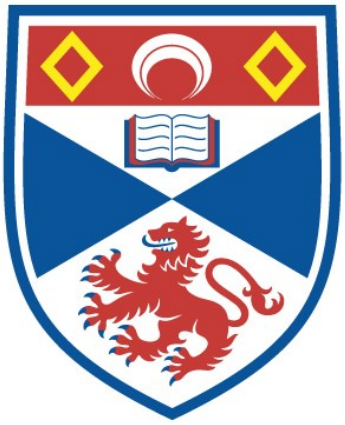
DESIGN AND SYNTHESIS OF INHIBITORS FOR THE HIV-1 PROTEASE

Nicholas Paul Camp

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



1994

DESIGN AND SYNTHESIS OF INHIBITORS FOR THE HIV-1 PROTEASE.

a thesis presented by

Nicholas Paul Camp

to the

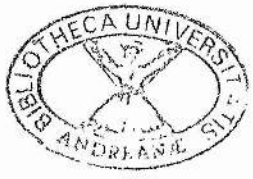
UNIVERSITY OF ST. ANDREWS

in application for

THE DEGREE OF DOCTOR OF PHILOSOPHY

St Andrews

January 1994



DECLARATION

I, Nicholas Paul Camp, hereby certify that this thesis has been composed by myself, that it is a record of my own work, and that it has not been accepted in partial or complete fulfiment of any other degree or professional qualification.

Signed .... Date 16.2.94

I was admitted to the Faculty of Science of the University of St. Andrews under Ordinance General No. 12 on 1st Oct 1990 and as a candidate for the degree of Ph.D. on 1st Oct 1990

Signed ..... Date 16.2.94

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate to the degree of Ph.D.

Signature of supervisor .....

Date .....

Copyright

DEDICATED TO

MY FAMILY.

"He who dares wins"

Del Boy Trotter.
(Only Fools and Horses 1980).

Acknowledgements.

Scientific.

I would like to thank my friend and supervisor, Professor David Gani, for the opportunity to study for a PhD and for providing such an excellent project. Thanks for giving me the freedom to develop this work and for helping me through the bad times. It was a pleasure to work with my HIV team mates and thank Paul Hawkins for his wit and intellect and David Perrey for his outstanding technical support in the early days. I wish the new team, namely Neil and Donald the best of luck and look forward to the publications. Thanks to Mahmoud (Akkars) for his wisdom and proof reading of this thesis (I know what it must have been like because I had to write it !) I am very grateful to the lovely Stacey for all of her help and efficiency and to Nigel (Dr. B.) for chemical discussions and for supporting Chelsea (somebody has to). Thanks also to the many project students who had to suffer my supervision and to the following for creating such a stimulating working environment, Johnnie (the one and only) Unikowski, Cam Brown, Julie, Neil, Merrence, Hawks, Donald, Morag and Lynda to name a few.

I would like to thank Ian Sadler for 600 MHz NMR spectra, Peter Hitchcock for the crystal structure, Swansea for their efficient mass spec service, Derek Kinchington at St. Bartholomew's Hospital for the *in vivo* assay results, BTG for funding some of this project and for providing a technician in the form of Mr Merry and finally the University of St. Andrews for funding.

Non-scientific.

I would like to raise my glass to the Southampton boys, Jimmy, Ted, Salvo, Andy and Richie Brown for the brilliant times and under whose supervision I obtained my fine degree result (cheers lads)!

Thanks to Arnie for the barbeques, impromptu parties and numerous beers. Thanks to all of the following extremely entertaining people, George (you know it makes sense) Cole, Graham, Johnnie (every day, in every way, I get a little better), Cammy boy, Merrence, Smokin 'MC', Jurgen and Mike (Fats) Bruce for the sessions in Kates bar. I will never forget the St. Andrews classics; the post graduate (it's a knockout) Xmas party, the Graduation Ball and food fight (special ticket holders only), the many wine receptions and the opening of the labs. All of these events starred my aforementioned colleagues and myself.

Back to reality, I thank the Nailsea boys for their guidance during my earlier academic days, especially to Dave Adams for his friendship and support during the financially difficult times. Thanks to the regulars of the Sawyers Arms for beers and encouragement and a special thanks to Uncle Marcus and Smithie for the building work during those summer holidays (If things don't shape up soon I might be coming to you for a labouring job!)

Finally, a special mention to my flatmate MC for putting up with me during the last few months and for his undisputed friendship during my stay in St. Andrews.

"Without humour there is nothing" !!

Abstract

A variety of phosphonamidate-containing peptides were synthesised as potential inhibitors of the HIV-1 protease. These transition state analogues were designed using known sequences from HIV-1 protease substrates and incorporated a unique Phe-Pro scissile bond mimic in an attempt to achieve selectivity over the mammalian aspartic proteases. Such compounds were found to be moderate inhibitors of the HIV-1 protease possessing IC_{50} values in the 1-100 μM range, both in *in vitro* and *in vivo* assays. However, the phosphonamidate methyl ester analogues showed a marked ability to enter cells and this feature was highlighted in the 1:1 ratio of *in vivo/ in vitro* IC_{50} values (generally for peptidic inhibitors, this ratio is 10-10000 fold higher, indicating poor cell uptake properties). Optimisation of the methyl ester analogues was attempted by alteration of the binding residues flanking either side of the phosphonamidate moiety. However, such alterations had only a small effect on inhibitor potency and the trends observed for the more potent hydroxyl-containing inhibitors were not seen with our compounds. These results suggest hydrogen-bond donating capacity as a key requirement for potent inhibition of the HIV-1 protease, a feature which the methyl ester analogues lack!

Due to the associated problems with peptidic inhibitors, the development towards two novel non-peptidic inhibitors of the HIV-1 protease was undertaken. Both cyclic inhibitors were designed to optimise the key interactions at the core of the active site of the enzyme in an attempt for selective, highly potent, low molecular weight inhibitors. Such inhibitors were designed to replace a structural water molecule in the flap region of the enzyme, whilst maintaining the key interactions with the catalytic aspartates. For this purpose cyclic compounds possessing both alcohol functionality and ring heteroatoms were synthesised. The first sulfur-containing analogue produced a moderate inhibitor (μM) of the HIV-1 protease and provided a lead compound for further development. The second seven membered ring analogue was designed on the basis of a recent literature compound and the synthesis incorporated a novel *bis*-ketone, derived from diethyl L-tartrate. This synthesis has yet to be completed and is currently under investigation in our group by Neil Piggot.

Contents

Abbreviation	Meaning
AIDS	acquired immune deficiency syndrome
AMB	2-(aminomethyl)-benzimidazole
AMP	2-(aminomethyl)-pyridine
AZT	3'-azido-3'-deoxythymidine
Boc	tertiarybutoxy carbonyl
CA	capsid protein
Chex	cyclohexyl
Cbz	carbobenzyloxy
ddI	2', 3'-dideoxyinosine
diast.	diastereomer
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DTT	D,L dithiothreitol
EDTA	ethylene diaminetetraacetic acid
ELISA	enzyme-linked immunosorbent assay
EPNP	1, 2-epoxy-3-(4-nitrophenoxy)propane
eq.	equivalents
FAB	fast atom bombardment
gp	glycoprotein
HIV	human immunodeficiency virus
HMPA	hexamethylphosphoramide
HPLC	high performance liquid chromatography
iBu	*iso*-butyl
IC_{50}	concentration required to reduce enzyme activity by 50%

IN	integrase
K_i	enzyme inhibition constant
K_M	Michaelis-Menten constant
MA	matrix protein
MHC	major histocompatibility complex
mRNA	messenger ribonucleic acid
N.D.	not determined
NMM	N-methylmorpholine
NMR	nuclear magnetic resonance
PBS	phosphate buffered saline
Pr	protease
Qua	quinoline-2-carbonyl
RF	replicative form
RN	ribonuclease
RNA	ribonucleic acid
RT	reverse transcriptase
SIV	simian immunodeficiency virus
TCID	tissue culture infective dose
THF	tetrahydrofuran
Thfg	tetrahydrofuranylglycine
TLC	thin layer chromatography
TMSBr	trimethylsilyl bromide
UV	ultraviolet

The one and three letter codes for the amino acids

Amino acid	Three letter code	Single letter code
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cyclohexylalanine	Cha	-
Cysteine	Cys	C
Glutamic acid	Glu	E
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

CHAPTER ONE

INTRODUCTION TO HIV-1

1.0 Introduction.

1.1 General introduction.

The search for a cure for the acquired immune deficiency syndrome (AIDS), continues to be one of the highest research priorities world wide. If current estimates of the dimension of the disease are accurate, millions of individuals will be in the terminal stages of the disease during the next decade. The immediate discovery of an effective therapy is therefore crucial.

AIDS is a disease that cripples the human immune system and the general consensus of opinion at present is that it is caused by the human immunodeficiency virus (HIV). Infection by the virus, which targets the CD4 receptor protein of human helper T-cells, eventually produces profound defects in cell-mediated immunity.[1] Over time, infection leads to a severe depletion of T-cells resulting in opportunistic infections and ultimately death.

The only two treatments for HIV infection so far approved are directed against the viral enzyme, reverse transcriptase (RT).[2] These are Zidovudine, (3'-azido-3'-deoxythymidine (AZT)), a 'suicide' precursor substrate and 2', 3'-dideoxyinosine (ddI), a chain termination inhibitor. Both of these therapies suffer from two major drawbacks; they are highly toxic[3] and resistant mutants have already appeared against AZT. The speed at which these mutants have appeared does not bode well for ddI. Additionally, some patients cannot tolerate AZT therapy at all.

1.2 The immune system.

The body is protected by a diverse army of cells and molecules that work in concert. Their ultimate target is an antigen, which is usually a foreign molecule from a bacterium or other invader. Specialised antigen-presenting cells, such as macrophages, roam the body, ingesting and fragmenting the antigens they find into antigenic peptides. Pieces of these peptides are joined to major histocompatibility complex (MHC) molecules and are displayed on the surface of the cell. Other white blood cells, called T-lymphocytes (T-cells), have receptor molecules that enable each of them to recognise a different peptide-MHC combination. T-cells activated by that recognition divide and secrete

lymphokines, or other chemical signals, that mobilise other components of the immune system. One set of cells that respond to these signals are the B-lymphocytes (B-cells), which can identify antigens free in solution. When activated, the B-cells divide and differentiate into plasma cells that secrete antibody proteins, which are soluble forms of their receptors. By binding to antigens that they find, the antibodies can neutralise them or precipitate their destruction by complementary enzymes or by scavenging cells. Some T- and B-cells become memory cells that persist in the circulation and boost the immune system's readiness to eliminate the same antigen if it presents itself in the future. Reduction in the number of T-cells can lead to a decline in immune function which can ultimately allow invasion by foreign molecules.

1.3 The human immunodeficiency virus.

The first viral species to be associated with AIDS was found both in the Western world and Central Africa and has now come to be known as the human immunodeficiency virus-1 (HIV-1).[4] Soon after the discovery of HIV-1, another major and distinct species of virus, called the HIV-2 was identified in West Africa.[5] HIV, a lentivirus, is a member of the family of retroviruses, *Retroviridae*.[6] The genomes of HIV-1 and HIV-2 are about 40% identical, but that of HIV-2 is about 80% identical to the genome of simian immunodeficiency virus (SIV). Although two related and pathogenic strains of HIV have been found, HIV-1 is the pathogen of greater consequence and therefore the following discussion will concentrate on this subtype.

The spherical HIV-1 (diameter approx. 1100 Å) comprises two portions, an envelope and a core.

One part of the envelope is a lipid bilayer that is similar to the plasma membrane of human cells. Embedded in this layer are three proteins that complete the envelope; the transmembrane protein (gp41), surface protein (gp120) and matrix protein (MA p17). The abbreviation for the protein (p) or glycoprotein (gp) is followed by a number that gives the approximate molecular size of the protein in kilodaltons as determined by electrophoresis.

A structural representation of HIV-1 is shown below (Figure 1.1).

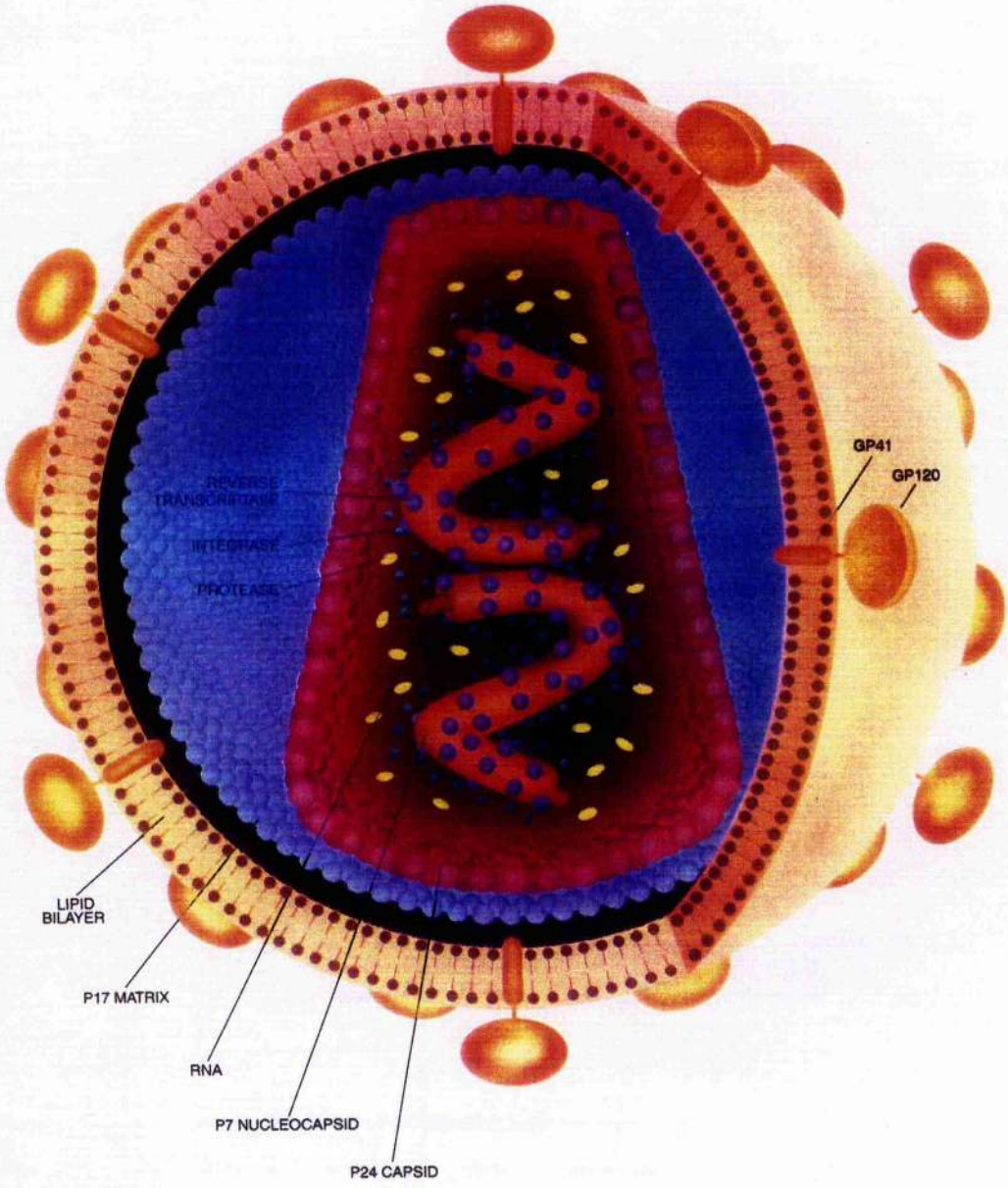


Figure 1.1. Schematic representation of infectious HIV.
(reproduced with the kind permission of *Scientific American*).

The core portion is roughly bullet-shaped, with an outer capsid protein (CA, p24). Inside, there are two molecules of RNA and four viral enzymes, the protease (Pr), reverse transcriptase (RT), ribonuclease H (RN) and integrase (IN).

1.3.1 The viral life cycle of HIV.

The primary cell type in which HIV-1 replicates is the human T4-lymphocyte.[7] However, many other cell types have been shown to be infected by the virus, including cells in the brain,[8] gastrointestinal tract,[9] kidney,[10] liver,[11] lung,[12] skin fibroblasts[13] and the nervous system.[14] Some of these targets do not display the receptor protein CD4, which suggests that CD4 is not the sole vehicle for viral entry into cells. There must also exist a second receptor which helps to expedite this process.

The replication cycle of HIV and other retroviruses within infected cells can be divided into two phases; the pre-integration and post-integration phases. The pre-integration phase is carried out almost exclusively by enzymes and proteins encoded within the infecting virion, while a number of steps in the post-integration phase, in which viral genes are expressed and new virions are assembled, are performed by cellular enzymes. The events of the life cycle can be summarised as follows (see overleaf, Figure 1.2):

i) The external glycoprotein portion of the viral envelope protein, gp120, recognises and binds with affinity to the extracellular component of the receptor protein CD4.

ii) The membrane portion of the viral envelope fuses to the lymphocyte membrane, followed by injection of the virion into the cell. This process is mediated by gp41.

iii) The virion is uncoated by a proteolytic event in which the HIV-1 protease may play a major role.[15] The viral RNA genome, the nucleocapsid proteins and the viral enzymes, reverse transcriptase (RT), integrase (IN) and protease (Pr) are released into the cytoplasm.

iv) The viral RNA is transcribed into DNA by reverse transcriptase.

v) The viral DNA then translocates to the cell nucleus where it is irreversibly integrated into the chromosomal DNA by the integrase enzyme. This forms the provirus which is now a permanent feature of the lymphocyte. These events

complete the pre-integration stage of the cycle.

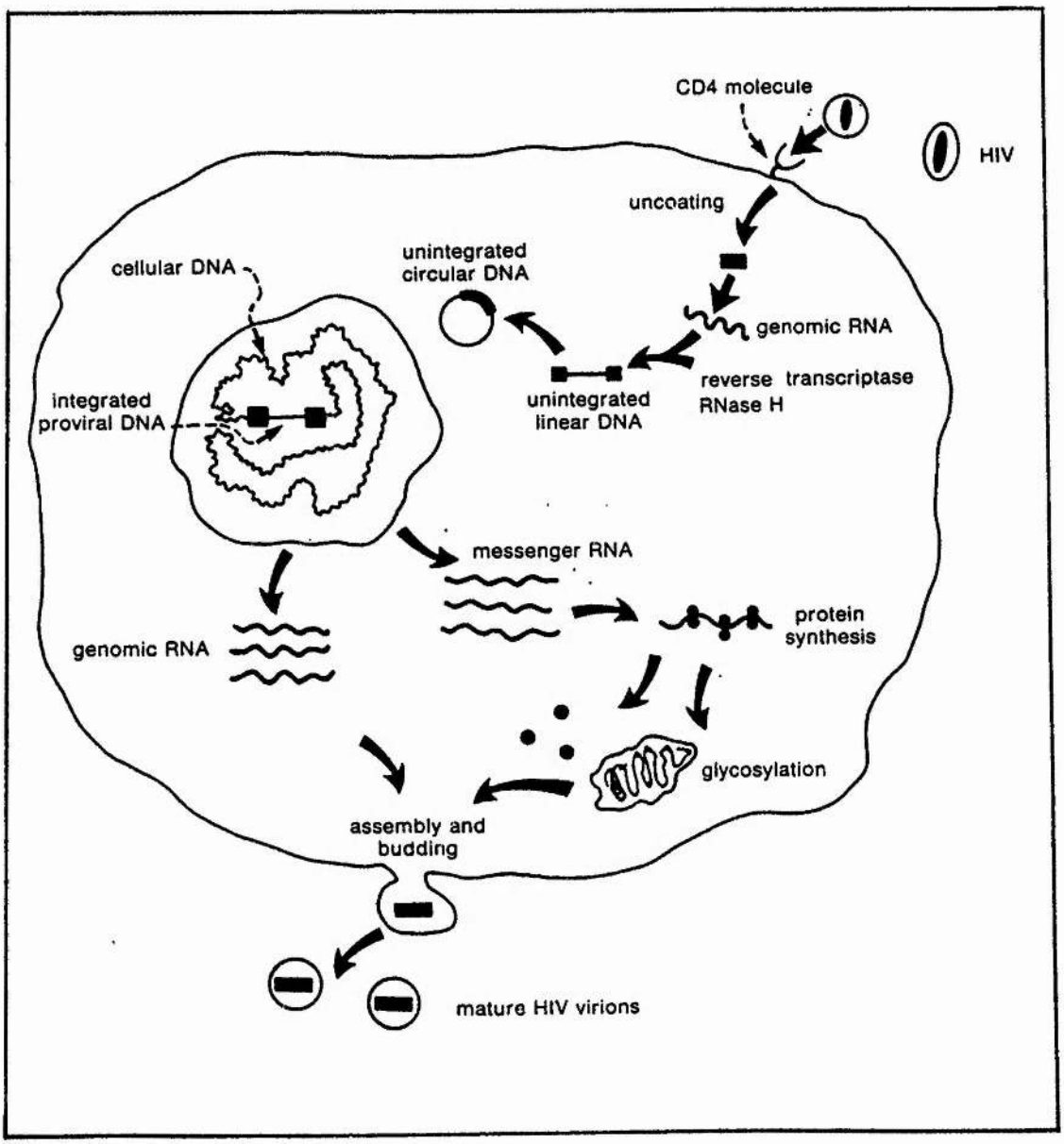


Figure 1.2. The life cycle of HIV-1 within a human T4-lymphocyte.

In the post-integration phase, the transcription and translation of the HIV provirus is subject to complex regulation by both *cis*-acting elements within the DNA genome and by several *trans*-acting viral proteins, such as *tat* and *rev*, which are unique to HIV.[16, 17]

vi) A primary RNA transcript is made by the host RNA polymerase II. This transcript provides both the messenger RNA (mRNA) for the synthesis of the viral polyproteins and the viral RNA which will be incorporated into the new virions.

vii) The host cellular ribosomes translate the proviral mRNA into viral polyproteins.

viii) Cellular enzymes carry out post-translational modification of the polyproteins such as glycosylation of the envelope proteins and myristoylation of the N-termini of the viral polyproteins. This latter modification appears to be essential for attachment of these proteins to the cell membrane and thus aid alignment of adjacent molecules.

ix) The glycoprotein envelope now displayed on the surface of the infected cell probably drives assembly and sequestration of the polyproteins and genomic RNA in an area of the host cell that is highly concentrated in viral elements. Non-infectious, 'immature' particles then leave the cell in a process often referred to as 'budding'.

x) Maturation of the fully-formed 'immature' virion particles which have budded from the cell is effected by the HIV-1 protease. The protease catalyses its own release and then specifically cleaves the viral polyproteins to yield the functional enzymes and structural proteins of the virion core. This processing completes the replication cycle of HIV-1. The resulting mature virion particles (see Figure 1.1.) are now able to promote a new infection in an adjacent T4-lymphocyte.

A number of steps in the life cycle of HIV-1 are potential targets for antiviral therapy. These include the cellular CD4 receptor, the enzymes reverse

transcriptase, ribonuclease H, integrase, protease, regulatory proteins needed for viral transcription, processing of viral mRNA, processing of viral proteins (e.g. glycosylation, myristoylation), assembly at the plasma membrane and budding of virus particles. Of these, the reverse transcriptase and protease enzymes have received the greater attention. The HIV-1 protease is of particular interest to our work and several excellent reviews discussing the HIV-1 protease as a possible therapeutic target for AIDS have been published.[18, 19, 20, 21]

The HIV-1 protease has an important role at the pre-integration stage of the cycle, such as destabilisation of the virion core and so inhibition of the protease may prevent proviral integration. The post-integration role of the protease involves the production of the mature infectious virion from the 'immature' virion and so inhibition of the protease may stop the spread of the virus from infected to non infected cells. The early and late role of the protease offers an immediate advantage over reverse transcriptase inhibitors, which act solely at a pre-integration stage. Inhibition of the protease should be effective at both pre- and post-integration stages. This theory has been supported by the fact that protease inhibitors have been found to block maturation of virions shed from chronically-infected T4-lymphocytes, as well as to attenuate acute infection of uninfected T4-lymphocytes, suggesting that the protease does indeed have an early and late role in the viral cycle.

However, inhibitors of reverse transcriptase have recently been shown to be more effective against an acute infection than in reducing the population of active virus in chronically-infected T4-lymphocytes.[22] This supports the pre-integration role and therefore drugs based on reverse transcriptase inhibitors should be more effective for patients who are at the earlier stages of the disease.

1.3.2 The HIV genome.

HIV-1 exhibits the same 5'-LTR-*gag-pol-env*-LTR-3' genomic organisation as all retroviruses and further encodes several proteins with regulatory or unknown function. The genomic organisation of HIV-1 is shown overleaf (Figure 1.3).

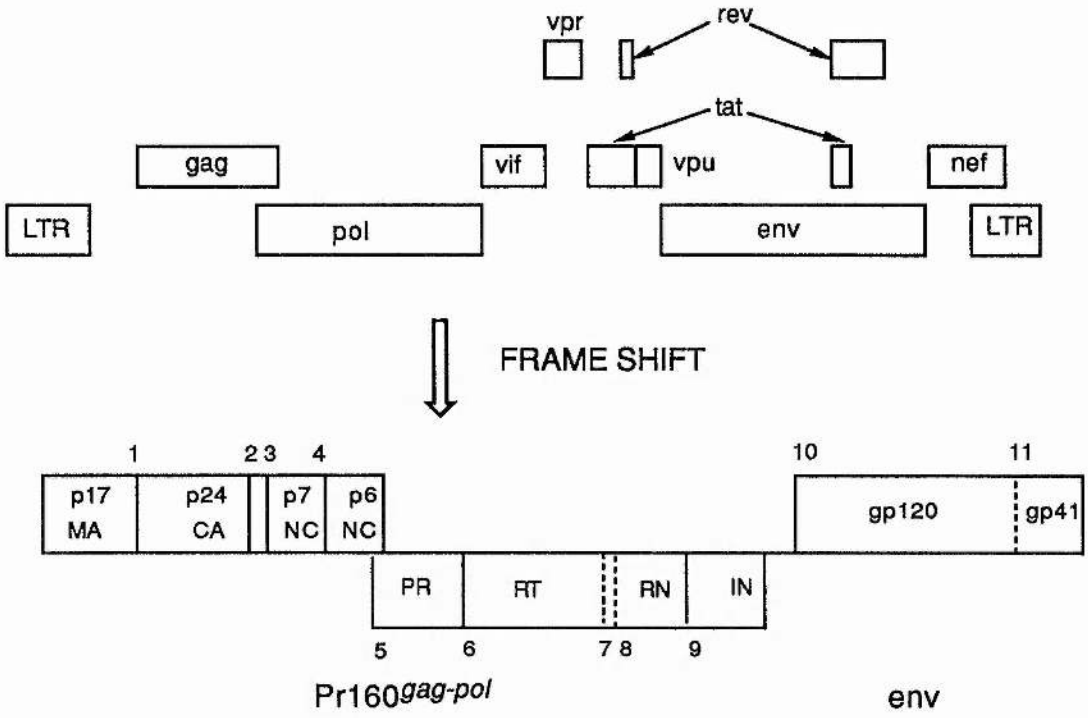


Figure 1.3. The HIV-1 genome; containing (i) LTR - long terminal repeat, (ii) vpr - activator of viral proteins, (iii) *rev* and *tat* - regulatory genes, (iv) vif - promotes viral infectivity, (v) vpu - promotes viral budding, (vii) nef - unknown function, (viii) components of the Pr160$^{gag\text{-}pol}$ will be defined in section 1.4.4.

The *gag* and *pol* regions are initially translated by the host ribosomes into large polyprotein precursors that are later processed to yield the mature structural and catalytic proteins found in the virion. Translation of the HIV-1 *gag* open reading frame results in a 55kD precursor Pr55^{gag}, which contains the viral structural proteins. The HIV-1 *pol* open reading frame is translated as a *gag-pol* fusion polyprotein, Pr160$^{gag\text{-}pol}$, as a result of a ribosomal frameshift from the *gag* to the *pol* frames at a specific site positioned before the *gag* termination codon. Pr160$^{gag\text{-}pol}$ contains the protease, reverse transcriptase, ribonuclease H and integrase. The proteolytic maturation of Pr55^{gag} and Pr160$^{gag\text{-}pol}$ is absolutely essential for the successful completion of the viral life cycle and the production of infectious virions (see section **1.3.1**).

1.4 The HIV-1 protease.

1.4.1 Classification.

From the genomic sequence of the virus, Ratner and co-workers postulated that the second open reading frame (*pol*) of HIV-1 encoded a protease analogous to those of other retroviruses.[6] More significantly, Toh *et al.* recognised that the highly conserved triad Asp-Thr-Gly (DTG), found in other retroviral sequences was homologous to the catalytic site of proteases belonging to the aspartic acid family and proposed that the viral enzymes were of this class.[23]

The aspartic proteases, so named because of the highly conserved pair of aspartic acid residues in their active site have long been recognised as enzymes from vertebrates, plants and fungi. Non-viral aspartic proteases, such as pepsin, renin, cathepsin D and chymosin have been extensively studied over the years and consequently much is known about their mechanism, amino acid sequence and structure. This has obvious benefits to the study of HIV-1 protease.

Further evidence confirmed the classification of the HIV-1 protease. First, the enzyme was susceptible to inhibition by the classical aspartic acid inhibitor, Pepstatin A (see Figure 1.4).[24] Second, the enzyme was irreversibly inactivated

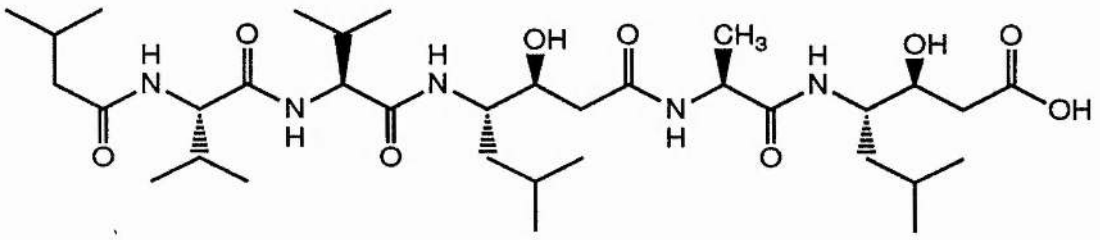


Figure 1.4. Pepstatin A.

by 1, 2-epoxy-3-(4-nitrophenoxy)propane (EPNP)[25] (see overleaf, Figure 1.5) which was previously shown to inhibit pepsin by esterification of one or both of its active site aspartate residues.[26] Third, site-directed mutagenesis of the active site Asp-25 to Asn-25 by Kohl *et al.* completely abolished enzyme activity.[27]

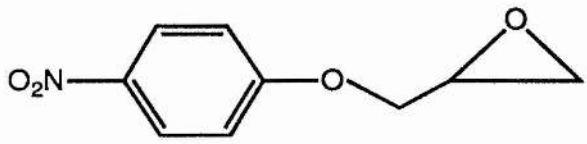


Figure 1.5. EPNP.

Ultimately, elucidation of the crystal structure of both native HIV-1 protease[28, 29, 30] and HIV-1 protease complexed with inhibitors (see section **1.4.8**) confirmed the classification of this enzyme.

1.4.2 Structural properties.

Despite the notable DTG sequence conservation, the retroviral proteases were found to have significant structural differences to their mammalian and fungal counterparts. For instance, HIV-1 protease contains only 99 amino acid residues, compared to over 200 residues for the non-viral proteases. The non-viral proteases are monomeric and contain the DTG sequence twice in their sequence. These bi-lobal proteins contribute one aspartate residue per lobe to the active site. However, the DTG sequence only appears once in the HIV-1 protease case. These observations led Pearl and Taylor [31] to suggest that the catalytically active form of HIV-1 protease was in fact a homodimer, with each monomer contributing one of the two conserved aspartates to the active site. This prediction has since been borne out by crystallographic studies (see Figure 1.15, section **1.4.8**). The solution of the three-dimensional structure of HIV-1 protease shows that the enzyme possesses a strict C-2 axis of symmetry, whereas the non-viral proteases have an approximate two fold axis. HIV-1 protease contains two identical 'flap' regions, analogous to the single flap found in the non-viral proteases, which form in part the substrate/ inhibitor binding sites. Another structural feature, which distinguishes the HIV-1 protease from the non-viral proteases, is that in all structures of HIV-1 protease-inhibitor complexes solved to date, within the 'flaps' is a water molecule which is hydrogen-bonded in a tetrahedral fashion to the amide hydrogens of Ile-50 and Ile-50' residues in the flaps. This water molecule appears to play an important role in binding to both the flaps and inhibitor. The C-2 nature of the enzyme and

the flap water molecule have prompted investigation into selective inhibitor design for the HIV-1 protease and will be discussed fully in sections **1.4.6 (viii)** and **2.6.2.**

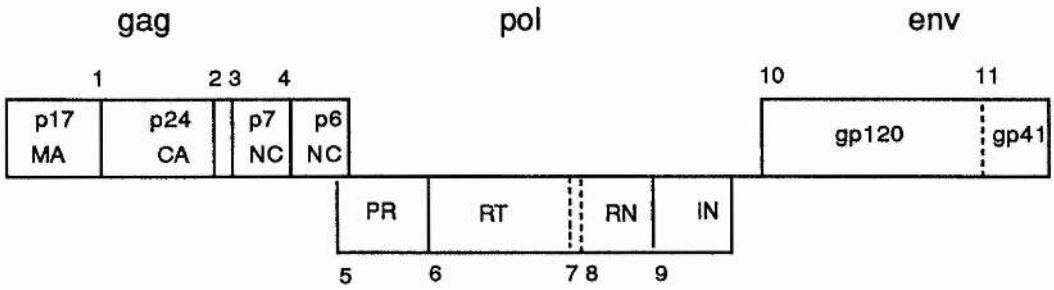
1.4.3 Preparation and purification.

An understanding of the protein chemistry and enzymology of the HIV-1 protease is an obvious prerequisite for inhibitor development and thus adequate amounts of HIV-1 protease are required. Although the natural source of HIV-1 protease is the virus itself, most researchers avoided this method of isolation. However, Lillehoj *et al.* purified minute quantities of protease directly from the virus and lived to tell the tale.[32] More popular methods of preparation include total synthesis and recombinant techniques. Schneider and Kent[33] synthesised an enzymatically active 99-residue protease based on the sequence of HIV-1. Nutt *et al.*[34] also prepared HIV-1 protease by total synthesis using solid-phase peptide methodology and demonstrated that refolding of the polyprotein yielded active enzyme that behaved as the expected dimer. Gel filtration chromatography confirmed the size of the protein by calibration against standard proteins. Many laboratories have cloned and expressed the HIV-1 protease in bacteria[27, 35] and yeast.[36] Expression in *Escherichia coli* of recombinant constructs containing the coding region of HIV-1 protease demonstrated that the enzyme is able to catalyse its own release from larger precursors. This may provide clues to the release of protease from the polyprotein in the maturation step of the viral life cycle (see section **1.3.1**).

A number of purification procedures have been reported, including affinity chromatography[37] and HPLC.[38]

1.4.4 Substrate specificity.

The HIV-1 protease when processing the viral polyproteins, $Pr55^{gag}$ and $Pr160^{gag\text{-}pol}$, is required to make highly specific cleavages. The cleavage sites of the HIV-1 protease within the *gag* and *pol* region are shown overleaf.



Pr160$^{gag\text{-}pol}$

where MA = matrix protein
CA = capsid protein
p7 NC = nucleocapsid protein
p6 NC = nucleocapsid protein
Pr = protease
RT = reverse transcriptase
IN = integrase
RN = ribonuclease

The cleavage sites (*) of the HIV-1 protease within the Pr160$^{gag\text{-}pol}$ are as follows:

	(P_4)-(P_1)		$(P_{1'})$-$(P_{4'})$
1.	Ser-Gln-Asn-Tyr	*	Pro-Ile-Val-Gln-
2.	Ala-Arg-Val-Leu	*	Ala-Glu-Ala-Met-
3.	Ala-Thr-Ile-Met	*	Met-Gln-Arg-Gly-
4.	Pro-Gly-Asn-Phe	*	Leu-Gln-Ser-Arg-
5.	Ser-Phe-Asn-Phe	*	Pro-Gln-Ile-Thr-
6.	Thr-Leu-Asn-Phe	*	Pro-Ile-Ser-Pro-
7.	Leu-Glu-Lys-Glu	*	Pro-Ile-Val-Gly-
8.	Ala-Glu-Thr-Phe	*	Tyr-Val-Asp-Gly-
9.	Arg-Lys-Ile-Leu	*	Phe-Leu-Asp-Gly-

The complete proteolytic processing of the Pr55^{gag} and Pr160$^{gag\text{-}pol}$ requires only eight discrete cleavages. However, the enzyme seems to tolerate a variety of amino acid residues in the P_1-$P_{1'}$ cleavage sites (the P_n, $P_{n'}$ terminology of Schechter and Berger[39] has been adopted and prescribes that amino acid residues are numbered P_1 P_2 ..... from the scissile bond to the amino terminus and $P_{1'}$ $P_{2'}$ ..... from the scissile bond towards the carboxy terminus). Three of the cleavage sites (1, 5 and 6), consist of aromatic amino acid-proline residues and this type of cleavage site is indeed unique to

retroviruses.[40] Therefore inhibitors which incorporate these P_1-$P_{1'}$ residues may deemed to be specific for the retrovirus. The other cleavage sites show a general pattern for hydrophobic residues in the P_1-$P_{1'}$ positions. Analysis of the proteolytic processing of HIV-1 protease led to the proposal that residues flanking the cleavage site could be assigned to one of three classes:

Class 1 contains Phe-Pro or Tyr-Pro at P_1-$P_{1'}$.
Class 2 contains Arg at P_4 and Phe-Leu at P_1-$P_{1'}$.
Class 3 contains Gln or Glu at $P_{2'}$.

In an effort to further characterise the enzyme and to provide information required for 'first generation' inhibitor design, the substrate specificity of HIV-1 protease was examined. Oligopeptides which incorporate the P_1-$P_{1'}$ cleavage site were synthesised and an examination of the preferred residues spanning P_4-$P_{3'}$ were tested as substrates. Darke *et al.* [41] showed that the HIV-1 protease was able to specifically cleave oligopeptides corresponding to all of the eight cleavage sites. Seven residues spanning P_4-$P_{3'}$ were required for efficient cleavage at the P_1-$P_{1'}$ amide bond. These findings are in agreement with crystallographic structures of HIV-1 protease-inhibitor complexes, where the enzyme may accommodate the equivalent of a heptapeptide, spanning the P_4-$P_{3'}$ residues. The relative rates of cleavage for polypeptides containing the natural cleavage sites may reflect a regulatory role.[42] Systematic replacement of the residues in Ser-Gln-Asn-Tyr*Pro-Ile-Val and Lys-Ala-Arg-Val-Leu*Ala-Glu-Ala heptapeptide substrates gave the following trends; (+) indicates a suitable residue and (-) indicates an unsuitable residue.

P_3 Gln, Cys, Tyr, Arg, Val, Asn, Asp and Glu all (+); Pro (-).
P_2 Val, Ile, Leu, Phe, Ala, Asp and Asn all (+); Gly and Pro (-).
P_1 Tyr, Phe, Met, Leu and norleucine all (+); Glu, Arg, Ile and Val (-).
$P_{2'}$ Ile, Ala and Leu all (+); Trp, Phe and Gly (-).

These data are useful for medicinal chemists engaged in inhibitor design. Inhibitors could be derived from the known HIV-1 substrates, by incorporation of the preferred residues into the P_4-$P_{3'}$ positions of the inhibitor. Replacement of the scissile amide bond at P_1-$P_{1'}$ with non-hydrolysable isosteres which mimic

the transition state or reaction intermediate of peptide cleavage can lead to tight binding inhibitors (see section **1.4.6** for a more detailed discussion).

The specificity of HIV-1 protease has also been studied in recombinant proteins.[43] Random mutagenesis of the P_1-P_1' positions at the three *gag-pol* cleavage sites produced proteins which showed variable susceptibility to cleavage.[44] Few or no changes were tolerated at the Pr-RT junction (site 6) and RT-IN (site 8); however the p6-Pr (site 5) accepted a variety of changes. Replacement of the natural Phe-Pro cleavage site with residues as diverse as Gly-His and Ala-Ile resulted in processing equivalent to that of the natural substrate. Therefore the HIV-1 protease is not restricted to the eight substrate cleavage sites and so inhibitor design has greater flexibility in the choice of residues for the P_1-P_1' positions. For instance, it may be possible to produce potent inhibitors that do not contain natural P_1-P_1' residues. Other regions that proved sensitive to change included residues in the active site area (Ala22-Leu33) and flap region (Ile47-Gly52). From crystal structures of protease-inhibitor complexes it can be seen that residues in these regions form an extensive network of hydrogen-bonds with bound inhibitors. Presumably, the same type of hydrogen-bonding occurs in the case of substrates and so disruption of this bonding network would cause a loss in binding potential. A third region included residues Thr74-Arg87. Selective modification of Arg-87 and Asn-88 destabilised the dimeric form of the protease and led to enzyme inactivation.[45]

1.4.5 Catalytic mechanism.

The nature of aspartic protease catalysis has been widely studied over the years. The large amount of sequence homology amongst the mammalian aspartic proteases and the sensitivity to the same inhibitors, indicates that they operate *via* the same mechanism. Initially, it was proposed that the aspartic proteases operated *via* a covalent mechanism, in which one of the active site aspartates reacted with the substrate to form an enzyme bound covalent intermediate. Two variations of the covalent mechanism were proposed; the amino-enzyme mechanism[46] and the acyl-enzyme mechanism.[47] These are outlined overleaf (Figures 1.6 and 1.7).

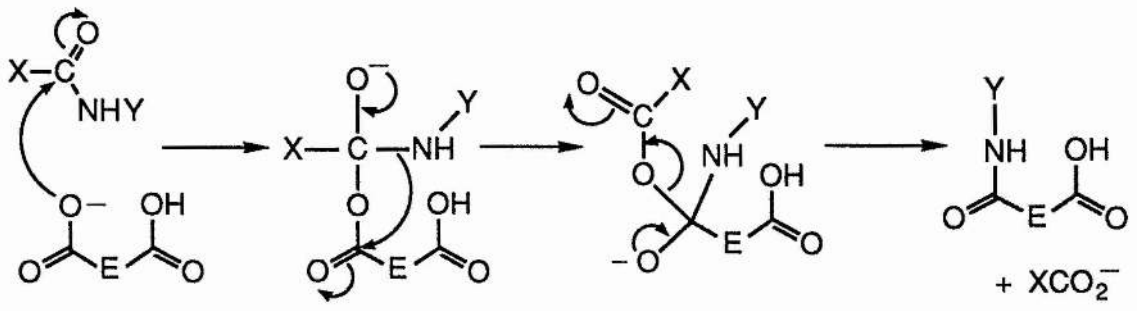


Figure 1.6. The amino mechanism where E represents enzyme.

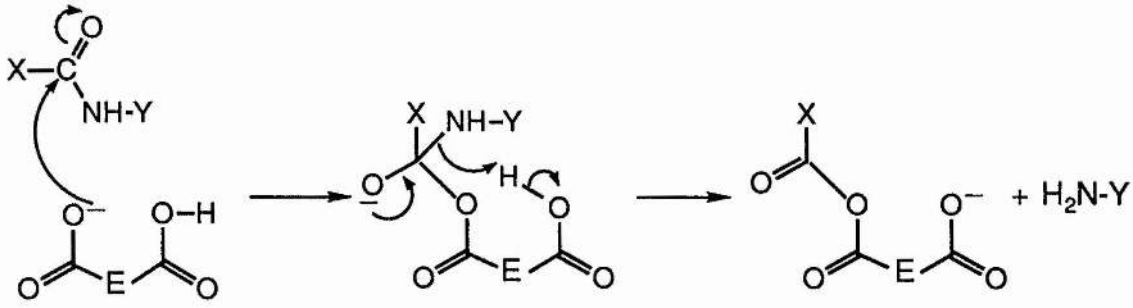


Figure 1.7. The acyl enzyme mechanism.

However, kinetic studies and examination of protease-inhibitor complexes has led to the proposal that the aspartic proteases operate *via* a general acid-general base mechanism, which involves both of the aspartates and a nucleophilic water molecule (Figure 1.8). The first key experiments to dispute the covalent mechanisms came from labelling studies. Antonov[48] carried out labelling experiments using $H_2{}^{18}O$ as solvent and found that ^{18}O-label was incorporated into uncleaved substrate. This could only be explained by a nucleophilic water mechanism.

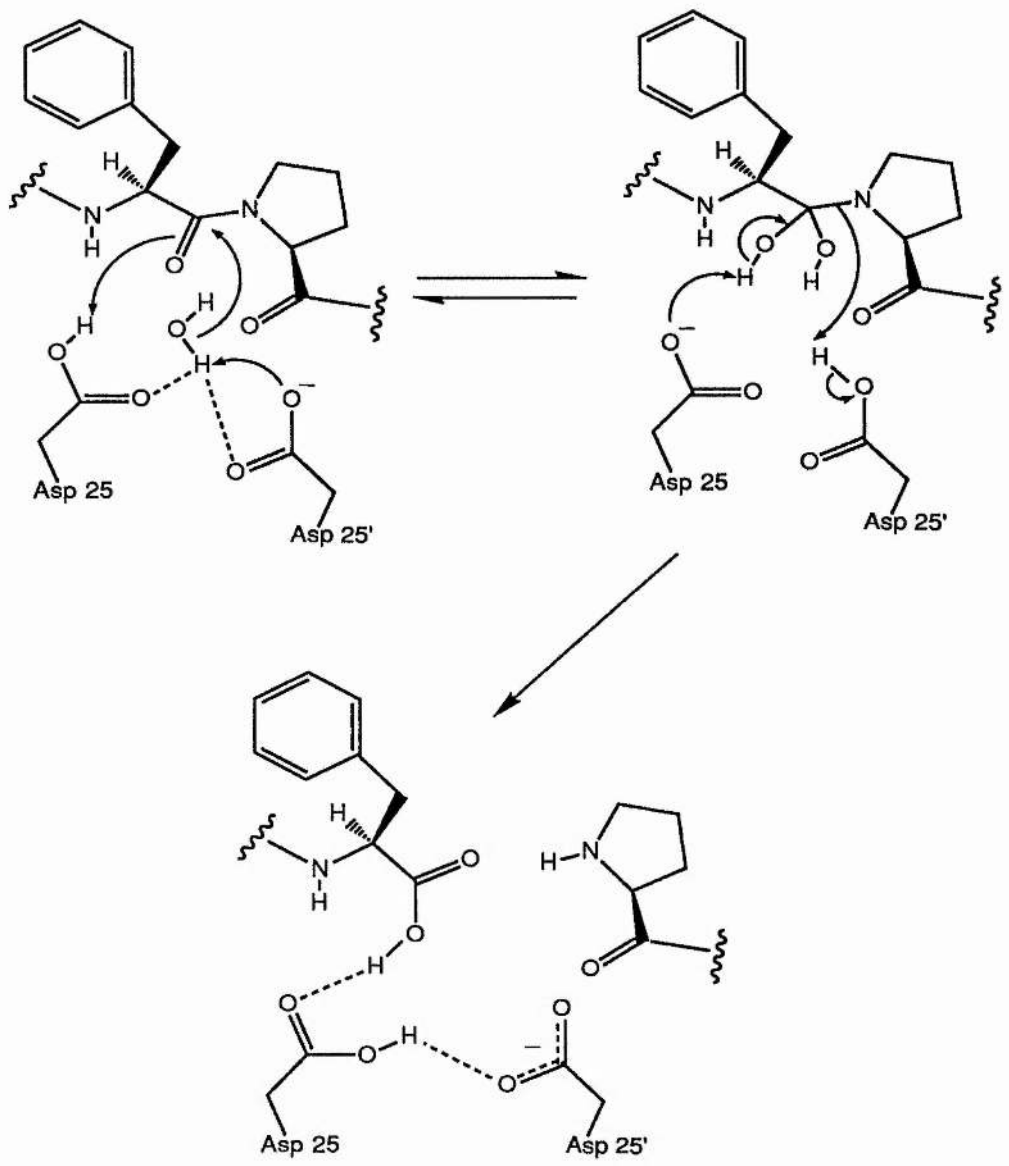


Figure 1.8. The general acid-general base mechanism.

In this mechanism, the reaction intermediate is a gem diol, which collapses to form products as shown. However, the gem diol intermediate could collapse to reform the substrate. The ^{18}O-label would be incorporated into the gem diol from the solvent $H_2{}^{18}O$ and on breakdown, the gem diol could produce ^{18}O-label incorporation into the carboxyl of the product or into the amide carbonyl of reformed substrate. ^{18}O-Labelled substrate cannot be explained *via* a covalent mechanism, as hydrolysis of any enzyme bound covalent intermediate with $H_2{}^{18}O$, would result in ^{18}O-label incorporation into the active site carboxyls; this is not observed. The analysis of protease-inhibitor complexes is consistent with

a general acid-general base mechanism. Inhibitors which mimic the gem diol intermediate, bind to both catalytic aspartates, one of which is protonated and one is not.

By virtue of the fact that the HIV-1 protease belongs to the aspartic acid family, one would expect that this enzyme operates by the same mechanism, as outlined in Figure 1.8. The key features of this mechanism are as follows; a water molecule is hydrogen-bonded to the Asp-25 and Asp-25' residues, of which, one is protonated. Binding of the substrate facilitates general acid protonation of the scissile amide carbonyl from Asp-25 and general base deprotonation of the lytic water molecule by Asp-25' to form the aforementioned gem diol tetrahedral intermediate. Synchronous deprotonation of one of the gem diol hydroxyls by Asp-25 and protonation of the departing prolyl nitrogen by Asp-25', renders the free carboxyl and amino fragments and restores the aspartyl residues to their initial protonation states.

Results from pH dependence studies of the kinetic constant k_{cat}/ K_M for several oligopeptide substrates show a 'bell-shaped' curve over the pH range 3.4-6.5 and indicate that an unprotonated group of pKa 3.3 and a protonated group of pKa 5.5-6.1 are required for catalysis.[49] This is in agreement with the general acid-general base mechanism. However, the mechanism shown in Figure 1.8, show only the main features. The position of the water molecule in the active site and the nature of the residues involved in catalysis have been the subject of debate. Several derivations of the general acid-general base mechanism have been proposed which share some central common features, such as the involvement of the aspartate diad and a nucleophilic water molecule. However, they differ in the following way:

i) the proposed protonation state of the tetrahedral intermediate,
ii) the timing of proton switches in the breakdown of this intermediate,
iii) the direction of binding of the scissile carbonyl with respect to the aspartate diad and
iv) the nature of the Brønsted/ Lewis acid component.

For a more detailed account, refer to the mechanisms of James,[50] Polgar,[51] Pearl,[52] Blundell[53] and Davies[54]. All of these mechanisms propose a gem diol/ gem diolate intermediate and so the use of transition state mimicry is valid for all

cases. The nature of the gem diol could be investigated using both neutral and negatively charged transition state mimics. For instance, if the mechanism proceeds *via* a gem-diolate intermediate, negatively charged inhibitors would be closer mimics and so should bind more tightly. On the other hand, if the mechanism proceeds *via* a neutral gem diol intermediate, negatively charged inhibitors would be poorer mimics and so may bind less tightly.

1.4.6 Inhibition of the HIV-1 protease.

Due to its critical role in the processing of mature virus, the HIV-1 protease has become a popular therapeutic target for AIDS. Early efforts in inhibitor design were based on transition state mimicry.[55] This involves replacement of the scissile amide bond at P_1-P_1' with non-hydrolysable functional groups which mimic the gem diol tetrahedral intermediate for the enzymic reaction. By binding the transition state of a reaction tightly, enzymes can lower the activation energy of a reaction and thus, mimicking the transition state with non-hydrolysable isosteres should lead to tight binding inhibitors. This strategy has proven successful for inhibitors for the aspartic protease, renin, an important enzyme involved in blood pressure regulation.[56] Examples of transition state analogues include reduced amides, statines, hydroxyethylenes, dihydroxyethylenes, hydroxyethylamines, difluoroketones and phosphinic acids, all of which are shown overleaf in Figure 1.9.

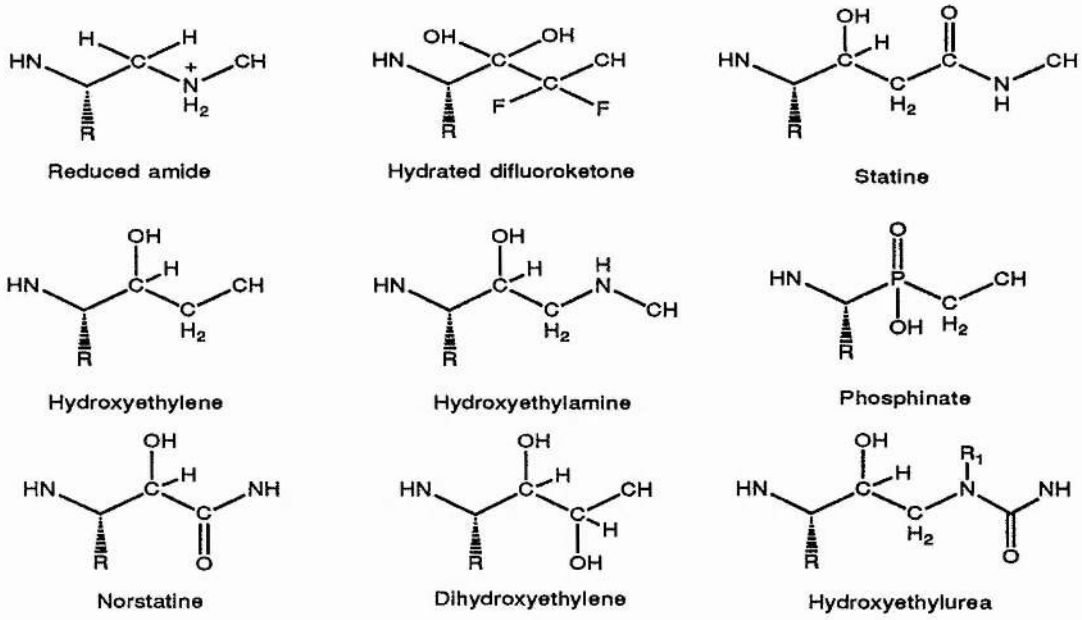


Figure 1.9. Transition state mimics for the HIV-1 protease.

Armed with the results from the study of renin inhibition and the fact that substrates of HIV-1 protease were known, rapid progress has been made in designing inhibitors for the HIV-1 protease. Incorporation of transition state mimics into substrates of HIV-1 protease has been successful in producing a variety of inhibitors. More significantly, the effectiveness of each transition state analogue has been compared. Dreyer *et al.*[57] incorporated each isostere into a Ser-Ala-Ala-Phe*Pro(Gly)-Val-Val-OMe substrate and found a wide range of inhibition (K_i= 62-19000 nM). These inhibitors can be ranked by potency in the order: reduced amide ((**1**) least potent) < phosphinic acid (**2**) < statine (**3**) < difluoroketone (**4**) < hydroxyethylene ((**5**) most potent) (shown overleaf in Figure 1.10). These trends have been reported by other groups in which peptide analogues of different substrate sequences have been studied.[28, 58, 59,

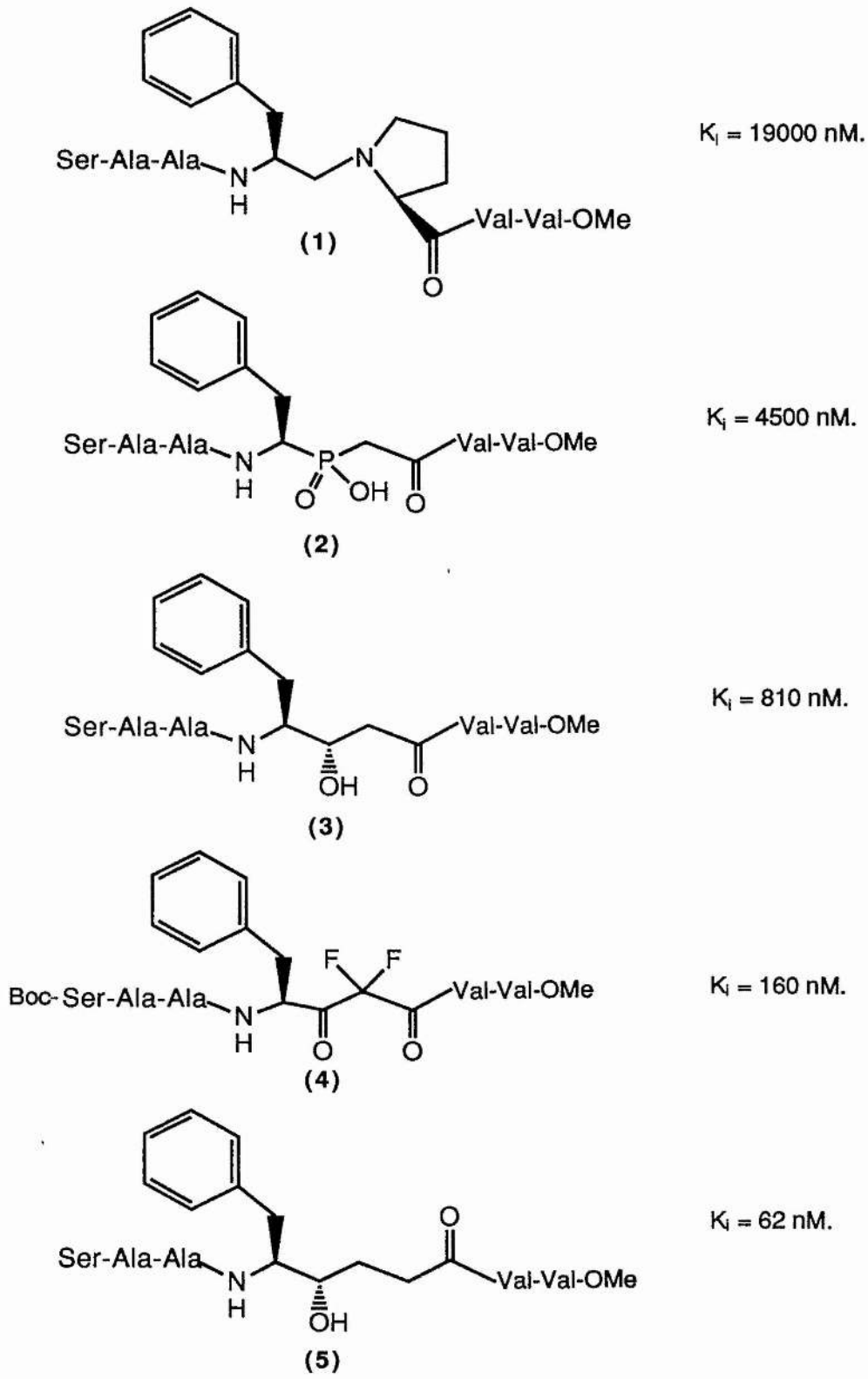
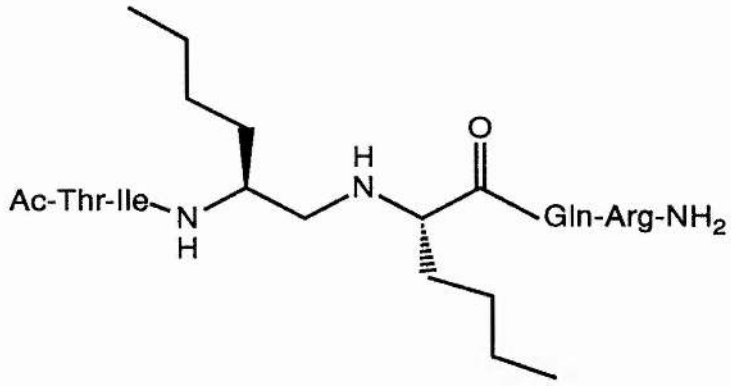


Figure 1.10. Comparison of dipeptide isosteres within a single peptide template.

A closer examination of these transition state mimics may shed some light on the observed trends. Each class of transition state isostere will be considered individually.

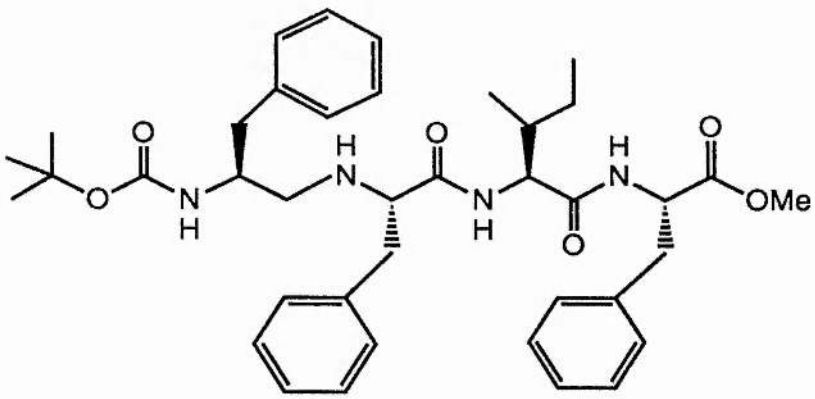
(i) Reduced amides.

The reduced amide concept was used by Szelke *et al.*[60, 61] to create potent inhibitors of renin. It was thought that the potency stemmed from an ionic interaction between the protonated secondary amine and the unionised aspartate in the active site. However, if true, this type of interaction cannot be considered to be one of transition state mimicry. It is not clear why the reduced amides are less potent towards HIV-1 protease, although a recent study by Hyland et al. [49] has shown that the pH dependence of a close analogue to (**1**) is consistent with binding of this cationic inhibitor to the protease only when *both* of the catalytic aspartates are unprotonated. However, at the pH optimum for the protease (pH 4.5-6.0), one of the aspartates is protonated. Considering the reduced amide isostere purely as a transition state mimic, may explain its lack of potency. Although tetrahedral in nature, this mimic does not possess either oxygen (or equivalent heteroatom) of the gem diol intermediate and so interactions with the catalytic aspartates should be poor. Another reduced amide inhibitor based on the Phe-Pro cleavage site was synthesised by Cushman *et al.*[62] A heptapeptide analogue, Thr-Leu-Asn-Phe-ψ-[CH_2N]-Pro-Ile-Ser was found to possess an IC_{50} of 1.4 μM (IC_{50} -the concentration of inhibitor required to reduce enzyme activity by 50%). More significantly, it was found that the enzyme only bound the (2S)-phenylalanine epimer, indicating the enzyme's preference for the natural amino acid form. The (2R)-phenylalanine epimer was found to be inactive. A more potent reduced amide inhibitor is compound (**6**) which was derived from a substrate sequence which incorporates the p1-p7 cleavage site within the Pr55^{gag} region. Replacement of Met at P_1-$P_{1'}$ with norleucine led to a substrate whose k_{cat}/ K_M value is among one of the most efficiently processed substrates of the HIV-1 protease.[41] Correspondingly, incorporation of this P_1-$P_{1'}$ site into compound (**6**) led to increased potency (K_i = 780 nM).[63]



(6); Norleucine based inhibitor (K_i = 780 nM).

Optimisation of the residues in the P_1-$P_{1'}$ sites can have a considerable influence on inhibitor potency. Urban and coworkers[64] investigated the nature of the residues in the P_1-$P_{3'}$ regions and optimised their reduced amide inhibitors. Their lead compound (**7**), Boc-Phe-ψ-[CH_2NH]-Phe-Ile-Phe-OMe possessed a K_i value of 23 nM, considerably more potent than any other reported inhibitor from this class. However, changing the Phe at $P_{1'}$ to a Gly residue causes an approximate 70 fold decrease in potency, with Boc-Phe-ψ-[CH_2NH]-Gly-Ile-Phe-OMe having a K_i value of 1650 nM. This result highlights the importance of the $P_{1'}$ residue and may explain the lower potency of compound (**1**) in the study of Dreyer *et al.*[57] In most cases, Phe-Phe at P_1-$P_{1'}$ produces more potent inhibitors than those based on Phe-Pro at P_1-$P_{1'}$. This is unfortunate, considering that inhibitors based on Phe-Pro were hoped to show selectivity towards the HIV-1 protease.



(7); A potent reduced amide inhibitor (K_i = 23 nM).

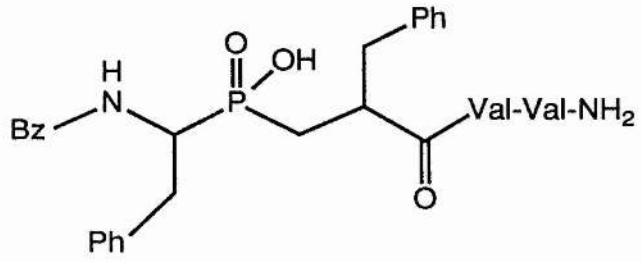
Further optimisation of compound (**7**) at the $P_{2'}$ site produced Boc-Phe-ψ-[CH_2NH]-Phe-Glu-Phe-OMe which gave a K_i value of 0.2 nM, the most potent of all the reduced amide inhibitors.

Although Dreyer's initial study on the comparison of transition state analogues within a single peptide template provided a useful guide to potency, perhaps a study based on optimised inhibitors would provide a stronger basis for direct comparison. The study of Urban clearly shows the sort of effect that optimisation in the P_1-$P_{3'}$ region can have. When comparing transition state analogues within unoptimised inhibitors, especially in the P_1-$P_{1'}$ region, it is difficult to tell whether the observed lack of potency is due to the transition state analogue or the lack of optimisation. By studying optimised inhibitors there can be no doubt.

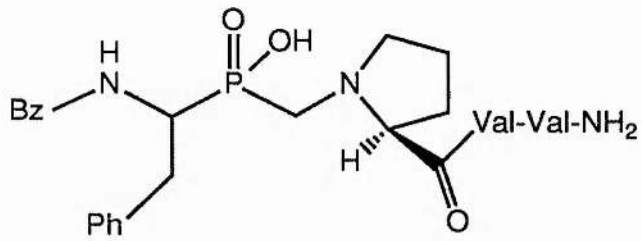
(ii) Phosphinic acids.

Bartlett and Kezer[65] designed phosphinic acids as inhibitors of the aspartic protease, pepsin. These inhibitors are amongst the most potent compounds reported for pepsin. The phosphorus atom provides the correct tetrahedral geometry and two oxygen atoms and so would be expected to be a good mimic of the gem diol intermediate. However, incorporation of the phosphinic acid moiety into an HIV-1 protease substrate gave compound (**2**) (see Figure 1.10), a moderate inhibitor (K_i = 4.5 μM). However, it must be noted that compound (**2**) is unoptimised in the $P_{1'}$ region. The poor inhibition of these compounds may be due to their anionic nature at the assay pH (6.0). The pKa of phosphinic acids[66] generally fall in the range 3.0-3.5 and so at the assay pH would be negatively charged. Literature precedent[67] indicates that a negatively charged species should be repelled by the negatively charged aspartate in the active site. The fact that phosphinic acids are potent inhibitors of pepsin may be due to the fact that pepsin exhibits maximum enzymic activity at a lower pH (< 4.0), at which a larger proportion of the inhibitor would exist in the neutral form. pH Dependence studies of a variety of phosphinic acids show the enzyme's preference for the unionised form. The optimisation of these compounds was undertaken by Grobelny *et al*,[68] who demonstrated that tetrahedral phosphinates incorporated into peptides spanning P_2-$P_{3'}$ function effectively as highly potent inhibitors of HIV-1 protease. As in the reduced amide case, affinity for the enzyme was

sensitive to substitution at the P_1 and P_1' sites with the Phe-Phe isostere providing superior binding to its Phe-Pro counterpart (see compounds (**8**) and (**9**) below). Phe-Pro isosteres were poor inhibitors, confirming earlier reports.[57] The binding affinity of phosphinate inhibitors was also sensitive to pH; a 500-fold increase in affinity was observed for some compounds as the pH was lowered from 6.5 to 4.5, thereby neutralising the charged phosphinate.



(8); Possessing a K_i value of 0.6 nM at pH 4.5.

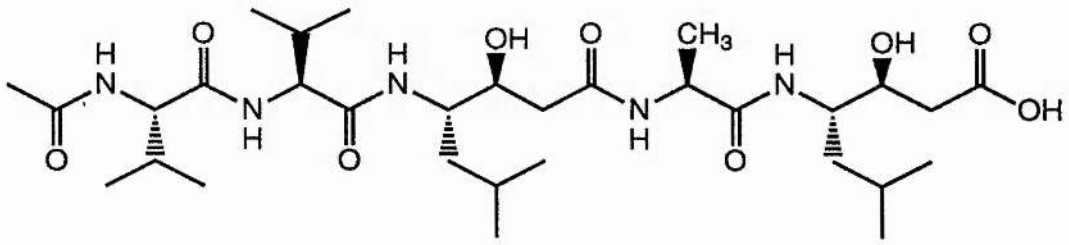


(9); Possessing a K_i value of 450 nM at pH 4.5.

However, this type of pH dependence is not definite. Recent results from Abdel-Meguid *et al.*[69] have shown potent inhibition with symmetrical phosphinic acids at pH 6.0, where the ionised form would exist. This suggests that either the intrinsic K_i is very small (picomolar) or that the ionised form is indeed a potent inhibitor. Their analysis of the pH dependence of K_i failed to unambiguously answer these questions, due to the similarity in pKa values of their inhibitor (pKa 3.1) and the active site aspartate residue (pKa 3.1-3.3) that must be protonated for an inhibitor to bind to HIV-1 protease. However, their crystallographic results suggest a protonated phosphinate in the enzyme complex.[69]

(iii) Statines and norstatines.

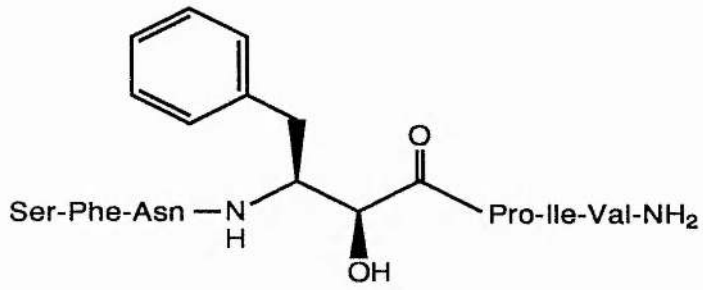
The natural product pepstatin (see Figure 1.4) containing the hydroxymethylene isostere (statine), helped to classify the HIV-1 protease as an aspartic protease. Pepstatin is a highly potent inhibitor of the mammalian aspartic proteases, often giving K_i values in the low nanomolar range. Pepstatin led the way in the search for aspartic protease inhibitors. The tetrahedral hydroxyl bearing carbon is the potent part of the inhibitor and is believed to have a dual role. First, the hydroxyl mimics to some extent the gem diol intermediate and second, it has been proposed[70] that the statine group may act as a 'bisubstrate', in which the binding sites of both the substrate and the nucleophilic water molecule are occupied. Structural analyses of aspartic protease-inhibitor complexes support these theories.[71, 72] Pepstatin A, is only a moderate inhibitor of HIV-1 protease (K_i ~ 1μM). Incorporation of the statine moiety into an HIV-1 protease substrate gave compound (**3**) which possesses a K_i value of 0.81 μM. This is a surprising result considering that the residues in pepstatin are not optimised for the HIV-1 protease and yet pepstatin is approximately equipotent to the substrate analogue. Richards *et al.*[73] reported that acetyl pepstatin, (see below compound (**10**)) is significantly more potent towards the HIV-1 protease (K_i = 20 nM). A seventy fold increase in potency for a simple modification to the P_4 residue is incredible, considering that the acetyl group is far removed from the active site.



(10); Acetyl pepstatin (K_i = 20 nM).

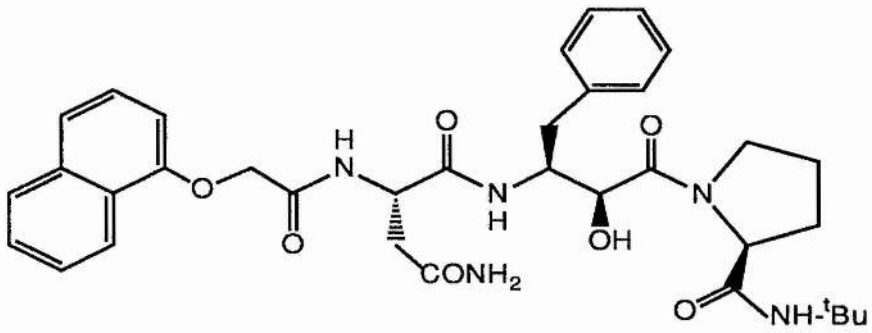
The norstatine based inhibitors incorporate the (hydroxymethyl)carbonyl isostere (see Figure 1.9). Mimoto *et al.*[74] reported potent inhibitors based on the Phe-Pro cleavage site. Their lead compound incorporated the unnatural amino

acid, phenylnorstatine ((2R,3S)-3-amino-2-hydroxy-4-phenylbutyric acid). However, the (2S,3S)-epimer was found to be 20 times more potent, showing a preference for (S)-stereochemistry at the hydroxyl bearing carbon. This allophenylnorstatine derivative (compound (**11**)) gave a K_i value of 5 nM.



(11); Allophenylnorstatine based inhibitor (K_i = 5 nM).

Krantz *et al.*[75] optimised their norstatine based inhibitors and produced their most potent inhibitor, compound (**12**), which gave a K_i value of 0.58 nM. The most active epimer also possesses the (S)-configuration at the hydroxyl bearing carbon.

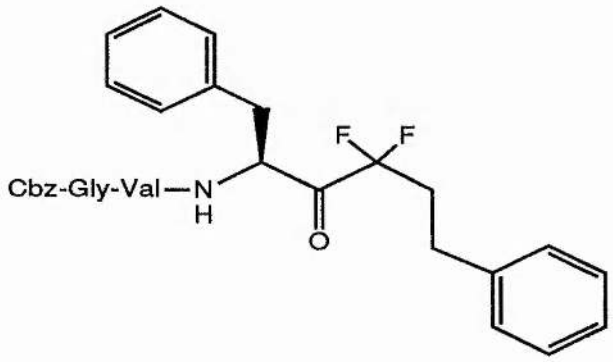


(12); Highly potent norstatine based inhibitor (K_i = 0.58 nM).

Optimisation of the P_3 residue was achieved using a (naphthyloxy)acetyl protecting group on the N-terminus. This type of unnatural amino acid analogue has become increasingly popular in inhibitor design. Often an increase in potency over the natural amino acids can be achieved. Second, these unnatural amino acid residues are often proteolytically stable and are therefore less susceptible to metabolic degradation. In some cases these residues improve solubility as well. Both of these factors are important in drug delivery.

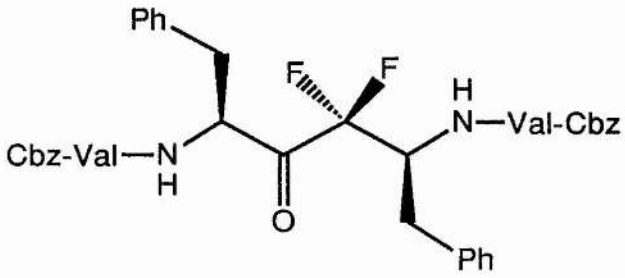
(iv) α-α-Difluoroketones.

α-α-Difluoroketones are potent inhibitors of several classes of enzymes by virtue of their propensity to form hydrates (see Figure 1.9), which converts the sp^2 hybridisation of the carbonyl to sp^3. In principle, these hydrates should be excellent mimics of the enzyme-generated tetrahedral intermediate. This is borne out by the finding that these compounds are very potent inhibitors of pepsin and renin.[76, 77] It was found that incorporation of the difluoroketone moiety into an HIV heptapeptide substrate gave compound (**4**) possessing a K_i value of 160 nM. Sham *et al.*[78] optimised the P_1, position with a phenylethyl substituent. Compound (**13**), although smaller than compound (**4**), gave an IC_{50} value of 1.0 nM.



(13); Possessing an IC_{50} value of 1.0 nM.

An extremely potent difluoroketone inhibitor was synthesised by Sham *et al.*[79] Compound (**14**) below gave a K_i value of 0.1nM. An interesting feature of this compound is that it is almost symmetrical.

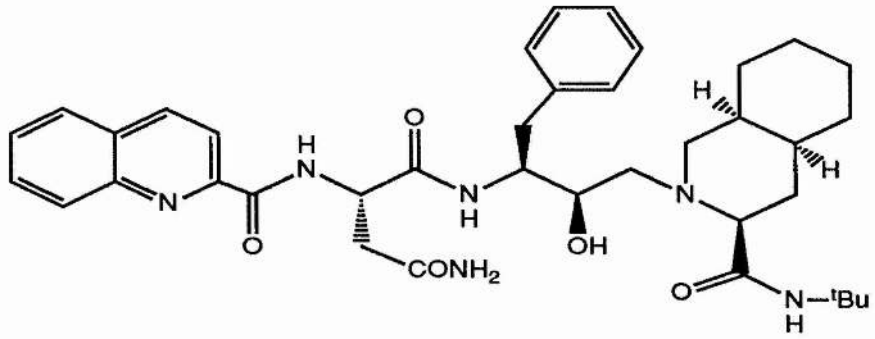


(14); Symmetry-based difluoroketone (K_i = 0.1 nM).

These type of C-2 symmetric inhibitors complement the C-2 homodimeric nature of the HIV-1 protease and represent an important advance in the development of inhibitors for this enzyme. For instance, their selectivity for HIV-1 protease over the momomeric aspartic proteases should be greatly enhanced. (For a fuller discussion of the C-2 symmetric inhibitors see section **1.4.6 (viii)**).

(v) Hydroxyethylamines.

The hydroxyethylamine moiety has produced some of the most potent HIV-1 protease inhibitors to date. Unlike most of the previously mentioned transition state analogues (with the exception of the reduced amides), the hydroxyethylamine group accommodates the incorporation of proline or its analogues into the inhibitor, thus allowing the scissile Phe-Pro bond to be replaced more exactly. However, the mimicry is not exact, as an additional methylene group is required for stability. Roberts and co-workers[80] (Roche, UK) reported the following compound (**15**) (Ro 31-8959) as an extremely potent inhibitor of HIV-1 protease with a K_i value of 0.12 nM. The preferred absolute configuration at the hydroxyl bearing carbon is (R). The norstatine based inhibitors (see compound (**12**)) show an identical preference for the relative stereochemistry of their hydroxyls, although, due to priority differences, their hydroxyls are assigned the (S)-configuration. The decahydroisoquinoline moiety contributes considerably to the potency of this compound, the corresponding derivative containing a prolyl group in this position is more than 50-fold less potent. This trend is reversed in the case of the norstatines[75] and so the conformation of the prolyl residue may be an important recognition element for the enzyme. The extra CH_2 in the hydroxyethylamine inhibitors will obviously alter the conformation of the prolyl residue, whereas the amide prolyl in the norstatine inhibitors is closer in representation to the prolyl residue in a Phe-Pro substrate.



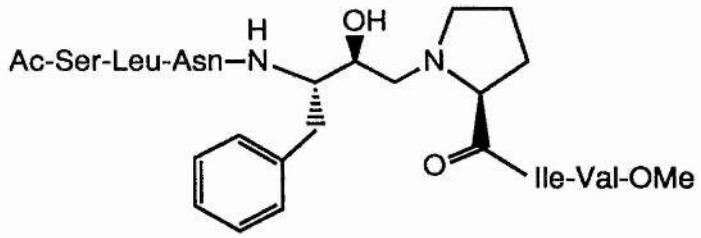
(15); Ro 31-8959, a highly potent inhibitor (K_i = 0.12 nM).

Roche were the first to incorporate unnatural amino acid residues into their design for HIV-1 protease inhibitors. Ro 31-8959 incorporates a quinoline 2-carbonyl protecting group as a P_3 substituent and the aforementioned decahydroisoquinoline moiety. Both of these residues contribute to the potency of this inhibitor, which is still amongst the most potent inhibitors of viral infectivity and is currently being evaluated in a clinical setting. Many of the recent inhibitors have been derived from this leading compound.

In contrast to the preference for the (R)-configuration for the hydroxyl bearing carbon in compound (**15**), Rich *et al.*[81] found a preference for the (S)-epimer in a series of hydroxyethylamine inhibitors. In compound (**16**), the (S)-epimer (K_i = 0.24 nM) was found to be 80-fold more potent than the (R)-epimer. However, it was found that when inhibitors lacked the P_3' substituent the (R)-epimer was more active than the (S)-epimer:

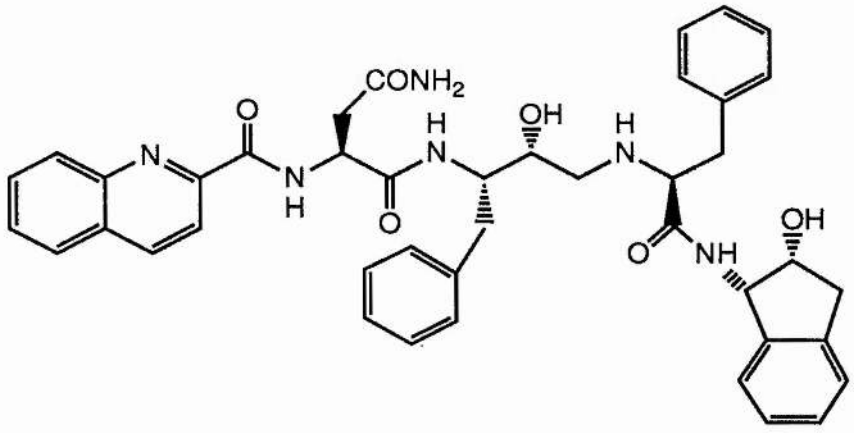
	P_1' P_2'		
Cbz-Asn-Phe-ψ[CH(OH)CH_2N]-Pro-O tBu.	IC_{50} = 450 nM	(S)-hydroxyl
Cbz-Asn-Phe-ψ[CH(OH)CH_2N]-Pro-O ^{t}Bu.	IC_{50} = 51 nM	(R)-hydroxyl

These results show that the substituents in the P_4-P_3' positions have a major influence on the binding of hydroxyethylamine inhibitors. There is a shift in preference from the (R)-epimer in smaller inhibitors to the (S)-epimer in longer inhibitors.



(16); (S)-Epimer (K_i = 0.24 nM).

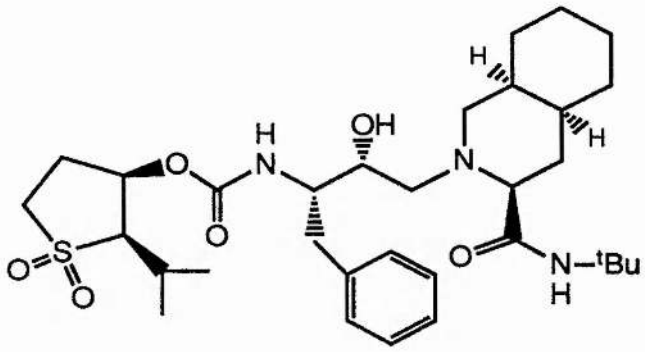
Merck (USA) have made an outstanding contribution to the development of HIV-1 protease inhibitors. Tucker *et al.*[82] reported an optimised hydroxyethylamine inhibitor based on the Phe-Phe cleavage site. Compound (**17**) (K_i value of 5 nM) incorporates the quinoline 2-carbonyl P_3 substituent and a chiral aminohydroxyindanol moiety as a $P_{2'}$ substituent in an effort to increase potency.



(17); Phe-Phe based hydoxyethylamine inhibitor (K_i = 5 nM).

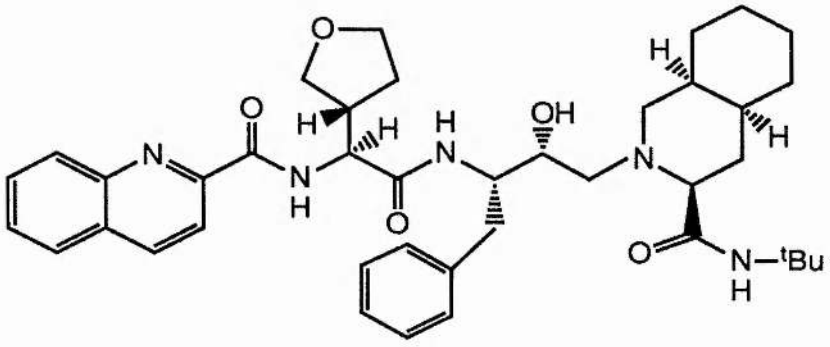
In an effort to develop the HIV-1 protease inhibitors as potential drugs, researchers have begun the search for lower molecular weight, non-peptidic inhibitors (see section **1.4.7** for a discussion on the problems associated with peptide-based drugs). With this goal in mind, medicinal chemists have begun the search for ligands that show comparable potencies to their peptide-based counterparts. Two recent reports highlight the progress already being made in this area. Ghosh *et al.*[83] have replaced the P_3 and P_2 residues in the Ro 31-8959 inhibitor by a single cyclic sulfolane residue. The conformationally

restricted *cis*-2-isopropylsulfolane ligand in compound (**18**) succeeds in producing a highly potent inhibitor (K_i value of 3 nM).



(18); Possessing a sulfolane ligand (K_i = 3 nM).

One explanation for the observed potency is that the sulfolane oxygens are making specific interactions in the S_2 binding domain, in a similar fashion to the hydrogen-bonding normally observed between the primary amide oxygen of the asparagine side chain and the NH residues of Asp-29 and Asp-30 in the S_2 subsite. Thompson *et al.*[84] (Merck, USA) replaced the P_2 Asn residue in Ro 31-8959 with a chiral tetrahydrofuranylglycine (Thfg) ligand. An IC_{50} value of 0.054 nM was achieved for compound (**19**), the inhibitor derived from the (2S,3'R) ligand, showing four times the potency of Ro 31-8959.



(19); Incorporating the (3'R)-Thfg ligand (IC_{50} = 0.054 nM).

In contrast, the inhibitor incorporating the (2S,3'S) Thfg ligand was 100-fold less potent than compound (**19**). This strong stereochemical preference supports the aforementioned hydrogen-bonding phenomenon as an important

factor for ligand binding in the S_2 subsite. A recent structure activity study of inhibitors incorporating these novel P_2 ligands as well as a novel pyramizine amide P_3 ligand has produced a number of subnanomolar inhibitors.[85]

(vi) Hydroxyethylenes.

To date, the hydroxyethylene isostere has been the most extensively utilised moiety in the design of HIV-1 protease inhibitors. Incorporation of the hydroxyethylene isostere into a heptapeptide analogue by Dreyer,[57] produced compound (**5**), the most potent inhibitor in the series (K_i value of 62 nM). Unlike statine based analogues, compound (**5**) and the hydroxyethylene series in general, preserve the atomic spacing of the dipeptide backbone. Compound (**5**) is an isosteric replacement for the dipeptide Phe-Gly. The stereochemical configuration of the 4-hydroxyl group is crucial for potent inhibition with the (4S)-epimer of compound (**5**) being 80 times more potent than the (4R)-epimer. Removal of the N-terminal in compound (**5**), led to an increase in affinity with Ala-Ala-Phe-ψ[CH(OH)CH_2]-Gly-Val-Val-OMe possessing a K_i value of 18 nM. Considering that Phe-Gly is not a characteristic cleavage site for the HIV-1 protease, these compounds provide promising leads. One drawback for this series is that exact mimics of the Phe-Pro cleavage sites cannot be achieved. Incorporation of a proline residue at P_1, would produce an amino-hemiacetal which would probably decompose in solution to produce an enamine (see below, Figures 1.11 and 1.12.).

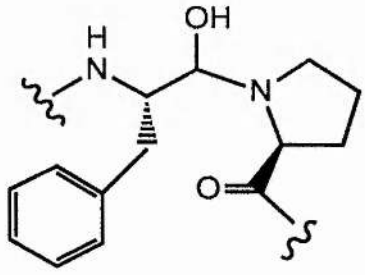


Figure 1.11. Amino-hemiacetal.

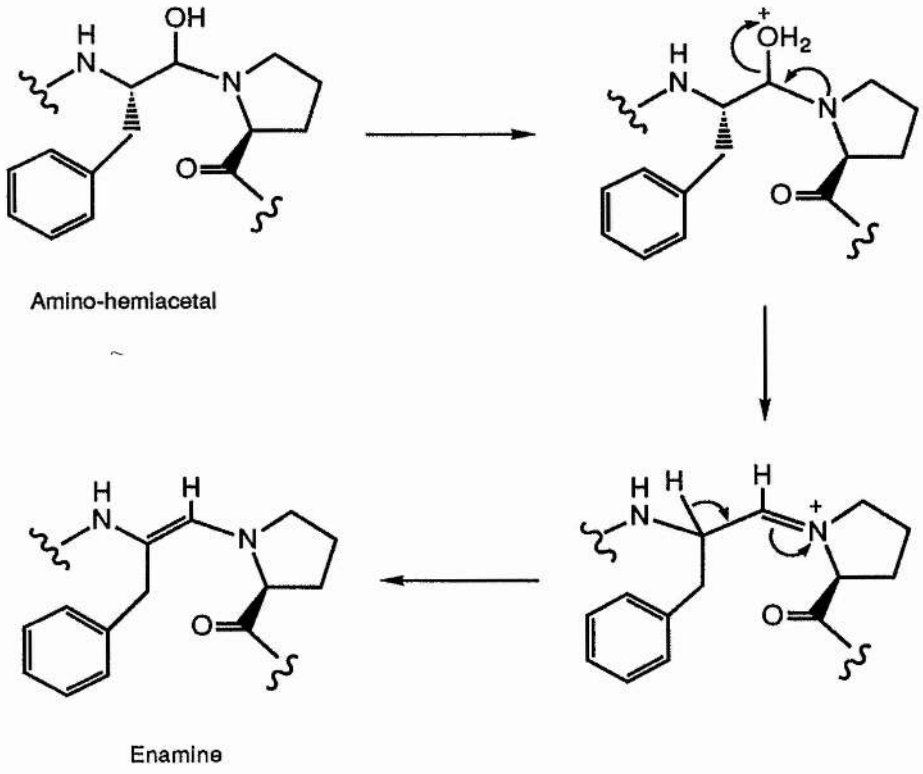
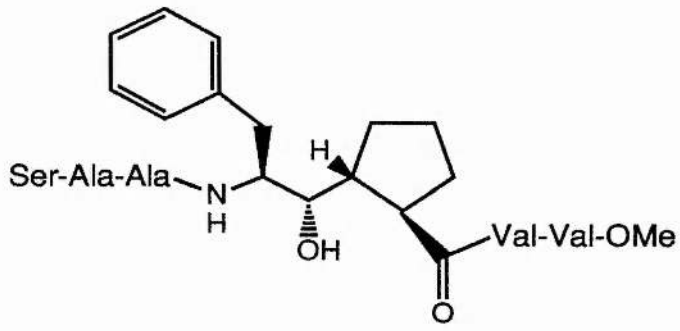


Figure 1.12. Probable fate of an amino-hemiacetal in solution.

However, Dreyer *et al.*[57] used a *trans*-cyclopentyl ring as a proline mimic in the P_1, position. Compound (**20**) was found to be a moderate inhibitor of HIV-1 protease (K_i = 500 nM).

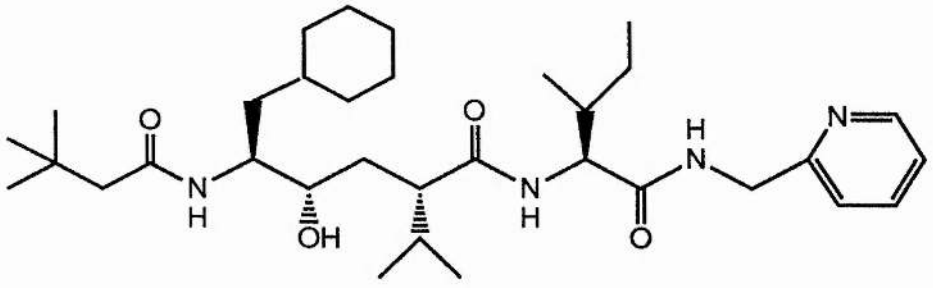


(20); Phe-Pro mimic (K_i = 500 nM).

The synthesis of two similar analogues based on Leu-Pro (Rich *et al.*)[86] and Tyr-Pro (Thompson *et al.*)[87] have also been reported with the latter analogue, possessing an IC_{50} value of 110 nM. No biological data was given for the Leu-

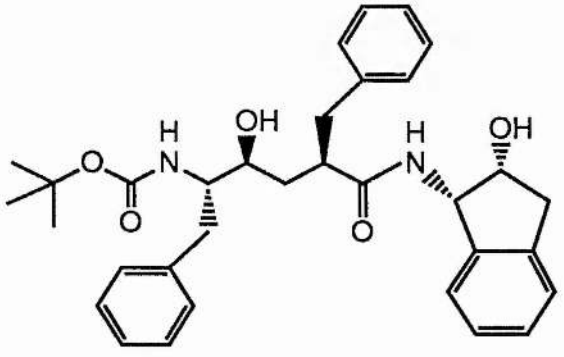
Pro analogue.

McQuade *et al.*[88] (Upjohn, USA) reported the pentapeptide inhibitor, compound (**21**) (K_i value of 70 nM), which incorporated the cyclohexylalanine-Val isostere. Cyclohexylalanine (Cha) residues are good replacements of Phe at P_1 and our own inhibitors have successfully incorporated the Cha residue in this position.



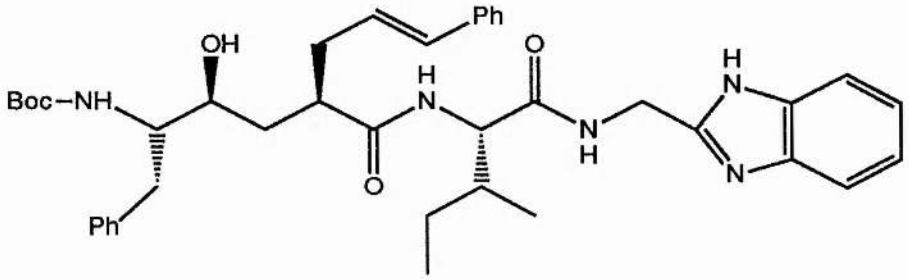
(21); Cha-Val hydroxyethylene inhibitor (K_i = 70 nM).

Aminomethylpyridine is a popular residue in the $P_{3'}$ position, providing a proteolytically stable C-terminus. Other optimised inhibitors include those based on the Phe-Phe cleavage site. Lyle *et al.*[89] (Merck, USA) developed the 1-amino-2-hydroxyindan (previously shown in compound (**17**)) ligand as an effective surrogate for the $P_{2'}$ residue and produced subnanomolar inhibitors that span only the P_2-$P_{2'}$ subsites. The hydroxyl group in the indan appears to replace the $P_{2'}$ oxygen carbonyl in a critical hydrogen-bond to the enzyme. Compound (**22**) possesses an IC_{50} value of 0.3 nM.



(22); Highly potent pentapeptide inhibitor (IC_{50} = 0.3 nM).

Vacca *et al.*[90] (Merck, USA) produced compound (**23**) which incorporates a styrene residue at $P_{1'}$ and an aminobenzimidazole moiety at $P_{2'}$. This compound is one of the most potent inhibitors of HIV-1 protease to date, giving an IC_{50} value of 0.03 nM.

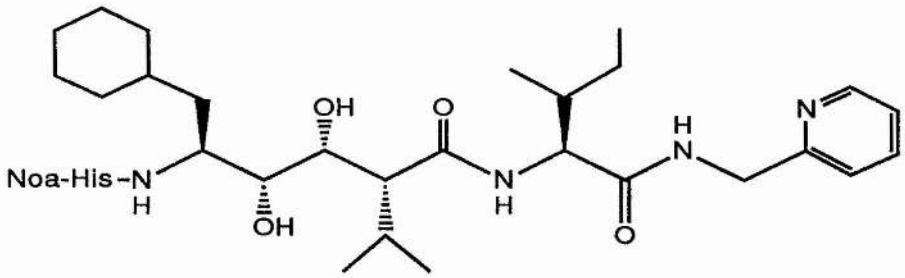


(23); The most potent inhibitor of HIV-1 protease to date.

These latter examples demonstrate that small, but potent, HIV-1 protease inhibitors of minimal peptide character are achievable. A significant amount of structure activity relationship data has now been compiled and many optimised hydroxyethylene inhibitors now exist. [91, 92, 93, 94,]

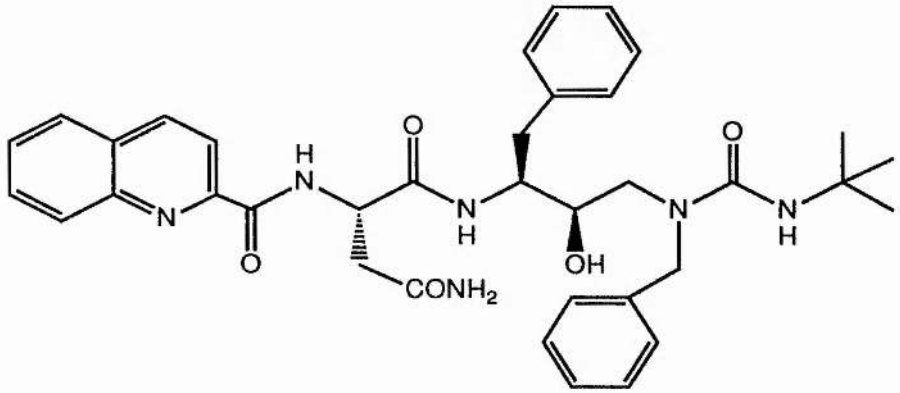
(vii) Modified hydroxyethylenes.

There have been several reports of inhibitors containing modified hydroxyethylene isosteres. Thaisrivongs and Ashorn *et al.*[95, 96,] (Upjohn, USA) have incorporated the dihydroxyethylene isostere (see Figure 1.9) into a series of inhibitors, of which compound (**24**), a Cha-Val analogue was the most potent in the series (K_i value < 1nM). Compound (**24**) also contains the naphthoxyacetyl (Noa) protecting group, which contributes to the potency of this inhibitor. More recently, an extensive structure activity study has been reported by the same researchers.[97]



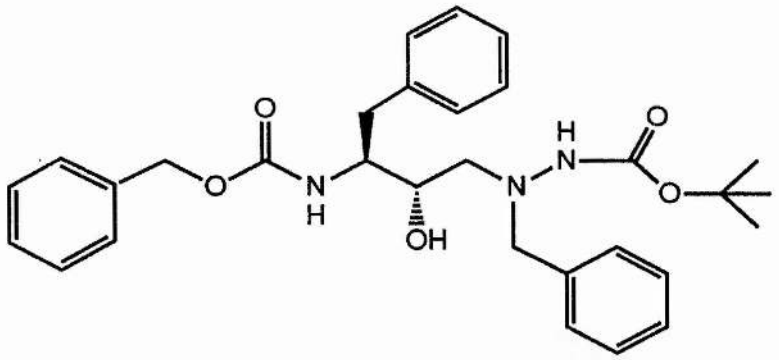
(24); Dihydroxyethylene inhibitor (K_i value < 1 nM).

Getman *et al.*[98] (Monsanto, USA) incorporated a hydroxyethylurea isostere into a series of compounds (see Figure 1.9). Compound (**25**) was found to be the most potent inhibitor in the series, possessing an IC_{50} value of 3 nM. This inhibitor has similar features to Ro 31-8959 and shows the same preference for the (R)-configuration at the hydroxyl bearing carbon.



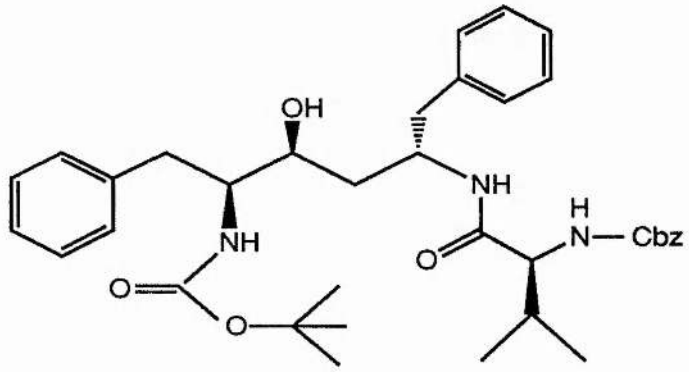
(25); Hydroxyethylurea inhibitor (IC_{50} = 3 nM).

Sham *et al.*[99] (Abbott, USA) have incorporated novel substituted hydrazines into a series of compounds and the Phe-Phe mimic, compound (**26**) produced an IC_{50} value of 5 nM.



(26); Incorporating a substituted hydrazine (IC_{50} = 5 nM).

Finally, Ghosh *et al.*[100] (Merck, USA) reported the stereoselective synthesis of a novel dipeptide mimic and its conversion into a potent HIV-1 protease inhibitor. Compound (**27**) possesses an IC_{50} value of 3 nM and shows a preference for the (S)-configuration at the hydroxyl bearing carbon.



(27); Possessing a novel dipeptide isostere (IC_{50} = 3 nM).

These last three examples show the very subtle changes being made by medicinal chemists in order to transform such molecules into prospective therapeutic agents.

(viii) C-2 Symmetric inhibitors.

The high degree of symmetry of the HIV-1 protease prompted Kempf and co-workers[101, 102] (Abbott, USA) in the design of C-2 symmetric inhibitors. They postulated that inhibitors based on symmetry might provide improvements over

traditional substrate-based inhibitors in terms of potency, specificity and reduced peptide character. For instance, symmetrical inhibitors should show selectivity for retroviral proteases as the active site of the mammalian aspartic proteases are less symmetric. The design of these inhibitors was achieved *via* an examination of the tetrahedral intermediate of peptide cleavage (see below).

(i) Placement of a C-2 axis of symmetry through the tetrahedral gem diol carbon led to the core structure (**A**).
(ii) Dissection of the scissile amide bond with a C-2 axis provided the diol (**B**).
(iii) A two fold rotation of the remaining N-terminus provides core structures (**A**) and (**B**) respectively (see below, Figure 1.13).

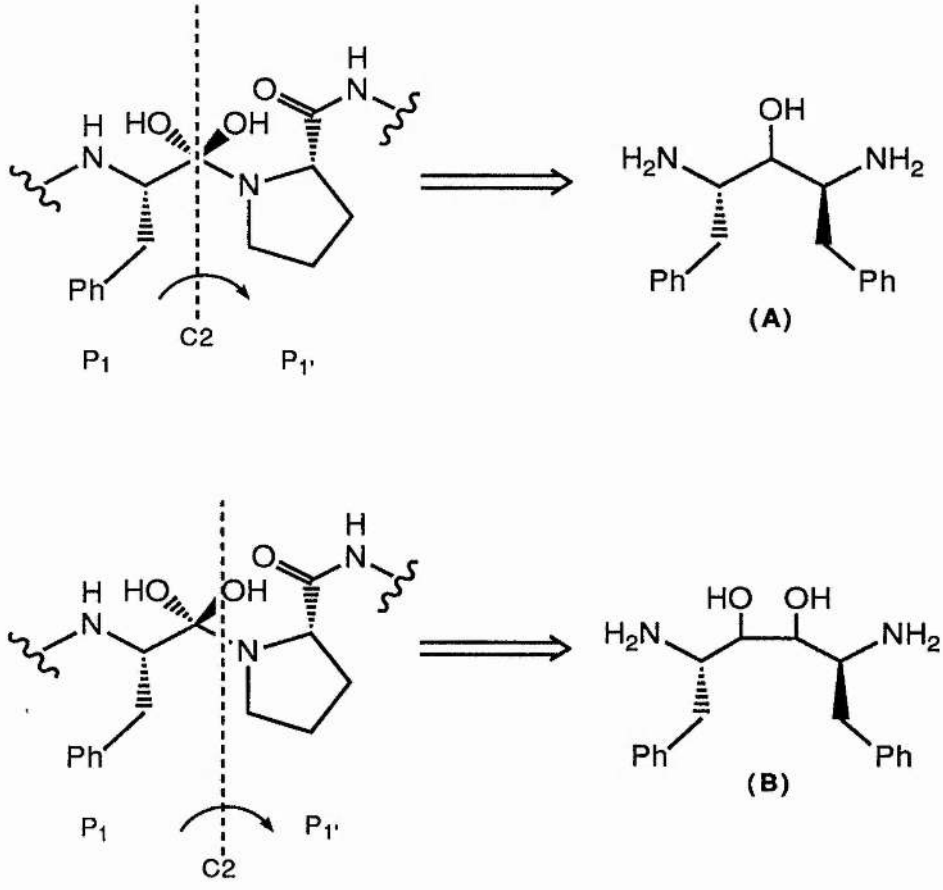
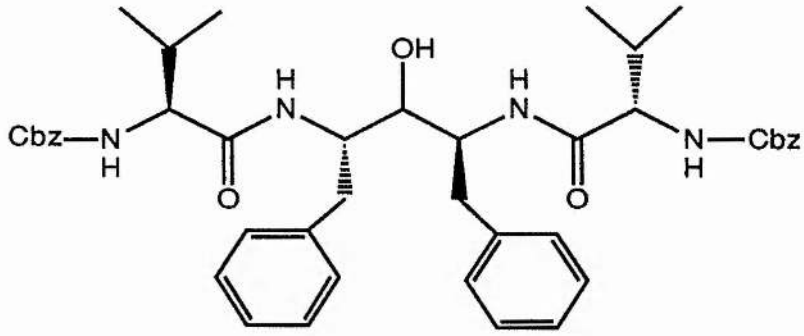


Figure 1.13. Design of C-2 symmetric inhibitors.

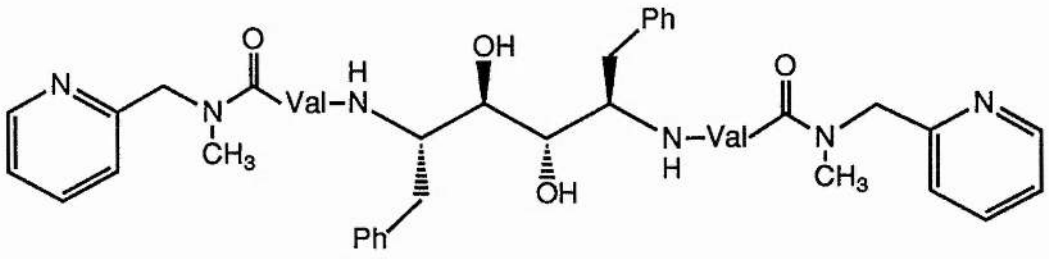
Extension of these core structures with suitable peptide residues resulted in nanomolar or subnanomolar inhibition. Compound (**28**) was derived from the

core structure (**A**) and possesses an IC_{50} value of 3 nM.



(28); C-2 Symmetric inhibitor (IC_{50} = 3 nM).

An identical Cbz-Val extension of core structure (**B**) produced a compound possessing an IC_{50} value of 0.22 nM. Interestingly, this diol was not highly sensitive to the stereochemistry of the hydroxyls with the (3R,4R)- and (3R,4S)-epimers being equipotent. Recently, both classes of inhibitor have been the subject of an extensive structure activity study and the same researchers have produced a large number of subnanomolar inhibitors.[103] Compound (**29**) (A-77003) provided the best balance of activity and solubility and is subsequently under clinical development (IC_{50} < 1 nM).



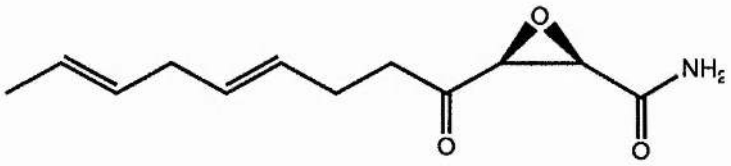
(29); (A-77003) C-2 Symmetric diol (IC_{50} < 1 nM).

The work of Kempf and colleagues has prompted the synthesis of C-2 symmetric inhibitors of a variety of other transition state analogues. These include phosphinic acids,[104, 69] difluoroketones,[79] sulfides, sulfoxides and sulfones [105] and dihydroxyethylenes. Several elegant syntheses of symmetrical dihydroxyethylenes have been reported starting from D-mannitol.[106, 107] However, no inhibition data has been reported for this symmetrical class of

inhibitor so far.

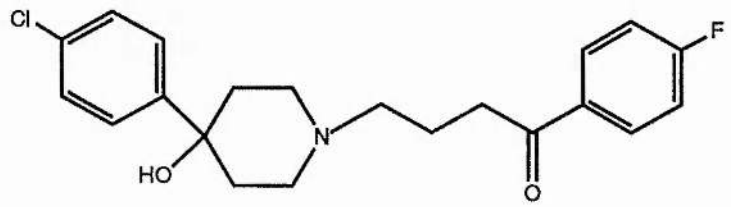
(ix) Non-peptide based inhibitors.

To date, very few non-peptide inhibitors of HIV-1 protease have been reported. The antifungal antibiotic cerulenin, compound (**30**) was found to be a weak inhibitor of HIV-1 protease (IC_{50} = 2.5 mM).[108]



(30); Antibiotic cerulenin (IC_{50} = 2.5 mM).

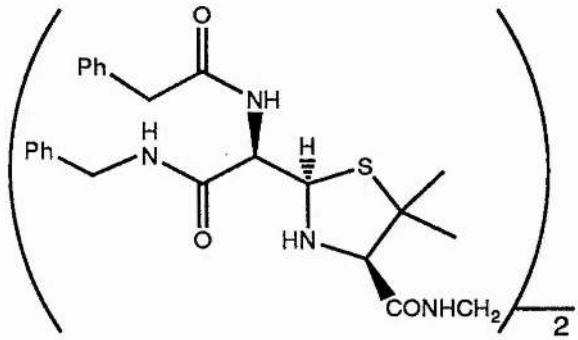
The chemical mechanism of inactivation of HIV-1 protease by cerulenin has not yet been elucidated, although a mechanism involving esterification of one of the catalytic aspartates by reaction with the epoxide is a possibility. Desjarlais *et al.*[109] have used a structure based, molecular modelling approach to discover novel inhibitors of HIV-1 protease. Compounds are docked into the active site of the enzyme and their steric complementarity with the three-dimensional structure of the active site is assessed. This method identified haloperidol, compound (**31**) as a weak inhibitor of the enzyme (K_i value of 100 μM). This technique should prove very useful in uncovering novel 3D pharmacophores from which inhibitors can be designed and synthesised.



(31); Haloperidol identified from a crystallographic data base (K_i = 100 μM).

A penicillin derived symmetrical dimer has been reported recently by

Humber and co-workers.[110] As an alternative to peptide-based inhibitors they embarked on a screening program in order to identify novel inhibitors for the HIV-1 protease. A crude sample of a penicillin dimer was shown to be active. However, when the sample was purified, the main component was inactive. The active component was shown to be a diester which was derived from the opening of the β-lactam ring of the penicillin dimer. By reacting the dimer with appropriate amines a series of active compounds were produced of which compound (**32**) was the most active (IC_{50} value of 0.9 nM).



(32); A penicillin derived symmetrical dimer (IC_{50} = 0.9 nM).

Various other non-peptide inhibitors of the HIV-1 protease have been reported including carboxylates,[111] didemnaketals,[112] flavones,[113] boronated porphyrins [114] and most recently even fullerene derivatives![115]

1.4.7 Effects of inhibitors against HIV-1 infected cell culture.

The K_i and IC_{50} values of all of the inhibitors mentioned so far are indicative of their effect against the HIV-1 protease in an *in vitro* assay. Therefore factors such as cell uptake and metabolic degradation do not need to be considered and these values provide an indication of affinity for the enzyme. These initial assays can identify potential problems such as aqueous insolubility. A more realistic guide to an inhibitor's potential as a therapeutic agent is provided in an *in vivo* or viral assay. In the case of the HIV-1 protease, the inhibitor is tested in two differing ways. First, the direct observation of inhibition of HIV-1 protease in chronically-infected T-lymphocytes is realised by treatment of the infected cell culture with inhibitors for a 24-48 hour period prior

to examination of the extent of processing of the viral polyproteins. The viral polyproteins are isolated by immunoprecipitation and the extent of polyprotein processing is determined by gel electrophoresis. The second type of study involves acute infection of virus-naive T-lymphocytes. Inhibitors and infected virus are added simultaneously and after several days, the extent of viral infectivity is determined by the quantification of three parameters: p24 *gag* antigen and reverse transcriptase activity released into the medium and the population of syncytia, the multinuclear giant T-lymphocytes that form during the infection period.

Both the chronic and acute infection assays were conducted on the hydroxyethylene inhibitors. Compound (**5**) and similar hexa and heptapeptide analogues were found to block $Pr55^{gag}$ and $Pr160^{gag\text{-}pol}$ processing in chronically-infected T-lymphocyte cultures and prevented the spread of HIV infection in lymphoid cell lines at 25-100 μM.[116] It should be noted that these values are representative of a particular strain of virus and the same inhibitors may have differing activities against other strains and so direct comparison of antiviral data is difficult. However, these antiviral results show significantly higher IC_{50} values than the *in vitro* results and indicate poor cell uptake, poor solubility or metabolic degradation of the inhibitors. As mentioned earlier, these problems are of particular concern for peptidic inhibitors. In fact, peptidic inhibitors longer than hexapeptides are generally inactive when tested against infected cells. For instance, compound (**16**), an octapeptide analogue (K_i = 0.24 nM) was found to be inactive at concentrations >10 μM i.e 40,000 times the K_i value.

Physical evidence for the effect of inhibitors comes in the form of electron microscopy. This technique has revealed that the virion particles produced by the inhibitor-treated lymphocytes contained apparently defective core structures; virions possessing "crescent-shaped" cores were preponderant while virions containing mature (see Figure 1.1), bullet-shaped cores represented smaller populations.[117, 118]

Inhibitors with decreasing peptide character have shown increasing antiviral activity. Compound (**21**) (K_i = 70 nM) was found to have an IC_{70} value of 1 μM in an antiviral assay.[88] The dihydroxyethylene inhibitor, compound (**24**) (Ki < 1 nM) gave an IC_{100} of 1 μM in a similar assay and showed an equivalent potency

to AZT.[95] Compound (**23**) (K_i = 0.03 nM) and compound (**22**) (K_i = 3 nM) gave IC_{100} values of 12 nM[90] and 200 nM[89] respectively, highlighting increasing antiviral activity for smaller, less peptidic inhibitors. The C-2 symmetric inhibitors, compounds (**27**) and the equivalent diol analogue produced IC_{50} values of 400 nM[101] and 20 nM[102] respectively. The hydroxyethylamine inhibitor, compound (**15**) Ro 31-8959, is one of the most potent inhibitors of antiviral infectivity to date (IC_{50} = 2 nM) and is currently being evaluated in a clinical setting. The remaining task is to convert potent antiviral inhibitors into safe, effective drugs for human administration.

1.4.8 Structural aspects of HIV-1 protease and its enzyme-inhibitor complexes.

Determination of the crystal structure of native HIV-1 protease confirmed the homodimeric architecture of the enzyme, with each monomer related to each other by a 2-fold rotation. The protein contains a β-sheet structure, similar to that found in mammalian and fungal proteases.[119] The active site triad Asp-25-Thr-26-Gly-27 occurs within a loop that forms a number of hydrogen-bonds with the corresponding loop of the other monomer in a fashion often described as the "fireman's grip".[120] The "fireman's grip" is a characteristic of the aspartic proteases and its effect within pepsin is shown below (Figure 1.14). The HIV-1 protease also contains two identical "flap" regions (see overleaf, Figure 1.15) which move as much as 7 Å upon binding of an inhibitor.[63]

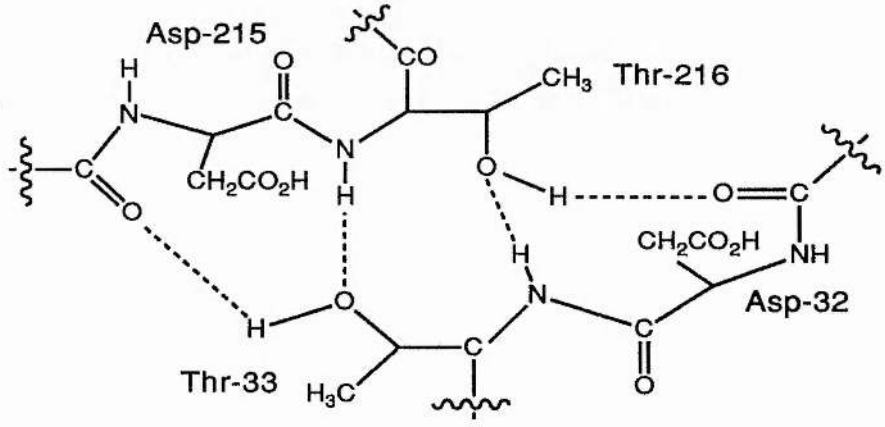


Figure 1.14. The "fireman's grip" in pepsin.

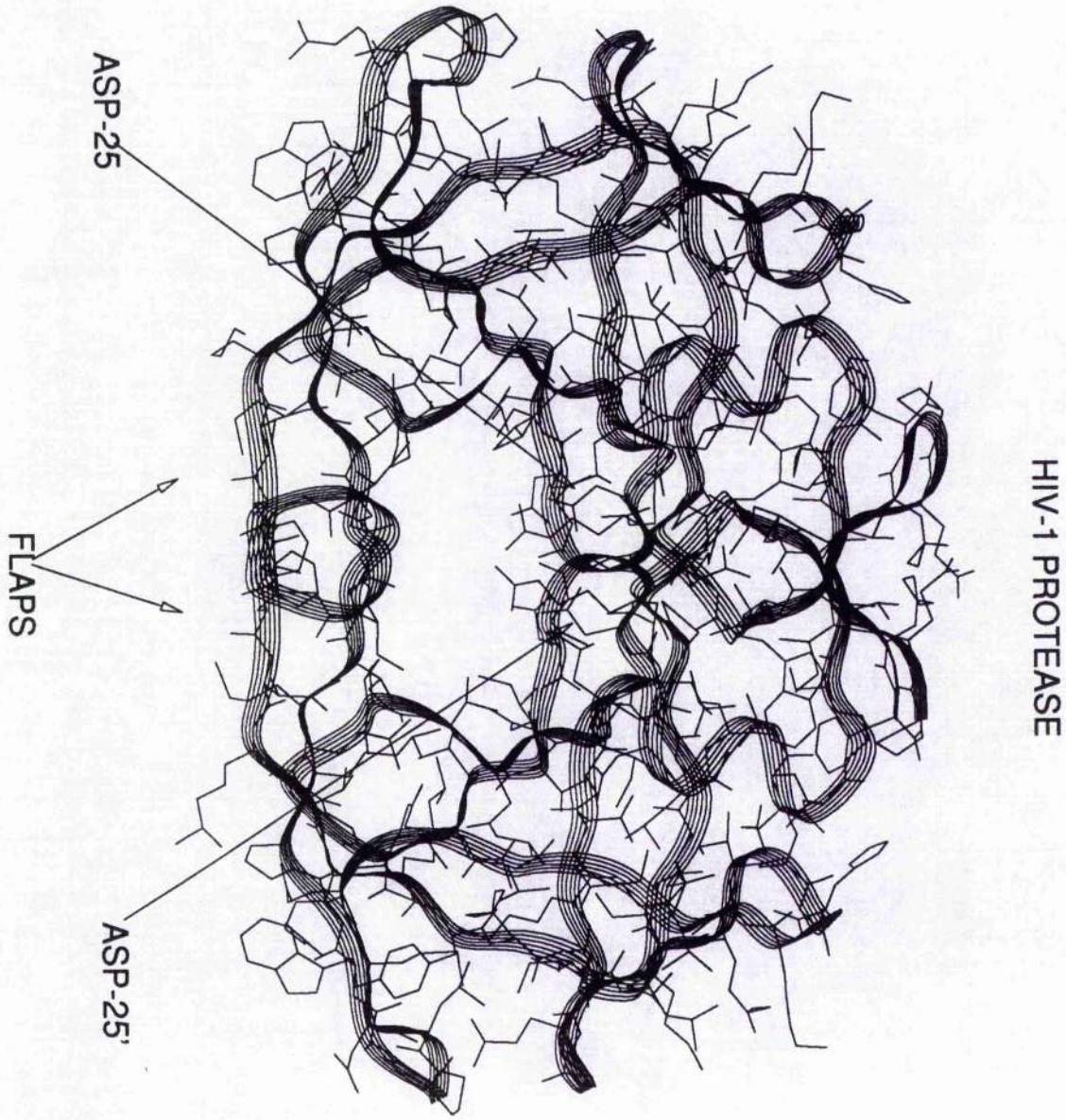


Figure 1.15. Overall fold of liganded HIV-1 protease.

The crystal structures of enzyme-inhibitor complexes have provided vital information to medicinal chemists involved in inhibitor design. Structures of the HIV-1 protease complexed with four inhibitors that differ in terms of structure and potency have been reported. These include acyl pepstatin (**10**),[121] the reduced amide (**6**),[63] the hydroxyethylamine (**16**)[122] and the C-2 symmetric inhibitor (**28**).[102] These inhibitor structures share several common attributes. Compounds (**10**), (**16**) and (**28**) all possess a secondary alcohol group and in all structures this hydroxyl group is positioned equally among the four oxygen atoms of the catalytic aspartates. This suggests that other substituents of these inhibitors may be of secondary importance and may explain the lower potency observed for classes of inhibitor which do not contain such a hydroxyl or hydrogen-bond donating species. One striking feature of all four inhibitor complexes is a tightly bound water molecule that bridges the two enzyme flaps to the inhibitor through hydrogen-bonds formed by the Ile-50/ Ile-50' amide hydrogens and the P_2 and $P_{1'}$ carbonyl oxygens of the inhibitor. Other crucial hydrogen-bonds include the P_2 / P_1 and $P_{1'}$ / $P_{2'}$ amide hydrogens of the inhibitors to the carbonyl groups of Gly-27 and Gly-27'. All of these important hydrogen-bonding interactions are shown schematically in Figure 1.16.

The presence of the flap water molecule and the symmetrical nature of the enzyme provide features which can be exploited in selective inhibitor design. Swain *et al.*[123] suggested that it should be possible to develop such inhibitors which incorporate these structural details. For instance, compounds which interact with or dislodge the water molecule should provide highly potent, specific inhibitors for the enzyme. With this objective in mind, we are currently investigating two types of novel non-peptidic inhibitors, which will be discussed in section **2.6.2**.

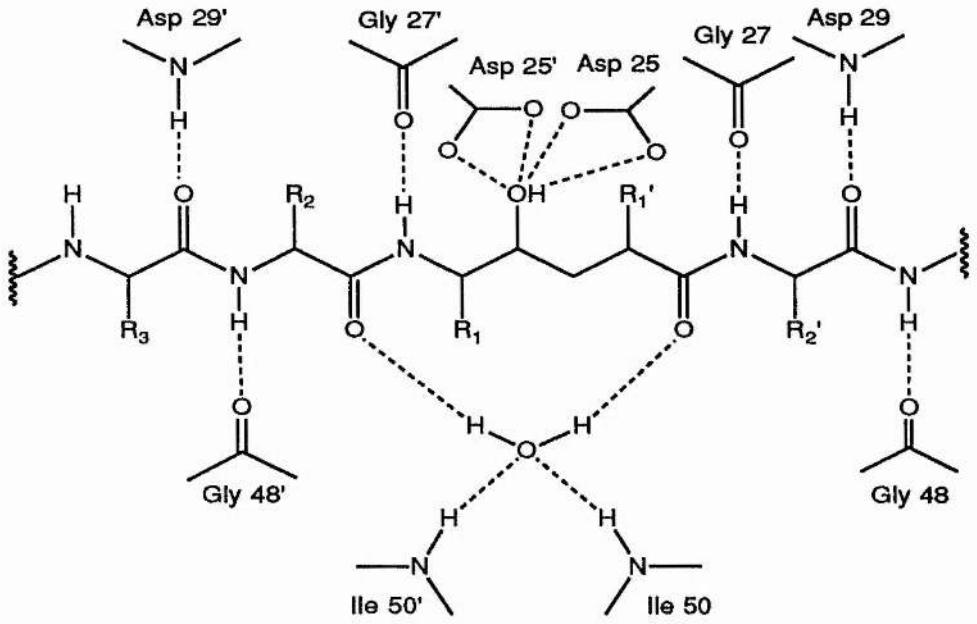


Figure 1.16. Key hydrogen-bonding interactions between the HIV-1 protease and a hydroxyethylene inhibitor.

1.4.9 Summary.

The development of inhibitors for the HIV-1 protease has escalated over the last few years. Armed with the known substrate sequences of the HIV-1 protease and the results from renin inhibition, medicinal chemists rapidly synthesised a range of classical transition state analogues, exemplified by compounds (**1-11**). These compounds provided the hydroxyethylene, hydroxyethylamine and the dihydroxyethylene isosteres as potent inhibitors for the enzyme. However, as the lesson of renin inhibition had shown, the use of peptidic analogues as drugs was problematic.[123, 124] First, these compounds are generally too large and are therefore prone to rapid elimination from the body by hepato-biliary extraction. Due to their abundance of peptide bonds, these inhibitors are prone to metabolic degradation by a host of the body's own proteolytic enzymes and so have a short duration of action. These compounds show poor aqueous solubility which contributes to their low oral bioavailability. Therefore, medicinal chemists are faced with the challenge of transforming these compounds into therapeutically useful agents. Researchers have already begun the task of replacing the peptide residues with low molecular weight ligands of equal or greater potency. Fortunately, the HIV-1 protease is

accommodating to a wide variety of residues and so this provides more flexibility in ligand design. Molecular modelling and the study of enzyme-inhibitor complexes has enabled the discovery of a number of novel ligands, some of which have been extremely successful in producing highly potent, small, minimal-peptide inhibitors. Compounds **17,18,19,22** & **23** are perfect examples. Compounds **25-27** highlight the subtle changes being made in order to improve the pharmacokinetics of these inhibitors. It remains to be seen if these changes are successful.

The work of Kempf and colleagues has provided an alternative direction in the design of novel inhibitors for the HIV-1 protease. These inhibitors, designed from the unique architecture of the enzyme offer selectivity over the mammalian aspartic proteases. Optimisation of these compounds has produced inhibitors with increasing activity towards viral infectivity.

From searches of crystallographic data banks, several non-peptidic inhibitors of the HIV-1 protease have been discovered. At present, this technique has only unmasked weak inhibitors for the enzyme, although the discovery of more potent leads is imminent. An alternative to random searching is one of computer aided design. With the key features for inhibitor potency being a preference for a hydroxyl bearing carbon to interact with the catalytic aspartates and a liking for Phe mimics at both P_1 and $P_{1'}$, novel non-peptide structures can be studied which incorporate these functionalities. Used in conjunction with the X-ray data from enzyme-inhibitor complexes, the important hydrogen-bonding interactions can be maintained and optimised.

CHAPTER TWO

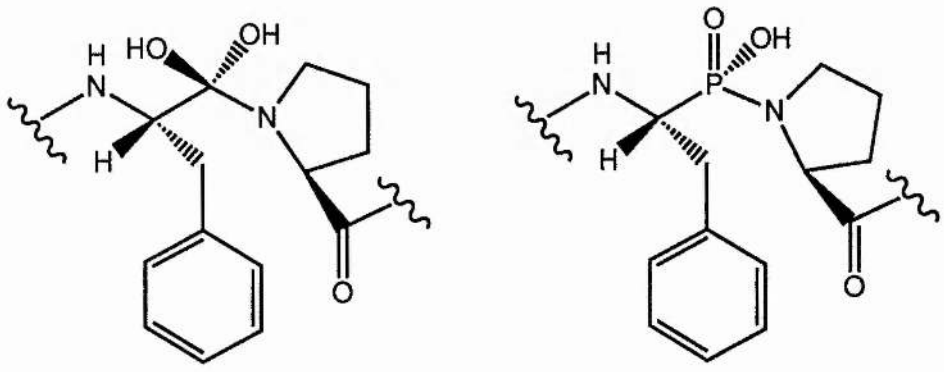
RESULTS AND DISCUSSION

2.0 Results and discussion.

2.1 Design of HIV-1 protease inhibitors.

In a continuing search for therapies for AIDS the HIV-1 protease, a member of the aspartic protease family, has been identified as a promising target. Protease activity is vital for maturation of the virus and several approaches have been used to produce inhibitors, based upon substitution of the scissile amide bond of HIV-1 protease substrates with non-cleavable analogues.[19]

We chose to employ the phosphonamidate moiety as a transition state mimic due to its very close electronic and geometric resemblance to the tetrahedral transition state/ high energy stable intermediate for peptide bond cleavage, see below (Figure 2.1).



Tetrahedral intermediate Phosphonamidate

Figure 2.1.

The phosphonamidate moiety fulfils many of the transition state requirements. For instance, this moiety provides the correct tetrahedral geometry, whilst providing both oxygen atoms of the gem diol tetrahedral intermediate. Phosphorus-oxygen single and double bonds are approximately 10-15% longer than the corresponding carbon-oxygen bonds.[125] This allows the P=O double bond in the phosphonamidate to mimic the lengthening of the carbonyl bond as it is attacked by the lytic water molecule. In contrast to many other transition state mimics, the phosphonamidate moiety also possesses a nitrogen atom at P_1, in the correct position. Potentially this could be important in

maintaining the hydrogen-bonding interactions in the enzyme's active site. In addition to potential as inhibitors, phosphonamidate analogues are suitable for probing the mechanism of the protease. For instance, phosphorus containing transition state mimics allow the use of ^{31}P NMR spectroscopy to study inhibitor-enzyme interactions in detail. However, in the short term we have only investigated these compounds as HIV-1 protease inhibitors.

As the pKa value for the phosphonamidate is approximately 3.0,[126] at the pH optimum for the HIV-1 protease (pH 5.5) it will exist predominantly as the anion. Thus, the phosphonamidate would not be a good mimic for a neutral gem-diol tetrahedral intermediate. However, as the exact nature of the stable intermediate in amide hydrolysis by the aspartic proteases is unknown, we decided to investigate the neutral phosphonamidate methyl ester as well. In this way we could provide two possible mimics for the stable intermediate and thus compare their interactions, see below (Figure 2.2).

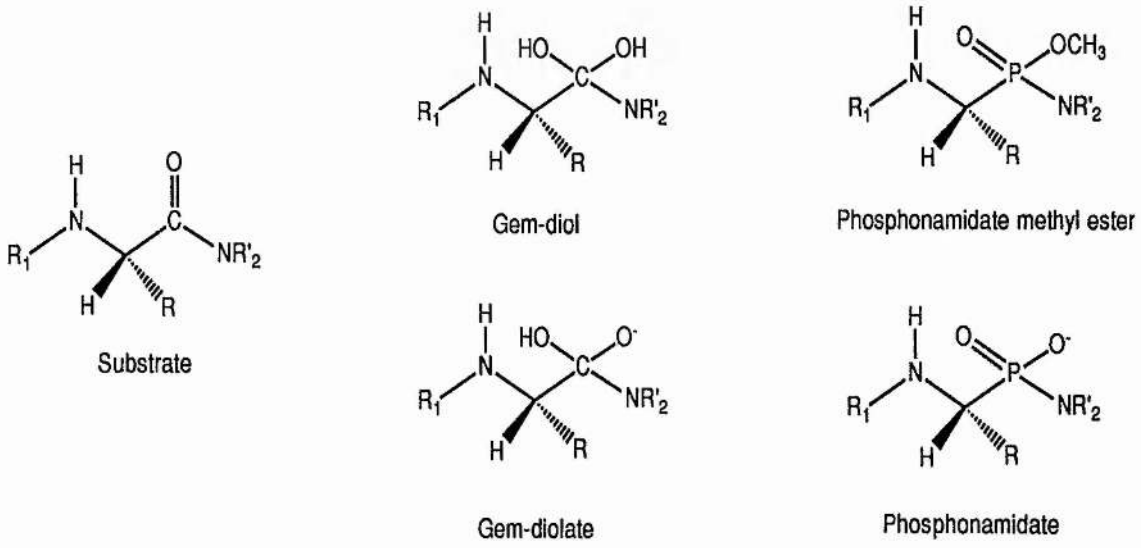


Figure 2.2. Comparison of the possible stable intermediates and phosphonamidate analogues.

The phosphonamidate methyl esters are the synthetic precursors to the phosphonamidates and so are readily available. The methyl esters have been found to be acid stable and so also offer an immediate advantage over the phosphonamidates. The phosphorous-nitrogen bond in phosphonamidates is acid labile and these compounds are known to have short half-lives in

solution.[127]

Despite lacking a hydrogen-bond donor, the phosphonamidate methyl ester has been shown to be a good mimic of the gem-diol intermediate in terms of charge density distribution. The charge congruency of the phosphonamidate methyl ester and the gem-diol tetrahedral intermediate has been proved by semi-empirical quantum mechanical calculations.[128] The results are given below in (Figure 2.3), with the partial charges residing on each atom shown.

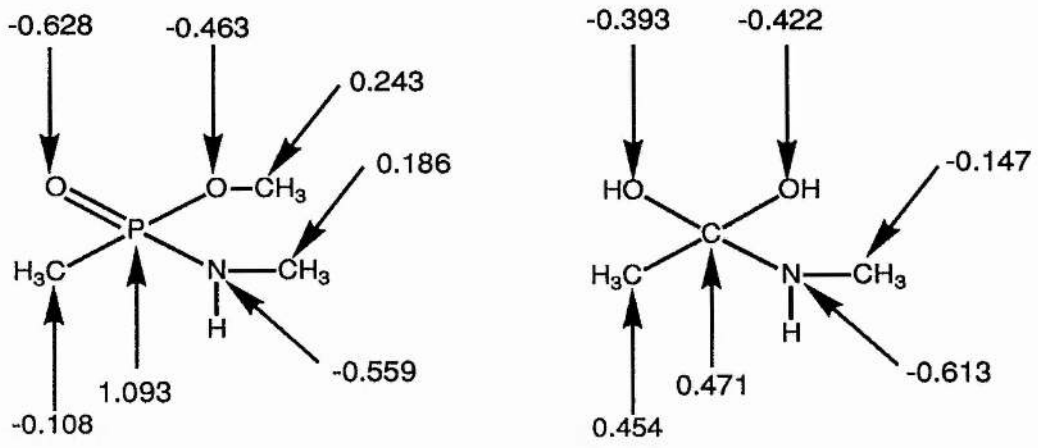
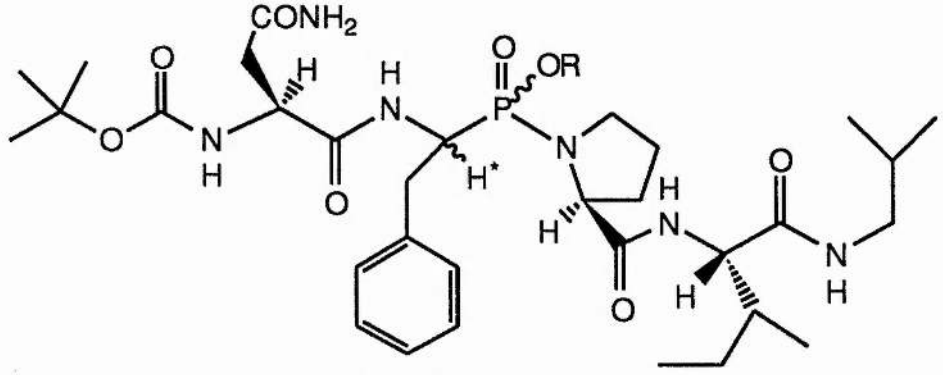


Figure 2.3. Charge density comparison of a model phosphonamidate methyl ester and a model amide hydrate.

The crucial points of similarity are the oxygens and the nitrogen of the hydrate and of the phosphonamidate methyl ester. The partial charges on the oxygens can be seen to be similar, showing again the close analogy between the two moieties. An *ab initio* calculation was recently carried out on phosphonamidates and similar results to those above were obtained.[129] In the same study it was found that the stable intermediate for acid-catalysed peptide bond hydrolysis is likely to be the neutral amide hydrate. Phosphonamidates were not good models for this stable intermediate but phosphonamidate methyl esters were found to be better. Although useful, these studies do not highlight the important hydrogen-bonding interactions between the enzyme and bound inhibitors. The fact that analogues containing hydroxyl groups tend to be the most potent inhibitors towards the HIV-1 protease suggests that both a hydrogen-bond acceptor and donor are required for optimisation of the interactions with the catalytic aspartates. The phosphonamidate methyl esters

possess hydrogen-bond acceptors but lack hydrogen-bond donating capacity.

Our initial targets (see below, Figure 2.4) were derived from the sequence of the protease-reverse transcriptase junction in the *gag-pol* polyprotein (Leu-Asn-Phe-Pro-Ile-Ser-).[130] This sequence has been used in the design of several successful inhibitors.[80]



Where R = CH_3 **Compound (44)**
R = Li **Compound (48)**

(Boc-(2S)-Asn-Phe-ψ[P(OR)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu.)

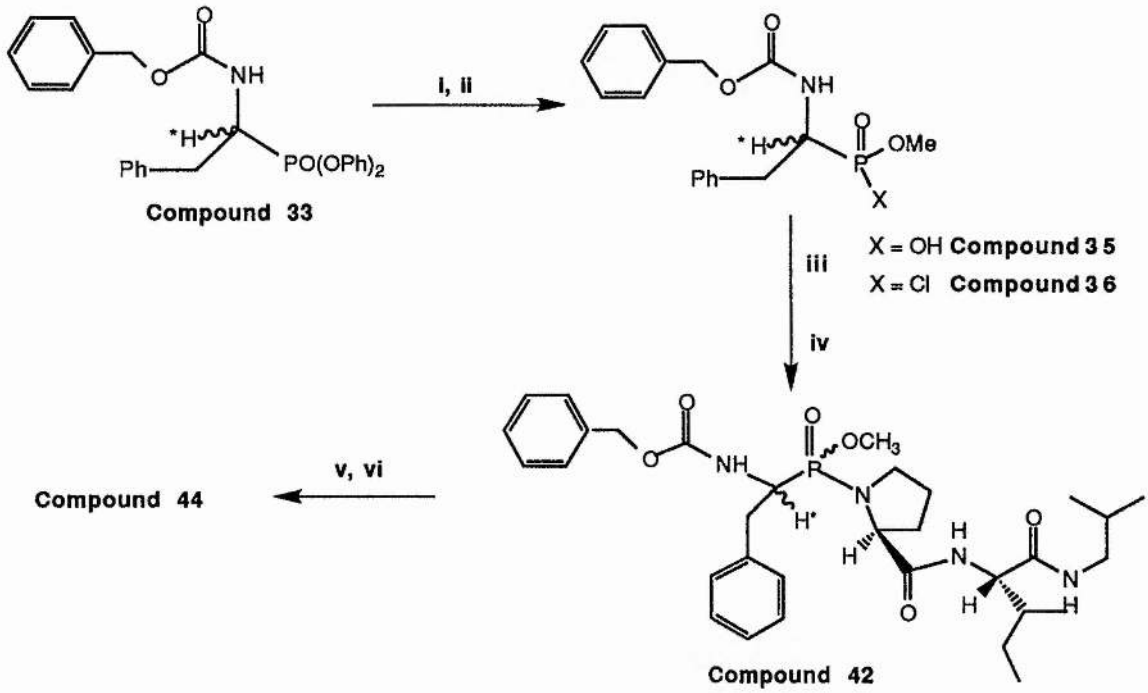
Figure 2.4.

Incorporation of the unique Phe-Pro scissile bond mimic into the design was expected to confer selectivity for the inhibition of the HIV-1 protease over the mammalian aspartic proteases.[131] For synthetic ease, the inhibitor also incorporates the Val surrogate, isobutyl amine, in the $P_{3'}$ position and the tertiarybutoxy carbonyl (Boc) protecting group as a P_3 substitute. Initially, the carbobenzyloxy (Cbz) protecting group was used in this position, but it was found that this group was unstable to the conditions required for the deprotection of the phosphonamidate methyl ester (see section **2.2 (iii)**).

2.2 Synthesis of phosphonamidate-containing peptide inhibitors.

The phosphonamidate analogues were synthesised by coupling activated aminophosphonic acid half esters (**35**) to various peptides, as outlined overleaf in Scheme 2.1. The half esters were activated by conversion to the appropriate

acid chloride, using thionyl chloride. Treatment of the resulting chlorophosphonate derivative (**36**) with a range of peptides gave the phosphonamidate methyl ester analogues in moderate to good yield (40-80%).



Scheme 2.1. The synthesis of compound 44.

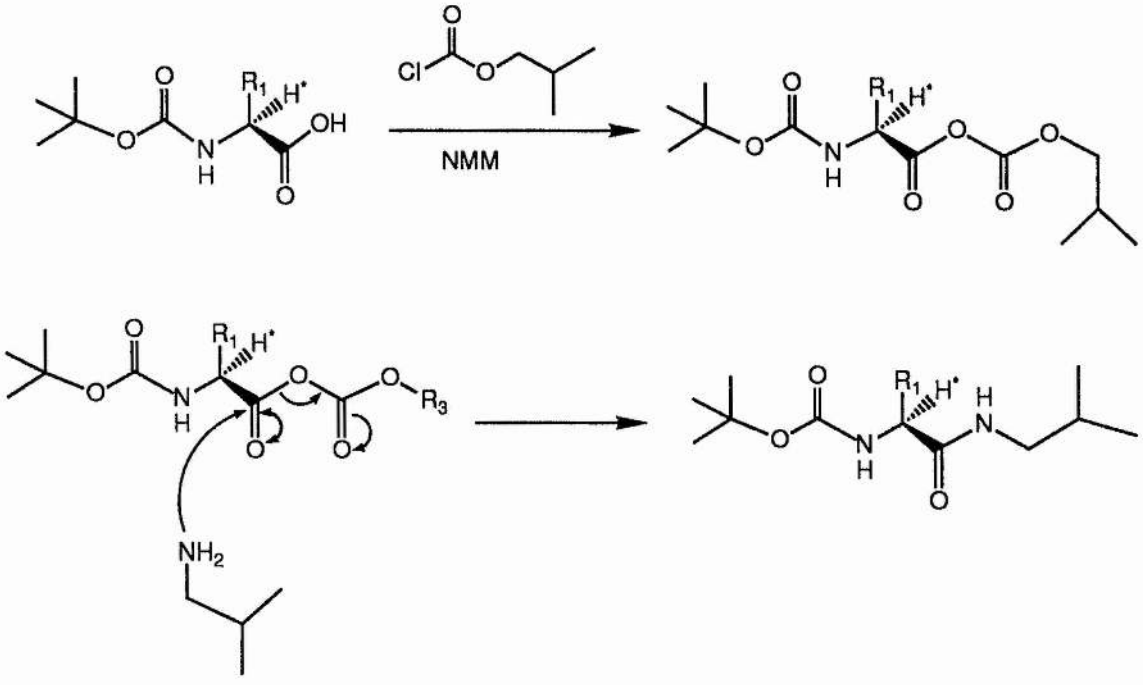
(i) 2.2 eq. NaOMe, MeOH, r.t.; (ii) 2.5 eq. NaOH, MeOH, r.t.; (iii) $SOCl_2$, CH_2Cl_2, r.t. then; (iv) HN-Pro-Ile-NH-iBu.HCl (**40**), NEt_3, CH_2Cl_2, r.t.; (v) Pd/C, H_2, MeOH, r.t.; (vi) Boc-Asn, iBuOCOCl, NMM, THF, -15 °C.

The phosphonamidate (**48**) was produced from the methyl ester (compound (**44**)) by treatment with 4 eq. lithium hydroxide in dioxan. The stereochemistry of the 4 diastereomers of compound (**44**), was assigned using a combination of X-ray crystallography on the synthetic intermediate (**42**), chemical degradation, ^{1}H, ^{13}C and ^{31}P NMR spectroscopy, as discussed in section **2.3**.

(i) Synthesis of peptide fragments.

The peptide fragments were synthesised using mixed anhydride coupling reactions. This method of peptide coupling was chosen because of its

advantages in speed, yield and ease of purification. Also, the enzyme has a binding preference for amino acid residues possessing the (S)-configuration and so racemisation of these stereogenic centres must be kept to a minimum during inhibitor synthesis. The mixed anhydride method uses alkyl chloroformates as peptide forming reagents. A suitably protected amino acid reacts with the alkyl chloroformate in the presence of a base, to form an anhydride. A second amino acid is then added, which reacts with the anhydride to form the peptide linkage (see below, Scheme 2.2).



Scheme 2.2. The mixed anhydride peptide coupling reaction.

The first step of the reaction is believed to be the attack of the tertiary base on the alkyl chloroformate to form a quaternary compound which in turn reacts with the carboxylic acid. The steric bulk of the base is important if racemisation is to be minimised and It has been found that methylamine reacts fastest, but causes maximum racemisation.[132] Base, solvent, temperature and alkyl chloroformate variations were examined and It was found that the best results were obtained using a combination of N-methylmorpholine (NMM), isobutylchloroformate and tetrahydrofuran (THF). [133] Consequently, this same combination was used in the course of this work.

The Pro-Ile-NH-iBu (**40**) tripeptide fragment was synthesised using Boc amino protection. The first step consisted of a coupling reaction between Boc-Isoleucine and isobutylamine (as exemplified in Scheme 2.2) to give the amino-protected dipeptide analogue, Boc-Ile-NH-iBu (**37**) in 85% yield. Subsequent deprotection using hydrogen chloride gas in ethyl acetate yielded Ile-NH-iBu hydrochloride (**38**) in 90%. This method of deprotection was chosen in preference to trifluoroacetic acid as the latter method led to problems during crystallisation. The second step involved condensing Boc-Pro with the dipeptide Ile-NH-iBu to afford the tripeptide analogue Boc-Pro-Ile-NH-iBu (**39**) in 91% yield. Subsequent deprotection in a similar manner gave the desired peptide fragment, Pro-Ile-NH-iBu (**40**) as the hydrochloride salt in 81% yield.

The tripeptide fragment provides the $P_{1'}$ to $P_{3'}$ residues of the inhibitor (see Figure 2.4). Variation of these residues is straight forward and during the course of this work, peptides incorporating (2S)-phenylalanine, (2R)-phenylalanine and glycine at $P_{1'}$ were also synthesised. 2-(Aminomethyl)-pyridine (AMP) and 2-(aminomethyl)-benzimidazole (AMB), popular residues in a host of HIV-1 protease inhibitors, were used to replace isobutylamine at the $P_{3'}$ position.

(ii) Synthesis of the aminophosphonic acid derivatives.

The methyl phosphonic acid derivatives (**35**) used in the chlorophosphonate coupling reactions were prepared by Paul Hawkins, using the method of Oleksyszyn and coworkers,[134] to yield the diphenyl phosphonate analogue (**33**) as shown in Scheme 2.1. This method involves condensing phenylacetaldehyde, benzyl carbamate and triphenyl phosphite in the presence of glacial acetic acid. The diphenylphosphonate analogue (**33**) was obtained in 31% yield. Transesterification of the diphenyl phosphonate using sodium methoxide provided the corresponding dimethyl phosphonate (**34**) in 72% yield. Subsequent saponification to the methyl phosphonic acid derivative (**35**), was achieved in 63% yield.

Although the three component reaction is inefficient, the reaction has been reproduced several times on a larger scale (30 g) and provides the analogue complete with suitable Cbz nitrogen protection. The Cbz group is easily removed by hydrogenolysis and allows extension in the P_2 and P_3 positions, as shown in Scheme 2.1.

Variation of the aldehyde in the three component coupling reaction allows a variety of P_1 residues to be examined. Hawkins[135] prepared analogues which incorporated phenylalanine, cyclohexylalanine, norleucine, leucine and valine respectively. Of these, the phenylalanine and cyclohexylalanine derivatives have been incorporated into a variety of phosphonamidate containing inhibitors. Incorporation of the other residues is presently being addressed in the laboratory by David Perrey.

A stereoselective route to optically active α-aminophosphonic acids would be useful and many methods are available (reviewed in[136]). These include resolution[137] and alkylation, using chiral auxiliaries.[138] For instance, the number of diastereomers produced after the chlorophosphonate-peptide coupling step would be reduced from four to two (see section **2.3 (iii)**). However, as the stereochemistry at phosphorus would be difficult to control, a chromatographic resolution step would still be required. Also, the inhibitory effect at P_1 with respect to stereochemical preference, was found to be small in the case of the phosphonamidate analogues (see section **2.5**) and so stereoselective syntheses were not pursued.

(iii) The chlorophosphonate-peptide coupling reaction.

The key step in this synthesis proved to be the coupling of the peptide fragment with the chlorophosphonate derivative (**36**). High yields for this two step reaction were required and so an examination of both the acid chloride and phosphonamidate formation steps were undertaken. Model coupling reactions using readily available isobutylamine were carried out prior to tripeptide couplings. Using a modification of the method of Bartlett,[139] the chlorophosphonate (**36**) was produced from (**35**) using 2 equivalents of thionyl chloride in anhydrous dichloromethane. Isobutylamine (2.5 equivalents) was added *in situ* to give Cbz-Phe-ψ[P(OMe)O-NH]-iBu (**41**) (see overleaf, Figure 2.5). The formation of the acid chloride was verified by ^{1}H NMR and a quantitative conversion was achieved within 4 hours as judged by the downfield shift for both the methyl phosphonate ester and the α CH (CH*) signals. Optimisation of the coupling reaction was achieved by repeatedly flushing away the gaseous by-products from the chlorination reaction using argon. In this

manner, the model phosphonamidate analogue (compound (**41**)) was obtained in a yield of 80% after one recrystallisation.

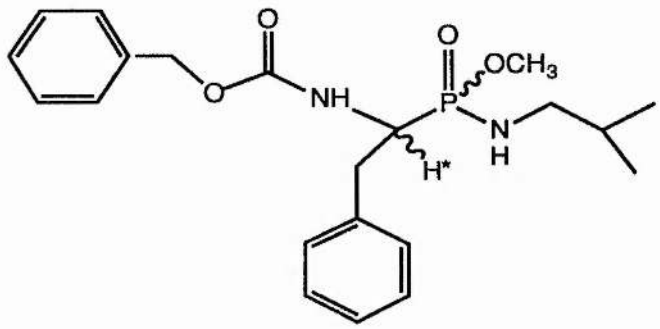


Figure 2.5. (41); Produced in the model coupling reaction.

In an analogous manner, proline methyl ester was condensed with the acid chloride (**36**) in a yield of 50%. The lower yield was attributed to the increased steric bulk of the secondary amino group. Rather than extending in a linear fashion from the proline methyl ester of this phosphonamidate analogue, it was decided to use a convergent synthesis. Thus, the tripeptide fragment, Pro-Ile-NH-iBu (**40**) was coupled to the acid chloride (**36**) in an optimised yield of 65%, to give the four expected diastereomers of compound (**42**);

i) Cbz-(**R**)-Phe-ψ[(**R**)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu.

ii) Cbz-(**R**)-Phe-ψ[(**S**)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu.

iii) Cbz-(**S**)-Phe-ψ[(**R**)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu.

iv) Cbz-(**S**)-Phe-ψ[(**S**)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu.

Four well resolved ^{31}P coupled doublets were observed in the ^{1}H NMR at δ 3.60, 3.66, 3.69 & 3.74 respectively, due to the methyl phosphonate ester group and are indicative of the number and amount of each diastereomer (see Figure 2.6). Interestingly, the diastereomers were produced in unequal amounts, with two diastereomers (δ 3.74 & 3.66) comprising 31% & 34% respectively and the other two (δ 3.69 & 3.60) comprising 19% & 16% respectively. The aforementioned ^{1}H NMR study of the acid chloride formation also showed no

stereochemical preference, as judged by equal ^1H NMR peak heights for the methyl phosphonate ester group. Thus, the diastereoselectivity must have arisen from the chlorophosphonate-peptide coupling reaction and may have been the result of a preferential attack at phosphorus.

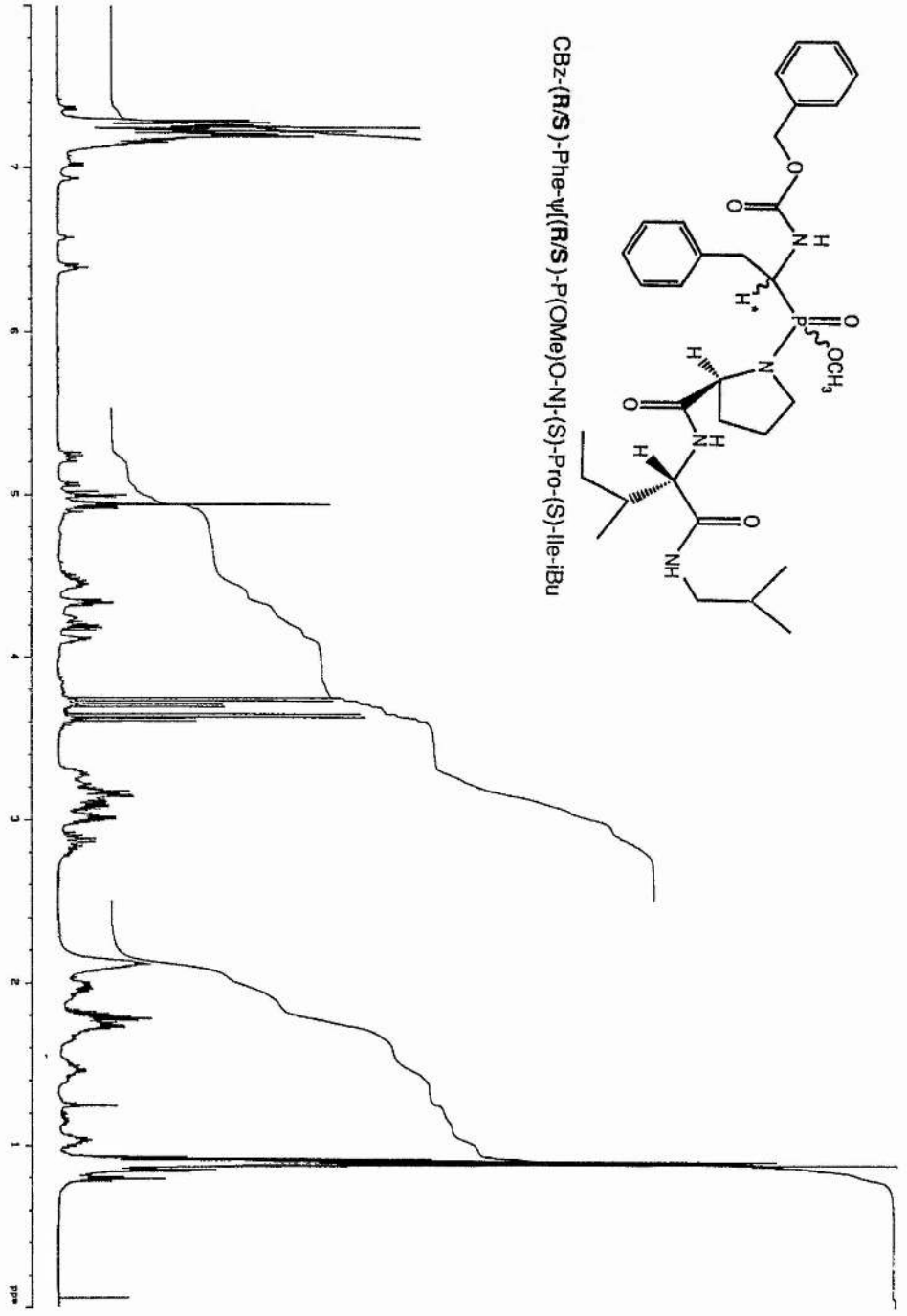


Figure 2.6. 600 MHz ^1H NMR showing 4 diastereomers of compound (42).

For instance, attack at phosphorus would be expected to go with inversion and so the preferred diastereomers may result from the attack of the incoming amino fragment from the opposite side of the benzyl chain (i.e. the least hindered side), thus producing a phosphonamidate analogue where the prolyl-peptide is *trans* with respect to the benzyl chain (see below, Figure 2.7). Using the same argument, the unfavoured diastereomers would result from an attack on the same side as the benzyl chain (i.e. the most hindered side), resulting in the prolyl-peptide being *cis* with respect to the benzyl chain.

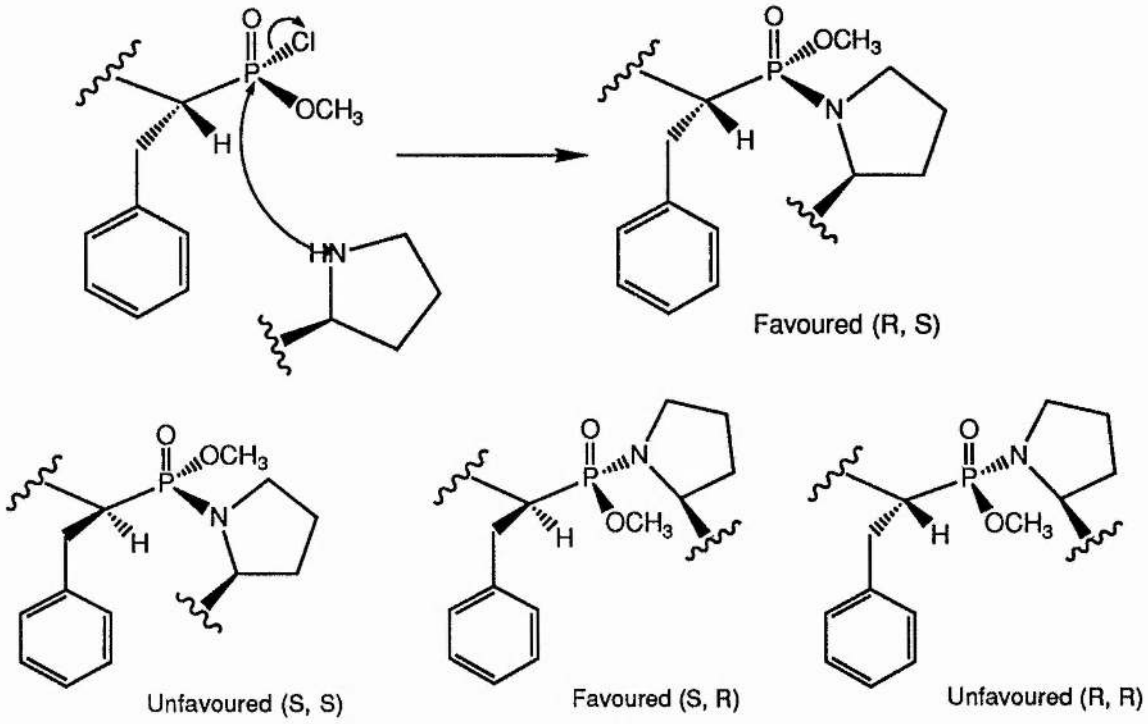


Figure 2.7. Theoretical explanation for observed diastereoselectivity.

Although speculative, this theoretical argument has implications with regard to the absolute stereochemical assignments and will be discussed in section **2.3.**

Each diastereomer of compound (**42**) was separated using column chromatography (EtOH: CH_2Cl_2 5:95), individually hydrogenolysed and then extended with Boc-(2S)-asparagine to yield the individual diastereomers of compound (**44**). Each diastereomer was tested as an inhibitor of the HIV-1

protease, the results of which will be discussed in section **2.5**.

The removal of the phosphonamidate methyl ester protection proved to be difficult. Our initial method used a reagent first prepared by Bartlett.[140] The lithium propane-1-thiolate reagent (**45**) in HMPA had been used successfully in the removal of hindered methyl esters. The thiolate anion effectively cleaves the alkyl-oxygen bond through nucleophilic displacement at carbon, to yield the lithium salt of the acid. The reagent was prepared initially in HMPA, as it had been shown that this solvent supported the greatest rates of cleavage for methyl esters.[141] In our case HMPA caused lots of problems. By TLC, the reaction proceeded readily, but the removal of the HMPA proved difficult. Also, due to the carcinogenic nature of HMPA, DMF was used as a suitable replacement. A model reaction with the isobutyl phosphonamidate methyl ester (compound (**41**)) produced the lithium salt of the phosphonamidate in reasonable yield. The result was encouraging and thus the demethylation to provide the target molecule (originally, (Cbz-(2S)-Asn-Phe-ψ[P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu.) was attempted. However, the thiolate anion readily attacked the benzyl portion of the Cbz group of the terminal asparagine group and competed with the removal of the phosphonamidate methyl ester even at 1.1 equivalents of reagent (see below, Figure 2.8).

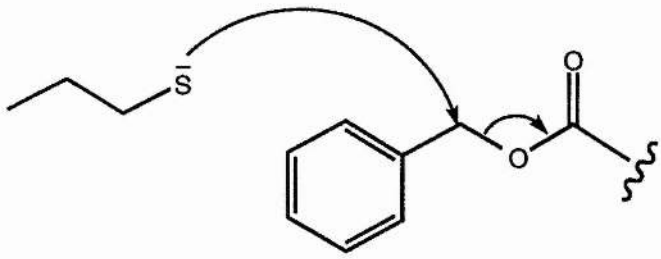
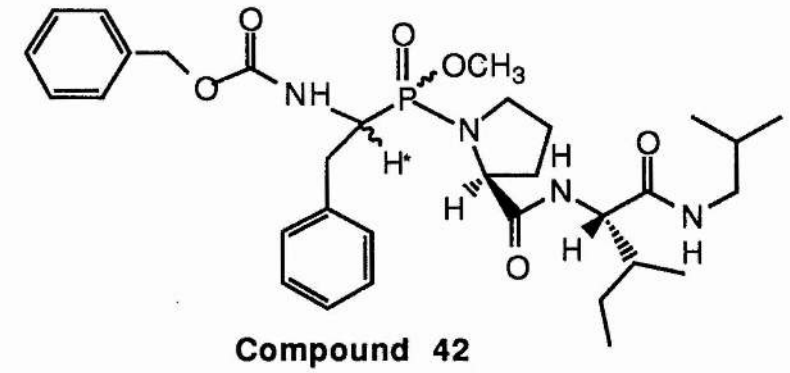


Figure 2.8. Possible decomposition of Cbz group by propane-1-thiolate.

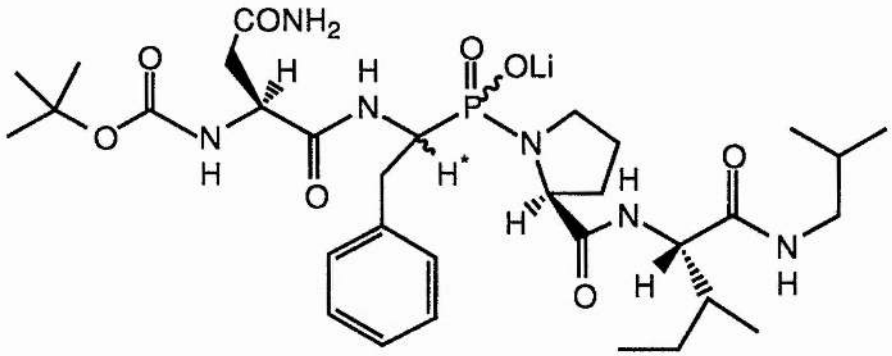
A closer examination of the literature, revealed an alternative method by the same author. Bartlett had used lithium hydroxide to remove the methyl ester group of his phosphonamidates.[139,142] He had used the thiol reagent successfully in the production of phosphinic acid inhibitors.[65] This could imply that the phosphonamidates are more stable as the lithium salt. Model reactions

of lithium hydroxide were successful, with the production of the lithium salt of the isobutyl phosphonamidate (compound (**46**)) in 77% yield. Following the success of the model phosphonamidate methyl ester deprotection using lithium hydroxide, This method was used in the deprotection of the phosphonamidate methyl ester of the target compound (**44**). However, our initial target, the Cbz-(2S)-Asn analogue of compound (**44**) could not be prepared due to the facile debenzylation of the Cbz group. This was surprising considering that the terminal Cbz group of the shorter model (compound (**41**)) remained untouched by both reagents. Therefore the amino terminal protecting group was altered to a base stable Boc group to facilitate selective phosphonamidate methyl ester deprotection. Hence, the new target became Boc-(2S)-Asn-Phe-ψ[PO_2-N]-(2S)-(Pro)-(2S)-Ile-NH-iBu (Compound (**48**)). The lithium hydroxide demethylation reaction proceeded very slowly in a variety of solvents (acetone, dioxan and acetonitrile), although complete cleavage was achieved using four equivalents of lithium hydroxide in dioxan at room temperature for 24 hours (indicated by the disappearance of the methyl ester doublets at ~ δ 3.6 in the ^{1}H NMR). This demethylation method for the deprotection of single diastereomers should be unsuitable, as the long reaction times and excess amounts of lithium hydroxide should facilitate racemisation. For this reason several other methods were investigated. The use of tertiarybutylamine, a very selective reagent used in the cleavage of phosphate and phosphonate esters[143] proved to be totally ineffective in phosphonamidate methyl ester cleavage. Bartlett and co-workers[144] had successfully used trimethylsilyl bromide (TMSBr) in the deprotection of the methyl esters of their phosphonate peptide inhibitors of pepsin. McKenna and co-workers[145] had previously used this reagent in the facile dealkylation of phosphinic acid esters and so this method of deprotection seemed promising. Deprotection of compound (**44**) was attempted and it was found that TMSBr was reactive enough to remove both the phosphonamidate methyl ester and the Boc group from the asparagine residue preferentially at 1.1 eq. This method appeared to be the most efficient way of phosphonamidate methyl ester deprotection but was not viable due to the accessibility and relative reactivity of the terminal protecting groups. Therefore our synthetic strategy was altered as follows (see overleaf, Scheme 2.3).



Compound 42

i, ii, iii



Compound 48

Scheme 2.3. Alternative phosphonamidate methyl ester deprotection.

(i) Pd/C, H_2, MeOH, r.t.; (ii) 2.2 eq. TMSBr, CH_2Cl_2, r.t.; (iii) (a) Boc-(2S)-Asn, iBuOCOCl, NMM, THF, -15 °C, (b) 1 eq. LiOH.

The phosphonamidate (**48**) was prepared by sequential removal of the benzyloxycarbonyl and phosphonamidate methyl ester protecting groups through catalytic hydrogenolysis followed by treatment of the free amine with trimethylsilyl bromide, prior to reaction with Boc-(2S)-asparagine using the mixed anhydride method. The use of TMSBr proved to be a facile method of phosphonamidate methyl ester deprotection. However, during the course of the TMSBr reaction the newly formed phosphonamidate was prone to cleavage at

the phosphorus-nitrogen bond. As this bond is known to be acid labile isobutylene was used to scavenge the acid produced during the reaction. This successfully prevented cleavage and the deprotected compound was extended with Boc-(2S)-asparagine in the usual manner (see Scheme 2.3).

The inherent problems associated with the phosphonamidates such as their ionised nature at the optimum pH of the HIV-1 protease, their acid instability and the difficulty of phosphonamidate methyl ester deprotection caused us to abandon them as possible therapeutic agents. Indeed the phosphonamidates were found to be moderate inhibitors of the HIV-1 protease (see section **2.5**). More significantly, the synthetic accessibility, the ease of resolution and the acid stability of their methyl ester counterparts made them far more attractive compounds and consequently were further developed in the course of this work. Similar phosphonamidate methyl ester analogues were prepared in order to optimise the inhibitory effect of this class of compounds. For instance, analogues incorporating Cha-ψ[P(OMe)O-N]-(2S)-Pro (**50**), Phe-ψ[P(OMe)O-NH]-(2S)-Phe (**51**), Phe-ψ[P(OMe)O-NH]-(2R)-Phe (**64**), Cha-ψ[P(OMe)O-NH]-(2S)-Phe (**52**) and Phe-ψ[P(OMe)O-NH]-Gly (**65**) were synthesised in a similar manner and tested for inhibition against the HIV-1 protease (see section **2.5**).

2.3 Absolute stereochemical assignment of compound 42.

The diastereomers of compound (**42**) were separated by silica chromatography and designated 1-4 in order of their elution. With these individual diastereomers in hand the task of assigning their absolute stereochemistry at both P_1 (CH*) and at phosphorus was undertaken. In this way each individual diastereomer could be tested for inhibition against the HIV-1 protease and the preferred stereochemistry at both carbon (CH*) and phosphorus examined.

After considerable effort, crystals were grown of the individual diastereomers of 3 and 4. Although both samples produced thin crystals, the X-ray crystal structure of diastereomer of 4 was solved from a very limited data set. The stereochemistry at carbon (CH*) was found to be (R) and the stereochemistry at phosphorus was also found to be (R) (The structure of diastereomer 4 is shown overleaf in Figure 2.9).

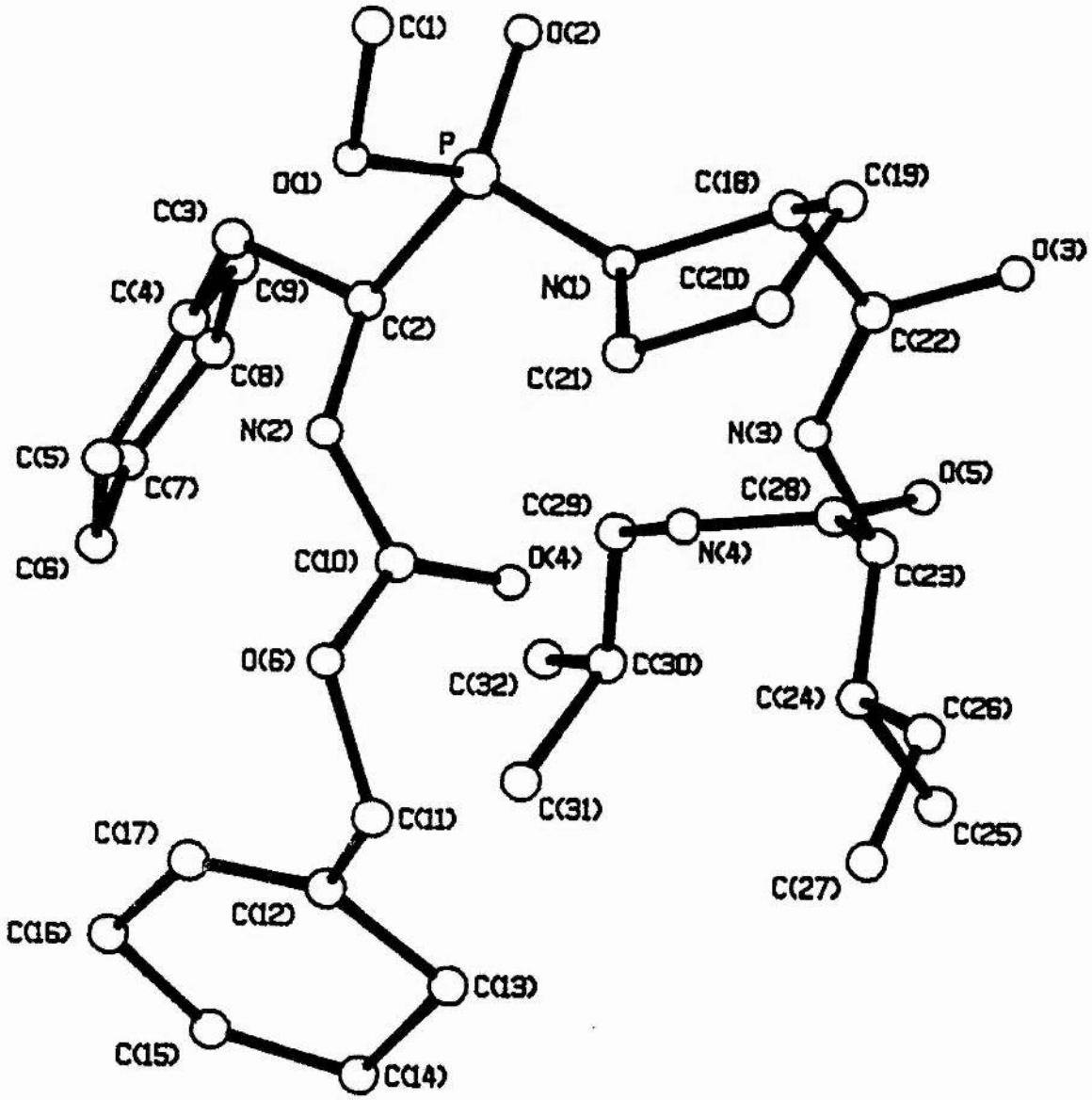


Figure 2.9. The crystal structure of Cbz-(R)-Phe-ψ[(R)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (42).

Further information was gained by studying the 600 MHz ^{1}H NMR spectra of each individual diastereomer of compound (**42**) (see overleaf, Figure 2.10). The chemical shifts of the α-hydrogen of the (2S)-Pro residue were δ 4.11, 4.10, 4.26, and 4.32 for diastereomers 1-4 respectively (all relevant chemical shift data is given in Table 2.1). After studying these values, it was predicted that D1 and D2 had the same stereochemistry at phosphorus, as it was assumed that the same stereochemistry at phosphorus would cause a similar effect on the chemical shift of the α-hydrogen of the (2S)-Pro residues. Similarly, it was predicted that D3 and D4 had the same stereochemistry at phosphorus (although different from D1 and D2). Therefore, using the known stereochemistry of D4 (**R**, (CH*), **R**, (Phos.)) the stereochemistry at phosphorus was assigned as follows:

D1 (x, (CH*), **S**, (Phos)),
D2 (x, (CH*), **S**, (Phos)),
D3 (x, (CH*), **R**, (Phos)),
D4 (**R**, (CH*), **R**, (Phos)).

The stereochemistry at carbon (CH*) was more difficult to predict using 600 MHz ^{1}H NMR alone as the chemical shifts of the (CH*) protons were very similar. An examination of the chemical shifts of both of the protons of the adjacent benzyl group (Ar$C\underline{H}_2$CH*) proved inconclusive as well. Therefore, a study of the 600 MHz ^{1}H NMR spectra of the individual diastereomers of compound (**44**) was undertaken. The α-hydrogen chemical shifts of the (2S)-Pro residue again indicated a (D1 & D2) and (D3 & D4) partnership with respect to the same stereochemistry at phosphorus. By extending compound (**42**) with (2S)-Boc-Asn a fixed stereogenic centre had been placed adjacent to the (CH*) and so it was hoped that an inductive effect, (similar to that for the (2S)-Pro α-hydrogen's) would be observed. Indeed, the values of the chemical shifts of the (CH*) protons for D2 and D3 were higher than those of D1 and D4. The chemical shifts of the asparagine CH_2 protons also indicated a (D1 & D4) and (D2 & D3) partnership with respect to the same stereochemistry at carbon.

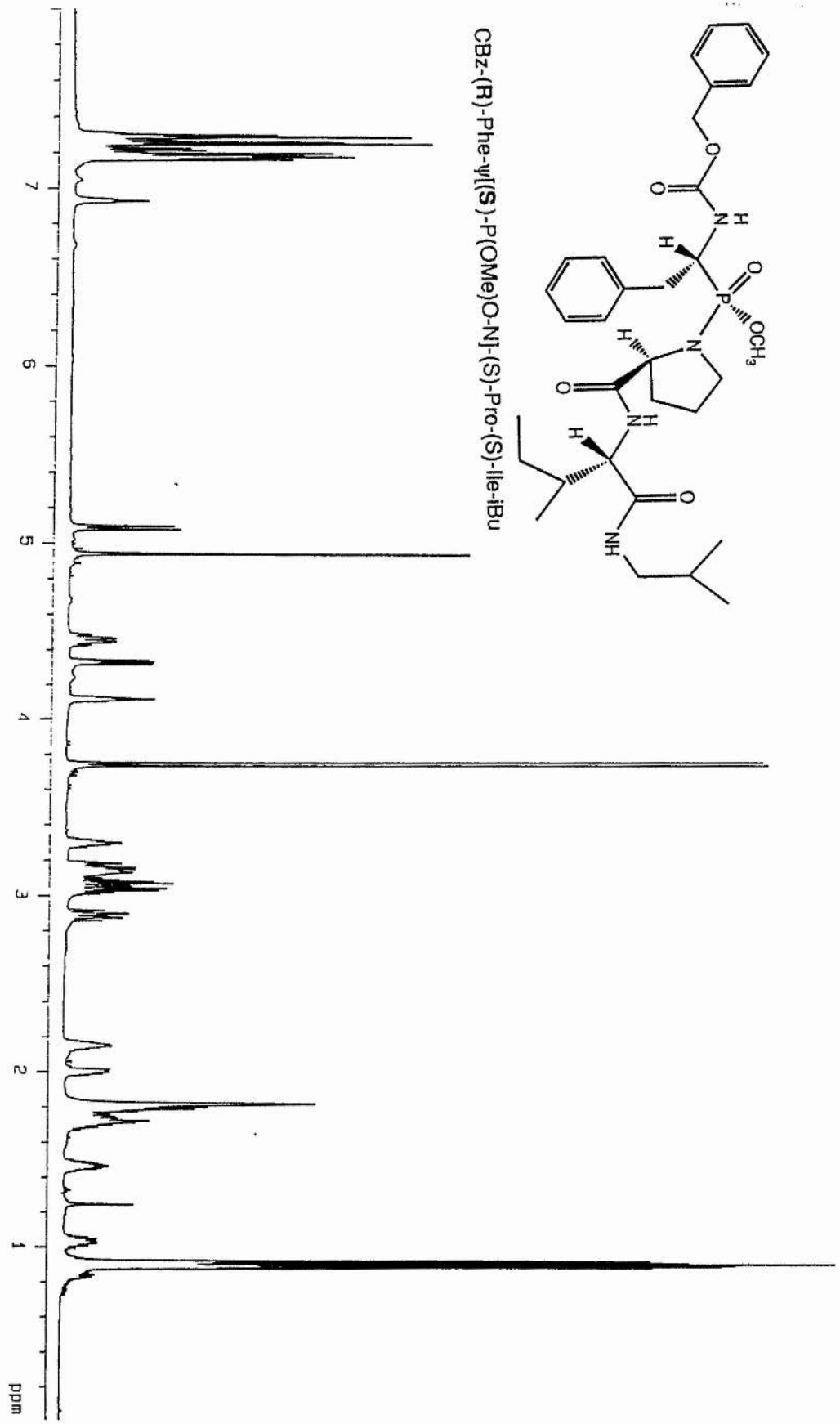


Figure 2.10. 600 MHz ^{1}H NMR of (R,S) diastereomer of compound (42).

Table 2.1, 600 MHz ^{1}H NMR data for compounds (42) and (44).

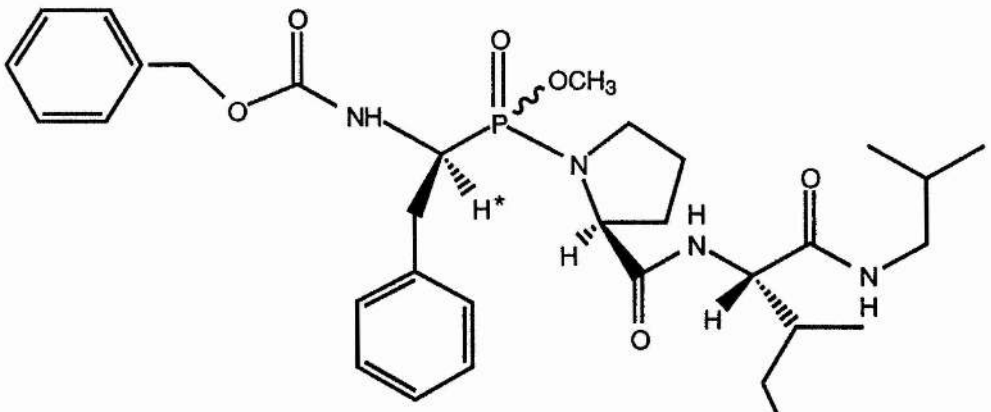
600 MHz ^{1}H NMR Data.

Compound (42)		Compound (44)	
D1.) α-H Pro	4.11 ppm.	**D1**.) α-H Pro	4.15 ppm.
$PO(OC\underline{H}_3)$	3.74	$PO(OC\underline{H}_3)$	3.75
CH*	4.46	CH*	4.74
$ArC\underline{H}_2$	2.88 & 3.18	$ArC\underline{H}_2$	2.95 & 3.17
		α-H Asn	4.32
		CH_2 Asn	2.46 & 2.62
D2.) α-H Pro	4.10 ppm.	**D2**.) α-H Pro	4.16 ppm.
$PO(OC\underline{H}_3)$	3.69	$PO(OC\underline{H}_3)$	3.76
CH*	4.49	CH*	4.81
$ArC\underline{H}_2$	2.84 & 3.21	$ArC\underline{H}_2$	2.80 & 3.27
		α-H Asn	4.35
		CH_2 Asn	2.18 & 2.46
D3.) α-H Pro	4.27 ppm.	**D3**.) α-H Pro	4.25 ppm.
$PO(OC\underline{H}_3)$	3.66	$PO(OC\underline{H}_3)$	3.66
CH*	4.46	CH*	4.78
$ArC\underline{H}_2$	2.89 & 3.18	$ArC\underline{H}_2$	2.86 & 3.16
		α-H Asn	4.55
		CH_2 Asn	2.07 & 2.21
D4.) α-H Pro	4.32 ppm.	**D4**.) α-H Pro	4.38 ppm.
$PO(OC\underline{H}_3)$	3.60	$PO(OC\underline{H}_3)$	3.62
CH*	4.45	CH*	4.66
$ArC\underline{H}_2$	2.80 & 3.28	$ArC\underline{H}_2$	2.84 & 3.28
		α-H Asn	4.36
		CH_2 Asn	2.45 & 2.64

From the X-ray structure and the 600 MHz ^{1}H NMR data, the absolute stereochemistry was tentatively assigned as follows:

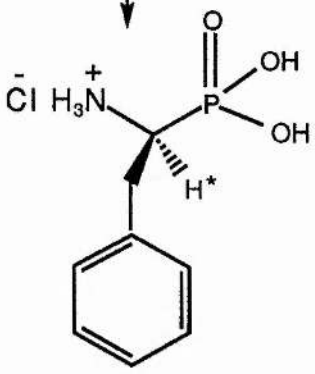
D1 (**R**, (CH*), **S**, (Phos)),
D2 (**S**, (CH*), **S**, (Phos)),
D3 (**S**, (CH*), **R**, (Phos)),
D4 (**R**, (CH*), **R**, (Phos)).

In order to unambiguously assign the stereochemistry at carbon, a series of acid hydrolysis experiments on the individual diastereomers of compound (**42**) (see overleaf, Scheme 2.4) was undertaken. Refluxing in a mixture of glacial acetic acid and 6 M HCl for two hours degraded compound (**42**) into a mixture of the hydrochloride salts of proline, isoleucine and isobutylamine and a phenylalanine phosphonic diacid derivative. Thin layer chromatography of the hydrolysis mixture gave spots with R_f values of 0.5, 0.67, 0.9 and 0.2 respectively. By isolating the phenylalanine derivative, a comparison of the direction of optical rotation with that of the known $[\alpha]_D$ values for the enantiomers of (1-amino-2-phenyl)ethyl phosphonic acid (**49**) could be undertaken.[146] In this way, the absolute stereochemistry of the phenylalanine carbon (CH*) in the original diastereomer could be deduced.



Compound 42

Reflux 6 M HCl / glacial acetic acid, 2h.



+ HCl salts of Pro, Ile, & iBu.

Compound 49

Scheme 2.4. Acid hydrolysis experiment.

Initial preparative TLC experiments were carried out on the hydrolysis mixture, using an isopropanol, conc. ammonia and water (26:6:5) solvent system. The methyl phosphonic acid derivative (**35**) was hydrolysed in an identical manner to compound (**42**) to give the required phosphonic diacid (**49**). This sample was used together with standard solutions of proline, isoleucine, and isobutylamine to identify which band of the TLC plate was required in order to isolate the required phenylalanine derivative. Initially, problems were found in separating the phenylalanine and proline derivatives, therefore the volume of ammonia in our system was doubled. This increased the separation and allowed the successful isolation of the phenylalanine derivative. It was found

that at least 50 mg of compound (**42**) was required in order to isolate sufficient phenylalanine derivative for the optical rotation experiment (100% recovery of the phenylalanine derivative would give 16 mg, however a proportion of this is lost.). The literature $[\alpha]_D$ values are -49° and +52° respectively (c 1 in 1 M NaOH) for the (R) and (S) enantiomers of (**49**).[146] Due to the small amounts of phenylalanine derivative recovered, we looked at the sign of optical rotation as a guide to the stereochemistry at carbon rather than the amount of rotation. Diastereomer 1 was degraded and the phenylalanine derivative isolated, after preparative TLC using the modified solvent system (isopropanol, conc. ammonia, water, 26:12:5). The optical rotation for D1 in 1 M NaOH was negative. This implied that the stereochemistry at carbon was (R). This was in agreement with the tentative assignment made on the basis of ^{1}H NMR. The sodium salt of the phenylalanine derivative (**49**) from D1 gave an identical ^{1}H NMR to the previously prepared racemic compound. The possibility of racemisation at (CH*) under the acidic conditions was of particular concern and so to test the validity of this experiment, the hydrolysis of D4 (stereochemistry known) was completed. The optical rotation was negative, which implied that the stereochemistry at carbon was (R) (in agreement with the X-ray structure). For completeness, the same experiments for D2 and D3 were carried out and positive optical rotations were observed for both D2 and D3. This implied that the stereochemistry at carbon was (S). Hence, the total stereochemical assignment had been achieved. Solving the absolute stereochemistry of the diastereomers of compound (**42**) (see below), automatically provided the absolute stereochemistry of the corresponding diastereomers of compound (**44**).

D1 (**R**, (CH*), **S**, (Phos)),
D2 (**S**, (CH*), **S**, (Phos)),
D3 (**S**, (CH*), **R**, (Phos)),
D4 (**R**, (CH*), **R**, (Phos)).

Interestingly, these stereochemical assignments are in accord with those predicted from the thereotical argument based on preferential attack at phosphorus (see Figure 2.7).

2.4 Phosphorus-based protease inhibitors.

Phosphorus-containing peptide analogues have been used extensively in the inhibition of the zinc and aspartic protease enzymes (reviewed in[67]). Indeed, these analogues have produced some of the most potent inhibitors of both classes. Although the zinc proteases and the aspartic proteases differ in many respects, they are related by mechanism; both catalyse the direct addition of water to the scissile carbonyl of the amide bond, with formation of a tetrahedral intermediate. Both mechanisms involve basic catalysis from a carboxylate side chain and acidic or Lewis acidic catalysis from another group in the active site. Zinc performs this role in the case of thermolysin and carboxypeptidase A.

The zinc proteases, thermolysin and the digestive enzyme carboxypeptidase A have received a great deal of attention. Early work in this area uncovered the naturally occurring phosphoramidon (see below, Figure 2.11), as a potent inhibitor of thermolysin.[147] The crystal structure of phosphoramidon bound to thermolysin was subsequently solved[148] and helped in the elucidation of the catalytic mechanism of the zinc proteases. Similar inhibitors for the mechanistically related enzyme carboxypeptidase A soon followed.[149]

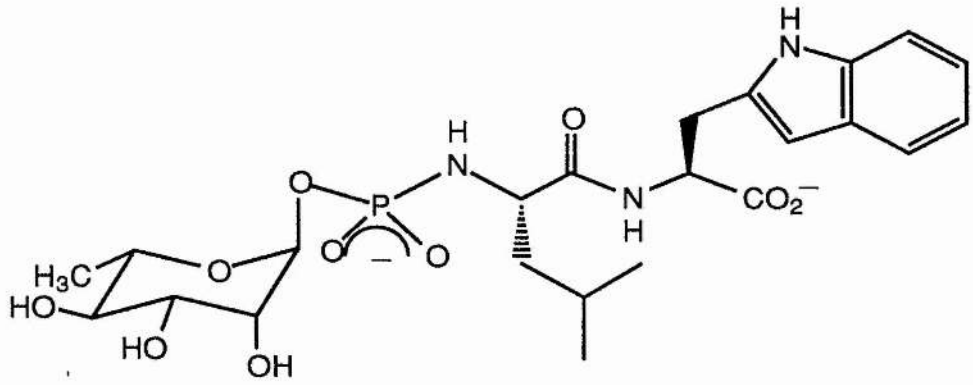


Figure 2.11. Phosphoramidon, the naturally occurring thermolysin inhibitor.

Bartlett and co-workers have made an outstanding contribution in the area of phosphorus-containing analogues. His group synthesised phosphonamidate-containing peptides as highly potent inhibitors of both thermolysin[142] and carboxypeptidase A.[139] These inhibitors are ideally suited due to their ionised

nature at the assay pH. In contrast to the aspartic proteases, the zinc proteases proceed *via* a gem diolate intermediate which coordinates to the zinc in the active site. Thus the ionised phosphonamidate is able to coordinate to the zinc in an analogous manner. This theory has been well supported by crystallographic studies.[150,151] Bartlett and Kezer[65] synthesised phosphinic acid analogues as inhibitors for the aspartic protease, pepsin. Their most potent inhibitor gave a K_i value of < 0.07 nM at pH 3.5. The aspartic proteases operate *via* a neutral gem diol intermediate (see Figure 1.8) and therefore this high potency is probably due to the fact that at pH 3.5 the phosphinic acid is predominantly protonated and thus binds tightly to the catalytic aspartates. These inhibitors are found to be less effective against the aspartic protease, penicillopepsin which operates at a higher pH. In fact, a pH dependence study has shown that K_i increases with increasing pH, suggesting that the enzyme binds only the neutral (protonated) form of the inhibitor.[67] Although mechanistically similar, the subtle differences between the zinc and aspartic proteases can have a profound effect on the design of inhibitors. Clearly, achieving high potency for one class of enzyme does not guarantee that the same level of potency will be achieved with the same inhibitor.

(i) Phosphorus-based inhibitors for the HIV-1 protease.

The use of phosphorus-based inhibitors for the HIV-1 protease has been investigated by a number of groups. The observation that phosphinic acids were highly potent inhibitors of the aspartic protease, pepsin, provided a good basis for HIV-1 protease inhibitors. Grobelny *et al.*[68] synthesised potent phosphinic acid inhibitors and observed a similar pH dependence to that observed by Bartlett (see section **1.4.6 (ii)**). Peyman and co-workers[104] synthesised the first symmetrical phosphinic acid inhibitors for the HIV-1 protease. The compound overleaf (see Figure 2.12) as a mixture of stereoisomers gave an IC_{50} value of 36 nM. Whilst optimising symmetry, this compound has the disadvantage of positioning the benzyl side chain at P_1' incorrectly (the backbone spacing of the peptide substrate is not preserved). More recently, the same compound was synthesised by Abdel-Meguid and co-workers.[69] They resolved the stereoisomers and the most potent isomer (the one which incorporates the natural amino acid stereochemistry) gave a K_i value of 3 nM at pH 6.

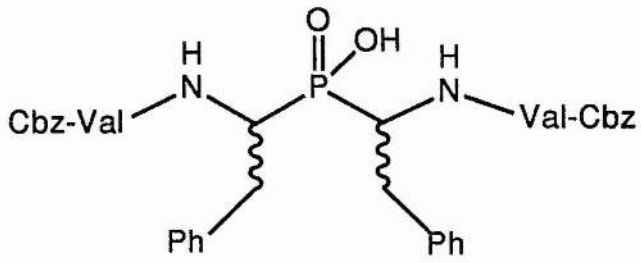


Figure 2.12. C-2 symmetric phosphinic acid inhibitor.

Despite its low K_i value this compound was found to be a weak inhibitor when tested in an antiviral assay, most probably due to poor cell penetration.[69] The methyl ester derivative of this compound was not an inhibitor of HIV-1 protease (Ki > 1mM), nor was it found to be active in an antiviral assay. This latter result is of particular interest to our work on the phosphonamidate methyl esters and its relevance will be discussed in section **2.5 (ii).**

During the course of this work the first phosphonamidate inhibitors for the HIV-1 protease were reported by McLeod and co-workers.[127] They showed that phosphonamidate methyl esters were as active against the HIV-1 protease, as their phosphonamidate counterparts (see below, Figure 2.13).

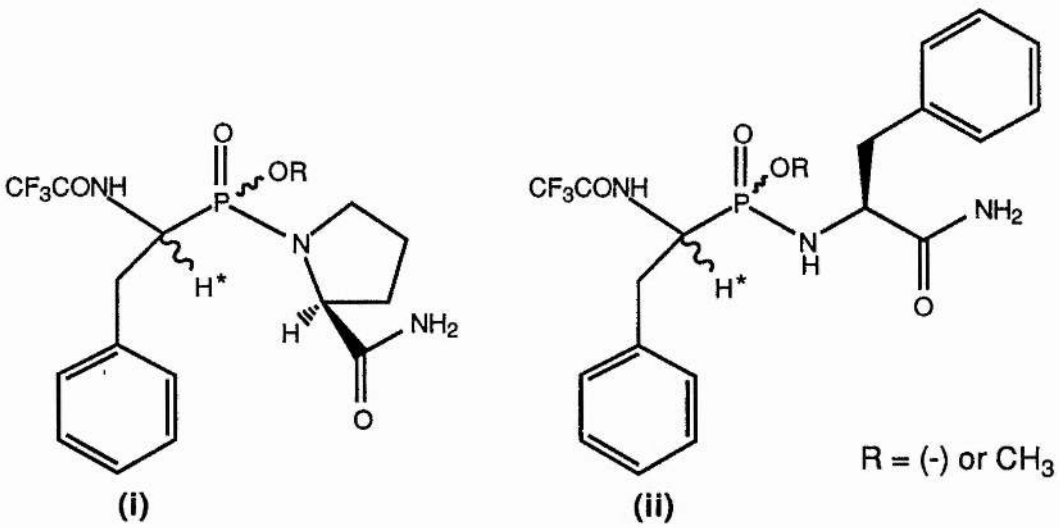


Figure 2.13. The phosphonamidate inhibitors of McLeod *et al.*

The phosphonamidate (R = (-)) of compound (**i**) gave a K_i value of 52 μM and the phosphonamidate methyl ester of (**i**) gave a K_i value of 44 μM, whilst the

phosphonamidate (R = (-)) of compound (**ii**) gave a K_i value of 45 μM and the phosphonamidate methyl ester of (**ii**) gave a K_i value of 80 μM. Under the assay conditions the phosphonamidates were reasonably stable ($t_{1/2}$ of (**i**) and (**ii**), 5 h. and 30 h. respectively), with the methyl esters being acid stable (less than 2 % decomposition over a week at pH 3.5). Whilst these compounds were only modest inhibitors of the HIV protease, the fact that they were inhibitors at all was encouraging, considering that these compounds were racemic at P_1 and were only dipeptide mimics. The HIV-1 protease is known to require hexapeptide substrates (see section **1.4.4**) and the shortest inhibitors so far published are tetrapeptide mimics (see section **1.4.6 (v)** and **(vi)**).

2.5 Results of inhibitor testing.

A variety of phosphonamidate methyl ester analogues were prepared and tested against the HIV-1 protease, in both *in vitro* and *in vivo* assays. The *in vitro* assays were performed at St. Andrews primarily by David Perrey and Paul Hawkins. These analogues were also tested in an *in vivo* assay by Derek Kinchington at St. Bartholomew's Hospital.

The *in vitro* assays were based on the method of Griffiths et al.[152] The substrate used (KARV-Nle-Phe$(NO)_2$-EA-Nle-G-NH_2) in our assays produced a K_M value of 33 μM[135] under the assay conditions, which compares well with the literature value of 35 μM.[152] The assay relies on monotoring a decrease in absorbance at 300 nm, due to the cleavage of the Nle-Phe(NO_2) bond. By measuring the initial rates at various concentrations of inhibitor (typically 0 to 100 μM), a Dixon plot[153] can be produced in which the IC_{50} value can easily be extracted..

Our initial target was tested both as the methyl ester (**44**) and the phosphonamidate (**48**). The phosphonamidate (**48**) was tested as a mixture of two epimers at the α-carbon (CH*) of the phosphophenylalanine residue and produced an IC_{50} value of 90 μM. At the assay pH (5.6) the phosphonamidate is negatively charged and so would repel the negatively charged aspartate residue in the active site.[67] This rather high value confirmed the expectation that the anionic compounds would be poor inhibitors and attention was turned to the

methyl esters. All four diastereomers of the methyl ester analogue (**44**) were tested individually and the results are shown below in Table 2.2.

Table 2.2 Test results for compound (44).

Compound	Configuration at (CH*,P)	IC_{50} (μM) *in vitro*	*in vivo*
Boc-N-F-[ψ]-P-I-NH-iBu.			
44 i	(R, R)	30	45
44 ii	(R, S)	80	>100
44 iii	(S, S)	100	>100
44 iv	(S, R)	>100	>100

(Where [ψ] = $PO(OCH_3)$)

These results show that the differences in potencies between the phosphonamidate and the methyl esters is very small. This is in agreement with results obtained by McLeod *et al.*[127] However, the results above show a preference for the (R)-configuration at both the α-carbon of the phosphophenylalanine residue (CH*) and at phosphorus. The (R)-configuration at carbon is expected due to its equivalence to the (S)-configuration in the natural substrate. The preferred configuration at phosphorus is difficult to interpret, especially as the methyl group on the phosphonamidate would not be expected to interact favourably with the aspartates. The relative potencies of each of the diastereomers was disappointing, considering the effort for both resolving and determining the absolute stereochemistry of each analogue. For instance, the results obtained by Cushman *et al.*[62] showed a strict preference for the (2S)-Phe epimer at P_1 (see section **1.4.6 (i)**).

As a matter of course, the individual diastereomers of the synthetic intermediate (compound (**42**)) were tested. In this shorter compound, the P_3 Boc group had been deleted and the P_2 asparagine residue replaced by a carbobenzyloxy moiety. These compounds were also tested in the *in vivo* assay and the results are shown overleaf in Table 2.3.

Table 2.3 Test results for compound (42).

Compound	Configuration at (CH*,P)	IC_{50} (μM) *in vitro*	IC_{50} (μM) *in vivo*
Cbz-F-[ψ]-P-I-NH-iBu.			
42 i	(R, R)	5	3
42 ii	(R, S)	25	25
42 iii	(S, S)	N.D.	25
42 iv	(S, R)	N.D.	>100

The results again highlighted the preference for the (R)-configuration at both CH* and at phosphorus. Surprisingly, these smaller inhibitors appeared to be more potent than the diastereomers of compound (**44**), despite lacking a P_3 residue. The *in vivo* result for the (S,S)-diastereomer was surprising, considering that this stereoisomer possessed the unnatural amino acid configuration at P_1 and so would be expected to bind less tightly than the (R,S)-diastereomer.

The importance of the P_1 residue in compound (**42**) was investigated by replacing Phe with a Cha (cyclohexylalanine) residue. Compound (**50**) was resolved into two pairs of diastereomers. The results for the most potent pair are given below in Table 2.4.

Table 2.4 Test results for compounds (50), (51), (52) & (55).

Compound	Configuration at (CH*,P)	IC_{50} (μM) *in vitro*	IC_{50} (μM) *in vivo*
Cbz-Cha-[ψ]-P-I-NH-iBu.			
50	(mix. 2 diast.)	16	8
Cbz-F-[ψ]-(2S)-F-I-NH-iBu.			
51	(mix. 4 diast.)	2	2
Cbz-Cha-[ψ]-(2S)-F-I-NH-iBu.			
52	(mix. 4 diast.)	2	5
H_2N-F-[ψ]-(2S)-F-I-NH-iBu.			
55	(mix. 4 diast.)	30	20

These results show that Cha can be used to replace the natural Phe residue in the P_1 position without a significant change in potency and compare to that predicted for a mixture of the (R,R)- and (R,S)-diastereomers of compound (**42**).

The literature indicated that a large number of the most potent inhibitors of the HIV-1 protease contained either a phenylalanine residue or a phenylalanine surrogate in the $P_{1'}$ position (for examples see section **1.4.6**). Thus, the (2S)-Pro residue at $P_{1'}$ in compound (**42**) was replaced with a (2S)-Phe residue. Compound (**51**) proved impossible to resolve using silica chromatography and was, therefore, initially tested as a mixture of four diastereomers (see Table 2.4). The small effect observed with respect to configuration for the resolved diastereomers did not warrant the resolution of further compounds prior to testing. Since the mixture contained four diastereomers, the most active diastereomer may have possessed an IC_{50} value of at most 500 nM. Even as a mixture, the Phe-Phe based inhibitor was more potent than the most active diastereomer of the corresponding Phe-Pro analogue (compound (**42 i**)). Similarly, the Phe residue at P_1 in compound (**51**) was replaced with a Cha residue and the result for compound (**52**) is given in Table 2.4. This result was in agreement with compound (**50**) in which Cha at P_1 was found to be equipotent to the Phe analogue in this position. Removal of the Cbz group in compound (**51**) gave compound (**55**) and led to a 10 fold reduction in potency (see Table 2.4) and showed that the Cbz group was compatible with the binding requirements of the S_2 pocket.

Table 2.5 Test results for compounds (56), (60) & (61).

Compound	**Configuration at (CH*,P)**	**IC_{50} (μM)**	
		in vitro	***in vivo***
Boc-N-F-[ψ]-(2S)-F-I-NH-iBu.			
56	(mix. 4 diast.)	100	>100
Qua-N-F-[ψ]-(2S)-F-I-NH-iBu.			
60	(mix. 4 diast.)	45	50
Qua-N-Cha-[ψ]-P-I-NH-iBu.			
61	(1 diast. unknown)	20	15

Compound (**55**) was extended with Boc-(2S)-asparagine, to give compound (**56**) (see Table 2.5), which again showed a reduction in potency, compared to the shorter analogue. Having observed this reduction in potency for P_3 containing inhibitors, the Boc group at this position was replaced. Literature precedent indicated that a suitable group in the P_3 position was the quinoline-2-carbonyl (Qua) moiety (see below, Figure 2.14). This residue was first used successfully in the highly potent Ro 31-8959 inhibitor (see section **1.4.6 (v) (15)**).

Figure 2.14. The quinoline-2-carbonyl (Qua) moiety.

Thus compound (**60**) (see Table 2.5) was prepared from compound (**55**) and Qua-(2S)-asparagine (**59**) using mixed anhydride chemistry. Qua-(2S)-asparagine (**59**) was synthesised in a simple one pot reaction from quinaldic acid chloride, using asparagine under modified Schotten-Baumann conditions.[154]

In accord with the observations reported in the literature, introduction of the Qua residue in the P_3 position caused an increase in potency compared to compound (**56**). However, compound (**60**) was 20 fold less potent than compound (**51**), again highlighting the preference for the shorter analogue. Clearly, the nature of the optimal P_3 residue is different for the phosphonamidate methyl ester analogues than for the more potent hydroxyl containing inhibitors. In some cases occupation of the S_3 pocket of the enzyme can cause conformational effects that can improve the key interactions at the core of the enzyme's active site. This may explain why longer hydroxyl containing inhibitors are more potent than those lacking such a P_3 residue, since their hydroxyl groups may be pushed closer to the active site aspartates, therefore, optimising the hydrogen-bonding interactions. Pushing the methyl group of the phosphonamidates closer to the catalytic aspartates would cause an adverse interaction, since the methyl esters lack a hydrogen-bond donor.

In a similar manner the Cbz group from compound (**50**) was removed and and the N-terminal extended with Qua-(2S)-asparagine (**59**). After chromatographic resolution a pair of diastereomers were obtained, with the most potent diastereomer possessing an IC_{50} value of 20 μM (compound (**61**), see Table 2.5). This result, in agreement with the finding for compound (**60**), showed that Qua improved inhibitor potency compared to the Boc group in the P_3 position, although the shorter compound (**50**) still maintained the greater potency.

At this point, it seemed that optimisation of the phosphonamidate methyl esters would be very difficult, as implemented alterations seemed to produce very small changes in potency. However, due to the close correlation of *in vivo* to *in vitro* IC_{50} values (the implications of which, will be discussed in section **2.5 (ii)**), the importance of the $P_{1'}$ position was investigated in more depth. Two approaches were pursued; first the (2S)-Phe residue in compound (**51**) was replaced with the unnatural (2R)-Phe residue (compound (**64**)) and second the benzyl side chain was removed by replacing the (2S)-Phe residue with a Gly residue (compound (**65**)).

Table 2.6 Test results for compounds (64), (65), (68) & (69).

Compound	**Configuration at (CH*,P)**	**IC_{50} (μM)**	
		in vitro	***in vivo***
Cbz-F-[ψ]-(2R)-F-I-NH-iBu.			
64	(mix. 4 diast.)	1	N.D.
Cbz-F-[ψ]-G-I-NH-iBu.			
65	(mix. 4 diast.)	>30	N.D.
Cbz-F-[ψ]-(2S)-F-NH-iBu.			
68	(mix. 4 diast.)	40	N.D.
Cbz-F-[ψ]-(2S)-F-(OMe).			
69	(mix. 4 diast.)	100	100

It was expected that a drop in potency would be observed when replacing the natural Phe residue with its unnatural counterpart (**64**). Surprisingly, this analogue was equally potent to compound (**51**). However, the literature

seemed to indicate that the $S_{1'}$ pocket of the enzyme, was accommodating to a variety of different residues of differing size. For instance, the highly potent hydroxyethylene inhibitor (compound (**23**)) discussed in section **1.4.6 (vi)** contained a styrene residue in this position. This particular residue is 2 carbon units longer than the benzyl side chain of the Phe residue and shows that the $S_{1'}$ subsite is large enough to bind groups that are bigger than the natural residue. In fact, the high potency of compound (**23**) probably arises from the increased hydrophobic interactions of the aromatic ring with those residues in the $S_{1'}$ subsite of the enzyme.

The removal of the benzyl side chain caused at least a 30 fold reduction in potency. An accurate IC_{50} value for compound (**65**) could not be obtained due to insolublity of the inhibitor at concentrations above 40 μM. This decrease in potency highlights the importance of the $P_{1'}$ residue and is in agreement with other findings (see section **1.4.6** for several examples). Finally, the effect of sequentially removing the $P_{3'}$ and $P_{2'}$ residues was examined. First, Ile-NH-iBu at $P_{2'}$ and $P_{3'}$ was replaced by isobutylamine, to give (**68**). This modification caused a 20 fold reduction in potency (see Table 2.6) compared to compound (**51**). Second, (2S)-Phe-Ile-NH-iBu was replaced by (2S)-Phe-(OMe) (**69**) and a 50 fold reduction in potency was observed (see Table 2.6) compared to compound (**51**). These results highlight the importance of the $P_{3'}$ and $P_{2'}$ residues for binding.

(i) Comparison of *in vitro* and *in vivo* assay results.

The most striking feature from the inhibitor assay results is the close 1:1 correlation of *in vivo* to *in vitro* IC_{50} values (see overleaf, Table 2.7). Generally, peptidic inhibitors lose potency in an *in vivo* assay due, primarily, to their inability to enter cells. The ratio of *in vivo* to *in vitro* IC_{50} values can be as great as 10,000:1 (see section **1.4.7** for a fuller discussion). The phosphonamidate methyl ester analogues show a marked ability to enter cells and this compensates for their moderate potency (μM) in the *in vitro* assay. Interestingly, the uptake properties appear to be totally independent of inhibitor size and composition. For instance, larger peptide inhibitors appear to be less effective than smaller peptide inhibitors in the *in vitro* assay (compare compounds (**44**) and (**42**)) and yet the 1:1 ratio is still observed for the larger analogues. For

other classes of inhibitor, the ratio of *in vivo* to *in vitro* IC_{50} values is much greater for larger peptide inhibitors than for smaller analogues (compare compounds (**16**) and (**22**) section **1.4.7**).

The only feature of our compounds that is different to the other more potent inhibitors is the transition state isostere. The residues chosen in the P_2-$P_{3'}$ positions are in common with many other inhibitors and so it appears that the phosphonamidate methyl ester group may be responsible for the uptake properties. This may explain why the 1:1 ratio is independent of size, as all of the analogues possess this functionality. Of greater relevance is the low anti viral activity of the aforementioned phosphinic acid and phosphinic acid ester inhibitors (see section **2.4 (i)**). These analogues are closest in terms of the transition state isostere to our compounds and yet do not exhibit our cell uptake properties. A study of the metabolism and mechanism of cell uptake would be needed in order to clarify this observation. For instance, metabolites could be isolated from cell lysates and characterised in order to check if the inhibitory effect was due to the inhibitor or a common metabolite. If the phosphonamidate methyl ester functionality was the key to efficient uptake, it would be worthwhile incorporating this group into the more potent hydroxyethylene inhibitors. In this way, improved uptake could be achieved whilst retaining the potency of the hydroxyethylene functionality.

Table 2.7. Summary of *in vitro* and *in vivo* inhibition data.

Compound	**Configuration at (CH*,P)**	**IC_{50} (μM)**	
		in vitro	***in vivo***
Boc-N-F-[ψ]-P-I-NH-iBu.			
44 i	(R, R)	30	45
44 ii	(R, S)	80	>100
44 iii	(S, S)	100	>100
44 iv	(S, R)	>100	>100
Cbz-F-[ψ]-P-I-NH-iBu.			
42 i	(R, R)	5	3
42 ii	(R, S)	25	25
42 ii	(S, S)	N.D.	25
42 iv	(S, R)	N.D.	>100

Cbz-Cha-[ψ]-P-I-NH-iBu.			
50	(mix. 2 diast.)	16	8
Cbz-F-[ψ]-(2S)-F-I-NH-iBu.			
51	(mix. 4 diast.)	2	2
Cbz-Cha-[ψ]-(2S)-F-I-NH-iBu.			
52	(mix. 4 diast.)	2	5
H_2N-F-[ψ]-(2S)-F-I-NH-iBu.			
55	(mix. 4 diast.)	30	20
Boc-N-F-[ψ]-(2S)-F-I-NH-iBu.			
56	(mix. 4 diast.)	100	>100
Qua-N-F-[ψ]-(2S)-F-I-NH-iBu.			
60	(mix. 4 diast.)	45	50
Qua-N-Cha-[ψ]-P-I-NH-iBu.			
61	(1 diast. unknown)	20	15
Cbz-F-[ψ]-(2R)-F-I-NH-iBu.			
64	(mix. 4 diast.)	1	N.D.
Cbz-F-[ψ]-G-I-NH-iBu.			
65	(mix. 4 diast.)	>30	N.D.
Cbz-F-[ψ]-(2S)-F-NH-iBu.			
68	(mix. 4 diast.)	40	N.D.
Cbz-F-[ψ]-(2S)-F-(OMe).			
69	(mix. 4 diast.)	100	100

(ii) Summary.

Although, phosphonamidate-containing peptide analogues have proved to be highly potent inhibitors of the zinc proteases, the same level of potency is not observed for the HIV-1 protease. This is probably due to their ionised nature at the assay pH (5.6), which causes a repulsive interaction between the negatively charged inhibitor and the negatively charged active site carboxylate. Phosphonamidates in general have an acid labile phosphorus-nitrogen bond and for this reason phosphonamidates offer no potential as therapeutic agents, due to their instability in the gut. Their ionised nature at physiological pH (7.4)

would prevent them from crossing cell membranes, as the lipid bilayers are highly impermeable to ionised molecules.

The problem of ionisation is not inherent in the phosphonamidate methyl ester analogues and consequently these neutral compounds have been shown to be moderate inhibitors (μM) of the HIV-1 protease. Due to their increased acid stability and unique cell uptake properties, these analogues were studied as potential anti-AIDS drugs. However, optimisation of the methyl ester analogues proved difficult, as the alteration of the side chain binding residues caused only small effects on the inhibitor potency. The stark trends observed for the more potent hydroxyethylene and hydroxyethylamine inhibitors were not followed by the phosphonamidate methyl esters. Optimisation of the hydroxyl-containing inhibitors is more easily achieved as these analogues possess the key features for inhibitor potency; hydrogen-bond donating and accepting ability. The phosphonamidate methyl ester analogues lack hydrogen-bond donating capacity and so optimisation towards the same level of potency as the hydroxyl-containing inhibitors, cannot be realised. However, these analogues may unveil a unique cell uptake mechanism which can be exploited in the development of future therapeutic agents and for this reason alone, warrant further investigation.

2.6 Non-peptidic inhibitors.

2.6.1 Introduction.

As an alternative to peptide-based inhibitors, the syntheses towards two novel non-peptidic HIV-1 protease inhibitors was undertaken. Selective, non-peptidic inhibitors of the HIV-1 protease offer clear advantages over peptide-based inhibitors (the problems associated with peptide-based inhibitors have been discussed in section **1.4.7**). By removal of the peptide bonds and reduction of the molecular weight of inhibitors, it was hoped that problems with short duration of action, rapid clearance and solubility could be eliminated. This task is far from easy considering that the actual peptide bonds, play a key role in the hydrogen-bonding between inhibitor and enzyme. However, it appears that the key interactions occur within the P_1 and $P_{1'}$ region and therefore, residues in the P_2, P_3, $P_{2'}$ and $P_{3'}$ seem to play only a secondary role in inhibitor binding.

2.6.2 Design of non-peptidic inhibitors.

X-ray crystallographic data have provided useful information with regard to the key hydrogen-bonding interactions within the active site of the HIV-1 protease (see Figure 1.16, section **1.4.8**). Studies of peptide-based inhibitors have identified the hydroxyethylene and hydroxyethylamine analogues as highly potent inhibitors of the HIV-1 protease. These inhibitors contain secondary hydroxyl groups which replace the nucleophilic water molecule (see Figure 1.8, section **1.4.5**) associated with the aspartic acid residues at the active site of the enzyme. These hydrogen-bonding interactions appear critical for inhibitor potency, as similar inhibitors that lack this hydroxyl tend to be poorer inhibitors of the HIV-1 protease. A second feature in the design of HIV-1 protease inhibitors first alluded to by Swain *et al.*,[122] involves the flap water molecule. This flap water molecule has a major role in hydrogen-bonding both to the substrate/ inhibitor and the enzyme flaps (see Figure 1.16, section **1.4.8**). Rationally replacing this flap-binding water molecule in the HIV-1 protease offers the opportunity to design selective inhibitors without relying on residues far removed from the central catalytic region of the active site.

In simple terms a compound which incorporated a hydroxyl group, residues to interact at the S_1 and $S_{1'}$ sites and a heteroatom to replace the flap water molecule was required. It became immediately apparent that a cyclic structure, which incorporated all of these features was necessary. This type of inhibitor should bind in the active site whilst maintaining interactions with the aspartates, in the S_1 and $S_{1'}$ sites and the flap region (see overleaf, Figure 2.15).

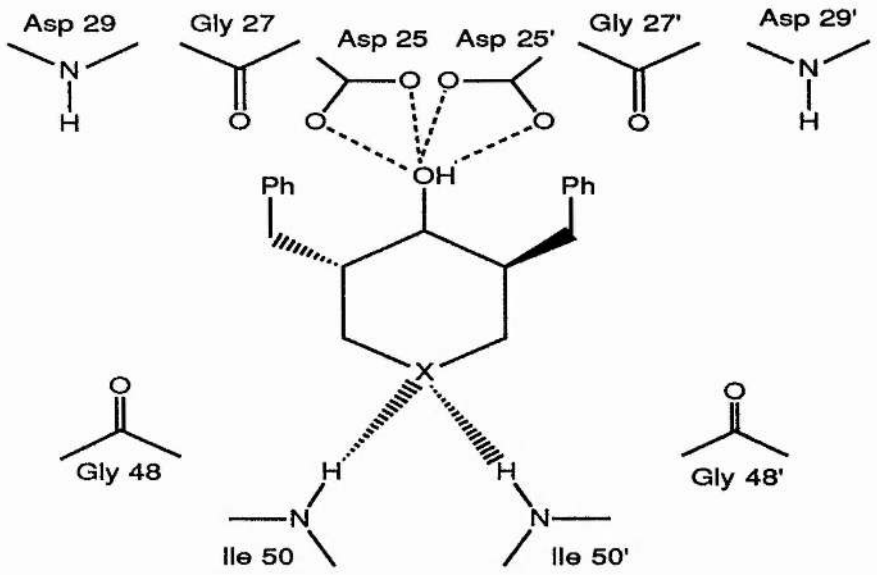


Figure 2.15. Proposed hydrogen-bonding interactions of the cyclic inhibitors.

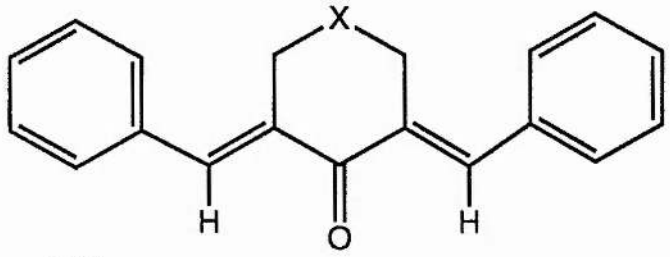
From a literature study and results from our own work on the phosphorus-based inhibitors, benzyl and cyclohexylalanine were identified as desirable residues in the P_1 and $P_{1'}$ positions. Additionally, the C-2 symmetry of the inhibitor was maintained as a requirement for selectivity (see section **1.4.6 (viii)**).

Incorporation of the heteroatom within the ring fixes the lone pair geometry of this atom and so enables the correct orientation for hydrogen-bonding to the flap amide hydrogens of Ile-50 and Ile-50' (as shown in Figure 2.15). The distance between the catalytic aspartates and the flap water molecule is approximately 5.5 Å and so places a constraint on inhibitor design, although some flexibility will be tolerated as the flap region is known to move as much as 7 Å upon binding of an inhibitor (see section **1.4.8**).

An obvious starting point would be to synthesise a tetrahydropyranol (X=O) analogue. However, modelling studies by Gani revealed that the methylene groups α to the oxygen atom would probably interact adversely (i.e. bang against) the flap residues. Replacing the oxygen atom with a sulphur atom seemed to overcome this problem, as the carbon-sulfur bonds (1.80 Å) were longer than the corresponding carbon-oxygen bonds (1.40 Å). Additionally, the the lone pairs on the sulfur atom are larger and so should be more accessible

for hydrogen-bonding to the flap amide hydrogens of Ile-50 and Ile-50'. However, molecular modelling can over estimate steric repulsion and so analogues in which X = carbon, oxygen and sulfur were synthesised.

Using simple base catalysed aldol chemistry (reviewed in[155]) adapted from an undergraduate procedure, analogues were prepared that contained Phe mimics at P_1 and $P_{1'}$. Reaction of tetrahydrothiopyran-4-one, tetrahydropyran-4-one and cyclohexanone with benzaldehyde under basic conditions yielded the the cyclic sulfur (**70**), oxygen (**71**) and carbon (**72**) ketone analogues in 62, 65 and 77% respectively (see below, Figure 2.16).

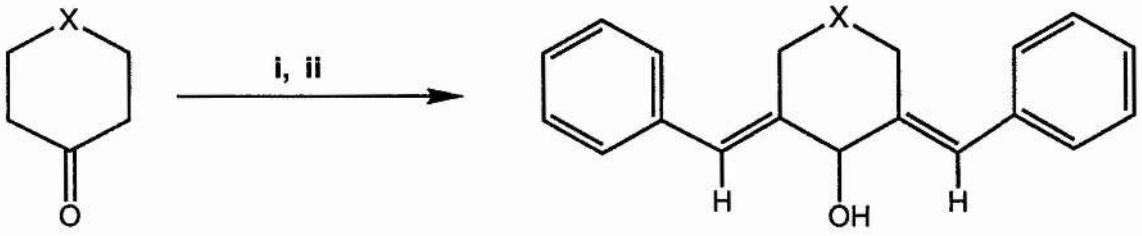


Where X = S, O and CH_2.

Figure 2.16. Cyclic ketone precursors (70)-(72).

The oxygen and sulfur ketone analogues were first prepared by Leonard *et al.*[156] and have subsequently been used in the synthesis of potential anti-inflammatory agents.[157] Our base catalysed method provides a milder alternative to the literature where these ketones have been synthesised using the same reagents under acidic conditions. Except in the case of the sulfur analogue our yields are better as well.

These ketones were selectively reduced with sodium borohydride to yield the respective alcohols (**73**) and (**74**) in good yield (> 90%) (see overleaf, Scheme 2.5).

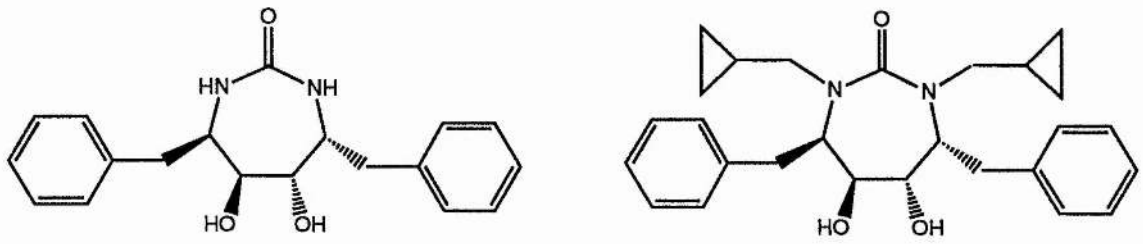


Scheme 2.5. Synthesis of the cyclic alcohol analogues (73) & (74).

(i) 2 eq. Benzaldehyde, NaOH, EtOH, r.t.; (ii) $NaBH_4$, EtOH, r.t. then H_2O.

Although these cyclic alcohols did not possess the desired sp^3 geometry at P_1 and $P_{1'}$ (see Figure 2.15) they provided useful leads. Reduction of the double bonds was attempted using hydrogen gas over a palladium/ carbon catalyst but produced a variety of isomers which proved impossible to separate in the short term. However, the sulfur analogue (**73**) (shown in Scheme 2.5) was found to be a moderate inhibitor of the HIV-1 protease (IC_{50} ~ 40 μM). An accurate IC_{50} value of (**73**) could not be achieved due to inhibitor insolubility at concentrations > 30 μM. A similar solubility problem was observed with the oxygen analogue (**74**) and so as of yet, no other IC_{50} values have been produced. However, these analogues are currently being fully investigated in the laboratory by Donald McNab. Future work includes oxidation of the sulfur, to provide sulfoxide and sulfone analogues. These analogues will be used to probe the flap interactions, as the additional distance provided by the oxygen atoms may improve hydrogen-bonding interactions with the flap. These analogues are likely to have improved aqueous solubility over the cyclic sulfide as well. Incorporation of sp^3 character, with the correct stereochemistry at P_1 and $P_{1'}$ will be attempted. Introduction of side chains α to the sulfur should improve inhibitor potency by binding in the S_2 and $S_{2'}$ subsites. Introduction of a hydrogen-bond donor on the aromatic ring (e.g OH or NH_2) of the P_1 and $P_{1'}$ residues may improve potency by hydrogen-bonding to the carbonyls of Gly-27 and Gly-27' (see Figure 2.15). Modelling studies are currently in progress to determine whether this type of interaction can be achieved and to predict at which position on the aromatic ring the hydrogen-bond donor needs to be located.

Towards the end of this work our attention was drawn to the following compounds[158] (see below, Figure 2.17). The core cyclic urea analogue (**iii**) was identified using crystallographic structural information from several HIV-1 protease-inhibitor complexes and 3D data base searching.



(iii) **(iv)**

Figure 2.17. Non-peptidic, cyclic urea HIV-1 protease inhibitors.

The core structure (**iii**) (K_i = 3 μM) was then optimised to yield a variety of highly potent inhibitors. The cyclic urea analogue (**iv**) (K_i = 3 nM) possessing cyclopropylmethyl side chains was found to be 100% orally bioavailable in rats and provides an extremely selective non-peptidic inhibitor of the HIV-1 protease. X-ray crystallographic studies revealed that these novel, cyclic inhibitors linked the aspartates to the flexible flaps *via* a hydrogen-bond network that did not include an intervening water molecule. Thus, these inhibitors had indeed displaced the flap water molecule and the X-ray study confirmed that these analogues were functioning in a manner, predicted in the design of our cyclic sulfur analogues.

Having examined these cyclic urea analogues, it appeared that these compounds needed to rotate by 90° in order to optimise the lone pair geometry of the urea carbonyl with respect to the flap amide hydrogens. Thus, the side chains appeared to be important factors in this process. On the basis of the cyclic urea analogues, the following inhibitor was designed (see overleaf, Figure 2.18). Introduction of sp^2 character within the ring served two purposes; First, the P_1 and P_1' residues need to rotate in order to bind in the S_1 and S_1' subsites, thus optimising the orientation of the lone pair geometry of the ketone carbonyl. Second, the double bonds within the ring should provide a

conformation in which both the ketone carbonyl and the hydroxyls lie approximately in the same plane. This is a requirement as a line through the nucleophilic water molecule and the flap-binding water molecule in the active site of the enzyme forms an axis of nearly C-2 symmetry.

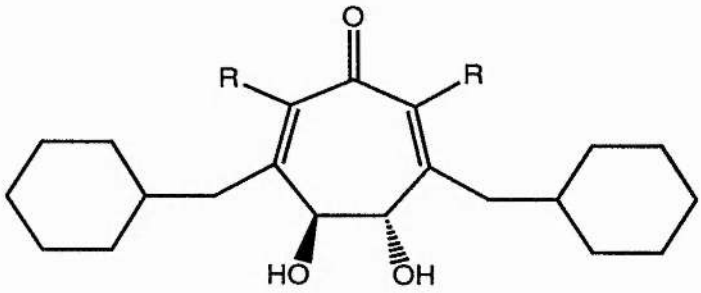
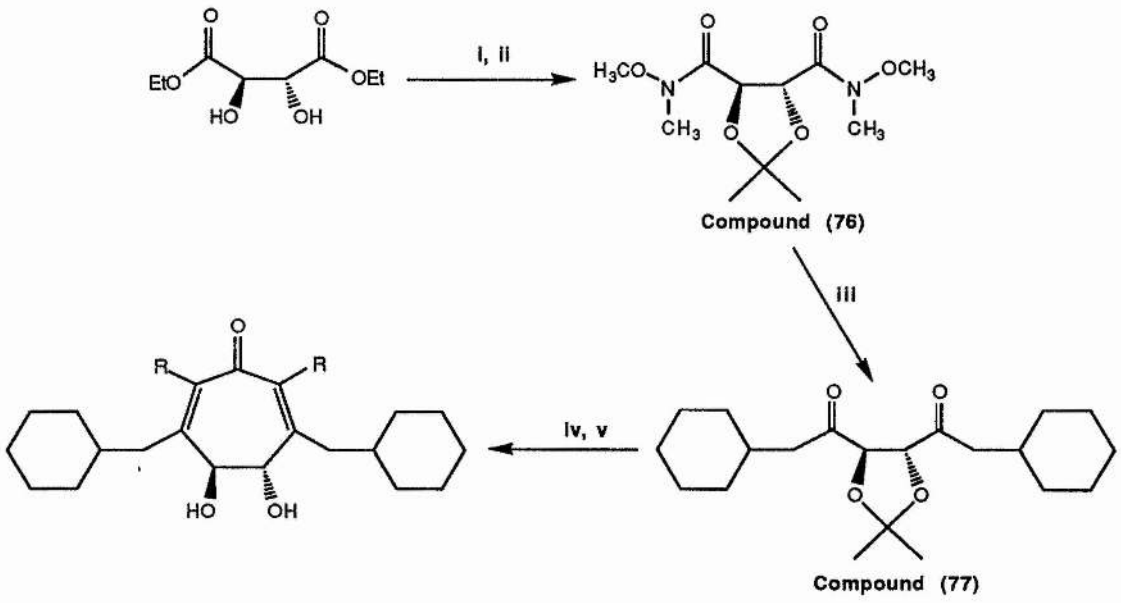


Figure 2.18. Novel, C-2 symmetrical non-peptidic inhibitor.

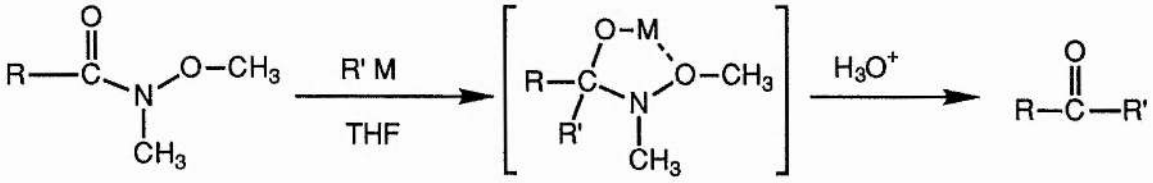
From retrosynthetic analyses of this inhibitor, several synthetic routes were identified. Of these, the shortest route containing arguably the most interesting chemistry was employed (see below, Scheme 2.6).



Scheme 2.6. Proposed synthetic route to cyclic inhibitor.

(i) Acetone, conc. H_2SO_4; (ii) $NH(OMe)CH_3$. HCl, NaOMe, MeOH; (iii) cyclohexylmethyl MgBr, ether, 0 °C; (iv) RCH_2COCH_2R, base, then diketone (**77**); (v) aqueous acid.

Due to the C-2 nature of the molecule, *bis*-chemistry could be used in order to minimise the number of reactions required. Diethyl L-tartrate already possesses the required (R,R)-absolute stereochemistry for the hydroxyls and so was identified as an ideal precursor. The ketal (**75**) was formed in 60% yield using acetone and a catalytic amount of conc. H_2SO_4 by the method of Ikemoto.[159] However, this extremely simple preparation was abandoned as optical rotation experiments revealed epimerisation. Similar methods of ketal protection had also produced epimerisation.[160] However, ketal protected dimethyl L-tartrate was available from Aldrich and so was subsequently used in the preparation of the diketone (**77**). The ketal protected *bis*-N-methoxy-N-methyl amide (**76**) was identified as an excellent precursor to the Grignard reaction. Weinreb[161] had first used these amides in the preparation of ketone and aldehydes. A major problem in the use of organometallics to form ketones, is the formation of unwanted tertiary alcohol. Due to the reactivity of the organometallic reagent and the electrophilic nature of the ketone, it is difficult to prevent further addition of the organometallic reagent to the ketone. However, the reaction involving the Weinreb amide proceeds through a very stable metal-chelated intermediate (see below, Scheme 2.7) and therefore even the use of excess Grignard does not result in tertiary alcohol formation.

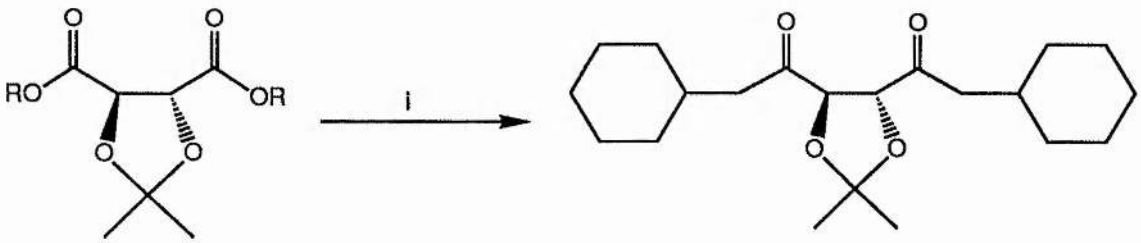


Scheme 2.7. Ketone formation *via* Weinreb amides.

Several routes towards the *bis*-amide (**76**) were attempted. These included *bis*-activation of L-tartaric acid *via* acid chloride and mixed anhydride formation. Both methods proved unsuccessful due primarily to insolubilty problems. Hence a method that used L-tartaric acid esters was needed. In this way the insolubility problems associated with the L-tartaric acid could be overcome. Seebach[162] had synthesised the *bis*-dimethyl amide from diethyl L-tartrate and so this method seemed promising. N, O-dimethylhydroxylamine should react in a similar manner. However, reaction of diethyl L-tartrate and ketal protected

diethyl L-tartrate (**75**) with N, O-dimethylhydroxylamine proved unsuccessful.

The synthesis of the ketal protected *bis*-cyclohexylmethyl ketone (**77**) was achieved by an alternative reaction using the method of Kikkawa *et al.*[163] This method involves the reaction of Grignard reagents with carboxylic esters, using triethylamine to prevent tertiary alcohol formation. The exact role of the triethylamine is not well understood, although it is believed to prevent tertiary alcohol formation by enolisation of the resultant ketone. Alternatively, the triethylamine may complex to the Grignard therefore lowering its reactivity. This reaction was attempted using ketal (**75**) and cyclohexylmethylmagnesium bromide in ether and after silica chromatography afforded the desired diketone (**77**) in 21% yield (see below, Scheme 2.8). An identical reaction in the absence of triethylamine, produced predominantly *bis*-tertiary alcohol (identified by ^{1}H NMR) and no diketone (**77**).



Compound (77)

Where R = Methyl or Ethyl

Scheme 2.8. Synthesis of the diketone (77).

(i) 2.5 eq. Cyclohexylmethyl MgBr, 6 eq. Et_3N, ether, 0 °C.

This reaction was successfully repeated on several occasions using previously prepared diethyl L-tartrate ketal (**75**) and the purchased dimethyl L-tartrate ketal. The best yield so far achieved was 32% after purification. The reaction was also successfully completed in THF, although this result has not been repeated by Neil Piggot. An insoluble solid which proved difficult to remove in the work up, appeared during the reaction in ether and therefore THF seemed a suitable alternative. However, Piggot has consistently produced the half tertiary alcohol, half ketone adduct probably due to the increased reactivity of the Grignard in THF. At present, a combination of THF/ ether is being used in

order to improve the yield of this step.

The key step of the synthesis has yet to be achieved and involves reaction of the novel diketone (**77**) to form the 7 membered ring system. Preformed enolates of activated ketones and tricarbonyl analogues were used in an attempted reaction with the diketone (**77**). If a reaction at one of the diketone carbonyls could be achieved, cyclisation, producing the 7 membered ring, should occur in an intramolecular fashion. Cyclohexylalanine was chosen in preference to phenylalanine residues due to the decreased acidity of the protons α to the carbonyl group. In this way the problem of intramolecular aldol-type condensation of the diketone could be avoided.

Reaction of the enolate of dimethyl acetonedicarboxylate (preformed using NaOMe) with the diketone (**77**) did not occur at room temperature. However, repeating the reaction under refluxing conditions produced a UV active spot by TLC. However, after purification and characterisation it was discovered that this reaction had produced a substituted cyclopentenone derivative (**78**) (see below, Figure 2.19). This was caused by the aforementioned intramolecular cyclisation of diketone (**77**).

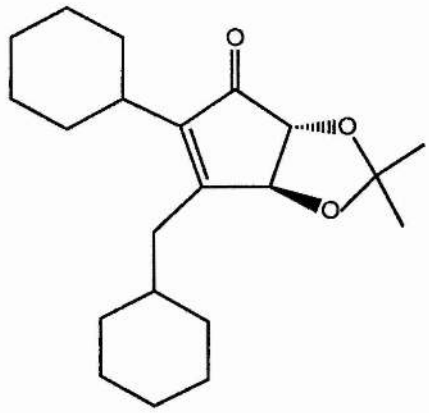


Figure 2.19. Substituted cyclopentenone (78), produced from the intramolecular cyclisation of diketone (77).

This reaction was a disappointment, although it highlights the synthetic utility of this type of diketone (**77**). For instance, this reaction gives an extremely short route to substituted cyclopentenones that contain a protected diol system with fixed absolute stereochemistry. The ketal can easily be removed to yield the subsequent diol. The absolute stereochemistry of the diol can be altered simply

by using dimethyl D-tartrate or *meso*-dimethyl tartrate and variation of the Grignard reagent can be used to alter the side chains.

An examination of the literature revealed the reaction of dimethyl acetonedicarboxylate with 1, 2-dicarbonyl compounds to yield cyclopentadienones. Cook and Weiss[164] have used this methodology in the preparation of polyquinanes and Eguchi *et al.*[165] have used dibenzyl ketone in the same type of condensation reaction. Although, 1, 2-dicarbonyl compounds would be more reactive towards this type of condensation reaction than our diketone (**77**), this reaction was attempted under the same conditions (NaOH and EtOH), using both dimethyl acetonedicarboxylate and dibenzyl ketone. After a few days, no desired cyclisation had occurred, except for a small amount of the cyclopentenone derivative (**78**) (see Figure 2.19).

Unfortunately, the total synthesis of our target molecule has not been achieved by myself in the short term. This molecule is currently being pursued by Neil Piggot in our laboratory and future work includes the use of more reactive dianions in an attempt to achieve cyclisation. Reaction of the dianion of the Wittig reagent, diethyl (2-oxopropyl)-phosphonate (shown below in Figure 2.20) with the diketone (**77**) will be attempted under kinetic conditions.

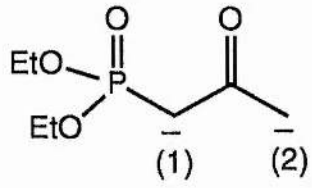


Figure 2.20 Dianion required for the cyclisation.

The reaction of this dianion should proceed by an initial attack of the more reactive anion (2), followed by the intramolecular reaction of (1). If successful, a reaction with the corresponding dianion of dibenzyl ketone will be attempted under the same conditions. This allows the introduction of side chains, which can bind in the S_2 and $S_{2'}$ subsites.

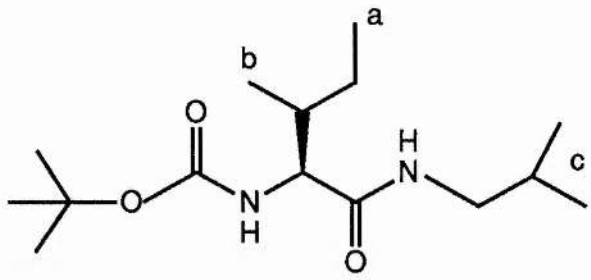
CHAPTER THREE

EXPERIMENTAL

3.0 Experimental.

NMR Spectra were recorded on a Bruker AM-300 f.t. spectrometer (^1H, 300 MHz; ^{13}C, 75 MHz; ^{31}P, 121.5 MHz) and a Varian Gemini f.t. spectrometer (^1H, 200 MHz; ^{13}C, 50 MHz). High field NMR spectra were obtained on a service basis at the University of Edinburgh (^1H, 600 MHz; ^{13}C, 150 MHz; ^{31}P, 243 MHz) ^1H NMR spectra were referenced on chloroform, methanol or DMSO, ^{13}C NMR spectra were referenced on chloroform, methanol or DMSO and ^{31}P spectra on external H_3PO_4. NMR spectra are described in parts per million downfield shift from TMS and are reported consecutively as position (δ_H or δ_C), relative integral, multiplicity (s-singlet, d-doublet, t-triplet, q-quartet, m-multiplet, d. of d.-doublet of doublets and br-broad), coupling constant ($J_{X,Y}$ Hz if applicable) and assignment. Infrared spectra were recorded using a Perkin Elmer 1420 ratio recording spectrometer and a Perkin Elmer 1710 f.t. i.r. spectrometer. The samples were prepared as nujol mulls or thin films between sodium chloride discs. Absorption maxima are given in wavenumbers (cm^{-1}) relative to a polystyrene standard. Melting points were measured using electrothermal melting point apparatus and are uncorrected. Optical rotations were measured on an Optical Activity Ltd. AA-100 polarimeter using 10 cm path length cells at room temperature. Mass spectra were recorded on a Kratos MS50 and obtained on a S.E.R.C. service basis at the University of Swansea using a VG ZAB E. Major fragments are given as percentages of the base peak intensity. Where appropriate, all solvents and reagents were freshly distilled prior to use. THF and ether were distilled from sodium/ benzophenone under a nitrogen atmosphere; DMF and CH_2Cl_2 were distilled from CaH_2. Flash chromatography was performed according to the procedure of Still[168] using Sorbisil C60 (40-60 μm) silica gel. Analytical thin layer chromatography was carried out on 0.25 mm precoated siliga gel plates (MN SIL G/UV_{254}) or on 0.1 mm precoated cellulose plates (CEL MN 300-10 /UV_{254}), and compounds were visualised by UV fluorescence, iodine vapour, ethanolic phosphomolybdic acid, aqueous potassium permanganate or ninhydrin.

Boc-(2S)-isoleucyl-isobutylamide (37).

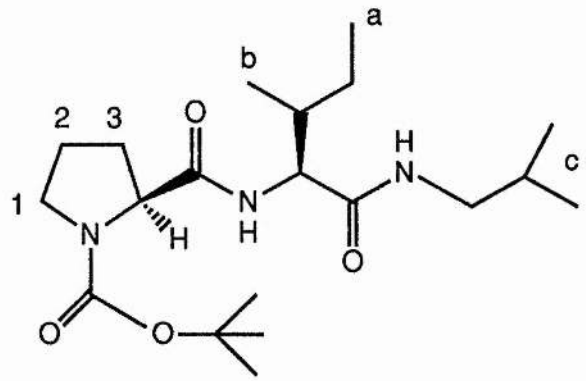


Boc-(2S)-isoleucine (4.62 g, 20 mmol) was dissolved in anhydrous THF (100 ml), N-methylmorpholine (2.18 ml, 20 mmol) was added and the solution cooled to -15 °C. Isobutylchloroformate (2.72 ml, 20 mmol) was added with stirring. The solution was left for 5 min. Isobutylamine (1.99 ml, 20 mmol) was added at -15 °C and the solution left for 5 min. The solution was allowed to stir at room temperature for a further 15 min. The hydrochloride salts were filtered off and the THF removed under reduced pressure to yield a white solid which was recrystallised from methanol/ water (4.85 g, 85%), m.p. 158-160 °C; (Found: C, 63.07; H, 10.57; N, 9.82. $C_{15}H_{30}N_2O_3$ requires C, 62.90; H, 10.56; N, 9.78%); *m/z* (Found: $[M + H]^+$ 287.2335. $C_{15}H_{31}N_2O_3$ requires 287.2335); $[\alpha]_D$ -27.1° (c 1 in MeOH); ν_{max} (nujol)/ cm^{-1} 3335 (NH str.), 1682 (CO str. (urethane)), 1655 (CO str. (amide)) and 1526 (NH bend); δ_H (200 MHz, C^2HCl_3) 0.9 (12H, m, CH_3's (a,b,c)), 1.10 (1H, m, 1 x CH_2 (Ile)), 1.42 (9H, s, Boc), 1.50 (1H, m, 1 x CH_2 (Ile)), 1.7-1.9 (2H, m, CH (iBu) & CH (Ile)), 2.98-3.20 (2H, m, CH_2 (iBu)), 3.84 (1H, d. of d., *J* 7.0, 9.0 Hz, α H), 5.08 (1H, d, *J* 8.6 Hz, NH (urethane)) and 6.10 (1H, t, *J* 5.0 Hz, NH (amide)). δ_C (50.3 MHz, C^2HCl_3) 12.0 ($\underline{C}H_3$ (a)), 16.2 ($\underline{C}H_3$ (b)), 21.0 ($\underline{C}H_3$'s (c)), 25.2 ($\underline{C}H_2$ (Ile)), 28.9 (Boc $\underline{C}H_3$'s), 29.0 ($\underline{C}H$ (iBu)), 37.9 ($\underline{C}H$ (Ile)), 47.4 ($\underline{C}H_2$ (iBu)), 60.2 (α $\underline{C}H$), 80.8 ($\underline{C}(CH_3)_3$), 156.6 ($\underline{C}O$ urethane) and 172.2 ($\underline{C}O$ amide); *m/z* (CI) 287 ($[M + H]^+$, 8%), 231 ($[M + H - (H_3C)_2CCH_2]^+$, 48), 213 ($[M + H - (H_3C)_2CCH_2 - H_2O]^+$, 42) and 187 ($[M + H - Boc]^+$, 100).

(2S)-Isoleucyl-isobutylamide hydrochloride (38).

HCl gas was bubbled into ethyl acetate (150 ml) at 0 °C for 1 h. Boc-(2S)-Ile-NH-iBu (**37**) (4.94 g, 17 mmol) was added to the acidic solution and left for 1.5 h. at room temperature. The solvent was removed under reduced pressure to yield an off white solid (3.45 g, 90%) that was used immediately in the preparation of Boc-(2S)-Pro-(2S)-Ile-NH-iBu (**39**), *m/z* (Found: [*M* + H - HCl]$^+$ 187.1810. $C_{10}H_{23}N_2O$ requires 187.1810); ν_{max} (nujol)/ cm^{-1} 3200 & 3080 (NH str.), 1668 (CO str.) and 1568 & 1520 (NH bend); δ_H (200 MHz, C^2HCl_3) 0.9 (12H, m, CH_3's (a,b,c)), 1.15 (1H, m, 1 x CH_2 (Ile)), 1.55 (1H, m, 1 x CH_2 (Ile)), 1.80 (2H, m, CH (iBu) & CH (Ile)), 2.82 & 3.10 (2H, m, CH_2 (iBu)), 3.65 (1H, d, *J* 4.8 Hz, α CH), 8.3 (3H, br, NH_3^+) and 8.64 (1H, t, *J* 5.4 Hz, NH (iBu)); δ_C (50.3 MHz, C^2HCl_3) 12.02 ($\underline{C}H_3$ (a)), 15.17 ($\underline{C}H_3$ (b)), 20.89 ($\underline{C}H_3$'s (c)), 25.42 ($\underline{C}H_2$ (Ile)), 28.69 ($\underline{C}H$ (iBu)), 37.19 ($\underline{C}H$ (Ile)), 47.86 ($\underline{C}H_2$ (iBu)), 58.48 (α $\underline{C}H$) and 168.91 ($\underline{C}O$ amide); *m/z* (CI) 187 ([*M* + H - HCl]$^+$, 100%).

Boc-(2S)-prolyl-(2S)-isoleucyl-isobutylamide (39).



Boc-(2S)-proline (430 mg, 2 mmol) was dissolved in anhydrous THF (20 ml), N-methylmorpholine (218 µl, 2 mmol) was added and the solution cooled to -15 °C. Isobutylchloroformate (272 µl, 2 mmol) was added with stirring and the solution left for 5 min. at -15 °C. (2S)-Ile-NH-iBu hydrochloride (**38**) (450 mg, 2 mmol) and N-methylmorpholine (218 µl, 2 mmol) were dissolved in THF and added to the cold solution. The reaction was left at -15 °C for 5 min. and a further 15 min. at room temperature. The hydrochloride salts were filtered off

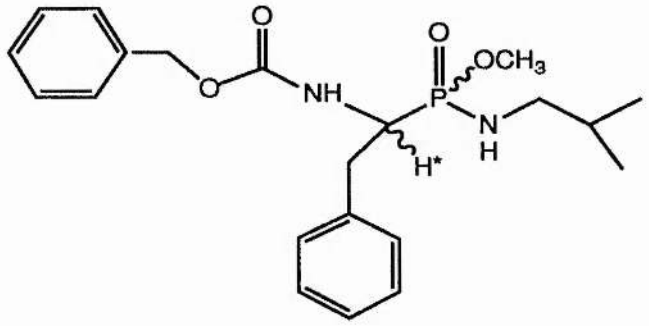
and the THF removed under reduced pressure to yield a white solid which was recrystallised from ethyl acetate/ light petroleum (698 mg, 91%), m.p. 126-128 °C; *m/z* (Found: $[M + H]^+$ 384.286. $C_{20}H_{38}N_3O_4$ requires 384.2862); $[\alpha]_D$ -85.5° (c 1 in MeOH); ν_{max} (nujol)/ cm^{-1} 3292 & 3089 (NH str.), 1706 (CO str. (urethane)), 1645 (CO str. (amide)) and 1553 (NH bend); δ_H (200 MHz, C^2HCl_3) 0.85 (12H, m, CH_3's (a,b,c)), 1.05 (1H, m, 1 x CH_2 (Ile)), 1.40 (1H, m, 1 x CH_2 (Ile)), 1.45 (9H, s, Boc), 1.7-2.3 (6H, m, CH (iBu), CH (Ile) & CH_2's (2) & (3) (Pro)), 2.85-3.2 (2H, m, CH_2 (iBu)), 3.45 (2H, br, CH_2 (1) (Pro)), 4.25 (2H, br, α CH's) and 6.67 & 6.94 (2H, br, NH's); δ_C (50.3 MHz, C^2HCl_3) 12.0 ($\underline{C}H_3$ (a)), 16.0 ($\underline{C}H_3$ (b)), 20.8 ($\underline{C}H_3$'s (c)), 24.6 ($\underline{C}H_2$ (2) Pro), 25.2 ($\underline{C}H_2$ (Ile)), 28.6 ($\underline{C}H$ (iBu)), 28.7 (Boc $\underline{C}H_3$'s), 29.3 ($\underline{C}H_2$ (3) Pro), 36.4 ($\underline{C}H$ (Ile)), 47.4 ($\underline{C}H_2$ (iBu)), 47.7 ($\underline{C}H_2$ (1) Pro), 58.6 (α $\underline{C}H$ (Ile)), 61.0 (α $\underline{C}H$ (Pro)), 81.3 ($\underline{C}(CH_3)_3$), 156.2 ($\underline{C}O$ urethane), 171.4 ($\underline{C}O$ (Ile)) and 172.4 ($\underline{C}O$ (Pro)); *m/z* (CI) 384 ($[M + H]^+$, 60%), 328 ($[M + H - (H_3C)_2CCH_2]^+$, 66) and 284 ($[M + H - Boc]^+$, 53).

(2S)-Prolyl-(2S)-isoleucyl-isobutylamide hydrochloride (40).

This deprotection was carried out in an identical manner to (**38**), using (**39**) (1.0 g, 2.6 mmol) in acidic ethyl acetate (50 ml) to yield a white solid, which was recrystallised from methanol/ ethyl acetate (675 mg, 81%), m.p. 220 °C (dec.); (Found: C, 56.29; H, 9.23; N, 12.93. $C_{15}H_{30}N_3O_2Cl$ requires C, 56.39; H, 9.47; N, 13.16%); *m/z* (Found: $[M + H - HCl]^+$ 284.2338. $C_{15}H_{30}N_3O_2$ requires 284.2338); $[\alpha]_D$ -66.1° (c 1 in MeOH); ν_{max} (nujol)/ cm^{-1} 3280 (NH str.), 1655 & 1637 (CO str. (amides)) and 1538 (NH bend); δ_H (600 MHz, d_6-DMSO) 0.85 (12H, m, CH_3's a,b,c), 1.10 (1H, m, 1 x CH_2 (Ile)), 1.45 (1H, m, 1 x CH_2 (Ile)), 1.67 (1H, m, CH (iBu)), 1.74 (2H, m, 1 x CH_2 (3) (Pro) & CH (Ile)), 1.84 (2H, m, CH_2 (2) (Pro)), 2.29 (1H, m, 1 x CH_2 (3) (Pro)), 2.80 (1H, m, 1 x CH_2 (iBu)), 2.95 (1H, m, 1 x CH_2 (iBu)) 3.15-3.3 (2H, m, CH_2 (1) (Pro)), 4.18 (1H, d. of d., *J* 7.80, 8.50 Hz, α CH (Ile)), 4.23 (1H, d. of d., *J* 6.90, 8.10 Hz, α CH (Pro)), 8.15 (1H, t, *J* 5.85 Hz, NH (iBu) and 8.65 (1H, d, *J* 8.50 Hz, NH (Ile)); δ_C (50.3 MHz, d_6-DMSO) 11.0 ($\underline{C}H_3$ (a)), 15.3 ($\underline{C}H_3$ (b)), 20.0 ($\underline{C}H_3$ (c)'s), 23.6 ($\underline{C}H_2$ (2) (Pro)), 24.4 ($\underline{C}H_2$ (Ile)), 28.0 ($\underline{C}H$ (iBu)), 29.9 ($\underline{C}H_2$ (3) (Pro)), 36.8 ($\underline{C}H$ (Ile)), 45.4 ($\underline{C}H_2$ (1) (Pro)), 45.9 ($\underline{C}H_2$

(iBu)), 57.8 (α CH (Ile)), 58.6 (α CH (Pro)), 168.2 (CO (Ile)) and 170.8 (CO (Pro)); *m/z* (CI) 284 ([*M* + H - HCl]$^+$, 100%).

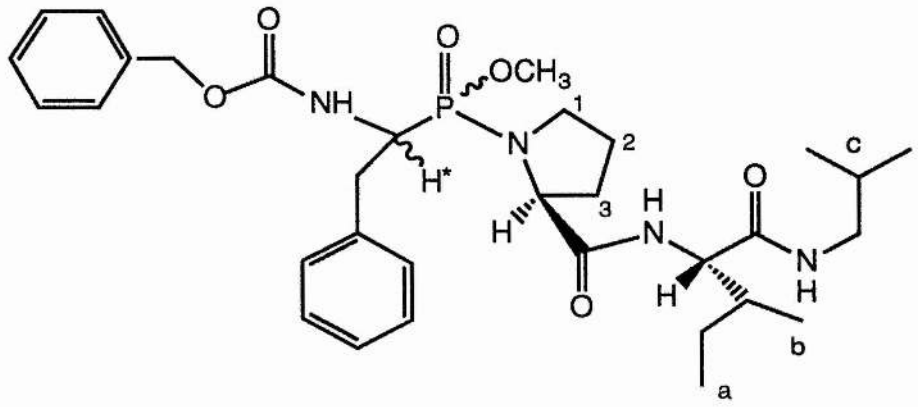
Cbz-Phe-ψ[P(OMe)O-NH]-iBu (41).



Using a modification of the method of Bartlett,[139] methyl hydrogen [1-(N-benzyloxycarbonyl)-amino]-benzylphosphonate (**35**) (698 mg, 2 mmol) was dissolved in anhydrous CH_2Cl_2 (10 ml). Freshly distilled thionyl chloride (300 µl, 4 mmol) was added and the solution flushed with argon for 1 min. After 4 h., the solution was flushed again and the solvent removed under reduced pressure. The colourless oil was rediluted with anhydrous CH_2Cl_2 (5 ml) and reflushed with argon. This process was repeated and the colourless oil redissolved in CH_2Cl_2 (20 ml). Isobutylamine (500 µl, 5mmol) was added to the chlorophosphonate at 0 °C and left to react at room temperature for 24 h. under an atmosphere of nitrogen. The organic solution was washed with 1 M HCl (2 x 30 ml), saturated sodium bicarbonate (2 x 30 ml) and water (30 ml) and then dried over anhydrous sodium sulphate. After filtration, the solvent was removed under reduced pressure to yield an off white solid, which was recrystallised from ethyl acetate/ light petroleum (650 mg, 80%), *m/z* (Found: [*M* + H]$^+$ 405.1943. $C_{21}H_{30}N_2O_4P$ requires 405.1943); δ_H (200 MHz, C^2HCl_3) 0.85 (6H, d, *J* 7.5 Hz, CH_3's), 1.60 (1H, m, CH (iBu)), 2.5-2.8 (2H, m, CH_2 (iBu)), 2.9 (1H, m, 1 x ArC$\underline{H}_2$C), 3.2 (1H, m, 1 x ArC$\underline{H}_2$C), 3.7 (3H, d, $J_{P,H}$ 10.6 Hz, PO(OCH_3)), 4.3 (1H, m, CH*), 4.95 (2H, s, ArC$\underline{H}_2$O), 5.40 (1H, d, *J* 8.5 Hz, NH) and 7.25 (10H, m, aromatic); δ_C (50.3 MHz, C^2HCl_3) 20.67 ($\underline{C}H_3$'s), 31.0 ($\underline{C}$H (iBu)), 36.5 (Ar$\underline{C}H_2$C), 49.41 ($\underline{C}H_2$ (iBu)), 49.36 & 51.27 (d, $\underline{C}$H* $J_{P,C}$ 143.5 Hz), 51.79 &

51.70 (d, PO(O<u>C</u>H$_3$), $J_{P,C}$ 7.2 Hz), 67.67 (Ar<u>C</u>H$_2$O), 127.5-130.12 (aromatic <u>C</u>H), 137.32 & 137.8 (quat. aromatic) and 155.8 (<u>C</u>O urethane); *m/z* (CI) 405 ([*M* + H]$^+$, 49%), 349 ([*M* + H - $(CH_3)_2CCH_2$]$^+$, 10), 332 ([*M* + H - $(CH_3)_2CHCH_2NH_2$]$^+$, 68) and 297 ([*M* + H - $ArCH_2OH$]$^+$, 20).

<u>Cbz-Phe-ψ[P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (42).</u>



The chlorophosphonate derivative was prepared in an identical manner to (**41**) using methyl hydrogen [1-(N-benzyloxycarbonyl)-amino]-benzylphosphonate (**35**) (1.22 g, 3.5 mmol). (2S)-Pro-(2S)-Ile-NH-iBu hydrochloride (**40**) (1.12 g, 3.5 mmol) in CH_2Cl_2 (20 ml) with triethylamine (1.05 ml, 7.5 mmol) was added dropwise at 0 °C. The reaction was left at room temperature for 24 h. under an atmosphere of nitrogen. The organic solution was washed with 1 M HCl (2 x 30 ml), saturated sodium bicarbonate (2 x 30 ml) and water (30 ml) and then dried over anhydrous sodium sulphate. After filtration, the solvent was removed under reduced pressure to yield an off-white solid (1.40 g, 65%). The diastereomers were resolved by silica chromatography (5% EtOH/ CH_2Cl_2);

Cbz-(R)-Phe-ψ[(S)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (a);
M.p. 82 °C (dec.); *m/z* (Found: [*M* + H]$^+$ 615.3311. $C_{32}H_{48}N_4O_6P$ requires 615.3311); $[\alpha]_D$ -129.8° (c 0.5 in MeOH); ν_{max} (nujol)/ cm^{-1} 3401 (NH str.), 1716 (CO str. (urethane)), 1669 & 1652 (CO str. (amide)), 1260 (P=O str.) and 1050 (P-OCH_3); δ_H (600 MHz, C^2HCl_3) 0.85 (12H, m, CH_3's (a,b,c)), 1.03 (1H, m, 1 x CH_2 (Ile)), 1.46 (1H, m, 1 x CH_2 (Ile)) 1.66-1.82 (4H, m, CH_2 (2) (Pro), CH (iBu) &

1 x CH_2 (3) (Pro)), 2.00 (1H, m, 1 x CH_2 (3) (Pro)), 2.15 (1H, m, CH (Ile)), 2.88 (1H, m, 1 x $ArC\underline{H}_2$), 3.00-3.20 (4H, m, CH_2 (iBu), 1 x CH_2 (1) (Pro) & 1 x $ArC\underline{H}_2$), 3.30 (1H, m, 1 x CH_2 (1) (Pro)), 3.74 (3H, d, $J_{P,H}$ 10.71 Hz, $PO(OCH_3)$), 4.11 (1H, m, α CH (Pro)), 4.33 (1H, d. of d., J 5.0, 9.3 Hz, α CH (Ile)), 4.46 (1H, m, CH*), 4.94 (2H, s, $ArC\underline{H}_2O$), 5.09 (1H, d, J 10.44 Hz, NH (urethane)), 6.93 (1H, t, J 5.50 Hz, NH (iBu)) and 7.18-7.32 (11H, m, aromatic & NH (Ile)); δ_C (74.76 MHz, C^2HCl_3) 11.54 ($\underline{C}H_3$ (a)), 15.71 ($\underline{C}H_3$ (b)), 20.03 ($\underline{C}H_3$ (c)), 24.52 ($\underline{C}H_2$ (Ile)), 25.24 ($\underline{C}H_2$ (2) (Pro)), 28.13 ($\underline{C}H$ (iBu)), 30.54 ($\underline{C}H_2$ (3) (Pro)), 35.69 ($Ar\underline{C}H_2C$), 36.32 ($\underline{C}H$ (Ile)), 46.69 & 48.60 (d, $\underline{C}H^*$, $J_{P,C}$ 144.2 Hz), 46.81 ($\underline{C}H_2$ (iBu)), 46.91 ($\underline{C}H_2$ (1) (Pro)), 50.95 & 51.05 (d, $PO(O\underline{C}H_3)$, $J_{P,C}$ 7.6Hz), 57.90 (α $\underline{C}H$ (Ile)), 62.59 (α $\underline{C}H$ (Pro)), 66.84 ($Ar\underline{C}H_2O$), 126.61-128.96 (aromatic $\underline{C}H$'s), 135.99 & 136.25 (quat. aromatic), 155.77 ($\underline{C}O$ urethane), 170.80 ($\underline{C}O$ (Ile)) and 172.84 ($\underline{C}O$ (Pro)); δ_P (121.5 MHz, C^2HCl_3) 31.14; *m/z* (FAB) 637 ([*M* + Na]$^+$, 32%), 615 ([*M* + H]$^+$, 22) and 332 ([*M* + H - Pro-Ile-iBu]$^+$, 37).

Cbz-(S)-Phe-ψ[(S)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (b);
M.p. 90 °C (dec.); *m/z* (Found: [*M* + H]$^+$ 615.3311. $C_{32}H_{48}N_4O_6P$ requires 615.3311); $[\alpha]_D$ -74.1° (c 0.5 in MeOH); ν_{max} (nujol)/ cm^{-1} 3247 (NH str.), 1721 (CO str. (urethane)), 1677 & 1645 (CO str. (amide)), 1224 (P=O str.) and 1039 (P-OCH_3); δ_H (600 MHz, C^2HCl_3) 0.85 (12H, m, CH_3's (a,b,c)), 1.01 (1H, m, 1 x CH_2 (Ile)),1.44 (1H, m, 1 x CH_2 (Ile)) 1.66 (2H, m, CH_2 (2) (Pro)), 1.80 (1H, m, CH (iBu)), 1.98 (1 x CH_2 (3) (Pro)), 2.13 (2H, m, 1 x CH_2 (3) (Pro) & CH (Ile)), 2.84 (1H, m, 1 x $ArC\underline{H}_2$)), 3.02 (1H, m, 1 x CH_2 (iBu)), 3.10 (2H, m, 1 x CH_2 (1) (Pro) & 1 x CH_2 (iBu)), 3.21 (2H, m, 1 x CH_2 (1) (Pro) & 1 x $ArC\underline{H}_2$)), 3.69 (3H, d, $J_{P,H}$ 10.62 Hz, $PO(OCH_3)$), 4.10 (1H, m, α CH (Pro)), 4.33 (1H, d. of d., J 4.95, 9.52 Hz, α CH (Ile)), 4.49 (1H, m, CH*), 4.98 (2H, AB_q, J 12.36 Hz, $ArC\underline{H}_2O$), 5.18 (1H, d, J 10.1 Hz, NH (urethane)) and 7.10-7.32 (12H, m, NH (iBu), NH (Ile) & aromatic); δ_C (74.76 MHz, C^2HCl_3) 11.48 ($\underline{C}H_3$ (a)), 15.87 ($\underline{C}H_3$ (b)), 20.07 ($\underline{C}H_3$ (c)), 24.36 ($\underline{C}H_2$ (Ile)), 25.08 ($\underline{C}H_2$ (2) (Pro)), 28.08 ($\underline{C}H$ (iBu)), 30.83 ($\underline{C}H_2$ (3) (Pro)), 35.93 ($Ar\underline{C}H_2C$), 36.12 ($\underline{C}H$ (Ile)), 46.86 ($\underline{C}H_2$ (iBu)), 47.88 & 49.74 (d, $\underline{C}H^*$, $J_{P,C}$ 139.9 Hz), 48.41 ($\underline{C}H_2$ (1) (Pro)), 51.56 & 51.66 (d, $PO(O\underline{C}H_3)$, $J_{P,C}$ 7.2 Hz), 57.96 (α $\underline{C}H$ (Ile)), 61.08 (α $\underline{C}H$ (Pro)), 66.88 ($Ar\underline{C}H_2O$), 126.76-129.03

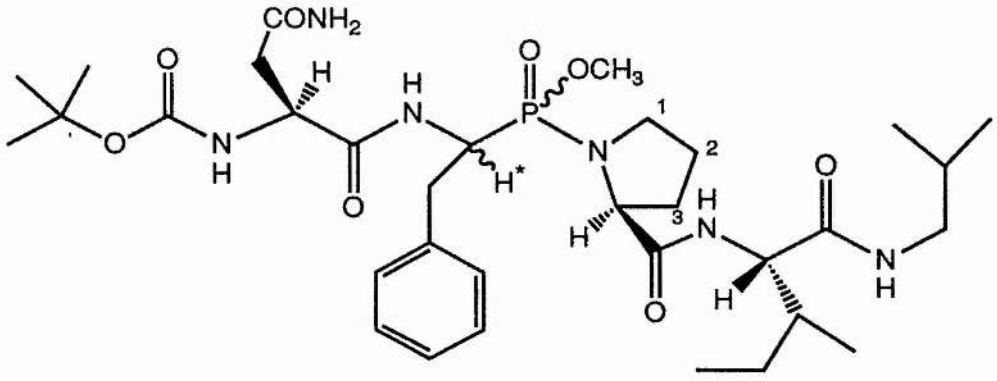
(aromatic CH's), 136.09 & 136.25 (quat. aromatic), 155.49 (CO urethane), 170.89 (CO (Ile)) and 172.62 (CO (Pro)); δ_P (121.5 MHz, C^2HCl_3) 30.19; *m/z* (FAB) 637 ([*M* + Na]$^+$, 10%), 615 ([*M* + H]$^+$, 12) and 332 ([*M* + H- Pro-Ile-iBu]$^+$, 21).

Cbz-(S)-Phe-ψ[(R)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (c);
M.p. 186-187 °C; (Found: C, 62.82; H, 7.88; N, 9.10. $C_{32}H_{47}N_4O_6P$ requires C, 62.53; H, 7.71; N, 9.11%); *m/z* (Found: [*M* + H]$^+$ 615.3311. $C_{32}H_{48}N_4O_6P$ requires 615.3311); $[\alpha]_D$ -30.1° (c 0.5 in MeOH); ν_{max} (nujol)/ cm^{-1} 3252 (NH str.), 1721 (CO str. (urethane)), 1677 & 1645 (CO str. (amide)), 1225 (P=O str.) and 1028 (P-OCH$_3$); δ_H (600 MHz, C^2HCl_3) 0.85 (12H, m, CH$_3$'s (a,b,c)), 1.06 (1H, m, 1 x CH$_2$ (Ile)), 1.42 (1H, m, 1 x CH$_2$ (Ile)), 1.74 (2H, m, CH$_2$ (2) (Pro)), 1.80 (1H, m, CH (iBu)), 1.92 (1H, m, CH (Ile)), 1.98 & 2.12 (2H, m, CH$_2$ (3) (Pro)), 2.89 (1H, m, 1 x ArCH$_2$)), 3.01 (1H, m, 1 x CH$_2$ (iBu)), 3.18 (3H, m, 1 x CH$_2$ (1) (Pro), 1 x CH$_2$ (iBu) & 1 x ArCH$_2$), 3.29 (1H, m, 1 x CH$_2$ (1) (Pro)), 3.66 (3H, d, $J_{P,H}$ 10.68 Hz, PO(OCH$_3$)), 4.22 (1H, d. of d., *J* 8.10, 9.30 Hz, α CH (Ile)), 4.27 (1H, m, α CH (Pro)), 4.46 (1H, m, CH*), 4.92 (2H, AB$_q$, *J* 12.36 Hz, ArCH$_2$O), 6.47 (2H, m, NH's (urethane & iBu)), 7.03 (1H, d, *J* 9.30 Hz, NH (Ile)) and 7.25 (10H, m, aromatic); δ_C (74.76 MHz, C^2HCl_3) 10.71 (CH$_3$ (a)), 15.40 (CH$_3$ (b)), 20.03 (CH$_3$ (c)), 24.62 (CH$_2$ (Ile)), 25.31 (CH$_2$ (2) (Pro)), 28.21 (CH (iBu)), 31.48 (CH$_2$ (3) (Pro)), 35.71 (ArCH$_2$C), 36.25 (CH (Ile)), 46.84 (CH$_2$ (iBu)), 47.52 (CH$_2$ (1) (Pro)), 48.16 & 50.12 (d, CH*, $J_{P,C}$ 148.3 Hz), 51.31 & 51.41 (d, PO(OCH$_3$), $J_{P,C}$ 7.3 Hz), 57.73 (α CH (Ile)), 61.37 (α CH (Pro)), 66.65 (ArCH$_2$O), 126.49-129.13 (aromatic CH's), 136.28 & 136.70 (quat. aromatic), 156.35 (CO urethane); 171.17 (CO (Ile)) and 173.34 (CO (Pro)); δ_P (121.5 MHz, C^2HCl_3) 31.64; *m/z* (FAB) 637 ([*M* + Na]$^+$, 23%), 615 ([*M* + H]$^+$, 27) and 332 ([*M* + H - Pro-Ile-iBu]$^+$, 28).

Cbz-(R)-Phe-ψ[(R)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (d);
M.p. 163-164 °C; (Found: C, 62.70; H, 7.92; N, 9.21. $C_{32}H_{47}N_4O_6P$ requires C, 62.53; H, 7.71; N, 9.11%); *m/z* (Found: [*M* + H]$^+$ 615.3311. $C_{32}H_{48}N_4O_6P$ requires 615.3311); $[\alpha]_D$ -47.0° (c 0.5 in MeOH); ν_{max} (nujol)/ cm^{-1} 3354 & 3209

(NH str.), 1695 (CO str. (urethane)), 1672 & 1652 (CO str. (amide)), 1252 (P=O str.) and 1036 (P-OCH_3); δ_H (600 MHz, C^2HCl_3) 0.85 (12H, m, CH_3's a,b,c), 1.14 (1H, m, 1 x CH_2 (Ile)), 1.50 (1H, m, 1 x CH_2 (Ile)), 1.76 (3H, m, CH_2 (2) (Pro) & CH (iBu)), 1.92 (1H, m, 1 x CH_2 (3) (Pro)), 2.05 (2H, m, 1 x CH_2 (3) (Pro)), CH (Ile)), 2.80 (1H, m, 1 x $ArC\underline{H}_2$)), 2.98 (1H, m, 1 x CH_2 (iBu)), 3.07 (2H, m, 1 x CH_2 (iBu) & 1 x CH_2 (1) (Pro)), 3.18 (1H, m, 1 x CH_2 (1) (Pro)), 3.28 (1H, m, 1 x $ArC\underline{H}_2$), 3.60 (3H, d, $J_{P,H}$ 10.53 Hz, $PO(OCH_3)$), 4.18 (1H, d. of d., *J* 7.95, 8.85 Hz, α CH (Ile)), 4.32 (1H, m, α CH (Pro)), 4.45 (1H, m, CH*), 4.97 (2H, AB_q, *J* 12.3 Hz, $ArC\underline{H}_2O$), 5.08 (1H, d, *J* 9.92 Hz, NH (urethane)), 6.61 (1H, t, *J* 5.85 Hz, NH (iBu)), 7.25 (10H, m, aromatic) and 7.45 (1H, d, *J* 8.85 Hz, NH (Ile)); δ_C (74.76 MHz, C^2HCl_3) 11.02 ($\underline{C}H_3$ (a)), 15.51 ($\underline{C}H_3$ (b)), 19.98 ($\underline{C}H_3$ (c)), 24.67 ($\underline{C}H_2$ (Ile)), 25.80 ($\underline{C}H_2$ (2) (Pro)), 28.18 ($\underline{C}H$ (iBu)), 31.12 ($\underline{C}H_2$ (3) (Pro)), 35.61 ($Ar\underline{C}H_2C$), 35.88 ($\underline{C}H$ (Ile)), 46.73 ($\underline{C}H_2$ (iBu)), 47.55 ($\underline{C}H_2$ (1) (Pro)), 49.46 & 50.40 (d, $\underline{C}H^*$, $J_{P,C}$ 146.5 Hz), 51.51 & 51.60 (d, $PO(O\underline{C}H_3)$, $J_{P,C}$ 7.4 Hz), 58.22 (α $\underline{C}H$ (Ile)), 62.23 (α $\underline{C}H$ (Pro)), 66.89 ($Ar\underline{C}H_2O$), 126.64-128.94 (aromatic $\underline{C}H$'s), 136.04 & 136.30 (quat. aromatic), 155.77 ($\underline{C}O$ urethane), 170.91 ($\underline{C}O$ (Ile)) and 173.31 ($\underline{C}O$ (Pro)); δ_P (121.5 MHz, C^2HCl_3) 28.90; *m/z* (FAB) 637 ([*M* + Na]$^+$, 27%), 615 ([*M* + H]$^+$, 18) and 332 ([*M* + H - Pro-Ile-iBu]$^+$, 13).

Boc-(2S)-Asn-Phe-ψ[P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (44).



Cbz-(R)-Phe-ψ[(S)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (**42a**) (70 mg, 110 μmol) was dissolved in methanol (5 ml) and palladium/ carbon catalyst (20 mg,

30% w/w) was added. The solution was stirred vigorously under a pressure of hydrogen gas (balloon) for 3 h. When the reaction had finished (by TLC), the catalyst was filtered off using celite. The colourless solution was concentrated under reduced pressure to yield H_2N-(R)-Phe-ψ[(S)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (**43**) as a colourless oil (47 mg, 86%).

Boc-(2S)-Asn (22 mg, 95 μmol) was dissolved in anhydrous THF (5 ml) with warming. N-methylmorpholine (10.5 μl, 95 μmol) was added and the solution cooled to -15 °C. Isobutylchloroformate (13 μl, 95 μmol) was added and the solution left for 5 min. H_2N-(R)-Phe-ψ[(S)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (47 mg, 98 μmol) (**43**) was dissolved in anhydrous THF (2 ml) and added dropwise to the solution at -15 °C. The reaction was left for 5 min. at this temperature and a further 25 min. at room temperature. The hydrochloride salts were filtered off and the THF removed under reduced pressure to yield an off-white solid which was purified by silica chromatography (10% EtOH/ CH_2Cl_2) to yield pure diastereomer (52 mg, 79%).

Boc-(2S)-Asn-(R)-Phe-ψ[(S)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (a); M.p. 120 °C (dec); *m/z* (Found: $[M + H]^+$ 695.387. $C_{33}H_{56}N_6O_8P$ requires 695.3897); $[\alpha]_D$ -127° (c 0.5 in MeOH); ν_{max} (nujol)/ cm^{-1} 3315 (NH str.), 1683 & 1652 (CO str., br), 1251 (P=O str.) and 1030 (P-OCH_3); δ_H (600 MHz, C^2HCl_3) 0.85 (12H, m, CH_3's (a,b,c)), 1.10 (1H, m, 1 x CH_2 (Ile)), 1.42 (9H, s, Boc CH_3's), 1.50 (1H, m, 1 x CH_2 (Ile)), 1.78 (1H, m, CH (iBu)), 1.88 (2H, m, CH_2 (2) (Pro)), 2.05 (3H, m, CH_2 (3) (Pro) & CH (Ile)), 2.46 (1H, ABX, *J* 5.77, 15.10 Hz, 1 x CH_2 (Asn)), 2.62 (1H, d, *J* 15.2 Hz, 1 x CH_2 (Asn)), 2.95 (2H, m, 1 x Ar$C\underline{H}_2$ & 1 x CH_2 (iBu)), 3.11 (1H, m, 1 x CH_2 (iBu)), 3.17 (2H, m, 1 x Ar$C\underline{H}_2$ & 1 x CH_2 (1) (Pro)), 3.31 (1H, m, 1 x CH_2 (1) (Pro)), 3.75 (3H, d, $J_{P,H}$ 10.80 Hz, $PO(OCH_3)$), 4.15 (1H, m, α CH (Pro)), 4.28 (1H, d. of d., *J* 6.23, 8.61 Hz, α CH (Ile)), 4.32 (1H, m, α CH (Asn)), 4.74 (1H, m, CH*), 5.68 (1H, s, 1 x $CONH_2$), 5.76 (1H, d, *J* 6.96 Hz, NH (urethane)), 5.99 (1H, s, 1 x $CONH_2$), 6.95 (1H, br, NH (iBu)), 7.18-7.28 (5H, m, aromatic), 7.32 (1H, d, *J* 8.79 Hz, NH (Ile)) and 7.44 (1H, d, *J* 8.79 Hz, NH (Asn)); δ_C (150 MHz, C^2HCl_3) 11.53 ($\underline{C}H_3$ (a)), 15.77 ($\underline{C}H_3$ (b)), 20.22 ($\underline{C}H_3$ (c)), 24.89 ($\underline{C}H_2$ (Ile)), 25.53 ($\underline{C}H_2$ (2) (Pro)), 28.35 (Boc $\underline{C}H_3$'s & $\underline{C}$H (iBu)), 31.13 ($\underline{C}H_2$ (3) (Pro)), 35.80 (Ar$\underline{C}H_2$C), 36.70 ($\underline{C}$H (Ile)), 37.22 ($\underline{C}H_2$ (Asn)), 45.55 &

46.5 (d, $\underline{C}H^*$, $J_{P,C}$ 135 Hz), 47.01 ($\underline{C}H_2$ (iBu)), 47.17 ($\underline{C}H_2$ (1) (Pro)), 51.45 ($PO(O\underline{C}H_3)$), 51.60 (α $\underline{C}H$ (Asn)), 58.30 (α $\underline{C}H$ (Ile)), 62.29 (α $\underline{C}H$ (Pro)), 80.48 ($\underline{C}(CH_3)_3$), 126.65-129.26 (aromatic $\underline{C}H$'s), 136.77 (quat. aromatic), 156.6 ($\underline{C}O$ urethane) and 171.13 & 173.56 ($\underline{C}O$ amides); δ_P (121.5 MHz, C^2HCl_3) 30.30; *m/z* (FAB) 717 ([*M* + Na]$^+$, 100%), 695 ([*M* + H]$^+$, 38) and 412 ([*M* + H - Pro-Ile-iBu]$^+$, 17).

Boc-(2S)-Asn-(S)-Phe-ψ[(S)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (b); (This was prepared in an identical manner to the (R, S) diastereomer); *m/z* (Found: [*M* + H]$^+$ 695.3897. $C_{33}H_{56}N_6O_8P$ requires 695.3897); $[\alpha]_D$ -54.5° (c 0.5 in MeOH); ν_{max} (nujol)/ cm^{-1} 3287 (NH str.), 1717 (CO str. (urethane), 1657 & 1652 (CO str., br), 1253 (P=O str.) and 1043 (P-OCH_3); δ_H (600 MHz, C^2HCl_3) 0.85 (12H, m, CH_3's (a,b,c)), 1.06 (1H, m, 1 x CH_2 (Ile)), 1.42 (9H, s, Boc CH_3's), 1.49 (1H, m, 1 x CH_2 (Ile)), 1.82 (3H, m, CH (iBu) & CH_2 (2) (Pro)), 2.06-2.20 (4H, m, CH_2 (3) (Pro), 1 x CH_2 (Asn) & CH (Ile)), 2.46 (1H, d, *J* 14.83 Hz, 1 x CH_2 (Asn)), 2.80 (1H, m, 1 x $ArC\underline{H}_2$), 3.02 (1H, m, 1 x CH_2 (iBu)), 3.12 (1H, m, 1 x CH_2 (iBu)), 3.27 (3H, m, 1 x $ArC\underline{H}_2$ & CH_2 (1) (Pro)), 3.76 (3H, d, $J_{P,H}$ 10.62 Hz, $PO(OCH_3)$), 4.16 (1H, m, α CH (Pro)), 4.35 (2H, m, α CH (Ile) & α CH (Asn)), 4.81 (1H, m, CH*), 5.50 (1H, s, 1 x $CONH_2$), 5.92 (1H, s, 1 x $CONH_2$), 5.97 (1H, d, *J* 7.87 Hz, NH (urethane)), 7.18-7.32 (7H, m, aromatic CH's & NH's (iBu & Ile)) and 7.47 (1H, d, *J* 8.06 Hz, NH (Asn)); δ_C (150 MHz, C^2HCl_3) 11.52 ($\underline{C}H_3$ (a)), 16.0 ($\underline{C}H_3$ (b)), 20.24 ($\underline{C}H_3$ (c)), 24.68 ($\underline{C}H_2$ (Ile)), 25.45 ($\underline{C}H_2$ (2) (Pro)), 28.32 (Boc CH_3's & $\underline{C}H$ (iBu)), 31.11 ($\underline{C}H_2$ (3) (Pro)), 36.02 ($Ar\underline{C}H_2C$), 36.26 ($\underline{C}H$ (Ile)), 36.99 ($\underline{C}H_2$ (Asn)), 46.75 & 47.55 (d, $\underline{C}H^*$, $J_{P,C}$ 120 Hz), 47.09 ($\underline{C}H_2$ (iBu)), 48.52 ($\underline{C}H_2$ (1) (Pro)), 51.15 (α $\underline{C}H$ (Asn)), 52.20 ($PO(O\underline{C}H_3)$), 58.32 (α $\underline{C}H$ (Ile)), 61.52 (α $\underline{C}H$ (Pro)), 80.3 ($\underline{C}(CH_3)_3$), 126.85-129.26 (aromatic $\underline{C}H$'s), 136.75 (quat. aromatic), 155.5 ($\underline{C}O$ (urethane)) and 170.93, 171.39 & 173.19 ($\underline{C}O$ amides); δ_P (121.5 MHz, C^2HCl_3) 30.43; *m/z* (FAB) 717 ([*M* + Na]$^+$, 31%), 695 ([*M* + H]$^+$, 22) and 412 ([*M* + H - Pro-Ile-iBu]$^+$, 13).

Boc-(2S)-Asn-(S)-Phe-ψ[(R)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (c); (This was prepared in an identical manner to the (R, S) diastereomer); *m/z* (Found: $[M + H]^+$ 695.3897. $C_{33}H_{56}N_6O_8P$ requires 695.3897); $[\alpha]_D$ -25.6° (c 0.5 in MeOH); ν_{max} (nujol)/ cm^{-1} 3233 (NH str.), 1700, 1684 & 1652 (CO str., br), 1224 (P=O str.) and 1045 (P-OCH_3); δ_H (600 MHz, C^2HCl_3) 0.85 (12H, m, CH_3's (a,b,c)), 1.13 (1H, m, 1 x CH_2 (Ile)), 1.39 (9H, s, Boc CH_3's), 1.51 (1H, m, 1 x CH_2 (Ile)), 1.82 (3H, m, CH (iBu), 1 x CH_2 (2) (Pro) & CH (Ile)), 1.91 (1H, m, 1 x CH_2 (2) (Pro)), 2.00 (1H, m, 1 x CH_2 (3) (Pro)), 2.07 (1H, br, 1 x CH_2 (Asn)), 2.21 (2H, m, 1 x CH_2 (3) (Pro) & 1 x CH_2 (Asn)), 2.86 (1H, m, 1 x Ar$C\underline{H}_2$), 3.02 (1H, m, 1 x CH_2 (iBu)), 3.16 (2H, m, 1 x Ar$C\underline{H}_2$ & 1 x CH_2 (1) (Pro)), 3.26 (1H, m, 1 x CH_2 (iBu)), 3.32 (1H, m, 1 x CH_2 (1) (Pro)), 3.66 (3H, d, $J_{P,H}$ 10.62 Hz, PO(OCH_3)), 4.25 (2H, m, α CH (Pro) & α CH (Ile)), 4.55 (1H, m, α CH (Asn)), 4.78 (1H, m, CH*), 5.57 (1H, s, 1 x $CONH_2$), 5.76 (1H, br., NH (urethane)), 6.01 (1H, s, 1 x $CONH_2$), 6.76 (1H, br., NH (iBu)), 7.18-7.28 (5H, m, aromatic), 7.34 (1H, br, NH (Ile)) and 7.83 (1H, d, *J* 8.06 Hz, NH (Asn)); δ_C (150 MHz, C^2HCl_3) 10.70 ($\underline{C}H_3$ (a)), 15.46 ($\underline{C}H_3$ (b)), 20.25 ($\underline{C}H_3$ (c)), 24.87 ($\underline{C}H_2$ (Ile)), 25.97 ($\underline{C}H_2$ (2) (Pro)), 28.28 (Boc CH_3's & $\underline{C}$H (iBu)), 31.69 ($\underline{C}H_2$ (3) (Pro)), 35.54 (Ar$\underline{C}H_2$C), 35.86 ($\underline{C}$H (Ile)), 37.94 ($\underline{C}H_2$ (Asn)), 45.85 & 46.80 (d, $\underline{C}$H*, $J_{P,C}$ 143 Hz), 47.09 ($\underline{C}H_2$ (1) (Pro)), 47.21 ($\underline{C}H_2$ (iBu)), 51.45 (α $\underline{C}$H (Asn)), 51.85 (PO($O\underline{C}H_3$)), 58.38 (α $\underline{C}$H (Ile)), 62.88 (α $\underline{C}$H (Pro)), 79.84 ($\underline{C}(CH_3)_3$), 126.73-129.44 (aromatic $\underline{C}$H's), 137.14 (quat. aromatic), 155.9 ($\underline{C}$O (urethane)) and 171.60, 172.51 & 174.02 ($\underline{C}$O's); δ_P (121.5 MHz, C^2HCl_3) 30.58; *m/z* (FAB) 717 ($[M + Na]^+$, 57%), 695 ($[M + H]^+$, 15) and 412 ($[M + H$ - Pro-Ile-iBu$]^+$, 10).

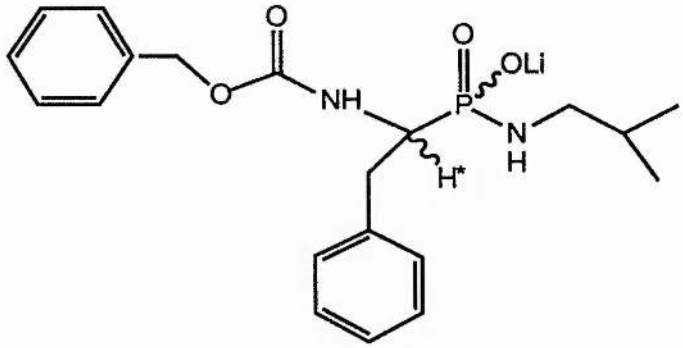
Boc-(2S)-Asn-(R)-Phe-ψ[(R)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (d); (This was prepared in an identical manner to the (R, S) diastereomer); *m/z* (Found: $[M + H]^+$ 695.390. $C_{33}H_{56}N_6O_8P$ requires 695.3897); $[\alpha]_D$ -61° (c 0.5 in MeOH); ν_{max} (nujol)/ cm^{-1} 3280 (NH str.), 1717 (CO str. (urethane)), 1683 & 1652 (CO str., br), 1251 (P=O str.) and 1024 (P-OCH_3); δ_H (600 MHz, C^2HCl_3) 0.85 (12H, m, CH_3's (a,b,c)), 1.14 (1H, m, 1 x CH_2 (Ile)), 1.43 (9H, s, Boc CH_3's), 1.53 (1H, m, 1 x CH_2 (Ile)), 1.77 (1H, m, CH (iBu)), 1.82-1.92 (2H, m, CH_2 (2) (Pro)), 1.95 (1H, m, 1 x CH_2 (3) (Pro)), 2.05 (1H, m, CH (Ile)), 2.12 (1H, m, 1 x

CH_2 (3) (Pro)), 2.45 (1H, ABX., J 6.64, 14.88 Hz, 1 x CH_2 (Asn)), 2.64 (1H, d, J 15.10 Hz, 1 x CH_2 (Asn)), 2.84 (1H, m, 1 x $ArC\underline{H}_2$), 2.99 & 3.11 (2H, m, CH_2 (iBu)), 3.23 (1H, m, 1 x CH_2 (1) (Pro)), 3.28 (2H, m, 1 x CH_2 (1) (Pro) & 1 x $ArC\underline{H}_2$)), 3.62 (3H, d, $J_{P,H}$ 10.62 Hz, $PO(OCH_3)$), 4.18 (1H, t, J 8.40 Hz, α CH (Ile)), 4.38 (2H, m, α CH (Pro) & α CH (Asn)), 4.66 (1H, m, CH*), 5.56 (1H, s, 1 x $CONH_2$), 5.76 (1H, d, J 7.78 Hz, NH (urethane)), 6.22 (1H, s, 1 x $CONH_2$), 6.81 (1H, br., NH (iBu)), 7.05 (1H, d, J 9.15 Hz, NH (Asn)), 7.18-7.28 (5H, m, aromatic) and 7.62 (1H, d, J 8.42 Hz, NH (Ile)); δ_C (74.76 MHz, C^2HCl_3) 10.95 ($\underline{C}H_3$ (a)), 15.58 ($\underline{C}H_3$ (b)), 20.04 ($\underline{C}H_3$ (c)), 24.84 ($\underline{C}H_2$ (Ile)), 25.75 ($\underline{C}H_2$ (2) (Pro)), 28.15 (Boc $\underline{C}H_3$'s), 28.27 ($\underline{C}H$ (iBu)), 31.35 ($\underline{C}H_2$ (3) (Pro)), 35.85 ($Ar\underline{C}H_2C$), 36.58 ($\underline{C}H$ (Ile)), 37.24 ($\underline{C}H_2$ (Asn)), 46.83 ($\underline{C}H_2$ (iBu)), 47.83 & 49.80 (d, $\underline{C}H^*$, $J_{P,C}$ 148 Hz), 47.83 ($\underline{C}H_2$ (1) (Pro)), 51.32 & 51.42 (d, $PO(O\underline{C}H_3)$, $J_{P,C}$ 7.3 Hz), 51.52 (α $\underline{C}H$ (Asn)), 58.58 (α $\underline{C}H$ (Ile)), 61.66 (α $\underline{C}H$ (Pro)), 80.26 ($\underline{C}(CH_3)_3$), 126.61-129.21 (aromatic $\underline{C}H$'s), 136.6 (quat. aromatic), 155.4 ($\underline{C}O$ (urethane)) and 171.09, 171.42, 172.64 & 173.70 ($\underline{C}O$ amides); δ_P (121.5 MHz, C^2HCl_3) 27.63; *m/z* (FAB) 717 ([*M* + Na]$^+$, 80%), 695 ([*M* + H]$^+$, 37) and 412 ([*M* + H - Pro-Ile-iBu]$^+$, 19).

Lithium propane-1-thiolate (45)

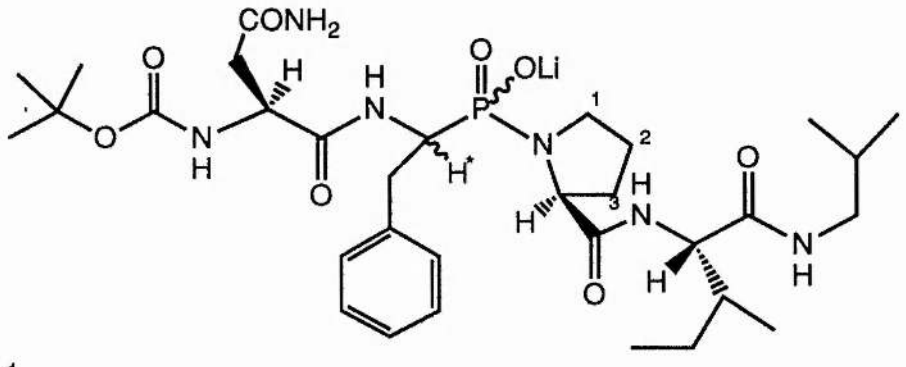
Using a modification of the method of Bartlett,[140] freshly distilled 1-propanethiol (5 ml, 55 mmol) was added to a suspension of finely ground lithium hydride (1.5 g, 189 mmol), in anhydrous oxygen-free DMF (30 ml) under an inert atmosphere. The reaction was left to stir at 25 °C for 1 h. followed by filtration without contact with air. The reagent was stored at 0 °C under argon pressure. The concentration of lithium propane-1-thiolate was typically 0.5-0.6 M.

Cbz-Phe-ψ[P(OLi)O-NH]-iBu (46).



A solution of Cbz-Phe-ψ[P(OMe)O-NH]-iBu (**41**) (50 mg, 124 μmol) in dioxan (1 ml) and 1 M lithium hydroxide (495 μl, 495 μmol) was refluxed for 24 h. The white paste was filtered off and washed with cold water (1 ml). After drying, the product was recrystallised from methanol/ ether to yield a white solid (38 mg, 77%), *m/z* (Found: [*M* + H]$^+$ 397.1869. $C_{20}H_{27}N_2O_4PLi$ requires 397.1869); δ_H (200 MHz, d_6-DMSO) 0.85 (6H, d, *J* 7.5 Hz, CH_3's), 1.60 (1H, m, CH (iBu)), 2.6-2.85 (3H, m, CH_2 (iBu) & 1 x ArC$\underline{H}_2$), 3.10 (1H, m, 1x ArC$\underline{H}_2$), 3.88 (1H, m, CH*), 4.95 (2H, s, ArC$\underline{H}_2$O), 5.40 (1H, d, *J* 8 Hz, NH) and 7.25 (10H, m, aromatic); *m/z* (FAB) 397 ([*M* + H]$^+$, 17%).

Boc-(2S)-Asn-Phe-ψ[P(OLi)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (48).



Method 1.

This was prepared in an identical manner to (**46**) using diastereomers (**44a**) and (**44b**) (50 mg, 72 μmol) and lithium hydroxide (1 M) (288 μl, 290 μmol) to

produce a white precipitate which was washed with acetone to remove any starting material (27 mg, 55%), *m/z* (Derivatised as methyl ester using CH_2N_2, Found: $[M + H]^+$ 695.3936. $C_{33}H_{56}N_6O_8P$ requires 695.3897) δ_H (200 MHz, d_4-MeOH) 0.90 (12H, m, CH_3's (a,b,c)), 1.22 (1H, m, 1 x CH_2 (Ile)), 1.45 (9H, s, Boc CH_3's), 1.50 (1H, m, 1 x CH_2 (Ile)), 1.7-2.4 (8H, m, CH (iBu), CH_2 (2) (Pro), CH (Ile), CH_2 (3) (Pro), CH_2 (Asn)), 2.75-3.5 (6H, m, Ar$\underline{C}H_2$, CH_2 (iBu), CH_2 (1) (Pro)), 3.85 (1H, m, α CH), 4.3 (2H, m, α CH's), 4.65 (1H, m, CH*) and 7.1-7.4 (5H, m, aromatic); δ_C (50.3 MHz, C^2HCl_3) 12.35 ($\underline{C}H_3$ (a)), 16.64 ($\underline{C}H_3$ (b)), 21.22 ($\underline{C}H_3$ (c)), 26.30 ($\underline{C}H_2$ (Ile)), 28.82 ($\underline{C}H_2$ (2) (Pro)), 29.33 (Boc $\underline{C}H_3$'s), 29.97 ($\underline{C}$H (iBu)), 32.84 ($\underline{C}H_2$ (3) (Pro)), 38.56-41.01 (Ar$\underline{C}H_2$C), ($\underline{C}$H (Ile)), ($\underline{C}H_2$ (Asn)), 48.22-50.16 ($\underline{C}H_2$ (iBu), ($\underline{C}$H*) & $\underline{C}H_2$ (1) (Pro)), 52.20 (α $\underline{C}$H (Asn)), 59.45 (α $\underline{C}$H (Ile)), 63.58 (α $\underline{C}$H (Pro)), 80.35 ($\underline{C}(CH_3)_3$), 128.17-130.89 (aromatic $\underline{C}$H's), 138.5 (quat. aromatic), 155.8 ($\underline{C}$O (urethane)) and 174.0-176.0 ($\underline{C}$O amides); δ_P (121.5 MHz, d_4-MeOH) 18.14, 16.33;

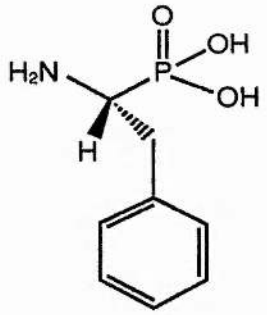
Method 2.

This was prepared using a modification of the method of Mckenna.[145] To a solution of isobutylene in anhydrous CH_2Cl_2 (5 ml) (prepared by sparging with isobutylene gas for 10 min.) was added freshly distilled trimethylsilylbromide (55 µl, 440 µmol) under an argon atmosphere.

H_2N-Phe-ψ[P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (**43**) ((**a**) and (**b**); 100 mg, 210 µmol) were dissolved in anhydrous CH_2Cl_2 (2 ml) and added to the TMSBr solution and the reaction left to stir at room temperature. The reaction was followed to completion by TLC (12 h.) and the CH_2Cl_2 was concentrated *in vacuo* to give a yellow oil that was rediluted and evaporated twice more with CH_2Cl_2 (2 x 2 ml). The resultant oil was taken up in acetone (2 ml) and water added (0.5 ml). Removal of solvent under reduced pressure yielded a white solid (86 mg, 88%) that was used immediately in the preparation of (**48**). This was prepared in an identical manner to (**44**) using H_2N-Phe-ψ[P(OH)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (**47**) (**a**) and (**b**) diastereomers; 62 mg, 130 µmol) and Boc-(2S)-asparagine (30 mg, 130 µmol) to produce an off-white solid (45 mg, 50%) that was dissolved in acetone and converted to (**48**) by the addition of 1 M

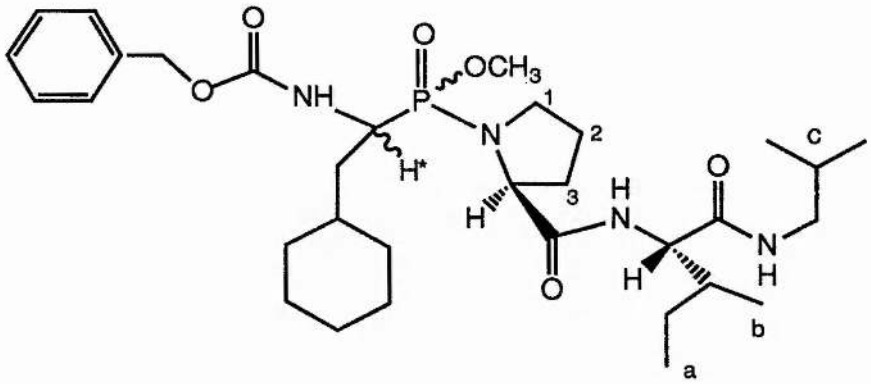
lithium hydroxide (70 µl, 70 µmol).

(1-Amino-(R)-2-phenyl)ethyl phosphonic acid (49).



Cbz-(R)-Phe-ψ[(S)-P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (**42a**) (50 mg, 80 µmol) was refluxed in a mixture of 6 M HCl and glacial acetic acid (2 ml, 1:1) for 2 h. The solvent was removed under reduced pressure and the solid redissolved in a minimum of water (0.5 ml). The hydrolysis mixture was purified by preparative TLC using a mixture of *iso*propanol, conc. ammonia and water (26:12:6). The band with an R_f value of 0.2 was removed and the cellulose stirred in a methanol/ water mixture (10 ml, 1:1) for 2 h. The cellulose was filtered off and washed with water. The water was removed under reduced pressure to yield a white solid which was dissolved in 1 M NaOH (2 ml), filtered through celite and the direction of optical rotation measured; $[\alpha]_D$ -0.046°; δ_H (200 MHz, ^2H_2O) (sodium salt of (**49**)) 2.30 & 2.68 (2H, m, ArC$\underline{H}_2$), 3.02 (1H, m, CH*) and 7.20 (5H, m, aromatic). (^{1}H NMR data was found to be identical to that of previously prepared racemic (1-amino-2-phenyl)ethyl phosphonic acid).

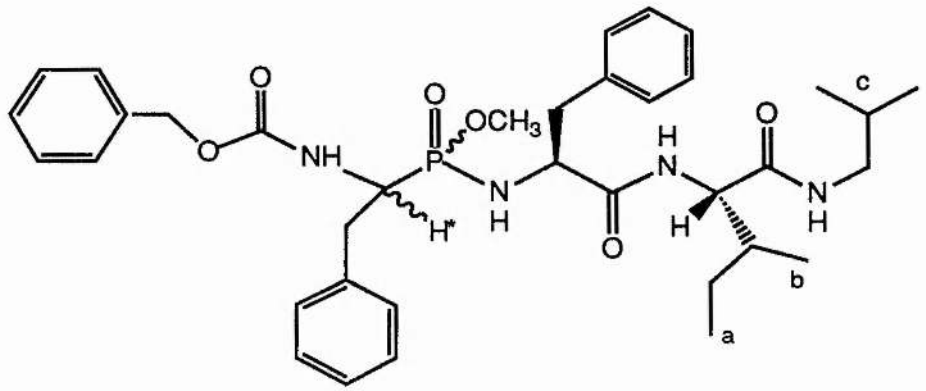
Cbz-Cha-ψ[P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (50).



This was prepared in an identical manner to (**42**), using methyl hydrogen [1-(N-benzyloxycarbonyl)-amino]-cyclohexylmethylphosphonate (1.07 g, 3 mmol) to give an off-white solid (1.08 g, 58%) which was purified by silica chromatography (6% EtOH/ CH_2Cl_2), to yield two pairs of two diastereomers; Diastereomers 3 & 4; *m/z* (Found: $[M + H]^+$ 621.3781. $C_{32}H_{53}N_4O_6P$ requires 621.3781); ν_{max} (nujol)/ cm^{-1}) 3259 (NH str. br.), 1724 & 1717 (CO str. (urethane)), 1660 & 1652 (CO str. (amide)), 1237 (P=O str.) and 1027 (P-OCH_3); δ_H (600 MHz, C^2HCl_3) 0.90 (12H, m, CH_3's (a,b,c)), 1.0-1.70 (17H, m, CH_2 (Ile), 5 x CH_2 (Cha), CH (Cha), CH_2 (2) (Pro), ChaCHC$\underline{H}_2$), 1.70-2.20 (4H, m, CH_2 (3) (Pro), CH (iBu), & CH (Ile)), 2.90-3.30 (4H, m, CH_2 (iBu), CH_2 (1) (Pro)), 3.61 & 3.64 (3H, d, $J_{P,H}$ 10.60 Hz, $PO(OCH_3)$), 4.12-4.34 (3H, m, α CH's & CH*), 5.10 (2H, m, Ar$\underline{C}H_2O$) and 7.25-7.40 (5H, m, aromatic); δ_C (150 MHz, C^2HCl_3) 10.89 & 11.40 ($\underline{C}H_3$ (a)), 15.56 & 15.86 ($\underline{C}H_3$ (b)), 20.17 ($\underline{C}H_3$ (c)), 24.83 ($\underline{C}H_2$ (Ile)), 25.56 & 26.0 ($\underline{C}H_2$ (2) (Pro)), 25.94-26.46 (5 x $\underline{C}H_2$ (Cha)), 28.33 ($\underline{C}H$ (iBu)), 31.63 & 31.83 ($\underline{C}H_2$ (3) (Pro)), 34.08 & 34.18 ($\underline{C}H$ (Cha)), 35.82 & 36.29 ($\underline{C}H$ (Ile)), 37.09 & 37.41 (Cha$\underline{C}H_2$), 45.00 & 45.99 (d, $\underline{C}H^*$, $J_{P,C}$ 148.4 Hz), 46.54 & 47.51 (d, $\underline{C}H^*$, $J_{P,C}$ 146.5 Hz), 46.99 ($\underline{C}H_2$ (iBu)), 47.72 ($\underline{C}H_2$ (1) (Pro)), 51.8 & 52.4 ($PO(O\underline{C}H_3)$), 57.96 & 58.48 (α $\underline{C}H$ (Ile)), 61.54 & 62.14 (α $\underline{C}H$ (Pro)), 67.06 & 67.32 (Ar$\underline{C}H_2O$), 127.5-128.5 (aromatic $\underline{C}H$'s), 136.21 & 136.48 (quat. aromatic), 155.98 & 156.54 ($\underline{C}O$ urethane), 170.90 & 171.22 ($\underline{C}O$ (Ile)) and 173.16 & 173.45 ($\underline{C}O$ (Pro)); δ_P (121.5 MHz, C^2HCl_3) 31.08 & 33.04; *m/z* (FAB) 643 ($[M + Na]^+$, 40%), 621 ($[M + H]^+$, 34) and 338 ([*M* + H- Pro-Ile-

iBu]$^+$, 61).

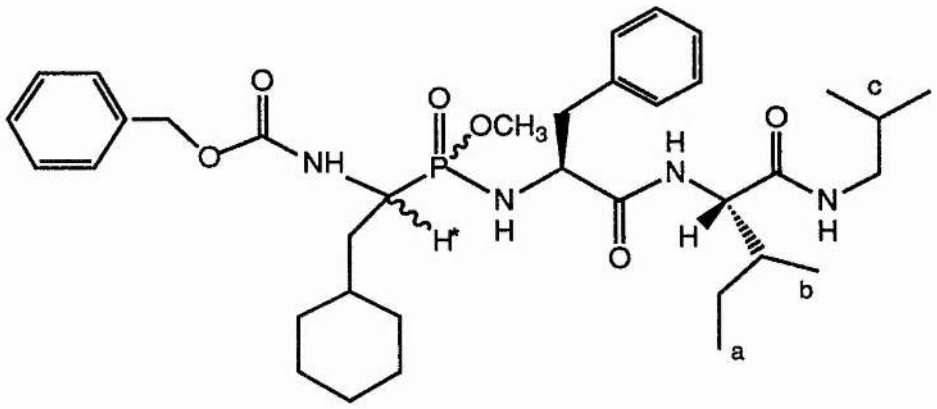
Cbz-Phe-ψ[P(OMe)O-N]-(2S)-Phe-(2S)-Ile-NH-iBu (51).



This was prepared in an identical manner to (**42**), using methyl hydrogen [1-(N-benzyloxycarbonyl)-amino]-benzylphosphonate (**35**) (1.05 g, 3 mmol) and (2S)-Phe-(2S)-Ile-iBu hydrochloride (1.11 g, 3 mmol) to yield an off-white solid (837 mg, 42%) which was purified by silica chromatography (10% EtOH/ CH_2Cl_2) prior to recrystallisation from methanol/ water to yield a mixture of four diastereomers, (Found: C, 64.96; H, 7.67; N, 8.27. $C_{36}H_{49}N_4O_6P$ requires C, 65.03; H, 7.43; N, 8.43%); *m/z* (Found: [*M* + H]$^+$ 665.3468. $C_{36}H_{50}N_4O_6P$ requires 665.3468); ν_{max} (nujol)/ cm^{-1} 3296 (NH str.), 1726 (CO str. (urethane)), 1694 & 1642 (CO str. (amide)), 1263 (P=O str.) and 1052 (P-OCH_3); δ_H (500 MHz, d_6-DMSO) 0.85 (12H, m, CH_3's (a,b,c)), 1.18 (1H, m, 1 x CH_2 (Ile)), 1.55 (1H, m, 1 x CH_2 (Ile)), 1.75 (1H, m, CH (iBu)), 1.82 (1H, m, CH (Ile)), 2.28 (1H, m, 1 x Ar$C\underline{H}_2$CH*)), 2.45 (1H, m, 1 x Ar$C\underline{H}_2$CH*)), 2.80 (1H, m, 1 x CH_2 (Phe)), 2.90 (1H, m, 1 x CH_2 (Phe)), 3.02 (2H, m, CH_2 (iBu)), 3.07, 3.24, 3.28 & 3.52 (3H, d, $J_{P,H}$ 10.60 Hz, PO(OCH_3)), 3.90-4.30 (3H, m, α CH's & CH*), 4.85-5.20 (2H, m, Ar$C\underline{H}_2$O), 7.10-7.50 (15H, m, aromatic) and 7.95-8.25 (4H, m, NH's); δ_C (74.76 MHz, d_6-DMSO) 10.92 ($\underline{C}H_3$ (a)), 15.27 ($\underline{C}H_3$ (b)), 19.98 ($\underline{C}H_3$ (c)), 24.30 ($\underline{C}H_2$ (Ile)), 27.80 ($\underline{C}$H (iBu)), 36.70 & 36.80 (Ar$\underline{C}H_2$C & $\underline{C}$H (Ile)), 40.0 ($\underline{C}H_2$ (Phe)), 46.0 ($\underline{C}H_2$ (iBu)), 49.0-52.0 ($\underline{C}$H* & PO($O\underline{C}H_3$)), 55.51 & 57.04 (α $\underline{C}$H's), 65.04 (Ar$\underline{C}H_2$O), 125.98-129.48 (aromatic $\underline{C}$H's), 137.98-138.13 (quat. aromatic), 155.60 ($\underline{C}$O urethane) and 170.70 & 172.50 ($\underline{C}$O (amides)); δ_P (121.5

MHz, d_6-DMSO) 29.62, 29.78, 29.94 and 30.08; *m/z* (FAB) 687 ([*M* + Na]$^+$, 20%) and 665 ([*M* + H]$^+$, 79).

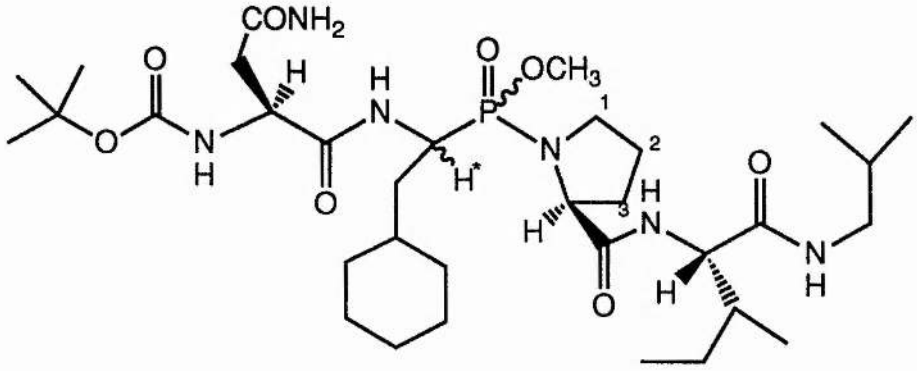
Cbz-Cha-ψ[P(OMe)O-N]-(2S)-Phe-(2S)-Ile-NH-iBu (52).



This was prepared in an identical manner to (**42**), using methyl hydrogen [1-(N-benzyloxycarbonyl)-amino]-cyclohexylmethylphosphonate (1.07 g, 3 mmol) and (2S)-Phe-(2S)-Ile-iBu hydrochloride (1.11 g, 3 mmol) to yield an off-white solid (1.46 g, 64%) which was purified by silica chromatography (5% EtOH/ CH_2Cl_2) prior to recrystallisation from methanol/ water to yield a mixture of four diastereomers, (Found: C, 64.41; H, 8.35; N, 8.32. $C_{36}H_{55}N_4O_6P$ requires C, 64.44; H, 8.27; N, 8.36%); *m/z* (Found: [*M* + H]$^+$ 671.3937. $C_{36}H_{56}N_4O_6P$ requires 671.3937); ν_{max} (nujol)/ cm^{-1} 3281 (NH str.), 1721 (CO str. (urethane)), 1665 & 1643 (CO str. (amide)) and 1263 (P=O str.); δ_H (600 MHz, C^2HCl_3) 0.85 (12H, m, CH_3's (a,b,c)), 1.0-2.20 (17H, m, CH_2 (Ile), 5 x CH_2 (Cha), CH (Cha), CH (Ile), CH (iBu) & ChaCHC$\underline{H}_2$)), 2.88-3.12 (4H, m, CH_2 (Phe) & CH_2 (iBu)), 3.35 (3H, d, $J_{P,H}$ 10.90 Hz, PO(OCH_3)), 3.43 (3H, d, $J_{P,H}$ 10.60 Hz, PO(OCH_3)), 3.47 (3H, d, $J_{P,H}$ 10.90 Hz, PO(OCH_3)), 3.59 (3H, d, $J_{P,H}$ 10.90 Hz, PO(OCH_3)), 4.02-4.28 (3H, m, α CH's & CH*), 5.00-5.16 (2H, m, ArC$\underline{H}_2$O) and 7.14-7.36 (10H, m, aromatic); δ_C (74.76 MHz, C^2HCl_3) 11.2 (C$\underline{H}_3$ (a)), 15.5 (C$\underline{H}_3$ (b)), 20.1 (C$\underline{H}_3$ (c)), 24.8 (C$\underline{H}_2$ (Ile)), 25.9, 26.3, 26.5 (5 x C$\underline{H}_2$ (Cha)), 28.4 (C$\underline{H}$ (iBu)), 31.7 (C$\underline{H}$ (Cha)), 34.0 (C$\underline{H}$ (Ile)), 36.5 (ChaCH C$\underline{H}_2$), 40.0 (C$\underline{H}_2$ (Phe)), 46.0-48.0 (C$\underline{H}$*), 47.0 (C$\underline{H}_2$ (iBu)), 51.6 (PO(OCH_3)), 56.5 & 56.8 (α C$\underline{H}$), 58.2 & 58.4 (α C$\underline{H}$), 67.2 (ArC$\underline{H}_2$O), 126.5-130 (aromatic C$\underline{H}$'s), 136.5-137.7 (quat. aromatic),

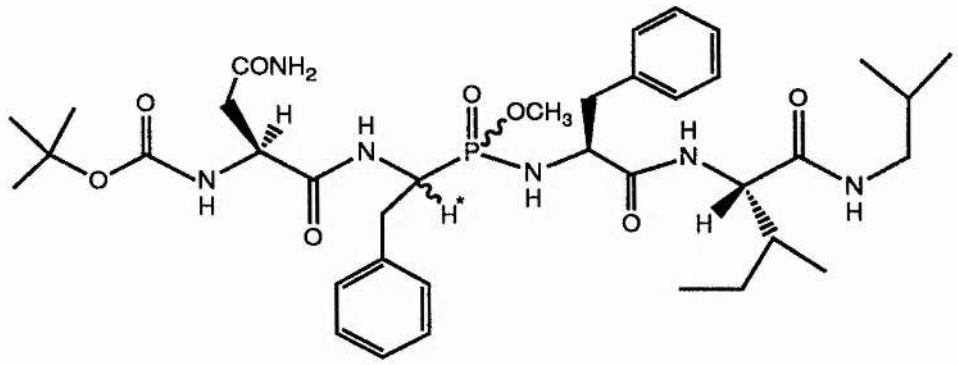
156.4 (CO urethane) and 171.0 & 173.0 (CO (amides)); δ_P (121.5 MHz, C^2HCl_3) 30.54, 31.21, 32.05 and 32.35; *m/z* (FAB) 693 ([*M* + Na]$^+$, 15%) and 671 ([*M* + H]$^+$, 22).

Boc-(2S)-Asn-Cha-ψ[P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (54).



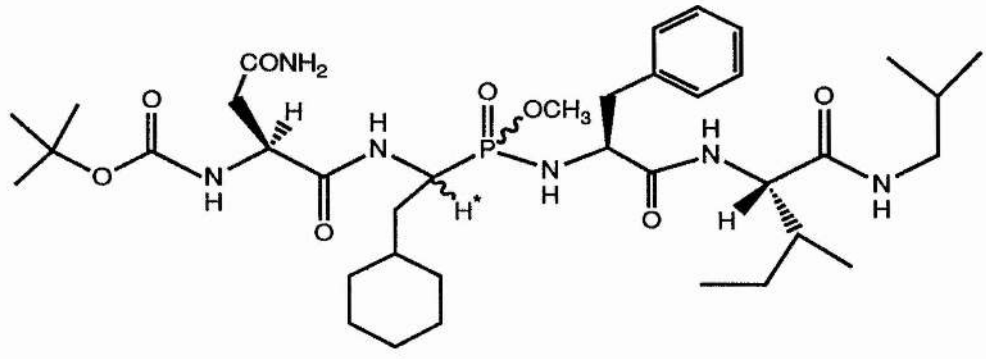
This was prepared in an identical manner to (**44**), using H_2N-Cha-ψ[(P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (**53**) (diastereomers 3 & 4, 75 mg, 154 μmol) and (2S)-Boc-Asn (36 mg, 155 μmol) to yield an off-white solid (80 mg, 74%) which was purified by silica chromatography (10% EtOH/ CH_2Cl_2), to yield two diastereomers; Diastereomer 4; *m/z* (Found: [*M* + Na]$^+$ 723.419. $C_{33}H_{61}N_4O_6PNa$ requires 723.4186); ν_{max} (nujol)/ cm^{-1} 3294 (NH str. br.), 1716 (CO str. (urethane)), 1652 (CO str., br. amide's), 1261 (P=O str.) and 1026 (P-OCH_3); δ_H (200 MHz, C^2HCl_3) 0.90 (12H, m, CH_3's a,b,c), 1.0-1.70 (17H, m, CH_2 (Ile), 5 x CH_2 (Cha), CH (Cha), CH_2 (2) (Pro) & ChaCHC$\underline{H}_2$), 1.50 (9H, s, Boc CH_3's), 1.70-2.20 (4H, m, CH_2 (3) (Pro), CH (iBu), & CH (Ile)), 2.50-3.30 (6H, m, CH_2 (iBu), CH_2 (1) (Pro), CH_2 Asn), 3.65 (3H, d, $J_{P,H}$ 10.60 Hz, $PO(OCH_3)$), 4.10-4.60 (4H, m, α CH's & CH*), 5.9 (1H, s, 1 x $CONH_2$), 6.7 (1H, s, 1 x $CONH_2$), 6.2-7.45 (NH's); δ_P (121.5 MHz, C^2HCl_3) 30.14; *m/z* (FAB) 723 ([*M* + Na]$^+$, 100%) and 701 ([*M* + H]$^+$, 25).

Boc-(2S)-Asn-Phe-ψ[P(OMe)O-N]-(2S)-Phe-(2S)-Ile-NH-iBu (56).



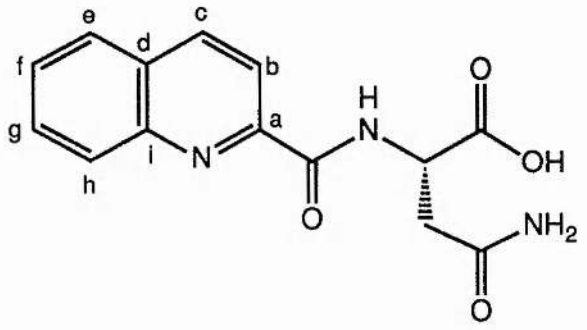
This was prepared in an identical manner to (**44**), using H_2N-Phe-ψ[(P(OMe)O-N]-(2S)-Phe-(2S)-Ile-NH-iBu (**55**) (180 mg, 340 μmol) and (2S)-Boc-Asn (79 mg, 340 μmol) to yield an off-white solid (180 mg, 71%) which was purified by silica chromatography (10% EtOH/ CH_2Cl_2), to yield four diastereomers; *m/z* (Found: [*M* + H]$^+$ 745.4054. $C_{37}H_{57}N_6O_8P$ requires 745.4054); δ_H (200 MHz, C^2HCl_3) 0.90 (12H, m, CH_3's (a,b,c)), 1.15 (1H, m, 1 x CH_2 (Ile)), 1.50 (10H, s, Boc CH_3's & 1 x CH_2 (Ile)), 1.80 (2H, m, CH (iBu) & CH (Ile)), 2.20-3.30 (8H, m, Ar$C\underline{H}_2$CH*, CH_2 (Phe)), CH_2 (iBu) & CH_2 (Asn)), 3.40 (3H, m, $PO(OCH_3)$), 3.90-4.60 (4H, m, α CH's & CH*), 5.95 (1H, s, 1 x $CONH_2$), 6.65 (1H, s, 1 x $CONH_2$) and 7.10-7.50 (10H, m, aromatic); *m/z* (FAB) 767 ([*M* + Na]$^+$, 5%) and 745 ([*M* + H]$^+$, 25).

Boc-(2S)-Asn-Cha-ψ[P(OMe)O-N]-(2S)-Phe-(2S)-Ile-NH-iBu (58).



This was prepared in an identical manner to (**44**), using H_2N-Cha-ψ[(P(OMe)O-N]-(2S)-Phe-(2S)-Ile-NH-iBu (**57**) in DMF (0.5 ml) (120 mg, 220 μmol) and (2S)-Boc-Asn (52 mg, 220 μmol) to yield an off-white solid (129 mg, 77%) which was purified by silica chromatography (10% EtOH/ CH_2Cl_2), to yield four diastereomers; *m/z* (Found: [*M* + H]$^+$ 751.4521. $C_{37}H_{64}N_6O_8P$ requires 751.4523); δ_H (200 MHz, C^2HCl_3) 0.90 (12H, m, CH_3's a,b,c), 1.0-2.0 (17H, m, CH_2 (Ile), 5 x CH_2 (Cha), CH (Cha), CH (Ile), CH (iBu) & ChaCHC$\underline{H_2}$)), 1.50 (9H, s, Boc CH_3's), 2.5-3.2 (6H, m, CH_2 (Phe), CH_2 (Asn) & CH_2 (iBu)), 3.32 & 3.48 (3H, m, $PO(OCH_3)$), 4.0-4.5 (4H, m, α CH's & CH*) and 7.14-7.36 (5H, m, aromatic); *m/z* (FAB) 773 ([*M* + Na]$^+$, 10%) and 751 ([*M* + H]$^+$, 100).

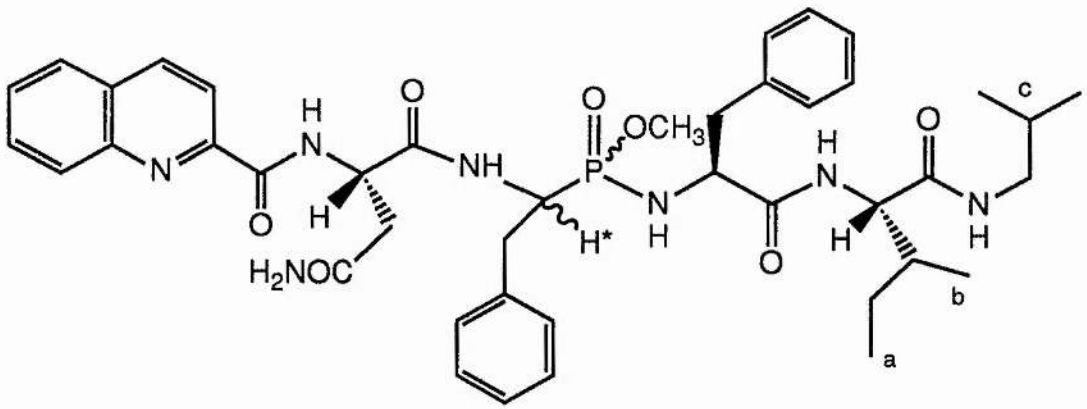
Qua-(2S)-asparagine (59).



Quinaldic acid (2 g, 11.5 mmol) was dissolved in anhydrous CH_2Cl_2 (50 ml). Thionyl chloride (1.75 ml, 23 mmol) was added and the white precipitate redissolved with triethylamine (1.60 ml, 11.5 mmol). After 3 h. the solvent was removed under reduced pressure and the acid chloride was redissolved in toluene (40 ml). To the reaction mixture was added a solution of asparagine (1.53 g, 11.5 mmol) and potassium carbonate (3.18 g, 23 mmol) in water (20 ml) and the solution stirred vigorously for 4 h. The aqueous layer was adjusted to pH 9 and separated. The aqueous layer was adjusted to pH 2 with 6 M HCl and the brown solid was filtered off, redissolved in ethanol and the solution decolourised with charcoal. Removal of solvent under reduced pressure yielded an off-white solid which was recrystallised from ethanol (2.05 g, 62%), m.p. 209 °C (dec.); (Found: C, 58.56; H, 4.32; N, 14.38. $C_{14}H_{13}N_3O_4$ requires C, 58.52; H, 4.56; N, 14.63%); (Found: [*M* + H]$^+$ 288.0984. $C_{14}H_{14}N_3O_4$ requires 288.0984);

$[\alpha]_D$ +23.3° (c 0.5 in MeOH); ν_{max} (nujol)/ cm^{-1} 3384 & 3365 (NH str. (amides)), 1710 (CO str. (acid)) and 1662 (CO str. (amide)); δ_H (300 MHz, d_6-DMSO) 2.80 (2H, ABX, *J* 15.97, 4.94 & 5.67 Hz, CH_2 (Asn)), 4.82 (1H, m, α CH (Asn)), 7.04 (1H, s, 1 x $CONH_2$), 7.54 (1H, s, 1 x $CONH_2$), 7.76 (1H, t, *J* 7 Hz, CH (f) (Qua)), 7.88 (1H, t, *J* 6.9 Hz, CH (g) (Qua)), 8.11 (2H, 2d, *J* 8.23 & 7.15 Hz, CH (b) & CH (e) (Qua)), 8.18 (1H, d, *J* 8.48 Hz, CH (h) (Qua)), 8.60 (1H, d, *J* 8.51 Hz, CH (c) (Qua)) and 9.18 (1H, d, *J* 8.56 Hz, NH†); δ_C (74.76 MHz, d_6-DMSO) 36.46 ($\underline{C}H_2$ (Asn)), 48.95 (α $\underline{C}$H (Asn)), 118.47 ($\underline{C}$H (b) (Qua)), 128.03 & 128.11 ($\underline{C}$H (e) & $\underline{C}$H (f) (Qua)), 128.88 ($\underline{C}$ (d) (Qua)), 129.14 ($\underline{C}$H (g) (Qua)), 130.53 ($\underline{C}$H (h) (Qua)), 137.94 ($\underline{C}$H (c) (Qua)), 145.90 ($\underline{C}$ (i) (Qua)), 149.47 ($\underline{C}$ (a) (Qua)), 163.37 ($\underline{C}$O amide), 171.69 ($\underline{C}ONH_2$) and 172.31 ($\underline{C}OOH$); *m/z* (CI) 288 ([*M* + H]$^+$, 12%). † identified by D_2O shake.

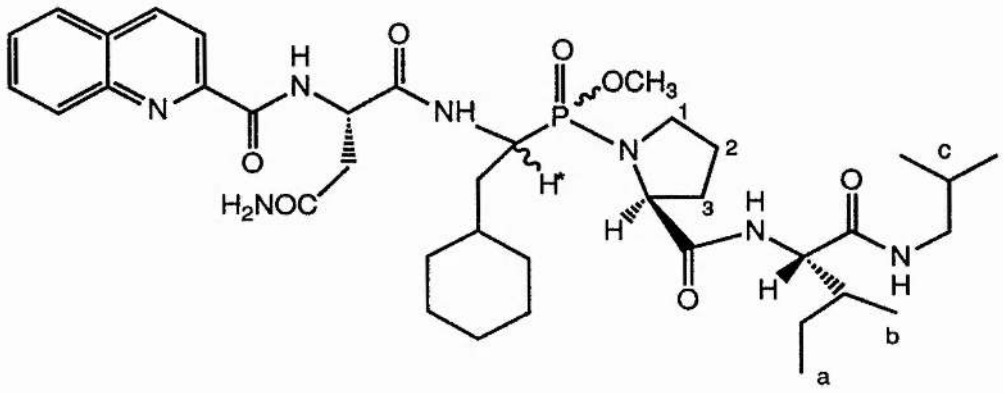
Qua-(2S)-Asn-Phe-ψ[P(OMe)O-N]-(2S)-Phe-(2S)-Ile-NH-iBu (60).



This was prepared in an identical manner to (**44**), using H_2N-Phe-ψ[(P(OMe)O-N]-(2S)-Phe-(2S)-Ile-NH-iBu (**55**) (40 mg, 76 μmol) and (**59**) (22 mg, 77 μmol) to yield an off-white solid (37 mg, 61%) which was purified by silica chromatography (10% EtOH/ CH_2Cl_2), to yield four diastereomers; *m/z* (Found: [*M* + H]$^+$ 800.3974. $C_{42}H_{55}N_7O_6P$ requires 800.3900); δ_H (200 MHz, d_4-MeOH) 0.80 (12H, m, CH_3's (a,b,c)), 1.0-2.0 (4H, m, CH_2 (Ile), CH (iBu) & CH (Ile)), 2.20-3.30 (8H, m, Ar$\underline{C}H_2$CH*, CH_2 (Phe), CH_2 (iBu), CH_2 (Asn)), 3.30 (3H, m, PO(O$\underline{C}H_3$)), 3.50-4.60 (4H, m, α CH's & CH*), 7.25 (10H, m, aromatic) and 7.6-

8.5 (6H, m, CH's (Qua)); *m/z* (FAB) 822 ([*M* + Na]$^+$, 100%) and 800 ([*M* + H]$^+$, 49).

Qua-(2S)-Asn-Cha-ψ[P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (61).



This was prepared in an identical manner to (**60**), using H_2N-Cha-ψ[(P(OMe)O-N]-(2S)-Pro-(2S)-Ile-NH-iBu (**53**) (two diastereomers; 140 mg, 290 μmol) and (**59**) (83 mg, 290 μmol) to yield an off-white solid (158 mg, 73%) which was purified by silica chromatography (8% MeOH/ CH_2Cl_2), to yield two diastereomers;

Diastereomer 3; ν_{max} (nujol)/ cm^{-1} 3282 (NH str. br.), 1700, 1696 & 1670 (CO str., br. amide's), 1223 (P=O str.) and 1039 (P-OCH_3); δ_H (200 MHz, C^2HCl_3) 0.80 (12H, m, CH_3's (a,b,c)), 1.0-2.20 (21H, m, CH_2 (Ile), 5 x CH_2 (Cha), CH (Cha), CH_2 (2) (Pro), ChaCHC$\underline{H}_2$, CH_2 (3) (Pro), CH (iBu), & CH (Ile)), 2.80-3.50 (6H, m, CH_2 (iBu), CH_2 (1) (Pro), CH_2 (Asn), 3.65 (3H, d, $J_{P,H}$ 10.60 Hz, $PO(OCH_3)$), 4.21 (2H, m, α CH's (Pro) & (Ile)), 4.60 (1H, m, CH*), 5.30 (1H, m, α CH (Asn)), 6.02 (1H, s, 1 x $CONH_2$), 6.85 (1H, s, 1 x $CONH_2$), 7.45 (1H, br., NH), 7.62 (1H, t, *J* 7 Hz, CH (f) (Qua)), 7.77 (1H, t, *J* 6.9 Hz, CH (g) (Qua)), 7.9 (2H, m, CH (b) & CH (e) (Qua)), 8.15 (1H, d, *J* 8.48 Hz, CH (h) (Qua)), 8.30 (1H, d, *J* 8.51 Hz, CH (c) (Qua)) and 9.15 (1H, d, *J* 8.6 Hz, NH).

Diastereomer 4; *m/z* (Found: [*M* + H]$^+$ 756.427. $C_{38}H_{59}N_7O_7P$ requires 756.4214); δ_H (200 MHz, C^2HCl_3) 0.80 (12H, m, CH_3's (a,b,c)), 1.0-2.20 (21H, m, CH_2 (Ile), 5 x CH_2 (Cha), CH (Cha), CH_2 (2) (Pro), ChaCHC$\underline{H}_2$, CH_2 (3) (Pro), CH (iBu), & CH (Ile)), 2.80-3.50 (6H, m, CH_2 (iBu), CH_2 (1) (Pro), CH_2 (Asn)),

3.64 (3H, d, $J_{P,H}$ 10.60 Hz, $PO(OCH_3)$), 4.21 (2H, m, α CH's), 4.60 (1H, m, CH*), 5.30 (1H, m, α CH (Asn)), 6.2 (1H, s, 1 x $CONH_2$), 6.7 (1H, s, 1 x $CONH_2$), 7.1 (1H, br, NH (iBu)), 7.32 (1H, br, NH), 7.6-7.9 (4H, m, CH (f) (Qua), CH (g) (Qua), CH (b) & CH (e) (Qua)), 8.15 (1H, d, *J* 8.48 Hz, CH (h) (Qua)), 8.20 (1H, d, *J* 8.51 Hz, CH (c) (Qua)) and 9.28 (1H, d, *J* 8.6 Hz, NH); *m/z* (FAB) 778 ([*M* + Na]$^+$, 94%) and 756 ([*M* + H]$^+$, 66).

(2R)-Phe-(2S)-Ile-NH-iBu hydrochloride (62).

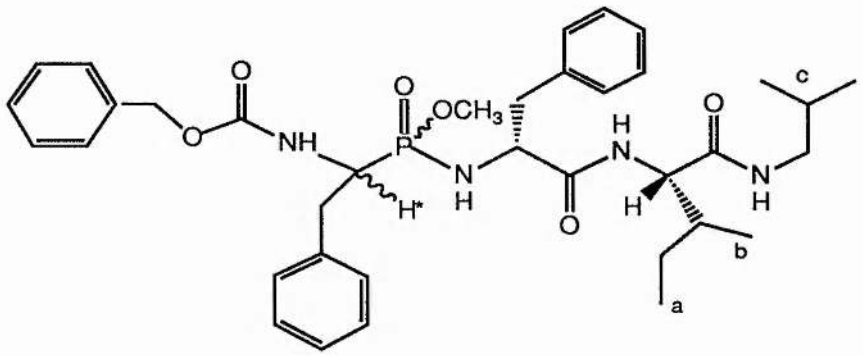
This was prepared in an identical manner to (**40**), using Boc-(2R)-Phe (1.0 g, 3.77 mmol) and (**38**) (840 mg, 3.77 mmol), followed by Boc deprotection, using HCl gas to yield a white solid, which was recrystallised from methanol/ ether (1.23 g, 88%), m.p. 252 °C (dec.); *m/z* (Found: [*M* + H - HCl]$^+$ 334.2495. $C_{19}H_{32}N_3O_2$ requires 334.2494); $[\alpha]_D$ -79.6° (c 0.5 in MeOH); ν_{max} (nujol)/ cm^{-1} 3350 & 3251 (NH str.) and 1679 & 1652 (CO str. (amides)); δ_H (200 MHz, d_6-DMSO) 0.63 (3H, d, *J* 7.6 Hz, CH_3 (b)), 0.75 (3H, t, *J* 7.5 Hz, CH_3 (a)), 0.87 (6H, d, *J* 8 Hz, CH_3 (c)), 1.02 (1H, m, 1 x CH_2 (Ile)), 1.18 (1H, m, 1 x CH_2 (Ile)), 1.70 (2H, m, CH (Ile) & CH (iBu)), 2.75-3.15 (4H, m, CH_2 (iBu) & CH_2 (Phe)), 4.13 (1H, t, *J* 7.0 Hz, α CH), 4.27 (1H, t, *J* 8.0 Hz, α CH), 7.30 (5H, s, Aromatic), 8.17 (1H, t, *J* 5.85 Hz, NH (iBu), 8.45 (3H, br, NH_3^+) and 8.6 (1H, d, *J* 8.76 Hz, NH (Ile)); δ_C (50.3 MHz, d_6-DMSO) 11.44 ($\underline{C}H_3$ (a)), 15.42 ($\underline{C}H_3$ (b)), 20.48 ($\underline{C}H_3$ (c)'s), 24.42 ($\underline{C}H_2$ (Ile)), 28.12 ($\underline{C}H$ (iBu)), 36.89 ($\underline{C}H_2$ (Phe)), 37.35 ($\underline{C}H$ (Ile)), 46.34 ($\underline{C}H_2$ (iBu)), 53.34 (α $\underline{C}H$ (Phe)), 57.54 (α $\underline{C}H$ (Ile)), 127.25, 128.66 & 129.75 (aromatic $\underline{C}H$), 135.22 (quat. aromatic) and 168.11 & 170.32 ($\underline{C}O$ amides); *m/z* (CI) 334 ([*M* + H - HCl]$^+$, 100%).

Gly-(2S)-Ile-NH-iBu hydrochloride (63).

This was prepared in an identical manner to (**40**), using Boc-Gly (875 mg, 5 mmol) and (**38**) (1.11 g, 5 mmol), followed by Boc deprotection, using HCl gas to yield a white solid, which was recrystallised from methanol/ ether (1.22 g, 87%), m.p. 165 °C (dec.); *m/z* (Found: [*M* + H - HCl]$^+$ 244.2025. $C_{12}H_{26}N_3O_2$

requires 244.2025); $[\alpha]_D$ -25.2° (c 0.5 in MeOH); ν_{max} (nujol)/ cm^{-1} 3412 & 3280 (NH str.) and 1674 & 1654 (CO str. (amides)); δ_H (200 MHz, d_6-DMSO) 0.85 (12H, m, CH_3's (a,b,c)), 1.12 (1H, m, 1 x CH_2 (Ile)), 1.45 (1H, m, 1 x CH_2 (Ile)), 1.70 (2H, m, CH (Ile) & CH (iBu)), 2.72-3.08 (2H, m, CH_2 (iBu)), 3.63 (2H, s, CH_2 (Gly)), 4.25 (1H, t, *J* 8.0 Hz, α CH (Ile)), 8.25 (4H, br., NH_3^+ & NH (iBu)) and 8.58 (1H, d, *J* 8.76 Hz, NH (Ile)); δ_C (50.3 MHz, d_6-DMSO) 11.38 ($\underline{C}H_3$ (a)), 15.64 ($\underline{C}H_3$ (b)), 20.44 ($\underline{C}H_3$'s (c)), 24.53 ($\underline{C}H_2$ (Ile)), 28.13 ($\underline{C}H$ (iBu)), 37.11 ($\underline{C}H$ (Ile)), 46.33 ($\underline{C}H_2$ (iBu)), 51.0 ($\underline{C}H_2$ (Gly)), 57.50 (α $\underline{C}H$ (Ile)) and 165.88 & 170.64 ($\underline{C}O$ amides); *m/z* (CI) 244 ([*M* + H - HCl]$^+$, 100%).

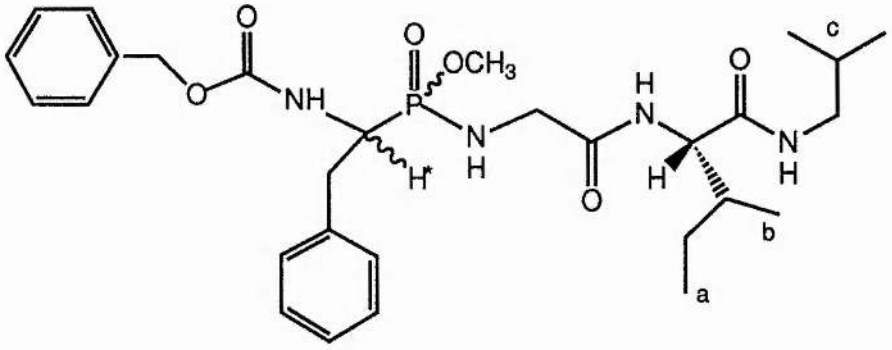
Cbz-Phe-ψ[P(OMe)O-N]-(2R)-Phe-(2S)-Ile-NH-iBu (64).



This was prepared in an identical manner to (**51**), using methyl hydrogen [1-(N-benzyloxycarbonyl)-amino]-benzylphosphonate (**35**) (349 mg, 1mmol) and (2R)-Phe-(2S)-Ile-iBu hydrochloride (**62**) (370 mg, 1 mmol) to yield an off-white solid (300 mg, 45%) which was purified by silica chromatography (10% EtOH/ CH_2Cl_2) prior to recrystallisation from methanol/ water to yield a mixture of four diastereomers, *m/z* (Found: [*M* + H]$^+$, 665.3456. $C_{36}H_{50}N_4O_6P$ requires 665.3468); ν_{max} (nujol)/ cm^{-1} 3269 (NH str.), 1703 (CO str. (urethane)), 1695 & 1654 (CO str. (amide)), 1261 (P=O str.) and 1049 (P-OCH_3); δ_H (200 MHz, d_6-DMSO) 0.80 (12H, m, CH_3's (a,b,c)), 1.05 (1H, m, 1 x CH_2 (Ile)), 1.32 (1H, m, 1 x CH_2 (Ile)), 1.7 (2H, m, CH (iBu) & CH (Ile)), 2.5-3.0 (6H, m, Ar$\underline{C}H_2$CH*, CH_2 (Phe) & CH_2 (iBu)), 3.07, 3.24, 3.28 & 3.52 (3H, d, $J_{P,H}$ 10.60 Hz, PO(OCH_3)),

3.90-4.30 (3H, m, α CH's & CH*), 4.85-5.20 (2H, m, ArC$\underline{H}_2$O), 7.10-7.50 (15H, m, aromatic) and 7.95-8.25 (NH's); δ_C (74.76 MHz, d_6-DMSO) 11.14 & 11.37 ($\underline{C}H_3$ (a)), 15.59 ($\underline{C}H_3$ (b)), 20.22 & 20.36 ($\underline{C}H_3$ (c)), 24.36 & 24.48 ($\underline{C}H_2$ (Ile)), 28.04 & 28.15 ($\underline{C}$H (iBu)), 35.19 (Ar$\underline{C}H_2$C), 38.60 ($\underline{C}$H (Ile)), 40.77 ($\underline{C}H_2$ (Phe)), 46.27 & 46.34 ($\underline{C}H_2$ (iBu)), 49.0-52.0 ($\underline{C}$H* & PO(O$\underline{C}H_3$)), 55.65, 55.81, & 57.11 (α $\underline{C}$H's), 65.01 & 66.33 (Ar$\underline{C}H_2$O), 126.19-129.88 (aromatic $\underline{C}$H's), 137.4-138.64 (quat. aromatic), 156.0 ($\underline{C}$O urethane) and 171.0 ,173.2 & 173.3 ($\underline{C}$O (amides)); δ_P (121.5 MHz, d_6-DMSO) 29.84, 30.10 and 30.24; *m/z* (FAB) 687 ([*M* + Na]$^+$, 36%) and 665 ([*M* + H]$^+$, 28).

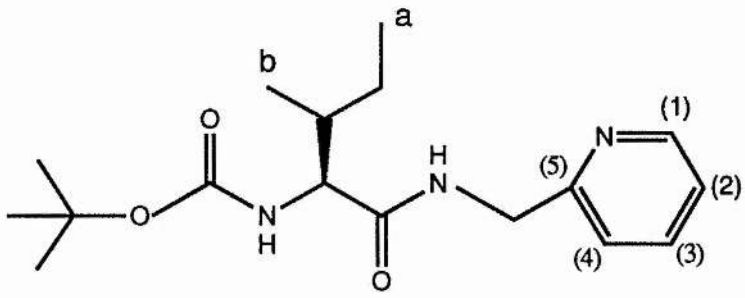
Cbz-Phe-ψ[P(OMe)O-N]-Gly-(2S)-Ile-NH-iBu (65).



This was prepared in an identical manner to (**42**), using methyl hydrogen [1-(N-benzyloxycarbonyl)-amino]-benzylphosphonate (**35**) (1.05 g, 3 mmol) and Gly-(2S)-Ile-iBu hydrochloride (**63**) (840 mg, 3 mmol) to yield an off-white solid (826 mg, 48%) which was purified by silica chromatography (10% EtOH/ CH_2Cl_2) prior to recrystallisation from methanol/ water to yield a mixture of four diastereomers, *m/z* (Found: [*M* + H]$^+$ 575.2979. $C_{29}H_{44}N_4O_6P$ requires 575.2998); ν_{max} (nujol)/ cm^{-1} 3297 (NH str.), 1704 & 1696 (CO str. (urethane)), 1684 & 1657 (CO str. (amide)), 1261 (P=O str.) and 1049 (P-OCH$_3$); δ_H (200 MHz, d_6-DMSO) 0.85 (12H, m, CH$_3$'s (a,b,c)), 1.1 (1H, m, 1 x CH$_2$ (Ile)), 1.45 (1H, m, 1 x CH$_2$ (Ile)), 1.7 (2H, m, CH (iBu) & CH (Ile)), 2.6-3.2 (4H, m, Ar$\underline{H}_2$CH* & CH$_2$ (iBu)), 3.62 (5H, m, CH$_2$ (Gly) & PO(OCH$_3$)), 3.90-4.30 (2H, m, α CH (Ile) & CH*), 4.85-5.05 (2H, m, Ar$\underline{H}_2$O), 7.10-7.4 (10H, m, aromatic) and 7.6-8.15 (NH's); δ_C (74.76 MHz, d_6-DMSO) 11.34 ($\underline{C}H_3$ (a)), 15.61 ($\underline{C}H_3$ (b)),

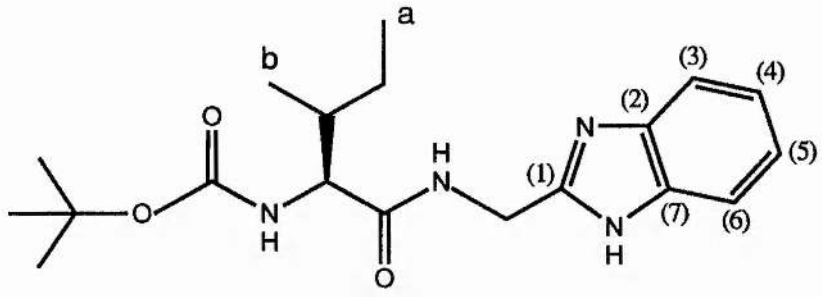
20.38 (<u>C</u>H_3 (c)), 24.56 (<u>C</u>H_2 (Ile)), 28.17 (<u>C</u>H (iBu)), 34.81 (Ar<u>C</u>H_2C), 37.23 (<u>C</u>H (Ile)), 46.31 (<u>C</u>H_2 (iBu)), 49.0-52.0 (<u>C</u>H* & PO(O<u>C</u>H_3)), 50.80 (<u>C</u>H_2 (Gly)), 57.05 (α <u>C</u>H (Ile)), 65.22 & 65.33 (Ar<u>C</u>H_2O), 126.4-129.38 (aromatic <u>C</u>H's), 137.4-138.64 (quat. aromatic), 156.04 (<u>C</u>O urethane) and 170.64 & 170.83 (<u>C</u>O (amides)); δ_P (121.5 MHz, d_6-DMSO) 31.01 and 31.44; *m/z* (FAB) 597 ([*M* + Na]$^+$, 100%) and 575 ([*M* + H]$^+$, 80).

<u>Boc-(2S)-isoleucyl-2-(aminomethyl)-pyridine (66)</u>



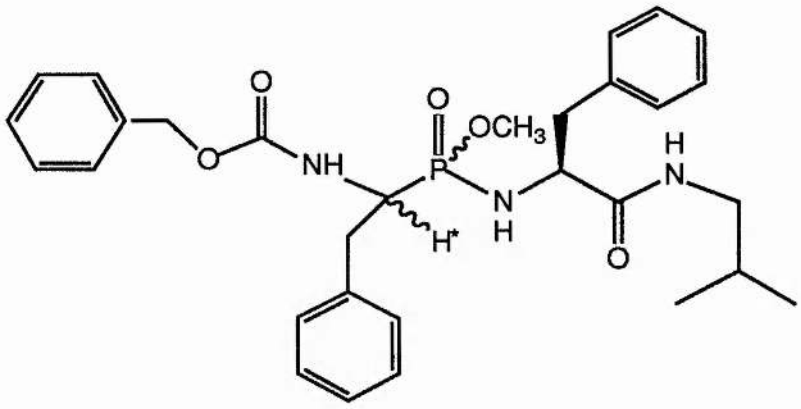
This was prepared in an identical manner to (**37**), using Boc-(2S)-Ile (463 mg, 2 mmol) and 2-(aminomethyl)-pyridine (210 µl, 2 mmol) to yield a yellow oil that was dissolved in an acetone/ water mixture (10 ml, 10:1) and left to crystallise overnight at room temperature, affording colourless needles (440 mg, 68%), *m/z* (Found: [*M* + H]$^+$ 322.2131. $C_{17}H_{28}N_3O_3$ requires 322.2131); ν_{max} (nujol)/ cm^{-1} 3335 (NH str.), 1682 (CO str. (urethane)), 1655 (CO str. (amide) and 1526 (NH bend); δ_H (200 MHz, C^2HCl_3) 0.9 (6H, m, CH_3's (a & b)), 1.10 (1H, m, 1 x CH_2 (Ile)), 1.42 (10H, m, Boc CH_3's & 1 x CH_2 (Ile)), 1.9 (1H, m, CH (Ile)), 3.84 (1H, t, *J* 7.0 Hz, α CH (Ile)), 4.55 (2H, d, *J* 5.0 Hz, CH_2 (Amp)), 5.25 (1H, d, *J* 8.6 Hz, NH (urethane)), 7.10-7.40 (3H, m, NH (amide) & CH (3) & (4)), 7.62 (1H, m, CH (2)) and 8.52 (1H, d, *J* 5.0 Hz, CH (1)); δ_C (50.3 MHz, C^2HCl_3) 11.97 (<u>C</u>H_3 (a)), 16.12 (<u>C</u>H_3 (b)), 25.18 (<u>C</u>H_2 (Ile)), 28.79 (Boc <u>C</u>H_3's), 37.91 (<u>C</u>H (Ile)), 44.85 (<u>C</u>H_2 (Amp)), 59.83 (α <u>C</u>H), 80.21 (<u>C</u>$(CH_3)_3$), 122.41 & 122.81 (<u>C</u>H (2) & (4)), 137.22 (<u>C</u>H (3)), 149.48 (<u>C</u>H (1)), 156.27 (<u>C</u>O urethane) and 156.81 (<u>C</u> (5)), 172.24 (<u>C</u>O amide); *m/z* (CI) 322 ([*M* + H]$^+$, 100%), 266 ([*M* + H - $(H_3C)_2CCH_2$]$^+$, 48), 213 ([*M* + H - $(H_3C)_2CCH_2$ - H_2O]$^+$, 40) and 187 ([*M* + H - Boc]$^+$, 92).

Boc-(2S)-isoleucyl-2-(aminomethyl)-benzimidazole (67)



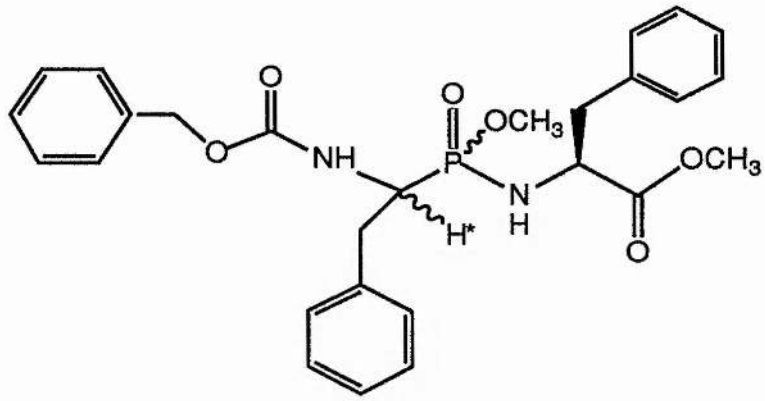
This was prepared in an identical manner to (**37**), using Boc-(2S)-Ile (463 mg, 2 mmol) and 2-(aminomethyl)-benzimidazole dihydrochloride (476 mg, 2 mmol) (preneutralised with triethylamine 560 ml, 4 mmol) to yield a yellow oil that was dissolved in an acetone/ water mixture (10 ml, 10:1) and left to crystallise overnight at room temperature, affording colourless needles (375 mg, 52%), *m/z* (Found: $[M + H]^+$ 361.2240. $C_{19}H_{29}N_4O_3$ requires 361.2239); ν_{max} (nujol)/ cm^{-1} 3344 & 3275 (NH str.), 1687 (CO str. (urethane)), 1656 (CO str. (amide)) and 1526 (NH bend); δ_H (200 MHz, C^2HCl_3) 0.9 (6H, m, CH_3's (a & b)), 1.15 (1H, m, 1 x CH_2 (Ile)), 1.50 (10H, m, Boc CH_3's & 1 x CH_2 (Ile)), 1.85 (1H, m, CH (Ile)), 3.96 (1H, t, *J* 6.0 Hz, α CH (Ile)), 4.5-5.0 (3H, m, CH_2 & NH), 5.28 (1H, m, NH), 7.20 (2H, m, Aromatic), 7.52 (2H, m, Aromatic) and 7.65 (1H, m, NH); δ_C (50.3 MHz, C^2HCl_3) 11.63 ($\underline{C}H_3$ (a)), 16.03 ($\underline{C}H_3$ (b)), 25.42 ($\underline{C}H_2$ (Ile)), 28.83 (Boc $\underline{C}H_3$'s), 37.51 ($\underline{C}H$ (Ile)), 38.14 ($\underline{C}H_2$ (Benz.)), 60.35 (α $\underline{C}H$), 80.72 ($\underline{C}(CH_3)_3$), 115.34 ($\underline{C}H$ (3) & (6)), 123.06 ($\underline{C}H$ (4) & (5)) 138.2 ($\underline{C}H$ (2) & (7)), 152.20 ($\underline{C}H$ (1)), 156.75 ($\underline{C}O$ urethane) and 174.32 ($\underline{C}O$ amide); *m/z* (CI) 361 ($[M + H]^+$, 62%), 287 ($[M + H - (H_3C)_2CCH_2 - H_2O]^+$, 50) and 261 ($[M + H - Boc]^+$, 58).

Cbz-Phe-ψ[P(OMe)O-N]-(2S)-Phe-NH-iBu (68).



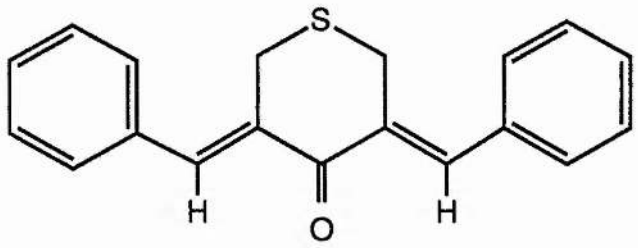
This was prepared in an identical manner to (**42**), using methyl hydrogen [1-(N-benzyloxycarbonyl)-amino]-benzylphosphonate (**35**) (698 mg, 2 mmol) and (2S)-Phe-NH-iBu (440 mg, 2 mmol) to yield, after work up, an off-white solid which was recrystallised from methanol/ water to yield a mixture of four diastereomers (580 mg, 53%), *m/z* (Found: $[M + H]^+$ 552.2627. $C_{30}H_{39}N_3O_5P$ requires 552.2627); ν_{max} (nujol)/ cm^{-1} 3296 (NH str.), 1695 (CO (urethane) 1653 (CO str. (amide)), 1215 (P=O str.) and 1056 (P-OCH_3); δ_H (200 MHz, d_6-DMSO) 0.85 (6H, m, CH_3's (iBu)), 1.65 (1H, m, CH (iBu)), 2.4-2.7 (6H, m, CH_2 (Phe), CH_2 (iBu) & Ar$\underline{C}H_2$CH*), 3.0 & 3.40 (3H, d, $J_{P,H}$ 10.60 Hz, PO(OCH_3)), 3.8-4.2 (2H, m, α CH (Phe) & CH*), 4.95 (2H, m, Ar$\underline{C}H_2$O), 5.0-5.3 (1H, m, NH's), 7.0-7.5 (15H, m, aromatic) and 8.15 (1H, br., NH); δ_C (74.76 MHz, d_6-DMSO) 20.33 ($\underline{C}H_3$ (iBu)), 28.17 & 28.26 ($\underline{C}H$ (iBu)), 33.86-34.52 (Ar $\underline{C}H_2$'s), 46.31 & 46.42 ($\underline{C}H_2$ (iBu)), 48.5-52.0 ($\underline{C}H$* & PO(O$\underline{C}H_3$)), 55.65, 55.97, 56.09 & 56.22 (α $\underline{C}H$ (Phe)), 65.01, 65.19, 65.31 & 65.40 (Ar$\underline{C}H_2$O), 126.23-128.88 (aromatic $\underline{C}H$'s), 137.33-138.63 (quat. aromatic), 155.95 & 156.05 ($\underline{C}O$ urethane) and 172.93 & 173.04 ($\underline{C}O$ (amide)); δ_P (121.5 MHz, d_6-DMSO) 29.44, 29.89, 30.04 and 30.14; *m/z* (FAB) 574 ($[M + Na]^+$, 10%) and 552 ($[M + H]^+$, 35).

Cbz-Phe-ψ[P(OMe)O-N]-(2S)-Phe-OMe (69).



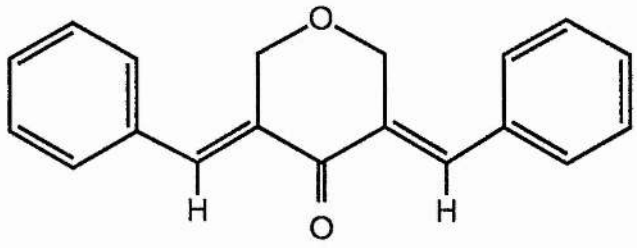
This was prepared in an identical manner to (**42**), using methyl hydrogen [1-(N-benzyloxycarbonyl)-amino]-benzylphosphonate (**35**) (349 mg, 1 mmol) and (2S)-Phe-OMe hydrochloride (323 mg, 1.5 mmol) to yield, after work up, an off-white solid which was recrystallised from acetone/ petrol to yield a mixture of four diastereomers (250 mg, 49%), *m/z* (Found: $[M + H]^+$, 511.1998. $C_{27}H_{32}N_2O_6P$ requires 511.1998); ν_{max} (nujol)/ cm^{-1} 3332 & 3185 (NH str.), 1750 (CO str. (ester)), 1695 (CO (urethane)), 1653 (CO str. (amide)), 1259 (P=O str.) and 1048 (P-OCH_3); δ_H (200 MHz, d_6-DMSO) 2.4-2.7 (2H, m, CH_2 (Phe)), 2.7-3.2 (2H, m, Ar$\underline{C}H_2$CH*), 3.10, 3.20, 3.35 & 3.50 (3H, d, $J_{P,H}$ 10.60 Hz, PO(OCH_3)), 3.8-4.2 (2H, m, α CH (Phe) & CH*), 4.90 (2H, m, Ar$\underline{C}H_2$O), 5.2-5.7 (1H, m, NH's) and 7.25 (15H, m, aromatic CH's); δ_C (74.76 MHz, d_6-DMSO) 34-34.25 (Ar $\underline{C}H_2$'s), 48.5-51.0 ($\underline{C}$H* & PO($O\underline{C}H_3$)), 51.97, 52.12 & 52.23 ($O\underline{C}H_3$ ester), 55.34, 55.58, & 55.68 (α $\underline{C}$H's), 65.11, 65.28 & 65.39 (Ar$\underline{C}H_2$O), 126.28-130.04 (aromatic $\underline{C}$H's), 137.56-138.83 (quat. aromatic), 155.86 ($\underline{C}$O urethane) and 173.86, 173.81 & 174.05 ($\underline{C}$O (ester)); δ_P (121.5 MHz, d_6-DMSO) 28.88 & 29.64; *m/z* (CI) 511 ($[M + H]^+$, 40%) and 403 ($[M + H - ArCH_2OH]^+$, 58).

Tetrahydro-3,5-bis(phenylmethylene)-4*H*-thiopyran-4-one (70).



A solution of sodium hydroxide (1 g, 25 mmol) in water (10 ml) and ethanol (10 ml) was prepared at 0 °C. Benzaldehyde (1.06 g, 10 mmol) was dissolved in a minimum volume of ethanol (5 ml) and added to a solution of tetrahydrothio-4*H*-pyran-4-one (580 mg, 5 mmol) in ethanol (5 ml). Half of this mixture was added dropwise to the alkali solution with vigorous stirring and the reaction mixture left for 15 mins at room temperature. The remainder of this mixture was then added and left for 1 h. The yellow precipitate was filtered and washed with water (100 ml) to remove excess alkali. Recrystallisation from acetone yielded yellow crystals (905 mg, 62%), m.p. 149-151 °C (lit. (i) 150-151 °C[156] (ii) 149-151 °C [166]); *m/z* (Found: M^+, 292.0922. $C_{19}H_{16}SO$ requires 292.0922); ν_{max} (nujol)/ cm^{-1} 1660 (CO str.) and 1600, 1577 & 1569 (C=C str.); δ_H (200 MHz, C^2HCl_3) 3.95 (4H, s, $C\underline{H}_2SC\underline{H}_2$), 7.20-7.55 (10H, m, Aromatic), 7.80 (2H, s, 2 x ArC<u>H</u>=C); δ_C (50.3 MHz, C^2HCl_3) 30.63 ($\underline{C}H_2S\underline{C}H_2$), 129.13, 129.49, 130.58 (Aromatic <u>C</u>H's), 134.30 & 135.03 (quat. aromatic & (ArCH=<u>C</u>)), 137.42 (Ar<u>C</u>H=C) and 189.51 (<u>C</u>O); *m/z* (CI) 293 ([*M* + H]$^+$, 100%).

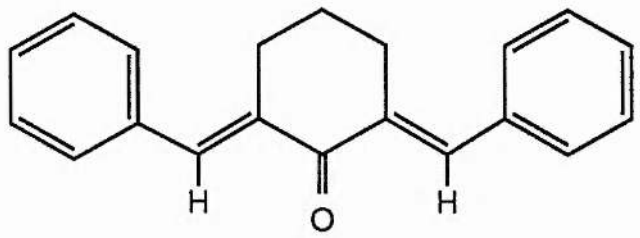
Tetrahydro-3,5-bis(phenylmethylene)-4*H*-pyran-4-one (71).



This was prepared in an identical manner to (**70**) using sodium hydroxide (1 g,

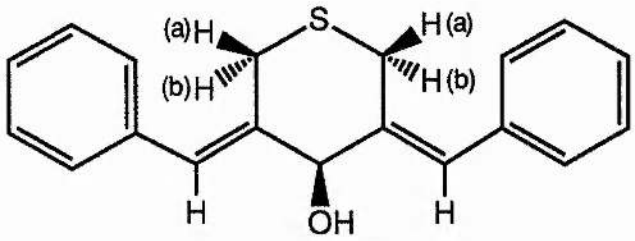
25 mmol), benzaldehyde (1.06 g, 10 mmol) and tetrahydro-4*H*-pyran-4-one (500 mg, 5 mmol), producing a yellow solid which was recrystallised from acetone to yield pale yellow needles (897 mg, 65%), m.p. 181-183 °C (lit.,[156] 185 °C); *m/z* (Found: [*M* + H]$^+$, 277.1229. $C_{19}H_{17}O_2$ requires 277.1228); ν_{max} (nujol)/ cm^{-1} 1668 (CO str.) and 1612 & 1582 (C=C str.); δ_H (200 MHz, C^2HCl_3) 5.00 (4H, s, $C\underline{H}_2OC\underline{H}_2$), 7.30-7.60 (10H, m, Ar C<u>H</u>'s) and 7.90 (2H, s, 2 x ArC<u>H</u>=C); δ_C (50.3 MHz, C^2HCl_3) 69.20 ($\underline{C}H_2O\underline{C}H_2$), 129.0, 129.8 & 131.0 (Aromatic <u>C</u>H's), 133.7 & 135.4 (quat. aromatic & (ArCH=<u>C</u>)), 137.2 (Ar<u>C</u>H=C) and 186.0 (<u>C</u>O); *m/z* (CI) 277 ([*M* + H]$^+$, 100%).

<u>2,6-bis(phenylmethylene)-cyclohexanone (72).</u>



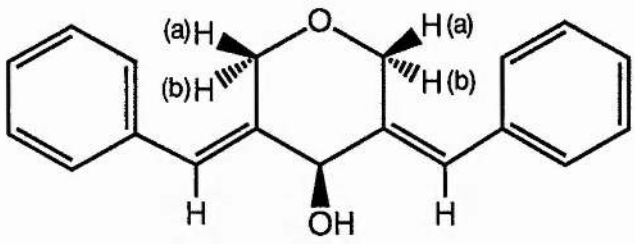
This was prepared in an identical manner to (**70**) using sodium hydroxide (1 g, 25 mmol), benzaldehyde (1.06 g, 10 mmol) and cyclohexanone (490 mg, 5 mmol), producing a yellow solid which was recrystallised from acetone to yield pale yellow needles (1.05 g, 77%), m.p. 116-118 °C (lit.[156] 117-118 °C); (Found: C, 87.42; H, 6.72; $C_{20}H_{18}O$ requires C, 87.56; H, 6.61%); *m/z* (Found: [*M* + H]$^+$ 275.1436. $C_{20}H_{19}O$ requires 275.1436); ν_{max} (nujol)/ cm^{-1} 1662 (CO str.) and 1606, 1576 & 1569 (C=C str.); δ_H (200 MHz, C^2HCl_3) 1.80 (2H, m, $CH_2C\underline{H}_2CH_2$), 2.95 (4H, m, $C\underline{H}_2CH_2C\underline{H}_2$), 7.30-7.60 (10H, m, Aromatic), 7.84 (2H, s, 2 x ArC<u>H</u>=C); δ_C (50.3 MHz, C^2HCl_3) 23.53 ($CH_2\underline{C}H_2CH_2$), 28.97 ($\underline{C}H_2CH_2\underline{C}H_2$), 128.90, 129.11, 130.88 (Aromatic <u>C</u>H's), 136.45 & 136.69 (quat. aromatic & (ArCH=<u>C</u>)), 137.44 (Ar<u>C</u>H=C) and 190.89 (<u>C</u>O); *m/z* (CI) 275 ([*M* + H]$^+$, 100%).

Tetrahydro-3,5-bis(phenylmethylene)-4*H*-thiopyran-4-ol (73).



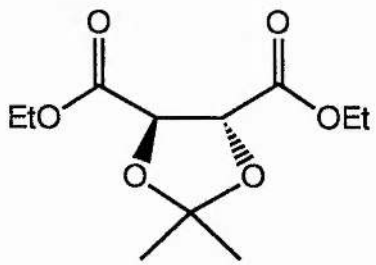
Tetrahydro-3,5-bis(phenylmethylene)-4*H*-thiopyran-4-one (**70**), (153 mg, 520 μmol) was suspended in ethanol (5 ml) and sodium borohydride (22 mg, 570 μmol) added portionwise over 15 min. at room temperature. The reaction was followed to completion both by colour change (yellow to colourless solution) and TLC. Water (5 ml) was added to the solution and the white precipitate filtered and washed with more water (10 ml). Recrystallisation from acetone/ water afforded colourless needles (145 mg, 94%), which were prone to decolourisation at room temperature and were stored at 0 °C in the dark; m.p. 117-119 °C; *m/z* (Found: M^+ 294.1078. $C_{19}H_{18}SO$ requires 294.1078); ν_{max} (nujol)/ cm^{-1} 3367 (OH str.) and 1598 & 1576 (C=C str.); δ_H (200 MHz, C^2HCl_3) 2.0 (1H, d, *J* 6 Hz, OH), 3.54 (2H, d, *J* 13.60 Hz, 2 x CH(b)), 3.86 (2H, d, *J* 13.60 Hz, 2 x CH(a)), 4.79 (1H, d, *J* 6 Hz, C<u>H</u>OH), 6.63 (2H, s, 2 x ArC<u>H</u>=C), 7.25-7.40 (10H, m, Aromatic); δ_C (50.3 MHz, C^2HCl_3) 26.81 ($\underline{C}H_2S\underline{C}H_2$), 81.04 (<u>C</u>HOH), 125.31 (Ar<u>C</u>H=C), 127.63, 128.89 & 129.47 (Aromatic <u>C</u>H's), 136.82, 140.82 (quat. aromatic & (ArCH=<u>C</u>)); *m/z* (EI) 294 (M^+, 8%), 278 ([*M* + H - OH]$^+$, 15) and 261 ([*M* - SH]$^+$, 12).

Tetrahydro-3,5-bis(phenylmethylene)-4*H*-pyran-4-ol (74).



This was prepared in an identical manner to (**73**) using (**71**) (168 mg, 610 μmol) and sodium borohydride (25 mg, 660 μmol) to yield a white solid which was recrystallised from acetone/ water to yield colourless needles (156 mg, 92%), which were prone to decolourisation at room temperature and were stored at 0 °C in the dark; m.p. 115 °C (dec.); *m/z* (Found: M^+, 278.1307. $C_{19}H_{18}O_2$ requires 278.1307); ν_{max} (nujol)/ cm^{-1} 3300 (OH str.), 1599 and 1575 (C=C str.); δ_H (200 MHz, C^2HCl_3) 2.48 (1H, d, *J* 6.2 Hz, OH), 4.55 (2H, d, *J* 13.55 Hz, 2 x CH(b)), 4.80 (2H, d, *J* 13.55 Hz, 2 x CH(a)), 4.93 (1H, d, *J* 6.2 Hz, C<u>H</u>OH), 6.68 (2H, s, 2 x ArC<u>H</u>=C) and 7.15-7.40 (10H, m, Aromatic); δ_C (50.3 MHz, C^2HCl_3) 67.25 ($\underline{C}H_2O\underline{C}H_2$), 76.27 (<u>C</u>HOH), 126.05 (Ar<u>C</u>H=C), 127.74, 128.88, 129.35 (Ar <u>C</u>H's), 136.58, 139.21 (quat. aromatic) and (ArCH=<u>C</u>); *m/z* (EI) 278 (M^+, 3%), 262 ([*M* + H - OH]$^+$, 2) and 248 ([*M* -CH_2O]$^+$, 4).

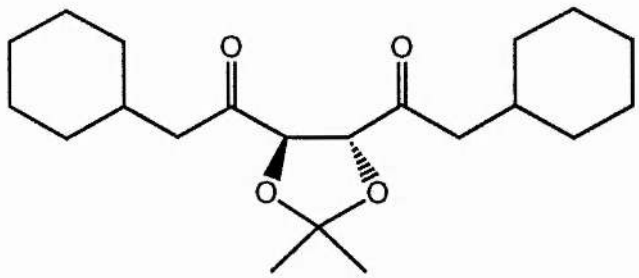
Diethyl 2,3-*O*-Isopropylidene-L-tartrate (75).



Using a modification of the method of Ikemoto,[159] diethyl L-tartrate (1 g, 4.85 mmol) was stirred for 24 h. at room temperature in a solution of concentrated H_2SO_4 (0.5 ml) in acetone (125 ml). The solution was cooled to 0 °C and

$NaHCO_3$ solution added until the solution reached pH 9. The acetone was removed under reduced pressure and the aqueous phase extracted with CH_2Cl_2 (3 x 50 ml). The organic layers were combined and washed with brine (2 x 30 ml), dried over anhydrous $MgSO_4$ and concentrated *in vacuo* to give a pale yellow oil which was purified by silica chromatography (2% EtOH, CH_2Cl_2) to yield pure ketal (717 mg, 60%), (Found: C, 53.46; H, 7.64. $C_{11}H_{18}O_6$ requires C, 53.65; H, 7.37%); *m/z* (Found: [*M* + H]$^+$ 247.1182. $C_{11}H_{18}O_6$ requires 247.1182); $[\alpha]_D$ -28.7° (c 1 in $CHCl_3$) (lit.[167] -40.2 (c 1 in $CHCl_3$)); ν_{max} (neat) 3468 (CH stretch) and 1758 (CO stretch); δ_H (200 MHz, C^2HCl_3) 1.32 (6H, t, *J* 7.6 Hz, 2 x $OCH_2C\underline{H}_3$), 1.50 (6H, s, $C(C\underline{H}_3)_2$), 4.28 (4H, q, *J* 7.6 Hz, 2 x $OC\underline{H}_2CH_3$) and 4.77 (2H, s, 2 x CH); δ_C (50.3 MHz, C^2HCl_3) 14.59 ($OCH_2\underline{C}H_3$), 26.86 ($C(\underline{C}H_3)_2$), 62.37 ($O\underline{C}H_2CH_3$), 77.64 & 77.68 ($\underline{C}$H's), 114.24 ($\underline{C}(CH_3)_2$) and 170.13 ($\underline{C}$O); *m/z* (CI) 247 ([*M* + H]$^+$, 100%).

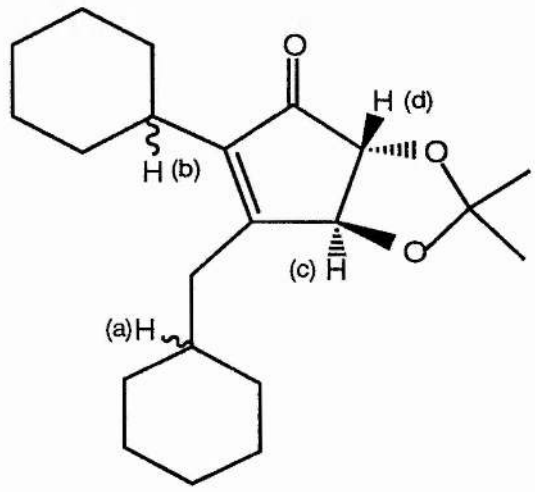
1,6-Dicyclohexyl-(3R,4R)-*O*-Isopropylidene-hexane-2,5-dione (77).



Using a modification of the method of Kikkawa,[163] magnesium turnings (243 mg, 10 mmol) and a few crystals of iodine were covered with anhydrous ether (10 ml) under an argon atmosphere. Freshly distilled cyclohexylmethyl bromide (2.21 g, 12.5 mmol) in anhydrous ether (20 ml) was added dropwise and the solution refluxed for 15 min. (until magnesium had dissolved) using a water bath. The solution was cooled to 0 °C and triethylamine (3.34 ml, 24 mmol) added, followed by dropwise addition of dimethyl 2,3-*O*-Isopropylidene-L-tartrate (872 mg, 4 mmol) in anhydrous ether (15 ml) over 30 min. The reaction mixture was left at 0 °C for 2 h. and then methanol (5 ml) added. The ether was removed under reduced pressure and the resultant oil taken up in ethyl acetate (100 ml), washed with 1 M HCl (3 x 25 ml), $NaHCO_3$ (1 x 25 ml), brine (2 x 25

ml) and dried over anhydrous $MgSO_4$. The ethyl acetate was concentrated *in vacuo* to give a pale yellow oil which was purified by silica chromatography (100% CH_2Cl_2) to yield pure diketone (448 mg, 32%), *m/z* (Found: $[M + H]^+$ 351.2535. $C_{21}H_{35}O_4$ requires 351.2535); $[\alpha]_D$ -5.5° (c 0.5 in MeOH); ν_{max} (neat) 3436 (CH stretch) and 1724 (CO stretch); δ_H (200 MHz, C^2HCl_3) 0.80-1.35 (10H, m, 5 x CH_2 (Chex)), 1.40 (6H, s, $C(C\underline{H}_3)_2$), 1.55-1.75 (10H, m, 5 x CH_2 (Chex)), 1.88 (2H, m, CH (Chex)), 2.50 (2H, d, *J* 6.86 Hz, Chex$C\underline{H}_2$CO) and 4.50 (2H, s, 2 x CH); δ_C (50.3 MHz, C^2HCl_3) 26.53 ($\underline{C}H_2$ (Chex)), 26.64 ($\underline{C}H_2$ (Chex)), 26.81 ($C(\underline{C}H_3)_2$), 33.59 ($\underline{C}H_2$ Chex), 33.65 ($\underline{C}H$ (Chex)), 33.68 ($\underline{C}H_2$ (Chex)), 47.03 (Chex$\underline{C}H_2$CO), 82.09 ($\underline{C}H$), 112.88 ($\underline{C}(CH_3)_2$) and 208.67 ($\underline{C}O$); *m/z* (CI) 351 ($[M + H]^+$, 100%) and 333 ($[M + H - H_2O]^+$, 60).

2-Cyclohexyl-3-cyclohexylmethyl-(4R,5R)-*O*-Isopropylidene-cyclopent-2-enone (78).



Isolated from the base catalysed cyclisation of (**77**), *m/z* (Found: M^+, 332.236 $C_{21}H_{32}O_3$ requires 332.235); ν_{max} (C^2HCl_3) 1713 (CO stretch) and 1624 (C=C stretch); δ_H (200 MHz, C^2HCl_3) 0.80-2.00 (21H, m, 10 x CH_2 (Chex) & CH (a) (Chex)), 1.30 (3H, s, 1 x $C(C\underline{H}_3)_2$), 1.40 (3H, s, 1 x $C(C\underline{H}_3)_2$), 2.30-2.60 (3H, m, CH (b) (Chex) & Cha$C\underline{H}_2$C), 4.37 (1H, d, *J* 6 Hz, CH (c)) and 5.02 (1H, d, *J* 6 Hz, CH (d)); δ_C (50.3 MHz, C^2HCl_3) 26.16, 26.69, 26.90, 26.99 & 27.18 ($\underline{C}H_2$ (Chex)), 27.37 & 28.31 ($C(\underline{C}H_3)_2$), 29.58, 29.88, 30.57, 33.63 & 33.70 ($\underline{C}H_2$

Chex), 36.57 (CH (a) (Chex)), 36.77 (CH (b) (Chex)), 36.93 (ChexCH_2C), 77.50 (CH (c)), 78.88 (CH (d)), 114.54 ($C(CH_3)_2$), 145.68 ($CH_2CR=CR'CO$), 169.13 ($CH_2CR=CR'CO$) and 204 (CO); *m/z* (EI) 332 (M^+, 3%), 317 ($[M - CH_3]^+$, 12), 302 ($[M - 2 \times CH_3]^+$, 2), 290 ($[M - C(CH_3)_2]^+$, 2) and 274 ($[M-(CH_3)_2CO]$+, 54).

HIV-1 protease assays (79).

***in vitro* assay.**

The assays are based upon the cleavage of the decapeptide Lys-Ala-Arg-Val-Nle-(NO_2)Phe-Glu-Ala-Nle-Gly-NH_2 between the norleucine and nitrophenylalanine residues to give a reduction in absorbance at 300 nm. The peptide is stored as a 10 mg ml^{-1} solution in water and 2 µl is added to 1 ml of the assay buffer. The assay buffer is;
100 mM NaOAc
200 mM NaCl
1mM DTT
pH 5.6

The inhibitors were made up as 1-10 mM stock solutions in methanol and were used at concentrations in the range 1-100 µM. Methanol was added to give a final concentration of 2% and the mixture was allowed to thermally equilibrate at 37 °C for 5 min. The reaction was then initiated by the addition of 15 µl of enzyme. The enzyme was stored as a 0.993 µM solution in 10 mM NaOAc, 0.05% 2-mercaptoethanol, 1 mM EDTA, 20% glycerol and 5% ethylene glycol. The procedure for pre-incubated assays was similar to the above, except that the enzyme and inhibitor were mixed and allowed to incubate together for 15 mins at 37 °C and the reaction was initiated by the addition of the substrate. The linear portion of the plot was used to give initial rates. Each assay was performed both in duplicate and on two separate occasions. The K_M for the substrate was determined to be 33 µM under the assay conditions.

***in vivo* assay.**

Antigen reduction assay.

C8166 T-lymphoblastoid cells were infected with 10 $TCID_{50}$ of HIV-1 of the RF stock ($HIV\text{-}1_{RF}$) and washed three times with PBS. Aliquots of 2 x 10^5 cells were resuspended in growth medium, RPMI 1640 supplemented by 10% fetal calf serum, with each compound to be tested at log dilutions from 0.001 to 100 μM. The cells were incubated at 37 °C for 72 h. under a 95% air/ 5% CO_2 atmosphere. The HIV antigen was measured in duplicate in the supernatants using a commercial ELISA assay.

Toxicity assay.

Aliquots of 2 x 10^5 uninfected C8166 cells were suspended in 1.5 ml of the above growth medium and cultured with the test compounds at the same log dilutions for 72 h. The cells were then washed with PBS and resuspended in 200 μl of growth medium containing ^{14}C protein hydrolysate. The cells were harvested after 16 h. and the ^{14}C incorporation measured. Untreated cells were used as controls.

APPENDIX A

Publications.

1. Camp, N.P., Hawkins, P.C.D., Hitchcock, P.B. and Gani, D., *Bioorg. Med. Chem. Lett.*, **2**, 1992, 1047.

2. Camp, N.P., Hawkins, P.C.D., Perrey, D.A. and Gani, D., in preparation.

APPENDIX B

Crystallographic data.

Crystallographic data have been deposited at the Cambridge Data Centre. $C_{32}H_{47}N_4O_6P$, M = 614.7. Orthorhombic, $P2_12_12_1$ (No. 19), a = 9.265 (5), b = 15.654 (10), c = 23.929 (11) Å; V = 3470.4 Å^3; Z = 4; d_c = 1.18 g cm^{-3}. Crystal size 0.25 x 0.15 x 0.02 mm. Total unique reflections = 3462, R = 0.118 for 738 reflections with $|F^2| > 2\sigma(F^2)$.

CHAPTER FOUR

REFERENCES

4.0 References

1. Bowen, D.L., Lane, H.C. and Fauci, A.S., *Ann. Intern. Med.*, 1985, **103**, 704.
2. Klecker, R.W., Collins, J.M., Yarchoan, R., Thomas, R., Jenkins, J.F., Broder, S. and Myers, C.E., *Clin. Pharmacol. Ther.*, 1987, **41**, 407.
3. Yarchoan, R., Klecker, R.W., Weinhold, K.J., Markham, P.D., Lyerly, H.K., Durack, D.T., Gelmann, E., Nusinoff-Lehrman, S., Blum., R.M., Barry, D.W., Shearer, G.M., Fischl, M.A., Mitsuya, H., Gallo, R.C., Collins, J.M., Bolognesi, D.P., Myers, C.E. and Broder, S., *Lancet* 1986, Vol. **I**, 575.
4. Gallo, R.C., Sarin, P.S., Gelmann, E.P., Robert-Guroff, M., Richardson, E., Kalyanaraman, V.S., Mann, D., Sidhu, G.D., Stahl, R.E., Zolla-Pazner, S., Leibowitch, J. and Popovic, M., *Science*, 1983, **220**, 865.
5. Clavel, F., Guetard, D., Brunvezinet, F., Charmaret, S., Rey, M.A., Santosferreira, M.O., Laurent, A.G., Dauguet, C., Katlama, C., Rouzioux, C., Klatzmann, D., Champalimaud, J.L and Montagnier, L., *Science*, 1986, **233**, 343.
6. Ratner, L., Haseltine, W., Pataraca, R., Livak, K.J., Starcich, B., Josephs, S.F., Doran, E.R., Rafalski, J.A., Whitehorn E.A., Baumeister, K., Ivanoff, L., Petteway, S.R., Jr., Pearson, M.L., Lautenberger, L.A., Papas, T., Ghrayeb, J., Chang, N.T., Gallo, R.C. and Wong-staal, F., *Nature*, 1985, **313**, 277.
7. Popovic, M., Read-connole, E. and Gallo, R.C., *Lancet* 1984, Vol. **II**, 1472.
8. Levy, J.A., Hollander, H., Shimabukuro, J., Mills, J. and Kaminsky, L., *Lancet* 1985, Vol. **II**, 586.
9. Nelson, J.A., Wiley, C.A., Reynolds-Kohler, C., Margaretten, W., Reese, C.E. and Levy J.A., *Lancet* 1988, Vol. **I**, 259.
10. Cohen, A.H., Sun, N., Shapshak, P. and Imagawa, D.T, *Mod. Pathol.*,1989, **2**, 125.
11. Cao., Y., Friedman-Klein., A., Huang, Y., Li, X, Mirabile, M., Moudgil, T., Zucker-Franklin, D. and Ho, D., *J. Virol.*, 1990, **64**, 2553.
12. Plata, F., Garcia-pons, F., Ryter, A., Lebargy, F., Goodenow, M.M, Dat., M.H., Autran, B. and Mayand, C., *AIDS Res. Hum. Retroviruses.*, 1990, **6**, 979.
13. Tateno, M., Gonzalez-Scarano, F. and Levy, J.A., *Proc. Natl. Acad. Sci. U.S.A.*, 1989, **86**, 4287.
14. Harouse, J.M., Kunsch, C., Hartle, H.T., Laughlin, M.A., Hoxie, J.A., Wigdahl, B. and Gonzalez-Scarano, F., *J. Virol.*, 1989, **63**, 2527.

15. Roberts, M.M. and Oroszlan, S., 1990, In Retroviral Proteases: Control of Maturation and Morphogenesis (L.H. Pearl, ed.), M. Stockton Press.
16. Sodroski, J., Patarca, R., Rosen, C. and Haseltine W.A., *Science*, 1985, **229**, 74.
17. Sodroski, J.G., Goh, W.C., Rosen, C., Dayton, A., Terwilliger, E. and Haseltine, W.A., *Nature*, 1986, **321**, 412.
18. Tomasselli, A.G., Howe, W.J., Sawyer, T.K., Wlodawer, A., and Heinrikson, R.L., *Chimica Oggi*, 1991, **5**, 6.
19. Huff, J.R., *J. Med. Chem.*,1991, **34**, 2305.
20. Debouk, C., *AIDS Res. Hum. Retroviruses*, 1992, **8**, 153.
21. Martin, J.A., *Antiviral Res.*, 1992, **17**, 265.
22. Poli, G., Orenstein, J.M., Kinter, A., Folks, T. and Fauci, A.S., *Science*, 1989, **244**, 575.
23. Toh, H., Ono, M., Saigo, K. and Miyata, T., *Nature*, 1985, **315**, 691.
24. Darke, P.L., Leu., C., Davis, L.J., Heimbach, J.C., Diehl, R.E., Hill, W.S., Dixon, R.A. and Sigal, I.S., *J. Biol. Chem.,* 1989, **264**, 2307.
25. Meek, T.D., Dayton, B.D., Metcalf, B.W., Dreyer, G.B., Strickler, J.E., Gorniak, J.G., Rosenberg, M., Moore, M.L., Magaard, V.W. and Debouck, C., *Proc. Natl. Acad. Sci. U.S.A.*, 1989, **86**, 1841.
26. Tang, J., *J. Biol. Chem.,* 1971, **246**, 4510.
27. Kohl, N.E., Emini, E.A., Schleif, W.A., Davis, L.J., Heimbach, J.C., Dixon, R.A.F., Scolnick, E.M. and Sigal, I.S., *Proc. Natl. Acad. Sci. U.S.A.*, 1988, **85**, 4686.
28. Navia, M.A, Fitzgerald, P.M.D., Mckeever, B.M., Leu, C.T., Heimbach, J.C., Herber, W.K., Sigal, I.S., Darke, P.L. and Springer, J.P., *Nature*, 1989, **337**, 615.
29. Wlodawer, A., Miller, M., Jaskolski, M., Sathyanarayana, B.K., Baldwin, E., Weber, I.T., Selk, L.M., Clawson, L., Schneider, J. and Kent, S.B.H., *Science*, 1989, **245**, 616.
30. Lappatto, R., Blundell, T., Hemmings, A., Overington, J., Wilderspin, A., Wood, S., Merson, J.R., Whittle, P.J., Danley, D.E., Geoghegan, K.F., Hawrylik, S.J., Lee, S.E., Scheld, K.G. and Hobart, P.M., *Nature*, 1989, **342**, 299.
31. Pearl, L. and Taylor, W., *Nature*, 1987, **329**, 351.
32. Lillehoj, E.P., Salazar, F.H., Mervis, R.J., Raum, M.G., Chan, H.W., Ahmed, N. and Venkatesan, S.J., *J. Virol.*, 1988, **62**, 3053.

33. Schneider, J. and Kent, S.B., *Cell*, 1988, **54**, 363.
34. Nutt, R.F., Brady, S.F., Darke, P.L., Ciccarone, T.M., Colton, C.D., Nutt, E.M., Rodkey, J.A., Bennett, C.D., Waxman, L.H., Sigal, I.S., Anderson, P.S. and Veber, D.F., *Proc. Natl. Acad. Sci. U.S.A.*, 1988, **85**, 7129.
35. Debouck, C., Gorniak, J.G., Strickler, J.E., Meek, T.D., Metcalf, B.W. and Rosenburg., *Proc. Natl. Acad. Sci. U.S.A.*, 1987, **84**, 8903.
36. Kramer, R.A., Schaber, M.D., Skalka, A.M., Ganguly, K., Wong-staal, F. and Reddy, E.P., *Science*, 1986, **231**, 1580.
37. Rittenhouse, J., Turon, M.C., Helfrich, R.J., Albrecht, K.S., Weigl, D., Simmer, R.L., Mordini, F., Erickson, J. and Kohlbrenner, W.E., *Biochem. Biophys. Res. Commun.*, 1990, **171**, 60.
38. Tomasselli, A.G., Olsen, M.K., Hui, J., Staples, D.J., Sawyer, T.K., Heinrikson, R.L. and Tomich, C.S., *Biochemistry*, 1990, **29**, 264.
39. Schechter, I. and Berger, A., *Biochem. Biophys. Res. Commun.*, 1967, **27**, 157.
40. Pearl, L. and Taylor, W., *Nature*, 1987, **328**, 482.
41. Darke, P.L., Nutt, R.F., Brady, S.F., Garsky, V.M., Cicarone, T.M., Leu, C.T., Lumma, P.K., Freidinger, R.M., Veber, D.F. and Sigal, I.S., *Biochem. Biophys. Res. Commun.*, 1988, **156**, 297.
42. Krausslich, H., Ingraham, R.H., Skoog, M.T., Wimmer, E., Pallai, P.V. and Carter, C.A., *Proc. Natl. Acad. Sci. U.S.A.*, 1989, **86**, 807.
43. Partin, K., Krausslich, H., Ehrlich, L., Wimmer, E. and Carter, C.A., *J. Virol.*, 1990, **64**, 3938.
44. Loeb, D.D., Hutchinson, C.A., Edgell, M.H., Farmerie, W.G. and Swanstrom, R., *J. Virol.*, 1989, **63**, 111.
45. Weber, I.T., *J. Biol. Chem.*, 1990, **265**, 10492.
46. Knowles, J.R., *Proc. Roy. Soc. B.*, 1970, **26**, 172.
47. Hunkapiller, M.W. and Richards, J.H., *Biochemistry*, 1972, **11**, 2829.
48. Antonov, V.K., Ginodman, L. M., Rumsh, L.D., Kapitannikov, Y.V., Barshevskaya, T.N., Yavashev, L.P., Gurova A.G. and Volkova, L.I., *Bioorgan. Khimia*, 1980, **6**, 436.
49. Hyland, L.J., Tomaszek, T.A. and Meek, T.D., *Biochemistry*, 1991, **30**, 8454.
50. James, M.G and Sielecki, A.R., *Biochemistry*, 1985, **24**, 3701.
51. Polgar, L., *FEBS Lett.*, 1987 **219**, 1.
52. Pearl, L., *FEBS Lett.*, 1987, **214**, 8.

53.. Blundell, T.L., Jenkins, J.A., Sewell, B.T., Pearl, L.H., Cooper, J.B., Tickle, I.J., Veerapandian B. and Wood, S.P., *J. Mol. Biol.*, 1990, **211**, 919.
54. Davies, D.R., *Ann. Rev. Biochem.*, 1991, **19**, 189.
55. Wolfenden, R., *Annu. Rev. Biophys. Bioeng.*, 1976, **5**, 271.
56. Fischer, G., *Nat. Prod. Rep.*, 1988, **5**, 465.
57. Dreyer, G.B., Metcalf, B.W., Tomaszek, T.A., Carr, T.J., Chandler, A.C., Hyland, L., Fakhoury, S.A., Magaard, V.W., Moore, M.L., Strickler, J.E., Debouck, C. and Meek, T.D., *Proc. Natl. Acad. Sci. U.S.A.*, 1989, **86**, 9752.
58. Billich, S., Knoop, T., Hansen, J., Strop, P., Sedlacek, J., Mertz, R. and Moelling, K., *J. Mol. Biol.*, 1988, **263**, 17905.
59. Moore, M.L., Bryan, W.M., Fakhoury, S.A., Maagard, V., Huffman, W.F., Dayton, B.D., Meek, T.D., Hyland, L., Dreyer, G.B., Metcalf, B.M., Strickler, J.E., Gorniak, J.G. and Debouck, C., *Biochem. Biophys. Res. Commun.*, 1989, **159**, 420.
60. Szelke, M., in *Aspartic proteinases and Their Inhibitors*; ed. Kosta, V., Walter de Gruyter: Berlin, 1985, 412.
61. Szelke, M., Leckie, B., Hallett, A., Jones, D.M., Sueiras, J., Atrash, B. and Lever, A.F., *Nature*, 1982, **299**, 555.
62. Cushman, M., Oh, J., Copeland, T.D., Oroszlan, S. and Snyder, S.W., *J. Org. Chem.*, 1991, **56**, 4161.
63. Miller, M., Schneider, J., Sathyanarayana, B.K., Toth, M.V., Marshall, G.R., Clawson, L., Selk, L., Kent, S.B. and Wlodawer, A., *Science*, 1989, **246**, 1149.
64. Urban, J., Konvalinka, J., Stehlikova, J., Gregorova, E., Majer, P., Soucek, M., Andreansky, M., Fabry, M. and Strop, P., *FEBS Lett.*, 1992, **298**, 9.
65. Bartlett, P.A. and Kezer, W.B., *J. Am. Chem. Soc.*, 1984, **106**, 4282.
66. Crofts, P.C. and Kosolapoff, G.M., *J. Am. Chem. Soc.*, 1953, **75**, 3379.
67. Bartlett, P.A., Marlowe, C.K., Giannousis, P.P. and Hansen, J.E., *Cold Spring Harbor Symp. Quant. Biol.*, 1987, **52**, 83.
68. Grobelny, D., Wondrak, E.M., Galardy, R.E. and Oroszlan, S., *Biochem. Biophys. Res. Commun.*, 1990, **169**, 1111.
69. Abdel-Meguid, S.S., Zhao, B., Murthy, K., Winborne, E., Choi, J.K., DesJarlais, R., Minnich, M., Culp, J., Debouck, C., Tomaszek, T., Meek, T.D. and Dreyer, G.B., *Biochemistry*, 1993, **32**, 7972.
70. Rich, D., *J. Med., Chem.*, 1985, **28**, 263.

71. James, M., Sielecki, A., Salituro, F., Rich, D. and Hofmann,T., *Proc. Natl. Acad. Sci. U.S.A.*, 1982, **79**, 6137.
72. Bott, R., Subramanian E. and Davies, D., *Biochemistry,* 1981, **21**, 6956.
73. Richards, A.D., Roberts, R.F., Dunn, B.M., Graves, M.C. and Kay, J., *FEBS Lett.*, 1989, **247**, 113.
74. Mimoto, T., Imai, J., Tanaka, S., Hattori, N., Takahashi, O., Kisanuki, S., Nagano, Y., Shintani, M., Hayashi, H., Sakikawa, H., Akaji, K. and Kiso, Y., *Chem. Pharm. Bull.*, 1991, **39**, 2465.
75. Tam, T., Carriere, J., Macdonald, D., Castelhano, A., Pliura, D., Dewdney, N., Thomas, E., Bach, C., Barnett, J., Chan, H. and Krantz, A., *J. Med., Chem.*, 1992, **35**, 1318.
76. Gelb, M.H., Svaren, J.P., and Abeles, R.H., *Biochemistry*, 1985, **24**, 1813.
77. Thaisirivongs, S., Pals, D.T., Kati, W.M., Turner, S.R., and Thomasco, L.M., *J. Med. Chem.*, 1985, **28**, 1553.
78. Sham, H.L., Betebenner, D.A., Wideburg, N.E., Saldivar, A.C., Kohlbrenner, W.E., Vasavanonda, S., Kempf, D.J., Norbeck, D.W., Zhao, C., Clement, J.J., Erickson, J. and Plattner, J.J., *Biophys. Res. Commun.*, 1991, **175**, 914.
79. Sham, H.L., Wideburg, N.E., Spanton, S.G., Kohlbrenner, W.E., Betebenner, D.A., Kempf, D.J., Norbeck, D.W., Plattner, J.J. and Erickson, J., *J. Chem. Soc., Chem. Commun.*, 1991, 110.
80. Roberts, N.A., Martin, J.A., Kinchington, D., Broadhurst, A.V., Craig, J.C., Duncan, I.B., Galpin, S.A., Handa, B.K., Kay, J., Krohn, A., Lambert, R.W., Merrett, J.H., Mills, J.S., Parkes, K.E.B., Redshaw, S., Ritchie, A.J., Taylor, D.L., Thomas, G.J. and Machin, P.J., *Science*, 1990, **248**, 358.
81. Rich, D., Sun, C., Prasad, J., Pathiasseril, A., Toth, M., Marshall, G., Clare, M., Mueller, R. and Houseman, K., *J. Med. Chem.*, 1991, **34**, 1222.
82. Tucker, T.J., Lumma, W.C., Payne, L.S., Wai, J.M., de Solms, J., Giuliani, E.A., Darke, P.L., Heimbach, J.C., Zugay, J.A., Schleif, W.A., Quintero, J.C., Emini, E.A., Huff, J.R. and Anderson, P.S., *J. Med. Chem.*, 1992, **35**, 2525.
83. Ghosh, A.K., Thompson, W.J., Lee, H.Y., McKee, S.P., Munson, P.M., Duong, T.T., Darke, P.L., Zugay, J.A., Emini, E.A., Schleif, W.A., Huff, J.R. and Anderson, P.S., *J. Med. Chem.*, 1993, **36**, 924.
84. Thompson, W.J., Ghosh, A.K., Holloway, K., Lee, H.Y., Munson, P.M., Schwering, J.E., Wai, J.M., Darke, P.L., Zugay, J.A., Emini, E.A., Schleif, W.A., Huff, J.R. and Anderson, P.S., *J. Am. Chem. Soc.*, 1993, **115**, 801.

85. Ghosh, A.K., Thompson, W.J., Holloway, K., McKee, S.P., Duong, T.T., Lee, H.Y., Munson, P.M., Smith, A.M., Wai, J.M., Darke, P.L., Zugay, J.A., Emini, E.A., Schleif, W.A., Huff, J.R. and Anderson, P.S., *J. Med. Chem.*, 1993, **36**, 2300.
86. Prasad, J. and Rich, D.H., *Tetrahedron Lett*, 1991, **32**, 5857.
87. Thompson, W.J., Ball, R.G., Darke, P.L., Zugay, J.A. and Thies, J.E., *Tetrahedron Lett*, 1992, **33**, 2957.
88. McQuade, T.J., Tomasselli, A.G., Liu, L., Karacostas, V., Moss, B., Sawyer, T.K., Heinrikson, R.L. and Tarpley, W.G., *Science*, 1990, **247**, 454.
89. Lyle, T.A., Wiscount, C., Guare, J., Thompson, W.J., Anderson, P.S., Darke, P.L., Zugay, J.A., Emini, E.A., Schleif, W.A., Quintero, J.C., Dixon, R., Sigal, I. and Huff, J.R., *J. Med. Chem.*, 1991, **34**, 1228.
90. Vacca, J.P., Guare, J.P., de Solms, J., Sanders, W.M., Giuliani, E.A., Young, S.D., Darke, P.L., Zugay, J.A., Sigal, I., Schleif, W.A., Quintero, J.C., Emini, E.A., Anderson, P.S. and Huff, J.R., *J. Med. Chem.*, 1991, **34**, 1225.
91. de Solms, J., Giuliani, E.A., Guare, J.P., Sanders, W.M., Graham, S.L., Wiggins, M., Darke, P.L., Sigal, I., Zugay, J.A., Emini, E.A., Schleif, W.A., Quintero, J.C., Emini, E.A., Anderson, P.S. and Huff, J.R., *J. Med. Chem.*, 1991, **34**, 2852.
92. Thompson, W.J., Fitzgerald, P.M., Holloway, K., Emini, E.A., Darke, P.L., McKeever, B.M., Schleif, W.A., Quintero, J.C., Zugay, J.A., Tucker, T.J., Schwering, J.E., Homnick, C.F., Nunberg, J., Springer, J.P. and Huff, J.R., *J. Med. Chem.*, 1992, **35**, 1685.
93. Young, S.D., Payne, L.S., Thompson, W.J., Gaffin, N., Lyle, T.A., Britcher, S.F., Graham, S.L., Schultz, T.H., Deana, A.A., Darke, P.L., Zugay, J.A., Schleif, W.A., Quintero, J.C., Emini, E.A., Anderson, P.S. and Huff, J.R., *J. Med. Chem.*, 1992, **35**, 1702.
94. Sawyer, T.K., Fisher, J.F., Hester, J.B., Smith, C.W., Tomasselli, A.G., Tarpley, W.G., Burton, P.S., Hui, J.O., McQuade, T.J., Conradi, R.A., Bradford, V.S., Liu, L., Kinner, J.H., Tustin, J., Alexander, D.L., Harrison, A.W., Emmert, D.E., Staples, D.J., Maggiora, L.L., Zhang, Y.Z., Poorman, R.A., Dunn, B., Rao, C., Scarborough, P.E., Lowther, W.T., Craik, C., DeCamp, D., Moon, J., Howe, W.J. and Heinrikson, R.L., *Bioorg. Med. Chem. Lett.*, 1993, **3**, 819.
95. Ashorn, P., McQuade, T.J., Thaisrivongs, S., Tomasselli, A.G., Tarpley, W.G. and Moss, B., *Proc. Natl. Acad. Sci. U.S.A.*, 1990, **87**, 7472.

96. Thaisrivongs, S., Tomasselli, A.G., Moon, J.B., Hui, J., McQuade, T.J., Turner, S.R., Strohbach, J.W., Howe, J., Tarpley, W.G. and Heinrikson, R.L, *J. Med. Chem.*, 1991, **34**, 2344.
97. Thaisrivongs, S., Turner, S.R., Strohbach, J.W., TenBrink, R,E., Tarpley, W.G., McQuade, T.J., Heinrikson, R.L, Tomasselli, A.G., Hui, J. and Howe, J., *J. Med. Chem.*, 1993, **36**, 941.
98. Getman, D.P., DeCrescenzo, G.A., Heintz, R.M., Reed, K.L., Talley, J.J., Bryant, M.L., Clare, M., Houseman, K., Marr, J.J., Mueller, R.A., Vazquez, M.L., Shieh, H-S, Stallings, W.C. and Stegeman, R.A., *J. Med. Chem.*, 1993, **36**, 288.
99. Sham, H.L., Betebenner, D.A., Zhao, C., Wideburg, N.E., Saldivar, A., Kempf, D.J., Plattner, J.J. and Norbeck, D.W., *J. Chem. Soc. Commun.*, 1993, 1053.
100. Ghosh, A.K., McKee, S.P., Thompson, W.J., Darke, P.L. and Zugay, J.A., *J. Org. Chem.*, 1993, **58**, 1025.
101. Kempf, D.J., Norbeck, D.W., Codacovi, L.M., Wang, X.C., Kohlbrenner, W.E., Wideburg, N.E., Paul, D.A., Knigge, M.F., Vasavanonda, S., Craig-Kennard, A., Saldivar, A., Rosenbrook, W., Clement, J.J., Plattner, J.J. and Erickson, J.J., *J. Med. Chem.*, 1990, **33**, 2687.
102. Erickson, J., Neidhart, D., Van Drie, J., Kempf, D.J., Wang, X., Norbeck, D., Plattner, J., Rittenhouse, J., Turon, M., Wideburg, N., Kohlbrenner, W., Simmer, R., Helfrich, R., Paul, D. and Knigge, M., *Science*, 1990, **249**, 527.
103. Kempf, D.J., Codacovi, L.M., Wang, X., Kohlbrenner, W., Wideburg, N., Saldivar, A., Vasavanonda, S., Marsh, K.C., Bryant, P., Sham, H.L., Green, B.E., Betebenner, D.A., Erickson, J.J. and Norbeck, D., *J. Med. Chem.*, 1993, **36**, 320.
104. Peyman, A., Budt, K.H., Spanig, J., Stowasser, B. and Ruppert D., *Tetrahedron Lett*, 1992, **33**, 4549.
105. Spaltenstein, A., Leban, J.J. and Furfine, E.S., *Tetrahedron Lett*, 1993, **34**, 1457.
106. Ghosh, A.K., McKee, S.P. and Thompson, W.J., *Tetrahedron Lett*, 1991, **32**, 5729.
107. Jadhav, P.K. and Woerner, F.J., *Bioorg. Med. Chem. Lett.*, 1992, **2**, 353.
108. Moelling, K., Schulze, T., Knoop, M-T., Kay, J., Jupp, R., Nicolaou, G. and Pearl, L., *FEBS Lett.*, 1990, **261**, 373.
109. Desjarlais, R.L., Seibel, G.L., Kuntz, I.D., Furth, P.S., Alvarez, J.C., Ortiz De Montellano, P.R., DeCamp, D.L., Babe, L.M. and Craik, C.S., *Proc. Natl.*

Acad. Sci. U.S.A., 1989, **86**, 9752.

110. Humber D.C., Cammack, N., Coates, J.A., Cobley, K.N., Orr, D.C., Storer, R., Weingarten, G.G. and Weir, M.P., *J. Med. Chem.*, 1992, **35**, 3081.
111. Brinkworth, R.I., Woon, T.C. and Fairlie, D.P., *Biophys. Res. Commun.*, 1991, **176**, 241.
112. Potts, B.C. and Faulkner, D.J., *J. Am. Chem. Soc.*, 1991, **113**, 6321.
113. Brinkworth, R.I., Stoermer, M.J. and Fairlie, D.P., *Biophys. Res. Commun.*, 1992, **188**, 631.
114. DeCamp, D.L., Babe, L.M., Salto, R., Lucich, J.L., Koo, M-S., Kahl, S.B. and Craik, C.S., *J. Med. Chem.*, 1992, **35**, 3426.
115. Friedman, S.H., DeCamp, D.L., Sijbesma, R.P., Srdanov, G., Wudl, F. and Kenyon, G.L., *J. Am. Chem. Soc.*, 1993, **115**, 6506.
116. Meek, T.D., Lambert, D.M., Dreyer, G.B., Carr, T.J. and Tomaszek, T.A., *Nature*, 1990, **343**, 90.
117. Petteway, S.R., Lambert, D.M. and Metcalf, B.W., *Trends Pharm. Sci.*, 1991, **12**, 28.
118. Meek, T.D., *J. Enzyme Inhibition*, 1992, **6**, 65; ref 146.
119. Blundell, T.L., Jenkins, J., Pearl, L., Sewell, T. and Pedersen, V., *Aspartic Proteinases and their Inhibitors*; Kostka, V., Ed.; Walter de Gruyter: Berlin, 1985; 151.
120. Pearl, L. and Blundell, T.L., *FEBS Lett.*, 1984, **174**, 96.
121. Fitzgerald, P.M., McKeever, B.M., VanMiddlesworth, J.F., Springer, J.P., Heimbach, J.C., Leu, C-T., Herber, W.K., Dixon, R.A. and Darke, P.L., *J. Biol. Chem.*, 1990, **265**, 14209.
122. Swain, A.L, Miller, M.M., Green, J., Rich, D.H., Kent, S.B. and Wlodawer, A., *Proc. Natl. Acad. Sci. U.S.A.*, 1990, **87**, 8805.
123. Greenlee, W.J., *Med. Res. Rev.*, 1990, **10**, 173.
124. Plattner, J.J. and Norbeck, D.W., Drug Discovery Tech., 1990; Clark, C.R., Moos W.H. Eds.; Ellis Horwood: Chichester, England; 92.
125. Sheldrick, W.S. and Morr, M., *Acta Crystallogr., Sect. B*, 1981, **B37**, 733.
126. Oney, I. and Caplow, M., *J. Am. Chem. Soc.*, 1967, **89**, 6972 .
127. McLeod, D.A., Brinkworth, R.I., Ashley, J.A., Janda, K.D. and Wirsching, P., *Bioorg. Med. Chem. Lett.*, 1991, **1**, 653.
128. Dewar, M.J.S., Zoebisch, E. G., Healy, E.F. and Stewart, J.J., *J. Am. Chem. Soc.*, 1985, **107**, 3902.

129. Teraishi, K., Saito, M., Fujii., I. and Nakamura, H., *Tetrahedron Lett.*, 1992, **33**, 7153.
130. Tomasselli, A.G., Hui., J.O., Sawyer, T.K., Staples, D.J., Borrow, C., Reardon, I.M., Have, W.J., DeCamp, D.L., Craik., C.S. and Heinrikson, R.L., *J. Biol. Chem.*, 1990, **265**, 14675.
131. Scharpe, S., De Meester, I., Hendriks, D., Vanhoof, G., van Sande, M. and Vriend, G., *Biochimie.*, 1991, **73**, 121.
132. Anderson, G.W., Zimmerman, J.E. and Callahan, F.M., *J. Am. Chem. Soc.*, 1967, **89**, 5012.
133. Chen, F.M.F., Lee, Y., Steinauer, R. and Benoiton, N.L., *Can. J. Chem.*, 1987, **65**, 613.
134. Oleksyszyn, J., Subotowska, L. and Mastalerz, P., *Synthesis*, 1979, 985.
135. Hawkins, P.C.D., Ph.D. Thesis, University of St. Andrews, 1993.
136. Dhawan, B. and Redmore, D., *Phosphorus and Sulfur*, 1987, **32**, 119.
137. Kowalik, J., Sanka-Dobrowolska, W. and Glowiak, T., *J. Chem. Soc.*, 1984, 446.
138. Schollkopf, U. and Schutze, R., *Liebigs Ann. Chem.*, 1987, 45.
139. Bartlett, P.A. and Jacobsen, N.E., *J. Am. Chem. Soc.*, 1981, **103**, 654.
140. Bartlett, P.A. and Johnson W.S., *Tetrahedron Lett.*, 1970, **46**, 4459.
141. Vaughan, W.R. and Baumann, J.B., *J. Org. Chem.*, 1962, **27**, 739.
142. Bartlett, P.A. and Marlowe, C.K., *Biochemistry*, 1983, **22**, 4618.
143. Gray, M.D. and Smith, D.J., *Tetrahedron Lett*, 1980, **21**, 859.
144. Bartlett, P.A., Hansen, J.E. and Giannousis, P.P., *J. Org. Chem.*, 1990, **55**, 6268.
145. McKenna, C.E., Higa, M.T., Cheung, N.H., and McKenna, M.E., *Tetrahedron Lett.*, 1977, 155; McKenna, C.E. and Schmidhauser, J.J., *J. Chem. Soc. Chem. Commun.*, 1979, 739.
146. Kafarski, P., Lejczak, B. and Szewczyk, J., *Can. J. Chem.*, 1983, **61**, 2425.
147. Komiyama, T., Suda, H., Aoyagi, T., Takeuchi, T., Umezawa, H., Fujimoto, K. and Umezawa, S., *Arch. Biochem. Biophys.*, 1975, **246**, 645.
148. Weaver, L.H., Kester, W.R. and Matthews, B.W., *J. Mol. Biol.*, 1977, **114**, 119.
149. Holmquist, B., *Biochemistry*, 1977, **16**, 4591.
150. Bolognesi, M.C. and Matthews, B.W., *J. Biol. Chem.*, 1979, **254**, 634.
151. Holmes, M.A. and Matthews, B.W., *Biochemistry*, 1981, **16**, 4591.
152. Griffiths, J.T., Phylip, L.H., Konvalinka, J., Strop, P., Gustchina, A., Wlodawer,

A., Davenport, R.J., Briggs, R., Dunn, B.M. and Kay, J., *Biochemistry*, 1992, **31**, 5193.

153. Dixon, M., *Biochem. J.*, 1953, **55**, 70.
154. Schotten, C., *Ber.*, 1884, **17**, 2544; Baumann, E., *Ber.*, 1886, **19**, 3218.
155. Nielsen, A.T. and Houlihan, W.J., *Org. Reactions*, 1968, **16**, 1.
156. Leonard, N.J. and Choudhury, D., *J. Am. Chem. Soc.*, 1957, **79**, 156.
157. Rovnyak, G.C. and Shu, V., *J. Org. Chem.*, 1979, **44**, 2518.
158. Lam. P., Eyermann, C.J., Jadhav, P.K., Hodge, C.N., Ru, Y., DeLucca, G.V., Bacheler, L.T., Meek, J.L., Otto, M.J., Rayner, M.M., Wong, N.Y., Chang, C.H., Weber, P.C., Jackson, D.A., Sharpe, T.R. and Erickson-Viitanen, S., presented at the IXth International Conference on AIDS, Berlin, June, 1993.
159. Ikemoto, N. and Schreiber, S.L., *J. Am. Chem. Soc.*, 1992, **114**, 2524.
160. Musich, J.A. and Rapoport, H., *J. Am. Chem. Soc.*, 1978, **100**, 4865.
161. Nahm, S. and Weinreb, S.M., *Tetrahedron Lett.*, 1981, **22**, 3815.
162. Seebach, D., Kalinowski, H.O., Langer, W., Crass, G. and Wilka, E-M., *Org. Synth.*, 1983, **61**, 24.
163. Kikkawa, I. and Yorifuji, T., *Synthesis*, 1980, 877.
164. Weiss, U. and Edwards, J.M., *Tetrahedron Lett.*, 1968, 4885; Oehldrich, J. and Cook, J.M., *Can. J. Chem.*, 1977, **55**, 82.
165. Eguchi, S., Ishiura, K., Noda, T. and Sasaki, T., *J. Org. Chem.*, 1987, **52**, 496.
166. Rovnyak, G.C., Shu, V. and Schwartz, J., *J. Heterocycl. Chem.*, 1981, **18**, 327.
167. Toda, F. and Tanaka, K., *Tetrahedron Lett.*, 1988, **29**, 551.
168. Still, W.C., Kahn, M. and Mitra, A., *J. Org. Chem.*, 1978, **43**, 2923.