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

Denise and John

and for Lucy
Acknowledgements

I would like to express my gratitude to Professor R. Ramage FRS for the provision of research facilities and his constant support and encouragement throughout the course of this work. I would also like to thank him for his understanding and patience when circumstances in my first two years led to somewhat erratic timekeeping.

I am grateful to Dr. G.J. Cotton for his advice and help on related, and unrelated topics. My thanks go to Dr. S.E.M Howie for her work on the binding studies and apoptosis and for her help in the immunological and biological aspects of the work. I would like to thank Dr. D. Harrison and Dr. J. Ross for their helpful discussions and M. Fernandez for his work on peptide binding and MIP-1α inhibition.

Special thanks go to the departmental technical staff. I am grateful to Kevin Shaw for his help with the peptide synthesis described herein, Brian Whigham for amino acid analysis and MALDI-TOF mass spectrometry, Alan Taylor for his work on FAB mass spectrometry and Dr. C. Hewage for his NMR work.

I would like to thank all my friends I have made in Edinburgh, both past and present, for making my time here so enjoyable and unforgettable. I would also like to thank Dr. N. Robertson for proof-reading this thesis, and reminding me how bad my English can be at times.

I wish to thank the Medical Research Council for their provision of the research grant.

Finally, my thanks go to Dr. W. Donovan at Edinburgh University Medical Centre, Mr Pritchett and all working in Ward 8 at South Tyneside General Hospital for giving me the chance to finish this work.
Abstract

The synthesis of IH1, a peptide designed to mimic a discontinuous epitope on HIV-I gp120, is reported. The aspartimide rearrangement inherent in this sequence, and in the peptide GC1, has been studied and reduced to low levels. The syntheses of variants of peptide GC1, containing differing number of residues found to be important for CD4 binding, have also been achieved. Thus peptides containing one, two, three and four residues necessary for high affinity binding have been synthesised. In addition a peptide has been synthesised which incorporates a synthetic β turn moiety other than the Cys-Val-Cys bridge present in GC1. Polyclonal sera raised to these peptides and their CD4 binding properties have been investigated.

IH1, the peptide containing four residues responsible for high affinity binding to CD4, has also been shown to interact with receptors on the surface of CD4+ cells. This non-CD4 recognition has been investigated utilising a photolabelled chemokine, MIP-1α. Results indicate that this binding involves interaction with CC-CKR5, a MIP-1α binding site thought to be involved in HIV-cell fusion.
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acm</td>
<td>Acetamidomethyl</td>
</tr>
<tr>
<td>acpc</td>
<td>cis-3-Aminocyclopentane-1-carboxylic acid</td>
</tr>
<tr>
<td>AIDS</td>
<td>Acquired immunodeficiency syndrome</td>
</tr>
<tr>
<td>Alloc</td>
<td>Allyloxycarbonyl</td>
</tr>
<tr>
<td>Boc</td>
<td>t-Butyloxycarbonyl</td>
</tr>
<tr>
<td>Bpoc</td>
<td>Biphenyloxycarbonyl</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>Bu'</td>
<td>Tertiary butyl</td>
</tr>
<tr>
<td>Bum</td>
<td>Butyloxymethyl</td>
</tr>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>CKR</td>
<td>Chemokine receptor</td>
</tr>
<tr>
<td>CTL</td>
<td>Cytotoxic T lymphocytes</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DCC</td>
<td>N,N′-dicyclohexylcarbodiimide</td>
</tr>
<tr>
<td>Dde</td>
<td>4,4-dimethyl-2,6-dioxocyclohex-1-ylidine</td>
</tr>
<tr>
<td>DIC</td>
<td>N,N′-diisopropylcarbodiimide</td>
</tr>
<tr>
<td>DIEA</td>
<td>Diisopropylethylamine</td>
</tr>
<tr>
<td>DMAP</td>
<td>4-(N,N′-dimethylamino)-pyridine</td>
</tr>
<tr>
<td>DMF</td>
<td>N-dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethylsulfoxide</td>
</tr>
<tr>
<td>DTNB</td>
<td>5,5′-dithiobis (2-nitrobenzoic acid)</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDT</td>
<td>Ethanedithiol</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme linked immunosorbent assay</td>
</tr>
<tr>
<td>FAB</td>
<td>Fast atom bombardment</td>
</tr>
<tr>
<td>FCS</td>
<td>Foetal calf serum</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein isothiocyanate</td>
</tr>
<tr>
<td>Fmoc</td>
<td>9-fluorenlymethoxycarbonyl</td>
</tr>
<tr>
<td>GdmCl</td>
<td>Guanidinium hydrochloride</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>gp</td>
<td>glycoprotein</td>
</tr>
<tr>
<td>GMEM</td>
<td>Glasgow's minimum essential medium</td>
</tr>
<tr>
<td>Hepes</td>
<td>N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency syndrome</td>
</tr>
<tr>
<td>Hmb</td>
<td>2-hydroxy-4-methoxybenzyl</td>
</tr>
<tr>
<td>HOEt</td>
<td>N-hydroxybenzotriazole</td>
</tr>
<tr>
<td>HOCl</td>
<td>Triazole</td>
</tr>
<tr>
<td>HOSu</td>
<td>N-hydroxysuccinimide</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>Ig</td>
<td>Immunoglobulin</td>
</tr>
<tr>
<td>IgG</td>
<td>Immunoglobulin class G</td>
</tr>
<tr>
<td>IgM</td>
<td>Immunoglobulin class M</td>
</tr>
<tr>
<td>il-2</td>
<td>Interleukin-2</td>
</tr>
<tr>
<td>mAb</td>
<td>Monoclonal antibody</td>
</tr>
<tr>
<td>MAP</td>
<td>Multiple antigenic peptide</td>
</tr>
<tr>
<td>Mbh</td>
<td>4,4-dimethoxybenzhydryl</td>
</tr>
<tr>
<td>MHC</td>
<td>Major histocompatibility complex</td>
</tr>
<tr>
<td>MMA</td>
<td>N-Methyl-mercaptoacetamide</td>
</tr>
<tr>
<td>NBOC</td>
<td>2-Nitrobenzyloxy carbonyl</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>Np</td>
<td>Nitrophenol</td>
</tr>
<tr>
<td>NVOC</td>
<td>6-Nitroveratryloxy carbonyl</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>PBMC</td>
<td>Peripheral blood mononuclear cells</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PE</td>
<td>Phycoerythrin</td>
</tr>
<tr>
<td>PGC</td>
<td>Porous graphitised carbon</td>
</tr>
<tr>
<td>Pfp</td>
<td>Pentafluorophenol</td>
</tr>
<tr>
<td>Phacm</td>
<td>Phenylacetamidomethyl</td>
</tr>
<tr>
<td>Pmc</td>
<td>2,2,5,7,8-Pentamethylchroman-6-sulfonyl</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>mRNA</td>
<td>Messenger RNA</td>
</tr>
<tr>
<td>rpm</td>
<td>revolutions per minute</td>
</tr>
<tr>
<td>SDF</td>
<td>Stromal derived factor</td>
</tr>
<tr>
<td>Silver triflate</td>
<td>Silver trifluoromethanesulfonate</td>
</tr>
<tr>
<td>TASP</td>
<td>Template assembled synthetic protein</td>
</tr>
<tr>
<td>Tbfmoc</td>
<td>17-tetrabenzo(a,c,g,i)fluorenylmethoxycarbonyl</td>
</tr>
<tr>
<td>Tcp</td>
<td>Trichlorophenol</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
</tr>
<tr>
<td>TFMSA</td>
<td>Trifluoromethanesulfonic acid</td>
</tr>
<tr>
<td>TMSBr</td>
<td>Trimethylsilyl bromide</td>
</tr>
<tr>
<td>TNF</td>
<td>Tumour necrosis factor</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>Trityl</td>
<td>Triphenylmethyl</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>Z</td>
<td>Benzyloxy carbonyl</td>
</tr>
</tbody>
</table>
The naturally occurring amino acids

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>3 letter code</th>
<th>1 letter code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alanine</td>
<td>Ala</td>
<td>A</td>
</tr>
<tr>
<td>Arginine</td>
<td>Arg</td>
<td>R</td>
</tr>
<tr>
<td>Asparagine</td>
<td>Asn</td>
<td>N</td>
</tr>
<tr>
<td>Aspartic acid</td>
<td>Asp</td>
<td>D</td>
</tr>
<tr>
<td>Cysteine</td>
<td>Cys</td>
<td>C</td>
</tr>
<tr>
<td>Glutamic acid</td>
<td>Glu</td>
<td>E</td>
</tr>
<tr>
<td>Glutamine</td>
<td>Gln</td>
<td>Q</td>
</tr>
<tr>
<td>Glycine</td>
<td>Gly</td>
<td>G</td>
</tr>
<tr>
<td>Histidine</td>
<td>His</td>
<td>H</td>
</tr>
<tr>
<td>Isoleucine</td>
<td>Ile</td>
<td>I</td>
</tr>
<tr>
<td>Leucine</td>
<td>Leu</td>
<td>L</td>
</tr>
<tr>
<td>Lysine</td>
<td>Lys</td>
<td>K</td>
</tr>
<tr>
<td>Methionine</td>
<td>Met</td>
<td>M</td>
</tr>
<tr>
<td>Phenylalanine</td>
<td>Phe</td>
<td>F</td>
</tr>
<tr>
<td>Proline</td>
<td>Pro</td>
<td>P</td>
</tr>
<tr>
<td>Serine</td>
<td>Ser</td>
<td>S</td>
</tr>
<tr>
<td>Threonine</td>
<td>Thr</td>
<td>T</td>
</tr>
<tr>
<td>Tryptophan</td>
<td>Trp</td>
<td>W</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>Tyr</td>
<td>Y</td>
</tr>
<tr>
<td>Valine</td>
<td>Val</td>
<td>V</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

1.1 The immune response

1.1.1 Synthetic peptides as antigens

1.1.2 Epitopes

1.1.3 Choice of sequence

1.2 Human Immunodeficiency Virus

1.2.1 HIV pathogenesis

1.2.2 HIV structure

1.2.3 Insertion and replication of the virion

1.2.4 Cytopathic effects of viral replication on host

1.2.5 Apoptosis

1.2.6 gp120 structure

1.2.6.1 The CD4 binding site

1.2.6.2 The V3 loop

1.2.6.3 Viral tropism

1.2.6.4 gp120 interaction with gp41

1.2.6.5 The gp120 binding site on CD4

1.2.6.6 The C2 region and non-CD4 entry

1.2.7 Co-receptors in HIV infection

1.3 Peptide synthesis

1.3.1 The Boc strategy

1.3.2 The Fmoc strategy

1.3.3 Activation of amino acids

1.3.3.1 Acyl chlorides

1.3.3.2 Acyl azides

1.3.3.3 Carbodiimides

1.3.3.4 Symmetrical anhydrides

1.3.3.5 Active esters

1.3.4 Side chain protection
1.3.5 Acid cleavage 45
1.3.6 Rearrangements 46
1.3.6.1 Histidine racemisation 46
1.3.6.2 Diketopiperazine formation 47
1.3.6.3 Aspartimide formation 47
1.3.7 Overall Fmoc synthesis 47

Chapter 2 : Results and discussion 49-103

2.1 Project overview 49
2.2 GC1 synthesis : the problems 52
2.2.1 The Asp-Gly rearrangement 52
2.2.2 Cysteine protection : The use of Cys-Phacm 57
2.2.3 Nα-Tbsfmc in purification of GC1 and its analogues 61
2.3 Optimised GC1 synthesis 63
2.4 Synthesis of GC1 analogues 66
2.4.1 GC1 mono-Ala 67
2.4.2 GC1 bis-Ala 69
2.5 GC1 cyclopentyl derivative 71
2.5.1 Synthesis of GC1.cyclopentyl 72
2.6 IH1 74
2.6.1 Orthogonal lysine protection 76
2.6.2 Synthesis of IH1 78
2.6.3 Circular dichroism of peptide IH1 85
2.7 Immunogenicity and epitope mapping studies of IH1 87
2.8 Immunogenicity of GC1, GC1.cyclopentyl, GC1 mono-Ala and GC1 bis-Ala 90
2.9 Binding of serum raised against GC1 and IH1 to all peptides 91
2.10 Binding of IH1 to CD4+ H9 cells 92
2.11 Binding of peptides to parent HeLa cell line or CD4\(^+\) transfected HeLa cells

2.12 Binding of IH1 to an area of the cell surface proximate to CD4

2.13 IH1 binds to the CDR2 region in domain 1 of CD4

2.14 Peptide inhibition of MIP-1\(\alpha\) binding

2.15 Induction of apoptosis

Chapter 3: Experimental

3.1 Notes

3.2 Solid phase peptide synthesis

3.3 Peptide GC1 Cys-Phenacm

3.4 Peptide GC1 : optimised synthesis

3.5 Peptide GC1 mono-Ala

3.6 Peptide GC1 bis-Ala

3.7 Peptide GC1.cyclopentyl

3.8 Peptide IH1

3.9 Peptide IH2

3.10 Immunisation with peptides

3.11 Immunogenicity of peptides

3.12 binding of \(\alpha\)-IH1 and \(\alpha\)-GC1 IgG to peptides

3.13 Protein G purification of antisera

Chapter 4: Appendices

4.1 Flow cytometry

4.2 Binding of peptides to CD4\(^+\) cells

4.3 Peptide blocking of \(\alpha\)-CD4 mAb binding to CD4

4.4 Peptide blocking of MIP-1\(\alpha\) binding to CC-CKR5\(^+\) cells
4.5 Localisation of IH1 binding to CD4^+ cells 137
4.6 *In vitro* induction of apoptosis 137

Chapter 5: References 139-154
Chapter 1: Introduction

1.1 The immune response

This project has been designed to obtain information which could be used towards the formation of a vaccine against AIDS. The work uses the humoral immune response to raise antibodies which will cross react with the virus coat protein. A brief introduction to the immune system is given below.

Protection against microbes in the body are afforded by two systems, natural and acquired immunity.

a) Natural immunity: includes physical barriers, phagocytic cells and various blood-borne molecules used in defence. These are present in the body before infection and are not specific for individual invaders.

b) Specific or acquired immunity: consists of defence mechanisms induced or stimulated by exposure to a foreign substance, or antigen. Acquired immunity is specific to a particular antigen, with the specific immune system 'remembering' each encounter so subsequent exposures stimulate increasingly efficient effector systems. Acquired immunity also amplifies the protective mechanisms of natural immunity, focusing them to sites of antigen entry, leading to faster elimination of the antigen. Features of natural and specific immunity are shown in Figure 1.1 below.
Specific immunity itself can be broken down into two different systems classed as humoral and cell-mediated immunity.

a) **Humoral immunity**: mediated by circulating proteins produced by B lymphocytes, termed antibodies or immunoglobulins, responsible for specific recognition and elimination of extracellular antigens.

b) **Cell-mediated immunity**: mediated by T lymphocytes and is effective against intracellular antigens.

<table>
<thead>
<tr>
<th></th>
<th>Natural</th>
<th>Specific (Acquired)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physiochemical barriers</td>
<td>Skin, mucous membrane, enzymes in mucosal secretions.</td>
<td>Cutaneous and mucosal immune systems; antibody in mucosal secretions.</td>
</tr>
<tr>
<td>Circulating molecules</td>
<td>Complement, cytokines.</td>
<td>Antibodies, cytokines.</td>
</tr>
<tr>
<td>Cells.</td>
<td>Phagocytes (macrophages, neutrophils), natural killer cells</td>
<td>Lymphocytes.</td>
</tr>
</tbody>
</table>

**Figure 1.1**

Features of natural and specific immunity

This project was designed to generate antibodies raised against synthetic peptides which were able to cross-react with HIV. These peptides were chosen to mimic regions of the outer coat protein of HIV, which are important in the lifecycle of the virus.
The processing of a foreign antigen in the initiation of a humoral immune response can be thought of in the following way (see Figure 1.2).

**Figure 1.2**
The humoral immune response
The initiation of a humoral immune response includes the following steps:

a) An antigen presenting cell (includes macrophages and dendritic cells) ingesting an invading antigen.
b) The antigen, in the above case a virion, is rapidly internalised via phagocytosis.
c) Internal enzymatic degradation breaks down the proteins of the virion into smaller peptides within the lysosomal compartment of the cell.
d) Peptides bound to specialised proteins called the Major Histocompatibility Complex class 2 (MHC-II).
e) The MHC-II / peptide complex is displayed on the surface of the cell.
f) When the T-cell receptor recognises the displayed peptide, it binds to the antigenic peptide with a surface receptor, CD4, also binding to the MHC-II.
g) interleukin-2 (IL-2), a T cell growth factor, is released, stimulating the T helper cells with bound antigen and resulting in proliferation and differentiation of the T helper cells.

The work in this thesis is concerned with the response of B lymphocytes to activation. Upon stimulation, the B lymphocytes are affected in one of two ways, either maturing to plasma cells or going on to form memory B lymphocytes. The plasma cells produce the antibodies necessary for efficient foreign body clearance. The only B lymphocytes affected are those with bound antigens to their surface receptors. These receptors have the same specificity as do the antibodies produced, hence bound antigen indicates the ability of the antibodies to bind to the foreign antigen. B and T lymphocyte memory cells produced in the humoral response are long lived cells which are easily stimulated to form plasma cells in the case of B lymphocytes, and cytokines in the case of T lymphocytes. These memory cells are the basis of the ability of the body to mount an improved response upon subsequent encounter with the same
antigen. B cell stimulation and changes in cytokine profile during the response leads to immunoglobulin class switching by the B lymphocytes. Secondary challenge with the same antigen then gives an enhanced response, with greater production of IgG relative to IgM. This ability to 'learn' from past challenges produces an antibody titre as shown in Figure 1.3.

![Figure 1.3](image)

A normal immune response to antigen A

In the past, there has been a tendency to use whole inactivated or live attenuated virus as a vaccine. These are likely to be very similar to the native virus, and hence the antibodies raised react with the pathogen. The use of either of these methods with HIV leads to grave concerns over safety as it is extremely difficult to ascertain whether all viral particles have been inactivated or attenuated. Live attenuated viruses are also difficult to use due to the variability of the virus. This leads to the chance that an attenuated virus could mutate back to a virulent form. The envelope glycoprotein of HIV, gp120, is extremely variable and mutates readily. Most of the neutralising antibodies are raised to gp120\textsuperscript{1}, hence vaccines using one isolate would be unlikely to cover sufficient of the viral population to lead to eradication.
1.1.1 Synthetic peptides as antigens

As stated earlier, gp120 is the surface protein to which most neutralising antibodies are raised. It has been shown that the actual major neutralising determinant in human infection is part of gp120 called the V3 loop\textsuperscript{2,3,4}. The high variability of this region leads to type specific antibodies, antibodies capable of neutralisation only one specific viral isolate. Antibodies raised to this area have also been shown to enhance infection\textsuperscript{5,6}.

Due to these problems, there is now a move towards the use of areas of the virus capable of neutralisation. Thus, parts of viral proteins can be introduced as peptides, allowing the priming of the immune system without the danger inherent in the use of viral particles. Also, areas found to be variable or capable of raising enhancing antibodies can be omitted. There are precedents for the use of peptides to raise antibodies capable of neutralising viruses, as discussed below.

Much work has been attempted on the use of synthetic peptides as antigens for the use in a new generation of synthetic peptide vaccines\textsuperscript{7}. Available information on the position of antigenic sites in toxins, viruses and parasites has stimulated the research\textsuperscript{8,9} and protective immunity has been elicited against foot and mouth disease\textsuperscript{10,11,12}, influenza\textsuperscript{13,14}, hepatitis B\textsuperscript{15,16,17} and cholera by immunisation with synthetic peptides.

The idea is that these synthetic peptides mimic sites important in the lifecycle of the virus in order that antibodies raised will not only flag virions for destruction, but could also block these important areas. Unfortunately this is not easy, with the sequence utilised needing to be chosen carefully. Linear peptides are rarely able to mimic desired sites on folded proteins to any great extent, as will be discussed below.
1.1.2 Epitopes

Epitopes are those parts of the antigen which are bound, or 'recognised', by the T or B antigen receptor. Thus, an epitope is a set of amino acids which bind to the antigen binding site of these antigen receptors. T lymphocyte epitopes differ from B lymphocyte epitopes in both amino acid sequence and length. T lymphocyte epitopes tend to be linear consisting of 11-23 amino acids. They need not lie on the surface of the molecule. B lymphocyte epitopes tend to be discontinuous and shorter than the T lymphocyte epitope, and are usually found on the surface of the molecule.

Epitopes can be split into two different classes, continuous and discontinuous areas of a particular protein.

a) **Continuous Epitopes**: These are epitopes defined by a stretch of amino acids in direct peptide linkage with distinctive conformational features allowing the binding of a specific antibody.

b) **Discontinuous Epitopes**: These consist of a group of amino acids which are not in direct peptide linkage but, as a consequence of the folding and three-dimensional structure of the protein, they are close enough in space to allow the binding of a specific antibody.

Most of the work up to now has concentrated on the synthesis of linear peptides to mimic continuous epitopes. However, it is well known that most of the epitopes on any native protein will be discontinuous to a certain extent. Thus, the use of linear peptides does not necessarily give antibodies capable of cross reacting with the native molecule, even if parts of the linear sequence of the protein are used.
1.1.3 Choice of sequence

The classical way of obtaining information on important parts of the viral protein of interest is the use of linear peptides to map the epitope of antibodies raised to the native protein. Peptides used to map the surface of proteins have been widely used\textsuperscript{19}, giving information on areas accessible to the immune system. From this work, a picture can be built up as to the important parts of the protein. The removal of whole domains of the protein can give ideas of the important areas in the molecule, but it is sometimes difficult to tell whether these removed domains are acting as scaffolding or have biological importance themselves. Point mutation studies, where certain amino acids are removed or replaced\textsuperscript{20}, can also give an idea of important sections, or even information on widely dispersed areas acting together\textsuperscript{21} and affect overall folding to a much smaller extent. Point mutation studies are also more likely to give information on discontinuous areas.

Structural information can be derived in a variety of ways including biological methods described above. Much information can be obtained from computer simulations of the peptide\textsuperscript{22,23}. This technique uses physical information from the peptide, and information from other known and related structures, to give an approximation of the protein structure. Unfortunately, due to the enormous amount of interactions in a folded protein, these results tend not to be very accurate. NMR studies\textsuperscript{24} also can be used to give structural information, along with crystallographic procedures. Both of these are extremely powerful tools, but are not suitable for all proteins.

Although relatively difficult, it is possible to predict the areas essential for either binding or post binding events. Peptides can then be designed to mimic this area. It is unlikely that a linear peptide will obtain the correct structure in solution in order to mimic the native protein. However, this may not be a problem providing sufficient similarity exists. There is evidence that the actual process of binding induces structure in the peptide. Thus, it may be that the peptide need only resemble the native protein
and contain some flexibility to give the required shape. The following work uses this premise, a linear peptide has been designed but a turn introduced to give structure. This turn also allows the peptide to present discontinuous epitopes, as will be discussed later.

The use of synthetic organic compounds as scaffolding for peptides has become another useful way to mimic native proteins. This can allow different areas of the protein to be bound together as they would in the folded protein. It also allows T cell epitopes from the protein to be included with B cell epitopes increasing the humoral response. The areas of interest include multiple antigenic peptides (MAP)\textsuperscript{25}, which have been used in the formation of B cell and T cell containing systems\textsuperscript{26,27,28}. The template assembled synthetic protein (TASP)\textsuperscript{29} can be used in the same way to design protein structure\textsuperscript{30,31,32}. Kemp's Triacid has also been used in similar ways\textsuperscript{33,34}. 
1.2 Human Immunodeficiency Virus (HIV)

In 1981, young homosexual men in the United States were presenting to doctors with diseases linked to immune suppression\(^{35}\). It soon became clear that this syndrome was due to a virus. A few years after the initial recognition of this new disease state, isolation and characterisation of the virus was achieved, and HIV was proven to be the etiological agent of Acquired Immunodeficiency Syndrome, AIDS\(^{36,37,38}\). The virus was identified as a member of the lentivirus family and as such, a retrovirus. Unlike other members of the lentivirus family, which affect T cells by immortalising them into continuous growth, one of the major pointers for HIV infection is the reduction in the number of T cells present in the body over time\(^{39,40}\).

1.2.1 HIV pathogenesis

HIV infection is highly variable between individuals\(^{41,42}\) but, generally events occurring after HIV infection can be described as follows, and schematically as in Figure 1.4.

![Figure 1.4](persistent_state.png)

**Features of HIV infection (from reference 60) showing**

- ARC (Aids related complex), (1) CD4\(^+\) cell count and (2) Viremia
Prior to seroconversion, when the body initiates production of antibodies against the virion, high levels of virus can be detected in the blood. Viremia is then reduced to a low level, where it generally stays, apart for episodic release of varying amounts of virus over time. It is thought that this suppression of the virus could be mediated by antiviral CD8+ cells\textsuperscript{43,44}. At the same time as clinical symptoms become apparent, viremia rises to a high level, where it stays throughout the final period of disease. Also at this time, the CD8+ anti-HIV response begins to decrease.

The CD4+ cell number decreases sharply during primary infection before rising to a level which is still below normal. A slow decrease in cell count then occurs throughout the asymptomatic state. This is thought to be due to destruction of cells by apoptosis\textsuperscript{45} (see later). Once symptoms occur, a marked decrease in CD4+ count can be seen at the same time as the increase in viremia.

1.2.2 HIV structure

Two closely related types of the virus have been identified, designated HIV-1 and HIV-2. They have 40% homology in their surface proteins and cause indistinguishable clinical syndromes\textsuperscript{46}. Like all viruses, HIV has two distinct stages in its lifecycle, the intra- and extracellular stage. A single virus particle, or virion, of HIV has a structure as shown below (Figure 1.5).
The virion has no biochemical pathways, the virus entering a host cell and subverting the cellular biochemistry in order to replicate. HIV is a retrovirus, hence the genetic information is carried as RNA, in this case two identical strands, each 9.2-9.7 kilobases long enclosed by the subunits of the p24 protein in the central capsid. The genetic material of HIV consists of env, encoding for the envelope glycoproteins, gag, for the structural core proteins, pol for the viral protease enzymes, and also vpr, tat, rev, nef and vpu whose products regulate viral reproduction.

Also present in the capsid is reverse transcriptase, the enzyme which transcribes the viral RNA into DNA before insertion into the host DNA. The whole virion is surrounded by a lipid bilayer. The envelope glycoprotein has two components gp120 and gp41 and these project out of the lipid bilayer. It is these that interact with the host cell prior to insertion of the genetic material. The envelope glycoproteins are
synthesised in the host cell as a precursor protein, gp160, which is proteolytically cleaved by a host-cell protease into gp120 and gp41\textsuperscript{47,48}. These two glycoproteins are then held together by non-covalent interactions\textsuperscript{19}.

1.2.3 Insertion and replication of the virion

Viral insertion and replication can be visualised as in Figure 1.6 below. The initial stage of insertion into the host cell is high affinity binding of gp120 to the host cellular receptor CD4\textsuperscript{49,50,51,52}. CD4 is not internalised\textsuperscript{53}, rather, structural changes occur to the gp120\textsuperscript{54,55,56} which allow presentation of gp41. Entry into the cell is pH independent\textsuperscript{57,58}, indicating that HIV fuses directly with the cell surface membrane. Insertion of the hydrophobic amino-terminus of gp41 into the cell membrane initiates fusion of the virus envelope with the cell surface membrane. The virion then uncleaks and enters the cell. Upon entry, reverse transcriptase is used to transcribe the viral RNA to DNA. This DNA is then circularised and inserted randomly into the host DNA by the viral integrase. The integrated DNA copy of the viral RNA genome is known as the provirus. Insertion of the viral genetic material can have several consequences. The cell can die from the presence of foreign DNA affecting the cellular pathways. The virus could lie dormant, with insertion of viral DNA not leading to replication, or the infected cell could go on to produce viral particles. The production of these particles from the integrated HIV provirus occurs only upon activation of the infected T lymphocyte. The host biochemistry is subverted, allowing the formation of viral RNA and RNA coding for viral proteins. Once formed, the proteins are enzymatically cleaved and the virion forms inside the cell. The formed virion moves to the surface of the cell where it buds, taking with it some of the host lipid bilayer and presenting both gp120 and gp41 on its surface.
1.2.4 Cytopathic effects of viral replication on host

As previously described, isolates of HIV recovered from the individual in advanced disease has a greater capacity for killing infected cells than strains isolated in the clinically latent period. These cytopathic effects of the virus on the host are many and varied. A major biological feature of HIV infection is the formation of multinucleated cells from the fusion of infected cells with uninfected CD4^+ cells. This effect is termed syncytia formation and has been shown to involve the envelope glycoproteins. Cells are seen to degenerate due to the presence of the virus affecting the permeability and integrity of the cellular membrane. Cytopathology can also be associated with the accumulation of viral DNA in the cytoplasm of the cell. The possible mechanisms of cytotoxicity by HIV or its proteins are described below (reviewed in ref. 60).
Postulated methods of viral cytotoxicity in HIV infection include:

a) Syncytium formation.
b) Accumulation of unintegrated viral DNA.
c) Virus release causing changes in membrane integrity.
d) Virus alteration of plasma membrane permeability to cations.
e) Decrease in synthesis of membrane lipids.
f) Decrease in "second messenger" (diacylglycerol) activity.
g) Interference with cellular heteronuclear RNA processing.
h) Degradation of cellular mRNA and reduction in cellular protein synthesis.
i) Induction of apoptosis (see below).
j) Release of cytotoxic cytokines by infected cells and/or uninfected cells.
k) Destruction by immunogenic responses.
l) Competitive inhibition by HIV of normal growth factors.

1.2.5 Apoptosis

One of the difficulties in understanding the complex pathology of HIV infection is to explain the progressive depletion of CD4 helper T cell population and the consequent destruction of the immune system. The view that HIV causes AIDS only by directly killing CD4 cells, for example by lysis, has been questioned by many workers in the field. Gougeon et al. noticed that peripheral blood lymphocytes from asymptomatic HIV-infected individuals are primed for apoptosis. Apoptosis is a form of programmed cell death and occurs through the activation of an intrinsic cell suicide mechanism (reviewed in ref. 63). During apoptosis, the nucleus and the cytoplasm condense, and the dying cell often fragments into membrane-bound apoptotic bodies that are rapidly phagocytosed and digested by macrophages or neighbouring cells. Apoptosis can be mediated by the interaction of a cellular surface receptor, Fas, with its ligand. Fas ligand is a membrane protein and a member of the tumour necrosis factor (TNF) family. It has been shown in mice with mutated Fas or Fas ligand that
the Fas system plays an important role in destruction of autoreactive immune cells and as an effector molecule of cytotoxic T cells.

It was known that cross-linking of CD4 molecules with CD4 antibodies or with gp120 led to upregulation of Fas. Also cross linking of the CD4 receptors with CD4 mAb's led to apoptosis, as long as Fas is present in a useable form. It was noted that if gp120 was used to effect the cross-linking, the cell was primed for apoptosis. It has also been shown that Fas was upregulated on the surface of T cells in HIV infected patients, along with Fas ligand and this was associated with apoptosis. From this and recent research by Silvestris et al., it has been shown that Fas overexpression parallels the progression of disease in AIDS. This would indicate that apoptosis and in particular the Fas mediated pathway, play major role in the depletion of CD4 helper T cells.

1.2.6 gp120 structure

The numbering of gp120 in this thesis refers to the BH10 clone of HIV from Ratner et al.

It has been shown that gp120 is a highly variable molecule which shows extensive genetic difference between isolates, and undergoes rapid evolution mainly due to the inefficient reverse transcription and no ability to proof read the viral RNA in the cell. More highly conserved amino acids include the cysteines used to form the interchain disulfide bonds, which are extremely conserved. gp120 is highly glycosylated, with about 50% of the mass of the molecule due to sugars. Glycosylation occurs at both N and O linked glycosylation sites. The sites are not randomly distributed, but are also highly conserved. It has been found that glycosylation is unnecessary for CD4 binding or antibody recognition, although it must be present for correct folding.
The amino acid sequence of gp120 from many different viral isolates has been sequenced and comparisons of these sequences have shown the presence of five linear regions that are highly conserved\(^8\). The structure can be visualised as in Figure 1.7.

<table>
<thead>
<tr>
<th>C1</th>
<th>V1</th>
<th>V2</th>
<th>C2</th>
<th>V3</th>
<th>C3</th>
<th>V4</th>
<th>C4</th>
<th>V5</th>
<th>C5</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>100</td>
<td>200</td>
<td>300</td>
<td>400</td>
<td>500</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1.7

gp120 showing areas of conserved and variable amino acids

Due to the size of the molecule, and its extensive glycosylation, the structural data is limited. The known data can be summarised as in Figure 1.9. The data discussed above is the limit of current knowledge as to structure of gp120. There have been attempts to model the structure\(^21,22\) (see Figure 1.8) but their validity has not been confirmed by structural studies, such as X-ray crystallography or NMR.

HIV-1 SU (gp120)

Figure 1.8

Computer Model of gp120 structure
Figure 1.9
Structure of gp120
1.2.6.1 The CD4 binding site

As with the overall structure of gp120, the CD4 binding site on gp120 is not clearly defined. Cordonnier et al. showed that four conserved regions of gp120 were needed for high affinity binding. It was found that an area of the C4 region was needed for binding, and that mAb's blocking CD4 interaction recognise epitopes in this region. Thus at the time this was thought to be the binding site. However, several mutation studies have shown that amino acids essential for correct binding of CD4 are spread throughout the conserved regions of gp120. Due to this, the binding site is thought to be discontinuous, with many areas of gp120 being involved. There is a maximum binding region, however, with a large part of the N terminus, some of the C terminus, and the variable regions V1, V2 and V3 found unnecessary for binding.

The mutation studies mentioned above have shown that there are five amino acids which play a large part in the high affinity binding of CD4. These are Thr-257 in the C2 region, Asp-368 and Glu-370 in the C3 region and Trp-427 and Asp-457 in the C4 region. These amino acids are absolutely conserved in all HIV-1 isolates. The amino acids are marked in Figure 1.8. All five of these amino acids have been shown to fall within the epitope for a CD4 blocking antibody, 1.5e, again suggesting they are close in space in the folded protein (see Figure 1.10).

![Figure 1.10](image-url)  
Antibody interaction with CD4 binding site
1.2.6.2 The V3 loop

The amino acids of the V3 loop lie within the disulfide bridge formed between two invariant cysteines at positions 303 and 338. As explained earlier, the V3 loop is the principal neutralising determinant, with over half of all antibodies produced recognising this one area. These antibodies are extremely potent neutralising agents, but are also very type specific. Most antibodies raised early in the disease tend to be against the V3 loop. The neutralising antibodies do not prevent gp120 binding to CD4, rather they block an unknown post-binding event. The V3 loop is highly variable, and increases in variability during progression of the disease. Progression also leads to the emergence of virus variants which resist neutralisation. Despite the high variability, the tip of the V3 loop has been shown to be relatively conserved, with the GPGR motif being extremely conserved. This area seems to be involved in some post binding event as mutations in the tip greatly reduce viral fusion. Clements et al. have shown the presence of potential proteolytic cleavage sites very near the tip of the loop. Research has shown the cleavage of the V3 loop with soluble CD4 and with thrombin. Purified Trypstatin, a serine protease inhibitor, reduces syncytium formation. Also Tryptase TL2, a serine esterase from CD4+ cells, was strongly inhibited with gp120. It was later shown by peptide studies that the V3 loop blocks the protease action, indicating the V3 loop has the ability to act as a ligand for a cellular bound protease.

1.2.6.3 Viral Tropism

The V3 loop also has a great effect in the tropism of a particular viral strain. It has been shown to be the primary determinant of viral tropism. Indeed, a single amino acid difference Asp-323 to Lys, leads to a change from M-tropism to T-tropism. Zhu et al. have shown that the viral phenotype at initial seroconversion was found to be uniformly macrophage tropic and non-syncytium-inducing. As disease progresses, an increased replicative capacity of the virus occurs in peripheral blood mononuclear cells (PBMC's) and T lymphocytes. Rapidly replicating variants
thus emerge during the asymptomatic period, and this appearance tends to indicate a worsening of the condition of the patient. Antibody studies tend to suggest two distinct conformations for the V3 loop, and these conformations elicit viral tropism. Gu et al have shown that the cleavage sites in the V3 loop, discussed above, also have an effect on viral tropism.

1.2.6.4 gp120 interaction with gp41

As stated earlier, the envelope glycoproteins gp120 and gp41 are formed as a precursor peptide gp160. This is processed in the cell to form the two proteins. The transmembrane protein, gp41, plays a key role in both virus-mediated cell-cell fusion and interaction by cell free virus. The N terminus of gp41 has been found to act as the fusogenic domain, whose function probably involves insertion into and disruption of the target cell membrane. The N terminus has been shown to have a leucine zipper motif, changes to which do not affect interaction with gp120 but has a major effect on the virus lifecycle. Mutagenesis studies have shown that amino acids away from the N terminus, a sequence running from 696 to 707, and in particular basic residues at 696 and 707, play a crucial role in fusion. The area running from 521-663 has been shown by antibody mapping to be covered during the non-covalent gp120-gp41 binding. This is in agreement with data showing the areas involved in gp120 binding to be 528-562 and 608-628, 560-561 and 571-582, 579-611. Interaction areas on gp120 have been found to be in the conserved, hydrophobic, regions at the N and C terminus (C1 and C5). As with gp120, structural information is limited, but comparisons with other lentiviruses give a predicted structure as in Figure 1.11.
Figure 1.11

Predicted structure for gp41

(adapted from ref. 125)
1.2.6.5 The gp120 binding site on CD4

The CD4 structure has been found to be made up of four domains as in Figure 1.12a below.

![Diagram of CD4 structure](image)

**Figure 1.12a**

Overall structure of CD4 suggested from the amino acid sequence

The binding site for HIV was initially found to be located somewhere on domain 1 of CD4\textsuperscript{126,127}. It was subsequently found that the binding site was discontinuous\textsuperscript{128}, containing regions from CDR2\textsuperscript{129,130} and CDR3\textsuperscript{131,132}. Mutation studies have shown there are several amino acids implicated in the high affinity binding, with Phe-43 and Arg-59 likely to contact gp120\textsuperscript{133}. Lys-35 and Lys-46 are also thought likely to contribute to binding\textsuperscript{134} (see Figure 1.12b, below). Research has also implicated domain 2 in post binding events, perhaps interfering with conformational changes in gp120/gp41 or CD4 that are necessary for fusion\textsuperscript{135}. 

23
1.2.6.6 The C2 region and non-CD4 entry

Although it is clear that CD4 is the principal cellular receptor for HIV, many cell lines of the nervous system do not express CD4, but can be infected. This infection is less efficient than through CD4, but it is possible with several strains. Fantini et al\textsuperscript{36} showed the level of galactosyl ceramide, expressed on the surface of particular cells, is associated with permissiveness to HIV infection. Harouse et al\textsuperscript{37} showed that antibodies against galactosyl ceramide inhibit uptake and infection of two neural cell lines. The area of gp120 capable of interaction with this cellular receptor has been mapped to the C2 region, between amino acids 206-275\textsuperscript{38}. 

Figure 1.12b
Domain 1 of CD4 showing residues involved in gp120 binding
1.2.7 Co-receptors in HIV infection

As described earlier, CD4$^+$ cells are the major target for HIV replication (for a review, see ref. 139). However it was clear that productive infection required more than the presence of CD4 on the cellular surface.

Expression of CD4 on the cell surface does not necessarily lead to membrane fusion. Indeed, HIV strains are generally classified as T-cell tropic (T-tropic) or macrophage tropic (M-tropic) depending upon their differing ability to infect CD4$^+$ target cells. It has been known for many years that the presence of the CD4 receptor on the surface of a murine cell leads to the ability to bind the virus, but not fusion and internalisation. Recent work has shown that the block to fusion, and hence infection, of a murine cell line is not due to the presence of an inhibitory component on the murine cell, but is due to an additional component found on the human cell.

Work several years ago has shown that CD8$^+$ lymphocytes suppress viral replication in PBMC's. It was later shown that the CD8$^+$ were actually secreting one or more soluble HIV suppressor factors. These suppression factors were finally shown to be the chemokines RANTES (regulated-upon-activation, normal T expressed and secreted), MIP-1α (macrophage inflammatory protein-1α), and MIP-1β. Both MIP-1α and MIP-1β have been shown to be released in early infection to allow the creation of an inflammation centre. It is unfortunate that this leads to the presence of T cells giving a fertile environment for the spread of HIV. Further proof came from the work of Broder et al. who showed that viral tropism restrictions occur at the membrane fusion level after the binding of gp120 to CD4, again pointing to a cellular component acting as a co-receptor. All the above evidence was pointing to the presence of a secondary (co-)receptor on the surface of the cell.

Epidemiological studies have shown that certain individuals do not become infected with HIV, even after multiple exposure from HIV infected partners. Further studies have shown that this relative resistance to HIV infection was only to macrophage
tropic viruses, and it was not due to an intrinsic resistance of cells from the exposed individual to HIV infection which were easily infected with T-cell tropic viruses\textsuperscript{148}.

The breakthrough came with the findings of Feng et al\textsuperscript{149}, who showed that T-cell tropic HIV was using a secondary receptor, variably named fusin, LESTR (leukocyte derived seven transmembrane domain receptor) or CXCR-4. The presence of fusin together with human CD4 on the surface of the cell, allowed non-human cells to support fusion and infection with T-cell tropic isolates. Indeed, the presence of antibodies to fusin lead to the blocking of cell fusion, but not binding of the gp120 to CD4 in those same cells. Continuing research showed that murine, feline, simian and quail cell lines could all support fusion and infection in the presence of both fusin and human CD4\textsuperscript{150}. The receptor has been known for several years\textsuperscript{151}, but the ligand was unknown. After the HIV findings, the ligand was identified as stromal cell-derived factor-1 (SDF-1)\textsuperscript{152,153}. The presence of this ligand has been shown to block T-cell tropic fusion with human PBMC's. The sequence of SDF-1 was also well known\textsuperscript{154} and was found to be a chemokine expressed in a broad range of tissues\textsuperscript{155}.

This work showed that T-cell tropic viruses used fusin as a co-receptor, thus it seemed logical that M-tropic HIV was highly likely to do the same by using a different receptor. Since M-tropic viruses could be blocked by the presence of the chemokines RANTES, MIP-1\textalpha or MIP-1\textbeta, a receptor to these three chemokines was the obvious candidate. A newly discovered receptor was known to bind these three chemokines exclusively, namely CC-CKR5 (for CC chemokine, chemokine receptor 5)\textsuperscript{156}. This receptor was then co-transfected and expressed with human CD4 on the surface of CC-CKR5 -ve cells. The presence of this receptor led to M-tropic viruses productively infecting non-macrophage and even non-human cells\textsuperscript{157,158,159,160,161}. It should be noted that this infection was almost totally retarded in the presence of the three chemokines named above. Work also showed that the patients who were far less susceptible to M-tropic HIV infection either had an overproduction of the chemokines secreted from CD8\textsuperscript{+} cells\textsuperscript{162}, or actually had a defective CC-CKR5 allele containing a deletion\textsuperscript{163}. Presumably this deletion leads to the inability of the virus to use the co-
receptor and hence disruption to fusion and replication. It should be noted that, although CC-CKR5 seems to be the major co-receptor to M-tropic HIV, there are indications that other chemokine receptors may be used although much less effectively.

Interestingly, the work by Doranz et al.\textsuperscript{143} showed that CC-CKR5 was the co-receptor for M-tropic HIV, and that a dual tropic HIV isolate could use both fusin and CC-CKR5 as entry cofactors. As the interaction of gp120 with the different cofactors changes the tropism of the virus, and the tropism is so markedly affected by the sequence of the V3 loop (see earlier), it is highly likely that the V3 loop is interacting in some way with the co-receptor. This can be visualised in Figure 1.13, below.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Interaction of Fusin with gp120 in T-cell tropic fusion}
\end{figure}

(Adapted from ref. 164).
1.3 Peptide synthesis

Peptides and proteins are an all pervasive class of organic molecules found in many areas of life. They are natural polymers, coded for by nuclear DNA. They are used by the body in widely differing ways, controlling neurological, immune, endocrine and synthesis processes. Biological catalysis is carried out by a specific set of proteins, known as enzymes. The study of the folding and interaction of these proteins can give an insight into many biological areas. For instance, there is a real opportunity to affect a disease state with the use of synthetic proteins, either as agonists or antagonists. Peptides can be designed such that the structure mimics the active site of a protein, allowing the introduction of such a site without the excess scaffolding usually accompanying it. This can then be used to raise antibodies which cross react with the native protein.

In nature, the protein is synthesised inside the cell using specific areas of DNA, the DNA having been previously transcribed to RNA for use as a template. It is formed by attachment of one amino acid to the next in the chain, the sequence running from the N terminus to the C terminus. The biological expression of proteins is an excellent way of obtaining proteins in relatively large amounts, especially with the use of overexpression technology. The great advantage of chemical synthesis of proteins over their expression is the ability in chemical synthesis to incorporate unnatural amino acids, i.e. ones not coded for by any DNA, and as such is an extremely flexible way into the enormous field of protein technology.

Peptide synthesis began using the techniques of the solution phase many years ago. It comes as a surprise to learn the first reported peptide synthesis was performed in 1901 with the formation of glycylglycine\textsuperscript{165}. The synthesis of peptides using the solution phase was extremely complex and fraught with problems, each of the intermediate compounds having to be isolated and purified. By the 1950's, peptides were readily being synthesised\textsuperscript{166,167,168}, but it was obvious to those involved that solution phase would be impossible to use for large proteins and synthesis of the
proteins utilising solid phase was likely to be the way forward\textsuperscript{169}. It was for this reason that Merrifield's work\textsuperscript{170} made such a huge impact. His simple and extremely elegant idea was to anchor the growing chain covalently to a solid resin particle, thereby allowing the removal of by-products and leading to the crystallisation of intermediates becoming increasingly unnecessary.

Peptides are biological polymers and can be thought of as a string of amino acids linked together using the amide bond (see Figure 1.14).

\begin{center}
\includegraphics[width=0.3\textwidth]{amide_bond.png}
\end{center}

Figure 1.14
The amide bond

Thus the synthesis of proteins and peptides can be thought of as the synthesis of amide bonds. Of course it is more complicated than that, with the bonding needing to be specifically between one amine and another carboxylic acid. The amino acids must be activated in order for them to form an amide bond. Protecting group strategies must be employed to ensure the correct bond is being formed, as in the following example.

Suppose two amino acids are used to form an amide bond. Both have amino and carboxylic acid groups. This leads to there being four ways for the amide bond to form, as shown in figure 1.15.
It should be noted that even this is a simplified version of the reactions occurring. Further reaction of the products above could lead to trimers, tetramers, etc. If the required product was R₁R₂, then the amino group of R₁ must be blocked, and specific activation of the carboxylic acid group of R₁ must be achieved, with the carboxylic acid group of R₂ being blocked. Not only that, many of the side chains of amino acids are also reactive in the conditions used for amide bond synthesis. This leads to the necessity for protection of these side chains as well.

In the initial paper by Merrifield, the solid phase synthesis of a tetrapeptide was reported. The synthesis used the Z group for protection, the group being removed with 30% HBr in acetic acid. It was found that the group was not fully removed and in later work Merrifield used Boc protection which cleaved easily, in TFA, due to the formation of the stable t-butyl carbocation (see Figure 1.16).
The overall synthetic pathway that Merrifield adopted is shown in Figure 1.17. Note that the $\text{Na}$ protection is acid cleavable, and the linker used is orthogonal in that it is base cleavable.

![Figure 1.17](image)

Merrifield initial approach to a dipeptide

The major advantages of solid phase synthesis can be seen from the above diagram. Between each cycle, the solid resin allows several washing steps to remove by-products and unreacted activated amino acid. Unfavoured coupling reactions can be forced by the use of large excesses of starting material. Also, the fact that the peptide is suspended on a resin means, due to the mutual solvating effect of the peptide chain and the polymer chain, the peptide is thermodynamically less likely to self aggregate$^{175}$.

Activation of the acid to allow coupling was carried out in the above example using the acid anhydride. This in turn is formed using $\text{N,N'}$ dicyclohexylcarbodiimide$^{176}$ (see Figure 1.18). These reactions will be discussed later.
1.3.1 The Boc strategy

This remains relatively unchanged from the initial work carried out by Merrifield. The support used is still styrene copolymerised with 2% divinylbenzene as a cross linking agent. This was chosen as it contains all the characteristics required in solid phase synthesis. These are outlined below.

a) The resin swells in organic solvents to up to five or six times its original volume, allowing free accessibility of the growing peptide chain to diffusing reagents, leading to fast reaction times.

b) The resin is stable to all conditions used in the synthesis

c) The resin is able to be functionalised, to allow attachment of the peptide.

d) As the peptide chain grows, the dry volume of the resin increases to accommodate the added mass.

The initial amino acid is coupled to the resin via its caesium salt by nucleophilic displacement of chlorine. Nα protection is afforded using the Boc group already mentioned. Removal of the Boc group prior to coupling is achieved using the relatively mild acid, TFA. The linker and side chain protecting groups are also acid cleavable, but are much less susceptible to cleavage. Thus, the TFA removal of Boc leaves the linker mainly unaffected. Neutralisation then leaves a free amino from which the further assembly can grow. The coupling of the subsequent Boc protected
amino acid is carried out using the carbodiimide grouping discussed above. The synthesis is cycled, with the peptide growing from the C to the N terminus, until the required peptide is present. HF is then used to cleave the peptide from the resin and deprotect the side chains.

Boc chemistry is an extremely powerful way to synthesise proteins and peptides, but it is not perfect. Improvements could come from the use of less drastic reagents, HF requires special glassware and very careful handling, and the side chain protection and the linker are not orthogonal to the N\(^\alpha\) protection. As they are all affected by acid, the constant cleaving of Boc with TFA, however relatively mild, will lead to loss of peptide from the resin. There have been attempts to synthesise more acid labile N\(^\alpha\) protecting groups to minimise this problem, an example being the Bpoc group \(^{177}\) shown in Figure 1.19. The Bpoc group was used for demanding syntheses, however the problems still arose from non orthogonality. This led to the attempt to find a base labile protecting group for solid phase synthesis and from there to the discovery and the use of Fmoc in solid phase peptide synthesis.

![Figure 1.19](image)

The Bpoc group

### 1.3.2 The Fmoc strategy

9- Fluorenylmethoxycarbonyl (Fmoc) group is a base labile protecting group used in solid phase peptide synthesis. Initial work by Bergmann and Zervas\(^{171}\), and later by McKay and Albertson\(^{174}\) showed the use of urethanes as protecting groups for amines in peptide synthesis. However, the groups that came from this work, including Boc, were cleaved under acid conditions to the carbamate, which then spontaneously decarboxylates to the free amine. Carpino and Han\(^{178}\), however, used the urethane
work and the process of β elimination to give a protecting group cleavable under extremely mild base conditions, namely the Fmoc group (see Figure 1.20).

![Figure 1.20](image)

**Figure 1.20**

An Fmoc protected amino acid

It was rapidly shown that the group was stable to acids and hydrogenation, and was easy to introduce using the chloroformate\(^{179}\), although the N-hydroxysuccinimide ester is now more usually used giving a cleaner reaction. The group gave total orthogonality between the linker and the amine protection.

This orthogonality allowed the introduction of acid labile protecting groups and linkers. t-Butyl ester and Boc groups could now be used as protection, together with the p-benzyloxybenzyl alcohol linker\(^{180}\). This is an extremely acid labile linker shown below (see Figure 1.21).

![Figure 1.21](image)

**Figure 1.21**

p-Benzyloxybenzyl alcohol (Wang) resin
The rapid cleavage of this linker by mild acid (e.g. TFA) can be explained by the mechanism shown in Figure 1.22 below.

![Figure 1.22 Acid cleavage of the Wang resin](image)

The relatively mild cleavage and deprotection conditions led to the synthesis of peptides whose purity and yield was much improved.\(^{181}\)

Cleavage of the Fmoc group occurred by \(\beta\)-elimination using an E1cB mechanism\(^{182,183,184}\), via the stabilised dibenzocyclopentadienide anion. The anion then breaks down to the dibenzofulvene and the carbamate of the peptide. The carbamate formed then readily decarboxylates to give the free amine. The dibenzofulvene reacts with piperidine to give the piperidide adduct. This adduct absorbs strongly in the UV,
with the isosbestic point at 302 nm\textsuperscript{185}. This can be used to follow the amount of amino acid coupled at each step and hence give a real time assessment of the progress of the synthesis\textsuperscript{186,187}. All syntheses discussed in this work have used this method for coupling assessment. The reaction mechanism is as shown in Figure 1.23.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure123.png}
\caption{\textit{\beta} elimination of Fmoc by piperidine}
\end{figure}

The use of this base labile protecting group was soon found to give comparable incorporation of the amino acids\textsuperscript{188} and indeed to be complementary to Boc chemistry. Thus this method was widely adopted and improved\textsuperscript{189}. 

\textsuperscript{36}
1.3.3 Activation of amino acids

In most cases in peptide synthesis, the formation of the amide bond is effected using a nucleophilic displacement reaction. This approach requires a more effective leaving group than the hydroxyl present on the carboxylic acid. The usual way to obtain this effective leaving group is to substitute the hydroxyl for a more electronegative substituent, $E^{190,191}$. This gives a reaction as in figure 1.24.

![Figure 1.24](image)

Activation and coupling of an amino acid.

(P = protecting group, E = electronegative substituent)

There are many ways to activate the amino acid. Comprehensive reviews on the subject have been written$^{192,193,194}$, and a full discussion is beyond the scope of this introduction. A few of the methods will be discussed below.

1.3.3.1 Acyl chloride

Early on in the synthesis of peptides, acyl chlorides were used extensively$^{195}$ however, formation of an acyl chloride is difficult in the presence of sensitive substrates and the Z amino acid chlorides are unstable. Fmoc amino acid chlorides, however, have been
shown to be stable, but are prone to racemisation by the mechanism shown in Figure 1.25.

![Racemisation of acyl chlorides](image)

**Figure 1.25**

Racemisation of acyl chlorides

This is a general rearrangement for all N^2 urethane protected, activated amino acids. The rearrangement can be reduced using less reactive leaving groups. It becomes a major problem with the acyl chlorides due to the extremely electronegative chloride ion. This leads to the formation of the oxazolone being more facile than reaction with the amine group.

1.3.3.2 Acyl azides

The use of acyl azides were also introduced at the turn of the century. The azide is less reactive than the acyl chloride and hence the formation of the oxazolone is disfavoured relative to amide formation. The major drawback of azide use is the rearrangement of the azide to form the isocyanate. Further reaction of the isocyanate is then possible as shown in Figure 1.26. This led to the azide method going out of favour when cleaner methods of synthesis became available. The azide method has recently enjoyed a resurgence due to its use in chemical ligation to form large proteins.
1.3.3.3 Carbodiimides

Ever since the initial report of the use of dicyclohexylcarbodiimide (DCC) in the synthesis of peptides\(^{176}\), the use of carbodiimides in peptide synthesis has become widespread. Solid phase synthesis has tended to favour diisopropylcarbodiimide (DIC)\(^{198}\) purely because the urea formed after reaction is more soluble in the DMF than the N,N'-dicyclohexylurea formed by DCC, allowing easy removal. Their use can be broken up into three distinct reaction pathways. These pathways are shown in Figure 1.27.

![Figure 1.26](image_url)

Curtius rearrangement of acyl azides

**Figure 1.26**

1.3.3.3 Carbodiimides
Addition of the amino acid to the DIC leads to the formation of an O-acylurea, giving an extremely potent acylating group. Direct coupling of a free amino group from another amino acid is possible to give the desired amide bond. More usually, the reactive O-acylurea is further reacted with either another carboxy group to give the symmetrical anhydride (R1=R2), or alternatively with a phenol to form an activated ester. Problems stem from the high reactivity of the O-acylurea since intramolecular acyl transfer can compete with the desired attack by external nucleophiles. This leads to the formation of the N-acylurea, coupling of which will lead to problems with separation post synthesis, see Figure 1.28.
Specific rearrangements occur to certain amino acids, such as dehydration of asparagine/glutamine, to give the nitrile, by the mechanism shown in Figure 1.29. This rearrangement can be reduced using trityl protection. To stop these rearrangement problems, and the extensive racemisation which can occur with susceptible carboxy components, the O-acylurea is reacted quickly with another equivalent of the amino acid to form the symmetrical anhydride, as described.
below. Alternatively, the O-acylurea can be reacted with a phenol to give a less reactive species, but still reactive enough to undergo aminolysis. The hydroxy compound must readily react with the O-acylurea, but also must be a good leaving group in its own right. These activated esters will be discussed later.

1.3.3.4 Symmetrical anhydrides

Symmetrical anhydrides are used fairly extensively in solid phase peptide synthesis in this laboratory. They are easily formed from two equivalents of the required amino acid, together with DIC, as described above. They lead to unambiguous aminolysis of the required amino group due to their reactivity. The major drawback in their use is the discarding of one equivalent of expensive amino acid per coupling cycle. Serious side reactions can occur with certain amino acids, the problems occurring with Asn and Gln have already been discussed.

1.3.3.5 Active esters

Active esters were initially synthesised to curtail the waste of reactants inherent in the use of symmetrical anhydrides. There are several different types of active esters, reviewed in reference\(^1\)\(^9\)\(^2\). The most important consideration in the use of active esters is the leaving ability of the ester group, as discussed earlier. Hence, some of the first compounds used were o- and p- nitrophenyl esters\(^2\)\(^0\). This has led to several esters still in use today, including the pentafluorophenol esters\(^2\)\(^1\) (see Figure 1.30).
Other esters have been introduced not linked to the nitrophenol ester, but they all have characteristics of being electron withdrawing. The ones most used are HOSu and HOBt, shown below (Figure 1.31).

HOBt reacts quickly with the O-acylurea, as shown in Figure 1.33. This leads to the formation of the HOBt ester which is much less reactive than the symmetrical anhydride. This reduces the side reactions inherent in the use of carbodiimide already discussed. HOBt ester formation also inhibits the rearrangement to the nitrile seen with Asn and Gln amino acids. More recently, there has been work on new activated esters using HOBt as a template. In particular, the work on HOCl has given rise to activated esters more reactive than their HOBt counterparts. Also, the smaller size of the HOCl alcohol (see Figure 1.32) allows quicker reaction with the O-acylurea and hence reduces racemisation due to oxazolone formation.
Active esters are mainly formed via DIC mediated coupling, reacting with the O-acylurea intermediate.

![Diagram of active ester formation](image)

**Figure 1.32**
HOCT

**Figure 1.33**
Formation of HOBT activated ester using DIC

1.3.4 Side chain protection

As stated earlier, the mild acid cleavage required for the Wang resin allows the use of more acid labile side chain protection. This protection is mostly via t-butyl esters or
ethers. Certain amino acids, such as cysteine, need to be orthogonally protected to allow for further reaction after cleavage. The most used protecting groups in Fmoc synthesis are shown in Figure 1.34.

<table>
<thead>
<tr>
<th>Amino acid</th>
<th>Side chain protection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp and Glu</td>
<td>OBut</td>
</tr>
<tr>
<td>Ser</td>
<td>Bu'</td>
</tr>
<tr>
<td>Thr</td>
<td>Bu'</td>
</tr>
<tr>
<td>Tyr</td>
<td>Bu'</td>
</tr>
<tr>
<td>Lys</td>
<td>Boc</td>
</tr>
<tr>
<td>Arg</td>
<td>Pmc&lt;sup&gt;208&lt;/sup&gt;</td>
</tr>
<tr>
<td>His</td>
<td>π-Bum&lt;sup&gt;209&lt;/sup&gt;, τ-Trt&lt;sup&gt;210&lt;/sup&gt;</td>
</tr>
<tr>
<td>Cys</td>
<td>Trt&lt;sup&gt;211&lt;/sup&gt;, Acm&lt;sup&gt;212&lt;/sup&gt;, SBut&lt;sup&gt;213&lt;/sup&gt;</td>
</tr>
<tr>
<td>Asn and Gln</td>
<td>Mbh&lt;sup&gt;214&lt;/sup&gt;, Trt&lt;sup&gt;199&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

Figure 1.34
The usual Fmoc side chain protecting groups

All other amino acids have no need for side chain protection, as they are unaffected by the conditions used for Fmoc coupling.

1.3.5 Acid Cleavage

Once chain assembly is completed, the peptide is cleaved from the resin at the same time as the side chains are deprotected. Usually TFA is used to affect this cleavage, although stronger acids such as TFMSA<sup>215</sup> have also been used. Deprotection of these side chains leads to the release of reactive carbocation intermediates. These can easily react with the peptide in an irreversible manner to form adducts which are difficult, if not impossible to break down. To counteract this effect, carbocation scavengers are added to the cleavage mix. These are small nucleophilic molecules, able to react with
the carbocations before they react with the peptide. The identity and mix of scavengers used depends upon the amino acids and hence the carbocations present. Scavengers used regularly include water, 1,2-ethanedithiol (EDT), thioanisole and \( m \)-cresol.

1.3.6 Rearrangements

Rearrangements specific to the formation of an activated amino acid have already been covered. Specific rearrangements due to the length of the peptide chain or the presence of specific amino acids or sequence of amino acids include the following.

1.3.6.1 Histidine racemisation

The basic ring of the side chain in histidine leads to the possibility of proton abstraction intramolecularly from the backbone, as shown in Figure 1.35. This leads to reprotonation from the top or bottom face and hence loss of chiral integrity.

![Histidine racemisation](image)

**Figure 1.35**

Histidine racemisation

Protection of the side chain is used to either reduce the basicity of the ring, as in τ-\( \text{Trt} \), or sterically blocking the π nitrogen, as in π-\( \text{Bum} \)^{269}. 

46
1.3.6.2 Diketopiperazine formation

Upon formation of a dipeptide and removal of the Fmoc protection in the usual way, the presence of certain amino acids in the dipeptide, mainly glycine, leads to attack of the peptide resin ester linkage by the free amine., as in Figure 1.36 below.

![Diketopiperazine formation](image)

Figure 1.36

Diketopiperazine formation

1.3.6.3 Aspartimide formation

This is a rearrangement which has caused major problems in the following work. It is discussed in detail later.

1.3.7 Overall Fmoc synthesis

The synthesis of peptides using the Fmoc strategy can thus be summarised as in Figure 1.37 below.
Figure 1.37

Fmoc synthesis
(a) attachment of first activated amino acid, (b) N- deprotection, (c) coupling of next activated amino acid in the chain, (d) repeat until the required sequence is present, (e) removal of final N- protection, (f) cleavage and deprotection of side chains
Chapter 2: Results and Discussion

2.1 Project overview

The impetus for this work comes from the desire to obtain information which may lead to the formation of a vaccine against AIDS, a disease fast becoming one of the most urgent medical emergencies to affect the human population. It took eight years for the first 100,000 cases of AIDS to develop in the US, the second 100,000 took just two years. HIV has now become the second biggest killer of twenty to thirty year olds in America (See Figure 2.1). Thankfully, after many years of little success, there have been great advances in the treatment of HIV infected individuals in the last six months. In particular, combination therapies using reverse transcriptase inhibitors and, more recently, the use of protease inhibitors have led to a relative lengthening of life expectancy in the developed world. Unfortunately, not only do these therapies lead to a relatively small increase in lifespan, they are also prohibitively expensive for use in countries where HIV has had the most devastating effect, namely Africa. The high mortality rate among young working adults there has caused an economic downturn and left more than a million children orphaned. It is therefore essential, especially in the African case, to find a preventative procedure rather than attempt to treat the subsequent illness.

![Figure 2.1](image-url)

**Figure 2.1**

Mortality rates per 100,000 people for 20-30 year old males in the USA
Vaccine technology has advanced greatly since its discovery by Jenner over two hundred years ago. This has led to the eradication of diseases which were the scourge of humanity for many centuries. However, the use of the usual vaccine treatments such as attenuated or whole killed viruses is out of the question with such a highly pathogenic virus as HIV. Not only that, along with areas of gp120 that are relatively conserved between different isolates, the virus contains areas which are highly variable. Antibodies to these areas are necessarily type specific and therefore useless for the broad neutralisation needed to eliminate the virus. The virus is in effect using its high replication rate and high variability to 'hide' from the immune system.

As stated earlier, the initial interaction of HIV with the target cell involves high affinity binding of the coat protein gp120 to the cell surface receptor CD4. This interaction is obviously crucial for the overall infection pathway of the virus and is thus an attractive target for intervention. Interference with binding will halt the whole process of infection in effect before it is even begun. Synthetic peptides are known to be capable of eliciting a specific immune response which will cross react with the intact native protein. Thus a peptide mimicking the 3D structure of the CD4 binding site on gp120 could initiate a humoral immune response leading to antibodies capable of cross reacting with the native virus. This would overcome many of the problems of a vaccine approach to HIV. In particular the CD4 binding site has been shown to be constant, leading to a vaccine effective across all viral isolates.

There have been a great number of gp120 peptides already synthesised and although many were able to raise antibodies which neutralised the virus in vitro, they were unable to neutralise in vivo. This is probably due to the antibodies being raised to peptides equating to linear epitopes. It has already been shown that all epitopes on proteins are discontinuous to a certain extent. Indeed the only characterised broadly neutralising antibodies raised in the course of HIV infection are raised to discontinuous epitopes, and the CD4 binding site itself is discontinuous.
This research described herein is a continuation from that of Dr G. J. Cotton\textsuperscript{218}, consisting of the synthesis of a peptide designed to mimic this discontinuous region of gp120. Olshevski et al\textsuperscript{89} have shown by point mutation studies that the CD4 binding site on gp120 has five amino acids essential for efficient interaction. Further confirmed that these five amino acids were indeed close in space in the folded protein. The peptide synthesised, named GC1, contains three of the five amino acids and a turn designed to present discontinuous along with linear epitopes. The Cys(378)-Val-Cys(445) turn was chosen as this had been shown to oxidise readily\textsuperscript{219,220}. GC1 is shown in Figure 2.2 below. The synthesis of GC1 will be improved and variants synthesised to study the affinity of binding both with soluble CD4, cells expressing CD4 on their surface and also to study the bodies immune response to the aforementioned peptides. The peptides are designed to mimic the discontinuous binding site of CD4 on gp120 to a greater or lesser degree.

(a) C3 and C4 regions of gp120 showing the closeness in space of the aforementioned amino acids. (b) GC1 showing how the peptide mimics the overall gp120 structure

(Adapted from Ref. 218)
2.2 GC1 synthesis: The problems

The coupling of each amino acid in the stepwise synthesis of GC1 is afforded in high yield (see Figure 2.3).

![Amino Acids](Image)

**Figure 2.3**
The coupling profile for a normal synthesis of GC1

Although the coupling is maximised, the synthesis of GC1 is not as straightforward as would be expected for a peptide of its size. The initial work in this project investigated maximising the yield, which was affected by two major factors. Firstly, the purification requires three separations by preparative HPLC. At each there is loss of around 50% of the peptide. The cumulative effect is a large loss of material and thus, the reduction in the number of purification steps is a priority. The major problem in the synthesis is the loss of material via a side reaction initiated by the base deprotection of the α-nitrogen, namely the Asp-Gly rearrangement.

2.2.1 The Asp-Gly rearrangement

The rearrangement of peptides containing the Asp-X motif, where X is a small amino acid, in the presence of acid or base has been known for many years\(^{221,222,223}\). It is
observed in both Fmoc and Boc chemistry and is a problem not completely solved. The rearrangement in base occurs by the mechanism shown in Figure 2.4.

The presence of base leads to an increased nucleophilicity of the backbone nitrogen, allowing cyclisation onto the carbonyl of the aspartic acid side chain. This attack and expulsion of the protecting t-butyl ester leads to the formation of the aspartimide. This can be ring opened by nucleophilic attack. Water present in the cleavage and deprotection procedures can open the aspartimide to form either the α or the β-peptide. Similarly, the piperidide base can ring open. However, due to steric considerations, the β-piperidide predominates.

Thus the presence of the Asp-Gly motif leads to rearrangement during the Naα-Fmoc base deprotection, before the coupling cycle. Cotton\textsuperscript{218} showed that changing the base
used for the deprotection, or the deprotection times did not alleviate the problem sufficiently.

The Asp-Gly motif in GC1 has led to enormous problems over the years due to the above rearrangement. After cleavage and deprotection of the peptide from the resin, three main products are discernible, instead of the one major product usually seen in a linear peptide of this length. The desired product, GC1 (Cys-Acm), was a small part of the overall crude yield (see Figure 2.5).

There have been several attempts at producing novel protecting groups which would negate this effect. The rearrangement is so facile when X is a small amino acid due to the ease in which the deprotonated backbone nitrogen can reach the protected side chain. The main attempts at reducing this rearrangement involve either increasing the steric interaction between the backbone nitrogen and the side chain, or removing the problem hydrogen via backbone protection.
Increasing the steric bulk of the side chain protecting group has been attempted\textsuperscript{224}. More recently, coupling of the troublesome glycine to the large 2-chlorotrityl resin\textsuperscript{225} has led to the same effect (see Figure 2.6(a)). This procedure necessitates the C-terminus of the synthesised peptide to be the glycine of the Asp-Gly motif, coupling the amino acids beyond to the C terminus in a later step. The use of amide bond protection using the 2-hydroxy-4-methoxybenzyl (Hmb) group\textsuperscript{226,227} has also been used to give a high degree of suppression of the rearrangement (see Figure 2.6(b)). Unfortunately, Hmb protected glycine is still very expensive and therefore unsuitable for regular use.

![Figure 2.6](image)

(a) The use of 2-chlorotrityl resin
(b) The use of Hmb backbone amide protection

Martinez and Bodansky\textsuperscript{222} mentioned an efficient suppression of aspartimide formation in Boc chemistry. They added an equimolar amount of tertiary amine to the TFA used to remove the Nα-Boc protection. It was presumed the base was acting by buffering the acid. Dolling et al.\textsuperscript{228} used this information in Fmoc chemistry. They found that the addition of 0.1M of either 2,4-dinitrophenol or HOBT to the Fmoc deprotection solution gave efficient suppression of the aspartimide formation, presumably by buffering the deprotection solution. This technique was found to be an efficient way of suppressing aspartimide formation to workable levels in GC1 and related peptides.
The synthesis of GC1 was repeated using 0.1M HOBt in the solution of 20% piperidine in DMF. The synthesis was monitored at 302nm in the usual way and was observed to give good coupling at each amino acid cycle (see Figure 2.7). Real time assessment of coupling efficiency was affected due to the UV activity of the HOBt molecule at $\lambda=302\text{nm}$. However, as the concentration of HOBT is constant throughout the three deprotections in each cycle, its activity could be removed to give an assessment of the concentration of piperidine / fulvene adduct.

The synthesis was completed and the peptide was cleaved from the resin in the usual way. This led to a crude peptide mix giving an HPLC trace as shown in Figure 2.8, indicating the rearrangement problem had been reduced to manageable levels.
A reduction in the number of preparative HPLC steps required was attempted. The initial work involved the use of different cysteine protection and the use of a novel N° protecting group, Tbfmoc, which has properties which facilitates purification.

2.2.2 Cysteine protection: The use of the cysteine-phenylacetamidomethyl (Phacm) group

Cysteine protection orthogonal to the usual acid deprotection was utilised throughout this project. It had been found earlier that there was a tendency for intermolecular oxidation during cleavage if the acid cleavable trityl group was used. Also, the use of Tbfmoc in the presence of free cysteines causes considerable problems. Cysteine protection utilised included Acm and variants of the phenacetyl group previously used in cysteine side chain protection, lysine side chain protection and the aminoprotection of purine nucleosides. Use of this protecting group would allow enzyme removal to be examined. The removal of Acm is usually achieved with silver triflate, mercury acetate or thallium trifluoroacetate. Such deprotection conditions led to the formation of the metallic...
salt of the peptide, the peptide itself being released by treatment with dithiothreitol. This reaction necessitates a further HPLC step to remove DTT salts and still protected peptide.

Penicillin acylase catalyses the hydrolysis of penicillin G into 6-aminopenicillanic acid and phenylacetic acid²³⁴ (see Figure 2.9).

![Chemical structure](image)

**Penicillin G**

**6-aminopenicillanic acid**

**Figure 2.9**

The hydrolysis of penicillin G

The enzyme is highly specific at the aromatic carbonyl side of the amide bond, yet highly tolerant of substrates variation at the amine side of the amide bond²³¹, allowing the use of a phenacetyl derivative on the side chain of lysine. This reaction was also seen to work with the sulfur of cysteine²²⁹, giving the protecting group shown in Figure 2.10. This enzyme can be attached to a solid support to allow for easy removal from the reaction mix without affecting substrate specificity²³⁵,²³⁶,²³⁷, perhaps allowing the removal of one of the HPLC steps.
Thus the Phacm group was used as the cysteine orthogonal protection in the synthesis of GC1 to evaluate its use and enzyme removal.

The synthesis of GC1 using Phacm protected cysteine was carried out in the usual way, with 0.1M HOBT present in the solution of 20% piperidine in DMF to reduce the aspartimide rearrangement to a minimum. The synthesis proceeded well to conclusion, with the coupling profile showing few problems (see Figure 2.11). Upon cleavage and deprotection, the protected peptide GC1 (Cys-Phacm) was relatively easily isolated using preparative HPLC.
Upon isolation of the cysteine-protected peptide, the removal of the Phacm protection was attempted with immobilised penicillin acylase. The peptide was suspended in buffer at pH 8 and incubated at 37°C for 24 hr with the resin-bound enzyme, the optimum conditions for its use. HPLC was used to monitor the reaction. Unfortunately, the Phacm protection was unaffected by these conditions, leaving the peptide in much the same state. Due to this peptide’s ability to undergo aspartimide rearrangement, the reaction could not be continued for any extended length of time at the optimum pH. It seems likely that the presence of two Phacm groups close together were blocking insertion of either of the groups into the sterically demanding active site of the enzyme. This has been shown to be a long cleft found at the centre of the kidney-shaped molecule. Thus, the Phacm protected cysteine could not be used in conjunction with enzyme deprotection in this case.

Removal of the Phacm protecting group using the usual deprotection protocols was attempted. Mercuric acetate and silver triflate were both used in an attempt to obtain the reduced, deprotected form of GC1. Mercuric acetate deprotection produced a peptide with mass as in Figure 2.12. The Phacm groups had been removed from the peptide, but upon DTT treatment to liberate the peptide from its mercuric salt, and gel column treatment, only a small amount of free peptide was recovered, the rest of the peptide binding mercury irreversibly.
Silver triflate deprotection was attempted using the usual protocol, as described in the experimental section. Complete removal of the Phacm group was not achieved using this technique and analytical HPLC showed a collection of peaks (see Figure 2.13).

Figure 2.13
HPLC trace of silver triflate deprotection of GC1 (Cys-Phacm)

2.2.3 Nα-Tbfmoc protection in purification of GC1 and its analogues

One of the main problems with any synthetic peptide synthesis is the separation of the final product from the truncated peptides present. Tbfmoc methodology was
introduced to utilise the fact the desired product should be the only peptide with an N terminus not acylated. Tbfmoc was synthesised first by Ramage and Raphy\textsuperscript{239} and improved upon by Brown \textit{et al}\textsuperscript{240}. It has been shown that this group could be used for the purification of oligonucleotides\textsuperscript{241} and peptides\textsuperscript{187}. The use of the Tbfmoc group was attempted in the synthesis of GC1.

The Tbfmoc group has the structure as shown in Figure 2.14. The large aromatic nature of the group leads to a high affinity for porous graphitised carbon (PGC)\textsuperscript{242}, and an increase in retention time on HPLC when attached to proteins. Fmoc is removed from the completed peptide whilst still on the resin using a solution of 20\% piperidine in DMF containing 0.1M HOBt. The chloroformate of the Tbf group is sonicated with the resin in dioxane in the presence of equimolar amounts of DIEA. The resin is washed and the fluorescence of the group at 364nm is used to determine the percentage loading of the Tbfmoc moiety. Once satisfactory loading has been achieved, approximately 100\%, the peptide is cleaved from the resin and deprotected. Scavengers are removed using gel chromatography, and the peptide loaded onto PGC. The peptide containing the Tbfmoc group is adsorbed onto PGC, and the truncated peptides removed by copious washing. The desired product is then released from the resin using a basic solution of guanidine hydrochloride. Upon removal of the salt, again using gel chromatography, the desired product can be recovered by lyophilisation.

![Thfmoc in peptide purification](image)

\textbf{Figure 2.14}

Tbfmoc in peptide purification
Tbfmoc loading to GC1 was never satisfactory, something which is presumably down to folding on the resin after removal of the Fmoc group. This led to only partial tagging of the desired product, and increased difficulty in purification and, after several attempts to improve the coupling of the Tbf chloroformate to the resin, this method was abandoned.

2.3 Optimised GC1 synthesis

The work carried out on cysteine protection and Tbfmoc assisted purification were not incorporated in the enhanced synthesis of GC1. The suppression of the aspartimide rearrangement, together with the silver triflate removal of Acm in the presence of scavengers, the isolation of the Acm protected peptide becoming unnecessary, has led to the synthesis of large amounts of the difficult peptide, GC1. Synthesis of GC1-Acm was carried out on a 0.5mmol scale using DIC to form the triazole activated ester. Histidine was coupled as the HOBt activated ester thus reducing racemisation. All amino acids were double coupled. Removal of the Acm group using silver triflate\textsuperscript{233} led to isolation of GC1.reduced, i.e. peptide prior to cystine bond formation (see Figure 2.15).
Oxidation and formation of cystine was carried out using 2.5% DMSO in TFA as in the protocol of Tam et al.\(^{243}\) (see Figure 2.16).

**Figure 2.15**
Removal of Acm with silver triflate in TFA

**Figure 2.16**
Cysteine oxidation using DMSO
Oxidation led to the formation of GC1, which was isolated using preparative HPLC. It gave an analytical HPLC trace, mass spec, Ellman's assay and amino acid analysis identical to that synthesised by Cotton\textsuperscript{218} (see Figure 2.17 and 2.18).

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Expected</th>
<th>Found</th>
<th>Amino Acid</th>
<th>Expected</th>
<th>Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asx</td>
<td>4.00</td>
<td>4.18</td>
<td>Val</td>
<td>2.00</td>
<td>1.79</td>
</tr>
<tr>
<td>Thr</td>
<td>3.00</td>
<td>3.18</td>
<td>Ile</td>
<td>2.00</td>
<td>1.66</td>
</tr>
<tr>
<td>Ser</td>
<td>5.00</td>
<td>4.67</td>
<td>Leu</td>
<td>3.00</td>
<td>3.29</td>
</tr>
<tr>
<td>Glx</td>
<td>1.00</td>
<td>0.98</td>
<td>Phe</td>
<td>1.00</td>
<td>0.98</td>
</tr>
<tr>
<td>Pro</td>
<td>1.00</td>
<td>0.92</td>
<td>His</td>
<td>1.00</td>
<td>1.11</td>
</tr>
<tr>
<td>Gly</td>
<td>5.00</td>
<td>5.22</td>
<td>Lys</td>
<td>1.00</td>
<td>1.38</td>
</tr>
<tr>
<td>Cys</td>
<td>2.00</td>
<td>0.67*</td>
<td>Arg</td>
<td>1.00</td>
<td>1.12</td>
</tr>
</tbody>
</table>

Asx = Asp and Asn, Glx = Glu and Gln. *Cysteine result is uncorrected.

Figure 2.17

Amino acid analysis of GC1

![HPLC trace](image1)

![Mass spectrum](image2)

Figure 2.18

Data obtained on GC1

65
2.4 Synthesis of GC1 analogues

As explained earlier, point mutation studies by Olshevsky et al.\(^8\) led to the realisation that there are five amino acids crucial for high affinity binding of gp120 to CD4. The peptide already synthesised, GC1, has been shown to mimic a discontinuous epitope with three of the amino acids present. To further expand this work, peptides were synthesised mimicking epitopes containing one and two of the necessary amino acids. Binding to CD4, and the ability to raise antibodies was studied. Selective removal of the important amino acids should lead to a reduction of binding of peptides to CD4, and antibodies raised to the native protein, gp120. The peptides were analogues of GC1 with one or both of the amino acids equating to Asp-457 or Asp-457 and Glu-370 substituted with alanine to give the structures shown in Figure 2.19.

![Figure 2.19](image)

The structure of the GC1 analogues, showing the amino acids critical for CD4 binding (boxed) (a) mono-Ala (2 critical amino acids) (b) Bis-Ala (1 critical amino acid)
2.4.1 GC1 mono-Ala

The peptide was synthesised on a 0.25 mmol scale using a single coupling method, with the triazole activated ester of the amino acid formed using DIC. The synthesis proceeded smoothly to give the coupling profile shown below (Figure 2.20).

![Coupling profile of peptide GC1 mono-Ala](image)

Figure 2.20

Coupling profile of peptide GC1 mono-Ala

Cysteine was incorporated as the Acm derivative as before. Differing from the synthesis of GC1, the Acm protected peptide was isolated using preparative HPLC. The Acm group was removed using silver triflate and the cystine bond formed using 2.5% DMSO in TFA. Analytical HPLC traces of the Acm protected, reduced and oxidised peptides are shown below (Figure 2.21).

Once isolated, the peptide gave an amino acid analysis corresponding to the required peptide (see Figure 2.22), Ellmans assay gave a result of 0.00 free cysteines and the accurate mass spectrum was correct to within 0.32 ppm.
Figure 2.21

HPLC traces of (a) Acm protected, (b) reduced and (c) oxidised mono-Ala peptide

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Expected</th>
<th>Found</th>
<th>Amino Acid</th>
<th>Expected</th>
<th>Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asx</td>
<td>3.00</td>
<td>2.99</td>
<td>Val</td>
<td>2.00</td>
<td>1.77</td>
</tr>
<tr>
<td>Thr</td>
<td>3.00</td>
<td>2.72</td>
<td>Ile</td>
<td>2.00</td>
<td>1.63</td>
</tr>
<tr>
<td>Ser</td>
<td>5.00</td>
<td>4.11</td>
<td>Leu</td>
<td>3.00</td>
<td>2.87</td>
</tr>
<tr>
<td>Glx</td>
<td>1.00</td>
<td>0.96</td>
<td>Phe</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Pro</td>
<td>1.00</td>
<td>1.05</td>
<td>His</td>
<td>1.00</td>
<td>0.92</td>
</tr>
<tr>
<td>Gly</td>
<td>5.00</td>
<td>4.82</td>
<td>Lys</td>
<td>1.00</td>
<td>0.90</td>
</tr>
<tr>
<td>Ala</td>
<td>1.00</td>
<td>1.19</td>
<td>Arg</td>
<td>1.00</td>
<td>1.27</td>
</tr>
<tr>
<td>Cys</td>
<td>2.00</td>
<td>1.01*</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Asx = Asp and Asn, Glx = Glu and Gln. *Cysteine result is uncorrected for oxidative effects.

Figure 2.22

Amino acid analysis of GC1 mono-Ala
2.4.2. GC1 bis-Ala

The above peptide was synthesised on a 0.25 mmol scale using a single coupling method, with the triazole activated ester of the amino acid formed using DIC. The synthesis proceeded relatively smoothly until the final two amino acids, when a large drop in coupling efficiency was observed, giving the profile shown below (Figure 2.23). Upon cleavage and deprotection, a fairly large amount of unwanted products were present. This can be seen in the analytical HPLC of the crude Acm protected material (see Figure 2.24), and in the amino acid analysis of the resin which indicated the presence of truncations.

![Figure 2.23](Image)

**Figure 2.23**

Coupling profile of GC1 bis-Ala

![Figure 2.24](Image)

**Figure 2.24**

Analytical HPLC of crude GC1 bis-Ala
Due to the relatively disappointing synthesis of GC1 bis-Ala and the presence of lysine at the N terminus of the completed peptide, ion exchange chromatography was attempted. pH 9 was found to be the ideal conditions for separation of the required peptide from its deletions. Unfortunately, upon scale up, it was noted that the peptide was not sufficiently soluble, even at relatively low concentrations, in the buffer for this method to be used. Tbfmoc separation in conjunction with porous graphitised carbon was also attempted in the same way as described for GC1 in section 2.2.3, with little success.

The Acm peptide was not separated, as for GC1. Removal of the Acm group was effected using the silver triflate method previously described (see section 2.3). Upon separation of the reduced peptide using HPLC, cystine formation was carried out using 2.5% DMSO in TFA, as for GC1. Upon separation of the peptide, the amino acid analysis and mass spectrum were in good agreement with expected results. Analytical HPLC showed the peptide to be >95% pure and Ellmans assay gave a result of 0.00 free cysteines, indicating correct oxidation (see Figure 2.25).

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Expected</th>
<th>Found</th>
<th>Amino Acid</th>
<th>Expected</th>
<th>Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asx</td>
<td>3.00</td>
<td>2.73</td>
<td>Val</td>
<td>2.00</td>
<td>1.64</td>
</tr>
<tr>
<td>Thr</td>
<td>3.00</td>
<td>2.74</td>
<td>Ile</td>
<td>2.00</td>
<td>1.58</td>
</tr>
<tr>
<td>Ser</td>
<td>5.00</td>
<td>4.13</td>
<td>Leu</td>
<td>3.00</td>
<td>2.99</td>
</tr>
<tr>
<td>Pro</td>
<td>1.00</td>
<td>0.90</td>
<td>Phe</td>
<td>1.00</td>
<td>0.89</td>
</tr>
<tr>
<td>Gly</td>
<td>5.00</td>
<td>4.85</td>
<td>His</td>
<td>1.00</td>
<td>0.91</td>
</tr>
<tr>
<td>Ala</td>
<td>2.00</td>
<td>2.15</td>
<td>Lys</td>
<td>1.00</td>
<td>0.79</td>
</tr>
<tr>
<td>Cys</td>
<td>2.00</td>
<td>0.57*</td>
<td>Arg</td>
<td>1.00</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Glx = Glu and Gln. *Cysteine result is uncorrected for oxidative effects.

(a) Amino acid analysis
To understand the role of the Cys(378)-Val-Cys(445) turn in both the presentation of the correct discontinuous epitopes, and its effect on overall antigenicity, a peptide was synthesised without this turn. The Cys(378)-Val-Cys(445) turn was replaced with an entirely synthetic moiety, 3-aminocyclopentane-1-carboxylic acid. (see Figure 2.26).

![cis-3-Aminocyclopentane-1-carboxylic acid](image)

This structure is thought to act as a β-turn mimic. The pseudo-diequatorial form of cis-aminopentane-1-carboxylic acid is energetically favoured relative to the diaxial form. However, upon addition of two peptide chains, the secondary structure
interactions can lead to the diaxial form being favoured. The peptide in question was synthesised by repeating the work of Urquhart\textsuperscript{244}, using the improved synthetic procedures to reduce aspartimide rearrangement.

2.5.1 Synthesis of GC1.cyclopentyl

The peptide was synthesised on a 0.25 mmol scale using a single coupling\textsuperscript{2} method, with the triazole activated ester of the amino acid formed using DIC. Automated synthesis was used apart from the coupling of the cyclopentyl moiety, which was single coupled manually using four fold excess of the triazole activated ester. Fmoc deprotections were carried out using a solution of 20\% piperidine in DMF buffered with 0.1M HOBt. The coupling profile showed no serious drops in efficiency (see Figure 2.27).

![Amino Acids](image)

**Figure 2.27**

Coupling profile of GC1.cyclopentyl

The synthesis led to a relatively clean peptide upon cleavage and deprotection. Few truncations or deletions were seen and there was little aspartimide rearrangement. As the Cys(378)-Val-Cys(445) moiety has been replaced, isolation of the major peak by HPLC led to the required peptide. Data collected on the peptide was found to be consistent with that already synthesised (see Figure 2.28).
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Expected</th>
<th>Found</th>
<th>Amino Acid</th>
<th>Expected</th>
<th>Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asx</td>
<td>4.00</td>
<td>3.92</td>
<td>Ile</td>
<td>2.00</td>
<td>1.67</td>
</tr>
<tr>
<td>Thr</td>
<td>3.00</td>
<td>2.78</td>
<td>Leu</td>
<td>3.00</td>
<td>2.97</td>
</tr>
<tr>
<td>Ser</td>
<td>5.00</td>
<td>4.23</td>
<td>Phe</td>
<td>1.00</td>
<td>0.88</td>
</tr>
<tr>
<td>Glx</td>
<td>1.00</td>
<td>1.04</td>
<td>His</td>
<td>1.00</td>
<td>0.92</td>
</tr>
<tr>
<td>Pro</td>
<td>1.00</td>
<td>0.75</td>
<td>Lys</td>
<td>1.00</td>
<td>0.90</td>
</tr>
<tr>
<td>Gly</td>
<td>5.00</td>
<td>5.12</td>
<td>Arg</td>
<td>1.00</td>
<td>0.95</td>
</tr>
<tr>
<td>Val</td>
<td>1.00</td>
<td>0.78</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Asx = Asp and Asn, Glx = Glu and Gln

(a) Amino acid analysis

(b) Analytical HPLC

(c) Mass Spectrum

Figure 2.28
Data collected on GC1.cyclopentyl
Once peptides containing one, two and three of the amino acids necessary for high affinity binding had been synthesised, a peptide was designed to integrate the fourth amino acid. Using the ideas perfected in the synthesis of GC1, a peptide was designed which could be formed linearly using solid phase peptide synthesis, but containing structural information allowing it to present both discontinuous and continuous epitopes.

The GC1 template afforded two potential strategies. Either the peptide containing the fourth amino acid could be attached to the ends of the GC1, either the C- or N-terminus or, with a slight alteration, it could be hung from the centre between the two cysteines. The latter was attempted as this was thought to be more likely to give the discontinuous epitopes required. The sequence of the template GC1 was changed to accommodate this strategy, incorporating a second N-terminus located between the two cysteines, by substituting the valine of the turn with a lysine to give the structure shown in Figure 2.29.

The extra amino acid was incorporated because of its involvement in high affinity binding\(^{89}\). However, this is not the only reason this area of gp120 is of interest. The amino acid Trp-427 was shown by Olshevsky to be extremely important for binding. Lasky et al\(^{87}\) also showed this, but seemed to find the whole region around Trp-427 critical for interaction with CD4. Deletion of Trp-427 results in a significant decrease in binding, whereas deletion of twelve amino acids from the area in question led to a complete loss of binding. Other work has also shown that this area is involved in a conformational switch of the whole gp120 molecule on moving from hydrophilic to a hydrophobic region, as happens when a virion encounters a cell surface\(^{245,246,247,248}\). The region has been shown to have the potential to form both \(\beta\)-sheet and \(\alpha\)-helix structure and sharply switches between the two. This area seems to be extremely important in binding and post-binding events and thus it was deemed to be a useful area to study.
To accomplish the synthesis, protection of the extra nitrogen of lysine needed to be orthogonal in relation to both base and acid, allowing selective deprotection when required. Hopp and Woods\textsuperscript{249} calculations carried out with the amino acids of GC1 (not shown) predicted a very hydrophobic area around the turn. IH1 was based
closely on the structure of GC1. Thus, the α-nitrogen was orthogonally protected and the peptide grown as normal from the Fmoc protected ε-nitrogen to reduce the chances of folding of the peptide on the resin obscuring the second N-terminus.

![NH3⁺ ε-nitrogen](image)

**Figure 2.30**
The amino acid lysine

### 2.6.1 Orthogonal lysine protection

There are several orthogonal protecting groups available for use with lysine. Enzyme cleavable protection, such as the Phenac group already mentioned, is unsuitable in this instance due to the swelling problems of the Wang resin in the aqueous solution required. Photosensitive protecting groups such as 6-nitroveratryloxycarbonyl (NVOC) and 2-nitrobenzyloxycarbonyl (NBOC) have been known for many years and have been used in varying ways, recently reviewed. One of the most useful has been shown to be the Alloc group first introduced for ε-amino protection by Stevens and Watanabe and later improved by Kunz and Crivici (see Figure 2.31(a)). Stable to both acids and bases, the Alloc group can be removed using several different palladium-catalysed deprotection conditions.
Figure 2.31

N(ε) protection of Lysine

The Dde group\textsuperscript{259} was found to be most convenient for removal under solid phase conditions (see Figure 2.32(b)). Dde is orthogonal to both Fmoc and Boc protection. It is stable to repeated Fmoc deprotections using piperidine, but is extremely labile using 2% v/v hydrazine in DMF, by the mechanism shown below (see Figure 2.32). The group is very versatile, with peptides containing Dde being used in template synthesis\textsuperscript{260}, in the synthesis of diepitopal multiple antigen peptides\textsuperscript{261} and cyclic peptides\textsuperscript{262}. The Dde group has also been used as a temporary protecting group\textsuperscript{263}, a useful property in relation to this research. The versatility and dependable nature of the group, together with its compatibility with solid phase synthesis, led to the groups use in the formation of IH1.
2.6.2 Synthesis of IH1

The peptide was synthesised on a 0.25 mmol scale using the single coupling method, with the triazole activated ester of the amino acid formed using DIC. The synthesis was carried out up to Cys-445 using normal automated synthesis. The Dde-Lys (Fmoc) OH was then manually coupled using four times excess of the activated triazole ester. This allowed expansion of the chain through the ε-nitrogen, with the introduction of the new amino acid sequence containing the fourth important amino acid, again synthesised using automation. Lys-433 was introduced using Fmoc-Lys (Fmoc)OH. Upon coupling, the Fmoc was removed and the two free NH₂ groups were capped using acetic anhydride. The Dde group located on lysine was then removed, Cys-378 was then coupled using automated synthesis and the peptide assembly was completed by growing from the α-nitrogen.
The synthesis proceeded relatively well, with a slight drop after the manual coupling of Dde protected lysine (see Figure 2.33(a)). After completion of the linear peptide, the Dde group was removed and the branched peptide completed. The addition of the rest of the amino acid sequence proceeded smoothly, as can be seen from the coupling profile (Figure 2.33(b)) (NB, the coupling percentages for the second part of the peptide, in Figure 2.33(b) are calculated from lysine between Cys-378 and Cys-445, with lysine taken as 100%).

(a) Coupling profile from Gly-445 to Lys-433 of IH1

(b) Coupling profile of second part of IH1

Figure 2.33

IH1 was isolated in a similar fashion to GC1. The main additional problem associated with IH1 was the presence of the amino acid methionine, which is prone to oxidation on cleavage and deprotection\textsuperscript{264,265} to form methionine sulfoxide. Cleavage and
Deprotection of the completed peptide was carried out in the usual way, with a small amount of thioanisole added in an attempt to suppress any acid oxidation of the vulnerable methionine. Analytical HPLC showed a reasonably clean peptide, with little of the oxidised product present (see Figure 2.34).

![Figure 2.34](image)

Analytical HPLC of crude IH1

The Acm protected peptide was not isolated, but deprotected immediately. This was first attempted using silver triflate in TFA, as for GC1. Analytical HPLC traces of the subsequent product showed a clean reaction. However, the product became more hydrophilic, contrary to that normally seen on removal of the Acm group (see Figure 2.35).
Analytical HPLC of the removal of Acm from IH1 using the silver triflate / TFA procedure (a) Acm protected IH1, (b) IH1 after silver triflate / TFA procedure

This shift to a shorter retention time indicated other reactions occurring at the same time as the Acm removal. Oxidation of methionine is known to lead to an increase in hydrophilic nature of the peptide due to the polarity of the S-O bond (see Figure 2.36)

Figure 2.36
Resonance structures of Met(O)
It seemed likely that the methionine was being oxidised during Acm removal. Further evidence of this was observed in the mass spectrum of the peptide, being 16 a.m.u. higher than that required for the reduced IH1.

Acm removal was then attempted utilising mercuric acetate in 50% acetic acid / water, this weaker acid should diminish the methionine oxidation. The reaction was initially carried out in air, leading to removal of the Acm groups, but relatively large amounts of oxidation was observed. Carrying out the deprotection under nitrogen reduced the amount of methionine oxidation in the above reaction, but it was still significant. Sonication of the acetic acid / water solution to remove dissolved oxygen, before its addition to the peptide, and a reduction in the amount of stirring given to the reaction, led to a great improvement in the amount of methionine oxidation (see Figure 2.39).

Acm removed IH1, methionine sulfoxide

Acm removed IH1

Figure 2.39
(a) Removal of Acm with mercury acetate under air and (b) after sonicating the solvent and under nitrogen

Removal of the Acm protection allowed the isolation of the reduced peptide by preparative HPLC. Once achieved, cystine formation was attempted. The most
successful method for formation of the bond in GC1 was with the use of DMSO in TFA. This gave short reaction times and good yield of the required product. Unfortunately, when attempting the same procedure with IH1, not only did cystine form, but the methionine was oxidised.

An oxidation procedure was required mild enough to allow cystine formation without methionine oxidation. Cotton$^{218}$ attempted to use air oxidation with GC1, but found the reaction slow and leading to side products. It was known that formation of the cystine bridge gives a characteristic peak shift with the reduced peptide eluting later than the oxidised. Air oxidation was attempted by dissolving the peptide in 50% acetonitrile / water and allowing to stand for several days. Analytical HPLC was used to follow the reaction. It was seen over time that most of the peak corresponding to the reduced peptide was converted to this new peak (see Figure 2.38).

![Figure 2.38](image-url)

Figure 2.38
Monitoring Disulfide bond formation by HPLC
Preparative HPLC was used to separate the peak. Amino acid analysis gave results corresponding to the correct peptide, along with Ellman's assay (expected = 0.00, obtained = 0.00) and mass spectrometry. Analytical HPLC showed >95% purity (see Figure 2.39)

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Expected</th>
<th>Found</th>
<th>Amino Acid</th>
<th>Expected</th>
<th>Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asx</td>
<td>5.00</td>
<td>5.10</td>
<td>Met</td>
<td>1.00</td>
<td>0.65</td>
</tr>
<tr>
<td>Thr</td>
<td>3.00</td>
<td>2.93</td>
<td>Ile</td>
<td>3.00</td>
<td>2.93</td>
</tr>
<tr>
<td>Ser</td>
<td>5.00</td>
<td>4.10</td>
<td>Leu</td>
<td>3.00</td>
<td>3.58</td>
</tr>
<tr>
<td>Glx</td>
<td>4.00</td>
<td>3.91</td>
<td>Phe</td>
<td>2.00</td>
<td>1.69</td>
</tr>
<tr>
<td>Pro</td>
<td>1.00</td>
<td>0.78</td>
<td>His</td>
<td>1.00</td>
<td>0.88</td>
</tr>
<tr>
<td>Gly</td>
<td>7.00</td>
<td>7.30</td>
<td>Lys</td>
<td>3.00</td>
<td>2.57</td>
</tr>
<tr>
<td>Cys</td>
<td>2.00</td>
<td>0.00*</td>
<td>Arg</td>
<td>1.00</td>
<td>1.02</td>
</tr>
<tr>
<td>Val</td>
<td>2.00</td>
<td>1.70</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Asx = Asp and Asn, Glx = Glu and Gln. *Cysteine result is uncorrected for oxidative effects.

(b) Analytical HPLC

(c) Mass Spectrum

Figure 2.39

Data obtained on the peptide IH1
2.6.3 Circular dichroism of peptide IH1

Circular dichroism is a technique to measure the optical activity and hence probe the structure of asymmetrical molecules in solution. CD signals are observed in the same UV region where the absorption of particular chromophores are found. In CD, a peptide solution is exposed to plane polarised light. If the peptide contains structural elements in which the individual amino acids are arranged non-randomly, then the light will be turned through an angle $\theta$ relative to the polarisation of the incident light. This angle gives information on the said structural elements. In general, the more secondary structure a protein contains, the larger the angle $\theta$ will be.

CD bands in the amide region contain information about peptide bonds and secondary structure. The far UV CD spectra was determined for IH1 at the Scottish circular dichroism facility at Stirling University\textsuperscript{267}. The secondary structure was then determined using the Contin secondary structure program\textsuperscript{268}, giving results as shown below.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>% $\alpha$-helix</th>
<th>% $\beta$-sheet</th>
<th>% remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>aqueous pH 7</td>
<td>3 ± 0.5</td>
<td>39 ± 0.5</td>
<td>58 ± 1.0</td>
</tr>
<tr>
<td>50% TFE</td>
<td>38 ± 1.1</td>
<td>29 ± 1.2</td>
<td>33 ± 2.0</td>
</tr>
</tbody>
</table>

The CD spectra obtained from IH1 showed low levels of $\alpha$-helical structure. Upon addition of trifluoroethanol, structure was induced corresponding to 38% $\alpha$-helix, with only a relatively small reduction in the amount of $\beta$-sheet present. The CD spectra is shown in Figure 2.40
Figure 2.40

CD spectrum of H1

6: ------ IH 3:7 50% TFE 0.02cm 4/10/95 CORROPT
5: ------ IH 3:7 0.5mg/ml 0.02cm 4/10/95 CORROPT
Note: In some of the following experiments serum obtained directly from the mice was used without further purification. In others, purified IgG was used. In each case, the experiment discussion makes clear which of the two was utilised. FMDV is a random peptide with 44 residues used in some of the following experiments as a negative control.

2.7 Immunogenicity and Epitope Mapping Studies of IH1

Investigation of the immunogenic properties of the peptides was carried out. 10µg of purified GC1, GC1 mono-Ala, GC1 bis-Ala, GC1 cyclopentyl or IH1 was injected interperitoneally into mice in Freund’s adjuvant. Boost injections of 10µg of peptide in saline were carried out every three weeks until the maximum endpoint titre was seen. Blood was recovered and allowed to clot, with centrifugation then allowing the isolation of the immune serum. Peptide bound to 96 well microtitre ELISA plates was incubated with serial dilutions of anti-serum, or normal mouse serum. Bound antiserum was detected using horseradish peroxidase conjugated goat anti-mouse Ig(G,A,M) followed by o-phenylenediamine. The colouration reaction was stopped by the addition of 2N HCl (50µl) and the optical density (OD) of the wells read at 490nm. This gave the graph shown in figure 2.41.
Titration of anti-IH1 antiserum against IH1. All values corrected for background absorbance of normal mouse serum at each dilution.

Epitope mapping studies investigated the presence of discontinuous epitopes on peptide IH1. Three linear peptides MRC 740-34 (KQSSGDPEIVTHSFNCGG), MRC 740-42 (GQIRCNSNTTGLTRDGQNS) and IH2 (KQFINMWQEVG) were used to map the continuous epitopes contained within the peptide. Serial dilutions of purified anti-peptide IgG and purified normal mouse IgG were incubated with the peptide and bound immunoglobulin detected as before. This gave results as shown in figure 2.42.
Comparison of antibody binding to peptide IH1 and linear peptides MRC 740-34 (KQSSGDPDEIVTHSFNCGGE), MRC 740-42 (GQIRCSSNITGLLTRDGQNS) and IH2 (KQFINMWQEVG). All values corrected for background absorbance of normal mouse serum at each dilution.

From the above, when a mixture of equimolar amounts of linear peptides was incubated with the purified serum, the OD values obtained were significantly lower than those to IH1 itself. These results show that component(s) of the IgG fraction recognise linear epitopes in both the C3 and C4 regions of gp120, but not to the same extent as the peptide IH1 is recognised. The peptide structure thus contains discontinuous epitopes, analogous to peptide GC1²¹⁸.
2.8 Immunogenicity of GC1, GC1.cyclopentyl, GC1 mono-Ala and GC1 bis-Ala

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Endpoint titre</th>
</tr>
</thead>
<tbody>
<tr>
<td>IH1</td>
<td>1:5000</td>
</tr>
<tr>
<td>GC1 mono-Ala</td>
<td>1:5000</td>
</tr>
<tr>
<td>GC1 bis-Ala</td>
<td>1:10000</td>
</tr>
<tr>
<td>GC1.cyclopentyl</td>
<td>no titre obtainable</td>
</tr>
<tr>
<td>GC1</td>
<td>1:150000</td>
</tr>
</tbody>
</table>

Figure 2.43

Endpoint titres of anti-peptide sera against peptides

Antiserum to peptide IH1 gives low cross binding. GC1 produced high titres of anti-peptide antibodies, confirming the results of Cotton et al\(^269\). This study showed the presence of a high level of class switched antibodies pointing to the presence of both B- and T-cell epitopes. Immunisation with GC1.cyclopentyl has been shown to raise negligible amounts of anti-peptide antibodies. Thus it seems likely that the Cys(378)-Val-Cys(445) turn is significant, either in presentation of the T-cell epitope, or residues in the turn are involved in the epitope itself with changes reducing the immunogenicity of the resultant peptide.

As can be seen from the results of serial dilutions of serum raised to peptides GC1 mono-Ala and GC1 bis-Ala, the removal of the charged residues from the peptide also affects immunogenicity. It is known the peptides found to bind to the MHC-I, the cellular protein used in the presentation of epitopes to T-lymphocytes, tend to be short peptides. These peptides are approximately eight to ten amino acids in length and generally contain charged residues\(^270\). It seems likely the removal of these charged amino acids in the mono and bis mutants would lead to disruption in binding of the T-cell epitope present to the MHC-I, and hence affect the efficiency of the humoral response to the peptide.
Antisera raised to GC1 and IH1 were tested for their ability to bind other peptides in the series. It was observed that in both cases there was maximum binding of the peptide the sera it was raised against. However, IH1 also bound equally well to α-GC1 IgG as GC1 itself, indicating IH1 has the ability to adopt conformations similar to those found in GC1. GC1 was not as well recognised by α-IH1 IgG as IH1 itself, suggesting there are epitopes present on IH1 not present on GC1. All other peptides bound with reduced ability to both sera, due to the structural differences present.
2.10 Binding of IH1 to CD4+ H9 cells

To demonstrate binding of IH1 to cells expressing CD4, IH1 or a random peptide FMDV were incubated with H9 cells (carried out by Dr. S.E.M Howie, Dept. of Pathology). Bound peptide was detected using either purified anti-GC1 IgG (shown to cross react with peptide IH1) or IgG from normal mouse serum, followed by incubation with phycoerythrin (PE) conjugated goat anti-mouse IgG. The +ve cells could then be visualised using the fluorescence of PE by flow cytometry. The results below (Figure 2.45) show binding of peptide at differing dilutions. Only in the presence of IH1 and anti-GC1 IgG was binding seen. The results are shown graphically in Figure 2.46.

![Flow cytometry results for H9 cells with IH1 and the random peptide FMDV](image)

**Figure 2.45**
Flow cytometry results for H9 cells with IH1 and the random peptide FMDV
Figure 2.46

Binding of IH1 and FMDV to CD4+ H9 cells at different molarities (all values corrected for background absorbance of normal mouse IgG at each dilution)
2.11 Binding of peptides to parent HeLa cell line, or CD4+, transfected HeLa cells

It was shown above, IH1 is binding to CD4+ H9 cells. Cotton et al\textsuperscript{269} showed GC1 was binding to CD4 on the surface of cells. To demonstrate binding of IH1 and GC1 variants to CD4, the peptides, along with the random peptide FMDV, were incubated with CD4+ HeLa cells and HeLa cells transfected to express surface CD4 (performed by Dr. S.E.M Howie, Dept. of Pathology). As above, bound peptide was detected using either purified anti-GC1 IgG or IgG from normal mouse serum, followed by
incubation with phycoerythrin (PE) conjugated goat anti-mouse IgG. It can be seen from the results above (Figure 2.47) that all peptides, except for the random peptide and IH1, bind to the CD4 transfected HeLa cells, but not to the parent cell line. The random peptide does not seem to bind to either cell lines. IH1 however, binds to the parent CD4⁻ cell line as well as the CD4⁺ transfected cells, but with greater affinity for the CD4⁺ cell line. This indicates binding to CD4 as well as surface receptors other than CD4.
2.12 Binding of IH1 to an area of the cell surface proximate to CD4

It had previously been shown that peptide GC1 binds to HeLa cells transfected to express CD4 on the cellular surface, but not the parent cell line\textsuperscript{269}. From the peptide binding studies to HeLa and CD4\textsuperscript{+} HeLa cells, it was evident all peptides were interacting with the cellular surface through the receptor CD4 to a certain extent. To further prove IH1 was binding to the CD4 surface receptor, bound IH1 was localised on the surface of cells using dual immunofluorescence studies (carried out by M. Fernandes, Dept. of Pathology). CD4\textsuperscript{+} MM6 (macrophage) cells were incubated with IH1 and then with antisera raised to the peptide GC1. Bound antisera was detected using TRITC-labelled goat antimouse immunoglobulin. The cells were also incubated with biotinylated α-CD4 IgG (MT310) and detected using FITC labelled avidin. Co-localisation of colouration on the surface of the cell showed binding of the peptide IH1 to the cell surface in an area very close to the CD4 receptor.

![Figure 2.48](image)

**Figure 2.48**

Binding of a) α-GC1 labelled with TRITC, b) α-CD4 labelled with FITC to cells pre-incubated with peptide IH1. Colouration shows both antibodies co-localised to the same region of the cell
2.13 IH1 Binds to the CDR2 Region in Domain 1 of CD4

The peptide IH1 and CD4 was shown to be co-localised on CD4⁺ cells above. To conclusively demonstrate that IH1 binds to CD4 and to identify the region of CD4 involved in the binding, IH1 was used to block the interaction of antibodies recognising different regions of CD4 (carried out by M. Fernandes, Dept. of Pathology). The results below (Figure 2.49) show that pre-incubation of CD4 expressing MM6 cells with IH1 blocks subsequent binding of antibodies directed against domain 1, and in particular the CDR2 region of CD4 (Q4120), but not antibodies against other regions such as domain 4 (L120). IH1 is binding to CD4, using the same area of the CD4 molecule as gp120, the CDR2 region, which itself indicates similarities in structure between the designed discontinuous epitope and the native protein.

![Bar graph showing binding of α-CD4 immunoglobulins](image)

**Figure 2.49**

Binding of α-CD4 immunoglobulins showing blocking of Q4120, an antibody binding to domain 1 of CD4, but not L120, binding to domain 4 on the surface of MM6 cells
2.14 Peptide inhibition of MIP-1\(\alpha\) binding

Studies above (Section 2.11, Figure 2.47) indicated binding of peptide IH1 to the CD4\(^+\) parent HeLa cell line. This unexpected result led to an attempt to pinpoint the alternate binding site. HeLa cells were known to express fusin on the cellular surface, the chemokine receptor shown to be involved in HIV fusion as a secondary co-receptor for T-cell tropic viruses\(^{149}\). Due to a lack of suitable reagents, the binding of IH1 to the fusin chemokine receptor could not be proven. CD4\(^+\) H9 cells were shown to express CC-CKR5 on its surface, the chemokine receptor known to act as a secondary co-receptor for macrophage-tropic HIV\(^{157,158,159,160,161}\). This chemokine receptor, CC-CKR5, is known to bind the chemokine MIP-1\(\alpha\) strongly and thus the possible interaction of the peptide IH1 with CC-CKR5 could be probed using the available photolabelled MIP-1\(\alpha\) reagent (R&D Systems, Europe Ltd.) (performed by M. Fernandes, Dept. of Pathology).

Before initiating the experiment, the hydrophobicity of the N-terminal areas of the chemokines MIP-1\(\alpha\), MIP-1\(\beta\) and RANTES known to bind to this receptor were compared with areas of IH1 using Hopp and Woods plots.
The similarity in the hydrophobicity values for IH1 with those of the N-termini of the chemokine showed IH1 could interact with the chemokine receptor CC-CKR5. To demonstrate the validity of this argument, the binding of photolabelled MIP-1α to its receptor in the presence of IH1 was studied. Cells were pre-incubated with the peptides at 0.01mM, and treated with the labelled MIP-1α. Reduction of binding of the chemokine would indicate peptide interaction with a surface component intimately involved with the chemokine binding site. MIP-1α binding was visualised using flow cytometry (see figure 2.53), giving the results shown below (figures 2.51) and tabulated in figure 2.52.
Figure 2.51

Binding of photolabelled MIP-1α to H9 cells after pre-incubation with 1) GC1 mono-Ala, 2) GC1 bis-Ala, 3) GC1, 4) IH1, 5) GC1.cyclopentyl, 6) IH2, 7) FMDV
Figure 2.52
Tabulation of inhibition of MIP-1α showing the greatest reduction with peptide IH1

Figure 2.53
Inhibition of MIP-1α binding with IH1, but not with peptide FMDV, pictured using flow cytometry
It can be seen from the results in Figure 2.52 that IH1 blocked interaction of MIP-1α with its receptor to the greatest extent. Further experiments were then attempted using both MM6 (macrophage) and H9 (T cell) and IH1 at different concentrations, to give the inhibition of MIP-1α binding results as shown in Figure 2.54. Pre-incubation with IH1 led to a reduction in the amount of photolabelled MIP-1α for both cell lines, with a negligible link with concentration at the concentrations used.

![Figure 2.54](image)

Inhibition of MIP-1α binding by pre-incubation of IH1 with a) MM6 (macrophage cell line), and b) H9 (T cell line)
2.15 Induction of apoptosis

The role of cross linked CD4 with α-CD4 immunoglobulins in the Fas driven apoptosis pathway is well documented. It has been proposed that the progressive depletion of CD4+ cells in HIV infection is also due to this upregulation of Fas induced by cross linking with gp120 and host α-gp120 immunoglobulin. Work by Cotton has shown GC1 and α-GC1 IgG can cause apoptosis in the CD4+ H9 cell line. The ability of peptide IH1 to induce apoptosis was investigated (performed by Dr. S.E.M Howie, Dept. of Pathology). The poor immunogenicity of the peptide led to the use of α-GC1 IgG, having previously shown recognition of IH1 using the said immunoglobulin, indicating the ability to cross link CD4 in the presence of bound peptide. Plates were coated with α-GC1 IgG or normal mouse IgG and H9 cells pre-incubated with peptide IH1 and a random peptide, FMDV, were added to the plate. Apoptosis was counted visually using a Lietz optical microscope. The results (figure 2.57) show IH1 does indeed cause increased apoptosis in the presence of α-GC1 IgG.

![Graph showing apoptosis induced in the presence of IH1 and α-GC1 IgG](image_url)

**Figure 2.55**

Apoptosis induced in the presence of IH1 and α-GC1 IgG
Chapter 3: Experimental

3.1 Notes

All amino acids were purchased from Bachem, except Dde-Lys(Fmoc)OH purchased from Novabiochem, and are of the L configuration. Dimethylformamide, 1,4-dioxane and piperidine were peptide synthesis grade and supplied by Rathburn Chemicals. Acetic anhydride, DMAP and HOBt were obtained from Aldrich. Peptide synthesis grade TFA was obtained from ABI. All other solvents were distilled before use. UV absorption spectra were recorded on a Varian Cary 210 spectrophotometer or a Perkin Elmer Lambda 11 in the indicated solvent. Sonication of solutions and resins were performed with a Decon FS300b sonic bath. Fast atom bombardment mass spectra (FAB), both high and low resolution, were measured on a Kratos MS50TC machine using thioglycerol matrix. Amino acid analysis was performed on a LKB 4151 alpha plus or a Pharmacia Biotech Biochrom 20 after Carius tube hydrolysis with constant boiling HCl at 110°C for the indicated time. Circular dichroism spectra were recorded on a JASCO J600 spectropolarimeter with the indicated path length and solvent. ELISA were performed on a 96-well Costar EIA / RIA microtitre plates and the optical densities measured using a Dynatech MR5000 microplate reader. Fluorescence measurements in flow cytometry were measured on a Coulter EPICS CS machine. In apoptosis experiments, cells were counted with a Lietz UV optical microscope. Gel filtration and ion exchange chromatography were performed using Pharmacia / LKB columns under the stated conditions together with two LKB 2138 UVICORDS (λ = 226 and 277 nm), a Pharmacia 2132 Microperpex peristaltic pump, an LKB 2112 redirac fraction collector and a Pharmacia GM gradient mixer. High performance liquid chromatography (HPLC) was carried out on one of two systems: an ABI 1783A programmable detector, two ABI 140A units for solvent delivery and an ABI 1480A injector, or a Gilson 305 control pump, a Gilson 306 slave pump, an ABI 785A programmable detector and a Gilson 811C dynamic mixer. The HPLC columns and gradients are given in the text.
3.2 Solid Phase Peptide Synthesis

All peptides were synthesised using amino acids N\textsuperscript{a} protected using the 9-
Fluorenylmethoxycarbonyl (Fmoc) group, with orthogonal, acid labile protection for
side chains where necessary, and an acid labile peptide-resin linker as described in the
text. Where more than one side chain protection is available, the text clarifies the one
used.

3.2.1 Preparation of Fmoc-amino acid resins

A solution of Fmoc-amino acid (3.52mmol) and N,N'-diisopropylcarbodiimide (DIC,
1.86mmol) in the minimum DMF (6ml) was sonicated for 15 min. This solution was
then added to p-benzyloxybenzyl alcohol (Wang) resin (functionalised at 0.88mmol/g,
1g, 0.88mmol) pre-swollen with 5ml DMF and treated with 4-(N,N'-dimethylamino)-
pyridine (DMAP, 30mg, 0.2mmol). The mixture was then sonicated for 1.5hr. The
sonication time was varied depending upon the required level of substitution. The
resin was collected by filtration, washed thoroughly with DMF, Dioxane and diethyl
ether and dried \textit{in vacuo} to yield the Fmoc-amino acid resin.

3.2.2 Determination of loading of Fmoc-amino acid resin

Samples of the thoroughly dried resin (1-3mg) were accurately weighed and added to
a solution of 20\% piperidine in DMF (10ml). The solution was then sonicated for 20
min. The UV absorbance of the solution was then measured at $\lambda=302$ nm and the
concentration of the piperidine-fulvene adduct calculated using the Beer-Lambert
Law, with the extinction coefficient of the adduct as 15,400dm\textsuperscript{3}mol\textsuperscript{-1}cm\textsuperscript{-1}. The resin
substitution could then be calculated from these values.
3.2.3 Automated Solid Phase Peptide Synthesis

All syntheses of peptides described used the following method of chain assembly. All the peptides were synthesised on an Applied Biosynthesis 430A automated peptide synthesiser fitted with an Applied Biosystems 757 UV monitoring system linked to a Hewlett Packard HP3396A integrator, allowing real time monitoring of coupling efficiencies. The syntheses were carried out on either 0.25 or 0.5 mmol scale.

Each amino acid cycle involved three basic steps, carried out in the following order.

1) capping any unreacted amino groups.
2) deprotection of the base labile α-amino terminus.
3) coupling of the next protected amino acid.

Each step of the cycle was followed by washing, and the cycle repeated until the required peptide sequence was obtained.

1) Capping
The resin was vortexed for 10 min. with a solution (10ml) containing acetic anhydride (0.5mol), DIEA (0.125mol) and HOBt (0.2%w/v) in a 1:1 solution of DMF/dioxane. The capping solution was drained and the resin washed in a 1:1 solution of DMF/dioxane (5 x 8ml).

2) Deprotection
Each deprotection cycle consisted of three deprotections, the resin vortexed with a solution of 20% piperidine in a 1:1 solution of DMF/dioxane (10ml) for 6, 1.5 and 1.5 min. Part of the deprotection solution was run to a UV monitor to allow real time assessment of coupling. Between these deprotections, the resin was washed with a 1:1 solution of DMF/dioxane (4 x 8ml). After the three deprotections, the resin was washed with a 1:1 solution of DMF/dioxane (8 x 8ml).
3) **Coupling**

This was carried out using double coupling with the HOBt activated ester, or single or double coupling using the triazole activated ester as described below.

**a) Double coupling with HOBt activated ester**

The initial coupling step utilised the symmetrical anhydride of the amino acid. The resin (0.25mmol) was vortexed with a solution of Fmoc amino acid pre-formed symmetrical anhydride (0.5mmol, formed from 1mmol Fmoc amino acid and 0.5mmol DIC reacted for 30 min in the activator vessel) for 30 min. The solution was removed and the resin washed with a 1:1 solution of DMF/dioxane (7 x 8ml). The resin was vortexed for a further 30 min with the Fmoc amino acid HOBt activated ester (0.5mmol, formed from 0.5mmol Fmoc amino acid, 0.5mmol HOBt and 0.5mmol DIC), drained and washed with a 1:1 solution of DMF/dioxane (7 x 8ml).

The exceptions to the above were asparagine, glutamine and histidine, which were coupled twice using their HOBt ester (0.5mmol, formed as above). Glycine was coupled singly as its symmetrical anhydride (0.5mmol, formed as above).

**b) Single or double coupling with triazole activated ester**

**Single coupling** - The resin (0.25mmol) was vortexed for 30 min. with a solution of pre-formed Fmoc amino acid triazole active ester (1mmol, formed from 1mmol Fmoc amino acid, 1mmol triazole and 1 mmol DIC, reacted for 20min.) before being drained and washed with a 1:1 solution of DMF/dioxane (7 x 8ml).

**Double coupling** - The resin (0.25mmol) was vortexed for 30 min. with a solution of pre-formed Fmoc amino acid triazole active ester (1mmol, formed from 1mmol Fmoc amino acid, 1mmol triazole and 1mmol DIC, reacted for 20min.) before being drained and washed with a 1:1 solution of DMF/dioxane (7 x 8ml). The resin was then vortexed for a second period (30min.) with a further amount of Fmoc amino acid triazole ester (1mmol, formed as above). The solution was drained and the resin washed with a 1:1 solution of DMF/dioxane (7 x 8ml).
3.2.4 Ellman's assay for free thiol groups

Quantitative Ellman's assay\textsuperscript{272} enables the number of free thiol groups per molecule to be calculated. Accurately weighed peptide (0.02-0.05\textmu mol) was dissolved in denaturing buffer consisting of 0.1M sodium phosphate buffer pH 7.3 (2.5ml), containing EDTA (0.01M) and GdmCl (6M). A fresh solution of 5.5'-dithiobis-(2-nitrobenzoic acid), (DTNB, 3mM) in sodium phosphate buffer pH 7.3 was prepared. Denaturing buffer (2.5ml) was placed in the reference cell and denaturing buffer plus peptide placed in the sample cell of a UV spectrometer. DTNB solution (100\textmu l) was added to each cell and mixed by inversion and left to stand for 5 min. The absorbance at 412nm was then recorded. The concentration of free thiols can then be calculated using the Beer-Lambert Law, with the extinction coefficient of the nitrothiobenzoate (NTB) formed being 13700M\textsuperscript{-1}cm\textsuperscript{-1} at 412nm in 6M GdmCl.

3.2.5 Initial acid cleavage of resin bound peptide

The scavenger mix specified in the text was added to dry peptide resin (50mg) and the mixture stirred for 5 min under nitrogen. The stated amount of TFA was then added and the mixture stirred at room temperature under nitrogen. Samples of the deprotection solution (approx. 0.5ml) were removed periodically and concentrated \textit{in vacuo}. The resulting residues were triturated with diethyl ether and the peptides collected by filtration. The peptide was washed with diethyl ether, dissolved in 20% acetic acid / water and analysed by HPLC. The optimum cleavage conditions were identified and scaled up to 200-400mg batches of peptide resin.

3.2.6 Attempted TMS Bromide deprotection of GC1 Cys-Acm

Resin bound peptide (200mg) was stirred at 0°C under nitrogen with EDT (500\textmu l), \textit{m}-cresol (100\textmu l) and thioanisole (1.17ml) for 5 min. TFA (7.5ml) was then added and the whole stirred for a further 10 min. TMS bromide (1.32ml) was added and the whole stirred for 3 hr. The resin was removed by filtration and washed with TFA. The
TFA was then reduced in vacuo. The peptide was triturated with diethyl ether (100ml) and collected by filtration, dissolved in acetic acid in water (100ml) and lyophilised to yield 85 mg of crude white solid. Analytical HPLC (Aquapore C18, 220×4.6mm, 7μm, 100μl loop, A = H2O, B = CH3CN, 0.1% TFA; 1 ml/min. 10-90% B over 35 min. λ = 214) gave a collection of products, which were difficult to separate.

3.2.7 Tbfmoc labelling of GC1 derivatives

Tbfmoc chloride (50mg, 0.12mmol) was taken and dissolved in the minimum of dioxane (6ml). Resin bound peptide, with the terminal Fmoc protecting group removed (0.11mmol/g, 200mg, 0.022mmol) along with DIEA (3μl, 0.025mmol) was then added. The mixture was sonicated in the dark for 2.5 hr and the resin thoroughly washed with dioxane (50ml, with sonication) and the resin dried in vacuo. A small portion of the Tbfmoc labelled resin (2.7mg) was then suspended in 20% piperidine in DMF and the solution sonicated for 20 min. The UV absorbance of the solution was then measured at λ=364 nm (2.68 Abs) and the concentration of the piperidine adduct calculated using the Beer-Lambert law, knowing the extinction coefficient. The resin substitution could then be calculated from these values (0.12mmol/g). The percentage loading of the Tbfmoc group could then be calculated (approximately 40%).

Resin bound peptide (310mg) was stirred at room temperature under nitrogen with EDT (2ml) and water (1ml) for 5 min. TFA (10ml) was then added and the whole stirred for a further 4 hr. The resin was then removed by filtration and washed with TFA. The TFA was then reduced in vacuo. The peptide was then triturated with diethyl ether (100ml) and left to stand over ether for 1 hr before being collected by filtration, dissolved in acetic acid in water (100ml) and lyophilised to yield 136 mg of crude white solid.
3.3 Peptide GC1 (KSSGDPEIVTHSFNCVCSSNITGGLLTLTRDGG) : Cysteine protected by Phenacm group.

3.3.1 Synthesis of bis-Phenacm GC1

Fmoc-Gly resin (0.66mmol/g, 0.38g, 0.25mmol) was formed as above. Cysteine was protected on the side chain as the Phenacm derivative\(^{273}\), aspartic acid as the t-butyl ester derivative, arginine as the Pmc derivative, threonine as the t-butyl ether derivative, serine as the t-butyl ether derivative, glutamic acid as the t-butyl ester derivative, histidine as the trityl derivative and lysine as the Boc protected derivative.

Asparagine was coupled unprotected on the side chain. All other amino acids were unprotected on the side chain. The synthesis was carried out on a 0.25mmol scale. Each amino acid was double coupled using initially the Fmoc amino acid symmetrical anhydride (0.5mmol, formed from 1mmol Fmoc amino acid and 0.5mmol DIC in the activator vessel) followed by the Fmoc amino acid HOBt activated ester (0.5mmol, formed from 0.5mmol Fmoc amino acid, 0.5mmol HOBt and 0.5mmol DIC during the Na\(^{+}\) deprotection, HOBt (0.1M) was present in the 1:1 solution of DMF/dioxane containing 20% piperidine. Amino acid analysis of the completed resin (24 hr hydrolysis) gave values as below.

Asx 4  2.89, Thr 3  2.56, Ser 5  2.98, Glx 1  0.92, Pro 1  1.13, Gly 5  5.13, Cys 2  0.12, Val 2  1.67, Ile 1  1.70, Leu 3  2.82, Phe 1  0.87, His 1  0.90, Lys 1  0.76, Arg 1  1.00

3.3.2 Cleavage and deprotection of bis-Phenacm protected GC1 from the resin

Resin bound peptide (286mg) was stirred at room temperature under nitrogen with EDT (0.5ml) and water (0.5ml) for 5 min. TFA (9.5ml) was added and the whole stirred for a further 4 hr. The resin was removed by filtration and washed with TFA. The TFA was then reduced in vacuo. The peptide was trituated with diethyl ether (100ml). The peptide was collected by filtration, dissolved in acetic acid in water (100ml) and lyophilised to yield 135 mg of crude white solid. The recovered peptide was gel filtered (G50 Sephadeex, 30% acetic acid in water, 30 ml/hr) and the required
fractions combined and lyophilised to give 55mg of white solid. Analytical HPLC (Vydac C18, 250×4.6mm, 5µM, 2ml loop, A = H2O, B = CH3CN, 0.1% TFA; 1 ml/min. 10-90% B over 35 min. λ = 214) gave a major peak, retention time 21.9 min., 60%B. The Phenac protected peptide was then isolated using preparative HPLC (Vydac C18, 250×22mm, 10µm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 9ml/min., 10-60%B in 30 min.

14mg of the title compound was isolated. Analytical HPLC (Vydac C18, 250×4.6mm, 5µm, 100µl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity; retention time 17 min., 49%B; m/z 3560.67859 (FAB), C134H242N42O51S2 requires 3560.71092; deviation = -9.90 ppm. Amino acid analysis (26 hr hydrolysis) gave values as below.

Asx 3.85, Thr3 2.63, Ser3 4.03, G1x1 1.25, Pro1 1.54, Gly5 5.03, Cys2 1.58, Val2 2.17, Ile2 1.59, Leu3 3.17, Phe1 1.03, His1 0.97, Lys1 0.98, Arg1 0.99.

3.3.3 Attempted enzyme removal of Phenacm group

GC1 Cys-Phenacm protected, from above (2.2mg, 0.6µmol) was taken and dissolved in ammonium acetate buffer (pH 8.5ml). A small amount of resin bound penicillin acylase (30mg) was added to the solution and the whole shaken at 37°C. Small portions of the solution were removed periodically and analysed by analytical HPLC (Vydac C18, 250×4.6mm, 5µm, 100µl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214). The desired removal of the Phenac group was not seen even after 14 days.

3.3.4 Attempted mercuric acetate removal of Phenacm group from bis-Phenacm protected GC1

GC1-Phenacm protected, from above (5mg, 1.4µmol) was dissolved in 30% acetic acid in water (2ml) and treated with 50 equivalents of mercury acetate (2.24mg, 7mmol) and the pH adjusted to 4. The whole was stirred in the dark for 24 hr. The solution was lyophilised and the resultant solid dissolved in GdnCl (6M, 10ml). 200
equivalents of dithiothreitol (DTT, 43mg, 0.28mmol) and the whole stirred for 24 hr in the dark. Small portions of the solution were removed periodically and analysed by analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 100μl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214). The white suspension of mercury / DTT salt was removed by filtration and the resultant solution lyophilised. After addition of a few drops of water to the resultant white suspension, the solution was gel filtered (G50 Sephadex, 30% acetic acid in water, 30ml/hr) and the desired fractions lyophilised to yield 2.2mg of white solid. Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 1 ml/min. 10-90% B over 35 min. λ = 214) gave a major peak at retention time 28.7 min, %B=75.5. The peak gave a mass consistent with mercury salt of GC1; m/z (laser desorption) 3464, C136H224N40O49S2Hg requires 3466.

3.3.5 Attempted silver trifluoromethanesulfonate removal of Phenacm group : formation of GC1 reduced

Phenacm protected GC1 above (10mg, 0.003mmol) was stirred at 4°C with TFA (5ml) and twenty equivalents of silver triflate (25mg, 0.11mmol) for 2 hr in the dark. The TFA was removed in vacuo and the silver salt of the peptide triturated with diethyl ether. The salt was dissolved in 50% acetic acid in water (4ml) and treated in the dark overnight at room temperature with 100 equivalents of DTT (56mg, 0.36mmol). The solution was diluted with water (6ml) and centrifuged (5000 rpm, 30 min). After centrifugation, the supernatant was decanted from the yellow pellet. The pellet was then washed with 20% acetic acid in water (5ml) and the supernatant collected after further centrifugation. The supernatants were combined. Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 100μl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) gave a major peak, retention time 17.5 min., 50%B; The reduced peptide was then isolated using preparative HPLC (Vydac C18, 250×22mm, 10μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 9 ml/min. Linear gradient 10%B for 10 min, 10-60%B in 30 min. The desired fractions
lyophilised to yield 2.4mg of the *title compound* as a white solid, found to be the reduced form of GC1, m/z (FAB) 3265 C_{136}H_{224}N_{40}O_{49}S_{2} requires 3265.
3.4 Peptide GC1 (KSSGDPVEIVTHSNCVCSSNITGLLLTRDGG) : Optimised synthesis

3.4.1 Synthesis of bis-Acm GC1

Fmoc-Gly resin (0.41mmol/g, 1.2g, 0.5mmol) was formed as above. Cysteine was protected on the side chain as the Acm derivative, aspartic acid as the t-butyl ester derivative, arginine as the Pmc derivative, threonine as the t-butyl ether derivative, asparagine as the trityl derivative, serine as the t-butyl ether derivative, glutamic acid as the t-butyl ester derivative, histidine as the trityl protected derivative and lysine as the Boc protected derivative. All other amino acids were unprotected on the side chain. The synthesis was carried out on a 0.5mmol scale. Each amino acid was double coupled using the Fmoc amino acid triazole activated ester (1mmol per coupling, formed from 1mmol Fmoc amino acid, 1mmol triazole and 1mmol DIC) as described above, except for histidine which was single coupled twice as the HOBt activated ester. During the Na deprotection, HOBt (0.1M) was present in the 1:1 solution of DMF/dioxane containing 20% piperidine. Amino acid analysis of the completed resin (24 hr hydrolysis) gave values as below.

Asx 4 3.07, Thr 3 2.40, Ser 3 3.30, Glx 1 0.78, Pro 1 1.15, Gly 5 4.89, Cys 2 0.23, Val 2 1.19, Ile 2 1.40, Leu 3 2.73, Phe 1 0.77, His 1 0.78, Lys 1 0.87, Arg 1 0.94

3.4.2 Cleavage and deprotection of bis Acm protected GC1 from the resin

Resin bound peptide (310mg) was stirred at room temperature under nitrogen with EDT (2ml) and water (1ml) for 5 min. TFA (10ml) was added and the whole stirred for a further 4 hr. The resin was removed by filtration and washed with TFA. The TFA was then reduced in vacuo. The peptide was triturated with diethyl ether (100ml) and left to stand over ether for 1 hr. The peptide was collected by filtration, dissolved in acetic acid in water (100ml) and lyophilised to yield 136mg of crude white solid. Analytical HPLC (Vydac C18, 250×4.6mm, 5μM, 100μl loop, A = H2O, B = CH3CN, 0.1% TFA; 1 ml/min. 10-90% B over 35 min. λ = 214) gave a major
peak, retention time 14.0 min., 42%B; m/z (laser desorption) 3408.3, \( \text{C}_{142}\text{H}_{233}\text{N}_{42}\text{O}_{51}\text{S}_{2} \) requires 3408.9.

### 3.4.3 Removal of Acm groups from crude bis-Acm protected peptide GC1: formation of GC1 reduced

Crude cleaved Acm protected GC1 above (50mg, approximately 0.01mmol) was stirred at 4°C with TFA (5ml) and silver triflate (75mg, 0.3mmol) for 2 hr in the dark. The TFA was removed \textit{in vacuo} and the silver salt of the peptide triturated with diethyl ether. The salt was dissolved in 50% acetic acid in water (10ml) and treated in the dark overnight at room temperature with DTT (230mg, 1.45mmol). The solution was diluted with water (10ml) and centrifuged (5000rpm, 30min). After centrifugation, the supernatant was decanted from the yellow pellet. The pellet was then washed with 20% acetic acid in water and the supernatant collected after further centrifugation. The supernatants were combined. Analytical HPLC (Vydac C18, 250\times4.6mm, 5µm, 100µl loop, A = H\textsubscript{2}O, B = CH\textsubscript{3}CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. \( \lambda = 214 \)) gave a major peak, retention time 14.7 min., 43.5%B; The reduced peptide was then isolated using preparative HPLC (Vydac C18, 250\times22mm, 10µm, 5ml loop, A = H\textsubscript{2}O, B = CH\textsubscript{3}CN, 0.1% TFA; 9ml/min. Linear gradient 10%B for 10 min, 10-37%B in 27 min., 37-42%B in 10 min. Three peaks were isolated and the appropriate fractions were lyophilised.

**Peak 1: GC1 reduced**

17mg of the \textit{title compound} isolated. Analytical HPLC (Vydac C18, 250\times4.6mm, 5µm, 100µl loop, A = H\textsubscript{2}O, B = CH\textsubscript{3}CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. \( \lambda = 214 \)) showed >95% purity; retention time 14.4 min., 43%B; m/z 3266.55776 (FAB), \( \text{C}_{136}\text{H}_{224}\text{N}_{40}\text{O}_{49}\text{S}_{2} \) requires 3266.57409; deviation = -5.00 ppm. Quantitative Ellman's assay gave a value of 1.58 (2.00) sulfhydryl groups per molecule. Amino acid analysis (26 hr hydrolysis) gave values as below.

Asx 4.14, Thr 3.06, Ser 3.70, Glx 1.03, Pro 1.30, Gly 5.00, Cys 1.13, Val 1.70, Ile 1.60, Leu 3.09, Phe 0.98, His 1.05, Lys 1.28, Arg 1.17.
Peak 2: GC1. reduced, aspartimide rearranged

4.1mg of the *title compound* isolated. Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 100μl loop, A = H₂O, B = CH₃CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) retention time 15.3 min., 45%B; m/z 3248.54652 (FAB), C₁₃₆H₂₂₂N₄₀O₄₈S₂ requires 3248.56352; deviation = 5.24 ppm. Quantitative Ellman's assay gave a value of 0.73 (2.00) sulfhydryl groups per molecule. Amino acid analysis (26 hr hydrolysis) gave values as below.

Asx₄ 3.52, Thr₃ 2.67, Ser₅ 4.27, Glx₁ 1.06, Pro₁ 1.00, Gly₅ 5.09, Cys₂ 0.15, Val₂ 1.65, Ile₂ 1.62, Leu₃ 3.07, Phe₁ 0.91, His₁ 0.98, Lys₁ 1.23, Arg₁ 1.11.

Peak 3: GC1. reduced, piperidide rearranged

1.0mg of the *title compound* isolated. Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 100μl loop, A = H₂O, B = CH₃CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) retention time 15.8 min., 46%B; m/z 3333.66800 (FAB), C₁₄₁H₂₃₃N₄₁O₄₈S₂ requires 3333.65267; deviation = 4.60 ppm. Quantitative Ellman's assay gave a value of 0.98 (2.00) sulfhydryl groups per molecule.

3.4.4 Oxidation of GC1. reduced: Formation of GC1

Pure reduced GC1 (10.7mg, 3.3μmol) was taken and stirred at room temperature with 2.5% DMSO in TFA (10ml) for 45 min. The TFA was reduced *in vacuo* and the peptide triturated with diethyl ether. The peptide was then collected by filtration and dissolved in acetic acid / water. Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 100μl loop, A = H₂O, B = CH₃CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity, retention time 14.4 min., 43%B. Analytical HPLC (Hichrom C18, 250×4.6mm, 5μm, 100μl loop, A = H₂O, B = CH₃CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed two peaks present.

The oxidised peptide was purified using preparative HPLC (Aquapore C18, 250×9.2mm, 10μm, 5ml loop, A = H₂O, B = CH₃CN, 0.1% TFA; 5ml/min, 10-90%B in 30 min). The peak was isolated and the appropriate fraction lyophilised to yield 4.6
mg of the title compound. Analytical HPLC (Hichrom C18, 250×4.6mm, 5μm, 100μl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity; retention time 14.9 min., 44%B. m/z 3263.54122 (FAB), C136H222N4049O49S2 requires 3263.55508; deviation = 4.25 ppm. Quantitative Ellman's assay gave a value of 0.00 (0.00) sulfhydryl groups per molecule. Amino acid analysis (25 hr hydrolysis) gave values as below.

Asx4 4.18, Thr3 3.18, Ser5 4.67, Glx1 0.98, Pro1 0.92, Gly5 5.22, Cys2 0.67, Val2 1.79, Ile2 1.66, Leu3 3.29, Phe1 0.98, His1 1.11, Lys1 1.38, Arg1 1.12.
3.5 Peptide GC1 mono-Ala (KSSGGDPEIVTHSFCVCSNITGLLDRAGG)

3.5.1 Synthesis of bis-Acm GC1 mono-Ala

Fmoc-Gly resin (0.566mmol/g, 440mg, 0.25mmol) was formed as above. Cysteine was protected on the side chain as the Acm derivative, aspartic acid as the t-butyl ester derivative, arginine as the Pmc derivative, threonine as the t-buty1 ether derivative, asparagine as the trityl derivative, serine as the t-butyl ether derivative, glutamic acid as the t-butyl ester derivative, histidine as the trityl protected derivative and lysine as the Boc protected derivative. All other amino acids were unprotected on the side chain. The synthesis was carried out on a 0.25mmol scale. Each amino acid was single coupled using the Fmoc amino acid triazole activated ester (1mmol per coupling, formed from 1mmol Fmoc amino acid, 1mmol triazole and 1mmol DIC) as described above, except for histidine which was single coupled as the HOBt activated ester. Amino acid analysis of the completed resin (26 hr hydrolysis) gave values as below.

Asx₃ 2.61, Thr₃ 3.14, Ser₅ 3.59, Glx₁ 0.93, Pro₁ 1.29, Gly₅ 4.63, Ala₁ 1.37, Cys₂ 0.47, Val₂ 1.65, Ile₂ 1.68, Leu₃ 3.15, Phe₁ 0.88, His₁ 1.00, Lys₁ 0.74, Arg₁ 1.33

3.5.2 Cleavage and deprotection of bis Acm protected GC1 mono-Ala from the resin

Resin bound peptide (65mg) was stirred at room temperature under nitrogen with EDT (1ml) and water (1ml) for 5 min. TFA (10ml) was added and the whole stirred for a further 4 hr. The resin was removed by filtration and washed with TFA. The TFA was then reduced in vacuo. The peptide was triturated with diethyl ether (100ml). The peptide was collected by filtration, dissolved in acetic acid in water (100ml) and lyophilised to yield 40 mg of crude white solid. The Acm protected peptide was isolated using preparative HPLC (Vydac C18, 250×22mm, 10μm, 5ml loop, A = H₂O, B = CH₃CN, 0.1% TFA; 9ml/min., linear gradient 10%B for 10 min, 10-55%B in 30 min).
3.5 Peptide GC1 mono-Ala (KSSGGDPEIVTHSFNCVCCSNITGLLLTRAGG)

3.5.1 Synthesis of bis-Acm, mono-Ala GC1

Fmoc-gly resin (0.566mmol/g, 440mg, 0.25mmol) was formed as above. Cysteine was protected on the side chain as the Acm derivative, aspartic acid as the t-butyl ester derivative, arginine as the PMC derivative, threonine as the t-butyl ether derivative, asparagine as the trityl derivative, serine as the t-butyl ether derivative, glutamic acid as the t-butyl ester derivative, histidine as the trityl protected derivative and lysine as the Boc protected derivative. All other amino acids were unprotected on the side chain. The synthesis was carried out on a 0.25mmol scale. Each amino acid was single coupled using the Fmoc amino acid triazole activated ester (1mmol per coupling, formed from 1mmol Fmoc amino acid, 1mmol triazole and 1mmol DIC) as described above, except for histidine which was single coupled as the HOBt activated ester. Amino acid analysis of the completed resin (26 hr hydrolysis) gave values as below.

Asx 3  2.61, Thr 3  3.14, Ser 5  3.59, Glx 1  0.93, Pro 1  1.29, Gly 5  4.63, Ala 1  1.37, Cys 2  0.47, Val 2  1.65, Ile 2  1.68, Leu 3  1.15, Phe 1  0.88, His 1  1.00, Lys 1  0.74, Arg 1  1.33

3.5.2 Cleavage and deprotection of bis Acm protected mono-Ala GC1 from the resin

Resin bound peptide (65mg) was stirred at room temperature under nitrogen with EDT (1ml) and water (1ml) for 5 min. TFA (10ml) was added and the whole stirred for a further 4 hr. The resin was removed by filtration and washed with TFA. The TFA was then reduced in vacuo. The peptide was triturated with diethyl ether (100ml). The peptide was collected by filtration, dissolved in acetic acid in water (100ml) and lyophilised to yield 40 mg of crude white solid. The Acm protected peptide was isolated using preparative HPLC (Vydac C18, 250×22mm, 10μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 9ml/min., linear gradient 10%B for 10 min, 10-55%B in 30 min.
4.6 mg of the title compound was isolated. Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity; retention time 20.8 min., 57.5%B; m/z 3364.67435 (FAB), C141H234N42O49S2 requires 3364.65849; deviation = 4.77 ppm. Amino acid analysis (24 hr hydrolysis) gave values as below.

Asx3 2.89, Thr3 3.08, Ser5 4.64, Glx1 1.00, Pro1 0.98, Gly5 4.84, Ala1 1.13, Cys2 1.21, Val2 1.65, Ile2 1.55, Leu3 2.98, Phe1 1.00, His1 1.00, Lys1 0.96, Arg1 1.02

3.5.3 Removal of Acm groups from bis-Acm protected peptide GC1 mono-Ala: formation of GC1 mono-Ala, reduced

Bis Acm protected GC1 mono-Ala above (15mg, 0.004mmol) was stirred at 4°C with TFA (5ml) and twenty equivalents of silver triflate (22.7mg, 0.10mmol) for 2 hr. in the dark. The TFA was removed in vacuo and the silver salt of the peptide triturated with diethyl ether. The salt was dissolved in 50% acetic acid in water (4ml) and treated in the dark overnight at room temperature with 100 equivalents of DTT (69mg, 0.44mmol). The solution was diluted with water (6ml) and centrifuged (5000rpm, 30min). After centrifugation, the supernatant was decanted from the yellow pellet. The pellet was then washed with 20% acetic acid in water (5ml) and the supernatant collected after further centrifugation. The supernatants were combined. The reduced peptide was then isolated using preparative HPLC (Vydac C18, 250×22mm, 10μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 9ml/min. Linear gradient 10%B for 10 min, 10-55%B in 45 min). The desired fractions were lyophilised to yield 3.9mg of the title compound as a white solid. Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity; retention time 22.3 min., 61%B; m/z (FAB) 3221.58694 C135H223N40O47S2 requires 3221.57643 deviation 3.26 ppm. Amino acid analysis (25 hr hydrolysis) gave values as below.

Asx3 2.98, Thr3 3.03, Ser5 4.60, Glx1 1.00, Pro1 0.88, Gly5 4.90, Ala1 1.10, Cys2 1.65, Val2 1.78, Ile2 1.61, Leu3 2.91, Phe1 0.98, His1 1.00, Lys1 0.94, Arg1 1.03
3.5.4 Oxidation of GC1 mono-Ala. reduced: Formation of GC1 mono-Ala

Pure reduced GC1 mono-Ala (5mg, 1.6μmol) was taken and stirred at room temperature with 2.5% DMSO in TFA (5ml) for 40 min. The TFA was reduced in vacuo and the peptide triturated with diethyl ether. The peptide was collected by filtration, dissolved in acetic acid / water and lyophilised. The peptide was then re-lyophilised from acetonitrile / water to yield 4.6mg of the title compound as a white solid. Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity, retention time 20.6 min., 57%B. m/z 3220.56756 (FAB), C135H222N40O47S2 requires 3220.56861; deviation = 0.32 ppm. Quantitative Ellman's assay gave a value of 0.00 (0.00) sulfhydryl groups per molecule. Amino acid analysis (26 hr hydrolysis) gave values as below.

Asx3 2.99, Thr3 2.72, Ser5 4.11, Glx1 0.96, Pro1 1.05, Gly5 4.82, Ala1 1.19, Cys2 1.01, Val2 1.77, Ile2 1.62, Leu3 2.87, Phe1 1.00, His1 0.92, Lys1 0.90, Arg1 1.27
3.6 Peptide GC1 bis-Ala (KSSGGDPAlVTHSFNCVCSSNITGLLLTRAGG)

3.6.1 Synthesis of bis-Acm GC1 bis-Ala

Fmoc-Gly resin (0.72mmol/g, 350mg, 0.25mmol) was formed as above. Cysteine was protected on the side chain as the Acm derivative, aspartic acid as the t-butyl ester derivative, arginine as the Pmc derivative, threonine as the t-butyl ether derivative, asparagine as the trityl derivative, serine as the t-butyl ether derivative, histidine as the trityl protected derivative and lysine as the Boc protected derivative. All other amino acids were unprotected on the side chain. The synthesis was carried out on a 0.25mmol scale. Each amino acid was single coupled using the Fmoc amino acid triazole activated ester (1mmol per coupling, formed from 1mmol Fmoc amino acid, 1mmol triazole and 1mmol DIC) as described above, except for histidine which was single coupled as the HOBt activated ester. Amino acid analysis of the completed resin (27 hr hydrolysis) gave values as below.

Asx 3  2.3  3, Thr 3  2.72, Ser 5  3.64, Pro 1.08, Gly 4.90, Ala 2.10, Cys 0.69, Val 1.41, Ile 1.44, Leu 2.87, Phe 0.85, His 0.89, Lys 0.49, Arg 1.06

3.6.2 Cleavage and deprotection of bis Acm protected GC1 bis-Ala from the resin

Resin bound peptide (330mg) was stirred at room temperature under nitrogen with EDT (2ml) and water (1ml) for 5 min. TFA (10ml) was added and the whole stirred for a further 4 hr. The resin was removed by filtration and washed with TFA. The TFA was then reduced in vacuo. The peptide was triturated with diethyl ether (100ml). The peptide was collected by filtration, dissolved in acetic acid in water (100ml) and lyophilised to yield 179mg of crude white solid. Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 100μl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214); major peak retention time 17.70 min., 50.5%B.
3.6.3 Removal of Acm groups from crude bis-Acm protected peptide GC1 bis-Ala: formation of GC1 bis-Ala. reduced

Crude bis Acm protected GC1 bis-Ala above (40mg, approx. 0.01mmol) was stirred at 4°C with TFA (5ml) and twenty equivalents of silver triflate (60mg, 0.28mmol) for 2 hr in the dark. The TFA was removed in vacuo and the silver salt of the peptide tritutated with diethyl ether. The salt was dissolved in 50% acetic acid in water (4ml) and treated in the dark overnight at room temperature with 100 equivalents of DTT (180mg, 1.2mmol). The solution was diluted with water (10ml) and centrifuged (5000rpm, 30min). After centrifugation, the supernatant was decanted from the yellow pellet. The pellet was then washed with 20% acetic acid in water (5ml) and the supernatant collected after further centrifugation. The supernatants were combined. The reduced peptide was then isolated using preparative HPLC (Vydac C18, 250×22mm, 10μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 9ml/min. Linear gradient 10%B for 10 min, 10-55%B in 45 min). The desired fractions were lyophilised to yield 3.5mg of the title compound as a white solid. Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity; retention time 21.7 min., 59.5%B; m/z (FAB) 3164.58188 C133H222N40O45S2 requires 3164.57878 deviation 0.98ppm. Quantitative Ellman's assay gave a value of 1.34 (2.00) sulfhydryl groups per molecule. Amino acid analysis (24 hr hydrolysis) gave values as below.

Asx 3 2.63, Thr 3 2.94, Ser 5 4.50, Pro 1 0.86, Gly 4.95, Ala 2 2.05, Cys 2 0.00, Val 2 1.73, Ile 2 1.56, Leu 3 3.05, Phe 1 0.97, His 1 0.94, Lys 1 0.85, Arg 1 1.15

3.6.4 Oxidation of GC1 bis-Ala. reduced: Formation of GC1 bis-Ala

Pure reduced GC1 bis-Ala (6mg, 1.6μmol) was taken and stirred at room temperature with 2.5% DMSO in TFA (6ml) for 40 min. The TFA was reduced in vacuo and the peptide tritutated with diethyl ether. The peptide was collected by filtration, dissolved in acetic acid / water and lyophilised. The peptide was then re-lyophilised from acetonitrile / water to yield 5.78mg of the title compound as a white solid. Analytical
HPLC (Vydac C18, 250×4.6mm, 5μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity, retention time 21.2 min., 58.5%B. m/z 3163.56217 (FAB), C133H221Na40O45S2 requires 3163.57095; deviation = 2.78 ppm. Quantitative Ellman's assay gave a value of 0.00 (0.00) sulfhydryl groups per molecule. Amino acid analysis (26 hr hydrolysis) gave values as below.

Asx 2.73, Thr 2.74, Ser 4.13, Pro 0.90, Gly 4.85, Ala 2.15, Cys 0.57, Val 1.64, Ile 1.58, Leu 2.99, Phe 0.89, His 0.91, Lys 0.79, Arg 1.01
3.7 Peptide GC1.cyclopentyl, incorporating cis-3-aminocyclopentane-1-carboxylic acid (acpc) (KSSGDPEIVTHSFN(acpc)SSNITGLLLLTRDGG)

3.7.1 Synthesis of GC1.cyclopentyl

Fmoc-Gly resin (0.54mmol/g, 460mg, 0.25mmol) was formed as above. Aspartic acid was protected on the side chain as the t-butyl ester derivative, arginine as the Pmc derivative, threonine as the t-butyl ether derivative, asparagine as the trityl derivative, serine as the t-butyl ether derivative, glutamic acid as the t-butyl ester derivative, histidine as the trityl protected derivative and lysine as the Boc protected derivative. All other amino acids were unprotected on the side chain. The synthesis was carried out on a 0.25mmol scale. Each amino acid, apart from the central cyclopentyl moiety was single coupled using automation, utilising the Fmoc amino acid triazole activated ester (1mmol per coupling, formed from 1mmol Fmoc amino acid, 1mmol triazole and 1mmol DIC) as described above, except for histidine which was single coupled as the HOBt activated ester. The peptide was completed up to the cyclopentyl moiety in the usual way. The terminal Fmoc was removed and cyclopentyl moiety manually coupled as the Fmoc triazole activated ester derivative (solution of 1mmol, formed from 1mmol cyclopentyl derivative, 1mmol triazole and 1mmol DIC). The solution was sonicated with the resin for 1 hr. The resin was then washed thoroughly and the synthesis completed automatically. Amino acid analysis of the completed resin (26 hr hydrolysis) gave values as below.

Asx 4  3.40, Thr 2.67, Ser 3.82, Pro 1  0.91, Gly 5  4.92, Val 1  1.22, Ile 2  1.70, Leu 3  3.21, Phe 1  1.07, His 0.98, Lys 0.97, Arg 1.11

3.7.2 Cleavage and deprotection of GC1. cyclopentyl peptide from the resin

Resin bound peptide (300mg) was stirred at room temperature under nitrogen with EDT (2ml) and water (1ml) for 5 min. TFA (10ml) was added and the whole stirred for a further 4 hr. The resin was removed by filtration and washed with TFA. The TFA was then reduced in vacuo. The peptide was triturated with diethyl ether...
(100ml), collected by filtration, dissolved in acetic acid in water (100ml) and lyophilised to yield 140mg of crude white solid. 100mg of the above crude cleaved peptide was then taken and the peptide isolated using preparative HPLC (Aquapore C18, 250×9.2mm, 10μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 5ml/min, 10-55%B in 30 min). The desired fractions were lyophilised to yield 35mg of the title compound as a white solid. Analytical HPLC (Hichrom C18, 250×4.6mm, 5μm, 100μl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity; retention time 14 min., 42%B; m/z (FAB) 3071.56578 C131H213N38O47S2 requires 3071.54789 deviation 5.82ppm. Amino acid analysis (24 hr hydrolysis) gave values as below.

Asx 3.92, Thr 2.78, Ser 4.23, Pro 1.04, Gly 5.12, Val 0.78, Ile 1.67, Leu 2.97, Phe 0.88, His 0.92, Lys 0.90, Arg 0.95
3.8 Peptide IH1 (KSSGDPEIVTHSFNCKCSSNITGLLLTRDGG)  
(KQFINMWQEVGG)

3.8.1 Synthesis of bis-Acm IH1

Fmoc-Gly resin (0.53mmol/g, 0.49g, 0.25mmol) was formed as above. Cysteine was protected on the side chain as the Acm derivative, aspartic acid as the t-butyl ester derivative, arginine as the Pmc derivative, threonine as the t-butyl ether derivative, asparagine as the trityl derivative, glutamine as the trityl derivative, tryptophan as the Boc derivative, serine as the t-butyl ether derivative, glutamic acid as the t-butyl ester derivative, histidine as the trityl protected derivative. Three orthogonal lysine derivatives were utilised, as described below. All other amino acids were unprotected on the side chain. The synthesis was carried out on a 0.25mmol scale. Each amino acid was double coupled using the Fmoc amino acid triazole activated ester (1mmol per coupling, formed from 1mmol Fmoc amino acid, 1mmol triazole and 1mmol DIC) as described above, except for histidine which was single coupled twice as the HOBt activated ester and the orthogonal lysine derivative described below. During the Na deprotection, HOBt (0.1M) was present in the solution of 20% piperidine in 1:1 DMF/dioxane.

KCSSNITGLLLTRDGG

The above peptide was synthesised using automated solid phase peptide synthesis in the usual way, described above. The underlined lysine above, located between the two cysteines was coupled as the N⁹(Dde)-N⁹(Fmoc) derivative. This was manually single coupled as the Fmoc triazole activated ester derivative (solution of 1mmol, formed from 1mmol lysine derivative, 1mmol triazole and 1mmol DIC). The solution was sonicated with the resin for 1 hr. The resin was then washed thoroughly with DMF and Dioxane and dried *in vacuo*. A small amount of resin (2.8mg) was taken and a load test was run as above, allowing the amount of coupling to be estimated (approximately 80%).
The peptide was continued after coupling of the lysine derivative, the chain being
grown from the N°- free amine using automated solid phase peptide synthesis in the
usual way to yield the above peptide. The N-terminal lysine was coupled as the
N°(Fmoc)-N°(Fmoc) derivative. Upon completion of coupling, the Fmoc groups were
removed using a solution of 1:1 DMF/dioxane containing 20% piperidine, with 0.1M
HOBt present in the usual way. The free amine groups were acylated using a solution
(10ml) containing acetic anhydride (0.5mol), DIEA (0.125mol) and HOBt (0.2%w/v)
in a solution of 1:1 DMF/dioxane. The resin was vortexed for 10 min then the capping
solution was drained and the resin washed with a solution of 1:1 DMF/dioxane (7 x
8ml).

The Dde group present on the underlined lysine was removed by sonication of the
resin with a solution of 2% hydrazine in DMF (10ml) for 20 min. The solution was
then removed and the resin washed with DMF(100ml) and dioxane (100ml). To
ensure the removal of the Dde group, a qualitative Kaiser test was used to indicate the
presence of free primary amines. A small sample of peptide resin was removed and
placed in a vial. Three solutions were added consecutively. Firstly, ninhydrin in
ethanol (5%w:v, 75μl) was added followed by phenol in ethanol (400%w:v, 100μl).
Finally a solution of aqueous KCN (1mM) in distilled pyridine (2%v:v, 75μl) was
added and the vial heated for 5 min at 110°C. A positive test was indicated by a blue
colouration. The peptide was then completed by coupling to the free α amine group
using automated solid phase peptide synthesis in the usual way to yield the above
peptide.

Amino acid analysis of the completed resin (32 hr hydrolysis) gave values as below.
Asx 4.72, Thr 3.52, Ser 5.89, Glx 4.93, Pro 0.87, Gly 7.33, Cys 0.76, Val 1.52,
Met 0.61, Ile 2.51, Leu 4.43, Phe 1.68, His 0.95, Lys 2.32, Arg 1.48
3.8.2 Cleavage and deprotection of peptide IH1 from the resin

Resin bound peptide (300mg) was stirred at room temperature under nitrogen with EDT (2ml), thioanisole (500μL) and water (1ml) for 5 min. TFA (10ml) was added and the whole stirred for a further 4 hr. The resin was removed by filtration and washed with TFA. The TFA was then reduced in vacuo. The peptide was trituated with diethyl ether (100ml) and left to stand over ether for 1 hr. The peptide was collected by filtration, dissolved in acetic acid in water (100ml) and lyophilised to yield 130mg of crude white solid. 20mg of the above crude peptide was taken and the Acm protected peptide isolated using preparative HPLC (Vydac C18, 250×22mm, 10μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 9ml/min., 10-30%B in 10 min, 30-50%B in 20 min.

Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 100μl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity; retention time 16 min., 46.5%B; m/z 4897.39078 (FAB), C210H333N60O49S3 requires 4897.36222; deviation = 5.83ppm.

3.8.3 Attempted silver trifluoromethanesulfonate removal of the Acm group

Crude Acm protected IH1 above (50mg, approximately 0.01mmol) was stirred at 4°C with TFA (5ml) and thirty equivalents of silver triflate (75mg, 0.33mmol) for 2 hr in the dark. The TFA was removed in vacuo and the silver salt of the peptide trituated with diethyl ether. The salt was dissolved in 50% acetic acid in water (4ml) and treated in the dark overnight at room temperature with 150 equivalents of DTT (230mg, 1.5 mmol). The solution was diluted with water (6ml) and centrifuged (5000rpm, 30min). After centrifugation, the supernatant was decanted from the yellow pellet. The pellet was then washed with 20% acetic acid in water (5ml) and the supernatant collected after further centrifugation. The supernatants were combined. The major peak was then isolated using preparative HPLC (Vydac C18, 250×22mm, 10μm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 9ml/min. Linear gradient 10%B for 10 min, 10-30%B in 10 min, 30-50%B in 20 min). The desired fractions were
lyophilised to yield 4.5mg of white solid. Analytical HPLC (Vydac C18, 250×4.6mm, 5μm, 5ml loop, A = H₂O, B = CH₃CN, 0.1% TFA; ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity; retention time 15.1 min., 44.5%B; m/z (FAB) 4769 C₂₀₄H₃₂₂N₅₈O₆₈S₃ (IH1, Acm removed, with methionine sulfoxide) requires 4769. Quantitative Ellman's assay gave a value of 1.76 (2.00) sulphydryl groups per molecule.

3.8.4 Mercuric acetate removal of Acm groups : formation of IH1, reduced

Crude Acm protected IH1 above (50mg, approximately 0.01mmol) was taken and dissolved in 50% acetic acid in water (2ml), this solution having been previously sonicated to remove any dissolved gases, and mercury acetate (12mg, 0.04mmol) added. The whole was stirred slowly under nitrogen at room temperature for 1 hr. DTT (100mg, 0.67mmol) was added and the whole stirred for a further 1.5 hr. The solution was diluted with water (10ml) and centrifuged (5000rpm, 30min). After centrifugation, the supernatant was decanted from the grey/white pellet. The pellet was then washed with 20% acetic acid in water (5ml) and the supernatant collected after further centrifugation. The supernatants were combined. The reduced peptide was isolated using preparative HPLC (Vydac C18, 250×22mm, 10μm, 5ml loop, A = H₂O, B = CH₃CN, 0.1% TFA; 9ml/min. Linear gradient 10-35%B in 20 min, 35-50%B in 20 min). The desired fractions were lyophilised to yield 6.9mg of the title compound as a white solid. Analytical HPLC (Hichrom C18, 250×4.6mm, 5μm, 100μl loop, A = H₂O, B = CH₃CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity; retention time 17.1 min., 49%B; m/z (FAB) 4756 C₂₀₄H₃₂₆N₅₈O₆₇S₃ requires 4756. Quantitative Ellman's assay gave a value of 1.60 (2.00) sulphydryl groups per molecule. Amino acid analysis (31 hr hydrolysis) gave values as below.

Asx 5.29, Thr₃ 3.20, Ser₃ 4.43, Glx₄ 4.10, Pro₁ 0.92, Gly₇ 7.25, Cys₂ 0.23, Val₂ 1.82, Met₁ 0.42, Ile₃ 2.65, Leu₃ 3.37, Phe₂ 1.75, His₁ 0.95, Lys₃ 2.68, Arg₁ 0.90
3.8.5 Attempted oxidation of IH1. reduced using DMSO / TFA

Pure reduced IH1 (3.5mg, 0.7µmol) was taken and stirred at room temperature with 2.5% DMSO in TFA (3.5ml) for 50 min. The TFA was reduced in vacuo and the peptide triturated with diethyl ether. The peptide was then collected by filtration and dissolved in acetic acid / water. Analytical HPLC (Vydac C18, 250×4.6mm, 5µm, 100µl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) gave a major peak at retention time 14.4 min., 43%B, corresponding to a methionine sulfoxide peptide. Quantitative Ellman's assay gave a value of 0.00 (2.00) sulfhydryl groups per molecule. This oxidation method also oxidised methionine to methionine sulfoxide as well as cystine oxidation.

3.8.6 Attempted N-methyl-mercaptoacetamide (MMA) reduction of methionine sulfoxide present in IH1

IH1 methionine sulfoxide (0.5mg, 0.1µmol) was taken and dissolved in 10% acetic acid in water (500µl). MMA (0.11mg, 0.1µl, 1.1µmol) was added and the solution incubated under nitrogen at 37°C for 5 days. Small portions of the solution were removed periodically and analysed by analytical HPLC (Vydac C18, 250×4.6mm, 5µm, 100µl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214). It was seen that the reaction was not giving quantitative reduction of the methionine sulfoxide. This method was abandoned.

3.8.7 Air oxidation of IH1. reduced

Pure reduced IH1 (10mg, 2.2µmol) was taken and dissolved in 50% acetonitrile in water (150µl) and left to stand for 7 days. The solution was then lyophilised to yield a white peptide. This was dissolved in 20% acetonitrile in water (10ml) and the oxidised peptide isolated using preparative HPLC (Aquapore C18, 250×9.2mm, 10µm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 5ml/min, 10-50%B in 40 min). The desired
fractions were lyophilised to yield 5.6 mg of the title compound as a white solid. Analytical HPLC (Hichrom C18, 250×4.6mm, 5μm, 100μl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity; retention time 16.8 min., 48.5%B; m/z (FAB) 4756.30508 C204H324N58O67S3 requires 4756.29581 deviation 1.95ppm. Quantitative Ellman's assay gave a value of 0.00 (2.00) sulphydryl groups per molecule. Amino acid analysis (32 hr hydrolysis) gave values as below.

Asx 5.10, Thr 3.93, Ser 4.10, Glx 3.91, Pro 0.78, Gly 7.30, Cys 2.00, Val 1.70, Met 0.65, Ile 2.93, Leu 3.58, Phe 1.69, His 0.88, Lys 2.57, Arg 1.02
3.9 Peptide KQFINMWQEVG: peptide IH2

3.9.1 Synthesis of IH2

Fmoc-Gly resin (0.52mmol/g, 480mg, 0.25mmol) was formed as above. Asparagine was protected as the trityl derivative, glutamine as the trityl derivative, tryptophan as the Boc derivative, serine as the t-butyl ether derivative, glutamic acid as the t-butyl ester derivative, and lysine as the Boc protected derivative. All other amino acids were unprotected on the side chain. The synthesis was carried out on a 0.25mmol scale. Each amino acid was single coupled using the Fmoc amino acid triazole activated ester (1mmol per coupling, formed from 1mmol Fmoc amino acid, 1mmol triazole and 1mmol DIC) as described above. Amino acid analysis of the completed resin (24 hr hydrolysis) gave values as below.

Asx 1 0.89, Glx 3 3.76, Gly 1 1.08, Val 1 1.00, Met 1 0.54, Ile 1 0.89, Phe 1 0.92, Lys 1 0.90

3.9.2 Cleavage and deprotection of peptide IH2 from the resin

Resin bound peptide (100mg) was stirred at room temperature under nitrogen with EDT (2ml), thioanisole (500µl) and water (1ml) for 5 min. TFA (10ml) was added and the whole stirred for a further 3 hr. The resin was removed by filtration and washed with TFA. The TFA was then reduced in vacuo. The peptide was triturated with diethyl ether (100ml), collected by filtration, dissolved in acetic acid in water (100ml) and lyophilised to yield 61mg of crude white solid. 20mg of the above crude cleaved peptide was then taken and the peptide isolated using preparative HPLC (Aquapore C18, 250×9.2mm, 10µm, 5ml loop, A = H2O, B = CH3CN, 0.1% TFA; 5ml/min, 10-35%B in 20 min). The desired fractions were lyophilised to yield 5.5 mg of the title compound as a white solid. Analytical HPLC (Hichrom C18, 250×4.6mm, 5µm, 100µl loop, A = H2O, B = CH3CN, 0.1% TFA; 1ml/min. 10-90% B over 35 min. λ = 214) showed >95% purity; retention time 11.6 min., 36.5%B; m/z (FAB) 1380.67920 C63H96N16O17S1 requires 1380.68601 deviation -4.90ppm. Amino acid analysis (25 hr hydrolysis) gave values as below.
Asx 1.03, Glx 2.93, Gly 1.07, Val 0.95, Met 0.64, Ile 0.93, Phe 0.93, Lys 0.93
3.10 Immunisation with peptides

Groups of five female BALB/C mice, 10 weeks old were immunised intraperitoneally with the peptide (10 μg per injection in 100 μl solution). The first injection was administered in either Freund's adjuvant or alum, followed by four further injections at three to four week intervals of peptide in PBS.

3.11 Immunogenicity of peptides

96-well Costar EIA/RIA microtitre plates were coated either overnight at 4°C, or for 2-3 hr at room temperature with peptide (1 μg/ml, 0.1 μg per well) in 0.1 M carbonate/bicarbonate buffer pH 9.6. The plate was then washed three times with phosphate buffered saline (PBS) containing 0.05% v:v Tween 20. Blocking was achieved by incubating with a solution of 1% m:v bovine serum albumen (BSA) in PBS for 1 hr at room temperature. After three further washes as described above, serially diluted serum (100 μl per well, diluted with 1% m:v bovine serum albumen (BSA) in PBS containing 0.05% Tween 20) from immunised and controlled untreated animals was added to the wells and incubated for 2 hr at room temperature. The plate was again washed and 1:750 dilution of goat antimouse IgG conjugated horseradish peroxidase (100 μl per well, diluted with of 1% m:v bovine serum albumen (BSA) in PBS containing 0.05% Tween 20) added and incubated for 1 hr. Unbound conjugate was removed by washing and α-phenylenediamine (0.4 mg/ml in phosphate / citrate buffer pH 5.0 containing 0.006% H₂O₂, 100 μl per well) added and the plate incubated for 10 min at room temperature. The colouration reaction was stopped by addition of 2N HCl (50 μl), and the optical density (OD) values were measured at 490 nm using a Dynatech MR5000 microplate reader.

3.12 Binding of anti-IH1 and anti-GC1 IgG to peptides

96-well Costar EIA/RIA microtitre plates were coated overnight at 4°C, with serially diluted peptide (100 μl/well) in 0.1 M carbonate/bicarbonate buffer pH 9.6. The plate
was then washed three times with phosphate buffered saline (PBS) containing 0.05% v:v Tween 20. Blocking was achieved by incubating with a solution of 1% m:v bovine serum albumen (BSA) in PBS for 1 hr at room temperature. After three further washes as described above, either anti-IH1 IgG, anti-GC1 IgG or IgG from normal mouse serum (100 μl per well, diluted with 1% m:v bovine serum albumen (BSA) in PBS containing 0.05% Tween 20) was added to the wells and incubated for 2 hr at room temperature. The plate was again washed and 1:750 dilution of goat `antimouse IgG conjugated horseradish peroxidase (100 μl per well, diluted with of 1% m:v bovine serum albumen (BSA) in PBS containing 0.05% Tween 20) added and incubated for 1 hr. Unbound conjugate was removed by washing and o-phenylenediamine (0.4 mg/ml in phosphate / citrate buffer pH 5.0 containing 0.006% H₂O₂, 100 μl per well) added and the plate incubated for 10 min at room temperature. The colouration reaction was stopped by addition of 2N HCl (50 μl), and the optical density (OD) values were measured at 490 nm using a Dynatech MR5000 microplate reader.

3.13 Protein G purification of antisera

A column containing Protein G Sepharose 4 fast flow (2g) was washed with 0.1M glycine, pH 2.5 (30ml). The column was then re-equilibrated with 20mM phosphate buffer, pH 7 (30ml). The supernatant was then added to the column and the unbound protein removed by elution with 20mM phosphate buffer (50ml). The bound IgG was then removed from the column with 0.1M glycine (10ml), collecting 1ml fractions. Each fraction was dotted onto nitrocellulose and allowed to dry at 37°C. The nitrocellulose was stained using coomasie blue (0.25g in 100ml of a solution of distilled water containing methanol (10%) and acetic acid (7%)). The fractions containing the peptide, appearing on the nitrocellulose as a blue stained central spot, were pooled.
Chapter 4: Appendices

4.1 Flow cytometry

Flow cytometry was carried out on a Coulter EPICS XL Flow Cytometer with a 15mW, single argon ion laser operating at 488nm. The percentages of labelled cells was established relative to background fluorescence of cells treated with fluorescent secondary antibody only.

4.2 Binding of peptides to CD4+ve cells

H9 cells, spun over lymphoprep density gradients to remove non-viable cells, were suspended in PBS containing 0.5% sodium azide and added to round bottomed 96-well tissue culture plates at a concentration of $3 \times 10^5$ cells/well and incubated on ice for 2 hr with increasing amounts of peptide in PBS (50 μl). The cells were washed with PBS containing 0.5% sodium azide (50 μl) and incubated on ice for 1 hr with either anti-GC1 IgG or IgG from normal mouse serum (2 μg). The cells were washed and phycoerythrin conjugated goat anti-mouse IgG added for a further 1 hr on ice. The cells were washed and resuspended in PBS containing 0.5% sodium azide for reading via flow cytometry.

4.3 Peptide blocking of anti-CD4 mAb binding to CD4

Viable CD4+ MM6 cells were isolated by gradient centrifugation at 1000 rpm for 25 min., washed with PBS (3x10ml) and pelleted in a 96-well microtitre plate at a concentration of $10^5$ cells per well. The cells were then incubated on ice for 2 hr with IH1 (0.4mM in PBS), washed with PBS and incubated with anti-CD4 mAb (10μl, 100μg/ml) on ice for 1 hr. The cells were washed with pre-chilled flow buffer (1% BSA in PBS with 0.05% w:v sodium azide) and bound mAb detected by incubation for 1 hr on ice with phycoerythrin labelled goat anti-mouse immunoglobulin at 1:100 concentration and counting positive cells by flow cytometry.
4.4 Peptide blocking of MIP-1α binding to CC-CKR5⁺ cells

Binding of biotinylated recombinant MIP-1α to H9 cells was analysed by flow cytometry following the manufacturers protocol, with viable cells being washed in PBS (3x10 ml) and pelleted in 96-well microtitre plates at a concentration of 10⁵ cells per well. The cells were incubated with or without peptide in 1% BSA in PBS (10μl) for 2 hr on ice, treated with biotinylated MIP-1α and detected using FITC labelled avidin. Positive cells were then counted using flow cytometry.

4.5 Localisation of IF1 binding to CD4⁺ cells

Viable cells were washed in PBS (3x10 ml) and pelleted in 96-well microtitre plates at a concentration of 3x10⁵ cells per well. The cells were incubated with IH1 (1μg/well, 21 nmol) in binding buffer (1mg/ml GMEM, 10% FCS, 1mg/ml Hepes in water) on ice for 2 hr and the plate washed with pre-chilled binding buffer and incubated with anti GC1 polyclonal sera (1:100 concentration in binding buffer) on ice for 1 hr. The cells were washed and any bound anti-GC1 polyclonal sera detected by incubation on ice for 30 min. and at room temperature for 15 min. with TRITC labelled anti-mouse immunoglobulin (1:50 concentration in binding buffer). Cells were washed with pre-chilled flow buffer (1% BSA in PBS, with 0.05% w:v sodium azide) and biotinylated mouse anti human CD4 mAb (clone MT310) was added at a concentration of 1:10 in flow buffer and incubated on ice for 30 min. Bound immunoglobulin was detected using FITC-labelled avidin at a concentration of 1:100 in flow buffer. Cells were then fixed with 0.4% formaldehyde and transferred to slides with a single drop of glycerol/PBS.

4.6 In vitro induction of apoptosis

a) Solid phase antibody:– Sterile IgG (2μg/well, 100μl) was coated onto tissue culture grade, 96-well microtitre plates in 0.05M carbonate/bicarbonate coating buffer, pH 9.6 overnight at 4°C. Immediately before use the wells were washed with cold PBS.
b) **Cells:** Viable H9 cells passaged in RPMI containing 3% foetal calf serum (3% RPMI) and separated over lymphoprep density gradients were washed with cold PBS (2x100μl) and counted in 0.5% trypan blue solution. Cells were spun down and incubated on ice for 2 hr with PBS (50μl) or PBS containing either IH1 or FMDV (50μl, 5μg/10⁶ cells). The cells were then washed with cold 3% RPMI, counted in tryphan blue and resuspended at 10⁵/ml in 3% RPMI. Quadruplicate sets of 2x10⁴ cells/well were then plated into an IgG coated microtitre plate. The plate was then spun gently (100rpm for 3 min.) and incubated for 5 hr at 37°C in 5% CO₂ in a humidified incubator.

c) **Counting apoptotic cells:** After 5 hr incubation at 37°C, contents of the wells were made into cytopins, air dried for 10 min., fixed in acetone for 10 min. and stained with haematoxylin. The percentage of cells showing apoptotic morphology was counted by light microscopy.
Chapter 5: References


13 Muller, S., Himmelspach, K. and Van Regenmortel, M.H.V., EMBO, 1982, 1, 421.


96 Nara, P.L., Smit, L., Dunlop, N., Hatch, W., Merges, M., Waters, D., Kelliher, J.,

J.A., Langlois, A.J., Dreesman, G.R., Boswell, R.N., Shadduck, P., Holley, L.H.,
1990, 249, 932.


99 Clements, G.J. Price-Jones, M.J., Stephens, P.E., Sutton, C., Schulz, T.F.,
Clapham, P.R., McKeating, J.F., McClure, M.O., Thomson, S., Marsh, M., Kay, J.,


48-52.


104 Murakima, T., Hattori, T. and Takasaki, K., *Biochimica et Biophysica Acta*, 1991,
1097, 279.

105 Avril, L-E., De Martino-Ferrer, M., Pignede, G., Séman, M. and Gauthier, F.,


107 Cann, A.J., Churcher, M.J., Boyd, M., O'Brien, W., Zhao, J-Q., Zack, J. and Chen,

108 Shioda, T., Oka, S., Ida, S., Nokihara, K., Toriyoshi, H., Mori, S., Takebe, Y.,

261, 1179.


111 Ebenbichler, C., Westervelt, P., Carillo, A., Henkel, T., Johnson, D. and Ratner, L.,


196 Curtius, T., *ibid.*, 1902, 35, 3226.


222 Martinez, J. and Bodansky, M., *ibid.*, 1978, 12, 277.


254 Crivici, A. and Lajole, G., Synthetic Comm., 1993, 23, 49-.


267 Price, N. and Kelly, S., Biochemistry Department, University of Stirling.


271 Cotton, G.J, personal communication.


273 Obtained from Brown, A.

274 Obtained from Urquhart, K.
Courses attended


"Medicinal chemistry", Professor R. Baker and colleagues, Merck Sharp and Dohme, University of Edinburgh, 1994-1996.

NMR spectroscopy, Drs. I. Sadler and P. Barlow, University of Edinburgh, 1996.


