The Chemical Synthesis of Proteins

by

Lynda Catherine Draffan

A thesis submitted for the
degree of Doctor of Philosophy

University of Edinburgh
December 1996

To my family

for always being there for me

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.

I would like to thank the Engineering and Physical Sciences Research Council for financial support.

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

I am grateful to Dr D. Campopiano for assistance with SDS-PAGE electrophoresis; to Helen Caldwell for help with Western blot analyses on the interferon-*gamma* products and to Dr S. Kelly at the University of Stirling for circular dichroism analyses.

Finally thanks to all my friends and colleagues in lab 29 who have made my time here so enjoyable. A special mention to Gail Morton for proof reading my thesis and to Robbie Stewart, Nicola Robertson, Helen McSparron, Louise Picken, Raymond Maguire and Ian Heslop for their friendship and support.

This thesis is submitted in part fulfilment of the requirements for 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.

University of Edinburgh

December 1996

Abstract

The development of protecting groups for the N^{ε} function of lysine have been investigated in order to prevent ambiguous coupling during fragment condensation of peptide segments. A number of enzyme cleavable protecting group candidates have been synthesised based on phenylacetyl derivatives and conditions for their enzymatic removal optimised. The syntheses of various test peptides utilising this class of protecting groups have been completed and shown to be suitable for the synthesis of small peptides but inappropriate for longer chain polypeptides such as the protein ubiquitin (76 amino acids).

An alternative chemical approach has been investigated with the synthesis, purification and deprotection of ubiquitin with six out of seven lysine residues protected as the tris(hydroxymethyl)nitromethyl carbamate. The ubiquitin recovered has been purified to homogeneity and characterised.

The total chemical synthesis of deglycosylated human interferon-*gamma* (dhIFN-γ) (143 residues) has been achieved by stepwise solid phase synthesis. The affinity for carbon of the N^{α}–protecting group tetrabenzo[a,c,g,i]fluorenyl-17-methoxycarbonyl (Tbfmoc) has been exploited to expedite the purification of dhIFN-γ. A purification protocol has been developed which afforded homogeneous material.

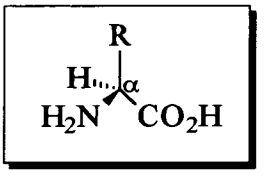
Abbreviations

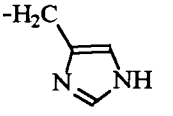
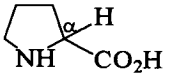
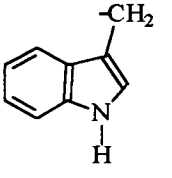
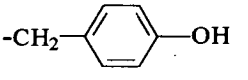
a.a	amino acid
A.A.A	amino acid analysis
ABI	applied biosystems
Acm	acetamidomethyl
AcOH	acetic acid
Ada	adamantyl
2-Adoc	2-adamantyloxycarbonyl
b	broad
Bepa	4-benzyloxyphenylacetyl
benzyl	benzyl
Boc	tert-butoxycarbonyl
BOP	benzotriazolyl-1-yl-oxy-tris-(dimethylamino)-phosphonium hexafluorophosphate
Bum	N(π)-*t*-butoxymethyl
Bupa	4-*t*-butoxyphenylacetyl
CD	circular dichroism
cDNA	complimentary deoxyribonucleic acid
CMP	chloromethylpolystyrene
COSY	correlated spectroscopy
d	doublet
Da	dalton
DBF	dibenzofuran
DCC	N,N'-dicyclohexylcarbodiimide
DCM	dichloromethane
Dde	1-(4,4-dimethyl-2,6-dioxocyclohex-1- ylidine)ethyl
DIC	N,N'-diisopropylcarbodiimide
DIEA	N,N-diisopropyl ethyl amine
dhIFN-γ	deglycosylated human interferon-gamma
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
FAB	fast atom bombardment
Fmoc	9-fluorenylmethyloxycarbonyl
FPLC	fast purification liquid chromatography
Gdm.Cl	guanidinium hydrochloride
HIC	hydrophobic interaction chromatography

hIFN-γ	human interferon-gamma
HMPA	hexamethylphosphoric triamide
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole
HOCt	ethyl 1-hydroxy-1H-1,2,3-triazole-4-carboxylate
HOOBt	3,4-dihydro-3-hydroxy-4-oxo-1,2,3-benzotriazine
Hopa	4-hydroxyphenylacetyl
HPLC	high performance liquid chromatography
HR	high resolution
IFN-γ	interferon-gamma
IL	interleukin
IR	infrared
k	kilo
m	meta
μ	micro
m	multiplet
MALDI	matrix assisted laser desorption ionisation
MBH	4,4'-dimethoxybenzhydryl
MeBepa	4-(4'-methylbenzyloxy)phenylacetyl
Mepa	4-methoxyphenylacetyl
MES	2-(N-morpholino)ethanesulfonic acid
min	minutes
MMA	methylmercaptoacetamide
Mpt	melting point
MS	mass spectrometry
Mtr	2,3,6-trimethyl-4-methoxybenzenesulphonyl
MWCO	molecular weight cut off
nm	nano metres
NMR	nuclear magnetic resonance
NOE	nuclear Overhauser effect
NOESY	Overhauser enhanced spectroscopy
OPfp	pentafluorophenyl
p	para
PAGE	polyacrylamide gel electrophoresis
PAM	*p*-aminomethylated
Pfp	pentafluorophenyl
PGC	porous graphitised carbon
Phenac	phenylacetyl
Pmc	2,2,5,7,8-pentamethylchroman-6-sulphonyl
PMSF	phenylmethylsulphonyl fluoride
PTH	phenylisothiocyanate
PVDF	polyvinyldifluoro
PyBOP	benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate
q	quartet
rhIFN-γ	recombinant human interferon-gamma

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	tris(hydroxymethyl)nitromethane derived protecting group
TOF	time of flight
TOCSY	total correlation spectroscopy
Tris	tris(hydroxymethyl)aminomethane
Trt	triphenylmethyl
Ub	ubiquitin
UV	ultraviolet
Z	benzyloxycarbonyl

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

Chapter two: N^{ε}-Lysine protecting groups

Chapter three: Stepwise Solid Phase Synthesis of deglycosylated human interferon-*gamma*

Chapter four: Experimental

Chapter 1: Introduction

1.1: Solid Phase Peptide Synthesis

1.1.1: General introduction

Demand is increasing for both the total synthesis and chemical modification of peptides and proteins for research, clinical and industrial purposes. This provides a strong incentive for continuous improvement of the general methodology for long chain polypeptide synthesis.

The conceptual breakthrough which stimulated an explosion in long chain polypeptide synthesis was the invention of the solid phase method by Merrifield.[1] In 1963 Merrifield demonstrated the feasibility of his revolutionary idea by synthesising a tetrapeptide through the sequential condensation of protected amino acids from the C-terminus with the first amino acid covalently bound to an insoluble solid polystyrene support. After stepwise assembly the peptide was cleaved from the solid support with simultaneous removal of side-chain protecting groups.

The development of solid phase peptide synthesis (SPPS) greatly facilitated the task of peptide formation. The techniques used in stepwise SPPS are simple, particularly the purification of resin-bound intermediates by filtration and washing. Also, the attachment of a protected peptide to a swollen resin support overcomes the poor solubility of protected peptide intermediates. The shortcomings of stepwise SPPS, including sequence-dependant difficulties in chain assembly, have to a large extent been overcome and are not obstacles to practical peptide synthesis.

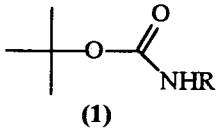
Since its introduction every facet of each stage of solid phase methodology has been subjected to detailed scrutiny with a view to optimisation. The resulting technological improvements in SPPS, together with a parallel refinement of analytical and purification techniques is the reason why the vast majority of all peptides are made by

this method. In addition, the basic principle of the solid phase method has been extended to other areas such as oligonucleotide synthesis and the ever expanding field of combinatorial libraries.

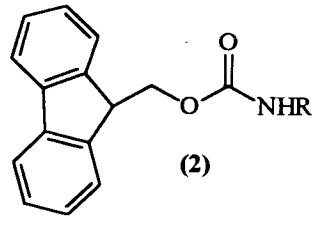
1.1.2: Protecting group tactics

The chemical synthesis of peptides requires the unambiguous formation of an amide bond between two amino acids. It is therefore necessary to protect reactive functionalities other than those involved in the peptide bond formation. Such protection must be reversible without damage to the assembled peptide chain. For stepwise SPPS at least two types of protection are necessary: temporary N^{α} protection, removable after formation of each peptide bond; and semi-permanent side-chain protection removable only after assembly of the complete peptide chain.

There are two main techniques in SPPS which are dictated by the protecting group strategy used. The first, is associated with Merrifield,[2] and involves N^{α} tert-butyloxycarbonyl (Boc) **(1)** protected amino acids in combination with benzyl-based side-chain protecting groups. Whereas the second employs N^{α} 9-fluorenylmethoxycarbonyl (Fmoc)[3] **(2)** protected amino acids and tert-butyl based side-chain protection.



(1)



(2)

The fundamental operations of the protecting group strategies in Boc and Fmoc SPPS are shown in Figure 1.

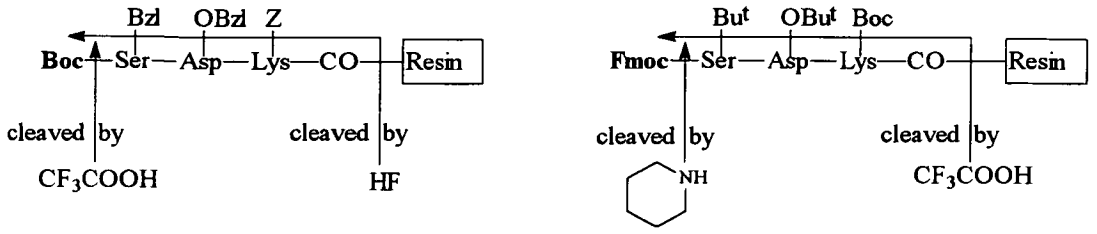


Figure 1 Protecting group strategies in Boc and Fmoc SPPS
Z group=benzyloxycarbonyl; Bzl=benzyl; Bu^t= tert-butyl; Boc= tert-butyloxycarbonyl and Fmoc= 9-fluorenylmethoxycarbonyl

The N^{α} Boc, benzyl-based side-chain protection strategy relies on the principle of graduated acid lability of the protecting groups. The Boc group is cleaved by trifluoroacetic acid while the more stable benzyl-based protecting groups require stronger acids such as hydrogen fluoride or trifluoromethanesulphonic acid. The main problems encountered with Boc SPPS include the incomplete orthogonality between the temporary protecting groups and the semi-permanent protecting groups (which include the peptide-polymer link, cleavage of which would result in loss of peptide). In addition, the use of strong acids not only requires specialised appartatus but may have deleterious effects on the peptide product.

The development of the base labile Fmoc group[3] and its incorporation into automated SPPS[4,5,6] provided an alternative to Boc synthesis and a truly orthogonal strategy. The Fmoc-based strategy uses different mechanisms for removal of the N^{α} and side-chain protecting groups. The N^{α} Fmoc group is removed with the mild base, piperidine and the tert-butyl based side-chain protecting groups are cleaved with trifluoroacetic acid. The general strategy for Fmoc SPPS is outlined in Figure 2.

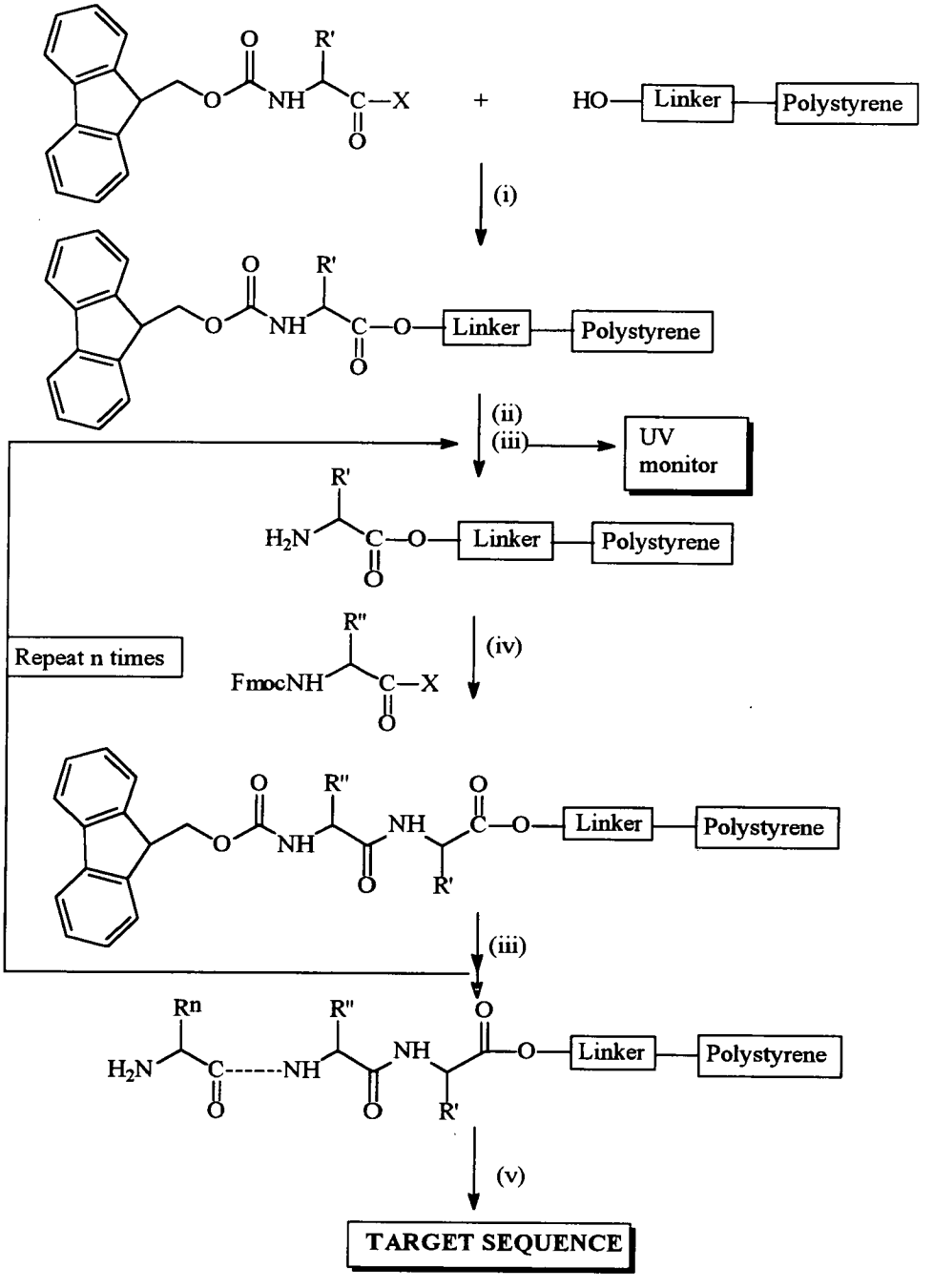


Figure 2 General strategy of N^{α} Fmoc SPPS: (i) attachment of first amino acid to the resin; (ii) capping with acetic anhydride; (iii) deprotection of N^{α} Fmoc with piperidine; (iv) coupling of next N^{α} protected amino acid and (v) simultaneous cleavage of protecting groups and peptide-resin linkage with aqueous trifluoroacetic acid

1.1.3: Role of the resin support

The term 'solid phase' is actually misleading since the reactions do not take place in, or on the surfaces of, solid phases but inside a swollen polymeric network. Several requirements for the solid support have been identified[7] and are summarised below.

The swollen polymer resin must:

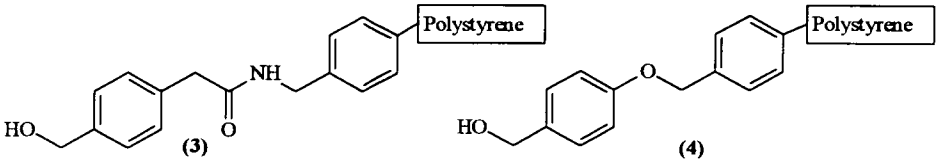
- possess functionality which allows attachment of the peptide chain
- be stable to all reaction conditions
- aid diffusion of reagents and removal of waste products
- exert a strong solvating effect on the attached peptide chains
- minimise aggregation between peptide chains

The solid support, introduced by Merrifield, is a unique copolymer of polystyrene-divinylbenzene which has been chloromethylated and is in the form of beads. The C-terminal N^{α} Boc protected amino acid salt of the target peptide is then attached to the polymer by nucleophilic displacement. The beads of resin are swollen in an organic solvent, such as DCM and peptide synthesis takes place inside the swollen polymeric matrix. The swollen peptide resin is approximately 80-90% solvent by volume, and there are on the order of 10^{12} growing peptide chains per polymer bead.

The Merrifield crosslinked-polystyrenes have found widespread use in both Fmoc and Boc SPPS. The main drawback with the Merrifield cross-linked polystyrenes is that the favoured solvent for swelling the polymeric matrix (DCM) does not fully solvate the conjugated peptide chain. Good solvation is crucial because it allows free access of reagents to reactive sites resulting in an increase in reaction rates and coupling yields for each step. With this objective Sheppard[8] reasoned that a resin that was more compatible with the polarity of the growing peptide chain would be beneficial. This led to the introduction of polyamide resins which are now employed in Fmoc-SPPS.[9,10]

1.1.4: Resin linkages

Refinements in SPPS have led to the introduction of specific linkers between the solid support and the peptide chain which can be exploited to suit the required synthetic strategy. In Fmoc SPPS a more acid labile peptide resin linkage can be used compared to the Merrifield PAM[11] resin (3) used in Boc chemistry. The most popular linker used in Fmoc SPPS is the 4-alkoxybenzylalcohol (Wang) linker[12] (4) which is cleaved in mild acid (50% trifluoroacetic acid in aprotic solvents).



The demand for fully protected peptide fragments for convergent synthesis has led to the introduction of a new class of linker in order to leave the t-butyl based protecting groups in place (Figure 3). To obtain the protected fragments these linkers are cleaved under extremely mild conditions such as very dilute trifluoroacetic acid[13] (5), acetic acid[14] (6) and fluoride ions[15] (7).

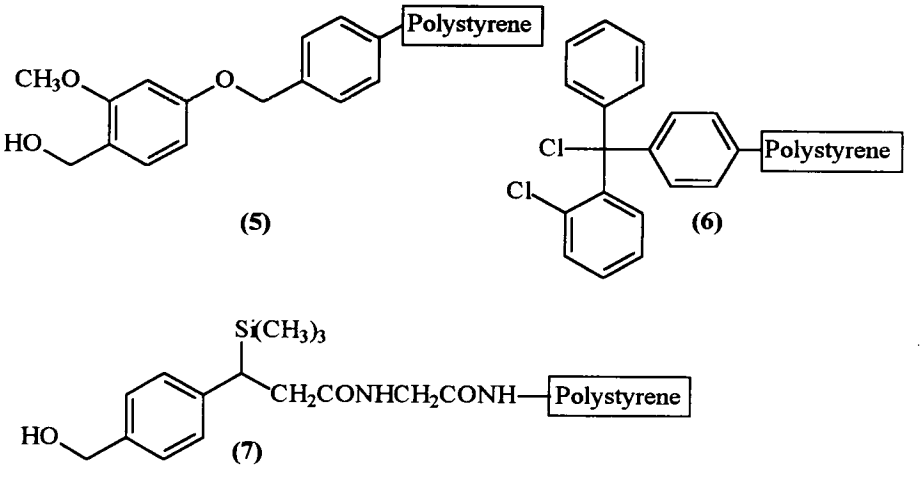
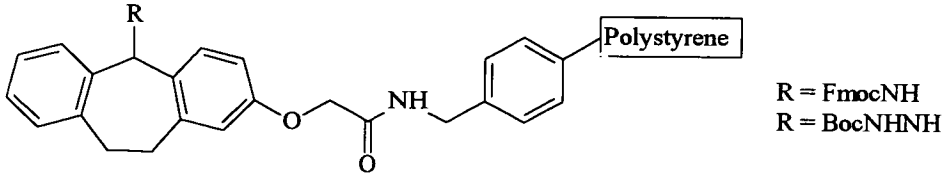


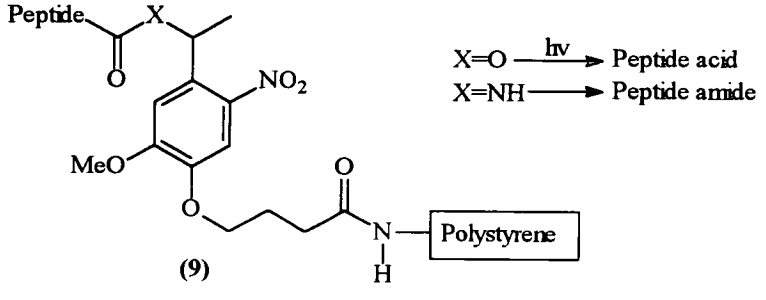
Figure 3 Resin linkers cleaved under mild conditions

The above resins all give the C-terminal carboxylic acid on cleavage, however alternative functionalities can be generated by modifying the linker. For example C-terminal amides[16] and hydrazides[16,17] are obtained by utilising a linker such as the 5 substituted 9-Fmoc aminodibenzocycloheptadiene derivative (**8**) which is cleaved with dilute trifluoroacetic acid. These C-terminal functionalities are important since many biologically significant peptides have a C-terminal amide and C-terminal hydrazides are required for convergent synthesis via azides.



(**8**)

Photolabile linkers have found widespread use throughout solid phase synthesis.[18] Photolytic release of peptides offers a method of orthogonal cleavage which complements the traditional use of TFA or HF. Recently a new class of ortho-nitrobenzyl photolabile supports (**9**) have been introduced with the improvement over previous linkers in that they are cleaved faster and liberate methionine-containing peptides with minimal oxidation.[19]



(**9**)

1.1.5: Activation and coupling: peptide bond formation

Efficient peptide bond formation in SPPS requires the generation and aminolysis of activated carboxy derivatives via the direct electrophilic acylation mechanism.[20,21] This requires the replacement of the carboxylic acid hydroxyl group with an

electronegative substituent (X), which enhances the electrophilicity of the carbonyl carbon thus facilitating attack by the amino component. Activation is necessary because ordinary carboxylic acids simply form salts with amines at ambient temperature; the transformation of these salts directly into amides requires severe heating which is quite incompatible with the presence of other functional groups.

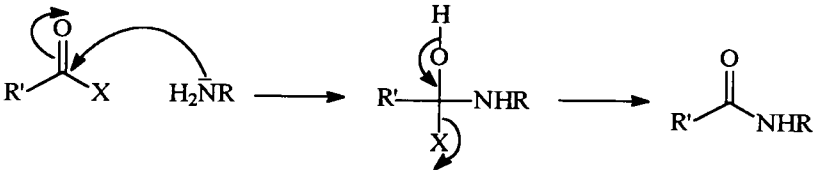


Figure 4 Direct electrophilic acylation mechanism[20,21]

Ideal chemistry would allow peptide bond formation to be carried out rapidly and quantitatively under mild conditions, avoiding side-reactions such as racemisation and generating only easily removed by-products. This challenge has stimulated great interest, resulting in many and diverse coupling reagents being proposed.[22] Although a large number of these reagents have shown subsequent feasibility only a few have been routinely applied in Fmoc SPPS and it is these coupling methods that are discussed here.

1.1.5.1: Racemisation during activation and coupling

A suitable activation and coupling reagent is one which retains the chiral integrity of each amino acid on coupling, thus avoiding racemisation. Racemisation is a highly undesirable side-reaction which has been closely studied with a view to defining conditions under which it is minimal. It is almost exclusively a base-induced reaction and in practice is only a matter for serious concern at the activation and coupling stages of a synthesis. Fortunately the conditions employed during Fmoc-SPPS mean that the risk of racemisation occurring by direct enolisation is very slight indeed. However, the main cause of racemisation in peptide synthesis is via the ozazolone mechanism [23] (Figure 5).

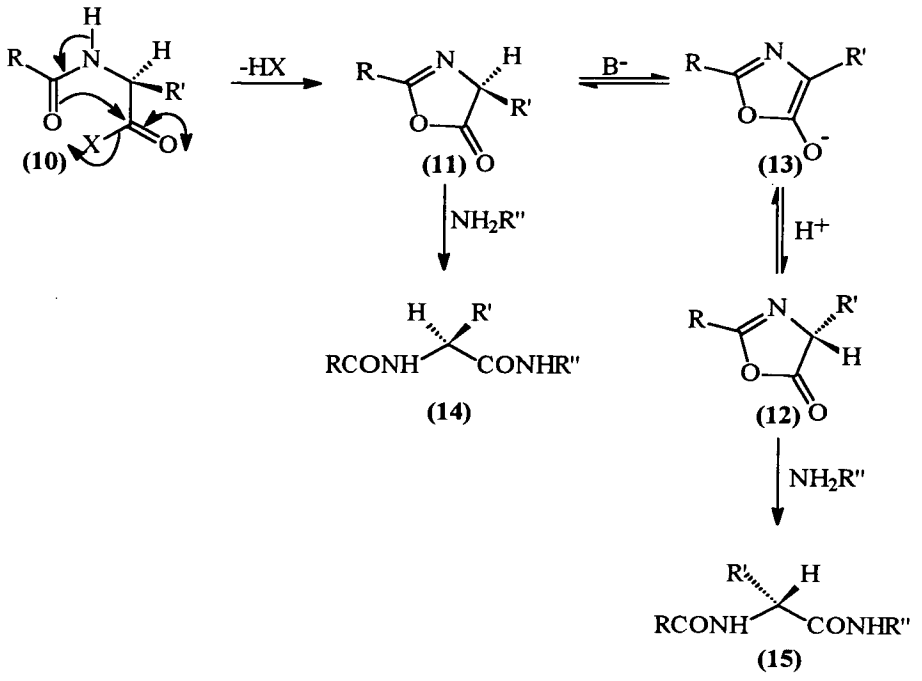


Figure 5 Base catalysed racemisation of acyl amino acids via the oxazolone mechanism[23]

Racemisation occurs by the cyclisation of activated acyl amino acids (**10**), on treatment with base, to give oxazolones (**11**) and (**12**). These oxazolones are themselves activated carboxylic acid derivatives and thus can react with an amine to form a peptide. However, racemisation via the stabilised anion (**13**) occurs more rapidly than aminolysis and both possible epimers (**14**) and (**15**) are formed. Oxazolone mediated racemisation is fastest with strongly electron withdrawing groups, X, and certain methods of coupling have been shown to have lower levels of racemisation than others.

1.1.5.2: Carbodiimides

Since the introduction in 1955 of N,N'-dicyclohexylcarbodiimide[24] (DCC) (**16**) and subsequently N,N'-diisopropylcarbodiimide[25] (DIC) (**17**) these carbodiimide reagents have been used extensively in peptide synthesis. Carbodiimides are very popular coupling agents for peptide bond formation either through *in situ* activation of the

carboxyl component in the presence of the amino component, or through the prior formation of a symmetric anhydride or an active ester.

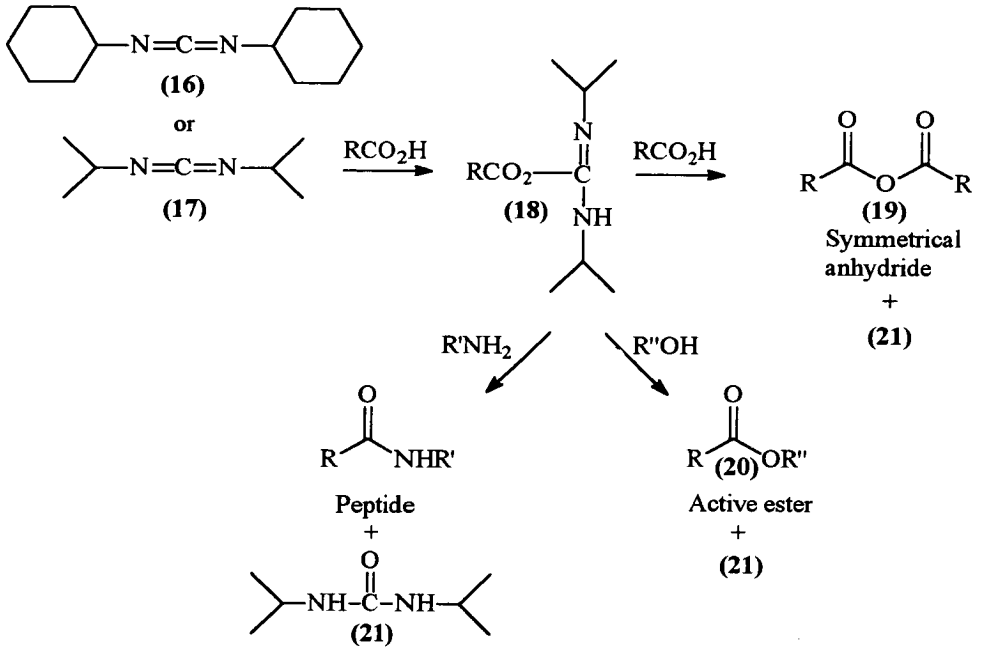


Figure 6 Use of carbodiimides for carboxyl activation

In all cases, the primary activating event is addition of the carboxyl group to the carbodiimide functionality to give an O-acylisourea (**18**) which is a potent acylating agent. O-Acylisourea formation is rapid, leading to peptide either by immediate aminolysis or via a symmetrical anhydride (**19**) or active ester (**20**), with the concomitant formation of a urea by-product. The N,N'-diisopropyl urea (**21**), by-product after DIC activation, is fully soluble in the solvents used in SPPS compared to the less soluble N,N'-dicyclohexylurea, thus DIC is the preferred reagent for automated synthesis.

The high reactivity of the O-acylisourea acylating agent can cause several serious side-reactions, such as racemisation (section 1.1.5.1), carboxamide dehydration of asparagine (**22**) (Figure 7) and glutamine and also the formation of N-acylureas (**23**) (Figure 8).

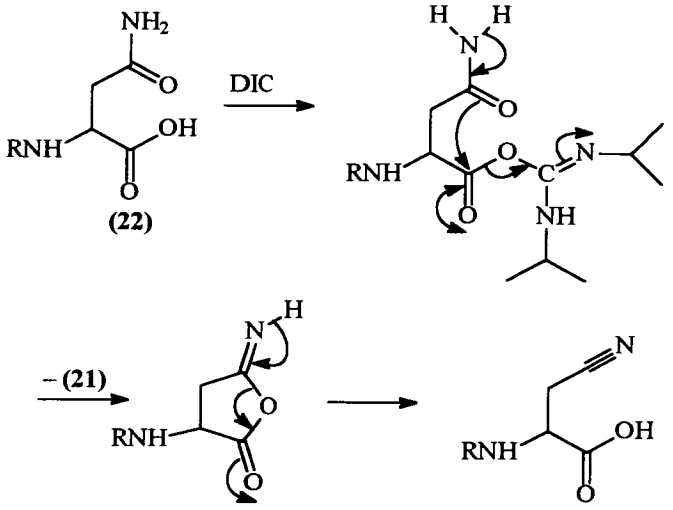


Figure 7 Carboxamide dehydration of asparagine

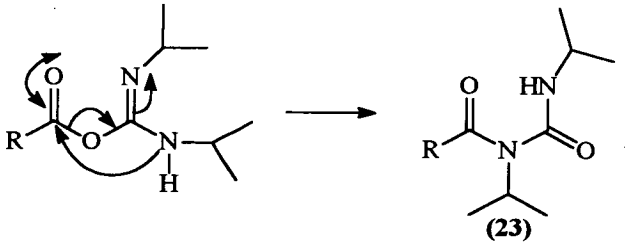


Figure 8 N-Acyl urea formation

These undesirable side-reactions can be alleviated by reacting the O-acylisourea with oxygen nucleophiles to form symmetrical anhydrides and active esters which are less reactive acylating agents.

1.1.5.3: Symmetrical anhydrides

A breakthrough in SPPS methodology occurred with the development of fully automated synthesis involving the separate formation of symmetric anhydrides or active esters immediately prior to use.[26] This provided a new dimension in flexibility of activation chemistry for SPPS; allowing the use of appropriately activated amino acid derivatives in optimal solvents.

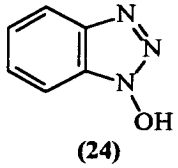
The advantages of symmetrical anhydrides (19) are that they are easily prepared (reaction of two equivalents of Fmoc-amino acid with one equivalent of carbodiimide)

and they offer unambiguous aminolysis. However the major drawbacks to their use are that firstly only half of the expensive Fmoc-amino acid is incorporated into the peptide and secondly carbodiimide promoted side-reactions, such as the dehydration of asparagine/glutamine (Figure 11) and histidine racemisation,[27] are still apparent.

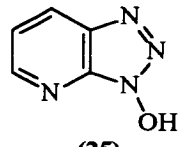
1.1.5.4: Active esters

The active esters (20) of N^{α} protected amino acids are easily prepared from one equivalent of Fmoc-amino acid and one equivalent of the appropriate alcohol in DIC. Thus with this coupling method there is no 'wastage' of amino acid. Also side-reactions, which normally accompany the coupling step, are minimised due to the activated esters being less reactive than the active intermediates discussed previously.

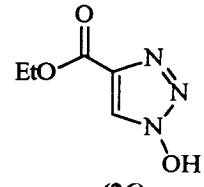
A commonly used and extremely effective coupling agent is 1-hydroxybenotriazole (HOBt) (24) which was introduced by Konig and Geiger to combat racemisation.[28] Analogues developed as a consequence of the highly efficient coupling of HOBt esters include, 1-hydroxy-7-azabenzotriazole (HOAt) (25)[29] and ethyl 1-hydroxy-1H-1,2,3-triazole-4-carboxylate (HOCt) (26).[30,31,32]



(24)



(25)



(26)

Of particular note is the HOCt analogue designed at Edinburgh by Davidson.[30] HOCt has been proven to be a superior coupling reagent to HOBt,[33] allowing the stepwise SPPS of previously unattainable large proteins such as interferon-γ (chapter 3) the obese gene[31] and erythropoietin.[33]

1.1.6: Fmoc SPPS side chain protection

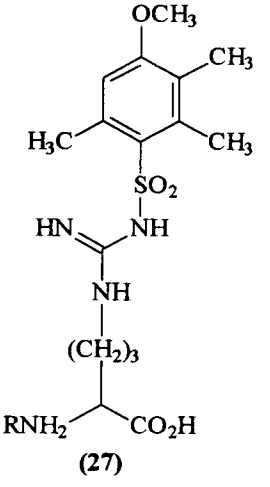
Fmoc-SPPS utilises side chain protection based on t-butyl based protecting groups with the notable exceptions of arginine, cysteine and histidine (Table1).

Amino acid	Protecting group
Arg	Mtr / Pmc
Asn / Gln	MBH / Trt
Asp / Glu	Bu^t esters
Cys	Acm / Trt / SBu^t
His	π-Bum / τ-Trt
Lys	Boc
Ser / Thr / Tyr	Bu^t ethers
Trp	Boc

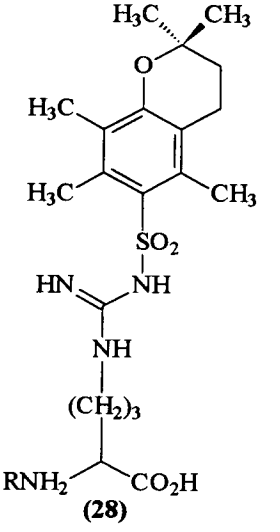
Table 1 Side chain protection in Fmoc-SPPS

1.1.6.1: Arginine protection

Protecting groups for the strongly nucleophilic guanidino side-chain group of arginine have been developed based on ring substituted arylsulphonyl derivatives, such as 4-methoxy-2,3,6-trimethylbenzenesulphonyl (Mtr)[34] **(27)** and 2,2,5,7,8-pentamethylchroman-6-sulphonyl (Pmc)[35,36] **(28)**.



(27)



(28)

The Pmc group is preferred since it is more acid labile. Although the Pmc group is cleaved faster, extended TFA cleavage times are still necessary especially in multiple arginine sequences. This conflicts with the desire to reduce cleavage times since prolonged treatment with TFA may cause shearing of the polypeptide chain. Research is currently underway at Edinburgh on the design of a new protecting group for arginine[37] which will be cleaved faster than Pmc.[35,36]

1.1.6.2: Histidine protection

Histidine (29) is especially prone to the problem of racemisation which can occur via the formation of an oxazolone,[23] (section 1.1.5.1) or through pathways involving the side chain imidazole ring of the residue.[27] Although the imidazole ring is only a weak base it is strong enough to cause intramolecular proton abstraction by the π nitrogen yielding (30) which can then form either the D (31) or L configuration (29).

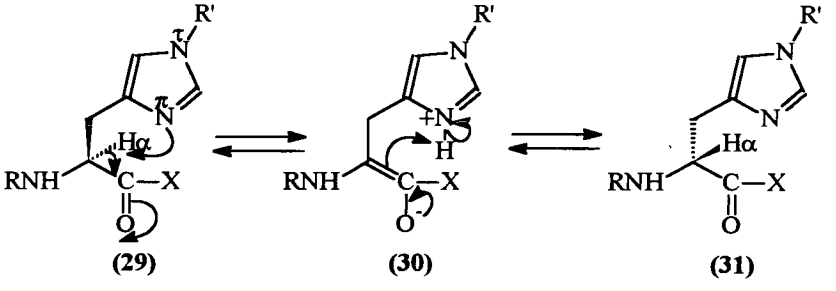
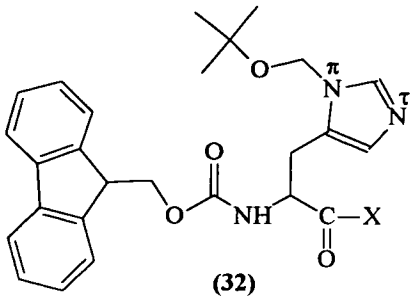
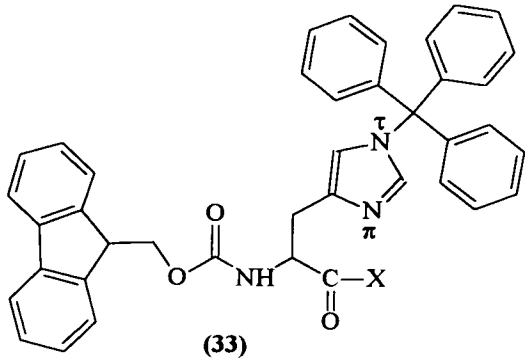


Figure 9 Racemisation of histidine

Blocking the π nitrogen with tert-butoxymethyl (Bum) (32) has been shown to virtually eliminate histidine racemisation.[38] Although Fmoc-His (Bum) is commercially available it is often impure and very expensive due to the difficulties in synthesising π-protected histidine. A cheaper alternative is to protect the τ-nitrogen with the bulky trityl[39] (33) group which significantly reduces the basicity of the π nitrogen and therefore suppresses proton abstraction from the α-carbon. Although, racemisation is not completely eliminated it can be further reduced by coupling with the mildly acidic HOBt active ester.



(32)



(33)

1.1.7: Fmoc SPPS: monitoring

In solid phase synthesis, purification consists only of filtration and washing of the resin-bound intermediates. Due to this, it is necessary to attempt to drive all reactions involved in the chain assembly to completion. Information on the coupling efficiencies at each cycle allow the chemistry to be improved in general and, if necessary, for a particular synthesis.

Deprotection of the N^{α} Fmoc protected amino acid (**34**) with base (20% piperidine/DMF) by a β-elimination mechanism results in a dibenzofulvene (**35**) which is quenched with excess piperidine giving a fulvene-piperidine adduct (**36**). This results in a strong UV absorbance at the isosbestic point of 302nm and thus monitoring of the deprotection solution provides an estimation of the % coupling of each cycle.[40]

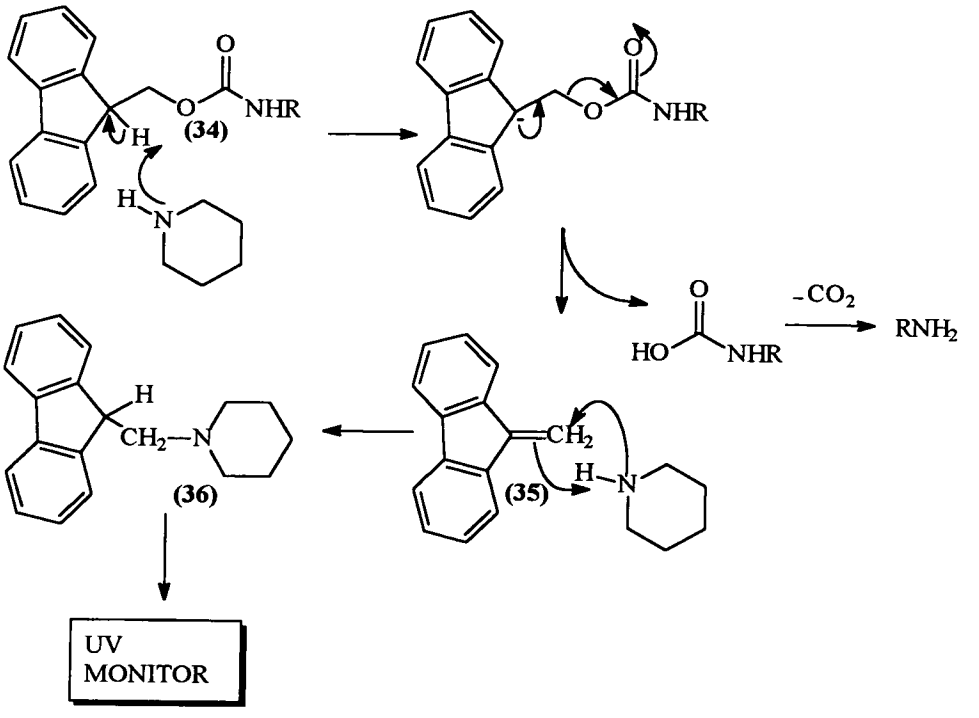


Figure 10 Mechanism of Fmoc deprotection

1.1.8: Fmoc SPPS: acidolytic cleavage

After the protected peptide chain has been assembled on the resin support, all the protecting groups must be removed and the peptide cleaved from the resin support releasing the free peptide into solution. Cleavage of peptides containing only the standard N^{α} Fmoc and t-butyl based side-chain protecting groups can be achieved by treating with 95% trifluoroacetic acid (TFA)/5% water.

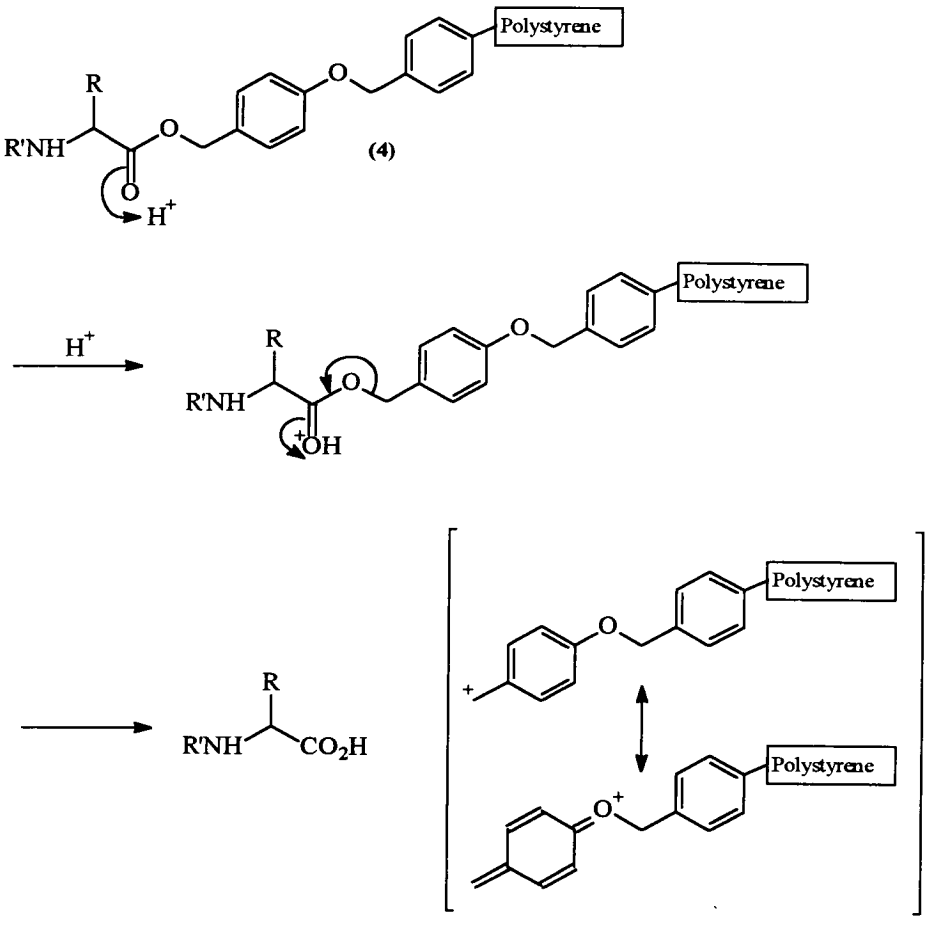


Figure 11 Acidolytic cleavage of peptide from linker

However, the cleavage reaction generates stable cations which can attack the electron rich side-chains of the amino acids tryptophan, tyrosine and methionine. Such side-reactions can be minimised by the addition of scavengers including water, 1,2-ethanedithiol (EDT), thioanisole, anisole and phenol. These reagents work by trapping the TFA liberated carbocations before they can interact with the susceptible side chains of α-amino acids in the peptide. The scavengers also act as nucleophiles in the cleavage step by shifting the mechanism of cleavage from S_N1 towards S_N2 for certain less acid labile protecting groups e.g. Pmc protected guanidino side chain of arginine (**28**). A possible mechanism is outlined in Figure 12.

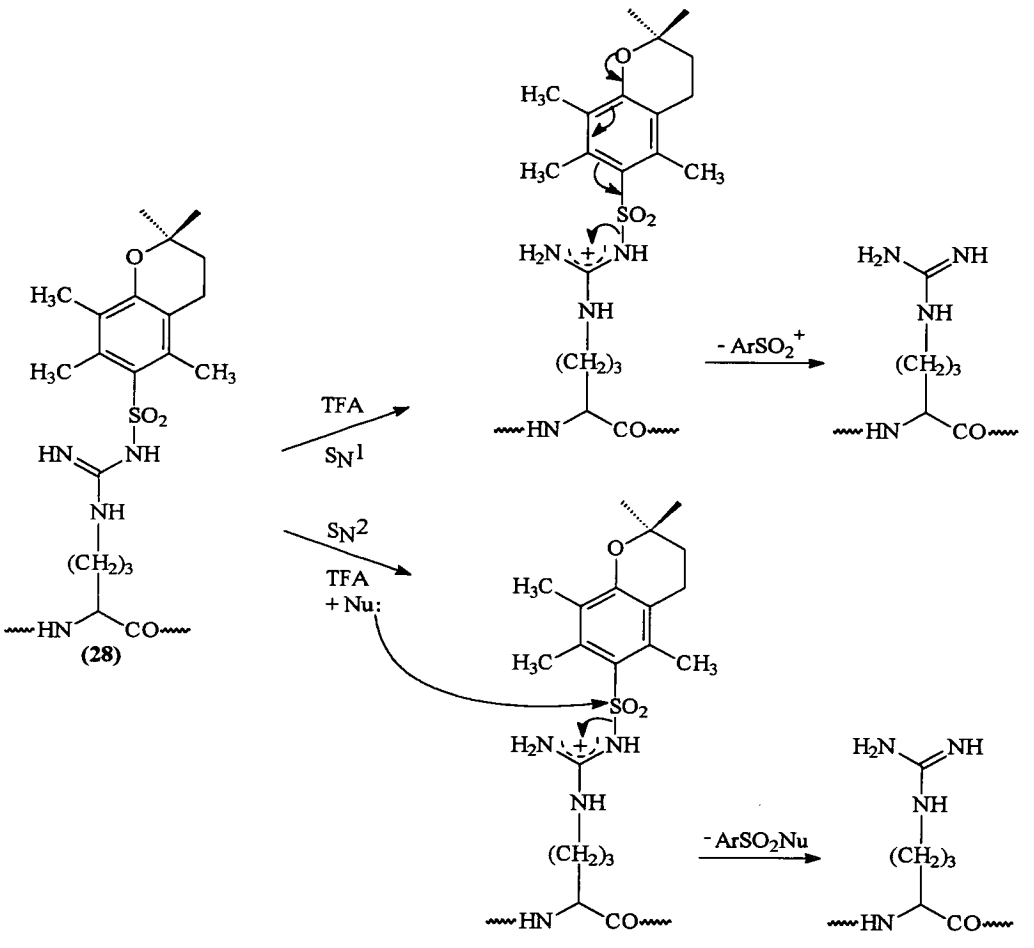


Figure 12 Acidolytic cleavage of Arg(Pmc)

1.2: Chemical Synthesis of Proteins

1.2.1: General introduction

Recent advances in recombinant DNA technology have made the synthesis of proteins possible using biological methods, but the ability to prepare such compounds using purely chemical methods remains extremely important. Chemical synthesis allows the replacement of key residues in a protein with unnatural amino acids whose chemical or structural properties can be tailored to probe specific aspects of protein function. In addition, isotopically labelled amino acids can be incorporated at one or more positions, thereby allowing highly focused structural studies of the molecule using

techniques such as multi-dimensional NMR spectroscopy. Fixed secondary structural elements, such as mimics of β-turns, can be introduced and the effects of folding, function and stability investigated. Most importantly chemical synthesis permits the systematic variation of structure to redesign known proteins with the aim of developing second generation products for therapeutic use. Clearly, the generation of such protein analogues requires a synthetic strategy over which the researcher has complete control at every step in the construction of the target molecule. Fortunately, total chemical synthesis meets this criterion and thus exists as a complementary approach to genetic engineering in the exploration of protein structure and function.

The fundamental synthetic protocols developed for the total chemical synthesis of proteins are combinations of steps selected from linear solid phase peptide synthesis or condensation reactions (either enzymatical or chemical) of presynthesised peptide fragments.

1.2.2: Stepwise synthesis of proteins

The stepwise synthesis of proteins by the solid-phase method has always been seen as a major objective in peptide chemistry.[41] However, the main obstacle to the success of such syntheses is the generation of unwanted truncated and deletion sequences and the subsequent inability to separate these from the desired target protein.[42] The physical and chemical similarities of the desired sequence and such impurities mean that conventional chromatographic techniques are not usually of high enough resolution to afford separation when used individually and lead to poor yields when used in combination. These apparent limitations to stepwise SPPS have been addressed at Edinburgh by combining (i) the effective coupling reagents DIC/HOCt, (ii) a capping step within each cycle using acetic anhydride and (iii) the introduction of a labile N^{α}-protecting group with enhanced chromatographic properties which permit affinity-based purification of the target sequence.

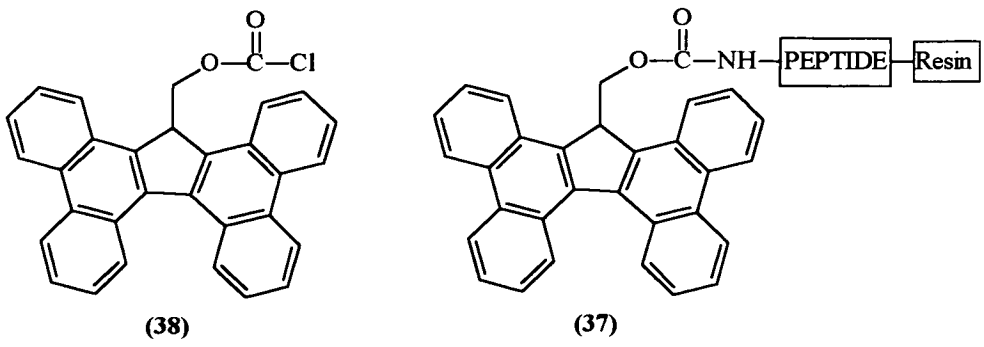


Figure 13 Tbfmoc chloroformate and Tbfmoc protected peptide

The base labile N^{α}-protecting group tetrabenzo[a,c,g,i]fluorenyl-17-methoxy carbonyl (Tbfmoc)[43] (**37**) was designed in this laboratory to simplify the purification of synthetic peptides and proteins produced using Fmoc/t-Butyl-based SPPS. The Tbfmoc group is incorporated onto the N-terminus of the resin-bound polypeptide sequence by reaction with the chloroformate (**38**). After cleavage of the Tbfmoc-labelled polypeptide from the resin, the desired sequence can then either be purified by exploiting the high affinity of the large, essentially planar, aromatic Tbfmoc group for porous graphitised carbon (PGC) or by using the hydrophobic nature and specific UV absorbance (364nm) of the Tbfmoc system to simplify the chromatographic purification by HPLC. Figure 14 outlines the general philosophy of an integrated synthetic/purification protocol for protein synthesis.

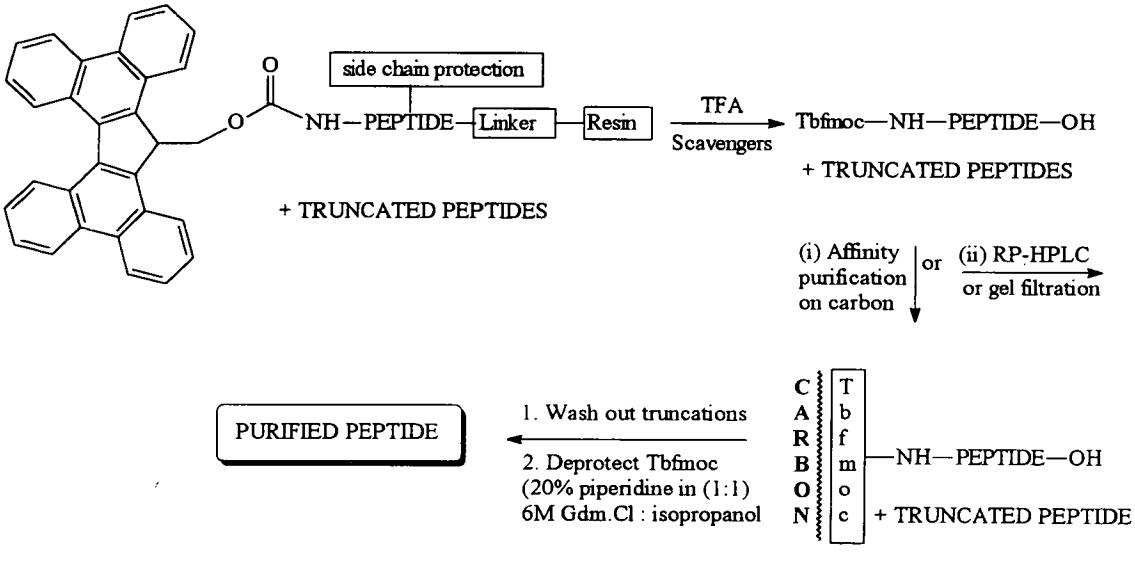


Figure 14 SPPS and Tbfmoc purification

Using this methodology stepwise SPPS has been reproducibly applied, in this laboratory, to the assembly of proteins including ubiquitin (76 residues)[44] and monocyte chemotactic protein-1 (MCP-1) (76 residues).[45] Larger proteins have been achieved with the syntheses of ribonuclease A (124 residues),[46] interleukin-3 (IL-3) (140 residues)[47] and, at Edinburgh, interferon-γ (143 residues) (chapter 3), the obese gene product (146 residues)[31] and erythropoietin (166 residues).[33] These synthetic achievments have demonstrated that the synthesis of large proteins is feasible with straightforward extensions of existing techniques. Thus the total chemical synthesis of other biologically significant proteins by stepwise SPPS is a practical and accessible goal.

1.2.3: Convergent synthesis of proteins

1.2.3.1: Fragment condensation of fully protected peptides

As outlined in section 1.2.2 improvements in stepwise SPPS have enabled the synthesis and purification of many biologically significant proteins. However, it would be misleading to believe that all large peptides or small proteins can now be obtained from automatic synthesisers with standardised reaction protocols. An alternative to stepwise construction of long peptide chains is to individually synthesise small protected peptide segments of the protein in question and then chemically link these together to produce the target sequence. This convergent procedure has the obvious advantage, in principle, that small peptide fragments are relatively straightforward to synthesise and purify. However, in order to achieve unambiguous reaction, protection of reactive side chain functionalities within each fragment is required. There are problems associated with protected peptide intermediates which detract from the advantages of this strategy. These are solubility difficulties,[48] propensity for racemisation during coupling and limited options for purification and characterisation of the fully protected fragments. Despite these problems, convergent condensation of fully-protected peptide segments has enjoyed notable successes with the synthesis of insulin,[49] ribonuclease A[50] and prothymosin α.[51]

A recent development in the attempt to overcome the insolubility of the protected peptides is the introduction, by Sakakibara, of a powerful solvent system which is a combination of TFE and chloroform or dichloromethane.[52] This solvent system not only dissolves almost all of the sparingly soluble protected peptides but also suppresses the racemisation during the coupling reaction.[52] The utility of this new solvent system was demonstrated by the synthesis of human midkine (121 residues) by the condensation of two fully protected intermediates using water soluble carbodiimide in the presence of HOOBt as coupling reagents.[53]

1.2.3.2: Fragment condensation of unprotected peptides

The use of highly selective chemistries to link together unprotected peptide fragments can avoid the requirement for side-chain protection. Such a concept was successfully demonstrated by Kent in the complete synthesis of an HIV-1 protease analogue (99 residues).[54] By replacing the peptide bond at the fusion site with a thioester linkage, unprotected synthetic fragments of the enzyme were linked together to directly yield a protein analogue which exhibited full enzymatic activity. The chemoselective ligation was achieved by a nucleophilic reaction at low pH (pH 3-5) between the thioacid-containing fragment (**39**) and the corresponding bromoacetylated fragment (**40**) (Figure 15).

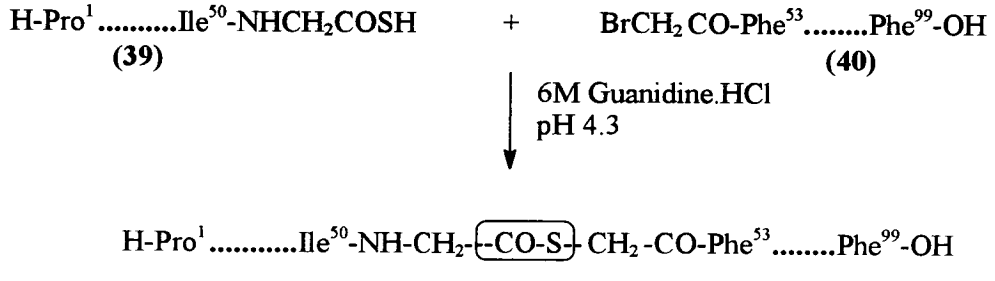


Figure 15 Chemoselective ligation of an HIV-1 protease analog

The main disadvantages associated with this technique are the instability of the thioester pseudo-peptide bond at physiological pH and the fact that this structural motif results in the introduction of an altered backbone into the protein. However,

this method of backbone engineering has been exploited for the rapid assembly of a four-helix bundle on a templating peptide.[55]

1.2.3.3: Enzyme catalysed ligation

Many of the difficulties associated with protein synthesis strategies can be circumvented by enzymatic coupling procedures. Proteolytic enzymes, particularly serine proteases, have been used extensively in peptide synthesis because of their selectivity, ability to function under mild reaction conditions and minimal requirements for protecting groups. Although it might seem counterintuitive to attempt to synthesise a protein using a protease, problems due to proteolytic degradation of the final peptide product by the enzyme can be minimised by manipulating the reaction conditions or through genetic engineering. An example of a suitable enzyme is lysine endoproteinase which preferentially cleaves at the carboxyl side of lysine residues.[56] Thus coupling of segments can only be carried out for carboxyl components with a lysine residue in the C-terminal position. To prevent the protease dismantling peptide bonds at another lysine sites in the peptide the other lysine residues in the unassociated fragments must be protected during the enzymatic synthesis.

Significant progress in the area of enzyme-driven ligation has been achieved by Nakatsuka through specific genetic engineering of the serine protease subtilisin.[57] A double mutation of subtilisin converts the protease into an acyl transferase called subtiligase which, importantly, has greatly diminished proteolytic activity compared with the wild-type enzyme. Replacement of the active-site serine with a cysteine residue results in efficient acylation of the enzyme by peptides esterified at their C-terminus with a glycolate-phenyl-alanine amide group (**41**).[58] A second active-site mutation relieves steric crowding, allowing subsequent deacylation of the enzyme-peptide thioester intermediate (**42**) by the N-terminus of the second peptide fragment.

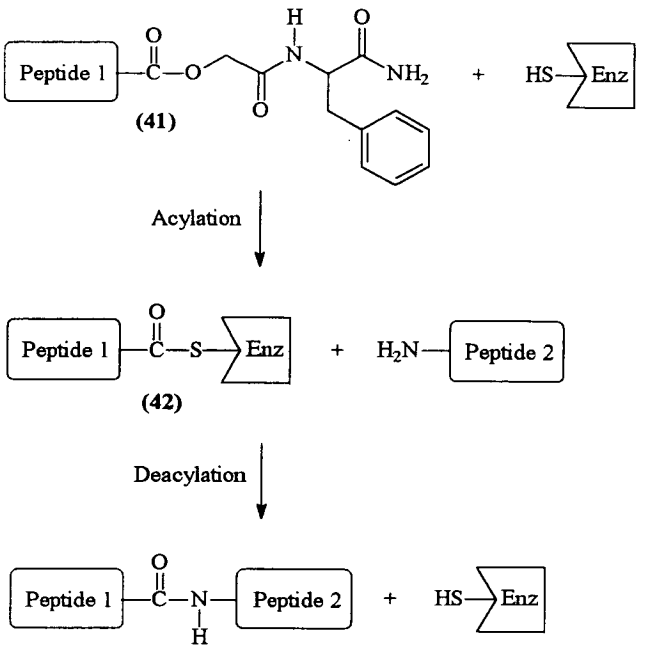


Figure 16 Subtiligase-catalysed ligation

Near quantitative yields have been obtained for reactions between model esters and dipeptide nucleophiles.[59] The total synthesis of ribonuclease A was achieved in good yield and high purity by subtiligase-catalysed ligation of six peptide segments ranging in length from 12 to 30 residues.[58] Ribonuclease A mutants containing the unnatural amino acid 4-fluorohistidine were also prepared allowing the mechanism of catalysis to be probed. Given that the incorporation of multiple unnatural amino acids in precise positions in a protein backbone is not currently possible using recombinant technologies, this application of subtiligase illustrates the potential utility of chemical approaches to total protein synthesis.

1.2.3.4: Native chemical ligation

An important extension of the chemical ligation approach has recently been introduced by Kent and coworkers that results in the straightforward generation of proteins with native backbone structures from fully unprotected peptide building blocks.[60] Referred to as native chemical ligation, the method utilises the reaction of a peptide bearing a thioester at its carboxyl terminus (**43**) with another unprotected

peptide segment containing an N-terminal cysteine residue (44) to give a thio ester-linked intermediate (45). Without change in the reaction conditions this intermediate undergoes spontaneous intramolecular reaction to form a native peptide bond at the reaction site. Consequently, the target full-length polypeptide is obtained in the desired final form without further manipulation.

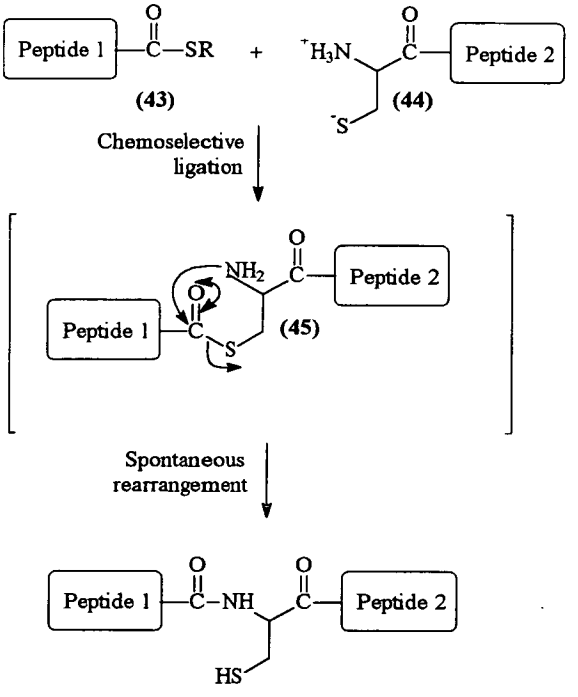


Figure 17 Native chemical ligation

As with their enzyme-driven counterparts, the native chemical ligation strategies can be performed in aqueous media, but have the added advantage of being tolerant to organic solvents[61] and to solubilising agents such as urea and guanidine.[54] The ligations can also be performed in the presence of all the functionalities commonly found in proteins, including free cysteine sulfhydryls. The only disadvantage of native chemical ligation is the necessity for a cysteine residue at the coupling site. Although this technique has been demonstrated to be successful a more general approach to fragment coupling would be desirable.

1.2.3.5: Azide method

One of the most successful methods of fragment condensation encompassing the minimal protection strategy and retaining a peptide amide bond is the azide procedure. Azide condensations are advantageous in that once the azide has been selectively introduced at the C-terminus it is only necessary to have the strongly nucleophilic N^{α} terminal amino function, the N^{ε} amino functions of lysine and the sulfhydryl functions of cysteine protected.[62] Azide activation is relatively racemisation free and the starting materials are easily prepared.

Acyl azides can be generated by the diazotisation of a hydrazide using nitrous acid or alkyl nitrites. This requires three consecutive steps for peptide bond formation: hydrazide formation (**46**), azide formation (**47**) and then aminolysis resulting in union of the two peptide fragments to give the target peptide.

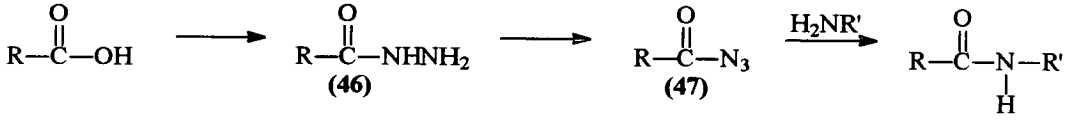


Figure 18 Peptide bond formation via azide

The formation of the hydrazide of a peptide has been achieved by the development of hydrazine resin linkers.[16,17] This has eliminated the need for hydrazinolysis and the associated problem of removal of Fmoc from the N-terminus by hydrazine. The hydrazine resin linker (**48**) should selectively incorporate the hydrazide to the C-terminus of the solid phase support, from which acidolysis with TFA will yield the deprotected hydrazide.

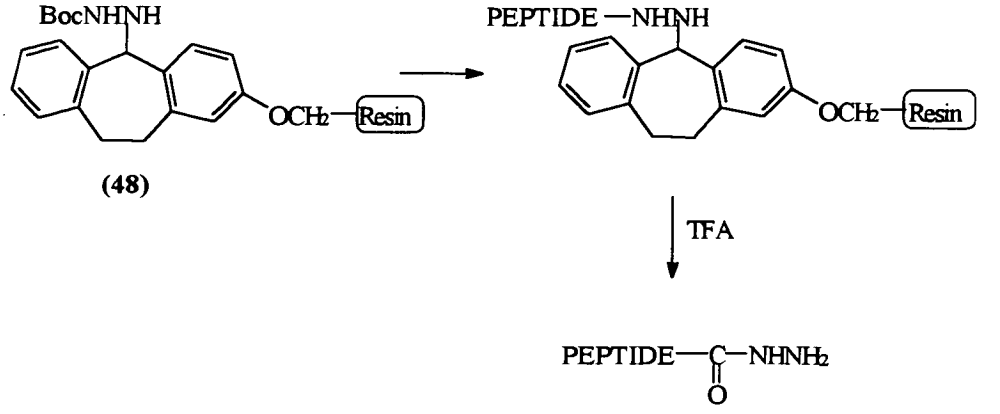


Figure 19 Formation of peptide hydrazide via a linker

The method of Honzl-Rudinger is applied to convert the hydrazide to the azide.[63] The peptide hydrazide (**49**) is reacted with an alkyl nitrite, such as t-butyl nitrite, generating an N-nitroso intermediate (**50**) which loses water to form the corresponding peptide azide (**51**).[64]

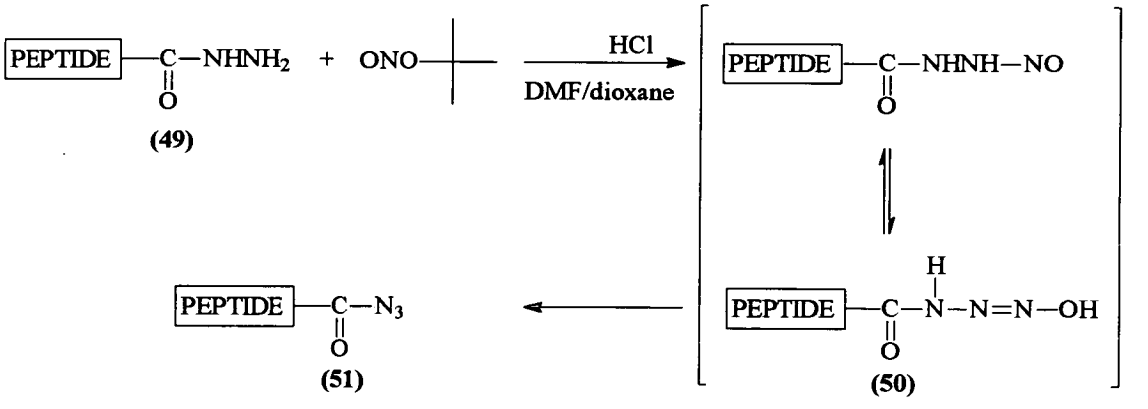


Figure 20 Azide formation via the Honzl-Rudinger method[63]

The azide (**51**) is normally not isolated and after neutralisation of the acid is immediately reacted with the amino component of the other peptide fragment. However, the azide itself is prone to side reactions such as amide formation which occurs when the N-nitroso intermediate (**52**) loses N_2O to form the unreactive amide (**53**).

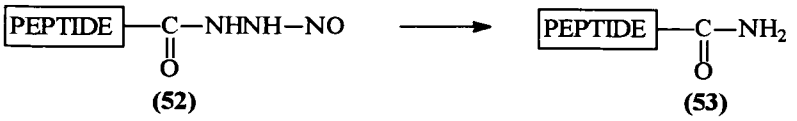


Figure 21 Rearrangement of the N-nitroso intermediate to form an amide

The other side reaction is the Curtius rearrangement in which the isocyanate (**54**) is formed which can then react with the amine component to form a urea derivative (**55**). These problematic side reactions can be suppressed by maintaining the temperature at -10^0C.[65]

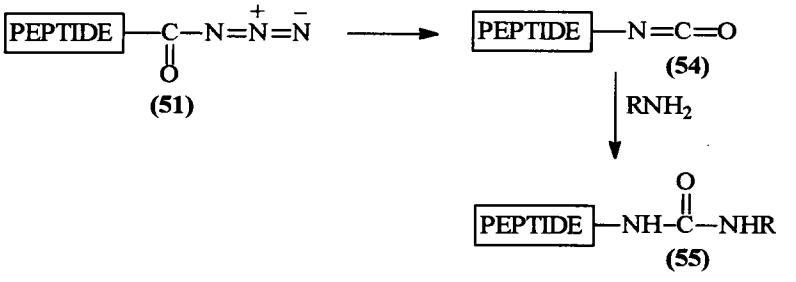


Figure 22 Urea formation via Curtius rearrangement of the azide

Notwithstanding these side reactions, the azide method is a very effective technique for the union of minimally protected peptide fragments. However, the successful marriage of Fmoc SPPS and azide condensations requires the development of a new protecting group methodology. It is necessary to protect the strongly nucleophilic N^{ε} function of lysine and the sulfhydryl function of cysteine. Protection is required both throughout the synthesis of the peptide fragment itself and during the condensation of the peptide fragments to form the desired protein. It would obviously be preferable to develop protecting groups that are introduced with the amino acids during peptide chain assembly. This is a difficult problem since the protecting groups must not only be orthogonal to the basic Fmoc removal conditions but must also be stable to the acidic resin peptide cleavage. In addition to these stringent requirements the

protecting groups should impart solubility to the fragments and ultimately be cleaved under mild conditions.

Chapter 2: N^{ε}-Lysine protecting groups

2.1: General introduction

Solid phase peptide synthesis, azide condensations and enzymatic fragment coupling all necessitate the protection of the N^{ε} amino function of lysine. Protection is required both throughout the synthesis of the peptide fragment itself and during coupling of the peptide fragments to form the desired protein. The requirements for the design of such a protecting group are stringent and may be summarised as follows. The group should:

- be orthogonal to acid labile side chain protecting groups and the C-terminal ester bond to the resin
- survive repeated Fmoc deprotections during peptide assembly on the resin
- attribute solubility to the synthetic peptide fragment and protein product
- be amenable to deprotection under mild conditions which are compatible with protein chemistry
- be straightforward to introduce onto the lysine side chain
- be relatively cheap to produce.

2.1.1: Objective

A specific goal requiring an exacting chemical strategy is the chemical synthesis of novel branched proteins such as ubiquitin conjugates. Ubiquitin is a highly conserved (76 amino acids, 8565 Da) protein found in all eukaryotes[66] both as a free protein and covalently attached to other proteins. In collaboration with Professor Mayer, at The University of Nottingham, interesting branched ubiquitin derivatives have been identified. The chemical synthesis of these derivatives is required to allow different aspects of ubiquitin metabolism to be investigated. In particular they could be used to probe the cellular roles played by the 26S proteasome complex. The 26S proteasome is a multisubunit protease which is a new therapeutic target for the pharmaceutical

industry with inhibitors having applications in the treatment of disorders such as metabolic acidosis-induced muscle wasting.[67]

Di-ubiquitin has been chosen as the first target molecule and is formed through an isopeptide bond between the ε amino group of lysine-48 and the C-terminal glycine of another ubiquitin molecule (Figure 23). To ensure selectivity of the amide bond to be formed between these two residues it is necessary to protect the other lysine residues in the unassociated fragments.

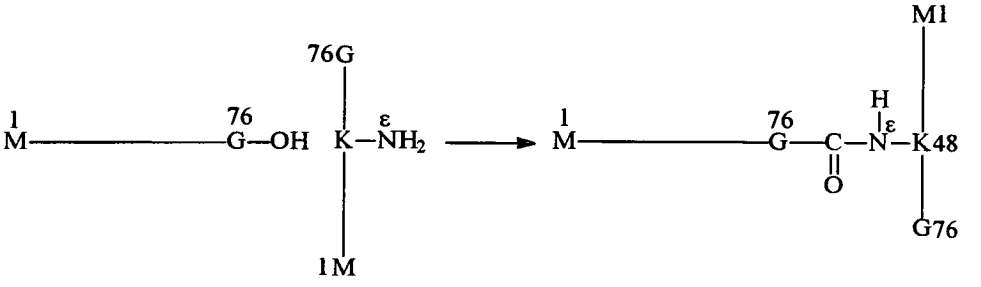


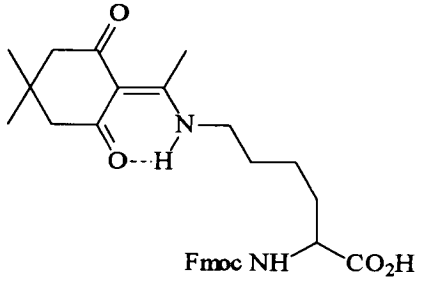
Figure 23 Synthesis of di-ubiquitin

The successful chemical synthesis of di-ubiquitin demands a protecting group for the N^{ε} amino function of lysine which allows 2nd order differentiation between different lysine residues. Therefore, the objective of this research was the development of a novel ε amino protecting group for lysine. This involved the discovery and synthesis of an appropriate lysine derivative and its incorporation into Fmoc SPPS. A test peptide was utilised for optimising the protection and deprotection conditions before moving to the large ubiquitin system.

2.1.2: Examples of existing N^{ε}-lysine protecting groups

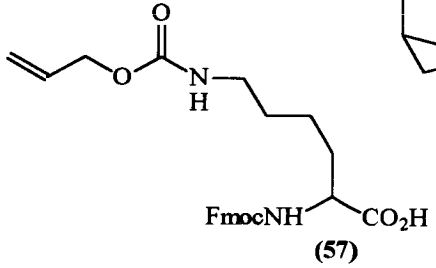
Notable accomplishments in protecting group strategy for lysine have been the development of groups such as the 1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidine)ethyl or Dde group (**56**).[68] The Dde group has been employed in the synthesis of branched peptides[69] and is efficiently removed by 2% hydrazine in DMF. However, when model peptides containing Lys(Dde) were subjected to 20% piperidine in DMF a 3.2% loss of Dde was observed after 2 hours and 6.6% after 4 hours. Although this

loss is considered sufficently low to validate this protection strategy for short chain peptide synthesis it may prove inappropriate in the synthesis of longer chain polypeptides (>100 residues).

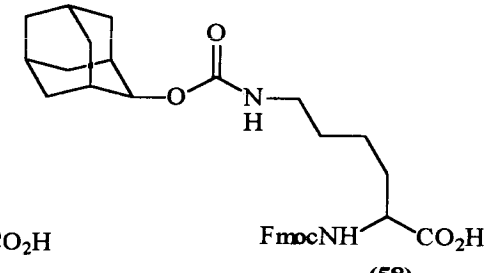


(56)

Another protecting group offering 2nd order differentiation of lysine is the Alloc group (**57**).[70] However, this group is deprotected by treatment with Pd(0) metal, a slow reaction requiring specialised argon equipment and often leaves the products contaminated with Pd.[69] An additional drawback is that the reaction cannot be monitored. The 2-adamantyloxycarbonyl (2-Adoc) group (**58**) has also been developed for ε-amino protection of lysine in convergent solid phase peptide synthesis.[71] This group is stable to both piperidine and TFA but is cleaved by the strong acids trifluoromethanesulfonic acid (TFMSA) or anhydrous HF in a few minutes at 0^0C.



(57)



(58)

The presence of aromatic and hydrophobic groups within post cleavage protecting groups is limited since they drastically reduce the solubility of the fragments they protect. This problem is particularly significant for large peptides which would be rendered insoluble by these groups making purification and further reactions exceedingly difficult. Although many protecting groups have been described in the

literature none adequately fulfil the aforementioned requirements for the successful synthesis and condensation of large peptide fragments.

2.2: Design of an enzyme cleavable lysine protecting group

Numerous chemical methods have been developed for the introduction of protecting groups which can be cleaved under the mildest of conditions.[72,73] Recently there has been renewed interest in using biocatalyts for functional group protection and deprotection in the field of peptide synthesis[74] and in other areas.[75] In addition to their stereodiscriminating properties, enzymes offer the opportunity to carry out highly chemoselective and regioselective transformations. They generally operate under physiological conditions and in many cases combine a high selectivity for the reactions they catalyse and the structures they recognise with a broad substrate tolerance. Therefore, the application of biocatalysts offers a viable alternative to classical methods of functional group protection and deprotection.

2.2.1: Penicillin G acylase

Penicillin G acylase (E.C. 3.5.1.11), from *Escherichia coli,* is an enzyme used industrially to hydrolyse benzylpenicillin (**59**) to yield phenylacetic acid (**60**) and 6-aminopenicillanic acid (**61**) (the β-lactam nucleus which is the basis for a wide range of semi-synthetic penicillins) (Figure 24).[76,77] The enzyme is highly specific at the aromatic carbonyl side of the amide bond, yet highly tolerant of substrates varying in the amine side of the amide bond.[78]

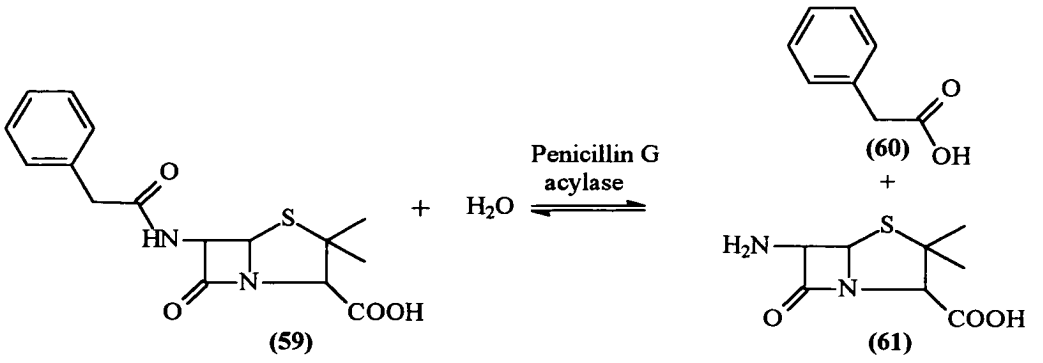


Figure 24 Penicillin acylase hydrolysis of penicillin G to form 6-aminopenicillanic acid and phenylacetic acid

Its potential in peptide synthesis was first demonstrated by Brtnik *et al,* who utilised the enzyme as a selective hydrolysis reagent for N^{ε}-phenylacetyl lysine in the solution synthesis of vasopressin (9 residues).[79] On completion of the synthesis the N^{ε}-phenylacetyl protected lysyl peptide (**62**) was deprotected by incubation with penicillin acylase to yield the lysyl derivative (**63**) (Figure 25). This synthesis demonstrated the possible utility of the enzyme in protecting group strategies and also the advantage of a protecting group connected by an amide bond (which should be as stable to acid and base as the other peptide bonds in the molecule).

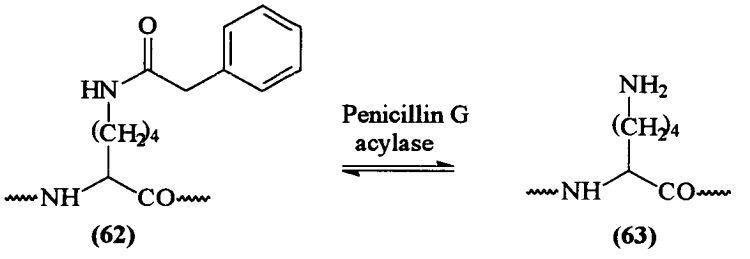


Figure 25 Penicillin acylase hydrolysis of N^{ε} phenylacetyl lysyl residue

A major problem associated with N^{ε} lysyl protection is the loss of protonation sites, and hence solubility. Previous work at Edinburgh established that the insolubility of peptides containing phenylacetyl protected lysine caused problems both in the purification and in the final enzymatic cleavage step.[80] The deprotection of the

phenylacetyl group was found to require several days incubation with the enzyme utilising conditions (40-50^{0}C, pH 7.1-8.5) which were not compatible with many peptides, especially those which readily hydrolyse or oxidise.

2.2.2: Structure and modelling studies of penicillin G acylase

One important reason for studying the enzyme is its extensive use by the pharmaceutical industry to produce large quantities of 6-aminopenicillanic acid (**61**), the starting point for the chemical manufacture of semi-synthetic penicillins. It has already been discovered that the substitution of a hydroxyl group in the 4 position of the benzene ring, to give 4-hydroxybenzyl penicillin, increases the rate of hydrolysis.[81] Knowledge of the enzyme's mechanism and the structural basis of its specificity may make it possible to further extend its range of substrates and to convert it to a synthase useful for enzymatic synthesis. With this impetus, the crystal structure of penicillin acylase has recently been resolved by Professor Dodson at York University.[82] The mature enzyme is a heterodimer of 20,500 and 69,000Da consisting of closely intertwined A and B chains (209 and 566 amino acids respectively).

The structures of the enzyme crystallised with phenylacetic acid (a competitive inhibitor) and phenylmethylsulphonyl fluoride (PMSF) have also been determined allowing the binding site for the side chain of the substrate to be located. The phenyl moiety of each of these inhibitors points towards the interior of the protein; into a mainly hydrophobic cleft lined with many aromatic residues and hydrophobic side chains. The increased rate of hydrolysis of substrates with a hydroxyl group at the 4 position of the benzene ring can be rationally explained by close examination of the structure of the active site. At the mouth of the binding pocket is a β1 serine residue which is responsible for the hydrolysis of the amide bond and at the back of the cleft there is a second serine (Figure 26). The side chain of this serine residue should interact favourably with the hydroxyl group, drawing the substrate into the active site and holding it there during hydrolysis.

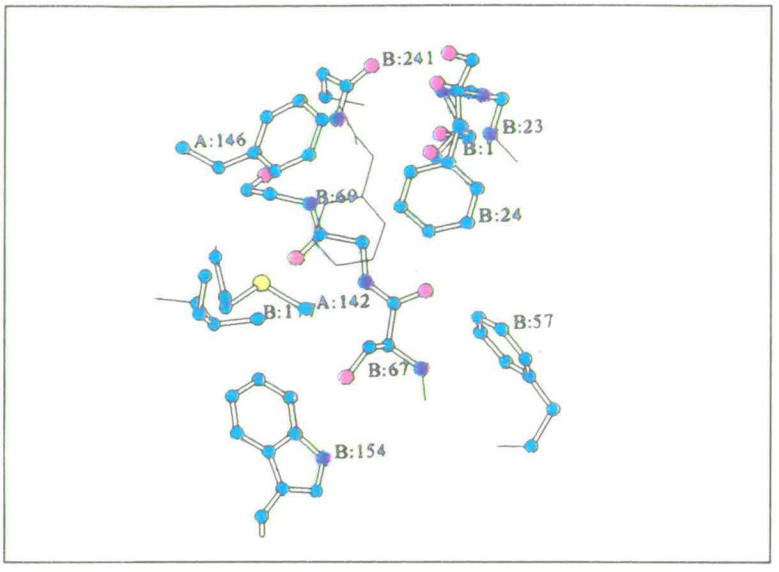
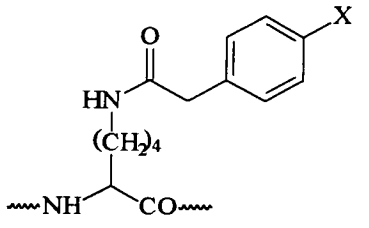


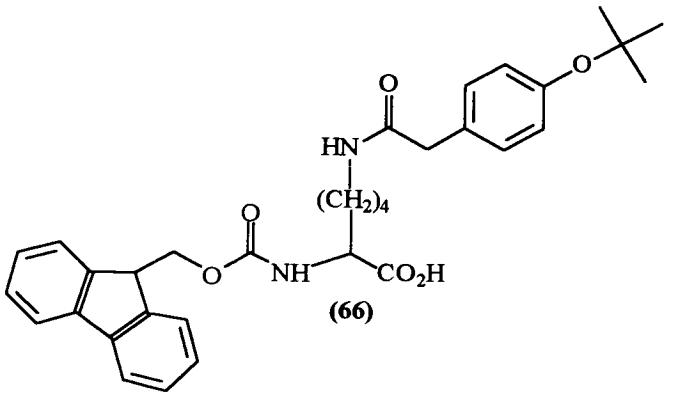
Figure 26 Molecular model of phenylacetic acid in the active site of penicillin G acylase. The phenylacetic acid molecule is shown as a line drawing and the nearby residues as 'ball-and-stick' representations.

Figure 26 shows a molecular model of phenylacetic acid in the active site illustrating that insertion of groups at positions other than the 4 position would interfere with the hydrophobic phenylalanine residues. This would cause a worse fit of the substrate into the active site and hence a dramatic reduction in the rate of hydrolysis. Consequently, the best substrate for the enzyme is one in which only the functional group at the 4 position is altered. The 4-hydroxybenzyl penicillin substrate could be modified for protecting group strategy to give the corresponding N^{ε}-4-hydroxyphenylacetyl lysine protecting group (Lys-Hopa) (**64**). The addition of a hydroxy group in the 4 position should improve the protecting group in two ways. Firstly, by making it more hydrophilic it should be more solubilising than the phenylacetyl group (Phenac) (**65**) and, secondly, by creating a better substrate for the enzyme it should be more susceptible to cleavage.



(64) X=OH
(65) X=H

Incorporation of the lysine derivative **(64)** into Fmoc SPPS methodology requires the phenolic group to be protected during peptide synthesis preventing it coupling *via* its hydroxy function to activated amino acids. The t-butyl ether derivative, used to protect tyrosine, was thought to be suitable due to its structural similarities with tyrosine. This would result in the amino acid derivative N^{α}-Fmoc-N^{ε}-4-t-butoxylphenylacetyl-lysine (Fmoc-Lys-Bupa) **(66)** being incorporated during peptide chain assembly. Subsequent cleavage of the peptide from the resin would regenerate the Hopa protected lysyl peptide which, after fragment condensation, could be fully deprotected by penicillin G acylase.

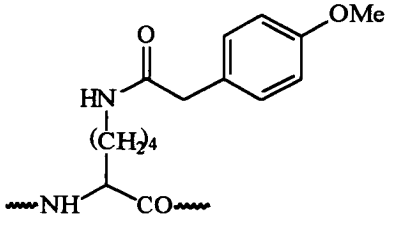


(66)

Pallin attempted to synthesise **(66)** and despite encountering synthetic problems a pentapeptide, H-Phe-Gly-Lys-Ala-Gly-OH, incorporating the Bupa lysyl derivative was synthesised.[80] This was tested against the same peptide sequence which contained the Phenac protected lysine **(65)**. The results obtained were very exciting since not only was the Hopa peptide more soluble but the Hopa group was

deprotected faster than the Phenac group. These results revealed that the Hopa protecting group possessed the desired features of increased solubility and fast removal required for fragment coupling reactions. Unfortunately, the synthesis of the Bupa lysine derivative proved irreproducible[33] and a new protecting group was required.

The subsequent proposal was to replace the t-butyl group with a methoxy group forming the 4-methoxyphenylacetyl protecting group (Mepa) (**67**). This group was chosen since the synthesis should be more accessible and molecular modelling studies had shown that it was structurally compatible with the active site. It was hoped that the oxygen at the 4 position would still be available for hydrogen bonding at the active site allowing the Mepa group to be hydrolysed as fast as the Hopa group.



(67)

In order to investigate the scope and limitations of the Mepa protecting group the same peptide sequence employed by Pallin was synthesised. As before the corresponding Phenac peptide was also synthesised as a control. The Phenac protected peptide (**68**) was deprotected much faster than the Mepa peptide (**69**) (approximate half-life of 6.5 hours versus 22 hours for the Mepa group). This disappointing result highlighted the substrate specificity of the enzyme and implied that the largest group tolerable in the 4 position of the phenylacetyl was a hydroxyl.

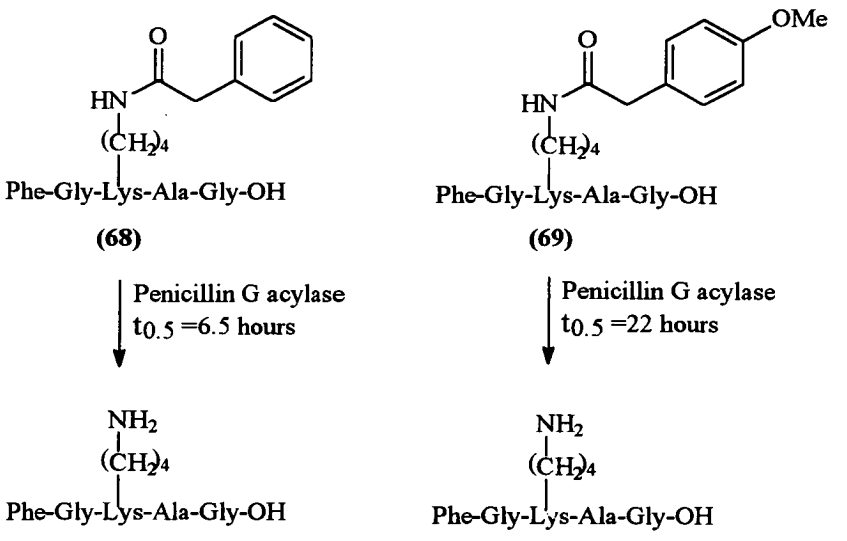


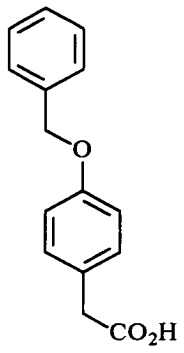
Figure 27 Penicillin G acylase deprotection of Phenac and Mepa protected lysyl pentapeptides

2.2.3: Strategy for development of a new enzyme cleavable N^{ε}-Lysine protecting group

The experiments utilising the Mepa group (67) and the findings by Robertson[33] and Pallin[80] indicated that the presence of the 4-hydroxyphenylacetyl (Hopa) moiety (64) was a crucial requirement. The Hopa group was desirable since it is the best substrate for the enzyme and also imparts the greatest solubility to the protected peptide. Accordingly the strategy adopted was to develop a protecting group which would yield the Hopa moiety after the peptide resin cleavage. The following reactions describe the efforts made to meet this strategy.

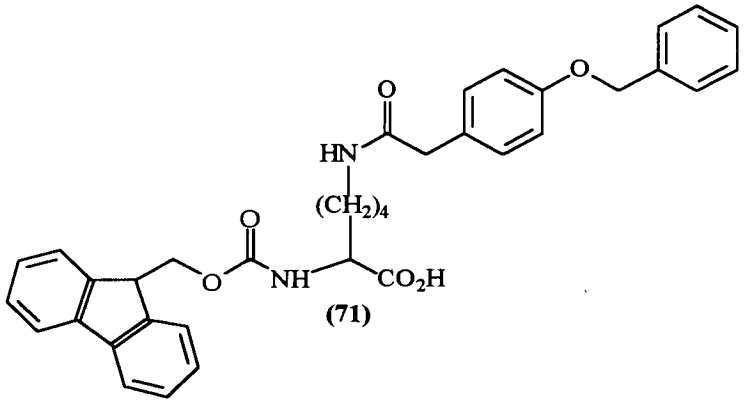
2.2.3.1: N^{ε}-4-benzyloxyphenylacetyl protected lysine

The use of benzyl ethers as protecting groups for phenols is well documented.[83] It was decided to employ a benzyl ether derivative to protect the hydroxy function thus allowing the Hopa derivative to be incorporated onto the side chain of lysine. Consequently the next target protecting group was a 4-benzyloxyphenylacetyl group (Bepa) (70).



(70)

The protected lysine residue should therefore be introduced as the amino acid derivative: N^{α}-9-Fmoc-N^{ε}-4-benzyloxyphenylacetyl-lysine (Fmoc-Lys-Bepa) (**71**).



(71)

This would result in a Bepa lysine (**72**) being incorporated into the peptide on the resin. Cleavage from the resin using TFA would yield the Hopa lysyl derivative (**64**), which could itself be cleaved using penicillin G acylase affording the N^{ε} amino deprotected peptide as in Figure 28.

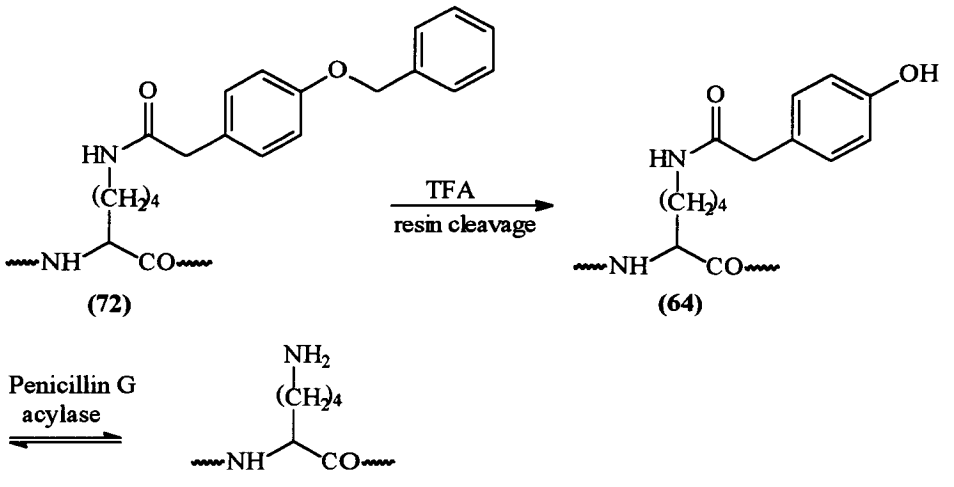


Figure 28 Two stage deprotection of Bepa protecting group

In general benzyl ethers are prepared from a phenol by treating an alkaline solution of the phenol with a benzyl halide.[84] Synthesis of the derivative was achieved by the following synthetic route. Firstly the 4-hydroxyphenylacetic acid methyl ester (**73**) was reacted with benzyl bromide to give the benzyl ether (**74**), which was then hydrolysed to remove the ester group yielding benzyloxyphenylacetic acid (**70**).

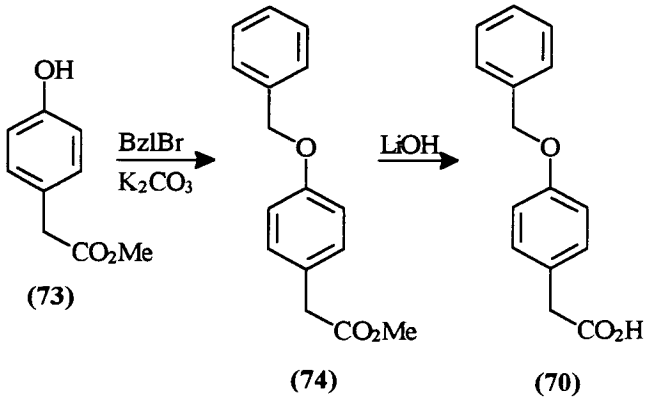


Figure 29 Synthesis of 4-hydroxyphenylacetic acid

Pallin[80] found that the most favourable synthetic route to Fmoc-Lys-Bupa (**66**) was one which involved the Bupa moiety being coupled to a copper lysine complex. This strategy was adopted for the preparation of the Bepa protecting group. Complexation

with copper means the two α-functionalities of lysine engage in chelate formation allowing specific reaction at the ε-amino group. Thus the Bepa moiety is coupled, *via* its acid chloride (**75**), to a copper lysine complex (**76**) resulting in copper[N^{ε}-4-benzyloxyphenylacetyl lysine]$_2$ (**77**) (Figure 30).

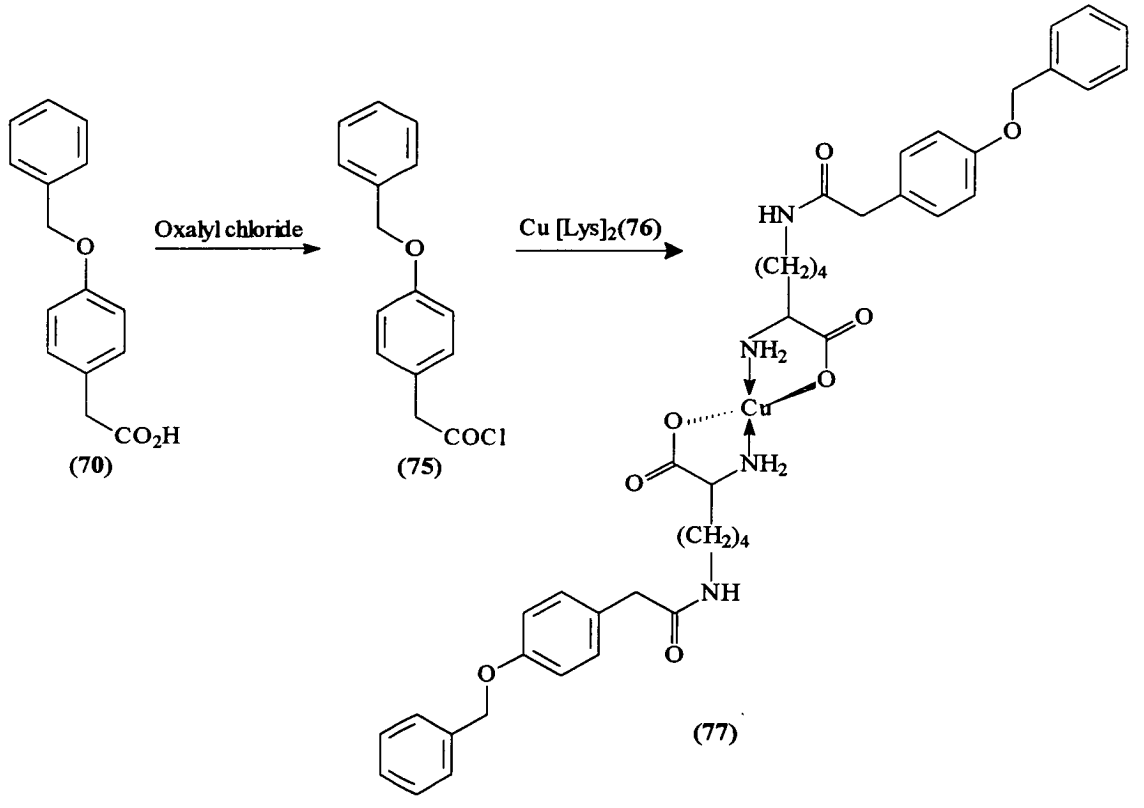
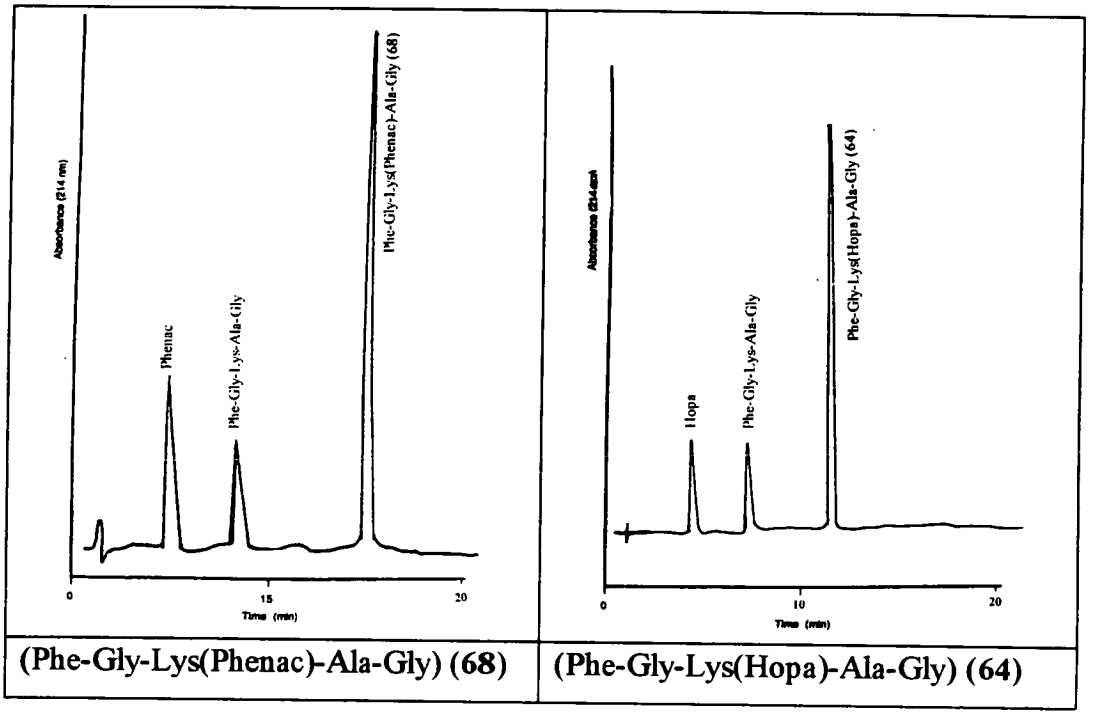


Figure 30 Synthesis of Copper[N^{ε}-4-benzyloxyphenylacetyl lysine]$_2$

Extraction of the copper from the insoluble complex was achieved by sonicating a mixture of (**77**) in a saturated solution of the disodium salt of ethylenediaminetetraacetic acid (EDTA).[85] The introduction of Fmoc using 9-fluorenylmethylsuccinimidyl carbonate (Fmoc-ONSu),[86] to give the Fmoc-Lys-Bepa (**71**) was reasonably straightforward.

To investigate the efficacy of the Bepa protecting group the same trial pentapeptide was synthesised so that a comparison of solubility and deprotection could be made with the Phenac and Mepa protecting groups.

When the pentapeptide with the Bepa protected lysine (**72**) was cleaved from the resin the desired Hopa protected peptide (**64**) was obtained as in Figure 28. After purification the (Phe-Gly-Lys(Hopa)-Ala-Gly) peptide (**64**) was dissolved in Tris buffer at pH 8.1, before the enzyme penicillin G acylase was added and the reaction incubated at 37°C with gentle agitation. The acylase (kindly supplied by SmithKline Beecham Pharmaceuticals Ltd.) was bound to a resin which enabled easy purification: the enzyme could simply be filtered from the reaction solution. The deprotection of the peptide to give the free N^{ε} amino lysyl derivative was monitored by analytical HPLC and was considerably faster than for the Phenac peptide (**68**). Figure 31 shows HPLC traces of the deprotection of the Phenac and Hopa protected peptides after one hour.



(Phe-Gly-Lys(Phenac)-Ala-Gly) (**68**) | (Phe-Gly-Lys(Hopa)-Ala-Gly) (**64**)

Figure 31 HPLC traces of penicillin acylase deprotection of Phenac and Hopa protected pentapeptides after 1 hour

An approximate half life of 4 hours was calculated compared to the 6.5 hours for the Phenac deprotection. In addition, the Hopa protected peptide eluted earlier than the

Phenac protected peptide. Thus not only is it deprotected faster but it is also more hydrophilic (Table 2).

Protecting group	Half life (hours)	Rt (mins), %B (See Figure 31)
Mepa	22	17, 49%
Phenac	6.5	17, 49%
Hopa	4	12, 38%

Table 2 Half lives and HPLC retention times for lysine protecting groups

This was a very promising result. However, the HPLC trace of the product after the TFA resin cleavage revealed three peaks (a, b & c (R_t =12, 18 & 20mins)) (Figure 32). These were trapped and submitted for MALDI MS. Peak (a) was the correct mass for the Hopa peptide and peaks (b) and (c) corresponded to the mass of the pentapeptide with the lysine Bepa protected.

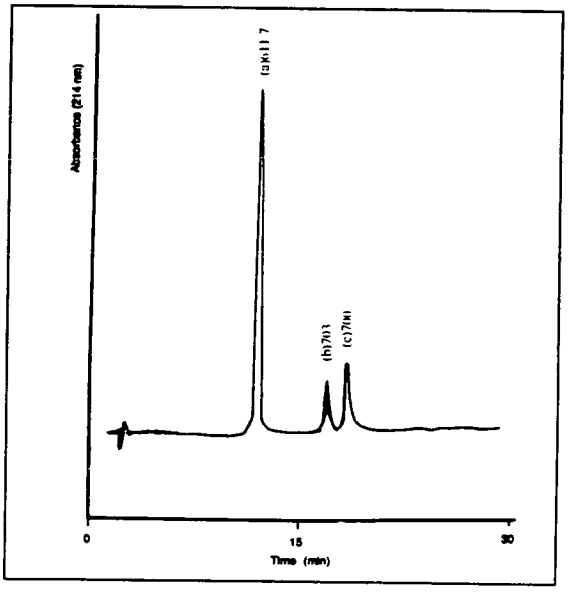


Figure 32 HPLC with MALDI MS results for peptides (a), (b) and (c) obtained after TFA resin cleavage

The peptides represented by peaks (b) and (c) must be isomers whose presence can be explained by examination of the mechanism of TFA cleavage. TFA-acidolysis of the benzyl ether (**78**) occurs *via* a mechanism involving alkyl-oxygen fission due to the stability of the intermediate carbocation thus formed (Figure 33). This process is hampered if the liberated benzyl carbonium ion (**79**) is involved in deleterious side reactions, the most obvious being recombination with the hydroxy function of Hopa (**80**) to reform the Bepa peptide (**78**). It is conceivable that the benzyl carbonium ion could also be attached at the 3 position on the benzene ring resulting in a compound with a different structure (**81**) but the same mass as (**78**).

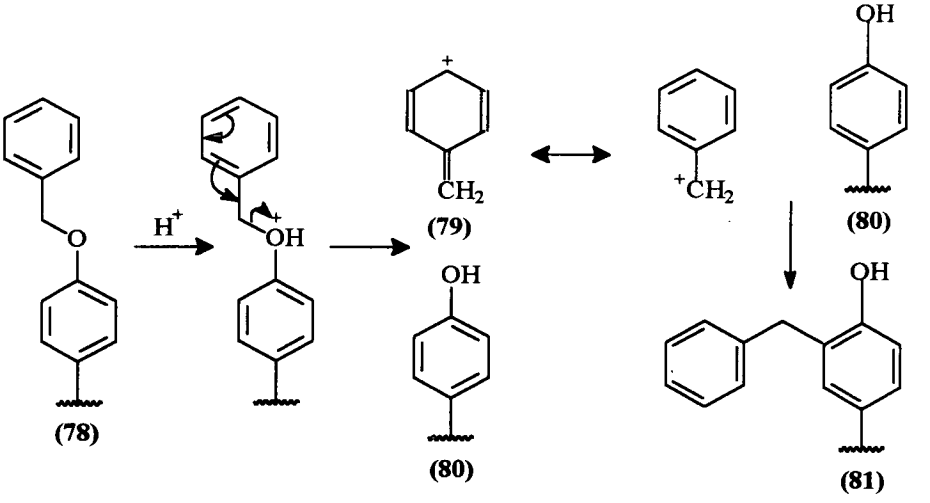
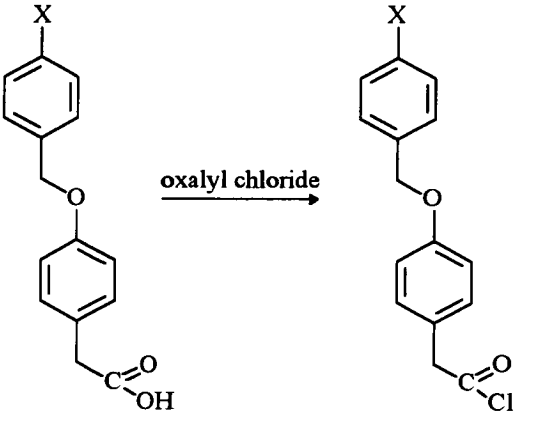


Figure 33 TFA acidolysis of benzyloxyphenyl acetyl group

To minimise this side reaction scavengers, and in particular phenol, must be present to quench the reactive benzylcarbonium ion. Various scavenger mixtures were investigated and the optimum cleavage and deprotection protocol was found to be: crystalline phenol, 1,2-ethanedithiol, thioanisole, water and TFA. Cleavage of the benzyl ether was followed by HPLC and was virtually complete after 3 hours.

It was proposed that modifying the protecting group itself would further improve the rate of cleavage of the benzyl ether. The introduction of a 4-methoxy or even a 4-methyl group should increase the stability of the benzyl cation. Since the cleavage of

the C-O bond is encouraged by the stability of the carbocation inclusion of such groups should further increase the cleavage rate.

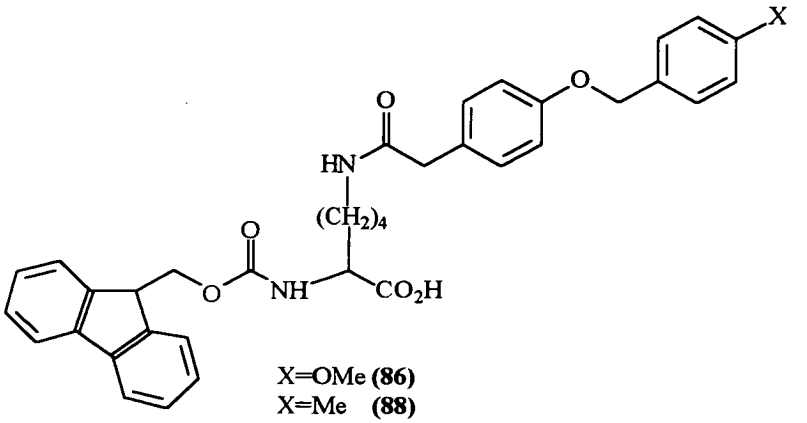


X=OMe **(82)**
X=Me **(83)**

X=OMe **(84)**
X=Me **(87)**

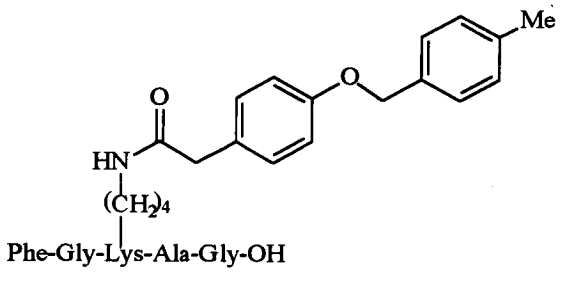
The next protecting group was therefore a 4-methoxy substituted benzyl ether **(82)**. The synthesis was analogous to the unsubstituted Bepa group. However, as with the Bupa group **(66)** problems arose during the synthesis of the acid chloride. The acid chloride **(84)** was prepared by reacting the 4-(4'-methoxybenzyloxy)phenylacetyl moiety **(82)** with oxalyl chloride in benzene. The reaction was followed by IR and when it had gone to completion the benzene was evaporated *in vacuo* to remove any unreacted oxalyl chloride and HCl formed. The 4-methoxy derivative was obviously unstable because a colour change was noted and by IR the acetyl chloride peak at 1800 cm^{-1} had disappeared. It was postulated that the HCl present in the reaction mixture breaks up the acid chloride **(84)** when the solution is concentrated to dryness. To overcome this problem the reaction mixture was reduced to half its original volume and reacted undiluted with the copper lysine **(76)** (Figure 30). Despite carrying out this reaction at pH 10, to neutralise any remaining HCl, the acid chloride did not successfully couple to the free N^{ε} amino of the copper lysine **(76)**. This was evident since, after extracting the copper, a white insoluble product was obtained which was identified as lysine.HCl (m.p. 245°C; no aromatics by NMR & no chromophore by U.V).

A small amount of copper[N^{ε}-4-(4'-methoxybenzyloxy)phenylacetyl lysine]$_2$ was obtained which was carried through to give Fmoc-Lys-(4-methoxy)Bepa (**86**) and characterised. However, it was concluded that the presence of the 4-methoxy group resulted in the destabilisation of the acid chloride (**84**) and the 4-methoxy substituted benzyl ether (**82**) was deemed not to be a viable protecting group.



X=OMe (**86**)
X=Me (**88**)

The 4-methyl substituted benzyl ether (**83**) formed a more stable acid chloride (**87**) and the synthesis and testing of Fmoc-Lys-(4-methyl)Bepa (**88**) was achieved.



(**89**)

The 4-methyl-Bepa pentapeptide (**89**) was synthesised and concurrent TFA cleavages, with identical scavenger mixtures, were carried out on 4-methyl-Bepa and Bepa protected pentapeptides. Samples of the cleavage mixture were removed after 1, 2 and 3 hours and analysed by HPLC. After 1 hour the formation of the desired Hopa group from both protected pentapeptides was observed by analytical HPLC (R_t 12mins., %B 38). The rate of cleavage was more than twice as fast for the 4-methyl-Bepa peptide (**89**). This was a promising result which agreed with our supposition. However, after 3 hours the Bepa group had almost fully cleaved, whereas for the

4-methyl-Bepa there was no further cleavage after 1 hour. It was speculated that although the initial cleavage was faster an equilibrium between the 4-methylbenzyl cation (90) and the 4-hydroxyphenylacetyl group (80) was being set up. Therefore, the 4-methylbenzyl cation (90) was in fact too stable and thus not effectively removed by the scavenger mixture. Figure 34 shows a comparison of the stability of the benzyl cations.

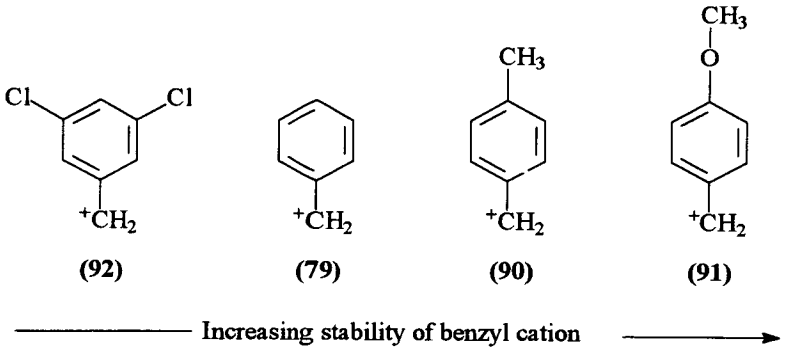


Figure 34 Comparison of stability of benzyl cations

The next proposal was to destabilise the benzyl cation by the addition of halogen atoms onto the benzene ring (92). This cation should be more easily removed by scavengers thus pushing the deprotection reaction to completion. Research using 2,6-dichlorobenzyl ethers to protect tyrosine had revealed only a low incidence of alkylation at the 3-position of tyrosine during acidolysis.[87] This hypothesis was tested by synthesising 4-(2,6-dichlorobenzyloxy)phenylacetyl protected lysine. Unfortunately, this molecule was very insoluble and therefore not compatible with automated solid phase peptide synthesis and thus this protecting group was abandoned.

From these investigations it was concluded that the Bepa group (70) was the most promising protecting group. This was due to its straightforward synthesis, compatibility with the automated peptide synthesiser and it gave the cleanest route to the Hopa protected lysine derivative. It was therefore decided to probe the scope and limitations of the Bepa protecting group by utilising it in the synthesis of the protein ubiquitin.

2.2.4: Incorporation of Bepa protected lysine into ubiquitin

As stated in Section 2.1.1 an application for the new N^{ε} lysine protecting group is to provide 2nd order differentiation between different lysine residues to permit the synthesis of di-ubiquitin. Consequently, to test the new protection strategy a ubiquitin derivative incorporating the Bepa protecting group was synthesised.

```
1                                   10                                  20
Met Gln ILe Phe Val Lys Thr Leu Thr Gly Lys Thr Ile Thr Leu Glu Val Glu Pro Ser
                                    30                                  40
Asp Thr Ile Glu Asn Val Lys Ala Lys Ile Gln Asp Lys Glu Gly Ile Pro Pro Asp Gln
                                    50                                  60
Gln Arg Leu Ile Phe Ala Gly Lys Gln Leu Glu Asp Gly Arg Thr Leu Ser Asp Tyr Asn
                                    70                      76
Ile Gln Lys Glu Ser Thr Leu His Leu Val Leu Arg Leu Arg Gly Gly
```

Figure 35 Amino acid sequence of bovine ubiquitin

Ubiquitin is a small protein (76 residues) containing seven lysine residues (Figure 35). In order to manipulate the orthogonality of the system the strategy was to protect six lysines with the Bepa group (70) and to protect the remaining lysine, at position 48, with Boc (1). After solid phase synthesis the ubiquitin derivative (93) could then be cleaved from the resin using TFA and scavengers. The TFA cleavage should remove the acid labile Boc group at lysine-48 and cleave the benzyl ether protecting the other six lysine residues leaving them Hopa protected to give [6,11,27,29,33&63-Lys(Hopa)]ubiquitin ($Ub(Hopa)_6$) (94).

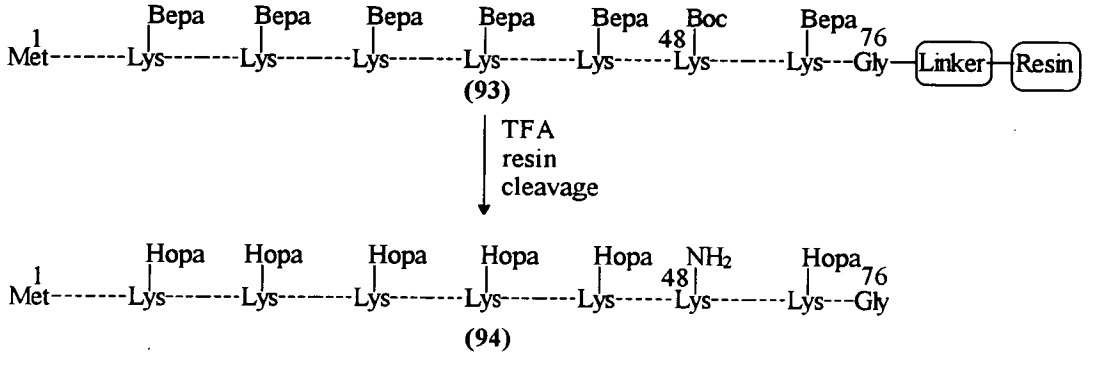


Figure 36 Schematic of the synthesis of [6,11,27,29,33&63-Lys(Hopa)]ubiquitin $Ub(Hopa)_6$

2.2.5: Chemical synthesis of $Ub(Hopa)_6$

The ubiquitin derivative was constructed as outlined in the experimental section without any drop in the coupling efficiency on introduction of Bepa protected lysine. On completion of the synthesis the resin-bound Fmoc protected $Ub(Bepa)_6$ (**95**) 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 the purification group Tbfmoc[43] (**37**) introduced at the N-terminus to give Tbfmoc-$Ub(Bepa)_6$-resin (**96**).

The next step was to cleave the protein from the resin and simultaneously remove all side chain protecting groups with the exception of the six Lys(Bepa) residues which were only partially cleaved to Lys(Hopa). The was achieved by utilising the TFA and scavenger cocktail selected from the studies on the Bepa pentapeptide (**72**) (Section 2.2.3). After 4.5 hours the resin was removed by filtration and the filtrate concentrated *in vacuo* to a yellow oil. The crude Tbfmoc-$Ub(Hopa)_6$ (**97**) was isolated from the oil by precipitaion with 2% v/v β-mercaptoethanol in diethyl ether at 0^0C.

2.2.6: Affinity purification of $Ub(Hopa)_6$ using Tbfmoc on PGC

During solid phase peptide synthesis acetylated truncates accumulate on the resin due to incomplete coupling steps and these need to be separated from the desired peptide sequence. This can be achieved by derivatising the N-terminus with the base labile Tbfmoc group prior to cleaving the peptide from the resin. Then exploiting the hydrophobicity of the Tbfmoc group to implement separation of the Tbfmoc-peptide from the truncates. Another advantage of using Tbfmoc for purification is that this group can also be utilised as a N^{α} protecting group for subsequent fragment coupling reactions.

As outlined in Section 1.2.2 two methods of purification have been developed for Tbfmoc-peptides. These are either affinity purification on porous graphitised carbon (PGC) or alternatively reverse phase HPLC which enables separation through the

increased retention time of the Tbfmoc-peptide compared to the truncates. For the purification of Tbfmoc-Ub(Hopa)$_6$ (**97**) the Tbfmoc/PGC purification was chosen in preference to reverse phase HPLC since studies by Wilken[88] had shown limited recovery of ubiquitin analogues when using the latter method.

Tbfmoc-Ub(Hopa)$_6$ (**97**) was solubilised in 6MGdm.Cl, diluted with isopropanol and centrifuged to remove a small amount of insoluble material. PGC was added to the supernatant and the mixture was vortexed for 10 minutes to allow (**97**) to be adsorbed onto the PGC. The PGC was washed to remove truncated peptides and finally the protein was cleaved from Tbfmoc by vortexing the PGC with a 10% piperidine solution and then centrifuging the PGC to a pellet. The supernatant was neutralised with acetic acid, the isopropanol removed *in vacuo* and the protein/guanidine solution was desalted by gel filtration to afford Ub(Hopa)$_6$ (**98**). Figure 37 shows the HPLC profiles of crude Tbfmoc-Ub(Hopa)$_6$ (**97**) and Ub(Hopa)$_6$ (**98**) obtained after the Tbfmoc/PGC purification.

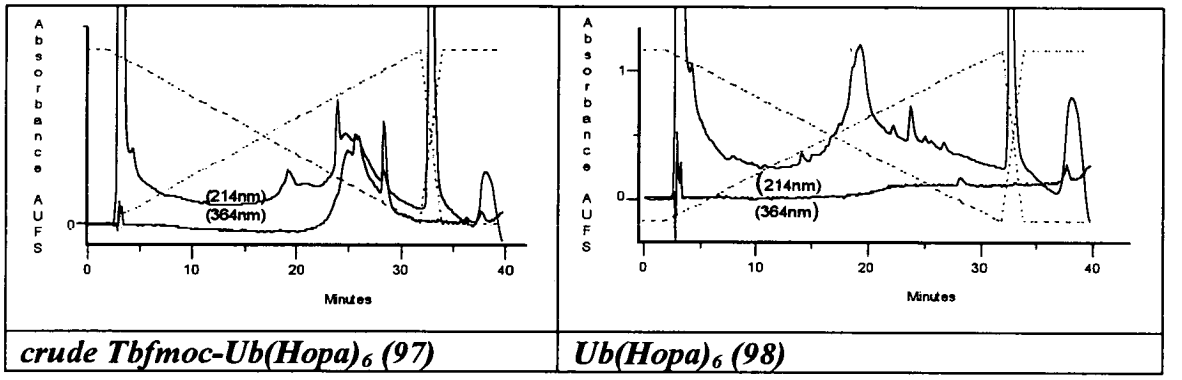


Figure 37 HPLC traces of crude Tbfmoc-Ub(Hopa)$_6$ (97) and Ub(Hopa)$_6$ (98)

The AAA, MALDI MS and N-terminal sequencing of (**98**) were as required, however the analytical RP HPLC was broad. This was probably due to the presence of the Hopa groups which would alter the conformation of the protein preventing it from folding in its native structure. In fact synthetic ubiquitin has to undergo a folding protocol before it is folded in its correct conformation.[88] Therefore the next goal was

to remove the Hopa groups, fold the protein and show that the ubiquitin obtained was analogous to native ubiquitin.

2.2.7: Attempted enzymatic deprotection of Ub(Hopa)$_6$ (98)

To confirm that the new protection/deprotection strategy can be successfully translated from small peptide models to protein fragments all the Hopa groups must be removed to generate ubiquitin. Hence Ub(Hopa)$_6$ **(98)** was incubated with immobilised penicillin G acylase utilising the conditions that had been successful for the trial pentapeptide (Tris buffer, pH 8.1 at 37^{0}C). Again the reaction was monitored by analytical HPLC but no deprotection was observed after 14 days.

Similar experiments by Brown, on an enzyme cleavable cysteine protecting group, had revealed that the protein could be adsorbed onto the solid support used to immobilise the enzyme.[89] This adsorption was due to the fact that the solid support was based on an ion exchange material which very quickly bound any free basic protein present in solution.[90] Although Ub(Hopa)$_6$ **(98)** is not basic, due to the lysine residues being capped, the free ubiquitin is basic with a pI of 8.5 and it was therefore surmised that any ubiquitin that had been deprotected or even partially deprotected would have been bound to the resin support. Therefore a second immobilised enzyme was obtained which utilised a polymeric solid support and the removal of the lysine protecting groups was attempted once more. HPLC monitoring indicated a slight shift in the retention time of the peak which remained broad. This peak was isolated and analysed by MALDI MS and gave essentially an identical mass spectra to Ub(Hopa)$_6$ **(98)**. Further attempts to deprotect the lysine residues by employing soluble enzymes also failed.

The optimum conditions for catalysis by immobilised penicillin acylase have been reported as pH 7-8.5 at 50^{0}C. Above 55^{0}C there is a rapid loss of activity and the soluble enzyme loses activity faster than the immobilised enzyme.[91] Thus the deprotection reactions were incubated at the elevated temperature of 50^{0}C, but with no success.

The fact that $Ub(Hopa)_6$ **(98)** was only sparingly soluble in the aqueous media employed would also limit the deprotection rates. In an attempt to overcome this disadvantage solubility-enhancing organic cosolvents, in particular DMF and DMSO were added. In their presence, however, the enzyme was denatured to a major extent and again no deprotection was achieved.

The failure of the enzyme mediated deprotection can be rationalised if the rate of the forward reaction, to produce the free amino function of lysine and 4-hydroxyphenylacetic acid **(99)**, is slower than the reverse reaction (Figure 38). The equilibrium of the hydrolysis reaction might not go to completion due to product inhibition. A possible explanation is that the macromolecules of $Ub(Hopa)_6$ **(98)** can not effectively interact with the active sites of penicillin acylase as effectively as the small molecules of 4-hydroxyphenylacetic acid **(99)**. The effective collision frequency of $Ub(Hopa)_6$ with penicillin acylase is thus much lower than 4-hydroxyphenylacetic acid with penicillin acylase. Therefore, the hydrolysis can only be completed when the 4-hydroxyphenylacetic acid is removed from the equilibrium.

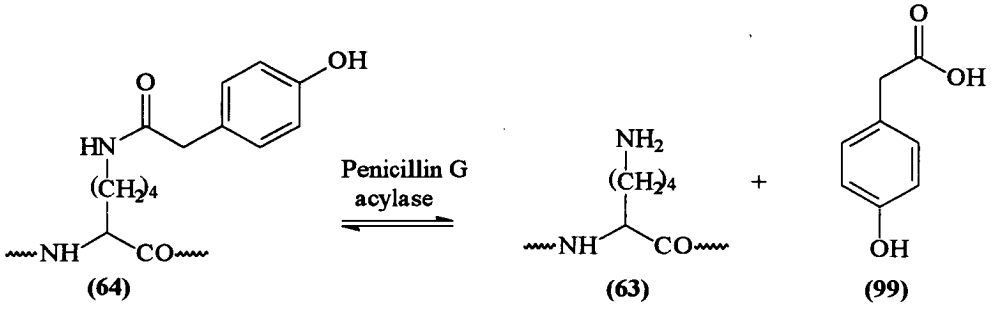


Figure 38 Hydrolysis equilibrium

This same difficulty was encountered by Wang in the deprotection of N-phenylacetylated insulin using immobilised penicillin G acylase.[92] However, Wang successfully removed the liberated phenylacetic acid shifting the hydrolysis reaction towards deprotection. This was achieved by the inclusion of a column of strong basic ion exchange resin connected to a column packed with the immobilised enzyme. The N-phenylacetylated insulin solution was then circulated through the two columns

which had been equilibrated with the same buffer at 37°C. The by-product, phenylacetic acid, was adsorbed by the basic resin allowing complete deprotection of the insulin. Attempts to reproduce these ingenious reaction conditions resulted in the Ub(Hopa)$_6$ **(98)** becoming adsorbed onto the ion exchanger and no removal of the Hopa protecting group was observed.

Another attempt to sequester the 4-hydroxyphenylacetic acid **(99)** and thus favour the forward reaction of the hydrolysis equilibrium was the addition of lysine to the reaction solution. Experiments utilising the Hopa protected pentapeptide **(64)** had revealed that the conditions which furnished the fastest deprotection were when the peptide was incubated with the immobilised enzyme in Tris buffer, pH 8.1 at 37°C with 50mM lysine.HCl. The lysine acts as a substrate for the reverse reaction i.e. between 4-hydroxyphenylacetic acid and an amine group. The small size of the lysine molecule compared to the protein means that it is more accessible to the active site and is thus more likely to react with 4-hydroxyphenylacetic removing it from the equation. However when these same reaction conditions were applied to the deprotection of the Hopa group in Ub(Hopa)$_6$ **(98)** they failed.

The 40 residue ubiquitin fragment [63-Lys(Hopa)]ubiquitin-(36-76) **(100)** was also incubated with penicillin G acylase in an attempt to deprotect just one Lys(Hopa) but those reactions also failed. However two Hopa protecting groups were incorporated into a peptide fragment for an azide fragment coupling reaction to obtain the 32 residue peptide hormone Salmon Calcitonin I (SCT).[33] Robertson then successfully deprotected the two Hopa groups using the conditions employed for the Hopa pentapeptide **(64)**. The deprotection was complete in 48 hours whereas there was no deprotection for Ub(Hopa)$_6$ **(98)** or [63-Lys(Hopa)]ubiquitin-(36-76) **(100)** using exactly the same conditions. Therefore Bepa is a suitable protecting group for use in the azide fragment condensation technique for small peptides (ca. 30 residues) but it is not applicable for larger peptide systems.

The failure to translate the enzymatic deprotection methodology from small peptides to proteins can be justified by examination of the crystal structure of penicillin G acylase (Figure 39). The three dimensional crystal structure of the enzyme shows that the catalytic serine residue, at the N-terminal of the B chain, is buried deep in the centre of the molecule. Clearly the relative inaccessibility of the active site means that whilst small unstructured molecules can easily access the binding pocket larger protein molecules (which will adopt some degree of structure under the reaction conditions) will have difficulty. The lower collision frequency of the protein with the active site of the enzyme will result in either incomplete deprotection or a shift in the equilibrium to favour the reverse reaction.

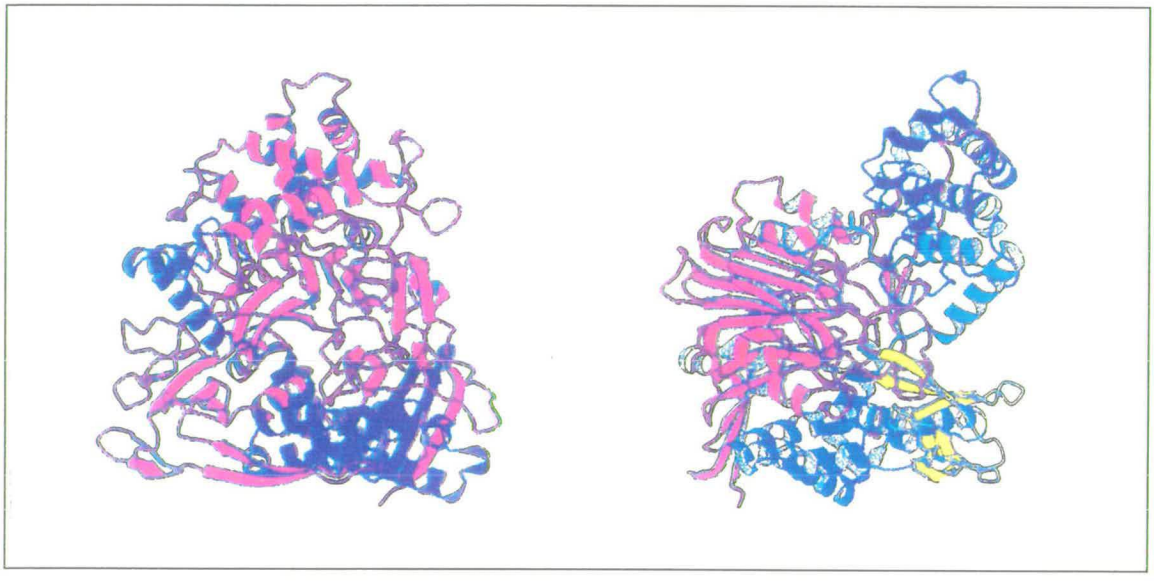


Figure 39 MOLSCRIPT representation of the three dimensional crystal structure of penicillin G acylase. (left) The A chain is shown in red wrapped around the B chain (blue), the side chain atoms of the active serine are shown at the apex of the deep depression at the centre and (right) an orthogonal view.

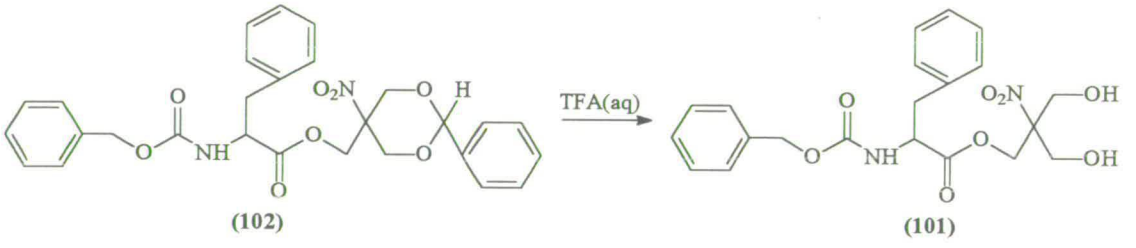
Variation of the structure of the substrate can significantly enhance the efficiency of penicillin acylase catalysed reactions. This approach of "substrate tuning" has been successfully employed for substrates of penicillin G acylase,[93] however it will not overcome the problems of steric hindrance encountered here. A strategy which might serve to conquer the problems would be the modification of the enzyme by genetic engineering. For example, if the active site could be made more accessible instead of

buried deep in the protein then the enzymatic deprotection of protected proteins may yet be possible. However, until such time as that, a new protecting group for lysine must be developed.

2.3: Design of a chemically cleaved N^{ε} lysine protecting group

2.3.1: Tris(hydroxymethyl)nitromethane (Tnm) derivatives as protecting groups in peptide synthesis

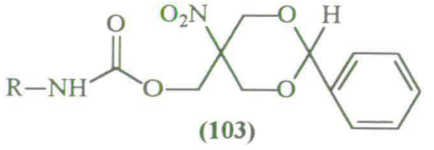
Research by Comer into a novel linker system for fragment condensation had lead to the evolution of a compound that was stable to Fmoc SPPS, partially cleaved by aqueous TFA and finally removed under basic conditions.[94] This tris(hydroxymethyl)nitromethane (Tnm) derivative (**101**) was then successfully applied as a protecting group for acid functionalities where the 1,3-diol would be protected as the acetal (**102**) during SPPS (Figure 40).



(102) (101)

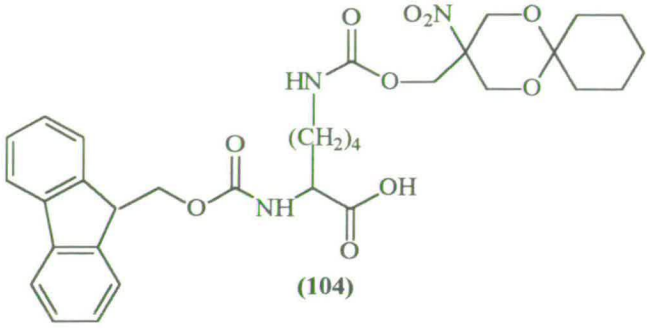
Figure 40 Tnm protection of acid functionality

After the success of the Tnm group in the protection of acid functions, it was decided to assess its potential as a protecting group for the N^{ε} amino function of lysine. The protection being in the form of the carbamate derivative of Tnm (**103**).



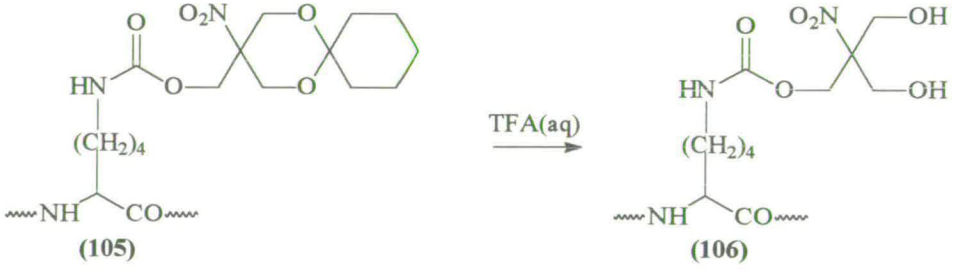
(103)

A comprehensive study by Comer incorporating different Tnm derivatives into model systems led to the optimisation of the lysine protecting group.[94] Consequently the new lysine derivative which is introduced during assembly of a peptide is N^{α}-9-Fmoc-N^{ε}-1:5-dioxaspiro-5:5-undecane-3-nitro-3-methoxycarbonyl-lysine (**104**).



(104)

As the acetal (**105**) this protecting group is stable to the basic deprotection conditions of Fmoc SPPS. After assembly of the peptide is complete the acetal is cleaved during the TFA resin cleavage to yield the 1,3-diol (**106**) which can easily be removed by treatment with base to give the amino function.



(105) **(106)**

Figure 41 TFA acidolysis of acetal (105) to 1,3-diol (106)

The synthesis of (**104**) was achieved by the synthetic route outlined in Figure 42. Cyclohexanone and tris(hydroxymethyl)nitromethane were refluxed in toluene using a Dean and Stark apparatus. The resulting alcohol (**107**) was attached to Fmoc-Lysine by firstly activating the alcohol as the chloroformate and then converting it to the more stable pentafluorophenylcarbonate (**108**).

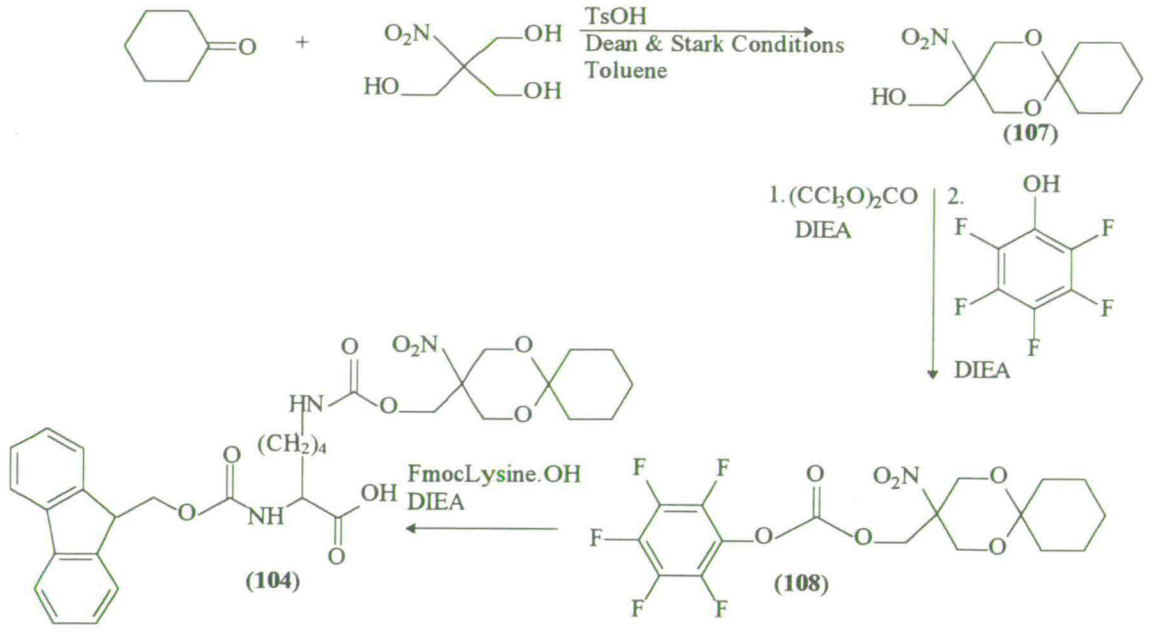


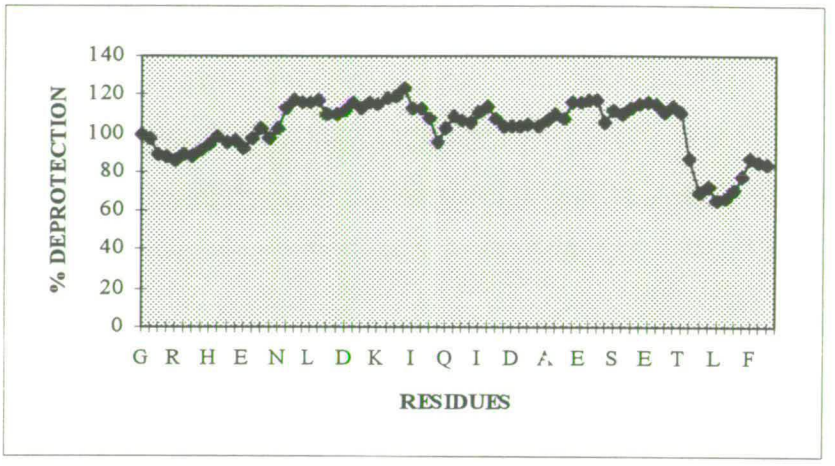
Figure 42 Synthetic route for (104)

2.3.2: Incorporation of Tnm protected lysine into ubiquitin

Comer successfully incorporated the Tnm group into model peptides and demonstrated its facile removal. The deprotection protocol utilised by Comer was to sonicate the Tnm protected peptide in DMSO/borate buffer pH 8.75 for thirty minutes.[94] These were very promising results involving one Tnm protected residue but the crucial test was the synthesis of the ubiquitin derivative with six lysines Tnm protected. If the Tnm protecting group could succeed where the enzyme-cleavable group had failed, its utility as a protecting group for fragment coupling experiments would be vindicated.

The ubiquitin derivative incorporating the six Tnm protected lysines and lysine-48 Boc protected was successfully constructed as outlined in the experimental section. An indication that the synthesis proceeded extremely well was obtained from the deprotection profile of the synthesis (Graph 1). The deprotection profile is not quantitative, it shows coupling percentages of greater than 100% which are probably associated with changes in the extent of resin swelling during the synthesis.[95] The

profile is analogous to those obtained from other syntheses of ubiquitin inferring that the lysine-Tnm derivative (**105**) is compatible with automated SPPS.



Graph 1 Deprotection profile for Ub(Tnm)$_6$

After solid phase synthesis a trial TFA resin cleavage was carried out to ensure that the acetal (**105**) was in fact cleaved to the 1,3-diol (**106**) (Figure 41). The product obtained was examined by MALDI MS and identified as the Tnm diol [6,11,27,29,33&63-Lys(Tnm)]ubiquitin (Ub(Tnm)$_6$) thus the first stage of the deprotection was successful. It was noted that this ubiquitin derivative was fully soluble compared to the Ub(Hopa)$_6$ (**98**) which was only partially soluble under the same conditions. Thus the Tnm group attributes solubility to the protected fragment which is an essential requirement for the lysine protecting group. However the principal demand is that the Tnm group can be easily removed under mild conditions. The next objective was therefore to remove the Tnm groups, fold the protein and show that the ubiquitin obtained was analogous to native ubiquitin.

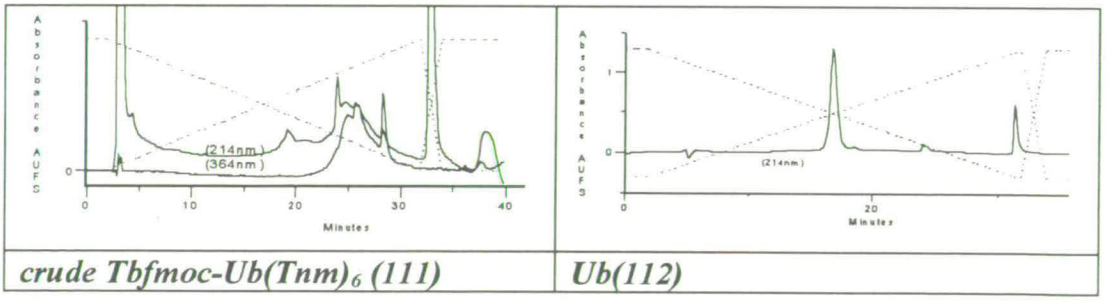
2.3.3: Purification of Ub(Tnm)$_6$ and deprotection of Tnm groups

2.3.3.1: Affinity purification of Ub(Tnm)$_6$ using Tbfmoc

The resin-bound Fmoc protected Ub(Tnm)$_6$ (**109**) was sonicated with capping solution before the N-terminal Fmoc group was removed and replaced with Tbfmoc (**37**) to give Tbfmoc-Ub(Tnm)$_6$-resin (**110**). The protein was then cleaved from the

resin with the simultaneous removal of all side chain protecting groups and the generation of six 1,3-diol Tnm protected lysines (**106**). The same cleavage conditions used for the Tbfmoc-Ub(Bepa)$_6$-resin (**96**) were employed and crude Tbfmoc-Ub(Tnm)$_6$ (**111**) was isolated as a white fluffy solid.

The Tbfmoc group can be deprotected by treating the Tbfmoc-peptide with 10% piperidine in 6MGdm.Cl/isopropanol. It was envisaged that these basic conditions would also deprotect the base labile Tnm group (**106**) making the Tbfmoc/PGC protocol a combined purification and deprotection step. The crude Tbfmoc tagged Ub(Tnm)$_6$ (**111**) was solubilised in the minimum amount of 6MGdm.Cl/isopropanol and vortexed with charcoal. Previous Tbfmoc purifications had used PGC as the adsorption medium but it was replaced by chromatographic grade charcoal which is much cheaper and equally effective. The Tbfmoc-Ub(Tnm)$_6$ (**111**) was adsorbed onto charcoal which was then washed to remove acetylated truncates. Once the truncated peptides had been removed the next step was to deprotect the Tbfmoc and the six Tnm groups. This was achieved by vortexing the charcoal with 10% piperidine in 6MGdm.Cl/isopropanol for ten minutes. After which the charcoal was centrifuged to a pellet the protein containing supernatant decanted off and neutralised to pH 4.5. The protein solution was reduced *in vacuo* to remove isopropanol and desalted by gel filtration and lyophilised to give Ub(**112**). Figure 43 shows HPLCs of Ub(**112**) and crude Tbfmoc-Ub(Tnm)$_6$ (**111**). The AAA and MALDI MS were as required for fully deprotected ubiquitin and analysis by HPLC showed that Ub(**112**) co-eluted with Ub(bovine). This was an excellent result which proved the efficacy of the Tnm protection/deprotection strategy.



crude Tbfmoc-Ub(Tnm)$_6$ (111) | *Ub(112)*

Figure 43 Analytical HPLCs of crude Tbfmoc-Ub(Tnm)$_6$ (111)and Ub(112)

Although it only takes ten minutes to deprotect the Tbfmoc and Tnm groups by vortexing with 10% piperidine in 6MGdm.Cl/isopropanol the pH may be as high as 13 or 14. Hence a milder method of purification and deprotection was also employed which involved separating the Tbfmoc-$(Tnm)_6$ **(111)** from any truncated peptides utilising size exclusion gel filtration with a Sephadex G50 (fine) column. The Tbfmoc containing fractions were pooled and lyophilised to give purified Tbfmoc-Ub$(Tnm)_6$ **(113)**. The non-Tbfmoc fractions **(114)** (truncations) were also pooled, lyophilised and examined by MALDI MS. This permitted the truncated peptide sequences to be identified and with this knowledge subsequent syntheses of ubiquitin may be improved.

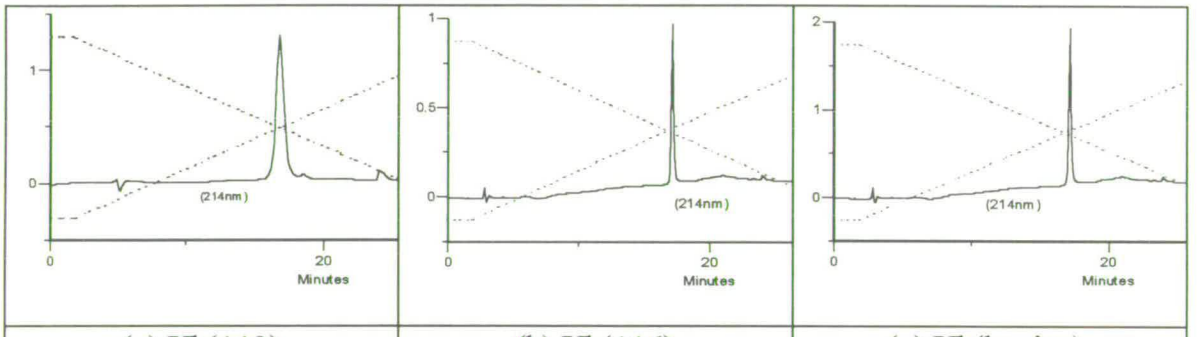
To deprotect the Tbfmoc and Tnm groups the purified Tbfmoc-Ub$(Tnm)_6$ **(113)** was dissolved in 6M Gdm.Cl/tris base pH 8.8 and sonicated at room temperature for 1.5 hours. The protein solution was then desalted by gel filtration and lyophilised to give Ub**(115)** which was equivalent to Ub**(112)**.

The HPLC profiles of Ub**(112)** and Ub**(115)** were broader than that of native ubiquitin. (Figure 44). This was explained by the presence of oxidised ubiquitin which eluted slightly earlier than native ubiquitin. Oxidised protein is formed if the methionine at position 1 is oxidised to the sulphoxide during the purification protocol. The broad HPLC profile was also because the synthetic ubiquitin was not folded correctly. It has been shown by Wilken that Tbfmoc purification causes ubiquitin to adopt a non-native structure.[88] However the native fold can be accessed *via* the use of cation exchange chromatography.

2.3.3.2: Reduction and folding of Ub(112)

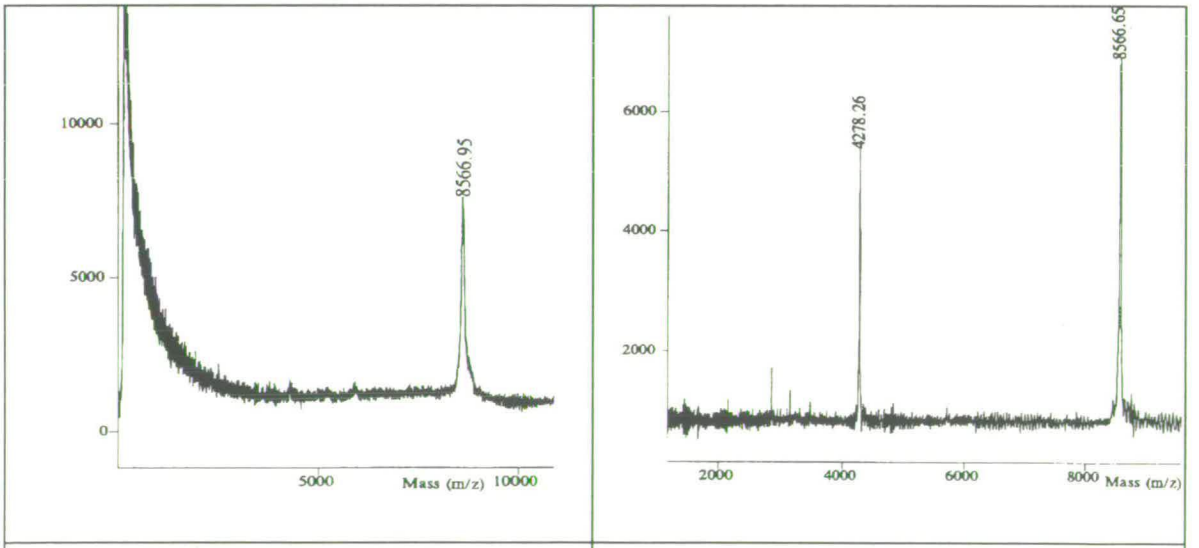
A portion of Ub**(112)** was dissolved in a 10% w/v solution of N-methylmercaptoacetamide (MMA) in aqueous acetic acid and incubated at 37^0C. The reduction reaction was monitored by HPLC and after 24 hours the protein solution was diluted and applied to a CM-Sepharose CL6B cation exchange column. The separation is based on the reversible adsorption of positively charged protein

molecules to a negatively charged matrix. The protein sample initially binds to the matrix and selective separation and elution of protein is achieved by application of a pH gradient followed by a salt gradient. This resulted in the isolation of Ub(**116**) which had a much sharper HPLC profile than Ub(**112**) before the cation exchange chromatography (Figure 44) thus indicating that cation exchange chromatography was crucial for the correct folding of ubiquitin. In addition, the HPLC profile and MALDI MS revealed that the ubiquitin obtained, Ub(**116**), was pure and indistinguishable from Ub (bovine).



(a) Ub(**112**) | (b) Ub(**116**) | (c) Ub(bovine).

Figure 44 Analytical HPLCs of (a) Ub(112), (b) Ub(116) and (c) Ub(bovine).



MALDI MS of Ub(116) | *MALDI MS of Ub(bovine)*

Figure 45 MALDI MS of Ub(116) and Ub(bovine)

The secondary structure content of Ub(**116**) was investigated using far UV circular dichroism (CD). Figure 46 shows the CD spectra for Ub(**116**) and Ub(bovine) and it is apparent that they have very similar profiles.

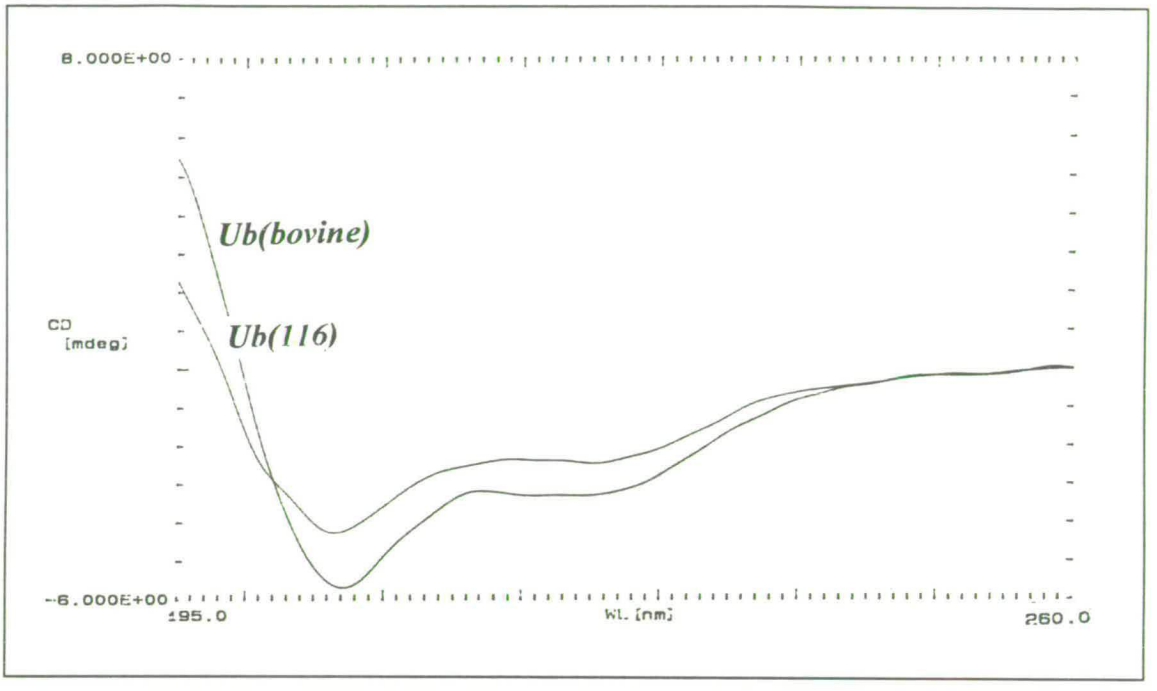


Figure 46 Far UV circular dichroism spectra for Ub(116) and Ub(bovine)

The CD analysis at pH 7 indicated both reduced α-helicity from 17% ± 1.6% in Ub(bovine) to 11% ± 1.5% in Ub(**116**), and an increased β-sheet of 60% ± 2.8% in Ub(**116**) *versus* 53% ± 2.6% in Ub(bovine). These results were comparable to other CD analysis obtained by Wilken for synthetic ubiquitin.[88] The deviations in the CD analysis between the synthetic material and Ub(bovine) may be largely due to a concentration effect since any salt present in the synthetic material would alter the dry weight and hence the concentration. The CONTIN procedure used to estimate the secondary structure is a limited program which is very dependent on concentration. However the CD analysis did indicate a similar secondary structure in Ub(**116**) and Ub(bovine).

2.3.3.3: Non-Tbfmoc purification protocol

To confirm the general application of the purification/deprotection strategy for Tnm protected peptides a non-Tbfmoc protocol was also investigated. The N^{α}-Fmoc group was removed from **(109)** and the protein cleaved from the resin using the same cleavage conditions employed for its Tbfmoc counterpart. The precipitated $Ub(Tnm)_6$ **(117)** was purified by size exclusion gel filtration using a Sephadex G50 (fine) column to afford $Ub(Tnm)_6$ **(118)** and the Tnm groups removed by sonicating **(118)** in 6M Gdm.Cl/trizma base pH 8.8 for 1.5 hours. The protein recovered contained fully deprotected ubiquitin, Ub**(119)**, but also present were truncated peptides of similar mass to ubiquitin. Thus gel filtration using a Sephadex G50 column was not sufficiently resolving to allow separation of ubiquitin from large truncated peptides.

It was hoped that FPLC size exclusion would afford better resolution. An aliquot of **(117)** was sonicated in 6MGdm.Cl, centrifuged and the supernatant loaded onto a Superdex™ 75 HR 10/30 column. However the $Ub(Tnm)_6$ **(120)** which eluted was identical to that obtained from gel filtration using the Sephadex column.

2.3.4 Utilisation of Tnm group in azide fragment condensation

The utility of the Tnm protecting group for fragment coupling reactions has recently been demonstrated by Robertson with the condensation of small fragments of the protein erythropoietin.[33] The hexapeptide azide fragment **(121)** was successfully coupled to a fragment containing a Tnm protected lysine residue **(122)**. The Tnm group was stable to the reaction conditions and is thus suitable for use in azide fragment couplings. The protected peptide **(123)** was isolated by preparative HPLC and the Tnm N^{ε} and Fmoc N^{α} protecting groups were removed simultaneously with a 10% piperidine solution to give the fully deprotected peptide **(124)**.

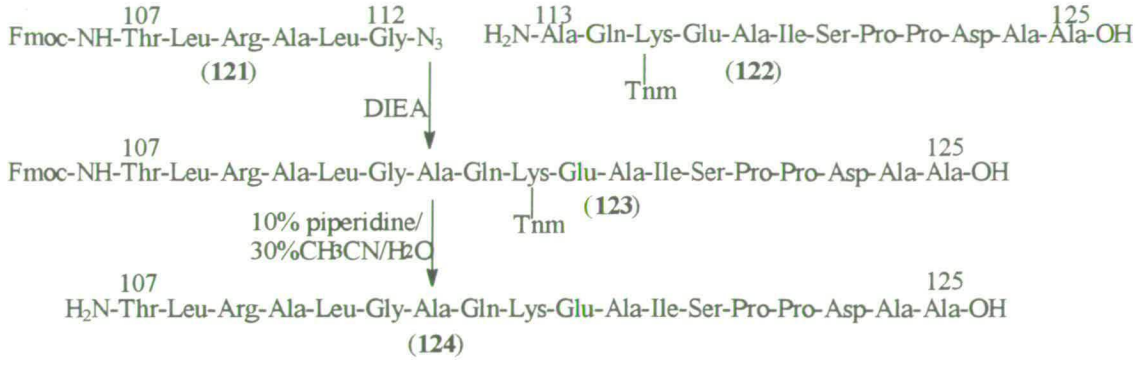


Figure 47 Azide fragment coupling of two peptide fragments utilising the Tnm protecting group for lysine

2.4: Conclusions and future work

Bepa was successfully incorporated into Fmoc SPPS of peptides and proteins and afforded the Lys-Hopa moeity after the TFA resin cleavage. However, although the Hopa moiety is a suitable protecting group for use in azide fragment condensations for small peptides (ca. 30 residues) it is not applicable for larger peptide systems.

It has been demonstrated that the Tnm group is a viable protecting group for the N^{ε} amino function of lysine. It attributes solubility to the protected peptide fragment, is stable to the conditions required for fragment condensations and is easily deprotected. From the investigations on $Ub(Tnm)_6$ it was apparent that derivatising the N-terminal with Tbfmoc facilitated purification of the protein. The Tbfmoc/charcoal protocol allowed the acquisition of synthetic ubiquitin which was homogeneous and indistinguishable from bovine ubiquitin. The Tbfmoc group also permitted the protein to be purified by gel filtration keeping the Tnm groups in place which would obviously be desirable for the coupling of purified protected protein fragments. The facile deprotection of the Tnm group was also achieved using basic solution.

Now that the Tnm protection/deprotection strategy has been validated then coupling of two ubiquitin fragments can be attempted. This work is presently being undertaken by Wang[96] at Edinburgh. The strategy for the synthesis of di-ubiquitin utilises the azide method (Section 1.2.3.5) for fragment condensation. This means that a second ubiquitin fragment would be synthesised with all seven lysine residues Tnm protected and with a hydrazide linker at the C-terminal. Thus after resin cleavage a Tnm protected ubiquitin with an activated acyl azide at the C-terminal could be produced. This would then be coupled to the first fragment *via* the free N^{ε} amino group at lysine-48 forming Tnm protected di-ubiquitin. The Tnm groups should then be removed by treatment with base to yield a branched protein with a peptide linkage involving the ε-amino group of lysine.

Chapter 3: Stepwise SPS of deglycosylated human interferon-*gamma*

3.1: General introduction

3.1.1: Protein chemistry

The upsurge of interest in protein chemistry, has to a large extent, been stimulated by the establishment of a biotechnology industry based on the production of genetically engineered proteins. Generally, biotechnology drugs (biopharmaceuticals) are molecules or derivatives of molecules that originate from living organisms and exert a biological effect,[97] for example growth hormone, the interferons and insulin. Since these drugs are large complex molecules that are often charged their delivery into the body at levels that are effective but non-toxic, is a demanding goal. In fact drug delivery of such substances is thought to represent the potential Achilles' heel for the biotechnology industry.[98] Currently, with few exceptions, drug delivery is by transdermal injection, however the skin is a formidable barrier, and some researchers remain sceptical about the usefulness of transdermal delivery routes. Nevertheless transdermal technology, namely sonophoresis, has been used to deliver therapeutic doses of proteins such as insulin, γ-interferon and erythropoeitin across the skin.[99,100] Sonophoresis, which involves applying low-frequency ultrasound, shows great promise however it requires further testing to optimise its performance.

The biotechnology industry has benefited from modern molecular biology techniques, such as cloning and cDNA sequencing of genes, which have made available amino acid sequence data of proteins. In many cases, the proteins themselves have never been isolated and their properties are largely or completely unknown. Chemical synthesis of peptide fragments, or of the whole molecule, can be used to explore the structure and function of such gene products. These studies may lead to the identification of the most active protein fragment which would enable proteins to be redesigned as second generation products for human pharmaceutical use.

3.1.2: Interferon-*gamma*

Human interferon-*gamma* (hIFN-γ) is a potent cytokine produced naturally by activated T cells and natural killer cells.[101,102] Although originally isolated based on its antiviral activity, hIFN-γ also exhibits powerful antiproliferative and immunomodulatory activities. These biological activities are essential for developing appropriate cellular defences against a variety of infectious agents. In addition hIFN-γ may play a key role in multiple sclerosis[103] and the onset of inflammatory reactions. Therefore hIFN-γ is an important therapeutic protein and thus an ideal synthetic target for SPS.

The amino acid sequence of hIFN-γ was originally deduced from a cDNA clone which predicted the existence of a signal sequence followed by 146 amino acids starting with the tripeptide sequence Cys-Tyr-Cys.[104] However, the amino-terminus of mature hIFN-γ, purified from induced peripheral blood lymphocytes, differed from the sequence predicted by cloning. In natural hIFN-γ the first three residues, Cys-Tyr-Cys, were missing to give a 143 amino acid polypeptide chain and the amino-terminus consisted of a pyroglutamate residue (Figure 48).[105] In addition, natural hIFN-γ was shown to consist of a mixture of molecules, all with identical amino-termini, but they displayed some variability in glycosylation and heterogeneity at the carboxyl-termini with the longest polypeptide being 138 amino acids.[105,106] The heterogeneity at the carboxyl-terminus was proposed to be the result of peptidase digestion either during or after secretion of the protein.[105]

```
 1                                          10                                          20
Gln Asp Pro Tyr Val Lys Glu Ala Glu Asn Leu Lys Lys Tyr Phe Asn Ala Gly His Ser
                   *                       30                                          40
Asp Val Ala Asp Asn Gly Thr Leu Phe Leu Gly Ile Leu Lys Asn Trp Lys Glu Glu Ser
                                           50                                          60
Asp Arg Lys Ile Met Gln Ser Gln Ile Val Ser Phe Tyr Phe Lys Leu Phe Lys Asn Phe
                                           70                                          80
Lys Asp Asp Gln Ser Ile Gln Lys Ser Val Glu Thr Ile Lys Glu Asp Met Asn Val Lys
                                           90                                  *       100
Phe Phe Asn Ser Asn Lys Lys Lys Arg Asp Asp Phe Glu Lys Leu Thr Asn Tyr Ser Val
                                          110                                         120
Thr Asp Leu Asn Val Gln Arg Lys Ala Ile His Glu Leu Ile Gln Val Met Ala Glu Leu
                                          130                                         140
Ser Pro Ala Ala Lys Thr Gly Lys Arg Lys Arg Ser Gln Met Leu Phe Arg Gly Arg Arg
         143
Ala Ser Gln
```

**** N-Linked glycosylation sites***

Figure 48 Amino acid sequence of hIFN-γ showing glycosylation sites

hIFN-γ has two potential N-linked glycosylation sites at positions Asn-25 and Asn-97 of its 143 amino acid long mature polypeptide chain (Figure 48). There is considerable heterogeneity in the core glycosylation of hIFN-γ with the molecules containing 0, 1 or 2 glycan residues.[107] The biologically active hIFN-γ is a homodimer. There are no cysteine residues in the mature hIFN-γ and thus the polypeptide lacks intramolecular disulphide bonds.

The expression of hIFN-γ in *Escherichia coli*[108,109] has resulted in the preparation of large quantities of highly purified recombinant hIFN-γ. As a result recombinant hIFN-γ has been fully characterised, is licensed for clinical use and detailed studies of structure-function relationships have been initiated. The three-dimensional structure of recombinant hIFN-γ has been resolved (Figure 49 (a)).[110] The protein is primarily α helical, with six helices in each subunit that comprise 62% of the structure; there is no β sheet. The six helices are labelled A-F from the N-terminus. Helix F is packed in a cleft created by helices A, B, C and D. With the exception of the AB loop, the helices are linked by short loops of 3 to 5 residues. The subunits are associated to each other in an antiparallel fashion and the molecules have multiple interactions along the polypeptide chains, forming a compact, relatively globular symmetrical dimeric structure.

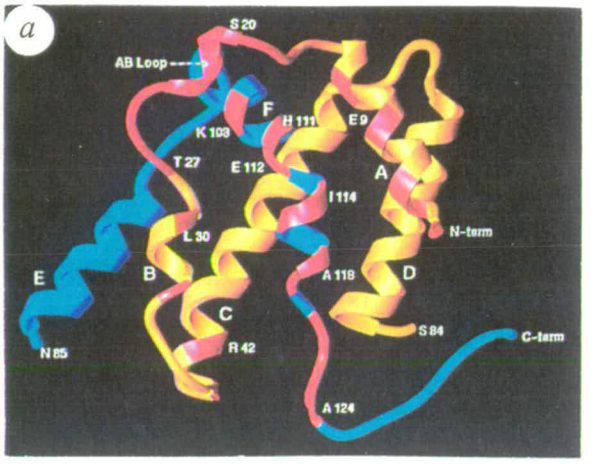


Figure 49(a) Domain structure of rhIFN-γ.
Residues 1-84 from chain 1 (yellow) and residues 85-132 (cyan) from chain 2.

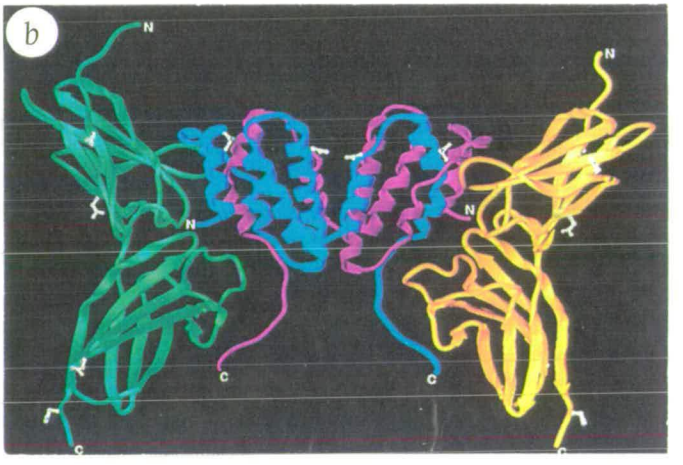


Figure 49(b) Complex between rhIFN-γ and its soluble high-affinity receptor.[111]
Receptor 1 (green) receptor 2 (yellow). Potential N-linked glycosylation sites are also shown.

Expression of biological activity appears to be mediated through binding to specific cell-surface receptors,[112] which have now been cloned.[113] The hIFN-γ dimer has two receptor binding-sites and it can simultaneously bind to two receptor molecules (Figure 49 (b)).[114] All four potential N-linked glycosylation sites are located on the surface of the hIFN-γ dimer.[110] Glycosylation is apparently not required for dimerisation of the subunits, since recombinant hIFN-γ produced in *Escherichia coli*

forms dimers which are biologically active.[115] Glycosylation may, however protect hIFN-γ against cellular proteases and prolong its survival time in the circulation.[116]

Both carboxyl-terminal truncation and glycosylation do not appear to affect either quaternary structure or biological activity of hIFN-γ.[110,117,118] In fact the removal of a limited number of carboxy-terminal residues dramatically enhances the biological activity.[119] Maximal enhancement was found with the deletion of 9 to 10 amino acids but the polypeptide is rendered inactive on the removal of more than 14 residues. From these results Döbeli proposed that hIFN-γ is synthesised as a preprotein and that its activity is modulated by proteolytic digestion.[119] Such post-translational modifications require further investigation to determine if there are other fragments with increased biological activity. The preparation of truncated hIFN-γ molecules with greater than 11 deleted residues requires an extreme expenditure of DNA synthesis, due to unfavourable cleavage sites for restriction enzymes.[119] An alternative route to obtain carboxy-terminally truncated hIFN-γ is proteolytic digestion by enzymes. However, even highly specific enzymes usually lead to a wide spectrum of fragments which are difficult to purify. Therefore, a major advantage of chemically produced hIFN-γ is that the amino acid sequence can be specifically modified to synthesise the most biologically active protein fragment and will avoid C-terminal heterogeneity.

3.2: Objectives

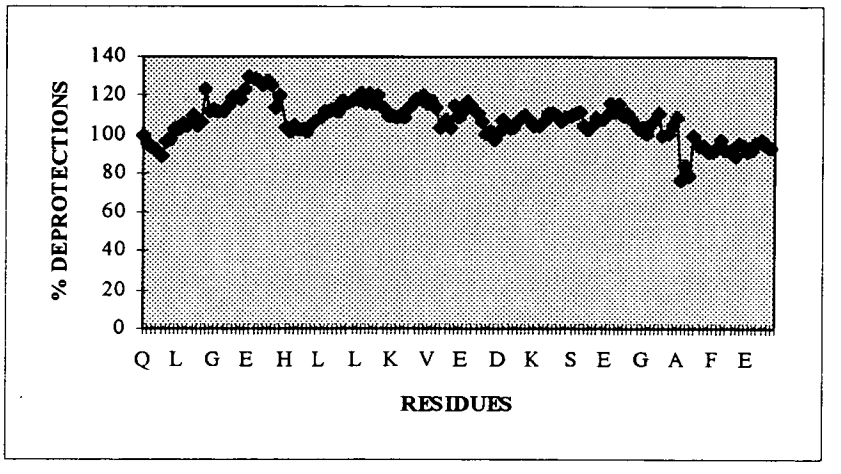
The aim of this project was to construct deglycosylated hIFN-γ (143 amino acids, 16.8kDa) by applying solid phase peptide synthesis. Gln hIFN-γ was synthesised because it is the first processed product which spontaneously cyclises to give the pyroglutamic N-terminus. On completion of the synthesis, the aim is to develop an effective purification protocol to obtain homogeneous synthetic protein. The attainment of pure protein will allow the biological activity to be determined and

should lead to crystallisation and structural data. Finally both the glycosylated form of hIFN-γ and stabilised, modified analogues could be synthesised.

3.3: Chemical synthesis of deglycosylated hIFN-γ (dhIFN-γ)

The pioneering stepwise syntheses of the obese gene product,[31] erythropoietin[33] and the catalytic domain of Stromelysin[120] have demonstrated that the methodology in this laboratory can be applied to larger polypeptide systems. The use of the superior coupling agent HOCt (**26**) was crucial, permitting the synthesis of proteins previously unattainable using HOBt (**24**).

The stepwise assembly of the dhIFN-γ sequence (Figure 48) was carried out on a 0.1mmol scale using Fmoc-Gln(Trt), 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 44 residues which were double coupled. It was hoped that double coupling from Ile-44 to Gln-1 would reduce the formation of large truncations and increase the yield of the desired sequence. From the deprotection profile the stepwise SPS of dhIFN-γ appeared to have proceeded extremely well with no obvious large drops in the percentage coupling of the amino acids (Graph 2).



Graph 2 Deprotection profile for dhIFN-γ

The success of the synthesis was confirmed by submitting resin-bound protein for amino acid analysis. The values obtained from the amino acid analysis were in good agreement with the expected composition (Table 3).

Amino acid	Found	Expected
Asx	22.4	20
Thr	5.4	5
Ser	10.0	11
Glx	21.0	18
Pro	3.7	2
Gly	5.3	5
Ala	7.7	8
Cys	0	0
Val	9.3	8
Met	2.3	4
Ile	7.9	7
Leu	12.0	10
Tyr	3.9	4
Phe	12.0	10
His	2.7	2
Lys	20.4	20
Arg	6.3	8

Table 3 A.A.A. of resin bound dhIFN-γ

The SPS of dhIFN-γ was carried out on only one occasion consequently limited material was available to determine the best purification strategy. The following sections describe the efforts made to characterise the product and to identify a suitable purification protocol to obtain homogeneous protein.

3.4: Purification of dhIFN-γ

As outlined in Section 1.2.2 the main obstacle to the success of stepwise SPS of proteins can be attributed to the generation and subsequent inability to separate closely related sequences from the target protein. Protracted chromatographic purification is usually required to isolate the protein, and it is sometimes even impossible to separate the desired protein from the physically and chemically similar truncated sequences. Derivatisation of the free N-terminus of the resin-bound polypeptide with the Tbfmoc group (37) allows differentiation between the target

sequence and any acetylated truncations. Tbfmoc methodology is one technique that can be utilised in the purification of dhIFN-γ but it was decided to examine the methods employed for the purification of natural and recombinantly produced hIFN-γ.

3.4.1: Literature purification of hIFN-γ

The partial purification of IFN-γ from human peripheral leukocytes was reported in 1981 by Yip using sequential chromatographic separations.[121] Subsequently natural hIFN-γ was purified to homogeneity utilising a concanavalin A-Sepharose column and RP HPLC.[105] Concanavalin A binds molecules which contain α-D-mannose and α-D-glucose therefore it is exploited as a means to separate and purify glycoproteins. In 1986 Miyata reported a purification procedure for natural hIFN-γ based on monoclonal antibody affinity chromatography which gave essentially quantitative recovery of the protein.[122] Both purification procedures resulted in two active species with apparent molecular weights of 20,000 and 25,000, as judged by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), corresponding to mono and di-glycosylated hIFN-γ respectively.

The purification of recombinant dhIFN-γ produced in *Escherichia coli* has also been achieved utilising antibody affinity chromatography.[123] However, Döbeli utilised sequential chromatographic techniques to purify dhIFN-γ and the carboxy-terminally truncated analogues.[119] First Döbeli applied the crude extract to a silica gel column then the eluate containing the dhIFN-γ was pumped through two hydrophobic interaction columns followed by a cation exchange column and finally a desalting step using a Sephadex column. Although this appears a somewhat lengthy procedure the protein obtained was homogeneous by SDS-PAGE and biologically active.

The protein produced by recombinant methods is not glycosylated and is therefore analogous to the dhIFN-γ which was synthesised by SPS. The objective of this research was not only to develop a purification protocol for dhIFN-γ but also one that could be applied to the general purification of other proteins. Consequently the

research of Döbeli was applicable since it did not involve specialised antibodies but instead employed standard purification techniques available in most laboratories.

3.4.2: Proposed strategy for the purification of dhIFN-γ

The proposed purification strategy for IFN-γ encompassed standard peptide purification techniques and those used to purify recombinantly produced dhIFN-γ and is outlined below:

- Attachment of Tbfmoc
- Cleavage from resin
- Tbfmoc/charcoal purification
- FPLC size exclusion
- Hydrophobic interaction chromatography (HIC)
- Isoelectric focusing (IEF)
- Denaturing and folding
- Ion exchange chromatography

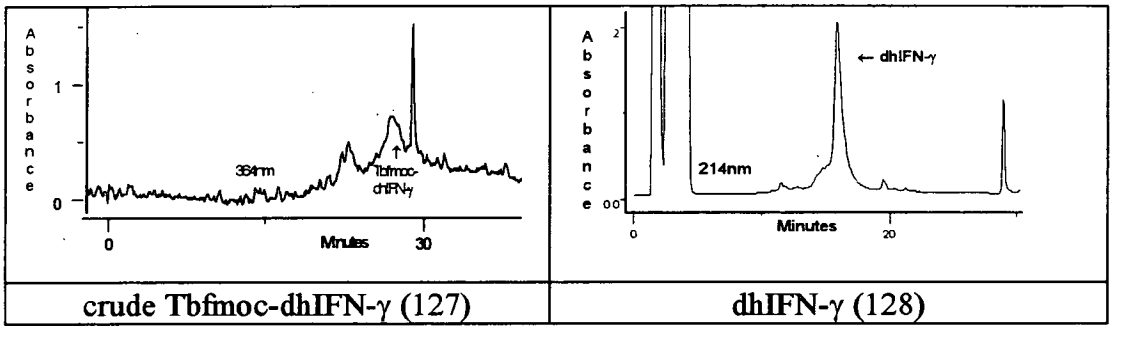
3.4.3: Affinity purification of dhIFN-γ using Tbfmoc on charcoal

On completion of the synthesis, the resin-bound Fmoc protected dhIFN-γ **(125)** was sonicated in the presence of acetic anhydride/HOBt/DIEA in DCM in order to cap any free amino groups. The N-terminal Fmoc group was then removed and Tbfmoc introduced at the N-terminus. The next step was to cleave the Tbfmoc labelled dhIFN-γ **(126)** from the resin and simultaneously remove all side chain protecting groups. It was desirable to ascertain the minimum time that afforded complete cleavage since subjecting the protein to TFA for a long period of time might cause cleavage of the polypeptide chain although this has not been proven. Therefore a trial resin cleavage was carried out in order to establish the optimum time. This was accomplished by stirring a portion of the resin-bound material in a scavenger cocktail of TFA, EDT, thioanisole, water and phenol. Samples of the cleavage mixture were analysed by HPLC after 3.5 hours and every half an hour after that until 6 hours had

passed. After 4.5 hours there was no further change in the HPLC profile and thus 4.5 hours was taken as the optimum cleavage time.

A large scale resin cleavage was then performed on (**126**) and after 4.5 hours the resin was removed by filtration and the crude protein was precipitated from the filtrate with diethyl ether. The crude Tbfmoc labelled protein (**127**) was isolated by centrifugation and lyophilisation.

Tbfmoc adsorption methodology, using charcoal as the solid phase, was applied as the primary purification of Tbfmoc-dhIFN-γ (**127**). Denaturing conditions (6MGdm.Cl) were used, as outlined in the experimental section. The Tbfmoc purified dhIFN-γ/guanidine solution (**128**) obtained was examined by RP-HPLC and observed as a single broad peak (Figure 50).



crude Tbfmoc-dhIFN-γ (127)	dhIFN-γ (128)

Figure 50 HPLC profiles of crude Tbfmoc-dhIFN-γ (127) and dhIFN-γ (128)

The dhIFN-γ (**128**) was characterised by N-terminal sequencing, amino acid analysis, MALDI MS and SDS-PAGE. These techniques confirmed that the dhIFN-γ sequence had been synthesised correctly and the latter two techniques allowed the purity of the material to be assessed. The MALDI MS and SDS-PAGE reproducibly revealed 4 components of approximate molecular weight 16.8, 14.3, 13.2 and 10.5kDa (Figure 51).

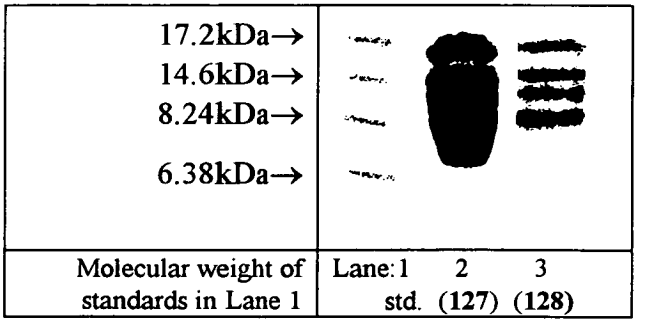


Figure 51 SDS-PAGE of Tbfmoc-dhIFN-γ (127) and dhIFN-γ (128)

It was proposed that the impurities revealed by the SDS-PAGE gel and mass spectra could have arisen from three sources. Firstly the acetylated truncated materials may not have been removed during the washing of the charcoal, perhaps because they themselves had bound to the charcoal via aromatic side chains. Or secondly the protein may have been sheared during the TFA resin cleavage. If the protein backbone was cut from the C-terminus then truncated fragments that were tagged with Tbfmoc would be carried through the purification protocol. The impurities could also have arisen from deletion sequences. Deletion peptides are similar to truncated peptides but, unlike truncated sequences they are not acetylated during the capping step. This could occur during the synthesis if the resin swelling is not sufficient to expose all the free amino functions for coupling. However, later in the synthesis the extent of resin swelling might alter and the amino functions would again become available for coupling. Thus deletions could be formed that are missing residues from the middle of the desired sequence. When Tbfmoc is incorporated onto the N-terminus of the target sequence the deletion sequences, which have not been capped, would also be labelled with Tbfmoc and carried through the purification protocol.

3.4.4: FPLC purification of dhIFN-γ (128) by gel filtration

The next purification technique employed was FPLC gel filtration adopting the principle of size exclusion (largest molecular weight elutes first). dhIFN-γ (**128**) was

loaded onto a Superdex™ 75 column which was then eluted isocratically with 6M Gdm.Cl. Fractions absorbing at 280nm were examined by analytical HPLC. Fractions 8-15 eluted with the correct retention time (R_t=17.2mins., 51%B) however fractions 8-10 gave the sharpest HPLC profiles (Figure 52).

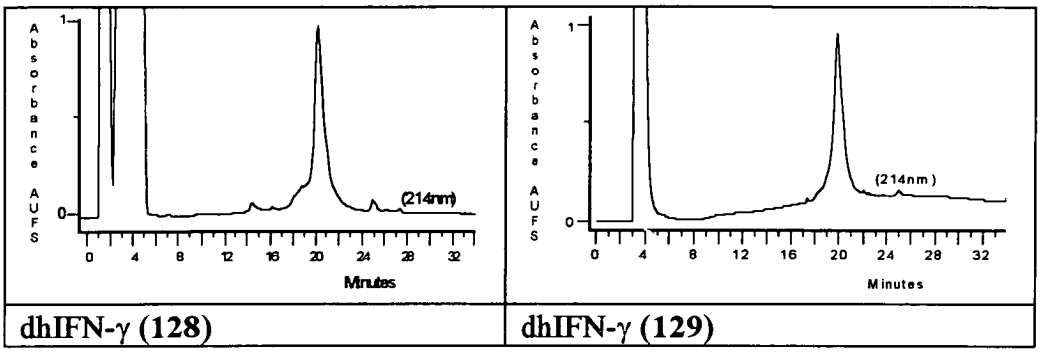


Figure 52 HPLCs of dhIFN-γ (128) (fraction 11) and dhIFN-γ (129) (fractions 8,9)

Individual fractions were desalted, examined by SDS PAGE PhastGel and mass spectra obtained. Fraction 8 and 9 were seen to consist of 3 components of approximate molecular weight 16.8, 14.3, 13.2kDa. Whilst the later fractions gave the same composition as dhIFN-γ (**128**). Fractions 8 and 9 were combined to give dhIFN-γ (**129**) and used for subsequent purification steps. Figure 53 shows the SDS-PAGE Phast gel of dhIFN-γ (**129**) (fractions 8,9) and dhIFN-γ (**128**) (fraction 11).

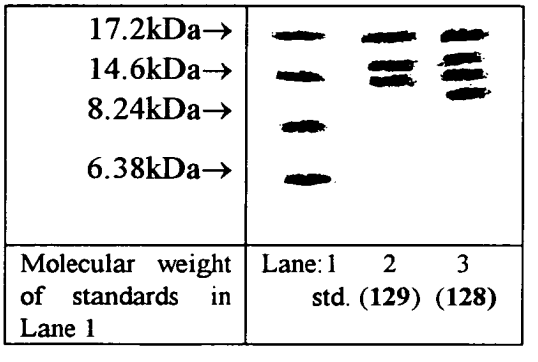


Figure 53 SDS-PAGE of dhIFN-γ (129) (fractions 8,9) and dhIFN-γ (128) (fraction 11)

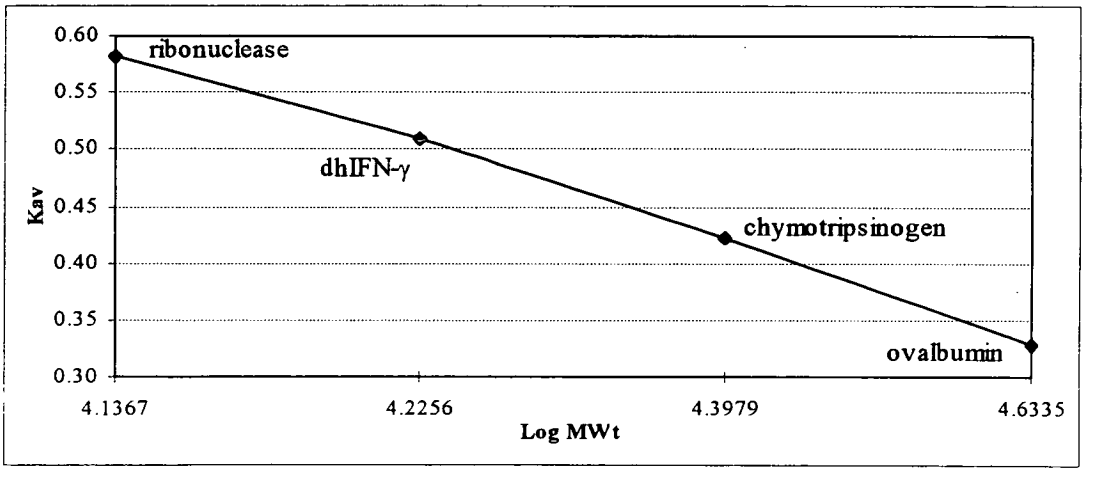
3.4.4.1: Molecular weight determination of dhIFN-γ (128) by FPLC

The FPLC experiments could also be used to determine the molecular weight of the eluted protein and hence gain further evidence that the correct sequence relating to dhIFN-γ had been synthesised. This was achieved by loading known molecular weight standards (Ovalbumin, Mwt 43, Chymotrypsinogen A, MWt 25 and Ribonuclease A, MWt 13.7kDa) onto the column and measuring their elution volumes. A graph was then plotted of the elution volume parameters (k_{AV}) of the standard *versus* the logarithm of their molecular weight. Where k_{AV} is calculated from the following equation:

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

where Ve = elution volume, Vt = column bed volume and Vo = column void volume

The k_{AV} value for dhIFN-γ was calculated from the FPLC experiments and the logarithm of its molecular weight was found from the graph (Graph 3) allowing the molecular weight of dhIFN-γ to be determined. The logMWt value calculated for dhIFN-γ was 4.2256 which compared well with the expected value of 4.2247.



Graph 3 Plot of logMwt vs k_{AV}

3.4.5: Folding of dhIFN-γ (129)

To be biologically active, all proteins must adopt specific folded three-dimensional structures. These thermodynamically favoured structures are maintained by relatively weak interatomic forces such as hydrogen bonding, hydrophobic interaction and ionic interactions.[124]

Many recombinant proteins have been successfully folded and their activity recovered using the following folding procedure. After isolating solid protein pellets from the cell, the protein is solubilised in a strong denaturant such as Gdm.Cl or urea. The denaturant concentration is then reduced by dilution or dialysis to allow the protein to refold to its native state. The protein can aggregate during refolding and, thereby, the recovery of active protein is reduced. Several approaches have been investigated to reduce aggregation, one of which is inclusion of folding aids in the dilution buffer.[125,126] The addition of the cosolvent, polyethylene glycol (PEG), to the dilution buffer has been shown to enhance refolding of active recombinant proteins including rhIFN-γ.[127] Cleland's results indicate that PEG is probably binding to specific segments of the protein to prevent aggregation and hence increase the recovery of active protein.[127]

The experimental conditions of Cleland were utilised for dhIFN-γ (**129**)[127] and the folding was monitored by analytical HPLC. It must be noted that the protein sample contained three components, as determined by SDS-PAGE and MALDI MS, and the purpose of the experiment was to test the literature folding protocol. The HPLC profile did become sharper which implied that the protein was folded (Figure 54). However conclusive evidence that the protein had folded in the correct configuration can only be obtained if the biological activity is tested.

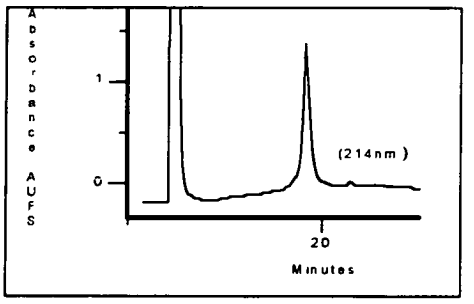


Figure 54 RP HPLC of dhIFN-γ (129) after folding

3.5: Characterisation of dhIFN-γ (129) (3 components 16.8, 14.3 and 13.2 kDa)

At this stage it was decided to fully characterise the protein, dhIFN-γ (129), obtained after FPLC size exclusion and folding to obtain evidence that the correct sequence had been constructed.

3.5.1: Tryptic digest and mass spectrometry data

The acquisition of MALDI MS for the entire sequence of synthetic proteins, such as erthropoietin and the obese gene has proved problematic however, mass spectra were obtained for dhIFN-γ. Mass spectrometry data for dhIFN-γ (16.8kDa) revealed the presence of three main components at 16.8, 14.3 and 13.2kDa (Figure 55).

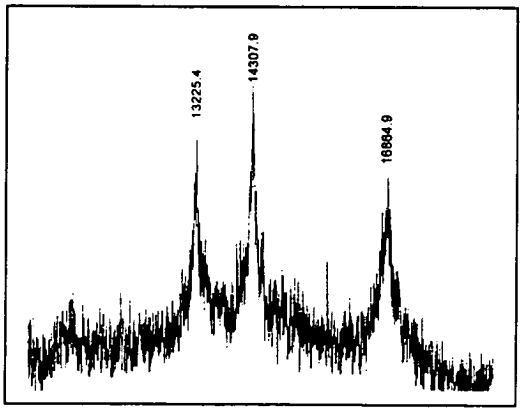


Figure 55 MALDI MS of dhIFN-γ (129)

A technique employed to acquire accurate mass spectra data is the hydrolysis of the intact protein into smaller fragments which can readily be examined by MALDL MS. The most suitable proteolytic enzyme was trypsin since it has only two cleavage sites (cuts at the C-terminus of Lys and Arg) which should minimise the number of fragments obtained. Also it should be able to withstand the denaturing conditions required for the reaction. The hydrolysis was carried out in 3M Gdm.Cl solution since dialysis of the protein into low salt had resulted in its precipitation.

Protein digestion followed by MS analysis of the mixture showed evidence for five main sites of cleavage. Not all possible cuts were observed but the molecular mass of some of the predicted fragments would have been too small to detect by MALDI MS. The results are shown in Table 4 below:

	Mass	
Fragment	**Found**	**Expected**
90-107	2160.8	2158.1
109-125	1822.4	1822.0
95-108	1675.1	1674.9 (Na^+)
43-55	1642.2	1642.8 (Na^+)
75-86	1496.6	1496.7 (Na^+)

Table 4 MALDI MS data obtained from tryptic digest

From the tryptic digest results there was no evidence for the presence of the N-terminal amino acids (1-12). 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

The intact protein was then sequenced from the N-terminus[128] in an attempt to characterise this part of dhIFN-γ. This was performed by Edman degradation labelling the N-terminus with phenyl isothiocyanate.[129]

A total of 15 rounds of Edman degradation were performed satisfactorily on dhIFN-γ **(129)**. Thus confirming the correct sequence of the N-terminal region and providing more evidence that the correct material had been synthesised.

3.5.3: Western blot analysis

Western blotting[130] using monoclonal mouse anti-hIFN-γ antibody (kindly supplied by Department of Pathology, University of Edinburgh) was carried out on dhIFN-γ **(129)**. The bands obtained were compared with a set of molecular weight markers and the gel showed that the antibody was active for the synthetic dhIFN-γ **(129)**. The bands were within the correct molecular weight for dhIFN-γ (~17kDa) (Figure 56). No lower molecular weight bands were present but a band at higher molecular weight was observed (~34kDa). This could be due to the presence of the dimer of dhIFN-γ which could have formed due to a greater concentration of dhIFN-γ compared to SDS.[131]

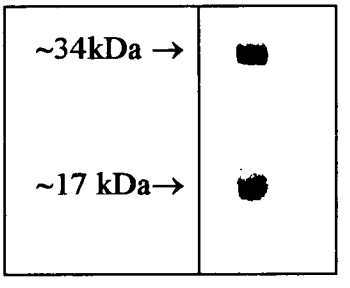


***Figure 56 Western blot analysis of dhIFN-γ* (129)**

3.5.4: Far UV circular dichroism

Crystallographic studies on human IFN-γ[110] and rabbit IFN-γ[132] confirmed the high α-helical content of 60-70% which had been predicted on the basis of UV circular dichroism.[133,134] The secondary structure content of dhIFN-γ **(129)** was investigated using far UV circular dichroism (Figure 57).

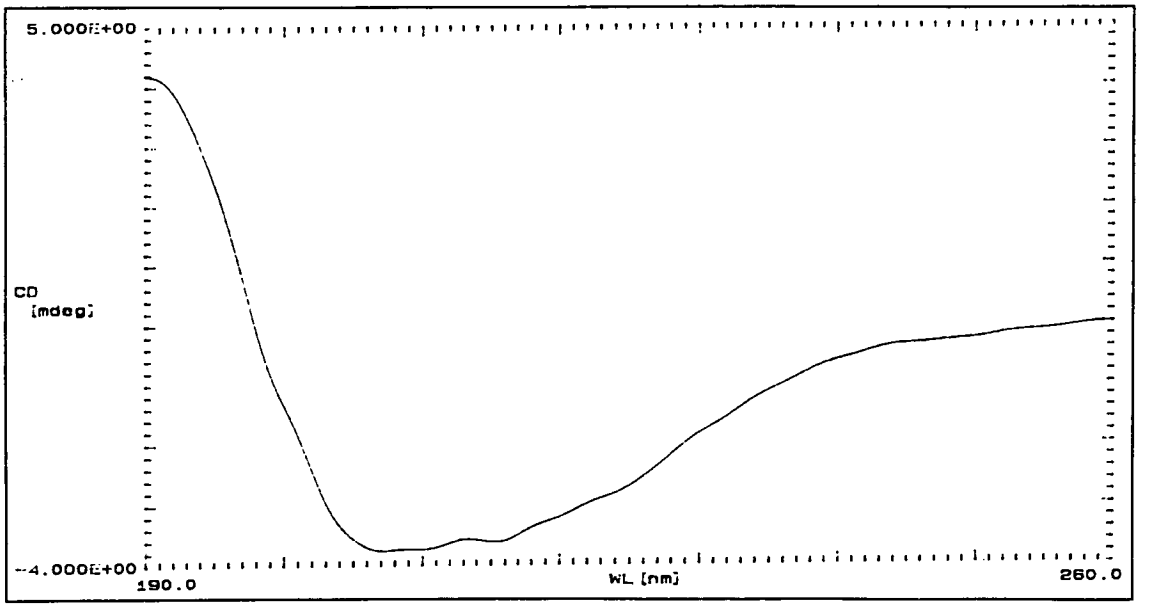


Figure 57 Far UV circular dichroism spectrum of dhIFN-γ (129)

The CD analysis at pH 7 indicated α-helicity of 44% ± 0.49. This result implied that the protein had folded and some secondary structure was present in the dhIFN-γ **(129)** sample.

3.6: Further purification of dhIFN-γ

Thus FPLC size exclusion had purified the dhIFN-γ from 4 components to 3 components. However, this technique was not of sufficient resolution to afford further separation. It was thought that although the 3 components were similar in size they might possess different charges and different hydrophobicities, therefore techniques such as IEF, ion exchange or hydrophobic interaction chromatography might be more applicable.

3.6.1: Hydrophobic interaction chromatography (HIC) of dhIFN-γ (128)

Hydrophobic interaction chromatography (HIC) is employed to separate substances on the basis of their varying strength of hydrophobic interaction with hydrophobic

groups attached to the uncharged gel matrix. For HIC the sample is applied in a solution of high ionic strength. Bound proteins are then eluted by reducing the hydrophobic interaction. This can be achieved by reducing the concentration of ions in the buffer with a decreasing salt gradient.

An aliquot of dhIFN-γ/6M GdmCl solution (**128**) was slowly diluted with the initial buffer: 0.7M ammonium sulfate in 10mM sodium borate buffer, pH8.0. The mixture was loaded onto a Phenyl Sepharose column and eluted with the initial buffer and then with successively reduced concentrations of ammonium sulphate and finally only 10mM sodium borate buffer. Fractions absorbing at 280nm were combined, concentrated and desalted. The material obtained, (dhIFN-γ (**130**)), was characterised and the composition of the protein found to be the same as dhIFN-γ (**128**) i.e 4 components were present.

3.6.2: Cation exchange chromatography of dhIFN-γ (129)

The purification protocol employed for recombinant dhIFN-γ involved using cation exchange chromatography.[116] Hence it was decided to apply this technique to the purification of the synthetic protein.

Problems were encountered with this technique due to precipitation of the protein in the low salt buffers required to load the protein onto the column. The presence of even 0.1M salt concentations resulted in the protein being washed off the column without binding.

As a general method the dhIFN-γ/6M Gdm.Cl (**129**) was slowly dialysed against native buffers (50mM MES buffer, pH 6.0 or 50mM phosphate buffer, pH 6.0) to fold the protein and desalt the solution to concentrations of ≤0.1M 6M GdmCl. The protein/buffer solutions were loaded onto a cation exchange column (Resource S) which had been equilibrated with the appropriate buffer. The column was then eluted with a linear NaCl-gradient of 0-1M NaCl with monitoring at 280nm. Fractions absorbing at 280nm were concentrated and both the filtrate and retentate were

examined by analytical RP HPLC, SDS PAGE and MALDI TOF MS. None of the cation exchange experiments afforded any protein pure or otherwise. It was inferred that the protein either bound irreversibly to the column or was not retained and washed off. It was noted that a significant amount of material was lost, due to precipitation when the protein was dialysed from the 6MGdm.Cl solution.

3.6.3: Isoelectric focusing (IEF) of dhIFN-γ (129)

Separation of proteins by isoelectric focusing is based on the fact that all proteins have a pH-dependent net charge. When a protein is electrophoresed through an established pH gradient, it will migrate until it reaches the pH where the net charge on the protein is zero; at that point it will stop migrating and is said to be focused at its isoelectric point or pI. The Rotofor® cell consists of a focusing chamber which is divided into 20 compartments each holding one fraction. After the protein has been focused, the solution in each compartment can be collected without mixing using the harvesting apparatus.

An aliquot of dhIFN-γ/6M Gdm.Cl **(129)** was slowly dialysed against 1M urea solutions before being loaded into the focusing chamber of the Rotofor® cell. The run was monitored by observing the voltage increase over time and was judged to be completed after 4 hours when the voltage stabilised. The twenty fractions obtained were examined by HPLC and the protein containing fractions were characterised by SDS-PAGE and MALDI MS. The material obtained, dhIFN-γ **(131)**, was characterised and the composition of the protein was found to be the same as dhIFN-γ **(129)** i.e 3 components were present.

3.6.4: Preparative electrophoresis of dhIFN-γ (128)

dhIFN-γ/6M GdmCl solution **(128)** (4 components) was desalted, lyophilised and dhIFN-γ (0.5mg) loaded onto a preparative polyacrylamide gel. Low molecular weight markers were loaded as a reference lane next to 10 sample elution lanes and the gel electrophoresed. After electrophoresis, the reference lane and the adjacent sample lane were cut out and stained, followed by destaining. The stained and

unstained gels were aligned and the band correspoding to dhIFN-γ (Mwt 16.8 kDa) was identified on the stained gel. The location of the dhIFN-γ band on the unstained gel was then estimated by extrapolating from the stained gel. The dhIFN-γ band was excised and eluted into 1ml elution buffer for 5 hours using the Model 422 Electro-Eluter.

After elution the sample was dialysed (centricon 3000) against 10mM sodium phosphate buffer, pH 7.5 to remove SDS and to concentrate the protein. The eluted protein dhIFN-γ (**132**) was examined by SDS-PAGE and a single band corresponding to a molecular weight of 16.8kDa was observed (Figure 58).

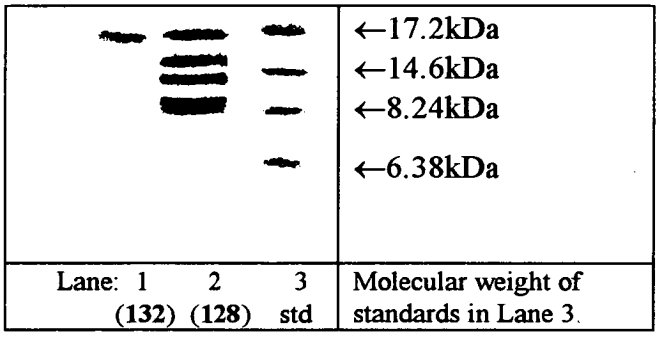


Figure 58 SDS-PAGE of dhIFN-γ (132) and dhIFN-γ (128)

The eluted protein dhIFN-γ (**132**) was submitted for N-terminal sequencing and the first two amino acids (Gln and Asp) were satisfactorily identified. Further rounds of Edman degradation were not obtained because of the low sequencing yield. This suggested that the N-terminus was blocked. The N-terminal sequence has previously been reported with a pyroglutamic residue at the N-terminus.[105] Thus these results are consistent with the presence of a N-terminal pyroglutamate since it is resistant to Edman degradation.

The isolation of purified dhIFN-γ, represented by a single clean band on the SDS-PAGE gel was an excellent result. Preparative electrophoresis is therefore effective for the purification of proteins with impurities of similar mass which cannot be

separated by standard chromatographic techniques. Manipulaton of this technique should result in milligram quantities of purified protein being obtained which could be used for biological testing and structural studies.

3.7: Conclusions and recommendations

The initial studies outlined in this text have revealed that the total synthesis of interferon-*gamma* is accessible by SPS. The protein has been successfully purified to homogeneity by a combination of chromatographic techniques and the application of preparative electrophoresis. The synthesis and purification of dhIFN-γ should be repeated to acquire sufficient pure protein for biological testing, crystallisation and the attainment of structural data.

Chapter 4: Experimental

4.1: Notes

All amino acids were purchased from either Bachem or Novabiochem or were synthesised as described in the text. All amino acids were of the L configuration. Bachem supplied the 4-alkoxybenzylalcohol functionalised polystyrene resin (Wang). Peptide synthesis grade dimethylformamide (DMF), 1,4-dioxan and piperidine were supplied by Rathburn Chemicals. N-N'-Diisopropylethylamine (DIEA) and trifluoroacetic acid (TFA) were peptide synthesis grade and purchased from Applied Biosystems (ABI). 1-Hydroxyl-4-ethoxycarbonyl-1,2,3-triazole (HOCt) and tetrabenzo[a,c,g,i]fluorenyl-17-methoxycarbonyl (Tbfmoc) were synthesised within the research group. Porous graphitised carbon (PGC) was supplied by Shandon Scientific. Ubiquitin (bovine) was purchased from Sigma.

Melting points were recorded in open capillaries using a Buchi 510 oil immersion melting point apparatus. Optical rotations were measured on a AA1000 polarimeter (Optical Activity Ltd.) using a 10.0cm cell in the solvent indicated in the text. Analytical thin layer chromatography (t.l.c.) was performed using aluminium sheets precoated with silica gel $60F_{254}$ (Merck) in the solvent systems indicated in the text. Compounds were visualised using suitable combinations of the following methods: ultraviolet (UV) absorption at 254nm; ninhydrin test for free amino groups; Mary's spray [4,4'-bis(dimethylamino)diphenylcarbinol] for free carboxylic acid functions.

Infrared spectra were recorded on a BIO-RAD SPC 3200 instrument in solution or as a liquid film as indicated. UV spectra were recorded on a Varian Cary 210 double beam spectrophotometer, a Unicam UV/vis double beam spectrophotometer or a Perkin Elmer single beam spectrophotometer in the solvents described in the text. A Corning ion analyser 150 was used for all pH measurements. Sonication was carried out in a Decon FS300b sonic bath. High and low resolution fast atom bombardment

mass spectra (FAB MS) were measured on a Kratos MS50TC instrument, using either thioglycerol, 3-nitrobenzyl alcohol or glycerol as matrix. Matrix assisted laser desorption ionisation mass spectra (MALDI) time of flight mass spectra (TOF MS) were recorded on a PerSeptive Biosystems Voyager™ Biospectrometry™ Workstation using either α-cyano-4-hydroxycinnamic acid or 3,5-dimethoxy-4-hydroxycinnamic acid as matrix. Calculated masses were based on average isotope composition and were derived using the program ADDMASS. Proton NMR spectra were recorded on either a Jeol FX-60 (60 MHz) or a Brucker WP-200 (200 MHz). Carbon-13 NMR spectra were recorded on a AC250 (60MHz) instrument in the solvents indicated in the text. Circular dichroism (CD) spectra were recorded at 20^0C on a JASCO J600 spectropolarimeter (University of Stirling) in the solvents stated.

Elemental analyses were performed on a Perkin-Elmer 2400 CHN elemental analyser. Amino acid analyses (AAA) were perfomed on a LKB 4150 alpha amino acid analyser on the hydrolysate obtained after heating samples in 6M HCl at 110^0C in sealed Carius tubes at the times indicated in the text. Protein sequencing was performed on an ABI 477A sequencer at the Welmet sequencing facility, University of Edinburgh. Polyacrylamide molecular weight gels were run using 20% homogeneous SDS gels (Pharmacia) on a LKB Pharmacia PhastGel electrophoresis system, and developed using coomassie blue staining. High performance liquid chromatography (HPLC) was carried out using either an ABI system comprising 2x1406A solvent delivery systems, a 1480 injector/mixer and a 1783 detector/controller, or a Gilson system comprising 2x306 solvent delivery systems, an 811C dynamic mixer, an 805 manometric module, a 119 UV/vis detector and a Gilson 715 software-driven gradient controller. Components were eluted from various columns, as described in the text, by a linear gradient of acetonitrile (far UV grade, Rathburn Chemicals) in Milli-Q grade water, where both solvents contained 0.1% v/v of HPLC grade TFA (Fisons). Fast protein liquid chromatography (FPLC) was carried out on a Pharmacia FPLC system with Liquid Chromatography Controller LCC-501 Plus using a Superdex™ 75HR 10/30 column and the conditions stated in the text. Gel filtration and cation exchange

chromatography were carried using Pharmacia/LKB apparatus comprising 2xLKB 2138 UVICORDS, a Pharmacia 2132 Microperpex peristaltic pump, an LKB 2112 redirac fraction collector and where appropriate a Pharmacia Gm gradient mixer.

4.2: Solid Phase Peptide Synthesis

All polypeptides 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.

4.2.1: Side chain protecting groups used in Fmoc SPPS

The side chain protecting groups used were as follows: tert-butyl (Bu^t) ethers for serine, threonine and tyrosine; Bu^t esters for aspartic and glutamic acid; τ-triphenylmethyl (Trt) for histidine; 2,2,5,7,8-pentamethylchroman-6-sulphonyl (Pmc) for arginine. The carboxamide side chains of asparagine and glutamine were incorporated as either the 4,4'-dimethoxybenzhydryl (MBH) or triphenylmethyl (Trt) derivatives as indicated in the text. Lysine was incorporated as either the tert-butyloxycarbonyl (Boc), phenylacetyl (Phenac), 4-methoxyphenylacetyl (Mepa), 4-benzyloxyphenylacetyl (Bepa), 4-(4'-methylbenzyloxy)phenylacetyl (MeBepa) or the N^{ε}-1,5-dioxaspiro-5:5-undecane-3-nitro-3-methoxycarbonyl (Tnm) derivatives as indicated in the text.

4.2.2: Coupling of the C-terminal amino acid onto 4-alkoxybenzylalcohol (Wang) resin

The C-terminal Fmoc-amino acid (6 equivalents) was dissolved in DMF (10ml) with N,N'-diisopropylcarbodiimide (DIC) (3 equivalents). The solution was sonicated for 15 minutes at room temperature before adding it to 4-alkoxybenzylalcohol functionalised polystyrene (Wang) resin which had been preswollen in a minimum

amount of DMF with a catalytic amount of 4-(N,N'-dimethylamino)-pyridine (DMAP). The mixture was sonicated for 2 hours at room temperature. The functionalised resin was then removed by filtration and sequentially washed with DMF, DCM and diethyl ether before being dried *in vacuo*.

If a low resin loading was required the procedure was modified as follows: 3 equivalents of the Fmoc-amino acid and 1.5 equivalents of DIC were used. After adding to the Wang resin and the DMAP, the mixture was allowed to stand at room temperature for 1.5 hours.

4.2.3: Determination of resin loading

The loading efficiency was determined by sonicating a sample of the dried Fmoc-amino acid resin (3-5mg) in 20% v/v piperidine/DMF in a 10ml volumetric flask for 20 minutes. The UV absorbance of the supernatant was then measured at 302nm and the loading calculated using the Beer-Lambert Law (ε_{302}=15400 for fulvene-piperidine adduct). The values for resin functionality (mmol/g) and percentage coupling obtained are listed in the text.

4.2.4: Automated SPPS

Synthetic procedures were preprogrammed into the ABI 430A synthesiser prior to the commencement of each synthesis. The first (C-terminal) amino acid was attached to the Wang resin manually (section 4.2.2) and then transferred to the reaction vessel for automated peptide synthesis. A synthetic cycle, resulting in the coupling of a single amino acid, involved recurring capping, deprotection and coupling steps with thorough washing of the resin after each step. Each cycle was then repeated in order to build up the desired sequence of amino acids, in a stepwise manner, from the C-terminus to the N-terminus. The preprogrammed synthetic cycles are summarised below:

1. Capping

Any unreacted amino functions were capped using acetic anhydride (0.5M), DIEA (0.125M) and HOBt (0.2% w/v) in DMF/1,4-dioxan (1:1, 10ml) with vortexing for 10 minutes. The reaction vessel was drained and the resin washed with six portions of DMF:1,4-dioxan (1:1).

2. Deprotection

The N^{α} Fmoc protecting group was cleaved using 20% v/v piperidine in DMF/1,4-dioxan (1:1, 10ml) for 6 minutes. An aliquot of the deprotection filtrate was passed through an UV detector with on line integration in order to quantify the amount of fulvene-piperidine adduct present at 302nm. Thus allowing the percentage coupling of each amino acid to be estimated. The deprotection step was then repeated a further two times, but with vortexing for 1.5 minutes to ensure complete removal of the Fmoc group. Finally the resin was washed with 6 portions of DMF/1,4-dioxan (1:1).

3. Coupling

The next amino acid residues were incorporated in one of two ways; either using the HOBt method or the HOCt method:

The HOBt method used double couple cycles. The first coupling cycle utilised a preformed symmetrical anhydride and the second 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 only once as its symmetrical anhydride. The symmetrical anhydride was formed, in the activator vessel, by reacting the Fmoc-amino acid (1mmol) with DIC (0.5mmol) in DMF/1,4-dioxan (1:1, 8ml) for 10 minutes before being transferred to the reaction vessel containing the resin. The mixture was vortexed for 30 minutes before the resin was drained and washed with four portions of DMF/1,4-dioxan (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)) before being

drained and washed with four portions of DMF/1,4-dioxan (1:1). For certain difficult couplings the coupling time was extended.

The HOCt method utilised the Fmoc-amino acid HOCt active ester in a single couple cycle. The only exception 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 except that a 1mmol scale was used.

4.2.5: Loading Tbfmoc onto the peptide-resin

Using the Fmoc loading test (4.2.3) the resin functionality (mmol/g) of the peptide-resin was calculated. Tbfmoc chloroformate (3 equivalents) was dissolved in DCM and the solution added to the dry resin followed by DIEA (3 equivalents). The mixture was sealed and sonicated in the dark for 3 hours with occasional mixing. The resin bound Tbfmoc-peptide was filtered, washed copiously with DCM then diethyl ether, dried *in vacuo* and a Tbfmoc load test carried out. The resin was swollen in 1,4-dioxan and stored in the freezer until required.

4.2.6: The Tbfmoc loading test

Dry Tbfmoc-peptide-resin (3-5mg) was sonicated in 20% v/v piperidine/1,4-dioxan in a 10ml volumetric flask for 15 minutes. The UV absorbance of the supernatant was measured at 364nm and the values for resin functionality (mmol/g) and percentage loading of Tbfmoc were calculated using the following equations:

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)

4.3: Experimental Details

4.3.1: Synthesis of enzyme cleavable lysine protecting groups

Copper[lysine]$_2$ (76)

Basic copper carbonate (66g, 0.3mol) was added to a boiling solution of lysine monohydrochloride (36.8g, 0.2mol) in water (500ml). The suspension was refluxed for 3 hours, hot filtered to remove excess insoluble copper carbonate and washed with cold water (30ml). The filtrate was evaporated *in vacuo* to yield the *title compound* as a blue solid (50g, 83%).

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

4-Methoxyphenylacetic acid (8g, 48mmol) was stirred in sodium dried benzene (45ml), at 0^{0}C, under an atmosphere of dry nitrogen. Freshly distilled oxalyl chloride (6.1g, 48mmol) was added followed by DMF (0.1ml). The reaction flask was fitted with a drying tube and the mixture stirred for 1 hour, until all the gas had evolved. The benzene was removed *in vacuo* to yield the acid chloride as a yellow oil which was not isolated. υ_{max} (C_6H_6)/cm^{-1} 3090 (ArCH), 1800 (COCl), 1611 & 1512 (aromatic C=C). The acid chloride (8.8g, 48mmol) was dissolved in 1,4-dioxan (10ml) and added dropwise over a period of 1 hour to a solution of copper[lysine]$_2$ **(76)** (7.2g, 20.4mmol) in water (50ml). Throughout the experiment the pH was kept constant at 10.0 with 10% sodium carbonate. The product precipitated from solution as the reaction progressed, after 1 hour the blue precipitate was removed by filtration, washed copiously with water and dried *in vacuo* to give the *title compound* as a blue solid (7.9g, 51%).

N$^{\varepsilon}$-4-Methoxyphenylacetyl lysine (67)

Dry, finely powdered, copper[N$^{\varepsilon}$-4-methoxyphenylacetyl lysine]$_2$ (5.7g, 8.8mmol) was suspended in a stirred solution of ethylenediaminetetraacetic acid disodium sal

precipitate of N^{ε}-4-methoxyphenylacetyl lysine was filtered from the blue, copper containing, solution. The precipitate was washed copiously with water and dried *in vacuo* to give the *title compound* as a white solid (3.4g, 67%), m.p. 241-242 °C; Found: C, 61.20; H, 7.53; N, 9.52; $C_{15}H_{23}N_2O_4$ requires C, 61.16; H, 7.63; N, 9.39%; m/z (Found: MH^+, 295.16562 ± 0.5 ppm; $C_{15}H_{24}N_2O_4$ requires 295.16578); υ_{max} ($CHBr_3$)/(cm^{-1}) 3271 (NH), 1668 (COOH), 1635, (CO amide I), 1540 (CO amide II), 1513 (aromatic C=C); λ_{max} (MeOH)/nm 283 (ε/$dm^3mol^{-1}cm^{-1}$ 2528), 276 (2940); δH (200MHz; D_2O) 1.11-1.69 (8H, m, CH_2 x 4, Lys), 2.85-2.9 (3H, m, CH_2Ar & αCH), 3.58 (3H, s, OCH_3), 6.69-7.09 (4H, m, aromatic); δC (200MHz; D_2O) 22.04, 31.45, 34.23, 40.06 (CH_2 x 4, Lys), 55.33 (OCH_3), 55.55 (αCH), 113.86 & 130.01 (aromatic CH), 129.66 (quat. aromatic), 156.91 (quat. aromatic C-OMe), 168.23 (amide CO); m/z (FAB) 295 (MH^+, 100%).

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

N^{ε}-4-Methoxyphenylacetyl lysine (3.1g, 10.5mmol) was stirred in water (50ml) and triethylamine (1.1g, 10.5mmol) in 1,4-dioxan (50ml) was added. To this solution was added 9-fluorenylmethylsuccinimidyl carbonate (3.4g, 10.1mmol) and the mixture stirred overnight. The reaction mixture was diluted with water (100ml) acidified to pH 2.0 using 2M HCl. The resulting precipitate was extracted into ethyl acetate, washed with water then brine, dried ($MgSO_4$) and the solvent removed *in vacuo* to give a white solid. Recrystallisation from ethyl acetate/diethyl ether gave the *title compound* as a white powder (3.7g, 68%); m.p. 144°C; t.l.c. (DCM/MeOH, 80:20) R_f 0.33; Found C, 69.56; H, 6.45; N, 5.38; $C_{30}H_{32}N_2O_6$ requires C, 69.75; H, 6.24; N, 5.42%; m/z (Found: MH^+ 517.23388 ± 1 ppm; $C_{30}H_{33}N_2O_6$ requires 517.23384); υ_{max} ($CHBr_3$)/(cm^{-1}) 3411, 3337 (NH), 1745 (urethane CO), 1692, (CO amide I), 1552 (CO amide II), 1600 & 1512 (aromatic C=C); λ_{max} (MeOH)/nm 302 (ε/$dm^3mol^{-1}cm^{-1}$ 6186), 290 (5928), 267 (20103); δH (200 MHz, d_6-DMSO) 1.37 (4H, m, CH_2 x 2, Lys), 1.62-1.70 (2H, m, $C\underline{H}_2CH$, Lys), 3.02-3.04 (2H, m, $C\underline{H}_2$-NH, Lys), 3.31 (2H, s,

COC$\underline{H}_2$-Ar), 3.70 (3H, s, OCH_3), 3.90-3.93 (1H, m, αCH, Lys), 4.23-4.31 (3H, m,OC$\underline{H}_2$ & CH, fluorenyl), 6.83-7.99 (14H, m, aromatics & NH x 2), 12.6 (1H, br., OH); δC (200 MHz, d_6-DMSO) 23.24, 28.80, 30.56, 38.59 (CH_2 x 4, Lys), 41.69 ($\underline{C}H_2CO$) 46.80 (αCH, Lys), 53.91 (CH, fluorenyl), 55.09 (OCH_3), 65.74 (CH_2, fluorenyl), 113.73, 120.24, 125.42, 127.20, 127.78, 130.02 (aromatic CH), 128.59 (quat. aromatic, $\underline{C}$-CH_2), 140.85, 143.94 (quat. fluorenyl, C's), 156.33, 157.96 ($\underline{C}O_2OH$ & quat. aromatic, C-OMe) 170.43 (amide CO), 174.14 (urethane CO); m/z (FAB) 517 (MH^+, 100%); $[\alpha]_D$=+7°, c=0.1 in MeOH.

Methyl 4-benzyloxyphenylacetate (74)

To a solution of methyl 4-hydroxyphenylacetate (**73**) (6.6g, 40mmol) and benzyl bromide (10.3g, 60mmol) in DMF (100ml) was added potassium carbonate (11.1g, 80mmol) and the resulting mixture was heated at 60°C for 7 hours and at room temperature overnight. The reaction mixture was filtered and the filtrate concentrated *in vacuo* to give a yellow oil. The residue was dissolved in diethyl ether, washed with a 1M HCl solution, then brine, dried ($MgSO_4$) and the solvent removed *in vacuo*. The orange residue was recrystallised from n-hexane to yield the *title compound* as a white powder (7.8g, 76%); m.p. 58-59°C (lit[135] 58.5-59°C); t.l.c. (DCM) R_f 0.5; Found C, 74.67; H, 6.58; calculated for $C_{16}H_{16}O_3$ C, 74.95; H, 6.29%; m/z (Found: MH^+ 257.11776; $C_{16}H_{17}O_3$ requires 257.11777); υ_{max} ($CHBr_3$)/cm^{-1} 3100 (ArCH), 1730 (CO_2Me), 1612 & 1511 (aromatic C=C); λ_{max} (MeOH)/nm 282 (ε/$dm^3mol^{-1}cm^{-1}$ 1997), 276 (2355); δH (200 MHz, $CDCl_3$) 3.57 (2H, s, CH_2CO), 3.68 (3H, s, OCH_3), 5.05 (2H, s, OCH_2), 6.92-7.44 (9H, m, aromatics); δC (200 MHz, $CDCl_3$) 40.17 ($\underline{C}H_2CO$), 51.88 (OCH_3), 69.88 (CH_2O), 114.81, 127.32, 127.85, 128.45, 130.16 (aromatic CH), 126.19 & 136.85 (quat. aromatic, $\underline{C}$-CH_2), 157.80 (quat. aromatic, C-O) 172.21 (CO_2Me); m/z (FAB) 256 (M^+, 100%).

4-Benzyloxyphenylacetic acid (70)

To a solution of methyl 4-benzyloxyphenylacetate (**74**) (7.0g, 27.3mmol) in acetone:water (50ml, 4:1) was added lithium hydroxide (1.7g, 41mmol) in water (30ml) and the mixture stirred for 2 hours. The white precipitate which formed was filtered off, washed with ethyl acetate suspended in water (100ml) and acidified to pH 3 using 2M HCl. The resulting precipitate was extracted into ethyl acetate, washed with water and brine, dried ($MgSO_4$) and the solvent removed *in vacuo* to yield a white solid. The *title compound* was recrystallised from ethyl acetate (5.2g, 79%), m.p. 124°C (lit[135] 114°C); t.l.c. (DCM/MeOH, 80:20) R_f 0.60; Found C, 74.48; H, 6.10; calculated for $C_{15}H_{14}O_3$ C, 74.36; H, 5.82%; m/z (Found: M^+ 242.09537; $C_{15}H_{14}O_3$ requires 242.09429); υ_{max} ($CHBr_3$)/cm^{-1} 3600-3300 (br. 0H), 3020 (ArCH), 1712 (CO_2H), 1512 (aromatic C=C); λ_{max} (MeOH)/nm 284 (ε/$dm^3mol^{-1}cm^{-1}$ 151), 276 (303); δH (200 MHz, $CDCl_3$) 3.59 (2H, s, CH_2CO), 5.05 (2H, s, OCH_2), 6.93-7.45 (9H, m, aromatics), 12.5 (1H, br., OH); δC (200 MHz, $CDCl_3$), 40.02 ($\underline{C}H_2CO$), 69.90 (CH_2O), 114.89, 127.34, 127.86, 128.47, 130.33 (aromatic CH), 125.47 & 136.81 (quat. aromatic, $\underline{C}$-CH_2), 157.98 (quat. aromatic, C-O) 177.98 (CO_2H); m/z (FAB) 242 (M^+, 100%).

4-Benzyloxyphenylacetyl chloride (75)

4-Benzyloxyphenylacetic acid (**70**) (4.7g, 19.4mmol) was stirred in sodium dried benzene (50ml) under an atmosphere of dry nitrogen. Freshly distilled oxalyl chloride (2.6g, 20.6mmol) was added followed by DMF (0.1ml). The reaction flask was fitted with a drying tube and the mixture stirred for 1 hour, until all the gas had evolved. The benzene was removed *in vacuo* to yield the acid chloride as a white solid which was not purified further. υ_{max} ($CHBr_3$)/cm^{-1} 3056 (ArCH), 1800 (COCl), 1611 & 1512 (aromatic C=C); δH (60 MHz, $CDCl_3$) 4.22 (2H, s, CH_2COCl), 5.22 (2H, s, OCH_2), 7.65-7.20 (9H, m, aromatics).

Copper[N^{ε}-4-benzyloxyphenylacetyl lysine]$_2$ (77)

Copper[lysine]$_2$ (**76**) (3.4g, 9.6mmol) was dissolved in water (50ml) and the pH adjusted to 10.0 using 10% sodium carbonate. 4-Benzyloxyphenylacetyl chloride (**75**) (5.0g, 19mmol) in 1,4-dioxan (10ml) was added dropwise over 0.5 hour keeping the pH constant at 10.0. The product precipitated from solution as the reaction progressed, after 1 hour the blue precipitate was removed by filtration, washed copiously with water and dried *in vacuo* to give the *title compound* as a blue solid (5.2g, 67%).

N^{ε}-4-Benzyloxyphenylacetyl lysine

Dry, finely powdered, copper[N^{ε}-4-benzyloxyphenylacetyl lysine]$_2$ (**77**) (4.0g, 5.0mmol) was suspended in a stirred solution of ethylenediaminetetraacetic acid disodium salt (6.5g, 17.5mmol) in water (100ml). The mixture was stirred for 2 days; after which time the precipitate was still a blue colour indicating that all the copper had not been removed. The mixture was then sonicated for a further 2 days. Then the white precipitate of N^{ε}-4-benzyloxyphenylacetyl lysine was filtered from the blue, copper containing, solution. The precipitate was washed copiously with water and dried *in vacuo* to give the *title compound* as a very pale blue solid (3.6g, 97%).

N^{α}-9-Fluorenylmethoxycarbonyl-N^{ε}-4-benzyloxyphenylacetyl lysine (71)

N^{ε}-4-Benzyloxyphenylacetyl lysine (3.3g, 8.9mmol) was stirred in water (50ml) and triethylamine (0.95g, 9.1mmol) in 1,4-dioxan (50ml) was added. To this suspension was added 9-fluorenylmethylsuccinimidyl carbonate (2.8g, 8.4mmol) and the mixture stirred overnight. The reaction mixture was filtered and the filtrate diluted with water (100ml), then acidified to pH 2.0 using 2M HCl. The resulting precipitate was extracted into ethyl acetate, washed with water then brine, dried ($MgSO_4$) and the solvent removed *in vacuo* to give a pale yellow solid. Recrystallisation from ethyl acetate/diethyl ether gave the *title compound* as a white powder (3.9g, 74%); m.p. 152°C; t.l.c. (DCM/MeOH, 80:20) R_f 0.38, (DCM/MeOH, 90:10) R_f 0.25; Found: C, 72.69; H, 6.40; N, 4.77; $C_{36}H_{36}N_2O_6$ requires C, 72.97; H, 6.12; N, 4.73%; m/z

(Found: MH^+ 593.26517; $C_{36}H_{37}N_2O_6$ requires 593.26514); υ_{max} ($CHBr_3$)/cm^{-1} 3362, 3325 (NH), 3020 (ArCH), 1736 (urethane CO), 1669 (CO_2H), 1648 (CO amide I), 1552 (CO amide II), 1511 (aromatic C=C); λ_{max} (MeOH)/nm 302 (ε/$dm^3mol^{-1}cm^{-1}$8876), 290 (7692), 266 (29586); δH (200 MHz, d_6-DMSO) 1.36 (4H, m, CH_2 x 2, Lys), 1.63-1.69 (2H, m, C$\underline{H}_2$CH, Lys), 3.01-3.03 (2H, m, C$\underline{H}_2$-NH, Lys), 3.30 (2H, s, COC$\underline{H}_2$-Ar), 3.89-3.92 (1H, m, αCH, Lys), 4.21-4.30 (3H, m, OC$\underline{H}_2$ & CH, fluorenyl), 5.05 (2H, s, OCH_2-Ar), 6.90-8.00 (19H, m, aromatics & NH x 2), 12.6 (1H, br., OH); δC (200 MHz, d_6-DMSO) 23.23, 28.79, 30.53, 30.82 (CH_2 x 4, Lys), 41.66 ($\underline{C}H_2CO$) 46.77 (αCH, Lys), 53.88 (CH, fluorenyl), 65.72 (CH_2, fluorenyl), 69.22 (CH_2O), 114.62, 120.24, 125.41, 127.19, 127.71, 127.87, 128.52, 130.03 (aromatic CH), 128.85 & 137.31 (quat. aromatic, $\underline{C}$-CH_2), 140.83, 143.92 (quat. fluorenyl, C's), 156.30, 157.02 ($\underline{C}O_2OH$ & quat. aromatic, C-OCH_2), 170.34 (amide CO), 174.14 (urethane CO); m/z (FAB) 593 (MH^+, 100%); $[\alpha]_D$=+6°, c=0.1 in MeOH.

4-Methoxybenzyl bromide

A mixture of 4-methoxybenzyl alcohol (10.0g, 72.4mmol) and concentrated hydrobromic acid (33.2g, 405mmol) was vigorously stirred for 15 minutes. The reaction mixture was then extracted with diethyl ether. The combined etheral extracts were washed with saturated sodium hydrogen carbonate then water, dried ($CaCl_2$) and the diethyl ether removed *in vacuo*. The *title compound* was obtained as a clear liquid which was used for the next reaction without further purification (13g, 89%).

Methyl 4-(4'-methoxybenzyloxy)phenylacetate

To a solution of methyl 4-hydroxyphenylacetate (7.2g, 43mmol) and 4-methoxybenzyl bromide (13g, 64.7mmol) in DMF (150ml) was added potassium carbonate (11.9g, 86.2mmol) and the resulting mixture was heated at 60°C for 7 hours and at room temperature overnight. The reaction mixture was filtered and diethyl ether (100ml) added to the filtrate. This organic solution was washed with 1M HCl then brine, dried

($MgSO_4$) and the solvent removed *in vacuo*. The resulting orange residue was recrystallised from n-hexane to yield the *title compound* as a white solid (10.5g, 85%); t.l.c. (DCM) R_f 0.70; υ_{max} ($CHBr_3$)/cm^{-1} 3040 (ArCH), 1730 (CO_2Me), 1610 & 1511 (aromatic C=C); λ_{max} (MeOH)/nm 283 (ε/$dm^3mol^{-1}cm^{-1}$ 1987), 279 (2340); δH (200 MHz, $CDCl_3$) 3.57 (2H, s, CH_2CO), 3.66 (3H, s, CO_2CH_3), 5.00 (2H, s, OCH_2), 6.82-7.42 (8H, m, aromatics); δC (200 MHz, $CDCl_3$) 40.17 ($\underline{C}H_2CO$), 51.92 ($CO_2\underline{C}H_3$), 55.10 (Ar-OCH_3), 69.70 (CH_2O), 113.89, 114.9, 129.09, 130.21 (aromatic CH), 125.84 & 128.85 (quat. aromatic, $\underline{C}$-CH_2), 157.85 & 158.03 (quat. aromatic, C-OMe &. C-O), 172.24 (CO_2Me).

4-(4'-Methoxybenzyloxy)phenylacetic acid (82)

To a suspension of methyl 4-(4'-methoxybenzyloxy)phenylacetate (10.0g, 35mmol) in acetone:water (50ml, 4:1) was added lithium hydroxide (2.2g, 52.5mmol) in water (30ml) and the mixture stirred for 1 hour. The white precipitate which formed was filtered off, washed with ethyl acetate, suspended in water (100ml) and slowly acidified to pH 3 using 2M HCl. The resulting precipitate was extracted into ethyl acetate, washed with water and brine, dried ($MgSO_4$) and the solvent removed *in vacuo* to give a white solid. The *title compound* was recrystallised from ethyl acetate (6.1g, 64%), m.p. 165°C; t.l.c. (DCM/MeOH, 80:20) R_f 0.60; Found C, 70.86; H, 6.15; $C_{16}H_{16}O_4$ requires C, 70.58; H, 5.92%; υ_{max} ($CHBr_3$)/cm^{-1} 3600-3300 (br. 0H), 3020 (ArCH), 1710 (CO_2H), 1611 & 1512 (aromatic C=C); λ_{max} (MeOH)/nm 283 (ε/$dm^3mol^{-1}cm^{-1}$ 1224), 274 (1496); δH (200 MHz, $CDCl_3$) 3.58 (2H, s, CH_2CO), 3.81 (3H, s, Ar-CH_3), 4.96 (2H, s,OCH_2), 6.89-7.36 (8H, m, aromatics), 12.5 (1H, br., OH); δC (200 MHz, $CDCl_3$) 39.92 ($\underline{C}H_2CO$), 55.20 ($ArOCH_3$), 69.71 (CH_2O), 113.90, 114.90, 129.11, 130.30 (aromatic CH), 125.43 & 128.84 (quat. aromatic, $\underline{C}$-CH_2), 158.03 & 159.35(quat. aromatic, C-O & C-OMe) 177.98 (CO_2H); m/z (FAB) 272 (M^+, 60%).

4-(4'-Methoxybenzyloxy)phenylacetyl chloride (84)

4-(4'-Methoxybenzyloxy)phenylacetic acid (**82**) (1g, 3.7mmol) was stirred in sodium dried benzene (50ml) under an atmosphere of dry nitrogen. Freshly distilled oxalyl chloride (0.56g, 4.4mmol) was added followed by DMF (0.1ml). The reaction flask was fitted with a drying tube and the mixture stirred for 1 hour, until all the gas had evolved. The yellow solution was reduced *in vacuo* to approximately 25ml and then used in the next reaction undiluted. υ_{max} (C_6H_6)/cm^{-1} 3050 (ArCH), 1799 (COCl), 1612 & 1511 (aromatic C=C).

Copper[N^{ε}-4-(4'-methoxybenzyloxy)phenylacetyl lysine]$_2$ (85)

Copper[lysine]$_2$ (**76**) (0.67g, 1.9mmol) was dissolved in water (25ml) and the pH adjusted to 10.0 using 10% sodium carbonate. 4-(4'-Methoxybenzyloxy)phenylacetyl chloride (**84**) (1.1g, 3.7mmol), in benzene, was added dropwise over 0.5 hour keeping the pH constant at 10.0. The product precipitated from solution as the reaction progressed, after 1 hour the blue precipitate was removed by filtration, washed copiously with water and dried *in vacuo* to give the *title compound* as a blue solid (0.7g, 47%).

N^{ε}-4-(4'-Methoxybenzyloxy)phenylacetyl lysine

Dry, finely powdered, copper[N^{ε}-4-(4'-methoxybenzyloxy)phenylacetyl lysine]$_2$ (**85**) (0.5g, 0.6mmol) was suspended in a saturated ethylenediaminetetraacetic acid disodium salt (2.2g, 6.0mmol) solution (water 25ml). The reaction mixture was then sonicated for 3 days; after which time the white precipitate of N^{ε}-4-(4'-methoxybenzyloxy)phenylacetyl lysine was filtered from the blue, copper containing, solution, washed copiously with water and dried *in vacuo* to give the *title compound* as a white solid (0.3g, 63%).

N^{α}-9-Fluorenylmethoxycarbonyl-N^{ε}-4-(4'-methoxybenzyloxy)-phenylacetyl lysine (86)

N^{ε}-4-(4'-Methoxybenzyloxy)phenylacetyl lysine (0.3g, 8.9mmol) was stirred in water (20ml) and triethylamine (0.08g, 0.75mmol) in 1,4-dioxan (20ml) was added. To this suspension was added 9-fluorenylmethylsuccinimidyl carbonate (0.24g, 0.71mmol) and the mixture stirred overnight. The reaction mixture was filtered and the filtrate diluted with water (50ml), then slowly acidified to pH 2.0 using 2M HCl. The resulting precipitate was extracted into ethyl acetate, washed with water then brine, dried ($MgSO_4$) and the solvent removed *in vacuo*. The pale yellow oil obtained was triturated with diethyl ether to yield an off white solid. The *title compound* was obtained, after recrystallisation from ethyl acetate, as a white solid (0.2g, 46%); m.p. 108-110°C; t.l.c. (DCM/MeOH, 90:10) R_f 0.37; Found C, 72.76; H, 6.39; N, 4.82; $C_{37}H_{38}N_2O_7$ requires C, 72.48; H, 5.17; N, 4.51%; m/z (Found: MH^+ 623.27583; $C_{37}H_{39}N_2O_7$ requires 623.27573); υ_{max} ($CHBr_3$)/cm^{-1} 3365, 3320 (NH), 3020 (ArCH), 1735 (urethane CO), 1668 (CO_2H), 1648 (CO amide I), 1550 (CO amide II), 1511 (aromatic C=C); λ_{max} (MeOH)/nm 302 (ε/$dm^3mol^{-1}cm^{-1}$7273), 290 (8485), 266 (23939); δH (200 MHz, d_6-DMSO) 1.34-2.75 (6H, m, CH_2 x 3, Lys), 3.01 (2H, m, $C\underline{H}_2$-NH, Lys), 3.29 (2H, s, $COC\underline{H}_2$-Ar), 3.74 (3H, s, OCH_3), 3.89-3.94 (1H, m, αCH, Lys), 4.22-4.26 (3H, m, OCH_2 & CH, fluorenyl), 4.96 (2H, s, OCH_2-Ar), 6.88-7.90 (18H, m, aromatics & NH x 2), 12.6 (1H, br., OH); δC (200 MHz, d_6-DMSO) 23.02, 28.82, 35.44, 38.59 (CH_2 x 4, Lys), 40.60 ($\underline{C}H_2CO$) 46.79 (αCH, Lys), 55.01

(CH, fluorenyl), 55.17 (OCH_3), 65.60 (CH_2, fluorenyl), 68.98 (CH_2O), 113.88, 114.63, 120.26, 129.49, 129.97 (aromatic CH), 125.27 & 128.69 (quat. aromatic, C-CH_2), 140.80, 143.94, 143.99 (quat. fluorenyl, C's & quat. aromatic C-OMe), 157.05, 159.02 (CO_2OH & quat. aromatic, C-OCH_2), 170.35 (amide CO), 173.90 (urethane CO); m/z (FAB) 623 (MH^+, 30%); $[\alpha]_D$=+10°, c=0.1 in MeOH.

4-Methylbenzyl bromide

A mixture of 4-methylbenzyl alcohol (10.0g, 82mmol) and concentrated hydrobromic acid (33.2g, 405mmol) was vigorously stirred for 2 hours. The reaction mixture was then extracted with diethyl ether. The combined etheral extracts were washed with saturated sodium hydrogen carbonate then water, dried ($CaCl_2$) and the diethyl ether removed *in vacuo*. The *title compound* was obtained as a clear liquid which was used for the next reaction without further purification (12g, 79%).

Methyl 4-(4'-methylbenzyloxy)phenylacetate

To a solution of methyl 4-hydroxyphenylacetate (7.2g, 43mmol) and 4-methylbenzyl bromide (12g, 65mmol) in DMF (150ml) was added potassium carbonate (11.8g, 86.1mmol) and the resulting mixture was heated at 60°C overnight. The reaction mixture was then filtered and diethyl ether (100ml) added to the filtrate. This organic solution was washed with 1M HCl then brine, dried ($MgSO_4$) and the solvent removed *in vacuo*. The resulting orange residue was recrystallised from n-hexane to yield the *title compound* as a white solid (9.3g, 80%); t.l.c. (DCM/MeOH, 80:20) R_f 0.80; υ_{max} ($CHBr_3$)/cm^{-1} 3019 (ArCH), 1730 (CO_2Me), 1612 & 1510 (aromatic C=C); δH (200 MHz, $CDCl_3$) 2.36 (3H, s, $ArCH_3$), 3.56 (2H, s, CH_2CO), 3.68 (3H, s, CO_2CH_3), 5.00 (2H, s, OCH_2), 6.91-7.30 (8H, m, aromatics); δC (200 MHz, $CDCl_3$) 21.07 ($ArCH_3$), 40.17 (CH_2CO), 51.88 (CO_2CH_3), 69.79 (CH_2O), 114.80, 127.46, 129.13, 130.14 (aromatic CH), 126.06, 137.60 (quat. aromatic, C-CH_2),

133.78 (quat. aromatic, C-CH_3), 157.85 (quat. aromatic, C-O), 172.24 (CO_2Me); m/z (FAB) 271 (MH^+, 70%).

4-(4'-Methylbenzyloxy)phenylacetic acid (83)

To a suspension of methyl 4-(4'-methylbenzyloxy)phenylacetate (8.8g, 33mmol) in acetone:water (40ml, 4:1) was added lithium hydroxide (2.0g, 49mmol) in water (24ml) and the mixture stirred for 1 hour. The white precipitate which formed was filtered off, washed with ethyl acetate, suspended in water (100ml) and slowly acidified to pH 3 using 2M HCl. The acidic mixture was extracted into ethyl acetate, washed with water and brine, dried ($MgSO_4$) and the solvent removed *in vacuo* to give a white solid. The *title compound* was recrystallised from ethyl acetate (6.0g, 71%), m.p. 155°C; t.l.c. (DCM/MeOH,80:20) R_f 0.60; Found C, 75.21; H, 6.13; $C_{16}H_{16}O_3$ requires C, 74.98; H, 6.29%; υ_{max} ($CHBr_3$)/cm^{-1} 3600-3300 (br. OH), 3020 (ArCH), 1702 (CO_2H), 1600 & 1512 (aromatic C=C); λ_{max} (MeOH)/nm 283 (ε/$dm^3mol^{-1}cm^{-1}$ 1446), 274 (1486); δH (200 MHz, $CDCl_3$) 2.48 (3H, s, CH_3), 3.68 (2H, s, CH_2CO), 5.22 (2H, s, OCH_2), 7.04-7.61 (8H, m, aromatics), 12.5 (1H, br., OH); δC (200 MHz, $CDCl_3$) 20.93 (Ar-CH_3), 40.09 ($\underline{C}H_2CO$), 69.71 (CH_2O), 114.94, 127.56, 129.23, 130.24 (aromatic CH), 126.26, 137.80 (quat. aromatics), 133.88 (quat. aromatic C-CH_3), 158.03, (quat. aromatic, C-O), 177.96 (CO_2H); m/z (FAB) 257 (MH^+, 100%).

4-(4'-Methylbenzyloxy)phenylacetyl chloride (87)

4-(4'-Methylbenzyloxy)phenylacetic acid (**83**) (6.0g, 23mmol) was stirred in sodium dried benzene (50ml) under an atmosphere of dry nitrogen. Freshly distilled oxalyl chloride (2.9g, 23mmol) was added followed by DMF (0.1ml). The reaction flask was fitted with a drying tube and the mixture stirred for 1 hour, until all the gas had evolved. The benzene was removed *in vacuo* to yield the acid chloride as a white solid which was not isolated (5.5g, 87%); υ_{max} ($CHBr_3$)/cm^{-1} 3042 (ArCH), 1798

(COCl), 1611 & 1510 (aromatic C=C); δH (200 MHz, $CDCl_3$) 2.34 (3H, s, CH_3), 4.34 (2H, s, CH_2COCl), 4.98 (2H, s, OCH_2), 6.88-7.27 (9H, m, aromatics).

Copper[N^{ε}-4-(4'-Methylbenzyloxy)phenylacetyl lysine]$_2$

Copper[lysine]$_2$ (**76**) (4.2g, 12mmol) was dissolved in water (75ml) and the pH adjusted to 10.0 using 10% sodium carbonate. 4-(4'-Methylbenzyloxy)phenylacetyl chloride (**87**) (5.5g, 20mmol) in 1,4-dioxan (20ml) was added dropwise over 0.5 hour keeping the pH constant at 10.0. The product precipitated from solution as the reaction progressed, after 1 hour the blue precipitate was removed by filtration, washed copiously with water and dried *in vacuo* to give the *title compound* as a blue solid (6.1g, 76%).

N^{ε}-4- (4'-Methylbenzyloxy)phenylacetyl lysine

Dry, finely powdered, copper[N^{ε}-4-(4'-methylbenzyloxy)phenylacetyl lysine]$_2$ (5.5g, 6.8mmol) was added to a saturated ethylenediaminetetraacetic acid disodium salt (12.5g, 34mmol) solution (water 100ml). The reaction mixture was then sonicated for 7 days; after which time the white precipitate of N^{ε}-4-(4'-methylbenzyloxy)phenylacetyl lysine was filtered from the blue, copper containing, solution, washed copiously with water and dried *in vacuo* to give the *title compound* as a white solid (4.5g, 87%); $[\alpha]_D$=+9°, c=0.1 in MeOH.

N^{α}-9-Fluorenylmethoxycarbonyl-N^{ε}-4-(4'-methylbenzyloxy)-phenylacetyl lysine (88)

N^{ε}-4-(4'-Methylbenzyloxy)phenylacetyl-lysine (4.2g, 10.9mmol) was stirred in water (75ml) and triethylamine (1.1g, 10.9mmol) in 1,4-dioxan (75ml) was added. To this suspension was added 9-fluorenylmethylsuccinimidyl carbonate (3.5g, 10.4mmol) and the mixture stirred overnight. The reaction mixture was filtered and the filtrate diluted with water (100ml), then slowly acidified to pH 2.0 using 2M HCl. The resulting precipitate was extracted into ethyl acetate, washed with water then brine, dried ($MgSO_4$) and the solvent removed *in vacuo* to give a pale yellow oil which was

triturated with diethyl ether to yield an off white solid. Recrystallisation from ethyl acetate gave the *title compound* as a white solid (3.2g, 49%); m.p. 125°C; t.l.c. (DCM/MeOH, 80:20) R_f 0.50; Found: C, 73.04; H, 6.45; N, 4.72; $C_{37}H_{38}N_2O_6$ requires C, 73.25; H, 6.31; N, 4.62%; m/z (Found 607.28356; $C_{37}H_{38}N_2O_6$ requires 60728281; υ_{max} ($CHBr_3$)/cm^{-1} 3412, 3309 (NH), 3018 (ArCH), 1720 (urethane CO), 1695 (CO_2H), 1645 (CO amide I), 1544 (CO amide II), 1510 (aromatic C=C); λ_{max} (MeOH)/nm 302 (ε/$dm^3mol^{-1}cm^{-1}$7273), 290 (8485), 266 (23939); δH (200 MHz, d_6-DMSO) 1.38 (4H, m, CH_2 x 2, Lys), 1.63-1.69 (2H, m, C<u>H</u>$_2$CH, Lys), 2.28 (3H, s, $ArCH_3$), 3.03-3.05 (2H, m, C<u>H</u>$_2$-NH, Lys), 3.31 (2H, s, COC<u>H</u>$_2$-Ar), 3.91-3.98 (1H, m, αCH, Lys), 4.23-4.31 (3H, m, OC<u>H</u>$_2$ & CH, fluorenyl), 5.00 (2H, s, OCH_2-Ar), 6.89-8.01 (18H, m, aromatics & NH x 2), 12.6 (1H, br., OH); δC (200 MHz, d_6-DMSO) 20.90 ($ArCH_3$), 23.27, 28.82, 30.59, 38.59 (CH_2 x 4, Lys), 41.72 (<u>C</u>H_2CO), 46.82 (αCH, Lys), 53.95 (CH, fluorenyl), 65.78 (CH_2, fluorenyl), 69.16 (CH_2O), 114.69, 120.26, 125.45, 127.23, 127.83, 129.09, 130.05 (aromatic CH), 128.80, 134.28, 137.12 (quat. aromatic), 140.87 & 143.95 (quat. fluorenyl, C's), 156.36 & 157.09 (<u>C</u>O_2OH & quat. aromatic, C-OCH_2), 170.45 (amide CO), 174.19 (urethane CO); m/z (FAB) 607 (MH^+, 100%); $[\alpha]_D$=+12°, c=0.1 in MeOH.

4.3.2: Synthesis of peptides incorporating enzyme cleavable lysine protecting groups

H-Phe-Gly-Lys(Mepa)-Ala-Gly-OH (69)

The initial resin was Fmoc-Gly-Wang resin (0.46g, 0.55mmol/g, 0.25mmol). The amino acids were coupled using the HOBt method and lysine was incorporated as the 4-methoxyphenylacetyl (Mepa) derivative. The peptide-resin was cleaved using a solution of TFA (9.5ml), water (0.5ml) for 1 hour and the resin was removed by filtration. The resin was washed with TFA (2ml), DCM (5ml) and the filtrate concentrated *in vacuo* to an oil. The peptide was precipitated from diethyl ether, filtered, dissolved in 20% aqueous acetic acid and lyophilised to yield the *title*

compound as a white fluffy solid (135mg, 86%); HPLC (vydac C_{18}, 250x4.6mm, 5μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) R_t 17 mins., %B 49; m/z (FAB) 627 (MH^+); AAA (12 hours) Gly_2 2.1, Ala_1 1.0, Phe_1 1.1, Lys_1 1.2.

H-Phe-Gly-Lys(Phenac)-Ala-Gly-OH (68)

The initial resin was Fmoc-Gly-Wang resin (0.46g, 0.55mmol/g, 0.25mmol). The amino acids were coupled using the HOBt method and lysine was incorporated as the phenylacetyl (Phenac) derivative. The peptide-resin was cleaved using a solution of TFA (9.5ml), water (0.5ml) for 1 hour and the resin was removed by filtration. The resin was washed with TFA (2ml), DCM (5ml) and the filtrate concentrated *in vacuo* to an oil. The peptide was precipitated from diethyl ether, filtered, dissolved in 20% aqueous acetic acid and lyophilised to yield the *title compound* as a white fluffy solid (130mg, 87%); HPLC (vydac C_{18}, 250x4.6mm, 5μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) R_t 17 mins., %B 49; m/z (FAB) 597 (MH^+); AAA (12 hours) Gly_2 2.2, Ala_1 1.0, Phe_1 1.1, Lys_1 1.1.

H-Phe-Gly-Lys(Hopa)-Ala-Gly-OH (64)

Route A: from H-Phe-Gly-Lys(Bepa)-Ala-Gly-OH (72)

Route B: from H-Phe-Gly-Lys(MeBepa)-Ala-Gly-OH (89)

Route A

The initial resin was Fmoc-Gly-Wang resin (0.42g, 0.59mmol/g, 0.25mmol). The amino acids were coupled using the HOBt method and lysine was incorporated as the 4-benzyloxyphenylacetyl (Bepa) derivative. The peptide-resin was cleaved using the following cleavage mixture for 3 hours: crystalline phenol (0.75g); 1, 2-ethanedithiol (EDT) (0.5ml), thioanisole (0.5ml); water (0.5ml) and TFA (10ml). The resin was washed with TFA (2ml), DCM (5ml) and the filtrate concentrated *in vacuo* to an oil. The peptide was precipitated from diethyl ether, filtered, dissolved in 20% aqueous acetic acid and lyophilised. HPLC (vydac C_{18}, 250x4.6mm, 5μm, A=H_2O, B=CH_3CN,

0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) 3 peaks (a, b & c (25:1:5)): (a) R_t 12 mins., %B 38; m/z (LDTOF) 611.7 (M^+, Phe-Gly-Lys(Hopa)-Ala-Gly-OH, requires 611.4); (b) R_t 18 mins., %B 52, m/z (LDTOF) 703 (M^+, Phe-Gly-Lys(Bepa)-Ala-Gly-OH, requires 702.5); (c) R_t 20 mins., %B 56, m/z (LDTOF) 700 (M^+, Phe-Gly-Lys(Bepa)-Ala-Gly-OH, requires 702.5). The peptides corresponding to peaks (b) & (c) must be isomers.

Route B

The initial resin was Fmoc-Gly-Wang resin (0.30g, 0.83mmol/g, 0.25mmol). The amino acids were coupled using the HOBt method and lysine was incorporated as the 4-(4'-methylbenzyloxy)phenylacetyl (MeBepa) derivative. The peptide-resin was cleaved using the following cleavage mixture for 3 hours: crystalline phenol (0.75g); EDT (0.5ml), thioanisole (0.5ml); water (0.5ml) and TFA (10ml). The resin was washed with TFA (2ml), DCM (5ml) and the filtrate concentrated *in vacuo* to an oil. The peptide was precipitated from diethyl ether, filtered, dissolved in 20% aqueous acetic acid and lyophilised. HPLC (vydac C_{18}, 250x4.6mm, 5μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) 3 peaks (a, d & e (25:1:5)): (a) R_t 12 mins., %B 38, m/z (LDTOF) 612.3 (MH^+, Phe-Gly-Lys(Hopa)-Ala-Gly-OH, requires 612.4); (d) R_t 17 mins., %B 52, m/z (LDTOF) 717.9 (MH^+, H-Phe-Gly-Lys(MeBepa)-Ala-Gly-OH, requires 717.6); (e) R_t 19 mins., %B 55, m/z (LDTOF) 718.9 (MH^+, H-Phe-Gly-Lys(MeBepa)-Ala-Gly-OH, requires 717.6). The peptides corresponding to peaks (d) and (e) must be isomers. Phe-Gly-Lys(Hopa)-Ala-Gly-OH: m/z (Found 613.29859; $C_{30}H_{41}N_6O_8$ requires 613.29856); AAA (12 hours) Gly_2 2.0, Ala_1 1.2, Phe_1 1.0, Lys_1 1.0.

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

Route A: from H-Phe-Gly-Lys(Mepa)-Ala-Gly-OH (**69**)

Route B: from H-Phe-Gly-Lys(Phenac)-Ala-Gly-OH (**68**)

Route C: from H-Phe-Gly-Lys(Hopa)-Ala-Gly-OH (**64**)

As a general method, a 13mg (0.2mmol) portion of each peptide was added to Tris.HCl buffer (50mM, 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 ^{0}C. The progress of each reaction was monitored by HPLC (vydac C_{18}, 250x4.6mm, 5μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-50% over 30 minutes, λ=214nm). The half lives for each deprotection reaction were estimated as follows: Route A: Mepa, 22 hours; Route B: Phenac, 6.5 hours and Route C: Hopa, 4 hours. On completion of the reactions the resin bound enzyme was removed by filtration and the peptides purified by semi-preparative HPLC (ABI C_{18} Aquapore, 250x25mm, 20μm, A=H_2O, B=CH_3CN, 0.1% TFA; 5ml/min., 10-50% over 30 minutes, λ=214nm) lyophilised to yield the *title compound* as a white fluffy solid (5mg, 52%); HPLC (vydac C_{18}, 250x4.6mm, 5μM, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-40% over 30 minutes, λ=214nm) R_t 8 mins., %B 10; m/z (FAB) 479 (MH^+); m/z (Found: MH^+ 479.26184; $C_{22}H_{35}N_6O_6$ requires 479.26179); AAA (12 hours) Gly_2 2.1, Ala_1 1.1, Phe_1 1.0, Lys_1 1.2.

4.3.3: Synthesis of ubiquitin incorporating Lys(Bepa)

Chemical synthesis of [6,11,27,29,33&63-Lys(Bepa)] ubiquitin (95)

The synthesis was performed on a 0.2mmol scale using the Fmoc-Gly functionalised Wang resin (0.49g, 0.41mmol/g). The side chain protecting groups used were as follows: tert-butyl (Bu^t) ethers for serine, threonine and tyrosine; Bu^t esters for aspartic and glutamic acid; τ-triphenylmethyl (Trt) for histidine; 2,2,5,7,8-pentamethylchroman-6-sulphonyl (Pmc) for arginine. The carboxamide side chains of asparagine and glutamine were incorporated as the 4,4'-dimethoxybenzhydryl (MBH) derivative. Lysine(48) was incorporated as the tert-butyloxycarbonyl (Boc) derivative and the other six lysine residues were protected as the 4-benzyloxyphenylacetyl (Bepa) (70) derivative. All amino acids were double coupled using the HOBt method

(0.5mmol DIC/1mmol HOBt) with the exception of glycine which was single coupled and asparagine, glutamine and lysine which were double coupled as the HOBt ester (0.5mmol HOBt/1mmol HOBt). Following completion of the synthesis the N-terminal Fmoc was left on and the resin bound product was capped by sonication for 30 minutes in acetic anhydride (0.5M), DIEA (0.125M) and HOBt (0.2% w/v) in DCM (20ml). The Fmoc-[6,11,27,29,33&63-Lys(Bepa)] ubiquitin-resin (**95**) was separated by filtration, washed with DMF and 1,4-dioxan and divided into four equal portions which were stored in 1,4-dioxan at 4^0C until required.

Tbfmoc loading

A 240mg portion of the Fmoc-[6,11,27,29,33&63-Lys(Bepa)] ubiquitin-resin (**95**) was sonicated in 20% v/v piperidine/DMF for 20 minutes to remove the Fmoc group and subsequently washed with DMF, DCM and diethyl ether and dried *in vacuo* for 30 minutes. Tbfmoc chloroformate (**38**) (12.4mg, 0.027mmol) was dissolved in DCM (10ml) and the solution added to the dry resin (240mg, 0.009mmol) followed by DIEA (4.7µl, 0.027mmol). The mixture was sealed and sonicated in the dark for 3 hours with occasional mixing. The Tbfmoc-[6,11,27,29,33&63-Lys(Bepa)] ubiquitin-resin (**96**) was filtered, washed copiously with DCM then diethyl ether and dried *in vacuo*. AAA (40hours) Asx_7 7.74, Thr_7 6.13, Ser_3 2.39, Glx_{12} 13.12, Pro_3 3.02, Gly_6 5.77, Ala_2 2.23, Val_4 3.54, Met_1 0.57, Ile_7 6.84, Leu_9 8.20, Tyr_1 0.71, Phe_2 2.0, His_1 0.86, Lys_7 6.23, Arg_4 3.17.

Cleavage of Tbfmoc-[6,11,27,29,33&63-Lys(Bepa)] ubiquitin-resin (96)

The next step was cleavage of the peptide from the resin and the simultaneous deprotection of all side chain protecting groups, with the exception of the six Lys(Bepa) residues which were only partially cleaved to give 4-hydroxyphenylacetyl lysine, Lys(Hopa). This was accomplished by firstly swelling the Tbfmoc-peptide-resin (**96**) in phenol (0.85g), EDT (0.5ml) and thioanisole (0.5ml) for 30 minutes under an atmosphere of dry nitrogen whilst protected from the light. Then water

(0.5ml) and TFA (10ml) were added and the cleavage mixture stirred for a further 4.5 hours. The resin was separated by filtration, washed with TFA (2ml) and DCM (5ml) and the filtrate concentrated *in vacuo* to a yellow oil. The crude Tbfmoc-peptide was precipitated from the oil by the addition of 2% v/v β-mercaptoethanol in diethyl ether (100ml) at 0^0C. The solid Tbfmoc-peptide was filtered, washed with diethyl ether, dissolved in 20% aqueous acetic acid and lyophilised to yield a white fluffy solid of Tbfmoc-[6,11,27,29,33&63-Lys(Hopa)] ubiquitin (Tbfmoc-Ub(Hopa)$_6$) (**97**) (125mg). AAA (40hours) Asx_7 7.64, Thr_7 6.34, Ser_3 2.22, Glx_{12} 13.36, Pro_3 2.78, Gly_6 5.69, Ala_2 2.23, Val_4 3.68, Met_1 0.55, Ile_7 6.83, Leu_9 8.48, Tyr_1 0.72, Phe_2 2.24, His_1 0.75, Lys_7 6.36, Arg_4 3.47.

Tbfmoc/PGC and gel filtration of Tbfmoc-[6,11,27,29,33&63-Lys(Hopa)] ubiquitin (97)

The crude Tbfmoc-[6,11,27,29,33&63-Lys(Hopa)] ubiquitin (**97**) (120mg) was solubilised in 6M GdmCl (10ml) and diluted with isopropanol (10ml) and centrifuged to remove a small amount of insoluble material. PGC (HPLC grade, 7μm, 600mg) was added and mixture vortexed for 10 minutes and centrifuged (15 minutes, 4000rpm) to a pellet. The supernatant was decanted and the adsorption of the Tbfmoc-Ub(Hopa)$_6$ (**97**) monitored by analytical HPLC (vydac C_8, 250x4.6mm, 5μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=364nm). No absorbance was visible at 364nm; thus it was assumed that all the Tbfmoc material had been adsorbed onto the PGC. The PGC was then washed to remove truncated peptides by vortexing with 6M GdmCl/isopropanol (1:1, 20ml) followed by centrifuging the PGC to a pellet. The wash protocol was repeated until no absorbance at 214nm was visible by HPLC. The protein was cleaved from the Tbfmoc by vortexing the PGC with 10% piperidine in 6M GdmCl/isopropanol (1:1, 20ml x 2) for 10 minutes after which the mixture was centrifuged to a pellet. The combined deprotection supernatants were neutralised to pH 4.5 with acetic acid and the isopropanol removed *in vacuo* prior to being loaded onto a Sephadex G50 (fine) column (660x30mm) which had been equilibrated with 30% aqueous acetic acid. The column was eluted at 30ml/h with monitoring at 226 and 277nm and collecting 20

minute fractions. The protein containing fractions were pooled and lyophilised. Pure 4-hydroxyphenylacetyl lysine protected ubiquitin Ub(Hopa)$_6$ (**98**) was obtained by semi-preparative HPLC (ABI aquapore RP300 C_8, 250x9.2mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 5ml/min., 10-60% over 30 minutes, λ=214nm) and lyophilised to yield a white fluffy solid of Ub(Hopa) (40mg); HPLC (ABI aquapore RP300 C_8, 220x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) R_t 16.4mins., 49%B; m/z (LDTOF) 9393.7 (Na adduct), $C_{426}H_{666}N_{105}O_{132}S_1$ requires 9370.7; AAA (40hours) Asx_7 7.54, Thr_7 6.55, Ser_3 2.60, Glx_{12} 12.87, Pro_3 2.56, Gly_6 5.77, Ala_2 219, Val_4 3.65, Met_1 0.68, Ile_7 6.59, Leu_9 8.36, Tyr_1 0.64, Phe_2 2.25, His_1 0.77, Lys_7 6.44, Arg_4 3.83. Sequencing analysis from N-terminus repetitive yield = 93% (Table 5).

Residue No.	1	2	3	4	5	6	7	8	9	10	11	12	13
Amino acid	Met	Gln	Ile	Phe	Val	Lys	Thr	Leu	Thr	Gly	Lys	Thr	Ile
Pmol	147	118	143	133	133	151	61	114	50	67	74	47	65

Table 5 N-Terminal sequencing results for Ub(Hopa)$_6$ (98)

Chemical synthesis of [63-Lys(Hopa)]ubiquitin-(36-76) (100)

The [63-Lys(Hopa)]ubiquitin-(36-76) peptide was obtained by the removal of approximately one third of the resin after the coupling of Ile 36 from the synthesis of [6,11,27,29,33&63-Lys(Bepa)] ubiquitin. The [63-Lys(Hopa)]ubiquitin-(36-76) (**100**) was purified by the same procedure as for [6,11,27,29,33&63-Lys(Hopa)] ubiquitin. HPLC (vydac C_{18}, 250x4.6mm, 5μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) R_t 20.8mins., 62%B; m/z (LDTOF) 4812.0, $C_{426}H_{666}N_{105}O_{132}S_1$ requires 4812.4; AAA (40hours) Asp_4 3.83, Thr_2 1.78, Ser_2 1.83, Glu_6 5.89, Pro_2 2.34, Gly_4 4.12, Ala_1 0.95, Val_1 0.89, Ile_3 2.37, Leu_7 7.32, Tyr_1 0.60, Phe_1 0.86, His_1 0.88, Lys_2 2.11, Arg_4 3.66.

Enzyme deprotection studies on 4-hydroxyphenylacetyl (Hopa) protected lysine peptides

Three Hopa protected lysine peptides: [6,11,27,29,33&63-Lys(Hopa)] ubiquitin (**98**), [63-Lys(Hopa)]ubiquitin-(36-76) (**100**) and H-Phe-Gly-Lys(Hopa)-Ala-Gly-OH (**64**) were individually incubated with penicillin acylase under the different conditions stated in table 4 and the reactions monitored by analytical HPLC. Three types of enzyme were employed in the deprotection experiments: (A) resin bound penicillin acylase (supplied by SmithKline Beecham Pharmaceuticals Ltd.), (B) resin bound penicillin acylase (supplied by Sigma) and (C) soluble penicillin acylase (supplied by Sigma). As a general method, 1mg portions of each peptide (**64**), (**98**) & (**100**) were dissolved in either 2ml of Tris.HCl buffer (50mM, pH 8.1) or ammonium acetate buffer (0.1M, pH 8.5). Penicillin acylase was added along with various reagents and the mixtures incubated at 37 ^{0}C. The progress of each reaction was monitored by HPLC (ABI aquapore RP300 C_8, 220x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) for [6,11,27,29,33&63-Lys(Hopa)] ubiquitin (**98**) or (vydac C_{18}, 250x4.6mm, 5μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-30% over 30 minutes, λ=214nm) for [63-Lys(Hopa)]ubiquitin-(36-76) (**100**) and the pentapeptide (**64**). Table 6 shows the enzyme deprotection experiments carried out on H-Phe-Gly-Lys(Hopa)-Ala-Gly-OH (**64**). Experiments 1-11 (Table 6) were repeated for [6,11,27,29,33&63-Lys(Hopa)] ubiquitin (**98**) and [63-Lys(Hopa)]ubiquitin-(36-76) (**100**) however no deprotection was observed for either protein in any of the reaction conditions listed.

substrate/ expt. no.	penicillin acylase	buffer	temp	variables	deprotection after 24hrs *
(64) expt. 1	A, B, C	Tris.HCl	37°C		7
(64) expt. 2	A, B, C	Tris.HCl	37°C	50mM Lys.HCl	1 *
(64) expt. 3	A	Tris.HCl	37°C	4M Gdm.HCl	no deprotection
(64) expt. 4	A	Tris.HCl	37°C	4M urea	no deprotection
(64) expt. 5	A, B, C	Tris.HCl	37°C	10%MeOH	6
(64) expt. 6	A, B, C	NH_4OAc	37°C	50mM Lys.HCl	2
(64) expt. 7	A, B, C	Tris.HCl	50°C	50mM Lys.HCl	3
(64) expt. 8	A, B, C	Tris.HCl	37°C	50mM Lys.HCl 0.1M LiCl	5
(64) expt. 9	A, B, C	Tris.HCl	37°C	50mM Lys.HCl 0.1M NaCl	4
(64) expt. 10	A	Tris.HCl	37°C	50mM Lys.HCl 50-80% DMSO	no deprotection
(64) expt. 11	A, B, C	Tris.HCl	37°C	50mM Lys.HCl 50% glycerol	8

* **(64)** in expt. 2 was deprotected fastest and was given a relative deprotection rate of 1 the other experiments were given rates with respect to expt. 1. **(64)** in expt. 11 was deprotected slowest.

***Table 6 Summary of enzyme deprotection experiments on H-Phe-Gly-Lys(Hopa)-Ala-Gly-OH* (64)**

4.3.4: Synthesis of chemically cleaved lysine protecting group

3-Hydroxymethyl-3-nitro-1,5-dioxaspiro-5:5-undecane[136] (107)

Cyclohexanone (40ml, 0.39mol) and tris(hydroxymethyl)nitromethane (58.9g, 0.39mol), plus a catalytic amount of toluene-4-sulphonic acid were refluxed in sodium dried toluene using a Dean Stark apparatus. After 24 hours, the solution was cooled to room temperature, washed with water (2x100ml), and the organic layer dried over $MgSO_4$. Concentration *in vacuo* gave a cream coloured solid which after

recrystallisation from hexane/ether gave the *title compound* as a white solid (37g, 41%); m.p. 99°C (lit.,[136] 97-98); t.l.c. (ethyl acetate/hexane, 30:70) R_f 0.57; Found: C, 52.19; H, 7.54; N, 5.65; calculated for $C_{10}H_{17}NO_5$ C, 51.96; H, 7.36; N, 6.06%; υ_{max} (nujol)/cm^{-1} 3415 (OH), 1545 (NO_2); δH (200 MHz, $CDCl_3$) 1.40-1.79 (10H, m, aliphatic ring CH_2), 2.09 (1H, t, J 6.38Hz, OH), 4.07 (2H, d, J 12.57Hz, axial dioxan ring CH_2) 4.09 (2H, d, J 6.33Hz, $C\underline{H}_2OH$) 4.36 (2H, d, J 12.56Hz, equatorial dioxan ring CH_2), δC (200 MHz, $CDCl_3$) 21.93-32.44 (CH_2 x 5, aliphatic ring), 60.19 (2xdioxan ring CH_2), 63.10 (CH_2OH), 86.28 (quat. C), 99.19 (ring junction quat. C); m/z (FAB) 232 (MH^+).

The pentafluorophenyl carbonate of 3-hydroxymethyl-3-nitro-1,5-dioxaspiro-5:5-undecane (108)

3-Hydroxymethyl-3-nitro-1,5-dioxaspiro-5:5-undecane (**107**) (1g, 5mmol), triphosgene (0.5g, 1.66mmol) and N-N'-diisopropylethylamine (DIEA) (0.9ml, 5mmol) were stirred in sodium dried toluene under an inert atmosphere of dry nitrogen. After 15 minutes a white precipitate had formed (DIEA.HCl), the first step of the reaction, formation of the chloroformate, was judged to have been completed (confirmed by IR: COCl 1779cm^{-1} in toluene). The chloroformate was reacted immediately *in situ* by the addition of a mixture of pentafluorophenol (1.01g, 5.5mmol) and DIEA (0.9ml, 5mmol), dissolved in the minimum amount of toluene. The reaction was then stirred for a further 15 minutes, (comparative IR indicated peak at 1779cm^{-1} replaced by COPFP 1787cm^{-1}). The reaction mixture was then rapidly washed twice with iced water, followed by brine solution, dried ($MgSO_4$), filtered and the solvent removed *in vacuo*. The *title compound* was obtained, after recrystallisation from hexane/ethyl acetate as a white solid (2.04g, 93%); m.p. 121-123°C; t.l.c. (ethyl acetate/hexane, 30:70) R_f 0.65; Found: C, 46.26; H, 3.62; N, 3.17; $C_{17}H_{16}NO_7$ requires C, 46.56; H, 3.72; N, 3.33%; m/z (Found: MH^+ 442.09070; $C_{17}H_{16}NO_7$ requires 442.09252); υ_{max} (nujol)/cm^{-1} 1787 (CO), 1529 (NO_2); δH (250 MHz, $CDCl_3$) 1.42-1.85 (10H, m, aliphatic ring CH_2), 4.12 (2H, d, J 12.66Hz, axial ring pair 2xCH), 4.45 (2H, d, J 12.63Hz, equatorial ring pair 2xCH), 4.85 (2H, s,

CH_2O-); δC (50 MHz, $CDCl_3$) 22.24-33.96 (CH_2x5, aliphatic ring), 60.42 (2xCH_2, dioxan ring), 68.24 (CH_2O-), 83.15 (quat. aliphatic C), 100.01 (ring junction quat. C), 150.45 (CO); m/z (FAB) 442 (MH^+).

N^{α}-9-Fluorenylmethoxycarbonyl-N^{ε}-1,5-dioxaspiro-5:5-undecane-3-nitro-3-methoxycarbonyl-lysine (104)

Fmoc-lysine (1.4g, 3.8mmol) was suspended in 1,4 dioxan/water (2:1, 60ml) and triethylamine (1.1ml, 7.6mmol) added. The solution was then cooled to 0°C and the pentafluorophenyl carbonate of 3-hydroxymethyl-3-nitro-1,5-dioxaspiro-5:5-undecane (**108**) (1.7g, 3.85mmol) added. The reaction mixture was allowed to warm to room temperature and stirred overnight. The reaction mixture was then concentrated *in vacuo* and the resulting residue was dissolved in ethyl acetate, washed with brine solution, water, dried ($MgSO_4$) and evaporated *in vacuo* to yield a white foam. This foam was redissolved in the minimum amount of DCM and purified by wet flash chromatography using DCM/MeOH as eluent. Fractions of R_f 0.3 (DCM/MeOH, 90:10) were combined and evaporated *in vacuo* to yield the *title compound* as a white solid (1.9g, 80%); m.p. 172°C; t.l.c. (DCM/MeOH, 90:10) R_f 0.30; Found: C, 61.10; H, 6.63; N, 6.41; $C_{32}H_{39}N_3O_{10}$ requires C, 61.40; H, 3.726.30; N, 6.71%; m/z (Found: M^+H^+ 626.27105; $C_{32}H_{39}N_3O_{10}$ requires 626.27137); υ_{max} (KBr)/cm^{-1} 3408 (NH), 2937 (CH), 1708 (CO), 1552 (NO_2); λmax (MeOH)/nm 265 (ε/$dm^3mol^{-1}cm^{-1}$ 19000), 290 (4800), 300(5700); δH (250MHz, d_6-DMSO) 1.23-1.84 (16H, m, CH_2), 2.91 (2H, s, $NHC\underline{H_2}$) 4.12- 4.23 (4H, m, axial ring pair 2xCH, αCH, Fmoc CH), 4.39-4.33 (6H, m, equatorial ring pair 2xCH, carbamate CH_2, Fmoc CH_2) 7.30-7.90 (8H, m, aromatic CH); (50 MHz, d_6-DMSO) 22.20, 25.03, 28.23, 38.60 (βCH_2 & ringCH_2), 30.82(αCH), 46.89(Fmoc CH), 60.50 (dioxan ring CH_2), 63.24 (ester CH_2), 65.45 (FmocCH) (8H, m, aromatic CH); 68.24 (CH_2O-), 83.15 (quat. aliphatic C), 100.01 (ring junction quat. C), 150.45 (CO); m/z (FAB) 626 (M^+H^+); $[\alpha]_D$=+4°, c=0.1 in MeOH.

4.3.5: Synthesis of ubiquitin incorporating Lys(Tnm)

Chemical synthesis of [6,11,27,29,33&63-Lys(Tnm)] ubiquitin

The synthesis was performed on a 0.13mmol scale using the Fmoc-Gly functionalised Wang resin (0.49g, 0.264mmol/g). The side chain protecting groups used were as follows: tert-butyl (Bu^t) ethers for serine, threonine and tyrosine; Bu^t esters for aspartic and glutamic acid; τ-triphenylmethyl (Trt) for histidine; 2,2,5,7,8-pentamethylchroman-6-sulphonyl (Pmc) for arginine. The carboxamide side chains of asparagine and glutamine were incorporated as the τ-triphenylmethyl (Trt) derivative. Lysine(48) was incorporated as the tert-butyloxycarbonyl (Boc) derivative and the other six lysine residues were protected as the N^{ε}-1,5-dioxaspiro-5:5-undecane-3-nitro-3-methoxycarbonyl (Tnm) derivative (**104**). The amino acids were single coupled using the HOCt method (1mmol) with the exception of histidine which was double coupled as the HOBt ester (1mmol each). The last 9 residues, Thr(9)-Met(1) and Gln(41)&(40) were double coupled as their HOCt active esters. Following completion of the synthesis the N-terminal Fmoc was left on and the resin bound product was capped by sonication for 30 minutes in acetic anhydride (0.5M), DIEA (0.125M) and HOBt (0.2% w/v) in DCM (20ml). The Fmoc-[6,11,27,29,33&63-Lys(Tnm)]ubiquitin-resin (**109**) was separated by filtration, washed with DMF and 1,4-dioxan and divided into four equal portions which were stored in 1,4-dioxan at 4^0C until required.

Tbfmoc loading

A 640mg portion of the Fmoc-[6,11,27,29,33&63-Lys(Tnm)]ubiquitin-resin (**109**) was sonicated in 20% v/v piperidine/DMF for 20 minutes to remove the Fmoc group and subsequently washed with DMF, DCM and diethyl ether and dried *in vacuo* for 30 minutes. Tbfmoc chloroformate (**38**) (37.1mg, 0.081mmol) was dissolved in DCM (10ml) and the solution added to the dry resin (600mg, 0.027mmol) followed by DIEA (14.1μl, 0.081mmol) . The mixture was sealed and sonicated in the dark for 3 hours with occasional mixing. The Tbfmoc-[6,11,27,29,33&63-Lys(Tnm)]ubiquitin-

resin (110) was filtered, washed copiously with DCM then diethyl ether and dried *in vacuo*. AAA (40hours) Asx_7 8.58, Thr_7 6.41, Ser_3 2.29, Glx_{12} 13.41, Pro_3 3.45, Gly_6 5.65, Ala_2 2.54, Val_4 3.62, Met_1 0.53, Ile_7 7.67, Leu_9 6.75, Tyr_1 0.68, Phe_2 2.45, His_1 0.73, Lys_7 7.35, Arg_4 2.95.

Cleavage of Tbfmoc-[6,11,27,29,33&63-Lys(Tnm)] ubiquitin-resin (110)

The next step was cleavage of the peptide from the resin and the simultaneous deprotection of all side chain protecting groups, with the exception of the six Lys(Tnm) residues which were only partially cleaved. This was accomplished by firstly swelling the Tbfmoc-peptide-resin in phenol (0.75g), EDT (2ml), thioanisole (0.5ml) and triisopropylsilane (TIS) (200µl) for 30 minutes under an atmosphere of dry nitrogen whilst protected from the light. Water (0.5ml) and TFA (10ml) were then added and the cleavage mixture stirred for a further 4 hours. The resin was separated by filtration, washed with TFA (5ml) and DCM (5ml) and the filtrate concentrated *in vacuo* to a yellow oil. The crude Tbfmoc-peptide was precipitated from the oil by the addition of 2% v/v β-mercaptoethanol in diethyl ether (100ml) at 0^0C. The diethyl ether was decanted off and the crude Tbfmoc-peptide dissolved in 20% aqueous acetic acid and lyophilised to yield a white fluffy solid of Tbfmoc-$Ub(Tnm)_6$ **(111)** (350mg). AAA (40hours) Asx_7 7.68, Thr_7 6.52, Ser_3 2.37, Glx_{12} 12.89, Pro_3 2.78, Gly_6 5.45, Ala_2 2.42, Val_4 3.65, Met_1 0.62, Ile_7 7.54, Leu_9 6.98, Tyr_1 0.59, Phe_2 2.37, His_1 0.73, Lys_7 7.42, Arg_4 3.50.

One portion of Tbfmoc-$Ub(Tnm)_6$ **(111)** was purified using the Tbfmoc/charcoal protocol while the second was purified without affinity purification onto charcoal.

Affinity purification on charcoal of Tbfmoc-[6,11,27,29,33&63-Lys(Tnm)] ubiquitin (111)

The crude Tbfmoc-[6,11,27,29,33&63-Lys(Tnm)] ubiquitin **(111)** (150mg) was solubilised in 6M GdmCl (15ml), diluted with isopropanol (15ml) and centrifuged to

remove a small amount of insoluble material. To the supernatant was added charcoal (1.5g) which had been prewashed with 10% piperidine in 6M GdmCl/isopropanol (1:1, 100ml) followed by 6M GdmCl/isopropanol (1:1, 5x100ml). The mixture was vortexed for 10 minutes and centrifuged (15 minutes, 4000rpm) to a pellet. The supernatant was decanted and the adsorption of the Tbfmoc-protein monitored by HPLC (ABI aquapore RP300 C_8, 220x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=364nm). No absorbance was visible at 364nm; thus it was assumed that all the Tbfmoc material had been adsorbed onto the charcoal. The charcoal was then washed to remove truncated peptides by vortexing with 6M GdmCl/isopropanol (1:1, 20ml) followed by centrifuging the charcoal to a pellet. The wash protocol was repeated until no absorbance at 214nm was visible by HPLC.

The protein was cleaved from the Tbfmoc by vortexing the charcoal with 10% piperidine in 6M GdmCl/isopropanol (1:1, 20ml x 3) for 10 minutes after which the mixture was centrifuged to a pellet. The combined deprotection supernatants were neutralised to pH 4.5 with acetic acid and the isopropanol removed *in vacuo* prior to being loaded onto a Sephadex G50 (fine) column (660x30mm) which had been equilibrated with 25% aqueous acetic acid. The column was eluted at 35ml/h with monitoring at 226 and 277nm and collecting 20 minute fractions. The protein containing fractions were pooled and lyophilised to yield fully deprotected Ub sulphoxide and Ub (labelled Ub(**112**)) as a white fluffy solid (80mg). AAA (40hours) Asx_7 7.80, Thr_7 6.31, Ser_3 2.69, Glx_{12} 13.52, Pro_3 3.25, Gly_6 5.34, Ala_2 2.66, Val_4 3.62, Met_1 0.53, Ile_7 7.29, Leu_9 6.82, Tyr_1 0.68, Phe_2 1.95, His_1 0.83, Lys_7 7.15, Arg_4 2.87.

HPLC (ABI aquapore RP300 C_8, 220x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) HPLC analysis showed two main peaks (1) R_t=16.4mins., 50%B and (2) R_t=16.8mins., 51%B; (bovine Ub, R_t=16.8mins., 51%B)

m/z (LDTOF) (1) 8581.8 ([MetS-oxide]Ub), $C_{378}H_{630}N_{105}O_{119}S_1$ requires 8581.8, (2) 8588.47 (Na adduct), $C_{378}H_{630}N_{105}O_{118}S_1$ requires 8565.9;
m/z (Electrospray) (2) 8557.0 (bovine Ub (Sigma) Found, 8558.0, $C_{378}H_{630}N_{105}O_{118}S_1$ requires 8565.9).

MMA reduction and cation exchange/CM-Sepharose CL6B of deprotected Ub(112)

Ub(**112**) (33mg) was dissolved in 10% aqueous acetic acid (10ml) and MMA (850.5µl, 9.6mmol) added giving a 10% w/v solution, which was purged with nitrogen and incubated at 37^{0}C for 24 hours. The reduction reaction was monitored by analytical RP HPLC at 214nm. The protein solution was diluted to 60ml with 50mM NH_4OAc, the pH adjusted to pH 4.5, and then loaded onto a CM-Sepharose CL6B column (280x16mm) which had been equilibrated with 50mM NH_4OAc pH 4.5. The column was then eluted with 50mM NH_4OAc pH 4.5 (90ml) at 30ml/h for 3 hours with monitoring at 226/277nm. A pH gradient of 50mM NH_4OAc at pH 4.5 to 50mM NH_4OAc pH 5.5 (500ml) was then run overnight. After which the column was run isocratically at 50mM NH_4OAc pH 5.5 (120ml) for 4 hours followed by a salt gradient of 50mM NH_4OAc pH 5.5 to 0.3M NH_4OAc at pH 5.5 (500ml) to elute the protein from the column. The protein containing fractions were pooled and lyophilised to give (8mg) of folded Ub. This material was further purified by semi-preparative HPLC (ABI aquapore RP300 C_8, 250x9.2mm, 7µm, A=H_2O, B=CH_3CN, 0.1% TFA; 5ml/min., 10-60% over 30 minutes, λ=214nm) and lyophilised to yield a white fluffy solid of pure folded Ub(**116**) (5mg); HPLC (vydac C_8, 250x4.6mm, 5µm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=364nm); R_t=17.4 mins., 53%B; m/z (LDTOF) 8566.95, $C_{378}H_{630}N_{105}O_{118}S_1$ requires 8565.85; AAA (40hours) Asx_7 7.37, Thr_7 7.03, Ser_3 3.09, Glx_{12} 12.49, Pro_3 3.52, Gly_6 5.93, Ala_2 2.07, Val_4 4.01, Met_1 0.63, Ile_7 6.66, Leu_9 8.57, Phe_2 1.99, His_1 0.91, Lys_7 6.60, Arg_4 3.60. UV circular dichroism: Ub(**116**) 0.5mgml^{-1} in 50mM sodium borate buffer, pH 7, α-helice 11% ± 1.5% and β-sheet of 60% ± 2.8%.

Gel filtration/Sephadex G50 of Tbfmoc-[6,11,27,29,33&63-Lys(Tnm)] ubiquitin (111)

A portion of the crude Tbfmoc-[6,11,27,29,33&63-Lys(Tnm)] ubiquitin (**111**) (200mg) was dissolved in 25% aqueous acetic acid (20ml) prior to being loaded onto a Sephadex G50 (fine) column (660x30mm) which had been equilibrated with 25% aqueous acetic acid. The column was eluted at 35ml/h with monitoring at 226 and 364nm and collecting 20 minute fractions. The Tbfmoc-Ub(Tnm)$_6$ (**113**) containing fractions were pooled and lyophilised (100mg); HPLC (ABI aquapore RP300 C_8, 220x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=364nm); R_t=26 mins., 76.6%B; AAA (44 hours) Asx_7 7.64, Thr_7 5.95, Ser_3 2.40, Glx_{12} 12.59, Pro_3 3.94, Gly_6 5.57, Ala_2 2.43, Val_4 3.74, Ile_7 7.07, Leu_9 8.05, Tyr_1 0.65, Phe_2 1.87, His_1 0.92, Lys_7 6.81, Arg_4 3.14.

Non-Tbfmoc material (**114**) (truncations) which eluted from the column were also pooled and lyophilised; HPLC (ABI aquapore RP300 C_8, 220x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=364nm); R_t=19.4 mins., 57%B; m/z (LDTOF) 8079.85 (acetyl Thr(12)-Gly(76) requires 8068.8), 5083.6 (acetyl Glu(34)-Gly(76) requires 5083.6), 4064.7 (acetyl Leu(43)-Gly(76) requires 4062.6).

Deprotection of Tnm group in Tbfmoc-[6,11,27,29,33&63-Lys(Tnm)] ubiquitin (113)

Tbfmoc-[6,11,27,29,33&63-Lys(Tnm)] ubiquitin (**113**) (5mg) was dissolved in 6M GdmCl/tris base pH 8.8 (5ml) and the solution sonicated at room temperature for 1.5 hours. The deprotection reaction was followed by analytical RP HPLC (ABI aquapore RP300 C_8, 220x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=364nm). The reaction mixture was then loaded onto a Sephadex G50 (fine) column (660x30mm) which had been equilibrated with 25% aqueous acetic acid. The column was eluted at 35ml/h with monitoring at

226 and 364nm and collecting 20 minute fractions. The protein containing fractions were pooled and lyophilised to yield fully deprotected Ub(**115**) as a white fluffy solid (3mg); HPLC (ABI aquapore RP300 C_8, 220x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) R_t=16.8 mins., 51%B; (bovine Ub, R_t=16.8mins., 51%B); m/z (LDTOF) 8601.61 ($C_{378}H_{630}N_{105}O_{118}S_1$ requires 8565.9); AAA (40hours) Asx_7 7.03, Thr_7 6.03, Ser_3 2.72, Glx_{12} 12.82, Pro_3 4.47, Gly_6 5.39, Ala_2 2.61, Val_4 3.44, Met_1 0.34, Ile_7 6.66, Leu_9 7.58, Tyr_1 0.42, Phe_2 1.83, His_1 0.68, Lys_7 6.49, Arg_4 3.27.

Non-Tbfmoc purification of [6,11,27,29,33&63-Lys(Tnm)] ubiquitin

A 225mg portion of the peptide-resin (**109**) was sonicated in 20% v/v piperidine/DMF for 20 minutes and subsequently washed with DMF, DCM and diethyl ether and dried *in vacuo* for 30 minutes. Cleavage of the peptide from the resin was accomplished by firstly swelling the peptide-resin in phenol (0.75g), EDT (2ml), thioanisole (0.5ml) and triisopropylsilane (TIS) (200μl) for 30 minutes under an atmosphere of dry nitrogen. Water (0.5ml) and TFA (10ml) were then added and the cleavage mixture stirred for a further 4 hours. The resin was separated by filtration, washed with TFA (5ml) and DCM (5ml) and the filtrate concentrated *in vacuo* to an oil. The crude peptide was precipitated from 2% v/v β-mercaptoethanol in diethyl ether (100ml). The diethyl ether was decanted off and the crude peptide dissolved in 20% aqueous acetic acid and lyophilised to yield a white fluffy solid of $Ub(Tnm)_6$ (**117**) (110mg).

Gel filtration/Sephadex G50 of [6,11,27,29,33&63-Lys(Tnm)] ubiquitin (117)

A portion of the crude $Ub(Tnm)_6$ (**117**)(100mg) was dissolved in 25% aqueous acetic acid (20ml) prior to being loaded onto a Sephadex G50 (fine) column (660x30mm) which had been equilibrated with 25% aqueous acetic acid. The column was eluted at 35ml/h with monitoring at 226 and 277nm and collecting 20 minute fractions. The protein containing fractions were pooled and lyophilised to yield $Ub(Tnm)_6$ (**118**) as a

white fluffy solid (90mg); HPLC (ABI aquapore RP300 C_8, 220x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) broad peak R_t=24mins., 71%B; m/z (LDTOF) 9491.22, ($C_{408}H_{672}N_{111}O_{154}S_1$ requires 9628.56); AAA (40hours) Asx_7 7.80, Thr_7 6.31, Ser_3 2.69, Glx_{12} 13.52, Pro_3 3.25, Gly_6 5.34, Ala_2 2.66, Val_4 3.62, Met_1 0.53, Ile_7 7.29, Leu_9 6.82, Tyr_1 0.68, Phe_2 1.95, His_1 0.83, Lys_7 7.15, Arg_4 2.87.

FPLC/Superdex™ 75 of [6,11,27,29,33&63-Lys(Tnm)] ubiquitin (117)

$Ub(Tnm)_6$ (**117**) (90mg) was sonicated in 6M GdmCl (6ml), centrifuged (15 mins., 4000rpm) and the supernatant loaded onto a Superdex™ 75 HR 10/30 column which had been equilibrated with 6M GdmCl. The column was eluted at 0.5ml/min with monitoring at 280nm and collecting 0.5ml fractions. Fractions 40-48 and fractions 75-80 absorbed at 280nm and after isolation by HPLC, gave identical mass spectra corresponding to [6,11,27,29,33&63-Lys(Tnm)] ubiquitin. All fractions containing protein were combined and extensively dialysed (MWCO 500, Spectra/Por® CE) against water and lyophilised to give $Ub(Tnm)_6$ (**120**) (80mg); HPLC (ABI aquapore RP300 C_8, 220x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) broad peak R_t=24mins., 71%B; m/z (LDTOF) 9448.44 ($C_{408}H_{672}N_{111}O_{154}S_1$ requires 9628.56); AAA (40hours) Asx_7 7.80, Thr_7 6.31, Ser_3 2.69, Glx_{12} 13.52, Pro_3 3.25, Gly_6 5.34, Ala_2 2.66, Val_4 3.62, Met_1 0.53, Ile_7 7.29, Leu_9 6.82, Tyr_1 0.68, Phe_2 1.95, His_1 0.83, Lys_7 7.15, Arg_4 2.87.

Deprotection of Tnm group in [6,11,27,29,33&63-Lys(Tnm)] ubiquitin (118)

[6,11,27,29,33&63-Lys(Tnm)] ubiquitin (**118**) (13mg) was dissolved in 6M GdmCl/tris base pH 8.8 (10ml) and the solution sonicated at room temperature for 1.5 hours. The deprotection reaction was followed by analytical HPLC (vydac C_8, 250x4.6mm, 5μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm). The reaction mixture was then loaded onto a Sephadex G50

(fine) column (660x30mm) which had been equilibrated with 25% aqueous acetic acid. The column was eluted at 35ml/h with monitoring at 226 and 377nm and collecting 20 minute fractions. The protein containing fractions were pooled and lyophilised to yield fully deprotected Ub(**119**) as a white fluffy solid (9mg); HPLC (vydac C_8, 250x4.6mm, 5μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) R_t=23.5 mins., 68%B; (bovine Ub, R_t=23.5mins., 68%B) m/z (LDTOF) 8582.4 ([MetS-oxide]Ub $C_{378}H_{630}N_{105}O_{119}S_1$ requires 8581.8). AAA (40 hours) Asx_7, Thr_7, Ser_3, Glx_{12}, Pro_3, Gly_6, Ala_2, Val_4, Met_1 , Ile_7, Leu_9, Tyr_1 ,Phe_2, His_1, Lys_7 , Arg_4.

MMA reduction and gel filtration/G50 Sephadex of Ub(119)

Ub(**119**) (9mg) was dissolved in 10% aqueous acetic acid (5ml) and MMA (464μl, 5.25mmol) added giving a 10% w/v solution, which was purged with nitrogen and incubated at 37^{0}C for 72 hours. The reduction reaction was monitored by analytical RP HPLC at 214nm. The protein solution was diluted with water and lyophilised three times. MMA was still present thus the reaction mixture was diluted with water and applied to a Sephadex G50 (fine) column (660x30mm) which had been equilibrated with 25% aqueous acetic acid. The column was eluted at 35ml/h with monitoring at 226 and 277nm and collecting 20 minute fractions. The protein containing fractions were pooled and lyophilised to yield Ub as a white fluffy solid (6mg); HPLC (vydac C_8, 250x4.6mm, 5μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ=214nm) R_t=23.5 mins., 68%B; (bovine Ub, R_t=23.5mins., 68%B) m/z (LDTOF) 8567.65, $C_{378}H_{630}N_{105}O_{118}S_1$ requires 8565.85; AAA (40hours) Asx_7 7.45, Thr_7 7.04, Ser_3 3.07, Glx_{12} 12.54, Pro_3 3.9, Gly_6 5.95, Ala_2 2.05, Val_4 4.01, Met_1 0.36, Ile_7 6.70, Leu_9 8.61, Phe_2 1.94, His_1 0.91, Lys_7 6.62, Arg_4 3.60.

4.3.6: Chemical synthesis and purification of dhIFN-γ

Chemical synthesis of dhIFN-γ

The synthesis was performed on a 0.1mmol scale using the Fmoc-Gln(Trt) Wang resin (0.72g, 0.14mmol/g). The side chain protecting groups used were as follows: tert-butyl (Bu^t) ethers for serine, threonine and tyrosine; Bu^t esters for aspartic and glutamic acid; τ-triphenylmethyl (Trt) for histidine; 2,2,5,7,8-pentamethylchroman-6-sulphonyl (Pmc) for arginine. The carboxamide side chains of asparagine and glutamine were incorporated as the τ-triphenylmethyl (Trt) derivatives. Lysine and tryptophan were protected as the tert-butyloxycarbonyl (Boc) derivatives. The amino acids were single coupled using the HOCt method (1mmol) with the exception of histidine which was coupled as the active HOBt ester. The last 44 residues, Gln(1)-Ile(44) were double coupled as their HOCt active esters. Approximately one third of the resin was removed after Asp(76) and the remaining resin used to continue the synthesis. On completion of assembly the N-terminal Fmoc group was left on and the resin bound product was capped by sonication for 30 minutes in acetic anhydride (0.5M), DIEA (0.125M) and HOBt (0.2% w/v) in DCM (20ml). The Fmoc-dhIFN-γ-resin (**125**) was separated by filtration, washed with DMF and 1,4-dioxan and divided into four equal portions which were stored in 1,4-dioxan at 4^0C until required.

Tbfmoc loading

A 640mg portion of the Fmoc-dhIFN-γ-resin (**125**) was sonicated in 20% v/v piperidine/DMF for 20 minutes to remove the Fmoc group and subsequently washed with DMF, DCM and diethyl ether and dried *in vacuo* for 30 minutes. Tbfmoc chloroformate (**38**) (29mg, 0.063mmol) was dissolved in DCM (10ml) and the solution added to the dry resin (540mg) followed by DIEA (11μl, 0.063mmol). The mixture was sealed and sonicated in the dark for 3 hours with occasional mixing. The Tbfmoc-dhIFN-γ-resin (**126**) was filtered, washed copiously with DCM then diethyl ether and dried *in vacuo*. AAA (40 hours) Asx_{20} 22.43, Thr_5 5.43, Ser_{11} 10.02, Glx_{18}

21.05, Pro_2 3.73, Gly_5 5.30, Ala_8 7.70, Val_8 9.31, Met_4 2.31, Ile_7 7.91, Leu_{10} 11.19, Tyr_4 3.91, Phe_{10} 12.04, His_2 2.77, Lys_{20} 20.43, Arg_8 6.30.

Cleavage of Tbfmoc-dhIFN-γ-resin (126)

To the resin bound Tbfmoc-protein (**126**) (540mg) were added phenol (0.75g), EDT (2ml), thioanisole (0.5ml) and allowed to remain for 30 minutes under an atmosphere of dry nitrogen whilst protected from the light. Water (0.5ml) and TFA (10ml) were then added and the cleavage mixture stirred for a further 4 hours. The resin was separated by filtration, washed with TFA (2ml) and the Tbfmoc-protein was precipitated from the filtrate by the addition of 2% v/v β-mercaptoethanol in diethyl ether (100ml) at 0^0C. The Tbfmoc-protein/diethyl ether mixture was centrifuged (10mins., 4000rpm) to give a pellet of Tbfmoc-protein which was dissolved in 20% acetonitrile in H_2O and lyophilised to yield a white fluffy solid of crude Tbfmoc-dhIFN-γ (**127**) (350mg), HPLC (ABI aquapore RP300 C_4, 100x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=364nm), R_t=19.2mins., 53%B and R_t=20.6 mins., 58%B ; AAA (40 hours) Asx_{20} 22.52, Thr_5 5.42, Ser_{11} 10.62, Glx_{18} 20.02, Pro_2 1.89, Gly_5 5.46, Ala_8 7.54, Val_8 9.51, Met_4 3.38, Ile_7 8.36, Leu_{10} 10.98, Tyr_4 4.64, Phe_{10} 10.95, His_2 2.49, Lys_{20} 21.61, Arg_8 6.55.

Affinity purification on charcoal of Tbfmoc-dhIFN-γ-resin (127)

The crude Tbfmoc-dhIFN-γ (**127**) ((300mg) was solubilised in 6M GdmCl (15ml), diluted with isopropanol (15ml) and centrifuged to remove a small amount of insoluble material. To the supernatant was added charcoal (3g) which had been prewashed with 10% piperidine in 6M GdmCl/isopropanol (1:1, 100ml) followed by 6M GdmCl/isopropanol (1:1, 5x100ml). The mixture was vortexed for 10 minutes and centrifuged (15 minutes, 4000rpm) to a pellet. The supernatant was decanted and the adsorption of the Tbfmoc-protein monitored by HPLC (ABI aquapore RP300 C_4,

100x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=364nm). No absorbance was visible at 364nm; thus it was assumed that all the Tbfmoc material had been adsorbed onto the charcoal. The charcoal was then washed to remove truncated peptides by vortexing with 6M GdmCl/isopropanol (1:1, 20ml) followed by centrifuging the charcoal to a pellet. The wash protocol was repeated until no absorbance at 214nm was visible by HPLC. The protein was cleaved from the Tbfmoc by vortexing the charcoal with 10% piperidine in 6M GdmCl/isopropanol (1:1, 20ml x 3) for 10 minutes after which the mixture was centrifuged to a pellet. The combined deprotection supernatants were neutralised to pH 4.5 with acetic acid and reduced *in vacuo* to remove the isopropanol. The protein solution (50ml) was filtered through celite and a 2μm filter, to remove residual charcoal, before being concentrated to a volume of 10ml (Amicon ultrafiltration, MWCO 5,000) of dhIFN-γ/guanidine solution (**128**). HPLC (ABI aquapore RP300 C_4, 100x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=280nm), R_t=17.2mins., 51%B; m/z (LDTOF) 16,909.1 ($C_{746}H_{1188}N_{211}O_{221}S_4$ requires 16,777.2), 14,377.3, 13,256.3 and 10,693; SDS PAGE 4 bands ≅16.8, 14.3, 13.2 and 10.5kDa; AAA (40 hours) Asx_{20} 22.46, Thr_5 5.60, Ser_{11} 10.30, Glx_{18} 19.74, Pro_2 1.20, Gly_5 5.64, Ala_8 7.80, Val_8 8.83, Met_4 3.49, Ile_7 8.65, Leu_{10} 10.36, Tyr_4 4.80, Phe_{10} 11.36, His_2 2.57, Lys_{20} 21.38, Arg_8 7.41.

Gel filtration/Sephadex G50 of dhIFN-γ (128)

A portion of the dhIFN-γ/6M GdmCl solution (**128**) (5ml) was loaded onto a Sephadex G50 (fine) column (660x30mm) which had been equilibrated with 25% aqueous acetic acid. The column was eluted at 35ml/h with monitoring at 226 and 277nm and collecting 20 minute fractions. The protein containing fractions were pooled and lyophilised to yield dhIFN-γ as a white fluffy solid (60mg); HPLC (ABI aquapore RP300 C_4, 100x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min.,

10-90% over 30 minutes, λ_1=214nm, λ_2=280nm). R_t=17.2mins., 51%B; m/z (LDTOF) 16,975.8 ($C_{746}H_{1188}N_{211}O_{221}S_4$ requires 16,777.2), 14,373.5, 13,292.1 and 10,696.9; SDS PAGE 4 bands ≅16.8, 14.3, 13.2 and 10.5kDa; AAA (40 hours) Asx_{20} 21.59, Thr_5 4.61, Ser_{11} 10.60, Glx_{18} 17.89, Pro_2 1.25, Gly_5 4.64, Ala_8 7.53, Val_8 8.09, Met_4 3.42, Ile_7 8.35, Leu_{10} 9.35, Tyr_4 3.95, Phe_{10} 11.03, His_2 2.12, Lys_{20} 19.24, Arg_8 7.33.

FPLC/Superdex™ 75 of dhIFN-γ (128)

A portion of the dhIFN-γ/6M GdmCl solution (**128**) (5ml) was loaded onto a Superdex™ 75 HR 10/30 column which had been equilibrated with 6M GdmCl. The column was eluted at 0.5ml/min with monitoring at 280nm and collecting 0.5ml fractions. Fractions absorbing at 280nm were examined by analytical RP HPLC (ABI aquapore RP300 C_4, 100x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=280nm). Fractions 8-15 eluted with the correct retention time (R_t=17.2mins., 51%B) however fractions 8-10 gave the sharpest HPLC profiles. Individual fractions were desalted examined by SDS PAGE PhastGel and mass spectra obtained; fraction 9: m/z (LDTOF) 16,884.9 ($C_{746}H_{1188}N_{211}O_{221}S_4$ requires 16,777.2), 14,307.9 and 13,225.4; SDS PAGE 3 bands ≅16.8, 14.3 and 13.2kDa; fraction 11: m/z (LDTOF) 16,909.1 ($C_{746}H_{1188}N_{211}O_{221}S_4$ requires 16,777.2), 14,294.2, 13,256.3 and 10,724.5; SDS gel 4 bands ≅16.8, 14.3, 13.2 and 10.5kDa.

Fractions 8 and 9 were combined to give dhIFN-γ (**129**) for subsequent purification reactions. AAA (40 hours) Asx_{20} 22.56, Thr_5 5.30, Ser_{11} 10.22, Glx_{18} 19.53, Pro_2 0.89, Gly_5 5.67, Ala_8 7.83, Val_8 8.83, Met_4 3.26, Ile_7 8.21, Leu_{10} 10.72, Tyr_4 4.68, Phe_{10} 11.02, His_2 1.78, Lys_{20} 20.96, Arg_8 7.25. UV circular dichroism: 0.5mgml^{-1} in 50mM sodium borate buffer, pH 7, α-helice 44 ± 0.49% and β-sheet of 25% ± .43%. Sequencing analysis of dhIFN-γ (**129**) from N-terminus repetitive yield = 99% (Table 7).

Residue No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Amino acid	Gln	Asp	Pro	Tyr	Val	Lys	Glu	Ala	Glu	Asn	Leu	Lys	Lys	Tyr	Phe
Pmol	310	152	350	288	426	352	171	315	178	251	355	381	519	252	364

Table 7 N-Terminal sequencing results for dhIFN-γ (129)

Hydrophobic Interaction chromatography/ Phenyl Sepharose of dhIFN-γ (128)

A 1ml portion of the dhIFN-γ/6M GdmCl solution (**128**) was slowly diluted to a volume of 6ml with 0.7M ammonium sulfate in 10mM sodium borate buffer, pH8.0. The mixture was loaded onto a Phenyl Sepharose 6 Fast Flow (high sub) column (supplied by Pharmacia Biotech) which had been equilibrated with 0.7M ammonium sulfate in 10mM sodium borate buffer, pH8.0. The column was eluted with 0.7M ammonium sulfate in 10mM sodium borate buffer, pH8.0 (20ml) at 30ml/h with monitoring at 226/277nm. The column was then eluted with successively reduced concentrations of ammonium sulphate: 0.35M(20ml), 0.175M(20ml) and finally only the10mM sodium borate buffer. Fractions absorbing at 280nm were combined, concentrated and desalted (Amicon ultrafiltration, MWCO 5,000) to give dhIFN-γ (**130**). HPLC (ABI aquapore RP300 C_4, 100x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=280nm) R_t=17.2mins., 51%B; m/z (LDTOF) 16,975.8 ($C_{746}H_{1188}N_{211}O_{221}S_4$ requires 16,777.2), 14,373.5, 13,292.1 and 10.696.9; SDS PAGE 4 bands ≅16.8, 14.3, 13.2 and 10.5kDa; AAA (40 hours) Asx_{20} 22.60, Thr_5 4.61, Ser_{11} 10.50, Glx_{18} 18.79, Pro_2 1.25, Gly_5 4.66, Ala_8

6.76, Val_8 8.09, Met_4 3.41, Ile_7 7.12, Leu_{10} 9.35, Tyr_4 3.95, Phe_{10} 11.03, His_2 2.12, Lys_{20} 19.24, Arg_8 7.33.

Cation exchange FPLC/Resource S of dhIFN-γ (129)

As a general method the dhIFN-γ/6M Gdm.Cl (**129**) was slowly dialysed against native buffers (50mM MES buffer, pH 6.0 or 50mM phosphate buffer, pH 6.0) to fold the protein and desalt the solution to concentrations of ≤0.1M 6M GdmCl. The protein/buffer solutions were loaded onto a cation exchange column (Resource S) which had been equilibrated with the appropriate buffer eluting at 4ml/min. The column was then eluted with a linear NaCl-gradient of 0-1M NaCl with monitoring at 280nm. Fractions absorbing at 280nm were concentrated (centricon ultrafiltration, MWCO 1000) and both the filtrate and retentate were examined by analytical RP HPLC, SDS PAGE and MALDI TOF MS. No protein was recovered from any of the cation exchange experiments.

Isoelectric focusing (IEF) of dhIFN-γ (129)

An aliquot of dhIFN-γ/6M Gdm.Cl (**129**) was slowly dialysed against 1M urea solutions to obtain a dhIFN-γ/1M urea solution (50ml). Ampholytes (pH3-10, 0.5ml) were added before the solution was loaded into the focusing chamber of the Rotofor® cell. The system was allowed to come to thermal equilibrium at the cooling temperature of 4^0C before connecting the power supply and running at 15W constant power. The run was monitored by observing the voltage increase over time and was judged to be completed after 4 hours when the voltage stabilised. At this point the run was allowed to continue for 30 minutes before harvesting. The twenty fractions obtained were examined by HPLC and the protein containing fractions (fractions 1 & 2, pH 4.5) were characterised by SDS-PAGE and MALDI MS and found to be indentical. dhIFN-γ (**131**), HPLC (ABI aquapore RP300 C_4, 100x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=280nm) R_t=17.2mins., 51%B; m/z (LDTOF) 16,908.5 ($C_{746}H_{1188}N_{211}O_{221}S_4$ requires 16,777.2), 14,378.3, 13.257 and 10593; SDS PAGE 4 bands ≅16.8, 14.3,

13.2 and 10.5kDa; AAA (40 hours) Asx_{20} 22.52, Thr_5 5.41, Ser_{11} 10.62, Glx_{18} 20.02, Pro_2 1.89, Gly_5 5.45, Ala_8 7.45, Val_8 9.50, Met_4 3.37, Ile_7 8.36, Leu_{10} 10.98, Tyr_4 4.64, Phe_{10} 10.94, His_2 2.49, Lys_{20} 21.61, Arg_8 7.54.

Tryptic digest of dhIFN-γ (129)

dhIFN-γ (**129**) (0.5mg) was dissolved in 3M GdmCl buffered Tris.HCL (50mM, pH 8, 37^{0}C). Trypsin (5% w/v) was added and the mixture incubated at 37^{0}C. After 2 and 6 hours aliquots were removed and the digestion stopped by the addition of 6MHCl (5μl). m/z (LDTOF) 2379.5 (M^+Lys(13)-Lys(34) requires 2380.2), 2160.8 (MH^+ Asp(90)-Arg(107) requires 2158.1), 1822.4 (MH^+ Ala(109)-Lys(125) requires 1822.0), 1675.1 (MH^+ Leu(95)-Lys(108), Na adduct, requires 1674.9), 1642.2 (MH^+ Lys(43)- Lys(55), Na adduct, requires 1642.8), 1496.6 (MH^+ Glu(75)-Lys(86), Na adduct, requires 1496.7); HPLC (ABI aquapore RP300 C_4, 100x4.6mm, 7μm, A=H_2O, B=CH_3CN, 0.1% TFA; 1ml/min., 10-90% over 30 minutes, λ_1=214nm, λ_2=280nm).

Glycine-sodium dodecyl sulphate-polyacrylamide gel electrophoresis (glycine-SDS-PAGE)

Analytical experiments were carried out using BIO-RAD Mini-PROTEAN™ vertical gel electrophoresis apparatus. Proteins were analysed by electrophoresis at 50mA per gel, at 25^{0}C under denaturing conditions (in the presence of sodium dodecyl sulphonate (SDS) and β-mercaptoethanol) using the Laemmli discontinuous buffer system.[137] Separating and stacking gels were made using Protogel (30% w/v acrylamide, 0.85 w/v N,N'-bis-methyleneacrylamide) and were polymerised by the addition of tetramethylethylenediamine (TEMED) and ammonium phosphate (APS). Separating gels (20% acrylamide) were prepared with 1.5M Tris.HCl, pH 8.8, 10% w/v SDS and stacking gels (4% acrylamide) were prepared with 125mM Tris.HCl, pH6.8, 10% w/v SDS. The gel running buffer contained 192mM glycine, 25mM Tris.HCl, pH 8.3 and 0.1% SDS.

Protein samples were prepared by heating the sample with the appropriate amount of 5 x SDS reducing buffer (62.5mM Tris.HCl, pH 6.8, 10% v/v glycerol, 2% v/v SDS, 5% v/v β-mercaptoethanol, 0.05% w/v bromophenol blue) at 90^{0}C for 5 minutes prior to electrophoresis. After electrophoresis, protein bands were visualised by staining with 10% v/v acetic acid, 40% v/v methanol, 0.1% w/v Coomassie Blue R-250 for 1 hour, followed by destaining in 10% v/v acetic acid, 40% v/v methanol for 2 hours or until the band of interest could be identified.

Western blot analysis of dhIFN-γ (128)

dhIFN-γ/6M GdmCl solution **(128)** was desalted, lyophilised and dhIFN-γ (0.5mg) and Multimark™ multi-coloured standards were loaded onto a 15% acrylamide SDS-gel. After electrophoresis the lane containing the molecular weight markers was cut out and kept for reference and the gel was allowed to equilibrate in transfer buffer (25mM Tris.HCl, pH 8.3, 187mM Glycine, 20% v/v Methanol) for 15 minutes. The proteins on the gel were then electrophoretically transferred to a nitrocellulose membrane using a BIO-RAD system for 1 hour at 100V. The nitrocellulose membrane was then blocked with blocking solution (5g Marvel in 100ml TBS: 20mM Tris, pH 7.5, 0.5M NaCl), for 1 hour at room temperature before being washed for 3 x 5 minutes with TTBS (20mM Tris, pH 7.5, 0.5M NaCl, 0.05% v/v Tween-20). The membrane was then incubated with monoclonal mouse anti-hIFN-γ antibody solution for 1 hour at room temperature. The TTBS wash procedure was repeated for 3 x 5 minutes before the membrane was incubated in biotinylated secondary antibody for 30 minutes. The membrane was then washed 3 x 5 minutes in TTBS followed by incubation in alkaline phosphatase conjugated ABC, diluted 1:30 in TTBS. After washing the membrane was incubated in the dark for 10-30 minutes with the enzyme substrate. The enzyme substrate was freshly prepared by adding 66μl NBT stock (Nitro Blue Tetrazolium, 50mg in 1ml 70% DMF) and 33μl BCIP stock (5 Bromo-4-Chloro-3-Indoyl Phosphate, 50mg in 1ml 100% DMF) to 10mls alkaline phosphatase buffer (0.1M Tris, pH 9.0, 0.1M NaCl and 50mM MgCl). After incubating for 15 minutes three dark bands had developed. Using the molecular weight marker lane

these bands were identified as approximately ~17kDa (dhIFN-γ requires 16.8kDa) ~34kDa (dimer of dhIFN-γ requires 33.6kDa).

Preparative glycine-SDS-PAGE of dhIFN-γ (128)

dhIFN-γ/6M GdmCl solution (**128**) was desalted and lyophilised and dhIFN-γ (0.5mg) dissolved in 5 x SDS reducing buffer (500μl). The sample was then heated at 90^{0}C for 5 minutes and loaded onto a 180mm x 150mm x 3mm 20% preparative polyacrylamide gel. Low molecular weight markers were loaded as a reference lane next to 10 sample elution lanes and the gel electrophoresed at 5mA at 4^{0}C overnight. After electrophoresis, the reference lane and the adjacent sample lane were cut out stained with Coomassie Blue R-250 (0.1% w/v in an aqueous solution of 40% v/v methanol, 10% v/v acetic acid) for 1 hour, followed by destaining in an aqueous solution of 40% v/v methanol, 10% v/v acetic acid for 1 hour. During this procedure the unstained gel was kept moist with a small amount of running buffer. The stained and unstained gels were aligned and the pertinent area identified on the stained gel. The location of the appropriate band on the unstained gel was then estimated by extrapolating from the stained gel. The band of interest was excised and eluted into 1ml elution buffer (192mM glycine, 25mM Tris.HCl, pH 8.3 and 0.1% SDS) for 5 hours at 4^{0}C using the Model 422 Electro-Eluter (Bio-Rad) according to the manufacturers instructions.

After elution the sample was dialysed (centricon 3000) against 10mM sodium phosphate buffer, pH 7.5 to remove SDS and concentrate the protein. The eluted protein dhIFN-γ (**132**) was examined by SDS-PAGE and a single band was visualised at 16.8kDa. Sequencing analysis from N-terminus: Found: residue(R)$_1$=Gln, R_2=Asp; dhIFN-γ requires R_1=Gln, R_2=Asp.

Chapter 5: References

[1] R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 3149.

[2] R. B. Merrifield, *Biochemistry,* 1964, **3**, 1385.

[3] L. A. Carpino and G. Y. Han, *J. Am. Chem. Soc.*, 1970, **92**, 5748.

[4] C. D. Chang and J. Meienhofer, *Int. J. Pept. Protein Res.*, 1978, **11**, 246.

[5] E. Atherton, D. Harkiss, C. J. Logan, R. C. Sheppard and B. J. Williams, *J. Chem. Soc. Chem. Commun.*, 1978, 537.

[6] E. Atherton, D. Harkiss and R. C. Sheppard, *J. Chem. Soc. Chem. Commun.*, 1978, 539.

[7] B. W. Erickson and R. B. Merrifield, *The Proteins,* eds. H. Neurath and R. L. Hill, vol. 2, 3rd edn., Academic Press, New York, 1976, pp255.

[8] E. Atherton and R. C. Sheppard, *Solid Phase Synthesis, A Practical Approach,* Oxford University Press, Oxford, 1989.

[9] E. Atherton, D. L. J. Clive and R. C. Sheppard, *J. Am. Chem. Soc.*, 1975, **97**, 6585.

[10] R. Ashardy, E. Atherton, D. L. J. Clive and R. C. Sheppard, *J. Chem. Soc. Perkin Trans. I,* 1981, 1151.

[11] A. R. Mitchell, B. W. Erickson, M. N. Ryabtsev, R. S. Hodges and R. B. Merrifield, *J. Am. Chem. Soc.*, 1976, **98**, 7357.

[12] S. S. Wang, *J. Am. Chem. Soc.*, 1973, **95**, 1328.

[13] M. Mergler, R. Tanner, J. Gosteli and P. Grogg, *Tetrahedron Lett.*, 1988, **29**, 4005.

[14] K. Barlos, D. Gatos, S. Kapolos, G. Paphotiu, W. Schafer and Y. Wenquing, *Tetrahedron Lett.*, 1989, **30**, 3947.

[15] R. Ramage, C. A. Barron, S. Bielecki and D. W. Thomas, *Tetrahedron Lett.*, 1987, **28**, 4105.

[16] R. Ramage, S. L. Irving and C. McInnes, *Tetrahedron Lett.*, 1993, **34**, 6599.

[17] R. Ramage, C. A. Barron, S. Beilecki and D. W. Thomas, *Tetrahedron Lett.*, 1992, **48**, 499.

[18] P. Lloyd-Williams, F. Albericio and E. Giralt, *Tetrahedron Lett.*, 1993, **49**, 11065.

[19] C. P. Holmes, D. G. Jones, B.T. Frederick and L. -C. Dong, *Peptides: Chemistry, structure and Biology,* eds. P.T.P. Kaumaya and R. S. Hodges, Mayflower Scientific Ltd., 1996, pp44.

[20] D. P. N. Satchell, *Chem. Ind.,* 1974, 683.

[21] J. H. Jones, *Chem. Ind.,* 1974, 723.

[22] E. Gross and J. Meienhofer, *The Peptides: Analysis, Synthesis, Biology,* Academic Press, New York, 1979, 11.

[23] M. Bodanszky, *Peptide Chem.,* Springer-Verlag, Berlin, 1988, pp115.

[24] J. C. Sheehan and G. P. Hess, *J. Am. Chem. Soc.,* 1955, **77**, 1067.

[25] D. Sarantakis, J. Teichman, E. L. Lien and R. L. Fenichel, *Biochem. Biophys. Res. Commun.,* 1976, **73**, 336.

[26] S. B. H. Kent, L. E. Wood, H. Beilan, S. Meister and T. Geiser, *Proc. 18th Eur. Pept. Symp.* Sweden, ed. U. Ragnarsson, Stockholm: Almqvist and Wiksell, 1984, pp185.

[27] H. Rink and B. Riniker, *Helv. Chim. Acta.,* 1974, **57**, 831.

[28] W. Konig and R. Geiger, *Ber. Dtsch. Chem. Ges.,* 1970, **103**, 788.

[29] L. A. Carpino, *J. Am. Chem. Soc.,* 1993, **115**, 4397.

[30] A. Davidson, PhD Thesis, University of Edinburgh, 1992.

[31] L. Jiang, PhD Thesis, University of Edinburgh, 1996.

[32] R. Ramage, G. W. Biggin, A. R. Brown, A. J. Comer, A. Davidson, L. C. Draffan, L. Jiang, G. H. Morton, N. Robertson, K. T. Shaw, G. Tennant, K. Urquart and J. Wilken, *Proc. 4th Int. Symp. in Solid Phase Synth. and Combinatorial Chemical Libraries,* Edinburgh, 1995.

[33] N. Robertson, PhD Thesis, Universiy of Edinburgh, 1996.

[34] M. Fujino, M. Wakimusa and C. Kitada, *Chem. Pharm. Bull.,* 1981, **29**, 2825.

[35] R. Ramage and J. Green, *Tetrahedron Lett.,* 1987, **28**, 2287.

[36] J. Green, O. Ogunjobi, R. Ramage and A. S. Stewart, *Tetrahedron Lett.,* 1988, **29**, 4341.

[37] G. Biggin, *Personal Communication*, University of Edinburgh.

[38] A. R. Fletcher, J. H. Jones, W. I. Ramage and A. V. Stachulski, *J. Chem. Soc. Perkin Trans. I,* 1979, 2261.

[39] R. Columbo, *J. Chem. Soc. Chem. Commun.*, 1984, 292.

[40] K. M. Otteson, R. L. Noble, H. Hoeprich, K. T. Shaw and R. Ramage, *Protein Analysis Renaissance,* 1993, 69.

[41] R. B. Merrifield, *Angew. Chem. Int.*, 1985, **24**, 799.

[42] G. Barany and R. B. Merrifield, *The Peptides, Analysis, Biology,* eds. E. Gross and E. Meienhofer, Academic Press, London, 1980, pp159.

[43] A. R. Brown, S. L. Irving and R. Ramage, *Tetrahedron Lett.*, 1993, **34**, 7129.

[44] R. Ramage, J. Green, T. W. Muir, O. M. Ogunjobi, S. Love and K. T. Shaw, *J. Biochem.*, 1994, **299**, 151.

[45] A. R. Brown, M. Covington, R. C. Newton, R. Ramage and P. Welch, *J. Pept. Sci.*, 1996, **2**, 40.

[46] B. Gutte and R. B. Merrifield, *J. Biol. Chem.*, 1971, **246**, 1922.

[47] I. Clark-Lewis, R. Aebersold, H. Ziltener, J. W. Schrader, L. E. Hood and S. B. H. Kent, *Science,* 1986, **231**, 134.

[48] K. Akaji, N. Fujii, H. Yajima, K. Hayashi and K. Mizuka, *Chem. Pharm. Bull.*, 1985, **33**, 184.

[49] P. Sieber, B. Kamber, A. Hartmann, A. Johl, B. Riniker and W. Rittel, *Helv. Chim. Acta.*, 1977, **60**, 27.

[50] H. Yajima and N. Fujii, *J. Am. Chem. Soc.*, 1981, **103**, 5867.

[51] K. Barlos, D. Gatos and W. Schafer, *Angew. Chem.*, 1991, **30**, 590.

[52] H. Kuroda, Y. -N. Chen, T. Kimura and S. Sakakibara, *Int. J. Pept. Protein Res.*, 1992, **40**, 294.

[53] T. Inui, J. Bodi, S. Kubo, H. Nishio, T. Kimura, S. Kojima, H. Maruta, T. Muramatsu and S. Sakakibara, *J. Pept. Sci.*, 1996, **2**, 28.

[54] M. Schnolzer and S. B. H. Kent, *Science,* 1992, **256**, 221.

[55] P. E. Dawson and S. B. H. Kent, *J. Am. Chem. Soc.*, 1993, **115**, 7263.

[56] P. A. Jekel, W. J. Weijer and J. J. Beintema, *Anal. Biochem.*, 1983, **134**, 347.

[57] T. Nakatsuka, T. Sasaki and E. T. Kaiser, *J. Am. Chem. Soc.*, 1987, **109**, 3808.

[58] D. Y. Jackson, J. Burnier, C. Quan, M. Stanley, J. Tom and J. A. Wells, *Science,* 1994, **266**, 243.

[59] L. Abrahmsen, J. Tom, J. Burnier K. A. Butcher, A. Kossiakoff and J. A. Wells, *Biochemistry,* 1991, **30**, 4151.

[60] P. E. Dawson, T. W. Muir, I. Clark-Lewis and S. B. H. Kent, *Science,* 1994, **266**, 776.

[61] T. W. Muir, M. J. Williams, M. H. Ginsberg and S. B. H. Kent, *Biochemistry,* 1994, **33**, 7701.

[62] R. G. Denkewalter, D. F. Veber, F. W. Holly and R. Hirschmann, *J. Am. Chem. Soc.,* 1969, **91**, 502.

[63] J. Honzl and J. Rudinger, *Coll. Czech. Chem. Commun.,* 1961, **26**, 2333.

[64] P. Sieber, B. Riniker, M. Brugger, B. Kamber and W. Kittel, *Helv. Chim. Acta.,* 1970, **53**, 2135.

[65] Y. S. Klausner and M. Bodanszky, *Synthesis,* 1974, **8**, 549.

[66] G. Goldstein, M. Scheid, U. Hammerling, EA. Boyse, D. H. Schlesinger and H. D. Niall, *Proc. Natl. Acad. Sci.,* USA, 1975, **72**, 11.

[67] S. P. Dawson, J. E. Arnold, N. J. Mayer, S. E. Reynolds, M. A. Billett, C. Gordon, L. Colleaux, P. M. Kloetzel, K. Tanaka and R. J. Mayer, *J. Biol. Chem.,* 1995, **270**, 1850.

[68] N. D. Hone, S. R. Chhabra, W. C. Chan and B. W. Bycroft, *Peptides,* eds. C. H. Schneider and A. N. Eberle, 1992, pp290.

[69] D. Lelievre, D. Daguet and A. Brack, *Tetrahedron Lett.,* 1995, **36**, 9317.

[70] R. Beugelmans, L. Neuville, M. Boischoussy, J. Chastanet and J. Zhu, *Tetrahedron Lett.,* 1995, **36**, 3129.

[71] Y. Nishiyama, N. Shintomi, Y. Kondo and Y. Okada, *J. Chem. Soc. Perkin Trans. I,* 1194, **21**, 3201.

[72] T. W. Greene and P. G. M. Wuts, *Protecting Groups in Organic Synthesis,* 2nd ed., Wiley and Sons, New York, 1991.

[73] H. Kunz and H. Waldmann, *Comprehensive Organic Synthesis,* eds. B. M. Trost, L. Flemming and E. Winterfeldt, Pergamon Press, Oxford, 1991, vol 6, pp 631-701.

[74] J. D. Glass, *The Peptides: Analysis, Synthesis and Biology,* eds. S. Udenfriend and J. Meienhofer, Academic Press, New York, 1987, vol. 9, pp 167-184.

[75] H. Waldmann and D. Sebastian, *Chem. Rev.,* 1994, **94**, 911.

[76] M. Cole, *Process Biochem.*, 1967, **2**, 35.

[77] M. Cole, *J. Biochem.*, 1969, **115**, 733.

[78] M. Cole and T. Savidge, *Meth. Enzymol.*, 1975, **43**, 705.

[79] F. Brtnik, T. Barth and K. Jost, *Coll. Czech. Chem. Commun.*, 1981, **46**, 1983.

[80] D. T. Pallin, *PhD thesis*, University of Edinburgh, 1993.

[81] H. T. Huang, T. A. Seto and G. M. Shull, *Appl. Microbiol.*, 1963, **11**, 1.

[82] H. J. Duggleby, S. P. Tolley, C. P. Hill, E. J. Dodson, G. Dodson and P. C. E. Moody, *Nature*, 1995, **373**, 264.

[83] T. W. Greene and P. G. M. Wuts, *Protecting Groups in Organic Synthesis*, 2nd ed., Wiley and Sons, New York, 1991, pp156.

[84] M. C. Venuti, B. E. Loe, G. H. Jones and J. M. Young, *J. Med. Chem.*, 1988, **31**, 2132.

[85] F. Albericio, E. Nicolas, J. Rizo, M. Ruiz-Gayer, E. Pedroso and E. Giralt, *Synthesis*, 1990, 119.

[86] R. E. Shute and D. H. Rich, *Synthesis*, 1987, 346.

[87] Y. Kiso, M. Satomi, K. Ukawa and T. Akita, *J. Chem. Soc. Chem. Commun., 1980, 1063.*

[88] J. Wilken, *PhD Thesis*, University of Edinburgh, 1995.

[89] A. Brown, *PhD Thesis*, University of Edinburgh, 1995.

[90] S. Love, *Personal Communication,* University of Edinburgh.

[91] Q. Wang, G. Liu and C. Shen, *Acta Biochim. Biophys. Sinica.*, 1980, **12**, 305.

[92] Q.-C. Wang, J. Fei, D. -F. Cui, S. -G. Zhu and L. -G. Xu, *Biopolmers*, 1986, **25**, s109.

[93] T. Pohl and H. Waldmann, *Tetrahedron Lett.*, 1995, **36**, 2963.

[94] A. R. J. Comer, *PhD Thesis*, University of Edinburgh, 1996.

[95] K. T. Shaw, *Personal Communication,* University of Edinburgh.

[96] P. Wang, *Personal Communication,* University of Edinburgh.

[97] S. S. Davis, *J. Pharm. Pharmacol.*, 1992, **44**, 186.

[98] B. M. Wallace and J. S. Lasker, *Science*, 1993, **250**, 912.

[99] K. Tachibana, *Pharm. Res.*, 1992, **9**, 952.

[100] S. Mitragotri, D. Blankschtein and R. Langer, *Science*, 1995, **269**, 850.

[101] M. A. Farrar and R. D. A. Schreiber, *Rev. Immun.,* 1993, **11**, 571.

[102] S. Pestka, J. A. Lager, K. C. Zoon and C. E. Samuel, *Rev. Biochem.,* 1987, **56**, 727.

[103] J. Link, *Acta Neurol. Scand. Suppl.,* 1994, **158**, 1.

[104] P. W. Gray and D. V. Goeddel, *Nature,* 1982, **298**, 859.

[105] E. Rinderknecht, B. H. O'Connor and H. Rodriguez, *J. Biol. Chem.,*1984, **259**, 6790.

[106] Y. -C. E. Pan, A. S. Stern, P. C. Familetti, F. R. Khan and R. Chizzonite, *Eur. J. Biochem.,* 1987, **166**, 145.

[107] Y. K. Yip, B. S. Barrowclough, C. Urban and J. Vilcek, *Proc. Natl. Acad. Sci. U.S.A.,* 1982, **79**, 1820.

[108] P. W. Gray, D. W. Leung, D. Pennica, E. Yelverton, R. Najarian, C. C. Simonsen, R. Derynck, P. J. Sherwood, D. M. Wallace, S. L. Berger, A. D. Levinson and D. V. Goeddel, *Nature,* 1982, **295**, 503.

[109] S. Tanaka, T. Oshima, K. Ohsuye, T. Ono, A. Mizono, A. Ueno, H. Nakazato, M. Tsujimoto, N. Higashi and T. Noguchi, *Nucleic Acids Res.,* 1983, **11**, 1707.

[110] S. E. Ealick, W. J. Cook, V. K. Senadhi, M. Carson, T. L. Nagabhushan, P. P. Trotta and C. E. Bugg, *Science*, 1991, **252**, 698.

[111] M. R. Walter, W.. T. Windsor, T.. L. Nagabhushan, D J. Lundell, C.. A. Lunn, P. J Zauodny and S. K Narula, *Nature,* 1995, **376**, 230.

[112] P. Anderson, Y. K. Yip and J. Vilcek, *J. Biol. Chem.,* 1982, **257**, 11301.

[113] M. Aguet, Z. Dembic and G. Merlin, *Cell,* 1988, **55**, 273.

[114] A. C. Greenlund, R. D. Schreiber, D. V. Goeddel and D. Pennica, *J. Biol. Chem.,* 1993, **268**, 18103.

[115] K. Alton, Y. Stabinsky, R. Richards, B. Ferguson, L. Goldstein, B. Altrock, L. Miller and N. Stebbing, *The Biology of Interferon System,* eds E. De Maeyer and H. Schellekens, Elsevier Science Publishers, Amsterdam, 1983, pp 119.

[116] K. Cantell, S. Hirvonen, T. Sareneva, J. Pirhonen and I. Julkunen, *J. Interferon Res.,* 1992, **12**, 175.

[117] L. E. Burton, P. W. Gray, D. V. Goeddel and E. Rinderknecht, *Biology of the Interferon System*, ed. H. Kirchner and H. Schellekens, Amsterdam, Elsevier, pp403.

[118] T. Arakawa, Y. -R. Hsu, C. G. Parker and P. -H. Lai, *J. Biol. Chem.*, 1986, **261**, 8534.

[119] H. Döbeli, R. Gentz, W. Jucker, G. Garotta, D. W. Hartmann and E. Hochuli, *J. Biotech.*, 1988, **7**, 199.

[120] G. H. Morton, *Personal Communication,* University of Edinburgh.

[121] Y. K. Yip, R. H. L. Pang, C. Urban and J. Vilcek, *Proc. Natl. Acad. Sci. U.S.A.*, 1981, **78**, 1601.

[122] K. Miyata, Y. Yamamoto, M. Ueda, Y. Kawade, K. Matsumoto and I. Kubota, *J. Biochem.*, 1986, **99**, 1681.

[123] H. -F. Kung, Y. -C. E. Pan, J. Moschera, K. Tsai, E. Bekesi, M. Chang, H. Sugino and S. Honda, *Meth. Enzym.*, 1986, **119**, 204.

[124] T.E. Creighton, WO 86/05809.

[125] J. L. Cleland and D. I. C. Wang, *Bio/Technology,* 1990, **8**, 1274.

[126] S. Tandon and P. M. Horowitz, *J. Biol. Chem.*, 1989,**262**, 4486.

[127] J. L. Cleland, S. E. Builder, J. R. Swartz, M. Winkler, J. Y. Chang and D. I. C. Wang, *Bio/Technology,* 1992, **10**, 1013.

[128] A.D. Cronshaw, Welmet Sequencing, University of Edinburgh.

[129] J. D Hayes, L. A. Kerr and A. D. Cronshaw, *J.Biochem.*, 1989, **264**, 437.

[130] H. Towbin, T. Staehelin and J. Gordon, *Proc. Natl. Acad. Sci. USA,* 1979, **76**, 4350.

[131] S. E. M. Howie, *Personal Communication,* University of Edinburgh.

[132] C. T. Samudzi, L. E. Burton and J. R. Rubin, *J. Biol. Chem.*, 1991, **266**, 21791.

[133] H. H. Hogrefe, P. McPhie, J. B. Bekisz, J. C. Enterline,D. Dyer, D. S. A. Webb, T. L. Gerrard and K. Zoon, *J. Biol. Chem.*, 1989, **264**, 12179.

[134] T. Arakawa, Y. -R. Hsu and D. A. Yphantis, *Biochemistry*,1987, **26**, 5428.

[135] P. Weiss, *J. Am. Chem. Soc.*, 1948, **70**, 4263.

[136] F. F. Blicke and E. L. Schumann, *J. Am. Chem. Soc.* 1954, **76**, 3153.

[137] U. K. Laemmli, *Nature*, 1970, **227**, 680.

Courses and conferences attended

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

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

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, St Andrews 1994 and Glasgow 1995, 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.

Twenty fourth International Symposium of the European Peptide Society, University of Edinburgh, 1996, various speakers.