The Chemical Synthesis of Proteins and Peptide *C*-Terminal Derivatives

by

Jennifer Ann Patterson

A thesis submitted for the
degree of Doctor of Philosophy

University of Edinburgh
1999

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 is original and has not been previously submitted in whole or in part, for any degree at this, or any other university.

To my parents and my brother, Richard

Acknowledgements

I would like to thank Professor Robert Ramage for the opportunity to continue my studies at Edinburgh University, and for his supervision, constant support and encouragement throughout the course of my PhD.

I am extremely grateful to Mr Kevin Shaw for assistance with the chemical synthesis of the peptides and proteins; Mr Brian Whigham for MALDI-TOF MS and amino acid analyses; Dr Andrew Cronshaw (Welmet Protein Characterisation Facility) for rapid and efficient *N*-terminal sequencing and Dr Emma Beatty (Edinburgh Centre For Protein Technology) for 600 MHz NMR.

Special thanks must go to Drs Dominic Campopiano, Gail Morton and Nicola Robertson and Miss Lisa McIver for teaching me the techniques necessary for my interferon-gamma work, and for the general advice offered. I am also indebted to Dr Martin Andrews and Craig Jamieson for proof reading this thesis.

Finally, I would like to thank the Ramage group past and present, and my friends and colleagues in Edinburgh for making my time here so memorable. I must especially thank Martin, Craig, Lorraine Bland, Carolyn Gordon and Dr Alastair Hay for providing unforgettable entertainment and support during the past three years.

Abstract

Methodology for the synthesis of peptide *C*-terminal aldehydes has been investigated. Modification of a linker system, based on a terabenzosuberyl construct, has been demonstrated to be suitable for the synthesis of peptide *C*-terminal semicarbazones. The route has been fully optimised to yield a series of peptide *C*-terminal semicarbazones and the corresponding peptide aldehydes.

The chemical synthesis of deglycosylated human interferon-gamma (143 residues) has been carried out. The purification of this protein has been investigated and a short purification protocol developed which is sufficiently general to allow application to other similar protein systems. Purification and characterisation of the synthetic interferon-gamma molecule has been completed and folding of the molecule attempted.

Abbreviations

Å	angstroms
AAA	amino acid analysis
ABI	Applied Biosystems
Abs	absorbance
Ac	acetyl
Ac_2O	acetic anhydride
AcOH	acetic acid
APS	ammonium persulphate
BBB	blood brain barrier
Boc	*tert*-butyloxycarbonyl
^0C	degrees centigrade
CAPS	3-[cyclohexylamino]-1-propanesulphonic acid
Cbz	benzyloxycarbonyl
CD	circular dichroism
cDNA	circular deoxyribonucleic acid
CE	cellulose ester
CGD	chronic granulomatous disease
cm	centimetres
CM	carboxymethyl
CNS	central nervous system
Cs_2CO_3	caesium carbonate
CsOH	caesium hydroxide
d	doublet
D	*dextro* rotatory
Da	daltons
DCC	*N, N*-dicyclohexylcarbodiimide
DCM	dichloromethane
DDT	dithiotheitol
dhIFN-γ	deglycosylated human interferon-gamma
diam	diameter
DIBAL	diisobutylaluminium hydride
DIC	*N, N*-diisopropylcarbodiimide
DIEA	*N, N*-diisopropylethylamine
dm	decimetres
DMAP	*N, N*-dimethylaminopyridine
DMF	*N, N*-dimethylformamide
DMSO	dimethylsulphoxide
DNA	deoxyribonucleic Acid
ECE	endothelin converting enzyme
E. Coli	*Eschericia Coli*
EDT	ethanedithiol
EDTA	ethylenediamine tetraacetic acid
EI	electron impact
eq	equivalents
ESI	electrospray ionisation

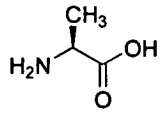
EtOH	ethanol
FAB	fast atom bombardment
FF	fast flow
Fmoc	9-fluorenylmethoxycarbonyl
FPLC	fast protein liquid chromatography
FT-IR	fourier transform infra red
Fuc	fucose
g	gram
Gal	galactose
GlcNAc	*N*-acetylglucosamine
HBr	hydrogen bromide
HCl	hydrogen chloride
HF	hydrogen fluoride
HOAt	1-hydroxy-7-azabenzotriazole
HOBt	1-hydroxybenzotriazole
HOCt	ethyl-1-hydroxy-1*H*-1,2,3-triazole-4-carboxylate
HPLC	high pressure liquid chromatography
HR	high resolution
IEF	isoelectric focusing
IFN	interferon
IFN-γR_{α}	interferon-gamma receptor extracellular domain
IFN-$\gamma R_{\beta 1}$	interferon-gamma receptor binding accessory factor
IPA	propan-2-ol
J	coupling constant
K	potassium
k_{AV}	elution volume parameters
KBr	potassium bromide
kDa	kilodalton
L	*levo* rotatory
L	litre
$LiBH_4$	lithium borohydride
m	multiplet
M	molar
MALDI	matrix assisted laser desorption ionisation
Man	mannose
MBHA	4-methylbenzhydrylamine
MeCN	acetonitrile
MeOH	methanol
mg	milligram
MHz	mega hertz
$MgSO_4$	magnesium sulphate
min	minute
ml	millilitre
mm	millimetre
mM	millimolar
mmol	millimole
mol	mole

Mpt	melting point
MS	mass spectrometry
mV	millivolt
MWt	molecular weight
MWtCO	molecular weight cut off
Na	sodium
NaCl	salt
NaHMDS	sodium hexamethyldisilamide
NaOAc	sodium acetate
NeuAc	*N*-acetylneuraminic acid
NK	natural killer
nm	nanometres
NMR	nuclear magnetic resonance
P	protecting group
PA	peptide aldehyde
PAGE	polyacrylamide gel electrophoresis
PAM	4-hydroxymethylphenylacetamidomethyl
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulphonyl
PEG	polyethylene glycol
PEGA	polyethylene glycol-dimethylacrylamide copolymer
PGC	porous graphitised carbon
$PhSO_3H$	phenylsulphonic acid
pI	isoelectric point
Pmc	2,2,5,7,8-pentamethylchroman-6-sulphonyl
pmol	picomole
ppm	parts per million
PVDF	polyvinyldifluoride
py	pyridine
rhIFN-γ	recombinant human interferon-gamma
rhIFN-γ1b	recombinant human interferon-gamma1b
R_f	retention factor
RP	reversed phase
rpm	revolutions per minute
RT	room temperature
s	singlet
sc	semicarbazone
SDS	sodium dodecylsulphate
SP	sulphonylpropyl
SPPS	solid phase peptide synthesis
t	triplet
t	tertiary
TBAF	tetrabutylammonium fluoride
Tbfmoc	17-tetrabenzo[*a, c, g, i*]fluorenylmethoxycarbonyl
^{t}Bu	*tertiary*-butyl
TEMED	*N, N, N, N*-tetramethylethylenediamine
TFA	trifluoroacetic acid
TFAA	trifluoroacetic acid anhydride

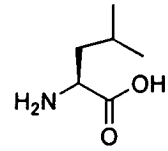
TFMSA	trifluoromethylsulphonic acid
THF	tetrahydrofuran
TIS	triisopropylsilane
tlc	thin layer chromatography
TMS	trimethylsilane
TOF	time of flight
Tris	tris(hydroxymethyl)aminomethane
Trt	trityl (triphenylmethyl)
UV	ultra violet
v	volume
Ve	elution volume
Vo	void volume
Vt	total volume
vis	visible
w	weight
W	watts
Z	benzyloxycarbonyl
λ	wavelength
µl	microlitre
µm	micro metre
µM	micro molar

Naturally Occurring Amino Acids

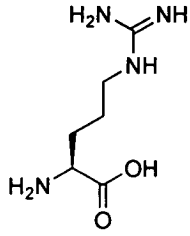
Alanine
Ala
A



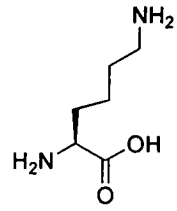
Leucine
Leu
L



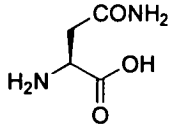
Arginine
Arg
R



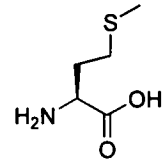
Lysine
Lys
K



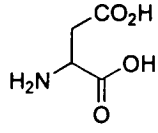
Asparagine
Asn
N



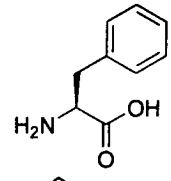
Methionine
Met
M



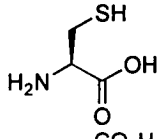
Aspartic Acid
Asp
D



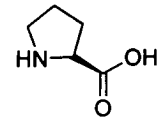
Phenylalanine
Phe
F



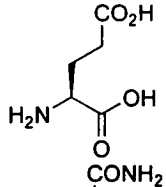
Cysteine
Cys
C



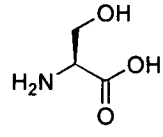
Proline
Pro
P



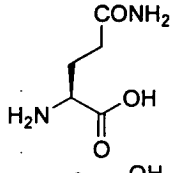
Glutamic Acid
Glu
E



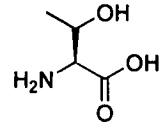
Serine
Ser
S



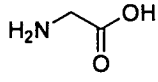
Glutamine
Gln
Q



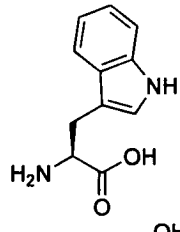
Threonine
Thr
T



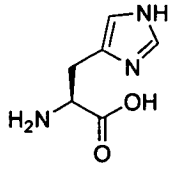
Glycine
Gly
G



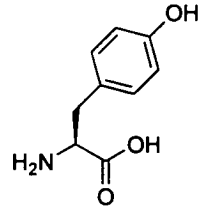
Tryptophan
Trp
W



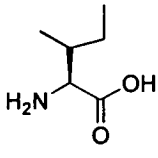
Histidine
His
H



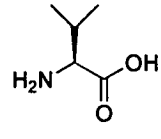
Tyrosine
Tyr
Y



Isoleucine
Ile
I



Valine
Val
V



Contents

Chapter 1: Solid Phase Peptide Synthesis

Chapter 2: The Solid Phase Synthesis of Peptide C-Terminal Semicarbazones and Aldehydes

Chapter 3: The Stepwise Chemical Synthesis of Deglycosylated Human Interferon-Gamma and its Purification

Chapter 4: Experimental

Chapter 1

Solid Phase Peptide Synthesis

1.1 Introduction

Peptides and proteins fulfil crucial functions in all biological processes, for example as hormones, enzymes, in cell-cell recognition and in the immune response. They are produced in nature by the cell's biosynthetic machinery at the site where the biological actions of the protein are required. Only small quantities of proteins are produced naturally due to the site directed synthesis and the potency of these molecules. As a result, there are only limited quantities of proteins available from natural sources, restricting the studies that can be carried out. Hence, there has been a powerful drive to perform the total chemical synthesis of peptides and proteins, rendering the production of sufficient material for study routine. Recent interest has been prompted by the use of peptides as antigens and synthetic vaccines, and the use of synthetic proteins to study folding phenomena.

1.1.1 Peptide Synthesis

Traditional solution phase synthesis of peptides is performed using a combination of classical coupling reagents, protecting group strategies and recrystallisation after each step to purify intermediates. This makes the synthesis of a relatively short peptide a laborious process, and often severe problems of insolubility may be encountered.

By consideration of the coupling of two amino acids, the fundamental issues of protein synthesis can be illustrated. A mixture of four different dimers may be obtained, **figure 1.1**. It is also impossible to prevent further chain elongation from the *N*- or *C*- termini of any of these dimers, as well as any reactive side chain functionality present. All these add to the accumulated by-products.

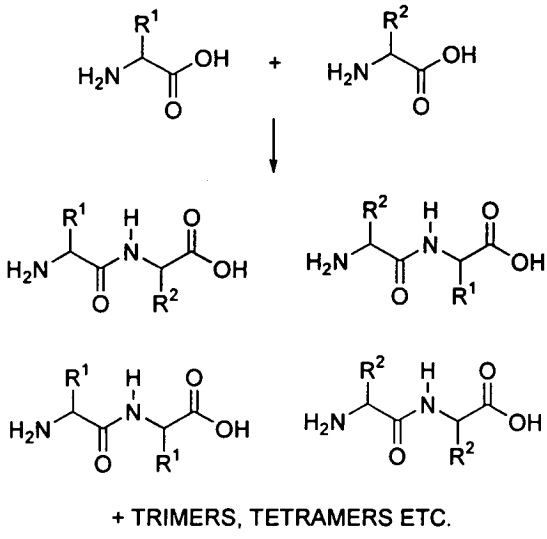
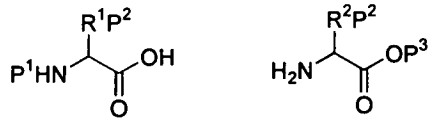


Figure 1.1: Coupling of Two Amino Acids

In order to achieve a unique dimeric product with the desired order of α-amino acid units, the non-participating amino and carboxylic acid functions, as well as any reactive side-chain functionality have to be masked. Hence, for a synthesis to be successful, a protecting group strategy must be invoked. The building blocks used now require three different levels of protection, P^1 (N^{α}-protection), P^2 (side chain functionality protection) and P^3 (carboxyl protection), **figure 1.2**.



P^1, P^2, P^3 = protecting groups

Figure 1.2: Suitably Protected Amino Acids
For Stereoselective Coupling

Development of protecting group strategies allowed the desired regiocontrol to be achieved, but did not remove the time consuming purification steps which are difficult to complete successfully. However, solution phase synthesis methodology successfully achieved the historically important and Nobel prize winning synthesis of oxytocin.[1,2]

In the 1950s, as the number of newly discovered peptides and proteins grew, so did the demand for synthetic peptides and peptide analogues. This in turn prompted a

ground breaking shift in peptide synthesis. In 1963, Bruce Merrifield described his simple, yet effective, new synthetic methodology, now termed Solid Phase Peptide Synthesis (SPPS).[3]

1.1.2 Solid Phase Peptide Synthesis

In SPPS, the first amino acid of the peptide backbone is attached *via* the carboxyl group to a solid support, thus providing P^3 of **figure 1.2**. This linkage must be stable to the conditions required for the removal of P^1. The side chain protection, P^2, should also be stable to the conditions employed for the removal of P^1. The resin bound amino acid is then condensed with a second amino acid, followed by a third, achieving stepwise elongation of the chain, **figure 1.3**. Today, SPPS is the method of choice for peptide synthesis, as it is a rapid, effective method of synthesis.

There are several advantages of SPPS which have resulted in the widespread use of this method. The first of these are that the main impurities, the unreacted amino acid and coupling reagents, can be simply removed by filtration and washing, avoiding purification procedures and giving easy isolation of intermediates. One consequence of this is that greater yields of product can be obtained by using vast excesses of the reagents during each cycle of the synthesis in order to drive each step to completion. This is not possible in solution phase synthesis, as it greatly complicates workup and purification procedures. Also, the solubility problems encountered in solution phase, particularly after the peptide chain has reached a certain length, are reduced. This is due to the lightly cross-linked polymer chains becoming intimately mixed with the peptide chains, exerting mutual solvating effects on each other. Due to the repetitive nature of the procedure, the method has been successfully developed into an automated process, and optimisation has allowed the routine synthesis of large polypeptides and proteins, inaccessible by the classical methods.

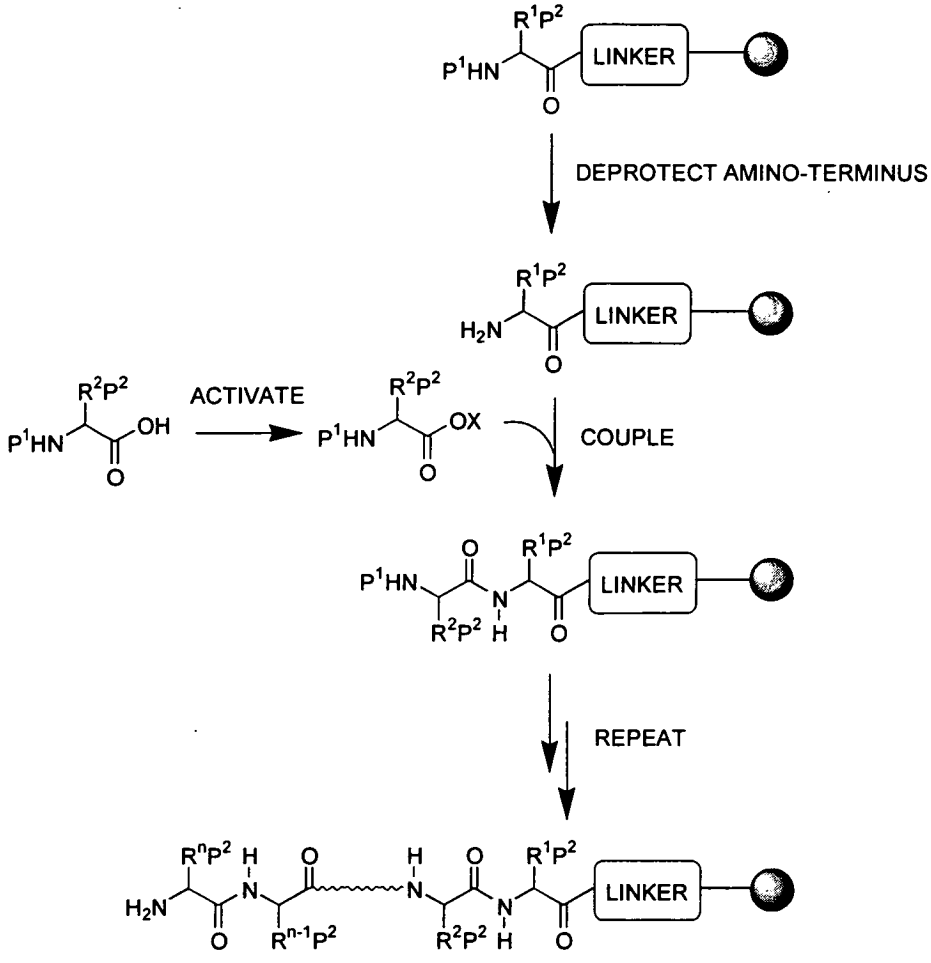


Figure 1.3: Schematic Representation of SPPS

1.2 Protecting Group Strategies

There are stringent requirements a protecting group must meet in order to be used successfully. It must suppress the reactivity of the functional group, without introducing any new and undesirable reactivity. It must be completely stable to the conditions used in each cycle of the peptide synthesis. It must also deprotect quantitatively and the deprotection must require conditions which do not damage the peptide chain

1.2.1 N^{α}-Protecting Group Strategies

The first group used for N^{α}-protection in SPPS was the benzyloxycarbonyl (Cbz or Z) group,[4] **figure 1.4**. This protecting group is removed using strong acid conditions (HBr in acetic acid), and was used in conjunction with acid labile side-chain

protection and anchor to the solid support. This group has gone out of favour since the repetitive acid cleavage steps required to release the *N*-terminus result in partial, slow deprotection of the side-chains and the loss of peptide from the solid support. This produces lower yields of the desired product, and again complicates purification due to the incidence of unwanted reactions at the side-chain functionalities.

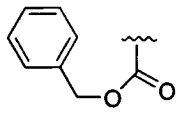


Figure 1.4: Z-Protecting Group

In subsequent syntheses, Merrifield replaced the Cbz group with the *tert*-butoxycarbonyl (Boc) group,[5, 6] **figure 1.5**, for his synthesis of Bradykinin.[7, 8] This group is cleaved using milder acid conditions (4 M HCl/dioxane or 50 % TFA/DCM) preventing cleavage of side-chain protecting groups and the link to the solid support, which are based on HF labile benzyloxycarbonyl ethers and esters. This strategy has been used extensively for SPPS; however, it is possible repetitive exposure to TFA solution can damage the peptide sequence by alteration of sensitive peptide bonds and initiation of acid catalysed side reactions.

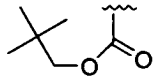


Figure 1.5: Boc-Protecting Group

The protection strategy currently favoured involves the use of 9-fluorenylmethoxycarbonyl (Fmoc) for *N*-terminal protection, **figure 1.6**. This group was originally developed by Carpino and Han,[9] and was first applied to SPPS by Meienhofer[10] and Sheppard.[11] It has the advantage that it offers a truly orthogonal protection strategy. The N^{α}-protecting group is cleaved under mildly basic conditions (20 % piperidine or morpholine in DMF), by a β-elimination mechanism. The side-chain protection and the linkage to the solid support thus may be labile to mild acid conditions (TFA/water) exposing the peptide to acid on only one occasion.

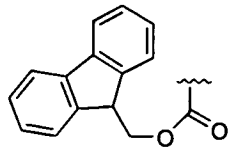


Figure 1.6: Fmoc-Protecting Group

Another advantage of this method is that the fulvene formed on cleavage of Fmoc and the adduct of the dibenzofulvene group and the excess piperidine is UV active, **figure 1.7**, with an isobestic point at 302nm. By UV examination of the deprotection solution in each cycle, an indication of the efficiency of the previous cycle can be obtained.[12]

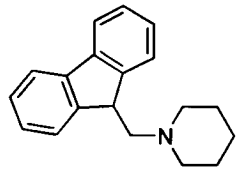


Figure 1.7: Fulvene-Piperidine Adduct

1.2.2 Fmoc-Compatible Side Chain Protection

As mentioned above, the side chain protecting groups may now be cleaved using mild acid, particularly TFA. The majority of amino acids are currently protected using ^{t}Bu ether or ester based groups. A variety of groups have been developed for amino acids such as arginine and cysteine which contain more problematic side-chain functionalities. **Table 1.1** shows the most commonly used Fmoc-compatible side-chain protecting groups.

Amino Acid	Protecting Group	Reference
Gly, Ala, Leu, Ile,Val, Phe,	not necessary	
Met, Pro		
Asp, Glu	^{t}Bu ester	13
Ser, Thr, Tyr	^{t}Bu ether	13
Lys, Trp	Boc	13, 14
Arg	Pmc or Pbf	15, 16, 17
Asn, Gln, His, Cys	Trt	18, 19, 20

Table 1.1: Fmoc-compatible side-chain protection

1.3 The Nature of the Solid Support

SPPS consists of a heterogeneous reaction mixture composed of an insoluble resin-bound peptide chain and a soluble activated amino acid derivative and solvent. However, it is important to remember that the reactions do not take place in, or on the surface of, the solid phase but in the swollen gel system produced by solvent penetration into the polymeric matrix. This produces a highly solution-like environment in which the synthesis occurs. As a result, the nature of the support chosen is crucial to the success of any synthesis. Many supports have been investigated, but only a few have met the stringent requirements and found widespread use.

To be suitable, the swollen polymer must be stable and inert in the reaction conditions employed at each stage of every cycle of the peptide synthesis. It must be sufficiently active to allow functionalisation with a linker moiety to which the peptide can be attached, grown upon and released from in good yield. The physical characteristics of the polymer are also important. The porosity and swelling properties must allow good penetration of the reagents and solvents to the internal active sites, allowing good contact between the growing peptide chains and reagents. In order to prevent peptide aggregation, the swollen polymer should also exert a strong solvating effect on the attached peptide chains, and have hydrophobicity properties which may minimise interactions between growing peptide chains.

1.3.1 Polystyrene Based Supports

In his ground breaking synthesis, Merrifield introduced a beaded form of a copolymer of polystyrene-divinylbenzene which contained reactive chloromethyl sites,[3] **figure 1.8**. This polystyrene resin was found to swell in a wide range of solvents including DMF, DCM and toluene, and as a result has found widespread use. The resin has the advantage that the reactive sites are uniformly distributed throughout the resin matrix, and not just on the surface of the beads, allowing high loadings of reaction sites to be obtained.

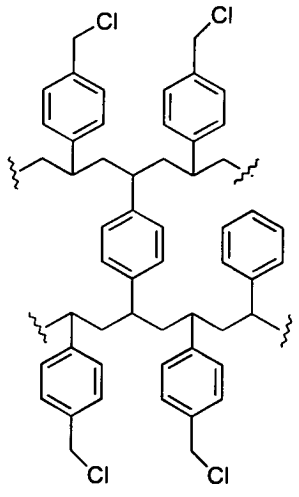


Figure 1.8: Divinylbenzene-Crosslinked Polystyrene Resin

1.3.2 Polyacrylamide Based Supports

Polyacrylamide resins, **figure 1.9**, were developed to be less hydrophobic than polystyrene, and to be more compatible with the polarity of the growing peptide chains.[21, 22] This was intended to ensure better solvation of larger peptide chains. Polyacrylamide resins were found to possess swelling characteristics complementary to those of the Merrifield polystyrene resin, with optimal swelling observed in DMF, and they have found widespread use in both Boc-SPPS and batch and continuous flow Fmoc-SPPS.

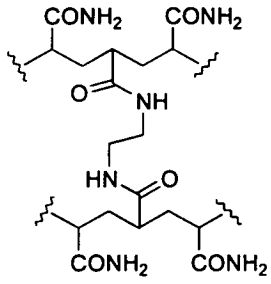
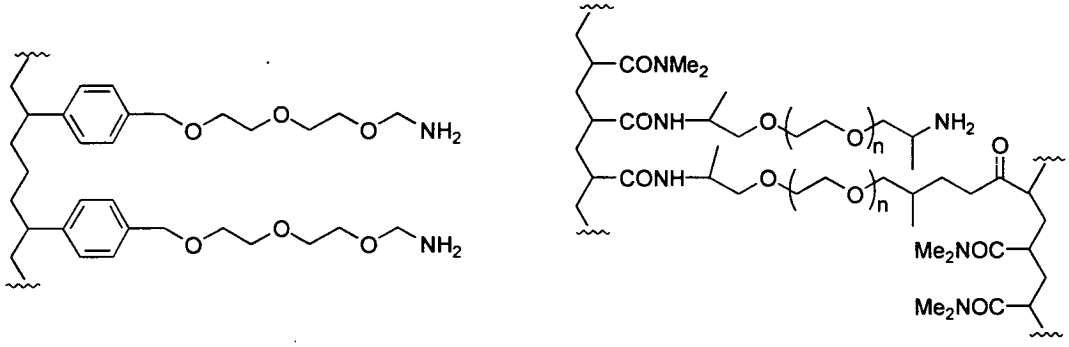


Figure 1.9: Polyacrylamide Resin

1.3.3 Supports Containing Polyethyleneglycol Grafts

These supports are prepared by grafting of sections of polyethyleneglycol molecules onto traditionally used resins, for instance polystyrene or polyacrylamide. These supports increase the solution-like nature of the environment the peptide chain is grown in, hence improving the solubility of the peptide chain. Examples of such supports are Tentagel®[23] and PEGA,[24] **figure 1.10**.



Tentagel® PEGA

Figure 1.10: Polyethyleneglycol Resins

1.3.4 Linker Groups

Linker molecules are designed to be incorporated between the solid support and the growing peptide chain. The impetus behind such research is that the peptide-support link can be fine tuned to make the anchor more sensitive to certain cleavage conditions. Linker molecules have been successfully designed to be cleaved by strong, medium or weak acids, bases or other nucleophiles, and by photolysis, hydrogenation or the use of other catalytic reagents. They have also been used to

incorporate *C*-terminal carboxylic acids, esters, amides or hydrazides on liberation from the solid support. Much of the recent research has centred on linkers which can be cleaved under very mild conditions, thus yielding peptides retaining side-chain protection for use in convergent protein synthesis. There is a wide range of linkers available for Fmoc-SPPS, summarised in **table 1.2**.

1.4 Activation and Coupling

In order to ensure to only the desired reaction occurs, the coupling step must be rapid and quantitative, even with hindered amino acid components. It must also proceed under mild conditions, avoiding side reactions and in particular changes in stereochemical integrity. In general, the incoming amino acid must be activated prior to coupling with the resin bound amino acid or peptide, in order to render it more susceptible to nucleophilic attack. Activating groups are electron withdrawing, in order to increase the electrophilicity of the carboxyl group and hence its reactivity. The incoming amino acid can be preactivated in a stable, crystalline form, or can be activated *in situ* using the appropriate reagents.

During activation and coupling, care must be taken to avoid deprotonation of the activated amino acid, **figure 1.11**. This could result in cyclisation to an oxazolone intermediate, **1**. Although the oxazolone system can be opened by the incoming amine to give the correctly formed product, **2**, it is possible that prior to ring opening an achiral oxazolone, **3**, can be produced. This would result in the incorporation of both the *L*- and the *D*-amino acid into the peptide, **2** and **5**.

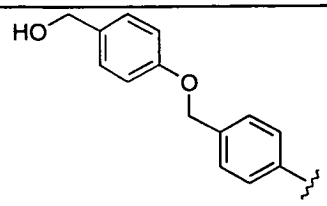
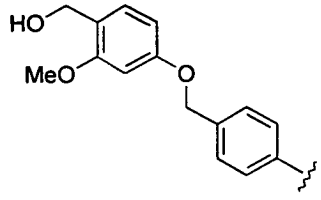
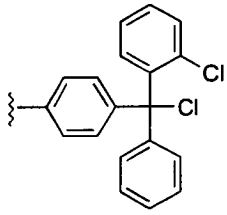
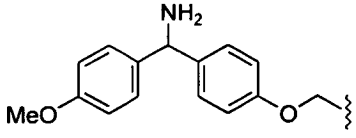
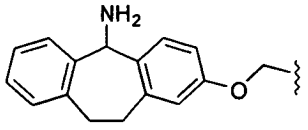
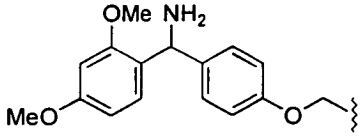
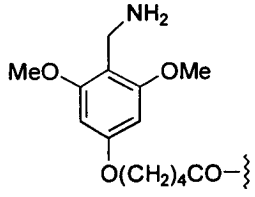
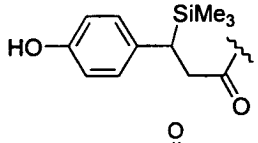
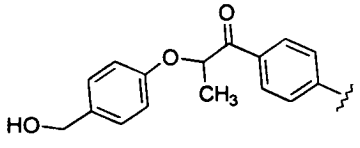
Linker	Cleavage Conditions	C-Terminal Derivative Produced	Reference
	50% TFA	carboxylic acid	25
	0.1-0.5 % TFA in DCM	acid with intact side-chain protection	26
	10 % AcOH	acid with intact side-chain protection	27
	90% TFA	amide	28
	50 % TFA in DCM	amide	29
	dilute TFA	amide with intact side-chain protection	30
	2% TFA in DCM	amide with intact side-chain protection	31
	TFA/DCM (7:3)	amide with intact side-chain protection	32
	TBAF	acid with intact side-chain protection	33
	hν (350nm)	acid with intact side-chain protection	34

Table 1.2: Linkers Commonly Used for Fmoc-SPPS

The susceptibility to this type of racemisation is dependent on the type of N^{α}-protection employed. Amino acids containing an N^{α}-amide group, for instance in the synthesis of peptides in the opposite direction, from the *N*-terminus to the *C*-terminus, are extremely likely to racemise in this manner. However, urethane protected monomers, for example Fmoc-amino acids are not susceptible to racemisation *via* the oxazolone mechanism.

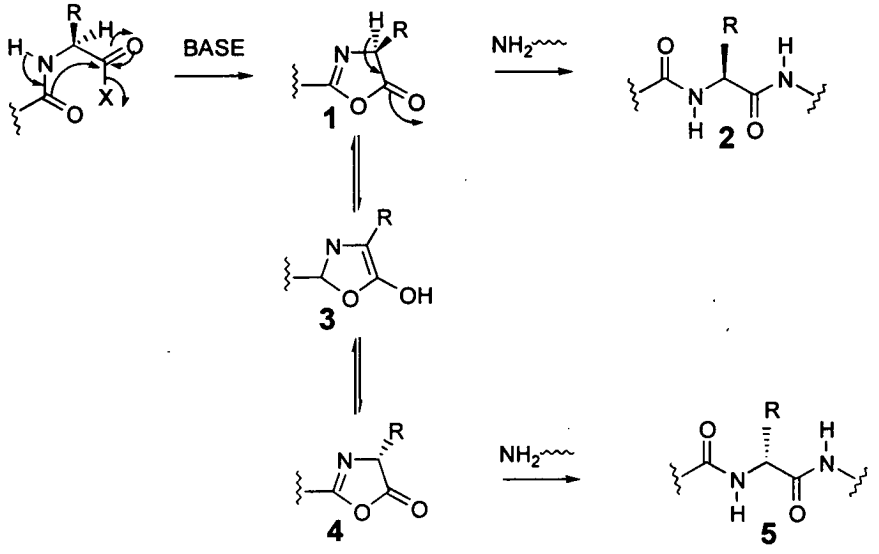
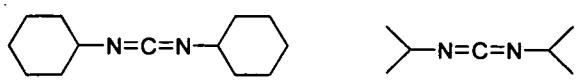


Figure 1.11: Racemisation *via* Oxazolone Mechanism

1.4.1 Carbodiimides

In the initial report of SPPS, Merrifield described the use of *N*, *N*-dicyclohexylcarbodiimide (DCC) as an activating agent,[3] **figure 1.12**. This agent was originally reported by Sheehan and Hess,[35] and has been found to be one of the most effective coupling reagents, especially in large scale synthesis. However, this reagent can cause racemisation of labile amino acids such as Phe. Insoluble urea by-products are also formed during the reaction, which can be problematic during SPPS. The use of *N*, *N*-diisopropylcarbodiimide (DIC),[36] which forms a soluble urea by-product, has overcome this problem.



DCC DIC

Figure 1.12: Carbodiimide Coupling Reagents

The carbodiimide reacts with the amino acid to form an *O*-acylisourea, **6**, from which there are three possible outcomes: reaction with the incoming amine to give the coupled product, **7**; reaction with another mole of the amino acid to give a symmetrical anhydride, **8**; and finally, reaction with a hydroxyl bearing molecule to yield an ester molecule, **9**, **figure 1.13**. The active ester, **9**, and the symmetrical anhydride, **8**, can also react with an incoming amino component to generate the coupled product, **7**.

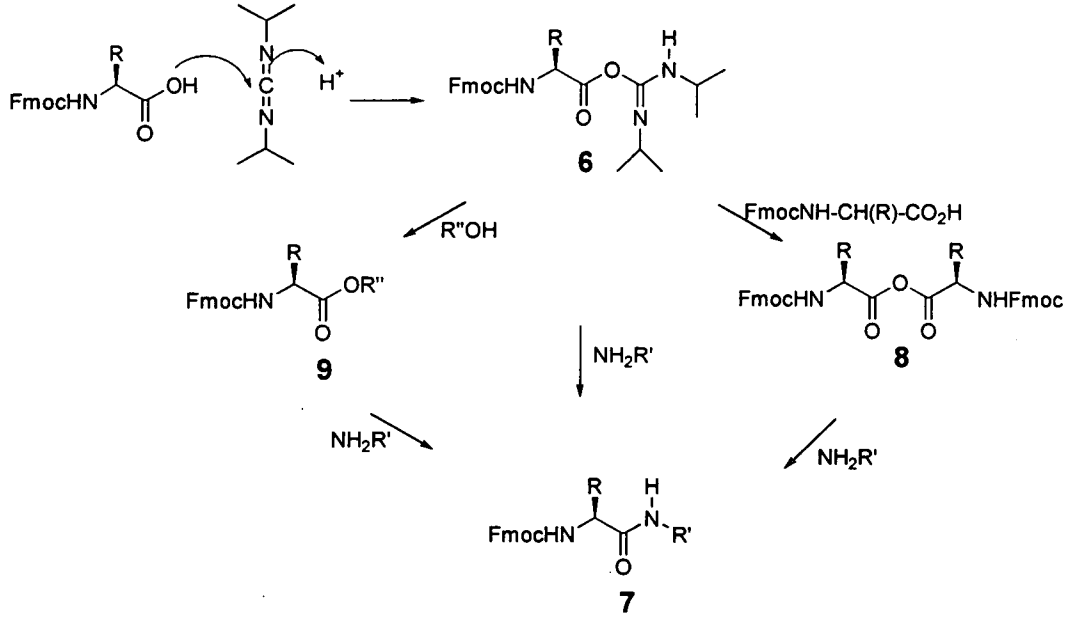


Figure 1.13: Activation and Coupling Methods Using DIC

1.4.2 Symmetrical Anhydrides

The coupling of symmetrical anhydrides, **figure 1.13**, is rapid and unambiguous. Symmetrical anhydrides themselves are easy to prepare, however, this method has the disadvantage that it is uneconomical, since one mole of amino acid is wasted per mole of symmetrical anhydride used.

1.4.3 Active Esters

Active esters were originally designed as coupling reagents in order to reduce the racemisation experienced using carbodiimide reagents alone. Reagents synthesised from *N*-hydroxytriazole molecules have found widespread use, due to their ease of preparation, efficiency in coupling and low occurrence of side reactions. Examples of commonly used reagents are 1-hydroxybenzotriazole (HOBt),[37] 1-hydroxy-7-azabenzotriazole (HOAt)[38] and ethyl-1-hydroxy-1*H*-1,2,3-triazole-4-carboxylate (HOCt),[39] **figure 1.14**.

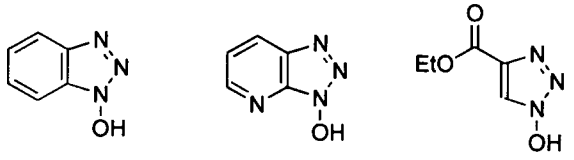


Figure 1.14: Triazole Based Coupling Reagents

The activation method of choice in Edinburgh is the *in situ* formation of the active HOCt esters using DIC. These activated amino acid derivatives have been found to be highly efficient, with minimal racemisation[40] observed. These have been used to prepare synthetic proteins of previously unattainable length, for example stromelysin catalytic domain[41] (173 amino acid residues), and deglycosylated human erythropoietin[42] (166 amino acid residues).

1.4.4 Capping Cycles

Although modern reagents regularly obtain near quantitative coupling yields, it is rare for every step to be 100%. To prevent the accumulation of incomplete peptide sequences, capping cycles are used. This method blocks any unreacted amino functionality by acylation using acetic anhydride after each coupling step, preventing any further reaction.

1.5 Problems Encountered During Synthesis

Occasionally, the synthesis of a larger polypeptide or protein sequence simply will not proceed. Generally, these difficult sequences show incomplete coupling reactions over several amino acids in a row, 5-15 residues from the resin and β-branched amino acids, such as Ile, Thr or Val, can exaggerate the poor couplings. In some cases, this cannot be explained, and may be a consequence of the sequence in question. However, in other cases, known factors play a role. For instance, there may be incompatibility between the chains of the solid support and the peptide chain, due to differences in hydrophobicity;[21] the peptide chains may aggregate with each other or with the chains of the polymer matrix[43] and finally, intermolecular β-sheet formation *via* hydrogen bonding can occur.[44] The above problems result in lower coupling efficiencies being obtained, due to the reduced nucleophilicity of the free amino functionalities, and becomes more pronounced as the peptide chain increases in length. UV monitoring can only show that the synthesis has already dropped in efficiency however, FT-IR monitoring of key frequencies may detect the formation of secondary structure elements, such as β-sheets, several steps before the yield drops.[45]

Sometimes, difficulties can be overcome by using different solvents, coupling conditions or solid supports, for example by changing to a polyamide based resin.[21] Lower loadings of the initial amino acid,[46] and using ultrasound to increase the yield of each coupling step[47] have also been found to improve the synthesis of difficult sequences. Addition of a chaotropic salt, such as LiBr, has been shown to improve the synthesis of certain difficult sequences.[48]

Finally, the incorporation of bulky groups to the growing peptide chain can prevent secondary structure formation. One method of achieving this is the incorporation of Hmb-amino acids,[49, 50] **figure 1.15**. These amino acids can be coupled with high efficiency and the Hmb group readily removed on completion of the synthesis *via* acidolysis.

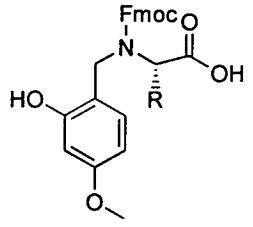


Figure 1.15: Fmoc-Amino Acid With Attached Hmb-Group

1.6 Cleavage Conditions

The exact conditions chosen for cleavage of the peptide from the solid support are dependant on the number and type of amino acids present, the sequence, the side-chain protection employed and the linker attachment. Optimum cleavage conditions can be determined using a small scale trial system of about 20-50 mg of resin bound peptide.

The linkers commonly used for Fmoc-SPPS are generally cleaved using TFA.[10, 11, 51] During the concomitant deprotection of side chain functionalities, ^{t}Bu cations are produced from the ^{t}Bu and Boc protecting groups. These can alkylate Trp, Tyr and Met residues present in the peptide chain.[52, 53, 54] As a result, a cocktail of scavengers must be added to the TFA. The most widely used cleavage mixture (TFA/water/thioanisole/phenol/EDT) is known as Reagent K.[55] Ethane-1,2-dithiol (EDT) is added to the cleavage mixture to scavenge ^{t}Bu cations and to protect Trp from acid catalysed oxidation. Thioanisole is added to prevent Met oxidation, and phenol is added to the cleavage of peptides containing multiple Trp or Tyr sequences or Arg(Pmc). Triisopropylsilane (TIS) is an odourless substitute for EDT,[56] and has also been shown to drive the reversible deprotection of trityl protected cysteine[57] to completion.

1.7 Purification and Folding

On completion of the synthesis, the correct primary sequence of amino acids must be purified from the mixture of chemically similar entities. With the well established methodology for SPPS this can often provide the first serious difficulties for the peptide chemist to overcome. Smaller peptides can be purified effectively in high yield using reversed phase high performance liquid chromatography (RP HPLC). However, larger peptides and proteins require the use of more specialised protein purification techniques including ion exchange, size exclusion and affinity chromatography methods, techniques which will be discussed in **section 3.2**.

Once the correct primary sequence has been purified, the amino acid backbone must be folded to attain its defined secondary and tertiary structure, **section 3.2.7**. This is the thermodynamically favoured three dimensional conformation of the protein, held together by relatively weak interactions including hydrogen bonds, hydrophobic interactions, salt bridges and disulfide bonds. The conformation adopted is determined by the amino acid sequence of the protein.[58] Although general and protein specific methods are available, the folding of synthetic proteins is often problematic and low yielding due to the formation of aggregate or misfolded molecules.

In general, synthetic proteins are more difficult to purify and fold than the naturally occurring or recombinant proteins simply because they are less soluble in the aqueous buffer systems preferred. This is a consequence of the protein being synthesised linearly in an organic environment with no secondary structure in place. On cleavage from the resin, a certain amount of misfolding may occur, resulting in hydrophobic surfaces of the protein being exposed and lowering the solubility if the molecule.

1.8 Fragment Coupling Strategies

Alternative routes to larger peptides and proteins are available which may avoid some of the problems encountered during stepwise construction. These involve building up the desired amino acid sequence by coupling peptide segments together.

The advantages of such a strategy are the fragments are more manageable, and may be purified and characterised prior to coupling to confirm their sequences. A greater theoretical yield of final product is possible and separation of the final product from the peptide fragments used to construct it should be relatively straightforward. Such a strategy however, may require more complicated protecting group strategies and due to sequence constraints fragments are most conveniently linked at certain sites.

Kent *et al.* have designed a number of chemoselective strategies for convergent synthesis of proteins. One highly powerful approach, referred to as native chemical ligation,[59] results in the straightforward generation of proteins with native backbone structures from fully unprotected peptide fragments. The initial step involves chemoselective reaction at an unprotected *C*-terminal thioester peptide with another unprotected fragment containing a *N*-terminal cysteine residue, **figure 1.16**. The initial product of this attack is a thioester which undergoes spontaneous intramolecular rearrangement to form the native amide bond with the regenerated cysteine side chain next to the site of ligation.

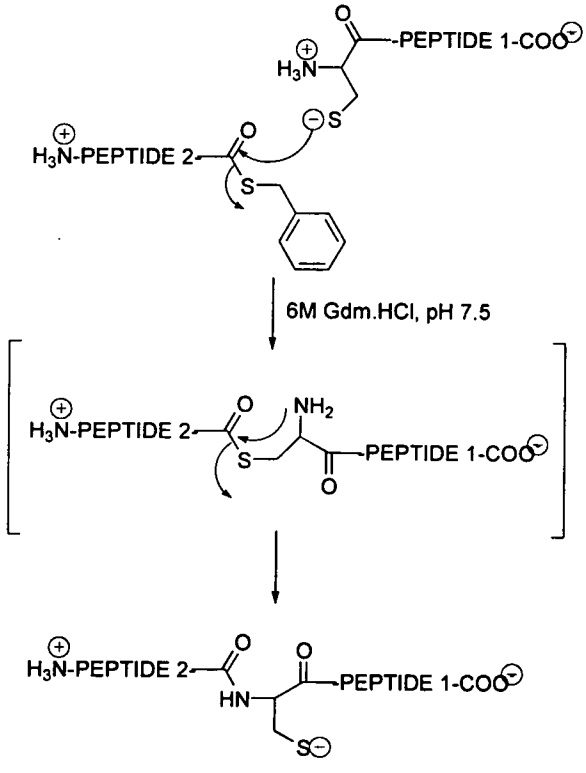


Figure 1.16: Native Chemical Ligation

Fragment condensation is also possible using traditional coupling methods and fully protected peptide fragments.

1.9 Summary and Outlook

As the previous sections have shown, the continuous development of new SPPS methodology has culminated in making the total chemical synthesis of large proteins viable. During the three and a half decades since the initial publication concerning SPPS, molecular biology has also made remarkable progress. It is now possible to clone and sequence DNA, producing a wealth of new protein information. Proteins can also be synthesised using recombinant technology.

A comparison of SPPS and recombinant technology shows that both methods play crucial, but complementary, roles in the elucidation of protein structure, function and biology. Compared to the traditional methods of isolation of the protein from natural sources, both methods can offer much larger yields of purified protein.

The strengths of SPPS lie in the complete control over amino acid sequence produced, which can sometimes prove ambiguous *via* genetic engineering. SPPS also allows incorporation of unnatural amino acids and NMR probe nuclei which are not routinely available *via* recombinant techniques. Finally, SPPS also permits *de novo* design of peptides with specific secondary structural units, predetermined conformation or tailor-made chemical or biological function.

1.10 References

1 V. du Vigneaud, C. Ressler, J. M. Swan, C. W. Roberts, P. G. Katsoyannis & S. Gordon, *J. Am. Chem. Soc.*, **1953**, *75*, 4879-4880.

2 V. du Vigneaud, C. Ressler, J. M. Swan, C. W. Roberts, P. G. Katsoyannis, *J. Am. Chem. Soc.*, **1954**, *76*, 3115-3121.

3 R. B. Merrifield, *J. Am. Chem. Soc.*, **1963**, *85*, 2149-2154.

4 M. Bergmann & L. Zervas, *Ber. Dtsch. Chem. Ges.*, **1932**, *65*, 1192-1201.

5 L. A. Carpino, *J. Am. Chem. Soc.*, **1957**, *79*, 98-101.

6 F. C. McKay & N. F. Albertson, *J. Am. Chem. Soc.*, **1957**, *79*, 4686-4690.

7 R. B. Merrifield, *J. Am. Chem. Soc.*, **1964**, *86*, 304-305.

8 R. B. Merrifield, *Biochemistry*, **1964**, *3*, 1385-1390.

9 L. A. Carpino & G. Han, *J. Org. Chem.*, **1972**, *37*, 3404-3409.

10 C.-D.Chang & J. Meienhofer, *Int. J. Peptide Protein Res.*, **1978**, *11*, 246-249.

11 E. Atherton, H. Fox, D. Harkiss C. J. Logan & R. C. Sheppard, *J. Chem. Soc., Chem. Commun.*, **1978**, 537-539.

12 K. M. Otteson, R. L. Noble, P. D. Hoeprich, K. T. Shaw & R. Ramage, *Applied Biosystems Research News*, **1993**, 1-12.

13 C.-D. Chang, M. Waki, M. Ahmad, J. Meienhofer, E. O. Lundell & J. D. Haug, *Int. J. Peptide & Protein Res.*, **1980**, *15*, 59-66.

14 P. White in "*Peptides: Chemistry, Structure & Biology*", Proc. 12th Am. Peptide Symp., J. A. Smith & J. E. Rivier eds., ESCOM, Leiden, **1992**, p537.

15 R. Ramage & J. Green, *Tetrahedron Lett.*, **1987**, *28*, 2287-2290.

16 J. Green, O. M. Ogunjobi, R. Ramage & A. S. J. Stewart, *Tetrahedron Lett.*, **1988**, *29*, 4341-4344.

17 L. A. Carpino, H. Schroff, S. A. Triolo, E.-S. M. E. Mansour, H. Wenschuh & F. Albericio, *Tetrahedron Lett.*, **1993**, *34*, 7829-7832.

18 P. Sieber & B. Riniker, *Tetrahedron Lett*, **1991**, *32*, 739-742.

19 P. Sieber & B. Riniker, *Tetrahedron Lett.*, **1987**, *28*, 6031-6034.

20 L. Zervas & D. M. Theodoropoulos, *J. Am. Chem. Soc.*, **1956**, *78*, 1359-1362.

21 R. C. Sheppard in "*Peptides 1971*", H. Nesvadba ed., North Holland Publishing, Amsterdam, **1973**, pp111-125.

22 E. Atherton, D. L. J. Clive & R. C. Sheppard, *J. Am. Chem. Soc.*, **1975**, *97*, 6584-6585.

23 E. Bayer & W. Rapp in "*Chemistry of Peptides and Proteins*", W. Voelter, E. Bayer, Y. A. Ouchinikov & V. T. Ivanov eds., de Gruyter, New York, **1986**, *3*, pp3-7.

24 M. Meldal, *Tetrahedron Lett.*, **1992**, *33*, 3077-3080.

25 S. S. Wang, *J. Am. Chem. Soc.*, **1973**, *95*, 1328-1333.

26 M. Mergler, R. Tanner, J. Gosteli, & P. Grogg, *Tetrahedron Lett.*, **1988**, *29*, 4005-4008.

27 K. Barlos, D. Gatos, S. Kapolos, G. Papaphotiu, W. Schafer & Y. Wenqing, *Tetrahedron Lett.*, **1989**, *30*, 3947-3950.

28 G. Briephol, J. Knoll & W. Stüber, *Tetrahedron Lett.*, **1987**, *28*, 5651-5654.

29 R. Ramage, S. L. Irving & C. McInnes, *Tetrahedron Lett.*, **1993**, *34*, 6599-6602.

30 H. Rink, *Tetrahedron Lett.*, **1987**, *28*, 3787-3790

31 P. Sieber, *Tetrahedron Lett.*, **1987**, *28*, 2107-2110.

32 F. Albericio & G. Barany, *Int. J. Peptide Protein Res.*, **1987**, *30*, 206-216.

33 R. Ramage, C. A. Barron, S. Bielecki & D. W. Thomas, *Tetrahedron Lett.*, **1987**, *28*, 4105-4108.

34 J. P. Tam, R. D. Dimarchi & R. B. Merrifield, *Int. J. Peptide Protein Res.*, **1980**, *16*, 412-425.

35 J. C. Sheehan & G. P. Hess, *J. Am. Chem. Soc.*, **1955**, *77*, 1067-1068.

36 D. Sarantakis, J. Teichman, E. L. Lien & R. L. Fenichel, *Biochem. Biophys. Res. Commun.*, **1976**, *73*, 336-342.

37 W. Konig & R. Geiger, *Chem. Berichte*, **1970**, *103*, 788-798.

38 L. A. Carpino, A. El-Faham, C. A. Minor & F. Albericio, *J. Chem. Soc., Chem. Commun.*, **1994**, 201-203.

39 L. Jiang, A. Davison, R. Ramage & G. Tennant, *Tetrahedron*, **1998**, *54*, 14233-14254.

40 N. Robertson, L. Jiang & R. Ramage, *Tetrahedron*, **1999**, *55*, 2713-2720.

41 G. H. Morton, *Ph. D. Thesis*, The University of Edinburgh, **1997**

42 N. Robertson & R. Ramage, *J. Chem. Soc., Perkin Trans. 1*, **1999**, 1015-1021.

43 J. Bedford, C. Hyde, T. Johnson, W. Jun, D. Owen, M. Quibel & R. C. Sheppard, *Int. J. Peptide Protein Res.*, **1992**, *40*, 300-307.

44 C. M. Deber, M. K. Lutek, E. P. Heimer & A. M. Felix, *Pept. Res.*, **1989**, *2*, 184-188.

45 B. Henkel & E. Bayer, *J. Pept. Sci.*, **1998**, *4*, 461-470.

46 S. B. H. Kent & R. B. Merrifield in "Peptides 1980", K. Brunfeldt ed., Scriptor, Copenhagen, **1981**, pp328-333.

47 J. Vágner, P. Kocna & V. Krhnák, *Pept. Res.*, **1991**, *4*, 284-288.

48 A. Thaler, D. Seebach & F. Cardinaux, *Helv. Chim. Acta*, **1991**, *74*, 617-627.

49 T. Johnson, M. Quibell, D. Owen & R.C. Sheppard, *J. Chem. Soc., Chem. Commun.*, **1993**, 369-372.

50 T. Johnson & M. Quibell, *Tetrahedron Lett.*, **1994**, *35*, 463-466.

51 E. Atherton, H. Fox, D. Harkiss & R. C. Sheppard, *J. Chem. Soc., Chem. Commun.*, **1978**, 539-540.

52 B. F. Lundt, N. L. Johansen, A. Vølund & J. Markussen, *Int. J. Peptide Protein Res.*, **1978**, *12*, 258-268.

53 B. Erikson & R. B. Merrifield, *J. Am. Chem. Soc.*, **1973**, *95*, 3750-3756.

54 P. Sieber, B. Riniker, M. Brugger, B. Kamber & W. Rittel, *Helv. Chim. Acta*, **1970**, *53*, 2135-2150.

55 D.S. King, G. C. Fields & G. B. Fields, *Int. J. Peptide Protein Res.*, **1990**, *36*, 255-256.

56 D. A. Pearson, M. Blanchette, M. L. Baker, C.A. Guindon, *Tetrahedron Lett.*, **1989**, *30*, 2739-2742.

57 I. Photaki, J. Taylor-Papadimitriou, C. Sakarellos, P. Mazarakis & L. Zervas, *J. Chem. Soc. (C)*, **1970**, 2683-2687.

58 C. B. Anfinsen, *Science*, **1973**, *181*, 223-230.

59 P.E. Dawson, T. W. Muir, I. Clark-Lewis & S. B. H. Kent, *Science*, **1994**, *266*, 776-779.

Chapter 2

The Solid Phase Synthesis of Peptide *C*-Terminal Semicarbazones and Aldehydes

2.1 Introduction

2.1.1 Biological Role of Peptide Derivatives

Many biological receptors are activated by peptides but, in general, they do not make attractive drug candidates. The major obstacles in the application of peptides as clinically useful drugs is their poor biomembrane penetration, rapid enzymatic degradation and short biological half lives. A possible approach to solve these delivery problems is derivatisation of the peptides to produce transport forms which are more lipophilic, less polar and more soluble in organic media than the parent peptides and are capable of protecting the peptide against degradation be enzymes present in the mucosal barrier or in the blood. It is, therefore, desirable to incorporate structural features into a peptide molecule which enhance lipid solubility and lower the polarity. It is postulated that incorporation of a *C*-terminal semicarbazone moiety could achieve this and consequently aid transport through lipid membranes, across the blood brain barrier (BBB) and into the central nervous system (CNS).

Other *C*-terminal modifications can prevent enzymatic degradation. For example, peptide *C*-terminal aldehydes (PAs) are an important class of transition state analogues, which have been extensively studied since they were first discovered as natural products.[1] PAs of various different structures have been found to be potent inhibitors of many enzymes implicated in a wide range of disease states. Proteolytic enzymes have been found to be the most susceptible to inhibition.

A widely studied PA is leupeptin[1] (N-acetyl-L-leucyl-L-leucyl-DL-arginal), **figure 2.1.1** which was the first PA to be isolated.

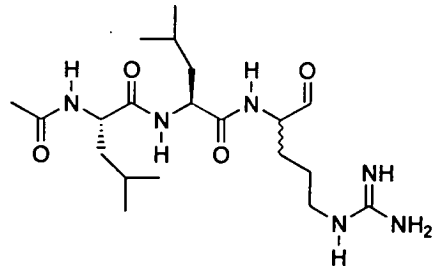


Figure 2.1.1: Leupeptin

Leupeptin has been shown to alter or suppress the symptoms of disease conditions such as rheumatoid arthritis,[2] muscular dystrophy,[3] allergic encephalomyelitis[4] and malaria.[5] It produces these effects *via* potent inhibition if a number of proteolytic enzymes. However, it is not selective among enzymes of similar substrate specificities, thus limiting its usefulness in the investigations of disease processes and as a therapeutic agent. These observations prompted the study of analogues of leupeptin and other peptide aldehydes as potential selective protease inhibitors.

Although the precise mode of action of PAs is unknown, a hydrolysis mechanism has been proposed,[6, 7] **figure 2.1.3**.

Due to the structural similarities between PAs and the natural substrate, the former can also participate in the hydrolysis mechanism of **figure 2.1.3**. However, an analogue of the usual tetrahedral intermediate will be formed,[8, 9] **figure 2.1.2**, which cannot participate any further in the hydrolysis mechanism.

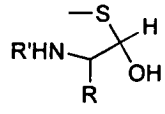


Figure 2.1.2: Tetrahedral Intermediate Generated From PA

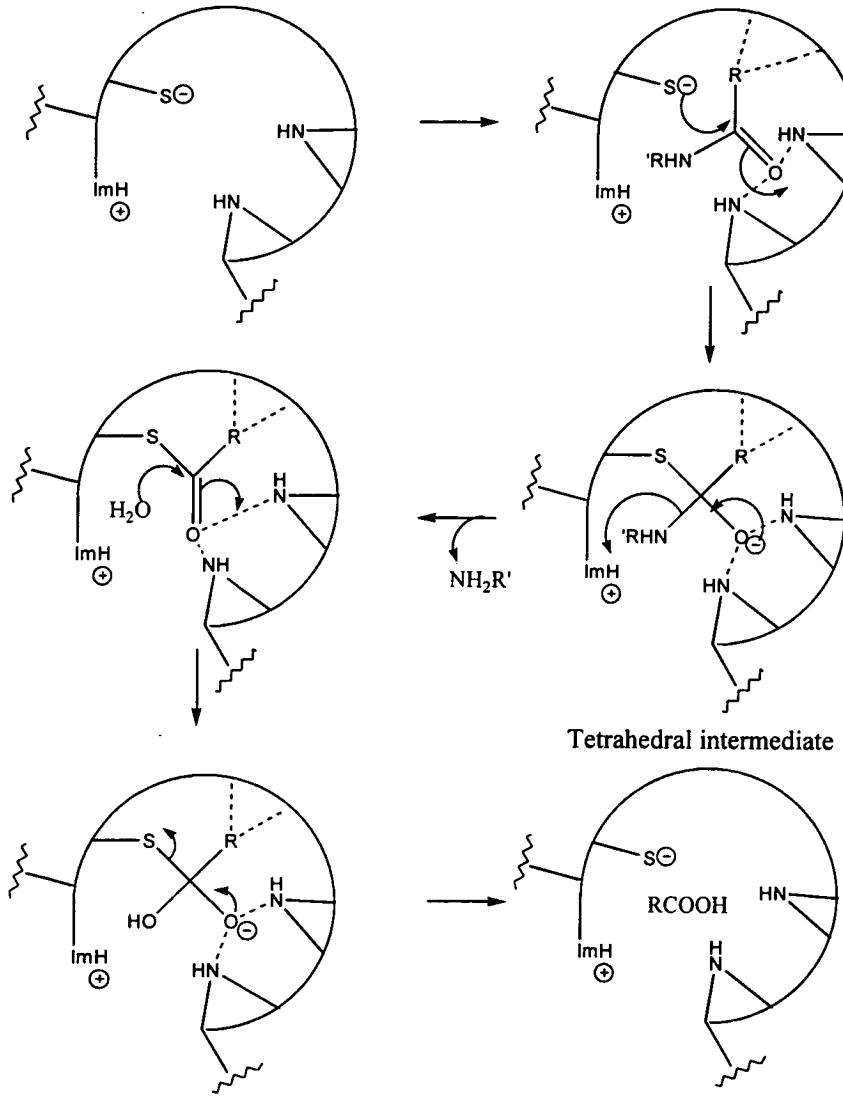


Figure 2.1.3: Proposed mechanism of protease hydrolysis[6, 7]

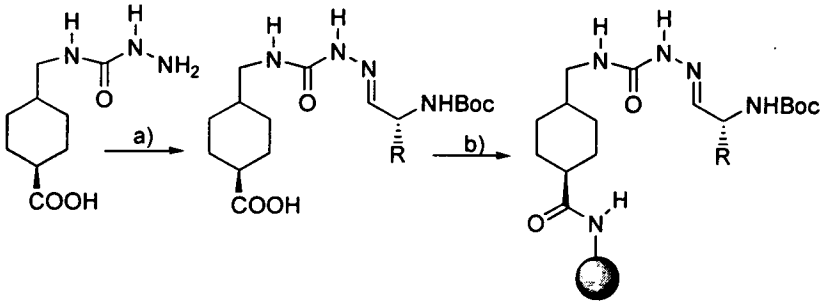
PAs have also been used in chemoselective peptide ligation,[10, 11] as affinity ligands for the purification of enzymes such as proteinases[12, 13] and for the synthesis of reduced peptide bond isosteres.[14, 15]

2.1.2 Solid Phase Synthesis of PAs

Until recently, only a few examples of the solid phase synthesis of PAs had been reported. However, in the last three years many new examples have been described.

The first example, by Webb,[16] relies on the protection of the aldehyde functional group as a stable semicarbazone. This strategy involves the synthesis of a linker

molecule, which is reacted with the aldehyde of the *C*-terminal residue, and is finally coupled onto the solid support, **figure 2.1.4**.

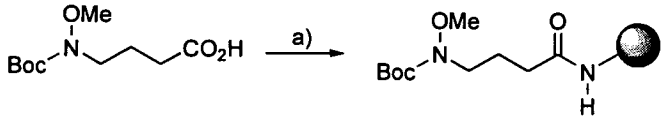


a) Boc-amino aldehyde/NaOAc/EtOH, b) MBHA resin

Figure 2.1.4: Webb Linker For PA Synthesis

Following Boc-SPPS, the free PAs are produced using catalytic hydrogenation. If protected PAs are required, cleavage using dilute aqueous acid/formaldehyde can be used.

Fehrentz and Martinez have developed several routes[17, 18, 19] to PAs on solid support. The first utilises a methyl hydroxylamine linker, which forms a Weinreb amide[20] moiety on loading of the *C*-terminal residue. The linker was prepared and attached to MBHA resin, **figure 2.1.5**, and chain elongation can be achieved using either Boc or Fmoc SPPS strategies, after deprotection of the initial Boc group.



a) MBHA resin, activation

Figure 2.1.5: Fehrentz Linker For PA Synthesis I

Reaction with lithium aluminium hydride after synthesis of the peptide yields the desired PAs. This linker also has the advantage that treatment with Grignard reagents produces peptide ketones.[17] The amount of lithium aluminium hydride used in cleavage must be increased with increasing chain length, and as a result the PAs produced may be contaminated with residual lithium salts, which may be toxic to sensitive biological assays. However, the method does have other advantages, in that

it can be used to generate large PAs with intact side-chain protection, which may be used for the chemical ligation of peptide fragments.

In 1997, Fehrentz and Martinez[18] published a second route to PAs using a phenyl ester linkage, **figure 2.1.6**.

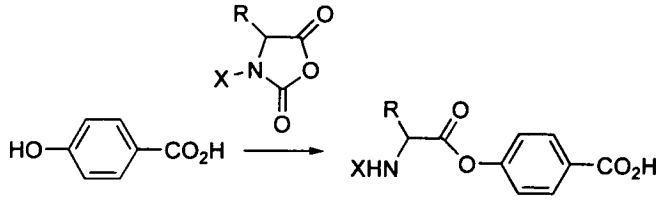
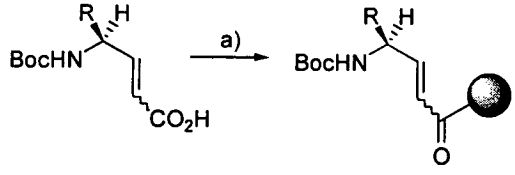


Figure 2.1.6: Fehrentz Linker For PA Synthesis II

Again, the linker is prepared and attached to MBHA resin. After SPPS, the PA can be released from the solid support using $LiAl(O^tBu)_3H$; however, over reduction to the alcohol has been observed.

A linker suitable for cleavage using ozonolysis has also been reported,[19] **figure 2.1.7**.

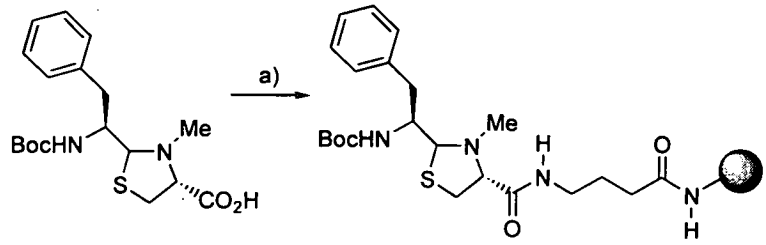


a) Merrified Resin, Cs_2CO_3

Figure 2.1.7: Fehrentz Linker For PA Synthesis III

The PAs isolated were found to be pure, and did not appear to have epimerised. This method has also been found to be suitable for sequences containing Asp or Glu residues. The synthesis of this linker was later modified[21] to eliminate the need to synthesise a different linker for each *C*-terminal amino acid.

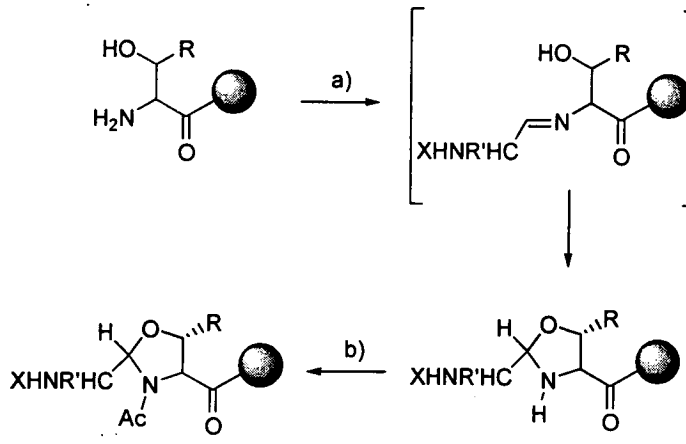
Galeotti *et al*[22] have used thiazolidine analogues of amino acids as the building blocks for the preparation of peptidyl aldehydes on solid phase, **figure 2.1.8**. For this method, cleavage is performed using a mixture of $CuCO_3$ and $CuCl_2$ in MeCN/water/DMF.



a) MBHA resin

Figure 2.1.8: Galeotti Linker For PA Synthesis

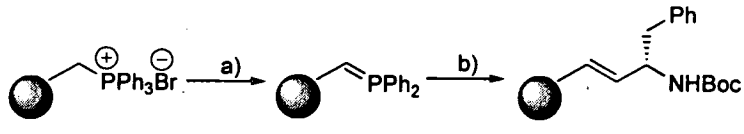
Ede and Bray[23] attached threonine to a solid support, which when treated with a dilute solution of the *N*-protected amino aldehyde, gives an imine intermediate. This can spontaneously cyclise to the stable oxazolidine moiety, which can then be elongated by SPPS, **figure 2.1.9**. Cleavage is achieved using TFA and AcOH. Omission of the TFA step yields protected PA fragments.



a) XNHR'CHCHO b) Ac_20

Figure 2.1.9: Ede Linker For PA Synthesis

Hall and Sutherland[24] have developed an olefinic linker using Wittig chemistry. This can be cleaved using ozonolysis and a reductive work up, **figure 2.1.10**.

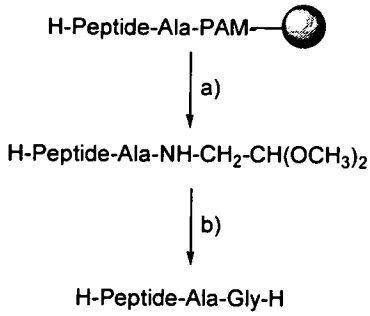


a) NaHMDS/THF b) Boc-amino aldehyde/THF

Figure 2.1.10: Hall Linker For PA Synthesis

This linker has been used for the combinatorial synthesis of an array of 27 tripeptide aldehydes.

Finally, Lelièvre and coworkers[25] have reported a method of generating PAs using Fmoc-^{t}Bu chemistry and the acid and base stable PAM linker, **figure 2.1.11**.



a) aminoacetaldehyde-dimethylacetal b)TFA

Figure 2.1.11: Lelièvre Synthesis of PAs

The advantages of this method are that no special linker is required, and the PA products are obtained in good yield (almost 100%). A masked aldehyde is generated which avoids the difficult purification of PAs, and the starting materials used are all cheap and commercially available. Finally this method generates PAs which can be used in fragment condensation. As this method was designed with the synthesis of fragments in mind, only *C*-terminal glycine aldehydes have been synthesised to date.

2.1.3 Research Overview

Previously, McInnes[26] synthesised the suberone molecule, **figure 2.1.12** which serves as the core for a number of different linker systems.

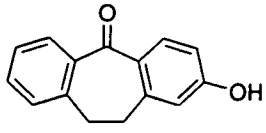
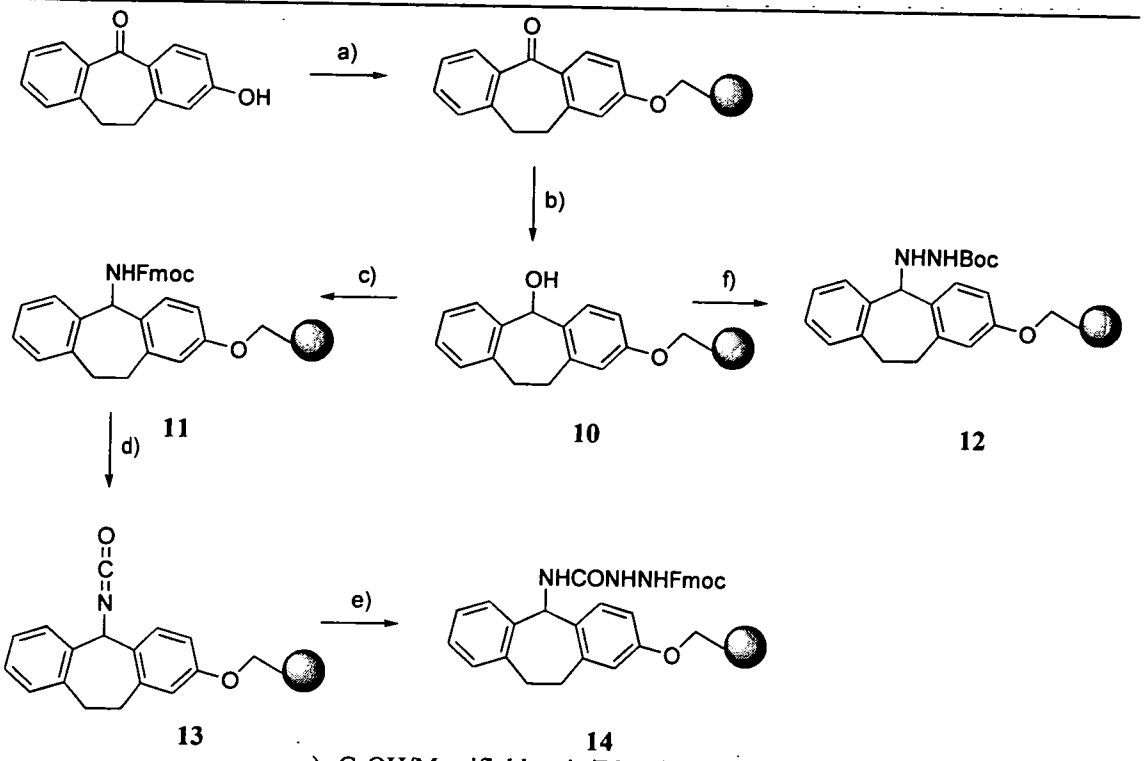


Figure 2.1.12: Suberone Linker

This precursor can be coupled to Merrifield resin, and the functionality manipulated in a few straightforward steps (**figure 2.1.13**) to give a series of linkers suitable for the SPPS of peptide derivatives.



a) CsOH/Merrifield resin/DMF b) $LiBH_4$/THF
b) $FmocNH_2$/$PhSO_3H$/DMF d) i) 20% piperidine/DMF
ii) DIEA, Triphosgene, DCM e) Fmoc-$NHNH_2$, DCM
f) $BocNHNH_2$, DIEA, DCM.

Figure 2.1.13: Synthesis of Linkers For Peptide Derivative Synthesis

The linkers shown in **figure 2.1.13** have been successfully used to prepare *C*-terminal peptide amides, **11**, hydrazides,[27] **12**, and aza-glycine peptides, **14**.[28] When linker **10**, **figure 2.2.13**, is used to synthesise a peptide chain with a *C*-terminal acid, the peptide-linker bond is very acid labile and can be cleaved in coupling steps by the weakly acidic HOBt.[28] This results in loss of peptide from the support, lowering the yield of peptide obtained.

Recently, Briggs[29] extended the usefulness of the semicarbazide linker, **14, figure 2.1.13**, further by demonstrating that an aldehyde, in this case dichlorobenzaldehyde, can be loaded successfully onto the linker by imine formation. Cleavage of the resin using TFA:water, purification and ozonolysis recovered the starting dichlorobenzaldehyde.

In this study, the suitability of the linker, **14, figure 2.1.13**, for the synthesis of PAs was assessed. The route adopted proceeds *via* the peptide *C*-terminal

semicarbazones, which may have interesting biological properties due to increased lipophilicity.

2.2 Results and Discussion

2.2.1 Synthesis of Fmoc-Hydrazine

The synthesis of the semicarbazone linker, **14**, requires Fmoc-hydrazine to be prepared in advance, due to the inherent instability of the isocyanate intermediate, **13**, **figure 2.2.13**, necessitating immediate trapping. Previously, Irving[28] used a route starting from the Fmoc-*N*-hydroxysuccinimide, **figure 2.2.1**. However, difficulties were encountered using this route, due to the instability of the product during column chromatography.

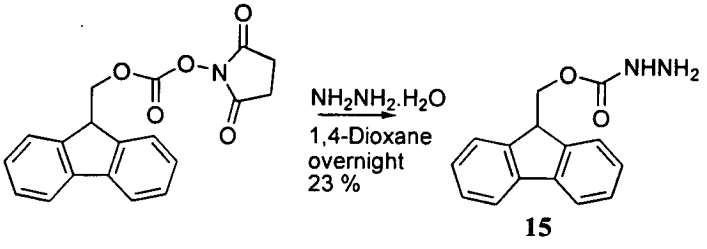


Figure 2.2.1: Synthesis of Fmoc-hydrazine I

Zhang and co-workers have synthesised Fmoc-hydrazine, **15**, for use as a fluorescent label[30] using Fmoc-chloroformate, **figure 2.2.2**. This method was found to be more effective, and was subsequently adopted.

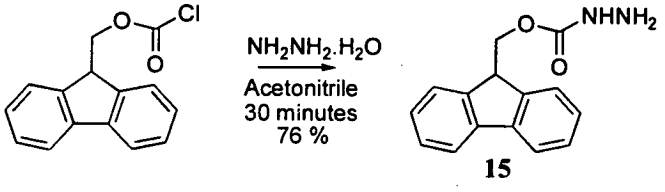


Figure 2.2.2: Synthesis of Fmoc-hydrazine II

2.2.2 Synthesis of the Linker

The linker, **14**, was synthesised as depicted in **figure 2.1.13**; from commercially available amide linker, **10**, however, the level of Fmoc-loading obtained was variable. Variation in temperature was found to have the greatest effect on the success of the linker synthesis. It was found that maintaining the reaction at constant temperature was essential. The parameters in each step, isocyanate generation and trapping the isocyanate with Fmoc-hydrazine, **15**, were then optimised. The results are depicted in **tables 2.2.1** and **2.2.2**.

Solvent	Base	Time	Conc(ml/g resin)	Eq of Triphosgene	Eq of Base
Toluene	DIEA	30 mins	10 ml	1 eq	1 eq
47 %	50 %	50 %	48.5 %	40 %	50 %
DCM	N,N-Dimethylaniline	1 hour	20 ml	2 eq	2 eq
50 %	45 %	56 %	50 %	48 %	46%
	2,4,6-Collidine	2 hours	30 ml	3 eq	
	40 %	40 %	40 %	50 %	
				4 eq	
				50 %	

Table 2.2.1: Parameters Affecting Isocyanate Generation[¶]

The conditions adopted for this step were sonication of the resin in DCM (20ml/g of resin), containing 1 equivalent of DIEA and 3 equivalents of triphosgene for one hour. Isocyanate formation could be monitored using FT-IR.

[¶] Loading levels in mmol/g are expressed as a percentage of the original resin loading level in mmol/g.

Solvent	Base	Time	Conc (ml/g resin)	Eq of Fmoc-$NHNH_2$
Toluene	With	30 mins	10 ml	1 eq
37 %	30 %	41 %	52 %	35 %
DCM	Without	1 hour	20 ml	2 eq
50 %	50 %	49 %	50 %	42 %
		2 hours	30 ml	3 eq
		50 %	55 %	50 %
		4 hours		4 eq
		50 %		48 %

Table 2.2.2: Parameters Affecting Isocyanate Trapping[†]

The optimum conditions found for trapping of the isocyanate species were sonication with 3 equivalents of Fmoc-hydrazine in DCM, in the absence of base, for 2 hours. Complete consumption of the isocyanate intermediate was checked using FT-IR, and the level of Fmoc-loading obtained measured *via* UV determination. Combination of these two optimised steps routinely produced resin loadings in the order of 0.2 - 0.3 mmol/g.

2.2.3 Synthesis of Fmoc-Amino Aldehydes

Having established the linker synthesis, it was necessary to synthesise the Fmoc-amino aldehydes which provide the *C*-terminal residue in any peptides synthesised.

α-Amino aldehydes are generally colourless crystals or oils, which are unstable chemically and configurationally, particularly in solution. As a result, they are generally used immediately after isolation. Ideally, purification should be avoided as this may induce racemisation. Exposure to silica is only possible if the aldehyde is first converted into the more stable semicarbazone derivative,[31] or if 0.1% pyridine is added to the eluent[32] to prevent racemisation.

Many synthetic routes to amino aldehydes have been devised, however, not all of these are suitable for the synthesis of derivatives of the naturally occurring amino acids. The methods that are suitable generally use the required amino acids as

[†] Loading level in mmol/g is expressed as a percentage of the original resin loading level in mmol/g.

starting materials, in order to incorporate the desired stereochemistry into the resulting amino aldehydes.

Methyl and ethyl esters may be reduced with DIBAL,[31] **figure 2.2.3**, and active amides, including imidazolidines and 3,5-dimethyl pyrazolidines, can be reduced to the aldehyde using lithium aluminium hydride.[33, 34]

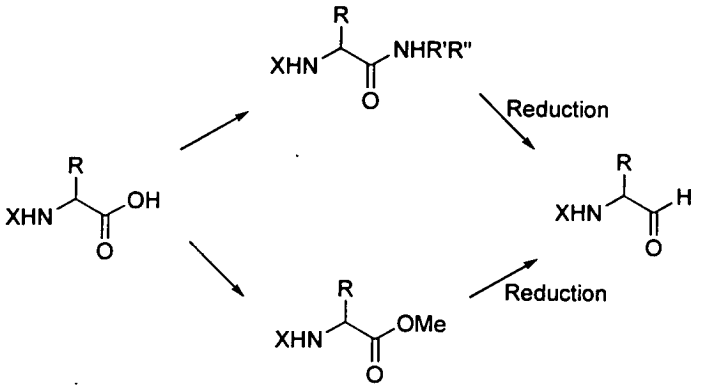


Figure 2.2.3: Synthesis of Amino Aldehydes *via* Reductive Methods

Oxidation/reduction procedures can also be employed, **figure 2.5.2**. α-Amino alcohols can be produced by borane-tetrahydrofuran reduction of the N^{α}-protected amino acids,[35] or by sodium borohydride reduction of the methyl ester.[36] Selective oxidation of the alcohol moiety to produce the α-amino aldehyde has then been achieved using Collins reagent (CrO_3/py),[37] DMSO activated with various reagents (SO_3.py,[38] oxalyl chloride,[39] TFAA[40] or DCC[41]), pyridinium chlorochromate[42] or pyridinium dichromate.[35]

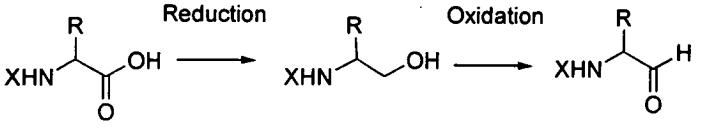
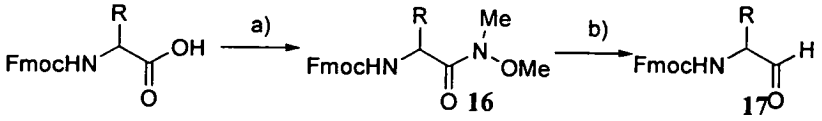


Figure 2.2.4: Synthesis of Amino Aldehydes *via* Reduction/Oxidation

The approach adopted was to form the Weinreb amide,[20] and subsequently reduce this species to the desired amino aldehyde, **figure 2.2.5**. This method was chosen because it is compatible with Fmoc-protecting group stategies, and does not cause

racemisation of the amino aldehyde.[43, 44] There is no evidence of this method causing over-reduction to the corresponding alcohol.



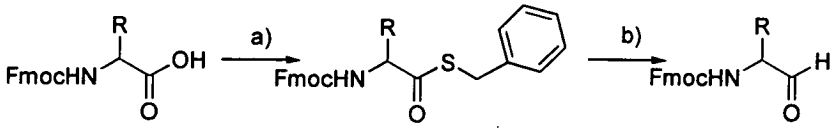
a) *N,O*-dimethylhydroxylaminehydrochloride/Ethylchloroformate/*N*-methylpiperidine/DCM
b) $LiAlH_4$/THF

Figure 2.2.5: Synthesis of Amino Aldehydes *via* Weinreb Amide

The Weinreb amides, **16**, were prepared from the corresponding Fmoc-protected amino acids in good yield. The purity of the compounds was checked by t.l.c. and analytical RP-HPLC, and each was estimated to be greater than 95 % pure, making chromatographic purification unnecessary.

The reduction to the aldehyde, **17**, was carried out using 1.3 equivalents of lithium aluminium hydride, except in the case of Fmoc-Asp(O^{t}Bu)-H,[45] which was prepared using only 1 equivalent of reducing agent.

Originally, an attempt was made to synthesise this amino acid derivative using the benzylthioester method of Ho and Ngu,[32] **figure 2.2.6** as this method has been reported to be sufficiently mild not to reduce the ester functionality of the side chain. However, the reduction was never observed to go to completion.



a) α-thiotoluene/DMAP/DCC/THF b) Triethylsilane/Pd-C/THF

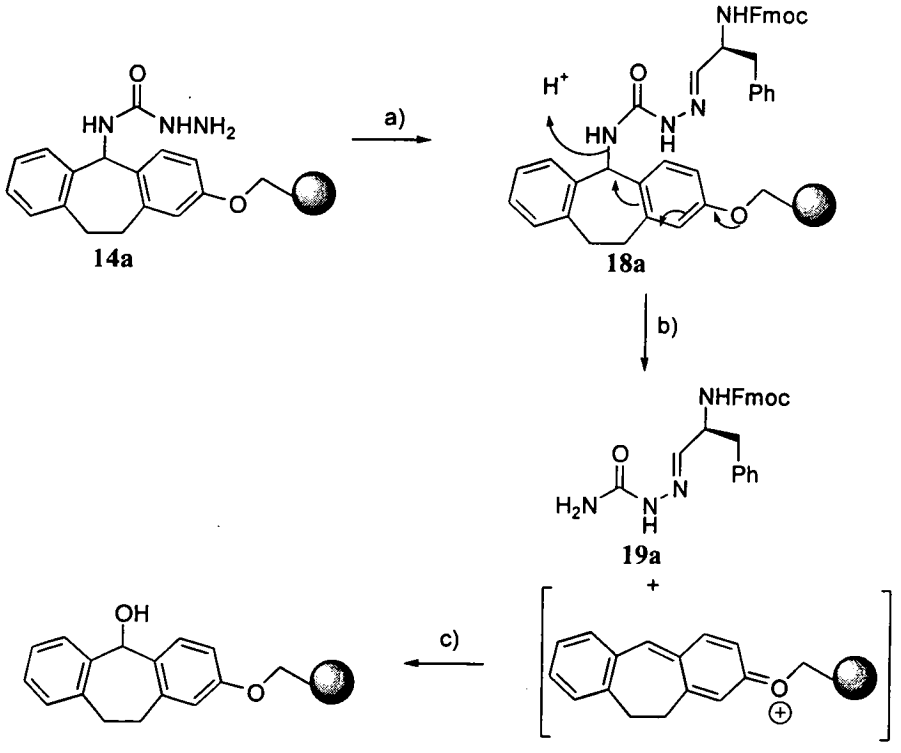
Figure 2.2.6: Synthesis of Amino Aldehydes *via* Benzylthioesters

In general, it was observed that the melting points measured for the Fmoc-amino aldehydes were not in agreement with the literature values recorded.[32, 44, 46] However, all other analytical data indicated that the correct molecules had been prepared.

2.2.4 Single Residue Studies

2.2.4.1 Fmoc-Phenylalaninal Semicarbazone

The most significant problem of PA synthesis, is avoiding epimerisation of the *C*-terminal (aldehydic) residue. In order to establish if racemisation occurred using the semicarbazide approach, a number of single residue studies were carried out. In these studies, Fmoc-Phe-H was used, as it is known to be highly susceptible to racemisation.[31] The amino aldehyde was loaded onto the linker, **14a**, in the presence of DIEA using sonication, followed by cleavage from the solid support and analysis, **figure 2.2.7**.



a) Fmoc-Phe-H, DIEA; DCM, b) TFA/H_2O; c) H_2O

Figure 2.2.7: Loading and Cleavage of The Linker

The material obtained, **19a**, was compared directly with that obtained by the synthesis of Fmoc-phenylalaninal semicarbazone, **19b**. The analytical data obtained was identical in all respects, with the exception of melting point and optical rotation measurements. These differences were attributed to the different work up procedures used. While the solution phase sample, **19b**, was precipitated, washed with ethyl

acetate and dried under vacuum, the solid phase sample, **19a**, was dissolved in aqueous acetonitrile and lyophilised. The freeze drying could have resulted in the incorporation of associated water molecules and TFA salts into this sample, hence, affecting the melting point and optical rotation measurements.

An alternative approach was adopted, with Fmoc-phenylalaninal semicarbazone, **19b**, prepared in solution, exposed to the conditions employed for cleavage of the samples from the resin (TFA/water for approximately 1.5 hours). Following precipitation from hexane, the sample, **19c**, was compared to that of the starting material, **19b**, and found to be identical, indicating that the cleavage conditions do not induce racemisation, **table 2.2.3**.

Compound		$[\alpha_D]$ (c g/100ml, DMF)	Mpt (^{0}C)
Fmoc-Phenylalaninal Semicarbazone (Solution Phase)	**19b**	-24.7 ^{0} (0.288)	144-145
Fmoc-Phenylalaninal Semicarbazone (Solid Phase)	**19a**	-10.0 (1.04)	136-137
Fmoc-Phenylalaninal Semicarbazone (TFA treated)	**19c**	-24.0 ^{0} (0.325)	143-145

Table 2.2.3: Comparison of Fmoc-Phenylalaninal Semicarbazone Samples

2.2.4.2 Pyruvic Acid Exchange

Following cleavage of peptide semicarbazones from the linker and subsequent purification, it is necessary to unmask the *C*-terminal aldehyde. There are many methods of converting semicarbazone molecules into aldehydes. Fehrentz and Martinez[19] have reported the use of ozone to cleave a PA linked to the solid support *via* a carbon-carbon double bond. However, although these conditions may be used to generate the PA from the semicarbazone, they are considered to be too harsh to expose peptide samples to. A number of milder reagents have been used to hydrolyse carbon-nitrogen double bonds, for example, phthalic anhydride,[47] levulinic acid/HCl,[48] HCl/formaldehyde[49] and aqueous acetic acid.[50]

The method used in this project was pyruvic acid exchange,[51] which has the advantage that it works efficiently due to a combination of effects. Both acid catalysed hydrolysis and carbonyl exchange trapping are possible. The reaction can

also be carried out in the presence of organic solvents, or simply by using excess pyruvic acid or water as solvent.

To test the suitability of this method, Fmoc-phenylalaninal semicarbazone produced in solution, **19b**, was hydrolysed back to the starting aldehyde, **20**, and compared to the compound produced by reduction of the Weinreb amide, **17a**, **table 2.2.4**.

Compound		$[\alpha_D]$ (c g/100ml, DMF)	Mpt (^{0}C)
Fmoc-Phenylalaninal (Reduction of Weinreb Amide)	**17a**	-43.4^0 (1.146)	100-102
Fmoc-Phenylalaninal (Pyruvic Acid Exchange)	**20**	-41.8^0 (0.467)	102-103

Table 2.2.4: Comparison of Fmoc-Phenylalaninal Samples

The results clearly indicate that pyruvic exchange does not appear to induce racemisation, and therefore was suitable for applying to peptide samples produced.

2.2.5 Synthesis of Test Peptides

2.2.5.1 Loading The Fmoc-Amino Aldehydes Onto The Linker

Fmoc-amino aldehydes prepared for the synthesis of test peptides were found to load onto the linker, in the presence of DIEA, in good yields based on starting and ending Fmoc-loading levels, **table 2.2.5**.

Loaded Compound		Loading Level (%)	Loading Time (Hours)
Fmoc-(*L*)Ala-H	**18b**	100	5
Fmoc-(*D*)Ala-H	**18c**	70	5
Fmoc-Phe-H	**18a**	90	5
Fmoc-Trp-H	**18d**	100	4
Fmoc-Asp(O^{t}Bu)-H	**18e**	90	5

Table 2.2.5: Loading of Amino Aldehydes[‡]

[‡] The obtained loading level in mmol/g is expressed as a percentage of the previous Fmoc-loading level in mmol/g.

2.5.5.2 Synthesis of Fmoc-Phe-Val-(L)Ala-H and Fmoc-Phe-Val-(D)Ala-H

The first sequence to be synthesised was that of Phe-Val-Ala, as the Boc-protected PA of this sequence was reported to be very stable and therefore easy to handle and purify.[18] Both the *L*- and *D*- analogues of the *C*-terminal residue were incorporated into the sequence to further investigate possible epimerisation. Fehrentz[52] has reported that examination of the aldehydic signal of a PA molecule containing three or more residues could be used to determine the extent of racemisation. If a single signal was observed in the aldehydic region, no racemisation had occurred. The syntheses of both tripeptides proceeded well, providing sufficient quantities of crude Fmoc-protected peptide semicarbazones for purification (RP HPLC), analysis and conversion to the PAs.

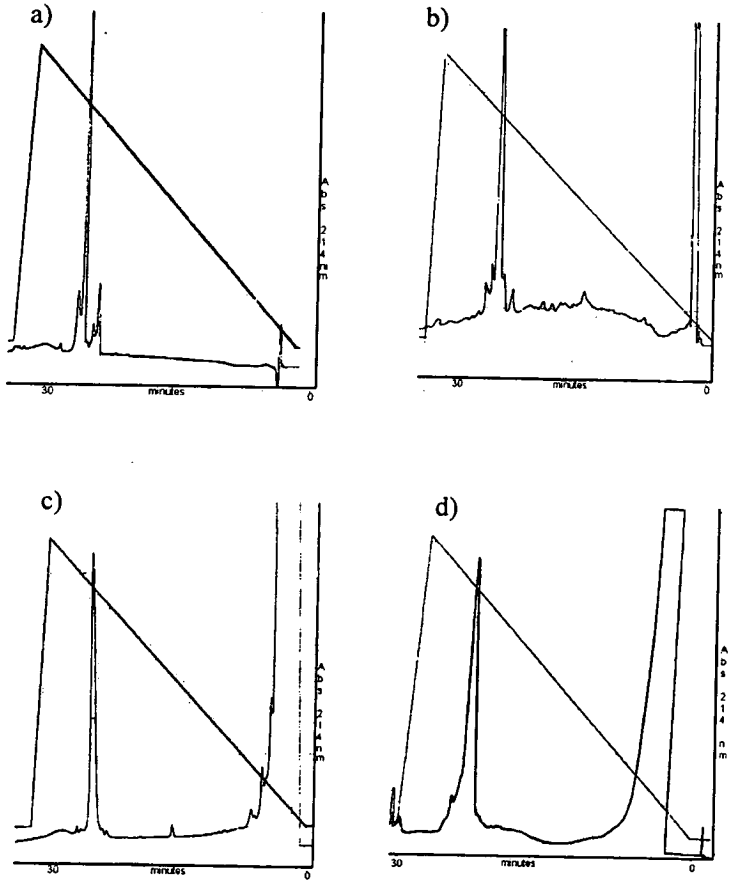


Figure 2.2.7: HPLC Analysis of a) Fmoc-Phe-Val-(L)Alaninal semicarbazone, **21**, b) Fmoc-Phe-Val-(D)Alaninal semicarbazone, **23**, c) Fmoc-Phe-Val-(L)Ala-H, **22**, d) Fmoc-Phe-Val-(D)Ala-H, **24**.

The Fmoc-protected peptide derivatives were obtained in reasonable levels of purity, as indicated by analytical RP HPLC, **figure 2.2.7**.

The peptide derivatives were analysed by MALDI-TOF mass spectrometry and amino acid analysis, **table 2.2.6**.

Sequence	Yield (%)	Mass (Found)	Mass (Requires)	AAA (24 Hours)
Fmoc-FV(*L*)A-**sc 21**	25	599.21 (MH^+)	599.71	$Phe_1$1.00, $Val_1$1.00
Fmoc-FV(*D*)A-**sc 23**	24	599.37 (MH^+)	599.71	$Phe_1$0.91, $Val_1$1.09
Fmoc-FV(*L*)A-H **22**	50	542.61 (MH^+)	542.65	$Phe_1$1.16, $Val_1$0.86
Fmoc-FV(*D*)A-H **24**	56	654.29 ($M^+CF_3CO_2^-$)	654.66	$Phe_1$1.13, $Val_1$0.87

Table 2.2.6: Analytical Data For Test Peptide Derivatives

(The suffix **-sc** has been adopted to depict a peptide *C*-terminal semicarbazone.)

The PA sequences were examined by 600 MHz NMR, and in each case, only one aldehydic signal was observed. The signals for each analogue, *L*- and *D*- were observed to exist at distinct chemical shifts, **figure 2.2.8**.

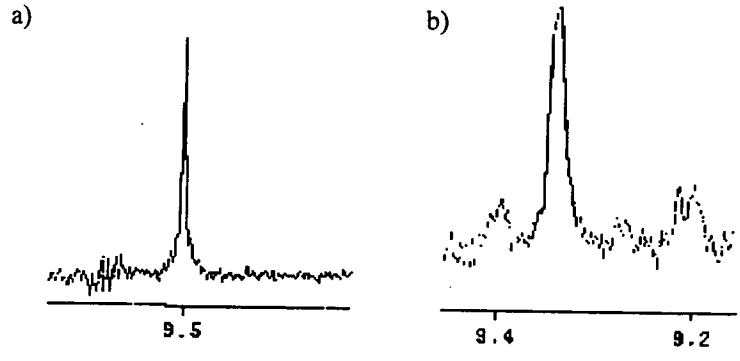


Figure 2.2.8: NMR Signals Of Aldehydic Signals For Test Peptides
a) Fmoc-Phe-Val-(*L*)-Ala-H, **22**, b) Fmoc-Phe-Val-(*D*)-Ala-H, **24**.

The studies carried out on the above test peptides effectively demonstrated that epimerisation of the *C*-terminal residue does not occur using this new method of PA

synthesis. It has also demonstrated that should epimerisation occur, the NMR method described by Fehrentz[52] would indicate this effectively.

2.2.5.3 Cleavage Studies

On cleavage of Fmoc-Phe-Val-(*L*)Alaninal semicarbazone from the linker, three components were observed by HPLC, **figure 2.2.9**. The components were identified as acylated *C*-terminal residue (peak 1), Fmoc-Phe-Val-(*L*)Ala-**sc**, **21** (peak 2) and Fmoc-Phe-Val-(*L*)Alaninal, **22** (peak 3).

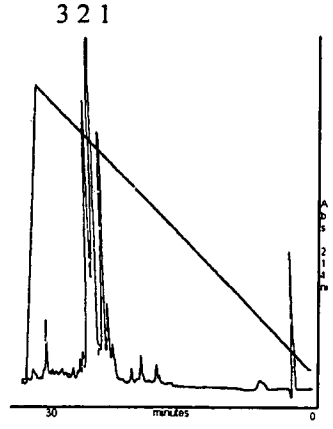


Figure 2.2.9: HPLC Profile of Crude Fmoc-Phe-Val-Alaninal Semicarbazone

Due to the apparent conversion of the semicarbazone to the PA during cleavage from the resin, a cleavage study was initiated. This was to ensure that complete conversion of the semicarbazone to the aldehyde did not occur during longer cleavage procedures (*ca.* 4-6 hours), which would be required if a sequence contained multiple arginine residues protected with the Pmc or Pbf groups (**section 1.2.2**). A sample of resin-bound Fmoc-Phe-Val-(L)Alaninal semicarbazone was stirred overnight in TFA/water (9:1). At timed intervals the mixture was examined by HPLC. The observations are given in **table 2.2.7**.

Time	Observation
30 min	peaks not resolved
1 hour	peaks not resolved
2 hours	semicarbazone predominant peak
3 hours	semicarbazone predominant peak
4 hours	equal amounts of semicarbazone and aldehyde
24 hours	equal amounts of semicarbazone and aldehyde

Table 2.2.7: Cleavage Study

From the table it would appear that, even after a 24 hour cleavage, it is still possible to isolate a reasonable quantity of peptide semicarbazone. This is advantageous, as these compounds are generally easier to purify and store than the corresponding PAs, and may be biologically interesting in their own right.

2.5.5.4 Synthesis of Further Sequences

Three further peptide sequences were studied by this method to demonstrate the versatility of the route. The first sequence synthesised was Fmoc-Gly-Ala-Lys-Gly-Phenylalaninal. This sequence was used to ensure that free amino groups, such as those in the side chain of the lysine residue, would not condense with the *C*-terminal aldehyde moiety. No problems were encountered with this peptide when maintained at acidic pH. The next sequence synthesised was Fmoc-His-Leu-Asp-Ile-Ile-Tryptophanal. This sequence is a fragment of the sequence of Endothelin,[53] and may prove to be an inhibitor of the Endothelin Converting Enzyme (ECE).[54] Finally, a known inhibitor[55, 56] of Caspase C was synthesised to test whether *C*-terminal aspartic acid aldehyde could be incorporated and also to see if the methodology could be extended to include longer sequences. These peptide derivatives were purified and analysed as previously, and the data obtained is summarised in **table 2.2.8** and **appendix A**.

Each of the PAs synthesised were examined by 600 MHz NMR, and were observed to have only one aldehydic proton signal at the chemical shifts given in **table 2.2.8**.

Compound	Yield (%)	Mass (Found)	Mass (Required)	AAA (24 Hours)	NMR (CHO)
Fmoc-GAKGF-sc **25**	40	763.46 $(M-H)Na^+$	763.83	$Gly_2$1.86, $Ala_1$1.09, $Lys_1$0.98	-
Fmoc-GAKGF-H **26**	62	761.74 $(M-H)K_2^+$	761.98	$Gly_2$1.86, $Ala_1$1.10, $Lys_1$0.97	9.34 ppm
Fmoc-HLDIIW-sc **27**	27	1082.04 (MNa^+)	1082.23	$Asp_1$1.03, $Ile_2$0.96, $Leu_1$1.09, $His_1$0.88	-
Fmoc-HLDIIW-H **28**	38	779.66 $(MH^+$-Fmoc)	779.94	$Asp_1$1.04, $Ile_2$0.96, $Leu_1$1.09, $His_1$0.88	9.29 ppm
Ac-AAVALLPAVL LALLAPDEVD-sc **29**	29	2079.52 (MNa^+)	2079.41	$Asp_1$1.14, $Glu_1$1.02, $Pro_2$1.92, $Ala_6$5.68, $Val_3$2.80, $Leu_6$5.99	-
Ac-AAVALLPAVL LALLAPDEVD-H **30**	50	1998.16 $(M-H)^-$	1998.35	$Asp_1$1.07, $Glu_1$1.05, $Pro_2$1.90, $Ala_6$5.61, $Val_3$3.10, $Leu_6$5.80	9.36 ppm

Table 2.2.8: Analytical Data For Peptide Derivatives

2.2.6 Summary and Outlook

The linker has been shown to be extremely effective for the synthesis of peptide *C*-terminal semicarbazones and aldehydes. The route proceeds in reasonable yield, and studies have shown there to be no loss of stereochemical integrity.

The methodology could be extended to include the synthesis of non-PAs. Peptide *C*-terminal ketones could be synthesised by loading amino ketones onto the linker in place of the amino aldehydes.

2.3 References

1 T. Aoyagi, T. Takeuchi, A. Matsuzaki, K. Kawamura, S. Kondo, M. Hamada, K. Maeda & H. Umezawa, *J. Antibiot.*, **1969**, *22*, 283-286.

2 T. Aoyagi, T. Takeuchi, A. Matsuzaki, K. Kawamura, S. Kondo, M. Hamada, K. Maeda & H. Umezawa, *J. Antibiot.*, **1969**, *22*, 183-186.

3 T. Aoyagi, T. Wada, Y. Ishikawa, F. Kojima, M. Nagai, T. Asani, Y. Nagai & H. Umezawa, *Exp. Neurol.*, **1984**, *84*, 326-337.

4 T. Aoyagi, T. Wada, M. Nagai, H. Sakaguchi, T. Osani, Y. Nagai & H. Umezawa, *Experimentia*, **1984**, *40*, 1405-1407.

5 L. W. Scheibel, E. Bueding, W. R. Fish & J. T. Hawkins, *Prog. Clin. Biol. Res.*, **1984**, *155*, 131-136.

6 K. Brocklehurst, F. Willenbrock & E. Salih in "*New Comprehensive Biochemistry*", A. Neuberger & K. Brocklehurst eds., Elsevier, Amsterdam, **1987**, vol *16*, pp36-158.

7 A. Storer & R. Ménard, *Methods Enzymol.*, **1994**, *244*, 486-500.

8 C. A. Lewis & R. Wolfenden, *Biochemistry*, **1977**, *16*, 4890-4895.

9 M. P. Gamcsik, J. P. G. Malthouse, W. U. Primrose, N. E. Mackenzie, A. S. F. Boyd, R. A. Russell & A. I. Scott, *J. Am. Chem. Soc.*, **1983**, *105*, 6324-6325.

10 J. P. Tam & J. C. Spetzler, *Biomedical Peptides, Proteins & Nucleic Acids*, **1995**, *1*, 123-132.

11 C. F. Lui, C. Rao & J. P. Tam, *J. Am. Chem. Soc.*, **1996**, *118*, 309-312.

12 A. Basak, X. W. Yuan, N. G. Seidah, M. Chretien & C. Lazure, *J. Chromatography*, **1992**, *581*, 17-29.

13 A. H. Patel, A. Ahsan, B. P. Suthar & R. M. Scultz, *Biochim. Biophys. Acta*, **1983**, *748*, 321-330.

14 J. Martinez, J. P. Bali, M. Rodriguez, B. Castro, R. Magous, J. Laur & M. F. Lignon, *J. Med. Chem.*, **1985**, *28*, 1874-1879.

15 Y. Sasaki & D. Coy, *Peptides*, **1987**, *8*, 119-121.

16 A. M. Murphy, R. Dagnino, P. L. Vallar, A. J. Trippe, S. L. Sherman, R. H. Lumpkin, S. Y. Tamura & T. R. Webb, *J. Am. Chem. Soc.*, **1992**, *114*, 3156-3157.

17 J.-A. Fehrentz, M. Paris, A. Heitz, J. Velek, C.-F. Lui, F. Winternitz & J. Martinez, *Tetrahedron Lett.*, **1995**, *36*, 7871-7874.

18 J.-A. Fehrentz, M. Paris, A. Heitz, J. Velek, F. Winternitz & J. Martinez, *J. Org. Chem.*, **1997**, *62*, 6792-6796.

19 C. Pothion, M. Paris, A. Heitz, L. Rocheblave, F. Rouch, J.-A. Fehrentz & J. Martinez, *Tetrahedron Lett.*, **1997**, *38*, 7749-7752.

20 S. Nahm & S. M. Weinreb, *Tetrahedron Lett.*, **1981**, *22*, 3815-3818.

21 M. Paris, A. Heitz, V. Guerlavais, M. Christau, J.-A. Fehrentz, & J. Martinez, *Tetrahedron Lett.*, **1998**, *39*, 7287-7290.

22 N. Galeotti, M. Giraud, & P. Jouin, *Lett. Peptide Sci.*, **1997**, *4*, 437-440.

23 N. J. Ede & A. M. Bray, *Tetrahedron Lett.*, **1997**, *38*, 7119-7122.

24 B. J. Hall & J. D. Sutherland, *Tetrahedron Lett.*, **1998**, *39*, 6539-6596.

25 D. Lelièvre, H. Chabane & A. Delmas, *Tetrahedron Lett.*, **1998**, *39*, 9675-9678.

26 C. McInnes, *Ph. D. Thesis*, The University of Edinburgh, **1990**.

27 R. Ramage, S. L. Irving & C. McInnes, *Tetrahedron Lett.*, **1993**, *34*, 6599-6602.

28 S. Irving, *Ph. D. Thesis*, The University of Edinburgh, **1994**.

29 K. L. Briggs, *Personal Communication*.

30 R.-E. Zhang, Y.-L. Cao & M. W. Hearn, *Anal. Biochem.*, **1991**, *195*, 160-170.

31 A. Ito, R. Takahashi & Y. Baba, *Chem. Pharm. Bull.*, **1975**, *23*, 3081-3087.

32 P. T. Ho & K.-Y. Ngu, *J. Org. Chem.*, **1993**, *58*, 2313-2316.

33 H. Zemlicka & M. Murata, *J. Org. Chem.*, **1976**, *41*, 3317-3321.

34 R. Nishizawa, T. Saino, T. Takita, H. Suda, T. Aoyagi & H. Umezawa, *J. Med. Chem.*, **1977**, *20*, 510-515.

35 C. F. Stanfield, J. E. Parker & P. Kanellis, *J. Org. Chem.*, **1981**, *46*, 4797-4800.

36 Y. Hamada & T. Shiori, *Tetrahedron Lett.*, **1982**, *23*, 1193-1196.

37 K. E. Rittle, C. F. Homnick, G. S. Ponticello & B. E. Evans, *J. Org. Chem.*, **1982**, *47*, 3016-3018.

38 Y. Hamada & T. Shiori, *Chem. Pharm. Bull.*, **1982**, *30*, 1921-1924.

39 J. R. Luly, J. F. Dellaria, J. J. Plattner, J. L. Soderquist & N. Li, *J. Org. Chem.*, **1987**, *52*, 1487-1492.

40 W. R. Ewing, B. D. Harris, K. L. Bhat & M. M. Joullie, *Tetrahedron*, **1986**, *42*, 2421-2428.

41 H. Seki, K. Koga & S. Yamada, *Chem. Pharm. Bull.*, **1972**, *20*, 361-367.

42 S. G. Pyne, M. J. Hensel & P. L. Fuchs, *J. Am. Chem. Soc.*, **1982**, *104*, 5719-5728.

43 J.-A. Fehrentz & B. Castro, *Synthesis*, **1983**, 676-678.

44 J. J. Wen & C. M. Crews, *Tetrahedron Asymmetry*, **1998**, *9*, 1855-1858.

45 M. Paris, C. Pothion, A. Heitz, J. Martinez & J.-A. Fehrentz, *Tetrahedron Lett.*, **1998**, *39*, 1341-1344.

46 J.-A. Fehrentz, C. Pothion, J.-C. Califano, A. Loffet & J. Martinez, *Tetrahedron Lett.*, **1994**, *35*, 9031-9034.

47 A. E. Gillam & T. F. West, *J. Chem. Soc.*, **1945**, 95-98.

48 C. H. De Puy & B. W. Ponder, *J. Am. Chem. Soc.*, **1959**, *81*, 4629-4631.

49 M. P. Cava, R. L. Little & D. R. Napier, *J. Am. Chem. Soc.*, **1958**, *80*, 2257-2263.

50 D. Taub, R. D. Hoffsommer, H. L. Slates, C.H. Kuo & N. L. Wendler, *J. Am. Chem. Soc.*, **1960**, *82*, 4012-4026.

51 E. B. Hershberg, *J. Org. Chem.*, **1948**, *13*, 542-546.

52 J.-A. Fehrentz, A. Heitz & B. Castro, *Int. J. Peptide Protein Res.*, **1985**, *26*, 236-241.

53 M. Yanagisawa, A. Inoue, T. Ishikawa, Y. Kasuya, S. Kimura, S.-I. Kumagaye, K. Nakajima, T. X. Watanabe, S. Sakakibura, K. Goto, T. Masaki, *Proc. Natl. Acad. Sci., USA*, **1988**, *85*, 6964-6967.

54 T. Sawamura, S. Kimura, O. Shinmi, Y. Sugita, M. Yanagisawa, K. Goto & T. Masaki, *Biochem. Biophys. Res. Commun.*, **1990**, *168*, 1230-1236.

55 D. W. Nicholson, A. Ali, N. A. Thornberry, J. P. Vaillancourt, C. K. Ding, M. Gallant, Y. Gareau, P. R. Griffin, M. Labelle, Y. A. Lazebnik, N. A. Munday, S. M. Raju, M. E. Smulson, T.-T. Yamin, V. L. Yu & D. K. Miller, *Nature*, **1995**, *376*, 37-43.

56 S. C. Wright, V. Schellenberger, H. Wang, D. H. Kinder, J. Talhouk & J. W. Larrick, *J. Exp. Med.*, **1997**, *186*, 1107-1117.

Chapter 3

The Stepwise Chemical Synthesis of Deglycosylated Human IFN–γ and its Purification

3.1 Introduction

In the late 1950s, both Isaacs and Lindenmann[1] and Nagano and Kojima[2] were studying the phenomenon of viral interference. Both groups demonstrated that cells exposed to various viruses or other substances (termed inducers) respond by production of a substance which can confer resistance to other cells subsequently exposed to the same or a related virus. This substance was designated as interferon. Isaacs and Lindenmann[1] characterised interferon as a protein with species specificity but which confers protection against a broad range of viruses.

Since then, alpha, beta, gamma, delta, tau and omega interferons have been identified. Some types, in particular interferons-α and -β, have been extensively studied and are well characterised.[3] Interferons delta,[4] tau[5] and omega[6] have been discovered more recently. All types of interferons have in common antiproliferative and immunomodulatory properties, as well as the antiviral activity which led to their initial discovery.[7]

3.1.1 Interferon-Gamma

Interferon-gamma (IFN-γ) was first isolated in 1965, when Wheelock demonstrated interferon-like activity in the supernatants of mononuclear cells after exposure to mitogen.[8]

The protein was originally called immune interferon, due to both its activity and its production by competent cell types possessing immune regulatory properties.[9] The protein was also called type II interferon, as it has different physicochemical[8, 9, 10] properties, molecular mechanism[11, 12, 13] and antigenicity[10, 14, 15] from IFN-α, -β, -δ, -τ and -ω, the type I interferons. Finally, in 1980 the protein was renamed IFN-γ.[7]

3.1.2 Biological Properties of IFN–γ

IFN-γ is a glycoprotein produced during immune response by activated T-lymphocytes[16] and natural killer (NK) cells.[17] Studies have shown that IFN-γ is primarily an immunomodulating agent[18, 19] and an effective antitumour agent and inhibitor of cell growth,[13, 20, 21] with antiproliferative effects on cells 10-100 times greater than those of IFN-α or -β.[10, 13, 20, 21] However, IFN-γ has been observed to induce the antiviral state in the host much more slowly than IFN-α or -β.[22] IFN-γ has been shown to potentiate the actions of IFN-α and -β when tested with these related proteins.[11, 23]

3.1.3 Clinical Applications of IFN-γ

Human clinical trials in various infectious disease indications have resulted in the licensing of Actimmune®, a *C*-terminally modified recombinant IFN–γ molecule (rhIFN-γ1b) for reduction of the life threatening infections associated with chronic granulomatous disease (CGD).[24]

Chronic granulomatous disease is a rare, inherited, pediatric immunodeficiency state. The white blood cells of the patient are unable to function normally to kill invading bacterial or fungal infections.

Until the early 1990s, therapy for the disease involved frequent doses of parenteral antibiotic to prevent and fight acquired infections. Aggressive surgical intervention was necessary in life threatening situations.

rhIFN–γ 1b has been found to reduce the frequency and severity of infection by greater than 70 % in CGD patients.[25] The mechanism of action of rhIFN-γ1b in the disease is undetermined. On termination of rhIFN-γ1b therapy following 1-3 years treatment, no increase in the frequency or severity of infection was observed. As a result most physicians now use rhIFN-γ1b in combination with antimicrobial agents in the management of this disease. This treatment of this condition is the most successful use of this protein clinically to date.

3.1.4 Characterisation of IFN-γ

Until the advent of recombinant protein technology, the most reliable method of production of IFN-γ was the stimulation of peripheral blood lymphocytes (PBL) by antigen or mitogen. However, only a small amount of material could be produced using this method, limiting the subsequent studies that could be carried out.

The cDNA of IFN-γ was first isolated in 1982 by Gray.[26] The gene was expressed in *Esherichia Coli* and monkey cells,[26] and although it coded for 166 amino acids, 20 amino acids were assumed to make up a signal sequence, leaving the mature protein at 146 residues long, commencing with the trio of residues, Cys.Tyr.Cys.[27] The molecular weight of the sequence was calculated to be 17110 Da.

The protein sequence was confirmed by a parallel cDNA study carried out by Devos[28] and by the production of a synthetic cDNA fragment using convergent synthesis[29] which was shown to express the correct protein in *E. Coli*. The IFN-γ gene has been found to be located on human chromosome 12.[30]

The sequence was determined from the natural protein by peptide mapping,[34] using a combination of tryptic digest, amino acid analysis, mass spectrometry and *N*-terminal sequencing. The sequence was almost identical to that determined by Gray,[26] except the *N*-terminus was found to be blocked, i.e. a pyroglutamate residue, not cysteine. Thus, the mature protein is 143 amino acids in length, **figure 3.1.1**. It is not clear if the Cys.Tyr.Cys residues are removed as part of the signal sequence or as a separate event.

Gln Asp Pro Tyr Val Lys Glu Ala Glu Asn Leu Lys Lys Tyr Phe Asn Ala Gly His Ser

Asp Val Ala Asp **Asn** Gly Thr Leu Phe Leu Gly Ile Leu Lys Asn Trp Lys Glu Glu Ser

Asp Arg Lys Ile Met Gln Ser Gln Ile Val Ser Phe Tyr Phe Lys Leu Phe Lys Asn Phe

Lys Asp Asp Gln Ser Ile Gln Lys Ser Val Glu Thr Ile Lys Glu Asp Met Asn Val Lys

Phe Phe Asn Ser Asn Lys Lys Lys Arg Asp Asp Phe Glu Lys Leu Thr **Asn** Tyr Ser Val

Thr Asp Leu Asn Val Gln Arg Lys Ala Ile His Glu Leu Ile Gln Val Mat Ala Glu Leu

Ser Pro Ala Ala Lys Thr Gly Lys Arg Lys Arg Ser Gln Met Leu Phe Arg Gly Arg Arg

Ala Ser Gln

Figure 3.1.1: The Amino Acid Sequence of human IFN–γ, the sites of glycosylation are indicated in bold

The amino acid sequence of the protein was found to be consistent with the observed hydrophobicity of natural IFN-γ.[10] There are 27 basic residues and 19 acidic residues, implying the protein will have a high isoelectric point (pI). The measured pI of the natural protein was found to be 8.6-8.7.[36]

The use of recombinant DNA technology to produce IFN-γ allowed large quantities of the protein to be isolated, and in turn structural and biological characterisation of this protein.

Natural IFN-γ has been found to be heterogeneous with respect to molecular weight[31, 32, 33, 34, 35] and pI.[36] Species of 15.5, 20 and 25 kDa have been observed to be equally active.[31, 32, 33, 34, 35] The heterogeneity is due to differences in core glycosylation,[35] *C*-terminal processing[35] and multimer formation.[31]

3.1.5 The C-Terminus of IFN-γ

Natural IFN-γ has a heterogeneous *C*-terminus, due to digestion by proteolytic enzymes.[37] Six different *C*-terminal species have been detected for the natural protein,[34] and up to 13 *C*-terminal amino acids have been found to have been removed.[34] The recombinant protein has also been found to be susceptible to proteolysis.[38, 39, 40] Proteins with truncated *C*-termini have been produced for studies by limited proteolysis[41, 42] and genetic engineering.[43]

There are conflicting opinions as to the importance of the *C*-terminus for the biological activity of the protein. However, the studies are in general agreement that the structure of the molecule is not affected by the removal of up to 20 amino acids, even though the activity varies.[39, 40, 41, 42] There is a small effect on the activity on removal of the first 5 residues, removal of 5-9 residues increases the activity and, on removal of subsequent residues, activity dimishes slowly.[44, 45]

3.1.6 Glycosylation of IFN-γ

On cloning the cDNA of the protein, Gray observed two potential sites for *N*-linked glycosylation, at positions 25 and 97 (indicated in bold in **figure 3.1.1**).[26] Glycosylation has been shown to be unimportant for biological activity[34, 46, 47, 48] or protein conformation,[46] since rhIFN-γ produced in *E. Coli* also forms biologically active dimers[49, 50, 51] and the di-, mono- and deglycosylated proteins all have similar circular dichroism (CD) spectra.[46]

The glycosylated protein is, however, relatively protease resistant when compared to the deglycosylated protein, which is rapidly inactivated. The glycoside chains cover relatively large areas of the surface of the dimers, sterically hindering protease action[49, 52] and increasing the circulatory lifetime of the protein.[51]

Two different studies[48, 53] have derived variations of the following structure, **figure 3.1.2**, for the glycoside chains of IFN-γ.

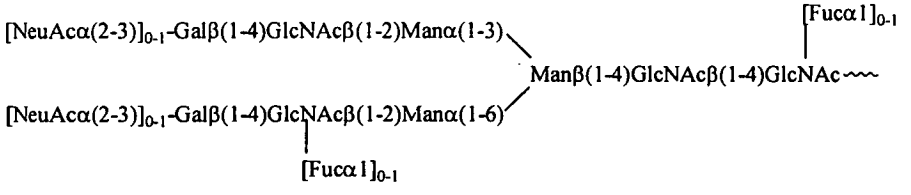


Figure 3.1.2: Proposed Glycan Structure For dhIFN-γ

3.1.7 Structure of IFN-γ

A high α-helical content was predicted for this protein using UVCD.[54, 55] This was confirmed when IFN–γ was crystallised[56] and its structure solved using x-ray diffraction, **figure 3.1.3**.[57] The protein was found to be a dimer, consisting of two identical subunits related by a 2-fold, non-crystallographic axis. Each subunit contains six α-helices, A-F (subunit 1) and A'-F' (subunit 2), which comprise 62 % of the amino acid backbone. These helices range in length from 9-21 residues. The subunits are intimately related, and held in an antiparallel fashion by a unique intertwining of the helical domains. There is no β-sheet within the subunits, or across the dimer interface.

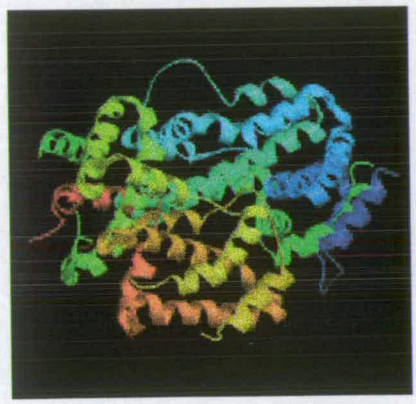


Figure 3.1.3: Ribbon Diagram of dhIFN-γ Dimeric Structure

The NMR structure of the molecule has also been determined, and is similar in all aspects, differing in the exact positions of the end points of the helices.[58]

3.1.8 IFN-γ Receptor and Receptor Binding

The IFN–γ receptor is distinct from the receptor for the type I interferons, and has been found to be species specific. The gene for the receptor is located on human chromosome 6,[59, 60] and the cDNA for the receptor has been cloned.[61] The extracellular domain of the receptor is 230 residues, is sufficient to bind IFN-γ. This

domain has been cloned, purified and named IFN-γR_{α}.[61, 62, 63] Ligand-receptor interactions were studied using IFN-γ bound to the extracellular domain, **figure 3.1.4**.

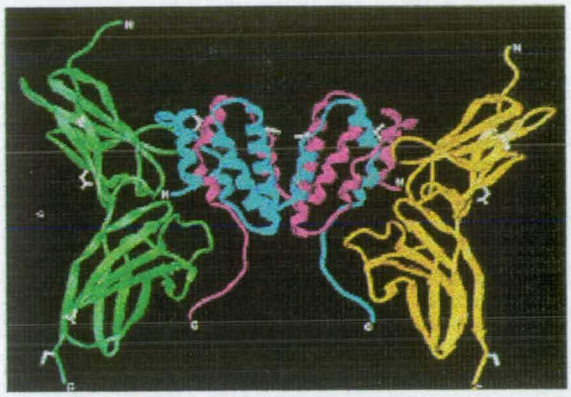


Figure 3.1.4: IFN-γ Bound to the Extracellular Domain of its Receptor

IFN–γ can cause dimerisation of the receptor, due to the inherent symmetry of the protein,[64] and a crystal structure of the 2:1 receptor:IFN-γ complex has been solved.[65] It is observed that the two IFN-γR_{α} chains are separate, and do not interact in the complex.[65]

The ligand binding surface has clusters of both acidic and basic amino acids against exposed aromatic residues. The binding interface encompases the amino terminus, helix A, the AB loop, helix B, helix F and the *C*-terminus (residues 128-132). It has been found that the residue ^{111}His is critical for maintaining the correct conformation of the protein for receptor binding.[66]

It has been observed that binding of IFN-γ to the receptor is not in itself sufficient to produce biological activity.[67, 68, 69, 70] An additional species specific protein, accessory factor 1 (IFN-$\gamma R_{\beta 1}$) is also required. The gene for this protein is located on human chromosome 21, and it has been cloned, purified and characterised.[71, 72] IFN-γ induced aggregation of IFN-γR_{α} and IFN-$\gamma R_{\beta 1}$ is sufficient to induce certain biological activities,[73] but, a third protein, as yet uncharacterised, is required to produce antiviral activity.[71, 74]

3.1.9 Research Overview

SPPS could play an important role in the story of this remarkable cytokine, supplementing the wealth of information already available. Production of IFN-γ using recombinant methods can yield post translational modifications which include heterogeneity in the *C*-terminus and sugar chains of the molecule. Such processing can not only complicate purification and analysis, but results in different research groups studying different forms of the protein, which can cause discrepencies on collation of information.

The chemical synthesis of IFN-γ using SPPS would ensure that sequences of definite length and amino acid composition could be produced for use in future studies. Currently, a programme is underway to develop methods to chemically glycosylate proteins such as IFN–γ and erythropoietin and, if successful, species with definite glycosylation patterns could be obtained.[75] These syntheses would allow the optimum form of IFN–γ to be determined and would test the limit of the currently available synthetic methodology.

As it is possible that the full sequence of IFN-γ is not necessary for biological activity, backbone epitopes could be synthesised using SPPS to probe this. Amino acid substitutions, and incorporation of unnatural amino acids such as isotopically labelled residues (for example, ^{15}N to allow further NMR studies), could also be achieved easily using SPPS.

With this in mind, a research programme was initiated to attempt to synthesise the complete amino acid backbone of IFN–γ. This initial study was intended to establish a purification protocol suitable for use in the purification of IFN-γ, and any analogues synthesised. Incorporated into this research programme was the aim to develop a generic purification protocol for synthetic proteins, one avoiding the use of monoclonal antibody affinity purification and relying solely on straightforward chromatographic procedures. Such a protocol could then be used in the purification of a wide range proteins of similar size and basicity.

The protein produced by SPPS can be directly compared to the protein produced recombinantly in *E. Coli,* since neither are glycosylated. This protein has been successfully purified; however, some protocols have relied on monoclonal antibody purification as the key step.[37, 38, 39, 76] There is literature precedent for stepwise chromatographic purification of dhIFN-γ.[45, 77] Other protocols involving a single chromatographic technique such as gel filtration[44, 78] or ion exchange chromatography[79] have also been reported.

The synthesis of dhIFN-γ was previously undertaken by Draffan.[80] The synthesis was observed to proceed well, but subsequent purification of the protein using Tbfmoc-affinity chromatography and gel filtration chromatography yielded a three component mixture, which could not be further resolved. It was anticipated that that the purification of IFN-γ could be completed using similar approaches to those detailed in the literature.

3.2 Results and Discussion

3.2.1 Stepwise Assembly of the dhIFN-γ Molecule

Prior to the synthesis of the protein, the deprotection profile, obtained previously,[80] was examined and double coupling cycles used to avoid the low yielding steps in the synthesis. The only significant drop in the assembly occurred after ^{22}Val had been coupled, 21 residues from completion of the synthesis. Therefore, with this information in hand, it was decided to proceed with the synthesis, commencing double coupling just before this residue, to overcome the drop in efficiency.

On completion of the synthesis, a deprotection profile was constructed, **figure 3.2.1**, and the synthesis was observed to have proceeded well, with no large drops in efficiency of coupling.

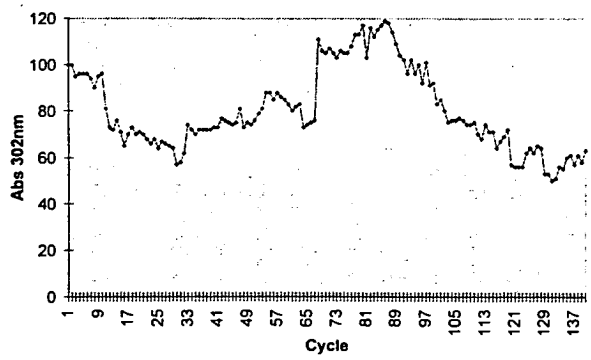


Figure 3.2.1: Deprotection Profile For Synthesis of dhIFN-γ

3.2.2 Tbfmoc-Charcoal Purification

The affinity purification of peptides using the tetrabenzo[*a, c, g, i*]fluorenyl-17-methoxycarbonyl (Tbfmoc) group[81] has been developed and applied to the purification of a number of synthetic proteins, for example deglycosylated human erythropoietin,[82] and ubiquitin.[83] This large, essentially planar group, **figure 3.2.2**, can be introduced at the *N*-terminus of a resin bound peptide or protein *via* the chloroformate, prior to cleavage from the solid support. Once in solution, the affinity of the Tbfmoc-group for carbon-based supports can be exploited to facilitate a primary purification of the peptide or protein. Another advantage of this system is the Tbfmoc-group has an isobestic point at 364 nm and increases the hydrophobicity of the sequence. This shifts the Tbfmoc-containing component away from the impurity molecules by HPLC as well as simplifying identification of the Tbfmoc-bearing component of the mixture.

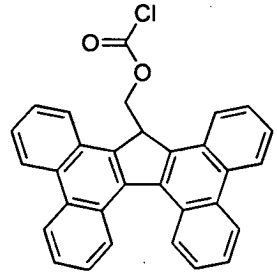


Figure 3.2.2: Tbfmoc-chloroformate

The carbon support initially investigated was porous graphitised carbon (PGC).[81] Although this support effectively separated Tbfmoc-labelled peptides from *N*-terminally capped deletion sequences, it had the disadvantage that the large quantities of PGC required to carry out each purification made the method very expensive. Activated animal charcoal was investigated as an alternative support and found to be less expensive and equally effective.

Tbfmoc-charcoal affinity purification was chosen as the primary purification step for dhIFN-γ, since Draffan[80] had found this step to be successful in removing most truncated sequences from the desired protein sequence. Hence, the resin bound Fmoc-dhIFN-γ sequence was capped, swelling the resin in DCM to ensure that all reactive functionalities which were not exposed when the resin was swollen in DMF, were blocked. The amino terminus was then deprotected, and the Tbfmoc group introduced as its chloroformate. The level of Tbfmoc-loading was checked using a method analogous to that used to check Fmoc-loading, but with UV analysis at 364 nm.

3.2.2.1 Cleavage Of Tbfmoc-dhIFN-γ From The Solid Support

At this point, a trial cleavage was carried out to investigate the optimum time required for complete removal of the protein from the solid support and the deprotection of the side chain functional groups. The cleavage was carried out on 50 mg of resin using the normal cleavage cocktail. At 30 minute intervals, a sample of the cleavage solution was removed, and the protein isolated by precipitation. HPLC analysis of the protein pellets obtained showed there to be no change in the profile after 4.5 hours. This is in good agreement with the results found by Draffan.[80]

The cleavage was then scaled up, the protein precipitated from cold diethyl ether, isolated by centrifugation and lyophilised to yield material for use in the Tbfmoc-affinity purification step. At this stage the crude Tbfmoc-dhIFN-γ was observed to contain many components by RP HPLC, **figure 3.2.3**.

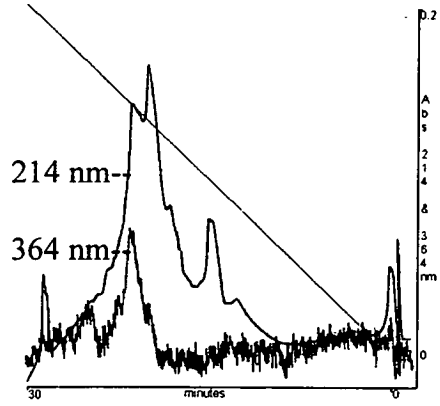


Figure 3.2.3: HPLC Profile of Crude Tbfmoc-dhIFN-γ

3.2.2.2 Tbfmoc-Charcoal Affinity Purification

Previously the charcoal system had been observed to be unstable to the basic cleavage conditions employed to release the purified peptide. This resulted in the production of fines, and contamination of the protein with a brown coloured species which could not be separated using centrifugation. As a result, it was necessary to thoroughly wash the charcoal, with all solutions employed in the purification prior to use, in an attempt to avoid contamination. The Tbfmoc-charcoal purification was carried out as depicted in **figure 3.2.4**, with monitoring by RP HPLC at 214 and 364 nm.

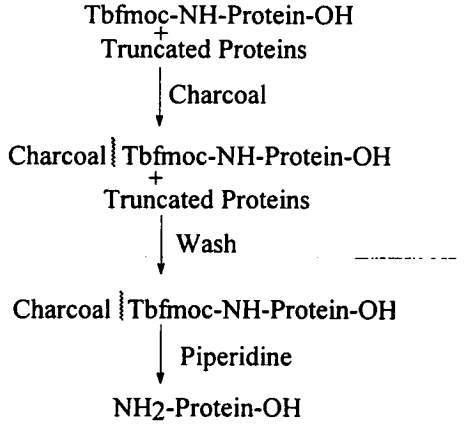


Figure 3.2.4: Tbfmoc-Charcoal Affinity Purification

Due to the hydrophobic nature of the IFN–γ sequence in general, and the extra hydrophobicity introduced to the system by the Tbfmoc-moiety, the purification in

the past had required the use of a 1:1 mixture of 6M guanidine hydrochloride and isopropyl alcohol to aid the solubility of the crude Tbfmoc-dhIFN-γ.[80] Problems have been experienced during removal of guanidine hydrochloride from protein samples due to interaction of the charged components with charged functionalities present in the side chains of the protein. To avoid this problem, urea solutions were used in place of the guanidine hydrochloride. It is usual to require an 8M solution of urea to solubilise proteins; however it was found that 6M urea completely dissolved the crude Tbfmoc-dhIFN-γ samples. On completion of the purification protocol, the protein solution was neutralised to pH 7 using acetic acid and examined by HPLC, **figure 3.2.5**.

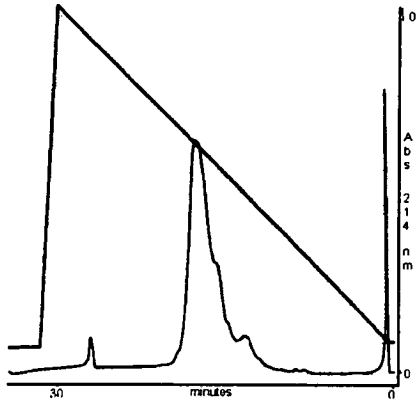


Figure 3.2.5: HPLC Profile after Tbfmoc-Charcoal Purification

The next step in the purification was FPLC size exclusion chromatography, which requires a highly concentrated protein solution to achieve the good resolution of protein components. Thus, several methods were investigated to determine the optimum method to yield a concentrated protein solution or solid protein. Firstly, ultrafiltration was tried, however, it was observed that in this case the protein was not retained by the membrane, hence no concentration was achieved.

The next method investigated involved the slow dialysis of the solution against an acetic acid solution. This effectively removed the urea and piperidine acetate from the protein solution, replacing them with acetic acid solution, which allowed the sample to be lyophilised to a solid. This meant that solutions of the required

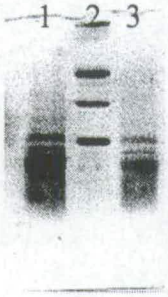
concentration for FPLC size exclusion chromatography could be prepared by redissolving the sample accordingly.

The final method investigated involved elution of the protein sample from a Sephadex G-50 column with acetic acid solution. Again, this effectively removed the urea and piperidine acetate salts rendering lyophilisation possible. When compared to dialysis of the sample, it was found that the desalting column allowed recovery of twice as much protein. This was attributed to the protein adhering to the dialysis membrane. The desalting column was also much faster than the dialysis; the process could be run overnight compared to one week, therefore this was adopted as the method of choice for future work.

3.2.2.3 Desalting Using a Sephadex G-50 Size Exclusion Column

The solution from Tbfmoc-charcoal affinity chromatography was loaded onto a Sephadex G-50 column equilibriated with 25 % acetic acid solution. Fractions eluted from the column were examined by UV for protein content. An aliquot of each UV active fraction was freeze dried before being examined by SDS-PAGE for protein content. No resolution of the protein bands was observed using this column. This was probably due to the large volume of solution loaded.

The protein material at this stage was observed to contain three components by SDS-PAGE, **figure 3.2.6**. When compared to the results obtained previously for this protein,[80] it can be seen that an improvement has occurred, since four components were obtained at this stage in the previous purification protocol.



Lane 1:	Crude Tbfmoc-dhIFN-γ
Lane 2:	Molecular Weight Markers (43, 25 & 13.7 kDa)
Lane 3:	dhIFN-γ after Tbfmoc-Charcoal Purification/Desalting

Figure 3.2.6: SDS-PAGE Profile Following Tbfmoc-Charcoal Purification

It was surprising that the impurity components persisted after the Tbfmoc-affinity purification protocol, since this method was designed specifically to remove them. There are several possibilities for their persistence.

Firstly, the acetylated truncates may not have been successfully washed away during the protocol, due to their binding to the carbon support *via* aromatic side chain functionalities of which there are 25 in the whole sequence. Indeed, the region ^{50}Val to ^{60}Phe contains 5 aromatic residues which on adopting an appropriate conformation would produce a hydrophobic surface.

Secondly, deletion molecules may have been produced during the synthesis whereby the *N*-terminus of the growing peptide chain has become buried due to changes in the resin swelling, and not participated in subsequent cycles of the synthesis. A further change in resin swelling would expose the free *N*-terminus, allowing it to continue in the synthesis, but effectively creating a molecule missing several residues from the middle of the sequence would be produced. This would result in several protein components all becoming tagged with the Tbfmoc-label on completion of the synthesis.

Finally, the protein may have sheared during acidiolytic cleavage. If this occurred at positions close to the *C*-terminus, large fragments would be produced which contained the Tbfmoc-label at the *N*-terminus. It has also been postulated that cleavage *via* attack of the softly nucleophilic thiols used as scavengers in the cleavage cocktail may be observed for susceptible sequences.[84] This possibility was ruled out by repeating the cleavage of the protein, omitting the thiol scavengers from the cocktail. However, an identical profile by SDS-PAGE was observed following Tbfmoc-affinity chromatography and desalting.

The centre band of approximate molecular weight 14 kDa on the SDS-PAGE, was electroblotted onto polyvinyldifluoride (PVDF) membrane and subjected to *N*-terminal sequencing. This showed the *N*-terminus of this component to be blocked, and also not be released following incubation with pyroglutamate aminopeptidase.

Hence indicating that the impurities are most likely to be capped truncates that have bound to the carbon support *via* aromatic side chain functionalities.

3.2.2.4 Tbfmoc-Polystyrene Purification

It was recently shown by Jamieson that chromatographic grade polystyrene could be used in place of the activated charcoal for Tbfmoc-affinity purification.[85] The advantage of this solid support is that extensive prewashing is no longer necessary. This has greatly reduced the time taken to complete an affinity purification from two days to half a day. The protein obtained on work up after desalting is also of higher quality, since it is no longer contaminated with charcoal fines produced during piperidine treatment. The polystyrene method should also be amenable to recycling, keeping the cost of the procedure in line with that using charcoal. Finally, it is anticipated that less polystyrene will be required, since the system should be more efficient due to the Tbfmoc-moiety interacting with the polystyrene in an end-on fashion as well there being the face-on, π-stacking interactions predicted for Tbfmoc and charcoal.

This new support was investigated for the purification of dhIFN-γ. The protocol followed is analogous to that followed for charcoal, **figure 3.2.4**, with the exception that acetic acid solution is used to load the protein and wash the polystyrene, and the purified protein is cleaved from the Tbfmoc-molecule using piperidine/acetonitrile solution. The HPLC profile following polystyrene purification, **figure 3.2.7**, is similar to that obtained following charcoal purification, **figure 3.2.5**.

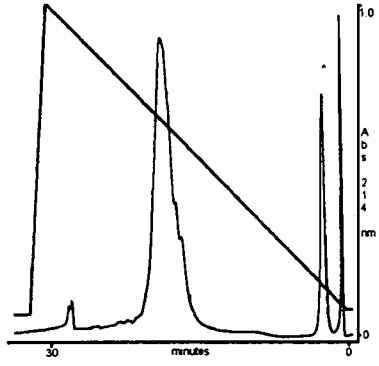
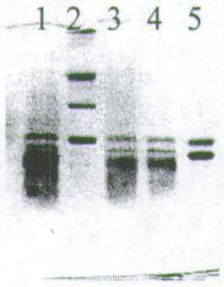


Figure 3.2.7: HPLC Profile After Tbfmoc-Polystyrene Purification

The protein solution was desalted using a Sephadex G-50 column, and on comparison of the solid material obtained using both methods by SDS-PAGE, no difference was seen, **figure 3.2.8**. Both methods produced protein material containing three components.



Lane 1: Crude Tbfmoc-dhIFN-γ
Lane 2: Molecular Weight Markers (43, 25 & 13.7 kDa)
Lane 3: Protein material following charcoal purification/desalting
Lane 4: Protein material following polystyrene purification/desalting
Lane 5: Molecular Weight Standards (17 & 13 kDa)

Figure 3.2.8: SDS-PAGE Profile Comparing Charcoal and Polystyrene

The one advantage polystyrene shown over charcoal in this purification is that a higher yield of the 3 component mixture was obtained, **table 3.2.1**.

Support	Amount of Crude Protein Purified	Amount of Support Used per mg of Crude Protein	Amount of 3 Component Mixture Recovered
Charcoal	100 mg	25 mg	25 mg
Polystyrene	100 mg	20 mg	40 mg

Table 3.2.1: Comparison of Charcoal and Polystyrene Supports

The effect of recycling the polystyrene support was also investigated. The polystyrene was used for a purification procedure and on completion, it was regenerated by agitation in warm (40^0C) toluene to desorb any Tbfmoc by-product still adsorbed to the support. The polystyrene was then thoroughly washed with methanol and allowed to dry *en vacuo*. The results for two cycles of this regeneration are shown in **table 3.2.2**.

Polystyrene	Amount Required per mg of Crude Protein
New	20 mg
After 1st Regeneration	20 mg
After 2nd Regeneration	20 mg

Table 3.2.2: Affect of Regeneration on Polystyrene Performance

From the results given in **table 3.2.2**, it is clear that regeneration of the polystyrene support is not detrimental to the success of subsequent purification procedures. In summary, these findings, combined with those of Jamieson,[85] have shown polystyrene to be a very suitable alternative to charcoal for Tbfmoc-affinity purification.

3.2.3 FPLC Size Exclusion

This technique, which separates on the basis of size, was used by Draffan in the original protocol, and was found to improve the purity of the protein sample from a four component mixture to a three component mixture. It has also been used for the purification of the recombinant protein.[77] As the current sample was already a three component mixture, it was hoped that this technique would be able to purify the dhIFN–γ further than had been previously possible.

The protein was dissolved in urea solution containing NaCl, buffered to pH 7.5. The salt was included to prevent any ionic interactions between the protein and the Superdex™ 75 matrix and the solution was buffered to a slightly basic pH as this protein has been found to be unstable in acidic media containing NaCl.[54] Fractions eluted from the column were examined by UV, at 280 nm, and aliquots of UV active fractions were desalted by dialysis against acetic acid solution, lyophilised and examined by SDS-PAGE for protein content. The profile of the protein elution is depicted in **figure 3.2.9**.

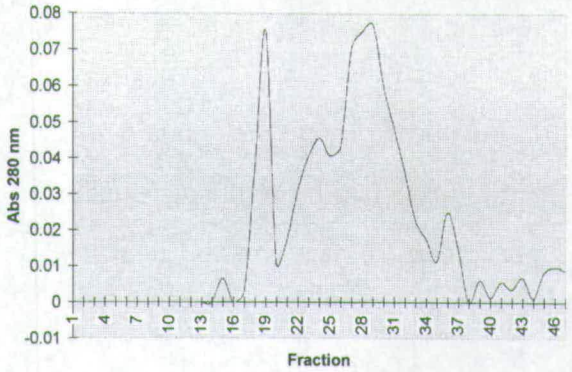


Figure 3.2.9: Elution Profile For FPLC Size Exclusion

Fractions 18-23 were found to consist of only two components by SDS-PAGE, **figure 3.2.10**. These fractions were combined, desalted by dialysis and lyophilised to yield solid material for analysis.

Lane 1: dhIFN-γ after Tbfmoc-affinity purification
Lane 2: dhIFN-γ after FPLC Size Exclusion
Lane 3: Molecular Weight Standards (43, 25 & 13.7 kDa)

Figure 3.2.10: SDS-PAGE Profile After FPLC Size Exclusion

Following FPLC size exclusion, the protein profile by HPLC is now a single peak, with a more symmetrical shape, consistent with the progress made in the purification of the protein, **figure 3.2.11**.

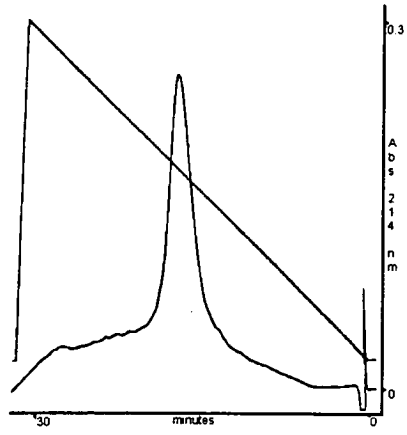


Figure 3.2.11: HPLC Profile of dhIFN-γ After FPLC Size Exclusion

Attempts to purify the protein material further using FPLC size exclusion, were not successful.

3.2.4 Further Purification of dhIFN-γ

In the previous study,[80] an attempt was made to purify dhIFN-γ using hydrophobic interaction and cation exchange chromatographies, and isoelectric focussing. The only successful method employed was the excising of the band containing the correct protein from a preparative SDS-PAGE gel. This method was however, very low yielding and hence, not a suitable preparative method.

Methods have been described in the literature for the use of nickel chelate and cation exchange chromatographies in the purification of recombinant IFN-γ.[77] Nickel chelate chromatography requires the molecule to be correctly folded in order to align the two histidine residues adjacent to one another, and so allow binding to the nickel column to take place. As the synthetic protein at this stage contains only irregular structure this technique was not attempted. Instead, an investigation of cation exchange chromatography was undertaken.

3.2.4.1 Cation Exchange Chromatography

An attempt was made to achieve separation based on the different charges held by the components in the protein mixture instead of using the difference in molecular weight.

Initial studies were performed using the weak cation exchanger, CM Sepharose CL-50 and crude dhIFN-γ containing 4 components obtained from fractions eluting later from the Sephadex G-50 desalting column. The correct pH for the chromatography was determined by equilibriating portions of the matrix to different pH's using various buffers. The protein was then introduced to the matrix as a urea solution, buffered appropriately for each portion. After vortexing for 2 minutes, the supernatants were examined by HPLC to establish whether protein binding had occurred or not. The results of this experiment are shown in **table 3.2.3**.

Tube	Buffer	pH	Binding
1	Acetate	5.5	Yes
2	Phosphate	6	Yes
3	Phosphate	6.5	Yes
4	Phosphate	7	Yes
5	Phosphate	7.5	Yes
6	Phosphate	8	No
7	Tris	8.5	No
8	Tris	9	No
9	Tris	9.5	No
10	Tris	10	No

Table 3.2.3: Determination of the correct pH for Cation Exchange

The results show that below pH 7.5 there are sufficient positive charges on the dhIFN-γ molecule to allow binding to the CM Sepharose CL-50.

Initially, it was hoped to incorporate a folding step with this purification method, by loading the solution onto the column in buffered urea, followed by slow removal of the urea in a preliminary gradient. This should result in refolding of the protein, assisted by the cation exchange medium. Finally, elution of the folded, purified protein from the cation exchange matrix can be achieved using a salt gradient. Attempts to achieve concomitant purification and refolding were unsuccessful, due to precipitation of the protein on the column on removal of the urea solution.

In subsequent studies, chromatography was carried out with urea present at all times to aid the solubility of the protein. As urea is an uncharged species it should not interfere with the progress of the purification. No separation was obtained when a salt gradient of 0-1 M NaCl over 5 column volumes was used. However, when this gradient was shallowed to 0-0.6 M urea over 8 column volumes, it was possible to separate the lowest band (approximately 8 kDa) from the remaining three. Encouraged by this result, attention was switched to the SP Sepharose FF matrix used by Zhang and co-workers[77] in their purification of the recombinant protein. This medium is a stronger cation exchange matrix than CM Sepharose CL-50, thus it was hoped that this would impart a better separation of the components on introduction of the salt gradient. Unfortunately this was not the case, and again, it was only possible to separate the smallest molecular weight component from the other three present.

The investigation of cation exchange chromatography has shown that it is possible to develop conditions for proteins with low solubility in aqueous buffers, avoiding protein precipitation. However, for this protein, the method proved to have no advantage over size exclusion chromatography. Attention was thus returned to this method of purification once more.

3.2.5 Sephadex G-75 Size Exclusion

In order to optimise separation in a size exclusion procedure there are several factors to be incorporated.[86] First, the sample must be loaded onto the column in a volume which does not exceed 1 % of the total column volume. Secondly, for optimal

separation, the ratio of the length to the diameter of the column should be in approximately 25:1. Finally, the flow rate of the column should be adjusted to be as low as possible.

Whilst Superdex^{TM} 75 offers the possibility for high loadings and fast chromatography, it was not useful in separating dhIFN-γ from its impurities. As a final attempt a Sephadex G-75 superfine size exclusion medium was examined. This gel offers the same separation range to Superdex^{TM} 75, but very much lower flow rates. There is literature precedent for using this medium in the purification of the recombinant IFN-γ.[78]

The protein was loaded onto the column at high concentrations. Buffered urea was used as solvent as, from experience, this was found to be the optimal solvent for dhIFN-γ. The protein was eluted using acetic acid solution, in order to permit lyophilisation of protein containing fractions following examination by UV. UV active fractions were examined by SDS-PAGE.

Significant separation was achieved using this column, and the protein material was found to contain only two components. Purification of the two component mixture was achieved by *via* reapplication of the 2 band material to the column. The elution profile is depicted in **figure 3.2.12**, and generally fractions 9 and 10 were found to contain purified protein, **figure 3.2.13**.

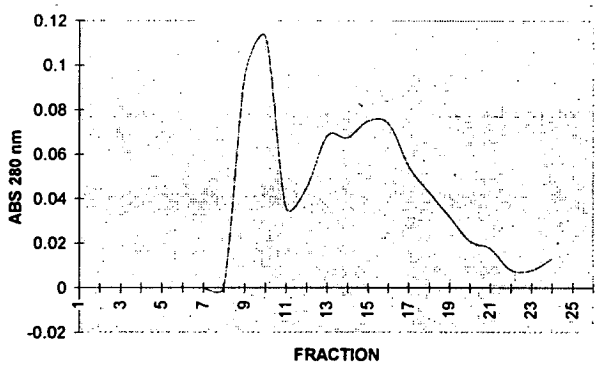
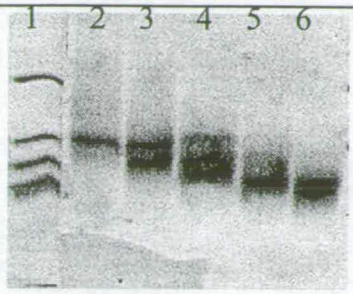


Figure 3.2.12: Elution Profile For Sephadex G-75 Size Exclusion



Lane 1: Molecular Weight Markers (43, 25, 17 & 13.4 kDa)
Lane 2: Fraction 10
Lane 3: Fraction 11
Lane 4: Fraction 12
Lane 5: Fraction 13
Lane 6: Fraction 14

Figure 3.2.13: SDS-PAGE Profile of Sephadex G-75 Size Exclusion

Using this method, 67 mg of three component dhIFN-γ has been purified, in portions, to yield 8 mg of pure protein for analysis and folding studies.

3.2.6 Characterisation of the Purified dhIFN-γ

3.2.6.1 HPLC

The HPLC trace obtained for the purified dhIFN-γ material consists of a relatively sharp, symmetrical peak, **figure 3.2.14**. This is generally indicative of a reasonable degree of purity.

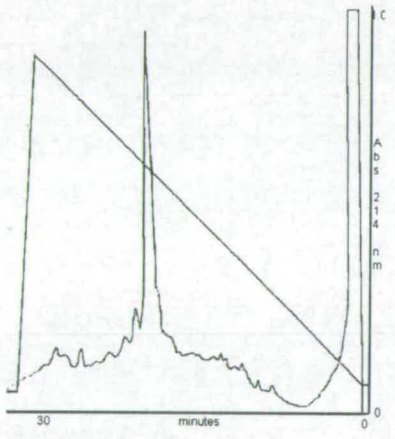


Figure 3.2.14: HPLC Profile of Pure dhIFN-γ

3.2.6.2 Amino Acid Analysis

Throughout the course of the purification the protein material has been examined by amino acid analysis, after hydrolysis of the material at 110 ^{0}C in HCl. The results are depicted in **table 3.2.4** and in general were found to be satisfactory.

	Resin-bound	Tbfmocd hIFN-γ-Resin	Tbfmoc-dhIFN-γ	After Charcoal/ G50	After Polystyrene/ G50	After FPLC	Pure dhIFN-γ
Asx_{20}	19.0	19.9	19.1	19.8	21.1	21.7	19.7
Thr_5	5.3	5.5	4.4	4.6	4.9	4.2	5.1
Ser_{11}	9.7	9.4	9.8	9.4	9.9	10.4	9.7
Glx_{18}	19.6	19.9	18.2	17.8	18.7	18.7	19.3
Pro_2	1.4	1.4	1.0	1.3	1.1	1.1	1.8
Gly_5	5.0	5.3	4.9	5.0	4.1	4.9	5.0
Ala_8	8.3	8.4	8.2	8.1	7.5	6.6	7.3
Val_8	8.2	8.0	8.4	8.9	10.3	8.5	8.6
Met_4	4.5	3.8	3.0	3.6	3.1	2.7	3.3
Ile_7	7.0	7.3	6.8	7.4	7.6	8.2	7.6
Leu_{10}	9.6	9.5	10.0	10.0	10.0	10.3	10.2
Tyr_4	3.3	3.3	2.9	3.3	1.5	2.6	3.0
Phe_{10}	10.3	10.7	10.6	10.8	11.6	13.2	12.2
His_2	2.1	1.6	2.4	2.6	2.0	1.7	2.0
Lys_{20}	20.0	19.9	19.1	19.8	22.2	22.3	20.7
Arg_8	9.8	9.6	9.9	8.5	8.1	7.0	7.8

Table 3.2.4: Amino Acid Analyses

Samples were analysed for their tryptophan content using the method described by Edelhoch.[87] By measuring the absorbance at 280 and 288 nm, of a solution of known protein concentration, it was possible to calculate the extinction coefficient for the solution at the two different wavelengths using **equation 3.1**.

$\varepsilon = A/C$ ε = molar extinction coefficient
A = UV absorbance
C = protein concentration

Equation 3.1: Calculation of Extinction Coefficient

This value could be converted into the corresponding number of moles of tryptophan per mole of protein using **equation 3.2.**

$$\varepsilon_{280} = 5690\ N_{Trp} + 1280\ M_{Tyr}$$

$$\varepsilon_{288} = 4815\ N_{Trp} + 385\ M_{tyr}$$

Equation 3.2: Calculation of Tryptophan Content

Where N and M = no of moles of Trp and Tyr per mole of protein respectively.

The average value for the tryptophan content was determined, using this method, to be 0.99 moles of tryptophan per mole of protein.

3.2.6.3 *N*-Terminal Sequencing

N-terminal sequencing was carried out[88] to ensure that the purified material was indeed the correct sequence, and not a large truncate molecule. This was performed by Edman degradation, labelling the *N*-terminus with phenylisothiocyanate.[89] Five cycles of Edman degradation showed the *N*-terminus to be intact, **table 3.2.5.**

Cycle	1	2	3	4	5
Expected	Gln	Asp	Pro	Tyr	Val
Found	Gln	Asp	Pro	Tyr	Val
Quantity	75.56pmol	79.23pmol	62.96pmol	52.37pmol	53.19pmol

Table 3.2.5: *N*-Terminal Sequencing Analysis

3.2.6.4 MALDI-TOF Mass Spectrometry

MALDI-TOF mass spectrometry was not possible until the dhIFN-γ was purified. Even the pure protein gave a poorly resolved peak, possibly due to the protein containing only irregular structure. The value obtained for the molecular weight of the protein was 16826 Da, **figure 3.2.15**, which is within 49 Da, or 0.3 %, of the theoretical value of 16777 Da.

Figure 3.2.15: MALDI-TOF Mass Spectrum of dhIFN-γ

3.2.6.5 Molecular Weight Determination By FPLC

To check the size of the purified protein further, the molecular weight was determined using FPLC size exclusion on a column calibrated with molecular weight standards.

For each standard, the elution volume parameter, k_{AV}, was calculated using **equation 3.3**, and the void volume was measured using blue dextrin.

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

Ve = elution volume
Vo = column void volume
Vt = total column volume

Equation 3.3: Determination of Elution Volume Parameters

The elution volume parameter for dhIFN-γ was determined, and a graph of k_{AV} versus the logarithm of the molecular weight of the standard was used to find the determine the molecular weight of dhIFN-γ , **figure 3.2.16**. This technique gave a determined molecular weight value of 16.7 kDa for dhIFN-γ.

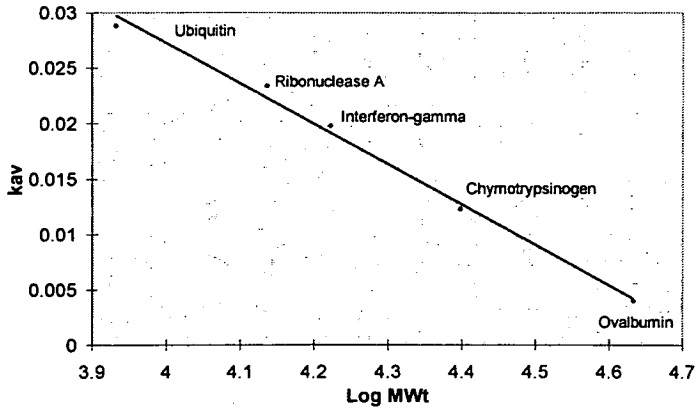


Figure 3.2.16: FPLC Determination of Molecular Weight

3.2.6.6 Tryptic Digest

This technique, hydrolysis of the protein into smaller fragments which can then be examined by MALDI-TOF mass spectrometry, was used to establish that the entire dhIFN-γ sequence had been correctly prepared. The enzyme used in the digest, Trypsin, cleaves only two sites; at the *C*-terminus of lysine and arginine. This, in theory, should produce a reasonable number of peptide fragments for mass spectrometric analysis. On completion of the digest, the peptide fragments were separated from the enzyme by RP HPLC. The first half of the digest mixture was collected *en masse* and lyophilised. An attempt was made however to collect individual HPLC peaks from the second half of the digest.

Mass spectrometric analysis, **table 3.2.6**, showed that not all possible cuts had occurred. However, a sufficient number of peptides were identified which accounted for the majority (approximately 70 %) of the protein sequence, thus establishing that the sequence had indeed been correctly synthesised.

Fragment	Found	Requires
^{59}Asn-^{61}Lys (Na+ salt)	430.419	430.459
^{39}Glu-^{42}Arg (Na+ salt)	657.389	657.591
^{139}Arg-^{143}Gln (H+)	674.639	674.738
^{74}Lys-^{69}Ser (K+ salt)	714.122	714.878
^{81}Phe-^{87}Lys (Na+ salt)	907.532	906.992
^{132}Ser-^{139}Arg	994.658	994.180
^{69}Ser-^{80}Lys (K+ salt)	1432.390	1431.685
^{56}Leu-^{68}Lys	1610.830	1610.680
^{13}Lys-^{42}Arg(Na+ salt)	3447.620	3448.752
^{14}Tyr-^{43}Lys (Na+ salt)	3447.620	3448.752
^{95}Leu-^{125}Lys (H+)	3455.010	3457.58

Table 3.2.6: Mass Spectrometry of Digest Fragments

No fragment corresponding to the *N*-terminus of the protein could be found. However, *N*-terminal sequencing has already confirmed the presence of these residues.

3.2.6.7 pI Determination by Isoelectric Focusing

Isoelectric focussing was carried out using a Rotofor® cell and ampholytes in the range pH 9-11. The focusing experiment was carried out using urea solution. This was to aid the solubility of the dhIFN-γ and because it is an uncharged species it does not interfere with the pH gradient. In this experiment, the pI was estimated to be 10.5. This compares favourably with the calculated value of 10.6.[90]

3.2.7 Folding of IFN–γ

The remarkable array of functions that proteins fulfil in the body require the protein to have adopted its native, fully folded structure. In this conformation, amino acids important for biological activity are held together, in perfect alignment, in order to interact with the complementary surface on the receptor molecule or enzyme.

IFN–γ has been observed to lose biological activity on exposure to acidic pH[8, 9, 10] or high temperatures.[91] Initially, it was assumed that once denatured by exposure to

acid, the biological activity of IFN–γ could not be regained. However, Arakawa and coworkers[54] have shown that in the absence of NaCl, IFN-γ can be refolded to its native state. However, during refolding aggregated or misfolded species may be formed containing the elements of the correct secondary structure, but the incorrect tertiary structure.[78] These studies show that although the synthetic dhIFN-γ sample has been exposed to acid on several occasions, it should be possible to fold the molecule into the biologically active form.

IFN–γ has been refolded by slow removal of denaturants *via* dialysis,[78] and by rapid dilution of a denatured solution into a large volume of stabilising buffer.[55, 79] However, both of these methods result in a low yield recovery of biologically active protein, due to the accumulation of inactive aggregate molecules.

Antibody assisted refolding[92] has been applied to IFN–γ, and significant refolding was observed.[93] Similarly, the protein has been refolded using the GroEL[94] and DnaK[93] molecular chaperone systems and excellent yields of biologically active protein and minimal aggregation have been reported.

The most useful method to date, in terms of cost implications and ease of use, is the polyethyleneglycol (PEG) assisted refolding reported by Cleland and Wang.[95] Again, this technique was found to minimise aggregate formation and enhance the recovery of active protein when applied to IFN–γ.[96] As a result, this method was adopted for the folding of the synthetic dhIFN-γ sample.

3.2.7.1 Folding of Purified dhIFN-γ

The method of Cleland[96] involved the denaturation of the protein in guanidine hydrochloride followed by rapid dilution using a buffered solution containing PEG. Extensive study of the parameters involved showed that optimal refolding was achieved when dilution occurred to yield a final guanidine concentration of 1M, protein concentration of 1mg/ml and a protein to PEG ratio of 1:2.[96] These conditions were applied to the purified dhIFN-γ sample with the substitution of urea in place of guanidine.

HPLC has been used in the past to monitor protein refolding, with diminishing area of the protein peak being observed during folding.[82] Monitoring the refolding of dhIFN-γ was more difficult, since HPLC could not be used due to the acidic and hence denaturing effect of the solvents involved. Acetonitrile has also been reported to interfere with the refolding of IFN–γ.[55] Instead, it was decided to attempt to isolate the refolded protein by removal of the PEG and residual urea on a Sephadex G-50 size exclusion column, which would allow the sample to be submitted for UV-CD analysis and the extent of folding determined. However, this procedure resulted in precipitation on the column, and no protein was recovered.

The method was repeated again, to produce a final protein concentration of 0.1 mg/ml, as the precipitation may have been caused by the high concentration of the protein. However, again no protein was recovered from the desalting column.

Finally, an attempt was made to remove the urea and PEG by dialysis, however over the time required (2 days) to complete this procedure the protein was observed to precipitate in the dialysis bag.

To investigate the extent of folding that had taken place, a sample of the folding buffer was removed after 24 hours and examined by non-reducing SDS-PAGE, **figure 3.2.17**. This indicated that only a small amount of dimer had formed from the purified dhIFN-γ.

Lane 1:	Folded dhIFN-γ
Lane 2:	Molecular Weight Markers (43, 25, 17 & 13 kDa)

Figure 3.2.17: SDS-PAGE Profile of dhIFN-γ Folding

3.2.8 Summary and Outlook

The chemical synthesis of deglycosylated IFN–γ has been succesfully completed, demonstrating that the methodology can be used to routinely produce large proteins. The purification protocol developed has also been applied to the purification of deglycosylated human erythropoietin,[97] establishing its general applicability of to basic proteins of a similar size to IFN-γ. In the future, this purification protocol could be applied to analogues of IFN–γ to obtain pure material for further study. Unfortunately, at this time, it has not been possible to isolate folded dhIFN-γ and obtain structural and biological characterisation. Molecular chaperone assisted refolding could be used to obtain sufficient folded protein for further study. The GroEL/ES chaperone system has been produced at Edinburgh,[98] and examined for folding of a synthetic protein.[85]

3.3 References

1 A. Isaacs & J. Lindenmann, *Proc. Royal Soc. London B Biol. Sci.*, **1957**, *147*, 258-267.

2 Y. Nagano & Y. Kojima, *Cr. Soc. Biol.*, **1958**, *152*, 1627-1629.

3 S. Pestka, J. A. Langer, K. C. Zoon & C. E. Samuel, *Ann. Rev. Biochem.*, **1987**, *56*, 727-777.

4 A. E. Whaley, C. S. R. Meka, L. A. Harbison, J. S. Hunt & K. Imakawa, *J. Biol. Chem.*, **1994**, *269*, 10864-10868.

5 R. M. Roberts, *Nature*, **1993**, *362*, 583.

6 G. R. Adolf, B. Früehbeis, R. Hauptmann, I. Kalsner, I. Maurer-Fogy, E. Ostermann, E. Patzelt, R. Schwendenwein, W. Sommergruber & A. Zoephel, *Biochim. Biophys. Acta*, **1991**, *1089*, 167-174.

7 W. E. Stewart, J. E. Blalock, D. C. Burke, C. Chang, J. K. Dunnik, E. Falcoff, R. M. Friedman, G. J. Galasso, W. K. Joklik, J. Vilcek, J. S. Youngner & K. C. Zoon, *Nature*, **1980**, *286*, 110.

8 E. F. Wheelock, *Science*, **1965**, *149*, 310-311.

9 R. Falcoff, *J. Gen. Virol.*, **1972**, *16*, 251-253.

10 M. P. Langford, J. A. Georgiades, G. J. Stanton, F. Dianzani & H. M. Johnson, *Infect. Immunol.*, **1979**, *26*, 36-41.

11 A. Zerial, A. G. Hovanessian, S. Stefanos, K. Huygen, G. H. Werner & E. Falcoff, *Antiviral Res.*, **1982**, *2*, 227-239.

12 C. E. Samuel & G. S. Knutson, *Virology*, **1983**, *130*, 474-484.

13 B. Y. Rubin & S. L. Gupta, *Proc. Natl. Acad. Sci., USA*, **1980**, *77*, 5928-5932.

14 M. De Ley, J. van Damme, H. Claeys, H. Weening, J. W. Heine, A. Billiau, C. Vermylen & P. de Somer, *Eur. J. Immunol.*, **1980**, *10*, 877-883.

15 M. Wiranowski-Stewart, L. S. Lin, I. A. Braude & W. E. Stewart, *Mol. Immunol.*, **1980**, *12*, 625-633.

16 I. L. Nathan, J. E. Groopman, S. G. Quan, N. Bersch & D. W. Golde, *Nature*, **1981**, *292*, 842-844.

17 G. Trinchieri, M. Matsumoto-Kobayashi, S. C. Clark, J. Seehra, L. London & B. Perussia, *J. Exp. Med.*, **1984**, *160*, 1147-1169.

18 M. Nakamura, T. Manser, G. D. N. Pearson, M. J. Daley & M. L. Gefter, *Nature*, **1984**, *307*, 381-382.

19 D. Wallach, M. Fellous & M. Revel, *Nature*, **1982**, *299*, 833-836.

20 S. B. Salvin, J. S. Youngner, J. Nishio & R. Neta, *J. Natl. Cancer Inst.*, **1975**, *55*, 1233-1236.

21 J. L. Crane, L. A. Glasgow, E. R. Kern & J. S. Youngner, *J. Natl. Cancer Inst.*, **1978**, *61*, 871-874.

22 F. Dianzani, M. Zucca, A. Scupham & J. A. Georgiades, *Nature*, **1980**, *283*, 400-402.

23 W. R. Fleischmann, J. A. Georgiades, L. C. Osborne & H. M. Johnson, *Infect. Immunol.*, **1979**, *26*, 248-253.

24 J. I. Gallin & H. L. Malech, *J. Am. Med. Assoc.*, **1990**, *263*, 1533-1537.

25 R. A. B. Ezekowitz, *N. Engl. J. Med.*, **1991**, *324*, 509-516.

26 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 & D. V. Goeddel, *Nature*, **1982**, *298*, 503-508.

27 P. W. Gray & D. V. Goeddel, *Nature*, **1982**, *298*, 859-863.

28 R. Devos, H. Cheroutre, Y. Taya, W. Degrave, H. Van Heuverswyn & W. Fiers, *Nucleic Acids Res.*, **1982**, *10*, 2487-2501.

29 S. Tanaka, T. Oshima, K. Ohsuye, T. Ono, A. Mizono, A. ueno, H. Nakazato, M. Tsujimoto, N. Higashi & T. Noguchi, *Nucleic Acids Res.*, **1983**, *11*, 1707-1723.

30 S. L. Naylor, A. Y. Sakaguchi, T. B. Shows, M. L. Law, D. V. Goeddel, P. W. Gray, *J. Exp. Med.*, **1983**, *57*, 1020-1027.

31 Y. K. Yip, B. S. Barrowclough, C. Urban & J. Vilcek, *Science*, **1982**, *215*, 411-413.

32 Y. K. Yip, B. S. Barrowclough, C. Urban & J. Vilcek, *Proc. Natl. Acad. Sci., USA*, **1982**, *79*, 1820-1824.

33 J. Friedlander, D. G. Fischer & M. Rubenstein, *Anal. Biochem.*, **1984**, *137*, 115-119.

34 E. Rinderknecht, B. H. O'Connor & H. Rodriguez, *J. Biol. Chem.*, **1984**, *298*, 859-863.

35 H. C. Kelker, J. Le, B. Y. Rubin, Y. K. Yip, C. Nagler & J. Vilcek, *J. Biol. Chem.*, **1984**, *259*, 4301-4304.

36 Y. K. Yip, R. H. L. Pang, C. Urban & J. Vilcek, *Proc. Natl. Acad. Sci., USA*, **1982**, *79*, 1820-1824.

37 Y.-C. E. Pan, A. S. Stern, P. C. Familletti, F. R. Khan & R. Chizzonite, *Eur. J. Biochem*, **1987**, *166*, 145-149.

38 H.-F. Kung, Y.-C. E. Pan, J. Moschera, K. Tsai, E. Bekesi, M. Chang, H. Sugino & S. Honda, *Methods Enzymol.*, **1986**, *119*, 204-210.

39 S. Honda, T. Asano, T. Kajio, S. Nakagawa, S. Ikeyama, Y. Ichimori, H. Sugino, K. Nara, A. Kakinuma & H.-F. Kung, *J. Interferon Res.*, **1987**, *7*, 145-154.

40 K. Kitano, S. Fujimoto, M. Nakao, T. Watanabe & Y. Nakao, *J. Biotech.*, **1987**, *5*, 77-86.

41 T. Arakawa, Y.-R. Hsu, C. G. Parker & P.-H. Lai, *J. Biol. Chem.*, **1986**, *261*, 8534-8539.

42 P. Leinikki, J. Calderon, M. H. Luquette & R. D. Schreiber, *J. Immunol.*, **1987**, *139*, 3360-3366.

43 G. Garotta, L. Ozmen, H. Dobeli, R. Gentz, S. Legrice, V. Bannwarth, E. Hochuli & K. Talmadge, *J. Interferon Res.*, **1987**, *7*, 685 (abstr).

44 D. Lundell, C. Lunn, D. Dalgarno, J. Fossetta, R. Greenberg, R. Reim, M. Grace & S. Narula, *Prot. Eng.*, **1991**, *4*, 335-341.

45 H. Döbeli, R. Gentz, W. Jucker, G. Garotta, D. W. Hartman & E. Hochuli, *J. Biotechnol.*, **1988**, *7*, 199-216.

46 T. Arakawa, Y.-R. Hsu, D. Chang, N. Stebbing & B. Altrock, *J. Interferon Res.*, **1986**, *6*, 687-695.

47 K. Miyata, Y. Yamamoto, M. Ueda, Y. Kawade, K. Matsumoto, I. Kubota, *J. Biochem.*, **1986**, *99*, 1681-1688.

48 J. H. G. M. Mutsaers, J. P. Kamerling, R. Devos, Y. Guisez, W. Fiers & J. F. G. Vliegenthart, *Eur. J. Biochem.*, **1986**, *156*, 651-654.

49 K. Cantell, S. Hirvonen, T. Sareneva, J. Pirhonen & I. Julkunen, *J. Interferon Res.*, **1992**, *12*, 177-183.

50 T. Sareneva, K. Cantell, L. Pyhala, J. Pirhonen & I. Julkunen, *J. Interferon Res.*, **1993**, *13*, 267-269.

51 V. Bocci, A. Pacini, G. P. Pessina, L. Paulescu, M. Muscettola & G. Lunghetti, *J. Gen. Virol.*, **1985**, *66*, 887-891.

52 T. Sareneva, J. Pirhonen, K. Cantell & I Julkunen, *Biochem. J.*, **1995**, *308*, 9-14.

53 S. Yamamoto, S. Hase, H. Yamauchi, T. Tanimoto & T. Ikenaka, *J. Biochem.*, **1989**, *105*, 1034-1039.

54 T. Arakawa, Y.-R. Hsu & D. A. Yphantis, *Biochemistry*, **1987**, *26*, 5428-5432.

55 H. H. Hogrefe, P. McPhie, J. B. Bekisz, J. C. Enterline, D. Dyer, D. S. A. Webb, T. L. Gerald & K. C. Zoon, *J. Biol. Chem.*, **1989**, *264*, 12179-12186.

56 S. Vijay-Kumar, S. E. Senadhi, S. E. Ealick, T. L. Nagabhushan, P. P. Trotta, R. Kosecki, P. Reichert & C. E. Bugg, *J. Biol. Chem.*, **1987**, *262*, 4804-4805.

57 S. Ealick, W. J. Cook, S. Vijay-Kumar, M. Carson, T. L. Nagabhushan, P. P. Trotta & C. E. Bugg, *Science*, **1991**, *252*, 698-702.

58 S. Grzesiek, H. Döbeli, R. Gentz, G. Garotta, A. M. Labhardt & A. Bax, *Biochemistry*, **1992**, *31*, 8180-8190.

59 A. Rashidbaigi, J. A. Langer, V. Jung, C. Jones, H. G. Morse, J. A. Tischfield, J. J. Trill, H.-F. Kung & S. Pestka, *Proc. Natl. Acad. Sci., USA*, **1986**, *83*, 384-388.

60 K. Pfizenmaier, K. Wiegmann, P. Scheurich, M. Krönke, G. Merlin, M. Aguet, B. B. Knowles & U. Ücer, *J. Immunol*, **1988**, *141*, 856-860.

61 M. Aguet, Z. Dembic & G. Merlin, *Cell*, **1988**, *55*, 273-280.

62 D. Novick, P. Orchansky, M. Revel & M. Rubenstein, *J. Biol. Chem.*, **1987**, *262*, 8483-8487.

63 J. Calderon, K. C. F. Sheehan, C. Chance, M. L. Thomas & R. D. Schreiber, *Proc. Natl. Acad. Sci., USA*, **1988**, *85*, 4837-4841.

64 A. C. Greenlund, R. D. Schreiber, D. V. Goeddel & D. Pennica, *J. Biol. Chem.*, **1993**, *268*, 18103-18110.

65 M. R. Walter, W. T. Windsor, T. L. Nagabhushan, D. J. Lundell, C. A. Lunn, P. J. Zauodny & S. K. Narula, *Nature*, **1995**, *376*, 230-235.

66 C. A. Lunn, J. Fossetta, D. Dalgarno, N. Murgolo, W. Windsor, P. J. Zavodny, S. K. Narula & D. Lundell, *Prot. Eng.*, **1992**, *5*, 253-257.

67 V. Jung, C. Jones, C. S. Kumar, S. Stefanos, S. O'Connell & S. Pestka, *J. Biol. Chem.*, **1990**, *265*, 1827-1830.

68 M. A. Farrar, J. Fernandez-Luna & R. D. Schreiber, *J. Biol. Chem.*, **1991**, *266*, 19625-19635.

69 T. Fischer, A. Rehm, M. Aguet & K. Pfizenmaier, *Cytokine*, **1990**, *2*, 157-161.

70 J. Soh, R. J. Donnelly, S. Kotenko, T. M. Mariano, J. R. Cook, N. Wang, S. Emanuel, B. Schwartz, T. Miki & S. Pestka, *Cell*, **1994**, *76*, 793-802.

71 J. Soh, R. J. Donnelly, T. M. Mariano, J. R. Cook, B. Schwartz & S. Pestka, *Proc. Natl. Acad. Sci., USA*, **1993**, *90*, 8737-8741

72 S. Hemmi, R. Böhni, G. Stark, F. Dimarco & M. Aguet, *Cell*, **1994**, *76*, 803-910.

73 M. Müller, J. Briscoe, C. Laxton, D. Guschin, A. Ziemiecki, O. Silvennoinen, A. G. Harpur, G. Barbieri, B. A. Witthuhn, C. Schindler, S. Pellegrini, A. F. Wilks, J. N. Ihle, G. R. Stark & I. M. Kerr, *Nature*, **1993**, *366*, 129-135.

74 V. Jung, A. Rashidbiagi, C. Jones, J. A. Tischfield, T. B. Shows & S. Pestka, *Proc. Natl. Acad. Sci., USA*, **1987**, *84*, 4151-4155.

75 F. Paoilini & E. Suarez, Unpublished Results.

76 S. K.-S. Luk, E. Jay & F. T. Jay, *J. Biol. Chem.*, **1990**, *265*, 13314-13319.

77 Z. Zhang, K.-T. Tong, M. Belew, T. Pettersson & J. C. Janson, *J. Chromatography*, **1992**, *604*, 143-155.

78 T. Arakawa, N. Kirby-Alton & Y.-R. Hsu, *J. Biol. Chem.*, **1985**, *260*, 14435-14439.

79 J. Haelewyn & M. De Ley, *Biochem. Mol. Biol. Int.*, **1995**, *37*, 1163-1171.

80 L. C. Draffan, *Ph. D. Thesis*, The University of Edinburgh, **1996**.

81 R. Ramage & G. Raphy, *Tetrahedron Lett.*, **1992**, *33*, 385-388.

82 N. Robertson & R. Ramage, *J. Chem. Soc. Perkin Trans.* 1, **1999**, 1015-1021.

83 J. Wilken, *Ph. D. Thesis*, The University of Edinburgh, **1995**.

84 R. B. Merrifield, *Personal Communication*

85 C. Jamieson, *Unpublished Results.*

86 D. M. Bollag, M. D. Rozycki & S. J. Edelstein in "*Protein Methods*", 2nd Ed., Wiley-Liss Inc., **1996**, p274.

87 H. Edelhoch, *Biochemistry*, **1967**, *6*, 1948-1955.

88 A. D. Cronshaw, Welmet Sequencing, University of Edinburgh.

89 J. D. Hayes, L. A. Kerr & A. D. Cronshaw, *Biochem. J.*, **1989**, *264*, 437-445.

90 Calculated using the computer programme PEPTIDE, Lighthouse Data.

91 M. G. Mulkerrin & R. Wetzel, *Biochemistry*, **1989**, *28*, 6556-6561.

92 J. D. Carlson & M. L. Yarmush, *Bio/Technology*, **1992**, *10*, 86-91.

93 K. Vandenbroeck & A. Billiau, *Biochimie*, **1998**, *80*, 729-737.

94 K. Vandenbroeck, E. Martens & A. Billiau, *Eur. J. Biochem.*, **1998**, *251*, 181-188.

95 J. L. Cleland & D. I. C. Wang, *Bio/Technology*, **1990**, *8*, 1274-1278.

96 J. L. Cleland, S. E. Builder, J. R. Swartz, M. Winkler, J. Y. Chang & D. I. C. Wang, *Bio/Technology*, **1992**, *10*, 1013-1019.

97 L. Bland, *Unpublished Results.*

98 C. Jamieson & D. Campopiano, *Unpublished Results.*

Chapter 4

Experimental

4.1 General

Chemicals were purchased from the Aldrich Chemical Company, Fisher (Acros) Scientific UK or Fisons. Inorganic reagents were purchased from BDH, and solvents from Prolabo or Rathburn Chemicals and were of analytical or HPLC grade where required. The purity was checked by means of melting point and/or proton NMR. Liquids were distilled before use and the boiling point checked. DCM was distilled from CaH_2 and THF distilled from sodium/benzophenone.

Melting points were recorded in open capillaries using a Buchi 510 oil immersion melting point apparatus and are not corrected. Optical Rotations were measured using a A1000 polarimeter (Optical Activity Ltd) using a 10.0 cm cell in the solvents indicated in the text. Analytical thin layer chromatography (t.l.c.) was performed using plastic sheets precoated with silica gel $60F_{254}$ (Merck) in the solvent systems described in the text. Compounds were visualised using absorption at 254 nm. Infrared spectra were recorded on a Bio Rad SPC3200 or a Perkin Elmer Paragon 1000 FT-IR Spectrometer, as the KBr disc of the solid. UV spectra were recorded on a Perkin Elmer single beam spectrophotometer in the solvents described in the text. Sonication was carried out in a Decon F5300b sonic bath. Proton NMR spectra were recorded on a Brucker WP-200 (200 MHz) or a Varian Gemini 200 instrument in the solvents indicated in the text referenced to TMS. Carbon-13 NMR spectra were recorded on a AC250 (60 MHz) instrument in the solvents described in the text. Elemental analyses were performed on a Perkin Elmer 2400 CHN elemental analyser.

High and low resolution fast atom bombardment mass spectra (FAB MS) were measured on a Kratos MS50TC instrument using thioglycerol, 3-nitrobenzyl alcohol or glycerol as matrix. Matrix assisted laser desorption ionisation (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. Electrospray mass spectrometry was performed using Micromass Platform II Mass Spectrometry. Calculated masses were based on average isotope composition and were derived using the program ADDMASS (Lighthouse Data).

Fmoc-protected amino acids were purchased from Nova Biochem, *p*-alkoxybenzylalcohol resin (Wang) and the amide resin [(2-copolystyrene-1%-divinylbenzene)methyl-aminocarbomethoxy-5-(9'-fluorenylmethoxycarbonyl)amino-dibenzocycloheptadiene resin] were purchased from Bachem. Piperidine (Rathburn Chemicals), acetic anhydride (Fluka), *N*,*N*-Diisopropylethylamine (DIEA) and 1,1,1-trifluoroacetic acid (TFA) (Applied Biosystems) used for the synthesis and cleavage of peptides were all peptide synthesis grade. The 1-hydroxy-4-ethoxycarbonyl-1,2,3-triazole (HOCt)[1] used in coupling procedures and the tetrabenzo[*a,c,g,i*]fluorenyl-17-methoxycarbonyl chloroformate (TbfmocCl)[2] was synthesised in these laboratories at Edinburgh University.

Amino acid analyses (AAA) were performed using a Pharmacia Biotech Biochrom 20 amino acid analyser or a LKB 4150 alpha amino acid analyser on the hydrolysate obtained after heating the sample in 6 M HCl at 110 °C in sealed Carius tubes for the times indicated in the text. After hydrolysis, HCl was removed on a Savant Speed Vac Plus SC110A for 2 hours. The residue was dissolved in 0.2 M citrate buffer pH 2.2 prior to analysis. High performance liquid chromatography (HPLC) was carried out using an ABI system comprising 2 x 1406A 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 manomeric module, a 119 UV/vis detector and a Gilson 715 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). Protein *N*-terminal sequencing was performed on an ABI 477A sequencer at the Welmet Protein Characterisation Facility (University of Edinburgh). This was performed by Edman degradation, labelling the *N*-terminus with phenyl isothiocyanate.[3] Fast protein liquid

chromatography (FPLC) was carried out on a Pharmacia FPLC system with Liquid Chromatography Controller LCC-501 Plus using a Superdex^{TM} 75 HR 10/30 or 26/60 column. Gel filtration (Sephadex G-50 and Sephadex G-75) and cation exchange (CM-Sephadex CL-50 and SP-Sepharose FF) media were purchased from Pharmacia. Column chromatography of proteins was carried out using Pharmacia LKB apparatus comprising 2xLKB 2138 UVCords, a Pharmacia 2132 Microperpex peristaltic pump, a LKB 2112 redirac fraction collector and a Pharmacia GM-1 gradient mixer where appropriate. Isoelectric focusing (IEF) was carried out using the Bio-Rad Rotofor system with ampholytes purchased from Bio-Rad or Fluka. Dialysis tubing, purchased from Spectrum, were Spectra/Por CE (cellulose ester) membranes with molecular weight cut-offs (MwtCO) of 10 000 Da. Samples were centrifuged using MSE Mistral 2000R (Sanyo). Sample concentration employed Amicon Ultrafiltration cell 8050 using Spectrum Molecular/Por membranes, MwtCO 5 000 Da. SDS-PAGE was carried out using the discontinuous buffer system of Laemmli[4] and a Bio-Rad Mini Protean II Cell. Each gel measured 7 cm (L) by 8 cm (W) by 75 mm (T) and had 10 wells. The stacking gel contained 4 % w/v acrylamide, 0.125 M Tris pH 6.8, 0.01 % w/v SDS and the separating gel 15 % w/v acrylamide, 0.375 M Tris pH 8.8, 0.01 % w/v SDS. The acrylamide was polymerised using TEMED and APS. Prior to electrophoresis the protein samples were denatured by heating in sample buffer (0.0625 M Tris pH 6.8, 2 % w/v SDS, 0.05 % w/v bromophenol blue in water/2-mercaptoethanol/glycerol (17:1:2)) for 10 minutes at 110 ^{O}C. Separation was achieved at 200 mV for 38 minutes using the normal running buffer (7.2 % w/v glycine, 0.5 % w/v SDS, 0.125 M Tris Base). Protein bands were visualised using 0.5 % w/v coomassie blue in acetic acid/methanol/water (1:4:5) and destained using acetic acid/methanol/water (1:4:5). Non-reducing PAGE was carried out using the following sample buffer; 0.031 M Tris, pH 6.8, 0.05 % bromophenol blue in water/glycerol (1:1). Samples were not heated prior to electrophoresis.

4.2 Solid Phase Peptide Synthesis

Peptides were synthesised on an ABI 430A automated peptide synthesiser with on line UV monitoring using an ABI 758A detector. All peptides were synthesised using the Fmoc strategy and acid labile peptide-resin linkers. The protected amino acids employed are depicted in **table 4.1**.

Amino Acid	Protecting Group
Gly/Ala/Leu/Ile/Val/Phe/Met/Pro	Not Necessary
Asp/Glu	Bu^t Ester
Ser/Thr/Tyr	Bu^t Ether
Lys/Trp	Boc
Arg	Pmc/Pbf
Asn/Gln	Trt
His	Trt
Cys	Trt

Table 4.1: Fmoc-Compatible Side Chain Protection

4.2.1 Determination of the Resin Loading

The loading efficiency was determined by sonicating a sample of the Fmoc-protected resin, accurately measured, (4-6 mg) in 20 % v/v piperidine/DMF (10 ml) for 10-15 minutes. The UV of the supernatant was then measured between 280 and 320 nm and the loading calculated using the Beer-Lambert law (ε_{302} = 15400 for fulvene piperidine adduct), **equation 4.1**.

$$\text{Functionality (mmol/g)} = \frac{10 \times \text{Abs at } 302\text{nm}}{9 \times \text{mass of resin in mg}}$$

Equation 4.1: Determination of Fmoc-Loading Level

The values for resin functionality (mmol/g) and percentage coupling obtained are listed in the text.

4.2.2 Automated SPPS

Synthetic procedures were programmed into the ABI 430A synthesiser prior to the commencement of each synthesis. The first (*C*-terminal) residue (the amino aldehyde in certain cases) was attached to the resin manually and then the resin transferred to the reaction vessel for automated peptide synthesisers. 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. The cycle was then repeated in order to build up the desired sequence of amino acids in a stepwise manner, from the *C*- to the *N*-terminus. The programmed synthetic cycles are summarised below:

1. Capping

Resin was treated with acetic anhydride (0.5 M), DIEA (0.125 M) and HOBt (0.2 % w/v) in DMF/1,4-dioxane (1:1, 10 ml) with vortexing for 10 minutes. The reaction vessel was drained and the resin washed with six portions of DMF/1,4-dioxane (1:1).

2. Deprotection

The N^{α}-Fmoc protecting group was cleaved using 20 % v/v piperidine in DMF/1,4-dioxane (1:1, 10 ml) for 4 minutes. Then an aliquot of the deprotection filtrate was passed through the UV detector to allow the percentage coupling of each amino acid to be estimated. The deprotection step was then repeated a further time with vortexing for 2.5 minutes to ensure complete removal of the Fmoc group. Finally the resin was washed with six portions of DMF/1,4-dioxane (1:1).

3. Coupling

Amino acid residues were coupled using the HOCt method. The HOCt active ester is preformed from the Fmoc-amino acid (1 mmol), HOCt (0.5 mmol) and DIC (0.5 mmol) and is coupled in a single cycle to the resin by vortexing for a 30 minute period. The reaction vessel is drained and the resin washed with four portions of

DMF/1,4-dioxane(1:1). The only exception is histidine which is coupled with HOBt (2 mmol) placed in the cartridge with histidine (1 mmol) prior to synthesis.

Note

On completion of the synthesis, the final Fmoc-amino acid loading was calculated using the Fmoc-loading test. The overall resin functionality was determined and compared with the theoretical final loading asuming 100% yield in every cycle of the synthesis (obtained using ABI synthesiser software). This gave an estimation of the overall success of the synthesis.

4.3 Experimental Details

4.3.1 The Solid Phase Synthesis of Peptide *C*-Terminal Semicarbazones and Aldehydes

9-Fluorenylmethoxycarbonyl hydrazine[5] (15)

9-Fluorenylmethoxycarbonyl chloroformate (1 g, 3.87 mmol) was dissolved in acetonitrile (150 ml) and added dropwise with stirring to hydrazine monohydrate (6 ml, 123 mmol). The reaction was darkened and stirred for 30 minutes before concentration by rotary evaporation. The white solid obtained was washed with cold ethanol and dried in a darkened vacuum desiccator to give the *title compound.*

Yield 0.75 g, 76 %; **m.p.** 174-176 °C (lit. 173-175[5]); **t.l.c.** (MeOH/CCl_3H/AcOH, 10:90:0.05) R_f 0.44; **CHN** Found C, 70.79 %; H, 5.75 %; N, 11.11 %; $C_{15}H_{14}N_2O_2$ requires C, 70.87 %; H, 5.51 %; N, 11.02 %; **m/z (HR FAB)** Found 255.1131 (MH^+); $C_{15}H_{15}N_2O_2$ requires 255.1134; ν_{max} **(KBr)/(cm^{-1})** 3317 (NH), 3027, 3018 & 2951 (CH), 1694 (CO); **δH (200 MHz, d_6-DMSO)** 4.17-4.36 (3H, m, CH & CH_2, Fmoc), 7.26-7.88 (10H, m, NH_2 & ArH), 8.35 (1H, s, NH); **δC (63 MHz, d_6-DMSO)** 46.0, 65.9, 119.8, 119.9, 120.0, 121.4, 126.8, 127.0, 127.1, 127.3, 137.5, 139.4, 140.4, 142.6, 146.5; λ_{max}**/nm (MeOH, ε/$dm^3mol^{-1}cm^{-1}$)** 265 (11300), 289 (3200) 300 (3400).

Synthesis of (2-copolystyrene-1%-divinylbenzene)methyl-aminocarbomethoxy-5-(9'-fluorenylmethoxycarbonyl)hydrazine-dibenzocycloheptadiene resin (14)

1. Modification of the Resin

Tricyclic amide resin (**11**) was sonicated in 20 % v/v piperidine/DMF (10 ml) for 30 minutes. The resin was filtered, washed exhaustively with DMF, 1,4-dioxane, DCM and dried.

2. Generation of the Isocyanate Intermediate (**13**)

The resin was swollen in the minimum amount of DCM (taken from 20 ml/g of resin) and DIEA (1 equivalent) added. The mixture was sonicated for 10 minutes. A solution of triphosgene (3 equivalents) in the remaining DCM was added, and the mixture sonicated at 10 ^{0}C for a further hour. The resin was filtered, washed exhaustively with DCM and dried.

ν_{max} **(KBr)/(cm^{-1})** 2250-2260 (N=C=O).

3. Trapping the Isocyanate Intermediate with 9-Fluorenylmethoxycarbonyl Hydrazine (**14**)

The isocyanate resin was swollen in the minimum amount of DCM (taken from 20 ml/g of resin). A solution of 9-fluorenylmethoxycarbonyl hydrazine (3 equivalents) in the remaining DCM was added and the mixture sonicated at 10 ^{0}C for 2 hours. The resin was filtered, washed exhaustively with DMF, 1,4-dioxane, DCM and dried. A KBr disc of the resin was made to ensure complete consumption of the isocyanate species. Resin functionality by UV typically 50-75 % coupling (0.20-0.30 mmol/g)

4. Capping the Free Amino Groups

The resin (**14**) was swollen in capping reagent (0.25 M acetic anhydride, 0.125 M DIEA, 0.2 % w/v HOBt in DMF/1,4-dioxane (1:1), 10 ml) and sonicated for 30 minutes. The resin was filtered and washed exhaustively with DMF, 1,4-dioxane, DCM and dried.

9-Fluorenylmethoxycarbonyl amino acid-*N*,*O*-dimethylhydroxylamides (16)

9-Fluorenylmethoxycarbonyl amino acid` (12 mmol) was suspended in DCM (50 ml) and stirred at 0 ^{0}C under dry nitrogen. *N*-methylpiperidine (1.35 ml, 11 mmol) was added, followed by ethyl chloroformate (1.00 ml, 10 mmol) and stirring continued for 3 minutes. A preformed solution of *N,O*-Dimethylhydroxylamine hydrochloride (11 mmol) and *N*-methylpiperidine (1.35 ml, 11 mmol) in DCM (30 ml) at 0 ^{0}C was added in a single portion whilst stirring at 0 ^{0}C. The mixture was stirred at 0 ^{0}C,

under dry nitrogen, for one hour before being allowed to rise to room temperature slowly overnight. The solvent was removed *in vacuo* and the residue dissolved in ethyl acetate. The insoluble precipitate was removed by filtration, and the filtrate washed with 20 % v/v citric acid (2 x 50 ml), saturated sodium bicarbonate solution (2 x 50 ml) and brine (50 ml). The organic layer was dried ($MgSO_4$) and the solvent evaporated to yield the *title compound.*

9- Fluorenylmethoxycarbonyl-L-Phenylalanine-*N,O*-dimethylhydroxylamide (16a)[6]

Yield 81 %; **mpt** 49-50 ^{0}C (no lit. value)[6]; **t.l.c.** (EtOAc/hexane, 2:1) R_f 0.48; **m/z (HR FAB)** Found 431.1958 (MH^+); $C_{26}H_{27}N_2O_4$ requires 431.1971; ν_{max} **(KBr)/(cm^{-1})** 3300 (NH), 2940 (CH), 1719 (C=O); **δH (200 MHz, d_6-DMSO)** 2.78-2.97 (2H, m, βCH_2, Phe), 3.11 (1H, s, NCH_3), 3.71 (1H, s, OCH_3), 4.10-4.24 (4H, m, CH & CH_2, Fmoc & αCH, Phe), 4.65-4.67 (1H, br, s, NH, Phe), 7.15-7.91 (13H, m, ArH); **δC (63 MHz, d_6-DMSO)** 14.1, 36.6, 46.6, 54.9, 61.2, 65.7, 120.1, 121.1, 121.4, 125.3, 126.4, 127.1, 127.4, 127.6, 128.0, 128.3, 128.9, 129.1, 129.3, 137.5, 138.0, 139.5, 140.7, 142.6, 143.8, 156.0; $[\alpha_D]^{22}$ -7.40 ^{0} (c 1.370, DMF); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 263 (21400), 289 (5700), 299 (6400).

9-Fluorenylmethoxycarbonyl-*L*-Alanine-*N,O*-dimethylhydroxylamide (16b)

Yield 78 %; **mpt** 130 ^{0}C; **t.l.c.** (EtOAc/hexane, 2:1) R_f 0.36; **m/z (HR FAB)** Found 355.1660 (MH^+); $C_{20}H_{23}N_2O_4$ Requires 355.1658; ν_{max} **(KBr)/(cm^{-1})** 3341 (NH), 3040, 2987, 2960, 2899 (CH), 1720 (C=O, urethane), 1667 (C=O, amide); **δH (200 MHz, d_6-DMSO)** 1.21 (3H, d, CH_3, J=7.1 Hz), 3.11 (3H, s, NCH_3), 3.72 (3H, s, OCH_3), 4.19-4.34 (3H, m, CH_2, Fmoc, & αCH, Ala), 4.50 (1H, t, CH, Fmoc, J=7.4 Hz), 7.29-7.90 (8H, m, ArH); **δC (63 MHz, d_6-DMSO)** 16.9, 17.3, 32.1, 46.7, 61.1, 65.7, 120.1, 121.4, 125.3, 127.1, 127.3, 127.7, 128.9, 137.5, 139.5, 142.7, 143.9, 143.9, 155.8, 173.1; $[\alpha_D]^{22}$ +6.65^0 (c 1.010, DMF); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 265 (25100), 289 (8400), 300 (8100).

9-Fluorenylmethoxycarbonyl-*D*-Alanine-*N,O*-dimethylhydroxylamide (16c)

Yield 72 %; **mpt** 119-120 ^{0}C; **t.l.c.** (EtOAc/hexane, 2:1) R_f 0.37; **m/z (HR FAB)** Found 355.1665 (MH^+); $C_{20}H_{23}N_2O_4$ Requires 355.1658; ν_{max} **(KBr)/(cm^{-1})** 3343 (NH), 3040, 2986, 2961, 2899 (CH), 1721 (C=O, urethane), 1668 (C=O, amide); **δH (200 MHz, d_6-DMSO)** 1.21 (3H, d, CH_3, J=7.1 Hz), 3.10 (3H, s, NCH_3), 3.72 (3H, s, OCH_3), 4.17-4.46 (3H, m, CH_2, Fmoc & αCH, Ala), 4.49 (1H, t, CH, Fmoc, J=7.1 Hz), 7.29-7.89 (8H, m, ArH); **δC (63 MHz, d_6-DMSO)** 16.9, 17.3, 32.0, 46.7, 61.2, 65.7, 120.2 (2C), 125.4 (2C), 127.2 (2C), 127.8 (2C), 140.8(2C), 143.9 (2C), 155.9, 173.2; $[\alpha_D]^{22}$ -5.58 ^{0} (c 1.112, DMF); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 265 (20400), 289 (5300), 300 (6200).

9-Fluorenylmethoxycarbonyl-*L*-Tryptophan-*N,O*-dimethylhydroxylamide (16d)

Yield 92 %; **m.p.** 74-75 °C; **t.l.c.** (EtOAc/Hexane, 2:1) R_f 0.28; **m/z (HR FAB)** Found 470.2085 (MH^+) ; $C_{28}H_{28}N_3O_4$ requires 470.2080; ν_{max} **(KBr)/(cm^{-1})** 3410, 3322 (NH), 3056, 2973, 2936 (CH), 1714 (C=O, urethane), 1654 (C=O, amide); **δH (200 MHz, d_6-DMSO)** 3.05 (2H, d, βCH$_2$, Trp, J=8.4 Hz), 3.10 (3H, s, NCH_3), 3.69 (3H, s, OCH_3), 4.05-4.13 (4H, m, CH & CH_2, Fmoc & αCH, Trp), 4.71 (1H, br, s, NH, Trp), 6.93-7.86 (13H, m, ArH), 10.85 (1H, br, s, NH, indole); **δC (63 MHz, d_6-DMSO)** 31.9, 46.7, 52.1, 59.9, 61.2, 65.7, 111.6, 118.0, 118.6, 120.2, 121.0, 121.5, 124.0, 124.2, 125.4, 127.1, 127.4, 127.7, 129.0, 110.2, 136.2, 140.8 (2C), 143.9 (2C), 156.1, 170.4, 172.7; $[\alpha_D]^{22}$ -20.70^0 (DMF, c 1.024); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 265 (14600), 290 (5700), 300 (3800).

9-Fluorenylmethoxycarbonyl-*L*-Aspartic Acid (^{t}Butyl Ester) -N,O-dimethyl-hydroxylamide (16e)

Yield 83 %; **m.p.** 59-60 °C; **t.l.c.** (EtOAc/Hexane, 2:1) R_f 0.50; **m/z (HR FAB)** Found 455.2203 (MH^+); $C_{25}H_{31}N_2O_6$ requires 455.2182; ν_{max} **(KBr)/(cm^{-1})** 3399, 3336 (NH), 3055, 2978, 2920 (CH), 1725 (C=O, urethane), 1657 (C=O, amide); **δH (200 MHz, d_6-DMSO)** 1.36 (9H, s, 3 x CH_3, ^{t}Bu), 2.73 (2H, d, βCH$_2$, Asp, J=6.1 Hz), 3.09 (3H, s, NCH_3), 3.66 (3H, s, OCH_3), 4.23-4.39 (4H, m, CH & CH_2, Fmoc & αCH, Asp), 4.82 (1H, d, NH, J=8Hz), 7.26-7.89 (8H, m, ArH); **δC (63 MHz, d_6-**

DMSO) 27.8 (3C), 32.2, 38.6, 46.7, 50.7, 61.2, 65.8, 80.4, 120.3 (2C), 125.3 (2C), 127.2 (2C), 127.8(2C), 140.9 (2C), 143.9 (2C), 156.0, 169.4, 176.6; $[\alpha_D]^{22}$ -9.08 ^{0} (DMF, c 1.090); λ_{max}**/nm (MeOH,** ε**/dm^{3}mol^{-1}cm^{-1})** 265 (17100), 289 (5100), 300 (5100).

9-Fluorenylmethoxycarbonyl Amino Acid Aldehydes (17)

9-Fluorenylmethoxycarbonyl amino acid Weinreb amide (2.4 mmol) was dissolved in dry, freshly distilled THF (20 ml) and cooled in a dry ice/acetone bath (-20 ^{0}C) under dry nitrogen. Lithium aluminium hydride (2.5-3.2 mmol) was added in small portions under dry nitrogen. The mixture was stirred for 10 minutes at -20 ^{0}C, and then allowed to warm to room temperature. The mixture was stirred for a further 30 minutes. The reaction was quenched by the addition of 0.6 M $KHSO_4$ (20 ml) and the pH of the solution adjusted to 3.0 by the addition of 1 M HCl (2 ml). The solution was diluted with ethyl acetate (25 ml) and the layers separated. The aqueous layer was extracted twice with ethyl acetate (25 ml) and the combined organic layers washed with 20 % citric acid (2 x 25 ml), saturated sodium bicarbonate solution (2 x 25 ml) and brine (2 x 25 ml). The solution was dried ($MgSO_4$) before the solvent was removed *in vacuo* to give the desired compound

9-Fluorenylmethoxycarbonyl-*L*-Phenylalaninal (17a)[6, 7]

Yield 72 %; **m.p.** 100-102 ^{o}C (lit. 106-108 ^{0}C[6] & 129-130 ^{0}C[7]); **t.l.c.** (EtOAc/hexane, 2:1) R_f 0.54; (DCM/MeOH/AcOH, 9:0.5:0.5) R_f 0.75; **m/z (HR FAB)** Found 372.1617; $C_{24}H_{22}NO_3$ requires 372.1600; ν **(KBr)/(cm^{-1})** 3333 (NH), 3061 (CH, Aromatic), 2961 & 2844 (CH), 1730 (C=O, urethane) 1690 (C=O, aldehyde); δ**H (200 MHz, d$_6$-DMSO)** 2.91 (2H, d, βCH_2, Phe, J=7.4 Hz), 4.11- 4.34 (4H, m, CH & CH_2, Fmoc & αCH, Phe), 7.16-7.89 (13H, m, ArH), 9.55 (1H, s, CHO); δ**C (63 MHz, d$_6$-DMSO)** 33.3, 46.7, 61.1, 65.5, 120.0, 120.2, 124.2, 125.2, 126.3, 126.8, 127.0, 127.1, 127.5, 127.7, 128.1, 128.3, 129.2, 137.8, 139.9, 140.8, 143.8, 148.7, 156.1, 200.6; $[\alpha_D]^{22}$ -43.30^0 (DMF, c 1.146); λ_{max}**/nm (MeOH,** ε**/dm^{3}mol^{-1}cm^{-1})** 265 (18000), 289 (5500), 300 (6500).

9-Fluorenylmethoxycarbonyl-*L*-Alaninal (17b)[7,8]

Yield 73 %; **m.p.** 140-142 °C (lit. 136-137 ^0C[7]& 145 ^0C[8]) ; **t.l.c.** (EtOAc/hexane, 2:1) R_f 0.71; **m/z (HR FAB)** Found 296.1299; $C_{18}H_{18}NO_3$ requires 296.1287; **ν (KBr)/(cm^{-1})** 3345 (NH), 3063, 2963, 2898, 2822 (CH), 1738 (C=O, urethane), 1684 (C=O, aldehyde); **δH (200 MHz, d_6-DMSO)** 1.42 (3H, d, CH_3 Ala, J=7.4 Hz), 4.19-4.37 (4H, m, CH & CH_2 Fmoc, αCH Ala), 7.15-7.92 (8H, m, ArH), 9.43 (1H, s, CHO); **δC (63 MHz, d_6-DMSO)** 13.9, 46.8, 55.2, 65.7, 120.1, 120.3, 124.3, 125.3, 126.8, 127.1, 127.2, 127.8, 140.0, 140.9, 143.9, 148.7, 156.2, 202.0; **$[\alpha_D]^{22}$** -11.57^O (DMF, c1.054); **λ_{max}/nm (MeOH, ε/$dm^3mol^{-1}cm^{-1}$)** 263 (16700), 265 (16800), 289 (5800), 300 (6000).

9-Fluorenylmethoxycarbonyl-*D*-Alaninal (17c)[8]

Yield 47 %; **m.p.** 136-138 °C (lit. 145 ^0C[8]); **t.l.c.** (EtOAc/hexane, 2:1) R_f 0.68; **m/z (HR FAB)** Found 296.1284; $C_{18}H_{18}NO_3$ requires 296.1287; **ν (KBr)/(cm^{-1})** 3342 (NH), 2965 (CH), 1737 (C=O, urethane), 1684 (C=O, aldehyde); **δH (200 MHz, d_6-DMSO)** 1.15 (3H, d, CH_3 Ala, J=7.4 Hz), 4.24-4.38 (4H, m, CH & CH_2 Fmoc, αCH Ala), 7.34-7.91 (8H, m, ArH), 9.44 (1H, s, CHO); **δC (63 MHz, d_6-DMSO)** 13.8, 46.8, 55.2, 65.6, 120.2 (2C), 125.2 (2C), 127.2 (2C), 127.7 (2C), 140.9 (2C), 143.9 (2C), 156.2 , 201.9; **$[\alpha_D]^{22}$** +8.89^O (DMF, c 0.990); **λ_{max}/nm (MeOH, ε/$dm^3mol^{-1}cm^{-1}$)** 265 (17700), 289 (5800), 300 (5600).

9-Fluorenylmethoxycarbonyl-*L*-Tryptophanal (17d)

Yield (0.92 g, 91 %); **m.p.** 79-82 °C; **t.l.c.** (EtOAc/hexane, 2:1) R_f 0.45; **m/z (HR FAB)** Found 411.1718; $C_{26}H_{23}N_2O_3$ requires 411.1709; **ν (KBr)/(cm^{-1})** 3404, 3344 (NH), 3056, 2952, 2923 (CH), 1699 (C=O, urethane), 1617 (C=O, aldehyde); **δH (200 MHz, d_6-DMSO)** 3.08 (2H, d, βCH_2, Trp, J=7Hz), 4.12-4.34 (4H, m, CH & CH_2, Fmoc & αCH, Trp), 6.96-7.89 (13H, m, ArH), 9.59 (1H, s, CHO), 10.87 (1H, s, NH, indole); **δC (63 MHz, d_6-DMSO)** 46.7, 59.8, 60.4, 65.7, 111.5, 118.2, 118.4, 120.0, 120.1, 120.8, 121.0, 123.7, 124.2, 125.2, 125.4, 127.1, 127.6, 109.7, 136.3, 140.0, 140.8, 143.8, 148.8, 156.2, 170.4, 201.3; **$[\alpha_D]^{22}$** -25.35 ^{0} (DMF, c 0.994); **λ_{max}/nm (MeOH, ε/$dm^3mol^{-1}cm^{-1}$)** 265 (20500), 289 (9100), 300 (7300).

9-Fluorenylmethoxycarbonyl-*L*-Aspartic Acid (^{t}Butyl Ester) Aldehyde (17e)[7]

Yield 74 %; **m.p.** 37-38 °C (lit. 61-62 °C[7]); **t.l.c.** (EtOAc/hexane, 2:1) R_f 0.62; **m/z (HR FAB)** Found 396.1800; $C_{23}H_{26}NO_5$ requires 396.1811; **ν (KBr)/(cm^{-1})** 3398, 3326 (NH), 3065, 2979, 2930 (CH), 1728 (C=O, urethane), 1699 (C=O, aldehyde); **δH (200 MHz, d_6-DMSO)** 1.36 (9H, s, 3 x CH_3 tBu), 2.76 (2H, d, βCH$_2$, Asp, J=8 Hz) 4.23-4.37 (4H, m, CH & CH_2, Fmoc, αCH, Asp), 7.27-7.93 (8H, m, ArH), 9.44 (1H, s, CHO); **δC (63 MHz, d_6-DMSO)** 18.2 (3C), 33.3, 46.8, 61.2, 65.7, 80.6, 120.3 (2C), 125.3 (2C), 127.2 (2C), 127.8 (2C), 140.9 (2C), 143.9 (2C), 156.3, 169.6, 200.3; **$[\alpha_D]^{22}$** -9.08 ^{0} (DMF, c 1.090); **λ_{max}/nm (MeOH, ε/dm^{3}mol^{-1}cm^{-1})** 265 (18700) 289 (7900), 300 (5400).

Loading 9-Fluorenylmethoxycarbonyl Amino Acid Aldehydes onto the Linker

1. Modification of the Resin (**14a**)

The resin (**14**) was prepared for use by sonication in 20 % v/v piperidine/DMF for 30 minutes. The resin was filtered, washed exhaustively with DMF, 1,4-dioxane, DCM and diethyl ether and dried.

2. Loading of 9-Fluorenylmethoxycarbony Amino Acid Aldehydes (**18**)

The resin was swollen in the minimum amount of DCM (taken from 20 ml/g of resin) and DIEA (1 equivalent) added. The mixture was sonicated for 10 minutes, then a solution of 9-fluorenylmethoxycarbonyl amino aldehyde (3 equivalents) in the remaining DCM was added, and sonication continued for a further 5 hours at RT. The resin was filtered, washed exhaustively with DMF, 1,4-dioxane, DCM and diethyl ether and dried.

3. Capping the Resin

The resin (**18**) was swollen in capping reagent (0.25 M acetic anhydride, 0.125 M DIEA, 0.2 % w/v HOBt in DMF/1,4-dioxane 1:1) and sonicated for 30 minutes. The resin was filtered, washed exhaustively with DMF, 1,4-dioxane, DCM and diethyl ether and dried.

Levels of Loading obtained by UV at 302nm 9-Fluorenylmethoxycarbonyl-*L*-Phenylalaninal 90 % (**18a**); 9-Fluorenylmethoxycarbonyl-*L*-Alaninal 100 % (**18b**); 9-Fluorenylmethoxycarbonyl-*D*-Alaninal 70 % (**18c**); 9-Fluorenylmethoxycarbonyl-*L*-Tryptophanal 100 % (**18d**); 9-Fluorenylmethoxycarbonyl-*L*-Aspartic Acid (^{t}Butyl Ester) Aldehyde 90 % (**18e**).

Cleavage of 9-Fluorenylmethoxycarbonyl-*L*-Phenylalaninal Semicarbazone From the Linker (19a)

Resin **18a** (0.393 mmol) was stirred under dry nitrogen in TFA/water (95:5, 10 ml) for 1 hour. The resin was filtered and the filtrate concentrated to a volume of 1 ml *in vacuo*. The residue was dissolved in water/acetonitrile/acetic acid (10:0.5:0.5, 20 ml) and lyophilised to yield a pale yellow solid.

Yield (106 mg, 63 %); **m.p.** 136-137 °C; **t.l.c.** (DCM/MeOH/AcOH, 9:0.5:0.5) R_f 0.63; **m/z (ES)** 429.2 (MH^+), 451.2 (MNa^+), 474.2 (MNa_2^{2+}); **νmax/cm^{-1} (KBr)** 3476, 3319 (NH), 3063, 2944 (CH), 1689 (C=O); **δH (200 MHz, d_6-DMSO)** 2.85 (2H, d, βCH_2, Phe, J=8Hz), 4.14-4.24 (4H, CH & CH_2, Fmoc & αCH, Phe), 7.14-7.92 (14H, ArH & HC=N), 9.93 (1H, NH, semicarbazone); **$[\alpha_D]^{22}$** -10.0 ° (c 1.04, DMF).

9-Fluorenylmethoxycarbonyl-*L*-Phenylalaninal Semicarbazone (19b)

9-Fluorenylmethoxycarbonyl-*L*-phenylalaninal **17a** (0.50 g, 1.4 mmol) was suspended in 50 % aqueous ethanol (25 ml). Semicarbazone hydrochloride (0.48 g, 4.3 mmol) and sodium acetate (0.49 g, 6.0 mmol) were added. The mixture was warmed under reflux and ethanol (15-20 ml) added to solubilise the aldehyde. The mixture was heated at reflux for 1 hour before the solvent was concentrated under reduced pressure to a volume of ~10 ml. 5 % sodium bicarbonate solution (10 ml) was added and a brown precipitate formed which was removed by filtration. The brown colour was removed by washing with ethyl acetate to yield a white solid

Yield (0.1472 g, 26 %); **m.p.** 143-145 °C; **t.l.c.** (DCM/MeOH/AcOH, 9:0.5:0.5) R_f 0.64; **m/z (HR FAB)** Found 429.1927; $C_{25}H_{25}N_4O_3$ requires 429.1927; **ν (KBr)/(cm^{-1})** 3476, 3319 (NH), 3062, 2947 (CH), 1689 (CO); **δH (200MHz, d_6-DMSO)** 2.87

(2H, d, βCH_2, Phe, J=8Hz), 4.15-4.18 (4H, m, CH & CH_2, Fmoc & αCH, Phe), 6.26 (2H, s, NH_2, semicarbazone), 7.20-7.99 (14H, m, ArH & HC=N), 9.93 (1H, s, NH, semicarbazone); **δC (63 MHz, d_6-DMSO)** 38.2, 46.8, 53.5, 65.6, 120.3, 125.3 (2C), 126.3 (2C), 127.8 (2C), 128.3 (2C), 129.3 (2C), 129.5 (2C), 138.2, 141.9 (2C), 143.9 (2C), 140.9, 156.9, 158.8; **$[\alpha_D]^{22}$** -24.7 ° (c 0.288, DMF); **λ_{max}/nm (MeOH, ε/$dm^3mol^{-1}cm^{-1}$)** 265 (17700), 289 (3600), 300 (4700).

Stability of 9-Fluorenylmethoxycarbonyl-*L*-Phenylalaninal Semicarbazone (19c)

9-Fluorenylmethoxycarbonyl-*L*-phenylalaninal semicarbazone **19b** (100 mg) was stirred in TFA/water (9:1, 1 ml) at RT for 1.5 hours. The TFA was removed *in vacuo* and the product isolated by precipitation from hexane.

Yield (0.095 g, 95 %); **m.p.** 143-145 °C; **t.l.c.** (DCM/MeOH/AcOH, 9:0.5:0.5) R_f 0.64; **m/z (HR FAB)** Found 429.1946; $C_{25}H_{25}N_4O_3$ requires 429.1927; **ν (KBr)/(cm^{-1})** 3321 (NH), 3063, 2948 (CH), 1689 (C=O); **δH (200MHz, d_6-DMSO)** 2.90 (2H, d, βCH_2, Phe, J=8Hz), 4.18-4.37 (4H, m, CH & CH_2, Fmoc & αCH, Phe), 6.26 (2H, s, NH_2, semicarbazone), 7.07-7.89 (14H, m, ArH & HC=N), 9.98 (1H, s, NH, semicarbazone); **δC (63 MHz, d_6-DMSO)** 38.2, 46.8, 53.4, 65.6, 120.2, 125.3 (2C), 126.3 (2C), 127.2 (2C), 127.7 (2C), 128.2 (2C), 129.3 (2C), 138.2, 141.9 (2C), 143.9 (2C), 140.8, 155.7, 156.9; **$[\alpha_D]^{22}$** -24.0 ° (c 0.325, DMF); **λ_{max}/nm (MeOH, ε/$dm^3mol^{-1}cm^{-1}$)** 265 (13500), 289 (4600), 300 (4800).

Hydrolysis of 9-Fluorenylmethoxycarbonyl-*L*-Phenylalaninal Semicarbazone (20)

9-Fluorenylmethoxycarbonyl-*L*-phenylalaninal semicarbazone **19b** (100 mg, 0.23 mmol) was dissolved in chloroform (8 ml). Pyruvic acid (2 ml) and water (0.3 ml) were added and the mixture stirred at room temperature for 24 hours. The mixture was diluted with chloroform (10 ml) and washed with water (3 x 2 ml), 5 % sodium bicarbonate solution (3 x 2 ml) and dried ($MgSO_4$). The solvent was evaporated to leave a white solid.

Yield (81 mg, 95 %); **m.p.** 102-103 °C; **t.l.c.** (DCM/MeOH/AcOH, 9:0.5:0.5) R_f 0.76; **m/z (HRFAB)** 372.1593(MH^+); $C_{24}H_{22}NO_3$ requires 372.1600; **νmax/cm^{-1}**

(KBr) 3331, 3321 (NH), 3061, 3020, 2958, 2843 (CH), 1726, 1690, 1638 (C=O); **δH (200MHz, d_6-DMSO)** 2.91 (2H, d, CH_2, Phe, J=8Hz), 4.06-4.36 (4H, m, CH & CH_2, Fmoc & αCH, Phe), 7.19-7.88 (13H, m, ArH), 9.51 (1H, s, CHO); **δC (63 MHz, d_6-DMSO)** 33.3, 46.8, 61.2, 65.6, 120.0, 120.2, 124.2, 125.2, 125.9, 126.4, 127.0, 127.1, 127.2, 127.7, 128.1, 128.3, 129.2, 137.9, 140.8, 143.7, 143.8, 143.9, 156.2, 200.9; **$[\alpha_D]^{22}$** -41.8 ^{0} (c 0.467, DMF); **λmax/nm (MeOH, ε/$dm^3mol^{-1}cm^{-1}$)** 265 (18400), 296 (3600), 300 (6100).

Synthesis of Fmoc-Phe-Val-(*L*)-Alaninal Semicarbazone (21)

The synthesis was carried out using resin **18b** (0.8530 g, 0.13 mmol/g, 0.11 mmol). On completion of the synthesis, the peptide was cleaved from the resin by sonication in TFA/water (95:5, 10 ml) for 1.5 hours. The resin was filtered and washed with TFA. The TFA was reduced *in vacuo* to 2 ml and cold ether (80 ml) was added dropwise to the yellow oil to yield a fluffy white precipitate which was collected by centrifugation. The precipitate was dissolved in acetonitrile/water (1:1, 30 ml) and lyophilised. The crude material was purified using preparative HPLC.

Final Fmoc-loading 0.079 mmol/g (62 %); **Crude Yield** 34 mg (102 %); **Purified Yield** 5 mg (25 %); **HPLC** (Vydac C_{18}, 250 x 4.6 mm, 5μm, A=H_2O, B=CH_3CN, 0.1 % TFA; 1 ml/min, 10-90 % over 30 minutes, λ=214 nm) R_t 22.4 min, 73.5 % B; **m/z (MALDI)** 599.206 (MH+), $C_{33}H_{39}N_6O_5$ requires 599.298; **AAA** Phe_1 1.0, Val_1 1.0.

Synthesis of Fmoc-Phe-Val-(*L*)-Alaninal (22)

Fmoc-Phe-Val-Alaninal semicarbazone **21** (1.4 mg) was dissolved in pyruvic acid/water (5:1, 300 μl) and sonicated for 1.5 hours. The product was isolated by HPLC.

Yield 0.65 mg (50 %); **HPLC** (Vydac C_{18}, 250 x 4.6 mm, 5μm, A=H_2O, B=CH_3CN, 0.1 % TFA; 1 ml/min, 10-90 % over 30 minutes, λ=214 nm) R_t 23.0 min, 76 % B; **m/z (MALDI)** 542.606 (MH+), $C_{32}H_{36}N_3O_5$ requires 542.655; **AAA** Phe_1 1.16, Val_1 0.86.

Synthesis of Fmoc-Phe-Val-(*D*)Alaninal Semicarbazone (23)

The synthesis was carried out using resin **18c** (0.900g, 0.10 mmol/g, 0.09 mmol). The peptide was cleaved from the resin by sonication in TFA/water (95:5, 10 ml) for 1.5 hours. The resin was filtered and washed with TFA. The TFA was reduced *in vacuo* to 2 ml and cold ether (80 ml) was added dropwise to the yellow oil to yield a fluffy white precipitate which was collected by centrifugation. The precipitate was dissolved in acetonitrile/water (1:1, 30 ml) and lyophilised. The crude material was purified using preparative HPLC.

Final Fmoc-Loading 0.053 mmol/g (55 %); **Crude Yield** 23.2 mg (84 %); **Purified Yield** 4 mg, 24 %; **HPLC** (Vydac C_{18}, 250 x 4.6 mm, 5μm, A=H_2O, B=CH_3CN, 0.1 % TFA; 1 ml/min, 10-90 % over 30 minutes, λ=214 nm) R_t 22.4 min, 73.5 % B; **m/z (MALDI)** 599.37 (MH+), $C_{33}H_{39}N_6O_5$ requires 599.30; **AAA** Phe_1 0.91, Val_1 1.09.

Synthesis of Fmoc-Phe-Val-(*D*)Alaninal (24)

Fmoc-Phe-Val-Alaninal semicarbazone **23** (1 mg) was dissolved in pyruvic acid/water (5:1, 300 μl) and sonicated for 1.5 hours. The product was isolated by HPLC.

Yield 0.5 mg (56 %); **HPLC** (Vydac C_{18}, 250 x 4.6 mm, 5μm, A=H_2O, B=CH_3CN, 0.1 % TFA; 1 ml/min, 10-90 % over 30 minutes, λ=214 nm) R_t 23.0 min, 76 % B; **m/z (MALDI)** 654.29 (CF_3COOM) $C_{34}H_{35}N_3O_7F_3$ requires 654.66; **AAA** Phe_1 1.13, Val_1 0.87.

Synthesis of Fmoc-Gly-Ala-Lys-Gly-Phenylalaninal Semicarbazone (25)

The synthesis was carried out using resin **18a** (1 g, 0.092 mmol/g, 0.092 mmol). On completion of the synthesis, the resin was stirred in a scavenger cocktail of EDT (0.20 ml), thioanisole (0.25 ml) TIS (0.25 ml) and phenol (0.75 g) for ten minutes, before the peptide was cleaved from the resin and deprotected by the addition of TFA/water (10:1, 11 ml). The mixture was stirred in the dark under dry nitrogen for 1.5 hours. The resin was filtered and washed with TFA. The TFA was reduced *in vacuo* to 2 ml and cold ether (80 ml) was added dropwise to the yellow oil to yield a

fluffy white precipitate which was collected by centrifugation. The precipitate was dissolved in acetonitrile/water (1:1, 30 ml) and lyophilised. The crude material (15 mg) was purified using preparative HPLC.

Final Fmoc-Loading 0.44 mmol/g (51 %); **Crude Yield** 49 mg (114 %); **Purified Yield** 6 mg (40 %); **HPLC** (ABI Aquapore C_{18}, 220 x 4.6 mm, 5μm, A=H_2O, B=CH_3CN, 0.1 % TFA; 1 ml/min, 10-90 % over 30 minutes, λ=214 nm) R_t 20.8 min, 65.5 % B; **m/z (MALDI)** 763.46 ($(M-H)^-Na^+$), $C_{38}H_{46}N_9O_7Na$ requires 763.82; **AAA** Gly_2 1.86, Ala_1 1.09, Lys_1 0.86.

Synthesis of Fmoc-Gly-Ala-Lys-Gly-Phenylalaninal (26)

Fmoc-Phe-Val-Alaninal semicarbazone **25** (5 mg) was dissolved in pyruvic acid/water (5:1, 1 ml) and stirred for 1.5 hours. The product was isolated by HPLC.

Yield 2.7 mg (62 %); **HPLC** (ABI Aquapore C_4, 100 x 4.6 mm, 5μm, A=H_2O, B=CH_3CN, 0.1 % TFA; 1 ml/min, 10-90 % over 30 minutes, λ=214 nm) R_t 17.2 min, 56.5 % B; **m/z (MALDI)** 761.74 $((M-H)K_2)^+$, $C_{37}H_{43}N_6O_7K_2$ requires 761.98; **AAA** Gly_2 1.86, Ala_1 1.10, Lys_1 0.97.

Synthesis of Fmoc-His-Leu-Asp-Ile-Ile-Tryptophanal Semicarbazone (27)

The synthesis was carried out using resin **18d** (0.60 g, 0.215 mmol/g, 0.129 mmol). On completion of the synthesis the resin was stirred in a scavenger cocktail of EDT (0.20 ml), thioanisole (0.25 ml) TIS (0.25 ml) and phenol (0.75 g) for ten minutes before the peptide was cleaved from the resin and deprotected by the addition of TFA/water (10:1, 11 ml). The mixture was stirred in the dark under dry nitrogen for 1.5 hours. The resin was filtered and washed with TFA. The TFA was reduced *in vacuo* to 2 ml and cold ether (80 ml) was added dropwise to the yellow oil to yield a fluffy white precipitate which was collected by centrifugation. The precipitate was dissolved in acetonitrile/water (1:1, 30 ml) and lyophilised. The crude material (15 mg) was purified using preparative HPLC.

Final Fmoc-Loading 0.134 mmol/g (74 %); **Crude Yield** 97 mg (114 %); **Purified Yield** 4 mg (27 %); **HPLC** (ABI Aquapore C_4, 100 x 4.6 mm, 5μm, A=H_2O, B=CH_3CN, 0.1 % TFA; 1 ml/min, 0-100 % over 30 minutes, λ=214 nm) R_t 21.1 min,

61 %B; **m/z (MALDI)** 1082.04 (MNa+), $C_{55}H_{71}N_{12}O_{10}Na$ requires 1082.23; **AAA** Asp_1 1.03, Ile_2 0.96, Leu_1 1.09, His_1 0.88.

Synthesis of Fmoc-His-Leu-Asp-Ile-Ile-Tryptophanal (28)

Fmoc-His-Leu-Asp-Ile-Ile-Tryptophanal semicarbazone **27** (1 mg) was dissolved in pyruvic acid/water (5:1, 1 ml) and stirred for 1.5 hours. The product was isolated by HPLC.

Yield 1.4 mg (38 %); **HPLC** (ABI Aquapore C_4, 100 x 4.6 mm, 5μm, A=H_2O, B=CH_3CN, 0.1 % TFA; 1 ml/min, 0-100 % over 30 minutes, λ=214 nm) R_t 18.9 min, 55% B; **m/z (MALDI)** 779.66 (M-Fmoc), $C_{39}H_{55}N_9O_3Na$ requires 779.94; **AAA** Asp_1 1.04, Ile_2 0.96, Leu_1 1.09, His_1 0.88.

Synthesis of the Semicarbazone of a Cathepsin-3 Inhibitor (29)

The synthesis was carried out using resin **18e** (0.893 g, 0.168 mmol/g, 0.15 mmol). The resin was stirred in a scavenger cocktail of EDT (0.20 ml), thioanisole (0.25 ml) TIS (0.25 ml) and phenol (0.75 g) for ten minutes before the peptide was cleaved from the resin and deprotected by the addition of TFA/water (10:1, 11 ml). The mixture was stirred in the dark under nitrogen for 1.5 hours then the resin was filtered and washed with TFA. The TFA was reduced *in vacuo* to 2 ml and cold ether (80 ml) was added dropwise to the yellow oil to yield a fluffy white precipitate which was collected by centrifugation. The precipitate was dissolved in acetonitrile/water (1:1, 30 ml) and lyophilised. The crude material (25 mg) was purified using preparative HPLC.

Final Fmoc-Loading 0.08 mmol/g (63 %); **Crude Yield** 93 mg (48 %); **Purified Yield** 7.25 mg (29 %); **HPLC** (Vydac C_{18}, 250 x 4.6 mm, 5μm, A=H_2O, B=CH_3CN, 0.1 % TFA; 1 ml/min, 40-100 % over 30 minutes, λ=214 nm) R_t 22.4 min, 85 %B; **m/z (MALDI)** 2079.52 ((M-H)$^-$Na$^+$), $C_{95}H_{160}N_{23}O_{27}Na$ requires 2079.44; **AAA** Asp_1 1.14, Glu_1 1.02, Pro_2 1.92, Ala_6 5.68, Val_3 2.80, Leu_6 5.99.

Synthesis of a Cathepsin-3 Inhibitor (30)

The Cathepsin-3 Inhibitor semicarbazone **29** (1 mg) was dissolved in pyruvic acid/water/acetonitrile (5:2:5, 1 ml) and sonicated for 2 hours. The product was isolated by HPLC.

Purified Yield 0.5 mg (50 %); **HPLC** (Vydac C_{18}, 250 x 4.6 mm, 5μm, A=H_2O, B=CH_3CN, 0.1 % TFA; 1 ml/min, 40-100 % over 30 minutes, λ=214 nm) R_t 30.5 min, 86 %B; **m/z (MALDI)** 1998.16 ($(M-H)^-$), $C_{94}H_{157}N_{20}O_{27}$ requires 1998.40; **AAA** Asp_1 1.14, Glu_1 1.02, Pro_2 1.92, Ala_6 5.68, Val_3 2.80, Leu_6 5.99.

4.3.2 The Stepwise Chemical Synthesis of Deglycosylated Human Interferon-Gamma and its Purification

HPLC of dhIFN-γ Samples

Analysis of all protein samples employed a C_4 column and the following conditions:

Aquapore C_4, 100 x 4.6 mm, 7 μm.

A = water, B = acetonitrile, 0.1 % TFA

2 ml loop, 1 ml/min

0-2 min 10 % B, 2-32 min 10-90 % B, 32-34 min 90-10 % B

λ = 214 nm

Loading of Fmoc-Gln(Trt)-OH onto Wang Resin

Fmoc-Gln(Trt)-OH (1 g, 1.64 mmol, 2 equivalents) and DIC (0.13 ml, 0.82 mmol, 1 equivalent) were dissolved in DMF (7 ml) and sonicated for 10 minutes. Wang Resin (1.08 g, 0.65 mmol/g) was swollen in the minimum amount of DMF, and a catalytic amount of DMAP (approx. 10 mg) added. The mixed anhydride was added to the resin, and the mixture allowed to stand for 20 mins. The resin was filtered, washed exhaustively with DMF, 1,4-dioxane, DCM and ether and allowed to dry *in vacuo*. The resin loading was checked by quantitative UV.

Resin Functionality 0.16 mmol/g

Chemical Synthesis of Fmoc-dhIFN-γ-Resin

The synthesis was carried out on a 0.112 mmol scale using Fmoc-Gln(Trt) functionalised Wang resin (700 mg, 0.116 mmol/g). Coupling was performed using the corresponding HOCt activated esters. The first amino acids were attached to the solid support *via* single coupling cycles with double coupling cycles being employed for the last 20 residues (with the exception of Gly). Approximately one third of the resin was removed after 71 cycles and a further one third was removed after 113 cycles. These portions were stored in 1,4-dioxane until the synthesis was recommenced using the resin, to yield a greater quantity of material. On completion

of the synthesis, the resin was washed exhaustively with DMF, 1,4-dioxane, DCM and ether and dried under vacuum. The resin from the synthesiser (1.667 g, 0.020 mmol/g, 0.033 mmol) was sonicated in capping reagent (0.5 M acetic anhydride, 0.125M DIEA, 0.2 % w/v HOBt in dry, freshly distilled DCM, 10 ml) for 30 minutes. The resin was filtered, washed with DCM and ether and dried *in vacuo.* The resin was stored at 4 ^{0}C in 1,4-dioxane until required.

Amount of Resin 1.667 g; **Final Functionality** 0.020 mmol/g; **AAA** (48 hours) Asx_{20} 19.0, Thr_5 5.3, Ser_{11} 9.7, Glx_{18} 19.6, Pro_2 1.4, Gly_5 5.0, Ala_8 8.3, Val_8 8.2, Met_4 4.5, Ile_7 7.0, Leu_{10} 9.6, Tyr_4 3.3, Phe_{10} 10.3, His_2 2.1, Lys_{20} 20.0, Arg_8 9.8.

Loading Tbfmoc to the dhIFN-γ-Resin

The *N*-terminal Fmoc group was removed from the resin (500 mg) by sonication in 20 % v/v piperidine in DMF (10 ml) for 30 minutes. The resin was filtered, washed exhaustively with DMF, 1,4-dioxane, DCM and ether and dried *in vacuo*. The Tbfmoc group was loaded onto the *N*-terminus by sonicating the resin, DIEA (5 µl, 0.03 mmol, 1 equivalent) and Tbfmoc chloroformate (41 mg, 0.09 mmol) in dry, distilled DCM (10 ml) for 3 hours in the dark. The Tbfmoc-loading was checked by UV. The resin was filtered, washed exhaustively with DCM and ether and dried *in vacuo*.

Amount of Resin 500 mg; **AAA** (48 hours) Asx_{20} 19.9, Thr_5 5.5, Ser_{11} 9.4, Glx_{18} 19.9, Pro_2 1.4, Gly_5 5.3, Ala_8 8.4, Val_8 8.0, Met_4 3.8, Ile_7 7.3, Leu_{10} 9.5, Tyr_4 3.3, Phe_{10} 10.7, His_2 1.6, Lys_{20} 19.9, Arg_8 9.6.

The Tbfmoc-Loading Test

The loading of the Tbfmoc functionalised resin was determined by treating an accurately weighed quantity of resin (8.0 mg) with 20 % v/v piperidine/1,4-dioxane (10 ml). After sonication for 10 minutes, the UV absorbance of the supernatant was recorded between 320 and 400 nm. The Tbfmoc-peptide-resin functionality was calculated from **equation 4.2**:

$$\text{Resin functionality (mmol/g)} = \frac{0.613 \times Abs_{364nm}}{\text{weight of resin (mg)}}$$

Equation 4.2: Determination of Tbfmoc-Loading Level

Tbfmoc-Loading Level by UV 0.0246 mmol/g (123 %).

Cleavage of the Tbfmoc-dhIFN-γ From The Resin

The resin (500 mg) was stirred in a cocktail of scavengers comprising 0.25 ml EDT, 0.5 ml thioanisole, 0.5 ml TIS and 0.75 g phenol in the dark, under dry nitrogen for 10 minutes. TFA/water (10:1, 11 ml) was added, and stirring in the dark under nitrogen was continued for a further 4.5 hours. The resin was filtered and washed with TFA (~ 2 ml). The filtrate was collected in ice cold ether (150 ml) which was allowed to stand in ice for 10 minutes to ensure complete precipitation of the protein. The white fluffy solid was collected by centrifugation (3000 rpm, 2 minutes). The pellet was washed with fresh portions of cold ether (3 x 40 ml) to remove residual scavengers, before being dissolved in acetonitrile/water (2:3, 50 ml) and lyophilised to a white solid.

Yield of protein 237 mg (141 %); **HPLC** 21.0 min, 66 % B; **AAA** (24 hours) Asx_{20} 19.1, Thr_5 4.4, Ser_{11} 9.8, Glx_{18} 18.2, Pro_2 1.0, Gly_5 4.9, Ala_8 8.2, Val_8 8.4, Met_4 3.0, Ile_7 6.8, Leu_{10} 10.0, Tyr_4 2.9, Phe_{10} 10.6, His_2 2.4, Lys_{20} 19.1, Arg_8 9.9; **SDS-PAGE** 4 components of approx. Mwt:17 kDa, 14 kDa, 11 kDa & 8 kDa

Tbfmoc-Charcoal Purification

Animal charcoal was prewashed using piperidine/6 M urea/IPA (2:9:9, 6 x 50 ml), and 6 M urea/IPA (1:1, 6 x 50 ml). Tbfmoc-dhIFN-γ (100 mg) was dissolved in 6M urea (20 ml), and a reference HPLC trace of the sample at 364 nm was obtained. The concentration of the solution was also checked by UV. Prewashed charcoal was added in small portions with vortexing (10 mins) and centrifugation (3500 rpm, 10 mins), until examination of the supernatant by HPLC at 364 nm indicated complete adsorption of the Tbfmoc-dhIFN-γ onto the charcoal had taken place. It was found that 25 mg of charcoal per mg of protein was sufficient for complete adsorption. The supernatant was removed and the pellet washed by vortexing with fresh solvent (6 M

urea/IPA, 3 x 25 ml) to ensure complete removal of deletions, truncates and any material still in solution. The Tbfmoc-protein bond was cleaved by vortexing the charcoal in 10 % piperidine in 6 M urea/IPA (1:1, 25 ml) for 10 mins. The supernatant was examined by HPLC at 214 nm to ensure the free dhIFN-γ was released into the solution. The IPA was removed *in vacuo* before the solution was neutralised to pH 6.0 using acetic acid (~ 3 ml).

HPLC 19.7 min, 58 % B.

Tbfmoc-Polystyrene Purification

Tbfmoc-dhIFN-γ (100 mg) was dissolved in 25 % acetic acid solution (25 ml), and a reference HPLC trace at 364 nm obtained. Chromatography grade polystyrene (30-75 μm mesh, 300 Å pore diameter) was added in small portions with vortexing (10 mins) and centrifugation (3500 rpm, 10 mins) until the supernatant, when examined by HPLC at 364 nm, indicated complete adsorption of the Tbfmoc-protein. It was found that 20 mg of polystyrene was required per mg of crude Tbfmoc-protein. The supernatant was decanted, and the polystyrene washed with acetic acid (2 x 25 ml) and water (25 ml) in order to remove all deletions, truncates and any material still in solution. The Tbfmoc-protein bond was cleaved by sonication of the polystyrene in 10 % piperidine in acetonitrile/water (2:3) (25 ml). This solution was examined by HPLC at 214 nm to ensure the dhIFN-γ had been released into the solution. The piperidine was neutralised to pH 6.0 by the addition of acetic acid (~ 3 ml). 25 % w/w acetic acid solution (15 ml) was added solubilise the protein.

HPLC 19.7 min, 58 % B.

Sephadex G-50 Gel Filtration of dhIFN-γ

Gel filtration was carried out using a Sephadex G-50 column (60 cm x 3 cm diam.), equilibrated with 25 % acetic acid solution. The dhIFN-γ solution from charcoal or polystyrene purification was loaded and the protein eluted using 25 % acetic acid solution. Fractions (20 ml) were collected and the purification monitored by UV (280 nm). The protein was found to elute in fractions 9 to 20. An aliquot of UV

active fractions was lyophilised and examined by SDS-PAGE for protein composition. The fractions with the highest purity protein profile were combined and lyophilised.

After Charcoal/Sephadex G-50

Yield of protein 25 mg; **HPLC** 19.5 min, 57 % B; **AAA** (24 hours) Asx_{20} 19.8, Thr_5 4.6, Ser_{11} 9.4, Glx_{18} 17.8, Pro_2 1.3, Gly_5 5.0, Ala_8 8.1, Val_8 8.9, Met_4 3.6, Ile_7 7.4, Leu_{10} 10.0, Tyr_4 3.3, Phe_{10} 10.8, His_2 2.6, Lys_{20} 19.8, Arg_8 8.5; **SDS-PAGE** 3 components of approx. Mwt: 17 kDa, 14 kDa & 11 kDa.

After Polystyrene

Yield of protein 40 mg; **HPLC** 19.7 min, 58 % B; **AAA** (24 hours) Asx_{20} 21.1, Thr_5 4.9, Ser_{11} 9.9, Glx_{18} 18.7, Pro_2 1.1, Gly_5 4.1, Ala_8 7.5, Val_8 10.3, Met_4 3.1, Ile_7 7.6, Leu_{10} 10.0, Tyr_4 1.5, Phe_{10} 11.6, His_2 2.0, Lys_{20} 22.2, Arg_8 8.1; **SDS-PAGE** 3 components of approx. Mwt: 17 kDa, 14 kDa & 11 kDa.

Electroblotting of dhIFN-γ onto PVDF Membrane[9]

dhIFN-γ was separated using SDS-PAGE and the separated bands transferred to PVDF membrane to allow analysis. Protein (0.5 mg) was dissolved in sample buffer (150 µl) and applied to 9 wells of the SDS-polyacrylamide gel. On completion of the run, the gel was stored in running buffer. The gel was soaked in (100 mM CAPS/ methanol/water 1:1:8) for 5 minutes. The blotting apparatus was assembled according to the manufacturers instructions and the cell run at 50 V for 2 hours at room temperature. The membrane was rinsed with water and stained with 0.5 % w/v coomassie blue in acetic acid/methanol/water (1:4:5) for 15 minutes. Destaining in acetic acid/methanol/water (1:4:5) was carried out at 37 ^{O}C overnight. The desired bands were cut from the PVDF membrane prior to analysis.

Digestion With Pyroglutamate Aminopeptidase

Protein (1 mg) was dissolved in 0.2 M ammonium bicarbonate, 10 mM EDTA, 5 mM DTT, 5 % v/v glycerol at pH 8 (240 µl) with vortexing. Pyroglutamate amino peptidase (1.5 mg) was added as a solution in the above buffer (250 µl). The digest

was incubated overnight at 37 ^{0}C, with gentle agitation. The protein material was isolated using RP HPLC and lyophilised.

Yield 0.6mg

FPLC Size Exclusion Purification of dhIFN-γ

Gel filtration was carried out using Superdex^{TM} 75 HR 26/60 column equilibrated with 6 M urea, 0.2 M NaCl, 0.1 M phosphate, pH 7.5 using a flow rate of 3 ml/min. The dhIFN-γ (5 mg) was dissolved in 6 M urea, 0.2 M NaCl, 0.1 M phosphate, pH 7.5 (2 ml) and loaded onto the column and a flow rate of 3ml/min applied. After 100ml of eluent had been passed through, fractions (6 ml) were collected. Progress was monitored using UV, and fractions absorbing at 280 nm were desalted by dialysis against acetic acid (25 % v/v), lyophilised and examined by SDS-PAGE. Fractions 18 to 23 were found to give the cleanest SDS-PAGE profile, these were combined, dialysed against 25 % acetic acid solution and lyophilised.

Yield 2 mg; **HPLC** 16.9 min, 55 % B; **AAA** (24 hours) Asx_{20} 21.7, Thr_5 4.2, Ser_{11} 10.4, Glx_{18} 18.7, Pro_2 1.1, Gly_5 4.9, Ala_8 6.6, Val_8 8.5, Met_4 2.7, Ile_7 8.2, Leu_{10} 10.3, Tyr_4 2.6, Phe_{10} 13.2, His_2 1.7, Lys_{20} 22.3, Arg_8 7.0; **SDS-PAGE** 2 components of approx Mwt: 17 kDa & 14 kDa.

Cation Exchange Chromatography Using CM Sephadex CL-50 (I)

CM Sephadex CL-50 (10 ml) was preswollen in 8M Urea, 80 mM Tris, pH 7.2. dhIFN-γ (5 mg, 3 component) was dissolved in 8 M urea, 80 mM Tris, pH 7.2 (5 ml), and was adsorbed to the gel by vortexing the solution with a small amount of the gel and the supernatant examined by HPLC for adsorption. It was found that 8 ml of ion exchange resin was required for complete adsorption. The gel was poured into a column (0.7 cm diam. x 21 cm high), and the urea removed over a 40 ml gradient. A sodium chloride gradient (100 ml, 0-1 M) in 80 mM Tris, pH 7.2 was introduced to elute the protein from the resin. The collected fractions (5 ml) were examined by HPLC for protein content but no protein was observed. The column was washed with 3M NaCl, 80 mM Tris, pH 7.2 (10 ml) to establish if the protein had bound

strongly to the support but again, no protein was observed. The column was then washed with 8 M urea, 80 mM Tris, pH 7.2 (15 ml). This eluted the protein from the column. The protein, on examination by SDS-PAGE following dialysis against 25 % acetic acid, was found to contain three bands of molecular mass 11, 14 & 17 kDa.

Cation Exchange Chromatography Using CM Sephadex CL-50 (II)

CM Sephadex CL-50 (10 ml) was preswollen in 7M Urea, 80 mM Tris, pH 7.2. dhIFN-γ (5 mg, 4 component) was dissolved in 7 M urea, 80 mM Tris, pH 7.2 (5 ml) and was adsorbed to the gel by vortexing the solution with a small amount of the gel and the supernatant examined by HPLC for adsorption. It was found that 5 ml of ion exchange resin was required for complete adsorption. The gel was poured into a column (0.7 cm diam x 14 cm high) and washed with 7 M urea, 80 mM Tris, pH 7.2 (5 ml). A sodium chloride gradient (25 ml, 0-1 M) in 7 M urea, 80 mM Tris, pH 7.2 was introduced to elute the protein from the resin. The collected fractions (1 ml) were examined by UV (280 nm) and HPLC for protein content. Protein containing fractions were dialysed individually against 25 % acetic acid solution and lyophilised. They were examined by SDS-PAGE for purity. Fractions 7-10 were found to contain protein bands of approx. mass 8, 11, 14 & 17 kDa.

Cation Exchange Chromatography Using CM Sephadex CL-50 (III)

CM Sephadex CL-50 (25 ml) was preswollen in 6M Urea, 50 mM Tris, pH 7.5. dhIFN-γ (10.6 mg, 4 components) was dissolved in 6 M urea, 50 mM Tris, pH 7.5 (20 ml) and was adsorbed to the gel by vortexing. The gel was poured into a column (2 cm diam x 13 cm high) and washed with 6M urea, 50 mM Tris, pH 7.5 (100ml). A sodium chloride gradient (320 ml, 0-0.6 M) was introduced to elute the protein from the resin. The collected fractions (5 ml) were examined by UV (280 nm) for protein content. Protein containing fractions were dialysed against 25 % acetic acid solution and lyophilised. On examination by SDS-PAGE, fractions 11-19 were found to contain protein bands of approx. mass 8, 11 & 14 kDa and fractions 21-28 were found to contain protein bands of approx. mass 11, 14 & 17 kDa.

Cation Exchange Chromatography Using SP Sepharose FF

SP Sepharose FF (25 ml) was equilibrated with 6M Urea, 20 mM Phosphate, pH 7.0. dhIFN-γ (10 mg, 4 components) was dissolved in 6 M urea, 50 mM Tris, pH 7.5 (20 ml) and was adsorbed to the gel by vortexing. The gel was poured into a column (2 cm diam x 13 cm high) and washed with 6M urea, 20 mM Phosphate, pH 7.0 (25 ml). A sodium chloride gradient (400 ml, 0-1 M) was introduced to elute the protein from the resin. The collected fractions (20 ml) were examined by UV (280 nm) for protein content. Protein containing fractions were dialysed against 25 % acetic acid solution and lyophilised. On examination by SDS-PAGE, fractions 7-14 were found to contain protein bands of approx. mass 11, 14 & 17 kDa.

Sephadex G-75 Gel Filtration

Gel filtration was carried out using a Sephadex G-50 column (90 cm x 3 cm diam.) equilibrated with 20 % acetic acid solution. The dhIFN-γ solution from charcoal or polystyrene purification was loaded in portions (7 mg) as a solution in 6 M urea, 0.2 M NaCl, 0.1 M sodium phosphate, pH 7.5 (400μl). The protein was eluted using 20 % acetic acid solution. Fractions were collected (60 minutes, approx 10ml) and examined by UV (280 nm). The protein was found to elute in fractions 7 to 18. These fractions were examined by SDS-PAGE, and the fractions which gave the cleanest protein profile of 2 bands (17kDa and 14 kDa) were combined and lyophilised and reapplied to the column as before. On examination of the second batch of fractions by SDS-PAGE, it was found that fractions 9 and 10 were generally pure protein.

Yield 0.75 - 1.5 mg; **HPLC** 19.5 min, 62 % B; **AAA** (24 hours) Asx_{20} 19.7, Thr_5 5.1, Ser_{11} 9.7, Glx_{18} 19.3, Pro_2 1.8, Gly_5 5.0, Ala_8 7.3, Val_8 8.6, Met_4 3.3, Ile_7 7.6, Leu_{10} 10.2, Tyr_4 3.0, Phe_{10} 12.2, His_2 2.0, Lys_{20} 20.7, Arg_8 7.8; **MALDI-TOF MS** 16826 Da, requires 16777 Da; **SDS-PAGE** one component of approx. MWt 17 kDa. **N-Terminal Sequencing** Gln (75.56 pmol), Asp (79.23 pmol), Pro (62.96 pmol), Tyr (52.37 pmol), Val (53.19 pmol)

Determination of Tryptophan[10]

Prior to the absorbance measurement, a blank sample (containing buffer solution) was run. From amino acid analysis, the tyrosine content of the protein is approximately 4, as expected. The protein solution had a concentration of 5.96 x 10^{-6} mmol/ml (in 6M urea solution). The average value for the tryptophan content was determined to be 0.99 moles of tryptophan per mole of protein, using the values depicted in **table 4.2** and **equations 3.1** and **3.2**.

Wavelength/nm	Abs	ε (calculated)	N_{Trp} (calculated)
280	0.63	10570	0.73
288	0.47	7886	1.24

Table 4.2: Data For Tryptophan Determination

FPLC Determination of Molecular Weight

Analysis was carried out using a Superdex^{TM} 75 HR 10/30 column which had been equilibrated with 6M urea using a flow rate of 0.5 ml/min. Blue Dextrin (Mwt 2 000 000) was loaded to measure the column void volume (V_o 8.05). 0.5 mg of protein was dissolved in 6M urea (200 µl), loaded onto the column and the elution volumes (V_e) recorded by monitoring at 280 nm. The molecular weight standards used were ovalbumin (Mwt 43 000, V_e 8.22), chymotrypsin (Mwt 25 000, V_e 8.57), ribonuclease A (Mwt 13 700, V_e 9.04) and ubiquitin (Mwt 8 600, V_e 9.27). For each standard, the elution parameter k_{AV} was calculated using **equation 3.3**. A graph, **figure 3.2.16**, of k_{AV} *versus* the logarithm of the molecular weight for each standard, was used to determine the molecular weight of dhIFN-γ. The elution volume for dhIFN-γ was determined to be 8.89 ml, giving a molecular weight of approximately 16 700 Da.

Tryptic Digest of dhIFN-γ

dhIFN-γ (0.5 mg) was dissolved in 6 M urea, 0.2 M phosphate pH 8.0 (200 µl) and trypsin (5 % w/w) added. The mixture was incubated at 37 ^{0}C, and after 2 hours and 6 hours, aliquots were removed and the digest stopped by the addition of 6 M HCl (5

μl). The samples were desalted on a RP HPLC Vydac C_{18} analytical column and the collected peaks examined by MALDI-TOF MS.

Found	430.419	^{59}Asn-^{61}Lys(Na^+ salt)	**Requires**	430.459
	657.389	^{39}Glu-^{42}Arg(Na^+ salt)		657.591
	674.639	^{139}Arg-^{143}Gln(H^+)	674.738	
	714.122	^{74}Lys-^{69}Ser(K^+ salt)		714.878
	907.532	^{87}Lys-^{81}Phe(Na^+ salt)		906.992
	994.658	^{132}Ser-^{139}Arg		994.180
	1432.390	^{69}Ser-^{80}Lys(K^+ salt)		1431.685
	1610.830	^{56}Leu-^{58}Lys		1610.680
	3447.620	^{13}Lys-^{42}Arg(Na^+ salt)		3448.752
		or ^{14}Tyr-^{43}Lys(Na^+ salt)		3448.752
	3455.010	^{95}Leu-^{125}Lys(H^+)		3457.580

Isoelectric Point Determination Using IEF

dhIFN-γ (10 mg) was dissolved in 4M urea (50 ml) and ampholytes (Fluka, pH range 9-11, 40 % w/v, 1 ml) were added. The solution was loaded into the focusing chamber of the Rotofor cell after mixing. The cell was rotated without applied power to allow the system to reach thermal equilibrium (~5 ^{O}C). After 10 minutes, the focusing was carried out at a constant power of 15 W. A large increase in voltage was initially observed which gradually stabilised over 3.5 hours. Once the voltage became constant, the 20 fractions were harvested and their individual pH recorded to ensure that a pH gradient had been established. All fractions were examined by HPLC to assess the protein content, **table 4.3**. The majority of protein was found in fractions 15 and 16, indicating a pI of approximately 9.5 (theoretical = 9.60[11]).

Fraction	pH	Fraction	pH
1	6.5	11	9.5
2	7.0	12	9.5
3	7.5	13	10.0
4	8.0	14	10.0
5	8.5	15	10.5
6	8.5	16	10.5
7	9.0	17	10.5
8	9.0	18	10.5
9	9.5	19	11.0
10	9.5	20	11.0

Table 4.3: pH Gradient Established by Isoelectric Focusing

Folding of dhIFN-γ[12]

dhIFN-γ (1 mg) was dissolved in 6M urea (167 µl) and stirred at 0^0C overnight. The solution was diluted to give a final concentration of 1mg protein per ml with 10mM ammonium acetate, pH 6.8, 0.38g/L PEG and stirring at 0^0C continued for a further 48 hours. Dimer formation was observed by non-reducing PAGE.

Non-Reducing PAGE: two components of approx. MWt 34 & 17 kDa.

4.7 References

1 L. Jiang, A. Davison, R. Ramage & G. Tennant, *Tetrahedron*, **1998**, *54*, 14233-14254.

2 R. Ramage & G. Raphy, *Tetrahedron Lett.*, **1992**, *33*, 385-388.

3 J. D. Hayes, L. A. Kerr & A. D. Cronshaw, *Biochem. J.*, **1989**, *264*, 437-445.

4 U. K. Laemmli, *Nature*, **1970**, *227*, 680-685.

5 R. E. Zhang, Y.-L. Cao & M. W. Hearn, *Anal. Biochem.*, **1991**, *195*, 160-170.

6 J.J. Wen & C. M. Crews, *Tetrahedron Asymmetry*, **1998**, *9*, 1855-1858.

7 P. T. Ho & K-Y Ngu, *J. Organic Chem.*, **1993**, *58*, 2313-2316.

8 J.A. Fehrentz, C. Pothion, J.-C. Califano, A. Loffet & J. Martinez, *Tetrahedron Lett.*, **1994**, *35*, 9031-9034.

9 P. Matsudaira, *J. Biol. Chem.*, **1987**, *261*, 10035-10038.

10 H. Edelhoch, *Biochemistry*, **1967**, *6*, 1948-1955.

11 Calculated using the program PEPTIDE, Lighthouse Data.

12 J. L. Cleland, S. C. Builder, J. R. Swartz, M. Winkler, J. Y. Chang & D. I. C. Wang, *Bio/Technology*, **1992**, *10*, 1013-1019.

Appendix

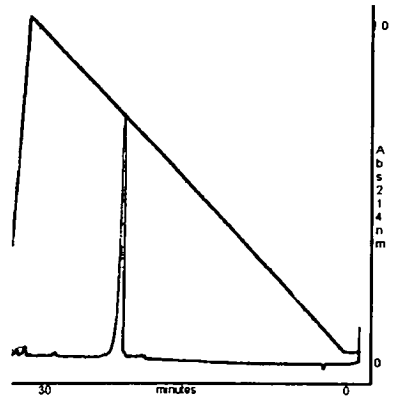


Figure A1: HPLC Profile For Fmoc-Gly-Ala-Lys-Gly-Phe-**sc**

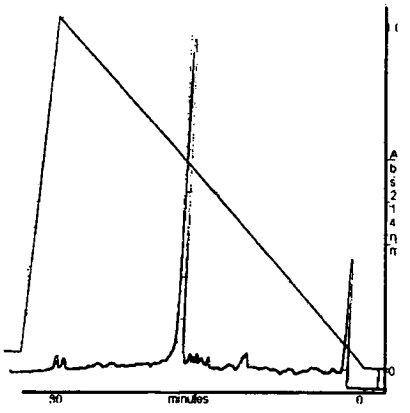


Figure A2: HPLC Profile For Fmoc-Gly-Ala-Lys-Gly-Phe-H

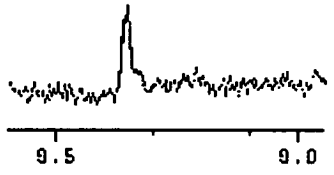


Figure A3: Aldehydic Proton NMR Signal For Fmoc-Gly-Ala-Lys-Gly-Phe-H

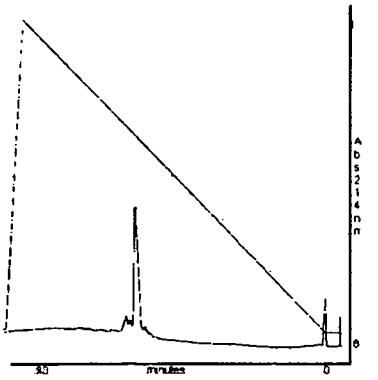


Figure A4: HPLC Profile For Fmoc-His-Leu-Asp-Ile-Ile-Trp-**sc**

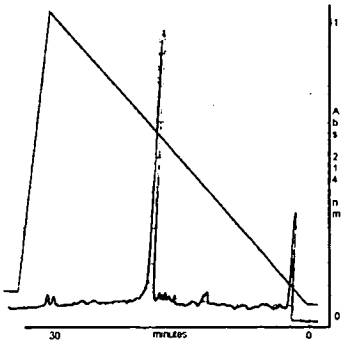


Figure A5: HPLC Profile For Fmoc-His-Leu-Asp-Ile-Ile-Trp-H

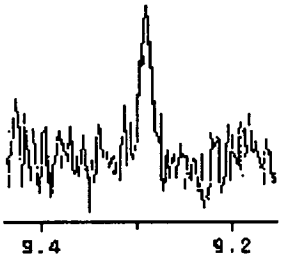


Figure A6: Aldehydic Proton NMR Signal For Fmoc-His-Leu-Asp-Ile-Ile-Trp-H

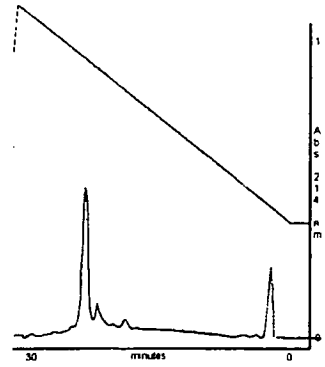


Figure A7: HPLC Profile For Caspase 3 Inhibitor Semicarbazone

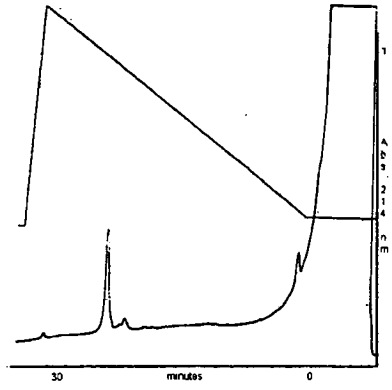


Figure A8: HPLC Profile For Caspase 3 Inhibitor

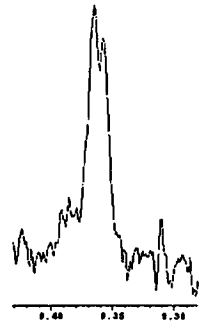


Figure A9: Aldehydic Proton NMR Signal For Caspase 3 Inhibitor

Lectures, Courses and Conferences Attended

"Effective Tutoring - Stage 1", organised by the Teaching and Learning Association, University of Edinburgh, 1996.

Organic Chemistry Seminars, University of Edinburgh, 1996-1999, various speakers.

Organic Chemistry Colloquia, University of Edinburgh, 1996-1999, various speakers.

Royal Society of Chemistry, Perkin Division, Scottish Meeting, University of Edinburgh 1996; Strathclyde University 1997; University of St Andrews 1998, various speakers.

"Protecting and Commercialising Inventions", organised by Zeneca Agrochemicals, University of Edinburgh, 1997, various speakers.

"Medicinal Chemistry", organised by Merck, Sharp & Dohme, University of Edinburgh, 1997, 1998, 1999, various speakers.

"Current Awareness in Organic Chemistry", sponsored by Zeneca, Grangemouth, University of Edinburgh, 1997, 1998, 1999, various speakers.

Edinburgh Centre for Protein Technology (ECPT) lectures, University of Edinburgh, 1997-1999, various speakers.

Solid Phase Synthesis and Combinatorial Chemical Libraries, 5^{th} International Symposium, London 1997; 6^{th} International Symposium, York 1999.

"Protein Folding", Prof. C. Dobson (Oxford University), Glasgow University, 1997.

Walker Memorial Lectures, University of Edinburgh, Prof B. F. G. Johnson (University of Cambridge) 1997, Prof J.-M. Lehn (Pasteur Institute, Strasbourg) 1998, Dr T. McKillop (Zeneca Pharmaceuticals) 1999.

"NMR Spectroscopy", University of Edinburgh, 1998, Drs Sadler, Barlow, Reed, Uhrin, Hewage and Parkinson.

"Safety Lectures", organised by Merck, University of Edinburgh, 1998, various speakers.

"Synthons in Organic Chemistry", Prof. E. Vilsmeier (University of Kaiserslautern), University of Edinburgh, 1998.

"From Proteomics to Proteins", organised by Perseptive Biosystems, Edinburgh 1998, 1999, various speakers.

Society of the Chemical Industry Graduate Symposium in Novel Organic Chemistry, University of Edinburgh 1998; University of Glasgow 1999, various speakers.

"Understanding Protein-Protein Interactions", organised by the Society of the Chemical Industry, London, 1998, various speakers.

Edinburgh Centre For Protein Technology Symposium, University of Edinburgh, 1998, various speakers.

25th International Symposium of the European Peptide Society, Budapest, 1998.

15th International Conference in Medicinal Chemistry, Edinburgh, 1998.

Royal Society of Chemistry Postgraduate Tour of Industry, Kent, 1998.

Romanes Symposium, University of Edinburgh, 1998, Prof. K. C. Nicolaou (Scripps Research Institute), Dr K Hale (University College, London), Prof. P. J. Kocienski (Glasgow University).

"Amino Acids in Organic Chemistry", Dr J. Podlech (University of Stuttgart), University of Edinburgh, 1999.

2nd International Conference in Biological Challenges For Organic Chemistry, St Andrews, 1999.

Presentations

"Studies Towards The Chemical Synthesis of Deglycosylated Human Interferon-Gamma", Organic Chemistry Seminars, 1997.

"New Aromatic Substitution Reactions", Lecture as part of the "Current Awareness in Organic Chemistry" series sponsored by Zeneca, Grangemouth, 1998.

"Novel Methods for the Modification of the C-Terminus of Peptides and Proteins", Lecture presentation at the Edinburgh Centre For Protein Technology Seminar, 1998.

"The Stepwise Solid Phase Synthesis of Deglycosylated Human Interferon-Gamma and Its Purification", Poster presentation at the 25th European Peptide Symposium, Budapest, 1998.

"The Chemical Synthesis of Proteins and Peptide C-Terminal Derivatives", Organic Chemistry Seminars, 1999.

"The Chemical Synthesis of Deglycosylated Human Interferon-Gamma and its Purification", Poster presentation at Biological Challenges For Organic Chemistry II, St Andrews, 1999.

"The Chemical Synthesis of Deglycosylated Human Interferon-Gamma and its Purification", Poster presentation at 6th International Conference in Solid Phase Synthesis and Combinatorial Chemical Libraries, York, 1999.

"The Chemical Synthesis of Deglycosylated Human Interferon-Gamma and its Purification", Lecture presentation at 6th International Conference in Solid Phase Synthesis and Combinatorial Chemical Libraries, York, 1999. This presentation was awarded a Young Scientific Investigator Award, sponsored by NovaBiochem.

Publications

J. A. Patterson, L. C. Draffan & R. Ramage in "*Peptides 1998*", Proc. 25th European Pept. Symp., S. Bajusz & F. Hudecz eds., Akadémiai Kidaó, Budapest, **1999**, pp162-163.

J. A. Patterson & R. Ramage, *Tetrahedron Lett.*, **1999**, *40*, 6121-6124.

J. A. Patterson & R. Ramage in "*Innovation and Perspectives in Solid Phase Synthesis: Peptides, Proteins and Nucleic Acids*", Proc. 6th Int. Symp. Solid Phase Synthesis & Combinatorial Chemistry Libraries, R. Epton ed., Mayflower Worldwide, **2000**, submitted.

Pergamon

Tetrahedron Letters 40 (1999) 6121–6124

TETRAHEDRON LETTERS

Solid Phase Synthesis of Peptide *C*-Terminal Semicarbazones and Aldehydes

Jennifer A. Patterson and Robert Ramage*

The Edinburgh Centre For Protein Technology, Department of Chemistry, The University of Edinburgh, West Mains Road, Edinburgh, EH9 3JJ, UK.

Received 27 May 1999; accepted 22 June 1999

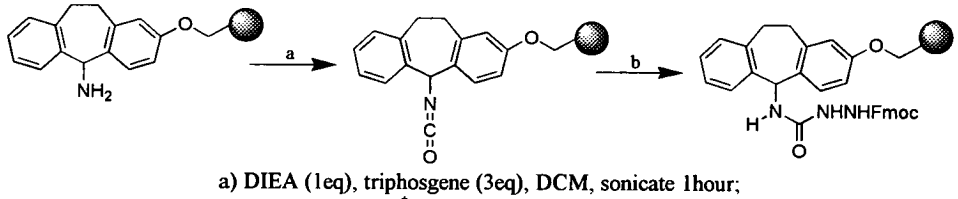
Abstract: A new linker based on the dibenzosuberyl system was developed in order to synthesise peptide *C*-terminal semicarbazones which can be readily converted into peptide *C*-terminal aldehydes. The method uses Fmoc-methodology and proceeds with no loss of stereochemical integrity. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: solid phase synthesis, peptide analogues/mimetics, enzyme inhibitors

Several examples of peptide aldehydes have been found to be potent inhibitors of enzymes including serine,[1-2] cysteine[3-4] and aspartyl[5-6] proteases and prohormone convertases.[7] Thus it is highly desirable to develop reliable routes to peptide *C*-terminal aldehydes which can also be used further in a wide range of chemistry including chemical ligation[8] and generation of reduced bond peptide isosteres.[9]

Although there are several methods for the synthesis of peptide aldehydes using SPPS,[10-18] we sought to extend the versatility of the dibenzocyclohept-1,4-diene (dibenzosuberyl) linker previously reported,[19] by introducing a semicarbazide moiety which would allow the synthesis and isolation of peptide *C*-terminal semicarbazones. Such derivatives are inherently more stable and easier to purify than the corresponding peptide aldehydes and, indeed, could have interesting biological properties. These peptide semicarbazones may be stored at 4°C until conversion into the peptide aldehyde is required.

The requisite semicarbazide linker can be synthesised from the corresponding amide linker in two simple steps (Figure 1) the course of which can be followed using IR and, in this way, routinely functionalities of 0.2-0.25 mmol/g (by UV determination of Fmoc) can be obtained.



a) DIEA (1eq), triphosgene (3eq), DCM, sonicate 1hour;
b) $FmocNHNH_2$† (2.5eq), DCM, sonicate 2 hours

Figure 1

* Corresponding Author: Tel +00 44 131 650 4720, Fax +00 44 131 667 7942, email R.Ramage@ed.ac.uk

† Fmoc-hydrazine is prepared according to the method outlined in Zhang, Z. E.; Cao, Y. L.; Hearn, M. W. *Anal. Biochem.* **1991**, *195*, 160-170.

0040-4039/99/$ - see front matter © 1999 Elsevier Science Ltd. All rights reserved.
PII: S0040-4039(99)01219-8

The *C*-terminal residues, Fmoc-protected amino aldehydes, were derived from Fmoc-amino acids as previously reported.[20-21] These were loaded onto the linker, in the presence of DIEA, in good yields based on start and end functionalities (Table 1).

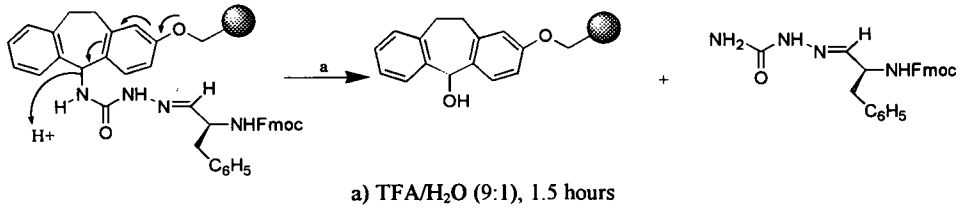
Compound	Loading Level (%)	Loading Time (Hours)
Fmoc-(L)Ala-H	100	5
Fmoc-(D)Ala-H	70	5
Fmoc-Phe-H	90	5
Fmoc-Trp-H	100	4
Fmoc-Asp(O^{t}Bu)-H	90	5

Table 1

To determine the extent, if any, of racemisation taking place under cleavage conditions (Figure 2), the stability of Fmoc-phenylalaninal semicarbazone to TFA treatment was checked (Table 2). It was proposed to use pyruvic acid exchange to convert the semicarbazones into aldehydes. This step was also examined for racemisation (Table 2). For each set of compounds identical tlc, MS, IR, ^{1}H and ^{13}C NMR were also obtained.

Compound	$[\alpha]_D$ (c g/100ml DMF)	Mpt (^{0}C)
Fmoc-Phenylalaninal Semicarbazone (Initially)	-24.7^0 (0.288)	144-145
Fmoc-Phenylalaninal Semicarbazone (TFA Treated)	-24.0^0 (0.325)	143-145
Fmoc-Phenylalaninal (Reduction of Weinreb Amide)	-43.3^0 (1.146)	100-102
Fmoc-Phenylalaninal (Pyruvic Acid Exchange)	-41.8^0 (0.467)	102-103

Table 2



a) TFA/H_2O (9:1), 1.5 hours

Figure 2

As a final check, the test peptide sequences Fmoc-Phe-Val-(L)Ala-H and Fmoc-Phe-Val-(D)Ala-H were synthesised. It has previously been reported[22] that racemisation in a peptide aldehyde, containing three residues, or more, will be indicated by more than one signal for the aldehydic protons in the NMR spectrum. For each of these peptides only one signal (at 9.50 ppm for the L-isomer and at 9.34 ppm for the D-isomer) was observed.

A series of test peptide semicarbazones have been prepared to assess the success of this method (Table 3) and these were subsequently converted to the corresponding peptide aldehydes (Table 4). Among the sequences prepared was an inhibitor of caspase 3[23-24] (final entry, Table 4).

Sequence‡	Yield§	Mass¶ (Found)	Mass (Calc)	AAA (24 Hrs)
Fmoc-FV(L)A-sc	25	599.21 (MH^+)	599.71	$Phe_1$1.00, $Val_1$1.00
Fmoc-FV(D)A-sc	24	599.37 (MH^+)	599.71	$Phe_1$0.91, $Val_1$1.09
Fmoc-GAKGF-sc	40	763.46 $(M-H)Na^+$	763.83	$Gly_2$1.86, $Ala_1$1.09, $Lys_1$0.98
Fmoc-HLDIIW-sc	27	1082.04 (MNa^+)	1082.23	$Asp_1$1.03, $Ile_2$0.96, $Leu_1$1.09, $His_1$0.88
Ac-AAVALLPAVL LALLAPDEVD-sc	29	2079.52 (MNa^+)	2079.41	$Asp_1$1.14, $Glu_1$1.02, $Pro_2$1.92, $Ala_6$5.68, $Val_3$2.80, $Leu_6$5.99

Table 3

Sequence	Yield^	Mass¶ (Found)	Mass (Calc)	AAA (24Hrs)	NMR†
Fmoc-FV(L)A-H	50	542.61 (MH^+)	542.65	$Phe_1$1.16, $Val_1$0.86	9.50 ppm
Fmoc-FV(D)A-H	56	654.29 ($M^+CF_3CO_2^-$)	654.66	$Phe_1$1.13, $Val_1$0.87	9.34 ppm
Fmoc-GAKGF-H	62	761.74 $(M-H)K_2^+$	761.98	$Gly_2$1.86, $Ala_1$1.10, $Lys_1$0.97	9.34 ppm
Fmoc-HLDIIW-H	38	779.66 (MH^+-Fmoc)	779.94	$Asp_1$1.04, $Ile_2$0.96, $Leu_1$1.09, $His_1$0.88	9.29 ppm
Ac-AAVALLPAVL LALLAPDEVD-H	50	1998.16 $(M-H)^-$	1998.35	$Asp_1$1.07, $Glu_1$1.05, $Pro_2$1.90, $Ala_6$5.61, $Val_3$3.10, $Leu_6$5.80	9.36 ppm

Table 4

In conclusion, this methodology is indeed very effective in producing peptide *C*-terminal semicarbazones and aldehydes with no epimerisation occurring at the *C*-terminal chiral centre.

‡ The suffix -sc has been adopted to indicate that the sequence is the semicarbazone of the *C*-terminal aldehyde.
§ Yield quoted is based on theoretical maximum based on Fmoc-loading on completion of the sequence and is calculated for isolated product after purification by preparative HPLC.
¶ All masses were determined using a Perseptive Biosystems Voyager™ MALDI-TOF mass spectrometer.
^ Yield quoted is for isolated product after purification by preparative HPLC.
† The signal quoted is that for the aldehydic proton and was the only signal observed in that region.

Acknowledgements
We are indebted to the Edinburgh Centre For Protein Technology and the Faculty of Science and Engineering, University of Edinburgh for financial provisions (to JAP). In addition we thank E. Beatty for 600MHz NMR, K. Shaw and B. Whigham for technical support and the BBSRC/DTI for funding.

References

1 McConnell, R. M.; York, J. L.; Frizzel, D.; Ezell, C. *J. Med. Chem.* **1993**, *36*, 1084-1089.
2 Bajusz, S.; Szell, E.; Bagdy, D.; Barabas, E.; Horvath, G.; Dioszegi, M.; Fittler, Z.; Szabo, G.; Juhasz, A.; Tomori, E.; Szilagyi, G. *J. Med. Chem.* **1990**, *33*, 1729-1735.
3 Chapman, K.T. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 613-618.
4 Graybill, T. L.; Dolle, R.E.; Helaszek, C. T.; Miller, R. E.; Ator, M. A. *Int. J. Peptide Protein Res.* **1994**, *44*, 173-182.
5 Sarubbi, E.; Seneci, P. F.; Angelestro, M. R.; Peet, N. P.; Denaro, M.; Islam, K. *FEBS Lett.* **1993**, *319*, 253-256.
6 Fehrentz, J. A.; Heitz, A.; Castro, B.; Cazaubon, C.; Nisato, D. *FEBS Lett*, **1984**, *167*, 273-276.
7 Basak, A.; Jean, F.; Seidah, N.G.; Lazure, C. *Int. J. Peptide Protein Res.* **1994**, *44*, 253-261.
8 Liu, C. F.; Rao, C.; Tam, J. P. *J. Am. Chem. Soc.* **1996**, *118*, 307-312.
9 Jacobson, K. A.; Marr-Leisy, D.; Rosenkranz, R.P.; Verlander, M. S.; Melmon, K. L.; Goodman, M. J. *J. Med. Chem.* **1983**, *26*, 492-499.
10 Lelievre, D.; Chabane, H.; Delmas, A. *Tetrahedron Lett.* **1998**, *39*, 9675-9678.
11 Paris, M.; Heitz, A.; Guerlavais, V.; Christau, M.; Fehrentz, J. A.; Martinez, J. *Tetrahedron Lett.* **1998**, *39*, 7287-7290.
12 Hall, B. J.; Sutherland, J. D. *Tetrahedron Lett.* **1998**, *39*, 6593-6596.
13 Fehrentz, J. A.; Paris, M.; Heitz, A.; Velek, J.; Winternitz, F.; Martinez, J. *J. Org. Chem.* **1997**, *62*, 6792-6796.
14 Pothion, C.; Paris, M.; Heitz, A.; Rocheblave, L.; Rouch, F.; Fehrentz, J. A.; Martinez, J. *Tetrahedron Lett.* **1997**, *38*, 7749-7752.
15 Ede, N. J.; Bray, A. M. *Tetrahedron Lett.* **1997**, *38*, 7119-7122.
16 Galeotti, N.; Giraud, M.; Jouin, P. *Lett. Peptide Sci.* **1997**, *4*, 437-440.
17 Fehrentz, J. A.; Paris, M.; Heitz, A.; Velek, J.; Liu, C. F.; Winternitz, F.; Martinez, J. *Tetrahedron Lett.* **1995**, *36*, 7871-7874.
18 Murphy, A. M.; Dagnino, R.; Vallar, P. L.; Trippe, A. J.; Sherman, S. L.; Lumpkin, R. H.; Tamura, S. Y.; Webb, T. R. *J. Am. Chem. Soc.* **1992**, *114*, 3156-3157.
19 Ramage, R.; Irving, S. L.; McInnes, C. *Tetrahedron Lett.* **1993**, *34*, 6599-6602.
20 Wen, J. J.; Crews, C. M. *Tetrahedron Asymmetry* **1998**, *9*, 1855-1858.
21 Paris, M.; Pothion, C.; Heitz, A.; Martinez, J.; Fehrentz, J. A. *Tetrahedron Lett.* **1998**, *39*, 1341-1344.
22 Fehrentz, J. A.; Heitz, A.; Castro, B. *Int. J. Peptide Protein Res.* **1985**, *26*, 236-241.
23 Nicholson, D. W.; Ali, A.; Thornberry, N. A.; Vaillancourt, J. P.; Ding, C. K.; Gallant, M.; Gareau, Y.; Griffin, P. R.; Labelle, M.; Lazebnik, Y. A.; Munday, N. A.; Raju, S. M.; Smulson, M. E.; Yamin, I. T.; Yu, V. L.; Miller, D. K. *Nature* **1995**, *376*, 37-43.
24 Wright, S. C.; Schellenberger, V.; Wang, H.; Kinder, D. H.; Talhouk, J. W.; Larrick, J. W. *J. Exp. Med.* **1997**, *186*, 1107-1117.