nm) and was thought to be complete after 36h due to the disappearance of the amine component (R_t = 18 minutes, 56%B). The reaction was stopped by addition of ether from which the reaction mixture precicpitated. The precipitate was filtered, dissolved in

30%AcOH/H_2O and finally lyophilised to a white fluffy solid. The crude peptide (120) (57.3mg, 24.7μmol) was dissolved in 10% piperidine/70%CH_3CN/H_2O to remove the Fmoc protecting group and simultaneously deprotect the side-chain of lysine. The solution was stored at room temperature overnight before acidifying with AcOH. The mixture was purified by preparative HPLC (ABI Aquapore RP300 C_8. 250 x 10mm, 20μ, A = H_2O, B = CH_3CN, 0.1% TFA, 5ml/min. 10-90%B over 30 minutes. λ = 214nm). The pure peptide was lyophilised to a white fluffy solid.

Yield 11mg, 14.3%.

HPLC (Vydac C_8. 250 x 2.6mm, 5μ, A = H_2O, B = CH_3CN, 0.1% TFA, 1ml/min. 10-90%B over 30 minutes. λ = 214nm). R_t = 20 minutes, 49%B. MS (MALDI TOF), m/z = 1919.56 (MH^+). Amino acid analysis (24h hydrolysis), $Asx_1$0.89, $Thr_1$1.09, $Ser_1$0.67, $Glx_1$1.01,$Gly_1$0.95, $Ala_5$4.96, $Pro_1$1.08, $Ile_1$0.71, $Leu_3$2.98, $Lys_1$0.9, $Arg_1$0.97.

8.4 References

(1) T. D. Pallin, *PhD Thesis*, Edinburgh University.

(2) S. L. Irving, *PhD Thesis*, Edinburgh University.

(3) C. M^{c}Innes, *PhD Thesis*, Edinburgh University.

(4) A. Davison, *PhD Thesis*, Edinburgh University.

(5) L. Jiang, *PhD Thesis*, Edinburgh University.

(6) J.-P. Boissel, W.-R. Lee, S. R. Presnell, F. E. Cohen and H. F. Bunn, *J. Biol. Chem.* 1993, 268, 15983-15993.

(7) H. Edeloch, *Biochemistry*, 1967, **6**, 1968.

(8) A. D. Cronshaw, *Wellmet Sequencing*, Edinburgh University.

Courses and Conferences Attended

Departmental Colloquia, University of Edinburgh, 1993-1996, various speakers.

Organic Research Seminars, University of Edinburgh, 1993-1996, various speakers.

NMR spectroscopy, Dr. I. H. Sadler and Mr. J. A. Parkinson, 1993-94.

Medicinal Chemistry, Professor R. Baker, Merck Sharpe & Dohme, Terling's Park, UK, 1993-96

"Chemical Development in the Pharmaceutical Industry", various speakers, Smithkline Beecham, UK, 1993-96.

Royal Society of Chemistry Perkin Division Scottish Meeting, Aberdeen 1993 and Glasgow 1995, various speakers.

Innovation and Perspectives in Solid Phase Synthesis and Complementary Technologies, Third International Symposium, University of Oxford, 1993, various speakers.

The Biochemical Society Meeting Nº. 651, University of Kent at Canterbury, UK, 1994, various speakers.

Fourth International Symposium Solid Phase Synthesis & Combinatorial Chemical Libraries, University of Edinburgh, 1995, various speakers.

SCI Graduate Symposium Novel Organic Chemistry, Herriot Watt University 1994 and St. Andrews 1996, various speakers: presented a seminar (St. Andrews 1996) on "The Chemical Synthesis of Proteins."

Preparative and Process Scale Liquid Chromatography, University of Cambridge, UK, 1996, various speakers.