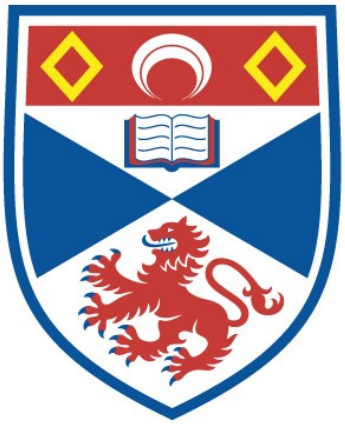
# AN INVESTIGATION INTO THE CATALYTIC MECHANISM OF THE ADENOVIRUS TYPE II PROTEINASE

Julie Anne Cornish

A Thesis Submitted for the Degree of Ph.D.
at the
University of St Andrews



1996

# An Investigation Into The Catalytic Mechanism Of The Adenovirus Type II Proteinase.

a thesis presented by

Julie Anne Cornish

to the

University of St Andrews

in application for

The degree of Doctor of Philosophy

St Andrews 1996



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

I, Julie Anne Cornish 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 fulfilment of any other degree or professional qualification.

signed Date 20/12/96

I was admitted to the Faculty of Science of the University of St Andrews under Ordinance General No. 12 on 01-10-92 and as a candidate for the degree of PhD. on 01-10-93.

signed Date 20/12/96

I herby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate to the degree of PhD.

signature of supervisor Date

## Copyright

# ACKNOWLEDGEMENTS

I would like to thank all the technical staff at St Andrews for their invaluable assistance over the 3 years, it has been greatly appreciated. I would also like to thank Dr Paul Hawkins, Dr Nick Camp, Dr Andy Cole, Dr Tony Maude and Dr John Wilkie for their friendship and many productive and sometimes heated discussions.

I am very grateful to Pigglet who created a wonderful lab. atmosphere to work in despite the fact that I'm from England! I am indebted to Kamal Badiani, Dr Basil Hartzoulakis, Phillipe Lasery, Donald McNab, Thomas Clifford, Dr Kim Haworth, Dr Richard Ireland, Dr Craig Anderson, John Morrison and Tracey Massil for their companionship and encouragement during my stay in St Andrews.

I would like to thank Dr John Walton, Dr John Wilkie and Dr Tony Maude for their check reading of my thesis, Dr Trevor Rutherford for his help with the protein NMR and Dr Heather Murray for her help with the proteinase assays.

I am eternally grateful to Dr Graham Kemp for his help with the proteinase assays and to my supervisor, Professor David Gani for his time and support during my 3 years. Lastly I would like to thank Andrew Taylor for his unquestioning support when times are hard, his wonderful company and his excellent sense of humor.

## THE MICROBE

The microbe is so very small
You cannot make him out at all,
but many sanguine people hope
To see him through a microscope.
His jointed tongue that lies beneath
A hundred curious rows of teeth;
His seven tufted tails with lots
of loverly pink and purple spots,
On each of which a pattern stands,
Composed of forty separate bands;
His eyebrows are of tender green.
All these have never yet been seen-
But scientists, who ought to know,
Assure us that they must be so...
Oh! Let us never, never doubt
What nobody is sure about!

*By Hilaire Belloc, reproduced with kind permission of the publishers, Random House, UK limited. Taken from complete Verse, Hilaire Belloc, Pimlico, London, p. 246.*

## Abstract

A series of $P_4$ (Cbz and t-Boc) N-protected potential substrates and inhibitors, containing the $P_1$ to $P_4$ substrate recognition sequence of the type 2 adenovirus proteinase (Leu-Ala-Gly-Gly) were prepared by solution phase peptide coupling techniques and tested for activity against the proteinase. The potential substrates contained the amide and ester moieties at the $P_1$ carbonyl position and the potential inhibitors contained the alcohol, acid, bromide, aldehyde, ketone, dimethylacetal, nitrile, alkenic, malonyl and epoxysuccinate moieties at the $P_1$ carbonyl position.

The esters, the t-Boc urethane and the *p*-nitroanilide moieties were substrates for the proteinase and the acid and the amides did not bind to the proteinase. Preliminary results show that the other inhibitors were mostly noncompetitive inhibitors for the adenovirus proteinase with approximate $K_i$'s between 15 and 200 $\mu mol\ dm^{-3}$.

The test results indicate that: the amides must contain a carbonyl group at $P_2$' to bind to the proteinase; the loss of the $P_1$' amine product is the rate limiting step for the hydrolysis of a substrate by the adenovirus proteinase; the P acid product leaves before the P' amine product, which is in complete contrast to classical cysteine proteinases such as papain; little protonation of the $P_1$'amide nitrogen or the $P_1$ carbonyl oxygen of the adenovirus proteinase-substrate complex occurs before the nucleophilic attack on the $P_1$ carbonyl carbon of the adenorirus proteinase-substrate complex.

# Abbreviations

| | |
|---|---|
| A | alanyl |
| Arg | arginyl |
| Asn | asparaginyl |
| Asp | aspartyl |
| B | asparaginyl or aspartyl |
| C | cysteinyl |
| Cbz | carbobenzyloxy |
| CE | capillary electrophoresis |
| $^{13}$C-NMR | carbon-13 nuclear magnetic resonance |
| Cys | cysteinyl |
| D | aspartyl |
| dCMP | deoxyribosyl cytosyl monophosphate residue |
| DMF | dimethyl formamide |
| DMSO | dimethyl sulphoxide |
| DNA | deoxyribosyl nucleic acid |
| DPF | diisopropyl fluorophosphate |
| DTT | dithiothreitol |
| E | glutamyl |
| F | phenylalanyl |
| G | glycyl |
| Gln | glutaminyl |
| H | histidinyl |
| His | histidinyl |
| $^{1}$H-NMR | proton nuclear magnetic resonance |

| | |
|---|---|
| I | isoleucyl |
| IBCF | *iso*-butyl chloroformate |
| IF | initiation factor |
| $iPr_2EtN$ | diisopropyl ethylamine |
| K | lysinyl |
| $K_M$ | Michaelis constant |
| L | leucyl |
| M | methionyl |
| MeOH | methanol |
| mRMA | messenger ribonucleic acid |
| N | asparaginyl |
| NEM | N-ethyl maleimide |
| NMM | N-methyl morpholine |
| NMM.HCl | N-methyl morpholine hydrochloride |
| P | prolinyl |
| PAH | phenylalanine hydroxylase |
| PCMP | parachloromercury benzoate |
| PK | protein kinase |
| PKU | phenylketonuria |
| PMSF | phenyl sulphonium methyl fluoride |
| POL | viral polymerase |
| pTP | preterminal protein |
| Q | glutaminyl |
| R | arginyl |
| RNA | ribosyl nucleic acid |
| S | serinyl |

| | |
|---|---|
| SBTI | soya bean trypsin inhibitor |
| Ser | serinyl |
| STEM | scanning transmission electron microscope |
| T | threonyl |
| t-Boc | *tertiary*-butoxycarbonyl |
| TBTU | O-benzotriazole-1-yl-N, N, N', N'-tetramethyluronium tetrafluoroborate |
| TFA | trifluoroacetyl |
| THF | tetrahydrofuran |
| V | valinyl |
| $V_{max}$ | maximal rate |
| W | tryptophanyl |
| X | any aminoacid residue |
| Y | tyrosinyl |
| Z | glutaminyl or glutamyl |

# Contents

## Introduction

## Results and Discussion



**Synthesis**

**Discussion**

## Experimental

# 1.0 The Adenovirus

## 1.0.1 Infection

In 1953 Rowe isolated a new cytopathogenic agent from the tissue culture fluid of spontaneously degenerating human adenoidial tissue.[1] The agent could infect other healthy cells and specific antibodies could be raised from rabbits to neutralize the agents activity, although the agent could not be grown on bacteriological media. Rowe concluded that the adenoidal degradation product was either a virus or a rickettsaie (parasitic microorganism, intermediate in structure between bacteria and viruses). Since then the cytopathogenic agent has been identified as a virus and 47 different human serotypes have been discovered. Serotypes 1-5 and 7 are endemic to the western world where they are a major factor in acute respiratory disease (serotypes 4 and 7),[2] pharyngoconjunctival fever (serotypes 3, 7, 14),[3] keratoconjunctivitus (serotypes 3, 7 and 8)[4] and gastroenteritis (serotypes 40 and 41).[3]

Adenovirus infections of immunocompromised patients are overwhelming[5] and are usually fatal. In these cases evidence of adenovirus has been found in the lungs, liver,[5, 6] kidneys[6] and the gastric tract. Indeed Rowe showed that human gamma-globulin could prevent the cytopathic effect of the adenovirus when added to infected cells.[1]

Childhood infections with adenovirus 12 give an enhanced risk of coeliac disease and dermatitis herpetiformis[7], while adenovirus 1 and 2 have been linked with infant intussusception of the bowel.[8] Adenovirus may also be one cause of persistent inflammatory arthritis[9] and childhood acute myocarditis.[10] One Adenovirus group (serotypes 12, 18 and 31) is highly oncogenic in animals[3] and there is some evidence that adenoviral DNA has been found in neurogenic tumors.[11]

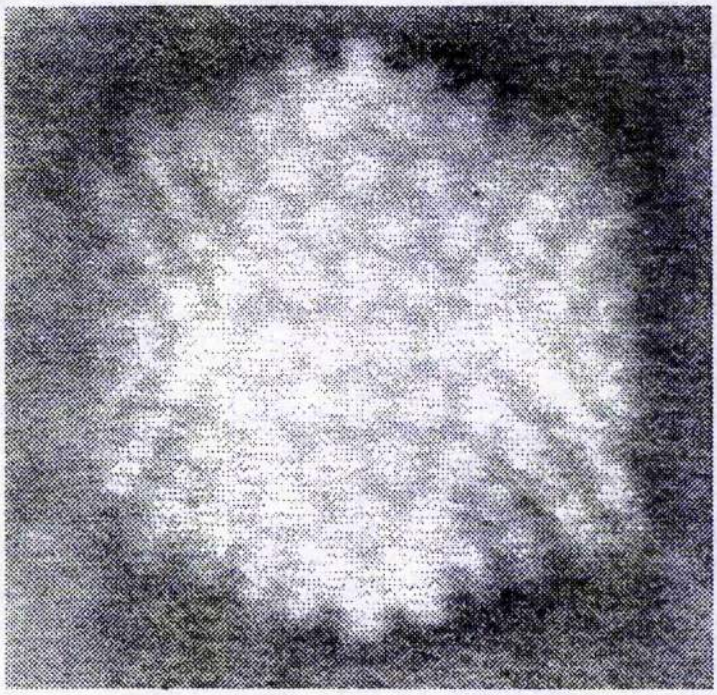
## 1.0.2 Structure

Adenovirus particles are non-enveloped, regular icosahedra of 1500Å diameter with a mass of 150 MDa. The virus particle contains a linear double stranded DNA genome and is built up from eleven different structural proteins.[3] The protein coat (capsid) consists of two hundred and forty hexons and twelve pentons. A penton is situated at each vertex, surrounded by five hexons. A fibre projects from each penton and its length depends on the serotype of the virus[3] (Fig. 1.1 a and b on page 3).[12]

Each hexon is a trimer of three identical polypeptide chains. The base of each monomer consists of a pair of antiparallel eight-stranded β-barrels, which lock together in a pseudo-hexagonal ring in the trimer. The hexons form a close-packed hexagonal array on the virus surface.[13] Protein IX links nine hexons in the centre of each face of the virus particle, while proteins VI and VIII bridge hexons between adjacent faces. The structure of the adenovirus 2 is shown in Fig. 1.2 on page 4.[13] The positions of the hexon (II), penton base (III), fibre (IV) and polypeptide IX are certain but the positions of polypeptides IIa, VI and VIII are tentatively assigned based on a 3-dimensional image construction of the intact virus.[13] The 3-dimensional image construction was based on the low resolution electron crystallographic structure of the hexon and multiple images from an image-enhanced scanning transmission electron microscope (STEM) of the virus in different orientations.[13]

The virus core is composed of the double stranded DNA genome and four different polypeptides. One of these polypeptides (the terminal protein) caps each end of the double stranded DNA. The protein-DNA links occur through serine hydroxyl groups and deoxyribose cytosyl monophosphate residues (Fig. 1.3 on page 5). This is called the pTP-dCMP complex. The other three polypeptides are involved in the positioning and packing of the DNA during viral particle assembly.[3]

a)



b)

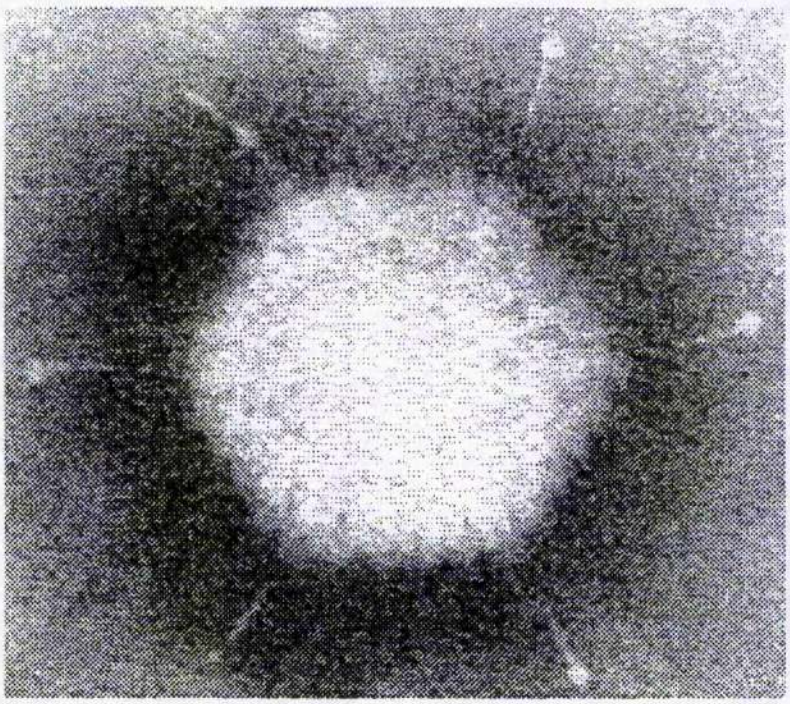


Fig.1.1: *Electronmicrographs of an adenovirus,*

*a) to show the icosohedral array and b) to show the fibres.*[a]

[a] reproduced with kind permission of Academic Press Ltd, London, England.

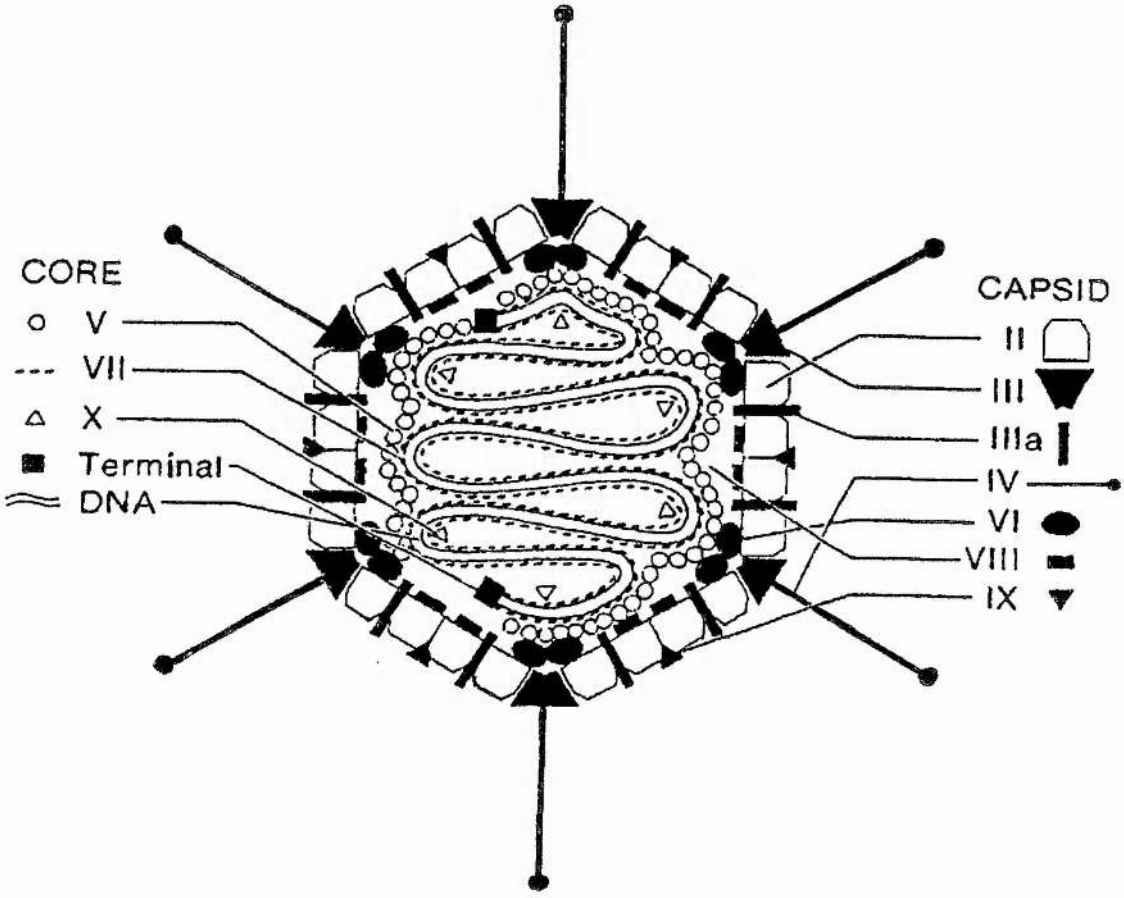



Fig. 1.2: *Schematic representation of an adenovirus.*

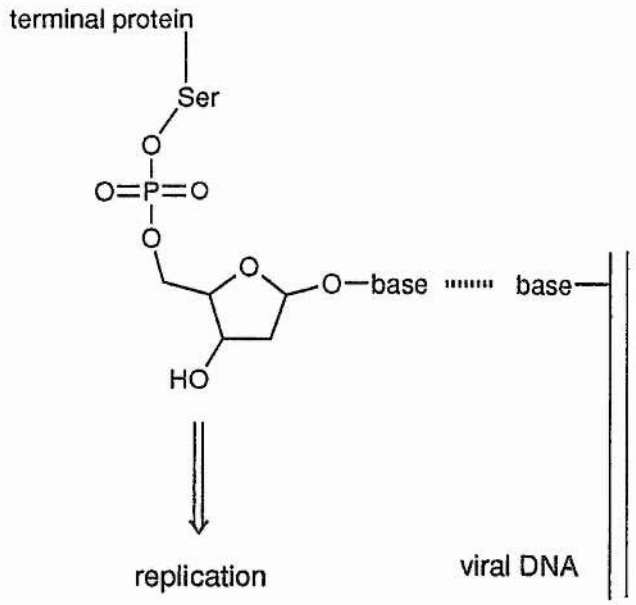



Fig. 1.3: *Positioning of the terminal protein on the viral DNA.*

## 1.0.3 Life cycle

The life cycle of the virus is divided into 2 phases; early phase (E), when virus penetration, transcription and translation occurs and late phase (L), during which duplication, and new virus assembly occurs.[3]

The adenovirus enters a cell through receptor mediated endocytosis.[14] Initially the head of the fibre attaches to a specific receptor on the cell's surface[15] and the vesicles enclose the virus.[16] Here the virus disrupts the membrane of the endocytic vesicles so that when the capsid is ruptured the viral contents are released into the cytoplasm.[16] This process is aided by the penton base,[17] which may bind to $\alpha_v$ integrin receptors,[18] and the acidic pH[19] of the endocytic vesicle.

The pentons are lost at this stage[20] and the virus is transported across the cytoplasm using the hexons which bind to the cell microtubules.[21] When the virus reaches the nucleus, viral DNA enters, leaving many of the viral proteins in the cytoplasm.[20]

Each region of the viral genome codes for a specific polypeptide, which is required for virus formation (Fig.1.4). The arrow in Fig. 1.4 indicates transcription direction and the functions of the genome regions are listed in Table 1.1.

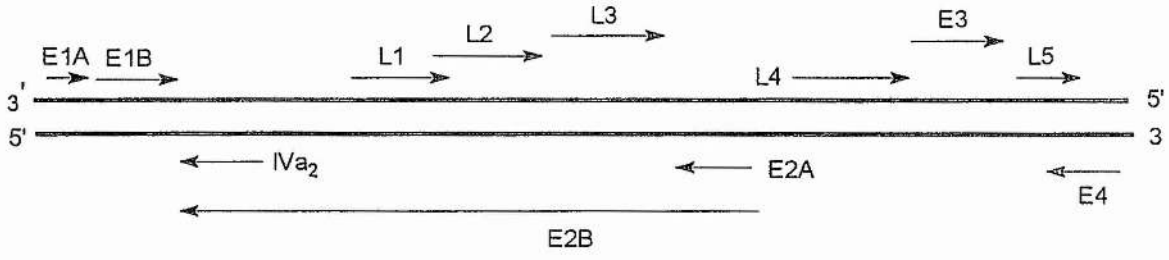



Fig 1.4: *The adenovirus type 2 genome.*

| **Region of Genome** | **mRNA codes for;** |
|---|---|
| E1A and E1B | proteins that control transformation function |
| E2A | DNA binding protein |
| E2B | DNA polymerase and terminal protein |
| E3 | protein that suppress the immune system and prevents host cell growth[22] |
| E4 | Viral assembly proteins |
| $IV_a$ | transcriptional activator[23] |
| L1 and L2 | penton proteins |
| L3 | hexon protein and a 23kDa proteinase[24] |
| L4 | hexon-stabilizing protein |
| L5 | fibre proteins |

Table 1.1: *The function of various regions of the adenovirus genome.*

Regulation of viral transcription (the process whereby mRNA is synthesized from a DNA template) is complex, and not fully resolved. However it is postulated that the E1A region is transcribed first, and that some of this mRNA codes for a protein that enables the other early regions, (given the prefix E in Fig. 1.3) to be transcribed. It is thought that a host transcription repressor for the other early regions exists, and the E1A protein either binds to the repressor or inhibits its synthesis, allowing transcription of the other early regions.[3]

The mRNA then leaves the nucleus through a nuclear pore and enters the cytoplasm where it is translated into polyprotein by host enzymes. Cellular protein synthesis is regulated by an initiation factor (IF), which must reach the nucleus before transcription can begin. When the mRNA reaches the cytoplasm translation occurs and new protein is produced.

Control occurs *via* a protein kinase (PK), which phosphorylates the IF, making it inactive. The PK is activated by the mRNA used to synthesize the protein. In this way genes can be turned on and off. The adenovirus inhibits PK, so that the mRNA that codes for a viral DNA polymerase (POL) can be made. Duplication can then commence. Viral DNA is a competitive inhibitor of PK, so IF is not inactivated and viral protein production can continue.[25]

Duplication of the viral DNA is also complex, and not fully resolved. The 3'-OH of the pTP-dCMP complex acts as a primer for the new strand of DNA, using one strand of viral DNA as the template.[26] Then POL binds to the pTP-dCMP complex, aided by a viral single-strand specific DNA-binding protein. The pTP-dCMP complex elongation is catalysed by POL and requires three host proteins; nuclear factor I, nuclear factor III and a topoisomerase (nuclear factor II).[27]

Meanwhile, in the cytoplasm hexons and pentons are assembled from their constituent monomers. A scaffold protein ensures correct association of the hexon monomers. The

capsid is produced from the hexons, pentons and other structural proteins, while the core polypeptides and the viral DNA are injected into the capsid. As the new virus matures, the proteins; pVI, pVII, 11kDa,[28] pTP,[29] pVIII, IIIa[30] and pMu[31] are cleaved by a 23kDa protease, encoded by the L3 region of the genome.[24]

The newly formed infectious viruses are then released by cell lysis. The virus has two strategies for inducing this; the first is to prevent cell growth, as the viral E3 protein removes epidermal growth factor receptors from the surface of the host cell,[22] preventing the host cell from receiving any growth factor. The second strategy, is to cleave the protein fibre network, that gives the cell structural integrity, with the 23kDa protease.[32]

## 1.0.4 Therapeutic uses

Recently there has been considerable interest in gene therapy to treat genetic disorders[33] such as cystic fibrosis[34] and cancer.[35] This involves the transfer and expression of DNA into affected cells by means of a vector, a task for which adenovirus may be suitable.

Replication-deficient adenovirus vectors have been used to transfer the *E. coli β-galactosidase* gene to rat brain cells both *in vivo*[36] and *in vitro*,[34] rat primary hepatocytes *in* vitro[35] and into fetal rabbits *in utero*.[37] This gene has also been transferred to the airway epithelial cells of sheep *in vivo*.[38]

Human pancreatic lipase has also been transferred to sheep's gall bladders *in vivo*[33] and human phenylalanine hydroxylase (PAH) gene has been transferred into PAH-deficient mice *in vivo* using adenovirus replication-deficient vectors. A lack of PAH causes phenylketouria (PKU) which produces severe mental retardation in children.[34]

Adenovirus vectors are particularly suitable as they can infect postmitotic cells[39] and do not incorporate DNA into the host genome, minimizing the risk of mutagenesis.[40] A clinical

trial is currently in progress to determine whether adenovirus can transfer the normal gene to the nasal epithelium of cystic fibrosis suffers.[40]

The E1A product of the adenovirus type 5 can act as a tumor suppressor gene for the *neu* gene in rats both *in vivo* and *in vitro*.[41] Over-expression of the human *neu* gene can result in many types of cancer *eg.* breast,[42] ovary,[43] lungs,[44] stomach[45] and colon.[46] The E1A region of the adenovirus may provide a useful tool for the development of therapeutic agents against these cancers.

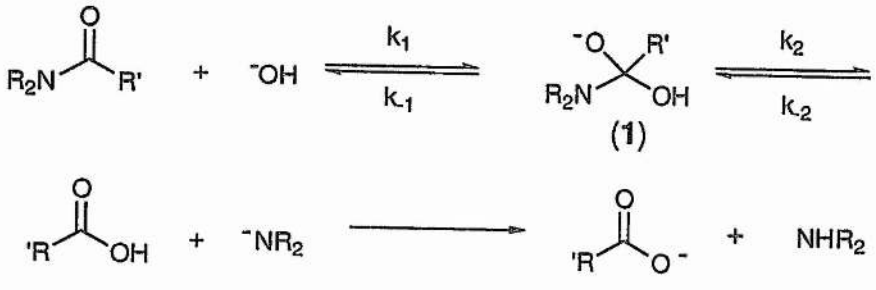
# 1.1 Solution mechanisms for amide bond hydrolysis

In solution there are three possible mechanisms to hydrolyse an amide bond; a) base catalysed, b) acid catalysed and c) strong acid catalysed.

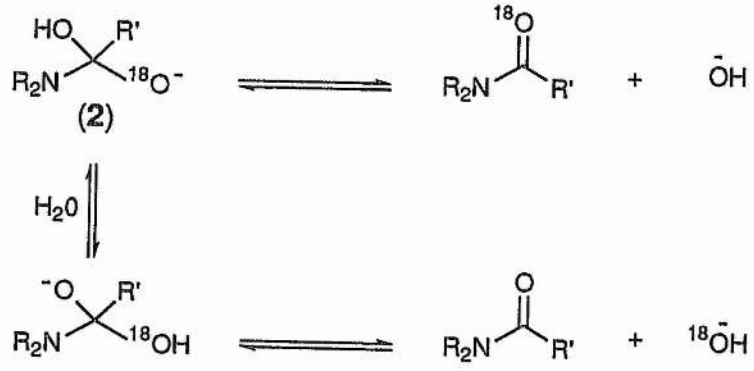
## 1.1.0 Base catalysed hydrolysis of an amide bond

In the base catalysed hydrolysis mechanism (Scheme 1.1) the nucleophilicity of the nucleophile (water in the uncatalysed hydrolysis) is increased by changing the nucleophile for a hydroxide ion, so that the rate of hydrolysis is speeded up. The use of $^{18}$O-labelled amides has shown that, during the base catalysed hydrolysis of primary and secondary amides, oxygen exchange in the starting material is rapid.[47] This fits with the observation that the amide anion is a poorer leaving group than the hydroxide ion,[48] so it is more likely that the hydroxide ion will leave before the amide anion in the tetrahedral intermediate (**1**) (Scheme 1.1). Therefore $k_{-1}$ is greater than $k_2$ and $k_2$ must be the rate determining step of the base catalysed hydrolysis.[47] In the case of an $^{18}$O-labelled amide, either the unlabelled attacking hydroxide, or labelled hydroxide, produced from proton exchange between the tetrahedral intermediate (**2**) and the solvent (Scheme 1.2) could leave before the amide anion. If the

labelled hydroxide leaves before the amide anion, the carbonyl oxygen of the amide starting material loses some of the $^{18}O$-label. Tertiary amides show no similar oxygen exchange and this has led to the suggestion that oxygen equilibration is actually mediated by a proton from the nitrogen (Fig. 1.5).[47]




Scheme 1.1: *The based catalysed hydrolysis of an amide bond.*




Scheme 1.2: *Hydroxide exchange during the based catalysed hydrolysis of an amide bond.*

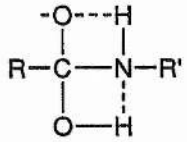
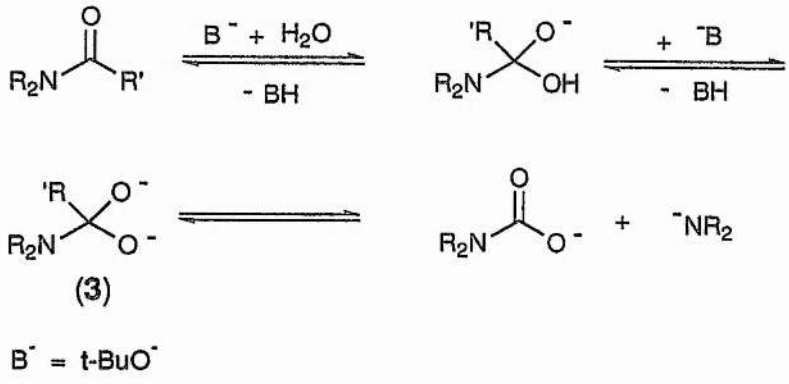


Fig. 1.5:

*The mechanism of oxygen equilibration in primary and secondary amide hydrolysis.*

As $k_{-1}$ is much greater than $k_2$ the base catalysed hydrolysis of amides in aqueous solution is not a very efficient process. However in non aqueous conditions *i.e.* using a small excess of water, excess potasium tertiary-butoxide as a base and ether as the solvent (Scheme 1.3), amide hydrolysis is rapid and the tetrahedral intermediate (**3**) (Scheme 1.3) formed is now a dianion which expels the negatively charged nitrogen rapidly.[49]

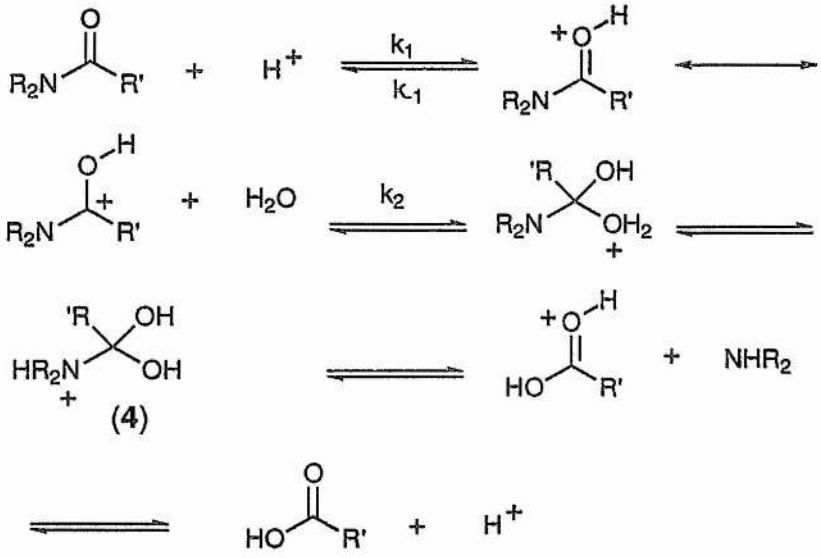


Scheme 1.3: *The non-aqueous base catalysed hydrolysis of an amide bond.*

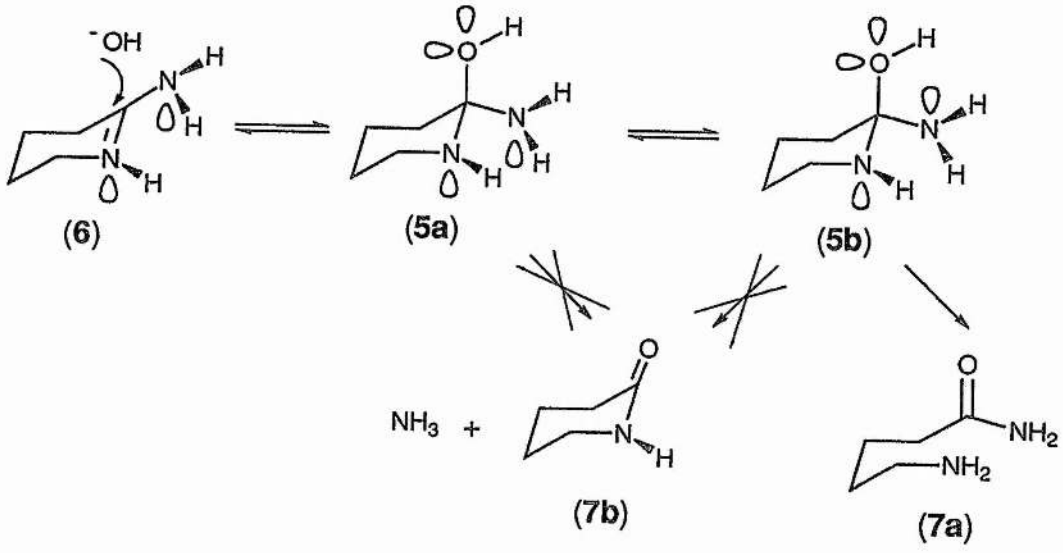
### 1.1.1 Acid catalysed hydrolysis of an amide bond

Acid catalysed hydrolysis of an amide bond is more complex as it is possible to polarise the carbon-oxygen bond by protonation of the carbonyl oxygen, or to polarise the carbon-nitrogen bond by protonation of the amide nitrogen. Both of these methods create a greater positive charge on the carbonyl carbon, compared to the unprotonated amide and enhance the electrophilicity of the carbonyl carbon. Oxygen protonation would result in an addition type hydrolysis mechanism, whereas nitrogen protonation would lead to as $S_N1$ type hydrolysis mechanism. In solution $^1$H-NMR spectroscopy has shown that oxygen protonation predominates in 100% acid solution. The methyl protons of dimethylformamide in sulphuric acid still appear as a doublet of singlets (the *cis* and *trans* methyl group are not equivalent as rotation of the carbon-nitrogen bond is restricted by the conjugation of the nitrogen lone pair and the carbonyl bond) so the carbon-nitrogen bond rotation is still restricted. If the amide nitrogen was protonated, the lone pair of the nitrogen would not be available for delocalisation with the carbonyl group and carbon nitrogen bond rotation would not be restricted so a singlet would be observed for the methyl signals of dimethylformamide.[50] The proton signal for a proton attached to the oxygen of an amide has also been observed at -90°C for dimethylformamide in fluorosulphuric acid, (10.40 ppm, referenced to trimethyl silane).[51] Acid hydrolysis usually occurs by the addition mechanism (Scheme 1.5).[52] No oxygen exchange occurs, unlike the basic hydrolysis mechanism as $k_2$ is greater than $k_{-1}$. The breakdown of the protonated nitrogen tetrahedral intermediate (**4**) (Scheme 1.4) should be fastest when one of the electron pairs of each oxygen is antiperiplanar to the lone pair in the amine leaving group.[53] The α-diamino alcohol (**5**) is an analogue of the tetrahedral transition state (**4**) and breaks down to give the primary amine (**7a**) and not the lactam (**7b**) (Scheme 1.5 on page 13).[54] The alcohol (**5**) is formed in conformation (**5a**) from the axial attack of

hydroxide at carbon-2 of the amidine (**6**). As both inversion of the ring and the secondary amine are slow compared to the breakdown of the alcohol (**5**), the secondary amine lone pair is unlikely to become antiperiplanar to the primary amine lone pair, so the lactam (**7b**) is not formed. By contrast both the primary amine and the alcohol groups rotate freely, allowing the alcohol (**5**) to adopt conformation (**5b**), where one of the alcohol lone pairs and the secondary amine lone pair are antiperiplanar to the primary amine lone pair so the primary amine (**7a**) is formed.



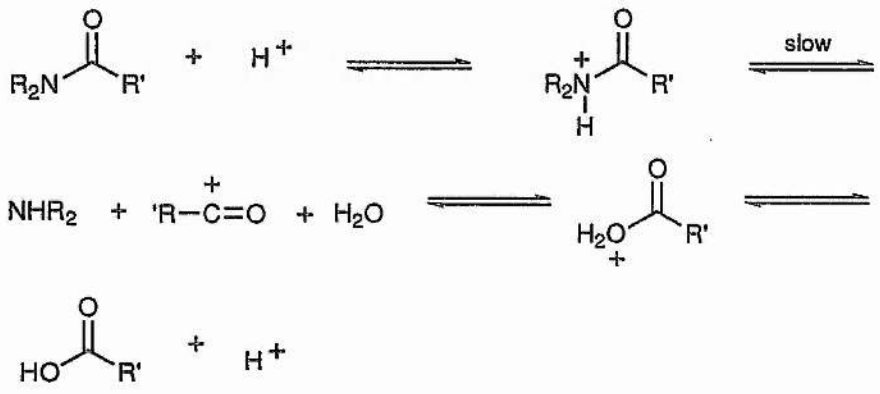

Scheme 1.4: *The acidic hydrolysis mechanism of an amide bond.*




Scheme 1.5: *The acid catalysed hydrolysis of a cyclic α-diamino alcohol (5).*

### 1.1.2 Strong acid catalysed hydrolysis of an amide bond

In above 80% acid solution, 4-nitroacetanilide,[55] 4-nitrobenzamide, benzamide and 4-methoxybenzamide[56] follow the $S_N1$ type mechanism (Scheme 1.6 on page 14). The amide nitrogen becomes protonated and very slowly breaks down to give a carbocation and the free amine. The carbocation is rapidly attacked by water to give the carboxylic acid, after deprotonation. This mechanism is probably observed for the benzamides, as the carbocation can be stabilised by the conjugating aromatic ring, which partially delocalises the positive charge, associated with the carbonyl carbon.




Scheme 1.6 : *The strong acid catalysed hydrolysis mechanism for an amide bond.*

## 1.2 The Mechanism and inhibition of cysteine proteinases

### 1.2.1 Classical cysteine proteinases

There are 3 classes of cysteine proteinases;[57] a) plant b) animal and c) bacterial. Plant tissue cysteine proteinases are the most studied as the enzymes are readily available from fruit. Papain is found in *Carica papaya* (tropical papaya fruit), actinidin is found in *Actinida chinensis* (kiwifruit) and bromelain is found in *Anaus comosus* (pineapple). These enzymes

have broad specificities and attack a wide range of substrates. Their function is probably to protect the ripening fruit from attack by insects, fungi *etc.*[58]

Animal tissue cysteine proteinases are found mostly in liver, spleen and kidney cell lysosomes. They are usually called cathepsins, although not all cathepsins are cysteine proteinases. These enzymes are involved in intracellular protein degradation and are named after the specific proteins they cleave. Cathepsin N acts on the N-terminal peptides of collagen while cathepsin P is thought to convert proinsulin to insulin.[57]

Bacterial cysteine proteinases such as the streptococcal proteinase isolated from hemolytic *streptococci,*[59] and clostripain, isolated from *Clostridium histolyticum,*[60] have very broad specificity and are most likely digestive enzymes.

Both plant and animal cysteine proteinases have a characteristic 2-domain structure as shown for papain in Fig.1.6a.[61] The sequence homology between the two groups of proteinase is poor (less than 50%), however there are only small differences in the polypeptide fold that contains the proteinase's active site.[57] The poor homology comes from the deletions and perturbations that give rise to a particular proteinase's substrate specificity.

The active site of papain (Fig. 1.6b) is in a deep extended cleft, between two domains. The side chains of a cysteine (Cys-25 in papain) and a histidine (His-159 in papain) are closely associated. In papain, the imidazole ring of His-159 is hydrogen bonded to the side chain of Asn-175 and can rotate by 30° about $C_\beta$-$C_\gamma$ in response to the changes in the protonation state of the $S_\gamma$ in Cys-25. This rotation is probably very important during catalysis.[57]

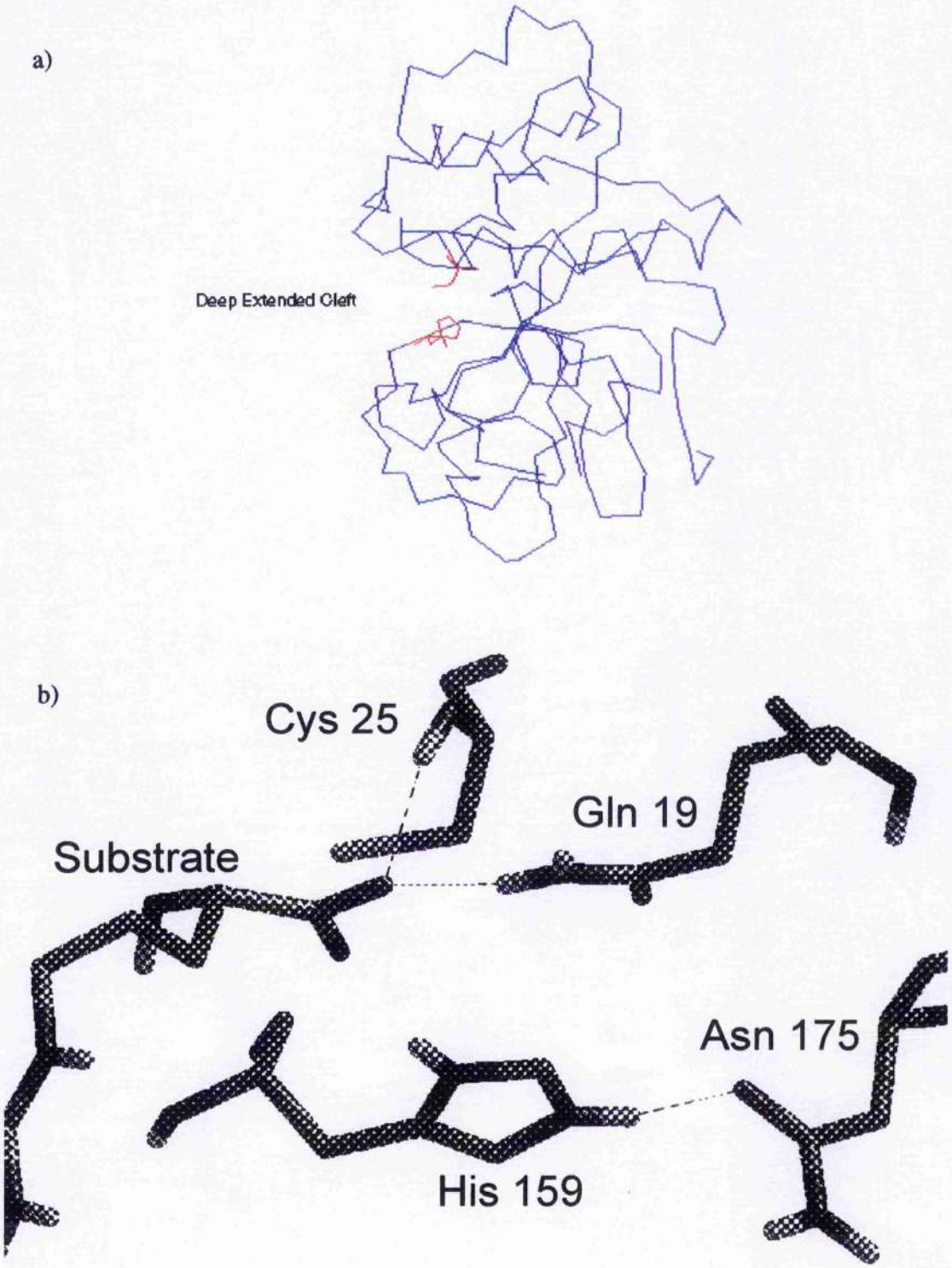



Fig 1.6: The structure of papain,[61] *a) the entire protein, b) the active site.*

In the thiol proteinases the essential nucleophile is provided by a cysteine residue. A thiol group is a poor nucleophile and a thiolate nucleophile is a good nucleophile. However the normal pKa of a thiol group is 9, so there should be little formation of a thiolate nucleophile at the pH of maximal activity (pH 4-7). However, various spectroscopic techniques (fluorescence,[62] UV[63] and NMR[64]) have indicated that the Cys-His couple is present as an ion pair in the pH range of the proteinase's activity. Three different states of the thiol-imidazole pair are possible depending on the pH (Fig. 1.7).

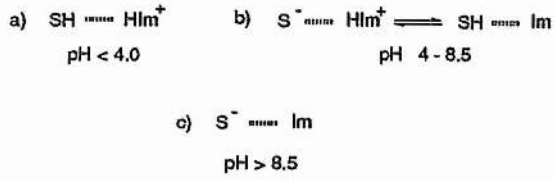



Fig. 1.7: *The possible protonation states of the thiol-imidazole pair: a) Protonated; b) Neutral and c) Deprotonated.*

At less than pH 4, the protonated form of the Cys-His pair (Fig. 1.7a) of papain predominates and papain is inactive. This suggests that the nucleophilicity of the thiol group is too low for proteolytic hydrolysis to occur. Above pH 8.5, the deprotonated form of the Cys-His pair of papain (Fig. 1.7c) predominates and papain is also inactive. The thiolate nucleophile by itself is unable to catalyse the hydrolysis of a substrate, suggesting that the imidazole may be present as an imidazolium ion in the active form of papain, as proton donation to the substrate would also be possible from the imidazolium ion. At the medium pH range, the Cys-His ion pair predominates (Fig. 1.7b) and papain is active. Both a thiolate nucleophile and a imidazolium, proton donor are required for proteolytic hydrolysis. Plots of the rate of acylation as functions of pH are bell shaped with a maxima at pH 6 and two pK

values at 4.0 and 8.5, corresponding to the formation and decomposition of the Cys-His ion pair, respectively.[57]

It is surprising that the thiolate-imidazolium ion pair exists at the middle pH range as the imidazole base (pKa of the conjugate acid 6.1) is usually considered as being too weak to deprotonate at thiol group (pKa 9). Energy calculations on model systems, suggested that the presence of the imidazole group of His-159 adjacent to Cys-25 increases the probability of ion pair formation.[65] More significantly, Cys-25 is situated at the end of the N-terminus of a long α-helix, where a slight positive charge exists at the N-terminus and a slight negative charge exists at the carboxy end terminus. Cys-25 is in a positively charged environment which favours the transfer of a proton from the thiol group of Cys-25 to the imidazole ring of His-159.[66] Molecular orbital calculations using models for Cys-25, His-159, Asn-175 and the α-helix also confirm this. [65, 67] When the contributions of all the surface groups to the overall electrostatic field of papain were found, the effect due to the α-helix accounted for half of the positive environment and the other half was due to Asp-158. The local electric field surrounding the Cys and the His residues in each particular proteinase would govern the exact protonation state of the Cys-His pair at any given pH and account for the differences in the pH range of the classical thiol proteinases.

The role of Asp-158 (papain) has been a source of controversy, in early studies it was assumed that a carboxyl group was responsible for the low pKa of 4, observed in pH activity studies.[68] More recently this pKa has been shown to be due to the thiolate imidazolium pair (see previous).

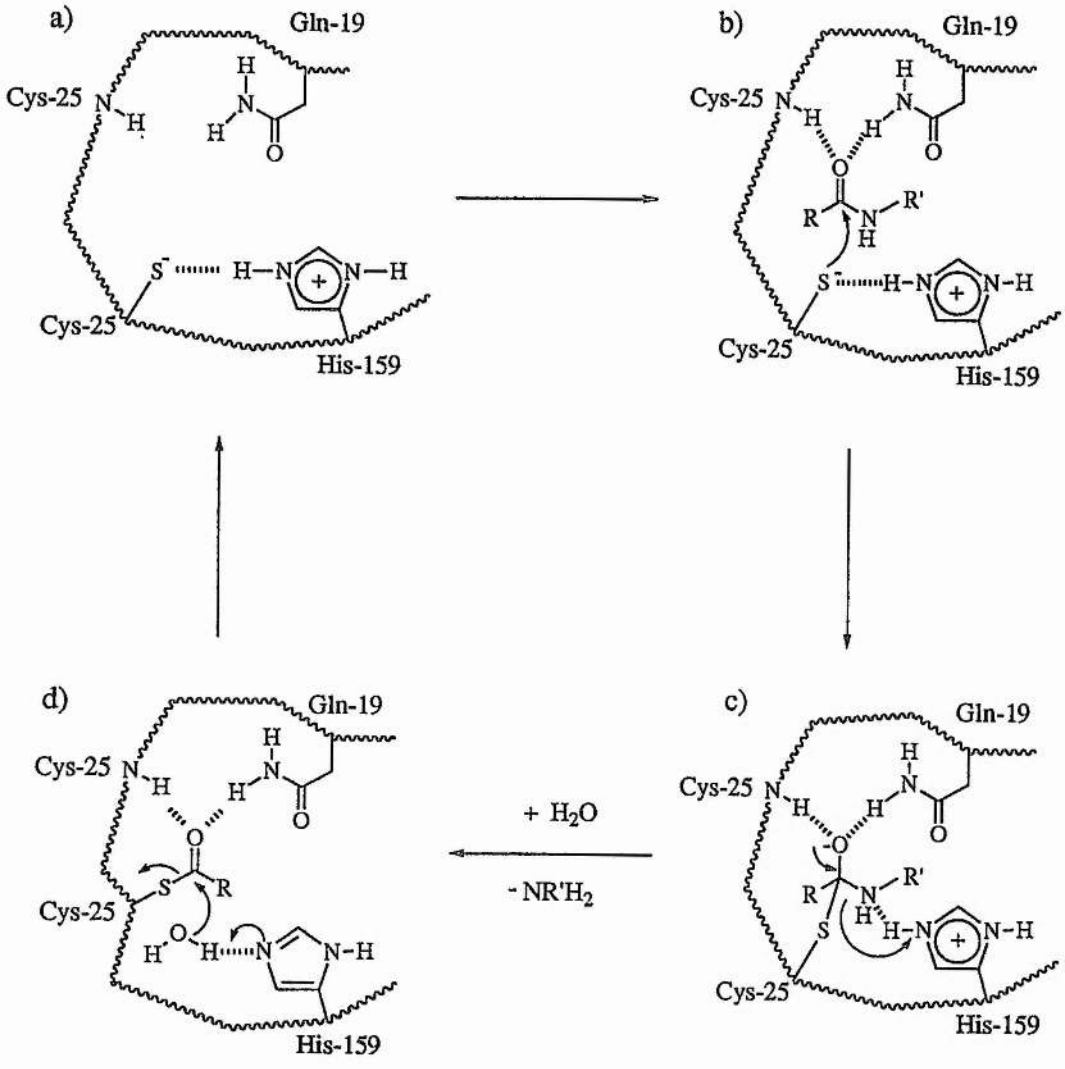
The crystal structures of both papain and actinidin show that the carboxyl group of Asp-158 (Asp-161 in actinidin) is turned away from the active site and forms a hydrogen bond to the main chain amido proton of residue 136 (139 in actinidin).[69] The Asp residue is

far too distant from the catalytic site (7.7Å away from Cys-25 and 5.5Å away from His-159 in papain) to be involved directly in catalysis. The crystal structures also show precisely the same arrangenent of the Asp and His residues, even though the crystals were grown at pH 9.3 and 8, respectively. If an Asp-His ion pair also formed part of the catalytic apparatus of the proteinase it should have been present at pH 6.0.

Actinidin and ficin have an Asp residue in a sequentially equivalent position to Asp-158 in papain and the corresponding residue in bromelain and cathepsin H is an Asn residue. However in cathepsin B, the corresponding residue to Asp-158 in papain is a Gly residue. There are no neighbouring substituents to compensate for such a change, so unless the mechanisms of these proteinases are different, which is unlikely, as they have broad structural and kinetic similarity, the direct involvment of a carboxyl group in catalysis can be ruled out and the Cys-His ion pair is the essential catalytic apparatus.[57]

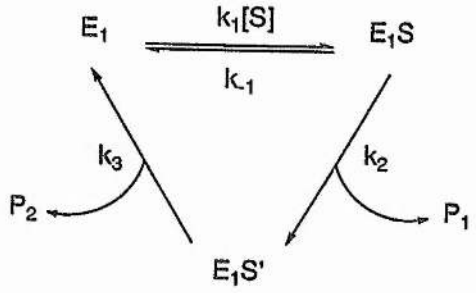
These findings lead to the proposed catalytic mechanism[57] of papain as given in Scheme 1.7. The formation of the enzyme substrate complex is followed by attack of the thiolate group of Cys-25 on the carbon of the $P_1$ carbonyl group (Scheme 1.7, structure b). The resultant negatively charged intermediate is stabilised by two hydrogen bonds; one between the oxyanion and the Cys-25 amido hydrogen and the other between the oxyanion and the side chain amido group of Gln-19 (Scheme 1.7, structure c).

The amide bond is cleaved by His-159 donating a proton to the $P_1$' amide nitrogen which now forms the leaving group. (Scheme 1.7, structure c) The free $P_1$' amine diffuses away from the active site and the $P_1$ remains attached to the proteinase as a thioester. The proteinase's Cys-His couple is reprotonated by nucleophilic attack of a water molecule on the thio-ester, releasing the $P_1$ acid from the enzyme's active site (Scheme 1.7, structure d).




Scheme 1.7: The catalytic mechanism of *papain.*

## 1.2.2 Substrate / enzyme kinetics




Scheme 1.8: The vector diagram of *papain's double intermediate mechanism.*

The vector diagram of the double intermediate mechanism of papain is given in Scheme 1.8, where $E_1$ is the free enzyme, $E_1S$ is the substrate enzyme complex, $P_1$ is the free amine product, $E_1S'$ is the enzyme product two complex and $P_2$ is the acid product. An expression for the reaction rate (v) can be found from using the King and Altman method[70] as follows;

Each enzyme species $[EX_i]$, i = (1 - n), can be written as shown in Table 1.2, where;

$[EX_i]$ = sum of the terms in the vector diagrams of the species $[EX_i]$ 1.2.1

| Enzyme species | Kinetic terms |
|---|---|
| $[E_1]$ | $k_2k_3 + k_{-1}k_3$ |
| $[E_1S]$ | $k_1k_3[S]$ |
| $[E_1S']$ | $k_1k_2[S]$ |

Table 1.2: *The kinetic terms of the enzyme species involved in the double intermediate mechanism.*

The rate is given by equation 1.2.2;

$$v = \frac{d[P_1]}{dt} = k_2[E_1S] \qquad 1.2.2$$

Substituting the value of $[E_1S]$ from Table 1.2 into 1.2.2 gives equation 1.2.3;

$$v = k_1k_2k_3[S] \qquad 1.2.3$$

The total enzyme concentration ($E_0$) is by equation 1.2.4;

$$[E_o] = [E_1] + [E_1S] + [E_1S'] \qquad 1.2.4$$

Dividing through equation 1.2.3 by $E_0$ on top and the expression for $E_0$ in 1.2.4 on the bottom results in equation 1.2.5 and factorising equation 1.2.5 gives equations 1.2.6 and 1.2.7;

$$v = \frac{k_1k_2k_3[S][E_o]}{k_2k_3 + k_{-1}k_3 + k_1k_3[S] + k_1k_2[S]} \qquad 1.2.5$$

$$v = \frac{k_1k_2k_3[S][E_o]}{k_3(k_2 + k_{-1}) + k_1[S](k_3 + k_2)} \qquad 1.2.6$$

$$v = \frac{k_2k_3(k_3 + k_2)^{-1}[S][E_o]}{k_3(k_2 + k_{-1})(k_1(k_3 + k_2))^{-1} + [S]} \qquad 1.2.7$$

The Michaelis Menten equation is defined in equation 1.2.8, where $V_{max}$ is the maximal velocity and $K_m$ is the combined equilibrium constant for an enzyme catalysed reaction (the Michaelis constant);

$$v = \frac{V_{max}[S]}{K_m + [S]} \qquad 1.2.8$$

Comparing equation 1.2.7 and 1.2.8 gives $K_m$ and $V_{max}$ for the double intermediate mechanism (equation 1.2.9).

$$V_{max} = \frac{k_2k_3[E_o]}{k_3 + k_2} \qquad K_m = \frac{k_3(k_2 + k_{-1})}{k_1(k_3 + k_2)} \qquad 1.2.9$$

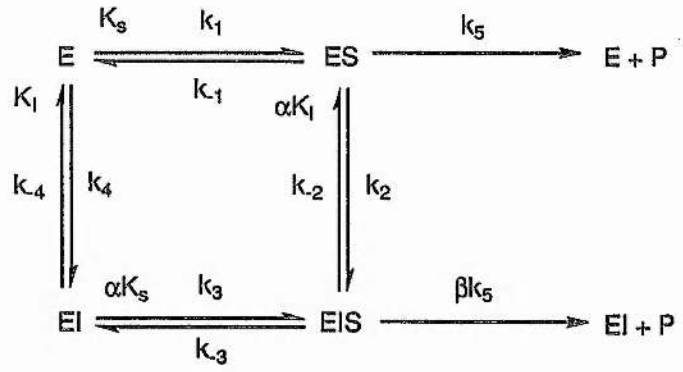
## 1.2.3 Enzyme Inhibition

In general terms, a proteinase inhibitor is a reagent that when added to a mixture of the proteinase and substrate, causes a decrease in rate of substrate cleavage. (Reagents such as guanidinium chloride, succinic anhydride, organic solvents also decrease the rate of activity as they cause perturbations in the enzyme's protein structure. These are usually considered as non-specific inactivators rather than inhibitors.)[71]

If proteolysis can be restored by dilution of the inhibitor solution or by gel filtration of the inhibitor, substrate and proteinase mixture, the inhibition is defined as reversible. If dilution fails ot restore proteolysis, the inhibition is defined as irreversible. On rare occasions some reversible inhibitors bind to their proteinase with such high affinities, that the inhibition is complete at equimolar concentrations in dilute solutions. Dilution of this type of mixture results in the inhibitor-proteinase complex dissociating too slowly to be observed. A reversible inhibitor that behaves in this manner is refered to as tight binding.[71]

### *1.2.3.1 Reversible inhibitors*

A general reversible inhibitor could interact with an enzyme in a number of ways; by competing with the substrate for the active site or by binding to one or more of the catalytic complexes. The substrate may or may not be able to interact with the enzyme / inhibitor complex (Scheme 1.9 on page 24). All possible interactions between the enzyme and inhibitor are shown in Scheme 1.9. If a particular inhibitor does not interact with an enzyme species, then the forward and reverse rate constants are zero. $K_s$ is the binding constant of the substrate to the free enzyme E and $K_i$ is the dissociation constant for the EI complex. The inhibitor changes the enzyme's affinity for S by a factor $\alpha$ and *vice versa,* while the rate of decomposition of EIS to the product is modified by a factor $\beta$.




Scheme 1.9: *The possible modes of inhibition as a reverible inhibitor.*

The effect of the inhibitor on $K_m$ and $V_m$ can be found as follows;

The various equilibrium constants for reversible inhibition are defined in Table 1.3.

| Equilibrium constant | Kinetic terms | Concentration terms |
|---|---|---|
| $K_i$ | $\frac{k_{-4}}{k_4}$ | $\frac{[E][I]}{[EI]}$ |
| $\alpha K_i$ | $\frac{k_{-2}}{k_2}$ | $\frac{[ES][I]}{[EIS]}$ |
| $K_s$ | $\frac{k_{-1}}{k_1}$ | $\frac{[E][S]}{[ES]}$ |
| $\alpha K_s$ | $\frac{k_{-3}}{k_3}$ | $\frac{[EI][S]}{[EIS]}$ |

Table 1.3: *The kinetic and concentration terms of the equilibrium constants involved in reversible inhibition.*

If [S] >> [I], the proportion of S turned over by the enzyme is negligible, so [I] and [S] are constant. The equilibrium concentrations of the species $[EX_i]$ (Table 1.4) are found from

the rate constants in Scheme 1.8 and Table 1.3 by the King Altman method[70] as mentioned previously in section 1.2.2 page 21.

| Enzyme species | Kinetic terms |
|---|---|
| [E] | $k_{-1}k_{-2}k_3[S] + k_{-1}k_{-2}k_{-4} + k_{-1}k_{-3}k_{-4} + k_2k_{-3}k_{-4}[I]$ |
| [ES] | $k_{-2}k_1k_3[S]^2 + k_{-3}k_1k_{-4}[S] + k_4k_3k_{-2}[I][S] + k_1k_{-2}k_{-4}[S]$ |
| [EIS] | $k_2k_1k_3[I][S]^2 + k_1k_2k_{-4}[S][I] + k_{-1}k_3k_4[I][S] + k_2k_3k_4[S][I]^2$ |
| [EI] | $k_2k_1k_{-3}[I][S] + k_{-1}k_{-2}k_4[I] + k_{-1}k_{-3}k_4[I] + k_2k_{-3}k_4[I]^2$ |

Table 1.4: *The equilibrium concentrations of the species [EX$_i$] in kinetic terms.*

From Table 1.4 it can be seen that;

$$(\alpha K_i)K_s = \frac{k_{-2}k_{-1}}{k_2k_1} \qquad 1.2.10$$

$$K_i(\alpha K_s) = \frac{k_{-4}k_{-3}}{k_4k_3} \qquad 1.2.11$$

As the rate of formation of EIS is not path dependent, equation 1.2.12 follows from Scheme 1.8 and equations 1.2.10 and 1.2.11;

$$\frac{k_{-4}k_{-3}}{k_4k_3} = \frac{k_{-2}k_{-1}}{k_2k_1} \qquad 1.2.12$$

$$k_{-4}k_{-3}k_2k_1 = k_{-2}k_{-1}k_4k_3 \qquad 1.2.13$$

Table 1.4 defines equation 1.2.14;

$$[E] = k_{-1}k_{-2}k_3[S] + k_{-1}k_{-2}k_{-4} + k_{-1}k_{-3}k_{-4} + k_2k_{-3}k_{-4}[I] \qquad 1.2.14$$

Dividing equation 1.2.14 by $k_1k_2k_3$ gives equation 1.2.15;

$$\frac{[E]}{k_1k_2k_3} = \frac{k_{-1}k_{-2}[S]}{k_1k_2} + \frac{k_{-1}k_{-2}k_{-4} + k_{-1}k_{-3}k_{-4}}{k_1k_2k_3} + \frac{k_{-3}k_{-4}[I]}{k_1k_3} \qquad 1.2.15$$

Substituting equilibrium constants in equation 1.2.16 and using the relationship in equation 1.2.13 gives;

$$[E] = k_1k_2k_3\alpha K_iK_s\{[S] + \frac{k_{-4}}{k_3} + \frac{k_3k_{-4}}{k_{-2}k_3} + \frac{k_4[I]\}}{k_1} \quad 1.2.16$$

Equation 1.2.16 can be simplified to equation 1.2.17.

$$E = \alpha YK_iK_s \quad 1.2.17$$

Where $$Y = k_1k_2k_3\{[S] + \frac{k_{-4}}{k_3} + \frac{k_3k_{-4}}{k_{-2}k_3} + \frac{k_4[I]\}}{k_1}$$

A similar procedure can be adopted for the other $[EX_i]$ species giving equations 1.2.18;

$$ES = \alpha Y[S]K_i \qquad EIS = Y[S][I] \qquad EI = \alpha Y[I]Ks \quad 1.2.18$$

The total enzyme concentration $E_o$ is given by equation 1.2.19;

$$E_o = E + ES + EI + EIS \quad 1.2.19$$

The maximal velocity is defined by equation 1.2.20;

$$V_{max} = k_5E_o \quad 1.2.20$$

The rate of reaction v is given by equation 1.2.21 from Scheme 1.8;

$$v = k_5[ES] + \beta k_5[EIS] \quad 1.2.21$$

Substituting equations 1.2.17 and 1.2.18 into equation 1.2.21 dividing by equation 1.2.19 and substituting in equation 1.2.20 gives equation 1.2.22;

$$v = \frac{V_{max}[S]\{\alpha K_i + \beta[I]\}}{[S][I] + \alpha\{K_iK_s + [S]K_i + K_s[I]\}} \quad 1.2.22$$

Inverting equation 1.2.22 gives equation 1.2.23;

$$\frac{1}{v} = \frac{[S][I] + \alpha\{K_iK_s + [S]K_i + K_s[I]\}}{V_{max}[S]\{\alpha K_i + \beta[I]\}} \quad 1.2.23$$

Scheme 1.8 and equation 1.2.23 can be used to define different types of inhibition (depending on the values of $\alpha$ and $\beta$) as follows;

### 1.2.3.1.1 Fully competitive inhibition

Fully competitive inhibition arises when $\alpha = \infty$ and $\beta = 0$. At a suitable [I], the bound inhibitor completely prevents the substrate binding to the free enzyme. Substituting these values of $\alpha$ and $\beta$ into equation 1.2.23 and assuming that $[S][I]\alpha^{-1} = 0$ gives equation 1.2.24;

$$\frac{1}{v} = \frac{1}{V_{max}} + \frac{K_m}{[S]V_{max}}\left\{1 + \frac{[I]}{K_i}\right\} \qquad 1.2.24$$

$(V_{max})^{-1}$ is unaltered and the apparent $(K_m)^{-1}$ is changed by a factor of $1 + [I](K_i)^{-1}$. A double reciprocal plot consists of a series of non-parallel lines that intersect the y axis at $(V_{max})^{-1}$ as shown in Fig. 1.8a on page 29. The inhibitor competes with the substrate for the enzyme's active site, so the inhibitor can be displaced at high [S]. $(V_{max})^{-1}$ can be found from the 1/v axis intercept of the double reciprocal plot. $K_i$ can be found from a graph of the double reciprocal plot gradients verses [I], where the new gradient is given by $K_m(V_{max}K_i)^{-1}$ and the new y intercept is given by $K_m V_{max}^{-1}$.

### 1.2.3.1.2 Uncompetitive inhibition

Uncompetitive inhibition arises when $\alpha < 1$ and $0 < \beta < 1$. I binds to ES preferentially over E, S binds to EI rather than E and EIS turns over less rapidly than ES. Substituting these values of $\alpha$ and $\beta$ into equation 1.2.23 and assuming that [I] is small compared to $K_i$ gives equation 1.2.25;

$$\frac{1}{v} = \frac{1}{V_{max}}\left\{1 + \frac{[I]}{\alpha K_i}\right\} + \frac{K_m}{V_{max}[S]} \qquad 1.2.25$$

$(V_{max})^{-1}$ is altered by a factor of $1 + [I](K_i)^{-1}$ and the apparent $(K_m)^{-1}$ is unchanged as the inhibitor does not affect the substrate binding to the enzyme. A double reciprocal plot consists of a series of parallel lines as shown in Fig. 1.8b. $V_{max}$ and $K_i$ can be found from a graph of the 1/v axis intercepts of the double reciprocal plot verses [I], where $(V_{max})^{-1}$ is given by the new intercept and $(\alpha V_{max} K_i)^{-1}$ is given by the new gradient.

#### 1.2.3.1.3 Non-competitive inhibition

Fully non-competitive inhibition arises when $\alpha = 1$ and $\beta = 0$. The bound inhibitor does not affect the binding of the substrate (and *vice versa*) but product formation is prevented as I has a similar affinity for E and ES. Substituting these values into equation 1.2.23 gives equation 1.2.26;

$$\frac{1}{v} = \frac{1}{V_{max}}\left\{1+\frac{[I]}{K_i}\right\} + \frac{K_m}{[S]V_{max}}\left\{1+\frac{[I]}{K_i}\right\} \qquad 1.2.26$$

$(V_{max})^{-1}$ and the apparent $(K_m)^{-1}$ are changed by a factor of $1 + [I](K_i)^{-1}$, as the inhibitor binds to both the free enzyme and another enzyme complex. If $\beta$ is non-zero the apparent $K_i$ is made up of two terms one deriving from the affect of $\alpha$ and the other arising from the effect of $\beta$. A double reciprocal plot consists of a series of non-parallel lines as shown in Fig. 1.8c. $V_{max}$ and $K_i$ can be found from a graph of the gradients of the double reciprocal plot verses [I]. $K_m V_{max}^{-1}$ is given by the new y axis intercept and $K_m(V_{max}K_i)^{-1}$ is given by the new gradient.

Partial non-competitive inhibition arises when $\alpha = 1$ and $0 < \beta < 1$. The bound inhibitor does not affect the substrate binding to the free enzyme but product formation from EIS is slower than from ES. Kinetically partial non-competitive inhibition is identical to full non-competitive inhibition.

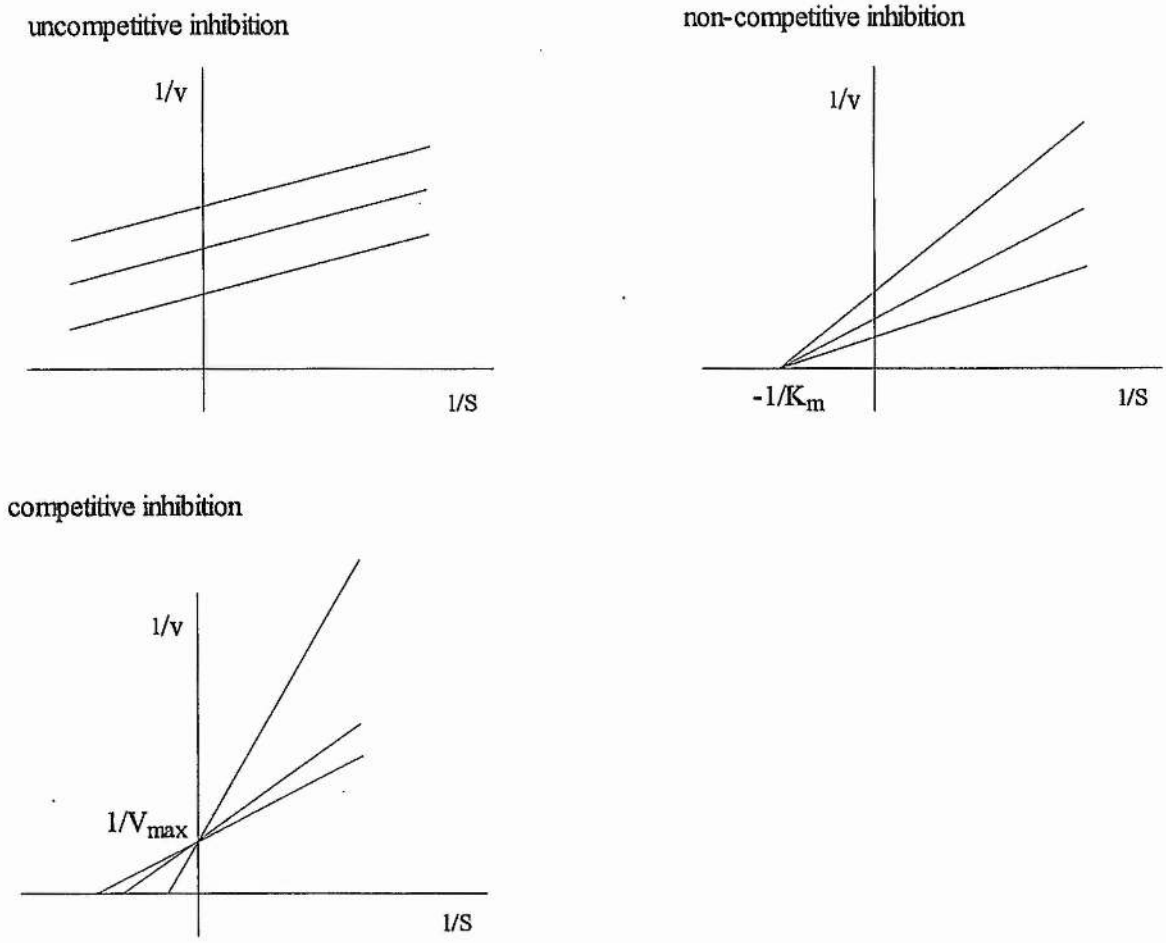



Fig. 1.8: *Double reciprocal plots for a) competitive inhibition, b) uncompetitive inhibition and c) non-competitive inhibition*

### 1.2.3.2 Irreversible inhibitors

Irreversible inhibitors act by binding tightly to the enzyme or one of the intermediates preventing catalytic activity. The binding is usually covalent and stoichiometric (often 1:1), so they have activity at a similar concentration to the enzyme and are therefore usually more potent inhibitors that reversible inhibitors. Irreversible inhibitors are characterized by their velocity of inactivation and can not be displaced by the substrate. Irreversible inhibitors are usually quite reactive molecules and tend to have short half lifes *in vivo*.

Some irreversible inhibitors (suicide inhibitors) are present in an inactive form which becomes activated by partial turn-over by the enzyme. Upon partial turnover a reactive, tight binding intermediate is formed that stops further turnover of the substrate. These have longer half lives *in vivo* as the inactive form of the molecule is usually not very reactive.

## 1.2.4 Thermodynamic analysis of proteinase catalysis

The hydrolysis of a substrate by a proteinase can be described by two equations (equation 1.2.27 and 1.2.28);

$$k = A \exp(-\Delta G^{\ddagger}/RT) \qquad 1.2.27$$

$$\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger} \qquad 1.2.28$$

Where $\Delta G^{\ddagger}$ is the change in Gibbs free energy of activation , $\Delta H^{\ddagger}$ is the change in enthalpy of activation, T is the temperature; $\Delta S^{\ddagger}$ is the change in entropy of activation; k is the rate constant; A is the frequency factor and R is the gas constant.

If the following reaction is considered:

$$S \rightleftharpoons (S...P)^{\ddagger} \rightleftharpoons P$$

where the substrate (S) is hydrolysed to form the products (P) via a transition state complex $(S...P)^{\ddagger}$. The rate of a proteinase catalysed reaction is faster than the same reaction uncatalysed. The energy profile of the uncatalysed reaction is shown in Fig. 1.9a on page 31.

It can be seen that if S has sufficient energy to overcome the barrier $\Delta G^{\ddagger}$, P is formed via the transition state $(S...P)^{\ddagger}$. The reverse reaction is also possible if P has sufficient energy to overcome $\Delta G^{\ddagger}$ and $\Delta G°$ and S is reformed. If the barrier, $\Delta G^{\ddagger}$ is lowered, the process will be speeded up at a given temperature, as a greater proportion of S has enough energy to overcome $\Delta G^{\ddagger}$ and a greater proportion of P has sufficient energy to overcome $\Delta G^{\ddagger}$ and $\Delta G°$. This can be achieved by either reducing $\Delta H^{\ddagger}$ or raising $\Delta S^{\ddagger}$.

a) b)

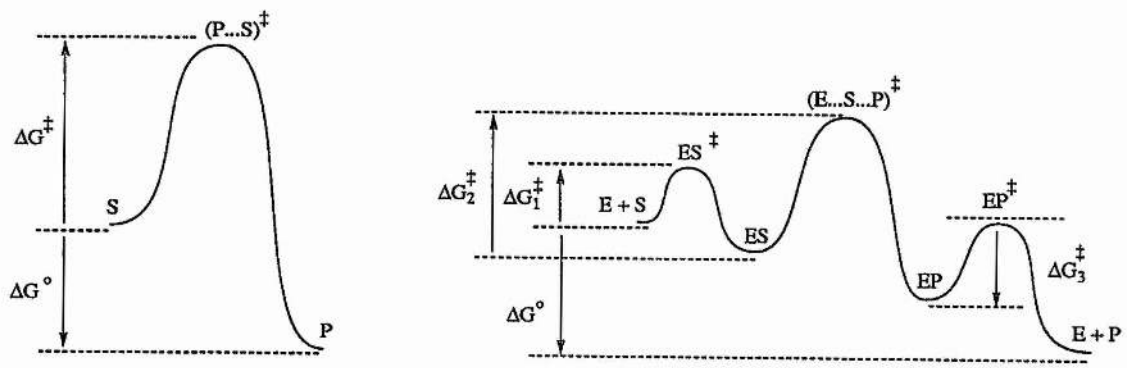



Fig. 1.9: *The energy profile of a) an uncatalysed and b) a proteinase catalysed hydrolysis.*

A proteinase can achieve this by preferentially binding to and stabilising $(S...P)^{\ddagger}$ relative to the reactants and products (Fig. 1.9b).

It should be noted that the standard Gibbs free energy change ($\Delta G°$) and the equilibrium constant ($K_{eq}$) are unchanged and the proteinase, unless adapted otherwise will catalyse both forward and reverse reactions until equilibrium is reached. Proteinases are very specific, usually catalysing only one type of hydrolysis, under very limited chemical and physical conditions. They are usually enantiomerically selective.

Proteinases are proteins with aminoacids linked in a specific sequence. They have a coiled structure folded into a specific shape and are stabilised by hydrogen bonds between amido hydrogens and carbonyl oxygens of the polypeptide chain, along with ionic and hydrophobic interactions between the side chains of aminoacid residues. Certain functional groups are brought together by the folding process to form the active site of the proteinase, and are responsible for the catalytic activity of the proteinase. These aminoacids are often found in different areas of the polypeptide chain. Occasionally the proteinase may require agents (cofactors) other that the substrate which supply specific chemical functions that are not possible with the proteinase alone.

## 1.2.5 The 23kDa adenovirus proteinase

### *1.2.5.1 Substrate specificity*

The adenovirus proteinase will cleave pTP, pVI, pVII,[28] PVIII,[27] pIIIa,[29] 11kDa[28] and pMu[30] proteins. These cleavages are vital for viral maturation and infectivity. A mutant virus with inactive proteinase can still produce these proteins and assemble virus particles, but the new virus particles are non-infectious.[24]

A proteinase catalyses the cleavage of specific peptide bonds in a polypeptide substrate (Fig. 1.10); where residues in the substrate as designated P and those of the proteinase are designated S. The residues are then numbered outwards from the cleaved bond and those residues on the carboxy side are designated prime.

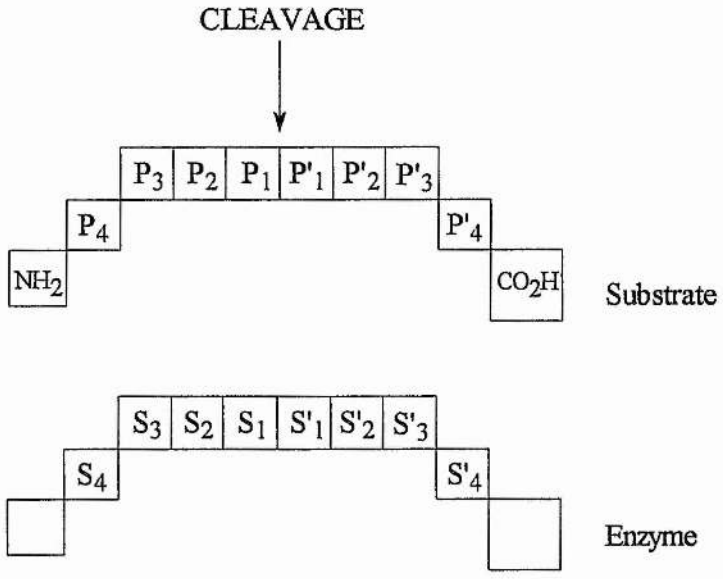



Fig. 1.10: *The cleavage of a substrate by a proteinase.*

Comparison of the cleavage sequences of these precursor proteins gave the following substrate recognition sequence for the enzyme[72] (Fig. 1.11) where usually Ala-Gly or Gly-Ala bonds are cleaved, although in special cases Gly-Phe, Arg-Gly and Phe-Gly bonds are also cleaved;[28]

| $P_4$ | $P_3$ | $P_2$ | $P_1$ | $P_1'$ | | $P_4$ | $P_3$ | $P_2$ | $P_1$ | $P_1'$ |
|---|---|---|---|---|---|---|---|---|---|---|
| (ML) | X | G | G | X | or | (ML) | X | G | X | G |

Fig. 1.11: *Substrate specificity of the adenovirus proteinase.*

A search of all the adenoviral proteins with this consensus sequence gave the known substrates for the proteinase as well as the penton, hexon, 52kDa, E21kDa, 14.5kDa and D-172 proteins. It is known that the mutant virus with inactive proteinase is defective at uncoating[73] so it may be that the proteinase cleaves the hexon and the penton *in vivo*.

### 1.2.5.2 Inhibitor studies

The proteinase is active at neutral pH and was unaffected by pepstatin,[74] a general aspartyl proteinase inhibitor, so is unlikely to be an aspartyl proteinase.[75] Pepstatin (3-methylbutanoyl-(2S)-valinyl-(2S)-valinyl-(3S,4S)-4-amino-3-hydroxy-6 methylheptanoyl alanyl (3S,4S)-4-amino-3-hydroxy-6 methylheptanoic acid) (**8**), shown in Fig. 1.12, is a tight binding inhibitor of aspartic proteinases (dissociation constant 45.7 pmol $dm^{-3}$ in the case of pepsin).[76] It is believed that pepstatin (**8**) mimics the tetrahedral intermediate for amide bond hydrolysis.[77]

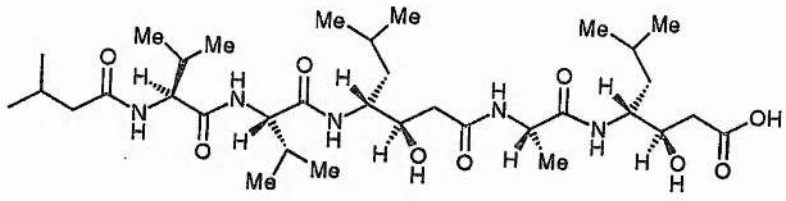


Fig. 1.12: *Pepstatin (8).*

The proteinase was also unaffected by ethylenediaminetetraacetic acid[80] so is unlikely to be a metalloproteinase.[78] This suggests that the adenovirus proteinase is either a cysteine or a serine proteinase. Usually these 2 classes of proteinase can be distinguished by the effects of a range of inhibitors given in Table 1.5 on page 36. However sequence alignments showing conserved active site residues are often required for absolute proof.[78]

The partially purified adenoviral proteinase is inhibited by diisopropylfluorophosphate (DFP) (**9**) and partially inhibited by benzylsulphonyl fluoride (PMSF) (**10**), thus implying that the adenovirus proteinase is a chymotrypsin-like serine proteinase.[79]

However upon further purification of the proteinase, it was found that general thiol attacking reagents; copper (II) ions, Zn (II) ions,[80] iodoacetate (**11**), N-ethylmaleimide (NEM) (**12**), dithiodipyridine (**13**) and parachloro mercury benzoate (PCMB) (**14**) all inhibit the enzyme.[80] Dithiodipyridine (**13**) also inhibited the proteinase at pH 5.0, suggesting that there is a thiol present with a very low pKa. Such a thiol group can be found in the active site of a cysteine proteinase.[78] At pH 3-5, dithiodipyridine (**13**) is regarded as a specific active site titrant for cysteine proteinases.[81] Soybean trypsin inhibitor (SBTI) (**15**) and chymostatin (**16**) [77] did not effect the enzyme.[80, 74]

However **E-64** (**17**) (see Section 2.2.1) and Leupeptin (**18**) (Fig. 1.13,[82] see Section 2.1.4 for inhibition of a cysteine proteinase by an aldehyde) also failed to inhibit the proteinase,[74] although they are general cysteine proteinase inhibitors. All these inhibitors have bulky residues in the $P_1$-$P_4$ binding sites and considering the binding requirements of the proteinase, it is suggested that these inhibitors may not fit into the proteinase's active site.

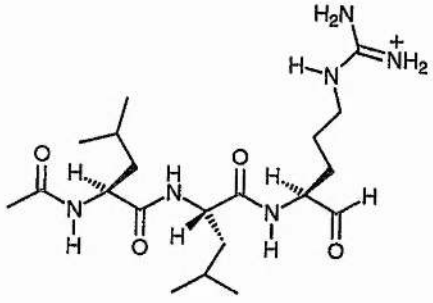


Fig. 1.13: *Leupeptin (18).*

Other cysteine proteinases can also be inhibited by PMSF (**10**),[80] although PMSF (**10**) is usually regarded as a more potent inhibitor of serine proteinases. Overall this evidence suggests that the adenovirus proteinase is a cysteine proteinase.[80]

| inhibitor | serine proteinase | cysteine proteinase |
|---|---|---|
| DFP (**9**) | inhibited | unaffected |
| PMSF (**10**) | inhibited | unaffected |
| SBTI (**15**) | inhibited | unaffected |
| chymostatin (**16**) | inhibited | unaffected |
| iodoacetate (**11**) | unaffected | inhibited |
| dithiodipyridine (**13**) | unaffected | inhibited |
| PCMB (**14**) | unaffected | inhibited |
| E-64 (**17**) | unafffected | inhibited |
| leupeptin (**18**) | unaffected | inhibited |
| optimum pH | 7-9 | 4-7 |

Table 1.5: *Specific inhibitors for cysteine and serine proteinases.*

### 1.2.5.3 Cofactor requirement

The gene encoding for the 23kDa proteinase was cloned and expressed into insect cells infected with *baculovirus*[83] and *E. coli.*[84] Upon purification to homogeneity the proteinase was found to be inactive. Activity could be recovered by the addition of heat-inactivated wild-type virus suggesting that a cofactor present in the heat-inactivated adenovirus was required for catalytic activity.[84] The residual fraction, left after complete adenovirus extraction , also restored catalytic activity. This fraction was analysed and found to contain a number of adenoviral structural proteins and adenoviral DNA. Purified adenoviral DNA on its own did not produce any activity.[84]

If dithiothreitol (DTT) (**19**), a reducing agent added to biological samples to prevent aerial oxidation, was added to the recombinant enzyme before the residual fraction, the proteinase was not activated. However if DTT (**19**) was added to a mixture of recombinant enzyme and the residual fraction after 15 min, the proteinase activity was enhanced. This suggested that the recombinant enzyme is activated by a factor present in the residual extract and that this activation is inhibited by reducing agents such as DTT (**19**).[72]

The proteins in the residual fraction were separated, purified and tested for proteinase activation. Two different fractions produced proteinase activation and were identified as the reduced and oxidised (cys-bridged dimer) form of an undecanomic peptide;[83]

GVQSLKRRRCF

Where G = glycyl; V = valinyl; Q = glutaminyl; S = serinyl; L = leucinyl; K = lysinyl; R = arginyl; C = cysteinyl; A = alanyl and F = phenylalanyl. This sequence is identical to the eleven C-terminal amino acid residues of the adenoviral structural protein pVI, a known substrate for the adenoviral proteinase. pVI also activates the proteinase. Neither GVQSLKRRRAF or KRRRCF could activate the proteinase, indicating that the Cys residue and the first five N-terminal amino acids were important for activation. The oxidised (dimer) form was found to be the actual cofactor for the proteinase. Activation by the reduced form of the peptide was due to traces of the dimer.[83]

It is believed that the cofactor is catalysing the oxidation / reduction of specific residues in the proteinase, and that a thiol disulphide interaction may be taking place. It is interesting to note that a catalytically inactive form of papain (propapain) exists where Cys-25 forms a disulphide bridge with Cys-22, leaving Cys-63 free. Propapain can be converted back to

papain by cleaving the Cys-22-Cys-25 link, forming a Cys-22-Cys-63 disulphide bond in place of the Cys-22-Cys-25 disulphide bond.[85] It could be that *in vivo* the newly translated proteinase folds into an inactive structure with the active site cysteine being involved in intramolecular disulphide bonding. During maturation the high concentration of pVI could activate the proteinase by disulphide interchange, leaving the active site cysteine free.[83] This represents a novel way of ensuring that the proteinase is only active when required.

The proteinase can be further activated by DNA, RNA, polyglutamic acid and polyaspartic acid, *ie.* polymers with a high negative charge density. *In vivo* viral DNA may act as a scaffold for the assembly of proteinase complexes, or the proteinase complex could use the DNA as a guide wire, travelling along the DNA as it cleaves the precursor proteins.[86] It is known that the proteinase is associated with viral cores.

### 1.2.5.4 Site directed mutagenesis

The proteinase has been sequenced in 12 different adenovirus serotypes and is highly conserved across serotypes, but contains no recognisable motifs in common with other known proteinases[73, 28, 87] The proteinase sequences for twelve adenovirus serotypes were aligned and the results are shown in Fig. 1.14 on page 39.[88] where N = asparginyl; D = aspartyl; E = glutamyl; H = histidinyl; I = isoleucinyl; M = methionyl; P = prolinyl; T = threonyl; W = tryptophanyl and Y = tryrosinyl.

| Type | 44 54 | ..95 104 | 122 | 182 |
|---|---|---|---|---|
| Avian 1 | ..TGSRASGGMH | ..TGL.RQPDRC | ..ACGLFSALIF | ..NSAYFRAHQE.. |
| Murine 1 | ..TASRETGGVH | ..SAVSSSSSKC | ..ACGLFCVLFL | ..NSLYFRLHVE.. |
| Canine 1 | ..TAGRETGGVH | ..SAIASTPDRC | ..ACGLFCCMFL | ..NSAYFRAHAA.. |
| Bovine 7 | ..TGPREQGGIH | ..SAL.SSPDKC | ..SCGLFCVFFL | ..KSLYFRQNEE.. |
| Bovine 3 | ..TAGRETGGVH | ..SALTSTADRC | ..ACGLFCCMFL | ..HSPHFVKHAA.. |
| Type 41 | TAGRETGGVH | ..SALASTPDHC | ..ACGLFCCMFL | ..HSPYFRRHRE.. |
| Type 40 | TAGRETGGVH | ..SALASTPDHC | ..ACGLFCCMFL | ..RSPYFQRHCE.. |
| Type 12 | ..TAGRETGGVH | ..SALAATKDRC | ..ACGLFCCMFL | ..LSPYFRHNRE.. |
| Type 2 | ..TAGRETGGVH | ..SAIASSPDRC | ..ACGLFCCMFL | ..HSPYFRSHSA.. |
| Type 3 | ..TAGRETGGEH | ..SALA.TKDRC | ..ACGLFCCMFL | ..HSSYFRSHSA.. |
| Type 4 | ..TAGRETGGEH | ..SALA.TKDRC | ..RCG.FSAACS | ..HSAYFRSHSA.. |
| Type 5 | ..TAGRETGGVH | ..SAIASSPDRC | ..ACGLFCCMFL | ..HSPYFRSHSA.. |

Fig. 1.14: *The sequence alignments of the twelve sequenced adenovirus proteinases.*

There are two conserved serine residues (Ser-95 and Ser-182, based on adenovirus 2 sequences), one conserved histidine residue (His-54) and two conserved cysteine residues (Cys-104 and Cys-122). Various mutants of the adenovirus 2 proteinase were made and are described in Table 1.16 on page 40.[88, 89]

Mutations of Ser-95 and Ser-182 to Ala (mutants 1 and 2) enhanced activity, so neither is important catalytically.[87] However when Cys-104 or Cys-122 were changed to Ala (mutants 3 and 4)[88, 89] and His-54 (mutant 5) was mutated to Arg[89] activity was lost. This leads to the

conclusion that the adenovirus proteinase is a cysteine proteinase as the only conserved residues that are essential for catalytic activity are His and Cys.[88, 89] However when Cys-122 was mutated to Gly[88] (mutant 6) catalytic activity remained, unlike mutant 7 where Cys-104 was changed to gly and no activity remained.[88] This suggests that Cys-104 is the active site nucleophile[88] and activation of the proteinase may involve a conformational change which is promoted by the interaction of the cofactor. In mutant 6 this interaction is still possible due to the small size of the glycine residue.[89]

| Mutant | residue altered | residue changed to | activity cf wild type |
|---|---|---|---|
| 1 | Ser-182 | Ala | 164% |
| 2 | Ser-95 | Ala | 110% |
| 3 | Cys-104 | Ala | 3% |
| 4 | Cys-122 | Ala | 0% |
| 5 | His-54 | Arg | 0% |
| 6 | Cys-122 | Gly | 81% |
| 7 | Cys-104 | Gly | 0% |

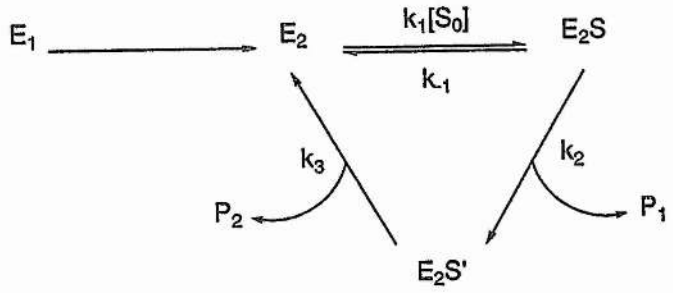
Table 1.6: Mutants of the adenovirus proteinase.

### 1.2.5.5 Possible mechanisms for the adenovirus proteinase

The adenovirus proteinase requires one histidine residue (His-54)[88, 89] and two cysteine residues (Cys-104 and Cys-122) for activity.[89] Cys-104 is thought to be the active site nucleophile,[88] while Cys-122 and the cofactor are involved in enzyme activation.[89] Inhibition by dithiopipyridine (**13**) indicated the presence of a thiol with a very low pKa, as would be

found in the active site of a cysteine proteinase.[80] It could be proposed that Cys-104 and His-59 form an ion pair as part of the active site of the adenovirus type 2 proteinase.

A summary of the possible double intermediate mechanism of the adenovirus proteinase is shown in Scheme 1.10.




Scheme 1.10: *The Adenoviral Proteinase's double intermediate mechanism.*

Initially, the cofactor binds to the inactive proteinase ($E_1$) changing its conformation and activating it to give $E_2$. The substrate then binds, it is not known whether acid interactions with the $P_1$ carbonyl, and / or base interactions with the $P_1$' amido nitrogen are important. The nucleophilic thiolate (Cys-104) then attacks the $P_1$ carbonyl carbon giving a stabilised tetrahedral intermediate ($E_2S$), which collapses upon attack of a water molecule to give two products, $P_1$ and $P_2$. The final step, involves product debinding, however it is not known which product *ie.* the P (free acid) or the P' (free amine) side, leaves first. The proteinase still hydrolyses a substrate by a double intermediate mechanism and the rate is still defined by equation 1.2.8 page 22, but $K_m$ and $V_{max}$ can not be defined in rate terms by the King Altman method[70] until the order of product debinding is known.

# 2.0 Aims

The adenovirus proteinase is a rather unusual enzyme. Site directed mutagenesis (Section 1.2.5.4) and inhibitor profile studies (Section 1.2.5.2) have shown that the proteinase belongs in the cysteine class of proteinases, however the adenovirus proteinase does not behave like a typical cysteine proteinase (Section 1.2.5).

The adenovirus proteinase is partially inhibited by some serine proteinase inhibitors (Section 1.2.5.2) but, remarkably is not inhibited by some typical cysteine proteinase inhibitors (Section 1.2.5.2). Site directed mutagenesis as described in Section 1.2.5.4 has shown that, of the two conserved cysteine residues (Cys-104 and Cys-122), and the one conserved histidine residue (His-54), each is required to be intact for normal levels of catalytic hydrolysis activity.[89, 90] Alteration of His-54 or Cys-104 residues, gives rise to completely inactive proteinase. However some activity remains if Cys-122 is replaced by a glycyl residue.[90] A thiol-containing cofactor (GVQSKRRRCF) is also required for proteinase activity, and once activated the proteinase can perform many catalytic cycles (Section 1.2.5.3).[84]

After the activation of the proteinase by the cofactor, the enzyme is able to catalyse the hydrolysis of substrates. This is believed to occur by way of a double intermediate type mechanism (Section 1.2.5.5), in which the order of product release is controlled by the proteinase but at the start of our studies, this was not known. The function of the three catalytically essential residues His-54, Cys-104 and Cys-122 is unknown. Furthermore, since there was no X-ray crystallographic data available for the proteinase, it was not known which other residues might participate in catalysis, or form hydrogen bonds or Van der Waals contacts with the intermediates involved in the hydrolysis of a substrate.

The major objective of our research was to probe the mechanism of the adenovirus proteinase by the use of mechanism-based inhibitors to help elucidate the active site structure of the proteinase.

## 2.1 The synthesis of reversible inhibitors

For a compound to serve as a mechanistic probe for an enzyme, the compound should specifically recognise the protein, bind to the active site and then further react in some way. Mechanistic probes can be substrates or inhibitors but in general, for proteinases, information on potential for binding has been derived from careful analysis of substrate structure-activity relationships. Such analyses reveal which substituents can interact strongly with polar, charged or hydrophobic groups at the active site of the enzyme. Where inhibitors are required, for biomedicinal and pharmacological studies, as is often the case, it is desirable that these substituents are placed in a framework that is inert towards proteolytic hydrolysis.

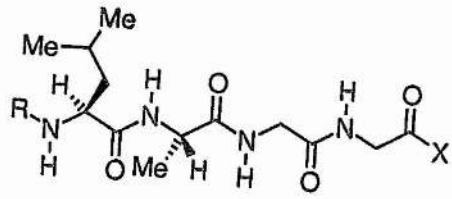
Prior to the work described here, no information or reports of the existence of specific inhibitors for the adenovirus proteinase had been published. We believed that the general binding requirements of the proteinase could be satisfied by the use of the substrate recognition sequence in the structure of the inhibitor and, therefore, chose to implement such a strategy.

The adenovirus proteinase has a very specific substrate recognition sequence; (M, L, I)$X^1$GG-$X^2$ (section 1.2.5.1), where $X^1$ and $X^2$ can be any residue and where the scissile bond is between the second glycyl residue and residue $X^2$.[29, 73] Thus it seemed sensible to use the sequence (2S)-leucyl (2S)-alanyl glycyl glycine, to satisfy the binding requirements for the proteinase in each series of potential inhibitor structures, where variations are introduced in place of the $X^2$ residue.

In order to measure the activity of the potential inhibitors it was necessary to consider possible assay procedures. One type of assay that had been employed in the Kemp group involved monitoring the change in absorbance of a chromophore that had reacted with the amino group produced from the hydrolysis of the substrate. So that this type of assay can be performed successfully, it is necessary to ensure that the leucine residue in the inhibitor is N-protected. To achieve this we chose to use carbobenzyloxy (Cbz), tertiary-butoxy carbonyl (t-Boc) and trifluoroacetyl (TFA) protection, as the starting reagents are cheap, readily available and reliable. Indeed, in the event, the N-protection of leucine with these protecting groups was achieved in high yields *ie.* 88%, 91% and 82% respectively. Upon testing the compounds with the proteinase, the chromophore assay procedure was found to be unreliable and the cleavage of the substrate was followed by capillary electrophoresis (CE) insead, so N-protection of the leucine residue was not actually necessary.

## 2.1.0 Rationale for the amide tetrapeptides

Using a series of simple amides in which the terminal leucyl amino group was either unprotected or N-protected with Cbz- or t-Boc-, the $P_1$ glycyl residue carboxy group was aminated to give a series of alkyl amides (Fig. 2.1). It was expected that it should be possible to determine the size of the S' binding pocket from structure-activity relationships data. This series of amides (Fig. 2.1) also possess different nitrogen leaving groups, so that if the rate determining step in the proteolytic hydrolysis of the amides involves breaking or forming a nitrogen bond, it was anticipated that different rates of proteolytic hydrolysis might be observed.



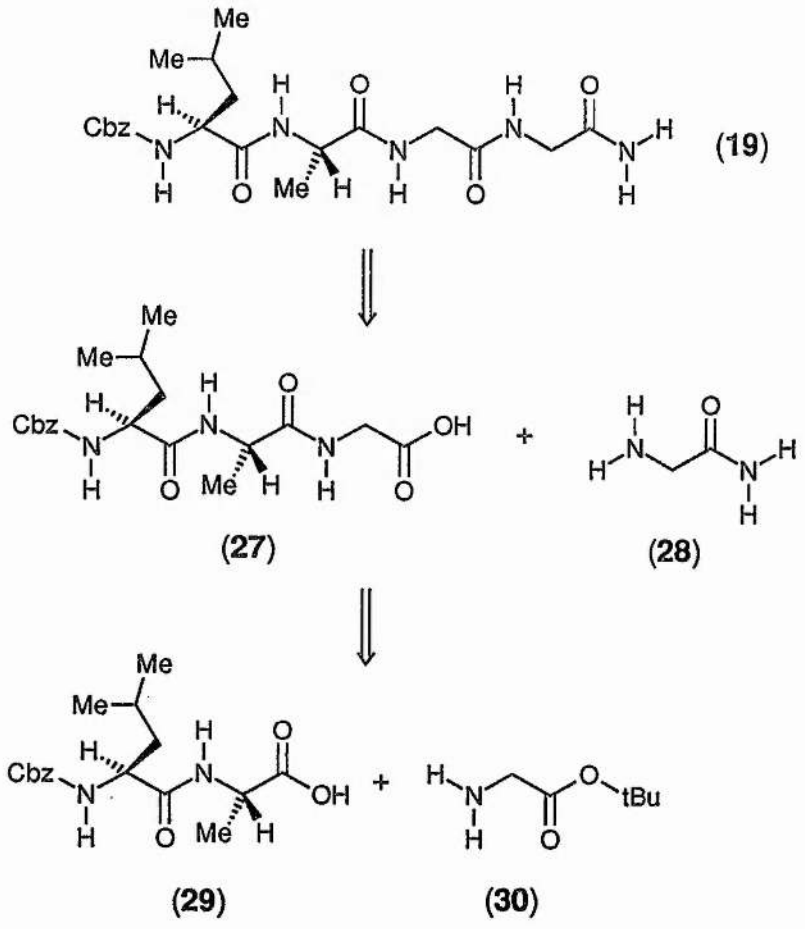
| R | X | Number |
|---|---|---|
| $-OCOCH_2C_6H_5$ | $-NH_2$ | **(19)** |
| $-OCOCH_2C_6H_5$ | -NHMe | **(20)** |
| $-OCOCH_2C_6H_5$ | $-NMe_2$ | **(21)** |
| $-OCOCH_2C_6H_5$ | -NMeEt | **(22)** |
| $-OCOCH_2C_6H_5$ | -N(OMe)Me | **(23)** |
| $-OC(CH_3)_3$ | -N-*p*$NO_2$-Bz | **(24)** |
| -H | $-NH_2$ | **(25)** |
| -H | -NHMe | **(26)** |

Fig. 2.1: *The series of simple amides.*

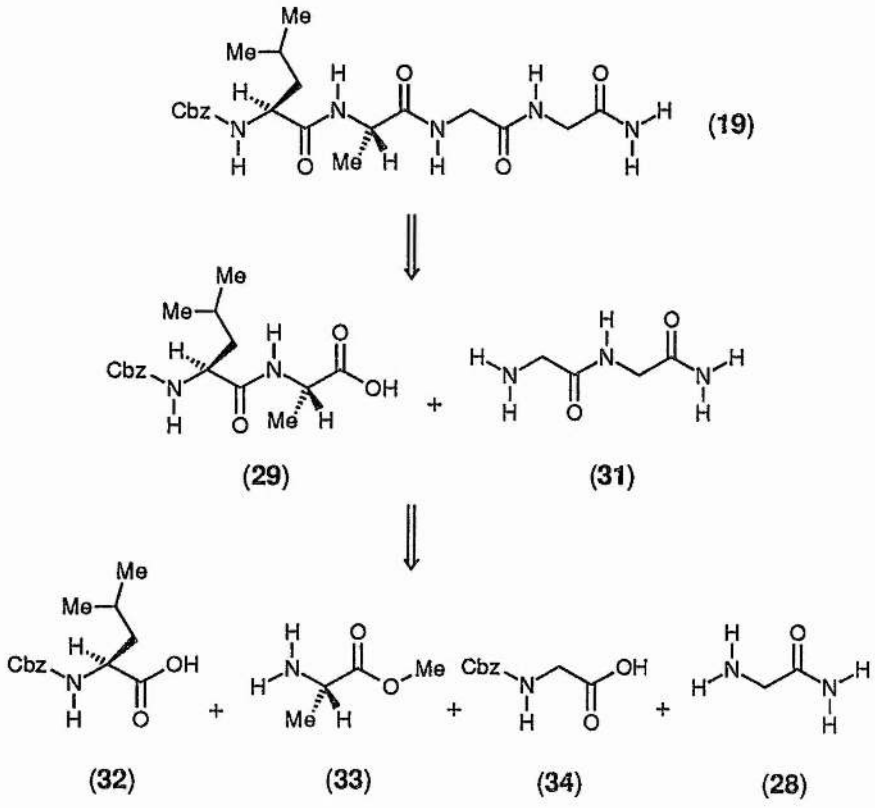
In considering the synthesis of the amides (Fig. 2.1) and taking carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycyl amide, Cbz-Leu-Ala-Gly-Gly-$NH_2$ **(19)**, as an example, the amide **(19)** can be disconnected in a stepwise manner (Scheme 2.1), or a convergent manner (Scheme 2.2). The stepwise disconnection of Cbz-Leu-Ala-Gly-Gly-$NH_2$ **(19)** results in carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycine **(27)** and glycyl amide **(28)**. It was envisaged that the tripeptide **(27)** could be prepared from carbobenzyloxy-(2S)-leucyl (2S)-alanine **(29)** and glycine t-butyl ester **(30)**, followed by carboxyl group deprotection.

The convergent disconnection of Cbz-Leu-Ala-Gly-Gly-$NH_2$ (**19**) results in Cbz-Leu-Ala (**29**) and glycyl-glycyl amide (**31**). It was expected that Cbz-Leu-Ala (**29**) could be prepared from carbobenzyloxy-(2S)-leucine (**32**) and (2S)-alanine methyl ester (**33**), followed by ester hydrolysis and Gly-gly amide (**31**) would be prepared from carbobenzyloxy-glycine (**34**) and glycyl amide (**28**), followed by N-deprotection.

Comparison of the convergent and the linear, stepwise approach, shows that if more than one tetrapeptide amide is prepared, the convergent approach requires the synthesis of several different glycyl glycyl amide dipeptides and the different glycyl amides, and the convergent approach requires the synthesis of the different glycyl amides only. As more than one amide would be prepared the stepwise method of synthesis appears to be better, as overall this method takes fewer steps than the convergent synthesis.




Scheme 2.1: *The stepwise disconnection of a tetrapeptide amide.*

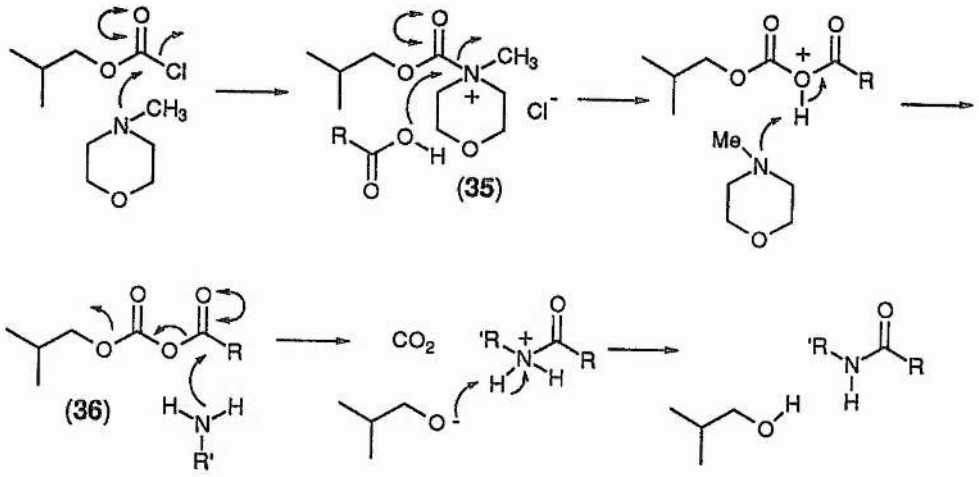



Scheme 2.2: *The convergent disconnection of a tetrapeptide amide.*

The key steps in the formation of the tetrapeptides consist of amide bond formation. There are numerous methods to form the amide bonds, such as the acid chloride,[90] azide,[91] activated ester,[92] anhydride[93] and carbodiimide routes.[94] We chose the mixed anhydride[95] approach where possible as usually, this method is quick, high yielding and the products do not require extensive purification. Racemisation of sensitive amino acids is also avoided by the use of a bulky base such as N-methylmorpholine (NMM) and a non-polar solvent such as THF or dichloromethane.[95]

The amide bond is formed by the use of an alkyl chloroformate such as *iso*-butyl chloroformate (i-BCF) and a base (NMM) which react to form an activated uronium species (**35**) (Scheme 2.3). The NMM is then displaced by an N-protected amino acid to give an

unsymmetrical, unstable anhydride (**36**) and N-methyl morpholine hydrochloride, after proton exchange. These two steps are very facile, and the formation of N-methyl morpholine hydrochloride is observed instantaneously upon the addition of the alkyl chloroformate, to a mixture of NMM and the N-protected amino acid in THF at -20°C. On addition of the O-protected amino acid, the mixed anhydride is attacked slowly to give the new peptide, carbon dioxide and *iso*-butyl alcohol.[95]




Scheme 2.3: *The mixed anhydride peptide coupling reaction.*

## *2.1.0.0 Synthesis of the tetrapeptide amides*

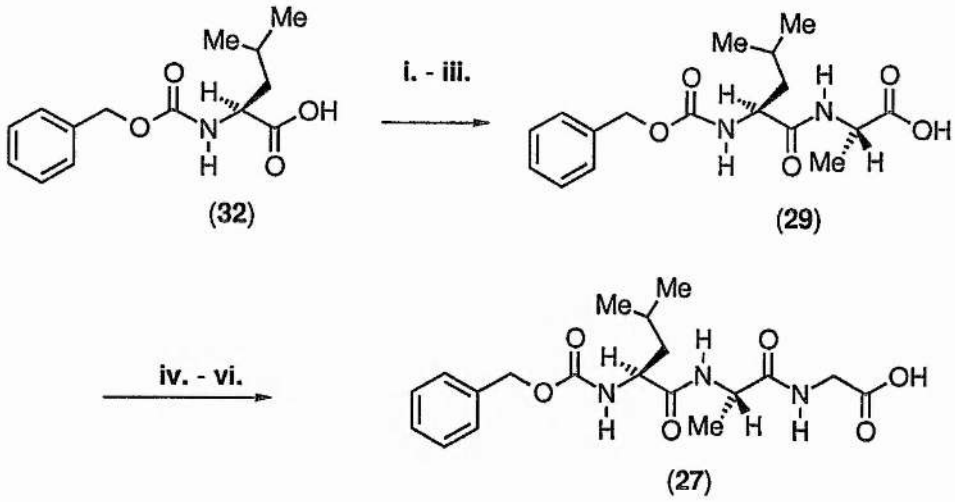
### 2.1.0.0.0 Synthesis of Carbobenzyloxy (2S)-leucyl (2S)-alanyl glycine (27)

Carbobenzyloxy (2S)-leucine (**32**)[96] was coupled with methyl (2S)-alaninate hydrochloride (**33**)[97] using the mixed anhydride procedure[95] to give methyl carbobenzyloxy-(2S)-leucyl (2S)-alaninate (**37**) in a yield of 95% after recrystallisation from ether / hexane (M.p.91°C; lit.,[98] 92-93°C) (Scheme 2.4). Treatment of the product (**37**) with 1 mol $dm^{-3}$ sodium hydroxide at 35°C for 3.5 hours and acidification gave carbobenzyloxy-(2S)-leucyl (2S)-alanine (**29**) in 88% yield after recrystallisation from ether / hexane (M.p. 146-147°C;

lit.,[99] 152-153°C) (Scheme 2.4). Both $^{1}$H- and $^{13}$C-NMR spectra showed the complete disappearence of the methyl ester signals.

Further extension of Cbz-Leu-Ala (**29**) was achieved using the mixed anhydride method[95] and t-butyl glycinate acetate salt (**30**) to give t-butyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycinate (**38**) in a yield of 88% after recrystallisation from acetone / water (M.p. 111-112°C). Cbz-Leu-Ala-Gly-OtBu (**38**) displayed the expected properties and showed the appearance of a singlet at 1.43 ppm coresponding to the t-butyl ester signal and a doublet at 3.98 ppm coresponding to the glycyl methylene group and a triplet at 6.94 ppm coresponding to the glycyl amide nitrogen proton in the $^{1}$H-NMR spectrum.

Removal of the t-butyl ester under acidic conditions[100] gave carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycine (**27**) in 83% yield after recrystallisation from acetone / ether (M.p. 131°C). The $^{1}$H and $^{13}$C NMR spectra showed the complete disappearance of the t-butyl ester signal.




**Reagents**
**i.** NMM, i-BCF in THF at -15°C, 5 min. **ii.** Ala Me ester.HCl (**33**) , NMM in DMF at -15°C, 5 min → r.t. 2 hours → (37). **iii.** NaOH in MeOH at 35°C, 41/2 hours; then HCl to pH 2. **iv.** NMM, i-BCF in THF at -15°C, 5 min. **v.** Gly t-Bu ester acetate, NMM in THF at -15°C, 5 min → r.t., 2 hours → (38). **vi.** 95% formic acid, 3 hours.

Scheme 2.4: *The synthesis of Cbz-(2S)-leucyl (2S)-alanyl glycine (27).*

Initially an attempt was made to prepare the Cbz-Leu-Ala-Gly (**27**) through the hydrolysis of ethyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycinate (**39**) which was itself prepared from the mixed anhydride coupling of ethyl glycinate hydrochloride (**40**) and Cbz-Leu-Ala (**2.11**) (M.p. 129°C). However Cbz-Leu-Ala-Gly-OEt (**39**) was resistant to base catalysed hydrolysis, under either aqueous methanol or anhydrous conditions (sodium t-butoxide and one equivalent of water).[101] Since the Cbz-group is not stable to acidic hydrolysis this route was abandoned in favour of the pathway to Cbz-Leu-Ala-Gly (**27**) depicted in Scheme 2.4.

### 2.1.0.0.1 Synthesis of t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycine (46)

t-Butoxycarbonyl (2S)-leucine (**41**)[102] was coupled with methyl (2S)-alaninate hydrochloride (**33**) using the mixed anhydride procedure[95] to give methyl t-butoxycarbonyl-(2S)-leucyl (2S)-alaninate (**42**) in 98% yield after recrystallisation from ether / hexane (M.p.110°C; lit.,[103] 114-116°C) (Scheme 2.5). Treatment of the product (**42**) with 1 mol $dm^{-3}$ sodium hydroxide at 35°C for 3.5 hours and acidification gave t-butoxycarbonyl-(2S)-leucyl (2S)-alanine (**43**) in 82% yield after recrystallisation from ether / petroleum ether (M.p. 135°C; lit.,[104] 132-133°C). Both the $^{1}H$ and $^{13}C$-NMR spectra showed the complete disappearence of the methyl ester signal at 3.75 ppm and 53.84 ppm respectively.

Further extension of t-Boc-Leu-Ala (**43**) was achieved using the mixed anhydride method[95] and benzyl glycinate tosylate salt (**44**)[105] to give benzyl t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycinate (**45**) as a monohydrate in 88% yield after recrystallisation from acetone / water (M.p. 68-69°C) ( Scheme 2.5). The $^{1}H$-NMR spectrum showed that the Gly α-protons were diastereotopic rather than equivalent and appeared as a two doublets at 4.08 and 4.03 ppm due to the effect of the chiral carbon centre in the adjacent Ala residue.

The benzyl ester was removed from t-Boc-Leu-Ala-Gly-OBz (**45**) by treatment with palladium (II) acetate, dimethylethylsilane and triethylamine;[106] then methanol to give

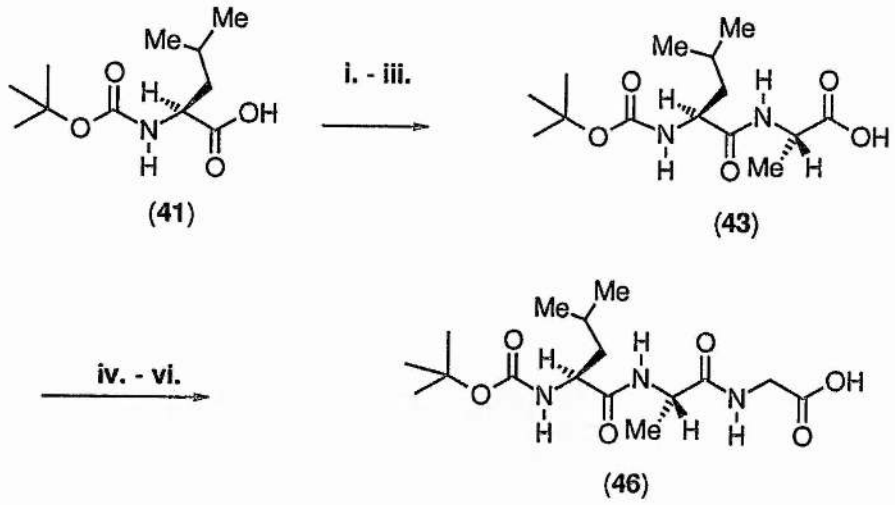
t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycine (**46**) in 96% yield after trituration from acetone / ether (M.p. 111-112°C).

The possible palladium (II) acetate mechanism is shown is Scheme 2.6 on page 53. Oxidative addition of dimethylethylsilane to palladium (II) acetate (16 electron species, oxidation state II) results in the cleavage of the silicon-hydrogen bond and the formation of complex (**47**) (18 electron species, oxidation state IV). The reductive elimination of acetic acid and the oxidative addition of the benzyl ester (**45**) to complex (**47**) results in the cleavage of the benzylic-oxygen bond of the benzyl ester (**45**) and the formation of complex (**48**) (18 electron species, oxidation state IV). Rotation of the groups attached to the metal centre allows the dimethylethylsilyl group and the peptidic carboxylate group to align next to one another so the reductive formation of the palladium bound silyl ester (**51**) and the addition of triethylamine results in the formation of complex (**49**) (18 electron species, oxidation state II). The loss af triethylamine and oxidative addition of acetic acid (binding in the X manner) results in the cleavage of the oxygen-proton bond of the acetic acid and the formation of complex (**50**) (18 electron species, oxidation state IV). The triethylamine may not necessarily bind to the palladium resulting in a 16 electron complex which could gain acetic acid in a similar manner to the 18 electron complex (**49**) to give complex (**50**). If the triethylamine does not interact with the palladium metal centre, the triethylamine could be added to prevent any acetic acid generated *in situ* from hydrolysing the palladium complexes formed during the catalytic cycle.

Loss of dimethylethylsilyl t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycinate (**51**) and the acetate binding in the X manner changing its binding mode to the LX manner results in the formation of complex (**52**). Rotation of the groups attached to the metal centre allows the hydrogen attom and the benzyl group to align next to one another so reductive elimination of toluene can occur from complex (**52**) to release palladium (II) acetate which is available for

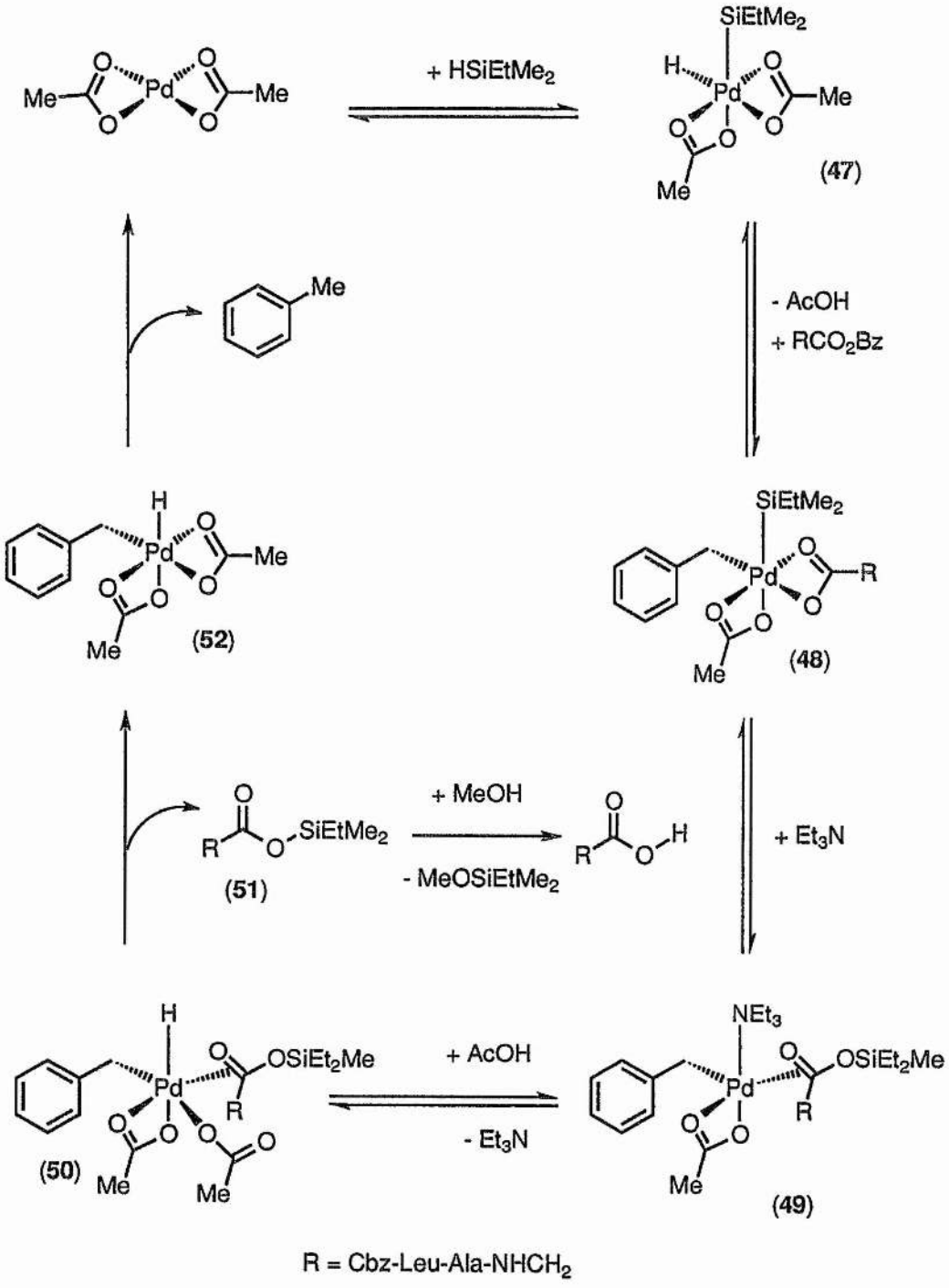
further catalytic cycles. Upon addition of methanol, the silyl ester (**51**) is attacked by the methanol to give t-Boc-Leu-Ala-Gly (**46**) and dimethylethylsilylmethoxide.

Catalytic hydrogenolysis of t-Boc-Leu-Ala-Gly-OBz (**45**) was also attempted in methanol but was accompanied by 10% transesterification to the methyl ester. Addition of small amounts of water to the methanol solvent to prevent the transesterification, resulted in the precipitation of the starting material. So this approach was abandoned and the method shown in Scheme 2.5 was used to prepare t-Boc-Leu-Ala-Gly (**46**). Both $^{1}$H- and $^{13}$C-NMR spectra showed the complete disappearence of the benzyl ester signals. In the $^{1}$H-NMR spectrum, the Gly α-protons remained non-equivalent and the signals appeared as two doublets at 3.96 ppm and 3.89 ppm. The structure of t-Boc-Leu-Ala-Gly (**46**) was also confirmed by IR spectroscopy, microanalysis and by mass spectrometry.




**Reagents**
**i.** NMM, i-BCF in THF at -15°C, 5 min. **ii.** Ala Me ester.HCl (**33**) , NMM in DMF at -15°C, 5 min → r.t. 2 hours → (42). **iii.** NaOH in MeOH at 35°C, 41/2 hours; then HCl to pH 2. **iv.** NMM, i-BCF in THF at -15°C, 5 min. **v.** Gly Bz ester tosylate (**44**), NMM in THF at -15°C, 5 min → r.t., 2 hours → (45). **vi.** $Pd(OAc)_2$, $Me_2EtSiH$, $Et_3N$ in acetone 5 min., then compound (**45**), 12 hours; then MeOH 3 hours.

Scheme 2.5: *The synthesis of t-Boc-(2S)-leucyl (2S)-alanyl glycine (2.24).*




Scheme 2.6:

*The benzyl ester deprotection mechanism of $Pd(OAc)_2$, dimethylethylsilane and methanol.*

### 2.1.0.0.2 Synthesis of the glycyl amides

Glycyl amide hydrochloride (**28**) and glycyl-N-methyl amide hydrochloride (**53**) were prepared by stirring methyl glycinate hydrochloride (**54**)[114] in a concentrated aqueous solution of the appropriate amine for 2 days.[107] The yields were 90% and 78% respectively, after recrystallisation from methanol (M.p. decomp. 182°C; lit.,[108] 200-203°C and 148°C; lit.,[109] 153.5-156°C respectively). In the case of the N-methyl amide (**53**), the signal for the methyl group attached to the nitrogen-atom appeared at 2.61 ppm and 28.56 ppm in the $^{1}$H- and $^{13}$C-NMR spectra.

It was envisaged that an analogous procedure could be used for the preparation of glycyl-N,N-dimethylamide (**55**) but after repeated attempts, only impure glycyl-N, N-dimethylamide (**55**) could be obtained and therefore an alternative route was investigated, as outlined below.

t-Butoxycarbonyl-glycine (**56**)[114] and aqueous dimethylamine were coupled using the mixed anhydride method[95] to give t-butoxycarbonyl-glycyl-N, N-dimethylamide (**57**), in 40% yield after recrystallisation from methanol (M.p. 74°C). The low yield from the reaction was probably due to the use of aqueous dimethylamine which presumably causes some hydrolysis of the mixed anhydride (Scheme 2.7). Acid catalysed hydrolysis[110] of the product (**57**) gave glycyl-N,N-dimethylamide (**55**), in 53% yield after recrystallisation from methanol / ether. Gly-N,N-dimethylamide (**55**) displayed the expected spectral signals and gave satisfactory elemental analysis. A sharp melting point for glycyl-N,N-dimethylamide (**55**) could not be obtained as the crystals were too hygroscopic.

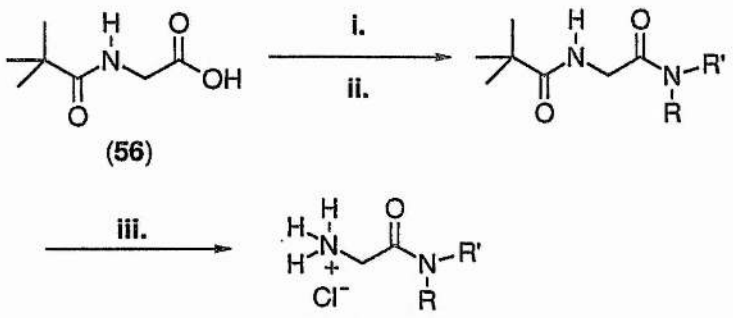
As expected, the $^{1}$H- and $^{13}$C-NMR spectra also showed that the two methyl groups attached to the amide nitrogen in both t-Boc-glycyl-N,N-dimethylamide (**57**) and glycyl-N, N-dimethylamide (**55**) were not equivalent. This arises as the carbon nitrogen bond obtains some double bond character from the delocalisation of the nitrogen lone pair within the

carbonyl bond. Rotation about the carbon nitrogen bond is restricted and when the methyl group is *cis* to the oxygen of the amide bond it is in a different enviroment, to that when the methyl group is *trans* to the oxygen of the amide bond. Thus two signals are observed, one for each methyl group in both the $^{1}H$ and $^{13}C$-NMR spectra.

An analogous procedure was used for the preparation of t-butoxycarbonyl-glycyl N-methyl,N-ethylamide (**58**) (M.p. 86°C) and t-butoxycarbonyl-glycyl N,O'-dimethylhydroxylamide (**59**) (M.p.101°C) in yields of 67% and 81%, respectively, after recrystallisation from methanol. Both of the t-Boc-protected amides showed the expected spectral characteristics and gave satisfactory microanalyses.

Acid catalysed hydrolysis of t-Boc-Gly-N-methyl,N-ethylamide (**58**) and t-Boc-Gly-N,O'-dimethylhydroxylamide (**59**) gave glycyl-N-methyl,N-ethylamide (**60**) and glycyl-N,O'-dimethylhydroxylamide hydrochloride (**61**) in a yields of 99% and 93% respectively. Both $^{1}H$- and $^{13}C$ NMR spectra, showed the complete loss of the t-butyl signal and the products of the hydrolysis were water soluble.

The $^{1}H$ and $^{13}C$-NMR spectra of t-Boc-Gly-N-methyl,N-ethylamide (**58**), t-Boc-Gly-N,O'-dimethylhydroxylamide (**59**), Gly-N-methyl,N-ethylamide, (**60**) and Gly-N,O'-dimethylhydroxylamide hydrochloride (**61**) showed that these compounds were present as *cis* and *trans* isomers (the alkyl amide signals and the α-proton / α-carbon signals were doubled). This is ascribed to the restricted rotation about the carbon-nitrogen bond of the amide which is very slow on the NMR time scale. The activation energy for rotation, which causes the loss of amide resonance is about 80 kJ $mol^{-1}$ and therefore both environments are observed by NMR spectroscopy (Fig. 2.2).




**Reagents**
**i.** N-methyl morpholine (NMM), i-BCF in THF at -15°C, 5 min. R = R' = Me; **ii.** $NHMe_2$ (aq) at -15°C, 5 min → r.t. 2 hours → **(57)**. R = Et, R' = Me; **ii.** NHMeEt.HCl, NMM in at -15°C, 5 min → r.t. 2 hours → **(58)**. R = Me, R' = OMe; **ii.** NHMe(OMe).HCl, NMM in at -15°C, 5 min → r.t. 2 hours → **(59)**. **iii.** Ethyl acetate saturated with HCl, -5°C, 1 hour; R = R' = Me → **(55)**, R = Et, R' = Me → **(60)** and R = Me, R' = OMe → **(61)**.

Scheme 2.7: *The preparation of various glycyl amides.*

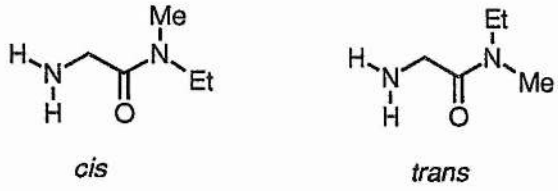
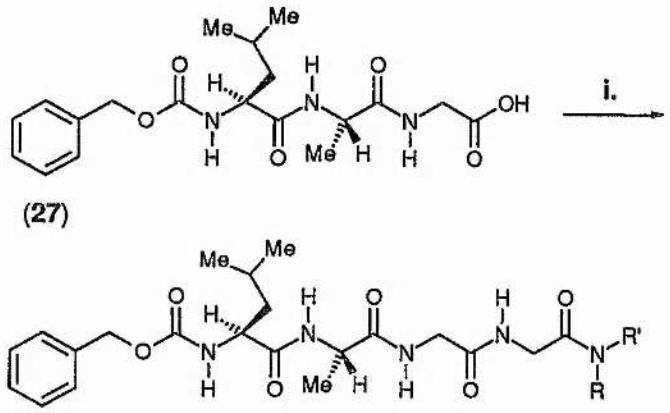



Fig. 2.2: *Cis and trans glycyl-N-methyl, N-ethyl amide, (2.36).*

### 2.1.0.0.3 Synthesis of the N-protected and unprotected tetrapeptide amides

The mixed anhydride procedure[95] failed to couple Cbz-Leu-Ala-Gly (27) (Section 2.1.0.0.0) to the glycyl amides (Section 2.1.0.0.2) in several instances, due to the low solubility of the reactants in THF at -20°C. So the carboxylic acid of Cbz-Leu-Ala-Gly (27) was activated with O-benzotriazole-1-yl-N, N, N', N'-tetramethyluronium tetrafluoroborate (TBTU)[111] in THF at 20°C, and then through reaction with the appropriate glycyl amide and diisopropylethylamine over night gave the following Cbz-protected tetrapeptide amides (Scheme 2.8 and Table 2.1), namely carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycyl-

amide (**19**), carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycyl-N-methyl-amide (**20**), carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl-N, N-dimethyl-amide (**21**), carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycyl-N-methyl N-ethyl amide (**22**) and carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl-N-, O'-dimethylhydroxylamide (**23**). Yields and melting points are given in Table 2.1 and Fig. 2.1. Full preparative details and spectral data are given in the experimental Section, pages 157-241.




**Reagents**

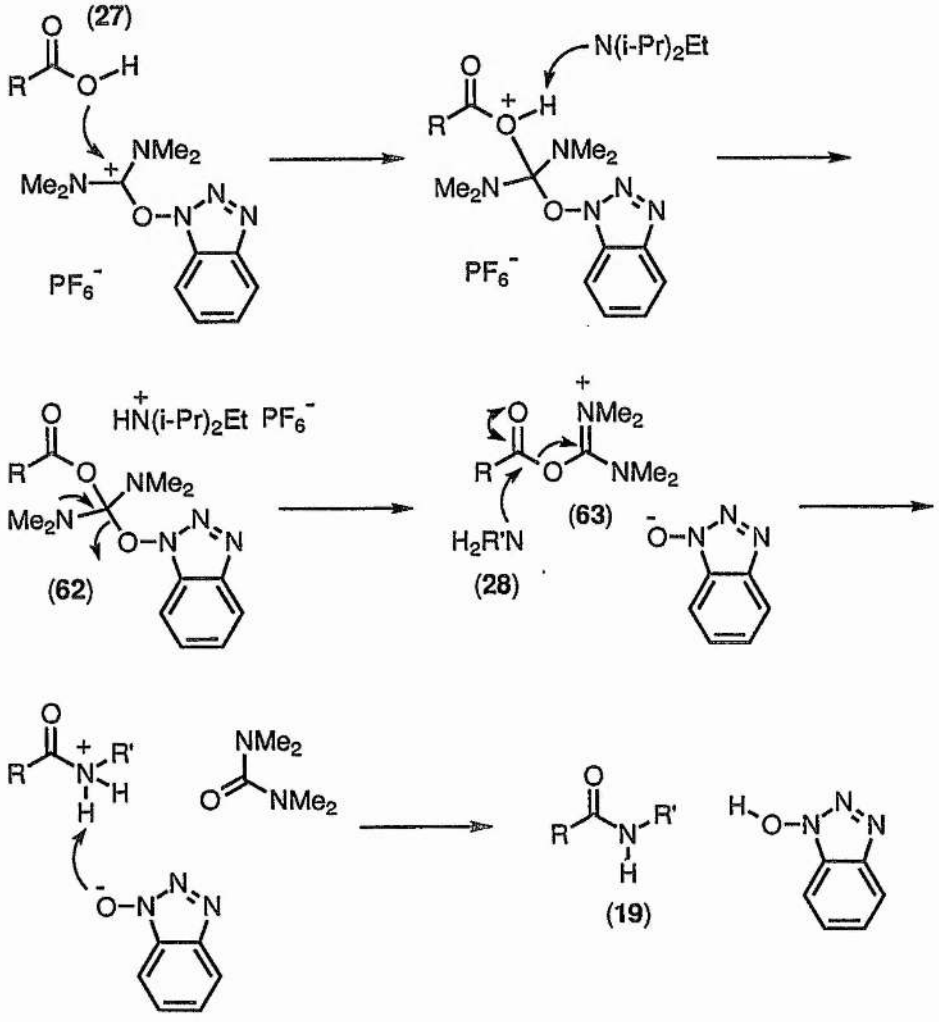
i.) TBTU, 2 eq. i-$Pr_2EtN$ and Gly-NRR'.HCl in THF, 12 hours.

Scheme 2.8: *The preparation of the Cbz-protected tetrapeptide amides.*

| Compound | R | R' | % yield | M.p / °C |
|---|---|---|---|---|
| **19** | H | H | 49 | 196-198 |
| **20** | H | Me | 68 | 213 |
| **21** | Me | Me | 34 | 163 |
| **22** | Me | Et | 64 | 170 |
| **23** | OMe | Me | 63 | 141-142 |

Table 2.1: *The percentage yields and melting points of the Cbz-protected tetrapeptide amides.*

The TBTU mechanism of amide bond formation is shown is Scheme 2.9. The carboxylic acid moiety of Cbz-Leu-Ala-Gly (**27**) attacks the TBTU to give the tetrahedral intermediate (**62**), after deprotonation. Elimination of hydroxybenzotriazole anion results in the resonance stabilised intermediate (**63**), which is nucleophilically atttacked by the amine functionality of the amide (in this case Gly-amide) to give N, N, N, 'N'-tetramethylurea, hydroxybenzotriazole and the amide bond after deprotonation.




Scheme 2.9: *The TBTU mechanism of amide bond formation.*

The carboxylic acid of t-Boc-Leu-Ala-Gly (**46**) (Section 2.1.0.0.1) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with glycyl N-*p*-nitro anilide hydrochloride (**64**) and diisopropylethylamine over night gave t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycyl glycyl-*p*-nitroanilide (**24**) in 83% yield after recrystallisation from methanol / ether (M.p. 188°C).

Unfortunately the Cbz-Leu-Ala-Gly-Gly-$NH_2$ (**19**) and the Cbz-Leu-Ala-Gly-Gly-NHMe (**20**) were too insoluble in water to be tested for biological activity. Therefore the N-protection was removed to make the amides water soluble. Hydrogenolysis of the Cbz-Leu-Ala-Gly-Gly-$NH_2$ (**19**) and the Cbz-Leu-Ala-Gly-Gly-NHMe (**20**) in acetic acid for 3 hours gave (2S)-leucyl (2S)-alanyl glycyl glycyl amide, acetate salt (**25**) (M.p. 126°C (decomp)) and (2S)-leucyl (2S)-alanyl glycyl glycyl-N-methyl amide, acetate salt (**26**) (M.p. 130°C) in 92% and 94% yields respectively. These amides were isolated as their acetate salts and their structures were confirmed by $^{1}$H-, $^{13}$C-NMR and IR spectroscopy and mass spectrometry.

## 2.1.1 Rationale for the ester tetrapeptides

Although esters are unable to provide a potentially key H-bonding donor, some proteinases are able to process ester substrates, for example chymotrypsin.[112] In the ester moiety, rotation about the carbon alkyl oxygen bond is not restricted as little delocalisation of the alkyl oxygen lone pair into the carbonyl bond occurs (unlike the carbon nitrogen bond in an amide, see Fig. 2.3). Thus the ester moiety is able to adopt more conformations than an amide bond.

It seemed reasonable to expect that the proteinase would bind to the tetrapeptide esters (Fig. 2.4 and Table 2.2), but whether the esters would act as substrates or inhibitors was unknown. In order to test this possibility some esters were prepared.

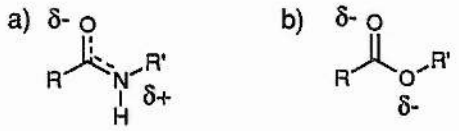



Fig. 2.3: *Comparison of a) The amide moiety and b) The ester moiety.*

## 2.1.1.0. Synthesis of the tetrapeptide esters

The carboxylic acid of Cbz-Leu-Ala-Gly (**27**) (section 2.1.0.0.1) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with the appropriate glycyl ester and diisopropylethylamine overnight gave the following Cbz-protected tetrapeptide esters; methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**), ethyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**66**) and t-butyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycinate (**67**). Yields and melting points are given in Table 2.2 and Fig. 2.4 on page 61, full preparative details and spectral data are given in the experimental Section, pages 157 - 241.

Initially a convergent strategy (Section 2.1.0, Scheme 2.2) was employed to prepare Cbz-Leu-Ala-Gly-Gly-OMe and (**65**) and Cbz-Leu-Ala-Gly-Gly-OEt (**66**); Carbobenzyloxy-glycine (**34**)[113] and methyl glycinate hydrochloride (**54**)[114] were coupled using the mixed anhydride procedure[95] to give methyl carbobenzyloxy-glycyl glycinate (**68**) in 61% yield after recrystallisation from ether / petroleum ether (M.p. 63°C; lit.,[115] 63-65°C) (Scheme 2.10). A similar procedure was used to prepare ethyl carbobenzyloxy-glycyl glycinate (**69**) from Cbz-glycine (**34**) and ethyl glycinate hydrochloride (**40**).[116] The compound was afforded in 59% yield after recrystallisation from ether / petroleum ether (M.p. 78°C; lit.,[117] 78-80°C). Removal of the Cbz-protection from Cbz-Gly-Gly-OMe (**68**) was achieved by catalytic hydrogenolysis in methanol for four hours to give methyl glycyl glycinate (**70**) in 81% yield

(M.p. (decomp.) 209°C) (Scheme 2.10). A similar procedure was used to prepare ethyl glycyl glycinate (**71**) from ethyl carbobenzyloxy-glycyl glycinate (**69**) and the product was obtained in 81% yield (M.p. 80°C) (Scheme 2.10). Unfortunately Gly-Gly-OMe (**70**) and Gly-Gly-OEt (**71**) were too insoluble in THF, dichloromethane or dimethylformamide to be coupled to Cbz-Leu-Ala (**29**).

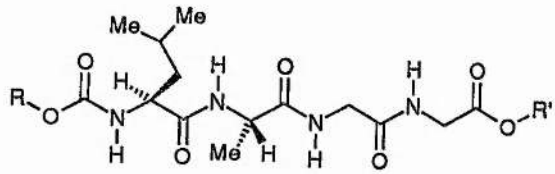
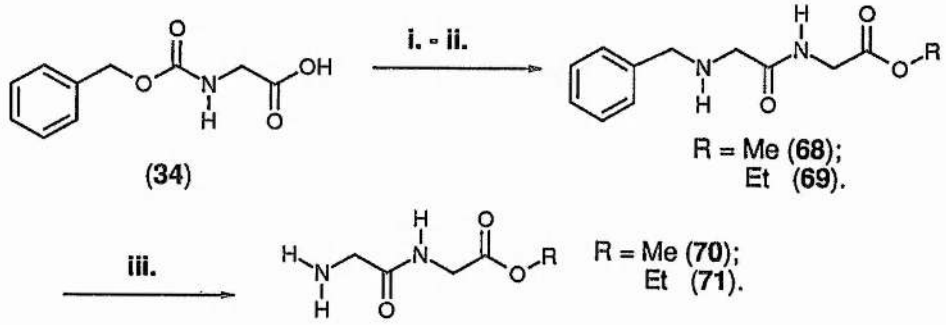


Fig. 2.4: *The N-protected tetrapeptide esters.*

| Compound | R | R' | % yield | M.p. / °C |
|---|---|---|---|---|
| **65** | Cbz | Me | 62 | 145-146 |
| **66** | Cbz | Et | 63 | 142-144 |
| **67** | Cbz | t-Bu | 72 | hygroscopic |
| **72** | t-Boc | Bz | 65 | hygroscopic |
| **73** | t-Boc | t-Bu | 75 | 153 |

Table 2.2:*The percentage yields and melting points of the N-protected tetrapeptide esters.*




**Reagents**
i. NMM, i-BCF in THF at -15°C, 5 min; R = Me; ii. NMM and Gly methyl ester HCl in THF at -15°C, 5 min → r.t., 2 hours. R = Et; ii. NMM and Gly ethyl ester HCl in THF at -15°C, 5 min → r.t., 2 hours. iii. $H_2$ Pd \ C 3 hours.

Scheme 2.10: *The preparation of Gly-Gly-OMe (70) and Gly-Gly-OEt (71).*

The carboxylic acid of t-Boc-Leu-Ala-Gly (**46**) (section 2.1.0.0.1) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with the appropriate glycyl ester and diisopropylethylamine over night gave the following t-Boc-protected tetrapeptide esters; benzyl t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycyl glycinate (**72**) and t-butyl t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycyl glycinate (**73**), see Fig. 2.4 and Table 2.2. All of the structures of the N-protected tetrapeptide esters (**65**) to (**67**) and (**72**) to (**73**) were confirmed by $^{1}$H-NMR, $^{13}$C-NMR, IR spectroscopy, mass spectrometry and microanalysis.

Ethyl (2S)-leucyl (2S)-alanyl glycyl glycinate (**74**) was prepared through catalytic hydrogenolysis of Cbz-Leu-Ala-Gly-Gly-OEt (**66**) in quantitative yield.

## 2.1.2 Rationale for the acid and alcohol tetrapeptides

N-(Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl)-1-amino, 2-hydroxy ethane (**75**) has the ability to accept hydrogen bonds from the $P_1$ carbonyl binding pocket or the $P_1$' amide binding pocket in the proteinase, depending on where the alcohol moiety binds to the proteinase (Fig. 2.5). Therefore, it was expected that inhibition of the proteinase by N-(Cbz-Leu-Ala-Gly)-1-amino-2-hydroxyethane (**75**) might indicate that hydrogen bonding interactions between the proteinase and the carbonyl oxygen at $P_1$ or the amide nitrogen at $P_1$' of a substrate are important.

Cbz-(2S)-leucyl (2S)-alanyl glycyl glycine (**76**) is the hypothetical acid hydrolysis product of the amide (section 2.1.0) and ester (Section 2.1.1) potential substrates. As Cbz-Leu-Ala-Gly-Gly (**76**) has similar binding requirements to the potential substrates, product inhibition of the proteinase may occur. In the case of papain, benzoyl glycine is a competitive inhibitor ($K_i$ 830 mmol dm$^{-3}$).[118]

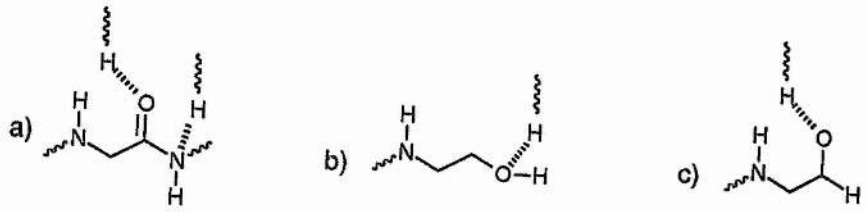



Fig. 2.5: *The possible interactions of Cbz-Leu-Ala-Gly-ethanolamine (75) with the $P_1$-$P_1$' binding pockets of the proteinase; a) substrate; b) alcohol binding in $P_1$' amide nitrogen pocket; c) alcohol binding in $P_1$ carbonyl oxygen pocket.*

## *2.1.2 Synthesis of the tetrapeptide alcohol and acid*

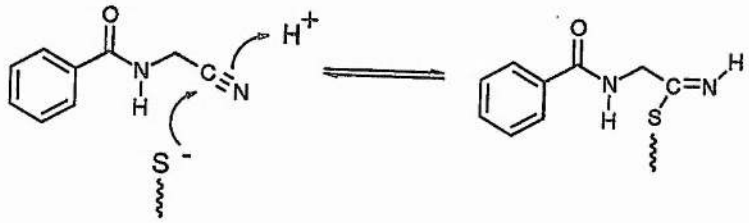
The carboxylic acid of carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycine (**27**) (section 2.1.0.0.1) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with ethanolamine and diisopropylethylamine overnight gave N-(Cbz-Leu-Ala-Gly)-2-amino-ethanol (**75**).

Sodium Cbz-Leu-Ala-Gly-Gly (**76**) was prepared by the base catalysed hydrolysis of Cbz-Leu-Ala-Gly-Gly-OMe (**65**) in methanol for 4.5 hours at 35°C. Both N-(Cbz-Leu-Ala-Gly)-1-amino 2-hydroxy ethane (**75**) and Cbz-Leu-Ala-Gly-Gly (**76**) showed the expected signals in their IR, $^1$H- and $^{13}$C-NMR spectra.

## 2.1.3 Rationale for the design of the nitrile potential inhibitors

N-acetyl-phenylalanyl aminoacetonitrile is a competitive inhibitor of papain ($K_i$ 0.73μM).[119] The nitrile functionality binds covalently to the active site of papain, forming a thioimidate adduct (Scheme 2.11). The presence of the thioimidate adduct was demonstrated in the $^{13}$C-NMR spectra of [1-$^{13}$C]-(N-benzoyl)-aminoacetonitrile (77).[120] When a 50% molar excess of the nitrile (77) was added to papain a major new signal appeared at 182 ppm, with

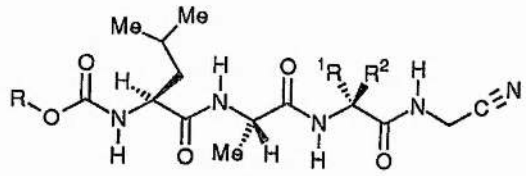
a corresponding loss in intensity of the nitrile signal at 118 ppm. The new signal's chemical shift is entirely consistent with the formation of the proposed thioimidate adduct (Fig. 2.6), since the signal falls between the range of thioamide carbons (200-210 ppm) and peptide carbons (160-170 ppm).[120] No evidence of nitrile hydrolysis by papain was observed.[121]



Scheme 2.11: *Formation of the thioimidate adduct from a nitrile in papains active site.*

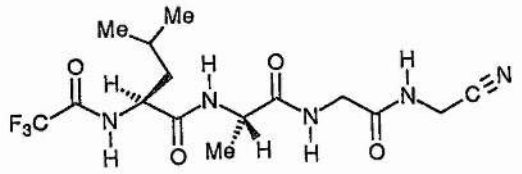
The nitrile functionality was included into the Leu-Ala-Gly-Gly sequence to give carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl aminoacetonitrile (**78**) and t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycyl aminoacetonitrile (**79**) (Fig 2.6 on page 65), in order to see if the adenovirus proteinase could be inhibited by the nitrile functionality. Carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2S)-alanyl aminoacetonitrile (**80**) and carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2R)-alanyl aminoacetonitrile (**81**) (Fig. 2.6) were also prepared to investigate whether the proteinase requires a glycyl residue at $P_2$ of the cyanide inhibitors.

Once it was established that Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) and t-Boc-Leu-Ala-Gly-aminoacetonitrile (**79**) inhibited the adenovirus proteinase, it was necessary to probe the binding interactions present between the P side of the tetrapeptide cyanide inhibitors and the S side of the proteinase. The series of trifluoroacetyl protected cyanides (Fig. 2.7 on page 65) were prepared to investigate the importance of the $P_4$, $P_3$ and $P_2$ residues in the binding of the tetrapeptide inhibitors.

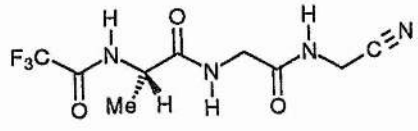


| Compound | R | $R^1$ | $R^2$ |
|---|---|---|---|
| **78** | Cbz | H | H |
| **79** | t-Boc | H | H |
| **80** | Cbz | H | Me |
| **82** | Cbz | Me | H |

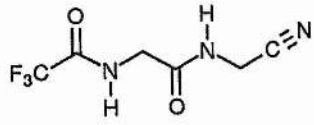
Fig. 2.6: *The Cbz- and t-Boc protected tetrapeptide nitrile inhibitors.*



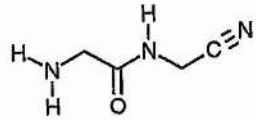
**(82)**



**(83)**



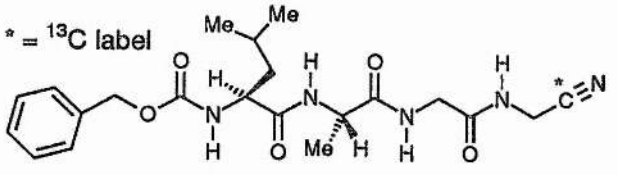
**(84)**



**(85)**

Fig 2.7: *The TFA-protected nitrile inhibitors.*

[1-$^{13}$C]-Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl aminoacetonitrile (**86**) (Fig 2.8) was prepared to establish whether Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) inhibited the adenovirus proteinase by forming a thioimidate adduct. Potentially, this could be achieved by examining mixtures of the activated proteinase and the $^{13}$C-labelled inhibitor by $^{13}$C-NMR spectroscopy.




**(86)**

Fig. 2.8: *The $^{13}$C-labelled cyanide inhibitor.*

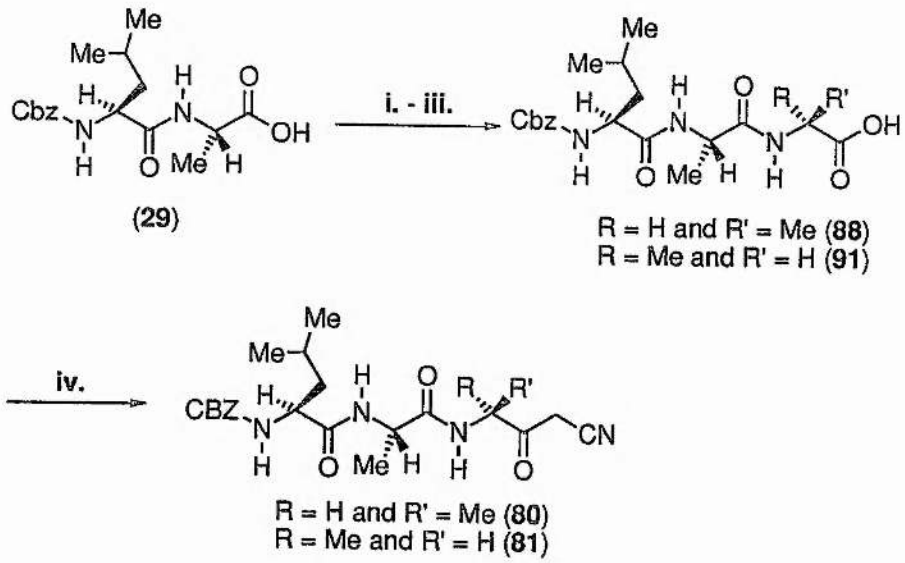
## *2.1.3.0 The synthesis of the potential nitrile inhibitors*

### 2.1.3.0.0 The synthesis of the Cbz- and t-Boc protected nitriles

The carboxylic acid of Cbz-Leu-Ala-Gly (**27**) (section 2.1.0.0.1) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with aminoacetonitrile hydrochloride (**85**) and diisopropylethylamine overnight afforded Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) in 78% yield after recrystallisation from methanol / ether (M.p. 159°C). Similarly t-Boc-Leu-Ala-Gly-aminoacetonitrile (**79**) was prepared from t-Boc-Leu-Ala-Gly (**46**) (section 2.1.0.0.1) and aminoacetonitrile hydrochloride (**85**) in 57% yield after recrystallisation from methanol / ether (M.p 162°C).

Carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2S)-alanyl aminoacetonitrile (**80**) was prepared as outlined below; Cbz-Leu-Ala (**29**) (Section 2.1.0.0.0) was coupled with methyl (2S)-alaninate hydrochloride (**33**) using the mixed anhydride procedure[95] to give methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2S)-alaninate (**87**) in 87% yield after

recrystallisation from acetone / water (M.p. 149-150°C; lit.,[122] 172-173°C) (Scheme 2.12). Alkaline hydrolysis of Cbz-Leu-Ala-Ala-OMe (**87**) in methanol gave carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2S)-alanine (**88**) in 71% yield after recrystallisation from methanol / ether. (M.p. 84-86°C). The $^{1}$H- and $^{13}$C-NMR spectra showed the complete disappearance of the methyl ester signal (Scheme 2.12). The carboxylic acid of Cbz-Leu-Ala-Ala (**88**) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with aminoacetonitrile hydrochloride (**85**) and diisopropylethylamine overnight gave carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2S)-alanyl aminoacetonitrile (**80**) in 66% yield after recrystallisation from methanol / ether (M.p. 161-162°C) (Scheme 2.12).




**Reagents**
**i.** NMM, i-BCF in THF at -15°C, 5 min. R = H and R' = Me; **ii.** NMM, Ala-Me ester. HCl (**33**), at -15°C,5 min → r.t., 2 hours → (**87**). **iii.** NaOH in MeOH at 35°C, 4 1/2 hours; then HCl pH 3 → (**88**). R = Me and R' = H; **ii.** NMM, (2R)-Ala-t-Bu ester. HOAc (**89**) at -15°C, 5 min → r.t., 2 hours → (**90**). **iii.** formic acid, 3 hours → (**91**). **iv.** i-$Pr_2$EtN, TBTU, aminoacetonitrile hydrochloride in THF, 12 hours.

Scheme 2.12: *The preparation of the nitriles containing alanine at the $P_2$ position.*

Carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2R)-alanyl aminoacetonitrile (**81**) was prepared as follows; Cbz-Leu-Ala (**29**) (Section 2.1.0.0.0) was coupled with t-butyl (2R)-alaninate hydrochloride (**89**) using the mixed anhydride procedure[95] to afford t-butyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2R)-alaninate (**90**) in 79% yield after recrystallisation from acetone / water (M.p. 121-122°C) (Scheme 2.12). Acidic hydrolysis of Cbz-Leu-Ala-(2R)-Ala-Ot-Bu (**90**) gave carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2R)-alanine (**91**) in 58% yield after recrystallisation from methanol / ether (M.p. 82-84°C). The $^{1}$H- and $^{13}$C-NMR spectra showed the complete disappearance of the t-butyl ester signal (Scheme 2.12). The carboxylic acid of Cbz-Leu-Ala-(2R)-Ala (**91**) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with aminoacetonitrile hydrochloride (**85**) overnight gave carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2R)-alanyl aminoacetonitrile (**81**) in 66% yield after recrystallisation from methanol / ether (M.p. 140°C) (Scheme 2.12).

All the structures of the N-protected tetrapeptide cyanides were confirmed by $^{1}$H-NMR, $^{13}$C-NMR and IR spectroscopy, microanalysis and mass spectrometry.

### 2.1.3.0.1 The synthesis of the TFA protected nitriles

Trifluoroacetyl-(2S)-leucine (**92**)[123] was coupled with benzyl (2S)-alaninate tosylate (**93**) using the mixed anhydride procedure[95] to give benzyl trifluoroacetyl-(2S)-leucyl (2S)-alaninate (**94**) in 77% yield (Scheme 2.13 on page 70). Treatment of TFA-Leu-Ala-OBz (**94**) with palladium (II) acetate, dimethylethylsilane and triethylamine[106] and then methanol gave trifluoroacetyl-(2S)-leucyl (2S)-alanine (**95**) in 83% yield after recrystallisation from dichloromethane / hexane (M.p. 92°C) (Scheme 2.13). Both the $^{1}$H- and $^{13}$C-NMR spectra showed the complete disappearence of the benzyl ester signals. The $^{1}$H- and $^{13}$C-NMR spectra also showed that TFA-Leu-Ala (**95**) was present in two

conformations, as most of signals for the leucine and the alanine residues were doubled. The spectra were complex at room temperature but at -50°C the signals of the two isomers could be distinguished.

Molecular dynamics at elevated temperature (500 K),[a] indicated that two different backbone conformations of the leucyl residue were possible. A 7-membered ring is formed from a hydrogen bond between the nitrogen hydrogen of the alanine residue and the carbonyl oxygen of the trifluoroacetyl group where the leucyl side chain is either equatorial ($C7_{eq}$ comformation, Fig. 2.9) or axial ($C7_{ax}$ conformation Fig. 2.9). Samples taken from the dynamics every ps were reoptimised using the AMBER force field and showed that the $C7_{eq}$ conformation has an energy range between -20.9 to -13.8 kJ $mol^{-1}$ and the the $C7_{ax}$ conformation has an energy range between -13.0 to -11.3 kJ $mol^{-1}$. There is a low energy difference between the two conformers (5.3 ± 4.5 kJ $mol^{-1}$) so it is reasonable to assume that at room temperature, both conformations would be populated.

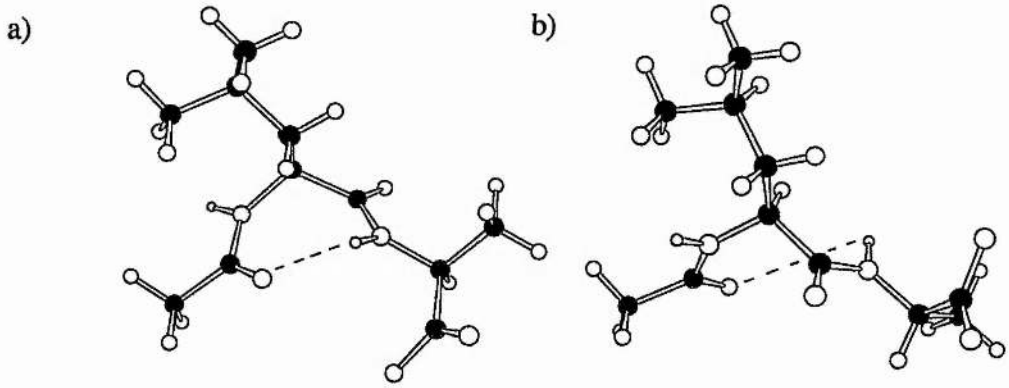
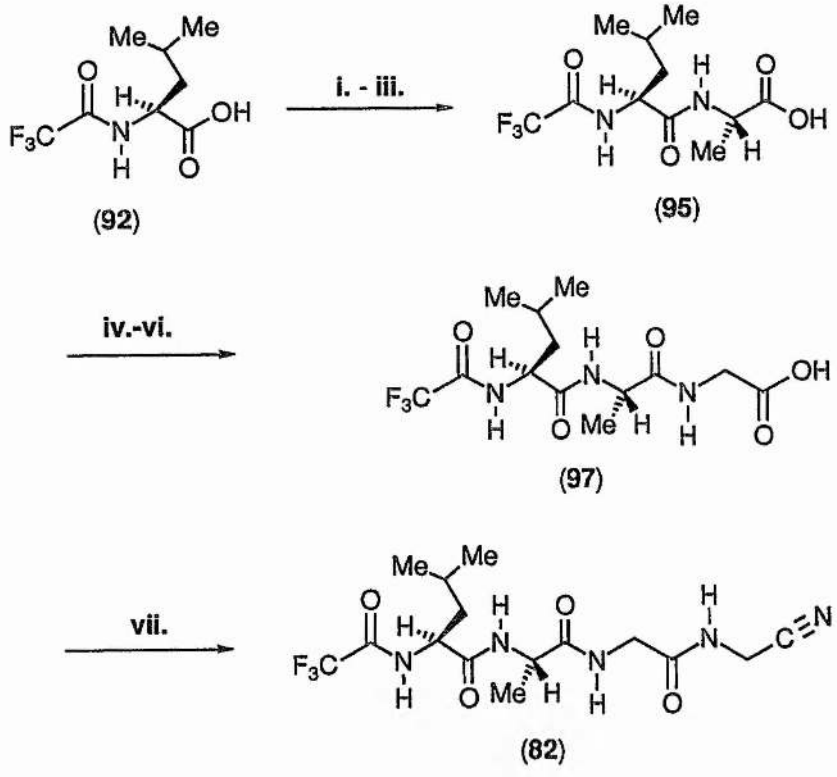



Fig. 2.9: *The two possible isomers of TFA-Leu-Ala (95);*

*a) The $C7_{eq}$ conformation and b) The $C7_{ax}$ conformation.*

[a] Molecular mechanics energies are relative energies compared with an arbitaey zero. These calculations were performed by Dr J. Wilkie.

Further extension of TFA-Leu-Ala (**95**) was achieved using the mixed anhydride method[95] and glycine benzyl ester tosylate (**44**) to give the monohydrate of benzyl trifluoroacetyl-(2S)-leucyl (2S)-alanyl glycinate (**96**) in 88% yield after recrystallisation from acetone / water (M.p. 121°C) (Scheme 2.13). In the $^{1}$H- and $^{13}$C-NMR spectra the same peak doubling as occured for TFA-Leu-Ala (**95**), also occurred for TFA-Leu-Ala-Gly-OBz (**96**).




(92) (95) (97) (82)

**Reagents**
**i.** NMM, i-BCF in THF at -15°C, 5 min; **ii.** Then NMM and Ala Bz ester. tosylate (**93**) in THF at -15°C, 5 min → r.t., 2 hours → (**94**). **iii.** $Pd(OAc)_2$, $Et_3N$, $Me_2EtSiH$ in dichloromethane 5 min, then compound (**94**), 12 hours; then MeOH, 3 hours. **iv.** NMM, i-BCF in THF at -15°C, 5 min; **v.**Then NMM and Gly Bz ester in THF at -15°C, 5 min → r.t., 2 hours → (**96**). **vi.** $Pd(OAc)_2$, $Et_3N$, $Me_2EtSiH$ in dichloromethane 5 min, then compound (**96**), 12 hours; then MeOH, 3 hours. **vii.** TBTU, i-$Pr_2EtN$ and aminoacetonitrile hydrochloride in THF, 12 hours.

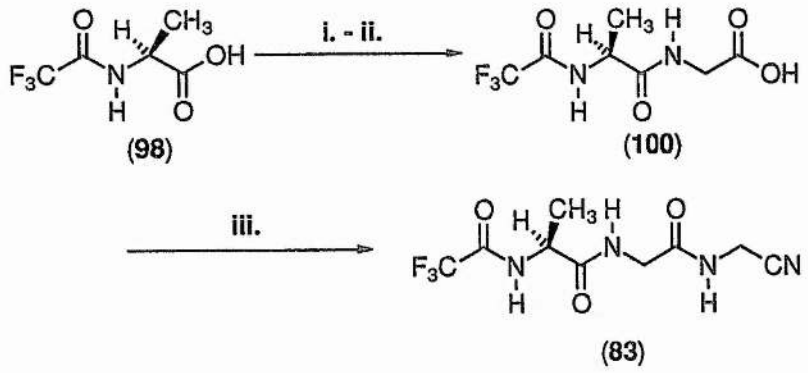
Scheme 2.13: *The synthesis of TFA-(2S)-leucyl (2S)-alanyl glycyl aminoacetonitrile (**82**).*

The benzyl ester was removed from benzyl TFA-leu-Ala-Gly (**96**) by treatment with palladium (II) acetate, dimethylethylsilane, triethylamine[106] and then methanol to give trifluoroacetyl-(2S)-leucyl (2S)-alanyl glycine (**97**) in 46% yield (Scheme 2.13 on page 70). Both $^{1}$H- and $^{13}$C-NMR spectra showed the complete disappearence of the benzyl ester signals.

The carboxylic acid of TFA-Leu-Ala-Gly (**97**) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with aminoacetonitrile hydrochloride (**85**) and diisopropylethylamine overnight to give trifluoroacetyl-(2S)-leucyl (2S)-alanyl glycyl aminoacetonitrile (**82**) in 82% yield after chromatography on silica eluting with dichloromethane / petroleum ether (3:1), then ethyl acetate (Scheme 2.13).

The structures of the TFA-protected compounds (**94**) to (**97**) and (**82**) were also confirmed by IR and mass spectrometry.

Trifluoroacetyl-(2S)-alanyl glycyl aminoacetonitrile (**83**) was prepared as described below; trifluoroacetyl alanine (**98**)[123] was coupled with benzyl glycinate tosylate (**44**) using the mixed anhydride procedure[95] to give benzyl trifluoroacetyl-(2S)-alanyl glycinate (**99**) in 96% yield after recrystallisation from dichloromethane / hexane (M.p. 84-86°C) (Scheme 2.14). Treatment of TFA-Ala-Gly-OBz (**99**) with palladium (II) acetate, dimethylethylsilane and triethylamine;[106] then methanol gave trifluoroacetyl-(2S)-alanyl glycine (**100**) in 63% yield (Scheme 2.14). Both $^{1}$H- and $^{13}$C-NMR spectra showed the complete disappearence of the benzyl ester signals. The carboxylic acid of TFA-Ala-Gly (**100**) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with aminoacetonitrile hydrochloride (**85**) overnight gave trifluoroacetyl-(2S)-alanyl glycyl aminoacetonitrile (**83**) in 35% yield after recrystallisation from dichloromethane / hexane (M.p. 149°C) (Scheme 2.14).

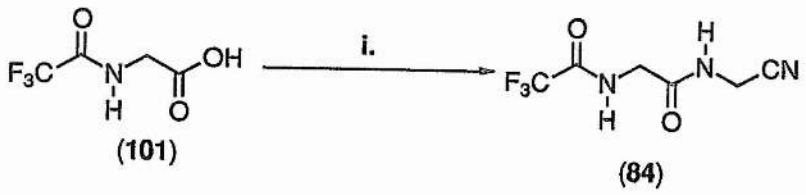



**Reagents**

**i.** NMM, i-BCF in THF at -15°C, 5 min; then NMM and Gly Bz ester. Tosylate (**44**) in THF at -15°C, 5 min → r.t., 2 hours → (**99**). **ii.** $Pd(OAc)_2$, $Et_3N$, $Me_2EtSiH$ in dichloromethane 5 min, then compound (**99**), 12 hours; then MeOH, 3 hours. **iii.** TBTU, i-$Pr_2EtN$ and aminoacetonitrile hydrochloride in THF, 12 hours.

Scheme 2.14: *The preparation of TFA-Ala-Gly-aminoacetonitrile (83).*

Trifluoroacetyl-glycyl aminoacetonitrile (**84**) was prepared by activating the carboxylic acid of TFA-Gly (**101**)[124] with TBTU[111] in THF at 20°C. Reaction of the activated ester with aminoacetonitrile hydrochloride (**85**) and diisopropylamine overnight gave trifluoroacetyl-glycyl aminoacetonitrile (**84**) in 57% yield after recrystallisation from methanol (M.p. 143°C) (Scheme 2.15). The structures of the TFA-protected compounds (**99**) to (**101**), (**83**) and (**84**) were also confirmed by IR and mass spectrometry.




**Reagents**

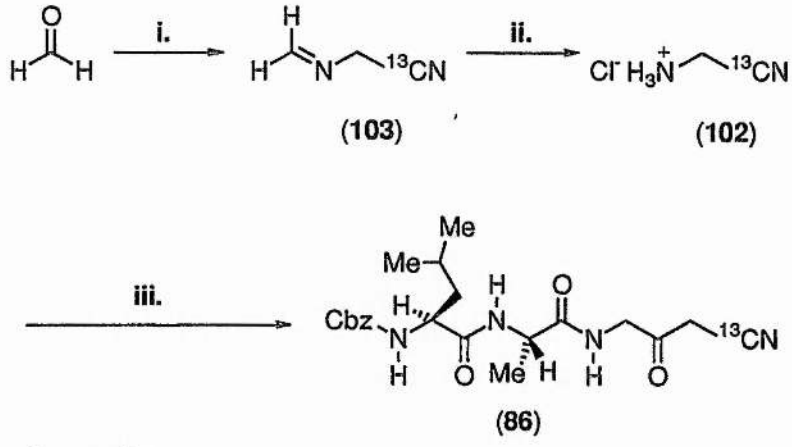
**i.** TBTU, i-$Pr_2EtN$ and aminoacetonitrile hydrochloride in THF, 12 hours.

Scheme 2.15: *The preparation of TFA-Gly-aminoacetonitrile (84).*

### 2.1.3.0.2 The synthesis of [1-$^{13}$C]-carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl aminoacetonitrile (86)

[1-$^{13}$C]-Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl aminoacetonitrile (**86**) was prepared to establish whether Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) inhibited the adenovirus proteinase by forming a hemithioimidate adduct, see Section 2.1.3 page 66.

[1-$^{13}$C]-Aminoacetonitrile (**102**) was prepared as described below; formaldehyde and ammonium chloride were carefully treated with [1-$^{13}$C]-potassium cyanide and acetic acid[125] to give [1-$^{13}$C]-methylene aminoacetonitrile (**103**) in 15% yield (Scheme 2.15). [1-$^{13}$C]-methylene aminoacetonitrile (**103**) was hydrolysed with ethanolic hydrochloric acid to give [1-$^{13}$C]-aminoacetonitrile hydrochloride (**102**) in 96% yield (Scheme 2.15). The procedures had been optimised first with unlabelled potassium cyanide to give methylene aminoacetonitrile (**104**) and aminoacetonitrile hydrochloride (**85**) in similar yields of 24% and 13% respectively.




**Reagents**
**i.** $K^{13}CN$, 2 $NH_4Cl$ in $H_2O$ at -5°C, then HOAc in $H_2O$ -5°C, 1 hour. **ii.** HCl in EtOH at -5°C, 1hour. **iii.** Cbz-Leu-Ala-Gly (**27**) TBTU and i-$Pr_2EtN$ in THF, 12 hours.

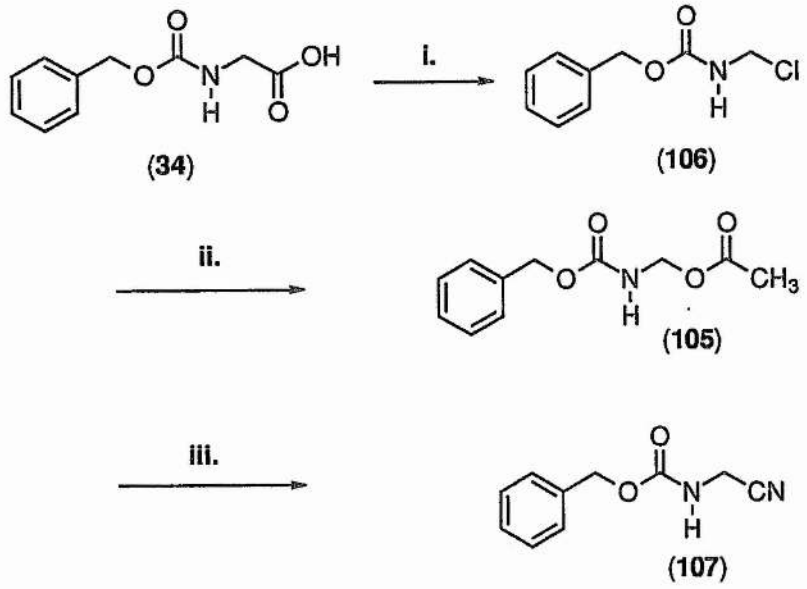
Scheme 2.16: *The preparation of [1-$^{13}$C]-Cbz-Leu-Ala-Gly-aminoacetonitrile (86).*

The carboxylic acid of Cbz-Leu-Ala-Gly (**27**) (section 2.1.0.0.1) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with [1-$^{13}$C]-aminoacetonitrile hydrochloride (**102**) and diisopropylethylamine overnight to give [1-$^{13}$C]-carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl aminoacetonitrile (**86**) in 90% yield after recrystallisation from methanol / ether (Scheme 2.16).

Initially the synthesis of aminoacetonitrile (**85**) was attempted from Cbz-Gly (**34**) by the use of halodecarboxylation. (Scheme 2.17 on page 75). This was in order to avoid the use of potassium cyanide under acidic conditions which liberates hydrogen cyanide. Cbz-glycine (**34**) was halodecarboxylated with 1.1 equivalents of lead (IV) tetracetate and 1.1 equivalents of lithium chloride. [126] Carbobenzyloxy-1-amino-1-acetoxymethane (**105**) was isolated in 67% yield (Scheme 2.17). The structure was confirmed by $^{1}$H-NMR where the glycyl α-protons appeared as a doublet at 5.29 ppm with a coupling constant of 7.6 Hz to the urethane proton and the acetoxy-protons appeared as a singlet at 2.05 ppm and by $^{13}$C-NMR which showed that the benzylic carbon and the glycyl α-carbon appeared at 67.82 and 67.12 ppm and the acetoxyl carbonyl carbon and methyl carbon appeared at 172.27 and 21.46 ppm respectively. Further characterisation was not possible as Cbz-1-amino-1-acetoxymethane (**105**) was too sensitive to moisture. The expected product of carbobenzyloxy-1-amino-1-chloromethane (**106**) was not obtained as acetolysis of Cbz-1-amino 1-chloromethane (**106**) occurs due to the presence of acetic acid formed *in situ*.

Treatment of Cbz-1-amino-1-acetoxymethane (**105**) with potassium cyanide in refluxing acetonitrile for 3 hours gave carbobenzyloxy-aminoacetonitrile (**107**) in 68% yield The structure was confirmed by $^{1}$H-NMR which showed that the glycyl α-protons occurred as a doublet at 4.04 ppm with a coupling constant of 7.6 Hz to the urethane proton and that the acetoxy-protons signal had completely disappeared. Deprotection of the Cbz group was

attempted by catalytic hydrogenolysis and by treatment with trimethylsilyl bromide[127] but both methods failed due to over reduction and lack of reaction respectively.



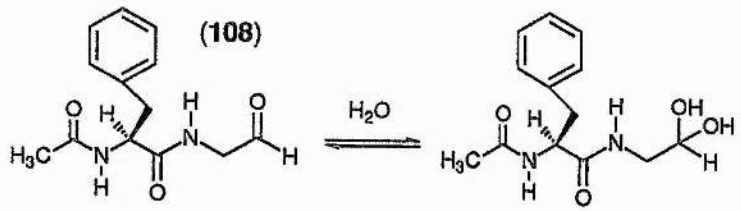

**Reagents**
**i.** $Pb(OAc)_4$ in MeCN, 10 min, then LiCl and heat until 80 -90°C, 2hours.
**ii.** AcOH *in situ.* **iii.** KCN MeCN at reflux, 3 hours.

Scheme 2.17: *The attempted preparation of aminoacetonitrile (85).*

## 2.1.4 Rationale for the aldehyde, methyl-ketones, dimethylacetal and alkenic potential inhibitors

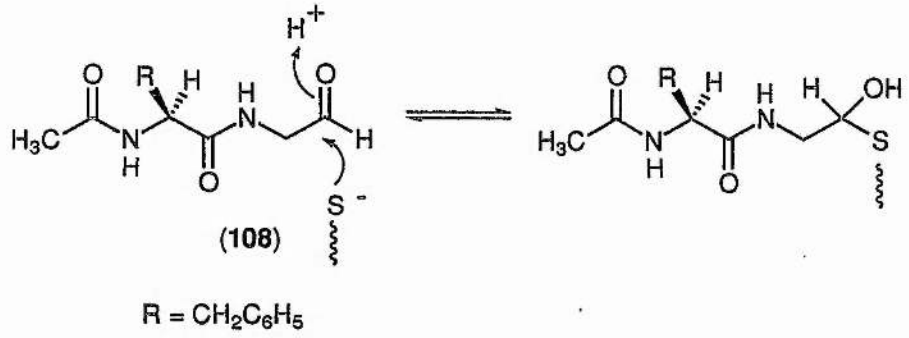
Acetyl-(2S)-phenylalanyl glycinal (**108**) is a potent competitive inhibitor of papain (5.2 nmol $dm^{-3}$).[119] 95% of the aldehyde (**108**) actually exists in the hydrated diol form (Scheme 2.18 on page 76) and it was unknown whether the aldehyde form or the hydrated diol form of the aldehyde was responsible for the inhibition of papain.[128]

It was suspected that the unhydrated, aldehyde form of acetyl phenylalanyl glycinal (**108**) bound covalently to the active site of papain, forming a hemithioacetal adduct (Scheme 2.19 on page 76).[128]




Scheme 2.18:

*The equilibrium between the aldehyde and the diol functionality of β-amido aldehydes.*



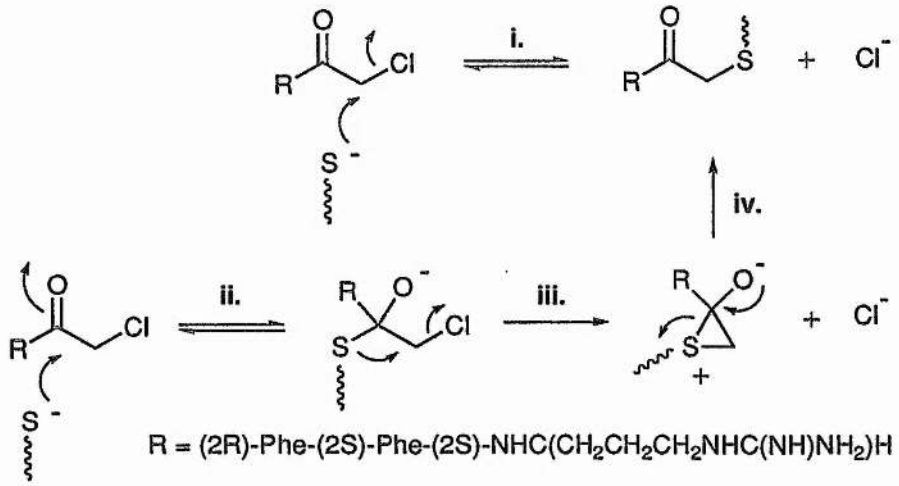

Scheme 2.19: *The formation of the hemithioacetal adduct between papain and acetyl phenylalanyl glycinal (108).*

The hemithioacetal adduct between papain and [1-$^{13}$C]-acetyl phenylalanyl glycinal (**109**) was observed from $^{13}$C-NMR spectroscopy of a mixture of [1-$^{13}$C]-acetyl phenylalanyl glycinal (**109**) and papain (both inhibitor and proteinase were at a final concentration of 70 mmol dm$^{-3}$). When the two solutions were mixed, the $^{13}$C-NMR spectrum of the $^{13}$C-labelled aldehyde (**109**) showed the presence of a major new signal at 74.9 ppm and the hydrate signal at 88.2 ppm and the aldehyde signal at 200.9 ppm had disappeared. The new resonance at 74.9 ppm was assigned as a hemithioacetal adduct, formed from papain and the aldehyde.[128] The asymmetric carbon signal of the hemithioacetal formed from acetaldehyde

and 1-mercaptoethanol occurs at 73.27 ppm, which is in close agreement to the new signal (74.9 ppm) observed when papain and [1-$^{13}$C]-acetyl-phenylalanyl glycinal (**109**) are mixed. The 13.3 ppm shift is too large to be accounted for by the diol form of [1-$^{13}$C]-acetyl phenylalanyl glycinal (**109**) binding to papain instead of hemithioacetal formation.

A methyl ketone can also inhibit a cysteine proteinase in a similar manner to an aldehyde by forming a hemithioketal adduct.[129] The methyl ketone does not usually bind as well as the aldehyde and a 100 fold fall in $K_i$ is usually observed for the methyl ketone compared to the aldehyde.[130] The electron donating capacity of the methyl group in the case of the methyl ketone reduces the electrophilicity of the carbonyl carbon compared to the aldehyde. So the thiol nucleophile does not attack the ketone carbonyl carbon as readily as the aldehyde carbonyl carbon. Therefore the equilibrium constant for the formation of the tetrahedral intermediate is smaller for the methylketone than for the aldehyde.

1-((2R)-phenylalanyl (2S)-phenylalanyl (2S)-arganyl)-chloromethane (**110**) irreversibly inhibits cathepsin B (beef spleen) ([I] = 2.5 nmol dm$^{-3}$, $t_{½}$ = 11.5 min).[131] There are two possible mechanisms for the inactivation (Scheme 2.20).




Scheme 2.20: *The possible mechanisms for the inactivation of a cysteine proteinase by a chloromethyl ketone.*

In the first mechanism the thiolate group is attacked directly by the chloromethyl group to give the acylated proteinase (step i, Scheme 2.20). In the second mechanism the thiolate group attacks the carbonyl carbon and forms a tetrahedral intermediate (step ii, Scheme 2.20). The sulphur then displaces chloride ion to give a sulphonium ion (step iii) which rearranges to form the acylated proteinase (step iv).

At the present time these possibilities have not been distinguished.[132] The second mechanism however, is more attractive as the proteinase treats the chloromethylketone as a substrate up to the inadvertent formation of the sulphonium ion.[132]

The dimethyl acetal functionality may have the capacity to mimic the amide bond, as one of the methoxy groups could replace the carbonyl group in the substrate and the other methoxy group could replace the nitrogen of the amide bond of the substrate (Fig. 2.10).

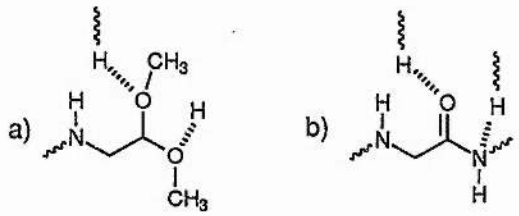



Fig. 2.10: *The possible binding interactions of the proteinase with: a) the dimethyl acetal functionality and b) the amide functionality.*

It may also be possible for carbon-carbon double bonds to mimic the amide bond and bind in the $P_1$ carbonyl binding pocket on the proteinase. Any proton donor present to polarise the carbonyl bond of the amide bond to be cleaved may also be able to polarise the carbon-carbon double bond (Fig. 2.11).

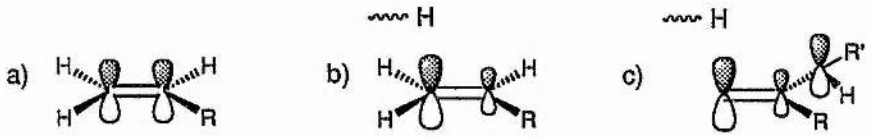
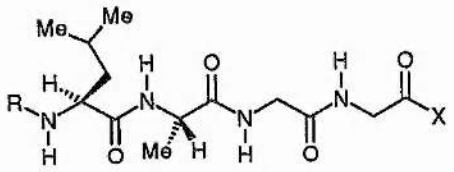



Fig. 2.11: *The possible interactions of an alkene with the active site of a proteinase; a) alkene; b) alkene in the active site and c) the amide bond in the active site.*

The aldehyde, methyl ketones, dimethyl acetal and alkene functionalities were included into the substrate recognition sequence of the adenovirus proteinase (Fig. 2.12), to give a series of potential inhibitors that would probe the electronic enviroment of the $P_1$-$P_1$' binding pocket of the proteinase.




| Compound | R | X |
|---|---|---|
| **111** | H | CHO |
| **112** | Cbz | CHO |
| **113** | t-Boc | $CH(OMe)_2$ |
| **114** | Cbz | $CH(OMe)_2$ |
| **115** | Cbz | $HC=CH_2$ |
| **116** | Cbz | COMe |
| **117** | Cbz | $COCH_2Cl$ |
| **118** | Cbz | $MeC=CH_2$ |

Fig 2.12: *The aldehyde, methyl ketones, dimethyl acetal and alkene potential inhibitors.*

### 2.1.4.0 The synthesis of the aldehyde methyl-ketones, dimethylacetal and alkenic potential inhibitors

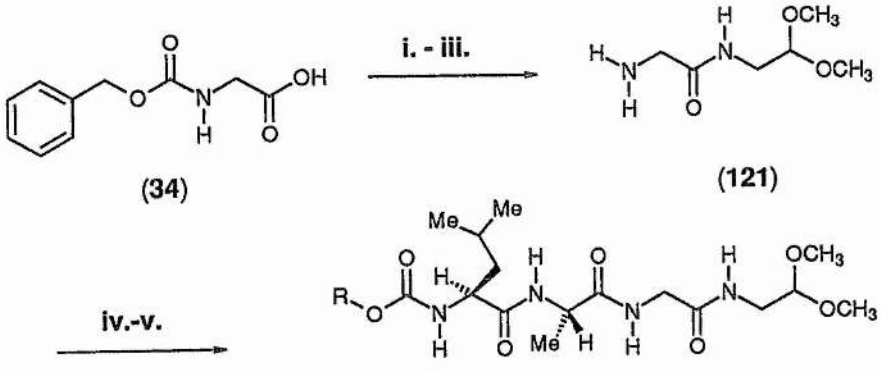
(2S)-Leucyl (2S)-alanyl glycyl aminoacetaldehyde (**111**) and carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl amino acetaldehyde (**112**) (Fig. 2.15) were initially prepared from the acidic hydrolysis of t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycyl amino acetaldehyde dimethyl acetal (**113**) and carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl amino acetaldehyde dimethyl acetal (**114**) respectively. However these methods failed due to polymerisation and dimerisation respectively (Section 2.1.4.0.0). Ozonolysis of carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl 1-amino prop-2-ene (**115**) proved to be the successful method of preparation of Cbz-Leu-Ala-Gly amino acetaldehyde (**112**) (Section 2.1.4.0.2).

Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl 3-aminopropanone (**116**) and carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl 3-amino-1-chloropropanone (**117**) were initially prepared from various diazoketones (Section 2.1.4.0.1) but these also failed due to polymerisation. Again, ozonolysis of carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl 1-amino 2-methyl prop-2-ene (**118**) proved to be the successful method of preparation of Cbz-Leu-Ala-Gly 3-amino propan-2-one (**116**) (Section 2.1.4.0.2).

#### 2.1.4.0.0 The unsucessful syntheses of tetrapeptide aldehydes

t-Butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycyl aminoacetaldehyde dimethyl acetal (**113**) was prepared as outlined below: Cbz-glycine (**34**) and aminoacetaldehyde dimethyl acetal (**119**) were coupled using the mixed anhydride procedure[95] to give carbobenzyloxy-glycyl-aminoacetaldehyde dimethyl acetal (**120**) in 88% yield (Found: $[M + H]^+$ 163.1083 $C_6H_{15}N_2O_3$ requires 163.1083) (Scheme 2.21). Hydrogenolysis of the Cbz-Gly-aminoacetaldehyde dimethyl acetal (**120**) for 3 hours gave glycyl-aminoacetaldehyde dimethyl acetal (**121**) and this was coupled to t-Boc-Leu-Ala (**43**)

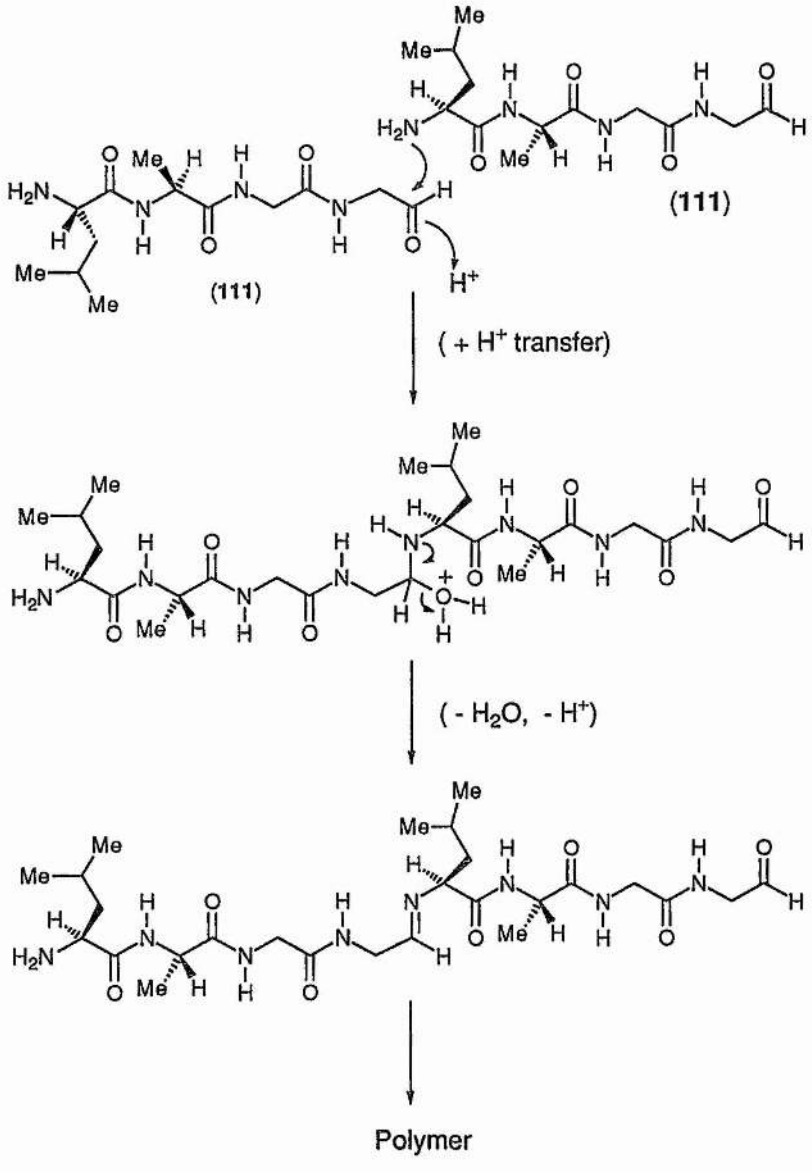
(Section 2.1.0.0.1) using the mixed anhydride method[95] to give t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycyl aminoacetaldehyde dimethyl acetal (**113**) in 83% yield after recrystallisation from methanol / ether (M.p. 156-158 C) (Scheme 2.21). Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl aminoacetaldehyde dimethyl acetal (**114**) was similarly prepared from Cbz-Leu-Ala (**29**) (Section 2.1.0.0.0) and Gly-aminoacetaldehyde dimethyl acetal (**121**) in 75% yield after recrystallisation from methanol / ether (M.p. 149-151°C) (Scheme 2.18). The assigned structures of t-Boc-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal (**113**) and Cbz-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal (**114**) were confirmed by $^{1}H$, $^{13}C$-NMR and IR spectroscopy, microanalysis and by their accurate mass and mass fragmentation patterns.




**Reagents**
**i.** NMM, i-BCF in THF at -15°C, 5 min; **ii.** aminoacetaldehyde (**119**) in THF at -15°C, 5 min → r.t., 2 hours → (**120**). **iii.** $H_2$ Pd \ C 3 hours. R = $C(CH_3)_3$; **iv.** t-Boc-Leu-Ala (**43**), NMM, i-BCF in THF at -15°C; **v.** Then (**121**) 5 min → r.t., 2 hours → (**113**). R = $CH_2C_6H_5$; **iv.**Cbz-Leu-Ala (**29**), NMM, i-BCF in THF at -15°C; **v.** Then (**121**) 5 min → r.t., 2 hours → (**114**).

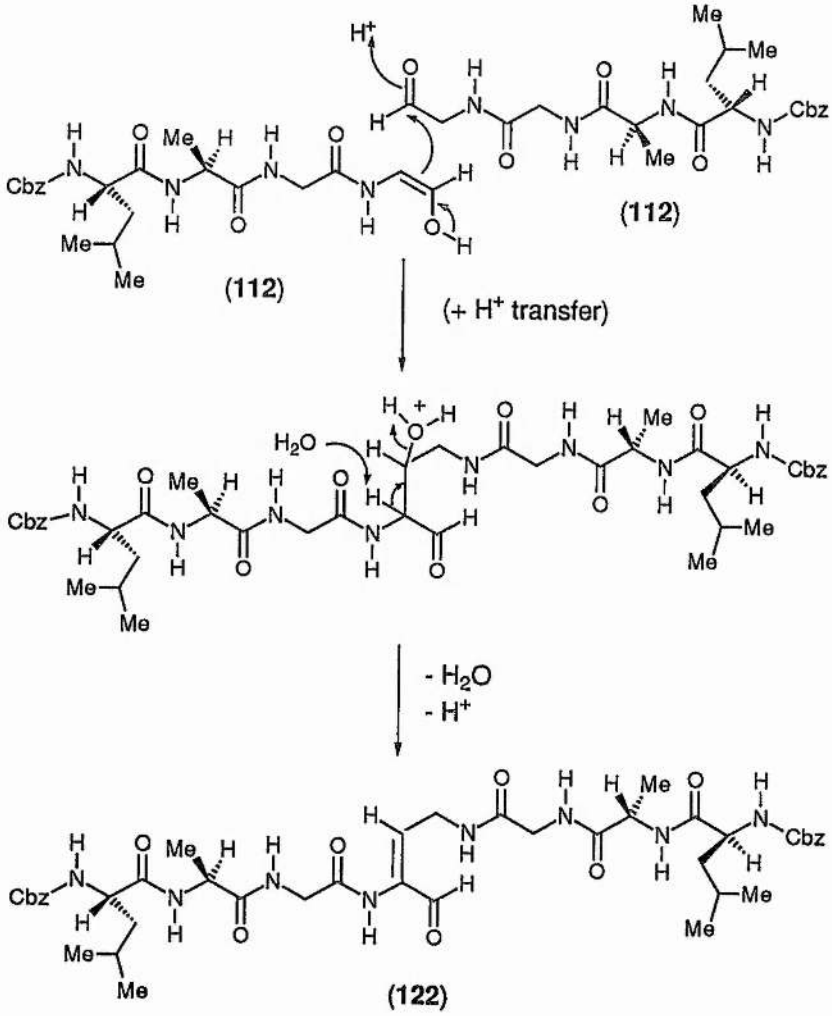
Scheme 2.21: *The preparation of the Cbz- and t-Boc protected tetrapeptide dimethyl acetals.*

Originally, preparation of (2S)-leucyl-(2S)-alanyl-glycyl-aminoacetaldehyde (**111**) was attempted by the acid catalysed hydrolysis of the urethane and dimethyl acetal moieties of t-Boc-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal (**112**) in aqueous methanol. However only polymeric material was isolated as determined by $^{1}H$- and $^{13}C$-NMR spectroscopy. It appeared that the aldehyde and amino functionalities had condensed (Scheme 2.22).




Scheme 2.22: *The self-condensation of the unprotected aldehyde.*

If the free amino group of Leu-Ala-Gly-aminoacetaldehyde (**111**) was N-protected the polymerisation problem may be avoided so Cbz-(2S)-leucyl (2S)-alanyl glycyl aminoacetaldehyde dimethylacetal (**112**) was treated with hydrochloric acid in methanol. However only impure N, N'-di(carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl)-2, 4-diamino-but-2-enal (**122**), as determined by $^{1}H$- and $^{13}C$-NMR spectroscopy and low resolution mass spectrometry, was isolated. Condensation had probably occured by the acid catalysed enolisation of the aldehyde moieties (Scheme 2.23).



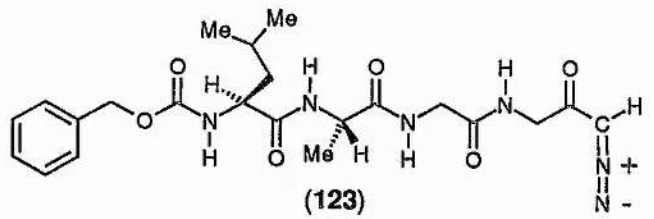

Scheme 2.23: *The condensation of the Cbz-protected aldehyde.*

Pure N, N'-di(carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl)-2, 4-diamino-but-2-enal (**122**) was prepared from the Cbz-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal (**114**), trimethylsilyl bromide and a propene acid trap in dry tetrahydrofuran in 45% yield after recrystallisation from methanol ether. The product displayed the correct mass, (FAB) 891 [M + H]$^+$ 100%). Trimethylsilyl iodide[133] also gave N, N'-di(carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl)-2, 4-diamino-but-2-enal (**122**) under similar conditions but the product was contaminated with iodine which formed from the decomposition of the trimethylsilyliodide under the reaction conditions used.

It is noted that the dimer, N, N'-di(carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl)-2, 4-diamino-but-2-enal (**122**) might still have the ability to interact with the proteinase either as for an aldehyde or as a Michael acceptor (Section 2.2.0).

### 2.1.4.0.1 The attempted syntheses of tetrapeptide ketones

Initially the mixed anhydride[95] of Cbz-Leu-Ala-Gly-Gly (**76**) was treated with diazomethane[134] to give Cbz-(2S)-leucyl (2S)-alanyl glycyl 1-amino 3-diazo propanone (**123**) (Fig. 2.13).

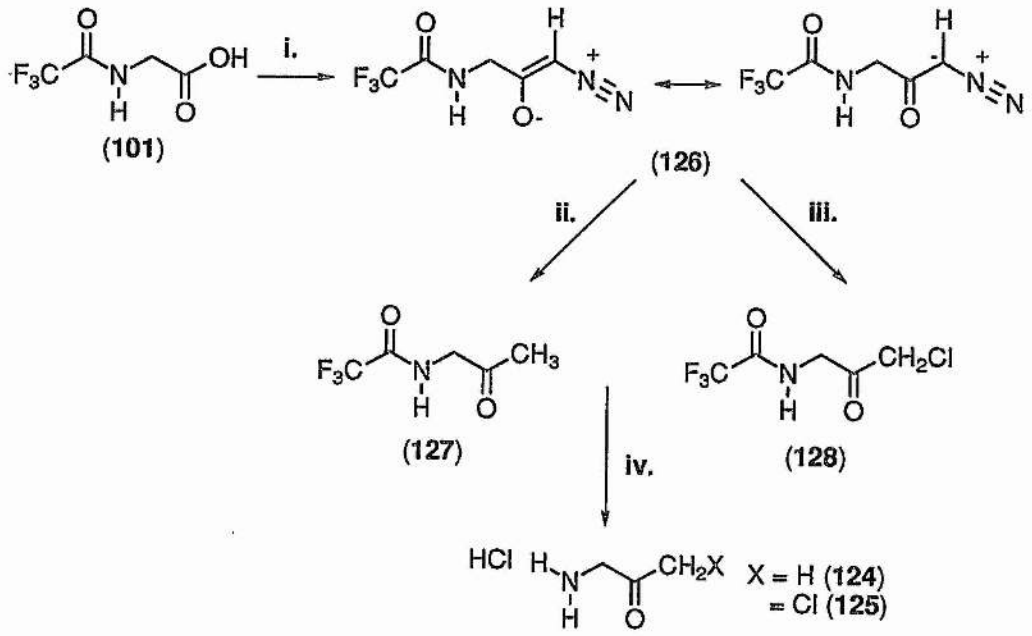



**(123)**

Fig. 2.13: *Cbz-Leu-Ala-Gly-1-amino 3-diazo propanone (123).*

The diazoketone (**123**) was then treated with excess hydroiodic acid to yield carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl 3-aminopropanone (**116**). However, a complex mixture of products resulted as determined by $^{1}$H- and $^{13}$C-NMR spectroscopy. It was believed that the formation of the impurities might be avoided by preparing 1-aminopropanone hydrochloride (**124**) and 1-amino 3-chloro-propanone hydrochloride (**125**) directly, as these compounds contain fewer functional groups to promote side reactions. 1-Aminopropanone hydrochloride (**124**) and 1-amino 3-chloro-propanone hydrochloride (**125**) could then be coupled to Cbz-Leu-Ala-Gly (**27**) to give carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl 3-amino propanone (**116**) and carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl 3-amino 1-chloro propanone (**117**) respectively.

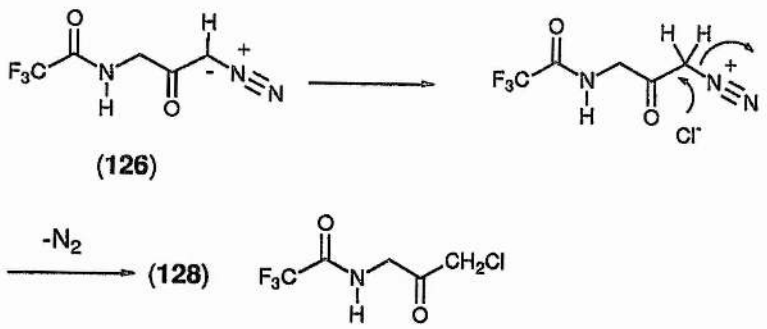
1-Aminopropanone hydrochloride (**124**) was prepared as described below; The mixed anhydride of TFA-glycine (**101**) and diazomethane were reacted to give trifluoroacetyl-1-amino- 3-diazo propanone (**126**) ($\upsilon_{max}$ 2970 and 2255 cm$^{-1}$ $CHN_2$) (Scheme 2.24). The diazoketone (**126**) was then directly treated with excess hydroiodic acid to give trifluoroacetyl-1-aminopropanone (**127**) in 56% yield (M.p. 57-58°C). TFA-deprotection of trifluoroacetyl-1-aminopropanone (**127**) in refluxing hydrochloric acid for 5 hours gave 1-amino propanone hydrochloride (**124**) in 91% yield (Found: $[M + H - HCl]^{+}$ 74.0601. $C_3H_8NO$ requires 74.0606). The diazoketone (**126**) was also treated with hydrochloric acid to give trifluoroacetyl-1-amino-3-chloro-propanone (**128**) in 57% yield (M.p. 53-55°C) (Scheme 2.24). TFA deprotection of trifluoroacetyl-1-amino-3-chloro-propanone (**128**) gave 1-amino-3-chloro-propanone hydrochloride (**125**). The chloromethylketone (**125**) was rather impure as judged by $^{1}$H- and $^{13}$C-NMR spectroscopy and further reaction with Cbz-Leu-Ala-Gly (**27**) was not attempted.



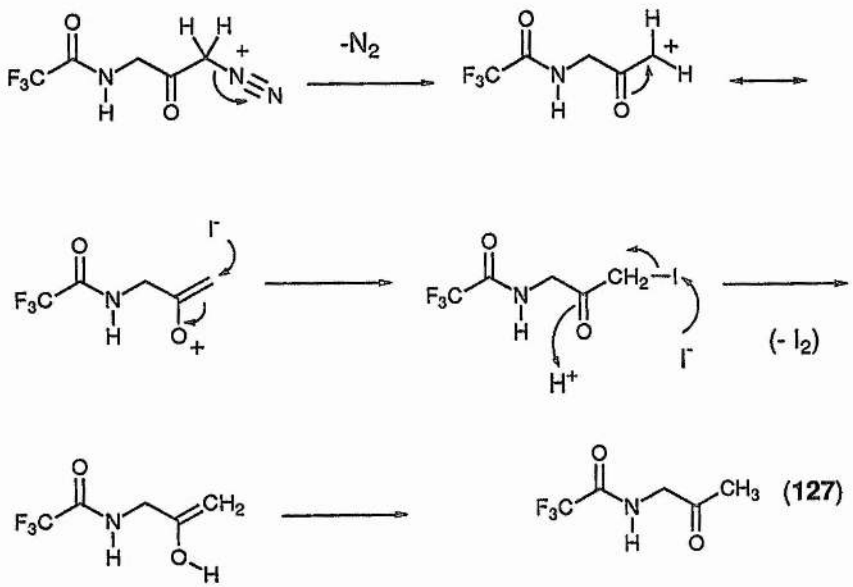

**Reagents**
**i.** i-BCF, NMM in THF at -15°C, 5 min, then10 $CH_2N_2$ in ether excluded from light at -15°C, 5 min → r.t., 2 hours . **ii.** 1.2 HI in dichloromethane excluded from the light, 2 hours. **iii.** HCl in dichloromethane excluded from the light, 2 hours. **iv.** 5M HCl reflux, 2 hours.

Scheme 2.24: *The preparation of 1-amino propanone hydrochloride (124) and 1-amino 3-chloro propanone hydrochloride (125).*

The difference in reactivity of hydrochloric acid and hydroiodic acid can be explained by the difference in charge distribution about the carbon-halogen bond. In the case of the chloro compound, chlorine is more electronegative than carbon so a partial bond polarisation exists with a partial negative charge residing on the chloride of the carbon-chloride bond. Therefore nucleophilic attack at the chloride is rather unlikely. For the iodo compound, iodine is less electronegative than carbon so a partial positive charge resides on the iodide of the carbon-iodide bond. This allows nucleophilic attack of the iodine of the carbon-iodide bond by another iodide from the solution giving the methyl ketone (Scheme 2.25 and 2.26).




Scheme 2.25: *The formation of a chloromethyl ketone.*




Scheme 2.26: *The formation of a methyl ketone.*

The mixed anhydride[95] of Cbz-Leu-Ala-Gly (**27**) (Section 2.1.0.0.0) was treated with pre-neutralised 1-aminopropanone hydrochloride (**124**) but only polymer and the tripeptide were isolated as determined by TLC, $^{1}$H- and $^{13}$C-NMR spectroscopy.

The preparation of carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl 3-amino propanone (**116**) was also attempted using Cbz-(2S)-leucyl (2S)-alanyl glycyl glycyl- N-, O'-dimethylhydroxylamide (**23**), but even using four equivalents of methyl magnesium bromide

in THF and refluxing for four hours gave no reaction. The N-, O'-dimethyl hydroxylamide (Weinreb amide) group allows only one equivalent of Grignard reagent to attack the amide carbonyl as the magnesium ion is co-ordinated to the methoxy group on the amide as well as the oxyanion forming a stable complex (Fig. 2.14).[135]

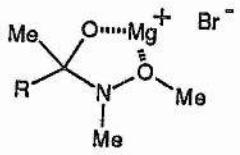


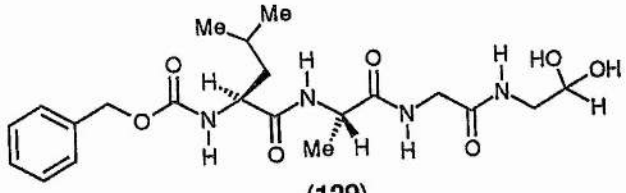

Fig 2.14: *The Weinreb amide complex formed with one equivalent to Grignard reagent.*

### 2.1.4.0.2 The successful synthesis of the tetrapeptide aldehyde and ketone

The carboxylic acid of carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycine (27) (section 2.1.0.0.1) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with allylamine and diisopropylethylamine overnight gave N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl)-1-amino prop-2-ene (**115**) in 40% yield after recrystallisation from acetone ether (M.p. 114°C). N-(Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl)-1-amino 2-methyl prop-2-ene (**117**) was similarly prepared from carbobenzyloxy-Leu-Ala-Gly (**27**) (Section 2.1.0.0.0) and 2-methyl allylamine in 32% yield after recrystallisation from acetone / ether (M.p. 148°C). The structures of the allyl amides (**115**) and (**117**) were confirmed by $^{1}$H-NMR, $^{13}$C-NMR, IR spectroscopy, microanalysis and mass spectrometry.

Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl aminoacetaldehyde (**112**) was prepared by ozonolysis[136] of Cbz-Leu-Ala-Gly-1-aminoprop-2-ene (**115**) in 70% yield (m/z 3%, $[M]^{+}$ 452). It should be noted that the aldehyde (**112**) exists primarily in the hydrated, diol

(**129**) form[119] (Fig. 2.15). Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl 1-amino-2-methyl propanone (**116**) was similarly prepared from N-(Cbz-Leu-Ala-Gly)-1-amino 2-methyl prop-2-ene (**118**) in 88% yield (Found: $[M + H]^+$ 449.2406. $C_{22}H_{33}N_4O_6$ requires 449.2400). The structures of the aldehyde (**112**) and methyl ketone (**116**) were confirmed $^1$H-NMR, $^{13}$C-NMR and IR spectroscopy, microanalysis and low resolution mass spectrometry.




**(129)**

Fig. 2.15: *The hydrated diol form (129) of Cbz-Leu-Ala-Gly-aminoacetaldehyde (112).*

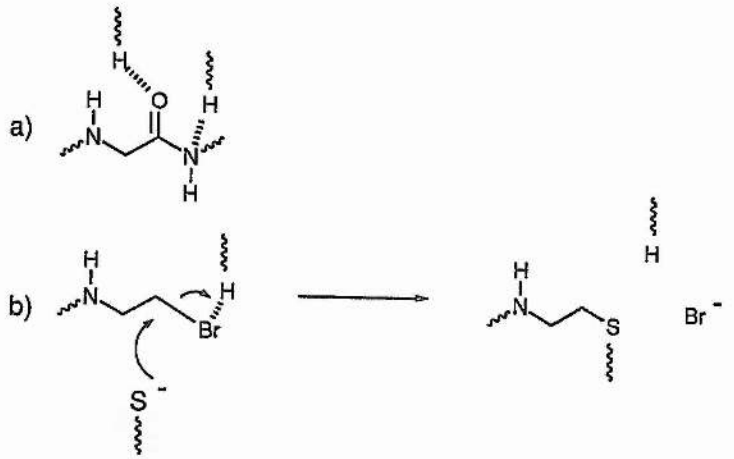
## 2.2 The synthesis of irreversible inhibitors

Irreversible inhibitors act by binding tightly to the enzyme or one of the intermediates, preventing catalytic activity. The binding is usually covalent and stoicheometric. See Introduction, Section 1.2.3.2.

### 2.2.0 Rationale for the design of potential inhibitors containing bromide and Michael acceptor moieties

It was expected that N-(Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl)-2-bromo ethylamine (**130**) might have the capacity to mimic a typical substrate for the adenovirus proteinase. The bromide might interact with a proton donor present in the $P_1$-$P_1$' amide nitrogen binding pocket (Scheme 2.27). If such an interaction was possible, the thiolate present in the proteinase's active site might nucleophilically displace the bromide from the carbon-bromide bond. This could lead to irreversible alkylation of the proteinase (Scheme

2.27). It is probably unlikely that the bromide could mimic the carbonyl group of the amide and bind in the $P_1$-$P_1$' amide carbonyl binding pocket as the carbon-bromide bond is too long. The Van der Waal's radius for bromine is very large (0.91 nm) compared to oxygen in an amide bond (0.15 nm). The average carbon-bromide bond length is 0.194 nm compared to the average carbon-oxygen bond length of 0.123 nm.

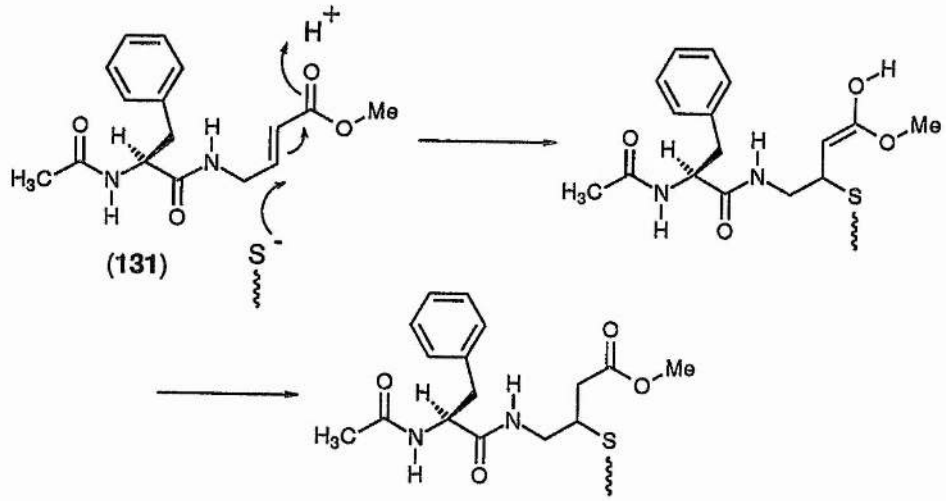



Scheme 2.27: *The tetrapeptide bromide's possible interaction with the proteinase; a) substrate; b) tetrapeptide bromide.*

The unsaturated ester, N-(acetyl-(2S)-phenylalanyl 2-carboxymethyl 1-amino but-2-ene (**131**) is a competitive irreversible inhibitor of papain ($k_2$ =1.82 mol $dm^{-3}$ $s^{-1}$).[137] , N-(acetyl-Phe-1-carboxymethyl 2-amino but-2-ene (**131**) slowly and irreversibly alkylates the active site thiolate of papain in a Michael fashion (Scheme 2.28 on page 91).

The unsaturated ester functionality of (**131**) was included in the substrate recognition sequence of the adenovirus proteinase to give O-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl-1-carboxymethyl-2-hydroxybut-2-ene (**132**) (Fig. 2.16). It was expected that O-(Cbz-Leu-Ala-Gly)-1-carboxymethyl-2-hydroxybut-2-ene (**132**) might inhibit the

adenovirus proteinase in a similar manner to N-(acetyl-Phe-1-carboxymethyl,2-aminobut-2-ene (**131**) inhibiting papain.




Scheme 2.28: *The irreversible alkylation of papain by the Michael acceptor (131).*

The unsaturated ester functionality of (**131**) was included in the substrate recognition sequence of the adenovirus proteinase to give O-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl-1-carboxymethyl-2-hydroxybut-2-ene (**132**) (Fig. 2.16). It was expected that O-(Cbz-Leu-Ala-Gly-1-carboxymethyl-2-hydroxybut-2-ene (**132**) might inhibit the adenovirus proteinase in a similar manner to N-(acetyl-Phe)-1-carboxymethyl,2-aminobut-2-ene (**131**) inhibiting papain.

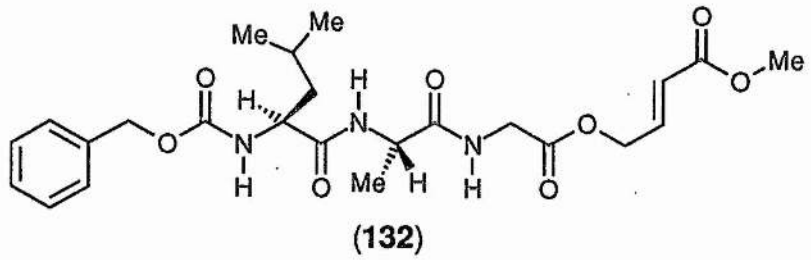



Fig. 2.16: *O-(Cbz-Leu-Ala-Gly)-1-carboxymethyl,2-hydroxybut-2-ene (132).*

### *2.2.0.0 The synthesis of bromide and Michael acceptor containing inhibitors*

The carboxylic acid of Cbz-Leu-Ala-Gly (**27**) (section 2.1.0.0.1) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with the 2-bromo-ethylamine and diisopropylethylamine overnight to give N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl)-2-bromoethylamine (**130**) as the monohydrate in 70% yield after recrystallisation from acetone / ether (M.p. 107°C). O-(Carbobenzyloxy-(2S)-leucyl-(2S)-alanyl-glycyl), 1-carboxymethyl-4-hydroxybut-2-ene (**132**) was prepared from Cbz-Leu-Ala-Gly (**27**) (Section 2.1.0.0.1), methyl γ-bromocrotonate and sodium hydrogen carbonate in dimethyl formamide[138] in 35% yield after chromatography on silica eluting with dichloromethane, then 1% methanol in dichloromethane through to 5% methanol in dichloromethane and recrystallisation from methanol / ether (M.p. 89°C). The structures of the bromide (**130**) and the Michael acceptor (**132**) were confirmed by H-NMR, $^{13}$C-NMR, IR spectroscopy, microanalysis and mass spectrometry.

## 2.2.1 Rationale for the epoxysuccinate based potential inhibitors

**E-64** ([N-(RR-3-*trans*-carboxyoxirane-2-carbonyl)-(2S)-leucyl]-amido(4-guanido) butane (**17**) (Fig. 2.17) was isolated[139] from cultures of *Aspergillus jamonicus* and is a potent highly selective irreversible inhibitor of cysteine proteinases[132] (rate constant for the inactivation of papain 6.38 x $10^5$ mol $dm^{-3}$ $s^{-1}$).[140] The (SS)-epoxy succinate analogue of **E-64** (**133**) (Fig. 2.17) is also a weaker irreversible inhibitor of cysteine proteinases (rate constant for the inactivation of papain 6.09 x $10^4$ mol $dm^{-3}$ $s^{-1}$).[132] (RR)-**E-64** (**17**) does not alkylate free thiols in solution at physiological pH[132] and (RR)-**E-64** (**17**) does not inhibit the adenovirus type 2 proteinase.[72] It should be noted that the peptide backbones of **E-64** (**17**) and the derivatives of **E-64** run from right to left rather than the usual peptide direction of left to right.

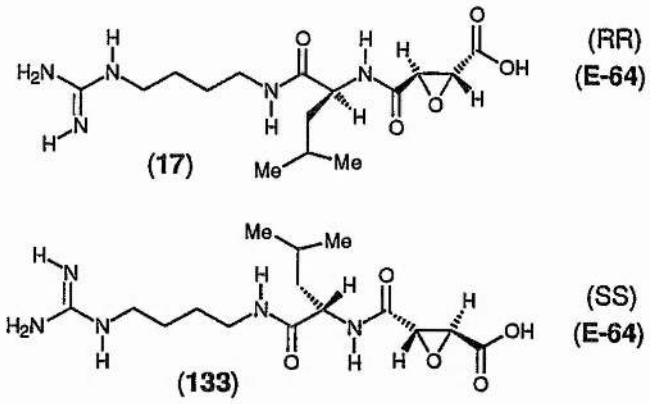
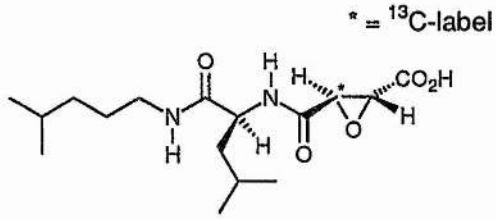



Fig. 2.17: *The (RR) isomer and the (SS) isomer of* ***E-64****.*

It was suggested that the nucleophilic thiolate of papain became alkylated by attacking the epoxysuccinyl moiety of (RR)-**E-64** (17).[141] However, the nucleophilic cysteine residue (Cys-25 in papain) could become alkylated by attacking either C-2 or C-3 of the epoxide ring. The mechanism of inactivation of papain by (RR)-**E-64** (17) was determined from $^{13}$C-NMR spectroscopy of a mixture of papain and [3-$^{13}$C]-**Ep-475** ([3-$^{13}$C]-[N-(RR-3-*trans*-carboxyoxirane-2-carbonyl)-(2S)-leucyl]-amido(3-methyl) butane) (**134**) (Fig. 2.18), a closely related derivative of (RR)-**E-64** (17).[141]




**Ep-475 (134)**

Fig. 2.18: *[3-$^{13}$C]-Ep-475 (134).*

When [3-$^{13}$C]-**Ep-475 (134)** was added to papain (both the concetrations of papain and **Ep-475 (134)** were at 1.7 mmol dm$^{-3}$), a new signal was observed at 76.6 ppm. This corresponded with the disappearence of the C-3 epoxide signal at 52.5 ppm.

If [3-$^{13}$C]-**Ep-475 (134)** was treated with benzyl mercaptan at pH 10, two products were observed, one corresponding to the attack of benzyl mercaptan anion at C-2 (**135**) and one from the attack of benzyl mercaptan anion at C-3 (**136**) (Fig. 2.24). The $^{13}$C signal of C-3 appeared at 53 ppm when the benzyl mercaptan anion has attacked the C-3 carbon to give a β-thioamide at C-3 (compound (**135**)) and the $^{13}$C signal of C-3 appears at 73 ppm when the benzyl mercaptan anion has attacked the C-2 carbon to give a β-hydroxyamide at C-3 (compound (**136**)).

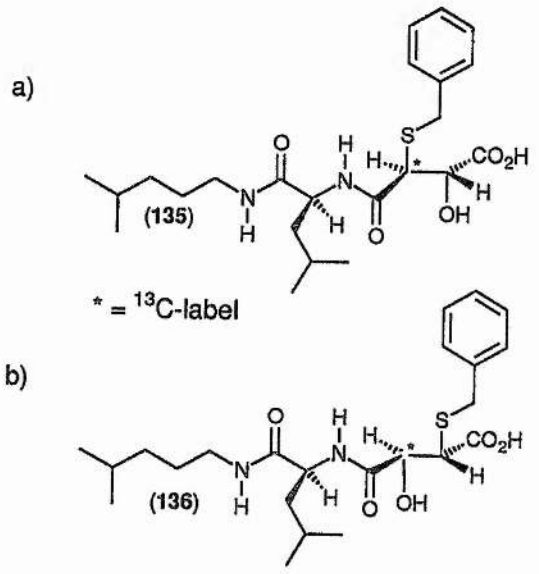
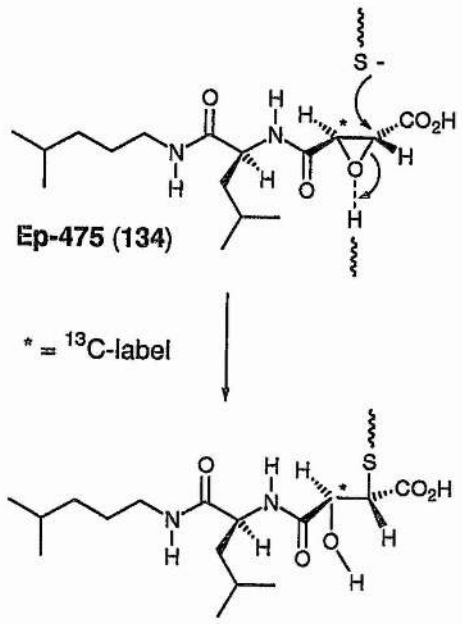



Fig. 2.19: *The two possible products from benzyl mecaptan anion attack on [3-$^{13}$C]-Ep-475 (134) ; a) C-3 attack and b) C-2 attack.*

By comparison of the new $^{13}$C signal observed when papain nucleophilically attacked [3-$^{13}$C]-**Ep-475 (134)** with the $^{13}$C signals observed when benzylmercaptan attacks [3-$^{13}$C]-**Ep-475 (134)**, it was concluded that papain attacked [3-$^{13}$C]-**Ep-475 (134)** at the C-2 carbon of the epoxide (Scheme 2.29).[141]




Scheme 2.29: *The alkylation of papain by [3-$^{13}$C]-Ep-475 (134).*

If the structure of **E-64** (**17**) could be modified to fit into the adenovirus proteinase's active site, a potentially very potent inhibitor for the adenovirus proteinase could be produced. Our first modifications of the **E-64** (**17**) structure were to include a substrate recognition sequence for the adenovirus proteinase and then simplify the trans epoxide such that the peptide backbone could be conserved. These considerations give (±)-O-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl)-1-hydroxy-2,3-oxirane propane (**137**) (Fig. 2.20) as a potential inhibitor. The stereochemistry of the adenovirus proteinase's active site is not known, so all isomers were considered as being worth synthesizing.

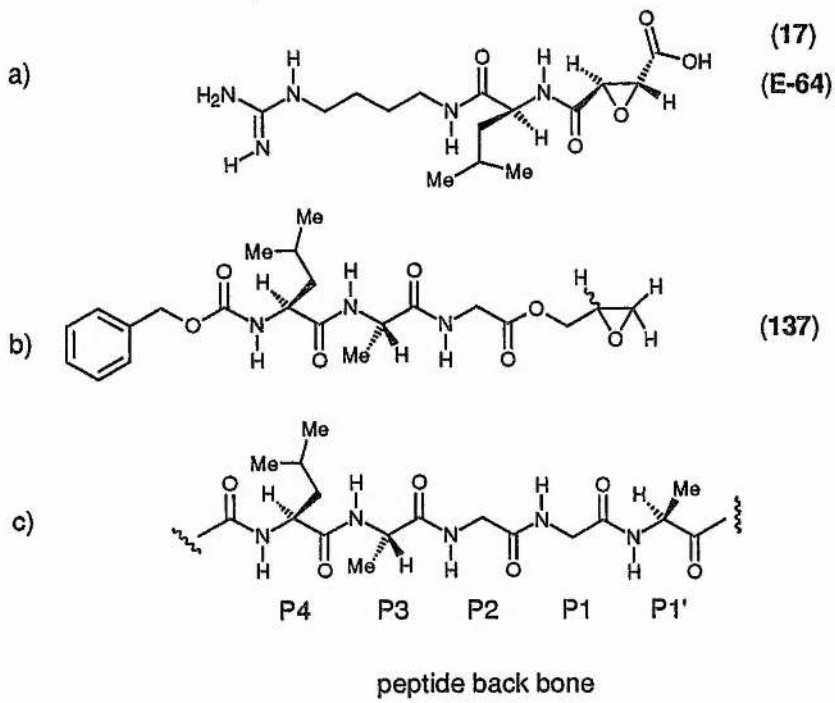



Fig. 2.20: *a) E-64 (17); b) (±)-O-(Cbz-Leu-Ala-Gly)-glycidol (137) and c) substrate.*

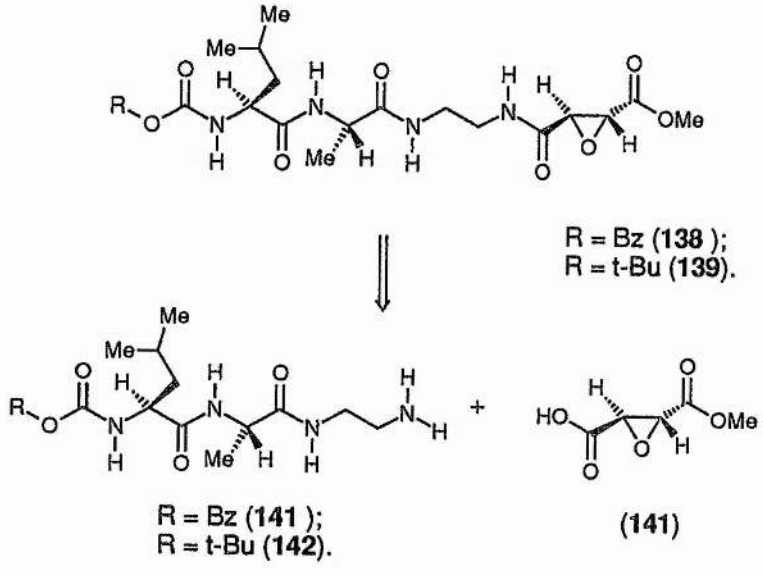
## *2.2.1.0 The attempted synthesis of (±)-O-(Cbz-Leu-Ala-Gly)-glycidol (137)*

Numerous attempts to couple Cbz-Leu-Ala-Gly (27) (Section 2.1.0.0.1) to (±)–glycidol via the mixed anhydride,[95] dicyclohexylcarbodiimide / p-dimethylaminopyridine[142] and trimethylsilylchloride / NMM[143] activated ester procedures failed, giving a complex mixture of products, including Cbz-Leu-Ala-Gly (27), the ring opened epoxide and (±)-glycidol. The acid chloride[90] and TBTU[111] activated ester couplings were not attempted as any hydrochloric acid or hydroxybenzotriazole produced *in situ* might cleave the epoxide moiety.[144]

### 2.2.1.1 Rationale for the N-protected N-((2S)-leucyl (2S)-alanyl)- N'-(3-trans-carboxyoxirane-2-methoxycarbonyl)-butane ethylene diamines

The **E-64** (**17**) structure was further modified by reintroducing the *trans*-epoxysuccinyl moiety and changing the position of the $P_2$-$P_1$ amide link by one C-atom towards the P' direction to give N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl)- N'-(3-trans-carboxyoxirane-2-methoxycarbonyl)-butane ethylene diamine (**138**) and N-(t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl)- N'-(3-trans-carboxyoxirane-2-methoxycarbonyl)-butane ethylene diamine (**139**) (Scheme 2.30).

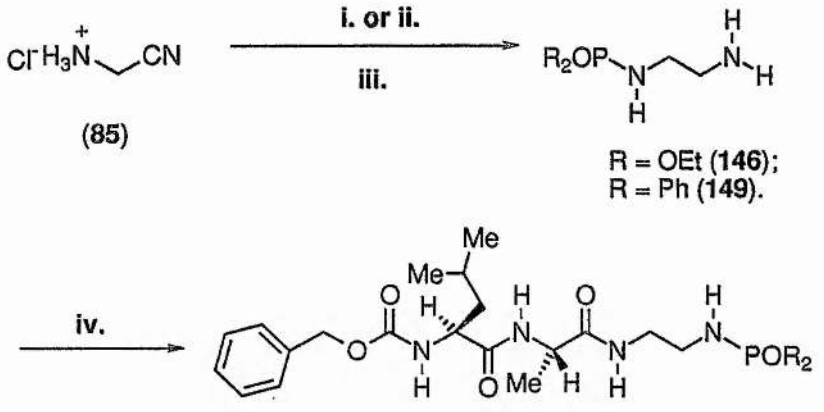
The tetrapeptide epoxides (**138**) and (**139**) were disconnected to give N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl)-ethylenediamine (**140**) and methyl (2R,3R)-2,3 epoxysuccinate (**141**) and N-(t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl)-ethylenediamine (**142**) and methyl (2R, 3R)-2, 3 epoxysuccinate (**141**), respectively (Scheme 2.30).




Scheme 2.30: *The disconnection of epoxides (138) and (139).*

### 2.2.1.1.0 The attempted synthesis of N-(Cbz-Leu-Ala)-ethylene diamine (140)

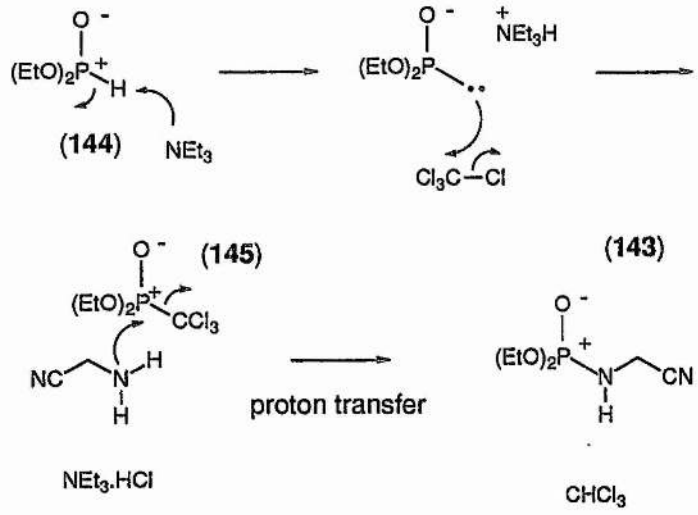
The first attempted synthesis of N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl)-ethylenediamine (**140**) was carried out using diethyl phosphonate protection. Aminoacetonitrile hydrochloride (**85**) was N-protected with diethyl phosphonate (**144**)[145] to give diethyl aminoacetonitrile phosphonate (**143**) in 98% yield (m/z (CI) 193 $[M + H]^+$, 100%) (Scheme 2.31).




**Reagents**
**i.** R = OEt; $CCl_4$, $HPO(OEt)_2$ (**144**) $Et_3N$, 0°C, 2 hours → (**143**). **ii.** R = Ph; $Ph_2POCl$, NMM in dichloromethane, 0°C, 2 hours → (**148**). **iii.** $H_2$, Pd / C in MeOH, R = OEt; 3 hours → (**146**), R = Ph; 2 days → (**149**). **iv.** Cbz-LA, NMM, i-BCF in THF at -15°C, 5 min. (**146**) or (**149**) in THF at -15°C, 5 min → r.t., 2 hours.

Scheme 2.31: *The attempted preparation of N-(Cbz-Leu-Ala)-ethylene diamine (140).*

The mechanism[145] for the formation of diethyl aminoacetonitrile phosphonate (**143**) is shown in Scheme 2.32. The diethyl phosphite (**144**) was deprotonated by triethylamine and nucleophilically attacked by carbon tetrachloride to give triethylamine hydrochloride and a trichloromethyl phosphonate (**145**). The amino group of aminoacetonitrile displaces the trichloromethyl group from the trichloromethyl phosphonate (**145**) which after proton exchange gives chloroform and diethyl aminoacetonitrile phosphonate (**143**).




Scheme 2.32: *The mechanism of the formation of aminoacetonitrile phosphonate (143)*

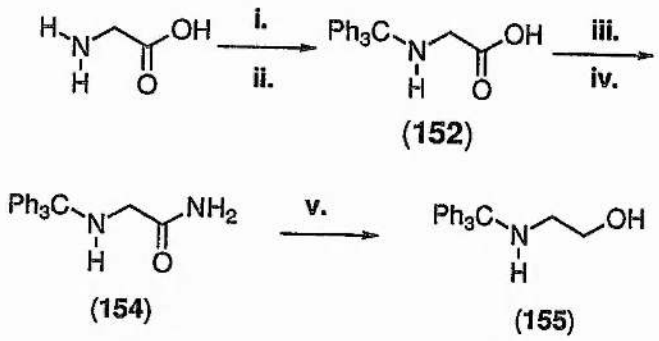
Reduction of the nitrile group of diethyl aminoacetonitrile phosphonate (**143**) with hydrogen and palladium dispersed in charcoal gave N-(diethyl phosphonate) ethylenediamine (**146**) in 90% yield (m/z (CI) 197 $[M + H]^+$, 100%). The mixed anhydride of Cbz-Leu-Ala (**29**) (Section 2.1.0.0.0) was treated with N-(diethyl phosphonate) ethylenediamine (**146**) to give N-(Carbobenzyloxy-(2S)-Leucyl (2S)-Alanyl)- N'-(diethylphosphonate)-ethylene diamine (**147**). However, a complex mixture of products resulted as observed by $^1$H- and $^{13}$C-NMR spectroscopy. It appeared that aminolysis of the ethoxy groups in the diethyl phosphonate protecting group might have occurred.

It was hoped that replacing the offending diethyl phosphonate protecting group with the diphenyl phosphine oxide protecting group might prevent the aminolysis.[146] N-(Diphenyl phosphine oxide)-aminoacetonitrile (**148**) was prepared from aminoacetonitrile hydrochloride (**2.59**), NMM and diphenylphosphoryl chloride in THF at 0°C in 52% yield after recrystallisation from methanol / ether (Found: $[M + H]^+$ 257.0838. $C_{14}H_{13}N_2OP$ requires 257.0844) (Scheme 2.32). Hydrogenation of N-(diphenyl phosphine oxide)-aminoacetonitrile

(**148**) for 2 days gave N-(diphenyl phosphine oxide)-ethylenediamine (**149**) in 98% yield (M.p. 58-59°C). Unfortunately the reaction of the mixed anhydride of Cbz-Leu-Ala (**29**) (Section 2.1.0.0.0) with N-(diphenyl phosphine oxide)-ethylenediamine (**149**) did not give N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl)-,N'-(diphenylphosphine oxide)-ethylenediamine (**150**), but a complex mixture of products as observed by $^{1}$H- and $^{13}$C-NMR spectroscopy.

The third attempted synthesis of N-(Cbz-Leu-Ala)-ethylenediamine (**140**) involved the coupling of N-trityl-ethylenediamine (**151**) to Cbz-Leu-Ala (**29**) (Section 2.1.0.0.0). It was hoped that N-trityl-ethylenediamine (**151**) could be prepared from N-trityl-glycyl amide (**154**) as described below; N-trityl-glycine (**152**) was prepared from glycine, trimethylsilylchloride, NMM and trityl chloride (trityl = triphenylmethyl) in 60% yield after recrystallisation from methanol (M.p. 179°C, lit.,[147] 178-179°C) (Scheme 2.33). Previous methods[147] gave mixtures of products as determined by TLC and $^{1}$H and $^{13}$C-NMR spectroscopy. N-trityl-glycine (**152**) was treated with diazomethane[148] to give methyl N-trityl-glycinate (**153**) in 98% yield after recrystallisation from methanol / ether (M.p. 99°C) (Scheme 2.33). Treatment of methyl N-trityl-glycinate (**153**) with sodamide in liquid ammonia[149] gave N-trityl-glycinamide (**154**) in 70% yield after recrystallisation from water (M.p. 179°C) (Scheme 2.33).

The reduction of N-trityl-glycinamide (**154**) proved to be very problematic; sodium borohydride / cobalt (II) chloride (hydrated),[150] sodium acylborohydride[151] and borontrihydride.THF complex[152] all caused over reduction of the trityl group as judged by $^{1}$H- and $^{13}$C-NMR spectroscopy. Lithium aluminium hydride[153] smoothly reduced the amide moiety to N-trityl ethanolamine (**155**) as determined by mass spectrometry, $^{1}$H- and $^{13}$C-NMR spectroscopy (Scheme 2.33).



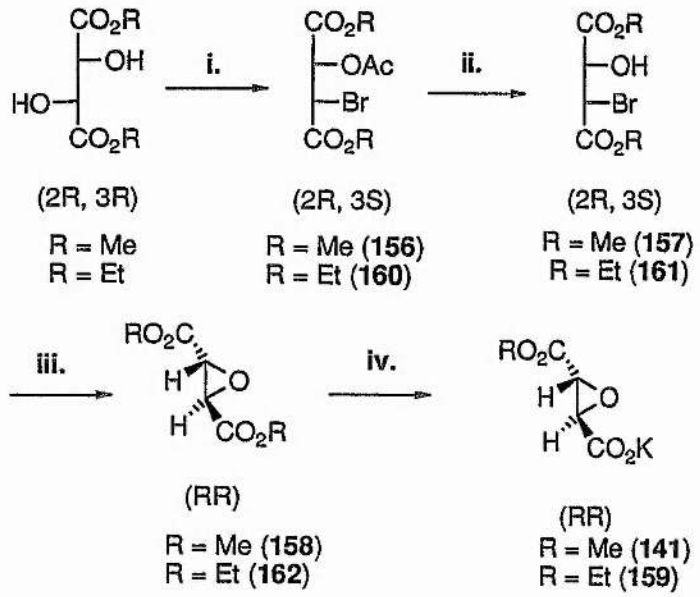

**Reagents**
**i.** 2.2 TMS-Cl, 2.2 NMM, 0°C, 10 min → reflux, 15 min. **ii.** NMM, trityl-Cl, reflux, 2 hours. Then MeOH 1 hour. **iii.** $CH_2N_2$, MeOH, 0°C, 2 hours → (153). **iv.** $NaNH_2$, liq. $NH_3$. -78°C,1 hour. **v.** $LiAlH_4$, 4 hours.

Scheme 2.33: *The attempted preparation of N-trityl-ethylenediamine (151).*

### 2.2.1.1.1 The synthesis of the epoxysuccinyl monoalkyl esters

Optically pure potassium methyl (2R,3R)-erythro 2,3-epoxysuccinate (**141**) was also prepared.[154] (2R,3R)-dimethyl tartrate was brominated with 30% hydrogen bromide in acetic acid to give (2R,3S) dimethyl-threo-2-acetoxy-3-bromosuccinate (**156**) in 32% yield (b.p. 136°C at 1.5 mm Hg; lit.,[155] 83-86°C at 0.1 mm Hg) (Scheme 2.28). The acetoxy group of (2R,3S)-dimethyl threo-2-acetoxy-3-bromosuccinate (**156**) was hydrolysed in methanol in the prescence of catalytic hydrogen bromide in acetic acid to give the (2R,3S)-dimethyl threo 3-bromo-2-hydroxysuccinate (**157**) in 31% yield (b.p. 145°C at 1.4 mm Hg; lit.,[155] 80-82°C at 0.03 mm Hg) (Scheme 2.34).

Deprotonation of the alcohol with sodium methoxide allowed the formation of (2R,3R) dimethyl epoxysuccinate (**158**) in a yield of 58% after recrystallisation from methanol (M.p. 72°C; lit.,[144] 77-78°C) (Scheme 2.28). (2R,3R) dimethyl epoxysuccinate (**158**) was partially saponified[154] to give potassium (2R,3R) methyl epoxysuccinate (**141**) in 94% yield (M.p. decomp. 148°C) (Scheme 2.34).



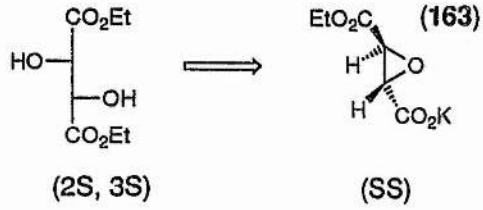

**Reagents**
**i.** HBr / HOAc, 0°C, 30 min, 2 hours r.t. **ii.** cat. HBr dry ROH, reflux, 2 hours. **iii.** NaOR, ROH, 0°C, 2 hours. **iv.** KOH, ROH 0°C, 2 hours.

Scheme 2.34:

*The preparation of potassium methyl (2R,3R)-erythro 2,3-epoxysuccinate (141).*

Potassium (2R,3R) ethyl epoxysuccinate (**159**)[144] (decomp. 133°C lit.,[144] decomp. 159°C) was prepared similarly from (2R,3R) diethyl tartrate giving (2R,3S)-erythro diethyl 2-acetoxy-3-bromosuccinate (**160**), (2R,3S)-erythro diethyl 2-hydroxy-3-bromosuccinate (**161**) and (2R,3R) ethyl epoxysuccinate (**162**) in yields of 50%, 72%, 75% and 44% respectively (Scheme 2.34).[144]

Potassium (2S,3S) ethyl epoxysuccinate (**163**) (Scheme 2.35) was prepared similarly from (2S,3S) diethyl tartrate giving (2S,3R)-threo diethyl 2-acetoxy-3-bromosuccinate (**164**), (2S,3R)-threo diethyl 2-hydroxy-3-bromosuccinate (**165**) and (2S,3S) ethyl epoxysuccinate (**166**) in yields of 45%, 79%, 84% and 73% respectively.[154] All the epoxysuccinates and their precursors gave satisfactory $^{1}$H-NMR, $^{13}$C-NMR, IR and low resolution mass spectrometry spectra.




Scheme 2.35: *Potassium (2S, 3S) ethyl epoxy succinate (163).*

A similar procedure was adopted to prepare a racemic mixture of (2R,3S) and (2S,3R) potassium ethyl epoxysuccinate (**167**) (Fig. 2.21) from (RS)-diethyl tartrate. However hydrolysis of the mixture of (2R,3R) and (2S,3S)-diethyl bromoacetoxysuccinate (**168**) produced only diethyl tartrate. A second procedure was also attempted, involving the formation of disodium *cis*-epoxysuccinate (**169**) from the sodium tungstate catalysed epoxidation of maleic acid with alkaline hydrogen peroxide.[156]

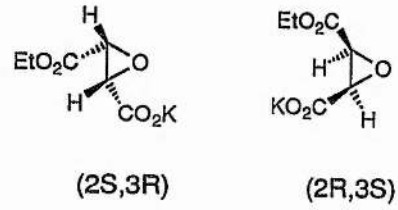


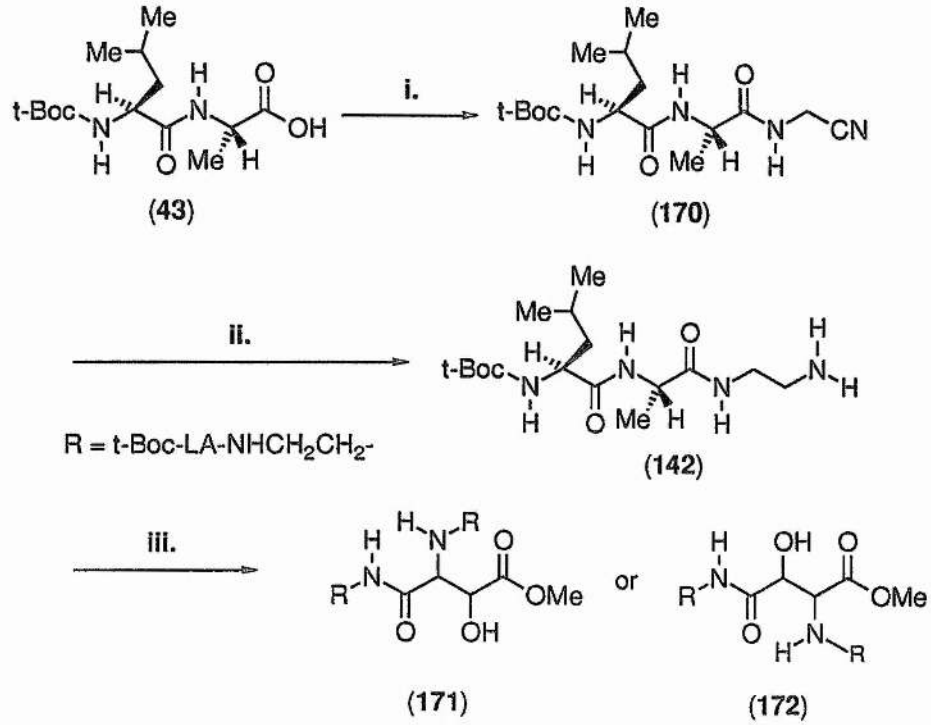

Fig: 2.21: *(±)-cis potassium ethyl epoxy succinate (167).*

Esterification of disodium *cis*-epoxy succinate (**169**) using either an acid catalysed[157] or a bromoalkyl[128] procedure failed, producing a mixture of diols as judged by $^{1}H$- and $^{13}C$-NMR spectroscopy. The (2R,3S) and the (2S,3R) potassium ethyl epoxysuccinates (**167**) were probably more unstable than potassium (2S,3S) ethyl epoxysuccinate (**163**) and potassium (2R,3R) ethyl epoxysuccinate (**159**). The carboxyl groups are on the same face of the molecule for the (2R,3S) and the (2S,3R) forms so that the opposite face to the carbonyl

### 2.2.1.1.2 The attempted synthesis of N-(t-Boc-Leu-Ala)-N'-(2S, 3S)-[(3-trans-carboxyoxirane-2-methoxycarbonyl) butane] ethylenediamine (139)

The synthesis of N-(t-Boc-Leu-Ala)-N'-(2S,3S)-[(3-trans-carboxyoxirane-2-methoxycarbonyl) butane] ethylenediamine (**139**) was attempted as follows; The carboxylic acid of t-Boc-Leu-Ala (**43**) (section 2.1.0.0.1) was activated with TBTU[111] in THF at 20°C. Reaction of the activated ester with aminoacetonitrile hydrochloride (**85**) and diisopropylethylamine overnight gave t-butoxycarbonyl-(2S)-leucyl-(2S)-alanyl aminoacetonitrile (**170**) in 84% yield after recrystallisation from methanol / ether (M.p. 134°C) (Scheme 2.36). Reduction of t-Boc-Leu-Ala-aminoacetonitrile (**170**) with hydrogen and palladium dispersed in carbon gave N-(t-butoxycarbonyl-(2S)-leucyl (2S)-alanyl)-ethylenediamine (**142**) in 98% yield (M.p. 99-102°C) (Scheme 2.36). $^{1}$H- and $^{13}$C-NMR spectra showed the complete disappearence of the nitrile group and the appearence of a mono acylated ethylene diamine moiety.

The final coupling of the N-(t-Boc-Leu-Ala)-ethylenediamine (**142**) to methyl (2R,3R)-2, 3-epoxysuccinate (**141**) was attempted using the mixed anhydride and the pentafluorophenyl[158] activated ester methods. However, both reactions gave the ring opened products (**171**) or (**172**), as determined by $^{1}$H- and $^{13}$C-NMR spectroscopy (Scheme 2.36). Even when three equivalents of methyl (2R,3R)-2, 3-epoxysuccinate (**141**) were used the same products resulted, indicating that the release in ring strain upon ring opening the epoxide is more favourable than amide formation.



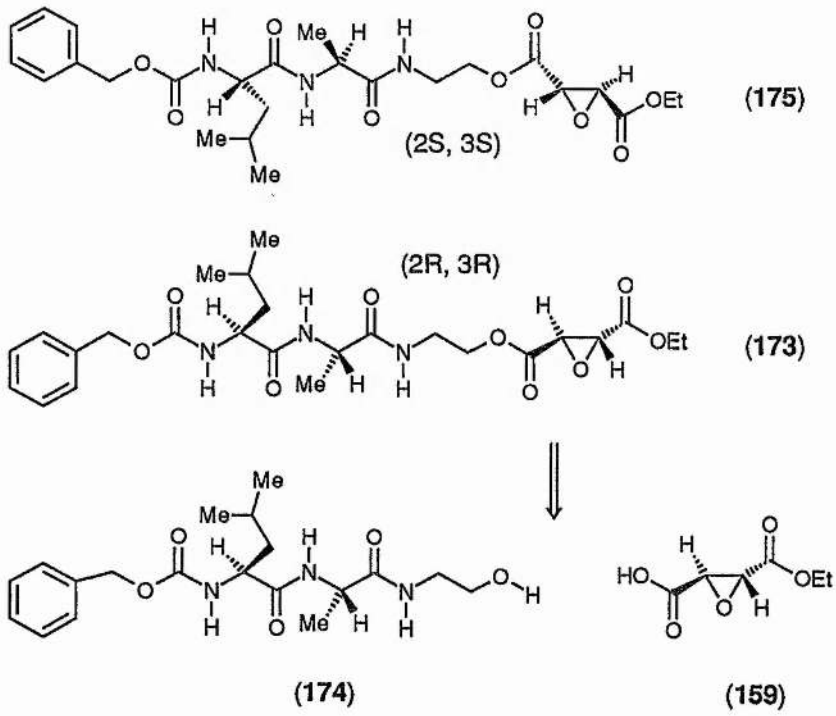

**Reagents**
**i.** i-$Pr_2$EtN, TBTU, aminoacetonitrile hydrochloride (**85**) in THF, 12 hours.
**ii.** $H_2$ / Pd / C in MeOH, 2 days. **iii.** Coupling to methyl (2R, 3R) epoxysuccinate (**141**).

Scheme 2.36: *The attempted preparation of N-(t-Boc-Leu-Ala)-N'-(2S,3S)-[(3-trans-carboxyoxirane-2-methoxycarbonyl) butane] ethylenediamine (139).*

### 2.2.1.2 Rationale for the N-protected O-((2S)-leucyl (2S)-alanyl)- N'-(3-trans-carboxyoxirane-2-methoxycarbonyl)-butane ethanolamines

It was expected that changing the $P_2$-$P_1$ bond from an amide to an ester bond (Scheme 2.37) to give (2R,3R)-O-[trans-carbonyl oxiran-2-yl 3-(ethoxycarbonyl) butane] N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl)-ethanolamine (**173**) (Scheme 2.37) as the new target compound, might prevent the ring opening of the epoxide moiety in section 2.2.1.1.2.

As the alcohol (**174**) (Scheme 2.37) required for the formation of the $P_2$-$P_1$ ester bond is less nucleophilic than the amine (**142**) (Scheme 2.36) from section 2.2.1.1 which ring opened the epoxide moiety during the formation of N-(t-Boc-Leu-Ala)-N'-(2S,3S)-[(3-*trans*-carboxyoxirane-2-methoxycarbonyl) butane] ethylenediamine (**139**). (2R,3R)-O-[*trans*-carbonyl oxiran-2-yl 3-(ethoxycarbonyl) butane] N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl)-ethanolamine (**173**) could be disconnected to give N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl)-ethanolamine (**174**) and (2R,3R) ethyl epoxysuccinate (**159**) (Scheme 2.37). (2S,3S)-O-[*trans*-carbonyl oxiran-2-yl 3-(ethoxycarbonyl) butane] N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl)-ethanolamine (**175**) (Fig. 2.29) could similarly be disconnected to give N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl)-ethanolamine (**174**) and (2S,3S) ethyl epoxy succinate (**163**).




Scheme 2.37: *Disconnection of epoxides (173) and (175).*

#### 2.2.1.2.0 Synthesis of the N-protected O-((2S)-leucyl (2S)-alanyl)- N'-(3-trans-carboxyoxirane-2-methoxycarbonyl)-butane ethanolamines

N-(Cbz-Leu-Ala)-ethanolamine (**174**) was prepared from the mixed anhydride of Cbz-Leu-Ala (**29**) (Section 2.1.0.0.0) and ethanolamine in 72% yield after recrystallisation from acetone / ether (M.p. 140°C). The mixed anhydride[95] of (2R,3R) ethyl epoxy succinate (**159**) was reacted with N-(Cbz-Leu-Ala)-ethanolamine (**174**) to give (2R,3R) O-[*trans*-carbonyl oxiran-2-yl 3-(ethoxycarbonyl) butane] N-(carbobenzyloxy-(2S)-leucyl-(2S)-alanyl)-ethanolamine (**173**) in 41% yield after chromatography on silica eluting with ethyl acetate (M.p. 128°C). (2S,3S) O-[*trans*-carbonyl oxiran-2-yl 3-(ethoxycarbonyl) butane] N-(carbobenzyloxy-(2S)-leucyl-(2S)-alanyl)-ethanolamine (**175**) was prepared in a similar manner from (2S,3S) ethyl epoxysuccinate (**163**) and N-(Cbz-Leu-Ala)-ethanolamine (**174**) in 37% yield after chromatography on silica eluting with ethyl acetate. (M.p. 150-151°C). The structures of N-(Cbz-Leu-Ala)-ethanolamine (**174**) and the epoxides (**173**) and (**175**) were confirmed by $^{1}$H-NMR, $^{13}$C-NMR, IR spectroscopy, microanalysis and mass spectrometry.

## 2.3 Results

### 2.3.0 The amide substrates

To assess the activity of the synthetic compounds, Kemp's assay procedures were used in which acetyl-leucyl (2S)-arginyl glycyl (2S)-alanyl glycyl (2S)-arginyl (2S)-serinyl (2S)-arginine ($K_M$ = 3.4 mmol dm$^{-3}$) was used as the substrate, see Section 3.1, pages 242 - 256. Other experiments utilised $^{1}$H-NMR spectroscopy to monitor activity, see Section 3.1.

None of the tetrapeptide simple amides (**19**) - (**23**) inhibited or were hydrolysed by the proteinase in 4 hours as judged by $^{1}$H-NMR spectroscopy (Table 2.3 on page 108). Cbz-Leu-Ala-Gly-Gly-$NH_2$ (**19**) and Cbz-Leu-Ala-Gly-Gly-NHMe (**20**) were not soluble in

the assay mixture so the N-protection from both of these amides was removed and Leu-Ala-Gly-Gly-$NH_2$ acetate salt (**25**) and Leu-Ala-Gly-Gly-NHMe acetate salt (**26**) were tested instead. However Leu-Ala-Gly-Gly-$NH_2$ acetate salt (**25**) and Leu-Ala-Gly-Gly-NHMe acetate salt (**26**) also failed to inhibit or be hydrolysed by the proteinase. The removal of the Cbz protection was not significant as Leu-Ala-Gly-Gly-OEt (**74**) was hydrolysed at a similar rate to Cbz-Leu-Ala-Gly-Gly-OEt (**66**).

| Compound | Reduction of initial rate / % [a] |
|---|---|
| Cbz-Leu-Ala-Gly-Gly-$NH_2$ (**19**) | not soluble |
| Cbz-Leu-Ala-Gly-Gly-NHMe (**20**) | not soluble |
| Cbz-Leu-Ala-Gly-Gly-$NMe_2$ (**21**) | 0 |
| Cbz-Leu-Ala-Gly-Gly-NMeEt (**22**) | 0 |
| Cbz-Leu-Ala-Gly-Gly-NMe(OMe) (**23**) | 0 |
| t-Boc-Leu-Ala-Gly-Gly-$NHC_6H_5$*p*-$NO_2$ (**24**) | 26.5[b] |
| Leu-Ala-Gly-Gly-$NH_2$ . HOAc (**25**) | 0 |
| Leu-Ala-Gly-Gly-NHMe . HOAc (**26**) | 0 |

Table 2.3: *Percentage inhibitions for various amides.*

[a] The inhibitors were tested at 1.35 mM in 0.6% DMSO solution, 680 μM substrate, 1.54 μM activating peptide and 200 nM proteinase, using 2 mM phosphate buffer with 100 mM NaCl at pH 7.8 at 37°C against a control that contained no inhibitor. These assays were performed by Dr. Heather Murray.

[b] The inhibitors were tested at 33.8 μM in 0.4% DMSO solution, 680 μM substrate, 1.54 μM activating peptide and 296 nM proteinase, using 2 mM tris buffer with 100 mM NaCl at pH 8.0 at 37°C against a control that contained no inhibitor. The t-Boc group was also hydrolysed and 33% of the urethane had been hydrolysed to the free amine t-BuOH and presumably carbon dioxide.

The simple amides (**19**) - (**23**), (**25**) and (**26**) therefore do not bind to the proteinase, despite possessing the target sequence Leu-Ala-Gly-Gly and additional interactions with the S' side of the proteinase are necessary for proteolytic hydrolysis of an amide substrate.

The more bulky and polar, t-Boc-Leu-Ala-Gly-Gly-$NHC_6H_5$-*p*-$NO_2$ (**2.40**) gave an initial substrate hydrolysis rate reduction of 27% and after 1 hour 67% of the *p*-nitro anilide functionality had been hydrolysed by the proteinase.[a] The rate of hydrolysis of the *p*-nitro anilide was slow compared to the natural substrates for the proteinase.

[a] The substrate was tested at 2.92 mmol $dm^{-3}$ in 0.4% $^2H_6$-DMSO solution, 283 μmol $dm^{-3}$ activating peptide and 2.47μmol $dm^{-3}$ proteinase, using 2 mmol $dm^{-3}$ phosphate buffer at pD 7.2 at 20°C.

## 2.3.1 The ester substrates

The proteinase is not specific for the amide bond and ester bonds are cleaved at comparable rates to the natural substrates (Table 2.4). Considering the Cbz-Leu-Ala-Gly-Gly-esters (**65**) - (**67**), as the size and hydrophobicity of the parent alcohol increases, the inhibition of the initial rate increases from 0% for Cbz-Leu-Ala-Gly-Gly-OMe (**65**) to 13% for Cbz-Leu-Ala-Gly-Gly-OtBu (**67**) (Table 2.4). The use of initial rates avoids complications due to turn over of the inhibitor, thus the degree of inhibition can give a measure of the binding constant for the inhibitor. From this it can be seen that the more hydrophobic esters bind more readily to the proteinase.

However, the rate of proteolytic hydrolysis follows the opposite trend and as the size and hydrophobicity of the parent alcohol increases, the rate of hydrolysis decreases from 100% in less than 10 min for Cbz-Leu-Ala-Gly-Gly-OMe (**65**) to 60% hydrolysis in 10 min for Cbz-Leu-Ala-Gly-Gly-OtBu (**67**). All the esters were hydrolysed to give Cbz-Leu-Ala-Gly-Gly (**76**) and the appropriate alcohol as judged by $^1$H-NMR spectroscopy.

t-Boc-Leu-Ala-Gly-Gly-OBz (**72**) did not inhibit the proteinase and was not hydrolysed by the proteinase after 3 hours. So t-Boc-Leu-Ala-Gly-Gly-OBz (**72**) does not bind to the adenovirus proteinase. Cbz-Leu-Ala-Gly-Gly (**76**), the acid hydrolysis product of an ester substrate, also failed to inhibit the proteinase. N-(Cbz-Leu-Ala-Gly)-ethanolamine (**75**) gave an initial rate percentage inhibition of 15% so the proteinase is able interact with an alcohol functionality in the $P_1$ carbonyl position or the $P_1$' amide nitrogen position.

| compound | Reduction of initial rate / %[a] | Ester hydrolysis rate as judged by $^1$H-NMR.[b] |
|---|---|---|
| Cbz-Leu-Ala-Gly-Gly-OMe (**65**) | 0 | 100% in under 10 min. |
| Cbz-Leu-Ala-Gly-Gly-OEt (**66**) | 5 | 98% in 10 min. |
| Cbz-Leu-Ala-Gly-Gly-OtBu (**67**) | 13 | 60% in 10 min. |
| t-Boc-Leu-Ala-Gly-Gly-OtBu (**73**) | 13 | 85% in 10 min.[c] |
| t-Boc-Leu-Ala-Gly-Gly-OBz (**72**) | 0 | 0 after 3 hours. |
| Cbz-Leu-Ala-Gly-Gly (**76**) | 0 | not applicable |
| Cbz--Leu-Ala-Gly-Gly-$NHCH_2CH_2OH$ (**75**) | 15 | not applicable |

Table 2.4: *Percentage inhibitions and hydrolysis rates of various esters.*

[a] The inhibitors were tested at 33.8 μmol dm$^{-3}$ in 0.4% DMSO solution, 680 μmol dm$^{-3}$ substrate, 1.54 μmol dm$^{-3}$ activating peptide and 296 nmol dm$^{-3}$ proteinase, using 2 mmol dm$^{-3}$ tris buffer with 100 mmol dm$^{-3}$ NaCl at pH 8.0 at 37°C against a control that contained no inhibitor.

[b] The inhibitors were tested at 2.92 mmol dm$^{-3}$ in 0.4% $^2H_6$-DMSO solution, 283 μmol dm$^{-3}$ activating peptide and .47 μmol dm$^{-3}$ proteinase, using 2 mmol dm$^{-3}$ phosphate buffer at pD 7.2 at 20°C.

c Partial t-Boc hydrloysis to t-BuOH and the free amine after 3 hours.

## 2.3.2 The t-butyl urethane substrates

If a t-Boc containing substrate *ie.* t-Boc-Leu-Ala-Gly-Gly-OtBu (**73**) or t-Boc-Leu-Ala-Gly-Gly-$NHC_6H_5$ *p*-$NO_2$ (**72**) the t-Boc-group is also hydrolysed to give the amine, t-butanol and presumably carbon dioxide.

The t-butyl urethane hydrolysis was only observed if the proteinase was present and the hydrolysis was not due to the hydrolysis products adversely affecting the buffer, as a pD of 7.2 was maintained throughout the hydrolysis period. A pH of below 3, or above 12 and heating at 150°C in solution would be required to hydrolyse the urethane group non-enzymatically.[159]

## 2.3.3 Reversible inhibitors

Preliminary results show that most of the inhibitors in Table 2.5 were noncompetitive inhibitors (Section 1.2.3.1.) of the adenovirus proteinase. The double reciprocal (Lineweaver Burke) graphs are given in Appendix A. The most potent noncompetitive inhibitor is Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) with an approximate $K_i$ of 15 μmol $dm^{-3}$. N'N'-(dicarboxybenzyl-Leu-Ala-Gly)-2,4-diamino-but-2-ene aldehyde (**122**) was the best competitive inhibitor for the adenovirus proteinase with an approximate $K_i$ of 6 μmol $dm^{-3}$ (Table 2.5 on page 112). Significantly, this inhibitor is also the only inhibitor that contains residues on the $P_1$' side. Cbz-Leu-Ala-Gly-aminoacetaldehyde (**112**) and O-(Cbz-Leu-Ala-Gly)-1-carboxymethyl,4-hydroxybut-2-ene (**132**) were also potent inhibitors for the adenovirus proteinase, both having estimated $K_i$'s of 10 μmol $dm^{-3}$. It was not possible to find the actual $K_i$'s for the aldehyde (**112**) and the Michael acceptor (**132**) from double reciprocal plots as the proteinase denatured, in several instances before the necessary measurements could be taken.

| X-Leu-Ala-Gly-Y | inhibition[a] / %[b] | $K_i$[c] | Type inhibition[f] |
|---|---|---|---|
| X = Cbz,<br>Y = $NHCH_2CN$ (**78**) | 52 | 15 | noncompetitive |
| X = t-Boc,<br>Y = $NHCH_2CN$ (**79**) | 46 | 20 | noncompetitive |
| X = Cbz,<br>Y = $NHCH_2CH{=}CH_2$ (**115**) | 17 | 130[d] | na[e] |
| X = Cbz,<br>Y = $NHCH_2CMe{=}CH_2$ (**118**) | 27 | 45 | noncompetitive |
| X = Cbz,<br>Y = $NHCH_2CH(OMe)_2$ (**114**) | 20 | 96 | noncompetitive |
| X = t-Boc,<br>Y = $NHCH_2CH(OMe)_2$ (**113**) | 12 | 200[d] | na[e] |
| X = Cbz,<br>Y = $NHCH_2CMeO$ (**116**) | 17 | 130[d] | na[e] |
| X = Cbz,<br>Y = $NHCH_2CHO$ (**112**) | 68 | 10[d] | na[e] |
| X = Cbz,<br>Y = $NHCR{=}CHCH_2$NH-GAL-Cbz,<br>R = CHO (**122**), | 89 | 6 | competitive |
| X = Cbz, (**132**)<br>Y = $OCH_2CH{=}CHCO_2Me$ (*E*) | 66 | 10[d] | na[e] |

Table 2.5: *Percentage inhibitions, $K_i$ and type for various reversible inhibitors.*

Footnotes for Table 2.5.

[a] Measured as a reduction in initial rate.

[b] The inhibitors were tested at 33.8 μmol $dm^{-3}$ in 0.4% DMSO solution, 680 μmol $dm^{-3}$ substrate, 1.54 μmol $dm^{-3}$ activating peptide and 296 nmol $dm^{-3}$ proteinase, using 2 mmol$dm^{-3}$ tris buffer with 100 mmol $dm^{-3}$ NaCl at pH 8.0 at 37°C against a control that contained no inhibitor.

[c] approximate $K_i$ measured in μmol $dm^{-3}$.

[d] Estimated $K_i$, assuming a linear relationship between initial-rate-%-inhibition and $K_i$.

[e] na = not applicable.

[f] determined from rate experiments using 5 different [S] and 3 different [I]. Each rate was determined once.

Unexpectedly, proteinase pre-incubated with the Michael acceptor, O-(carboxybenzyl-Leu-Ala-Gly)-1-carboxymethyl,4-hydroxybut-2-ene (**132**), showed no inactivation with preincubation times of 3, 5 and 7 min. So the Michael acceptor (**132**) was also a reversible inhibitor, rather than an irreversible inhibitor as expected from inhibitor studies with other cysteine proteinases (Section 2.2.0).

The Cbz-protected tetrapeptide inhibitors, *ie.* Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) and Cbz-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal (**114**) were better than their corresponding t-Boc-protected analogues, ie. t-Boc-Leu-Ala-Gly-aminoacetonitrile (**79**) and t-Boc-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal (**113**) (Table 2.5).

### *2.3.3.0 Nitriles*

The nature of the $P_4$ N-protection affects the degree of binding of Leu-Ala-Gly-aminoacetonitrile inhibitors to the proteinase (Table 2.6 on page 114, compounds (**78**), (**79**) and (**82**)). The proteinase prefers a polarisable π-electron system in the $P_5$-$P_6$ position of a substrate to a lipophilic, t-butyl group, or an electron rich

trifluoromethyl group. No hydrolysis of Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) was observed at low proteinase concentration by $^1$H-NMR spectroscopy.[c]

Successive removal of the $P_4$ and $P_3$ residues of TFA-Leu-Ala-Gly-aminoacetonitrile (**78**) resulted in the total loss of inhibition. The $P_4$ leucine residue is responsible for most of the binding interactions of the nitrile inhibitor with the proteinase and removal of the $P_4$ leucine residue results in a 90% loss of inhibition (Table 2.6).

| Compound | reduction of initial rate % [a] | reduction of initial rate %[b] |
|---|---|---|
| Cbz-Leu-Ala-Gly-$NHCH_2CN$ (**78**) | 52 | 95 |
| t-Boc-Leu-Ala-Gly-$NHCH_2CN$ (**79**) | 46 | 54 |
| Cbz-Leu-Ala-Ala-$NHCH_2CN$ (**80**) | 26 | not applicable |
| Cbz-Leu-Ala-(2R)-Ala-$NHCH_2CN$ (**81**) | 16 | not applicable |
| TFA-Leu-Ala-Gly-$NHCH_2CN$ (**82**) | not applicable | 50 |
| TFA-Ala-Gly-$NH_2CH_2CN$ (**83**) | not applicable | 5 |
| TFA-Gly-$NH_2CH_2CN$ (**84**) | 0 | 0 |
| $NH_2CH_2CN$.HCl (**85**) | 0 | 0 |

Table 2.6 : *Percentage inhibitions for various nitriles.*

[c] The inhibitors were tested at 2.92mmol dm$^{-3}$ in 0.4% $^2H_6$-DMSO solution, 283μmol dm$^{-3}$ activating peptide and 2.47μmol dm$^{-3}$ proteinase, using 2mmol dm$^{-3}$ phosphate buffer at pD 7.2 at 20°C.

Footnotes for Table 2.6.

[a] The inhibitors were tested at 1.35 mmol $dm^{-3}$ in 0.6% DMSO solution, 680 μ mol $dm^{-3}$ substrate, 1.54 μmol $dm^{-3}$ activating peptide and 200 nmol $dm^{-3}$ proteinase, using 2 mmol $dm^{-3}$ phosphate buffer with 100mmol $dm^{-3}$ NaCl at pH 7.8 at 37°C against a control that contained no inhibitor.

[b] The inhibitors were tested at 33.8 μmol $dm^{-3}$ in 0.4% DMSO solution, 680μ mol $dm^{-3}$ substrate, 1.54 μmol $dm^{-3}$ activating peptide and 296 nmol $dm^{-3}$ proteinase, using 2 mmol $dm^{-3}$ tris buffer with 100 mmol $dm^{-3}$ NaCl at pH 8.0 at 37°C against a control that contained no inhibitor by Dr Heather Murray.

Replacement of the $P_2$ glycine residue of Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) with (2S)-alanyl, *ie.* Cbz-Leu-Ala-Ala-aminoacetonitrile (**80**) or (2R)-alanyl, *ie.* Cbz-Leu-Ala-(2R)-Ala-aminoacetonitrile (**81**) results in a 50% and 70% loss of binding to the proteinase, respectively. So although, either a (2S) or (2R)-alanyl residue can replace the $P_2$ glycyl residue in the nitrile inhibitor, the $P_2$ binding site binds a glycine residue preferentially. Unfortunately there is no data available on the binding preferences of the $P_2$ binding position of typical substrates for the proteinase.

### 2.3.3.1 The $^{13}$C-labelled nitrile / proteinase $^{13}$C-NMR experiments

Mixtures of the activated proteinase and [1-$^{13}$C]-Cbz-Leu-Ala-Gly-aminoacetonitrile (**86**) where studied by $^{13}$C-NMR spectroscopy to establish whether the nitrile inhibitor (**86**) and the adenovirus proteinase form a thioamidate adduct in a similar manner to papain (Section 2.1.3).

The first $^{13}$C-NMR experiment (conditions and concentrations in Table 2.7 on page 116) of the proteinase and [1-$^{13}$C]-Cbz-Leu-Ala-Gly-aminoacetonitrile (**86**) showed that [1-$^{13}$C]-Cbz-Leu-Ala-Gly-aminoacetonitrile (**86**) formed a thioamidate adduct with the mercaptoethanol, that had been added to stabilise the proteinase. Any thioamidate adduct

formed between the proteinase and the $^{13}$C-labelled inhibitor would be masked by the mercaptoethanol-$^{13}$C-labelled inhibitor adduct.

The experiment was repeated (experiment 2, conditions and concentrations in Table 2.7), using slightly more proteinase and less inhibitor so the ratio between the bound and unbound forms of the $^{13}$C-labelled inhibitor (**86**) was lower, and omitting the mercaptoethanol. However, after 3 hours of acquisition, only the signal due to the unbound $^{13}$C-labelled inhibitor (**86**) was present. No evidence of a thioamidate adduct was observed. A small sample of proteinase, taken from the experiment before the addition of the $^{13}$C-labelled inhibitor (**86**) and incubated for an hour at 25°C, before proteolytic assay, showed no activity. Further studies carried out by Dr Kemp at the St Andrews Biochemistry department, showed that the proteinase was not active in $D_2O$.

| Experiment | 1 | 2 | 3 |
|---|---|---|---|
| Proteinase | 23.9 nmol | 52.2 nmol | 217 nmol |
| Activating peptide | 2.39 μmol | 6.40 μmol | 21.7 μmol |
| DMSO | 2.2% | 0.6% | 0.04% |
| 1% mercaptoethanol solution | 8.6% | 0 | 0 |
| $^{13}$C-labelled inhibitor | 2310 nmol | 694 nmol | 868 nmol |
| Buffer | 50 mmol dm$^{-3}$ Tris pD 8.0 | 50 mmol dm$^{-3}$ Tris pD 8.0 | 50 mmol dm$^{-3}$ phosphate pD 7.2 |
| Solvent | $D_2O$ | $D_2O$ | $H_2O$ + 10% $D_2O$ |

Table 2.7: *The moles of reagents used in the $^{13}$C-NMR experiments 1-3 to determine whether or not the proteinase forms a thioamidate adduct with the nitrile inhibitors.*

A third $^{13}$C-NMR experiment (experiment 3, concentrations and conditions in Table 2.7) was attempted using a 10 mm NMR tube, to increase the sensitivity of the $^{13}$C-NMR experiment, 10% $D_2O$ in $H_2O$ was used as the solvent and phosphate buffer instead of Tris buffer to avoid extra signals in the $^{13}$C-NMR spectrum. However, after 3 hours of acquisition, only the signal due to the unbound $^{13}$C-labelled inhibitor (**86**) was present. No evidence of a thioamidate adduct was observed. A small sample of proteinase, taken from the experiment before the addition of the $^{13}$C-labelled inhibitor (**86**) and incubated for 30 min. at 25°C, before the proteolytic assay was carried out, again showed no activity. It would appear that the proteinase is inactivated by one or more of the conditions used in the NMR assay. Consequentially the proteinase was too unstable, to see whether or not a hemithioamidate adduct formed between the nitrile inhibitor and the proteinase by $^{13}$C-NMR techniques.

## 2.3.4 Irreversible inhibitors

The epoxides (**173**) and (**175**) and the bromide (**130**) were moderate inhibitors of the proteinase (Table 2.8 on page 118). Both epoxides were stable in the assay buffer, and after they had been in the buffer for 1 hour, showed no signs of ring opening detectable by $^1$H NMR spectroscopy.

It was possible that the epoxides (**173**) and (**175**) and the bromide (**130**) irreversibly alkylated the proteinase, but bound poorly to the proteinase. However, preincubation of these inhibitors with the proteinase for 3, 5 or 7 min prior to the addition of the substrate did not result in any significant inactivation of the proteinase with time. Therefore the epoxides (**173**) and (**175**) and the bromide (**130**) do not irreversibly inhibit the adenovirus type 2 proteinase.

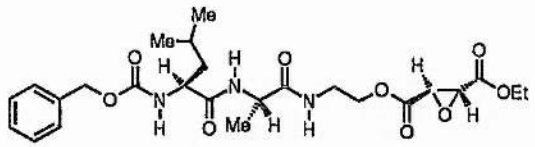
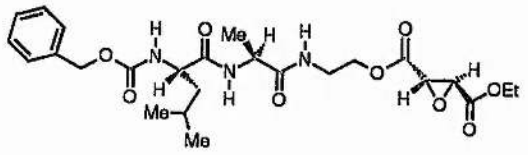
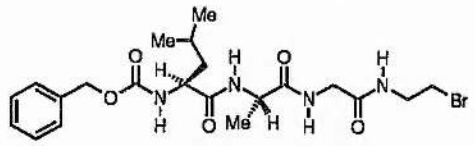
| inhibitor | | Reduction of initial rate / %[a] |
|---|---|---|
|  | **(173)** | 18 |
|  | **(175)** | 11 |
|  | **(130)** | 9 |

Table 2.8: *percentage inhibitions of various potential irreversible inhibitors.*

[a] The inhibitors were tested at 33.8 μmmol $dm^{-3}$ in 0.4% DMSO solution, 680 μmol $dm^{-3}$ substrate, 1.54 μmmol $dm^{-3}$ activating peptide and 296 nmol $dm^{-3}$ proteinase, using 2 mmol $dm^{-3}$ tris buffer with 100 mmol $dm^{-3}$ NaCl at pH 8.0 at 37°C against a control that contained no inhibitor.

# 2.4 Discussion

## 2.4.0 The simple amides

None of the simple amides (Table 2.3, Section 2.3.0) bound to the proteinase as no inhibition or hydrolysis was observed. However t-Boc-Leu-Ala-Gly-Gly-*p*-nitroanilide (**24**) bound to the proteinase (26.5%-initial-rate percentage-inhibition) and was hydrolysed fairly slowly (67% of the *p*-nitroanilide functionality was hydrolysed after 1 hour) compared to the esters (Table 2.4, section 2.3.1) and the natural substrates for the proteinase. The lack of hydrolysis of Leu-Ala-Gly-Gly-$NH_2$ (**25**) was particularly interesting, as papain hydrolyses the amide functionality (benzoyl-Gly-amide (**176**), $K_s$ = 0.2 mol $dm^{-3}$).[160]

All the simple amides contain the $P_1$-$P_4$ sequence of residues, but there are no residues on the P' side of a substrate. The proteinase, does not require a specific sequence of amino acids on the P' side of a substrate[29, 73], but may require further interactions with the peptide backbone of the P' side of a substrate for it to bind. The simple amides do not possess the necessary P' amide bond and so cannot bind to the proteinase.

t-Boc-Leu-Ala-Gly-Gly- *p*-nitroanilide (**24**) contains a large, polar *p*-nitroanilide group corresponding to both the $P_1$' and $P_2$' residues and is hydrolysed by the proteinase. The *p*-nitroanilide group is therefore able to fulfill the role of the P'-binding group.

Comparison of the amide back bone of a substrate with the *p*-nitroanilide group (Fig. 2.22) shows that one of the oxygens of the *p*-nitro group of the *p*-nitroanilide lies in a similar position to the carbonyl oxygen of the $P_2$' residue. The dihedral angles of the substrate are not known and therefore the exact positions of the carbonyl groups of the substrate cannot be known.

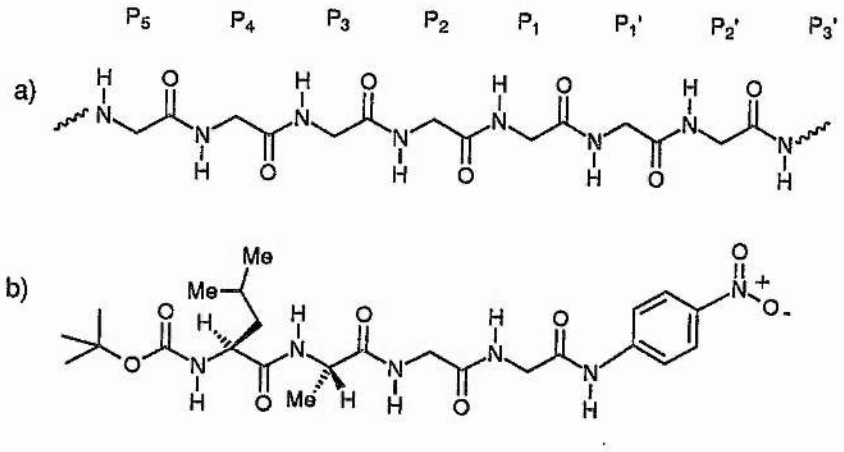



Fig. 2.22:

*A comparison of a) the amide backone of a substrate with b) the p-nitroanilide group.*

This suggests that the proteinase requires a carbonyl group at the $P_2$' position to successfully bind and hydrolyse an amide substrate. The other simple amides lack any such functional group and so cannot bind to the proteinase and function niether as substrates or inhibitors.

*p*-Nitroaniline may also be able to fulfill this requirement in a similar manner to the *p*-nitroanilide (**24**). This could cause amine product inhibition of the proteinase, particularly as *p*-nitroaniline is rather insoluble in water and would account for the slow rate of hydrolysis of the *p*-nitroanilide group.

## 2.4.1 The ester substrates

In contrast to the amides all the esters (Table 2.4, section 2.3.1), except Cbz-Leu-Ala-Gly-Gly-OBz (**72**) bound to the proteinase. However the ester functionality can adopt more conformations than the amide functionality as rotation about the oxygen-alkyl bond is not restricted. In an amide bond the carbonyl bond and the amide nitrogen lone

pair are in conjugation so the nitrogen-carbon and the carbon-oxygen bonds are coplanar to obtain maximal π-orbital over-lap.

The adenovirus proteinase may bind an amide substrate in a conformation that strains the $P_1$-$P_1$' amide bond so the $P_1$ carbonyl group and the $P_1$' amide nitrogen are no longer co-planar. The $P_1$' amide electrons would relocalise on the nitrogen, destabilising the amide bond and facilitating its hydrolysis. A proton situated in close proximity to the $P_1$' amide nitrogen would also achieve the same result by polarising the delocalised amide nitrogen lone pair back towards the nitrogen in the amide bond.

The simple amides can not adopt this 'strained conformation' as they do not interact with the $P_2$' carbonyl binding pocket. An ester faces no energy penalty if the $P_1$ carbonyl carbon and the alkyl-oxygen are not co-planar, so the ester can adopt the 'binding conformation' without the need for the $P_2$' carbonyl interaction with the proteinase.

As the size and hydrophobicity of the alcohol increases, the degree of binding to the proteinase also increases (Table 2.4, section 2.3.1). The binding order of the esters could be due to solubility effects, if the active site of the proteinase is hydrophobic. The more lipophilic the the ester, the less soluble the ester will be in water so the more likely the ester will bind to the proteinase to avoid the polar environment in solution. Alternatively the proteinase usually binds a glycine or alanine residue at $P_1$'[73] so the binding order of the esters could also be due to possible binding interactions with the $S_1$' side of the proteinase. The t-butyl ester and to a lesser extent the ethyl ester are able to mimic the size and shape of the glycine or alanine residue at $P_1$' whereas the methyl ester is too small (Fig. 2.23).

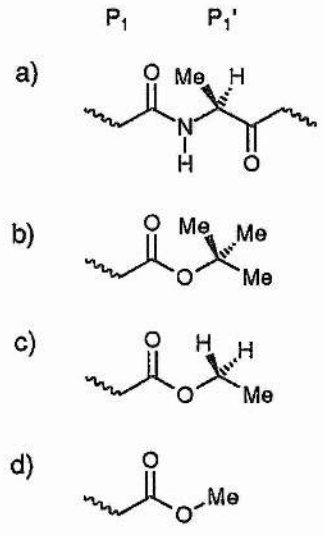



Fig. 2.23:

*Comparison of a) the substrate; b) t-butyl ester; c) ethyl ester and d) methyl ester.*

t-Boc-Leu-Ala-Gly-Gly-OBz (72) failed to bind to the proteinase. While the benzyl ester is of similar size to the *p*-nitroanilide group (Fig. 2.24) it does not occupy the same space or possess the same charge distribution. It therefore seems likely that the benzyl ester cannot fit into the active site of the proteinase.

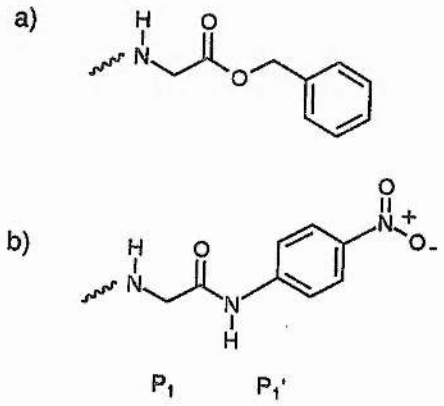



Fig. 2.24:

*The comparison of a) The benzyl ester and b) The p-nitroanilide group.*

The order of the rates of hydrolysis of the esters by the proteinase (Table 2.4, Section 2.3.1) follows the opposite trend to the binding order of the esters. If neither of the products of the ester hydrolysis bound to the proteinase, the substrate was saturating the enzyme's active site and the binding of the substrate was the rate limiting step then the order of hydrolysis would follow the opposite trend to that observed. The better the substrate binds to the proteinase (Fig. 2.25a on page 124), the lower the energy of the transition state complex, $[ES]^{\ddagger}$ compared to the unbound substrate and enzyme, [E + S] and the lower the change in Gibbs free energy of activation ($\Delta G_2^{\ddagger} < \Delta G_1^{\ddagger}$). As the change in activation energy is inversely proportional to the rate of reaction (Section 1.2.4, equation 1.2.27), the better the substrate binds to the proteinase, the faster the reaction rate.

If the substrate was not saturating the proteinase's active site and the substrate binding was the rate limiting step, the order of ester hydrolysis would also follow the opposite trend to that observed, as is the case when the substrate is saturating the active site of the enzyme. However two effects are responsible; the substrate binds better to the proteinase so the activation energy falls and [ES] increases. Both these effects increase the rate of hydrolysis as the substrate binds better to the enzyme.

If the substrate was saturating the enzyme's active site and the binding of the substrate was not the rate limiting step (Fig. 2.25b), the rate of hydrolysis of the esters would follow the observed trend. The better the substrate binds to the enzyme, the lower in energy the enzyme-substrate complex, [ES] and the higher the change in Gibbs free energy of activation ($\Delta G_4^{\ddagger} > \Delta G_3^{\ddagger}$) and the slower the reaction rate.

If the substrate was not saturating the proteinase's active site and the substrate binding was not the rate limiting step, the kinetics are rather complex, As $K_M$ falls the [ES] increases so the rate of reaction is increased, however ES also becomes lower in energy so the barrier height for the reaction increases, slowing the reaction rate down. These two effects are

antagonistic and which dominates depends on how close the substrate is to saturating the enzyme's active site. As the active site of the enzyme becomes saturated, the degree of binding of the substrate to the enzyme is less important and the overall reaction rate is dominated by the barrier height, to an increasing extent. The binding of the esters to the proteinase does not appear to be rate limiting so the active site of the enzyme must be close to saturation with the substrate or saturated with substrate.

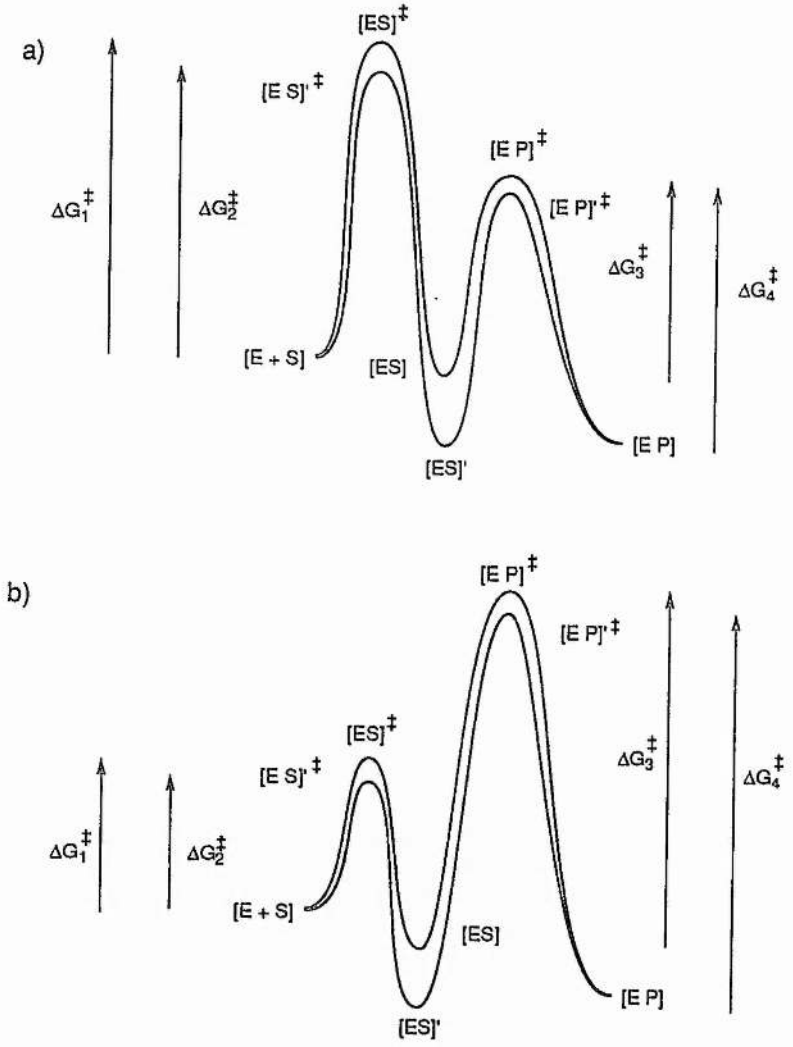



Fig. 2.25: *A comparison of the energy diagrams for two substrates with different binding affinities for a proteinase a) when the enzyme is saturated by the substrate and b) when the enzyme is not saturated by the substrate.*

On hydrolysis, each of the esters, Cbz-Leu-Ala-Gly-Gly-OMe (**65**), Cbz-Leu-Ala-Gly-Gly-OEt (**66**) and Cbz-Leu-Ala-Gly-Gly-O-tBu (**67**) gives the same acid hydrolysis product (Cbz-Leu-Ala-Gly-Gly (**76**)), and the appropriate alcohol.

Cbz-Leu-Ala-Gly-Gly (**76**) does not inhibit the adenovirus proteinase (Table 2.4, section 2.3.1). This is in contrast to papain where the acid hydrolysis product (N-benzoyl-glycine (**177**), $K_i$ 17 mM) does inhibit the proteinase.[122] The lack of inhibition of Cbz-Leu-Ala-Gly-Gly (**76**) could be due to a charge repulsion effect between the negatively charged carboxylate acid functionality and either the proteinase or the amine product, as the acid will not be protonated at pH 8.

The proteinase may bind the alcohol product significantly and slow down departure from the active site following hydrolysis of the ester bond. The rate limiting step in the proteolytic hydrolysis of the ester bond could change from the binding of the substrate to the loss of the alcohol product and the observed rate of hydrolysis would also be followed.

The observed changes in the rates of ester hydrolysis could also be due to an electronic effect encountered during the rate limiting step of the hydrolysis mechanism of the proteinase. If the breaking of the carbon-nitrogen bond was the rate limiting step, as is the case for papain,[161] then the nature of the leaving group may play a major role in determining the reaction rate.

The rate of hydrolysis of the esters would decrease with the size of the alkyl chain of the alcohol, if the oxygen leaving group left as an alkoxide. The electron donating capacity of an alkyl chain increases with the size of an alkyl chain, so an alkoxide would be progressively destabilised if the size of the alkyl chain increases, as more negative charge would be pushed onto the already negatively charged oxygen. This situation would result in a slower rate of hydrolysis as the size of the alcohol side chain increases, which is in agreement with the rates of ester hydrolysis presented in proteinase Table 2.4, Section 2.3.1.

The effect of the leaving group on the rate of hydrolysis is more complex, if the oxygen leaving group is an oxonium species, two opposite effects control the basicity of the alkyl oxygen as the size of the alkyl chain increases. The basicity of the alkyl oxygen would increase as the size of the alkyl chain increases due to the increasing capacity of the alkyl chain to donate electrons to the alkyl oxygen, however as the alkyl chain increases in size, it starts to hinder the approach of the in coming proton so after a certain size, the basicity of the alkyl oxygen will start to decrease. The point at which the size of the ester might start to hinder the in coming proton would depend on the direction of alkyl oxygen protonation by the proteinase, which is unknown. However, the fact that t-butyl esters are hydrolysed at similar rates to the methyl and ethyl esters argues against rate limiting steric effects in attacking the carbonyl group.

## 2.4.2 The t-butyl-urethane substrate

The proteinase hydrolysed the t-Boc-nitrogen protecting group from t-Boc-Leu-Ala-Gly-Gly- *p*-nitroanilide (**24**) and t-Boc-Leu-Ala-Gly-Gly-OtBu (**73**) at a slow rate to give the free amine, t-butanol and presumably carbon dioxide (Section 2.3.2).

The urethane hydrolysis was not due to an acidic environment in the $P_5$ binding pocket, as the acid product of the hydrolysis does not bind to the proteinase (section 2.4.1), so the t-butyl urethane hydrolysis would finish at the same time as the hydrolysis of the $P_1$-$P_1$' bond. This was not observed as t-butyl urethane hydrolysis continued after the hydrolysis of the $P_1$-$P_1$' bond.

The t-butyl urethane hydrolysis was also not due to the hydrolysis products adversely affecting the buffer, as a pD of 7.2 was maintained throughout the 3 hour hydrolysis procedure and the controls, where the proteinase was omitted showed no hydrolysis of the t-butyl urethane.

The residues in the t-Boc-protected substrate may shift towards the S' side of the proteinase, the leucine residue would occupy the $P_1$' position and the t-butyl urethane would occupy the $P_1$ position. The t-butyl urethane's size is not very different to the glycine or alanine residue, that usually occupys the $P_1$ position (Fig. 2.26). However, the electrostatic requirements of the urethane are different to the $P_1$ residue as an oxygen takes the place of the α-carbon of the aminoacid residue. Once the t-butyl urethane has bound in the $P_1$-$P_1$' position, the proteinase could then presumably hydrolyse the t-butyl urethane functionality in a similar manner to the amide and ester bond.

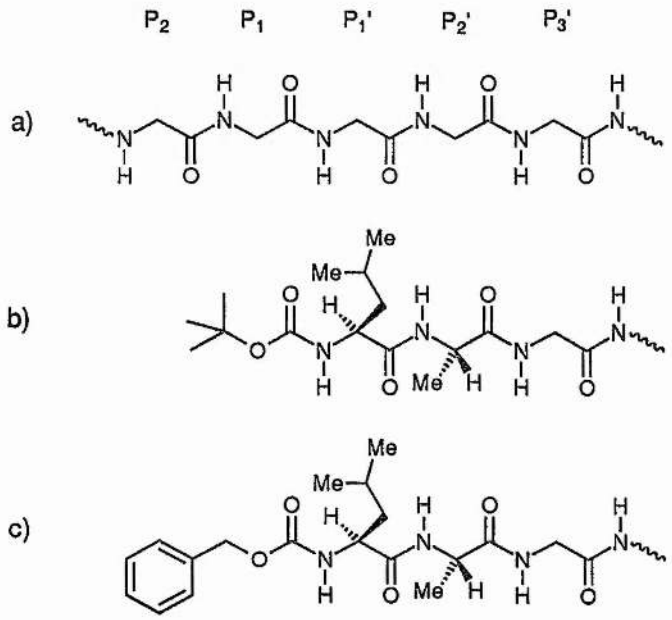
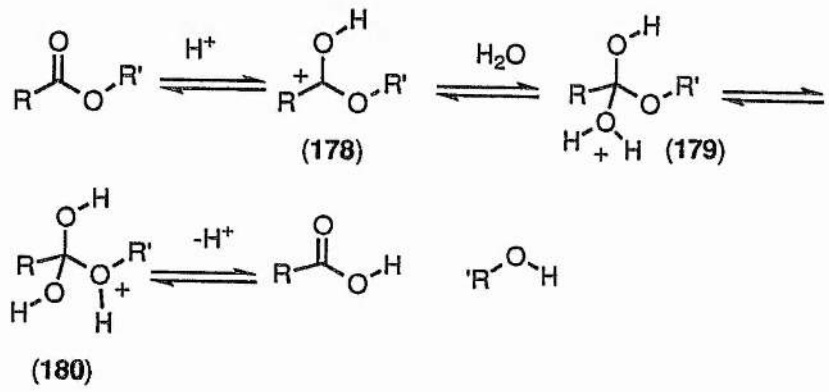



Fig 2.26: *A comparison of a) The substrate; b) The t-Boc-protected substrate and c) The Cbz-protected substrate.*

It would also be possible for the Cbz-protected substrates to shift towards the S' side of the proteinase, however in this case the benzyl group may be too big to fit into $S_1$ so no hydrolysis of the benzyl urethane was detected.

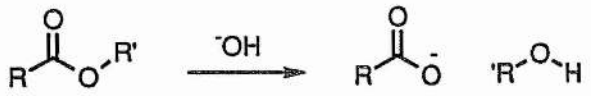
## 2.4.3 The possible mechanism of hydrolysis of the esters

If the possible solution mechanisms and products of ester hydrolysis in solution are compared to the products of proteolytic ester hydrolysis, it may be possible to gain an insight into the mechanism of the adenovirus proteinase. There are six observed solution mechanisms of ester bond hydrolysis; $A_{AC}2$ (Scheme 2.38), $B_{AL}2$ (Scheme 2.39), $A_{AC}1$ (Scheme 2.40), $B_{AC}2$ (Scheme 2.41), $B_{AL}1$ (Scheme 2.42) and $A_{AL}1$ (Scheme 2.43). The $A_{AC}2$ mechanism (Scheme 2.38) involves the protonation of the carbonyl oxygen to give a carbocation (**178**) which is nucleophilically attacked by water to give the tetrahedral intermediate (**179**). Proton exchange gives the tetrahedral intermediate (**180**) which eliminates the alcohol to give the carboxylic acid after deprotonation.




Scheme 2.38: *The $A_{AC}2$ mechanism of ester hydrolysis.*

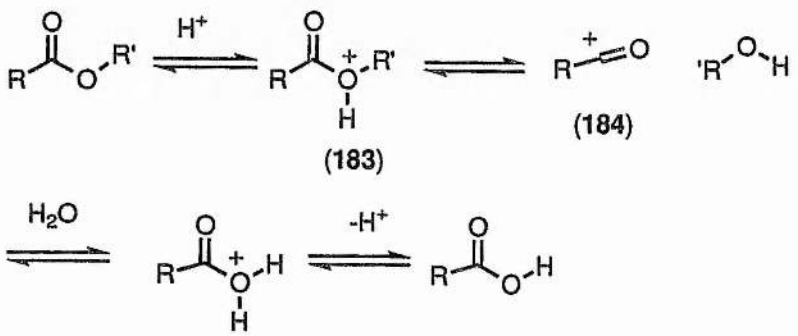
The $B_{AL}2$ mechanism (Scheme 2.39) is rather rare as the carbonyl group must be extremely hindered as the hydroxide ion directly attacks the alkyl group of the ester to give the alcohol and the carboxylate as the products. This mechanism has been observed for the hydrolysis of β-butyrolactone (**181**) and β-propiolactone (**182**) under neutral conditions.[162] The other mechanisms are considered below.



Scheme 2.39: *The $B_{AL}2$ mechanism of ester hydrolysis.*

It is likely that the proteolytic hydrolysis of the ester and amide bonds proceeds by the same mechanism, so only the ester hydrolysis mechanisms involving acyl-oxygen bond cleavage are considered as alkyl-nitrogen bond cleavage does not occur during amide hydrolysis *ie.* the $A_{AC}1$, $A_{AC}2$ and $B_{AC}2$ mechanisms. These mechanisms are very similar to the solution methods of amide hydrolysis discussed previously in Section 1.1.

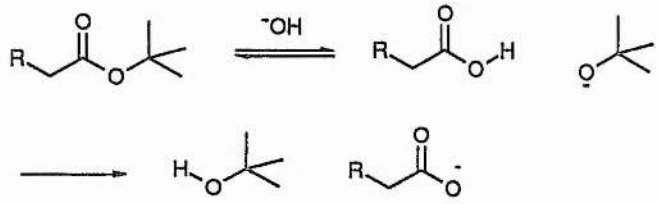
The $A_{AC}1$ mechanism (Scheme 2.40) is rather rare in solution, and occurs when the water nucleophile is prevented from attacking the carbonyl carbon by a bulky acyl group eg 2, 4, 6-trimethyl benzyl.[163] The alkyl oxygen of the ester becomes protonated to give the carbocation (**183**) and the alcohol is eliminated to give the carbocation (**184**) which is hydrated and deprotonated to give the carboxylic acid.



Scheme 2.40: *The $A_{AC}1$ ester hydrolysis mechanism.*

In solution, the acid ($A_{AC}2$ mechanism, Scheme 2.38) or the base catalysed ($B_{AC}2$ mechanism, Scheme 2.41) hydrolysis of Cbz-Leu-Ala-Gly-Gly-OMe (**65**) and Cbz-Leu-Ala-Gly-Gly-OEt (**66**) would give rise to Cbz-Leu-Ala-Gly-Gly (**76**) and the appropriate alcohol, the same products as observed for the proteolytic hydrolysis of these compounds. However, in solution both the acid and base catalysed hydrolysis of Cbz-Leu-Ala-Gly-Gly-OtBu (**67**) give different alkyl hydrolysis products (t-butanol in strong base, isobutene in weakly basic and neutral solution and t-butanol in strongly acidic solution), as the t-butyl group can form a stabilised carbocation in solution.[164] The proteolytic hydrolysis of Cbz-Leu-Ala-Gly-Gly-OtBu (**67**) gives t-butanol as the only alkyl hydrolysis product.

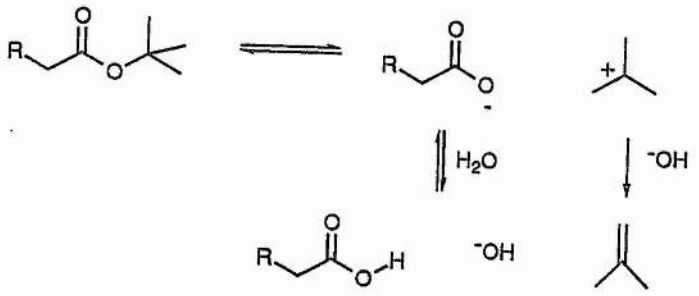
In a strongly basic solution, the t-butyl ester is hydrolysed by a $B_{AC}2$ mechanism to give t-butanol and the carboxylic acid as products (Scheme 2.41). The t-butoxide anion is nucleophilically displaced from the t-butyl ester by the hydroxide ion to give the carboxylic acid, which is deprotonated by the t-butoxide anion to give t-butanol and the carboxylate ion (Scheme 2.41).




Scheme 2.41: *The $B_{AC}2$ hydrolysis mechanism.*

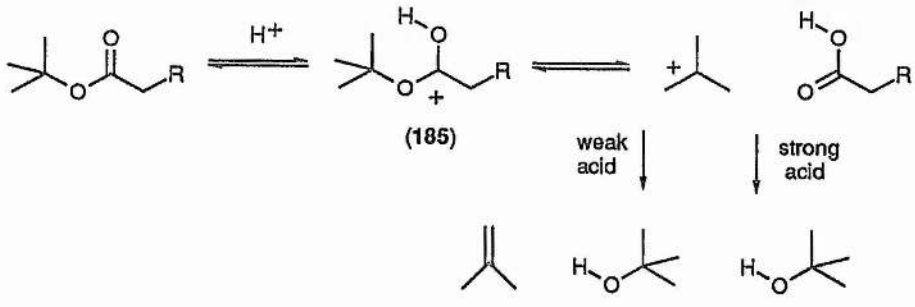
However, in a neutral, or a weakly basic solution a t-butyl ester is hydrolysed by a $B_{AL}1$ mechanism (Scheme 2.42). The solvation of the t-butyl ester results in alkyl-oxygen bond

cleavage to give the carboxylate anion and the t-butyl cation which is deprotonated by a hydroxide ion to give isobutene.[165]




Scheme 2.42: *The $B_{AL}1$ hydrolysis mechanism.*

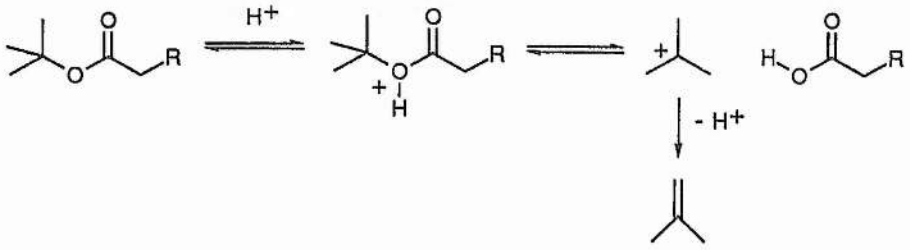
In acidic solution, t-butyl ester hydrolysis occurs via an $A_{AL}1$ mechanism (Scheme 2.43). Protonation of the carbonyl oxygen results in a carbocation (**185**) which in turn eliminates a t-butyl carbocation leaving the carboxylic acid. In strongly acidic solution, the t-butyl carbocation becomes hydrated, giving t-butanol,[165] as it is unlikely that one of the methyl groups of the t-butyl carbocation will lose a proton in strongly acidic solution. However in dilute acid the t-butyl carbocation will deprotonate more readily giving isobutene. Some t-butanol will be produced as this process occurs in competition with hydration described above.




Scheme 2.43: *The $A_{AL}1$ hydrolysis mechanism.*

The proteinase may possess any of the catalytic functionalities involved in the solution hydrolysis, as the proteinase can have both 'acidic' and 'basic' enviroments together in its active site. In addition, the proteolytic hydrolysis mechanism can proceed in either a sequential, or a concerted manner.

Alkyl oxygen protonation may also be possible in the proteinase active site, as the proteinase has the capacity to protonate specific groups on a substrate in a regioselective manner and they need not be the most basic site in solution. Alkyl oxygen protonation of a t-butyl ester would result in the rapid elimination of the t-butyl carbocation, leaving the carboxylic acid. This t-butyl carbocation could either deprotonate, giving isobutene, or be hydrated to give t-butanol and the catalytic hydroxonium ion (Scheme 2.44).




Scheme 2.44: *Alkyl oxgen protonation of a t-butyl ester to effect hydrolysis.*

The proteinase catalysed the hydrolysis of Cbz-Leu-Ala-Gly-Gly-OtBu (**67**) and t-Boc-Leu-Ala-Gly-Gly-OtBu (**73**) to give Cbz-Leu-Ala-Gly-Gly (**76**) or t-Boc-Leu-Ala-Gly-Gly (**186**) and t-butanol. If the first step in the hydrolysis of the ester bond involved significant protonation of the carbonyl carbon or protonation of the alkyl oxygen, the t-butyl ester would decompose to give the carboxylic acid and the t-butyl carbocation, before the thiolate nucleophile could attack the carbonyl carbon. The carboxylic acid would then diffuse away from the active site of the proteinase, but the fate of the t-butyl carbocation is more complex. The carbocation could either react further with the proteinase or be released into

the solvent. The proteinase should not stabilise the carbocation as that particular binding pocket is suited to binding an amide nitrogen and is likely to involve hydrogen bond donors or Lewis acids, both of which will greatly destabilise any carbocation present. The carbocation could also alkylate a susceptable group within the proteinase but this would probably lead to inactivation of the proteinase. However as Cbz-Leu-Ala-Gly-Gly-OtBu (**67**) and t-Boc-Leu-Ala-Gly-Gly-OtBu (**73**) are substrates alkylation of the proteinase does not occur. The carbocation has no alternative but to diffuse away from the active site of the proteinase into the solution, where at the weakly alkaline pH the t-butyl carbocation would be deprotonated to give isobutene.

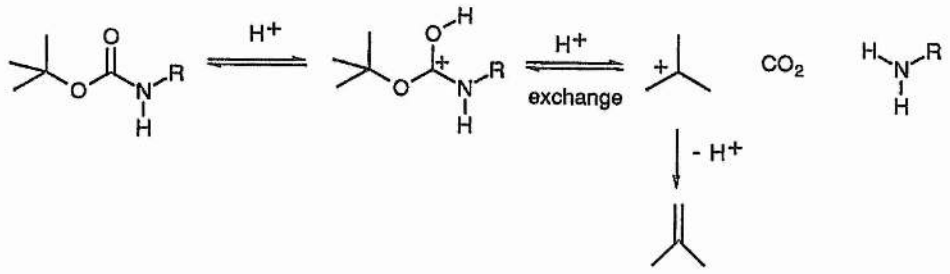
As no significant amounts of isobutene were observed (detected by the loss of the t-butyl signals in the $^1$H-NMR spectra) during the hydrolysis of Cbz-Leu-Ala-Gly-Gly-OtBu (**67**) and t-Boc-Leu-Ala-Gly-Gly-OtBu (**73**), the proteinase is unlikely to catalyse the hydrolysis of an ester bond and consequently an amide bond, by significant protonation of the carbonyl oxygen or the leaving group in the first step.

As the adenovirus proteinase should catalyse the hydrolysis of all its substrates by the same mechanism, the only mechanism which would produce t-butanol is the strong base catalysed hydrolysis mechanism ($B_{AC}2$) (Scheme 2.41).

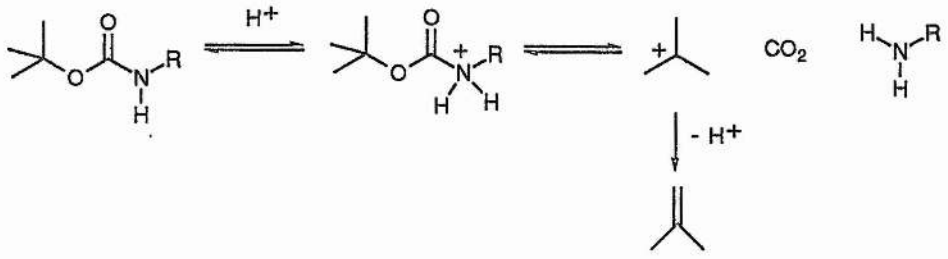
## 2.4.4 The hydrolysis of the t-butyl urethane

A predominantly base catalysed hydrolysis mechanism as described above is consistent with the proteolytic hydrolysis of the t-butyl urethane substrate, which resulted in the free amine, t-butanol and carbon dioxide (section 2.3.2).

As is the case for the t-butyl ester (section 2.4.3), protonation of the carbonyl oxygen (Scheme 2.45) or the urethane nitrogen (Scheme 2.46) would result in the production of isobutene rather than the observed hydrolysis product of t-butanol.




Scheme 2.45: *The hydrolysis of a t-butyl urethane by protonation of the carbonyl oxygen.*

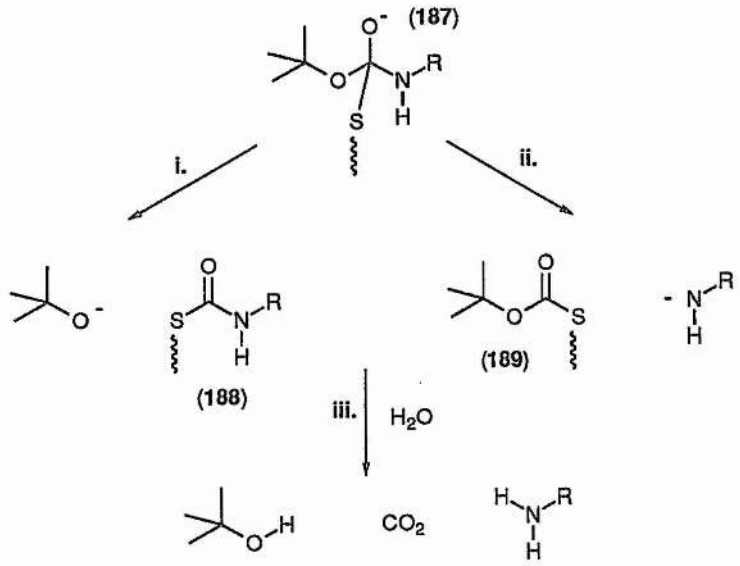



Scheme 2.46: *The hydrolysis of a t-butyl urethane by protonation of the urethane nitrogen.*

In a predominantly base catalysed proteolytic mechanism, a thiolate nucleophile attacks the urethane giving a tetrahedral intermediate (**187**) (Scheme 2.47). The tetrahedral intermediate (**187**) can either eliminate t-butoxide leaving a thiourethane (**188**) (step i, Scheme 2.47) or an amide anion leaving a thioester (**189**) (step ii, Scheme 2.47). Energetically step i (Scheme 2.47) would be favoured over step ii (Scheme 2.47) as t-butoxide is a better leaving group than an amide anion.

A thiourethane bond is very similar to an amide bond and water is not a sufficiently strong nucleophile to hydrolyse it. The ability of the nitrogen to act as a leaving group must be enhanced, for example by protonation. In this case the thiourethane (**188**) would be attacked

by water or hydroxide (from the protonation of the t-butoxide ion to give t-butanol) and decompose to give carbon dioxide, the free enzyme and the amine (step iii, Scheme 2.47). Enhancing the ability of the nitrogen as a leaving group, may cause it to leave in preference to the oxygen leaving group. The thioformate (**189**) could then be hydrolysed by water or hydroxide (from the protonation of the amide anion to give the amine) to give carbon dioxide, t-butanol and the free enzyme (step iii, Scheme 2.47).




Scheme 2.47: *The possible methods of hydrolysis of a t-butyl urethane.*

Thus elimination of either oxygen or nitrogen as the first leaving group, gives rise to the same final products, so it is not possible to distinguish these two mechanisms. Both mechanisms require that the urethane nitrogen, and consequently the amide nitrogen, in the case of a natural substrate, to be protonated at some point during the hydrolysis pathway. This N-protonation cannot occur prior to the attack of the thiolate nucleophile as the t-butyl urethane and the t-butyl ester do not decompose to give isobutene instead of t-butanol.

## 2.4.5 The epoxide inhibitors

(2R,3R)-O-[trans-carbonyl oxiran-2-yl 3-(ethoxycarbonyl) butane] N-(Cbz-Leu-Ala-)-ethanolanine (**173**) and (2S,3S)-O-[trans-carbonyl oxiran-2-yl 3-(ethoxycarbonyl) butane] N-(Cbz-Leu-Ala-)-ethanolanine (**175**) are weak reversible inhibitors for the adenovirus proteinase and no evidence of irreversible inhibition could be found (Section 2.3.4).

Both (RR) **E-64** (**17**) and (SS) **E-64** (**133**) (Fig 2.24) are potent irreversible inhibitors of papain (Section 2.2.1). This difference in activity towards epoxides between papain and the adenovirus proteinase could represent a major difference in mechanism between papain and the adenovirus proteinase.

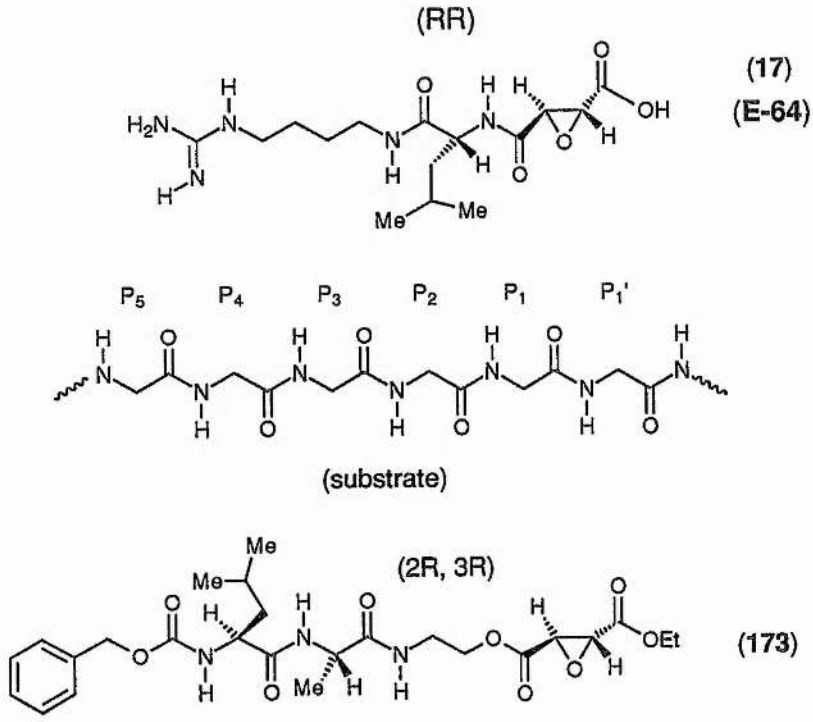



Fig. 2.27:

*A comparison of a) E-64 (17); b) The substrate back-bone and c) The epoxide (173).*

Comparing the structure of the epoxide (**173**) to that of the substrate (Fig. 2.27), shows that the epoxide (**173**) has a carbonyl group corresponding to both the $P_1$ and $P_1$' α-carbons. While the $P_2$ carbonyl group has been exchanged for a methylene group. **E-64** (**17**) (Fig. 2.27) similarly has carbonyl groups at the $P_1$ α-carbon position and the $P_1$' α-carbon position. However the peptide backbone is completely disrupted from the $P_2$ residue onwards. These differences are not important for papain and **E-64** (**17**) binds irreversibly to papain.

The loss of the $P_2$ carbonyl from the epoxides (**173**) and (**175**) is unlikely to be significant, as the $P_2$-$P_1$ residues are not sufficient, by themselves to bind to the proteinase (Section 2.3.3.0).

The hydrolytic mechanism of papain (Section 1.2.1) involves the protonation of the carbonyl oxygen. The epoxide oxygen of E-64 (**17**) and Ep-475 (**134**) (Section 2.2.1) binds in the $P_1$ carbonyl position so the epoxide oxygen is protonated, prior to attack by the thiolate nucleophile. Alkylation of the nucleopilic thiolate is rapid as the protonated oxygen is a good leaving group.[166] The hydrolysis of both the t-butyl ester (Section 2.4.3) and the hydrolysis of the t-butyl urethane (Section 2.4.4) to give t-butanol indicated, that nucleophilic attack of the $P_1$ carbonyl carbon occured at the same time or after N- or O-protonation of a substrate.

The epoxides (**173**) and (**175**) bound weakly to the proteinase (Section 2.3.4). So, assuming that the epoxide oxygen binds in the $P_1$ carbonyl binding pocket, the thiolate nucleophile cannot open the epoxide ring, unless protonation of the epoxide oxygen occurs in conjunction with or before the attack of the thiolate nucleophile. Since an alkoxide leaving group is a poor leaving group compared to a thiolate leaving group.

The lack of irreversible inhibition of the epoxides (**173**) and (**175**) also suggests that O-protonation of the $P_1$ carbonyl oxygen of the substrate does not occur to any significant extent.

## 2.4.6 The nitrile inhibitors

As mentioned previously (Section 2.3.3.0), the nature of the $P_4$ N-protection affects the degree of binding to the proteinase of the Leu-Ala-Gly-aminoacetonitrile inhibitors. N-Protecting groups with a polarisable π-electron system in the $P_5$-$P_6$ positions, such as a benzyl group, tend to bind more strongly to the proteinase.

This evidence suggests that the proteinase has residues that can interact with a π-electron system, such as a tryptophan or a phenylalanyine residue, π-stacking between the benzyl group and the side chain of the aminoacid residue (Fig. 2.29), would enhance the binding of the Cbz-protected nitrile, compared with TFA- or t-Boc- protected nitriles.

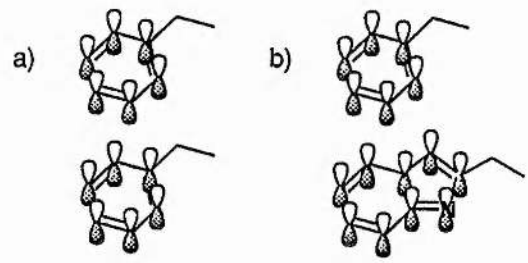



Fig. 2.28: *π-stacking interactions between a benzyl group and;*

*a) a phenylalanyl residue and b) a tryptophanyl residue.*

Both Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) and t-Boc-Leu-Ala-Gly-aminoacetonitrile (**79**) were reversible noncompetitive inhibitors of the adenovirus proteinase (Section 2.3.3.0) with approximate $K_i$'s of 15 μmol $dm^{-3}$ and 20 μmol $dm^{-3}$ respectively. Similar nitriles *ie.*

acetyl-Phe-aminoacetonitrile (**188**) ($K_i$ 0.73 μmol $dm^{-3}$)[167] and benzoyl-aminoacetonitrile (77) ($K_i$ 0.38 mmol $dm^{-3}$),[168] are competitive inhibitors of papain.

Noncompetitive inhibition arises when an inhibitor binds to an enzymatic intermediate species and the free enzyme to the same extent (Section 1.2.3.1.3) and the level of inhibition varies with substrate concentration. It is therefore possible that the substrate could bind at the active site of the proteinase and the nitrile inhibitors bind to the proteinase, elsewhere, modifying the conformation of the proteinase and reducing the rate of proteolytic hydrolysis. Alternatively the inhibitor could bind at the active site of a proteinase-substrate intermediate as well as the free enzyme.

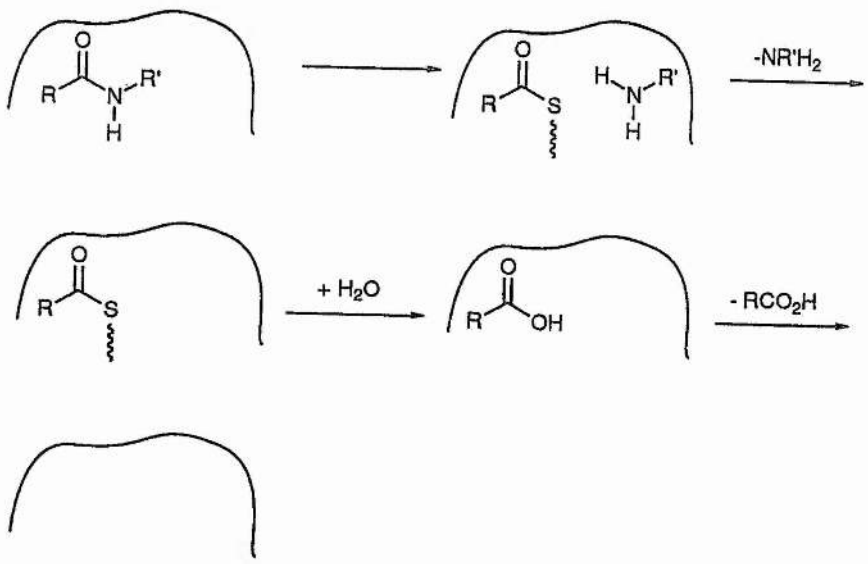
From the series of TFA-protected nitriles (Table 2.6, Section 2.3.3.0), the nitrile inhibitors show the same residue sequence specificity as the natural substrates. The removal of the $P_4$ leucyl residue results in an 80% loss of initial-rate-percentage-inhibition, and the loss of the $P_3$ alanyl residue results in the 100% loss of initial-rate-percentage-inhibition. This implies that the nitrile inhibitors bind in the P side substrate binding pocket, as the probability of another region of the proteinase requiring the same sequence specificity as the active site is extremely low.

Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) and t-Boc-Leu-Ala-Gly-aminoacetonitrile (**79**) do not contain any residues in the P' positions. It may be possible for the nitriles to mimic the P side acid leaving group, after the cleavage of the $P_1$-$P_1$' amide bond. It seems likely that the nitriles can bind to the free enzyme, mimicing the substrate and bind to the proteinase after the P side acid group has left while the P' side amine product is still attached to the proteinase. The nitriles would be binding at the active site of the proteinase and behave as a noncompetitive inhibitor which fits the observed results.

From this it is clear that the order of product release and the rate determining step differs between papain and the adenovirus type 2 proteinase. Papain initially cleaves the $P_1$-$P_1$'

amide bond by the formation of a thioester and the P' side amine product (step i, Scheme 2.48). The amine product diffuses away from the the active site of papain (step ii, Scheme 2.48) and the thioester is hydrolysed by water to give the His-Cys ion pair and the P side acid product (step iii, Scheme 2.48). In papain cleavage of the carbon-nitrogen bond is the rate determining step.[161] The P side acid product then diffuses away from the active site of papain, leaving papain free for further catalytic cycles (step iv, Scheme 2.48).

P side mimics, such as acetyl-Phe-aminoacetonitrile (**188**) and benzoyl-aminoacetonitrile (77) are only able to inhibit the free enzyme, as at all other steps during the catalytic cycle the S side of the proteinase is blocked by the P side of the substrate.

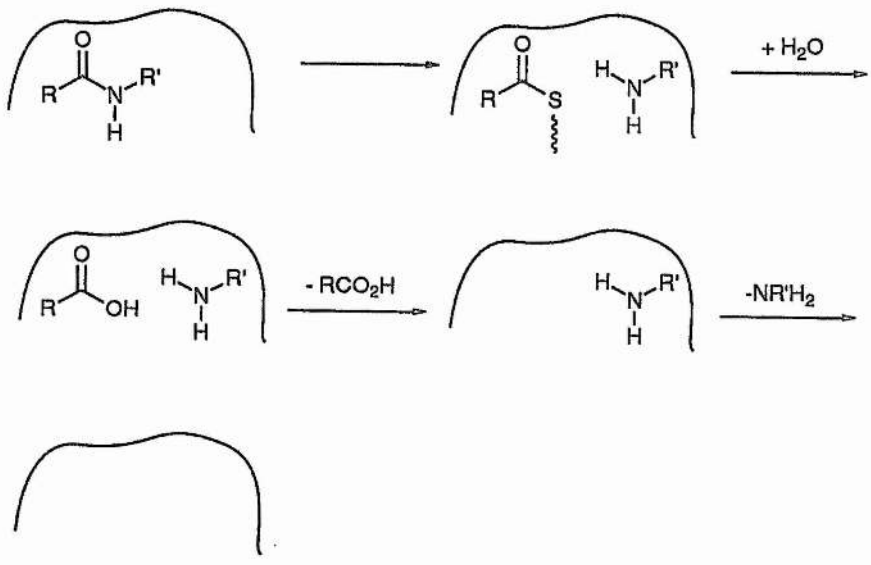



Scheme 2.48: *The mechanism of papain.*

The noncompetitive nature of the nitriles suggests that the adenovirus proteinase behaves differently, as the hydrolysis products leave in a different order compared to papain. Initially the adenovirus proteinase cleaves the $P_1$-$P_1$' amide bond by the formation of a thioester and the P' side amine product (step i, Scheme 2.49). However the amine does not leave the

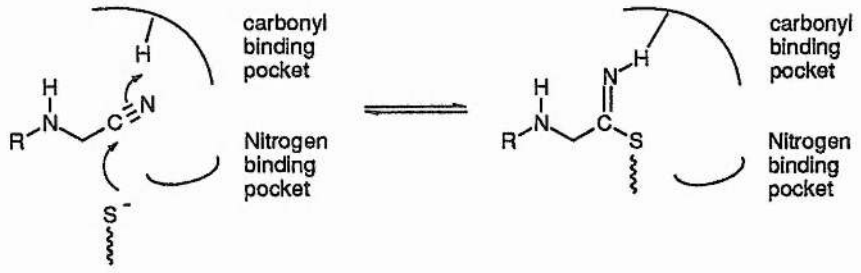
active site of the proteinase and the thioester is hydrolysed by water to give the His-Cys ion pair and the P side acid product (step ii, Scheme 2.49). The P side acid product then diffuses away from the active site of papain, leaving the P' side amine still attached to the proteinase (step iii, Scheme 2.49). The P' side amine product then diffuses away from the active site of the proteinase, leaving the adenovirus proteinase free of further catalytic cycles (step iv).

For the adenovirus proteinase, inhibited by the nitriles, the loss of the P' amine product must be the rate determining step in the hydrolysis reaction, otherwise an inhibitor binding to the enzyme-P' amine complex would have no effect on the overall reaction rate. The lack of hydrolysis of the simple amides (Section 2.4.0) has shown that if the carbon-nitrogen bond is planar and rotation of the carbon-nitrogen bond is restricted, additional interactions between the P' side of the substrate and the proteinase are required for the potential substrate to bind to the proteinase. As the cyanides contain no P' residues it is reasonable to suppose that the nitrile functionality is able to adopt the amide binding conformation without the need for P' interaction.




Scheme 2.49: *The possible mechanism for the adenovirus proteinase.*

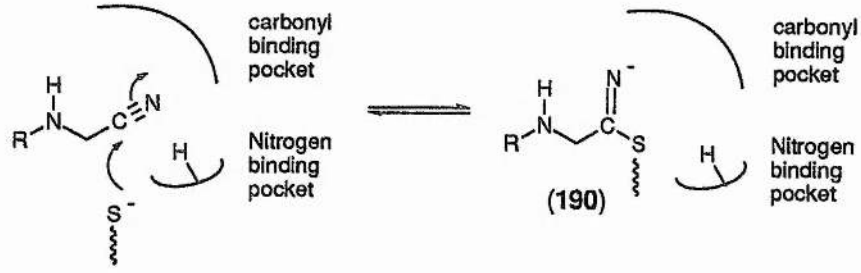
The mode of binding of Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) and t-Boc-Leu-Ala-Gly-aminoacetonitrile (**79**) is also rather intriguing. The nitrile binds to free papain allowing the thiolate nucleophile to attack the nitrile carbon. The sp hydridised nitrogen becomes an $sp^2$ hybridised thioamidate nitrogen and moves up into the $P_1$ carbonyl oxygen pocket (Scheme 2.50). The $P_1$ carbonyl oxygen binding pocket contains hydrogen bond donors, which normally polarise the $P_1$ carbonyl bond. In this case the protons also act to stabilise the thioamidate adduct between papain and the nitrile inhibitor (Scheme 2.50). The formation of the adduct is reversible and deprotonation of the thioamidate nitrogen results in the reformation of the nitrile inhibitor and papain again.




Scheme 2.50: *The formation of a thioamidate adduct between papain and a nitrile inhibitor.*

In the case of the adenovirus proteinase, hydrolysis of the t-butyl ester (Section 2.4.3) and the t-butyl urethane (Section 2.4.4) along with the nature of inhibition by the epoxides (Section 2.4.5), alcohol and the bromide (Section 2.4.7), all suggest that there is no significant protonation of the carbonyl oxygen during proteolytic hydrolysis by the adenovirus proteinase. When the nitrile inhibitor binds to the adenovirus proteinase-amine complex, the nitrile carbon can again mimic the carbonyl carbon of the amide substrate and is attacked by the thiolate nucleophile. However when the nitrogen moves up into the $P_1$ carbonyl oxygen binding pocket, there is no significant protonation, so a thioamidate anion

(**190**) is formed (Scheme 2.51). The thioamidate anion (**190**) is not as stable as a thioamidate and would tend to expel the thiolate nucleophile, reforming the nitrile.




Scheme 2.51: *The formation of a thioamidate adduct (190) between the adenovirus proteinase and a nitrile inhibitor.*

As no significant $P_1$ carbonyl oxygen protonation appears to occur during the proteolytic hydrolysis mechanism of the adenovirus proteinase, the nitrile would be a less potent inhibitor than for papain where significant protonation of the carbonyl oxygen occurs.

However direct comparison of the $K_i$ values of inhibitors for the adenovirus proteinase and papain is not possible, as the inhibitors also have different peptide sequences. There are two effects on the affinity of the proteinase for the inhibitor: One is due to the cyanide functionality and the other, to the peptidyl sequences. These effects can not be separated and direct comparison of the $K_i$'s for inhibitors is not useful.

Retention of the P' amine product in the proteinase's active site is enhanced by the formation of a partial bond between the thioamidate group's carbon and the P' amine's nitrogen (Fig. 2.29a). In this case, the thioamidate (**190**)-P' amine complex (**191**) could mimic the substrate (Fig. 2.29c) or the transition state complex where the carbon-nitrogen bond of the amide is breaking while the thiolate-carbonyl carbon and the nitrogen-proton bonds are forming (**192**) (Fig. 2.29b).

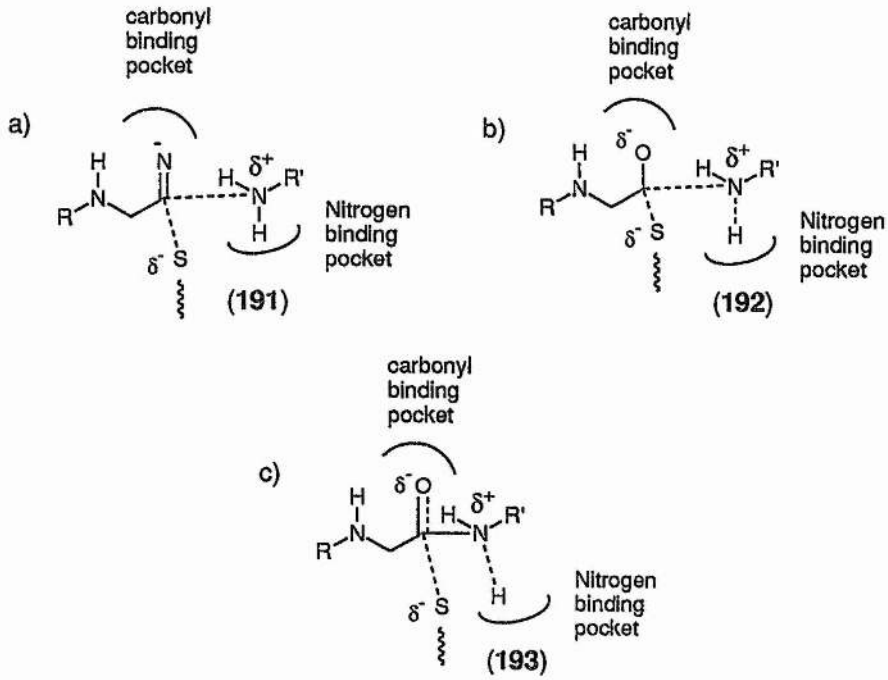
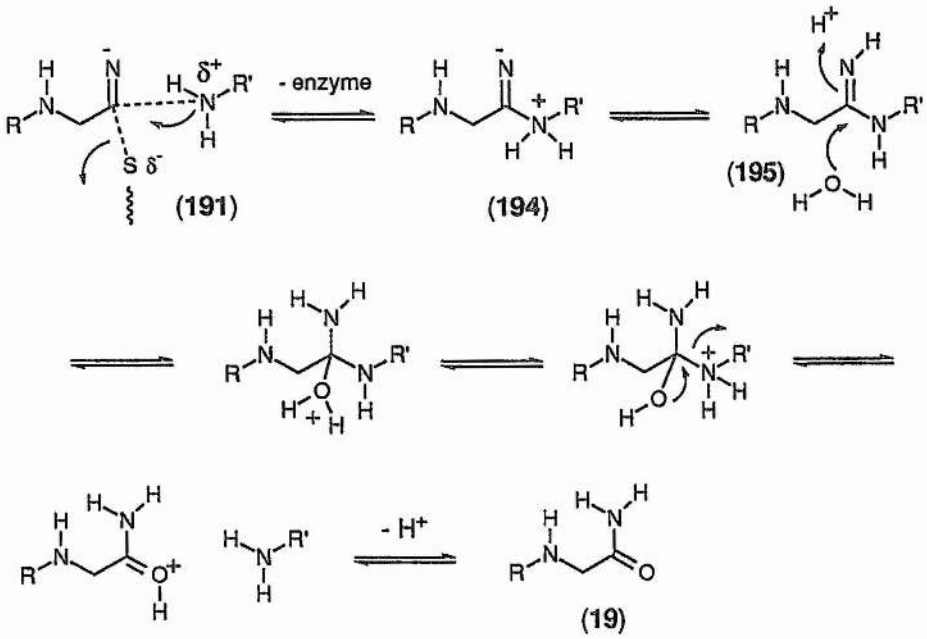



Fig. 2.29:

*a) The possible bonding interactions of the thioamidate anion (**190**) and the P' side amine; b) the substrate transition state and c) substrate.*

This leads to the possibility that the proteinase could catalyse the formation of the amide analogues (**194**) and (**195**) by the reverse reaction of amide bond hydrolysis (Scheme 2.52). After debinding from the proteinase the amide analogue (**195**) would be hydrolysed to give the simple amide (**19**) and the P' amine product. The nitrile has effectively been hydrated to an amide.

Hydration of the nitrile by the proteinase could be confirmed by the loss of inhibition with time as the simple amide (**19**) did not inhibit the proteinase, and by the presence of the amide (**19**) in the hydrolysis mixture.




Scheme 2.52: *The possible formation of an amide from a nitrile by the proteinase.*

## 2.4.7 The alcohol and bromide inhibitors

N-(Cbz-Leu-Ala-Gly)-ethanolamine (**75**) and N-(Cbz-Leu-Ala-Gly)- 2-bromo-ethylamine (**130**) were weak reversible inhibitors of the adenovirus proteinase, reducing the initial rates by 15 and 18% respectively (Sections 2.3.1 and 2.3.4). The $K_i$'s of the alcohol (**75**) and the bromide (**130**) were estimated as 1.2 mmol $dm^{-3}$ and 1.1 mmol $dm^{-3}$ respectively. No evidence of irreversible alkylation of the proteinase by the bromide (**130**) was observed.

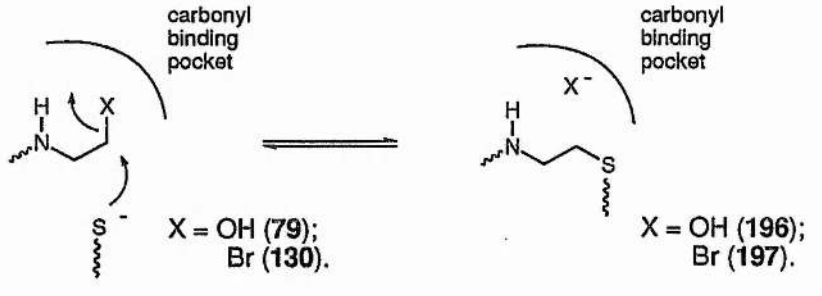
Inhibition of the adenovirus proteinase by the alcohol (**75**) is in complete contrast to papain where the $K_i$ for N-benzoyl-ethanolamine (**187**) is greater than or equal to 1 mol $dm^{-3}$,[160] suggesting significant differences in mechanism between the two proteinases. As mentioned previously (Section 2.1.2), the alcohol functionality can bind at either the $P_1$ carbonyl oxygen position or the $P_1$' amide nitrogen position. The bromide is likely only to

bind in the $P_1$' amide nitrogen position, as the carbon-bromide bond is much longer than the carbon-oxygen double bond (Section 2.2.0).

If both the alcohol (**75**) and the bromide (**130**) are competitive inhibitors it is likely that the alcohol and the bromide functionalities both bind in the $P_1$' amide nitrogen binding pocket. This also fits with the observation that the bromide (**130**) is a slightly better inhibitor than the alcohol (**75**), as the proteinase may bind a transition state in which the $P_1$-$P_1$' carbon-nitrogen bond is weakened and is thus longer than a normal amide bond.

The bromide (**130**) and alcohol (**75**) inhibitors and, the ester, t-butyl urethane and amide substrates all contain a group at the $P_1$' amide nitrogen position that can accept a proton donor. Therefore the adenovirus proteinase could partially protonate the $P_1$' amide nitrogen, at some point during the proteolytic hydrolysis mechanism.

If the alcohol (**75**) and the bromide (**130**) are noncompetitive inhibitors of the proteinase, the hydroxy group of the alcohol (**75**) could bind in the $P_1$ carbonyl binding pocket of the free enzyme and the thiolate nucleophile could displace the hydroxyl group of the alcohol to give alkylate proteinase (**196**) (Scheme 2.53). As the alcohol is a reversible inhibitor of the proteinase, the hydroxide is unable to leave the active site of the proteinase and attacks the alkylated proteinase to give the alcohol again.




Scheme 2.53: *The reversible formation of alkylated proteinase from the alcohol (75) and bromide (130) inhibitors.*

As in the case for the nitrile (Section 2.4.6) and the dimethyl acetal (Section 2.4.8) the P' amine product could also be trapped in the alkylated enzyme (Fig. 2.30a) and mimic the transition state as in Fig. 2.30b and the substrate (Fig. 2.29c).

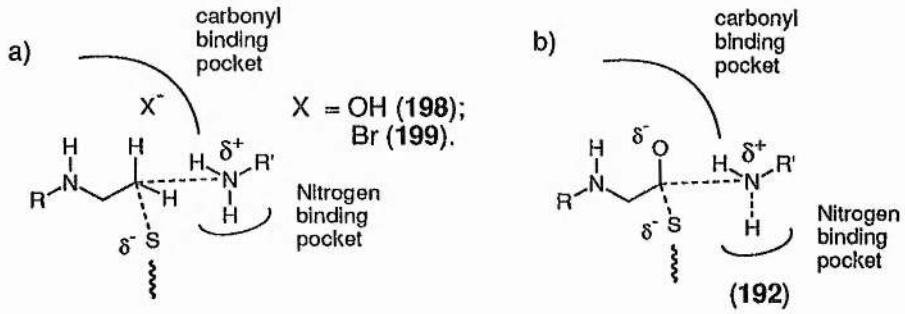



Fig 2.29 *a) The possible bonding interactions of the alkylated proteinase (191) and the P' side amine and b) the substrate transition state.*

If the bromide group is not too big for the carbonyl binding pocket, the bromide (**130**) could also act as a noncompetitive inhibitor, binding to the free enzyme and the P' amine-enzyme complex in a similar manner to the alcohol (Scheme 2.53 and Fig. 2.30).

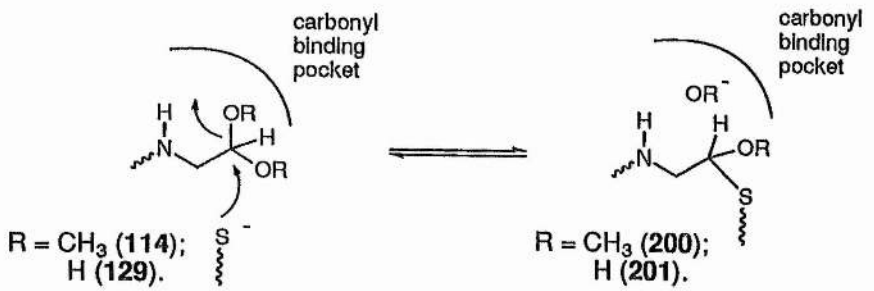
Unfortunately the bromide (**130**) and the alcohol (**75**) were not sufficiently soluble in water for their $K_i$'s and mode of inhibition to be determined, so their binding modes are unknown.

## 2.4.8 The dimethyl acetal, aldehyde and methyl ketone inhibitors

Cbz-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal (**114**) is a reversible noncompetitive inhibitor of the adenovirus proteinase, reducing the initial rate by 68% and having an approximate $K_i$ of 96 $\mu moldm^{-3}$ (Section 2.3.3). The dimethyl acetal inhibitor (**114**) binds to the same enzyme species as the nitrile inhibitors (Section 2.4.7) *ie.* the free enzyme and the amine proteinase complex. The dimethyl acetal (**114**), does not require P' interactions to

bind the free enzyme, unlike the simple amides (Section 2.4.0). One of the methoxy groups of the dimethyl acetal (**114**) probably binds in the $P_1$ carbonyl binding pocket but the location of the second methoxy group of the dimethyl acetal in the active site of the proteinase is unknown. The second methoxy group does not bind in the $P_1$' amide nitrogen binding pocket as the dimethylacetal is a noncompetitive inhibitor so this position is already occupied by the P' side amine product in the enzyme-P' amine complex.

In both enzyme species, The dimethylacetal (**114**) could be attacked by the thiolate nucleophile to give a hemithioacetal (**200**) and methoxide (Scheme 2.54). No evidence of hydrolysis of the dimethyl acetal functionality was observed by $^1$H-NMR spectroscopy after the dimethyl acetal (**114**) was incubated with the proteinase for 3 hours. Therefore, if the methoxide is produced in the carbonyl binding pocket it is unable to leave the active site of the proteinase and attacks the hemithioacetal (**200**) to give the dimethylacetal (**114**) again.




Scheme 2.54: *The reversible formation of a hemithioacetal (200) (or hemithiodiol (201)) from the dimethylacetal (114) (or diol (129)) inhibitors.*

The P' amine could again be trapped in the hemithioacetal (**200**) to give transition state and substrate mimic as in Fig. 2.31 and Fig. 2.29c, respectively in a similar manner to the thioimidate anion (**190**) (Section 2.4.6).

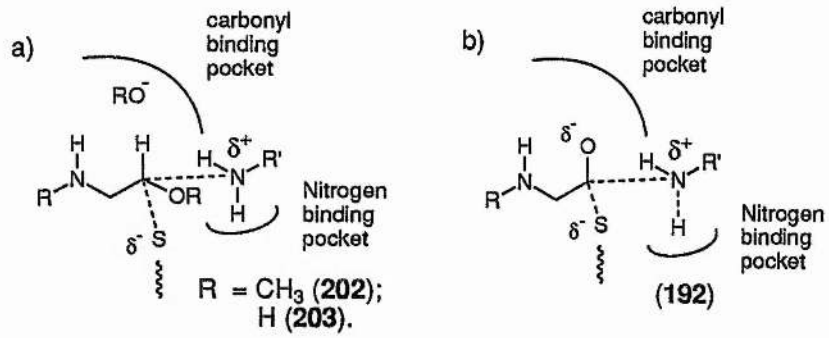


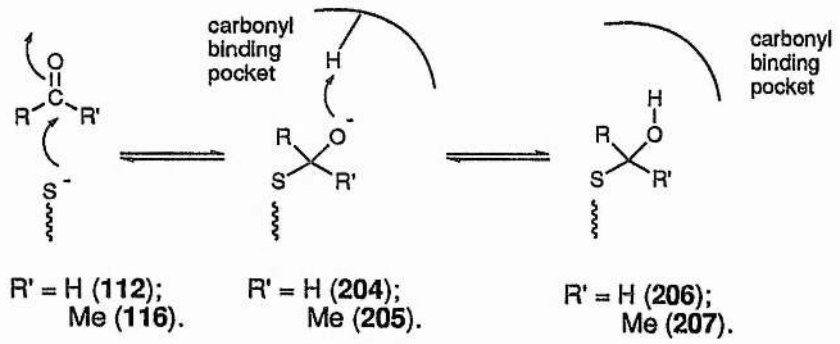

Fig. 2.31: *a) The possible bonding interactions of the hemithioacetal (192) and the P' side amine and b) the substrate transition state.*

Cbz-Leu-Ala-Gly-aminoacetaldehyde (**112**) is a powerful reversible inhibitor of the adenovirus proteinase which can reduce the initial rate by 68% and an has an estimated $K_i$ of 10 μmol $dm^{-3}$. Cbz-Leu-Ala-Gly-1-amino propanone (**116**) is a much weaker reversible inhibitor, only reducing the initial rate by 17% and has an estimated $K_i$ of 130 μmol $dm^{-3}$ (Section 2.3.3).

The aldehyde (**112**) and methyl ketone (**116**) could interact with the proteinase in a number of ways. The aldehyde exists in a hydrated diol form (**129**) (Section 2.1.4.0) and may interact with the proteinase in a similar manner to the dimethyl acetal (**114**) (Fig. 2.31 and Scheme 2.54). Less than 1% of the methyl ketone (**116**) exists in the hydrated diol form and this correlates well with the low level of inhibition by the methyl ketone (**116**) compared to the aldehyde (**112**).

However the diol functionality is a lot more polar than the dimethyl acetal functionality so the 70% loss of activity in methylation the hydroxyl groups of the diol could be due to the proteinase preferring a polar functionality at the $P_1$ carbonyl binding pocket.

The aldehyde (**112**) and ketone (**116**) could also form a hemithioacetal (**206**) or a hemithioketal (**207**) adduct with the proteinase (Section 2.1.4) (Scheme 2.55).




Scheme 2.55: *The formation of a hemithioacetal and hemithioketal adduct.*

The hemithioketal adduct is more difficult to form than the hemithioacetal adduct, as the methyl group of the methyl ketone donates negative charge to the carbonyl carbon, reducing it's electrophilicity. This fits with the methyl ketone (**116**) being a worse inhibitor than the aldehyde (**112**).

As there is little O-protonation of the $P_1$ carbonyl group a hemithioacetal (or ketal) adduct formed between an aldehyde (or ketone) would exist as an hemithioacetal anion (**204**) (or hemithioketal anion (**205**)) rather than the neutral hemithioacetal (**206**) (or hemithioketal (**207**)). If this is the case the hemithioacetal anion (**204**) (or hemithioketal anion (**205**)) would not be very stable and collapse back rapidly to give the proteinase-bound aldehyde (or ketone).

At present it is not known whether the aldehyde (**112**) and ketone (**116**) form covalent adducts with the proteinase or not. Papain certainly does form covalent adducts with aldehyde and ketone based inhibitors (Section 2.1.4).

If the proteinase was more stable, the use of [1-$^{13}$C]-NMR on labelled aldehydes and ketones could establish the nature of the binding interactions between the aldehyde and ketone with the proteinase.

## 2.4.9 Alkenic inhibitors

Preliminary results show that Cbz-Leu-Ala-Gly-1-amino prop-2-ene (**115**) and Cbz-Leu-Ala-Gly-1-amino 2-methyl prop-2-ene (**118**) are both noncompetitive reversible inhibitors of the adenovirus proteinase reducing the initial rate by 17 and 27% respectively and have estimated $K_i$ of 130 μmol dm$^{-3}$ and an approximate $K_i$ of 45 μmol dm$^{-3}$ respectively (Section 2.3.3). The simple alkene inhibitors (**115**) and (**118**) bind to both the amine-proteinase complex and free enzyme.

There are two possible bound conformations for the simple alkene inhibitors (**115**) and (**118**). In the extended form (Fig. 2.32a), the alkene moiety binds across the $P_1$-$P_1$' amide bond, while in the compact form (Fig. 2.32b) the alkene moiety binds in the $P_1$ carbonyl binding pocket of the proteinase. In the extended conformation, the 1-amino 2-methyl prop-2-ene inhibitor (**118**) would place a lipophilic methyl group in the polar $P_1$ carbonyl binding pocket. This would be a rather unfavourable interaction, and suggest that the 1-amino 2-methyl prop-2-ene inhibitor (**118**) would be a worse inhibitor of the proteinase than the 1-amino prop-2-ene inhibitor (**115**).

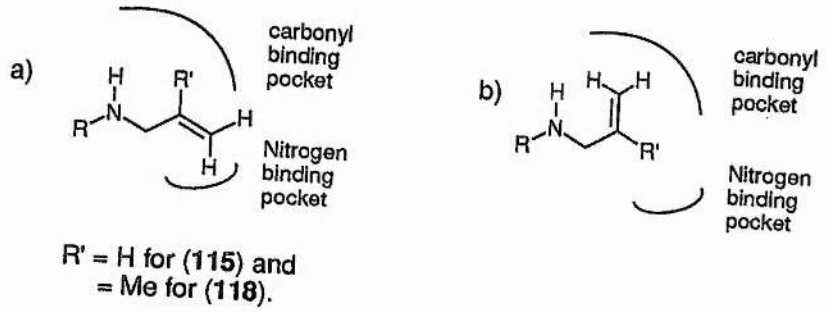



Fig. 2.32: *The two possible conformations of the simple alkene inhibitors; a) extended and b) compact.*

However the alkene (**118**) is the more potent inhibitor of the two, suggesting that they bind in the compact conformation. In this case the alkene moiety is positioned in the $P_1$ carbonyl binding pocket of the proteinase. The $P_1$ carbonyl binding pocket would be probably slightly positively charged to attract the carbonyl oxygen and this positively charged environment would polarise the $\pi$-electrons of the alkene moiety (Fig. 2.33). The close proximity of the thiolate nucleophile and perhaps the P' amine product would also polarise the $\pi$-electrons of the alkene.

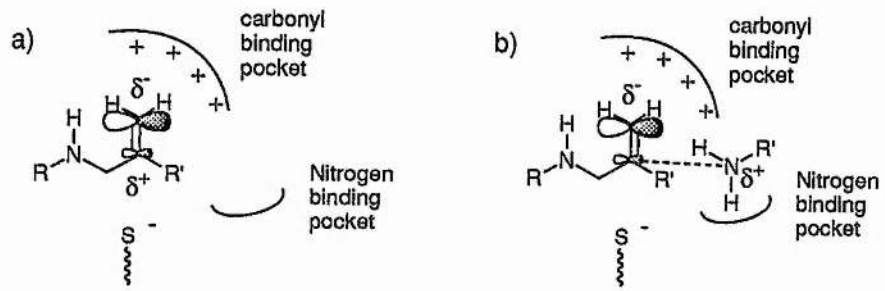



Fig. 2.33: *The polarisation of an alkene in the $P_1$ carbonyl binding pocket: a) of the substrate and b) of the P'-amine enzyme complex.*

An electron donating group such as a methyl group at R' would enhance this polarisation, by counteracting the accompanying increase in positive charge on the carbon attached to R'. The 1-amino,2-methylprop-2-ene inhibitor (**118**) would be expected to be the better inhibitor, as is the case experimentally.

O-(Cbz-Leu-Ala-Gly)-1-carboxymethyl,3-hydroxyprop-2-ene (**132**) is a reversible inhibitor of the adenovirus proteinase which reduces the initial rate by 66% and has an estimated $K_i$ 10 $\mu$mol $dm^{-3}$ (Section 2.3.3). Such Michael acceptors are however potent competitive, irreversible inhibitors of papain (Section 2.2.0). The Michael acceptor (**132**),

like the simple alkenes (**115**) and (**118**) could bind in two possible conformations to the proteinase (Fig. 2.34).

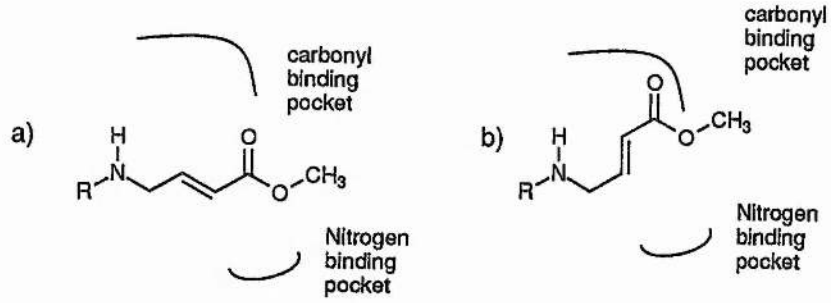



Fig. 2.34: *The two possible conformations of the Michael acceptor; a) extended and b) compact.*

If the Michael acceptor (**132**) were to bind in the compact form, the carboxymethyl group would be required to enter the $P_1$ carbonyl binding pocket of the proteinase. This is unlikely owing to the size of the carboxymethyl group, suggesting that the Michael acceptor (**132**) binds in the extended conformation.

## 2.4.10 The aldehyde dimer

N,N'-di-(Cbz-Leu-Ala-Gly)- 2,4-diaminobut-2-ene aldehyde (**122**) is the most potent competitive reversible inhibitor found. The initial rate is reduced by 89% and the approximate $K_i$ is 6 μmol $dm^{-3}$ (Section 2.3.3). The aldehyde dimer (**122**) binds to the free enzyme as the P' side of the inhibitor prevents the P' side amine product from being bound at the same time as the inhibitor.

The aldehyde dimer (**122**) could in principle bind in either the aldehyde or diol forms or as a Michael acceptor (Fig. 2.35).

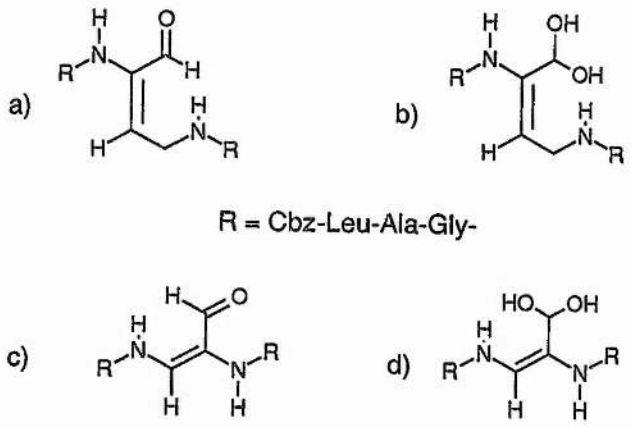



Fig. 2.35: *The possible binding modes of the aldehyde dimer inhibitor; a) as an aldehyde; b) as a diol; c) as a Michael acceptor and an aldehyde and d) as a Michael acceptor and a diol.*

It would be difficult to determine which conformation the aldehyde dimer (**122**) bound in, as both the Michael acceptor (**132**) and the aldehyde (**112**) have similar $K_i$'s and no irreversible inhibition of the aldehyde dimer (**122**) was detected.

The aldehyde dimer is able to bind to the free enzyme, like the natural substrates, the esters (Section 2.4.1) and the *p*-nitroanilide (**45**) (Section 2.4.0) substrates, as the aldehyde dimer (**122**) contains the necessary P' side residue(s) to interact with proteinase. Since the residues in the amide backbone of the P' side of the aldehyde dimer (**122**) are reversed, it follows that only a limited number of P' side backbone interactions are required for binding to the proteinase. If the entire P' amide backbone was required for inhibitor binding to the proteinase, the aldehyde dimer (**122**) would find it impossible to present all the required functionalities in the correct sequence.

The P' side of the aldehyde dimer (**122**) is sufficiently flexible to allow the necessary P' functionalities to interact with the S' side of the proteinase. So the aldehyde dimer (**122**) can bind to the free proteinase.

## 2.4.11 Summary

A series of $P_4$ (Cbz and t-Boc) N-protected potential substrates and inhibitors, containing the $P_1$ to $P_4$ substrate recognition sequence of the type 2 adenovirus proteinase (Leu-Ala-Gly-Gly) were prepared by solution phase peptide coupling techniques and tested for activity against the proteinase. The potential substrates contained the amide and ester moieties at $P_1$ carbonyl position and the potential inhibitors contained the alcohol, acid, bromide, aldehyde, ketone, dimethylacetal, nitrile, alkenic, malonyl and epoxysuccinate moieties at $P_1$ carbonyl position.

The esters, the t-Boc urethane and the *p*-nitroanilide moieties were substrates for the proteinase and the acid and the amides did not bind to the proteinase. Preliminary results suggest that the other inhibitors were mostly noncompetitive inhibitors for the adenovirus proteinase with approximate $K_i$'s between 15 and 200 μmol $dm^{-3}$.

The test results indicate that;

- the amides must contain a carbonyl group at $P_2$' to bind to the proteinase, the *p*-nitroanilide (**24**) was the only amide to be hydrolysed;
- $P_2$ carbonyl binding pocket may be essential for the $P_1$-$P_1$'amide bond to adopt a non-coplanar conformation to allow catalysis, as the ester moiety is turned over without the need for $P_2$' interaction.
- the loss of the $P_1$' amine product is the rate limiting step for the hydrolysis of a substrate by the adenovirus proteinase, as the nitriles, dimethyl acetals and alkenic inhibitors were able to bind to both the enzyme-P' amine product complex and the free enzyme, inhibiting in a noncompetitive manner;
- the P acid product leaves before the P' amine product, which is in complete contrast to classical cysteine proteinases such as papain;

- little protonation of the $P_1$'amide nitrogen or the $P_1$ carbonyl oxygen of the adenovirus proteinase-substrate complex occurs before the nucleophilic attack of the $P_1$ carbonyl carbon of the adenorirus proteinase-substrate complex, as the t-butyl esters and the t-Boc urethane were hydrolysed to give t-butanol and the epoxysuccinates (**173**) and (**175**) and the bromide (**130**) failed to inhibit the adenovirus proteinase irreversibly.

# 3.0 Compound synthesis

## 3.0.0 General procedures for reagent and solvent preparation.

Dry solvents were prepared as follows: acetone was dried over $CaSO_4$ for 4 hours under a $CaSO_4$ drying tube, refluxed for 2 h, distilled and stored under argon. The dry acetone was used as soon as possible.[169] Acetonitrile was stirred over $CaH_2$ under a silica drying tube until no more $H_2$ was evolved. Then fresh $CaH_2$ was added and the acetonitrile refluxed under $N_2$ for 2 h, distilled and then stored under argon.[170] The dry acetonitrile was used as soon as possible. Dichloromethane was dried over $CaH_2$ for 2 h, distilled under $N_2$ and then stored under argon.[169] The dry dichloromethane was used as soon as possible. N,N-Dimethylformamide (DMF) was stirred overnight with activated Linde 4A molecular sieves under a silica drying tube and distilled at reduced pressure onto freshly activated Linde 4A molecular sieves.[169] The dry N,N-dimethylformamide was used within 4 weeks. Ethanol and methanol were dried using magnesium turnings and a catalytic amount of iodine under a silica drying tube, refluxed for 1 h, distilled and then stored under argon.[171] The dry alcohol was used within 3 weeks. Diethyl ether and tetrahydrofuran were dried with KOH, then treated with sodium and benzophenone until a purple / blue colour was observed. The solvent was distilled under nitrogen, stored under argon and used, as soon as possible.[172]

Ethereal diazomethane solution was prepared from the action of KOH on N-methyl, N-nitrosotoluene-*p*-sulphonamide.[173] If dry ethereal diazomethane solution was required the solution was carefully dried for 1 h over KOH pellets at -5°C. Iodine was removed from hydriodic acid by washing 3 times with a suspension of Amberlite LA-2 in petrol (4%, 10 $cm^3$ per 100 $cm^3$ of acid)[174] and after separation from the organic layer, the aqueous phase was distilled at reduced pressure. Triethylamine was stirred for 3 h over $CaH_2$ under a drying tube, refluxed for 2 hours under $N_2$, distilled and stored under argon.[169] The dry triethylamine

was used on the same day. N-methyl morpholine was stirred with ninhydrin over a silica drying tube for ½ h, then distilled and used within a month. Thionyl chloride was refluxed under $N_2$ with sulphur and distilled twice.[169]

## 3.0.1 Compound characterisation

Melting points were measured on an electrothermal melting point apparatus and are uncorrected. Optical rotations were measured at room temperature using an optical activity AA-1000 polarimeter, and are given in $10^{-1}$ g $cm^{-3}$ and the solution concentration is given in g $cm^{-3}$.

Elemental analyses were performed by Silvia Smith at the St Andrews University microanalytical service and are quoted to the nearest 0.05%. Accurate mass determination and low resolution mass spectroscopy were carried out by the SERC service at Swansea using a VG AZB-E instrument or by Colin Millar at the St Andrews University mass spectrometry service using a Kratos MS-50 or a Fisons VG autospec instrument. The major fragments are given as percentage of the most abundant fragment.

Infrared specra were recorded on a Perkin-Elmer 1710 Fourier transform infrared spectrophotometer. If solid, the samples were prepared as Nujol mulls, or as thin films between sodium chloride disks, or as dispersions in potassium bromide disks, or as chloroform solutions between sodium chloride disks. If oils, the samples were prepared as thin films between sodium chloride discs. Absorption maxima are given in wave numbers ($cm^{-1}$) and are referenced to a polystyrene standard at 1603 $cm^{-1}$.

NMR spectra were recorded using a Varian Gemini 200 (200 MHz $^1H$ and 50.31 MHz $^{13}C$), or a Bruker AM-300 (300 MHz $^1H$ and 74.76 MHz $^{13}C$) or the SERC service at Warwick (400 MHz $^1H$ and 100 MHz $^{13}C$) NMR spectrometers. The chemical shifts are

given in parts per million and the splitting constants are given in Hz and the following abbreviations are used s, singlet; d, doublet; t, triplet;, q, quartet; quin, quintette; sex, sextet and sep. septet. For $^{1}$H-spectra the chemical shifts were referenced to the $^{1}$H in the remaining monoprotonated solvent, taken relative to TMS as follows; in $^{2}$H-chloroform to $C\underline{H}Cl_3$ (7.23 ppm), in $^{2}H_2$-water to $^{2}HO\underline{H}$ (4.90 ppm), in $^{2}H_6$-dimethylsulphoxide to $\underline{H}^{2}H_2CSOC^{2}H_3$ (4.60 ppm), in $^{2}H_4$-methanol to $\underline{H}^{2}H_2CO^{2}H$ (3.35 ppm), in $^{2}H_3$-acetonitrile to $C\underline{H}^{2}H_2CN$ (1.95 ppm) and in $^{2}H_6$-acetone to $C\underline{H}^{2}H_2COC^{2}H_3$ (2.15 ppm). For $^{13}$C-spectra the chemical shifts were referenced to the $^{13}$C in the deuterated solvent, taken relative to TMS as follows; in $^{2}$H-chloroform to $\underline{C}^{2}HCl_3$ (77.0 ppm), in $^{2}H_2$-water to 2,2-dimetyl,5-sulphonic acid-sila-pentane (0 ppm), in $^{2}H_6$-dimethylsulphoxide to $^{2}H_3\underline{C}SOC^{2}H_3$ (39.5 ppm), in $^{2}H_4$-methanol to $^{2}H_3\underline{C}O^{2}H$ (49.0 ppm), in $^{2}H_3$-acetonitrile to $\underline{C}^{2}H_3CN$ (1.3 ppm) and in $^{2}H_6$-acetone to $C^{2}H_3CO\underline{C}^{2}H_3$ (30.6 ppm).

## 3.0.2 Synthetic methods and compound characterisation.

Aminoacetonitrile hydrochloride (**85**) was purchased from the Aldrich chemical company, glycyl *p*-nitroanilide hydrochloride (**64**) was purchased from the Sigma chemical company and t-butyl (2R)-alaninate acetate (**89**) and t-butyl glycinate acetate (**30**) were purchased from the Nova Biochemistry company.

### *3.0.2.0 Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinamide (19)*

Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycine (**27**) (0.397 g, 1.01 mmol) was dissolved in dry THF (25 ml). i-$Pr_2EtN$ (0.35 $cm^3$, 2.02 mmol), glycinamide hydrochloride (**28**) (112 mg, 1.01 mmol) and TBTU (324 mg, 1.01 mmol)[111] were added and the suspension stirred overnight. The solvent was removed at reduced pressure to give a white solid which was dispersed in water and filtered. The solid was washed with water (20 $cm^3$), sodium hydrogencarbonate (saturated, 5 $cm^3$), water (5 $cm^3$), citric acid (10%, 5 $cm^3$), water (5 $cm^3$) and brine (saturated, 5 $cm^3$) to give a white solid (221 mg, 49%), m.p. 196-198°C; (Found: C, 56.4; H, 6.8; N, 15.65. $C_{21}H_{31}N_5O_6$ requires C, 56.1; H, 6.95; N, 15.60%; *m/z* (Found: $[M + H]^+$, 450.2353. $C_{21}H_{32}N_5O_6$ requires 450.2353); $[\alpha]_D$ -36.4 (*c* 0.25 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3475, 3425, 3335 (3 x NH), 3085 (aromatic CH), 2960 (aliphatic CH) and 1680, 1670, 1655, 1630, 1615 ( 4 x CONH and $CO_2NH$); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.54 (5H, m, $C_6H_5$), 5.28 (2H, s, $C_6H_5C\underline{H}_2$), 4.46 (1H, q, $J_{CH_3}$ 7.06, $C\underline{H}CH_3$), 4.33 (1H, t, $J_{CH_2}$ 7.50, $NHC\underline{H}CH_2$), 4.12 (1H, d, $J_{H_b}$ 16.86, $C\underline{H}_aH_b$), 4.08 (1H, d, $J_{H_a}$ 17.07, $CH_a\underline{H}_b$), 4.00 (1H, d, $J_{H_d}$ 17.10, $C\underline{H}_cH_d$), 3.99 (1H, d, $J_{H_c}$ 16.97, $CH_c\underline{H}_d$), 1.90 (1H, quin, $J_{CH}$ 6.71, $C\underline{H}(CH_3)_2$), 1.74 (2H, t, $J_{CH}$ *2* 7.31, $CHC\underline{H}_2CH$), 1.57 (3H, d, $J_{CH}$ 7.14, $CHC\underline{H}_3$) and 1.13

(6H, t, *J* 6.69, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz, $C_2H_3O_2H$) 176.24, 176.14, 174.70, 172.38, (4 x CO), 159.13 ($CO_2NH$), 138.44 (*i*-$C_6H_5$), 129.79 (*m*-$C_6H_5$), 129.34 (*p*-$C_6H_5$), 129.10 (*o*-$C_6H_5$), 68.13 ($\underline{C}H_2C_6H_5$), 55.33 ($\underline{C}HCH_2$), 51.29 ($\underline{C}HCH_3$), 44.13, 43.50 ($\underline{C}H_2CO$), 42.11 ($CHC\underline{H}_2$), 26.14 ($\underline{C}H(CH_3)_2$), 23.73, 22.14 ($CH(\underline{C}H_3)_2$) and 17.55 ($CH\underline{C}H_3$); *m/z* (FAB) 472 (100%, $[M + Na]^+$), 450 (12, $[M + H]^+$), 413 (21 $[M + Na - CH_3CONH_2]^+$) and 391 (13, $[M + H - CH_3CONH_2]^+$).

## *3.0.2.1 Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl, N'-methyl-glycinamide (20)*

This compound was prepared in a manner identical to that described for carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinamide (**19**) using N-methyl-glycinamide hydrochloride (**55**) (126 mg, 1.01 mmol) to give a white solid (316 mg, 68%). m.p. 213°C; (Found: C, 57.0; H, 7.3; N, 14.95. $C_{22}H_{33}N_5O_6$ requires C, 57.0; H, 7.2; N, 15.10%); *m/z* (Found: $[M + Na]^+$, 486.2320. $C_{22}H_{33}N_5O_6Na$ requires 486.2329); $[\alpha]_D$ -28.8 (*c* 0.13 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3290 (NH), 3090 (aromatic CH), 2955, 2870 (2 x aliphatic CH), 1690 ($CO_2NH$) and 1675, 1660, 1645, 1630 (4 x CONH); $\delta_H$ (300 MHz; $^2H_3CO_2H$) 7.34 (5H, m, $C_6H_5$), 5.09 (1H, d, $J_{H_b}$ 12.5, $C_6H_5C\underline{H}_aH_b$), 5.06 (1H, d, $J_{H_a}$ 12.5, $C_6H_5CH_a\underline{H}_b$), 4.26 (1H, q, $J_{CH_3}$ 7.2, $C\underline{H}CH_3$), 4.13 (1H, t, $J_{CH_2}$ 7.70, $NHC\underline{H}CH_2$), 3.91 (1H, d, $J_{H_b}$ 16.90, $C\underline{H}_aH_b$), 3.87 (1H, d, $J_{H_a}$ 16.90, $CH_a\underline{H}_b$), 3.79 (1H, d, $J_{H_d}$ 16.90, $C\underline{H}_cH_d$), 3.78 (1H, d, $J_{H_c}$ 16.90, $CH_c\underline{H}_d$), 2.71 (3H, s, $NCH_3$), 1.70 (1H, m, $J_{CH}$ 6.80, $C\underline{H}(CH_3)_2$), 1.53 (2H, t, $J_{CH_2}$ 7.30, $CHC\underline{H}_2CH$), 1.37 (3H, d, $J_{CH}$ 7.20, $CHC\underline{H}_3$), 0.94 (3H, t, *J* 6.70, $CH(C\underline{H}_3)_2$) and 0.92 (3H, t, *J* 6.60, $CH(CH_3)_2$); $\delta_C$ (75.5 MHz, $C^2H_3O^2H$) 175.80, 172.03 (4 x CO), 158.77 ($CO_2NH$), 138.10 (*i*-$C_6H_5$), 129.44 (*m*-$C_6H_5$), 129.01 (*p*-$C_6H_5$), 129.76 (*o*-$C_6H_5$), 67.85 ($\underline{C}H_2C_6H_5$), 55.10 ($\underline{C}HCH_2$), 51.00 ($\underline{C}HCH_3$), 43.94, 43.55 ($NHCH_2$), 41.88 ($CHC\underline{H}_2$), 26.33 ($\underline{C}H(CH_3)_2$),

25.87 ($NCH_3$), 23.45, 21.87 ($CH(\underline{C}H_3)_2$) and 17.27 ($CH\underline{C}H_3$); *m/z* (FAB) 486 (100%, $[M + Na]^+$), 464 (10, $[M + H]^+$), 444 (29 $[M + Na - CH_2CHCH_3]^+$), 352 (12, $[M + H + Na - CO_2CH_2C_6H_5]^+$) and 330 (10, $[M + 2H - CO_2CH_2C_6H_5]^+$).

### *3.0.2.3 Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl, N', N-dimethyl-glycinamide (21)*

Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycine (**27**) (0.142 g, 0.36 mmol) was dissolved in dry THF (25ml). i-$Pr_2EtN$ (0.968 $cm^3$, 0.72 mmol), N', N-dimethyl glycinamide hydrochloride (**55**) (50 mg, 0.36 mmol) and TBTU (0.116 g, 0.36 mmol) were then added and stirred overnight. The solvent was removed under reduced pressure and the oil was partitioned between ethyl acetate (40 $cm^3$) and water (40 $cm^3$), washed with water (2 ×50 $cm^3$), citric acid (10%, 30 $cm^3$), $NaHCO_3$ (saturated, 30 $cm^3$), brine (saturated, 20 $cm^3$) and dried ($MgSO_4$). The solvent was removed under reduced pressure and the last traces of ethyl acetate were azeotropically removed with methylene chloride to give an off white solid which was recrystalised from acetone / ether to give a white solid (59 mg, 34%), m.p. 163°C; *m/z* (Found: $[M + H]^+$, 478.2663. $C_{23}H_{36}N_5O_6$ requires 478.2666); $[\alpha]_D$ -28.2 (*c* 0.09 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3295 (NH), 3065 (aromatic CH), 2960 (aliphatic CH), 1720 ($CO_2NH$) and 1690, 1670, 1650, 1635 (4 x CONH); $\delta_H$ (300 MHz; $C^2HCl_3$) 7.60, 7.45 (2H, t, $N\underline{H}CH_2$), 7.40 (1H, d, $J_{CH}$ 7.40, $N\underline{H}CH_3$), 7.31 (5H, m, $C_6H_5$), 5.99 (1H, d, $J_{CH}$ 8.13, $NHCO_2$), 5.10 (1H, d, $J_{H_b}$ 12.5, $C_6H_5C\underline{H}_aH_b$), 5.00 (1H, d, $J_{H_a}$ 12.5, $C_6H_5CH_a\underline{H}_b$), 4.73 (1H, t, $J_{CH_2}$ 7.20, $NHC\underline{H}CH_2$), 4.38 (1H, q, $J_{CH_3}$ 6.63, $C\underline{H}CH_3$), 4.14, 4.06 (4H, t, $NHC\underline{H}_2$ and $NHC\underline{H}_2$), 2.91, 2.90 (6H, s, $NCH_3$), 1.65 (1H, m, $J_{CH_3}$ 6.57, $J_{CH}$ 5.97, $C\underline{H}(CH_3)_2$), 1.54 (2H, q, $J_{CH}$ 5.94, 7.17, $CHC\underline{H}_2CH$), 1.38 (3H, d, $J_{CH}$ 6.75, $CHC\underline{H}_3$) and 0.89 (6H, d, $J_{CH}$ 6.33,

$CH(C\underline{H}_3)_2$; $\delta_C$ (75.5 MHz, $C^2HCl_3$) 172.57, 172.44, 168.56, 167.66 (4 x CO), 156.56 ($CO_2NH$), 136.43 (*i*-$C_6H_5$), 128.48 (*m*-$C_6H_5$), 128.11 (*p*-$C_6H_5$), 127.98 (*o*-$C_6H_5$), 66.92 ($\underline{C}H_2C_6H_5$), 53.61 ($\underline{C}HCH_2$), 48.87 ($\underline{C}HCH_3$), 43.00 ($CHC\underline{H}_2$), 42.02, 41.41 ($NHCH_2$), 35.93, 35.60 ($NCH_3$), 24.75 ($\underline{C}H(CH_3)_2$), 23.03, 21.94 ($CH(\underline{C}H_3)_2$) and 18.81 ($CH\underline{C}H_3$); *m/z* (CI) 478 (100%, $[M + H]^+$), 433 (9 $[M + H - HN(CH_3)_2]^+$) and 370 (30, $[M - OCH_2C_6H_5]^+$).

## *3.0.2.4 Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl, N'-ethyl, N'-methyl-glycinamide (21)*

This compound was prepared in a manner identical to that described for carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl, N', N-dimethyl-glycinamide (**21**) using N'-ethyl, N'-methyl-glycinamide hydrochloride (**60**) (116 mg, 0.76 mmol) to give a white solid which was recrystalised from acetone / ether (238 mg, 64%), m.p. 170°C; (Found: C, 58.5; H, 7.8; N, 14.20. $C_{24}H_{37}N_5O_6$ requires C, 58.6; H, 7.6; N, 14.30%); *m/z* (Found: $[M + H]^+$, 492.2817. $C_{24}H_{38}N_5O_6$ requires 492.2822); $[\alpha]_D$ -29.7 (*c* 0.33 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3300 (NH), 3100, 3060 (2 x aromatic CH), 2955, 2930, 2850 (3 x aliphatic CH), 1715 ($CO_2NH$) and 1695, 1650, 1635, 1625 (4 x CONH); $\delta_H$ (300 MHz; $C^2HCl_3$) 7.91, 7.79 (3H, m, 2 x $N\underline{H}CH_2$ and $N\underline{H}CH_3$), 7.29 (5H, m, $C_6H_5$), 6.38 (1H, d, $J_{CH}$ 6.15, $NHCO_2$), 5.07 (1H, d, $J_{H_b}$ 27.09, $C_6H_5C\underline{H}_aH_b$), 5.03 (1H, d, $J_{H_a}$ 27.3, $C_6H_5CH_a\underline{H}_b$), 4.92 (1H, m, *J* 6.09, $NHC\underline{H}CH_2$), 4.49 (1H, m, $C\underline{H}CH_3$), 4.19 (4H, m, $NHC\underline{H}_2$ and $NHC\underline{H}_2$), 3.36 (3H, sex, *J* 7.5, $NCH_2$), 3.24 (3H, q, *J* 6.77, $NCH_2$), 2.86, 2.82 (3H, s, $NCH_3$), 1.66 (1H, m, *J* 6.48, $C\underline{H}(CH_3)_2$), 1.54 (2H, t, $J_{CH}$ 6.89, $CHC\underline{H}_2CH$), 1.37 (3H, d, $J_{CH}$ 6.57, $CHC\underline{H}_3$) 1.10 (3H, t, *J* 7.02, $NCH_2C\underline{H}_3$), 0.99 (3H, t, *J* 6.90, $NCH_2C\underline{H}_3$), and 0.877 (6H, d, $J_{CH}$ 6.33, $CH(C\underline{H}_3)_2$ (*cis* and *trans* isomers possible); $\delta_C$ (75.5 MHz, $C^2HCl_3$) 172.79, 172.43, 168.58, 167.16, 167.11 (5 x

CO), 156.32 ($CO_2NH$), 136.67 (*i*-$C_6H_5$), 128.37 (*m*-$C_6H_5$), 127.92 (*p*-$C_6H_5$), 127.82 (*o*-$C_6H_5$), 66.60 ($\underline{C}H_2C_6H_5$), 53.45 ($\underline{C}HCH_2$), 48.64 ($\underline{C}HCH_3$), 43.27 ($CHC\underline{H}_2$), 43.04, 42.60 ($N(CH_3)\underline{C}H_2$), 42.76 ($NHCH_2$), 41.67, 41.25 ($NHCH_aH_b$), 33.29, 32.71 ($NCH_3$), 24.72 ($\underline{C}H(CH_3)_2$), 22.99, 22.07 ($CH(\underline{C}H_3)_2$), 19.42 ($CH\underline{C}H_3$) and 13.19, 12.08 ($NHCH_2\underline{C}H_3$); *m/z* (CI) 492 (80%, $[M + H]^+$), 420 (9, $[M - HNC(CH_3)CH_2CH_3]^+$) and 384 (100 $[M - OCH_2C_6H_5]^+$).

## *3.0.2.5 Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl, N', O-dimethyl-glycinamide (23)*

This compound was prepared in a manner identical to that described for carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl, N', N-dimethyl-glycinamide (**21**) using N, O, dimethyl glycinamide hydrochloride (**61**) (158 mg, 1.02 mmol) to give an off white solid which was recrystalised from acetone / ether (315 mg, 63%), m.p. 141-142°C; (Found: C, 55.95; H, 7.15; N, 14.20. $C_{23}H_{35}N_5O_7$ requires C, 55.95; H, 7.15; N, 14.20%); *m/z* (Found: $[M + H]^+$, 494.2630. $C_{23}H_{36}N_5O_7$ requires 494.2615); $[\alpha]_D$ -28.1 (*c* 0.19 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3300 (NH), 3085 (aromatic CH), 2960, 2875 (2 x aliphatic CH), 1720 ($CO_2NH$) and 1705, 1680, 1655, 1635 (4 x CONH); $\delta_H$ (300 MHz; $C^2HCl_3$) 7.59 (3H, s, 2 x $N\underline{H}CH_2$ and $N\underline{H}CH_3$), 7.30 (5H, m, $C_6H_5$), 6.05 (1H, s, $NHCO_2$), 5.08 (1H, d, $J_{H_b}$ 18.5, $C_6H_5C\underline{H}_aH_b$), 5.04 (1H, d, $J_{H_a}$ 18.6, $C_6H_5CH_a\underline{H}_b$), 4.74 (1H, m, $NHC\underline{H}CH_2$), 4.40 (1H, m, *J* 6.6, $C\underline{H}CH_3$), 4.20 (2H, d, $J_{NH}$ 3.99, $NHC\underline{H}_2$) 4.15 (2H, d, $J_{NH}$ 6.00, $NHC\underline{H}_2$), 3.65 (3H, s, $OCH_3$), 3.12 (3H, s, $NCH_3$), 1.65 (1H, m, *J* 6.49, $C\underline{H}(CH_3)_2$), 1.54 (2H, m, *J* 6.21, 6.63, $CHC\underline{H}_2CH$), 1.36 (3H, d, $J_{CH}$ 6.75, $CHC\underline{H}_3$) and 0.88 (6H, d, $J_{CH}$ 6.33, $CH(C\underline{H}_3)_2$; $\delta_C$ (75.5 MHz, $C^2HCl_3$) 177.32, 172.74, 172.54, 168.92 (4 x CO), 156.37 ($CO_2NH$), 136.50 (*i*-$C_6H_5$), 128.46 (*m*-$C_6H_5$), 128.06 (*p*-$C_6H_5$), 127.93 (*o*-$C_6H_5$), 66.85 ($\underline{C}H_2C_6H_5$), 61.46 ($OCH_3$),

53.54 (CHCH$_2$), 48.85 (CHCH$_3$), 43.08 (NHCH$_2$), 42.14 (CHCH$_2$), 40.84 (NHCH$_2$), 32.33 (NCH$_3$), 24.72 (CH(CH$_3$)$_2$), 23.03, 22.00 (CH(CH$_3$)$_2$) and 18.98 (CHCH$_3$); *m/z* (CI) 494 (54%, [M + H]$^+$), 433 (28, [M - N(OCH$_3$)CH$_3$]$^+$), 376 (25 [M - NHCH$_2$CON(CH$_3$)OCH$_3$]$^+$) and 319 (22, [M - CH(CH$_3$)CONHCH$_2$CON(CH$_3$)OCH$_3$]$^+$]).

### *3.0.2.6 Tertiary butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycyl glycyl p-nitro anilide (24)*

This compound was prepared in a manner identical to that described for carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl-, N', N-dimethyl-glycinamide (**21**) using t-Boc-(2S)-leucyl (2S)-alanyl glycinate (**46**) (200 mg, 0.557 mmol) and glycine *p*-nitroanilide hydrochloride to give a pale yellow solid which was recrystallised from MeOH / ether to give a white solid (248 mg, 83%). m.p. 188°C; (Found: C, 53.4; H, 6.45; N, 15.55. $C_{24}H_{36}N_6O_8$ requires C, 53.7; H, 6.75; N, 15.65%); *m/z* (Found: ([M + H]$^+$, 537.2680). $C_{24}H_{37}N_6O_8$ requires 537.2680); $[\alpha]_D$ -45.0 (*c* 0.10 in MeOH); $\upsilon_{max}$ (KBr disk)/cm$^{-1}$ 3380, 3300 (NH), 3180, 3110, 3050 (3 x aromatic CH), 2990, 2900, 2870 (3 x aliphatic CH), 1680 ($CO_2NH$) and 1720, 1650, 1640, 1635 (4 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 8.25 (2H, d, $J_{CH}$ 9.27, CHCNO$_2$), 7.99 (2H, d, $J_{CH}$ 9.25, CHCNHCO), 4.33 (1H, q, $J_{CH_3}$ 7.14, CHCH$_3$), 4.16 (1H, d, $J_{H_b}$ 17.09, CH$_a$H$_b$), 4.10 (1H, m, NHCHCH$_2$), 4.05 (1H, d, $J_{H_a}$ 17.06, CH$_a$H$_b$), 4.02 (1H, d, $J_{H_d}$ 16.89, CH$_c$H$_d$), 3.89 (1H, d, $J_{H_c}$ 16.90, CH$_c$H$_d$), 1.67 (1H, m, *J* 6.70, CH(CH$_3$)$_2$), 1.46 (11H, s, C(CH$_3$)$_3$ and CHCH$_2$CH), 1.45 (3H, d, $J_{CH}$ 7.02, CHCH$_3$), 0.89 (3H, d, *J* 6.55, CH(CH$_3$)$_2$ and 0.85 (3H, t, *J* 6.53, CH(CH$_3$)$_2$); $\delta_C$ (75.5 MHz, $C^2H_3O^2H$) 176.71, 176.62, 172.57, 170.48 (4 x CO), 158.48 ($CO_2NH$), 146.00, 144.99 (*i*-$C_6H_5$), 126.05, 120.84 (*m*-$C_6H_5$ and *o*-$C_6H_5$), 80 99 (C(CH$_3$)$_3$), 54.64 (CHCH$_2$), 51.50 (CHCH$_3$), 44.60,

44.44 ($NHCH_2$), 42.23 ($CHCH_2$), 29.04 ($C(CH_3)_3$), 26.06 ($CH(CH_3)_2$), 23.82, 22.04 ($CH(CH_3)_2$) and 17.74 ($CHCH_3$); *m/z* (CI) 537 (6%, $[M + H]^+$), 437 (68 $[M + 2H - CO_2C(CH_3)_3]^+$) and 185 (100, $[NHCHCH_3CONHCH_2CONHCH_2CO]^+$).

### *3.0.2.7 (2S)-Leucyl (2S)-alanyl glycyl glycinamide acetate (25)*

To a stirred solution of carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinamide (**19**) (86 mg, 0.191 mmol) in glacial acetic acid (5 $cm^3$) was added hydrogen and the hydrogen atmosphere was maintained for 3 h. The reaction mixture was filtered through celite and the solvent was removed under reduced pressure to give a white solid (65 mg, 92%), m.p. 126°C (decomp.); *m/z* (Found: $[M + H - HOAc]^+$ 316.1980, $C_{13}H_{26}N_5O_4$ requires 316.1985); $[\alpha]_D$ -9.23 (*c* 0.26 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3600 - 2800 (NH and CH), 1705, 1685, 1670, 1655 (4 x CONH) and 1630 ($CO_2$); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 4.38 (1H, q, $J_{CH_3}$ 6.85, $CHCH_3$), 4.00 (1H, d, $J_{H_b}$ 16.83, $CH_aH_b$), 3.91 (2H, s, $NHCH_2$), 3.87 (1H, d, $J_{H_a}$ 20.04, $CH_aH_b$), 3.80 (1H, m, $NHCHCH_2$), 1.99 (5H, s, $CONH_2$ and $COCH_3$), 1.70 (3H, m, $CH_2CH(CH_3)_2$), 1.45 (3H, d, $J_{CH}$ 6.96, $CHCH_3$) and 1.03 (6H, t, *J* 5.50, $CH(CH_3)_2$); $\delta_C$ (75.5 MHz, $C^2H_3O^2H$) 175.45, 174.38, 173.09, 172.01 (4 x CO), 53.31 ($CHCH_2$), 51.14 ($CHCH_3$), 43.87, 43.22 ($CH_2CO$), 42.70 ($CHCH_2$), 25.53 ($CH(CH_3)_2$), 22.38 ($CO_2CH_3$), 23.29, 22.18 ($CH(CH_3)_2$) and 17.52 ($CHCH_3$); *m/z* (CI) 316 (100%, $[M + H - HOAc]^+$) and 185 (23, $[M - HOAc - NHCH_2CONHCH_2CONH_2]^+$).

### *3.0.2.8 (2S)-Leucyl (2S)-alanyl glycyl glycinamide-N'-methyl acetate (26)*

This compound was prepared in a manner identical to that described for (2S)-leucyl (2S)-alanyl glycyl glycinamide acetate (**25**) using carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl, N-methyl-glycinamide (**20**) (60 mg, 0.13 mmol) to give a white solid (49 mg, 94%), m.p. 130°C; *m/z* (Found: $[M + H - HOAc]^+$ 330.2144. $C_{14}H_{27}N_5O_4$ requires 330.2141); $[\alpha]_D$ -4.0 (c 0.08 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3600 - 3000 (NH), 2960 (aliphatic CH), 1670, 1665, 165, 1650 (CONH) and 1640 ($CO_2$); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 4.37 (1H, q, $J_{CH_3}$ 7.10, C<u>H</u>$CH_3$), 3.99 (1H, d, $J_{H_b}$ 16.8, C<u>H</u>$_a$H$_b$), 3.89 (2H, s, NHC<u>H</u>$_2$), 3.86 (1H, d, $J_{H_a}$ 16.4, CH$_a$<u>H</u>$_b$), 3.74 (1H, m, NHC<u>H</u>$CH_2$), 2.78 (3H, s, $NCH_3$), 1.98 (3H, s, $CO_2CH_3$), 1.68 (3H, m, C<u>H</u>$_2$C<u>H</u>$(CH_3)_2$), 1.45 (3H, d, CHC<u>H</u>$_3$) and 1.02 (6H, t, *J* 6.00, CH(C<u>H</u>$_3$)$_2$); $\delta_C$ (75.5 MHz, $C^2H_3O^2H$) 175.58, 174.09, 172.11, 172.04 (4 x CO), 53.41 (<u>C</u>HCH$_2$), 51.01 (<u>C</u>HCH$_3$), 43.87 (NH<u>C</u>H$_a$H$_b$), 43.47 (NHCH$_2$), 43.20 (CHC<u>H</u>$_2$), 26.35 ($NCH_3$), 25.50 (<u>C</u>H$(CH_3)_2$), 22.35, 22.10 (CH(<u>C</u>H$_3$)$_2$) and 17.45 (CH<u>C</u>H$_3$); *m/z* (CI) 330 (100%, $[M + H - HOAc]^+$) and 185 (20, $[M - HOAc - HNCH_2CONHCH_2CONHMe]^+$).

### *3.0.2.9 Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycine (27)*

Tertiary-butyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycinate (**38**) (2.07 g, 4.61 mmol) was dissolved in formic acid (95%, 30 $cm^3$) and stirred for 3 h.[100] The solution was neutralised to pH 8 with NaOH (1 $moldm^{-3}$) and reacidified to pH 3 with HCl (1 $moldm^{-3}$). The mixture was extracted with ethylacetate (3 x 30 $cm^3$) and the combined organic extracts washed with brine (saturated, 40 $cm^3$), dried ($MgSO_4$) and the solvent was removed under reduced pressure. The last traces of ethylacetate were azeotropically removed with

methylene chloride to give a yellow solid which was recrystalised from acetone / ether. Usually a gel resulted from recrystalisation, which was triturated with ether to give a white solid (1.49 g, 83%), m.p. 131 °C; (Found: C, 57.8; H, 6.7; N, 10.50. $C_{19}H_{27}N_3O_6$ requires C, 58.0; H, 6.9; N, 10.70 %); *m/z* (Found: $[M + H]^+$, 394.1978. $C_{19}H_{28}N_3O_6$ requires 394.1978); $[\alpha]_D$ -50.3 (*c* 0.32 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3307 (NH), 3035 (aromatic CH), 2960, 2870 (aliphatic CH), 1735, 1690 ($CO_2NH$ and $CO_2H$) and 1652, 1623 (2 x CONH); $\delta_H$ (400 MHz; $C^2H_3O^2H$) 7.39 (5H, m, $C_6H_5$), 5.15 (1H, d, $J_b$ 12.56, $C_6H_5C\underline{H}_aH_b$), 5.11 (1H, d, $J_a$ 12.44, $C_6H_5CH_a\underline{H}_b$), 4.42 (1H, q, $J_{CH_3}$ 7.13, $C\underline{H}CH_3$), 4.20 (1H, dd, *J* 6.32, 8.76, $NHC\underline{H}CH_2$), 3.94 (1H, s, $NHC\underline{H}_aH_b$), 3.92 (1H, s, $NHCH_a\underline{H}_b$), 1.75 (1H, sep, $J_{CH}$ 6.71, $C\underline{H}(CH_3)_2$), 1.61 (2H, m, $CHC\underline{H}_2CH$), 1.41 (3H, d, $J_{CH}$ 7.16, $CHC\underline{H}_3$), 0.99 (3H, d, $J_{CH}$ 6.67, $CH(C\underline{H}_3)_2$) and 0.97 (3H, d, $J_{CH}$ 6.75, $CH(C\underline{H}_3)_2$); $\delta_C$ (100 MHz; $C^2H_3OH$) 175.28, 175.03, 175.96 (2 x NHCO and $CO_2H$), 158.76 ($CO_2NH$), 138.14 (*i*-$C_6H_5$), 129.48 (*m*-$C_6H_5$), 129.03 (*p*-$C_6H_5$), 128.86 (*o*-$C_6H_5$), 67.80 ($C_6H_5\underline{C}H_2$), 41.88 ($\underline{C}HCH_3$), 41.88 ($NHCH_2$ and $\underline{C}H_2CH$), 25.85 ($\underline{C}H(CH_3)_2$), 23.46, 21.80 ($CH(\underline{C}H_3)_2$) and 18.00 ($CH\underline{C}H_3$); *m/z* (CI) 411 (12 %, $[M + H + NH_3]^+$), 394 (40, $[M + H]^+$), 319 (10, $[M - CH_3 - CH_2CO_2H]^+$), 303 (38, $[M + H - C_6H_5CH_2]^+$) and 286 (100, $[M - OCH_2C_6H_5]^+$).

### *3.0.2.10 Glycyl-amide hydrochloride (28)*[108]

Methyl glycinate hydrochloride (**54**) (1.00 g, 8.00 mmol) was dissolved in $NH_3$ solution (conc, 30 $cm^3$) and stirred for 2 days.[107] The ammonia was flushed from the solution with $N_2$ for 10 min and the solvent was removed under reduced pressure to give a white solid which was recrystallised from methanol. (0.78 g, 89%), decomp. 182°C (lit.,[108] 200-203°C); *m/z* (Found: $[M - Cl]^+$, 75.0557. $C_2H_6N_2O$ requires 75.0558); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3300, 3200, 3120, (3 x NH), 1710 (CONH); $\delta_H$ (200 MHz; $^2H_2O$) 3.71 (2H, s, $CH_2$); $\delta_C$ (50.3 MHz;

$^2H_2O$) 172.13 ($CO_2$) and 42.91 ($CH_2$); *m/z* (EI) 74 (4%, $[M - HCl]^+$), 44 (12, $[M - NH_2CH_2]^+$) and 30 (100, $[M - HCl - CONH_2]^+$).

### *3.0.2.11 Carbobenzyloxy-(2S)-leucyl-(2S)-alanine (29)*[99]

To a solution of methyl carbobenzyloxy-(2S)-leucyl (2S)-alaninate (**37**) (2.0 g, 5.7 mmol) in methanol (30 $cm^3$), NaOH (1 $moldm^{-3}$, 6.85 $cm^3$, 6.85 mmol) was added at 35°C. The reaction mixture was followed to completion by silica TLC eluting with 5% methanol / 95% methylene chloride (4½ hours).

The methanol was removed under reduced pressure and water (20 $cm^3$) was added to the aqueous phase. This was acidified to pH 2 using HCl (4 $moldm^{-3}$). The aqueous layer was extracted with chloroform (3 x 100 $cm^3$) and the combined chloroform extracts were washed with brine (saturated, 50 $cm^3$), dryed ($MgSO_4$) and the chloroform was removed under reduced pressure to give a white solid which was recrystalised from ether / hexane to give white crystals (1.68 g, 88%), m.p. 146-147°C (lit.,[99] 152-153°C); (Found: C, 60.7; H, 7.15; N, 8.40. $C_{17}H_{24}N_2O_5$ requires C, 60.7; H, 7.2; N, 8.35%) (lit.,[99] Found: C, 60.4; H, 7.2; N, 8.44.); *m/z* (Found: $[M + H]^+$, 337.1763. $C_{17}H_{25}N_2O_5$ requires 337.1763); $[\alpha]_D$ -30.8 (*c* 0.32 in MeOH); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3290, 3280 (2 x NH), 1720 ($CO_2H$), 1710 ($CO_2NH$) and 1700 (CONH); $\delta H$ (300 MHz; $C_2HCl_3$) 7.26 (5H, m, $C_6H_5$), 6.85 (1H, d, $J_{CH}$ 6.66, OCONH), 6.35 (1H, s, $CO_2H$), 5.03 (2H, s, $C_6H_5C\underline{H}_2$), 5.51 (1H, d, *JCH* 8.67, $N\underline{H}CHCH_3$), 4.48 (1H, sep, $C\underline{H}CH_3$), 4.23 (1H, m, $NHC\underline{H}CH_2$), 1.52 (3H, m, $C\underline{H}2C\underline{H}(CH_3)_2$), 1.35.(3H, d, $J_{CH}$ 7.05, $CHC\underline{H}_3$) and 0.85 (6H, d, $J_{CH}$ 2.91, $CH(C\underline{H}_3)_2$); $\delta C$ (50.3 MHz, $C_2HCl_3$) 176.15 ($CO_2H$), 173.14 (CONH), 157.01 (0CONH), 136.51 (*i*-$C_6H_5$), 129.04 (*m*-$C_6H_5$), 128.75 (*p*-$C_6H_5$), 128.51 (*o*-$C_6H_5$), 67.72 ($C_6H_5C\underline{H}_2$), 53.85 ($NH\underline{C}HCH_2$), 48.68 ($C\underline{H}CH_3$), 41.80 ($CH\underline{C}H_2CH$), 25.06 ($\underline{C}H(CH_3)_2$), 23.39, 22.47 ($CH(\underline{C}H_3)_2$) and 18.36 ($CH\underline{C}H_3$); *m/z* (CI)

337 (7%, $[M + H]^+$), 293 (3, $[M + H - CO_2]^+$), 246 (26, $[M + H - C_6H_5CH_2]^+$), 229 (100, $[M - OCH_2C_6H_5]^+$) and 203 (10, $[M + 2H - C_6H_5CO_2]^+$).

### *3.0.2.12 Carbobenzyloxy-(2S)-leucine (32)*[98]

(2S)-Leucine (2.97 g, 22.6 mmol) was dissolved in NaOH (1 moldm$^{-3}$, 25 cm$^3$) and the solution cooled to -5°C. NaOH (1 moldm$^{-3}$, 25 cm$^3$) and benzylchloroformate (4.24 g, 24.9 mmol) were added in alternate portions over 30 min in 6 aliquots with vigorous at room temperature. After for 2 h, the solution was washed with ether (2 x 25 cm$^3$) and the aqueous layer was acidified to pH 4 with HCl (5 moldm$^{-3}$) and then to pH 2 with HCl (2 moldm$^{-3}$) at 0°C. The acidified aqueous layer was extracted with ether (3 x 50 cm$^3$) and the combined ethereal extracts were washed with HCl (1 moldm$^{-3}$, 50 cm$^3$), brine (saturated, 50 cm$^3$) and dryed ($MgSO_4$). The solvent was removed under reduced pressure to give a colourless oil (5.29 g, 88%), *m/z* (Found: $[M]^+$, 265.1321. Calc. for $C_{14}H_{19}NO_4$ 265.1314); $[\alpha]_D$ -19.2 (*c* 0.37 in MeOH) (lit.,[98] -7.86 (*c* 0.31 in MeOH)); $\upsilon_{max}$ (thin film)/cm$^{-1}$ 3320, 3120 (OH and NH), 3090, 3070, 3040 (2 x aromatic CH), 2960, 2870, 2290 (3 x aliphatic CH), 1720, 1710 ($CO_2NH$ and $CO_2$); $\delta_H$ (200 MHz; $C^2HCl_3$) 9.5 (1H, s, $CO_2H$), 7.36 (5H, m, $C_6H_5$), 6.05 (1H, d, NH), 5.13 (2H, s, $C\underline{H}_2C_6H_5$), 4.38 (1H, m, NCH), 1.65 (3H, m, $C\underline{H}_2C\underline{H}(CH_3)_2$) and 0.97 (6H, d, $J_{CH}$ 3.94, $CH(C\underline{H}_3)_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 178.71 ($CO_2H$), 156.71 (OCONH), 136.59 (*i*-$C_6H_5$), 129.04 (*m*-$C_6H_5$), 128.74 (*p*-$C_6H_5$), 128.62 (*o*-$C_6H_5$), 67.69 ($C_6H_5\underline{C}H_2$), 52.88 (NCH), 41.94 ($CH\underline{C}H_2CH$), 25.25 ($\underline{C}H(CH_3)_2$) and 23.34, 22.20 ($CH(\underline{C}H_3)_2$); *m/z* (EI) 265 (2%, $[M]^+$), 174 (1, $[M - CH_2C_6H_5]^+$), 130 (3, $[M - CO_2CH_2C_6H_5]^+$) and 91 (100, $[C_6H_5CH_2]^+$).

### 3.0.2.13 Methyl (2S)-alaninate hydrochloride (33)[97]

To a suspension of (2*S*)-alanine (6.75 g, 75.8 mmol) in dry methanol (50 $cm^3$) was added dropwise $SOCl_2$ (12 $cm^3$, 0.165 mol) at 0°C. The reaction mixture was refluxed for 2 h. Dry ether was then added dropwise until the solution went cloudy and the product was allowed to crystallise out at -20°C. After 2 days the white needle shaped crystals were filtered under $N_2$. The very hydroscopic methyl (2*S*)-alanate hydrochloride was dried under vacuum (9.45 g, 89%), m.p. 110-112°C (lit.,[97] 110.5-111°C); (Found: C, 34.35; H, 7.3; N, 10.3. Calc. for $C_4H_{10}NO_2Cl$ C, 34.4; H, 7.2; N, 10.05%); *m/z* (Found $[M - Cl]^+$, 104.0712. $C_4H_{11}NO_2$ requires 104.0712); $[\alpha]_D$ +7.59 (*c* 0.29 in MeOH) (lit.,[97] +5.8 (c 0.29 in MeOH)); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3430 (NH) and 1740 (CO); $\delta_H$ (200 MHz; $^2H_2O$) 4.10 (1H, q, $J_{CH_3}$ 7.30, CH), 3.74 (3H, s, $OCH_3$) and 1.53 (3H, d, $J_{CH}$ 7.30, $CH_3$); $\delta_C$ (50.3 MHz; $^2H_2O$) 174.13 ($CO_2$), 56.43 ($OCH_3$), 51.84 (CH) and 17.95 ($CH_3$); *m/z* (CI) 121 (10%, $[M + NH_3 - Cl]^+$), 104 (100, $[M - Cl]^+$), 88 (2, $[M - HCl - CH_3]^+$) and 30 (44, $[M - Cl - CH_3 - CO_2CH_3]^+$).

### 3.0.2.14 Carbobenzyloxy-glycine (34)[96]

This compound was prepared in a manner identical to that described for the carbobenzyloxy-(2S)-leucine (**32**) using glycine (3.00 g, 40.0 mmol). The work up was modified as follows; after neutralisation the white solid was filtered, washed with water, dried and recrystallised from methylene chloride / petrol to give white crystals (7.31 g, 87%), m.p. 118-119°C (lit,[96] 121°C); (Found: C, 57.4; H, 5.35; N, 6.60. Calc. for $C_{10}H_{11}NO_4$ C, 57.4; H, 5.3; N, 6.7%); *m/z* (Found: $[M + H + NH_3]^+$, 227.1032. $C_{10}H_{15}N_2O_4$ requires 227.1032); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3330 and 3030 (OH and NH), 1730, 1710 (2 x $CO_2NH$) and 1690, 1680 (2 x $CO_2H$); $\delta_H$ (200 MHz; $(C^2H_3)_2SO$) 7.61 (1H, t, $J_{CH_2}$ 5.90, NH), 7.38 (5H, m, $C_6H_5$), 5.06 (2H, s, $C_6H_5C\underline{H}_2$) and 3.69 (2H, d, $J_{NH}$ 6.2, $NCH_2$); $\delta_C$ (50.3 MHz; $(C^2H_3)_2SO$)

171.88 ($CO_2H$), 156.78 (CONH), 137.28 (*i*-$C_6CH_5$), 128.64 (*m*-$C_6H_5$), 128.10 (*p*-$C_6H_5$), 128.01 (*o*-$C_6H_5$), 65.73 ($C_6H_5\underline{C}H_2$) and 42.40 ($NCH_2$); *m/z* (CI) 227 (23%, $[M + H + NH_3]^+$), 210 (10, $[M + H]^+$), 166 (18, $[M + H - CO_2]^+$), 119 (7, $[M + H - C_6H_5CH_2]^+$), 104 (100, $[M + 2H - C_6H_5CH_2O]^+$), 91 (12, $[M - CO_2NHCH_2CO_2H]^+$) and 76 (27, $[M + 2H - C_6H_5CH_2O_2C]^+$).

### 3.0.2.15 Methyl carbobenzyloxy-(2S)-leucyl alaninate (37)[98]

Carbobenzyloxy-(2S)-leucine (3.00 g, 11.3 mmol) was dissolved in dry THF (50 $cm^3$) with stirring and NMM (1.24 $cm^3$, 11.3 mmol) was added. The mixture was stirred for 5 minutes and cooled to -15 °C. i-BCF (1.48ml, 11.3 mmol) was added and a white precipitate of NMM.HCl formed. Methyl alaninate hydrochloride (1.58 g, 11.3 mmol) was dispersed in dry DMF (6ml) and NMM (1.24ml, 11.13 mmol) was added to the DMF solution. When a clear solution formed (heat may be necessary) this was added to the THF solution 5 min. after the i-BCF and the reaction mixture was allowed to stir for 5 min. at -15°C and then for 2 h at room temperature.

The NMM salts were filtered from the reaction mixture (sometimes they dissolved in the DMF / THF solution) and the THF was removed under reduced pressure.[95] Water was added to the DMF solution so an azeotrope formed and the solvent was removed under reduced pressure. This was repeated several times to give a white solid which was dissolved in ether (70 $cm^3$), washed with citric acid (10%, 50 $cm^3$), $NaHCO_3$ (saturated, 50 $cm^3$) and brine (saturated, 30 $cm^3$). The organic layer was dried ($MgSO_4$) and the solvent was removed under reduced pressure to give a white powder which was recrystallised from ether / hexane to give a white solid (3.76 g, 95%), m.p. 91°C, (lit.[98] 92-93C); (Found: C, 61.5; H, 7.4; N, 8.25. Calc. for $C_{18}H_{26}N_2O_5$ C, 61.7; H, 7.5; N, 8.00%); *m/z* (Found: $[M + H]^+$, 351.1920.

Calc. for $C_{18}H_{27}N_2O_5$ 351.1920); $[\alpha]_D$ -41.1 (*c* 0.28 in MeOH); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3300 (NH), 1740 ($CO_2Me$), 1690 ($CO_2NH$) and 1650 (CONH); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.35 (5H, m, $C_6H_5$), 6.58 (1H, d, $J_{CH}$ 5.32, CONH), 5.26 (1H, d, $J_{CH}$ 7.96, $CO_2NH$), 5.11 (2H, s, $C_6H_5C\underline{H}_2$), 4.56 (1H, sep, $J_{CH}$ 7.16, $C\underline{H}CH_3$), 4.20 (1H, m, $J_{CH}$ 5.14, $NHC\underline{H}CH_2$), 3.75 (3H, s, $OCH_3$), 1.62 (3H, m, $C\underline{H}_2C\underline{H}(CH_3)_2$), 1.39 (3H, d, $J_{CH}$ 7.04, $CHC\underline{H}_3$) and 0.94 (6H, d, $J_{CH}$ 6.60, $CH(C\underline{H}_3)_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 173.71 ($\underline{C}O_2CH_3$), 172.54 (NHCO), 156.77 ($CO_2NH$), 136.71 (*i*-$C_6H_5$), 129.00 (*m*-$C_6H_5$), 128.64 (*p*-$C_6H_5$), 128.51 (*o*-$C_6H_5$), 67.49 ($C_6H_5\underline{C}H_2$), 53.84 ($OCH_3$), 52.94 ($NH\underline{C}HCH_2$), 48.48 ($\underline{C}HCH_3$), 42.10 ($CH\underline{C}H_2CH$), 25.07 ($\underline{C}H(CH_3)_2$), 23.41, 22.47 ($CH(\underline{C}H_3)_2$) and 18.56 ($CH\underline{C}H_3$); *m/z* (CI) 351 (23%, $[M + H]^+$), 243 (100, $[M - OCH_2C_6H_5]^+$), 183 (2, $[M - C_6H_5CH_2CO_2 - CH_3OH]^+$) and 155 (2, $[M + H - C_6H_5CH_2CO_2 - CO_2CH_3]^+$).

### *3.0.2.16 Tertiary-butyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycinate (38)*

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl (2S)-alaninate (**15**) using carbobenzyloxy-(2S)-leucyl (2S)-alanine (**29**) (1.76 g, 5.24 mmol) and tertiary-butyl glycinate acetate (**30**) (1.00 g, 5.24 mmol) and THF instead of DMF. The work up was modified as follows; the crude product was extracted into ethyl acetate before washing and the remaining ethyl acetate was azeotropically removed with methylene chloride. The white solid was recrystalised from acetone / water to give white needle shaped crystals (2.07 g, 88%). m.p. 111-112°C; (Found: C, 61.8; H, 7.95; N, 9.25. $C_{23}H_{35}N_3O_6$ requires C, 61.45; H, 7.85; N, 9.35%); *m/z* (Found: $[M + H]^+$, 450.2604. $C_{24}H_{35}N_3O_6$ requires 450.2604); $[\alpha]_D$ - 48.5 (*c* 0.33 in MeOH); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3300 (NH), 1755 ($CO_2C(CH_3)_3$), 1733 ($CO_2NH$) and 1690, 1640 (2 x CONH);

$\delta_H$ (200 MHz; $C^2HCl_3$) 7.35 (5H, m, $C_6H_5$), 6.94 (1H, t, $N\underline{H}CH_2$), 6.80 (1H, d, $J_{CH}$ 7.28, $N\underline{H}CHCH_3$), 5.52 (1H, d, $J_{CH}$ 8.04, $N\underline{H}CHCH_2$), 5.12 (1H, d, $J_{CH_b}$ 12.16, $C_6H_5C\underline{H}_aCH_b$), 5.11 (1H, d, $J_{CH_a}$ 12.28, $C_6H_5CH_aC\underline{H}_b$), 4.61 (1H, quin, $J_{NH,}$ 7.14, $J_{CH_3}$ 6.76, $C\underline{H}CH_3$), 4.23 (1H, q, $J_{NH}$ 8.21, $J_{CH_2}$ 4.0, $NHC\underline{H}CH_2$), 3.89 (2H, d, $J_{NH}$ 5.0, $NHC\underline{H}_2$), 1.65 (1H, m, $J_{CH_3}$ .6.4, $J_{CH_2}$ 7.84, $C\underline{H}(CH_3)_3$), 1.58 (1H, m, $J_{CH}$ 56, $J_{H_a}$ 9.36, $J_{CH_3}$ 5.6, $CHCH_a\underline{H}_bCH$), 1.51 (1H, m, $J_{CH}$ 4.06, $J_{H_b}$ 9.28, $J_{CH_3}$ 5.54, $CHC\underline{H}_aH_bCH$), 1.43 (9H, s, $C(CH_3)_3$), 1.36 (3H, d, $J_{CH}$ 6.84, $CHC\underline{H}_3$) and 0.90 (6H, d, $J_{CH}$ 6.4, $CH(C\underline{H}_3)_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 172.13, 171.99, 168.61 ($NHC\underline{O}$), 156.22 ($CO_2NH$), 136.09 (*i*-$C_6H_5$), 128.39 (*m*-$C_6H_5$), 128.05 (*p*-$C_6H_5$), 127.94 (*o*-$C_6H_5$), 82.18 ($\underline{C}(CH_3)_3$), 66.95 ($C_6H_5\underline{C}H_2$), 53.47 ($\underline{C}HCH_2$), 48.59 ($\underline{C}HCH_3$), 41.90 ($NHCH_2$), 41.56 ($\underline{C}H_2CH$), 27.88 ($C(\underline{C}H_3)_3$), 24.55 ($\underline{C}H(CH_3)_2$), 22.85, 21.73 ($CH(\underline{C}H_3)_2$) and 18.27 ($NHCH\underline{C}H_3$); *m/z* (CI) 467 (4%, $[M + H + NH_3)]^+$), 450 (13, $[M + H]^+$), 394 (88, $[M + H - C(CH_3)_3]^+$), 376 (5, $[M + H + NH_3 - C_6H_5CH_2]^+$), 342 (15, $[M - OCH_2C_6H_5]^+$), 303 (15, $[M + 2H - C(CH_3)_3 - C_6H_5CH_2]^+$) and 286 (100, $[M + H - C_6H_5CH_2O - C(CH_3)_3]^+$).

### *3.0.2.17 Ethyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycinate (39)*

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl (2S)-alaninate (**37**) using carbobenzyloxy-(2S)-leucyl (2S)-alanine (**29**) (1.76 g, 5.24 mmol) and ethyl glycinate hydrochloride (**40**) (0.73 g, 5.24 mmol). The work up was modified as follows; the crude product was extracted into ethyl acetate before washing, and the remaining ethyl acetate was azeotropically removed with methylene chloride from the white solid, which were recrystalised from acetone / water to give white crystals (1.54 g, 70%). m.p. 129°C; *m/z* (Found: $[M + H]^+$, 422.2291. $C_{21}H_{32}N_3O_6$ requires

422.2290); $\upsilon_{max}$ (nujol)/cm$^{-1}$ 3460 (NH), 3270 (aromatic CH), 2500 (aliphatic CH), 1750 ($CO_2Et$), 1710 ($CO_2NH$) and 1680, 1640 (2 x CONH); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.34 (5H, m, $C_6H_5$), 6.68 (1H, t, $J_{CH}$ 3.8, $N\underline{H}CH_2$), 6.71 (1H, d, $J_{CH}$ 6.8, $N\underline{H}CHCH_3$), 5.33 (1H, d, $J_{CH}$ 7.8, $N\underline{H}CHCH_2$), 5.11 (2H, $C_6H_5C\underline{H}_2$), 4.56 (1H, m, $J_{NH,}$ 6.8, $C\underline{H}CH_3$), 4.19 (3H, m, $J_{CH_3}$ 7.10, $OCH_2$ and $NHC\underline{H}CH_2$), 4.0 (2H, d, $J_{NH}$ 4.8, $NHC\underline{H}_2$), 1. 66 (3H, m, $C\underline{H}_2C\underline{H}(CH_3)_3$), 1.38 (3H, d, $J_{CH}$ 7.0, $CHC\underline{H}_3$) 1.27 (3H, t, $J_{CH_2}$ 7.1, $OCH_2C\underline{H}_3$) and 0.93 (6H, d, $J_{CH}$ 5.8, $CH(C\underline{H}_3)_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 172.84, 172.76, 170.17 ($NH\underline{C}O$), 156.89 ($CO_2NH$), 136.64 (*i*-$C_6H_5$), 129.06 (*m*-$C_6H_5$), 128.69 (*p*-$C_6H_5$), 128.52 (*o*-$C_6H_5$), 67.61 ($C_6H_5\underline{C}H_2$), 61.98 ($OCH_2$), 54.17 ($\underline{C}HCH_2$), 49.21 ($\underline{C}HCH_3$), 42.03 ($NHCH_2$), 41.82 ($\underline{C}H_2CH$), 25.17 ($\underline{C}H(CH_3)_2$), 23.45, 22.30 ($CH(\underline{C}H_3)_2$), 18.67 ($NHCH\underline{C}H_3$) and 14.60 ($OCH_2\underline{C}H_3$); *m/z* (CI) 439 (3%, $[M + H + NH_3)]^+$), 422 (9, $[M + H]^+$), 331 (10, $[M + H - C_6H_5CH_2]^+$), 314 (100, $[M - OCH_2C_6H_5]^+$).

### *3.0.2.18 Ethyl glycinate hydrochloride (40)*[116]

This compound was prepared in a manner identical to that described for the methyl alaninate hydrochloride (**33**) using glycine (5.69 g, 75.8 mmol) and dry ethanol and the work up was modified as follows; The solvent was removed under reduced pressure to give white needle shaped crystals which were recrystallised from ethanol / ether to give white needle shaped crystals (8.05 g, 76%), m.p. 143°C (lit.,[116] 145-148°C); (Found: C, 34.55; H, 7.25; N, 9.9. Calc. for $C_4H_{10}NO_2Cl$ C, 34.4; H, 7.2; N, 10.05%); *m/z* (Found: $[M - Cl]^+$, 104.0712. $C4H_{10}NO_2$ requires 104.0712); $\upsilon_{max}$ (Nujol)/cm$^{-1}$ 3480 (NH), 1750 (CO); $\delta_H$ (200 MHz; $^2H_2O$) 4.27 (2H, q, $J_{CH_2}$ 7.17, $OCH_2$), 3.88 (2H, s, $NHCH_2$), 1.26 (3H, t, $J_{CH_2}$ 7.18, $CH_3$); $\delta_C$ (50.3 MHz, $^2H_2O$) 171.04 (CO), 66.21 ($OCH_2$), 43.12 ($CH_2NH$) and 16.11 ($CH_3$); *m/z* (CI) 104 (100%, $[M - Cl]^+$), 76 (2, $[M - Cl - CO]^+$) and 45 (3, $[M - Cl - NH_2CH_2CO]^+$).

## 3.0.2.19 Tertiary-butoxycarbonyl-(2S)-leucine.$H_2O$ (41)[102]

(2S)-Leucine (3.68 g, 2.81 mol) was dissolved in NaOH (1 $moldm^{-3}$, 30 $cm^3$) with stirring. Di-tertiary-butyl dicarbamate (7.94 g, 3.38 mol) was dissolved in tertiary-butyl alcohol (30 $cm^3$) and added to the solution which was stirred vigorously for 2 h. The tertiary-butyl alcohol was removed under reduced pressure and the remaining solution was washed with ether (2 x 25 $cm^3$). The aqueous layer was acidified at 0°C to pH 2 with $H_2SO4$ (2 $moldm^{-3}$). The white crystals were filtered, washed with distilled water (50 $cm^3$ and dried to constant mass in a $CaCl_2$ desiccator (6.40 g, 91%), m.p. 82-84°C (lit.,[102] 83-84°C); (Found: C, 52.1; H, 9.3; N, 5.45. Calc. for $C_{11}H_{23}NO_5$ C, 52 0; H, 9.3; N, 5.60%); *m/z* (Found: $[M - OH]^+$ Calc. for 232.1549. $C_{11}H_{21}NO_4$ 232.1549.); $[\alpha]_D$ -40.0 (*c* 0.14 in MeOH); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3500 (OH), 3380 (NH), 2740 (alkyl CH), 1720 ($CO_2NH$) and 1670 ($CO_2H$); $\delta_H$ (200 MHz; $C^2HCl_3$) 8.30 (1H, s, $CO_2H$), 4.90 (1H, d, *JCH* 8.34, NH), 4.32 (1H, m, $CH_2C\underline{H}$), 1.54 (3H, m, $C\underline{H}_2C\underline{H}(CH_3)_2$), 1.45 (9H, s, $C(CH_3)_3$ and 0.96 (6H, d, *JCH* 6.26, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.74 MHz; $C^2HCl_3$) 177.78 ($CO_2H$), 155.64 (CONH), 80.07 ($O\underline{C}(CH_3)_3$), 53.95 (NHCH), 41.32 ($CH_2$), 28.18 ($C(\underline{C}H_3)_3$), 24.67 ($\underline{C}H(CH_3)_2$) and 22.73, 21.66 ($CH(\underline{C}H_3)_2$); *m/z* (CI) 249 (10%, $[M + H + NH_3]^+$), 232 (30, $[M + H]^+$), 193 (100, $[M + NH_3 + 2H - C(CH_3)_3]^+$), 176 (65, $[M + 2H - C(CH_3)_3]^+$), 149 (6, $[M + NH_3 + 2H - CO_2C(CH_3)_3]^+$), 132 (100, $[M + 2H - CO_2C(CH_3)_3]^+$), 116 (4, $[M - CH(CH_2CH(CH_3)_2)CO_2H]^+$), 86 (98, $[M - CO_2 - CO_2C(CH_3)_3]^+$) and 70 (4, $[M - NHCO_2C(CH_3)_3 - CO_2H]^+$).

## *3.0.2.20 Methyl tertiary-butoxycarbonyl-(2S)-leucyl (2S)-alaninate (42)*[103]

This compound was prepared in a manner identical to that described for the methyl benzoxycarbonyl-(2S)-leucyl (2S)-alaninate (**37**) using anhydrous tertiary-butoxycarbonyl-(2S)-leucine which was prepared as follows; tertiary-butoxycarbonyl-(2S)-leucine.$H_2O$ (**41**) (6.38 g, 2.55 mmol) was dissolved in methylene chloride (100 $cm^3$) and dried ($Na_2SO_4$). The solution was filtered and the solvent removed at reduced pressure to give a yellow oil. The crude dipeptide (**42**) was recrystallised from ether / hexane to give white crystals (7.92 g, 98%), m.p. 110°C, (lit.,[103] 114-116°C); (Found: C, 56.9; H, 9.05; N, 8.65. $C_{15}H_{28}N_2O_5$ requires C, 56.95; H, 8.9; N, 8.85%); *m/z* (Found: $[M + H]^+$, 317.2076. $C_{15}H_{29}N_2O_5$ requires 317.2076); $[\alpha]_D$ -65.2 (*c* 0.25 in MeOH); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3320, 3250 (2 x NH), 3080 (CH), 1740 ($CO_2CH_2C_6H_5$), 1680 ($CO_2NH$) and 1640 (CONH); $\delta_H$ (200 MHz; $C^2HCl_3$) 6.79 (1H, d, $J_{CH}$ 6.4, N<u>H</u>CHCH$_3$), 5.03 (1H, d, $J_{CH}$ 8.2, N<u>H</u>CHCH$_2$), 4.54 (1H, pent, $J_{CH_3}$ 7.2, C<u>H</u>CH$_3$), 4.18 (1H, m, NHC<u>H</u>CH$_2$), 3.72 (3H, s, $OCH_3$), 1.54 (3H, m, C<u>H</u>$_2$C<u>H</u>(CH$_3$)$_2$), 1.37 (3H, d, $J_{CH}$ 7.2, CHC<u>H</u>$_3$), 1.42 (9H, s, $C(CH_3)_3$) and 0.91 (6H, m, CH(C<u>H</u>$_3$)$_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 173.66, 172.82 (2 x CO), 156.17 (<u>C</u>O$_2$N), 80.41 (<u>C</u>(CH$_3$)$_3$), 53.35 ($OCH_3$), 52.87 (NH<u>C</u>HCH$_2$), 48.39 (<u>C</u>HCH$_3$), 41.81 ($CH_2$), 28.76 (C(<u>C</u>H$_3$)$_3$), 25.11 (<u>C</u>H(CH$_3$)$_2$), 23.43, 22.42 (CH(<u>C</u>H$_3$)$_2$) and 18.59 (CH<u>C</u>H$_3$); *m/z* (CI) 317 ($[M + H]^+$, 23%), 275 (18, $[M + 2H - CH(CH_3)_2]^+$), 261 (82, $[M + 2H - C(CH_3)_3]^+$), 243 (34, $[M + H - CH(CH_3)_2 - OCH_3]^+$), 231 (38, $[M + 2H - CH(CH_3)CO_2CH_3]^+$) and 217 (100, $[M + 2H - CO_2C(CH_3)_3]^+$).

### 3.0.2.21 Tertiary-butoxycarbonyl-(2S)-leucyl-(2S)-alanine (43)[104]

This compound was prepared in a manner identical to that described for the carboxy benzyl-(2S)-leucyl (2S)-alaninate (**29**) using methyl tertiary-butoxy carbonyl-(2S)-leucyl (2S)-alaninate (**42**) (6.00 g, 19.0 mmol) except that the hydroscopic tertiary-butoxycarbonyl-(2S)-leucyl (2S)-alanine crystals were filtered under nitrogen to give white crystals (4.73 g, 82%), m.p. 135°C (lit.,[104] 132-133°C); (Found: C, 55.5; H, 8.4; N, 9.10. Calc. for $C_{14}H_{26}N_2O_5$ C, 55.6; H, 8.65; N, 9.25%); *m/z* (Found: $[M + H]^+$, 303.1920. Calc. for $C_{14}H_{27}N_2O_5$ 303.1920); $[\alpha]_D$ -32.2 (*c* 0.34 in MeOH) (lit.,[104] -32.6 (*c* 1.01 in MeOH)); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3370, 3330 (2 x NH), 3090 (CH), 1730 ($CO_2NH$), 1670 ($CO_2H$) and 1640 (CONH); $\delta_H$ (400 MHz; $C^2HCl_3$) 7.80 (1H, s, $CO_2H$), 7.02 (1H, d, $J_{CH}$ 6.72, N<u>H</u>C$CH_3$), 5.22 (1H, d, $J_{CH}$ 8.36, N<u>H</u>CHC$H_2$), 4.54 (1H, q, C<u>H</u>$CH_3$), 4.21 (1H, dd, NHC<u>H</u>$CH_2$), 1.63 (1H, m, C<u>H</u>$(CH_3)_2$), 1.48 (2H, m, CHC<u>H</u>$_2$CH, 1.44 (3H, d, CHC<u>H</u>$_3$), 1.42 (9H, s, $C(CH_3)_3$) and 0.91 (6H, m, CH$(C\underline{H}_3)_2$); $\delta_C$ (54.3 MHz; $C^2HCl_3$) 176.03 ($CO_2H$), 173.45 (CONH), 156.53 . ($CO_2NH$), 80.88 ($\underline{C}(CH_3)_3$), 53.39 (NH<u>C</u>HC$H_2$), 48.62 (<u>C</u>H$CH_3$), 41.64 (CH<u>C</u>H$_2$CH), 28.77 ($C(\underline{C}H_3)_3$), 25.10 ($\underline{C}H(CH_3)_2$), 18.41 (CH<u>C</u>H$_3$) and 23.30, 22.48 ($CH(\underline{C}H_3)_2$); *m/z* (CI) 303 (10%, $[M + H]^+$), 246 (33, $[M + H - C(CH_3)_3]^+$), 229 (100, $[M - C(CH_3)_3 - OH]^+$), 203 (40, $[M + 2H - CO_2C(CH_3)_3]^+$), 185 (5, [M - H - $CONHCH(CH_3)CO_2H]^+$) and 86 (20, $[(NH_2)CH(CH_2CN(CH_3)_2]^+$).

### 3.0.2.22 Benzyl glycinate tosylate salt (44)[105]

Glycine (8.22 g, 112 mmol) was dispersed in toluene (150 $cm^3$) and benzyl alcohol (55 $cm^3$, 512 mmol) and *p*-toluene sulphonic acid.$H_2O$ (127.77 g, 134 mmol) were added. The reaction was refluxed under Dean and Stark conditions for 5 h. On cooling yellow needle shaped crystals formed which were filtered and washed with ether (150 $cm^3$) and

recrystalisation from methanol / ether gave white needle shaped crystals (30.58 g, 81%), m.p. 123°C (lit.,[105] 132-134°C); (Found: C, 57.25; H, 5.5; N, 4.20. Calc. for $C_{16}H_{19}NO_5S$ C, 56.95; H, 5.7; N, 4.15%); *m/z* (Found: $[M - C_7H_7SO_3 + H + NH_3]^+$ Calc. for 183.1134. $C_9H_{15}N_2O_2$ 183.1143); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 1750 (CO); $\delta_H$ (200 MHz; $^2H_2O$) 7.62 (2H, d, $J_{CH}$ 8.02, C<u>H</u>$CSO_3$), 7.37 (5H, s, $C_6H_5$), 7.29 (2H, d, $J_{CH}$ 4.72, C<u>H</u>$CCH_3$), 5.22 (2H, s, $C_6H_5C\underline{H_2}$), 3.88 (2H, s, $NCH_2$) and 2.31 (3H, s, $CH_3$); $\delta_C$ (50.3 MHz; $^2H_2O$) 170.85 (CO), 145.32, 142.22, 147.50 (*i*-$C_6CH_5$), 132.29, 131.81, 131.72, 131.40, 128.20 ($C_6H_5$), 71.20 ($C_6H_5\underline{C}H_2$), 43.00 ($NCH_2$) and 23.33 ($CH_3$); *m/z* (CI) 183 (17%, $[M + H + NH_3 - C_7H_7SO_3]^+$), 166 (100, $[M + H - C_7H_7SO_3]^+$) and 108 (18, $[M + H - COCH_2NH_2 - C_7H_7SO_3]^+$).

### *3.0.2.23 Benzyl tertiary-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycinate (45)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alaninate (**37**) using butoxycarbonyl-(2S)-leucyl (2S)-alanine (**29**) (2.50 g, 8.27 mmol), benzyl glycinate tosylate (**44**) (2.79 g, 8.27mol), THF instead of DMF. The work up was modified as follows; the crude product was extracted into ethyl acetate, before washing and the remaining ethyl acetate was azeotropically removed with methylene chloride to yield a clear, colourless hydroscopic oil which was stored under $N_2$. The oil was recrystalised from acetone / water to give mono hydrated white needle shaped crystals. (3.27 g, 88%). m.p. 68-69°C; (Found: C, 61.8; H, 8.15; N, 9.35. $C_{23}H_{35}N_3O_6$ requires C, 61.45; H, 7.85; N, 9.35%); *m/z* (Found: $[M + H]^+$, 450.2604). $C_{23}H_{36}N_3O_6$ requires 450.2604); $[\alpha]_D$ -41.5 (*c* 0.27 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3490 ($H_2O$), 3120 (aromatic CH), 2980, 2965, 2936 (3 x aliphatic CH), 1751, 1700 ($CO_2NH$ and

$CO_2H$),) and 1655, 1651 (2 x CONH); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.34 (5H, m, $C_6H_5$), 6.82 (1H, s, $N\underline{H}CHCH_3$), 6.57 (1H, d, $J_{CH}$ 7.44, $N\underline{H}CHCH_2$), 5.17 (1H, d, $J_{CH_b}$ 12.28, $C_6H_5C\underline{H}_aH_b$), 5.14 (1H, d, $J_{CH_a}$ 12.48, $C_6H_5CH_a\underline{H}_b$), 4.86 (1H, m, $C\underline{H}CH_2$), 4.51 (1H, quin., $J_{NH}$ 7.16, $J_{CH}$ 7.26, $C\underline{H}CH_3$), 4.08 (1H, dd, $J_{NH}$ 5.46, $J_{CH_b}$ 17.90, $NHC\underline{H}_aH_b$), 4.03 (1H, dd, $J_{NH}$ 5.20, $J_{CH_a}$ 8.20, $NHCH_a\underline{H}_b$), 1.67 ( 2H, m, $CHC\underline{H}_2CH$), 1.61 (2H, s, $H_2O$), 1.49 (1H, m, $C\underline{H}(CH_3)_2$), 1.42 (9H, s, $C(CH_3)_3$), 1.39 (3H, d, $J_{CH}$ 7.12, $CHC\underline{H}_3$), 0.94 (3H, d, $J_{CH}$ 6.04, $CH(C\underline{H}_3)_2$) and 0.92 (3H, d, $J_{CH}$ 5.76, $CH(C\underline{H}_3)_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 172.42, 172.07, 169.28 (NHCO), 155.90 ($CO_2NH$), 135.08 (*i*-$C_6H_5$), 128.52 (*m*-$C_6H_5$), 128.41 (*p*-$C_6H_5$), 127.28 (*o*-$C_6H_5$), 80.47 ($\underline{C}(CH_3)_3$), 67.06 ($C_6H_5\underline{C}H_2$), 53.32 ($\underline{C}HCH_2$), 48.60 ($\underline{C}HCH_3$), 41.23 ($NHCH_2$), 40.76 ($\underline{C}H_2CH$), 28.15 ($C(\underline{C}H_3)_3$), 24.66 ($\underline{C}H(CH_3)_2$), 22.87, 21.68 ($CH(\underline{C}H_3)_2$) and 17.69 ($CH\underline{C}H_3$); *m/z* (CI) 467 (11%, $[M + H + NH_3]^+$), 450 (30, $[M + H]^+$), 411 (20, $[M + 2H + NH_3 - C(CH_3)_3]^+$), 394 (60, $[M + 2H - C(CH_3)_3]^+$), 376 (13, $[M + H + NH_3 - C_6H_5CH_2]^+$) and 350 (100, $[M + 2H - CO_2C(CH_3)_3]^+$).

### *3.0.2.24 Tertiary butoxycarbonyl-(2S)-leucyl-(2S)-alanyl-glycine (46)*

To a solution of palladium (II) acetate (5 mg, 0.022 mmol) in dry acetone (20 $cm^3$), in a flame dried flask, under argon, was added dimethylethylsilane (0.88 $cm^3$, 0.668 mmol) and then dry triethylamine (10 $mm^3$, 71.2 μmol). The solution immediately went black.

Benzyl t-butoxycarbonyl-leucyl-alanyl-glycinate.$H_2O$ (**45**) (200 mg, 0.445 mmol) was dissolved in dry dichloromethane, dried ($MgSO_4$), filtered and the solvent removed under reduced pressure to give a yellow oil. The dried benzyl tertiary butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycinate was dissolved in acetone under argon and after 5 min, was added to the reaction mixture and stirred for 12 h (yielding the silyl ester).[106] The resulting solution was

filtered through celite and methanol (3 $cm^3$) was added to the filtrate, which was stirred for a further 3 h. The solvent was removed under reduced pressure to give a yellow oil which was recrystallised from AnalaR acetone / ether to give a gel which was triturated to give a white solid (153 mg, 96%), m.p. 111-112°C; (Found: C, 53.25; H, 8.45; N, 11.65. $C_{16}H_{29}N_3O_6$ requires C, 53.45; H, 8.15; N, 11.70%); *m/z* (Found: $[M + H]^+$, 360.2135. $C_{16}H_{29}N_3O_6$ requires 360.2135); $[\alpha]_D$ -58.3 (*c* 0.18 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3344, 3314 (2 x NH), 2966 (aliphatic CH), 1735 ($CO_2H$), 1700 ($CO_2NH$) and 1681, 1653 (2 x CONH); $\delta_H$ (300 MHz; $C^2H_3O^2H$) 4.46 (1H, q, $J_{CH_3}$ 7.06, C<u>H</u>$CH_3$), 4.18 (1H, dd, *J* 5.69, 9.26, C<u>H</u>$CH_2$), 3.96 (1H, d, $J_{H_b}$ 17.94, NHC<u>H</u>$_a$H$_b$), 3.89 (1H, d, $J_{H_a}$ 17.85, NHCH$_a$<u>H</u>$_b$), 1.72 (1H, sep, $J_{CH_3}$ 6.76, C<u>H</u>$(CH_3)_2$), 1.59 (2H, dd, *J* 5.22, 2.55, CHC<u>H</u>$_2$CH), 1.48 (9H, s, $C(CH_3)_3$), 1.42 (3H, d, $J_{CH}$ 7.14, CHC<u>H</u>$_3$), 0.99 (3H, d, $J_{CH}$ 6.40, CH(C<u>H</u>$_3$)$_2$) and 0.97 (3H, d, $J_{CH}$ 6.06, CH(C<u>H</u>$_3$)$_2$); $\delta_C$ (75.6 MHz; $C^2H_3O^2H$) 175.80, 175.24, 173.33 (2 x NH<u>C</u>O and $CO_2H$), 158.44 ($CO_2NH$), 81.05 (<u>C</u>$(CH_3)_3$), 54.98 (<u>C</u>H$CH_2$), 50.37 (<u>C</u>H$CH_3$), 42.31 ($NHCH_2$), 42.19 (<u>C</u>$H_2$CH), 29.00 (C(<u>C</u>H$_3$)$_3$), 26.19 (<u>C</u>H$(CH_3)_2$), 23.77, 22.14 (CH(<u>C</u>H$_3$)$_2$) and 18.44 (NHCH<u>C</u>H$_3$); *m/z* (CI) 377 (10%, $[M + H + NH_3]^+$), 360 (47, $[M + H]^+$), 321 (18, $[M + NH_3 + H - C(CH_3)_3]^+$), 304 (40, $[M + H - C(CH_3)_3]^+$) and 260 (100, $[M + 2H - CO_2C(CH_3)_3)]^+$).

### *3.0.2.25 Glycyl-N', methylamide hydrochloride (53)*[109]

This was prepared in a manner identical to that described for the glycinamide hydrochloride (**28**) using methylamine solution (saturated) to give white needle shaped crystals (0.94 g, 78%), m.p.148°C (lit.,[109] 153.5-156°C); (Found C, 28.65; H, 7.5; N, 22.3. Calc. for $C_3H_8N_2OCl$ C, 28.95; H, 7.3; N, 22.5%); *m/z* (Found: $[M - HCl]^+$, 88.0635. Calc. for $C_3H_8NO_2$ 88.0636); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3250, 3050 (2 x NH) and 1662 (CONH); $\delta_H$ (200

MHz, $^2H_2O$) 3.62 (2H, s, $CH_2$) and 2.61 (3H, s, $CH_3$); $\delta_C$ (50.3 MHz, $^2H_2O$) 170.12 (CO), 43.12 ($CH_2$) and 28.56 ($CH_3$); *m/z* (EI) 88 (8%, $[M - HCl]^+$) and 58 (9, $[M - HCl - NHCH_3]^+$).

### 3.0.2.26 Methyl glycinate hydrochloride (54)[114]

This compound was prepared in a manner identical to that described for the methyl alaninate hydrochloride (**33**) using glycine (5.69 g, 75.8 mmol) and the work up modified as follows; On cooling white needle shaped crystals formed, which were filtered and recrystallised from methanol / ether to give white needle shaped crystals (8.87 g, 94%), m.p.178°C (lit.,[114] 175°C); (Found C, 34.35; H, 7.3; N, 10.3. Calc. for $C_3H_8NO_2Cl$ C, 34.5; H, 7.25; N, 10.05%); *m/z* (Found: $[M - Cl]^+$, 90.0555. Calc. for $C_3H_8NO_2$ 90.0555); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3310 (NH), 2690, 2640 (2 x aliphatic CH) and 1750 (CO); $\delta_H$ (200 MHz, $^2H_2O$) 3.85 (2H, s, $CH_2$) and 3.76 (3H, s, $OCH_3$); $\delta_C$ (50.3 MHz, $^2H_2O$) 171.52 ($CO_2$), 49.92 ($CH_3$) and 56.22 ($CH_2$); *m/z* (EI) 90 (52%, $[M - Cl]^+$), 75 (100, $[M - Cl - CH_3]^+$) and 57 (30, $[M - HCl - H_2O - CH_3]^+$).

### 3.0.2.27 Glycyl- N', N', dimethylamide hydrochloride (55)

Ethyl acetate (30 $cm^3$) was saturated with HCl at -5°C for 30min. Then a solution of tertiary-butyl-glycyl N', N', dimethylamide (57) (0.35 g, 1.73 mmol) in ethyl acetate (5 $cm^3$) was added and the solution was stirred at -5°C for one hour.[110] The HCl gas was flushed from the solution with $N_2$ for 10 min and the solvent was removed under reduced pressure to give a white solid which was recrystallised from methanol / ether to give hydroscopic white needle shaped crystals (125 mg, 53%), (Found: C, 34.35; H, 7.95; N, 20.10. $C_4H_{11}N_2OCl$ requires C, 34.65; H, 8.6; N, 20.20%; *m/z* (Found: $[M - Cl]^+$, 103.0871. $C_4H_{11}N_2O$ requires

103.0871); $\upsilon_{max}$ (KBr disk)/cm$^{-1}$ 3420 (NH) and 1660 (CONH); $\delta_H$ (200 MHz, $^2H_3CO^2H$) 3.99 (2H, s, $CH_2$) and 3.05, 3.01 (6H, ds, $CH_3$); $\delta_C$ (50.3 MHz, $^2H_3CO^2H$) 167.48 (CO), 41.47 ($CH_2$) and 36.88, 36.22 ($CH_3$); *m/z* (CI) 205 (4%, $[2M - 2Cl - H]^+$), 103 (100, $[M - Cl]^+$), 72 (4, $[M - HCl - CH_2NH_2]^+$) and 70 ($M - HCl - NH_3 - CH_3]^+$).

### 3.0.2.28 Tertiary-butoxycarbonyl-glycine (56)[114]

This compound was prepared in a manner identical to that described for the tertiary-butoxycarbonyl-(2S)-leucine (**41**) using glycine (1 g, 13.0 mmol),with the following modifications; After neutralisation the product was extracted into ether (3 x 50 cm$^3$), washed with brine (saturated, 50 cm$^3$) and dried (MgSO4). The solvent removed under reduced pressure to give a clear oil, which was recrystallised from ether / hexane to give white needle shaped crystals (1.92 g, 85%), m.p. 86°C (lit.,[114] 85-89°C); $\upsilon_{max}$ (Nujol)/cm$^{-1}$ 3400 (OH), 3340 (NH), 1740 ($CO_2NH$) and 1670 ($CO_2H$); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.02 (1H, d, $J_{CH}$ 2.4, NH), 5.22 (1H, d, $J_{CH}$ 4.2, NH), 4.01 (2H, d, $J_{NH}$ 5.6, $CH_2$), 3.95 (2H, d, $J_{NH}$ 5.2, $CH_2$), 1.50 (9H, s, $CH_3$) and 1.47 (9H, s, $CH_3$); $\delta_C$ (75.74 MHz; $C^2HCl_3$) 175.48, 174.47 ($CO_2H$), 157.89, 156.52 (CONH), 82.34 ($O\underline{C}(CH_3)_3$), 43.91, 42.73 ($CH_2$) and 28.79, 28.50 ($CH_3$).

### 3.0.2.29 N-Tertiary-butoxycarbonyl-N', N'-Dimethyl glycinamide (57)

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl alaninate (**37**) using tertiary-butoxycarbonyl glycine (**56**) (0.40 g, 2.28 mmol) and dimethylamine solution (0.34 cm$^3$, 22.8 mmol). The work up was modified as follows; the mixture was extracted into ethyl acetate and the solvent removed at reduced pressure to give a white powder which was recrystallised from methanol to give white crystals (186 mg, 40%), m.p. 74°C; (Found: C, 53.8; H, 9.35; N, 13.9. $C_9H_{18}N_2O_3$ requires

C, 53.45; H, 8.95; N, 13.85%); *m/z* (Found: $[M]^+$ 202.1324. $C_9H_{18}N_2O_3$ requires 202.1317); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3274 (NH), 3040, 3019, 3010, 2978, 2939 (5 x aliphatic CH), 1720 ($CO_2NH$) and 1641 (CONH); $\delta_H$ (200 MHz; $C_2HCl_3$) 5.52 (1H, s, NH), 3.93 (2H, d, $J_{NH}$ 4.4, $CH_2$), 3.97, 2.95 (6H, ds, $CH_3$) and 1.43 (9H, s, $CH_3$); $\delta_C$ (50.3 MHz; $C_2HCl_3$) 168.73 (CO), 156.33 ($CO_2NH$), 80.04 ($\underline{C}(CH_3)_3$), 42.75 ($CH_2$), 36.31, 36.06 ($NCH_3$) and 28.84 ($C(\underline{C}H_3)_3$); *m/z* (EI) 202 (1%, $[M]^+$), 146 (23, $[M + H - C(CH_3)_3]^+$), 129 (14, $[M - OC(CH_3)_3]^+$), 102 (3, $[M + H - CO_2C(CH_3)_3]^+$), 72 (83, $[M - CH_2NHCO_2C(CH_3)_3]^+$) and 28 (100, $[CNH_2]^+$).

### *3.0.2.30 N-Tertiary butoxycarbonyl-, -N'-ethyl, N'-methyl glycinamide (58)*

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl-alaninate (**37**) using tertiary-butoxycarbonyl glycine (**56**) (2.00 g, 11.4 mmol) and ethylmethylamine (1.04 $cm^3$, 11.4 mmol.). The work up was modified as follows; the mixture was extracted into ethyl acetate and the solvent removed under reduced pressure to give a white powder which was recrystallised from methanol to give white needle shaped crystals (1.66 g, 67%), m.p. 86°C; (Found C, 55.75; H, 9.6; N, 13.05. $C_{10}H_{20}N_2O_3$ requires C, 55.5; H, 9.35; N, 12.95%); *m/z* (Found: 217.1548 $[M + H]^+$, $C_{10}H_{21}N_2O_3$ requires 217.1552); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3274 (NH), 2968, 2938 (2 x aliphatic CH), 1718 ($CO_2NH$) and 1635 (CONH); $\delta_H$ (200 MHz; $C_2HCl_3$) 5.54 (1H, s, NH), 3.94, 3.90 (2H, dd, $J_{NH_a}$ 4.0, $J_{NH_b}$ 14.8, $CH_2$), 3.42, 3.25 (2H, dq, $J_{CH_3}$ 7.2, $C\underline{H}_2CH_3$), 2.92, 2.90 (3H, ds, $NCH_3$), 1.42 (9H, s, $CH_3$) and 1.12 (3H, dt, $J_{CH_2}$ 7.2 $CH_2C\underline{H}_3$); $\delta_C$ (50.3 MHz; $C_2HCl_3$) 168.15 (CO), 156.29 ($CO_2NH$), 79.99 ($\underline{C}(CH_3)_3$), 43.75, 43.23, 42.90, 42.47 ($CH_2$), 33.78, 33.30 ($NCH_3$), 28.84 ($C(\underline{C}H_3)_3$) and 13.75, 12.76 ($CH_2\underline{C}H_3$).

### 3.0.2.31 N-Tertiary butoxycarbonyl-N',O-dimethylglycyl-hydroxylamide (59)

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl-alaninate (**37**) using tertiary butoxycarbonyl glycine (**56**) (1.00 g, 5.7 mmol) and N', O-dimethyl glycine hydroxylamine. The work up was modified as follows; the mixture was extracted into ethyl acetate and the solvent was removed under reduced pressure to give a white solid which was recrystallised from methanol to give white crystals (1.00 g, 81%), m.p. 101°C; (Found: C, 49.75; H, 8.65; N, 12.9. $C_9H_{18}N_2O_3$ requires C, 49.5; H, 8.3; N, 12.85%; *m/z* (Found: 219.1339 $[M + H]^+$, $C_9H_{19}N_2O_4$ requires 219.1345); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3288 (NH), 3042, 3006, 2988, 2976, 2963, 2918 (6 x aliphatic CH), 1717 ($CO_2NH$) and 1696, 1658 (2 x CONH); $\delta_H$ (200 MHz; $C_2H_3O_2H$) 4.00 (2H, s, $CH_2$), 3.77 (3H, s, $OCH_3$), 3.20 (3H, s, $CH_3$) and 1.46 (9H, s, $CH_3$); $\delta_C$ (75.5 MHz; $C_2H_3O_2H$) 174.50 (CONH), 158.90 ($CO_2NH$), 80.85 ($\underline{C}(CH_3)_3$), 62.16 ($OCH_3$), 42.55 ($CH_2$), 32.97 ($NCH_3$) and 28.99 ($C(\underline{C}H_3)_3$); *m/z* (CI) 219 (20%, $[M + H]^+$), 163 (100, $[M + 2H - C(CH_3)_3]^+$), 157 (10, $[M - HNCH_3(OCH_3)]^+$), 145 (10, $[M - OC(CH_3)_3]^+$) and 119 (53, $[M + 2H - CO_2C(CH_3)_3]^+$).

### 3.0.2.32 Glycyl-N',N',ethylmethylamide hydrochloride (60)

This compound was prepared in a manner identical to that described for the glycyl-N,N,dimethylamide hydrochloride (**55**) using tertiary butoxycarbonyl-glycyl-*N'*,*N'*-ethylmethylamide (**58**) (1.00 g, 4.63 mmol) to give a yellow oil (781 mg, 99%), *m/z* (Found: $[M + H - Cl]^+$, 117.1028. $C_5H_{13}N_2O$ requires 117.1023); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3480, 3430 (2 x NH) and 1730, 1650 (2 x CONH); $\delta_H$ (300 MHz, $C^2H_3O^2H$) 4.04, 4.00 (2H, s, $CH_2CO$), 3.50, 3.40 (2H, q, $J_{CH_3}$ 7.7, $CH_2CN$), 3.06, 3.01 (3H, s, $NCH_3$) and 1.27, 1.17 (3H, t, $J_{CH_2}$

7.2, $NCH_2C\underline{H}_3$); $\delta_C$ (73.5 MHz, $C^2H_3O^2H$) 166.60, 166.50, (2 x CO), 44.58, 43.92 (2 x $N\underline{C}H_2CO$), 41.27, 40.98 (2 x $N\underline{C}H_2CH_3$), 34.09, 33.17 (2 x $NCH_3$) and 13.34, 12.40 (2 x $CH_2\underline{C}H_3$); *m/z* (CI) 117 (100%, $[M - HCl]^+$), 99 (20, $[M - Cl - NH_3]^+$) and 70 (8, $[M + H - Cl - NH_3 - 2CH_3]^+$).

### 3.0.2.33 Glycyl-N',O'-dimethylhydroxylamide hydrochloride (61)

This was prepared in a manner identical to that described for the glycyl-N',N'-dimethylamide hydrochloride (**55**) using tertiary butoxycarbonyl-glycyl N',O'-dimethylhydroxylamide (**59**) (0.52 g, 2.38 mmol) to give a clear oil (0.34 g, 93%), *m/z* (Found: $[M - Cl]^+$, 119.0825. $C_4H_{13}N_2O_2$ requires 119.0821); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3440 (NH) and 1670 (CONH); $\delta_H$ (300 MHz, $^2H_3CO^2H$) 7.80 (3H, s, $NH_3^+$), 3.56 (10.5H, s, $CH_2$), 3.49 (3H, s, $OCH_3$), 3.32 (15.75H, s, $OCH_3$), 2.87 (2H, s, $CH_2$), 2.77 (15.75H, s, $NCH_3$) and 2.51 (3H, s, $NCH_3$) (*cis* and *trans* isomers); $\delta_C$ (73.5 MHz, $^2H_3CO^2H$) 169.47, 167.84 (2 x CO), 62.19, 62.00 (2 x $OCH_3$), 40.83, 40.71 (2 x $CH_2$) and 35.53, 32.48 (2 x $NCH_3$); *m/z* (CI) 119 (100, $[M - Cl]^+$), 89 (6, $[M - Cl - 2CH_3]^+$), 72 (7, $[M - Cl - CH_3 - NH_3]^+$) and 62 (15, $[M - Cl - NH_3CH_2 - CO]^+$).

### 3.0.2.34 Methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (65)

To a solution of carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycine (**27**) (0.30 g, 0.762 mmol) in dry THF (25ml). i-$Pr_2EtN$ (0.265 $cm^3$, 1.52 mmol), methyl glycinate hydrochloride (**54**) (96 mg, 0.762 mmol) and TBTU (0.245 g, 0.76 mmol) were added and stirred overnight.[111] The solvent was removed under reduced pressure and the oil was partitioned between ethyl acetate (40 $cm^3$) and water (40 $cm^3$), washed with water (2 ×50 $cm^3$), citric

acid (10%, 30 $cm^3$), $NaHCO_3$ (saturated, 30 $cm^3$), brine (saturated, 20 $cm^3$) and dried ($MgSO_4$). The solvent was removed under reduced pressure and the last traces of ethyl acetate were azeotropically removed with methylene chloride to give a white solid which was recrystalised from methanol to give white crystals (219 mg, 62%), m.p. 145-146°C; (Found: C, 57.05; H, 6.9; N, 12.1. $C_{22}H_{32}N_4O_7$ requires C, 56.85; H, 6.95; N, 12.05%); *m/z* (Found: $[M + H]^+$, 465.2350. $C_{22}H_{33}N_4O_7$ requires 465.2350); $[\alpha]_D$ -45.5 (*c* 0.17 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3300 (NH), 3065 (aromatic CH), 2955, 2875 (2 x aliphatic CH), 1750, 1685 ($CO_2NH$ and $CO_2R$), and 1665, 1655, 1635 (3 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.38 (5H, m, $C_6H_5$), 5.12 (2H, s, $C_6H_5C\underline{H}_2$), 4.34 (1H, q, $J_{CH_3}$ 7.22, $C\underline{H}CH_3$), 4.19 (1H, t, $J_{CH_2}$ 7.52, $NHC\underline{H}CH_2$), 4.00 (2H, s, $NHC\underline{H}_2$), 3.94 (2H, q, $J_{CH}$ 16.98, *J* 49.77, $NHC\underline{H}_2$), 3.73 (3H, s, $OCH_3$), 1.74 (1H, sep, $C\underline{H}(CH_3)_2$), 1.60 (2H, t, $J_{CH}$ 6.81, $J_{CH}$ 7.56, $CHC\underline{H}_2CH$), 1.42 (3H, d, $J_{CH}$ 7.14, $CHC\underline{H}_3$) and 0.98 (6H, m, $CH(C\underline{H}_3)_2$); $\delta_C$ (150 MHz, $C^2H_3O^2H$) 175.77, 175.28, 172.09 (3 x CO), 158.76 ($CO_2NH$), 138.10 (*i*-$C_6H_5$), 129.43 (*m*-$C_6H_5$), 128.99 (*p*-$C_6H_5$), 128.82 (*o*-$C_6H_5$), 67.82 ($\underline{C}H_2C_6H_5$), 55.01 ($\underline{C}HCH_2$), 52.63, 52.60 ($OCH_3$), 50.87, 50.85 ($\underline{C}HCH_3$), 43.35, 43.32 ($\underline{C}H_2CO$), 41.81 ($CHC\underline{H}_2$), 25.85 ($\underline{C}H(CH_3)_2$), 23.52, 23.49, 23.46, 21.87, 21.82 ($CH(\underline{C}H_3)_2$) and 17.40 ($CH\underline{C}H_3$); *m/z* (CI) 482 (47%, $[M + H + NH_3]$), 465 (100, $[M + H]^+$), 374 (22, $[M + H - CH_2C_6H_5]^+$) 357 (27, $[M - OCH_2C_6H_5]^+$) and 331 (18, $[M + 2H - CO_2CH_2C_6H_5]^+$).

### *3.0.2.35 Ethyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (66)*

This compound compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) ethyl glycinate hydrochloride (**40**) to give a white solid which was recrystallised from methanol to give

white crystals (228 mg, 63%), m.p. 142-144°C; (Found: C, 55.58; H, 7.11; N, 11.37. $C_{23}H_{34}N_4O_7$ requires C, 55.62; H, 7.31; N, 11.29%); *m/z* (Found: $[M + H]^+$, 479.2506. $C_{23}H_{35}N_4O_7$ requires 479.2506); $[\alpha]_D$ -30.8 (*c* 0.13 in MeOH); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3400 (NH), 3300 (aromatic CH), 3330 (aliphatic CH) and 1740 ($CO_2CH_3$), 1700 ($CO_2NH$) and 1680, 1660, 1630 (3 x CONH); $\delta_H$ (400 MHz; $C^2HCl_3$) 7.33 (1H, t, $J_{CH_2}$ 7.16, N<u>H</u>$CH_2CO$), 7.32 (5H, m, $C_6H_5$), 7.19 - 7.15 (2H, m, N<u>H</u>$CHCH_3$ and N<u>H</u>$CH_2CO_2$), 5.62 (1H, d, $J_{CH}$ 7.48, $CO_2NH$<u>C</u>H), 5.08 (2H, s, $C_6H_5C\underline{H}_2$), 4.53 (1H, m, $J_{NH}$ 7.00, C<u>H</u>$CH_3$), 4.30 (1H, m, $J_{NH}$ 7.48, NHC<u>H</u>$CH_2$), 4.15 (2H, q, $J_{CH_2}$ 7.16, $OCH_2$), 4.10 (2H, d, $J_{NH}$ 7.20, NHC<u>H</u>$_2CO_2CH$), 4.00 (2H, d, $J_{NH}$ 5.36, NHC<u>H</u>$_2$), 3.73 (3H, s, $OCH_3$), 1.74 (1H, sep, C<u>H</u>$(CH_3)_2$), 1.60 (2H, t, $J_{CH}$ 6.81, $J_{CH}$ 7.56, NHC<u>H</u>$_2CO$), 1.58 (3H, m, C<u>H</u>$_2CH(C\underline{H}_3)_2$) 1.23 (3H, t, $J_{CH_2}$ 7.16, $OCH_2C\underline{H}_3$), 1.36 (3H, d, $J_{CH}$ 7.04, $CHCH_3$) and 0.91 (6H, m, $CH(CH_3)_2$); $\delta_C$ (100 MHz; $C^2HCl_3$) 172.55, 172.37, 169.54, 168.95 (NHCO), 156.35 ($CO_2NH$), 135.95 (*i*-$C_6H_5$), 128.38 (*m*-$C_6H_5$), 128.10 (*p*-$C_6H_5$), 127.85 (*o*-$C_6H_5$), 67.03 ($C_6H_5C\underline{H}_2$), 61.31 (<u>C</u>$H_2CH_3$), 53.59 (<u>C</u>$HCH_2$), 49.08 (<u>C</u>$HCH_3$), 42.79 ($C\underline{H}_2CO_2CH_2$), 41.38 (NH<u>C</u>$H_2CO$), 41.12 (CH<u>C</u>$H_2CH$), 24.54 (<u>C</u>$H(CH_3)_2$), 22.78, 21.72 ($CH(\underline{C}H_3)_2$), 18.01 (CH<u>C</u>$H_3$) and 13.93 ($CH_2\underline{C}H_3$); *m/z* (CI) 496 (4%, $[M + H + NH_3]$), 479 (47, $[M + H]^+$), 388 (17, $[M + H - CH_2C_6H_5]^+$), 371 (100, $[M - C_6H_5CH_2O]^+$) and 345 (16, $[M + 2H - CO_2CH_2C_6H_5]^+$).

### *3.0.2.36 t-Butyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (67)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycine (**27**) (0.150 g, 0.381 mmol) and tertiary-butyl glycinate acetate (73 mg, 0.381 mmol) to give a white solid (228 mg, 72%), *m/z* (Found: $[M + H]^+$, 507.2808.

$C_{25}H_{39}N_4O_7$ requires 507.2819); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.38 (5H, m, $C_6H_5$), 5.14 (1H, d, $J_{CH_b}$ 12.46, $C_6H_5C\underline{H}_aH_b$), 5.11 (1H, d, $J_{CH_a}$ 7.39, $C_6H_5CH_a\underline{H}_b$), 4.36 (1H, q, $J_{CH_3}$ 6.64, $C\underline{H}CH_3$), 4.21 (1H, t, $J_{CH_2}$ 7.45, $NHC\underline{H}CH_2$), 4.01 (1H, d, $J_{CH_b}$ 16.74, $NHC\underline{H}_aH_bCO$), 3.88 (3H, m, $NHCH_a\underline{H}_bCO$ and $C\underline{H}_2CO_2$), 1.74 (1H, sep, $J$ 6.43, $C\underline{H}(CH_3)_2$), 1.61 (2H, t, $J_{CH}$ 6.57, $CHC\underline{H}_2$), 1.50, 1.49 (9H, ds, $C(CH_3)_3$), 1.42 (3H, d, $J_{CH}$ 7.00, $CHC\underline{H}_3$), 0.98 (3H, d, $J_{CH}$ 5.56, $CH(C\underline{H}_3)_2$) and 0.97 (3H, d, $J_{CH}$ 5.94, $CH(CH_3)_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 175.67, 175.21, 171.94, 170.24 (4 x CO), 158.71 ($CO_2NH$), 138.08 (*i*-$C_6H_5$), 129.44 (*m*-$C_6H_5$), 129.00 (*p*-$C_6H_5$), 128.82 (*o*-$C_6H_5$), 82.82 ($\underline{C}(CH_3)_3$), 67.79 ($C_6H_5\underline{C}H_2$), 54.98 ($\underline{C}HCH_2$), 50.76 ($\underline{C}HCH_3$), 43.32 ($\underline{C}H_2CO_2CH_2$), 42.84 ($\underline{C}H_2CO$), 42.72 ($CH\underline{C}H_2CH$), 28.28 ($C(\underline{C}H_3)_3$), 25.80 ($\underline{C}H(CH_3)_2$), 23.49, 21.87 ($CH(\underline{C}H_3)_2$) and 17.48 ($CH\underline{C}H_3$); *m/z* (CI) 507 (4, $[M + H]^+$), 451 (30, $[M + 2H - C(CH_3)_3]^+$) and 136 (100, $[C_6H_5CH_2OCHO]^+$).

### *3.0.2.39 Methyl carbobenzyloxy-glycyl glycinate (68)*[115]

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl-alaninate (**29**) using carbobenzyloxy-glycine (**34**) (3.51 g, 16.8 mmol) and methyl glycinate hydrochloride (**54**) to give a white solid (2.86 g, 61%), m.p. 63°C (lit.,[115] 63-65°C), *m/z* (Found: $[M + H]^+$, 281.1137. Calc. for $C_{13}H_{17}N_2O_5$ 281.1137); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3350, 3330 (2 x NH), 3050 (aromatic CH), 1730 ($CO_2Me$) and 1710 ($CO_2NH$); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.33 (5H, m, $C_6H_5$), 6.66 (1H, t, NHCO), 5.60 (1H, t, $NHO_2C$), 5.13 (2H, s, $C_6H_5C\underline{H}_2$), 4.04 (2H, d, $J_{NH}$ 5.12, $C\underline{H}_2CO_2Me$), ), 3.92 (2H, d, $J_{NH}$ 5.78, $CONHC\underline{H}_2$), 3.75 (3H, s, $OCH_3$); $\delta_C$ (74.8 MHz; $C^2HCl_3$) 170.08, 169.46 (2 x CO), 156.56 ($CO_2NH$), 135.95 (*i*-$C_6H_5$), 128.37 (*m*-$C_6H_5$), 128.06 (*p*-$C_6H_5$), 127.88 (*o*-$C_6H_5$), 67.03 ($C_6H_5\underline{C}H_2$), 52.22 ($OCH_3$), 44.17 ($\underline{C}H_2CO_2Me$), 40.89 ($NH\underline{C}H_2CO$); *m/z* (CI) 281

(100%, $[M + H]^+$), 237 (50, $[M + H - CO_2]^+$), 209 (5, $[M + H - CO_2 - CO]^+$), 190 (20, $[M + H - C_6H_5CH_2]^+$), 173 (3, $[M - C_6H_5CH_2O]^+$) and 147 (31, $[M + 2H - C_6H_5CH_2O_2C]^+$).

### 3.0.2.38 Ethyl carbobenzyloxy-glycyl glycinate (69)[117]

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl-alaninate (**29**) using carbobenzyloxy-glycine (**34**) (3.51 g, 16.8 mmol) and ethyl glycinate hydrochloride (**40**) to give a white solid (2.93 g, 59%), m.p. 78°C (lit.,[117] 78-80°C), (Found: C, 56.95; H, 6.0; N, 9.4. Calc. for $C_{14}H_{18}N_2O_5$ C, 57.1; H, 6.15; N, 9.5%); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3300, 3250 (NH), 3080 (aromatic CH), 1750, 1740 (NHCO and $CO_2Et$), and 1690 (CONH); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.34 (5H, m, $C_6H_5$), 6.88 (1H, t, NHCO), 5.81 (1H, t, $NHO_2C$), 5.12 (2H, s, $C_6H_5C\underline{H}_2$), 4.18 (2H, q, $J_{CH_3}$ 7.11, $OCH_2$), 4.00 (2H, d, $J_{NH}$ 5.42, $C\underline{H}_2COEt$), ), 3.91 (2H, d, $J_{NH}$ 5.58, $CO_2NHC\underline{H}_2$) and 1.26 (3H, t, $J_{CH_2}$ 7.14, $CH_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 170.31, 170.04 (CONH and $CO_2Me$), 157.21 ($CO_2NH$), 136.63 (*i*-$C_6H_5$), 129.04 (*m*-$C_6H_5$), 128.72 (*p*-$C_6H_5$), 128.57 (*o*-$C_6H_5$), 67.69 ($C_6H_5\underline{C}H_2$), 62.10 ($OCH_2$), 44.85 ($\underline{C}H_2CO_2CH_2$), 41.72 ($O_2CNH\underline{C}H_2$) and 14.60 ($CH_3$); *m/z* (CI) 295 (29%, $[M + H]^+$), 269 (100, $[M - C_2H_2]^+$), 249 (17, $[M - OEt]^+$) and 114 (49, $[M + H - OEt - C_6H_5CH_2O_2C]^+$).

### 3.0.2.39 Methyl glycyl glycinate (70)

To a stirred solution of methyl carbobenzyloxy-glycyl glycinate (**68**) (1.77 g, 6.3 mmol) in methanol (20 ml) was added wet carbon/palladium catalyst (5%, 185 mg, 10% Pd by mass). The flask was sealed, flushed with argon, then hydrogen and the mixture was stirred under an atmosphere of hydrogen with vigorous stirring. The reaction was followed to completion by silica chromatography eluting with 5% methanol / 95% methylene chloride (4½ h). The

reaction mixture was filtered through celite and the methanol removed under reduced pressure to give a white solid (0.74 g, 81%), m.p. 198°C (sublimes), 209 (decomp.); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3300, 3170 (NH), 3050 (aromatic CH), 1710, 1700, 1680 (2 x CONH and $CO_2Me$); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 4.05 (2H, s, $C\underline{H}_2CO_2Me$), 3.47 (2H, s, $NHC\underline{H}_2CO$) and 3.44 (3H, s, $OCH_3$); *m/z* (EI) 146 (25, $[M]^+$), 115 (4, $[M - OMe]^+$) and 86 (6, $[M - CO - MeOH]^+$).

## *3.0.2.40 Ethyl glycyl glycinate (71)*

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-glycyl glycinate (**70**) using ethyl carbobenzyloxy-glycyl glycinate (**69**) (1.86 g, 6.3 mmol) to give a white solid (0.82 g, 81%), m.p. 80°C (sublimes at 77°C); *m/z* (Found: $[M]^+$, 160.0848 $C_6H_{12}N_2O_3$ requires 160.0848); $\upsilon_{max}$ (Nujol)/$cm^{-1}$ 3400, 3310 (2 x NH), 1740 ($CO_2Et$) and 1650 (NHCO), $\delta_H$ (200 MHz; $^2H_3CCN$) 4.15 (2H, q, $J_{CH_3}$ 7.12, $OCH_2$), 3.92 (2H, d, $J_{NH}$ 6.02, $C\underline{H}_2COEt$), ), 3.23 (2H, s, $NH_2C\underline{H}_2$) and 1.24 (3H, t, $J_{CH_2}$ 7.13, $CH_3$); $\delta_C$ (100 MHz; $^2H_3CCN$) 174.47, 171.07 (2 x CO), 61.79 ($OCH_2$), 45.25 ($\underline{C}H_2CO_2CH_2$), 41.45 ($O_2CNH\underline{C}H_2$) and 14.44 ($CH_3$); *m/z* (CI) 161 (100%, $[M + H]^+$), 146 (2, $[M + H - Me]^+$), 132 (2, $[M + H - OEt]^+$), 115 (2, $[M + H - Et - NH_3]^+$) and 104 (5, $[M + 2H - NH_2CH_2CO]^+$).

### 3.0.2.41 Benzyl tertiary-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycyl glycinate (72)

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl-glycinate (**65**) using tertiary-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycine (**46**) (0.200 g, 0.557 mmol) and benzyl glycinate tosylate (**44**) (197 mg, 0.557 mmol) to give a white solid (189 mg, 65%), *m/z* (Found: $[M + H]^+$, 507.2826. $C_{25}H_{39}N_4O_7$ requires 507.2819); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.38 (5H, m, $C_6H_5$), 5.20 (2H, s, $C_6H_5C\underline{H}_2$), 4.36 (1H, q, $J_{CH_3}$ 7.06, $C\underline{H}CH_3$), 4.12 (1H, t, $J_{CH_2}$ 7.32, $NHC\underline{H}CH_2$), 4.44 (2H, s, $C\underline{H}_2CO_2$), 4.02 (1H, d, $J_{CH_b}$ 22.56, $NHC\underline{H}_aH_bCO$), 3.88 (1H, d, $J_{CH_a}$ 16.96, $NHCH_a\underline{H}_bCO$), 1.72 (1H, sep, *J* 6.51, $C\underline{H}(CH_3)_2$), 1.56 (2H, t, $J_{CH}$ 6.80, $CHC\underline{H}_2$), 1.47 (9H, ds, $C(CH_3)_3$), 1.42 (3H, q, $J_{CH}$ 7.14, $CHCH_3$), 0.98 (3H, d, $J_{CH}$ 6.49, $CH(CH_3)_2$) and 0.97 (3H, d, $J_{CH}$ 9.32, $CH(CH_3)_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 176.02, 175.94, 175.26, 127.01 (4 x CO), 158.09 ($CO_2NH$), 137.10 (*i*-$C_6H_5$), 129.52 (*m*-$C_6H_5$), 129.24 (*p*-$C_6H_5$), 129.15 (*o*-$C_6H_5$), 80.68 ($\underline{C}(CH_3)_3$), 67.82 ($C_6H_5\underline{C}H_2$), 54.49 ($\underline{C}HCH_2$), 50.84 ($\underline{C}HCH_3$), 43.34 ($C\underline{H}_2CO_2CH_2$), 42.02 ($\underline{C}H_2CO$), 41.87 ($CH\underline{C}H_2CH$), 28.71 ($C(\underline{C}H_3)_3$), 25.81 ($\underline{C}H(CH_3)_2$), 23.48, 21.89 ($CH(\underline{C}H_3)_2$) and 17.59 ($CH\underline{C}H_3$); *m/z* (CI) 507 (5%, [M + H]) 451 (17, $[M + 2H - C(CH_3)_3]^+$), 433 (38, $[M - OC(CH_3)_3]^+$) and 407 (100, $[M + 2H - CO_2C(CH_3)_3]^+$).

### 3.0.2.42 Tertiary butyl tertiary-butoxycarbonyl-(2S)-leucyl-(2S)-alanyl-glycyl glycinate (73)

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl-(2S)-alanyl-glycyl-glycinate (**65**) using tertiary butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycine (**46**) (400 mg, 1.11 mmol) and tertiary butyl glycinate acetate

(212 mg, 1.11 mmol). Recrystallisation from methanol / ether gave white crystals (197 mg, 75%), m.p. 153°C; (Found: C, 55.7; H, 8.55; N, 11.7. $C_{22}H_{40}N_4O_7$ requires C, 55.9; H, 8.55; N, 11.85%); *m/z* (Found: $[M + H]^+$, 473.2975. $C_{22}H_{41}N_4O_7$ requires 473.2975); $[\alpha]_D$ -27.3 (*c* 0.15 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3400, 3305, 3086 (3 x NH), 2980, 2925, 2860 (3 x aliphatic CH), 1735, 1710 ($CO_2NH$ and $CO_2t$-Bu), and 1690, 1655, 1635 (3 x CONH); $\delta_H$ (300 MHz; $C^2HCl_3$) 7.70 (2H, m, $N\underline{H}CH_2$ and $N\underline{H}CH_3$), 7.55 (1H, t, $N\underline{H}CH_2$), 6.23 (1H, d, $J_{CH}$ 7.00, $CO_2NH$), 4.37 (1H, sex, $J_{CH_2}$ 6.09, $NHC\underline{H}CH_2$), 4.14 (1H, q, *J* 7.16, $C\underline{H}CH_3$), 3.88 (4H, qd, $J_{CH}$ 13.04, $J_{CH}$ 3.99, $NHC\underline{H}_2$), 1.74 (1H, sep, $C\underline{H}(CH_3)_2$), 1.60 (2H, m, $CHC\underline{H}_2CH$), 1.44 (9H, s, $NCO_2C(CH_3)_3$), 1.41 (9H, s, $CO_2C(CH_3)_3$), 1.35 (3H, d, $J_{CH}$ 7.08, $CHC\underline{H}_3$) and 0.918 (6H, q, *J* 5.30, $CH(C\underline{H}_3)_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 174.45, 173.88, 170.29, 169.88 (4 x NHCO), 157.19 ($CO_2NH$), 81.04, 79.89 ($CO_2C(\underline{C}H_3)_3$ and $NCO_2\underline{C}(CH_3)_3$), 54.42 ($\underline{C}HCH_2$), 50.92 ($\underline{C}HCH_3$), 43.62, 42.84, 42.18, ($C\underline{H}_2CONHC\underline{H}_2$, $CH\underline{C}H_2CH$), 28.84 ($NHCO_2C(\underline{C}H_3)_3$), 28.57 ($CO_2C(\underline{C}H_3)_3$), 25.76 ($\underline{C}H(CH_3)_2$), 23.88, 22.32 ($CH(\underline{C}H_3)_2$) and 18.37 ($CH\underline{C}H_3$); *m/z* (CI) 490 (20%, $[M + H + NH_3]$), 473 (50, $[M + H]^+$), 434 (40, $[M + H + NH_3 - C(CH_3)_3]^+$), 417 (100, $[M + H - C(CH_3)_3]^+$), 361 (34, $[M + 2H - 2C(CH_3)_3]^+$) and 317 (50, $[M + 2H - 2C(CH_3)_3 - CO_2]^+$).

### *3.0.2.43 Ethyl (2S)-leucyl (2S)-alanyl glycyl glycine (74)*

This compound was prepared in a manner identical to that described for the the methyl carbobenzyloxy-glycyl glycinate (**70**) using ethyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycine (**66**) (400 mg, 0.762 mmol) to give a yellow oil (287 mg, 100%), *m/z* (Found: $[M + H]^+$, 345.2141. $C_{15}H_{28}N_4O_5$ requires 345.2138); $\delta_H$ (200 MHz; $^2H_3CO^2H$) 4.31 (1H, q, $J_{CH_3}$ 7.13, $C\underline{H}CH_3$), 4.18 (2H, q, $J_{CH_3}$ 7.11, $OCH_2$), 4.00 (1H, d, $J_{H_b}$ 12.45, $NHC\underline{H}_aH_b$), 3.94 (2H, s, $C\underline{H}_2CO_2Et$), 3.84 (1H, d, $J_{H_a}$ 17.05, $NHCH_a\underline{H}_b$), 3.45 (1H, t, $J_{CH_2}$ 6.70, $NHC\underline{H}CH_2$),

1.71 (1H, sep, *J* 6.59, C<u>H</u>(CH$_3$)$_2$), 1.53 (2H, t, $J_{CH}$ 7.08, CHC<u>H</u>$_2$CH), 1.40 (3H, d, $J_{CH}$ 6.15, CHC<u>H</u>$_3$) 1.29 (3H, t, 7.57 CH$_2$C<u>H</u>$_3$), 0.97 (3H, d, *J* 6.15, CH(C<u>H</u>$_3$)$_2$) and 0.94 (3H, d, *J* 5.78, CH(C<u>H</u>$_3$)$_2$); $\delta_C$ (50.3 MHz, $C^2H_3O^2H$) 177.99, 175.73, 172.39, 171.49 (4 x CO), 62.62 (OCH$_2$), 54.49 (NH<u>C</u>HCH$_2$), 51.15 (<u>C</u>HCH$_3$), 45.22, 43.69 (<u>C</u>H$_2$CO), 42.27 (CHC<u>H</u>$_2$), 25.96 (<u>C</u>H(CH$_3$)$_2$), 23.80, 22.67, (CH(<u>C</u>H$_3$)$_3$), 17.89 (CH<u>C</u>H$_3$) and 14.30 (CH$_2$<u>C</u>H$_3$); *m/z* (CI) 345 (5%, $[M + H]^+$), 288 (39, $[M + H - CH_2CH(Me)_2]^+$) and 231 (100, $[M + H - COCH(NH_2)CH_2CH(Me)_2]^+$).

## *3.0.2.44 N-(Carbobenzyloxy-(2S)-leucyl-(2S)-alanyl-glycyl)-ethanolamine (75)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using ethanolamine hydrochloride (99 mg, 1.02 mmol) to give a clear yellow oil (169 mg, 33%), $[\alpha]_D$ -32.3 (*c* 0.28 in MeOH); $\upsilon_{max}$ (thin film)/cm$^{-1}$ 3300 (OH), 3100 (aromatic CH), 2800 (aliphatic CH), 1700 ($CO_2NH$) and 1680, 1630, 1620 (4 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$), 7.37 (5H, m, C$_6$H$_5$), 5.12 (2H, s, C$_6$H$_5$C<u>H</u>$_2$), 4.29 (1H, q, *J* 6.29, C<u>H</u>CH$_3$), 4.20 (1H, t, NHC<u>H</u>CH$_2$), 3.96 (1H, d, $J_{CH}$ 16.83, NHC<u>H</u>$_a$H$_b$), 3.78 (1H, d, $J_{CH}$ 16.79, NHCH$_a$<u>H</u>$_b$), 3.64 (2H, t, $J_{CH_2}$ 5.19, OC<u>H</u>$_2$), 3.36 (2H, t, $J_{CH_2}$ 3.24, NC<u>H</u>$_2$CH$_2$), 1.74 (1H, sep, *J* 6.07, C<u>H</u>(CH$_3$)$_2$), 1.61 (2H, d, *J* 6.21, CHC<u>H</u>$_2$CH), 1.41 (3H, d, $J_{CH}$ 6.37, CHC<u>H</u>$_3$), 0.99 (3H, d, $J_{CH}$ 5.79, CH(C<u>H</u>$_3$)$_2$ and 0.97 (3H, d, $J_{CH}$ 6.12, CH(C<u>H</u>$_3$)$_2$); $\delta_C$ (75.5 MHz, $^2H_3CO^2H$) 175.88, 175.33, 171.70 (3 x CO), 158.69 ($CO_2NH$), 138.11 (*i*-C$_6$H$_5$), 129.47 (*m*-C$_6$H$_5$), 129.08 (*p*-C$_6$H$_5$), 128.80 (*o*-C$_6$H$_5$), 67.77 (<u>C</u>H$_2$C$_6$H$_5$), 61.44 (CH$_2$O), 54.98 (<u>C</u>HCH$_2$), 51.07 (<u>C</u>HCH$_3$), 43.56 (NHC<u>H</u>$_2$), 43.21 (NHC<u>H</u>$_2$CH$_2$), 43.10 (CH<u>C</u>H$_2$), 25.79 (<u>C</u>H(CH$_3$)$_2$), 23.44, 21.90 (CH(<u>C</u>H$_3$)$_2$) and 17.26 (CH<u>C</u>H$_3$).

### *3.0.2.45 Sodium carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (76)*

This compound was prepared in a manner identical to that described for the carbobenzyloxy-(2S)-leucyl (2S)-alanine (**29**) using methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) (1.32g, 2.83 mmol) without acidification to give a white solid (1.32 g, 98%), m.p. 140°C; $\upsilon_{max}$ (Nujol)/cm$^{-1}$ 3400 (NH), 2550 (aliphatic CH), 1730 ($CO_2H$), 1705 ($CO_2NH$), 1650, 1670, 1680 (3 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.38 (5H, m, $C_6H_5$), 5.16 (1H, d, $J_{H_b}$ 12.51, $C_6H_5C\underline{H}_aH_b$), 5.16 (1H, d, $J_{H_a}$ 12.42, $C_6H_5CH_a\underline{H}_b$), 4.39 (1H, q, $J_{CH_3}$ 7.14, $C\underline{H}CH_3$), 4.20 (1H, t, $J_{CH_2}$ 7.05, $NHC\underline{H}CH_2$), 3.99 (1H, d, $J_{H_b}$ 16.77, $NHC\underline{H}_aH_b$), 3.86 (1H, d, $J_{H_a}$ 16.74, $NHCH_a\underline{H}_b$), 3.85 (1H, d, $J_{H_d}$ 17.3, $NHC\underline{H}_cH_d$), 3.74 (1H, d, $J_{H_c}$ 17.07, $NHCH_c\underline{H}_d$), 1.74 (1H, sep, $J$ 6.61, $C\underline{H}(CH_3)_2$), 1.61 (2H, t, $J_{CH}$ 7.17, 7.41, $CHC\underline{H}_2CH$), 1.43 (3H, d, $J_{CH}$ 7.14, $CHC\underline{H}_3$) and 0.98 (6H, t, $J$ 6.39, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz, $C^2H_3O^2H$) 176.53, 175.84, 175.58, 171.37 (4 x CO), 159.00 ($CO_2NH$), 138.43 (*i*-$C_6H_5$), 129.77 (*m*-$C_6H_5$), 129.30 (*p*-$C_6H_5$), 129.14 (*o*-$C_6H_5$), 68.09 ($\underline{C}H_2C_6H_5$), 55.28 ($\underline{C}HCH_2$), 51.01 ($\underline{C}HCH_3$), 44.67, 43.88, ($\underline{C}H_2CO$), 42.14 ($CHC\underline{H}_2$), 26.13 ($\underline{C}H(CH_3)_2$), 23.81, 22.12, ($CH(\underline{C}H_3)_3$) and 17.97 ($CH\underline{C}H_3$).

### *3.0.2.46 Carbobenzyloxy-(2S)-leucyl-(2S)-alanyl-glycyl-amino acetonitrile (78)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using amino acetonitrile hydrochloride (**85**) (94 mg, 1.02 mmol). Recrystallisation from MeOH / ether gave white crystals (370 mg, 78%), m.p. 159°C; (Found: C, 58.55; H, 6.5; N, 16.2. $C_{21}H_{29}N_5O_5$ requires C, 58.45; H, 6.8; N, 16.25%); *m/z* (Found: $[M + H + NH_3]^+$, 449.2510. $C_{21}H_{33}N_6O_5$ requires 449.2510); $[\alpha]_D$ -24.6 (*c* 0.13 in MeOH); $\upsilon_{max}$ (chloroform solution)/$cm^{-1}$ 3330 (NH), 3020 (aromatic CH), 2400 (CN) and 1735, 1701, 1685, 1676, 1642 ($CO_2NH$ and 4 x CONH); $\delta_H$ (400 MHz; $^2H_3COC^2H_3$) 7.97 (1H, t, N<u>H</u>$CH_2$), 7.88 (1H, d, $J_{CH}$ 4.32, N<u>H</u>$CH_3$), 7.73 (1H, t, N<u>H</u>$CH_2$), 7.37 (5H, m, $C_6H_5$), 6.65 (1H, d, $J_{CH}$ 7.44, $CO_2NH$), 5.10 (1H, d, $J_{H_b}$ 20.43, C<u>H</u>$_a$H$_b$C$_6$H$_5$), 5.07 (1H, d, $J_{H_a}$ 20.46, CH$_a$<u>H</u>$_b$C$_6$H$_5$), 4.27 (2H, m, C<u>H</u>$CH_3$ and NHC<u>H</u>$CH_2$), 4.17 (2H, d, $J_{NH}$ 5.76, NHC<u>H</u>$_2$), 3.95 (1H, dd, $J_{NH}$ 6.56, $J_{CH_b}$ 10.48, NHC<u>H</u>$_a$H$_b$), 3.79 (1H, dd, $J_{NH}$ 6.22, $J_{CH_a}$ 11.44, NHCH$_a$<u>H</u>$_b$), 1.75 (1H, sep, $J_{CH_3}$ 6.71, C<u>H</u>$(CH_3)_2$), 1.62 (2H, t, $J_{CH_2}$ 7.26, CHC<u>H</u>$_2$CH), 1.34 (3H, d, $J_{CH}$ 7.04, CHC<u>H</u>$_3$) and 0.92 (6H, m, CH(C<u>H</u>$_3$)$_2$); $\delta_C$ (100 MHz; $(^2H_3C)_2O$) 174.46, 173.37, 170.39 (3 x NHCO), 157.40 ($CO_2NH$), 138.01 (*i*-$C_6H_5$), 129.19 (*m*-$C_6H_5$), 128.65 (*p*-$C_6H_5$), 128.59 (*o*-$C_6H_5$), 117.52 (CN), 66.94 ($C_6H_5$<u>C</u>H$_2$), 54.53 (<u>C</u>H$CH_2$), 50.73 (<u>C</u>H$CH_3$), 43.21 (C<u>H</u>$_2$CONHC<u>H</u>$_2$), 41.76 (CH<u>C</u>H$_2$CH), 27.68 (<u>C</u>H$_2$CN), 25.31 (<u>C</u>H$(CH_3)_2$), 23.37, 21.89 (CH(<u>C</u>H$_3$)$_2$) and 17.26 (CH<u>C</u>H$_3$); *m/z* (CI) 449 (57%, [M + H + $NH_3$]), 341 (100, $[M + NH_3 - C_6H_5CH_2O]^+$), 324 (23, $[M - OCH_2C_6H_5]^+$) and 297 (30, $[M + H - CO_2CH_2C_6H_5]^+$).

### *3.0.2.47 Tertiary-butoxycarbonyl-(2S)-leucyl-(2S)-alanyl-glycyl amino acetonitrile (79)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using tertiary butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycine (**46**) (200 mg, 0.556 mmol) and aminoacetonitrile hydrochloride. Recrystallisation from MeOH / ether gave white crystals (125 mg, 56%), m.p. 162°C; (Found: C, 54.35; H, 7.7; N, 17.35. $C_{18}H_{31}N_5O_5$ requires C, 54.4; H, 7.85; N, 17.65%); *m/z* (Found: $[M + H]^+$, 398.2403. $C_{18}H_{32}N_5O_5$ requires 397.2325); $[\alpha]_D$ -29.3 (*c* 0.15 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3320 (NH), 3065 (aromatic CH), 2965, 2875 (2 x aliphatic CH), 2280 (CN), 1700 ($CO_2NH$) and 1685, 1675, 1650 (3 x CONH); $\delta_H$ (300 MHz; $^2H_3COC^2H_3$) 8.01 (1H, t, N<u>H</u>$CH_2$), 7.87 (1H, d, N<u>H</u>$CH_3$), 7.79 (1H, t, N<u>H</u>$CH_2$), 6.30 (1H, d, $J_{CH}$ 7.50, $CO_2NH$), 4.26 (1H, m, $J_{CH_2}$ 6.78, $J_{NH}$ 6.39, NHC<u>H</u>$CH_2$), 4.18 (2H, d, $J_{NH}$ 5.97, C<u>H</u>$_2$CN), 4.12 (1H, m, $J_{NH}$ 7.0, C<u>H</u>$CH_3$), 3.94 (1H, dd, $J_{NH}$ 6.41, $J_{CH_b}$ 10.74, NHC<u>H</u>$_a$H$_b$), 3.79 (1H, dd, $J_{NH}$ 5.79, $J_{CH_a}$ 11.28, NHCH$_a$<u>H</u>$_b$), 1.734 (1H, sep, $J_{CH_3}$ 6.65, C<u>H</u>$(CH_3)_2$), 1.60 (2H, dd, $J_{CH_2}$ 6.69, 7.25, CHC<u>H</u>$_2$CH), 1.40 (9H, s, $C(CH_3)_3$), 1.35 (3H, d, $J_{CH}$ 7.11, CHC<u>H</u>$_3$), 0.93 (3H, d, *J* 6.45, CH(C<u>H</u>$_3$)$_2$) and 0.91 (3H, d, *J* 6.12, CH(C<u>H</u>$_3$)$_2$); $\delta_C$ (75.3 MHz, $(^2H_3C)_2OC$) 175.28, 173.90, 170.84 (3 x NHCO), 157.34 ($CO_2NH$), 117.94 (CN), 80.04 (<u>C</u>$(CH_3)_3$), 54.53 (<u>C</u>H$CH_2$), 51.13 (<u>C</u>H$CH_3$), 43.62 (C<u>H</u>$_2$CONHC<u>H</u>$_2$), 42.04 (CH<u>C</u>H$_2$CH), 28.93 (C(<u>C</u>H$_3$)$_3$), 28.11 (<u>C</u>H$_2$CN), 25.73 (<u>C</u>H$(CH_3)_2$), 23.79, 22.35 (CH(<u>C</u>H$_3$)$_2$) and 17.77 (CH<u>C</u>H$_3$); *m/z* (CI) 415 (24%, $[M + H + NH_3]^+$), 398 (9, $[M + H]^+$), 359 (22, $[M + NH_3 + H - C(CH_3)_3]^+$), 341 (8, $[M + H - C(CH_3)_3]^+$), 297 (30, $[M + H - CO_2C(CH_3)_3]^+$) and 298 (100, $[M + 2H - CO_2C(CH_3)_3]^+$).

### *3.0.2.48 Carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2S)-alanyl aminoacetonitrile (80)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2S)-alanine (**88**) (200 mg, 0.49 mmol) and aminoacetonitrile hydrochloride (45 mg, 0.49 mmol) to give a white solid which was recrystallised from MeOH / ether to give white crystals (144 mg, 66%). m.p.161-162°C; (Found: $[M + H]^+$, 446.2410. $C_{22}H_{32}N_5O_5$ requires 446.2417); $[\alpha]_D$ -50.8 (*c* 0.13 in MeOH); $\upsilon_{max}$ (KBr disc)/$cm^{-1}$ 3290 (NH), 3060 (aromatic CH), 2930, 2910, 2880, 2850 (4 x aliphatic CH), 2260 (CN) and 1710, 1680, 1650, 1640 ($CO_2NH$ and 3 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.38 (5H, m, $C_6H_5$), 5.14 (2H, m, $C_6H_5C\underline{H}_2$), 4.35 (2H, q, $J_{CH_3}$ 7.11, $C\underline{H}CH_3$), 4.19 (2H, s, $CH_2CN$), 4. 19 (1H, m, $NHC\underline{H}CH_2$), 1.71 (1H, sep, *J* 6.74, $C\underline{H}(CH_3)_3$), 1.60 (2H, t, *J* 7.26, $CHC\underline{H}_2CH$), 1.41 (3H, d, $J_{CH}$ 5.16, $CHC\underline{H}_3$), 1.39 (3H, d, $J_{CH}$ 5.39, $CHC\underline{H}_3$), 1.00 (3H, d, $J_{CH}$ 6.88, $CH(C\underline{H}_3)_2$) and 0.97 (3H, d, $J_{CH}$ 7.04, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 175.78, 175.05, 174.68 (3 x $NH\underline{C}O$), 158.70 ($CO_2NH$), 138.13 (*i*-$C_6H_5$), 129.47 (*m*-$C_6H_5$), 129.00 (*p*-$C_6H_5$), 128.71 (*o*-$C_6H_5$), 117.42 (CN), 67.76 ($C_6H_5\underline{C}H_2$), 54.24 ($NH\underline{C}HCH_2$), 50.66, 50.28 ($\underline{C}HCH_3$), 41.85 ($\underline{C}H_2CH$), 28.03 ($\underline{C}H_2CN$), 25.84 ($\underline{C}H(CH_3)_2$), 23.38, 21.92 ($CH(\underline{C}H_3)_2$) and 17.57, 17.63 ($CH\underline{C}H_3$); *m/z* (CI) 446 (50%, $[M + H]^+$), 408 (45, $[M + H - CH_2CN]^+$), 338 (100, $[M - OCH_2C_6H_5)]^+$) and 312 (30, $[M + 2H - O_2CCH_2C_6H_5)]^+$).

### 3.0.2.49 Carbobenzyloxy-(2S)-leucyl-(2S)-alanyl (2R)-alanyl-aminoacetonitrile (81)

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2R)-alanine (**91**) (200 mg, 0.49 mmol) and aminoacetonitrile hydrochloride (**85**) (45 mg, 0.49 mmol) to give a white solid which was recrystalised from MeOH / ether to give white crystals (144 mg, 66%), m.p.140°C; (Found: $[M + H]^+$, 446.2408. $C_{22}H_{32}N_5O_5$ requires 446.2417); $[\alpha]_D$ -16.6 (*c* 0.20 in MeOH); $\upsilon_{max}$ (KBr disc)/$cm^{-1}$ 3300 (NH), 3030 (aromatic CH), 2920, 2900, 2850 (3 x aliphatic CH), 2250 (CN) and 1690, 1670, 1650, 1640 ($CO_2NH$ and 3 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.36 (5H, m, $C_6H_5$), 5.16 (1H, d, $J_{H_b}$ 12.44, $C_6H_5C\underline{H}_aH_b$), 5.16 (1H, d, $J_{H_a}$ 12.45, $C_6H_5CH_a\underline{H}_b$), 4.38 (1H, q, $J_{CH_3}$ 7.23, $C\underline{H}CH_3$), 4.28 (1H, q, $J_{CH_3}$ 7.21, $C\underline{H}CH_3$), 4.20 (1H, d, $J_{H_b}$ 7.47, $C\underline{H}_aH_bCN$), 4.17 (1H, m, $NHC\underline{H}CH_2$), 4.13 (1H, d, $J_{H_a}$ *8.46*, $CH_a\underline{H}_bCN$), 1.74 (1H, sep, *J* 6.69, $C\underline{H}(CH_3)_2$), 1.60 (2H, t, *J* 7.27, $CHC\underline{H}_2CH$), 1.43 (3H, d, $J_{CH}$ 7.36, $CHC\underline{H}_3$), 1.40 (3H, d, $J_{CH}$ 7.15, $CHC\underline{H}_3$), 0.98 (3H, d, $J_{CH}$ 5.97, $CH(C\underline{H}_3)_2$) and 0.96 (3H, d, $J_{CH}$ 6.26, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 176.17, 175.59, 175.23 (3 x $NH\underline{C}O$), 159.03 ($CO_2NH$), 138.44 (*i*-$C_6H_5$), 129.77 (*m*-$C_6H_5$), 129.32 (*p*-$C_6H_5$), 129.15 (*o*-$C_6H_5$), 117.73 (CN), 68.11 ($C_6H_5\underline{C}H_2$), 55.23 ($\underline{C}HCH_2$), 51.20, 50.63 ($\underline{C}HCH_3$), 42.20 ($\underline{C}H_2CH$), 28.45 ($\underline{C}H_2CN$), 26.14 ($\underline{C}H(CH_3)_2$), 23.71, 22.21 ($CH(\underline{C}H_3)_2$) and 17.77, 17.45 ($CH\underline{C}H_3$); *m/z* (CI) 446 ($[M + H)]^+$, 100%), 338 (15, $[M - OCH_2C_6H_5)]^+$) and 312 (18, $[M + 2H - O_2CCH_2C_6H_5)]^+$).

### *3.0.2.50 Trifluroacetyl-(2S)-leucyl-(2S)-alanyl-glycyl-amino acetonitrile (82)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using trifluoroacetyl-(2S)-leucyl (2S)-alanyl glycine (**97**) (172 mg, 0.484 mmol) and amino acetonitrile hydrochloride (**85**) (45 mg, 0.484 mmol) to give an orange oil which was purified by silica chromatography eluting with 3:1 dichloromethane / 40 / 60 petroleum ether until the impurities eluted off and then elution with ethylacetate to give a pale yellow oil (156 mg, 82%), *m/z* (Found: $[M + H]^+$, 394.1711. $C_{15}H_{23}N_5O_4F_3$ requires 394.1702); $[\alpha]_D$ -11.9 (*c* 0.16 in MeOH); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3295, 3075 (2 x NH), 2960, 2875 (2 x aliphatic CH), 2280 (CN) 1710 ($CO_2NH$) and 1680, 1675, 1665 (3 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 4.54 (1H, q, $J_{CH_3}$ 4.53, C$\underline{H}$CH$_3$), 4.33 (1H, sep, *J* 6.94, NHC$\underline{H}$CH$_2$), 3.96 (1H, d, $J_{CH_b}$ 12.39, NHC$\underline{H}_a$H$_b$), 3.89 (1H, d, $J_{CH_a}$ 9.8, NHCH$_a\underline{H}_b$), 1.84 (1H, m, C$\underline{H}$(CH$_3$)$_2$), 1.72 (2H, t, $J_{CH_2}$ 9.03, CHC$\underline{H}_2$CH), 1.43 (3H, d, $J_{CH}$ 7.15, CHC$\underline{H}_3$), 1.03 (3H, d, *J* 6.44, CH(C$\underline{H}_3$)$_2$) and 1.01 (3H, d, *J* 3.30, CH(C$\underline{H}_3$)$_2$); $\delta_C$ (75.5 MHz; $^2H_3COC^2H_3$) 175.40, 173.85, 173.68, 171.90 (4 x NHCO), 159.23 (q, $J_{CF_3}$ 36.96, $COCF_3$), 109.25 ($CF_3$), 109.14 (CN), 53.86, 53.29 ($\underline{C}$HCH$_2$), 51.11, 51.06 ($\underline{C}$HCH$_3$), 43.33 (C$\underline{H}_2$CONHCH$_2$), 41.28, 41.23 (CH$\underline{C}$H$_2$CH), 25.97, 25.91 ($\underline{C}$H$_2$CN), 24.77, 24.74 ($\underline{C}$H(CH$_3$)$_2$), 23.35, 23.29, 21.94, 21.73 (CH($\underline{C}$H$_3$)$_2$) and 19.35, 17.29 (CH$\underline{C}$H$_3$); *m/z* (CI) 394 (100%, $[M + H]^+$), 376 (14, $[M + H - H_20]^+$), 355 (50, [M - HF - F]$^+$) and 281 (65, $[M - NHCOCF_3]^+$).

### *3.0.2.51 Trifluoroacetyl-(2S)-alanyl-glycyl-amidoacetonitrile (83)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using trifluoroacetyl-(2S)-alanyl glycine (**100**) (181 mg, 0.748 mmol) and aminoacetonitrile hydrochloride (69 mg, 0.748 mmol). A white solid resulted which was recrystallised from dichloromethane / hexane to give white crystals (72 mg, 35%), m.p. 149°C; *m/z* (Found: $[M + H]^+$, 281.0860. $C_9H_{12}N_4O_3F_3$ requires 281.0862); $[\alpha]_D$ -41.3 (*c* 0.16 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3370, 3300 (2 x NH), 3060, 2990, 2950 (3 x aliphatic CH), 2255, 2240 (2 x CN) and 1720, 1710, 1695, 1675, 1665, 1655 (6 x CONH); $\delta_H$ (300 MHz, $C^2H_3O^2H$) 4.50 (1H, q, $J_{CH_3}$ 7.19, C$\underline{H}$CH$_3$), 4.22 (2H, s, CH$_2$), 3.98 (1H, d, $J_{H_b}$ 16.86, C$\underline{H_a}$H$_b$), 3.91 (1H, d, $J_{H_a}$ 16.87, CH$_a\underline{H_b}$) and 1.49 (3H, d, $J_{CH}$ 7.17, CH$_3$); $\delta_C$ (73.5 MHz, $C_2H_3O_2H$) 174.35, 171.99 (2 x CO), 159.24 (q, $J_{CF_3}$ 37.5, $\underline{C}$OCF$_3$), 117.66 (CN), 117.61 (q, $J_{F_3}$ 286, CF$_3$), 51.30 (CH), 43.55 ($\underline{C}$H$_2$CO), 28.23 ($\underline{C}$H$_2$CN) and 17.65 (CH$_3$); *m/z* (CI) 281 (100%, $[M + H]^+$), 260 (22, $[M - HF]^+$) and 225 (31, $[M + H - H_2NCH_2CN]^+$).

### *3.0.2.52 Trifluoroacetyl-glycyl-amidoacetonitrile (84)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using trifluoroacetyl-glycine (**98**) (0.41 g, 2.41 mmol) and aminoacetonitrile hydrochloride (**85**) (0.22 g, 2.41 mmol) to give an off white solid which was recrystalised from methanol to give white needle shaped crystals (283 mg, 57%), m.p. 143°C; (Found: C, 34.4; H, 2.6; N, 19.85. $C_6H_6N_3O_2F_3$ requires C, 34.45; H, 2.9; N, 20.10%; *m/z* (Found: $[M + H]^+$, 210.0491. $C_6H_7N_3O_2F_3$ requires 210.0491); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3300, 3100 (2 x NH), 2990, 2940, 2795, 2710 (4 x aliphatic CH), 2270 (CN) and 1730, 1670 (2 x CONH); $\delta_H$ (200 MHz, $^2H_3CO^2H$) 4.20 (2H,

s, $CH_2CON$) and 3.99 (2H, s, $CH_2CN$); $\delta_C$ (73.5 MHz, $_2H_3CO_2H$) 170.54 (CO), 159.60 (q, $J_{CF_3}$ 148.2, $\underline{C}OCF_3$), 117.65 (q, $J_{F_3}$ 570.0, $CF_3$), 115.80 (CN), 43.27 ($CH_2$) and 28.26 ($CH_2CN$); *m/z* (CI) 210 (100%, $[M + H]^+$).

### *3.0.2.53 Aminoacetonitrile hydrochloride (85)*

Methyleneaminoacetonitrile (40 mg, 0.74 mmol) was dissolved in ethanolic HCl ($1.0moldm^{-3}$, 0.88 $cm^3$) at -5°C and stirred for 1 h. The solvent was removed under reduced pressure and the white light sensitive solid was recrystallised from ethanol / ether to give white crystals (8.5 mg, 13%), $\delta_H$ (300 MHz, $^2H_3CO^2H$) 4.13 (s, $CH_2$); $\delta_C$ (75.5 MHz, $^2H_3CO^2H$) 115.08 (CN) and 28.31 ($CH_2$).

### *3.0.2.54 [1-$^{13}$C]-Carbobenzyloxy-(2S)-leucyl-(2S)-alanyl-glycyl-amino acetonitrile (86)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using [1-$^{13}$C]-aminoacetonitrile hydrochloride (**102**) (39 mg, 0.417 mmol). Recrystallisation from MeOH / ether gave white crystals (162 mg, 90%), *m/z* (Found: $[M + H]^+$, 433.2279. $C_{20}{}^{13}CH_{30}N_5O_5$ requires 433.2280); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.38 (5H, m, $C_6H_5$), 5.15 (1H, d, $J_{H_b}$ 13.45, $C\underline{H}_aH_bC_6H_5$), 5.11 (1H, d, $J_{H_a}$ 12.70, $CH_a\underline{H}_bC_6H_5$), 4.32 (1H, q, $J_{CH_3}$ 6.98, $C\underline{H}CH_3$), 4.19 (2H, d, $J_C$ 8.23, $C\underline{H}_aH_bCN$ and $NHC\underline{H}CH_2$), 4.19 (1H, d, $J_C$ 7.49, $CH_a\underline{H}_bCN$), 4.00 (1H, d, $J_{CH_b}$ 17.05, $NHC\underline{H}_aH_b$), 3.83 (1H, d, $J_{CH_a}$ 17.08, $NHCH_a\underline{H}_b$),), 1.76 (1H, sep, $J_{CH_3}$ 6.68, $C\underline{H}(CH_3)_2$), 1.61 (2H, t, $J_{CH_2}$ 5.71, $CHC\underline{H}_2CH$), 1.44 (3H, d, $J_{CH}$ 7.11, $CHC\underline{H}_3$), 1.00 (3H, d, $J_{CH}$ 6.38, $CH(C\underline{H}_3)_2$) and 0.98 (3H, d, $J_{CH}$ 6.27, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$)

175.96, 175.47, 171.99 (3 x NHCO), 158.20 ($CO_2NH$), 138.13 (*i*-$C_6H_5$), 129.48 (*m*-$C_6H_5$), 129.03 (*p*-$C_6H_5$), 128.81 (*o*-$C_6H_5$), 117.40 (CN), 67.80 ($C_6H_5\underline{C}H_2$), 54.98 ($\underline{C}HCH_2$), 50.92 ($\underline{C}HCH_3$), 43.30 ($NHC\underline{H}_2CO$), 41.81 ($CH\underline{C}H_2CH$), 27.94 (d, $J_C$ 129, $\underline{C}H_2CN$), 25.82 ($\underline{C}H(CH_3)_2$), 23.46, 21.84 ($CH(\underline{C}H_3)_2$) and 17.26 ($CH\underline{C}H_3$); *m/z* (CI) 433 (25%, $[M + H]^+$) and 325 (100, $[M - OCH_2C_6H_5]^+$).

## *3.0.2.55 Methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2S)-alaninate (87)*[122]

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl (2S)-alaninate (**37**) using carbobenzyloxy-(2S)-leucyl (2S)-alanine (**29**) (0.88 g, 2.62 mmol), methyl (2S)-alaninate hydrochloride (**33**) (360m g, 2.62 mmol) and the work up was modified as follows; ethyl acetate was used as a solvent rather than ether, the remaining ethyl acetate was azeotropically removed with methylene chloride from the white solid which was then recrystallised from acetone / water to give white crystals (1.27 g, 87%). m.p. 149-150°C, (lit.,[122] 172-173°C); $[\alpha]_D$ -64.1 (*c* 0.37 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3290 (NH), 3080 (aromatic CH), 2980, 2960, (2 x aliphatic CH), 1750 ($CO_2Me$), 1710 ($CO_2NH$) and 1650 (CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.38 (5H, m, $C_6H_5$), 5.13 (2H, s, $C_6H_5C\underline{H}_2$), 4.45 (1H, q, $J_{CH_3}$ 7.06, $C\underline{H}CH_3$), 4.40 (1H, q, $J_{CH_3}$ 7.19, $C\underline{H}CH_3$), 4.20 (1H, t, $J_{CH_2}$ 7.45, $NHC\underline{H}CH_2$), 3.75 (3H, s, $OCH_3$), 1.72 (1H, sep, *J*.6.58, $C\underline{H}(CH_3)_2$), 1.59 (2H, t, *J* 7.03, $CHC\underline{H}_2CH$), 1.42 (3H, d, $J_{CH}$ 6.61, $CHC\underline{H}_3$), 1.40 (3H, d, $J_{CH}$ 5.25, $CHC\underline{H}_3$), 0.99 (3H, d, $J_{CH}$ 6.28, $CH(C\underline{H}_3)_2$) and 0.97 (3H, d, $J_{CH}$ 6.22, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 175.13, 174.55, 174.50 (3 x $NH\underline{C}O$), 158.56 ($CO_2NH$), 136.19 (*i*-$C_6H_5$), 129.47 (*m*-$C_6H_5$), 128.99 (*p*-$C_6H_5$), 128.77 (*o*-$C_6H_5$), 67.76 ($C_6H_5\underline{C}H_2$), 54.91 ($\underline{C}HCH_2$), 52.72 ($OCH_3$), 50.00, 49.43 ($\underline{C}HCH_3$), 42.03 ($\underline{C}H_2CH$), 25.21 ($\underline{C}H(CH_3)_2$),

23.44, 21.86 ($CH(\underline{C}H_3)_2$) and 18.07, 17.36 ($CH\underline{C}H_3$); *m/z* (CI) 422 (100%, $[M + H]^+$) and 314 (42, $[M - OCH_2C_6H_5)]^+$).

### 3.0.2.56 Carbobenzyloxy-(2S)-leucyl-(2S)-alanyl (2S)-alanine (88)

This compound was prepared in a manner identical to that described for the carbobenzyloxy-(2S)-leucyl (2S)-alaninate (**29**) using methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2S)-alaninate (**87**) (1.00 g, 2.37 mmol) to give a white solid which was recrystallised from MeOH / ether to give white crystals (661 mg, 69%), m.p. 84-86°C; $[\alpha]_D$ -80.5 (*c* 0.22 in MeOH); $\upsilon_{max}$ (KBr disc)/$cm^{-1}$ 3110 (NH), 3050, 3030, 3000 (3 x aromatic CH), 2920, 2890, 2850 (3 x aliphatic CH), 1720, 1710 ($CO_2NH$ and $CO_2H$), and 1680, 1640 (2 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.36 (5H, m, $C_6H_5$), 5.13 (2H, s, $C_6H_5C\underline{H}_2$), 4.43 (2H, q, $J_{CH_3}$ 7.10, 6.93, 2 x $C\underline{H}CH_3$), 4.22 (1H, t, $J_{CH_2}$ 8.22, 6.49, $NHC\underline{H}CH_2$), 1.72 (1H, sep, *J*.6.61, $C\underline{H}(CH_3)_2$), 1.61 (2H, t, *J* 6.76, $CHC\underline{H}_2CH$), 1.42 (6H, t, $J_{CH}$ 7.57, 8.06, 2 x $CHC\underline{H}_3$), 0.99 (3H, d, $J_{CH}$ 6.33, $CH(C\underline{H}_3)_2$) and 0.97 (3H, d, $J_{CH}$ 6.09, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 175.68, 175.11, 174.37 (3 x NHCO), 158.51 ($CO_2NH$), 136.16 (*i*-$C_6H_5$), 129.45 (*m*-$C_6H_5$), 128.97 (*p*-$C_6H_5$), 128.76 (*o*-$C_6H_5$), 67.74 ($C_6H_5\underline{C}H_2$), 54.85 ($\underline{C}HCH_2$), 50.02 (2 x $\underline{C}HCH_3$), 42.04 ($\underline{C}H_2CH$), 25.84 ($\underline{C}H(CH_3)_2$), 23.47, 21.94 ($CH(\underline{C}H_3)_2$) and 18.06, 17.65 ($CH\underline{C}H_3$); *m/z* (CI) 408 (100%, $[M + H]^+$), 390 (22, $[M - OH]^+$), 364 (14, $[M + H - CO_2]^+$), 319 (17, $[M - NHCHCH_3CO_2H]^+$) and 274 (10, $[M + 2H - O_2CCH_2C_6H_5]^+$).

### *3.0.2.57 Teritary butyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2R)-alaninate (90)*

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl (2S)-alaninate (**37**) using carbobenzyloxy-(2S)-leucyl-(2S)-alanine (**29**) (0.88 g, 2.62 mmol), tertiary butyl (2R)-alaninate hydrochloride (**89**) and the work up was modified as follows; ethyl acetate was used as the extraction solvent, the remaining ethyl acetate was azeotropically removed with methylene chloride from the white solid which was recrystallised from acetone / water to give white needle shaped crystals (962 mg, 79%). m.p. 121-122°C; *m/z* (Found: $[M + H]^+$, 464.2764. $C_{24}H_{38}N_3O_6$ requires 464.2761); $[\alpha]_D$ -23.0 (*c* 0.10 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3300 (NH), 3050, 3030 (2 x aromatic CH), 2950, 2930, 2900, 2820 (4 x aliphatic CH), 1720 ($CO_2R$), 1700 ($CO_2NH$) and 1640 (CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.36 (5H, m, $C_6H_5$), 5.17 (1H, d, $J_{CH_b}$ 12.63, $C_6H_5C\underline{H}_aH_b$), 5.12 (1H, d, $J_{CH_a}$ 12.42, $C_6H_5CH_a\underline{H}_b$), 4.46 (1H, q, $J_{CH_3}$ 7.05, $C\underline{H}CH_3$), 4.29 (1H, q, $J_{CH_3}$ 7.21, $C\underline{H}CH_3$), 4.16 (1H, t, $J_{CH_2}$7.45, $NHC\underline{H}CH_2$), 1.74 (1H, sep, *J*.6.73, $C\underline{H}(CH_3)_2$), 1.59 (2H, t, *J* 7.30, $CHC\underline{H}_2CH$), 1.41 (3H, d, $J_{CH}$ 5.38, $CHC\underline{H}_3$), 1.39 (3H, d, $J_{CH}$ 5.22, $CHC\underline{H}_3$), 0.99 (3H, d, $J_{CH}$ 6.70, $CH(C\underline{H}_3)_2$) and 0.97 (3H, d, $J_{CH}$ 6.75, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 175.26, 174.33, 173.49 (3 x $NH\underline{C}O$), 158.78 ($CO_2NH$), 138.09 (*i*-$C_6H_5$), 129.49 (*m*-$C_6H_5$), 129.05 (*p*-$C_6H_5$), 128.89 (*o*-$C_6H_5$), 82.79 ($C_6H_5\underline{C}H_2$), 67.88 ($C_6H_5\underline{C}H_2$), 55.36 ($\underline{C}HCH_2$), 50.28, 50.06 (2 x $\underline{C}HCH_3$), 41.93 ($\underline{C}H_2CH$), 28.22 ($C(\underline{C}H_3)_3$), 25.89 ($\underline{C}H(CH_3)_2$), 23.43, 21.89 ($CH(\underline{C}H_3)_2$) and 17.57, 18.02 ($CH\underline{C}H_3$); *m/z* (CI) 464 (100%, $[M + H]^+$) and 408 (77, $[M + 2H - C(CH_3)_3]^+$).

### 3.0.2.58 Carbobenzyloxy-(2S)-leucyl (2S)-alanyl (2R)-alanine (91)

This compound was prepared in a manner identical to that described for the Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycine (**27**) using tertiary butyl carbobenzyloxy-(2S)-leucyl-(2S) alanyl (2R)-alaninate (**90**) (800 mg, 1.73 mmol) to give a white solid which was recrystallised from MeOH / ether to give white crystals (408 mg, 58%). m.p. 82-84°C; $[\alpha]_D$ -32.1 (*c* 0.23 in MeOH); $\upsilon_{max}$ (KBr disc)/cm$^{-1}$ 3500 - 3200 (OH), 330 (NH), 3020, 2940 (2 x aromatic CH), 2800 (aliphatic CH), 1720, 1710 ($CO_2H$ and OCONH) and 1680, 1640 (2 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.39 (5H, m, $C_6H_5$), 5.17 (1H, d, $J_{H_a}$ 12.48, $C_6H_5CH_a\underline{H}_b$), 5.11 (1H, d, $J_{H_b}$ 12.38, $C_6H_5C\underline{H}_aH_b$), 4.47 (1H, q, $J_{CH_3}$ 7.15, C$\underline{H}$CH$_3$), 4.40 (1H, q, $J_{CH_3}$ 7.24, C$\underline{H}$CH$_3$), 4.16 (1H, t, $J_{CH_2}$ 7.48, NHC$\underline{H}$CH$_2$), 1.76 (1H, sep, *J*.6.64, C$\underline{H}$(CH$_3$)$_2$), 1.61 (2H, dd, *J* 7.14, 7.31, CHC$\underline{H}_2$CH), 1.46 (3H, d, $J_{CH}$ 7.31, CHC$\underline{H}_3$), 1.39 (3H, d, $J_{CH}$ 7.13, CHC$\underline{H}_3$), 0.99 (3H, d, $J_{CH}$ 676, CH(C$\underline{H}_3$)$_2$) and 0.97 (3H, d, $J_{CH}$ 6.68, CH(C$\underline{H}_3$)$_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 176.00, 175.60, 174.55 (3 x NH$\underline{C}$O), 159.10 (OCONH), 138.35 (*i*-$C_6H_5$), 129.76 (*m*-$C_6H_5$), 129.33 (*p*-$C_6H_5$), 129.22 (*o*-$C_6H_5$), 68.20 ($C_6H_5\underline{C}H_2$), 55.69 ($\underline{C}$HCH$_2$), 50.36 ($\underline{C}$HCH$_3$), 42.15 ($\underline{C}$H$_2$CH), 26.17 ($\underline{C}$H(CH$_3$)$_2$), 23.70, 22.15 (CH($\underline{C}$H$_3$)$_2$) and 18.23, 18.07 (CH$\underline{C}$H$_3$); *m/z* (CI) 408 ([M + H]$^+$, 88%), 390 (14, [M + H - $H_2O$]$^+$), 320 (60, [M + H - NHCHCH$_3$CO$_2$H]$^+$) and 264 (100, [M + H - CHCH$_3$CONHCHCH$_3$CO$_2$H]$^+$).

### 3.0.2.59 Trifluoroacetyl-(2S)-leucine (92)[123]

To a stirred solution of (2S)-leucine (4.00 g, 30.0 mmol) in trifluoroacetic acid (15 cm$^3$) was added trifluoroacetic anhydride (5.2 cm$^3$, 39.0 mmol) was added dropwise at -10°C with stirring. After 5 min the salt / ice bath was removed and the reaction was stirred at room temperature for 30 min. The trifluoroacetic acid was removed by distillation under reduced

pressure to give a white amorphous solid which was recrystallised from dichloromethane to give white plate like crystals (5.60 g, 82%), m.p. 71-72°C (lit.,[123] 71-73°C); (Found: C, 42.55; H, 5.5; N, 6.60. Calc. for $C_8H_{12}NO_3F_3$ C, 42.3; H, 5.35; N, 6.20%); *m/z* (Found: $[M + NH_3 + H]^+$, 245.1115. Calc. for $C_8H_{12}N_2O_3F_3$ 245.1113); $[\alpha]_D$ -39.9 (*c* 0.6 in EtOH); $\upsilon_{max}$ (nujol)/$cm^{-1}$ 3290, 3110 (NH and OH), 1730 ($CO_2NH$) and 1710 ($CO_2H$); $\delta_H$ (200 MHz; $C^2HCl_3$) 10.42 (1H, s, $CO_2H$), 6.97 (1H, d, $J_{CH}$ 8.14, NH), 4.67 (1H, m, $CH_2C\underline{H}$), 1.72 (3H, m, $C\underline{H}_2C\underline{H}(CH3)_2$) and 0.97 (6H, d, $J_{CH}$ 4.68, $CH(C\underline{H}_3)_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 177.09 ($CO_2H$), 157.77 (q, $J_{CF_3}$ 38.0, $COCF_3$), 116.06 (q, $J_{CF_3}$ 288, $CF_3$), 51.57 (NHCH), 41.18 ($CH_2$), 25.31 ($\underline{C}H(CH_3)_2$) and 23.08, 22.05 ($CH(\underline{C}H_3)_2$); *m/z* (CI) 245 (100%, $[M + H + NH_3]^+$), 225 (4, $[M + NH_3 - F]^+$), 182 (3, $[M + NH_3 - F - CO_2]^+$), 168 (4, $[M - CO_2 - CH_3]^+$) and 114 (2, $[M + 2H - CH(CH_2CH(CH_3)_2)CO_2H]^+$).

### *3.0.2.60 Benzyl (2S)-alaninate tosylate salt (93)*

(2*S*)-Alanine (5.00 g, 56.1 mmol) was dispersed in toluene (100 $cm^3$) and benzyl alcohol (30 $cm^3$, 280 mmol) and *p*-toluene sulphonic acid.$H_2O$ (13.95 g, 0.067mol) were added. The reaction was refluxed under Dean and Stark conditions for 5 h. On cooling yellow needle shaped crystals formed which were filtered and washed with ether (150 $cm^3$). Recrystallisation from methanol / ether gave white needle shaped crystals (18.45 g, 89%), m.p. 198-199°C; $[\alpha]_D$ +5.8 (*c* 0.36 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 1750 (CO); $\delta_H$ (300 MHz, $(^2H_3C)_2SO$) 8.38°(3H, s, $NH_3^+$), 7.51 (2H, d, $J_{CH}$ 8.0, $C_6H_4$), 7.40 (5H, m, $C_6H_5$), 7.13 (2H, d, $J_{CH}$ 7.86, $C_6H_5$), 5.23 (2H, s, $C\underline{H}_2C_6H_5$), 4.19 (1H, m, $C\underline{H}CH_3$), 2.29 (3H, s, $C_6H_4C\underline{H}_3$) and 1.41 (3H, d, $J_{CH}$ 7.2,$CHC\underline{H}_3$); $\delta_C$ (73.5 MHz, $^2H_3\underline{C}SOC^2H_3$) 170.00 (CO), 145.64, 137.95, 135.41 (*i*-$C_6H_5$ and *i*-$C_6H_4$), 128.68, 128.53, 128.24, 125.67 ($C_6H_5$, $C_6H_4$), 67.20 ($\underline{C}H_2C_6H_5$), 48.14 (CH), 20.96 ($\underline{C}H_3C_6H_4$) and 15.87 ($CH\underline{C}H_3$).

### 3.0.2.61 Benzyl trifluoroacetyl-(2S)-leucyl-alaninate (94)

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl alaninate (**37**) using trifluoroacetyl-(2S)-leucine(**92**) (2.79 g, 12.3 mmol) and benzyl alaninate tosylate salt (**93**) (4.54 g, 12.3 mmol) to give an orange oil which solidified to a white waxy solid on standing (3.87 g, 77%), *m/z* (Found: $[M + H]^+$, 389.1693. $C_{18}H_{23}N_2O_4F_3$ requires 389.1688); $[\alpha]_D$ -34.1 (*c* 0.34 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3290 (NH), 3070 (aromatic CH), 2961, 2875 (2 x aliphatic CH), 1745 ($CO_2CH_2C_6H_5$) and 1715, 1700, 1660 (3 x NHCO); $\delta_H$ (300 MHz, $C^2H_3O^2H$) 7.34 (5H, m, $C_6H_5$), 5.18 (2H, m, $C\underline{H}_2C_6H_5$), 4.57 (2H, m, $C\underline{H}CH_3$ and $C\underline{H}CH_2$), 1.63 (3H, m, $C\underline{H}_2C\underline{H}(CH_3)_2$), 1.42 (3H, d, $J_{CH}$ 7.26, $CHC\underline{H}_3$) and 0.923 (6H, m, $CH(C\underline{H}_3)_2$); $\delta_C$ (70.3 MHz, $C^2H_3O^2H$) 172.34, 172.25, 170.76, 170.96 ($\underline{C}O_2CH_2$ and CONH), 156.96 (q, $\underline{C}OCF_3$), 135.27, 135.21 (*i*-$C_6H_5$), 128.57 (*o*-$C_6H_5$), 128.41 (*p*-$C_6H_5$), 128.09 (*m*-$C_6H_5$), 115.81 (q, $J_{F_3}$ 288, $CF_3$), 67.19, 67.14 ($\underline{C}H_2C_6H_5$), 52.09 ($\underline{C}HCH_2$), 48.28 ($\underline{C}HCH_3$) 41.28, 41.25 ($\underline{C}H_2CH(CH_3)_2$) 24.73, 24.60 ($\underline{C}H(CH_3)_2$), 22.74, 22.65, 21.95, 21.90 ($CH(\underline{C}H_3)_2$) and 17.86, 17.60 ($CH_3$); *m/z* (CI) 389 (100%, $[M + H]^+$), 371 (18, $[M - OH]^+$) and 280 (17, $[M - H_2O - CH_2C_6H_5]^+$).

### 3.0.2.62 Trifluoroacetyl-(2S)-leucyl-alanine (95)

This compound was prepared in a manner identical to that described for tertiary-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycine (**46**) using benzyl trifluoroacetyl-(2S)-leucyl (2S)-alanine (0.51 g, 1.27 mmol) and dry dichloromethane as the solvent to give a yellow oil which solidified to a white solid on standing (299 mg, 83%), m.p. 92°C; m/z (Found: $[M + H]^+$, 299.1224. $C_{11}H_{18}N_2O_4F_3$ requires 299.1219); $[\alpha]_D$ -39.5 (*c* 0.21 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3280, 3150 (NH and OH), 2965 (aliphatic CH), 1720 ($CO_2H$), and 1670,

1660 (2 x CONH); $\delta_H$ (300 MHz, $C^2H_3O^2H$, -50°C) 4.82, 4.73 (1H, dd, C<u>H</u>$CH_2$), 4.52, 4.50 (1H, q, C<u>H</u>$CH_3$), 1.81 (3H, m, C<u>H</u>$_2$C<u>H</u>$(CH_3)_2$), 1.60 (3H, dd, $J_{CH}$ 6.70, 6.46, CHC<u>H</u>$_3$) and 1.13 (6H, t, 8.84, 4.28, CH(C<u>H</u>$_3)_2$); $\delta_C$ (70.3 MHz, $C^2H_3O^2H$) 176.27, 176.21, 173.92, 173.74, ($CO_2H$ and CONH), 158.80 (m, $COCF_3$), 117.69, 117.64 (q, $J_{F_3}$ 286.2, 286.5, $CF_3$), 53.08, 52.84 (<u>C</u>H$CH_2$), 49.70 (<u>C</u>H$CH_3$) 41.79, 41.25 (<u>C</u>H$_2$CH$(CH_3)_2$) 26.11, 25.98 (<u>C</u>H$(CH_3)_2$), 23.90, 23.86, 21.73, 21.57 (CH(<u>C</u>H$_3)_2$) and 17.17, 17.50 ($CH_3$); m/z (CI) 299 (100%, $[M + H]^+$) and 281 (22, $[M - OH]^+$).

## *3.0.2.63 Benzyl trifluoroacetyl-(2S)-leucyl-(2S)-alanyl-glycinate (96)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alaninate (**37**) using trifluoroacetyl-(2S)-leucyl (2S)-alanine (**94**) (136 mg, 0.456 mmol), benzyl glycinate tosylate (**44**) (154 mg, 0.456 mmol) and the work up was modified as follows; the crude product was extracted into ethyl acetate before washing and the remaining ethyl acetate was azeotropically removed with methylene chloride from the clear oil to give a white solid (187 mg, 92%), m.p. 121°C; *m/z* (Found: $[M + H]^+$, 446.1902). $C_{20}H_{27}N_3O_5F_3$ requires 446.1903); $[\alpha]_D$ -27.2 (*c* 0.18 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3310, 3300 (2 x NH), 3090, 3040 (2 x aromatic CH), 2960, 2880 (2 x aliphatic CH), 1740 ($CO_2CH_2C_6H_5$) and 1720, 1700, 1690, 1675, 1655, 1640 (6 x CONH); $\delta_H$ (300 MHz; $C^2H_3O^2H$, -50°C) 7.57 (5H, m, $C_6H_5$), 5.33 (2H, s, $C_6H_5C$<u>H</u>$_2$), 4.71 (1H, m, C<u>H</u>$CH_2$), 4.52 (1H, q, $J_{CH}$ 6.98, C<u>H</u>$CH_3$), 4.27 (2H, d, $J_{CH_b}$ 17.80, NHC<u>H</u>$_a$H$_b$), 4.12 (1H, d, $J_{CH_b}$ 14.08, NHC<u>H</u>$_a$H$_b$), 4.07 (1H, d, $J_{CH_a}$ 14.22, NHCH$_a$<u>H</u>$_b$), 3.04 (1H, q, $J_{CF_3}$ 19.47, $NHCOCF_3$), 1.83 (3H, m, CHC<u>H</u>$_2$C<u>H</u>$(CH_3)_2$), 1.57 (3H, d, $J_{CH}$ 5.39, CHC<u>H</u>$_3$), 1.37 (3H, d, $J_{CH}$ 5.39, CH(C<u>H</u>$_3)_2$) and 1.34 (3H, d, $J_{CH}$ 5.39, CH(C<u>H</u>$_3)_2$); $\delta_C$ (73.5 MHz; $C^2H_3O^2H$) 176.12, 175.92, 173.88, 170.93 ($CO_2H$ and 2 x NHCO), 159.17 (m, <u>C</u>$OCF_3$), 137.36

(*i*-$C_6H_5$), 19.92 (*m*-$C_6H_5$), 129.75 (*p*-$C_6H_5$, *o*-$C_6H_5$), 117.66 (q, $J_{F_3}$ 286.6 $CF_3$), 117.61 (q, $J_{F_3}$ 285.8, $CF_3$), .68.16 ($C_6H_5\underline{C}H_2$), 53.71, 53.07 ($\underline{C}HCH_2$), 50.75, 50.56 ($\underline{C}HCH_3$), 42.02 ($NHCH_2$), 41.28, 41.09 ($\underline{C}H_2CH$), 26.12, 26.04 ($\underline{C}H(CH_3)_2$), 23.94, 23.60 ($CH(\underline{C}H_3)_2$) and 18.24 ($CH\underline{C}H_3$); *m/z* (CI) 446 (100%, $[M + H]^+$) and 281 (11, $[M - NHCH_2CO_2CH_2C_6H_5]^+$).

### *3.0.2.64 Trifluoroacetyl-(2S)-leucyl (2S)-alanyl-glycine (97)*

Trifluoroacetyl-(2S)-leucyl (2S)-alanyl glycine (**97**) was prepared in a manner identical to that described for tertiary-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycine (**46**) using benzyl trifluoroacetyl-(2S)-leucyl (2S)-alanyl glycinate (**96**) (472 mg, 0.106 mmol) and dry dichloromethane to give a yellow oil (172 mg, 46%), *m/z* (Found: $[M]^+$, 355.1349. $C_{13}H_{20}N_3O_5F_3$ requires 355.1355); $[\alpha]_D$ -24.4 (*c* 0.39 in MeOH); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3360 (OH), 3295, 3080 (2 x NH), 2960, 2875 (2 x aliphatic CH), 1735 ($CO_2H$) and 1720, 1665 (2 x CONH); $\delta_H$ (300 MHz; $C^2H_3O^2H$) 4.53 (1H, m, $C\underline{H}CH_2$), 4.45 (1H, q, $J_{CH_3}$ 7.00, $C\underline{H}CH_3$), 3.95, 3.93 (2H, s, $NHC\underline{H}_2$), 1.62 (3H, m, $C\underline{H}_2C\underline{H}(CH_3)_2$), 1.43 (3H, d, $J_{CH}$ 7.22, $CHC\underline{H}_3$) and 0.98 (6H, m, $CH(C\underline{H}_3)_2$); $\delta_C$ (73.5 MHz; $C^2H_3O^2H$) 174.94, 173.07, 172.64 (2 x $NH\underline{C}O$ and $CO_2H$), 53.96, 53.50 ($\underline{C}HCH_2$), 50.34, 50.21 ($\underline{C}HCH_3$), 41.74 ($NHCH_2$), 41.32 ($\underline{C}H_2CH$), 24.73 ($\underline{C}H(CH_3)_2$), 23.31, 23.17, 21.76, 22.02 ($CH(\underline{C}H_3)_2$) and 17.88 ($CH\underline{C}H_3$); *m/z* (CI) 355 (70 %, $[M]^+$), 281 (15, $[M - NHCH_2CO_2H]^+$) and 176 (100, $[M - HF - 2CH_3 - CHCH_3CONHCH_2CO_2H]^+$).

### 3.0.2.65 Trifluoroacetyl-(2S)-alanine (98)[123]

This compound was prepared in a manner identical to that described for trifluoroacetyl-(2S)-leucine (**92**) using (2*S*)-alanine (2.00 g, 22.4 mmol) to give a yellow oil which solidified on standing overnight to oily white needle shaped crystals (3.55 g, 93%), (Found: C, 32.8; H, 3.35; N, 7.70. Calc. for $C_5H_6NO_3F_3$ C, 32.45; H, 3.25; N, 7.55%); *m/z* (Found: $[M + H]^+$, 186.0379. Calc. for $C_5H_7NO_3F_3$ 186.0378); $[\alpha]_D$ -46.7 (*c* 0.14 in MeOH); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3304, 3100 (2 x NH), 1720 ($CO_2H$) and 1670, 1655 (2 x CONH); $\delta_H$ (300 MHz, $C^2HCl_3$) 11.20 (1H, s, $CO_2H$), 7.19 (1H, d, $J_{CH}$ 7.19, NH), 4.63 (1H, q, $J_{CH_3}$ 7.25, CH) and 1.55 (3H, d, $J_{CH}$ 7.26, $CH_3$); $\delta_C$ (73.5 MHz, $C^2HCl_3$) 176.30 (CO), 157.26 (q, $J_{CF_3}$ 38.0, $\underline{C}OCF_3$), 115.5 (q, $J_{CF_3}$ 287.0, $CF_3$), 48.57 ($CH_2$) and 17.26 ($CH_3$); *m/z* (CI) 186 (100%, $[M + H]^+$) and 168 (36, $[M - OH]^+$).

### 3.0.2.66 Benzyl trifluoroacetyl-alanyl glycinate (99)

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl alaninate (**37**) using trifluoroacetyl alanine (**98**) (0.50 g, 2.70 mmol) and benzyl glycinate tosylate salt (**44**) (0.99 g, 2.7 mmol) to give a white solid which was recrystallised from dichloromethane / hexane to give white crystals (0.86 g, 96%), m.p. 84-86°C; (Found: C, 50.65; H, 4.5; N, 8.15. $C_{14}H_{15}N_2O_4F_3$ requires C, 50.6; H, 4.55; N, 8.45%; *m/z* (Found: $[M + H]^+$, 333.1066. $C_{14}H_{16}N_2O_4F_3$ requires 333.1062); $[\alpha]_D$ -54.2 (*c* 0.23 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3320 (NH), 3090, 3040, 3010, (3 x aromatic CH), 2985, 2950, 2890 (3 x aliphatic CH), 1750 ($CO_2CH_2$) and 1720, 1670 (2 x NHCO); $\delta_H$ (300 MHz, $C^2H_3O^2H$) 7.39 (5H, m, $C_6H_5$), 5.20 (2H, s, $C\underline{H}_2C_6H_5$), 4.53 (1H, q, $J_{CH_3}$ 7.13, CH), 4.10 (1H, d, $J_{CH_b}$ 17.63, $C\underline{H}_aH_b$), 3.98 (1H, d, $J_{CH_a}$ 17.67, $CH_a\underline{H}_b$), 2.90 (1H, q, $J_{CF_3}$ 17.5,

NHCHCH$_3$) and 1.46 (3H, d, $J_{CH}$ 7.16, CH$_3$); $\delta_C$ (70.3 MHz, C$^2$H$_3$O$^2$H) 174.00, 170 89 (CO$_2$CH$_2$ and CONH), 158.67 (q, $J_{CF_3}$ 39.0, COCF$_3$), 137.12 (*i*-C$_6$H$_5$), 129.54 (*o*-C$_6$H$_5$), 129.31 (*p*-C$_6$H$_5$), 129.28 (*m*-C$_6$H$_5$), 67.94 (CH$_2$C$_6$H$_5$), 117.33 (q, $J_{F_3}$ 287.0, CF$_3$), 50 64 (CH), 42.14 (CH$_2$) and 17.67 (CH$_3$); *m/z* (CI) 333 ([M +H]$^+$, 100%) 315 (5, [M - OH]$^+$) and 243 (12, [M + 2H - CH$_2$C$_6$H$_5$]$^+$).

### 3.0.2.67 Trifluoroacetyl-alanyl-glycine (100)

This compound was prepared in a manner identical to that described for tertiary butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycine (**46**) using benzyl trifluoroacetyl-(2S)-alanyl glycinate (**99**) (0.40 g, 1.20 mmol) and dichloromethane as the solvent to give a yellow oil (182 mg, 63%), *m/z* (Found: [M + H]$^+$, 243.0593. C$_7$H$_{10}$N$_2$O$_4$F$_3$ requires 243.0593); $[\alpha]_D$ -27.7 (c 0.07 in MeOH); $\upsilon_{max}$ (KBr disk)/cm$^{-1}$ 3395 (NH and OH), 2960 (aliphatic CH), 1720 (CO$_2$H) and 1710, 1670 (2 x CONH); $\delta_H$ (300 MHz, C$^2$H$_3$O$^2$H) 4.53 (1H, q, $J_{CH_3}$ 7.17, CH), 3.94 (1H, d, $J_{CH_b}$ 34.72, CH$_a$H$_b$), 3.89 (1H, d, $J_{CH_a}$ 27.67, CH$_a$H$_b$), and 1.49 (3H, d, $J_{CH}$ 7.18, CH$_3$); $\delta_C$ (70.3 MHz, C$^2$H$_3$O$^2$H) 173.84, 172.69 (CO$_2$H and CONH), 158.67 (q, $J_{CF_3}$ 37.51, COCF$_3$), 117.35 (q, $J_{F_3}$ 286.5, CF$_3$), 50 65 (CH), 41.80 (CH$_2$) and 17.66 (CH$_3$); *m/z* (CI) 243 (100%, [M + H]$^+$) and 225 (14, [M - H$_2$O]$^+$).

### 3.0.2.68 Trifluoroacetyl-glycine (101)[124]

This compound was prepared in a manner identical to that described for trifluoroacetyl-(2S)-leucine (**92**) using glycine (5.00 g, 66.7 mmol) to give white plate like crystals (9.99 g, 88%), m.p. 115-117°C (lit.,[124] 114-116.4°C); (Found: C, 28.25; H, 2.05; N, 8.15. Calc. for C$_4$H$_4$NO$_3$F$_3$ C, 28.05; H, 2.35; N, 8.15%); *m/z* (Found: [M]$^+$ 171.0135, C$_4$H$_4$NO$_3$F$_3$ requires

171.0143); $\upsilon_{max}$ (Nujol) /cm$^{-1}$ 3300, 3110 (2 x NH), 2590 (OH), 1825 ($CO_2NH$) and 1705 ($CO_2H$); $\delta_H$ (200 MHz, $^2H_2O$) 4.08 (2H, s, $CH_2$); $\delta_C$ (50.3 MHz, $^2H_2O$) 174.73 ($CO_2H$), 162.22 (q, $J_{CF_3}$ 37.52, $COCF_3$) 118.58 (q, $J_{F_3}$ 285.75, $CF_3$) and 43.82 ($CH_2$); *m/z* (EI) 171 (2%, $[M]^+$), 126 (81, $[M - CO_2H]^+$), 102 (4, $[M - CF_3]^+$), 97 (6, $[M - NHCH_2CO_2H]^+$) and 28 (100, $[CNH_2]^+$).

### *3.0.2.69 [1-$^{13}$C]-Aminoacetonitrile hydrochloride (102)[120]*

This compound was prepared in a manner identical to that described for aminoacetonitrile (**85**) using [1-$^{13}$C]-methylene aminoacetonitrile (**103**) (30 mg, 0.43 mmol) to give white crystals (39 mg, 96%), $\delta_H$ (300 MHz, $^2H_3CO^2H$) 4.18 (d, $J_C$ 9.03, $CH_2$); $\delta_C$ (75.5 MHz, $^2H_3CO^2H$) 114.80 ($^{13}$CN) and 28.01 (d, $J_C$ 64.4, $CH_2$).

### *3.0.2.70 [1-$^{13}$C]-Methylene aminoacetonitrile (103)[120]*

Formaldehyde solution (0.57 cm$^3$, 6.72 mmol) and $NH_4Cl$ (0.21 g, 3.84 mmol) were placed in a flask at -5°C with vigorous stirring. [1-$^{13}$C]-KCN (0.25 g, 3.84 mmol) was dissolved in water (2 cm$^3$) and added dropwise to the formaldehyde / $NH_4Cl$ solution over 20 min. After 10 min acetic acid (0.15 g, 2.53 mmol) in water (1 cm$^3$) was also added dropwise to the formaldehyde $NH_4Cl$ / KCN solution over 10 min and stirring continued for 1 h. Removal of the solvent (3 cm$^3$) under reduced pressure gave a white crystaline precipitate which was filtered and dried to give white crystals (35 mg, 15%), (Found: $[M + H]^+$, 70.0489. $^{13}C_1C_2H_5N_2$ requires 70.0486); $\delta_H$ (200 MHz; $^2H_3CO^2H$) 3.81 (2H, d, $J_C$ 7.51, $CH_2CN$) and 3.70 (2H, s, $NCH_2$); $\delta_C$ (50.3 MHz; $^2H_3CO^2H$) 116.76 ($^{13}$CN), 72.32 ($NCH_2$) and 40.51 (d, $J_C$ 56.14, $CH_2CN$); *m/z* (CI) 208 (3%, $[3M + H]^+$), 180 (33, [3M + H -

$NCH_2)]^+$), 139 ([2M + H]$^+$, 42), 111 (100, [2M + H - $NCH_2$)]$^+$), 70 (33, [M +H - $NCH_2$)]$^+$) and 42 (12, [M + H - $NCH_2$]).

### 3.0.2.71 Methylene aminoacetonitrile (104)[125]

This compound was prepared in a manner identical to that described for methylene aminoacetonitrile (**103**) using unlabelled KCN to give white crystals (54 mg, 24%), (Found: [2M + H]$^+$, 137.0825. $C_6H_9N_4$ requires 137.0823); $\delta_H$ (200 MHz; $^2H_3CO^2H$) 4.87 (2H, s, $CH_2CN$) and 3.66 (2H, s, $NCH_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 116.81 (CN), 72.36 ($NCH_2$) and 40.57 ($\underline{C}H_2CN$); *m/z* (CI) 341 (3%, [5M + H]$^+$), 273(15, [4M + H]$^+$), 205 (15, [3M + H)]$^+$), 137 (77, [2M + H)]$^+$), 110 (100, [2M + 2H - $NCH_2$]$^+$) and 69 (33, [M +H)]$^+$).

### 3.0.2.72 1-Acetoxy 1-carbobenzyloxy-amino methane (105)

To a stirred solution of carbobenzyloxy glycine (**34**) (750 mg, 3.59 mmol) in dry acetonitrile (40 cm$^3$) under argon in a 2 necked flask was added $Pb(OAc)_4$ (1.75 g, 3.95 mmol) and the mixture stirred for 10 min, to give a yellow solution and a black ppt. LiCl (167 mg, 3.95 mmol) was then added and the mixture was kept at 80-90°C for 2 h to give a clear solution and a white ppt. The solution was filtered and the solvent was distilled off at reduced pressure to give a black solid. The solid was washed with ether (3 x 100 cm$^3$), and after filtering, the solvent removed under reduced pressure to give a colourless unstable, moisture sensitive oil (490 mg, 62%), $\delta_H$ (200 MHz; $C^2HCl_3$) 7.35 (5H, m, $C_6H_5$), 6.10 (1H, t, $J_{CH}$ 5.6, NH), 5.19 (2H, d, $J_{NH}$ 7.6, $NCH_2$); 5.13 (2H, s, $C_6H_5C\underline{H}_2$) and 2.05 (3H, s, $CH_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 172.27 (CO), 156.34 (CONH), 136.28 (*i*-$C_6CH_5$), 129.06 (*m*-$C_6H_5$), 128.85 (*p*-$C_6H_5$), 128.75 (*o*-$C_6H_5$), 67.82, 67.12 ($C_6H_5\underline{C}H_2$ and $NCH_2O$) and 21.46 ($CH_3$).

### 3.0.2.73 Carbobenzyloxy-aminoacetonitrile (107)

1-Acetoxy 1-carbobenzyloxy-aminomethane (**106**) (468 mg, 2.10 mmol) was dissolved in dry acetonitrile under argon and KCN (150 mg, 2.31 mmol) was added and the solution was refluxed for 3 h. The solvent was removed under reduced pressure to give a white solid which was dissolved in ethyl acetate (40 $cm^3$), washed with water (3 x 30 $cm^3$), brine (saturated, 20 $cm^3$), dried ($MgSO_4$) and the solvent removed under reduced pressure to give a colourless oil (270 mg, 68%), $\delta_H$ (200 MHz; $C^2HCl_3$) 7.35 (5H, m, $C_6H_5$), 5.10 (1H, s, NH), 5.14 (2H, m, $C_6H_5C\underline{H}_2$) and 4.04 (2H, d, $J_{NH}$ 7.6, $NCH_2$); *m/z* (CI) 191 (9, $[M + H]^+$), 152 (15, $[M + 2H - CH_2CN]^+$) and 57 (100, $[M + 2H - CO_2CH_2C_6H_5]^+$).

### 3.0.2.74 Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycylaminoacetaldehyde (112)

N-(Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl) 1-aminoprop-2-ene (**115**) (30 mg, 69.4μmol) was dissolved in ethyl acetate (3 $cm^3$) in a 2 necked flask with heating and at -5-0°C ozonized oxygen in oxygen (0.3%) was bubbled through the solution until a KI / HOAc solution indicator (5%, 10%, 100 $cm^3$) went red (15 min). The ozone was flushed from the flask with $O_2$ and $Me_2S$ (12 $mm^3$, 166.6 μmol) was added and stirred for 1 h at room temperature.[136] The solvent was removed under reduced pressure and the resulting oil was dissolved in ethyl acetate (5 $cm^3$), washed with water (3 x 3 $cm^3$), brine solution (saturated, 3 $cm^3$), dried ($MgSO_4$) and the solvent removed under reduced pressure to give an unstable, hydroscopic, waxey white solid (22 mg, 70%), $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3260 (NH), 3000 (aromatic CH), 2960, 2900, 2810 (3 x aliphatic CH), 1720 ($CO_2NH$) and 1710, 1700, 1690, 1650, 1630 (5 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.38 (5H, m, $C_6H_5$), 5.53 (2H, s, $C\underline{H}_2C_6H_5$), 4.40 (2H, m, $C\underline{H}(OH)_2$ and $C\underline{H}CH_3$), 4.20 (1H, t, *J* 7.50, $NHC\underline{H}CH_2$), 3.96 (1H,

d, $J_{CH_b}$ 18.80, NHC<u>H</u>$_a$H$_b$), 3.92 (1H, d, $J_{CH_a}$ 14.86, NHCH$_a$<u>H</u>$_b$), 1.74 (1H, sep, $J_{CH_3}$ 6.70, C<u>H</u>(CH$_3$)$_2$), 1.63 (2H, t, $J_{CH_2}$ 7.84, CHC<u>H</u>$_2$CH), 1.42 (3H, d, $J_{CH}$ 7.06, CHC<u>H</u>$_3$), 0.99 (3H, d, $J_{CH}$ 6.38, CH(C<u>H</u>$_3$)$_2$) and 0.97 (3H, d, $J_{CH}$ 6.25, CH(C<u>H</u>$_3$)$_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 175.74, 175.31, 171.99 (3 x NHCO), 158.78 ($CO_2NH$), 138.14 (*i*-$C_6H_5$), 129.47 (*m*-$C_6H_5$), 129.01 (*p*-$C_6H_5$), 128.83 (*o*-$C_6H_5$), 90.81 ($CH(OH)_2$), 67.80 ($C_6H_5\underline{C}H_2$), 54.99 (NH<u>C</u>HCH$_2$), 51.19 (<u>C</u>HCH$_3$), 43.36 (NHCH$_2$), 41.82, 41.75 (CH<u>C</u>H$_2$CH and <u>C</u>H$_2$CH(OH)$_2$), 25.84 (<u>C</u>H(CH$_3$)$_2$), 23.46, 21.84 (CH(<u>C</u>H$_3$)$_2$) and 17.47 (CH<u>C</u>H$_3$); *m/z* (EI) 452 (7%, $[M]^+$), 434 (7, $[M - H_2O]^+$), 409 (47, $[M - CHO]^+$) and 394 (100, $[M - NHCH_2CHO]^+$).

## *3.0.2.75 Tertiary-butoxycarbonyl-(2S)-leucyl (2S)-alanyl glycyl glycinal dimethyl acetal (113)*

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl-alaninate (**37**) using tertiary-butoxycarbonyl-(2S)-leucyl (2S)-alanine (**43**) (2.71 g, 8.96 mmol) and glycylaminoacetaldehyde dimethyl acetal (**121**) (1.45 g, 8.96 mmol) to give a white solid which was recrystallised from methanol / ether to give white crystals(4.00 g, 83%). m.p. 156-158°C; (Found: C, 54.05; H,8.6; N 12.55. $C_{20}H_{38}N_4O_7$ requires C, 53.78; H, 8.58; N, 12.55%); *m/z* (Found: $[M + H]^+$, 447.2820. $C_{20}H_{39}N_4O_7$ requires 447.2819); $[\alpha]_D$ - 26.5 (*c* 0.22 in MeOH); $\upsilon_{max}$ (Nujol)/cm$^{-1}$ 3340, 3280, 3090 (3 x NH), 1710 ($CO_2NH$), 1690, 1630 (2 x CONH); $\delta_H$ (400 MHz, $C^2HCl_3$) 7.21 (1H, t, CHCH$_3$CON<u>H</u>), 6.75 (1H, d, $J_{CH}$ 6.68, CHCH$_3$N<u>H</u>), 6.61 (1H, t, CH$_2$CON<u>H</u>), 5.05 (1H, d, $J_{CH}$ 5.8, $CO_2N\underline{H}$), 4.41 (2H, t, and m, $J_{CH}$ 5.52, C<u>H</u>(OCH$_3$)$_2$ and C<u>H</u>CH$_2$CH(CH$_3$)$_2$), 4.08 (1H, m, C<u>H</u>CH$_3$), 4.00 (1H, dd, $J_{NH}$ 6.0, $J_{CH}$ 16.9, NHC<u>H</u>$_a$H$_b$CO), 3.86 (1H, dd, $J_{NH}$ 5.4, $J_{CH}$ 16.9, NHCH$_a$<u>H</u>$_b$CO), 3.55 (2H, m, C<u>H</u>$_2$CH(OCH$_3$)$_2$), 3.39 (6H, m, OC<u>H</u>$_3$), 1.59

(2H, m, C$\underline{H}_2$CH(CH$_3$)$_2$), 1.46 (1H, m, C$\underline{H}$(CH$_3$)$_2$), 1.42 (9H, s, C(C$\underline{H}_3$)$_3$), 1.40 (3H, d, $J_{CH}$ 7.2, CHC$\underline{H}_3$), 0.93 (6H, t, CH(C$\underline{H}_3$)$_2$); $\delta_C$ (74.8 MHz; $C^2HCl_3$) 172.95, 172.39, 168.90 (3 x CONH), 155.99 ($CO_2$NH), 102.24 (C(OCH$_3$)$_2$), 80.36 (C(CH$_3$)$_3$), 53.98, 53.83 (2 x OCH$_3$), 53.43 (NHCHCH$_2$), 49.23 (CHCH$_3$), 42.95 (NHCH$_2$CO), 41.07, (CH$_2$CH(CH$_3$)$_2$), 40.84 (CH2CH(OCH$_3$)$_2$), 28.17 (C(CH$_3$)$_3$), 24.63 (CH(CH$_3$)$_2$), 22.86, 21.74 (CH(CH$_3$)$_2$, 17.97 (CHCH$_3$); *m/z* (FAB) 447 (1%, $[M + H]^+$), 429 (4, $[M - OH]^+$), 415 (37, $[M - OMe]^+$), 341 (100, $[M - H - NHCH_2CH(OCH_3)_2]^+$), 315 (10, $M + H - OMe - CO_2C(CH_3)_3]^+$) and 285 (2, $[M - NHCH_2CH(OCH_3)_2 - C(CH_3)_3]^+$).

### *3.0.2.76 Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl amino acetaldehyde dimethyl acetal (114)*

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl (2S)-alaninate (**37**) using carbobenzyloxy-(2S)-leucyl (2S)-alanine (**29**) (841 mg, 2.50 mmol) and glycyl-aminoacetaldehyde dimethyl acetal (**121**) (405 mg, 2.50 mmol) to give a white solid which was recrystallised from acetonitrile / 40/ 60 petrol to give white crystals (894 mg, 75%), m.p. 149-151°C; (Found: C, 57.15; H, 7.5; N, 11.45. $C_{23}H_{36}N_4O_7$ requires C, 57.47; H, 7.55; N, 11.66%); *m/z* (Found: $[M + H]^+$, 481.2680. $C_{23}H_{37}N_4O_7$ requires 481.2662); $[\alpha]_D$ - 31.5 (*c* 0.17 in MeOH); $\upsilon_{max}$ (Nujol)/cm$^{-1}$ 3330, (NH), 3091 (aliphatic CH), 1720, 1680, 1660, 1630 (4 x CO); $\delta_H$ (300 MHz, $^2H_3CO^2H$) 7.51 (5H, m, $C_6H_5$), 5.27 (2H, s, $C_6H_5C\underline{H}_2$), 4.63 (1H, t, $J_{CH_2}$ 5.48, C$\underline{H}$(OCH$_3$)$_2$), 4.47 (1H, q, $J_{CH_3}$ 7.14, C$\underline{H}$CH$_3$), 4.34 (1H, t, $J_{CH_2}$ 7.51, C$\underline{H}$CH$_2$), 4.10 (1H, d, $J_{CH_b}$ 16.89, C$\underline{H}_a$H$_b$CO), 3.95 (1H, d, $J_{CH_a}$ 16.89, CH$_a\underline{H}_b$CO), 3.53 (6H, s, OCH$_3$), 3.50 (2H, d, $J_{CH}$ 5.01, C$\underline{H}_2$C(OCH$_3$)$_2$), 1.90 (1H, m, C$\underline{H}$(CH$_3$)$_2$), 1.76 (2H, t, C$\underline{H}_2$CH(CH$_3$)$_2$), 1.56 (3H, d, $J_{CH}$ 7.14, CHC$\underline{H}_3$), 1.14 (6H, t, CH(C$\underline{H}_3$)$_2$); $\delta_C$ (50.3 MHz, $^2H_3CO^2H$ ) 175.97, 175.58, 171.99,

159.09 (CO), 138.42 (*i*-$C_6H_5$), 129.77 (*m*-$C_6H_5$), 129.34 (*p*-$C_6H_5$), 103.92 ($\underline{C}H(OCH_3)_2$), 68.11 ($C_6H_5\underline{C}H_2$), 55.34 ($NH\underline{C}HCH_2$), 54.81, 54.57 (2 x $OCH_3$), 51.17 ($\underline{C}HCH_3$), 43.72 ($NH\underline{C}H_2CO$), 42.43 ($\underline{C}H_2CH(OCH_3)_2$), 42.13 ($CH\underline{C}H_2CH$), 26.13 ($\underline{C}H(CH_3)_2$), 23.77, 22.16 ($CH(\underline{C}H_3)_2$) and 17 .70 ($CH\underline{C}H_3$); *m/z* (FAB) 503 (48%, $[M + Na]^+$), 481 (16, $[M + H]^+$) and 449 (100, $[M - OMe]^+$).

## *3.0.2.77 N-(Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl) 1-aminoprop-2-ene (115)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using allyl amine (58 $mm^3$, 1.02 mmol) to give a white solid which was recrystallised from acetone / ether to give white crystals (132 mg, 34%), m.p. 114°C; (Found: C, 61.05; H, 7.6; N, 12.85. $C_{22}H_{32}N_4O_5$ requires C, 61.1; H, 7.45; N, 12.95%); *m/z* (Found: $[M + H]^+$, 433.244). $C_{22}H_{33}N_4O_5$ requires 433.2451); $[\alpha]_D$ -36.0 (*c* 0.05 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3280 (NH), 3050 (aromatic CH), 2890, 2850, 2800 (3 x aliphatic CH), 1710 (OCONH) and 1680, 1660, 1630 (3 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.38 (5H, m, $C_6H_5$), 5.43 (2H, dt, *J* 1.55, $C\underline{H}_2CH$), 5.23 (1H, tt, *J* 5.87, $CH_2C\underline{H}$), 5.13 (1H, s, $C\underline{H}_2C_6H_5$), 4.30 (1H, q, $J_{CH_3}$ 7.0, $C\underline{H}CH_3$), 4.17 (1H, quin, *J* 4.70, 2.76, $NHC\underline{H}CH_2$), 3.97 (1H, d, $J_{CH_b}$ 16.85, $NHC\underline{H}_aH_b$), 3.85 (2H, dt, $J_{CH}$ 5.34, $J_{CH_2}$ 1.82, $C\underline{H}_2CH(CH_2)$), 3.80 (1H, d, $J_{CH_a}$ 17.05, $NHCH_a\underline{H}_b$), 1.75 (1H, sep, $J_{CH_3}$ 6.67, $C\underline{H}(CH_3)_2$), 1.60 (2H, t, $J_{CH_2}$ 7.13, 7.44, $CHC\underline{H}_2CH$), 1.41 (3H, d, $J_{CH}$ 7.17, $CHC\underline{H}_3$), 1.00 (3H, d, $CH(C\underline{H}_3)_2$) and 0.97 (3H, d, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 175.80, 175.36, 171.35 (3 x NHCO), 158.81 ($CO_2NH$), 138.13 (*i*-$C_6H_5$), 135.18 ($(CH_2)\underline{C}HCH_2$), 129.47 (*m*-$C_6H_5$), 129.03 (*p*-$C_6H_5$), 128.82 (*o*-$C_6H_5$), 116.22 ($(\underline{C}H_2)CHCH_2$), 67.83

($C_6H_5\underline{C}H_2$), 55.08 ($NH\underline{C}HCH_2$), 50.99 ($\underline{C}HCH_3$), 43.53 ($\underline{C}H_2CH(CH_2)$), 42.68 ($NHCH_2$), 41.80 ($CH\underline{C}H_2CH$), 25.82 ($\underline{C}H(CH_3)_2$), 23.42, 21.89 ($CH(\underline{C}H_3)_2$) and 17.28 ($CH\underline{C}H_3$); *m/z* (CI) 433 (14%, $[M + H]^+$) and 325 (100, $[M - C_6H_5CH_2O]^+$).

## *3.0.2.78 Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl 3-aminopropan-2-one (116)*

This compound was prepared in a manner identical to that described for Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycylaminoacetaldehyde (**112**) using N-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl) 1-amino 2-methyl prop-2-ene (**118**) (35 mg, 78.4 μmol) to give an, unstable white solid (31 mg, 88%), (Found: $[M + H]^+$, 449.2406. $C_{22}H_{33}N_4O_6$ requires 449.2400); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3300 (NH), 2950, 2920, 2850 (3 x aliphatic CH), 1710, 1680, 1650, 1640 (4 x CO); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.38 (5H, m, $C_6H_5$), 5.15 (1H, d, $J_{CH_b}$ 12.67, $C\underline{H}_aH_bC_6H_5$), 5.11 (1H, d, $J_{CH_a}$ 13.61, $CH_a\underline{H}_bC_6H_5$), 4.33 (1H, m, $C\underline{H}CH_3$), 4.16 (1H, m, $NHC\underline{H}CH_2$), 4.07 (2H, s, $CH_2COMe$), 4.00 (1H, d, $J_{CH_b}$ 16.89, $NHC\underline{H}_aH_b$), 3.86 (1H, d, $J_{CH_a}$ 16.85, $NHCH_a\underline{H}_b$), 2.16 (3H, s, $COCH_3$), 1.74 (1H, sep, $J_{CH_3}$ 6.60, $C\underline{H}(CH_3)_2$), 1.59 (2H, t, $J_{CH_2}$ 7.00, $CHC\underline{H}_2CH$), 1.42 (3H, d, *JCH* 7.10, $CHC\underline{H}_3$), 0.99 (3H, d, $J_{CH}$ 6.38, $CH(C\underline{H}_3)_2$) and 0.97 (3H, d, $J_{CH}$ 6.29, CH(C<u>H</u>3)2); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 205.78 (COMe), 175.80, 175.36, 171.96 (3 x NHCO), 158.76 (OCONH), 138.14 (*i*-$C_6H_5$), 129.46 (*m*-$C_6H_5$), 129.01 (*p*-$C_6H_5$), 128.82 (*o*-$C_6H_5$), 67.80 ($C_6H_5\underline{C}H_2$), 54.99 ($NH\underline{C}HCH_2$), 50.86 ($\underline{C}HCH_3$), 50.23 ($\underline{C}H_2COMe$), 43.39 ($NHCH_2$), 41.82 ($CH\underline{C}H_2CH$), 25.82 ($\underline{C}H(CH_3)_2$), 23.46, 21.83 ($CH(\underline{C}H_3)_2$) and 17.40 ($CH\underline{C}H_3$); *m/z* (CI) 449 (100%, [M + H]), 431 (10, $[M - OH]^+$), 393 (8, $[M + 2H - CH_2COCH_3]^+$) and 336 (10, $[M + 2H - CH_2CONHCH_2COCH_3]^+$).

## *3.0.2.79 N-(Carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl) 1-amino 2-methyl prop-2-ene (118)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl-(2S)-alanyl-glycyl-glycinate (**65**) using 1-amino 2-methyl prop-2-ene hydrochloride (90 mg, 1.02 mmol) to give a white solid which was recrystallised from acetone / ether to give white crystals (146 mg, 32%), m.p. 148°C; (Found: C, 61.85; H, 7.75; N, 12.50. $C_{23}H_{35}N_4O_5$ requires C, 61.85; H, 7.65; N, 12.55%); *m/z* (Found: $[M + H]^+$, 447.2609. $C_{23}H_{36}N_4O_5$ requires 447.2607); $[\alpha]_D$ - 26.92 (*c* 0.26 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3300 (NH), 3100, 3080 (2 x aromatic CH), 2950, 2940, 2880 (3 x aliphatic CH), 1710 (OCONH) and 1680, 1660, 1640 (3 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.38 (5H, m, $C_6H_5$), 5.12 (2H, s, $C\underline{H}_2C_6H_5$), 4.85, 4.82 (2H, ds, CMe($C\underline{H}_2$)), 4.30 (1H, q, $J_{CH_3}$ 7.16, $C\underline{H}CH_3$), 4.17 (1H, t, *J* 7.52, NHC$\underline{H}$CH$_2$), 3.96 (1H, d, $J_{CH_b}$ 16.88, NHC$\underline{H}_a$H$_b$CMe), 3.81 (1H, d, $J_{CH_a}$ 17.20, NHCH$_a\underline{H}_b$CMe), 3.83 (1H, d, $J_{CH_b}$ 15.73, NHC$\underline{H}_a$H$_b$CO), 3.74 (1H, d, $J_{CH_a}$ 15.78, NHCH$_a\underline{H}_b$CO), 1.75 (4H, m, $CH_2CC\underline{H}_3$ and C$\underline{H}$(CH$_3$)$_2$), 1.59 (2H, t, $J_{CH_2}$ 6.74, CHC$\underline{H}_2$CH), 1.41 (3H, d, $J_{CH}$ 7.15, CHC$\underline{H}_3$), 0.99 (3H, d, *J* 6.71, CH(C$\underline{H}_3$)$_2$) and 0.97 (3H, d, *J* 6.76, CH(C$\underline{H}_3$)$_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 175.85, 175.42, 171.46 (3 x NHCO), 158.83 ($CO_2$NH), 142.90 (($\underline{C}H_2$)CHCH$_3$), 138.14 (*i*-$C_6H_5$), 129.48 (*m*-$C_6H_5$), 129.04 (*p*-$C_6H_5$), 128.84 (*o*-$C_6H_5$), 111.19 ((CH$_2$)$\underline{C}$HCH$_3$), 67.84 ($C_6H_5\underline{C}H_2$), 55.07 (NH$\underline{C}$HCH$_2$), 51.02 ($\underline{C}$HCH$_3$), 45.75 ($\underline{C}H_2$CH(CH$_2$)), 43.55 (NHCH$_2$), 41.77 (CH$\underline{C}H_2$CH), 25.81 (C$\underline{C}$H$_3$), 23.45 ($\underline{C}$H(CH$_3$)$_2$), 21.86, 20.46 (CH($\underline{C}$H$_3$)$_2$) and 17.26 (CH$\underline{C}$H$_3$); *m/z* (CI) 447 (100%, [M + H]) and 339 (53, $[M - C_6H_5CH_2O]^+$).

### *3.0.2.80 Carbobenzyloxy-glycyl glycinal dimethyl acetal (120)*

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl-alaninate (**37**) using carbobenzyloxy-glycine (**34**) (3.51 g, 16.8 mmol) and aminoacetaldehyde dimethyl acetal (**119**) (1.77 g, 16.8 mmol) and THF instead of DMF to give a yellow oil, *m/z* (Found: $[M + H]^+$, 297.1450) $C_{14}H_{21}N_2O_5$ requires 297.1450); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3400, 3330 (2 x NH), 3190, 3150, 3130 (3 x aromatic CH), 2940 , 2840 (2 x aliphatic CH), 1730 ($CO_2NH$ ) and 1670 (CONH); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.40 (5H, m, $C_6H_5$), 6.33 (1H, t, N<u>H</u>$CH_2CO$), 5.60 (1H, t, $J_{CH_2}$ 5.60, N<u>H</u>$CH_2CH$), 5.12 (2H, s, $C_6H_5C\underline{H}_2$), 4.36 (1H, t, $J_{CH_2}$ 5.13, C<u>H</u>$(OCH_3)_2$), 3.86 (2H, d, $J_{NH}$ 5.78, NHC<u>H</u>$_2CO$), 3.38 (2H, m, $J_{NH}$ 5.58, C<u>H</u>$_2CH(OCH_3)_2$) and 3.37 (6H, s, $OCH_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 169.72 (OCONH), 157.10 (CONH), 136.63 (*i*-$C_6H_5$), 129.04 (*m*-$C_6H_5$), 128.74 (*p*-$C_6H_5$), 128.58 (*o*-$C_6H_5$), 102.92 (<u>C</u>H$(OCH_3)_2$), 67.63 ($C_6H_5\underline{C}H_2$), 54.89 ($OCH_3$), 44.93 (<u>C</u>$H_2CH$) and 41.39 (NH<u>C</u>$H_2CO$);. *m/z* (CI) 297 (4%, $[M + H]^+$), 279 (4, $[M - OH]^+$), 265 (100, $[M - OMe]^+$), 189 (13, $[M - C_6H_5CH_2O]^+$), 142 (31, $[M - C_6H_5CH_2O - MeOH - CH_3]^+$) and 75 (17, $[M - C_6H_5CH_2OCONHCH_2CONHCH_2]^+$).

### *3.0.2.81 Glycyl-aminoacetaldehyde dimethyl acetal (121)*

This was prepared in an identical manner to methyl glycyl glycinate (**70**) using carbobenzyloxy-glycyl aminoacetaldehyde dimethyl acetal (**121**) (1.87 g, 6.3 mmol) to give a pale yellow oil (1.94 g, 95%), *m/z* (Found: $[M + H]^+$, 163.1083 $C_6H_{15}N_2O_3$ requires 163.1083); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3300 ($NH_2$, NH), 2940, 2840 (2 x aliphatic CH), 1660 (CONH);. $\delta_H$ (300 MHz; $C^2HCl_3$) 7.53 (1H, t, $J_{CH_2}$ 5.48, NH), 4.40 (1H, t, $J_{CH_2}$ 5.28, C<u>H</u>$(OCH_3)_2$), 3.46 (2H, s, C<u>H</u>$_2CH$), 3.41 (2H, t, $J_{NH_2}$ 5.52, $NH_2C\underline{H}_2$), 3.37 (6H, s, 2 x

$OC\underline{H}_3$), 3.17 (2H, s, $N\underline{H}_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 172.39 (CO), 102.95 ($\underline{C}(OCH_3)_2$), 54.60 ($OCH_3$), 44.60 ($\underline{C}H_2NH_2$), 40.96 ($NH\underline{C}H_2$); *m/z* (CI) 163 (10%, $[M + H]^+$), 131 (100, $[M - OMe]^+$), 116 (100, $[M - OMe - CH_3]^+$), 99 (7, $[M - OMe - MeOH]^+$), 88 (2, $[M + H - CH(OCH_3)_2]^+$), 75 (64, $[M - CH_2NHCOCH_2NH_2]^+$), 58 (7, $[M - NHCH_2CH(OCH_3)_2]^+$).

## *3.0.2.82 N,N'-Di-(carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl) 2, 4-diamino-but-2-ene aldehyde (122)*

Propene was bubble through dry dichloromethane (30 $cm^3$) for 15 min in a flame dried flask. Carbobenzyloxy-(2S)-leucyl-(2S)-alanyl-glycyl-amino acetaldehydedimethyl acetal (**112**) (152 mg, 316 mmol) dissolved in dry dichloromethane (3 $cm^3$) was added and the solution was stirred for 2 h under a propene atmosphere. The propene was displaced from solution with $N_2$ and the solvent was removed under reduced pressure to give a yellow soild which was recrystallised from MeOH / ether to give yellow crystals (62 mg, 45%), (Found: $[½M + H]^+$, 435.2248. $C_{21}H_{31}N_4O_6$ requires 435.2248); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.35 (5H, m, $C_6\underline{H}_5$), 6.59 (1H, d, $J_{CH}$ 5.32, NH), 6.50 (1H, d, $J_{CH}$ 5.88, NH), 5.92 (1H, t, $J_{CH_2}$ 4.46 =CH), 5.08 (2H, s, $C\underline{H}_2C_6H_5$), 4.91 (1H, t, $J_{CH_2}$ 7.17, $CH(OH)_2$), 4.44 (2H, m, $NHC\underline{H}CH_2$ and $C\underline{H}CH_3$), 4.25 (3H, m, $C\underline{H}_2CH{=}C$, $NHC\underline{H}_aH_b$), 4.18 (1H, d, $J_{CH_a}$ 17.31, $NHCH_a\underline{H}_b$), 3.0 (2H, s, OH), 1.84 (1H, sep, $C\underline{H}(CH_3)_2$), 161 (2H, q, $CHC\underline{H}_2CH$), 1.31 (3H, m, $CHC\underline{H}$) and 0.98 (6H, m, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz; $^2H_3CO^2H$) 174.52, 172.98, 171.39, 165.64 (4 x NHCO), 157.36 ($CO_2NH$), 138.56 (*i*-$C_6H_5$), 130.37, 129.90, 129.54, 128.95, 128.90 (5 x $C_6H_5$), 112.17 ($C{=}\underline{C}C(OH)_2$), 108.82 ($\underline{C}{=}CC(OH)_2$), 71.74 ($CH(OH)_2$), 67.08 ($C_6H_5\underline{C}H_2$), 54.85, 54.67 ($NH\underline{C}HCH_2$), 49.86, 49.10 ($\underline{C}HCH_3$), 46.65, 46.58 ($NHCH_2$), 42.62, 42.52 ($CH\underline{C}H_2CH$), 25.72, ($\underline{C}H(CH_3)_2$), 23.83, 23.66, 22.23 ($CH(\underline{C}H_3)_2$) and 18.39, 18.29

($CH\underline{C}H_3$); *m/z* (FAB) 891 (100%, $[M + H]^+$), 457 (52, $[½M + Na]^+$), 435 (5, $[½M + H]^+$) and 391 (32, $[M - CH_2CHO]^+$).

### 3.0.2.83 3-Aminopropanone hydrochloride (124)

Trifluoroacetyl 3-aminopropane (**127**) (137 mg, 0.81 mmol) was dissolved in HCl (5 $moldm^{-3}$, 3 $cm^3$) and refluxed for 2 h. The HCl was flushed from solution with $N_2$ and the solution was lyophilised to give a yellow, light sensitive, hydroscopic solid (88 mg, 91%), (Found: $[M - Cl]^+$, 74.0601. $C_3H_8NO$ requires 74.0606); $\delta_H$ (200 MHz; $^2H_2O$) 4.08 (2H, s, $CH_2$) and 2.28 (3H, s, $CH_3$); $\delta_C$ (50.3 MHz; $^2H_2O$) 206.9 (CO), 50.70 ($CH_2$) and 29.63 ($CH_3$); *m/z* (CI) 90 (8%, $[M + NH_3 - HCl]^+$) and 74 (100, $[M - Cl)]^+$).

### 3.0.2.84 Trifluoroacetyl 1-amino-3-diazo propanone (126)

This compound was prepared in a manner identical to that described for the methyl carbobenzyloxy-(2S)-leucyl (2S)-alaninate (**37**) using trifluoroacetyl glycine (**101**) (210 mg, 1.22 mmol) and dry etherial diazomethane solution (12.2 mmol).[134] The following precautions were also followed; friction free glass joints were used, the reaction mixture was excluded from light and the temperature was not allowed to rise above 0°C. No attempt was made to purify the diazoketone and after allowing the excess diazomethane to evaporate from the reaction vessel for an hour, the NMM.HCl salts were removed by filtration and the solvent was removed under reduced pressure with a cold water bath to give an unstable, reddy orange, light sensitive solid, (Found: $[M - N_2]^+$, 167.0201. $C_5H_4NO_2F_3$ requires 167.0194); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3390 (NH), 2970, 2255 ($CHN_2$), 1730 ($CO_2NH$) and 1650 ($CON_2$); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.35 (1H, s, NH), 5.42 (1H, s, $CHN_2$) and 4.13 (2H, d, $J_{NH}$

5.00, $CH_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 187.07 ($\underline{C}OCHN_2$), 157.26 (q, $J_{CF}$ 37.5, $\underline{C}OCF_3$), 115.67 (q, $J_{F_3}$ 284.9, $CF_3$), 66.75 ($CHN_2$) and 26.14 ($CH_2$); *m/z* (EI) 167 (15%, $[M - N_2]^+$), 126 (77, $[M - COCHN_2]^+$) and 69 (100, $[CF_3]^+$).

### *3.0.2.85 Trifluoroacetyl 3-amino propanone (127)*

To a stirred solution of trifluoroacetyl 3-amino propane diazoketone (**126**) (476 mg, 2.44 mmol) in dichloromethane (30 $cm^3$) was added constant boiling HI (2.60 $moldm^{-3}$, 2.35 $cm^3$, 6.1 mmol) and the reaction was stirred for 2 h, excluded from light.[134] The solvent was removed under reduced pressure to give a yellow light sensitive solid (232 mg, 56%), m.p.57-58°C; $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3400 (NH), 2400, (aliphatic CH), 1725 (CONH) and 1700 (COMe); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.25 (1H, s, NH), 4.25 (2H, d, $J_{NH}$ 4.40, $CH_2$) and 2.28 (3H, s, $CH_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 198.49 ($\underline{C}OCH_3$), 154.82 (q, $^2J_{CF}$ 37.8, $\underline{C}OCF_3$), 113.37 (q, $^1J_{F_3}$ 287.2, $CF_3$), 47.11 ($CH_2$) and 24.90 ($CH_3$); *m/z* (EI) 169 ($[M]^+$, 2%), and 126 (20, $[M - COCH_3]^+$).

### *3.0.2.86 Trifluoroacetyl 3-amino-1-chloro-propanone (128)*

To a solution of trifluoroacetyl 3-aminopropane diazoketone (**126**) (476 mg, 2.44 mmol) in dichloromethane (30 $cm^3$), was added HCl (1.00 $moldm^{-3}$, 2.68 $cm^3$, 2.68 mmol) and the reaction was stirred for 2 h, excluded from light. The solvent was removed under reduced pressure to give a yellow light sensitive solid (497 mg, 57%), m.p.53-55°C; $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3230 (NH), 1740 (CONH) and 1700 ($COCClH_2$); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.22 (1H, s, NH), 4.47 (2H, d, $J_{NH}$ 4.84, $CH_2$) and 4.20 (2H, s, $CClH_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$)

197.32 ($\underline{C}OCH_2Cl$), 157.89 (q, $J_{CF}$ 38.1, $\underline{C}OCF_3$), 116.01 (q, $J_{F_3}$ 287.2, $CF_3$), 47.44, 46.40 ($CH_2$ and $CH_2Cl$).

### *3.0.2.87 2-Bromo, carbobenzyloxy-(2S)-leucyl-(2S)-alanyl-glycyl-aminoethane (130)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using 2-bromo, aminoethane hydrochloride (164 mg, 1.02 mmol) to give a white solid which was recrystallised from acetone / ether to give white monohydrated crystals (388 mg, 70%), m.p. 107°C, (Found: C, 48.7; H, 6.5; N, 10.80. $C_{21}H_{31}N_4O_5Br$ requires C, 49.0; H, 6.25; N, 10.85%); $[\alpha]_D$ -32.8 (*c* 0.18 in MeOH); $\upsilon_{max}$ (KBr disc)/$cm^{-1}$ 3450-3200 ($H_2O$), 3100, 3050 (2 x aromatic CH), 2980, 2960 (2 x aliphatic CH), 1750 (OCONH) and 1730, 1700, 1650 (3 x CONH); $\delta_H$ (300 MHz; $_2H_3CO^2H$), 7.38 (5H, m, $C_6H_5$), 5.14 (2H, s, $C_6H_5C\underline{H}_2$), 4.38 (3H, m, *J* 5.40, $CH_2Br$ and $C\underline{H}CH_3$), 4.18 (1H, t, *J* 7.48, $NHC\underline{H}CH_2$), 4.04 (1H, d, $J_{CH}$ 17.74, $NHC\underline{H}_aH_b$), 3.97 (1H, d, $J_{CH}$ 17.92, $NHCH_a\underline{H}_b$), 3.29 (2H, t, $J_{CH_2}$ 5.00, $NC\underline{H}_2CH_2Br$), 1.75 (1H, sep, *J* 6.55, $C\underline{H}(CH_3)_2$), 1.61 (2H, d, *J* 7.19, $CHC\underline{H}_2CH$), 1.42 (3H, d, $J_{CH}$ 7.15, $CHC\underline{H}_3$), 1.00 (3H, d, $J_{CH}$ 6.79, $CH(C\underline{H}_3)_2$ and 0.97 (3H, d, $J_{CH}$ 6.78, $CH(C\underline{H}_3)_2$; $\delta_C$ (75.5 MHz, $_2H_3CO^2H$) 173.65, 173.51, 168.46 (3 x CO), 156.67 (OCONH), 135.93 (*i*-$C_6H_5$), 127.36 (*m*-$C_6H_5$), 126.93 (*p*-$C_6H_5$), 126.65 (*o*-$C_6H_5$), 65.65 ($\underline{C}H_2C_6H_5$), 60.05 ($\underline{C}HCH_2$), 53.00 ($\underline{C}HCH_3$), 48.30 ($CH_2Br$), 39.89 ($NHC\underline{H}_2$), 39.56 ($NHC\underline{H}_2CH_2$), 37.60 ($CH\underline{C}H_2$), 23.54 ($\underline{C}H(CH_3)_2$), 21.16, 19.64 ($CH(\underline{C}H_3)_2$) and 15.38 ($CH\underline{C}H_3$).

### *3.0.2.88 O-(Carbobenzyloxy-(2S)-leucyl-(2S)-alanyl-glycyl) 1-carboxy methyl 2-hydroxy but-2-ene (132)*

To a solution of carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycine (**27**) (300 mg, 0.76 mmol) was dissolved in DMF (20 $cm^3$) was added methyl γ-bromocrotonate (683 mg, 3.82 mmol) then $NaHCO_3$ (128 mg, 1.52 mmol). The solution was then stirred for 12 h.[138] The solvent was removed under reduced pressure and the crude product was purified by silica chromatography, eluting with a gradient of dichloromethane through to 1:19 MeOH / dichloromethane to give a yellow oil which was recrystallised from MeOH / ether to give white crystals (130 mg, 35%), m.p. 89°C; (Found: C, 58.7; H, 7.0; N, 8.55. $C_{24}H_{33}N_3O_8$ requires C, 58.65; H, 6.75; N, 8.55%); *m/z* (Found: $[M + H]^+$, 492.2340. $C_{24}H_{34}N_3O_8$ requires 492.3346); $[\alpha]_D$ - 41.7 (*c* 0.24 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3300 (NH), 3050, 3020 (2 x aromatic CH), 2950, 2940, 2880 (3 x aliphatic CH), 1760 ($CO_2Me$), 1740 ($CO_2NH$) and 1700, 1680, 1640 (3 x CONH); $\delta_H$ (300 MHz; $_2H_3CO_2H$) 7.36 (5H, m, $C_6H_5$), 6.99 (1H, dt, *J* 4.53, HC=C<u>H</u>$CH_2$), 6.11 (1H, dt, *J* 1.91, 15.80, <u>H</u>C=C$HCH_2$), 5.13 (1H, s, C<u>H</u>$_2C_6H_5$), 4.86 (1H d, $J_{CH_b}$ 1.88, OC<u>H</u>$_a$H$_b$), 4.85 (1H d, $J_{CH_a}$ 1.88, OCH$_a$<u>H</u>$_b$), 4.45 (1H, q, $J_{CH_3}$ 7.12, C<u>H</u>$CH_3$), 4.19 (1H, d, $J_{CH_b}$ 17.85, NHC<u>H</u>$_a$H$_b$), 4.06 (2H, s, $NHCH_2$), 4.03 (2H, s, $NHCH_2$), 3.76 (3H, s, $CH_3$), 1.73 (1H, sep, *J* 6.78, C<u>H</u>$(CH_3)_2$), 1.60 (2H, t, *J* 6.85, CHC<u>H</u>$_2$CH), 1.42 (3H, d, $J_{CH}$ 7.15, CHC<u>H</u>$_3$), 1.00 (3H, d, *J* 6.25, CH(C<u>H</u>$_3$)$_2$) and 0.97 (3H, d, *J* 6.11, CH(C<u>H</u>$_3$)$_2$); $\delta_C$ (75.5 MHz; $_2H_3CO_2H$) 175.31, 175.28, 170.54, 167.84 (4 x NHCO), 158.73 (OCONH), 143.04 (C=<u>C</u>CO), 138.15 (*i*-$C_6H_5$), 129.48 (*m*-$C_6H_5$), 129.09 (*p*-$C_6H_5$), 128.82 (*o*-$C_6H_5$), 122.56 (<u>C</u>=CCO), 67.78 ($C_6H_5$<u>C</u>$H_2$), 64.26 ($OCH_2$), 55.05 (NH<u>C</u>$HCH_2$), 52.19 ($OCH_3$), 50.12 (<u>C</u>$HCH_3$), 41.90 (<u>C</u>$H_2$CH($CH_2$)), 25.85 (<u>C</u>H$(CH_3)_2$), 23.48, 21.82 (CH(<u>C</u>H$_3$)$_2$) and 18.00 (CH<u>C</u>H$_3$); *m/z* (CI) 492 (100%, $[M + H]^+$) and 384 (9, $[M - C_6H_5CH_2O]^+$).

### *3.0.2.89 Potassium (2R,3R)-methyl 2,3-epoxysuccinate (141)*

To a stirred solution of (2R,3R)-dimethyl 2,3 epoxysuccinate (**158**) (220 mg, 1.17 mmol) in dry MeOH (3 $cm^3$) was added KOH (85%, 76 mg, 1.35 mmol) at 0°C. The suspension was stirred for 2 h and the product was precipitated from solution by the addition of ether (10 $cm^3$).[144] The suspension was centrifuged (4000 rpm) for 10 min and the supernatant was decanted from the solid (203 mg, 94.3%), m.p. 148°C (decomp.); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3022, 2980 (2 x aliphatic CH) and 1750 (CO); $\delta_H$ (200 MHz; $^2H_2O$), 3.73 (3H, s, $OCH_3$), 3.55 (1H, d, $J_{CH}$ 2.07, CH) and 3.44 (1H, d, $J_{CH}$ 1.95, CH); $\delta_C$ (50.3 MHz; $^2H_2O$) 176.27, 173.27 (2 x CO), 56.94, 56.05 (2 x CH) and 54.43 ($OCH_3$).

### *3.0.2.90 N-(tertiary-Butoxycarbonyl-(2S)-leucyl-(2S)-alanyl ethylenediamine (142)*

This compound was prepared in a manner identical to that described for the methyl glycyl glycinate (**70**) using tertiary-butoxycarbonyl-(2S)-leucyl (2S)-alanyl aminoacetonitrile (**170**) (220 mg, 0.647 mmol) and carbon / palladium catalyst (10%, 44 mg, 20% Pd by mass). The reaction was complete after 2 days to give a white solid (219 mg, 98%), m.p. 99-102°C; $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3500, 3400, 3200 (3 x NH), 2900, 2850, 2790 (3 x aliphatic CH), 1700, 1690 (OCONH and CONH); $\delta_C$ (75.5 MHz; $C^2H_3O^2H$) 175.61, 175.08 (2 x NHCO), 158.13 (OCONH), 80.70 ($\underline{C}(CH_3)_3$), 54.68 ($\underline{C}HCH_2$), 50.63 ($\underline{C}HCH_3$), 41.93 ($NHCH_2$), 39.44 ($\underline{C}H_2CH$), 30.72 ($CH_2NH_2$), 28.74 ($C(\underline{C}H_3)_3$), 25.87 ($\underline{C}H(CH_3)_2$), 23.51, 21.89 ($CH(\underline{C}H_3)_2$) and 17.99 ($NHCH\underline{C}H_3$); *m/z* (CI) 345 (4%, $[M + H]^+$), 327 (14, $[M + H - H_2O]^+$), 245 (12, $[M + 2H - CO_2C(CH_3)_3)]^+$) and 71 (100, $[C_3H_5NO]^+$).

### 3.0.2.91 Diethylaminoacetonitrile phosphonate (143)

Aminoacetonitrile hydrochloride (**85**) (2.00 g, 21.6 mmol) was dispersed in $CCl_4$ (30 $cm^3$) and triethylamine (6.02 $cm^3$, 43.2 mmol) was added. At 0°C diethylphosphite dissolved in $CCl_4$ (10 $cm^3$) was added dropwise over 20 min and a white ppt formed. The reaction was stirred for 2 h and the solvent was removed under reduced pressure to give a brown solid.[145] The solid was partitioned between ethyl acetate (100 $cm^3$) and water (100 $cm^3$), washed with water (2 x 50 $cm^3$), brine (saturated, 40 $cm^3$), dried ($MgSO_4$) and the solvent was removed under reduced pressure to give a yellow oil (2.47 g, 60%), $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3570, 3440, 3200 (3 x NH), 2990, 2970, 2900 (3 x aliphatic CH) and 2200 (CN); $\delta_H$ (200 MHz; $C^2HCl_3$), 4.65 (1H, dt, NH), 4.09 (4H, dt, *J* 7.53, 7.08, $CH_2O$), 3.80 (2H, dd, *J* 7.46, 6.35, $CH_2CN$) and 1.33 (6H, dt, *J* 7.04, 7.08 $CH_3$); $\delta_C$ (75.5 MHz; $C^2HCl_3$) 117.89 (CN), 63.54 (d, *J* 5.3, $CH_2O$), 30.19 ($\underline{C}H_2CN$) and 16.56 (d, *J* 7.2, $CH_3$); *m/z* (CI) 193 (100%, $[M + H]^+$).

### 3.0.2.92 N-(Diethyl phosphonate) ethylenediamine (146)

This compound was prepared in a manner identical to that described for the methyl glycyl glycinate (**70**) using diethylaminoacetonitrile phosphonate (**145**) (150 mg, 0.781 mmol) and carbon / palladium catalyst (10%, 23 mg, 15% Pd by mass). The reaction was complete after 3 h to give a white solid (137 mg, 90%), $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3500, 3300 (2 x NH) and 3000, 2870, 2810 (3 x aliphatic CH); $\delta_H$ (300 MHz; $C^2HCl_3$), 4.04 (4H, m, $CH_2O$), 3.70 (1H, s, NH), 3.20 (2H, s, $NH_2$), 2.99 (2H, q, *J* 2.99, $C\underline{H}_2NH$) 2.71 (2H, t, *J* 5.55 $CH_2NH_2$), and 1.28 (6H, dt, *J* 7.07, 7.08 $CH_3$); $\delta_C$ (75.5 MHz; $C^2HCl_3$) 62.25 (d, *J* 5.09, $CH_2O$), 49.77 ($CH_2NH$), 40.45 (d, *J* 7.19 $NHCH_2$) and 16.15 (d, *J* 7.19, $CH_3$); *m/z* (CI) 197 (100%, $[M + H]^+$).

### 3.0.2.93 N-(Diphenylphosphine)-amino acetonitrile (148)

Aminoacetonitrile hydrochloride (**85**) (200 mg, 2.16 mmol) was dispersed in dry dichloromethane (20 $cm^3$) and at 0°C diphenylphosphinyl chloride (0.41 $cm^3$, 2.16 mmol) was added dropwise over 10 min. The reaction was stirred for 2 h and the solvent removed under reduced pressure to give a white solid.[146] The solid was partitioned between water (20 $cm^3$). The ethyl acetate was washed further with water (2 x 20 $cm^3$), citric acid solution (10%, 10 $cm^3$), brine solution (saturated, 10 $cm^3$) and dried ($MgSO_4$). The solvent was removed under reduced pressure to give a white solid which was recrystallised from MeOH / ether to give white needle shaped crystals (146 mg, 52%), (Found: $[M + H]^+$, 257.0838. $C_{14}H_{14}N_2O_1P_1$ requires 257.0844); $\delta_H$ (200 MHz; $^2H_3CO^2H$) 7.90, 7.55 (5H, dm, $C_6H_5$) and 3.93 (2H, d, $J_P$ 12.29, $CH_2$); $\delta_C$ (50.3 MHz; $^2H_3CO^2H$) 134.84, 134.20, 130.47, 133.46, 133.27, 130.39, 130.21, 132.75, 132.68, 132.83 (12 x $C_6H_5$), 119.29 (CN) and 29.35 ($CH_2$); *m/z* (CI) 257 (100%, $[M+H]^+$).

### 3.0.2.94 N-(Diphenyl phosphine oxide)-ethylenediamine (149)

This compound was prepared in a manner identical to that described for the methyl glycyl glycinate (**70**) using N-(diphenylphosphine)-amino acetonitrile (**148**) (1.00 g, 3.91 mmol) and carbon / palladium catalyst (10%, 200 mg 20% Pd by mass). The reaction was complete after 2 days to give a white solid (1.00 g, 98%), m.p. 58-59°C; $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3400, 3200 (2 x NH), 3010 (aromatic CH) and 2920, 2750 (2 x aliphatic CH); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.85, 7.40 (5H, dm, $C_6H_5$), 3.04 (2H, dt, $J_P$ 11.19, $J_{CH_2}$ 6.41, $CH_2$) and 2.74 (2H, t, $J_{CH_2}$ 5.36 $C\underline{H}_2NH_2$); $\delta_C$ (75.5 MHz; $C^2HCl_3$) 133.39, 132.18, 132.05, 131.67, 128.57, 128.4 (12 x $C_6H_5$), 49.82 (d, $J_P$ 4.95, $NHCH_2$) and 40.39 ($NH_2CH_2$); *m/z* (CI) 261 (27%, $[M]^+$) and 218 (100, $[M + H - (CH_2)_2NH_2]^+$).

### 3.0.2.94 N-(Triphenyl methyl)-glycine (152)

Under argon, glycine (270 mg, 3.6 mmol) was dispersed in dry dichloromethane (30 $cm^3$) in a 2 necked flask fitted with a stopper and a reflux condenser. Then at 0°C trimethylsilyl chloride (1.37 $cm^3$, 10.8 mmol) and dry NMM (1.18 $cm^3$, 10.8 mmol) were added and the mixture was refluxed until the glycine dissolved (15 min). Triphenylmethyl chloride (1.00 g, 3.6 mmol) dissolved in warm dry dichloromethane (20 $cm^3$) was added along with dry NMM (0.39 $cm^3$, 3.6 mmol) and the solution was refluxed for 2 h. MeOH (5 $cm^3$) was added and the reaction was stirred for a further 1 h before filtering and removing the solvent under reduced pressure and the resultant pale yellow solid was recrystallised from MeOH to give a white crystals (688 mg, 60%), m.p. 179°C (lit.,[147] 178-179°C); (Found: C, 79.3; H, 5.85; N, 4.30. Calc. for $C_{21}H_{19}N_1O_2$ C, 76.45; H, 6.05; N, 4.40%); *m/z* (Found: $[M]^+$, 317.1409. Calc. for $C_{21}H_{19}N_1O_2$ 317.1411); $\upsilon_{max}$ (chloroform)/$cm^{-1}$ 3060, 3010 (2 x aromatic CH), 2830 (aliphatic CH) and 1640 ($CO_2H$); $\delta_H$ (200 MHz; ($C^2HCl_3$) 7.32 (15H, m, $C_6H_5$), 5.60 (2H, s, NH and OH) and 3.19 (2H, s, $CH_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 174.91 ($CO_2H$), 144.44 (*i*-$C_6CH_5$), 129.07 (*m*-$C_6H_5$), 128.69 (*p*-$C_6H_5$), 127.60 (*o*-$C_6H_5$), 72.19 (($C_6H_5)_3\underline{C}$) and 46.72 ($NCH_2$); *m/z* (EI) 317 (10%, $[M]^{\cdot}$), 299 (15, $[M - H_2O]^+$), 258 (30, $[M - CH_2CO_2H]^+$) and 240 (100, $[M - C_6H_5]^+$).

### 3.0.2.95 Methyl N-(triphenyl methyl)-glycinate (153)

Triphenylmethyl-glycine (**152**) (1.00 g, 3.15 mmol) was dissolved with heating in MeOH (100 $cm^3$) and at 0°C an etherial diazomethane solution (30 $cm^3$, 0.35 $moldm^{-3}$) was added until a yellow colour persisted. The solution was stirred for 2 h and acetic acid was added dropwise until the yellow solution went colourless.[148] The solvent was removed under reduced pressure to give a cream solid which was recrystallised from MeOH / ether to give

white needle shaped crystals (1.02 g, 98%), m.p. 99°C; (Found: C, 79.55; H, 6.45; N, 4.14. $C_{22}H_{21}NO_2$ requires C, 79.75; H, 6.4; N, 4.25%); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3400, 3300 (2 x NH), 3020, 3000 (2 x aromatic CH), 2850, 2800 (2 x aliphatic CH) and 1740 (CO); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.45 (15H, m, $C_6H_5$), 3.62 (3H, s, $OCH_3$), 3.14 (2H, s, $CH_2$) and 2.45 (1H, s, NH); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 173.45 (CO), 145.83 (*i*-$C_6CH_5$), 129.13 (*m*-$C_6H_5$), 128.46 (*p*-$C_6H_5$), 127.03 (*o*-$C_6H_5$), 71.18 (($C_6H_5)_3\underline{C}$), 52.29 ($OCH_3$) and 46.28 ($NCH_2$); *m/z* (CI) 332 (0.2%, $[M + H]^+$) and 243 (100, $[M - NHCH_2CO_2Me]^+$).

### *3.0.2.96 N-(Triphenylmethyl)-glycinamide (154)*

Ammonia (30 $cm^3$) was condensed into a 3-necked flask under $N_2$ at -78°C and sodium (33 mg, 1.53 mmol) was added in small pieces. Iron(III)chloride hexahydrate (5 mg) was added and the solution was stirred until the blue solution went grey (1h). Methyl N-(triphenyl methyl)-glycinate (**153**) (300 mg, 0.905 mmol) was added and stirred for a further 1 h.[149] The ammonia was allowed to evaporated and the solid residue was extracted into MeOH (2 x 40 $cm^3$), filtered and the solvent removed under reduced pressure to give a pale yellow solid which was dissolved in water (10 $cm^3$) washed with ethyl acetate (5 $cm^3$) and recrystallised from the aqueous phase to give white crystals (200 mg, 70%), m.p. 179°C; (Found: C, 79.9; H, 6.35; N, 8.85. $C_{21}H_{20}N_2O_1$ requires C, 79.7; H, 6.35; N, 8.85%); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3410, 3300 (2 x NH), 3200 (aromatic CH), 2850, 2830 (2 x aliphatic CH) and 1680, 1640 (2 x CO); $\delta_H$ (200 MHz; $^2H_3CO^2H$) 7.32 (15H, m, $C_6H_5$) and 2.97 (2H, s, $CH_2$); $\delta_C$ (50.3 MHz; $^2H_3CO^2H$) 179.08 (CO), 145.33 (*i*-$C_6CH_5$), 128.51 (*m*-$C_6H_5$), 127.51 (*p*-$C_6H_5$), 129.04 (*o*-$C_6H_5$), 77.00 (($C_6H_5)_3\underline{C}$) and 48.29 ($NCH_2$); *m/z* (CI) 317 (0.1%, $[M + H]^+$) and 243 (100, $[M + H - NHCH_2CO_2NH_2]^+$).

### *3.0.2.97 (2R,3S)-Dimethyl threo-2-acetoxy-3-bromosuccinate (156)*[155]

HBr in acetic acid (30% wt / vol, 120 $cm^3$, 0.87 mol) was added dropwise to (2R, 3R)-dimethyl tartrate (50 g, 0.28 mol) at 0°C over ½ h and stirred for 2 h at room temperature. The solution was poured onto ice (250 $cm^3$) and extracted into ether (3 x 100 $cm^3$), washed with brine (saturated, 100 $cm^3$) and dried ($MgSO_4$). The solvent was removed under reduced pressure to give an orange oil.[154] Kugelrohr distillation gave compound (**156**) as a straw coloured oil (11.59 g, 32%), b.p. 136°C at 1.5 mmHg (lit.,[155] 83-86C at 0.1 mmHg); $\delta_H$ (200 MHz; $C^2HCl_3$) 5.59 (1H, d, $J_{CH}$ 5.5 CHOAc), 4.79 (1H, d, $J_{CH}$ 5.5 CHBr), 3.81, 3.79 (3H, ds, $OCH_3$) and 2.17 (6H, s, $CCH_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 169.92, 167.39, 166.77 (3 x CO), 72.99 (CHOAc), 54.04 ($COCH_3$), 43.67 ($OCH_3$) and 20.91 ($CCH_3$); *m/z* (CI) 285, 283 (24%, 25, bromine isotopes, $[M + H]^+$), 145 (82, $[M + H - Br - CO_2CH_3]^+$) and 99, 97 (31, 100, bromine isotopes, $[BrOH_2]^+$).

### *3.0.2.98 (2R,3S)-dimethyl threo-2-hydroxy-3-bromosuccinate (157)*[155]

(2R,3S)-Dimethyl threo-2-acetoxy-3-bromosuccinate (**156**) (10 g, 36.0 mmol) was dissolved in dry MeOH (40 $cm^3$) and HBr in acetic acid (30% wt / vol, 12.7 $cm^3$, 1.37 mol) was added and the solution was refluxed under $N_2$ for 2 h. The solvent was removed under reduced pressure to give a red oil.[154] Kugelrohr distillation give compound (**157**) as straw coloured oil (3.15 g, 31%), b.p. 145°C at 1.4 mmHg (lit.,[155] 80-82°C at 0.03 mmHg); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3470 (OH), 2960 (aliphatic CH) and 1740 (CO) (lit.,[154] 3600-3100, 2985 and 1750); $\delta_H$ (200 MHz; $C^2HCl_3$),4.71 (1H, d, $J_{CH}$ 4.7 CHOH), 4.66 (1H, d, $J_{CH}$ 4.6 CHBr), 3.82, 3.80 (6H, ds, $OCH_3$) and 3.55 (1H, s, OH); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 171.11, 167.81 (2

x CO), 72.96 (CHOH) 54.04 (CHBr) and 53.61, 47.04 (2 x $CH_3$); *m/z* (CI) 243, 241 (97, 100, bromine isotopes, $[M + H]^+$) and 229, 227 (15, 14, bromine isotopes, $[M - CH_3]^+$).

### 3.0.2.99 (2R,3R)-Dimethyl 2,3-epoxysuccinate (158)[155]

A solution of (2R,3S)-dimethyl threo-2-hydroxy-3-bromosuccinate (**157**) (4.5 g, 19.0 mmol) in dry MeOH (20 $cm^3$) was added dropwise to a stirred solution of sodium (0.48 g, 20.9 mmol) in dry MeOH (20 $cm^3$), under nitrogen at 0°C over 20 min. The solution was stirred at 0°C for 2 h and the excess sodium was quenched with a drop of acetic acid. The solvent was removed under reduced pressure using a cold water bath and the white solid was partitioned between water (50 $cm^3$) and ethyl acetate (50 $cm^3$).[154] The ethyl acetate layer was washed with brine (saturated, 20 $cm^3$), dried ($MgSO_4$) and the solvent removed under reduced pressure to give a white solid which was recrystallised from MeOH to give white needle shaped crystals (1.75 g, 58%), m.p. 72°C (lit.,[155] 77-78°C); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3070, 3010, 2920 (3 x aliphatic CH) and 1740 (CO); $\delta_H$ (200 MHz; $C^2HCl_3$) 3.68 (2H, s, CH) and 3.81 (6H, s, $OCH_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 197.95 (CO), 53.87 ($OCH_2$) and 52.34 ($OCH_3$); *m/z* (CI) 161 (100%, $[M + H]^+$) and 145 (7, $[M - CH_3]^+$).

### 3.0.2.100 Potassium (2R,3R)-ethyl 2,3 epoxysuccinate (159)[144]

This compound was prepared in a manner identical to that described for the potassium (2R,3R)-methyl 2,3-epoxysuccinate (**117**) using (2R,3R)-diethyl 2,3-epoxysuccinate (**162**) (3.00 g, 16.0 mmol), dry ethanol and KOH (85%, 1.03 g, 18.4 mmol) to give a white solid (1.39 g, 44%), m.p. 133°C (decomp.), (lit.,[144] 159°C decomp.); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3040, 3000 (2 x aliphatic CH) and 1760, 1750 (2 x CO); $\delta_H$ (200 MHz; $^2H_2O$), 4.20 (2H, q, $J_{CH_3}$ 7.18, $CH_2$), 3.45 (1H, d, $J_{CH}$ 2.16, CH), 3.45 (1H, d, $J_{CH}$ 2.16, CH) and 1.22 (3H, t, $J_{CH_2}$

7.15 $CH_3$); $\delta_C$ (50.3 MHz; $^2H_2O$) 171.51, 168.11 (2 x CO), 61.24 ($OCH_2$), 49.87, 52.22 (2 x CH) and 11.24 ($CH_3$).

### 3.0.2.101 (2R,3S)-Diethyl threo-2-acetoxy-3-bromosuccinate (160)[154]

This compound was prepared in a manner identical to that described for the (2R,3S)-dimethyl threo-2-acetoxy-3-bromosuccinate (**156**) using (2R,3R)-diethyl tartrate (30 g, 0.145 mol) to give an orange oil which was purified by silica chromatography eluting with ether to give a straw coloured oil (22.6 g, 50%), $\upsilon_{max}$ (thin film)/$cm^{-1}$ 2990 (aliphatic CH) and 1750 (CO) (lit.,[144] 2985, 1756); $\delta_H$ (200 MHz; $C^2HCl_3$) 5.57 (1H, d, $J_{CH}$ 5.5 CHOAc), 4.77 (1H, d, $J_{CH}$ 5.5, CHBr), 4.23 (4H, m, $OCH_2$), 2.15 (3H, s, $OCCH_3$) and 1.28 (6H, t, $J_{CH_2}$ 7.17, $CH_2C\underline{H}_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 169.84, 166.72, 166.10 (2 x CO), 73.16 (<u>C</u>HOAc), 62.77, 63.46 (2 x $OCH_2$), 44.36 (CHBr), 20.83 ($CH_3$) and 14.47, 14.39 (2 x $CH_2C\underline{H}_3$); *m/z* (CI) 313, 311 (25%, 26, bromine isotopes, $[M + H]^+$), 271, 269 (99, 100, $[M + 2H - COCH_3]^+$) and 233 (14, $[M + 2H - Br]^+$).

### 3.0.2.102 (2R,3S)-Diethyl threo-2-hydroxy-3-bromosuccinate (161)[154]

This compound was prepared in a manner identical to that described for the (2R,3S)-dimethyl threo-2-hydroxy-3-bromosuccinate (**157**) using (2R,3S)-diethyl threo-2-acetoxy-3-bromosuccinate (**160**) (20 g, 64 mmol) and dry ethanol (100 $cm^3$) to give a yellow oil (12.3 g, 72%), $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3470 (OH), 2985 (aliphatic CH) and 1740 (CO) (lit.,[144] 3480, 2985, 1740); $\delta_H$ (200 MHz; $C^2HCl_3$) 4.69 (1H, d, $J_{CH}$ 3.74 CHOH), 4.65 (1H, d, $J_{CH}$ 4.36 CHBr), 4.30 (4H, q, *J* 7.05, $OCH_2$), 3.47 (1H, s, OH) and 1.29 (6H, t, $J_{CH_2}$ 6.82, $CH_2C\underline{H}_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 170.76, 167.13 (2 x CO), 73.00 (CHOH),

63.33, 63.08 (2 x $OCH_2$), 48.15 (CHBr) and 14.53, 14.41 (2 x $CH_2\underline{CH_3}$); *m/z* (CI) 271, 269 (60%, 63, bromine isotopes, $[M + H]^+$), 191 (13, $[M + 2H - Br]^+$), 173 (36, $[M + 2H - Br - H_2O]^+$) and 117 (100, $[M + H - Br - CO_2CH_2CH_3]^+$).

### *3.0.2.103 (2R,3R)-Diethyl 2,3 epoxysuccinate (162)*[154]

This compound was prepared in a manner identical to that described for the (2R,3R)-dimethyl 2,3-epoxysuccinate (**158**) using dry EtOH and (2R,3S)-diethyl threo-2-hydroxy-3-bromosuccinate (**161**) (6.00 g, 23.0 mmol) to give a yellow oil (3.35 g, 77%), *m/z* (Found: $[M + H]^+$, 189.0760. Calc. for $C_8H_{12}O_5$ 189.0763); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 2990 (aliphatic CH) and 1740 (CO) (lit.,[154] 2986, 1748); $\delta_H$ (200 MHz; $C^2HCl_3$), 4.27 (2H, q, $J_{CH_3}$ 7.14, $CH_2$), 4.25 (2H, q, $J_{CH_3}$ 7.14, $CH_2$), 3.66 (2H, s, CH) and 1.31 (6H, t, $J_{CH_2}$ 7.18, $CH_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 167.24 (CO), 62.69 ($OCH_2$), 52.49 (CH) and 14.50 ($CH_3$); *m/z* (CI) 189 (50%, $[M + H]^+$) and 73 (100, $[CO_2Et]^+$).

### *3.0.2.104 Potassium (3S,3S)-ethyl 2,3-epoxysuccinate (163)*

This compound was prepared in a manner identical to that described for the potassium (2R,3R)-methyl 2,3 epoxysuccinate (**117**) using (2R,3R)-Diethyl-2,3-epoxysuccinate (**162**) (3.00 g, 16.0 mmol), dry ethanol and KOH (85%, 1.03 g, 18.4 mmol) to give a white solid (1.41 g, 45%), m.p. 133°C (decomp.); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3010, 3000 (2 x aliphatic CH) and 1750, 1730 (2 x CO); $\delta_H$ (200 MHz; $^2H_2O$), 4.22 (2H, q, $J_{CH_3}$ 7.17, $CH_2$), 3.56 (1H, d, $J_{CH}$ 2.00, CH), 3.46 (1H, d, $J_{CH}$ 1.80, CH) and 1.23 (3H, t, $J_{CH_2}$ 7.15 $CH_3$); $\delta_C$ (50.3 MHz; $^2H_2O$) 171.51, 168.09 (CO), 61.24 ($OCH_2$), 49.87, 52.26 (CH) and 11.25 ($CH_3$).

### 3.0.2.105 (2S,3R)-Diethyl erythro-2-acetoxy-3-bromosuccinate (164)[154]

This compound was prepared in a manner identical to that described for the (2R,3S)-dimethyl threo-2-acetoxy-3-bromosuccinate (**156**) using (2S,3S)-diethyl tartrate (30 g, 0.145 mol) to give an orange oil which was purified by silica chromatography eluting with ether to give a straw coloured oil (35.7 g, 79%), *m/z* (Found: $[M + H]^+$, 311.0123. $C_{10}H_{15}O_6Br$ requires 311.0130); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 2990 (aliphatic CH) and 1750 (CO) (lit.,[144] 3000 and 1765); $\delta_H$ (200 MHz; $C^2HCl_3$),5.57 (1H, d, $J_{CH}$ 5.41 CHOAc), 4.76 (1H, d, $J_{CH}$ 5.41 CHBr), 4.24 (4H, m, $OCH_2$), 2.15 (3H, s, $OCCH_3$) and 1.28 (6H, t, $J_{CH_2}$ 7.04, $CH_2C\underline{H}_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 169.84, 166.72, 166.10 (3 x CO), 73.16 (CHOAc), 62.77, 63.46 (2 x $OCH_2$), 44.36 (CHBr), 20.83 ($CH_3$) and 14.47, 14.39 (2 x $CH_2\underline{C}H_3$); m/z (CI) 313, 311 (46%, 48, bromine isotopes, $[M + H]^+$), 271, 269 (3, 4, $[M + 2H - COCH_3]^+$) and 75 (100, $[H_2CO_2Et]^+$).

### 3.0.2.106 (2S,3R)-Diethyl erythro-2-hydroxy-3-bromosuccinate (165)[154]

This compound was prepared in a manner identical to that described for the (2R,3S)-dimethyl threo-2-hydroxy-3-bromosuccinate (**157**) using (2S,3R)-diethyl threo-2-acetoxy-3-bromosuccinate (**164**) (20 g, 64 mmol) and dry ethanol (100 $cm^3$) to give a yellow oil (14.50 g, 84%), $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3475 (OH), 2985 (aliphatic CH) and 1745 (CO) (lit.,[144] 3500 and 3000, 1750); $\delta_H$ (200 MHz; $C^2HCl_3$) 4.71 (1H, d, $J_{CH}$ 4.23, CHOH), 4.66 (1H, d, $J_{CH}$ 4.19, CHBr), 4.28 (4H, m, $OCH_2$), 3.40 (1H, s, OH) and 1.30 (6H, t, $J_{CH_2}$ 7.14, $CH_2C\underline{H}_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 170.79, 167.16 (2 x CO), 73.01 (CHOH),

63.34, 63.08 (2 x $OCH_2$), 48.13 (CHBr) and 14.52, 14.41 (2 x $CH_2\underline{C}H_3$); *m/z* (CI) 271, 269 (8%, 9, bromine isotopes, $[M + H]^+$), 173 (12, $[M + 2H - Br - H_2O]^+$) and 117 (100, $[M + H - Br - CO_2CH_2CH_3]^+$).

### 3.0.2.107 (2S,3S)-Diethy 2,3-epoxysuccinate (166)[154]

This compound was prepared in a manner identical to that described for the (2R,3R)-dimethyl 2,3-epoxysuccinate (**158**) using dry EtOH and (2S,3R)-diethyl threo-2-hydroxy-3-bromosuccinate (**165**) (6.00 g, 23.0 mmol) to give a yellow oil (3.14 g, 73%), $\upsilon_{max}$ (thin film)/$cm^{-1}$ 2990 (aliphatic CH) and 1740 (CO) (lit.,[144] 3020, 2960 and 1760); $\delta_H$ (200 MHz; $C_2HCl_3$), 4.27 (2H, q, $J_{CH_3}$ 7.19 $CH_2$), 4.26 (2H, q, $J_{CH_3}$ 7.10 $CH_2$), 3.67 (2H, s, CH) and 1.32 (6H, t, $J_{CH_2}$ 7.15, $CH_3$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 167.23 (CO), 62.73 ($OCH_2$), 52.50 (CH) and 14.51 ($CH_3$); *m/z* (CI) 189 (50%, $[M + H]^+$) 161 (7, $[M + 2H - Et]^+$), 117 (15, $[M + 2H - CO_2Et]^+$) and 73 (100, $[CO_2Et]^+$).

### 3.0.2.108 Disodium cis epoxysuccinate (169)[156]

Maleic acid (29 g, 0.25 mol) was dissolved in water (75 $cm^3$) and NaOH (15 g, 0.38 mol) in water (25 $cm^3$) was added dropwise over 20 min so the temperature did not exceed 60°C. At 60°C sodium tungstate dihydrate (1.65 g, 7.6 mmol) was added and $H_2O_2$ solution (30%, 32 $cm^3$, 0.30 mol) was added dropwise so the temperature did not exceed 60°C. Aliquots of NaOH solution (2 $cm^3$ in total, 5 $moldm^{-3}$) were added periodically to maintain the pH above pH 4. The solution was stirred for 1 hour at 60°C, cooled to 40°C and the remaining NaOH solution added. The solution was concentrated under reduced pressure to 75 $cm^3$ and ethanol (100 $cm^3$) was added to give a white ppt was filtered and dried,[156] m.p. 215°C (decomp.);

$\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 171.75 (CO); $\delta_H$ (200 MHz; $^2H_2O$) 3.65 (1H, s, CH); $\delta_C$ (50.3 MHz; $^2H_2O$) 171.75, (CO), 53.36 (CH).

### *3.0.2.109 Tertiary-butoxycarbonyl-(2S)-leucyl-(2S)-alanyl-aminoacetonitrile (170)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl (2S)-alanyl glycyl glycinate (**65**) using tertiary-butoxycarbonyl-(2S)-leucyl (2S)-alanine (**43**) (200 mg, 0.661 mmol) and aminoacetonitrile hydrochloride (61 mg, 0.661 mmol). Recrystalisation from MeOH / ether gave white crystals (188 mg, 84%), m.p. 134°C; *m/z* (Found: $[M + H]^+$, 341.2197. $C_{16}H_{29}N_4O_4$ requires 341.2189); $[\alpha]_D$ -53.0 (*c* 0.25 in MeOH); $\upsilon_{max}$ (thin film)/$cm^{-1}$ 3380 (NH), 2950 (CH), 2210 (CN), 1680, 1670, 1650 ($CO_2NH$ and 2 x CONH); $\delta_H$ (200 MHz; $C^2HCl_3$) 7.57 (1H, t, $N\underline{H}CH_2$), 6.57 (1H, d, $J_{CH}$ 7.33, $N\underline{H}CH_3$), 4.95 (1H, d, $J_{CH}$ 5.13, $N\underline{H}CHCH_2$), 4.51 (1H, quin, *J* 7.31, $NHC\underline{H}CH_2$), 4.30 (1H, d, $J_{CH_b}$ 6.27, $C\underline{H}_aH_bNH$), 4.22 (1H, d, $J_{CH_a}$ 6.15, $CH_a\underline{H}_bNH$), 4.02 (1H, m, $C\underline{H}CH_3$), 1.66 (3H, m, $C\underline{H}_2C\underline{H}(CH_3)_2$), 1.47 (9H, s, $C(CH_3)_3$) 1.43 (3H, d, $J_{CH}$ 4.04, $CHC\underline{H}_3$), 0.98 (3H, d, $J_{CH}$ 4.40, $CH(C\underline{H}_3)_2$) and 0.95 (3H, d, $J_{CH}$ 4.15, $CH(C\underline{H}_3)_2$); $\delta_C$ (50.3 MHz; $C^2HCl_3$) 176.09, 175.36 (2 x CO), 158.57 (OCONH), 117.74 (CN), 81.12 ($\underline{C}(CH_3)_3$), 54.95 ($NH\underline{C}HCH_2$), 50.48 ($\underline{C}HCH_3$), 42.20 ($CH_2$), 29.04 ($C(\underline{C}H_3)_3$, 28.33 ($\underline{C}H_2CN$), 26.16 ($\underline{C}H(CH_3)_2$), 23.79, 22.17 ($CH(\underline{C}H_3)_2$), and 18.11 ($CH\underline{C}H_3$); *m/z* (CI) 341 (29%, $[M + H]^+$), 285 (100, $[M - NHCH_2CN]^+$), 267 (4, $[M - OC(CH_3)_3]^+$) and 241 (19, $[M + 2H - CO_2C(CH_3)_3]^+$).

### *3.0.2.110 O-(2R,3R)-[trans-Carbonyl oxiran-2-yl 3-(ethoxycarbonyl) butane] N-(carbobenzyloxy-leucyl alanyl)-ethanolamine (173)*

This compound was prepared in a manner identical to that described for the methyl benzoxycarbonyl-(2S)-leucyl (2S)-alaninate (**29**) using N-(Carbobenzyloxy-(2S)-leucyl (2S)-alanyl)-ethanolamine (**174**) (93 mg, 0.245 mmol) and pre-neutralised potassium (2R,3R)-ethyl 2,3 epoxysuccinate (**159**) (52 mg, 0.245 mmol) which was prepared as follows;[144] potassium (2R,3R)-ethyl 2,3 epoxysuccinate (**156**) (80 mg) was dissolved in $NaHSO_4$ solution (5%, 4 $cm^3$) and NaCl was added until the saturation point was reached. The (2R,3R)-ethyl 2,3 epoxysuccinate was then extracted into ethyl acetate (3 x 10 $cm^3$), dried ($MgSO_4$) and the solvent removed under reduced pressure to give a yellow oil. The crude oil (**173**) was purified by silica chromatograpy using ethyl acetate to give a white solid (58 mg, 41%), m.p. 128°C; (Found: C, 57.8; H, 6.85; N, 8.00. $C_{25}H_{35}N_3O_9$ requires C, 57.55; H, 6.75; N, 8.05%); *m/z* (Found: $[M + H]^+$, 522.2455. $C_{25}H_{36}N_3O_9$ requires 522.2452); $[\alpha]_D$ -5.89 (*c* 0.82 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3300 (NH), 3100, 3070, 3040 (3 x aromatic CH), 2970, 2930, 2880 (3 x aliphatic CH), 1750, 1730, 1710, 1680, 1640 (5 x CO); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.39 (5H, m, $C_6H_5$), 5.14 (2H, s, $C\underline{H}_2C_6H_5$), 4.30 (5H, m, $OC\underline{H}_2CH_3$, $OC\underline{H}_2CH_2$ and $C\underline{H}CH_3$), 4.16 (1H, t, $J_{CH_2}$ 7.52, $NHC\underline{H}CH_2$), 3.74 (1H, d, $J_{CH}$ 5.19, CHO), 3.74 (1H, d, $J_{CH}$ 5.20, CHO), 3.49 (2H, m, $CH_aH_bN$), 1.74 (1H, sep, $J_{CH_3}$ 6.66, $C\underline{H}(CH_3)_2$), 1.60 (2H, t, $J_{CH_2}$ 7.15, $CHC\underline{H}_2$), 1.37 (3H, t, $J_{CH_2}$ 6.26, $CH_2C\underline{H}_3$), 1.31 (3H, d, $J_{CH}$ 7.12, $CHC\underline{H}_3$), 0.92 (3H, d, $J_{CH}$ 6.50, $CH(C\underline{H}_3)_2$) and 0.97 (3H, d, $J_{CH}$ 6.44, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz, $^2H_3CO^2H$) 175.26, 175.12 (2 x CONH), 168.45, 168.40 (2 x $CO_2R$), 158.76 (OCONH), 138.99 ($i$-$C_6H_5$), 129.49 ($m$-$C_6H_5$), 129.03 ($p$-$C_6H_5$), 128.78 ($o$-$C_6H_5$), 67.89 ($C_6H_5C\underline{H}_2$), 65.40 ($O\underline{C}H_2CH_2$), 63.20 ($O\underline{C}H_2CH_3$), 55.21 ($NH\underline{C}HCH_2$), 53.09, 53.00 (2 x

CHO), 50.47 (C<u>H</u>CH$_3$), 41.84 (CH$_2$N), 39.24 (CH<u>C</u>H$_2$CH), 25.86 (<u>C</u>H(CH$_3$)$_2$), 23.46, 21.88 (CH(<u>C</u>H$_3$)$_2$), 17.91 (CH<u>C</u>H$_3$) and 14.34 (CH$_2$<u>C</u>H$_3$); *m/z* (CI) 522 (5%, $[M + H]^+$) and 362 (100, $[M + H - CO_2CH(O)CHCO_2Et]^+$).

## *3.0.2.111 N-(Carbobenzyloxy-(2S)-leucyl (2S)-alanyl) ethanolamine (174)*

This compound was prepared in a manner identical to that described for methyl carbobenzyloxy-(2S)-leucyl-(2S)-alanyl-glycyl-glycinate (**65**) using carbobenzyloxy-(2S)-leucyl (2S)-alanine (**29**) (300 mg, 0.893 mmol) and ethanolamine hydrochloride (87 mg, 0.893 mmol) to give a white solid which was recrystallised from acetone / ether to give white crystals (244 mg, 72%), m.p. 140°C; (Found: C, 60.4; H, 7.5; N, 11.05. $C_{19}H_{29}N_3O_5$ requires C, 60.15; H, 7.70; N, 11.05%); $[\alpha]_D$ -49.1 (*c* 0.16 in MeOH); $\upsilon_{max}$ (thin film)/cm$^{-1}$ 3240 (OH), 3050, 3040, 3010 (3 x aromatic CH), 2980, 2900, 2850 (3 x aliphatic CH), 1700, 1650 (2 x CONH); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.41 (5H, m, $C_6H_5$), 5.25 (1H, d, $J_{CH_b}$ 12.75, C<u>H</u>$_a$H$_b$C$_6$H$_5$), 5.18 (1H, d, $J_{CH_a}$ 12.40, CH$_a$<u>H</u>$_b$C$_6$H$_5$), 4.48 (1H, q, $J_{CH_3}$ 6.74, C<u>H</u>CH$_3$), 4.30 (1H, t, $J_{CH_2}$ 7.11, NHC<u>H</u>CH$_2$), 3.70 (2H, t, $J_{CH_2}$ 5.62, CH$_2$O), 3.40 (2H, t, $J_{CH_2}$ 5.54, CH$_2$N), 1.80 (1H, sep, $J_{CH_3}$ 6.43, C<u>H</u>(CH$_3$)$_2$), 1.86 (2H, t, $J_{CH_2}$ 7.10, CHC<u>H</u>$_2$), 1.45 (3H, d, $J_{CH}$ 7.01, CHC<u>H</u>$_3$), 1.05 (3H, d, $J_{CH}$ 6.52, CH(C<u>H</u>$_3$)$_2$) and 1.03 (3H, d, $J_{CH}$ 6.49, CH(C<u>H</u>$_3$)$_2$); $\delta_C$ (75.5 MHz, $^2H_3CO^2H$) 175.08, 174.83 (2 x CONH), 158.52 (OCONH), 137.99 (*i*-C$_6$H$_5$), 129.39 (*m*-C$_6$H$_5$), 128.94 (*p*-C$_6$H$_5$), 128.71 (*o*-C$_6$H$_5$), 67.68 (C$_6$H$_5$C<u>H</u>$_2$), 61.39 (OCH$_2$), 54.97 (NH<u>C</u>HCH$_2$), 50.28 (<u>C</u>HCH$_3$), 42.90 (CH$_2$N), 41.86 (CH<u>C</u>H$_2$CH), 25.74 (<u>C</u>H(CH$_3$)$_2$), 23.44, 21.87 (2 x CH(<u>C</u>H$_3$)$_2$) and 18.15 (CH<u>C</u>H$_3$); *m/z* (CI) 380 (100%, $[M + H]^+$) and 272 (58, $[M - C_6H_5CO]^+$).

### *3.0.2.112 O-(2S,3S)-[trans-Carbonyl oxiran-2-yl 3-(ethoxycarbonyl) butane] N-(carbobenzyloxy-leucyl-alanyl)-ethanolamine (175)*

This compound was prepared in a manner identical to that described for the methyl benzoxycarbonyl-(2S)-leucyl (2S)-alaninate (**29**) using N-(Carbobenzyloxy-(2S)-leucyl (2S)-alanyl)-ethanolamine (**174**) (93 mg, 0.245 mmol) and pre-neutralised potassium (2S,3S)-ethyl-2,3-epoxysuccinate (**163**) (56 mg, 0.264 mmol) which was prepared as follows;[144] potassium (2S,3S)- ethyl-2,3-epoxysuccinate (**163**) (80 mg) was dissolved in $NaHSO_4$ solution (5%, 4 $cm^3$) and NaCl was added until the saturation point was reached. The (2S,3S)-ethyl threo 2,3 epoxysuccinate was then extracted into ethyl acetate (3 x 10 $cm^3$), dried ($MgSO_4$) and the solvent removed under reduced pressure to give a yellow oil. The crude oil(**175**) was purified by silica chromatograpy using ethyl acetate to give a white solid (51mg, 37%), m.p. 150-151°C; *m/z* (Found: $[M + H]^+$, 522.2456. $C_{25}H_{36}N_3O_9$ requires 522.2452); $[\alpha]_D$ +1.80 (*c* 0.89 in MeOH); $\upsilon_{max}$ (KBr disk)/$cm^{-1}$ 3300 (NH), 3050, 3025, 3010 (3 x aromatic CH), 2930, 2920, 2880 (3 x aliphatic CH), 1750, 1730, 1710, 1690, 1650 (5 x CO); $\delta_H$ (300 MHz; $^2H_3CO^2H$) 7.39 (5H, m, $C_6H_5$), 5.14 (2H, s, $C\underline{H}_2C_6H_5$), 4.75 (4H, m, $OC\underline{H}_2CH_3$ and $OC\underline{H}_2CH_2$), 4.33 (1H, q, $J_{CH_3}$ 5.73, $C\underline{H}CH_3$), 3.95 (1H, t, $J_{CH_2}$ 5.96, $NHC\underline{H}CH_2$), 3.75, 3.73 (2H, ds, CHO), 3.47 (2H, m, $CH_a\underline{H}_bN$), 3.46 (2H, dt, *J* 10.66, 20.38, $C\underline{H}_aH_bN$), 1.78 (1H, sep, $J_{CH_3}$ 6.62, $C\underline{H}(CH_3)_2$), 1.60 (2H, t, $J_{CH_2}$ 7.4, $CHC\underline{H}_2$), 1.38 (3H, d, $J_{CH}$ 7.19, $CHC\underline{H}_3$), 1.33 (3H, t, $J_{CH_2}$ 7.14, $CH_2C\underline{H}_3$), 0.94 (3H, d, $J_{CH}$ 6.51, $CH(C\underline{H}_3)_2$) and 0.97 (3H, d, $J_{CH}$ 6.43, $CH(C\underline{H}_3)_2$); $\delta_C$ (75.5 MHz, $^2H_3CO^2H$) 178.77, 178.61 (2 x CONH), 175.32, 175.17 (2 x $CO_2R$), 158.52 (OCONH), 138.23 (*i*-$C_6H_5$), 129.54 (*m*-$C_6H_5$), 129.08 (*p*-$C_6H_5$), 128.83 (*o*-$C_6H_5$), 67.85 ($C_6H_5C\underline{H}_2$), 65.45 ($OC\underline{H}_2CH_2$), 63.26 ($O\underline{C}H_2CH_3$), 55.26 ($NH\underline{C}HCH_2$), 53.15, 53.06 (CHO), 50.53 ($C\underline{H}CH_3$), 41.89

($CH_2N$), 39.29 ($CH\underline{C}H_2CH$), 25.91 ($\underline{C}H(CH_3)_2$), 23.52, 21.93 ($CH(\underline{C}H_3)_2$), 17.97 ($CH\underline{C}H_3$) and 14.40 ($CH_2\underline{C}H_3$); *m/z* (CI) 522 (7%, $[M + H]^+$) and 362 (100, $[M - CO_2CH(O)CHCO_2Et]^+$).

# 3.1 Biological assays

## 3.1.0 Initial methods to determine potential inhibitor activity

Three different methods were used to determine the activity of the potential inhibitors. The three assay methods were fairly similar, except that different preparations of proteinase were used between each of the methods. The following tables show which methods were used to test the various potential inhibitors mentioned in the discussion.

| Compound | Number |
|---|---|
| Cbz-Leu-Ala-Gly-Gly-$NH_2$ | **19** |
| Cbz-Leu-Ala-Gly-Gly-NHMe | **20** |
| Cbz-Leu-Ala-Gly-Gly-$NMe_2$ | **21** |
| Cbz-Leu-Ala-Gly-Gly-NMeEt | **22** |
| Cbz-Leu-Ala-Gly-Gly-NMeOMe | **23** |
| Leu-Ala-Gly-Gly-$NH_2$ . HOAc | **25** |
| Leu-Ala-Gly-Gly-NHMe . HOAc | **26** |
| Cbz-Leu-Ala-Gly-Gly | **76** |
| Cbz-Leu-Ala-Gly-aminoacetonitrile | **78** |
| t-Boc-Leu-Ala-Gly-aminoacetonitrile | **79** |
| TFA-Leu-Ala-Gly-aminoacetonitrile | **82** |
| TFA-Ala-Gly-aminoacetonitrile | **83** |
| TFA-Gly-Gly-aminoacetonitrile | **84** |
| aminoacetonitrile hydrochloride | **85** |

Table 3.1: *The potential inhibitors tested by method 2.*

| Compound | Number |
|---|---|
| t-Boc-Leu-Ala-Gly-Gly-*p*-nitroanilide | **24** |
| Cbz-Leu-Ala-Gly-Gly-OMe | **65** |
| Cbz-Leu-Ala-Gly-Gly-OEt | **66** |
| t-Boc-Leu-Ala-Gly-Gly-OtBu | **73** |
| N-(Cbz-Leu-Ala-Gly)-ethanolamine | **75** |
| Cbz-Leu-Ala-Gly-aminoacetonitrile | **78** |
| t-Boc-Leu-Ala-Gly-aminoacetonitrile | **79** |
| Cbz-Leu-Ala-Ala-aminoacetonitrile | **80** |
| Cbz-Leu-Ala-(2R)-Ala-aminoacetonitrile | **81** |
| t-Boc-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal | **113** |
| Cbz-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal | **114** |
| Cbz-Leu-Ala-Gly-1-amino prop-2-ene | **115** |
| Cbz-Leu-Ala-Gly-1-amino 2-methyl prop-2-ene | **118** |
| Cbz-Leu-Ala-Gly-2-bromoethylamine | **130** |
| N, N' (di(Cbz-Leu-Ala-Gly))- 2, 4-diamino but-2-ene aldehyde | **122** |
| O-(Cbz-Leu-Ala-Gly)- 1-carboxymethyl, 4-hydroxy but-2-ene | **132** |
| (2R,3R)-O-[trans-carbonyl oxiran-2-yl 3-(ethoxycarbonyl) butane] N-(Cbz-Leu-Ala-)-ethanolamine | **173** |
| (2S,3S)-O-[trans-carbonyl oxiran-2-yl 3-(ethoxycarbonyl) butane] N-(Cbz-Leu-Ala-)-ethanolamine | **175** |

Table 3.2: *The potential inhibitors tested by method 1.*

| Compound | Number |
|---|---|
| Cbz-Leu-Ala-Gly-Gly-OtBu | **67** |
| t-Boc-Leu-Ala-Gly-Gly-OBz | **72** |
| t-Boc-Leu-Ala-Gly-aminoacetonitrile | **79** |
| Cbz-Leu-Ala-Gly-aminoacetaldehyde | **112** |
| Cbz-Leu-Ala-Gly-1-aminopropanone | **116** |

Table 3.3: *The potential inhibitors tested by method 3.*

### *3.1.0.0 Method 1*

All solvents used were HPLC grade. Each set of assays consisted of a control (4% DMSO solution) and 4 potential inhibitors dissolved in DMSO and diluted to 4% DMSO with water. Each assay was offset by 1 min compared to the previous assay, so overall one reading was taken every min.

Each potential inhibitor assay was conducted as follows; proteinase solution (20 $mm^3$, 3.87 $\mu mol\ dm^{-3}$, 77.4 nmol), activating peptide (GVQSLKRRRCF) solution (20 $cm^3$, 10 $\mu mol\ dm^{-3}$, 0.2 pmol) and tris / NaCl buffer (50 $cm^3$, 2 $mmol\ dm^{-3}$, tris and 100 $mmol\ dm^{-3}$ NaCl at pH 8) where mixed in the order shown and incubated at 37°C for 15 min.

Substrate (Acetyl-LRGA-GRSR) solution (25 $mm^3$, 4.42 $mmol\ dm^{-3}$, 9.77 pmol) and potential inhibitor 4% DMSO solution (25 $mm^3$, 0.22 $mmol\ dm^{-3}$, 4.4 pmol) were mixed and also incubated at 37°C for 10 min. After the proteinase had been incubated for 15 min, the inhibitor / substate mixture (40 $mm^3$) was added to the activated proteinase solution, mixed and incubated at 37°C.

After 5 min from the addition of the substrate to the proteinase, an aliquot (20 $mm^3$) of the first assay solution was added to a trifluoroacetic acid solution (50 $mm^3$, 1%) and mixed. Further aliquots (20 $mm^3$) of the assay solution were quenched with trifluoroacetic acid solution at 10, 15, 20, 25 min time intervals.

Each aliquot was further diluted with water (80 $mm^3$) and filtered through a micropore filter at 13000 rpm for 3 min. A clean filter was used for each assay, filtering the first aliquot first and reusing the filter for the later aliquots. Each sample was then analysed by capillary electrophoresis using a Bio-rad, BioFocus™ 3000 capillary electrophoresis system by Dr Kemp at the University of St Andrews Biochemistry department

### *3.1.0.1 Method 2*

The assay was carried out in a manner identical to that described for method 1, using proteinase solution (1.74 μmol $dm^{-3}$, 34.8 nmol) and potential inhibitor 6% DMSO solution of (8.8 mmol $dm^{-3}$, 0.176 μmol). All method 2 assays were performed by Dr Heather Murray at the University of St Andrews Biochemistry department.

### *3.1.0.2 Method 3*

The assay was carried out in a manner identical to that described for method 1, except that all the volumes used were halved, aliquots were taken at 6, 12 and 18 min time intervals and using proteinase solution (2.17 μmol $dm^{-3}$, 21.7 nmol).

### *3.1.0.4 Treatment of the capillary electrophoresis results*

The output from the capillary electrophoresis system is depicted in Fig. 3.1, showing the substrate (S) and product ($P_1$ and $P_2$) peaks.

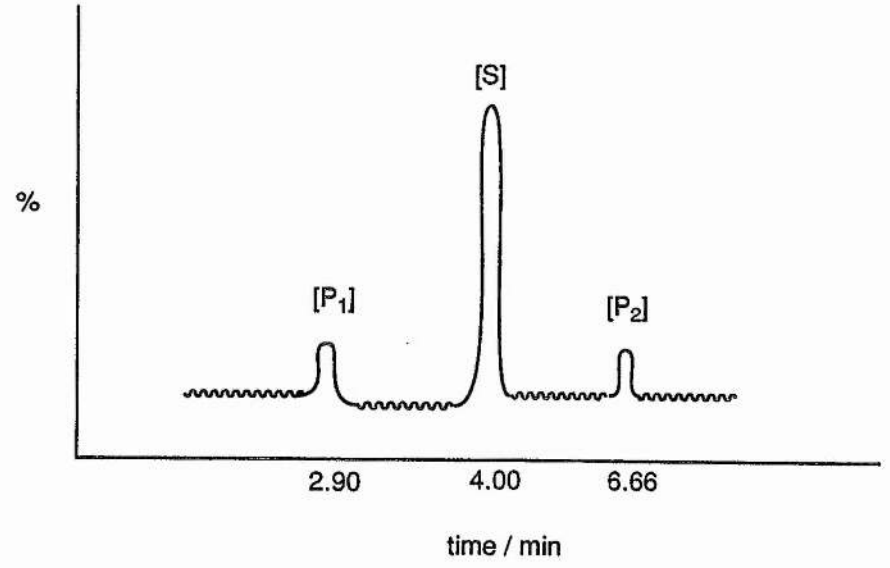



Fig. 3.1: *The output from a capillary electrophoresis system.*

Integration gives the size of each peak as a percentage of the total area of all the peaks. The initial rate is determined from the plot of the size of $P_1$ from the integration versus time in seconds. Provisionally the initial rates were determined by hand, but then by using the Enzfitter program[175] as the 2 methods of data analysis were identical. Enzfitter was set to first-order-rate-plus-offset. The initial rate is given in equation 3.1.

The first order rate equation is given by;

$$[S]_t = [S]_o [1 - e^{kt}] \qquad 3.1$$

Where the concentration of the substrate at time, t is given by $S_t$, the concentration of the substrate at t = 0 (the limit in the Enzfitter program) is given by $S_0$ and the rate constant is given by k. Differentiation of equation 3.1 gives;

$$\frac{d[S]_t}{dt} = -[S]_0 k e^{kt} \qquad 3.2$$

But as;

$$\frac{d[S]_t}{dt} = \frac{-[P]_t}{dt} \qquad 3.3$$

Therefore;

$$\frac{d[P]_t}{dt} = [S]_0 k e^{kt} \qquad 3.4$$

Substituting t = 0 in equation 3.4 gives;

$$\text{rate at } t = 0 = [S]_0 k \qquad 3.5$$

## 3.1.1 Inactivation studies

### *3.1.1.0 Method 1*

Proteinase solution (0.1845 $cm^3$, 3.87 μmol $dm^{-3}$, 71.4 nmol), activating peptide (GVQSLKRRRCF) solution (0.1845 $cm^3$, 10 μmol $dm^{-3}$, 1.85 pmol) and tris / NaCl buffer (0.461 $cm^3$, 2 mmol $dm^{-3}$, tris and 100 mmol $dm^{-3}$ NaCl at pH 8) where mixed in the order shown and incubated at 0°C for 15 min. This solution was separated out into 2 portions (400 $cm^3$) and after 1 min, N' N'-dicarboxybenzyl-Leu-Ala-Gly)- 2, 4-diamino-but-2-ene aldehyde (**122**) 4% DMSO solution (88.9 $mm^3$, 4.4 μmol $dm^{-3}$, 0.391 nmol) was added to one portion and 4% DMSO solution (88.9 $mm^3$) was added to the other.

After a further 3 min the inhibitor / enzyme solution (0.111 $cm^3$) was added to the substrate (Acetyl-LRGA-GRSR) solution (0.020 $cm^3$, 4.42 mmol $dm^{-3}$, 88.4 nmol) at 37°C and mixed. Then after 1, 2, 3, 6.30 and 10 min from the addition of the proteinase solution to the substrate solution, aliquots (20 $mm^3$) of the assay solution were quenched with trifluoroacetic acid solution (50 $mm^3$, 1%). A similar procedure was repeated for the control / enzyme solution after the first assay was complete. A further 3 sets of assays for the control / proteinase solution and the blank / proteinase were similarly taken after the first set.

The aliquots were diluted, filtered and analysed as for Section 3.0.2.0.0, method 1. This procedure was only successful for a very robust proteinase preparation. One attempt out of six gave no proteinase activity in the control.

### *3.1.1.1 Method 2*

Proteinase solution (10 $mm^3$, 2.17 $\mu mol\ dm^{-3}$, 21.7 pmol), activating peptide (GVQSLKRRRCF) solution (0.010 $cm^3$, 10 $\mu mol\ dm^{-3}$, 10 nmol), tris / NaCl buffer (25 $mm^3$, 2 $mmol\ dm^{-3}$, tris and 100 $mmol\ dm^{-3}$ NaCl at pH 8) and the 4% DMSO inhibitor solution (10 $mm^3$, 0.22 $mmol\ dm^{-3}$, 2.2 pmol) were mixed in the order shown and incubated at 37°C for 3 min.

Then substrate (Acetyl-LRGA-GRSR) solution (10 $mm^3$, 4.42 $mmol\ dm^{-3}$, 4.42 pmol) was added, mixed and aliquots (10 $mm^3$) were removed and quenched with trifluoroacetic acid solution (0.025 $cm^3$, 1%) after 5, 10 and 15 min from the addition of the substrate to the proteinase solution. The same procedure was repeated for a control of 4% DMSO, off setting the additions and the aliquots by 1 min from the inhibitor assay.

This procedure was repeated twice more, leaving the inhibitor / control solutions for 5 and 7 min with the proteinase, before the addition of the substrate. The quenched aliquots were diluted with water (40 $mm^3$), filtered and analysed as in Section 3.0.2.0.0 method 1.

## 3.1.3 Double reciprocal (Lineweaver Burke) plots and $K_i$ determination

The aim of these plots is to give an estimation of $K_i$ and to give an indication as to the type of inhibition observed. The concentrations of inhibitor chosen are critical to the accuracy of the plots. If too large concentrations are chosen, the gradient of the double reciprocal plots are very large and a small error in any of the points, particularly the lower concentrations of substrate which are usually the most prone to error will significantly alter the 1/v intercept value. If too small inhibitor concentrations are chosen the gradients of the double reciprocal plots are too small and a small error in any of the points, particularly the lower concentrations of substrate which are usually the most prone to error will significantly

alter the gradient. These problems can usually be avoided by chosing inhibitor concentrations of about $K_i$, *ie* ½ $K_i$, $K_i$ and twice $K_i$.

A rough estimate of $K_i$ can be determined from the concentration of inhibitor required to give 50% inhibition of the proteinase for competitive inhibition. ½ $K_i$ gives around 25% inhibition and twice $K_i$ gives approximately 75% inhibition. However saturation of the enzyme's active site by the inhibitor starts to affect the linear relationship between inhibition and concentration of inhibitor and 87% inhibition is given by a concentration of inhibitor of about four times $K_i$. Estimations of $K_i$ for the various inhibitors, the concentrations of inhibitors used to determine $K_i$ and the assay method used for each determination are give in Table 3.4.

| 34 μM of inhibitor represents; | Inhibitor number | Proteinase conc. | Concentrations used /mmol $dm^{-3}$ | Assay method |
|---|---|---|---|---|
| $Ki_I$ | **(78)** | 1.04 μmol $dm^{-3}$ | 0.00, 0.11, 0.22, 0.44 | 1 |
| ½$K_i$ | **(118)** | 0.35 μmol $dm^{-3}$ | 0.00, 0.22, 0.44, 0.55 | 1 |
| $K_i$ | **(79)** | 1.04 μmol $dm^{-3}$ | 0.00, 0.11, 0.22, 0.44 | 1 |
| 1/3 $K_i$ | **(114)** | 1.04 μmol $dm^{-3}$ | 0.00, 0.22, 0.44, 0.55 | 1 |
| 6$K_i$ | **(122)** | 2.17 μmol $dm^{-3}$ | 0.00, 0.031, 0.044, 0.073 | 1 |

Table 3.4: *Concentrations tested and assay methods for various inhibitors for double reciprocal plots and $K_i$ determination.*

### 3.1.3.0 method 1

Each concentration of inhibitor, *ie.* zero, ½ $K_i$ and twice $K_i$ as shown in Table 3.2 were assayed for proteinase activity at five different substrate concentrations; 3.87, 4.42, 5.52, 7.73 and 11.05 mmol $dm^{-3}$ as follows;

Ten assays were carried out, each staggered by 1 min time. Each assay was conducted as follows; proteinase solution (20 $mm^3$), activating peptide (GVQSLKRRRCF) solution (10 $mm^3$, 20 μmol $dm^{-3}$, 0.2 pmol) and tris / NaCl buffer (15 $mm^3$, 2 mmol $dm^{-3}$, tris and 100 mmol $dm^{-3}$ NaCl at pH 8) were mixed in the order shown and incubated at 37°C for 10min.

Substrate (Acetyl-LRGA-GRSR) solution (15 $mm^3$) and potential inhibitor 4% DMSO solution (15 $mm^3$) were mixed and also incubated at 37°C. After 10 min the inhibitor / substrate solution (0.020 $cm^3$) was added to the activated proteinase solution, mixed and aliquots were taken at 10 min intervals for 50 min and quenched in TFA solution (25 $mm^3$, 1%). Each aliquot was diluted with water (40 $mm^3$), filtered and analysed as in method 1 (Section 3.3.0.0).

### 3.1.3.1 Treatment of results

The initial rates were determined from the capillary electrophoresis as before (Section 3.3.0.4). However the concentrations of substrate were different between each assay, so the $P_1$ percentage peak area was multiplied by the concentration of substrate in mol $dm^{-3}$ to give the actual concentrations of $P_1$.

The type of inhibition observed for each inhhibitor and the $K_i$ was obtained from the double reciprocal plot (Lineweaver-Burke plot) of each inhibitor concentration *ie* 1/v verses 1/S as described in Section 1.2.3.

# 3.2 NMR Spectroscopy

## 3.2.0 $^1$H-NMR Assay's of potential Substrates

It was possible that the proteinase could hydrolyse the amides, the esters, the cyanides and the dimethylacetal functionalities. Proteolytic assays with the potential substrates were conducted in $D_2O$ or phosphate buffer and the course of the reaction followed by $^1$H-NMR to detect any hydrolysis products. If hydrolysis was observed *ie.* the $^1$H-NMR spectrum showed the disappearence of the ester signals and the appearence of the product signals, an effort was made to identify the new signals. Controls for the proteolytic assays were conducted where $D_2O$ or phosphate buffer was added instead of the proteinase solution and any change in the substrate was monitored by $^1$H-NMR spectroscopy at 300MHz.

Two assays were employed to determine if hydrolysis of a potential substrate occurred, differing only in the concentration of proteinase used for each assay. Aproximately enough proteinase to hydrolyse 1 mg of substrate (2 μmol) ten times over in 3 hours was added in each case.

### *3.2.0.0 $^1$H-NMR assay 1*

The following potential substrates were tested by this method (Table 3.5). The potential substrate (1mg, 2 μmol) was dissolved in $C^2H_3SOC^2H_3$ (4 mm$^3$) and then the solution of potential substrate was added to $D_2O$ (365 mm$^3$) after vigorous shaking and heating most of the potential substrate that had precipitated out redissolved. The solution was allowed to cool to room temperature and $D_2O$ (366.5 mm$^3$) was added.

The proteinase solution (57 mm$^3$, 34.5 μmol dm$^{-3}$, 1.97 nmol) was added to a mixture of activating peptide (GVQSLKRRRCF) solution (4.6 mm$^3$, 43.5 mmol dm$^{-3}$, 200 nmol) and 1% mercaptoethanol solution (2.9 mm$^3$) and incubated at 0°C for 15 min.

The activated proteinase solution was added to the potential substrate solution with vigorous mixing and the solution was placed in an NMR tube and a $^1$H-NMR acquisition at 300MHz was carried out as soon as possible. Further acquisitions were carried out after one, two, three and four hours from the addition of the proteinase solution to the potential substrate solution.

If the $^1$H-NMR spectrum had changed compared to the starting material, the nature of the products were determined as follows; The tetrapeptide product was compared with the $^1$H-NMR spectrum spectra of Cbz-Leu-Ala-Gly-Gly (**76**) and N-(Cbz-Leu-Ala-Gly)-ethanolamine (**75**). The nature of the alcohol product (in the case of the esters) and the amine product (in the case of the amides) was determined from adding a drop of the relevant alcohol or amine to the assay mixture and running another $^1$H-NMR spectrum to make sure the alcohol or amine product signals appeared at the same chemical shift as the authenic material.

| Compound | Number |
|---|---|
| Cbz-Leu-Ala-Gly-Gly-NMeEt | **(22)** |
| Leu-Ala-Gly-Gly-NHMe. HOAc | **(26)** |
| Cbz-Leu-Ala-Gly-Gly-OMe | **(65)** |
| Cbz-Leu-Ala-Gly-Gly-OEt | **(66)** |
| t-Boc-Leu-Ala-Gly-Gly-OtBu | **(73)** |
| Cbz-Leu-Ala-Gly-aminoacetonitrile | **(78)** |
| Cbz-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal | **(114)** |

Table 3.5: *Potential substrates for the Adenovirus tested by NMR method 1.*

Two controls were also carried out. The first control consisted of the same materials as the assay but the potential substrate was missed out. The second controls contained the same materials as the assay but $D_2O$ was added instead of the proteinase solution and the $^1$H-NMR spectra were recorded at one, two, three and four hour time intervals.

### 3.2.0.1 $^1$H-NMR assay 2

The following potential substrates were tested by this method (Table 3.6);

| Compound | Number |
|---|---|
| t-Boc-Leu-Ala-Gly-Gly-*p*$NO_2$ anilide | **(24)** |
| Cbz-Leu-Ala-Gly-Gly-OMe | **(65)** |
| Cbz-Leu-Ala-Gly-Gly-OtBu | **(66)** |
| t-Boc-Leu-Ala-Gly-Gly-OBz | **(67)** |
| t-Boc-Leu-Ala-Gly-Gly-OtBu | **(73)** |

Table 3.6: *Potential substrates for the adenovirus tested for method 2.*

The potential substrate (1mg, 2 μmol) was dissolved in $C^2H_3SOC^2H_3$ (4 $mm^3$) and then the solution of potential substrate was added to phosphate buffer (327 $mm^3$, 50 mmol $dm^{-3}$, pD 7.6) and with a lot of vigorous shaking so that most of the potential substrate that had precipitated out redissolved. The solution was allowed to cool to room temperature and phosphate buffer (0.327 $cm^3$, 50 mmol $dm^{-3}$, pD 7.6) was added.

The proteinase solution (50 $mm^3$, 37.5 μmol $dm^{-3}$, 1.88 nmol) was added to a mixture of activating peptide (GVQSLKRRRCF) solution (50 $mm^3$, 4.3 mmol $dm^{-3}$, 215 nmol) and 1% mercaptoethanol solution (3 $mm^3$) and incubated at 0°C for 15 min.

The activated proteinase solution was added to the potential substrate solution with vigorous mixing and the solution was placed in an NMR tube and a $^1$H-NMR acquisition at 300MHz was carried out as soon as possible. Further acquisitions were carried out after one, two, three and four hours from the addition of the proteinase solution to the potential substrate solution.

If the $^1$H-NMR spectrum had changed compared to the starting material, the nature of the products were determined as follows; The tetrapeptide product was compared with the $^1$H-NMR spectrum spectra of Cbz-Leu-Ala-Gly-Gly (**76**) and N-(Cbz-Leu-Ala-Gly)-ethanolamine (**75**). The nature of the alcohol product (in the case of the esters) and the amine product (in the case of the amides) was determined from adding a drop of the relevant alcohol or amine to the assay mixture and running another $^1$H-NMR spectrum to make sure the alcohol or amine product signals appeared at the same chemical shift as the authenic material.

Three controls were also carried out. The first control was to check that any new signals came from the hydrolysis of a substrate rather than the activated proteinase solution and consisted of the same materials as the assay, but the potential substrate was missed out. $^1$H-NMR spectra were recorded at one and four hour time intervals. The second controls were to check that the observed hydrolysis was due to the proteinase and contained the same materials as the assay, but $D_2O$ was added instead of the proteinase solution. The $^1$H-NMR spectra were recorded at one, two, three and four hour time intervals. The third control was to check that the proteinase was active under the assay conditions used.

## 3.2.0 $^{13}$C-NMR of The $^{13}$C-labelled Nitrile

As mentioned in the discussion (Section 2.3.3.1) three different $^{13}$C-NMR experiments were carried out with adenovirus proteinase solution and [1-$^{13}$C]-Cbz-Leu-Ala-Gly-aminoacetonitrile (**86**). The experiments were carried out using the same method, as outlined below. The concentrations and volumes of the reagents are given in Tables 3.7, 3.8 and 3.9.

Proteinase, activating peptide, 1% mercaptoethanol solution (if used) and buffer (or water) were mixed together and the $^{13}$C-NMR spectrum was recorded at 125.80 MHz at 30°C for 1 hour. Then the inhibitor DMSO solution was added and another $^{13}$C-NMR spectrum was recorded over 3 hours.

A control spectrum was also recorded over 3 hours, which contained the same amounts of reagents as used in the adduct experiment, however the appropriate buffer was added instead of the proteinase solution.

| Reagent | Volume / $cm^3$ | Conc. / mol $dm^{-3}$ | moles |
|---|---|---|---|
| proteinase in $D_2O$ | 0.55 | $4.35 \times 10^{-5}$ | $2.39 \times 10^{-8}$ |
| activating peptide (GVQSLKRRRCF) | 0.089 | $2.44 \times 10^{-3}$ | $2.39 \times 10^{-6}$ |
| $d^6$-DMSO + (**86**) | 0.016 | 0.14 | $2.31 \times 10^{-6}$ |
| mercaptoethanol solution | 0.062 | 1% | not applicable |

Table 3.7:

*Concentrations and volumes of the reagents used in the first $^{13}$C-NMR experiment.*

| Reagent | Volume / $cm^3$ | Conc. / mol $dm^{-3}$ | moles |
|---|---|---|---|
| proteinase in $D_2O$ | 0.400 | $1.30 \times 10^{-4}$ | $5.22 \times 10^{-8}$ |
| activating peptide (GVQSLKRRRCF) | 0.200 | $3.20 \times 10^{-2}$ | $6.40 \times 10^{-6}$ |
| $d^6$-DMSO + (**86**) | 0.004 | 0.174 | $6.94 \times 10^{-7}$ |
| Tris buffer[a] | 0.100 | not applicable | not applicable |

Table 3.8:

*Concentrations and volumes of the reagents used in the second $^{13}C$-NMR experiment.*

| Reagent | Volume / $cm^3$ | Conc. / mol $dm^{-3}$ | moles |
|---|---|---|---|
| proteinase in $H_2O$ | 2.5 | $8.70 \times 10^{-5}$ | $2.17 \times 10^{-7}$ |
| activating peptide (GVQSLKRRRCF) | added as solid | 22mg | $2.17 \times 10^{-5}$ |
| DMSO + (**86**) | 0.011 | 0.079 | $8.68 \times 10^{-7}$ |
| Posphate buffer[b] | 0.2 | not applicable | not applicable |
| $D_20$ for locking | 0.29 | 10% of final vol. | not applicable |

Table 3.9:

*Concentrations and volumes of the reagents used in the third $^{13}C$-NMR experiment.*

a 50 mmol $dm^{-3}$ Tris, 0.2 mol $dm^{-3}$ NaCl, 10 mmol $dm^{-3}$ ethylenediaminetetraacetic acid in $D_2O$ at pH 8.

b 50 mmol $dm^{-3}$ pospahte buffer at pH 7.6.

# Appendix 1

## Preliminary double reciprocal plots

## and

## approximate $K_i$ determinations for

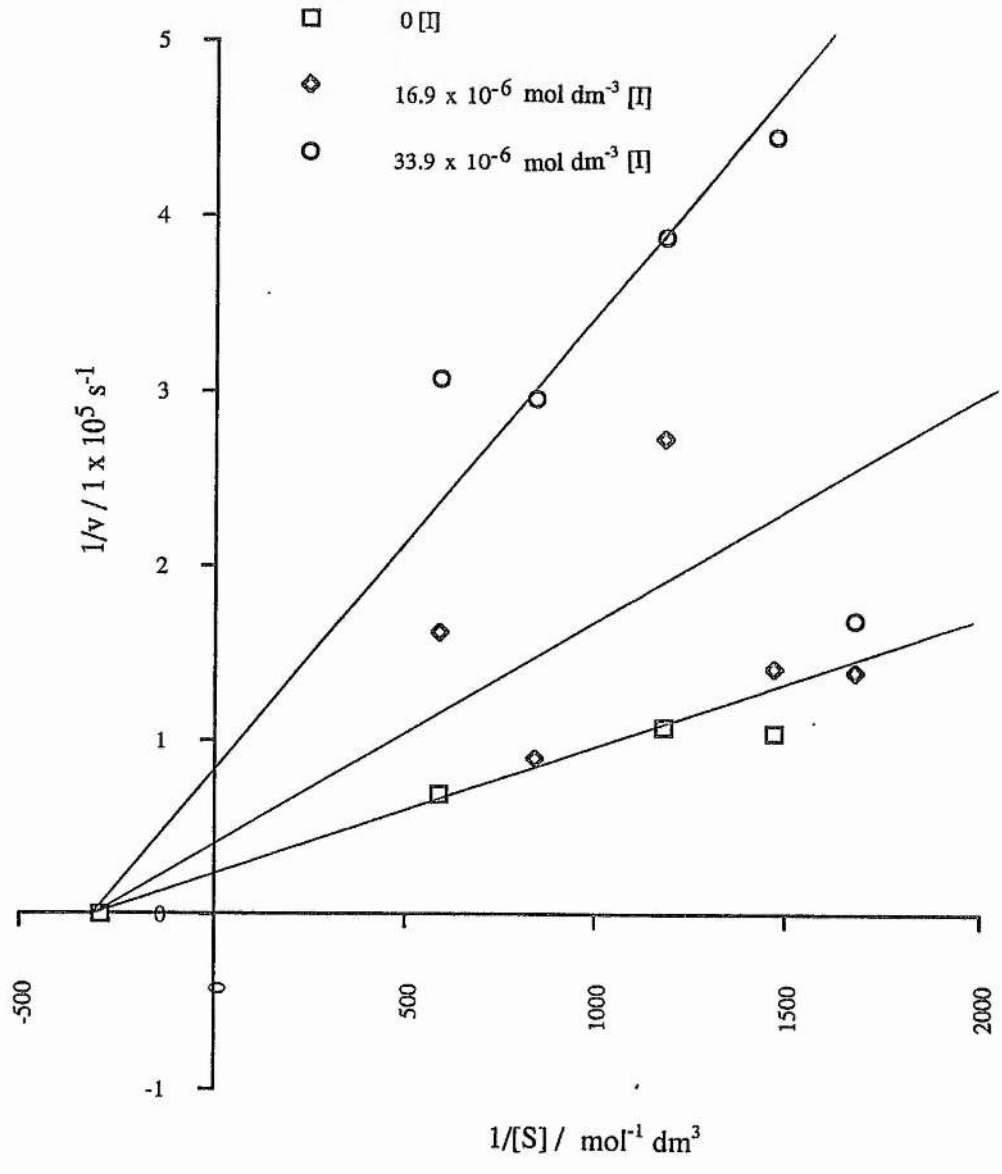
## the

## Inhibitors tested

## Preliminary double reciprocal plots for Cbz-Leu-Ala-Gly-aminoacetonitrile (78)

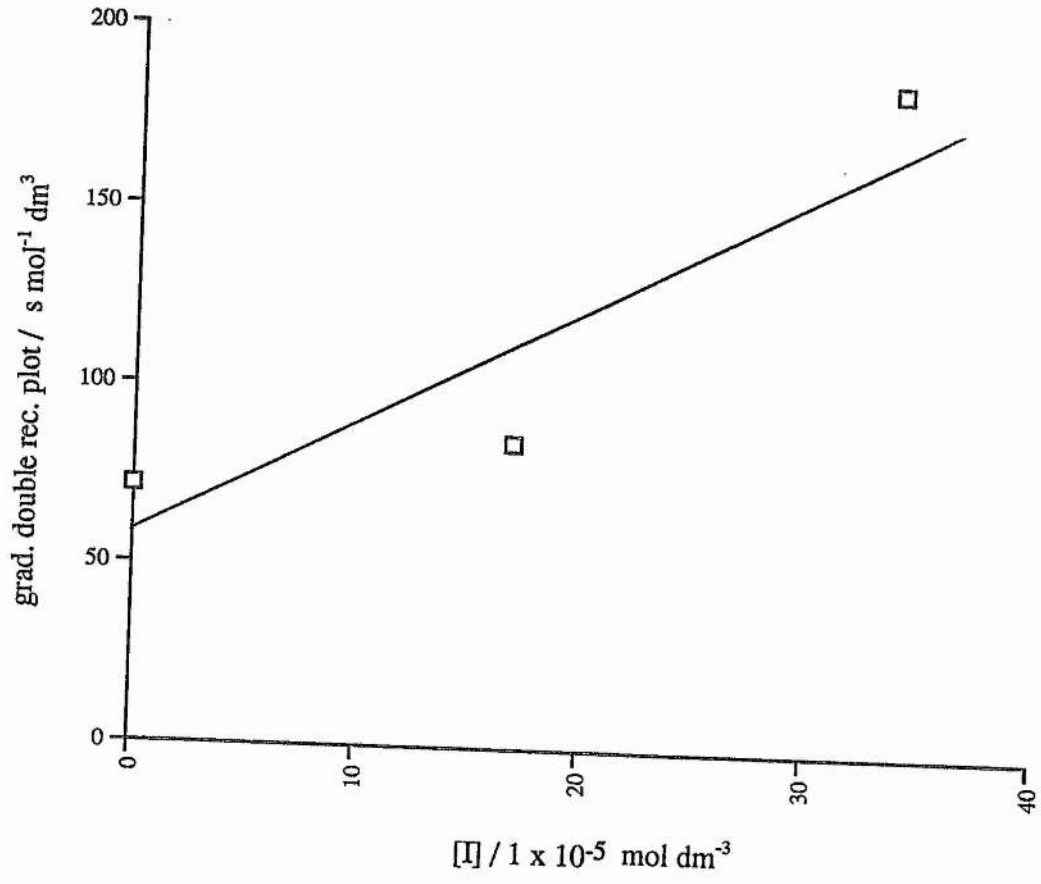
The double reciprocal plots are graphical representations of the mode of inhibition, for each inhibitor. The data points shown have been calculated from the rate of appearance of product, which were obtained from non-linear regressional analysis of the [P] verses time data, using the enzfitter program (Section 3.1.3, Experimental).

Each rate was determined once as problems were encountered with the proteinase preparation which was unstable and denatured readily. The best fit lines for the estimated type of inhibition are shown on each plot.

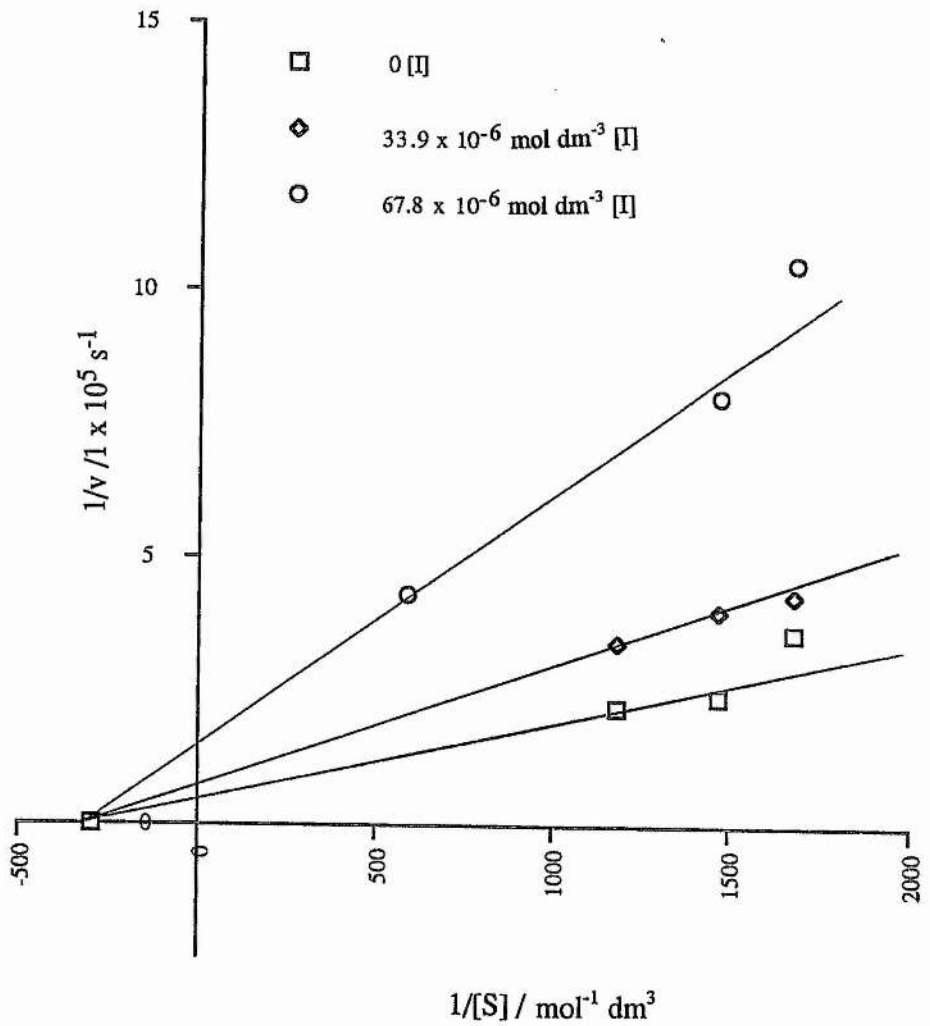


## Approximate $K_i$ determination for Cbz-Leu-Ala-Gly-aminoacetonitrile (78)

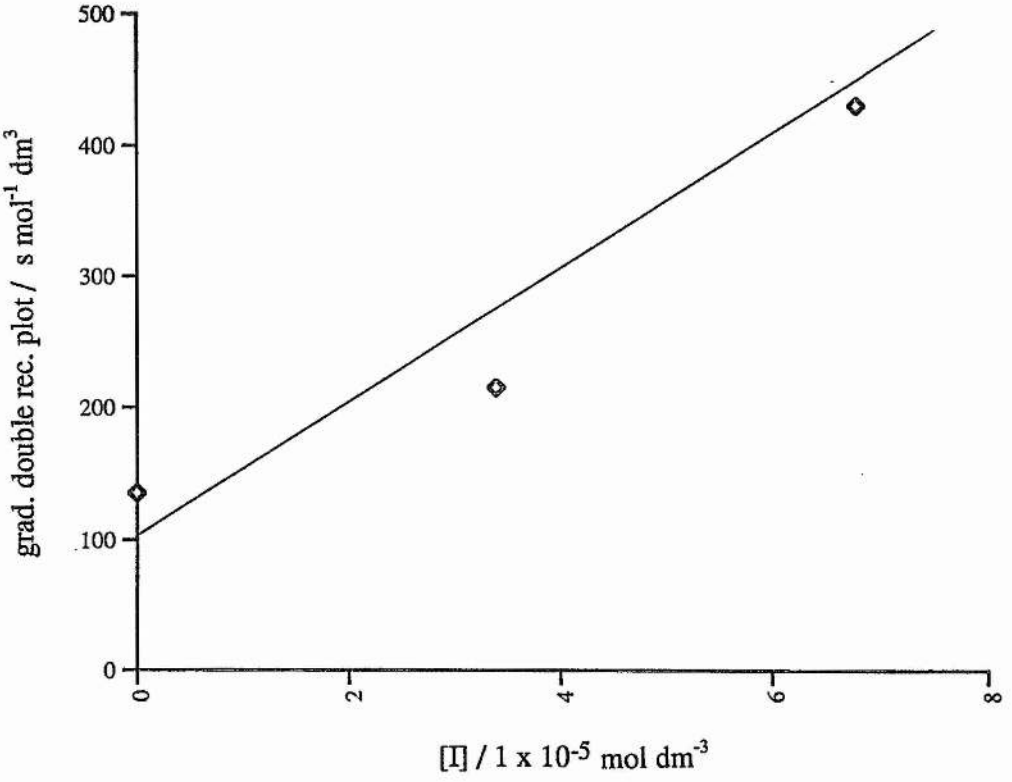
Dr Heather Murray also carried out similar rate determination experiments and also found that Cbz-Leu-Ala-Gly-aminoacetonitrile (**78**) was a noncompetitive inhibitor of the adenovirus proteinase, with a $K_i$ of 16 μmoldm$^{-3}$.[176] A $K_i$ of 15 μmoldm$^{-3}$ was determined from this experiment.



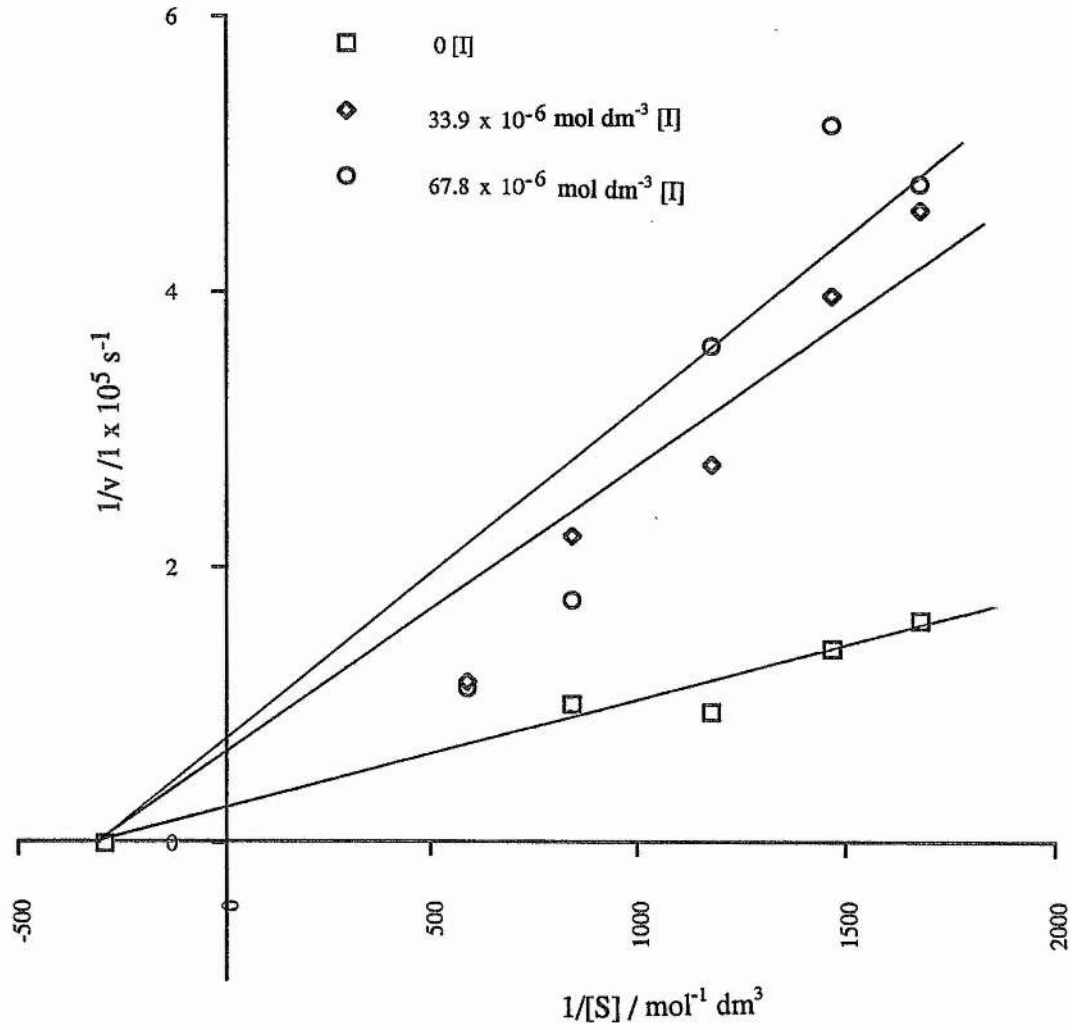
## Preliminary double reciprocal plots for t-Boc-Leu-Ala-Gly-aminoacetonitrile (79)



## Approximate $K_i$ determination for t-Boc-Leu-Ala-Gly-aminoacetonitrile (79)

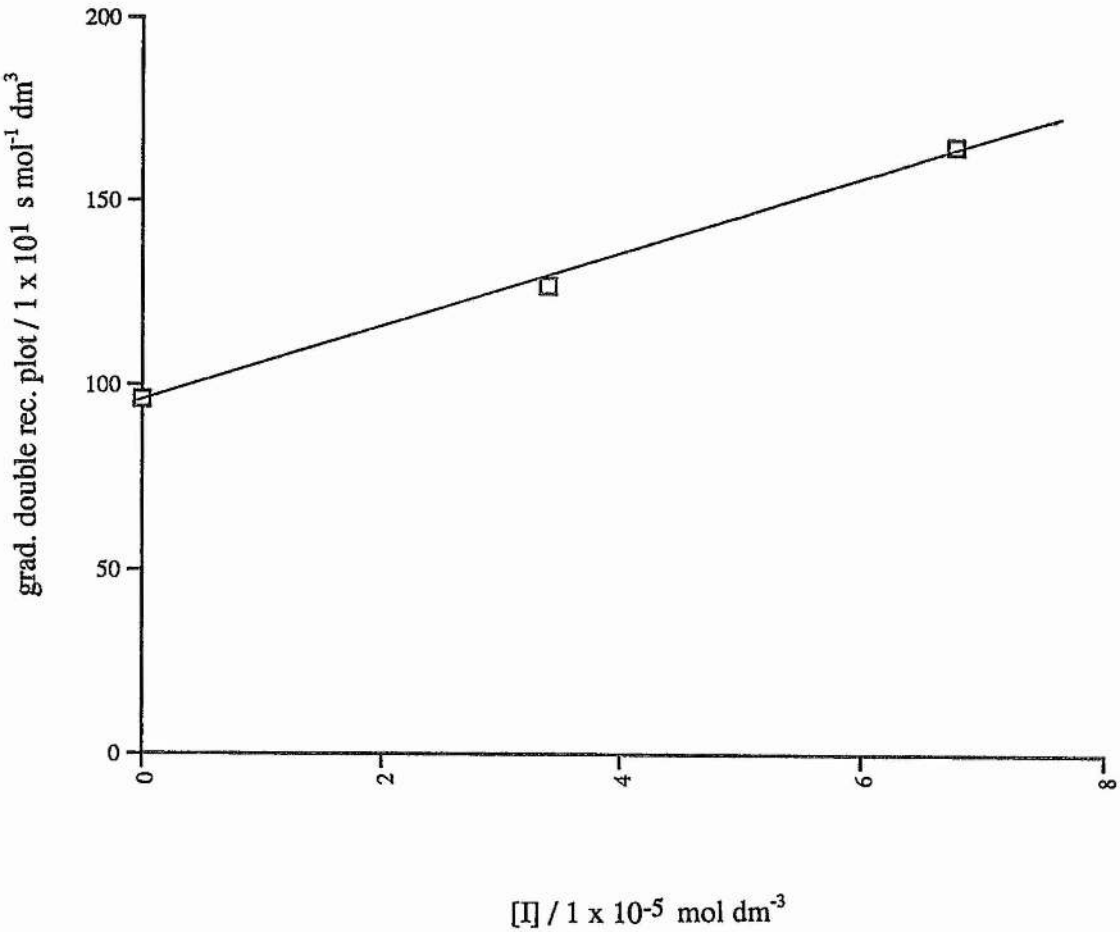


**Preliminary double reciprocal plots for Cbz-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal (114)**

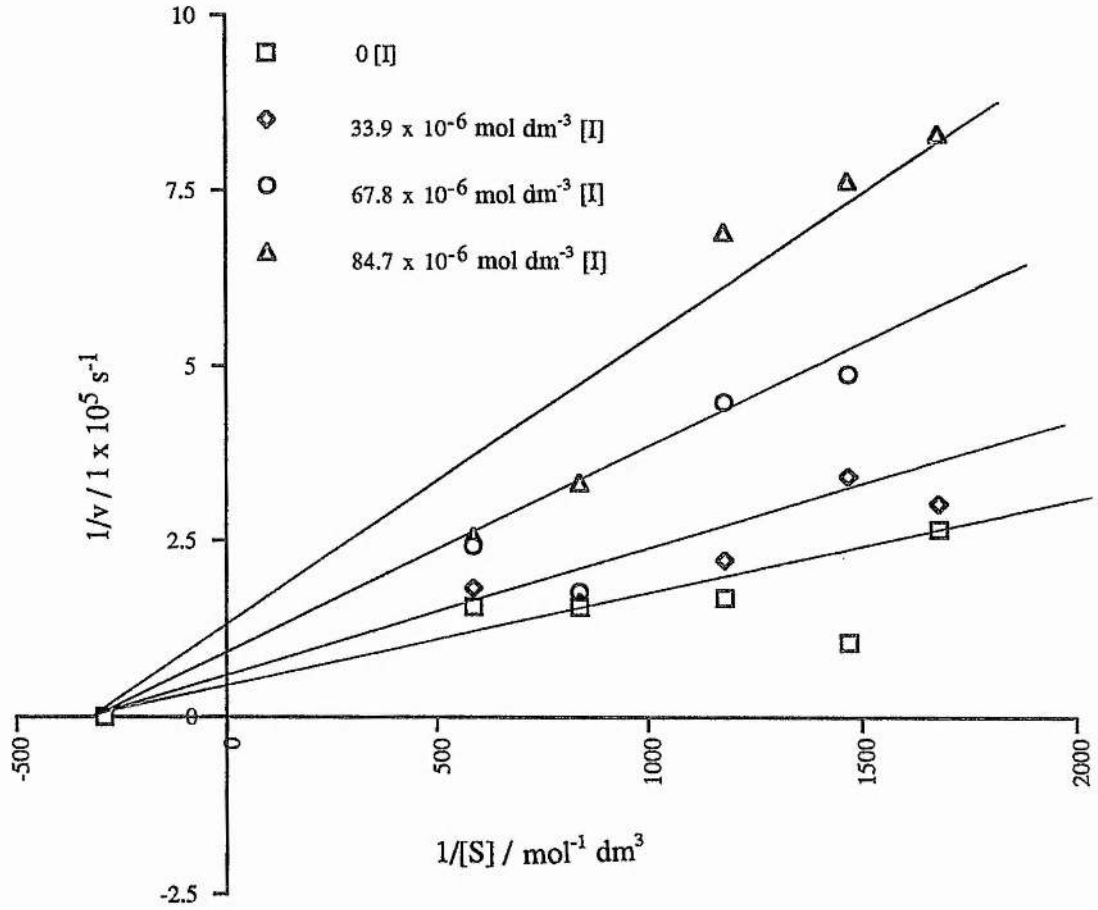


## Approximate $K_i$ determination for Cbz-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal (114)

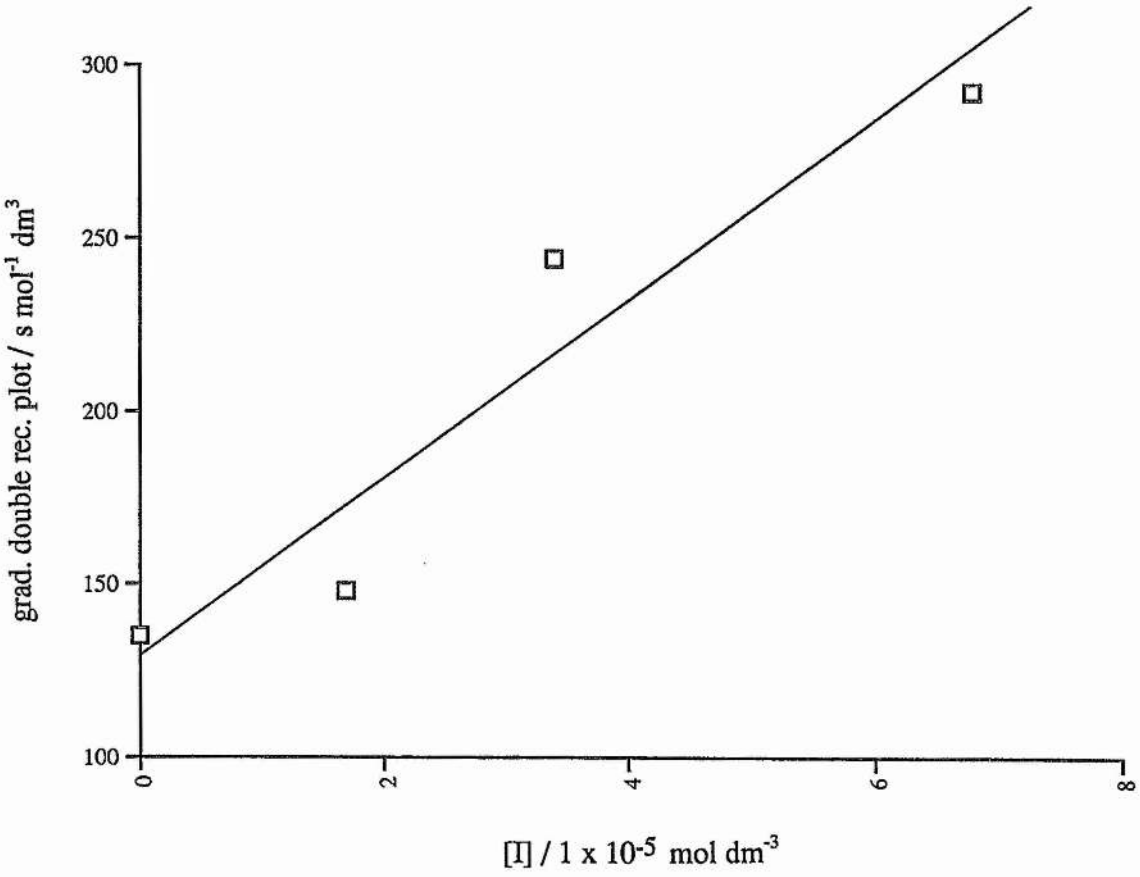
Dr Heather Murray also carried out similar rate determination experiments and also found that Cbz-Leu-Ala-Gly-aminoacetaldehyde dimethyl acetal (**114**) was a noncompetitive inhibitor of the adenovirus proteinase, with a $K_i$ of 380 μmoldm$^{-3}$.[176] A $K_i$ of 96 μmoldm$^{-3}$ was determined from this experiment.



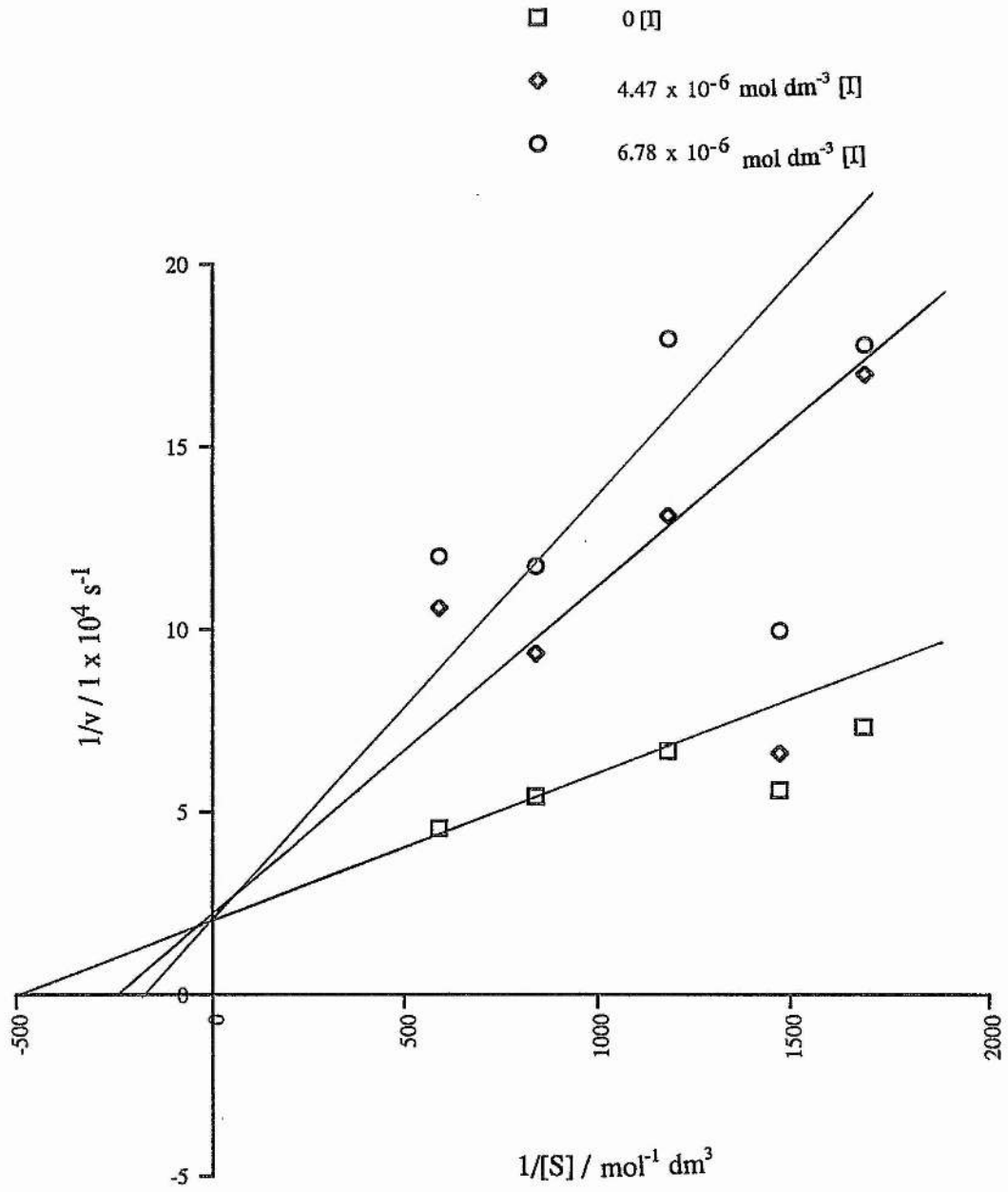
**Preliminary double reciprocal plots for Cbz-Leu-Ala-Gly-1-amino 2-methyl prop-2-ene (118)**



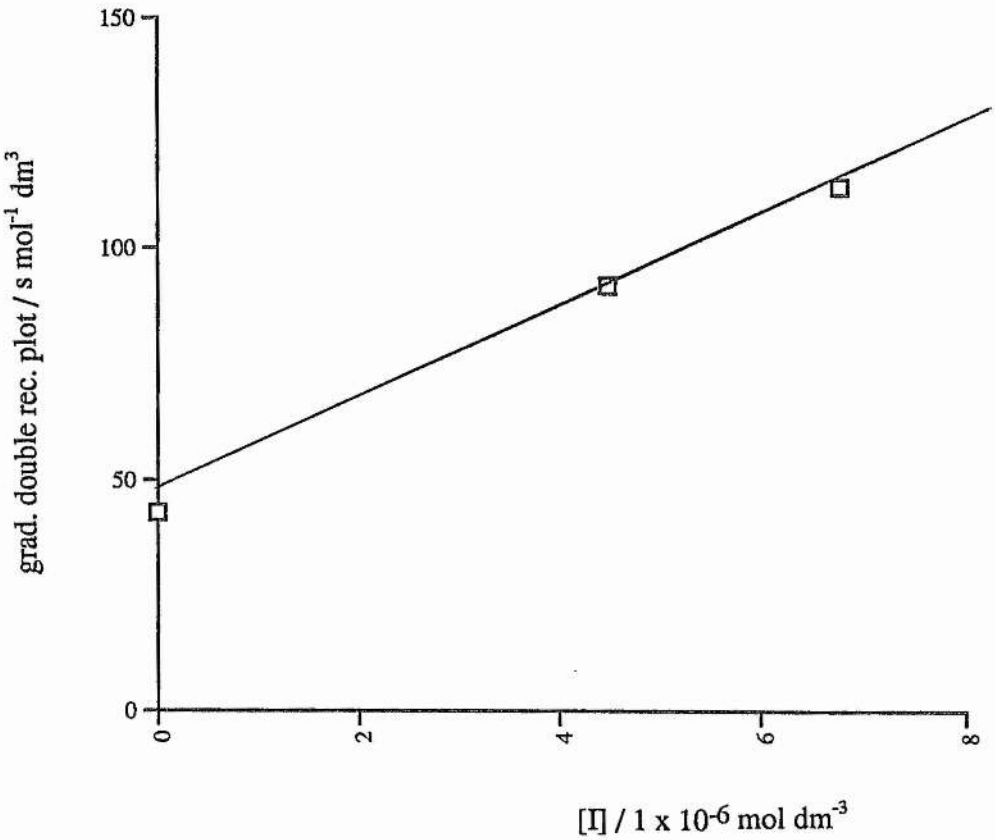
## Approximate $K_i$ determination for Cbz-Leu-Ala-Gly-1-amino 2-methyl prop-2-ene (118)



**Preliminary double reciprocal plot for N,N' (di(Cbz-Leu-Ala-Gly))-2,4-diamino but-2-ene aldehyde (122)**



## Approximate $K_i$ determination for N,N' (di(Cbz-Leu-Ala-Gly))-2,4-diamino but-2-ene aldehyde (122)



# 4.0 References


[1] W. P. Rowe, R. J. Huebner, L. K. Gilmore, R. H. Parrott and T. G. Ward, *Pro. Soc. Exp. Biol. Med.*, 1953, **84**, 570-573.

[2] W. J. Mogabgab, *Am. Rev. Respir. Dis.*, 1968, **97**, 345-358.

[3] M. S. Horwitz, *Virology*, ed. B. N. Fields and D. M. Knipe, Raven Press Ltd., New York, 1990, ch. 60-61, pp. 1679-1740.

[4] E. S. Murray, R. S. Chang, S. D. Bell jr., M. L. Tarizzo and J. C. Synder, *Am. J. Ophthalmol.*, 1956, **43**, 32; F. M. Bennet, B. M. Law, W. Hamilton and A. MacDonald, *The Lancet*, 1957, 670-673.

[5] J. M. Zahradnik, M. J. Spencer and D. M. Porter, *Am. J. Med.*, 1980, **68**, 725-732.

[6] W. R. Green, W. L. Greaves, W. R. Frederick and L. Taddesse-Heath, *Clinical Infectious Diseases*, 1994, **18**, 989-991.

[7] M. -L. Lähdeaho, M. Lehtinen, H.-R. Rissa, H. Hyöty, T. Reunala and M. Mäki, *Int. Arch. Allergy Immunol*, 1993, **101**, 272 - 276.

[8] H. J. Porter, C. J. H. Padfield, L. C. Peres, L. Hieschowitz and P. D. Berry, *J. Clin. Pathol.*, 1993, **46**, 154-158.

[9] D. K. Ford, H. B. Stein, M. Sculzer, E. D. Bateman, J. C. Hogg and S. Hayashi, *J.Rheum.*, 1993, **20**, 310-313.

[10] A. B. Martin, S. Webber, F. J. Fricker, R. Jaffe, G. Demmler, D. Kearney, Y-H. Zhang, J. Bodurtha, B. Gelb, J. Ni, J. T. Bricker and J. A. Towbin, *Circulation*, 1994, **90**, 330-339.

[11] H. Ibelgaufts, K. W. Jones, N. Maitland, N. F. Shaw, *Acta Neuropathol. (Berl)*, 1982, **56**, 113-117.

[12] R. C. Valentine and H. G. Pereira, *J. Mol. Biol.*, 1965, **13**, 13-20.

[13] P. L. Stewart and R. Burnett, *J. Appl. Phys.*, 1993, **32**, 1342-1347.

[14] U. Swensson, *J. Virol.*, 1985, **55**, 442-449.

[15] N. Louis, P. Fender, A. Barge, P. Kitts and J. Chroboczek, *J. Virol.*, 1994, **68**, 4104-4106.

[16] D. J. P. Fitzgerald, R. Padmanabhan, I. Pastan and M. C. Willingham, *Cell*, 1983, **32**, 607-617.

[17] P. Seth, H. Fitzgerald, M. Ginsberg, M. Willingham and I. Pastan, *Mol. Cell. Biol.*, 1984, **4**, 1528-1533.

[18] T. J. Wickham, P. Mathias, D. A. Cheresh and G. R. Nemerow, *Cell*, 1993, **73**, 309-319.

[19] P. Seth, H. Fitzgerald, M. Ginsberg, M. C. Willingham and I. Pastan, *J. Virol.*, 1984, **51**, 650-655.

[20] L. Philipson, K. Lonberg-Holm and U. Pettersson, *J. Virol.*, 1968, **2**, 1064-1075.

[21] R. B. Luftig and R. R. Weihing, *J. Virol.*, 1975, **16**, 696-706.

[22] P. Hoffman and C. Carlin, *Mol. Cell. Biol.*, 1994, **14**, 3695-3706.

[23] C. Tribouley, P. Lutz, A. Staub and C. Kedinger, *J. Virol.*, 1994, **68**, 4450-4457.

[24] J. Weber, *J. Virol.*, 1976, **17**, 462-471.

[25] T. Pe'ery, K. H. Mellits and M. B. Mathews, *J. Virol.*, 1993, **67**, 3534-3543.

[26] N. D. Stow, *EMBO J.*, 1982, **1**, 863-880; R. T. Hay, N. D. Stow and I. M. McDougall, *J. Mol. Biol.*, 1984, **175**, 493-510.

[27] R. J. Hay and W. C Russell, *Biochem.J.*, 1989, **258**, 3-16.

[28] C. W. Anderson, P. R. Baum and R. F. Gesteland, *J. Virol.*, 1973, **12**, 241-252.

[29] M. L. Tremblay, C. V. Déry, B. G. Tolbot and J. Weber, *Biochem. Biophys. Acta,* 1983, **743**, 239-245.

[30] M. -L. Boudin, J. -C. D'Halluin, C. Cousin and P. Boulanger, *Virol.,* 1980, **101**, 144-156.

[31] P. Freimuth and C. W. Anderson, *Virol.,* 1993, **193**, 348-355.

[32] Y. Zhang and R. J. Schneider, *J. Virol.*, 1994, **68**, 2544-2555.

[33] H. Maeda and R. G. Crystal, *Clinical Res.*, 1993, **41**, 206A.

[34] R. J. Cristiano, L. C. Smith and S. L. Woo, *Proc. Natl. Acad. Sci. USA,* 1993, **90**, 2122-2126.

[35] K. N. Wills and P. Menzel, *J. Cell. Biochem.*, 1993, **17E**, 204.

[36] S. Akli, C. Caillaud, E. Vigne, L. D. Stratford-Perricaudet, L. Poenaru, M. Perricaudet, A. Kahn and M. R. Peschanski, *Nature Genetics,* 1993, **3**, 224-228.

[37] E. M. Toloza, A. R. Miller, W. Guo and K. Rhoades, *J. Cell. Biochem.*, 1994, **18A**, 229.

[38] P. Lemarchand, M. Jones, C. Daniel, I. Yamada, A. Mastrangeli and R. G, Crystal, *Am. Rev. Respir. Dis.*, 1993, **147**, 28.

[39] B. C. Trapnell, N. Mittereder, S. Yei, Y. -N. Chang and P. Tolstoshev, *J. Cell. Biochem.*, 1993, **17E**, 204.

[40] *Nature Genetics,* 1993, **3**, 1-2.

[41] D. Yu, J. K. Wolf, M. Scanion, J. E. Price and M. C. Hung, *Cancer Res.*, 1993, **53**, 891-898.

[42] D. J. Slamon, G. M. Clark, S. G. Wong, W. J. Levin, A. Ullrich and W. McGuire, *Science,* 1987, **235**, 177-182.

[43] D. J. Slamon, W. Godolphin, J. A. Holt, S. G. Wong, D. E. Kieth, W. J. Levin, S. G. Stewart, J. Udove, A. Ullrich and W. McGuire, *Science,* 1989, **244**, 707-712.

[44] M. P. Schneider, M. C. Hung, S. M. Chiocca, J. Manning, X. Zhao, K. Fang and J. A. Roth, *Cancer Res.*, 1989, **49**, 4968-4971.

[45] J. Yokota, T. Yamamoto, N. Miyajima, K. Toyoshima, N. Nomura, H. Sakamoto, T. Yoshida, M. Terada annd T. Sugimura, *Oncogene*, 1988, **2**, 263-287.

[46] J. D'Emilia, K. Bulovas, K. D'Ercole, B. Wolf, G. Steele jr. and I. C. Summerhayes,

[47] C. A. Burton, B. Nayak and C. J. O'Conner, *J. Org. Chem.*, 1968, **33**, 572-575.

[48] C. D. Ritchie, *J. Am. Chem. Soc.*, 1975, **97**, 1170-1179.

[49] P. G. Gassman, P. K. Hodgson and R. L. Balchunis, *J. Am. Chem. Soc.*, 1976, **98**, 1275-1276.

[50] A. R. Katritzky and R. A. Y. Jones, *Chem. Ind.*, 1961, 722-727.

[51] R. J. Gillespie and T. Birchall, *Can. J. Chem.*, 1963, **41**, 148-155.

[52] K. Yates and T. A. Modro, *Acc. Chem. Res.*, 1978, **11**, 190-196; R. A. Cox and K. Yates, *Can. J. Chem.*, 1981, **59**, 2853-2863.

[53] P. Deslongchamps, *'Stereoelectronic Effects in Organic Chemistry'*, Pergamon Press, Oxford, 1983, pp. 101-118.

[54] C. L. Perin and G. M. Arrhenius, *J. Am. Chem. Soc.*, 1982, **104**, 2839-2842.

[55] J. T. Edwards, G. D. Derdall and S. C. Wong, *J. Am. Chem. Soc.*, 1978, **100**, 7023-7027.

[56] J. T. Edwards and S. C. R. Meacock, *J. Chem. Soc.*, 1957, 2000-2007.

[57] E. N. Baker and J. Drenth, in *Biological Macromolecules and Assemblies*, ed. F. A. Jurnak and A. McPherson, Wiley, New York, 1986, vol. 3, pp. 313-368.

[58] J. R. Whittaker and J. Perez-Villasenor, *Arch. Biochem. Biophys.*, 1968, **124**, 70-78.

[59] T. -Y. Liu and S. D. Elliot, in *The Enzymes 3*, ed. P. D. Boyer, Academic Press, New York, 3rd ed., vol 5, pp. 609-647.

[60] W. M. Mitchell and W. F. Harrington, in *The Enzymes 3*, ed. P. D. Boyer, Academic Press, New York, 3rd ed., vol 5, pp. 699-711.

[61] I. G. Kaphuis, K. H. Kalk, M. B. A. Swarte and J Drenth, *J. Mol. Biol.*, 1984, **179**, 233; Protein data bank Brookhaven national laboratory entry 9PAP.

[62] F. A. Johnston, S. D. Lewis and J. A. Shafer, *Biochemistry*, 1981, **20**, 52-58.

[63] L. Polgar, *FEBS Lett.*, 1974, **47**, 15-18

[64] F. A. Johnson, S. D. Lewis and J. A. Schafer, *Biochemistry*, 1981, **20**, 44-46.

[65] P. T. Van Duijnen, B. T. Thole and W. G. J. Hol, *Biophys. Chem.*, 1979, **9**, 273-280.

[66] W. G. J. Hol, P. T. Van Duijnen and H. J. C. Berendsen, *Nature*, 1978, **273**, 443-446.

[67] P. T. Van Duijnen, B. T. Thole, R. Broer and W. C. Nieuwpoort, *Int. J. Quantum Chem.*, 1980, **17**, 651-671.

[68] E. L. Smith, *J. Biol. Chem.*, 1958, **233**, 1392-1397.

[69] E. N. Baker in *Structural Studies on molecules of Biological Interest*, eds. G. G. Dodson, J. P. Glusker and D. Sayre, Clarendon Press, Oxford, Ch. 31, pp. 339-349.

[70] D. V. Roberts, *Enzyme Kinetics*, Cambridge University Press, Cambridge, 1977, pp. 53-42.

[71] C. G. Knight, *Research Monographs In Cell and Tissue Physiology*, 1986, **12**, 23-51.

[72] A. Webster, S. Russell, P. Talbot, W. C. Russell and G. D. Kemp, *J. Gen. Virol*, 1989, **70**, 3225-3234.

[73] B. D. Miles, R. B. Luftig, J. A. Weatherbee, R. R. Weihing and J. Weber, *Virol.*, 1980, **105**, 265-269.

[74] A. Webster and G. Kemp, *J. Gen. Virol*, 1993, **74**, 1415-1420.

[75] A. J. Barrett, *Mammalian Cells and Tissues*, ed. A. J. Barrett, Amsterdam, North-Holland, Amsterdam, 1977, pp. 1-55.

[76] R. J. Work and D. W. Burkitt, *Arch. Biochem. Biophys.*, 1979, **194**, 157-164.

[77] J. Marciniszyn, J. A. Hartsuck and J. Tang, *J. Biol. Chem.*, 1976, **251**, 7088-7094.

[78] J. S.Bond and P. E. Butler, *Ann. Rev. Biochem.*, 1987 **56**, 333-364

[79] A. R. Bhatti and J. Weber, *Virol.*, 1979, **96**, 478-485.

[80] A. Webster, W. C. Russell and G. D. Kemp, *J. Gen Virol.*, 1989, **70**, 3215-3223.

[81] F. Willenbrock and K. Brocklehurst, *Biochem. J.*, 1985, **227**, 521-528.

[82] T. Saino, T. Someno, H. Miyazaki and S.-I. Tshii, *Chem. Pharm. Bull.*, 1982, **30**, 2329-2325.

[83] A. Webster, R. T. Hay and G. Kemp, *Cell*, 1993, **72**, 97-104.

[84] C. W. Anderson, *Prot. Expr. Purif.*, 1993, **4**, 8-15.

[85] K. Brocklehurst and M. P. J. Kirstan, *Nature, New Biology*, 1973, **242**, 167-170.

[86] W. F. Mangel, R. J. McGrath, D. L. Toledo and C. W. Anderson, *Nature*, 1993, **361**, 274-275.

[87] F. Cai, D. Tang and J. M. Weber, *Biochim. Biophys. Acta*, 1992, **1129**, 339-341

[88] A. W. Grierson, R. Nicholson, P. Talbot, A. Webster and G. Kemp, *J. Gen. Virol.*, 1994, 75, 1-4.

[89] C. Rancourt, K. Tihanyi, M. Bourbonniere and J. M. Weber, *Proc. Natl. Acad. Sci, USA*, 1994, **91**, 844-847.

[90] E. Fischer, *Chem. Ber.*, 1903, **36**, 2094-2106.

[91] T. Curtius, *Chem.* Ber., 1902, **35**, 3226-3228.

[92] M. Bodansky, *Principles of Peptide Synthesis*, 1984, Springer-Verlang, Berlin, pp. 28-36.

[93] T. Wieland, W. Kern and R. Schering, *Ann. Chem.*, 1950, **569**, 117-122.

[94] J. C. Sheehan and G. Hess, *J. Am. Chem. Soc.*, 1955, 77, 1067-1068.

[95] G. W. Anderson, J. E. Zimmerman and F. M. Callahan, *J. Am. Chem Soc.*, 1967, **89**, 5012-5017.

[96] E. Kleiger, E. Schröder and H. Gibian, *Liebigs Ann. Chem.*, 1961, **640**, 157-167.

[97] S. Goldscmidt and K. Gupta, *Chem. Ber.*, 1965, **98**, 2831-2836.

[98] M. Bergmann, L. Zervas, J. S. Fruton, F. Schnider and H. Schleinch, *J. Biol. Chem*, 1949, **109**, 325-346.

[99] W. J. Polglase and E. L. Smith, *J. Am. Chem. Soc.*, 1949, **71**, 3081-3084.

[100] S. Chandrasekaran, A. F. Klunge and J. A. Edwards, *J. Org. Chem.*, 1977, **42**, 3973-3974.

[101] P. G. Gassman and W. N. Schenk, *J. Org. Chem.*, 1977, **42**, 918-920.

[102] W. S. Broadbent, J. S. Morley and B. E. Stone, *J. Chem. Soc. C*, 1967, 2632-2636.

[103] N. Fujii, A. Otaka, S. Furakoshi, M. Nomaizu, K. Akaji, H. Yajima, K. Kitagawa, T. Akita, K. Ando and T. Kawamoto, *Chem. Pharm. Bull.*, 1986, **34**, 603-612.

[104] K. Esko and S. Karlsson, *Acta Chem. Scand.*, 1970, **24**, 1415-1422.

[105] L. Zervas, M. Winitz and J. P. Greenstein, *J. Org. Chem.*, 1957, **22**, 1521-1515.

[106] M. Sakaitani, N. Kurokawa and Y. Ohfune, *Tetrahedron Lett.*, 1986, **27**, 3753-3754.

[107] E. T. Roe, J. T. Scanlan and D. Swern, *J. Am. Chem. Soc.*, 1949, **71**, 2215-2218.

[108] H. E. Johnson and D. G. Crosby, *J. Org. Chem.*, 1962, **27**, 798-802.

[109] C. S. Marvel, J. R. Elliot, F. E. Boettner and H. Yuska, *J. Am. Chem. Soc.*, 1946, **68**, 1681-1686.

[110] G. L. Stahl, R. Walter and C. W. Smith, *J. Org. Chem.*, 1978, **43**, 2285-2286.

[111] V. Dourtoglou and B. Gross, *Synthesis*, 1984, 572-573.

[112] W. C. Schubert, *J. Mol. Biol.*, 1972, **67**, 151-157.

[113] G. Losse and E. Demuth, *Chem. Ber.*, 1961, **94**, 1762-1766.

[114] T. Curtius and F. Goebel, *J. Prakt. Chem.*, 1888, **37**, 150.

[115] N. F. Albertson and F. C. McKay, *J. Am. Chem. Soc.*, 1953, **75**, 5323-5326.

[116] T. Wieland and B. Heinke, *Liebigs Ann. Chem.*, 1956, **599**, 70-80.

[117] R. W. Chambers and F. H. Carpenter, *J. Am. Chem. Soc.*, 1955, 77, 1522-1526.

[118] L. A. A.E. Sluyterman, *Biochim. Biophys. Acta*, 1964, **85**, 316-321.

[119] C. A. Lewis and R. Wolfenden, *Biochemistry*, 1977, **16**, 4890-4895.

[120] J. B. Moon, R. S. Coleman and R. P. Hanzlik, *J. Am. Chem. Soc.*, 1986, **108**, 1350-1351.

[121] L. A. A.E. Sluyterman and J. Widjenes, *Biochem. Biophys. Acta*, 1973, **302**, 95-101.

[122] I. Tomida, T. Nunosawa and N. Mochizuki, *Agric. Biol. Chem.*, 1977, **41**, 2407-2412.

[123] F. Weygand and R. Geiger, *Chem. Ber.*, 1956, 89, 647-656.

[124] E. E. Schallenberg and M. Calvin, *J. Am. Chem. Soc.*, 1955, 77, 2779-2783.

[125] R. Jay and T. Curtius, *Chem. Ber.* 1894, **27**, 59-62.

[126] J. K. Kochi, *J. Org. Chem*, 1965, **30**, 3265-3271; J. K. Kochi, *J. Am. Chem. Soc*, 1965, **87**, 2500-2502.

[127] M. E. Jung and M. A. Lyster, *J. Chem. Soc., Chem. Comm,*. 1978, 315-316.

[128] M. P. Gamcsik, J. P. G. Malthouse, W. U. Primrose, N. E. Mckenzie, A. S. F. Boyd, R. A. Russell and A. I. Scott, *J. Am. Chem. Soc.*, 1983, **105**, 6324-6325.

[129] M. R. Bendall, I. L. Cartwright, P. I. Clark, G. Lowe and D. Nurse, *Eur. J. Biochem*, 1977, **79**, 201-209.

[130] G. Lienhard and W. J. Jencks, *J. Am. Chem. Soc.*, 1966, **88**, 3982-3995.

[131] E. Shaw and C. Ketther, *Acta Biol. Med. Ger.*, 1981, **40**, 1503-1511.

[132] D. H. Rich in *Proteinase Inhibition*, eds. A. J. Barett and G. Salvesen, Elsever Science Publishers BV (Biomedical Division), Amsterdam, 1986, Ch. 4, pp. 153-178.

[133] M. E. Yung, W. A. Andrus and P. L. Ornstein, *Tetrahedron Lett.* 1977, **48**, 4175-4178.

[134] D. Gani and D. Young, *J. Chem. Soc,. Perkin Trans. I*, 1983, 2393-2398.

[135] S. Nahm and S. M. Weinreb, *Tetrahedron Lett.*, 1981, **22**, 3815-3818.

[136] F. P. Cossio, B. Lecea and C. Palomo, *J. Chem. Soc., Chem. Commun.*, 1987, 1743-1744; D. Gupta, R. Soman and S. Dev, *Tetrahedron*, 1982, **38**, 3013-3018.

[137] R. P. Hanzlik and S. A. Thompson, *J. Med. Chem.*, 1984, **27**, 711-712.

[138] V. Bocchi, G. Casnati, A. Dossena and R. Marchelli, *Synthesis*, 1979, 961-962.

[139] K. Hanada, M. Tamai, S. Ohmura, J. Sawada and I. Tanaka, *Agric. Biol. Chem.*, 1978, **42**, 523-528.

[140] A. J. Barrett, A. A. Kembhavi, M. A. Brown, H. Krischke, C. G. Knight, M. Tamai and K. Hanada, *Biochem. J.* 1982, **201**, 189-198.

[141] Y. Yabe, D. Guillaume and D. H. Rich, *J. Am. Chem. Soc.*,1988, **110**, 4043-4044.

[142] A. Hassner and V. Alexanian, *Tetrahedron Lett.*, 1978, **46**, 4475-4478.

[143] M. A. Brook and T. H. Chan, *Synthesis*, 1983, 201-203.

[144] A. Korn, S. Rudolph-Böhner and L. Moroder, *Tetrahedron*, 1994, **50**, 8381-8392.

[145] F. R. Atherton, H. T. Openshaw and A. R. Todd, *J. Chem. Soc., Trans.*, 1945, 660-663.

[146] R. Ramage, D. Hopton, M. J. Parrot , G. W. Kenner and G. A. Moore, *J. Chem. Soc., Perkin Trans. I*, 1984, 1357 - 1370.

[147] L. Zervas and D. M. Theodoropoulos, *J. Am. Chem. Soc.*, 1956, **78**, 1359-1363.

148 A. I. Kosak, C. W. Huntington and I. R. MacGregor, *J. Am. Chem. Soc.*, 1954, **76**, 4481-4482.

149 C. F. Huebner, R. Lucas, H. B. MacPhillamy and H. A. Troxell, *J. Am. Chem. Soc.*, 1955, 77, 469-471.

150 T. Satoh and S. Suzuki, *Tetrahedron Lett.*, 1969, **52**, 4555-4558.

151 N. Umino, T. Iwakuma and N. Itoh, *Tetrahedron Lett.*, 1976, **10**, 763-766.

152 H. C. Brown and P. Heim, *J. Org. Chem.*, 1973, **38**, 912-916.

153 S. Iriuchijima, *Synthesis,* 1978, 684-685.

154 K. Mori and H. Iwasawa, *Tetrahedron,* 1980, **36**, 87-90.

155 M. Akhtar, Ph.D. Thesis, University of Southampton, 1988.

156 G. B. Payne and P. H. Williams, *J. Org. Chem.*, 1959, **24**, 54-55.

157 M. Tamai, T. Adachi, K. Oguma, S. Morimoto, K. Hanada, S. Ohmura and M. Ohzeki, *Agric. Biol. Chem.*, 1981, **45**, 675-679.

158 I. Kisfaludy, J. E. Roberts, R. H. Johnson, G. L. Mayers and J. Kovacs, *J. Org. Chem.*, 1970, **35**, 3563-3565.

159 T. W. Greene and P. G. M. Wuts, *Protective Groups in Organic Synthesis*, John Wiley and Sons, Inc., 1991, 2nd edn., p. 442.

160 G. Lowe and A. Williams, *Biochem. J.*, 1965, **96**, 199-204; E. L. Smith, V. J. Chavra and J. M. Parker, *J. Biol. Chem.*, 1958, **230**, 283-293.

161 M. H. O'Leary and M. D. Kluetz, *J. Am. Chem. Soc.*, 1972, **94**, 665-666; M. H. O'Leary, M. Urberg and A. Young, *Biochemistry*, 1974, **13**, 2077-2081.

162 F. A. Long and M. Purchase, *J. Am. Chem. Soc,* 1950, **73**, 3267-3273; A. R. Olson and R. J. Miller, *J. Am. Chem. Soc.*, 1938, **60**, 2687-2697.

163 M. L. Bender, H. Ladenheim and M. C. Chen, *J. Am. Chem. Soc.*, 1961, **83**, 123-127.

164 C. K. Ingold, *Structure and Mechanism in Organic Chemistry*, Cornell University, Ithaca, 1969, 2nd edn., pp. 1137-1142.

[165] C. A. Bunton, A. E. Comyns, J. Graham and J. R. Quayle, *J. Chem. Soc.*, 1955, 3817-3824.

[166] A. Albeck, R. Persky and S. Kliper, *Bioorg. Med. Chem. Lett.*, 1995, **5**, 1767-1772.

[167] J. O'C Westerik, Ph. D. Thesis, 1974, University of North Carolina, Chapel Hill.

[168] E. C. Lucas and A. Williams, *Biochemistry*, 1969, **8**, 5125-5135.

[169] D. D. Perrin and W. L. F. Armarego, *Purification of Organic Chemicals*, Pergamon Press, Oxford, 1988, 3rd edn.

[170] J. F. Coetzee, *Pure and Appl. Chem.*, 1966, **13**, 429-435.

[171] H. Lund and J. Bjerrum, *Chem. Ber*, 1931, **64**, 210-214.

[172] J. F. Coetzee and T. -C. Chang, *Pure and Appl. Chem.*, 1985, **57**, 633-638.

[173] *Vogel's Text Book of Practical Organic Chemistry*, ed. B. S. Furniss, A. J. Hannaford, P. W. G. Smith and A. R. Tatchell, Longman Scientific and Technical, Harlow, 5th edn, method 3, p. 432.

[174] C. M. Davidson and R. F. Jameson, *Chem. Ind. (Lond.)*, 1963, 1686-1687.

[175] Robin J. Leatherbarrow (Biosoft, 222 Hills Road, Cambridge, CB2 1JP).