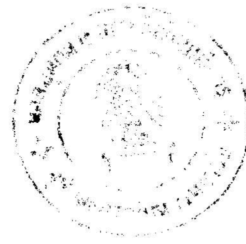


CHAPTER I

THEORY AND BACKGROUND

1.1 The organic chemistry of peptides

Peptides are composed of α-amino acids joined together by amide bonds. The simplest of these compounds posseses only two amino acid units per molecule, whereas more complex substances may incorporate a large number of amino acids. Peptides of relatively low molecular weight (M< ~10,000) are called polypeptides, the higher molecular weight materials, are proteins. Small polypeptides, sometimes referred to as oligopeptides, are given names which indicate the number of amino acid units in the molecule; dipeptides posses two such units; tripeptides three; tetrapeptides four; and so on.[1]

At first, it may seem surprising that compounds which are based upon such a simple structural unit can fulfil so wide a diversity of functions. However, the formulation of peptides as polyamides of α-amino acid is substantiated by a large body of evidence, which involves the accumulated results of physical, degradative and synthetic investigations. Similarly, it might be thought that the organic chemistry of peptides will be somewhat limited and uninformative, but this is not the case. The apparent simplicity of the system is deceptive. It is true that peptide chemistry is relatively specialized, but this impression is mainly dued to the peculiar technical difficulties involved in the formation of peptide bonds. These difficulties are mainly dued to the

process of racemization. Investigations of peptide structures and functions call for a sound appreciation of the basic principles of organic chemistry, and studies of peptides may be used with singular success to illustrate the importance of these principles.

1.1.1 Peptide synthesis

Emil Fischer was the first who synthesized the simplest dipeptide, glycylglycine, in 1901. Even that time, Fischer appreciated the significance of being able to synthesize peptides of defined sequence.[2] Max Bergmann and Leonidas Zervas [3] introduced to peptide synthesis the carbobenzoxy derivative for the protection of the α-amino group of amino acids and peptides in 1932.

However, chemists have since developed two new types of methods for synthesizing a peptide with the desired sequence of amino acids.

1.1.1.1 Classical method

The classical method or the liquid method is still one of the most widely used methods of peptide synthesis.[4] Reactive groups of the peptides that do not participate in the coupling reaction are protected. The unprotected groups of the components are coupled, either by prior activation of the carboxyl group or by the use of various coupling reagents.

When the two amino acids have been joined together, one or more of the protecting groups is removed, and the next, partially protected amino acid is coupled to the dipeptide. The deprotection/protection process is replaced and the desired amino acids are added to the growing chain until the desired kind and number of amino acids have been joined together in the required sequence.

1.1.1.2 Solid phase method

The introduction of solid phase synthesis in 1963 by R.B. Merrifield proved to be a giant step forward.[5] In this technique, the -COOH terminal of an amino acid is attached to polystyrene resin beads. Then the next, protected amino acid is coupled to the first one and the protective group is removed. After each amino acid is added, excess reagents are washed away and the necessity of isolating each intermediate is eliminated. When the chain is complete, the peptide is cleaved from the resin. The entire process has been automated. The great advantages of the solid phase technique are the ease of operation and the high overall yield.

For this study, all peptide syntheses were carried out by the classical method.

1.1.2 Peptide bond formation

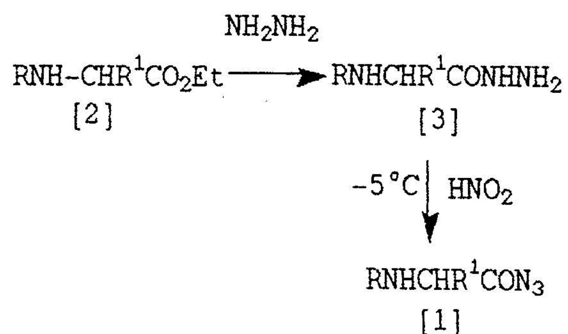
Some methods of peptide bond formation are shown below:

1.1.2.1 Acid chloride method

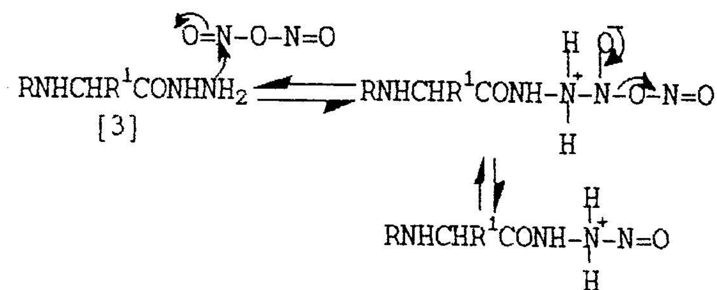
Fischer and Otto[6] used acid chlorides for peptide synthesis in 1903. Ethyl esters were used for carboxyl protection. Most amino acid chlorides are unstable, so now this method is seldom used.

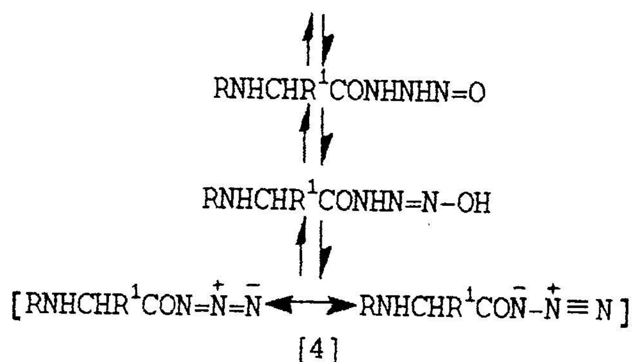
1.1.2.2 The azide coupling method

This method was introduced by Curtius[7] in 1902 and is still used widely. Azides [1] can be produced readily from α-amino acid esters [2] by the reaction of hydrazine followed by the action of a nitrosating agent.

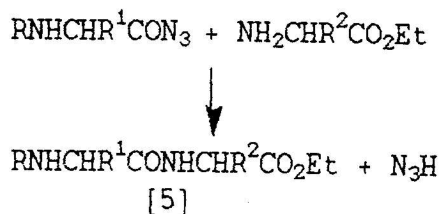


Azide formation presumably occurs the following way:

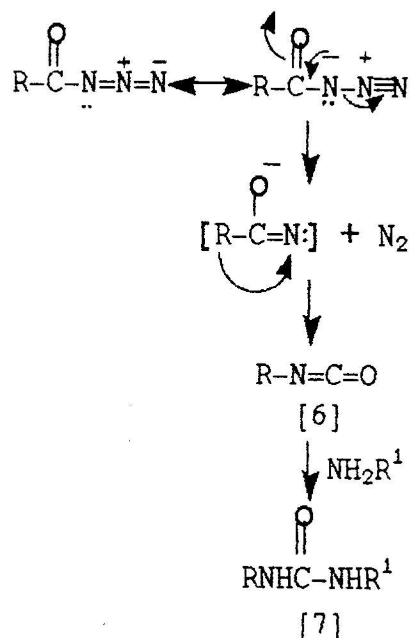




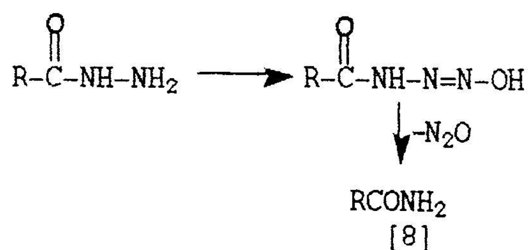
The intermediate hydrazides [3] are usually quite stable and even crystalline, whereas the acyl azides [4] are unstable and are not isolated. They react with the amino group to yield the required peptide [5] and hydrazoic acid. The latter is such a weak acid (pK_a=5) that it is not necessary to add excess base during the coupling to ensure its neutralization.



The azide method of coupling is subject to several undesirable side reactions. The azide undergoes a Curtius rearrangement to produce the corresponding isocyanate [6]. Of course, isocyanates in general are susceptible to nucleophilic addition and this can occur intermolecularly or intramolecularly. Thus, urea derivatives [7], which are often similar in properties to the required peptide and therefore difficult to separate from it, may be produced by the attack of the amino groups on the isocyanate. The Curtius rearrangement occurs when acyl azides are warmed and hence it is important to keep the reaction mixture cold.



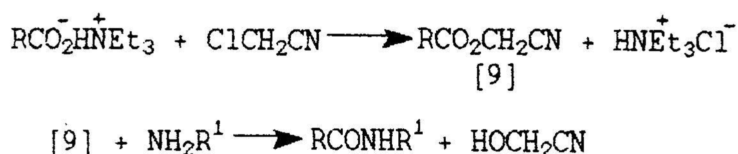
Another impurity, which is sometimes formed during azide coupling reactions, is the primary amide [8] analogous to the acyl hydrazide. It is thought that the amide does not arise due to a rearrangement of the azide itself. If the azide is prepared from the acid chloride and sodium azide, amide formation will not be detected. More likely, the amide is formed directly from the hydrazide. Amide formation is at a minimum when the water content of the reaction mixture is low. Consequently, t-butyl nitrate and nitrosyl chloride, which can be used in anhydrous solvents, have been advocated as preferred nitrosating reagents.



The azide coupling remains important in spite of its several shortcomings because it seems to be entirely free from racemization. The azide procedure is the method of choice for the preparation of histidyl and seryl peptides.(8-11)

.1.1.2.3 The active ester method

Since peptide bond formation, even from amino acid methyl esters, is thermodynamically favoured, it might be expected that this type of reaction would be even more feasible if the methyl moiety were replaced by stronger electron-withdrawing groups. Thus, cyanomethyl esters [9] readily formed from a tertiary base salt of the carbonyl component and α-chloroacetonitrile in an inert solvent, are aminolyzed rapidly and have been employed in peptide synthesis. Esters like this, which have enhanced rates of aminolysis, are referred to, somewhat loosely, as active esters.



The preparation of active esters derived from acidic phenols is best accomplished by the use of a diimide reagent. Esters from the less acidic phenols can be prepared by the action of the phenol on a mixed anhydride of the N-protected amino acid, although other approaches may be preferred. One of

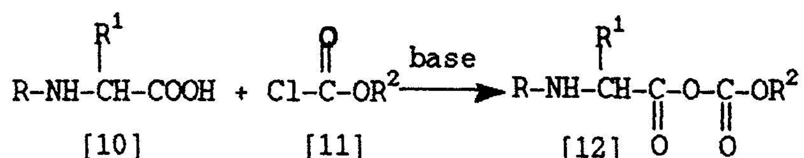
the attractions of the active ester approach is that the esters can be prepared and isolated in an optically pure, generally crystalline form quite independently of the coupling reaction.

1.1.2.4 The mixed anhydride method

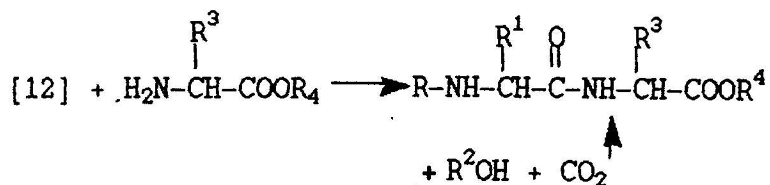
Wieland and Sehring(12) introduced the mixed anhydride coupling in 1950. The anhydrides with monoesters of carbonic acid and with organic acids will be discussed here.

1.1.2.4.1 Mixed anhydrides with monoesters of carbonic Acid

The anhydrides are formed by the reaction (13) of N-protected amino acids [10] with various esters of chloroformic acid [11].



Aminolytic cleavage of the anhydride yields only CO_2 and an alcohol, in additon to the peptide.



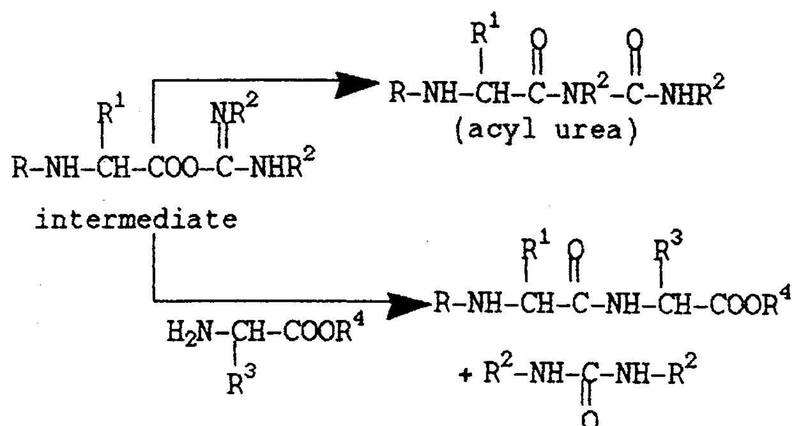
The condensation of the acylamino group and chloroformate require only a few minutes at -5 to -10 °C.[14,15] With temperatures as high as 10 °C, the reaction times might be up to 30 minutes.[16] The base, triethylamine, is used to neutralize the HCl formed. Few side reactions are encountered with this method of bond formation. Aminolytic cleavage of the anhydride at the wrong place may occur on rare occasions.[17,18] No racemization occurs when protected amino acids are coupled.

1.1.2.4.2 Mixed anhydrides with organic acids

The organic acids that have been used with good results have been very limited in number. The sterically hindered isovaleric[19] and pivalic acid[20] have been the few examples known.

1.1.2.5 The carbodiimide method

Sheeha and Hess[21] introduced N,N'-dicyclohexyl carbodiimde coupling in 1955. The intermediate formed will be changed either to the desired peptide or to an acyl urea.[22]



Acyl urea formation is the side reaction in this method. The use of strong bases also favors this side reaction. However, the use of low temperature; (0 °C) can reduce this side reaction. Dicyclohexylurea,the by-product of N,N-dicyclohexylcarbodiimide formation, is only slightly soluble in most solvents. Removing the last traces of the urea from the product is sometimes difficult.

The method used for peptide bond formation in this study was the mixed anhydride method, using monoesters of carbonic acid.

1.2 Human polymorphonuclear leukocyte proteolytic enzymes

The granules of human polymorphonuclear leukocytes contain various enzymes.[23] The three most abundant and well defined neutral proteinases of polymorphonuclear leukocytes are elastase, cathepsin G and collagenase. This study concerns the first two enzymes only.

Human leukocyte elastase (HLE) is a serine proteinase located in the azurophil granules.[23] The enzyme is a

neutral proteinase and exists as at least three isoenzymes that are immunologically indistinguishable and all have molecular weights of about 30,000 Daltons[24] with approximately 218 amino acid residues.[25,26] They can be termed glycoproteins since about twenty percent of the total weight of the enzyme is composed of saccharide units bound to the protein.[24] In addtion, HLE is an endopeptidase, able to hydrolyze the amide bonds of many proteins, cell component and various tissue types, including the oxidized β-chain of insulin[27] ,the structural proteins of glomerular basement membrane[28], components of the human component system[29],fibrinogen[30],fibronectin[31], α_2-plasma inhibitor[32], coagulation factors[33], proteoglycans[34], collagen types I,II,IV[27] and III[35],as well as elastin.[36] Furthermore, HLE activity against elastin has been shown to be stimulated 5-fold by human leukocyte cathepsin G.

Cathepsin G is a serine proteinase present in azurophil granules.[23] The enzyme is a chymotrypsin-like neutral proteinase, with a M_r of approximately 28,000 Daltons.[37] It is able to hydrolyze cartilage proteoglycans and insoluble collagen.[38] The enzyme also causes conversion of angiotensin I to angiotensin II that is associated with the inflammatory process[39]and conversion of angiotensinogen to angiotensin II.[40]

1.3 Inflammatory diseases, pulmonary emphysema and arthritis

The enzymes under study are not synthesized on demand, rather, they are stored in the cell in an inactivated form, i.e. complexed with a specific, natural inhibitor. This is known as α_1-proteinase inhibitor(α_1-PI) or antitrypsin(α_1-AT). This inactivation[41] is absolutely necessary for the cell not to be lysed by the enzyme, which, in spite of its name, is not entirely specific. α-PI contains a number of free -SH groups, which are very sensitive to oxidation[42], particularly by reaction with oxidative free radicals. The inflammatory process is a very complicated cascade of events. While it is, as yet, not completely understood, there is ample evidence to show the formation of various peroxide and hydroperoxide radicals. Since inhaled oxidizing agents, such as cigarette smoke or urban pollution have been shown to oxidise those -SH groups in α_1-PI which are associated with the maintenance of its inhibitory conformation, it is not surprising to observe the liberation of those proteolytic enzymes under those conditions, ultimately leading to the progressive destruction of the air-ways of the lung. This is the process which is associated with emphysema. It should be noted, that a significant portion of the human population is known to be deficient in α_1-PI[43] and therefore particularly prone to destructive disease. This deficiency is genetically determined, and currently there are

development of α_1-PI replacement therapy. Unfortunately, α_1-PI is available in practical quantities only from animal sources, and the immunological problems have not, as yet, been solved.

In arthritis, the degradative processes are similar in that the presence of oxidizing free radicals have been conclusively shown in inflamed joints and tissues. It should be pointed out that arthritis is one of the most common diseases of mankind and, while it is not lethal, the misery and loss of mobility with crippling pain results in a lowering of the quality of life for many.

1.4 Synthetic inhibitors of human proteolytic enzymes

There are many classes of compounds reported in the literature, which inhibit human leukocyte elastase and cathepsin G. Some of these are listed in table 1.1.

Table 1.1 Some HLE and HLC-G inhibitors.

Class of compound	References
aminoacylchloromethanes	46
sulfonyl fluorides	47
imidazole-N-carboxamides	48
azapeptides	49,50
cis-unsaturated fatty acids	51
chloromethyl ketones	46
heterocyclic transition-state analogs and acylating agents	52,53
diisopropylfluorophosphate	54,55
tosylphenylchloromethane	54,55
substituted isocoumarins	56
cyclohexylamide	57
gold thiomalate	58

All the compounds in Table 1.1 have one or more reactive functional groups. They are difficult, if not impossible, to target onto receptor sites, as they are likely to react with the many components of the body available between the point of administration and the target at the receptor site. In addition, some of these compounds, such as diisopropylfluorophosphate, are rather poor inhibitors of

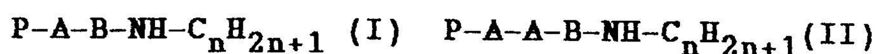
cathepsin G. Tosylphenylalanylchloromethane also inhibits cathepsin G much more slowly than it does chymotrypsin, so that the inhibition often appears incomplete.[54,55] Of course, these compounds are highly toxic and only serve as inhibitors *in vitro* for specific, non-therapeutic purposes. However, oleic acid, a cis-unsaturated carboxylic acid, has been shown to be a reasonably good specific inhibitor of HLE, but not of PPE (porcine pancreatic elastase), trypsin, chymotrypsin and cathepsin G, which has indicated that the principal difference between HLE and the other serine proteases could be due to the different hydrophobic character of a site near the active site.[46]

1.5 Hydrophobic inhibitors of proteolytic enzymes

Previous studies at La Trobe and Chulalongkorn Universities[59] resulted in a series of simple derivatives of short peptides which were potent specific inhibitors of human leukocytic elastase (HLE), collagenase and cathepsin G. These components had the advantage in therapy that they were composed of natural components of the body (i.e. amino acids and fatty acid residues) and were non-toxic either in their original form or when metabolized. The general requirement was a bulky protective group, one to three amino acids, which provided the required specificity, and a hydrocarbon chain of *ca.* 10-14 carbon length. Since the enzymes we sought to

019433

inhibit were members of the serine protease family, we also included trypsin and chymotrypsin in our study. The main reason for this was our desire to design and synthesise specific inhibitors, since these latter two enzymes are of importance to the normal digestive process and hence their inhibition would result in unwanted effects. The general formulae are I and II.



$$\text{P-A-B-NH-}C_nH_{2n+1}\ (I) \quad \text{P-A-A-B-NH-}C_nH_{2n+1}\ (II)$$

where P is benzyloxycarbonyl(Z-) or benzoyl(BZ-) or tertiary butyloxycarbonyl(BOC-)

A is valine

B is tyrosine or phenylalanine

n is 10 or 12

The individual compounds synthesized and tested are shown in Table 1.2.

Table 1.2 Synthesized and tested compounds.

No.	Name	Short formula[a]
1	Benzoyl-L-valine-L-tyrosylamido-decane	BZ-V-T-NH-C_{10}
2	Benzoyl-L-valine-L-tyrosylamido-dodecane	BZ-V-T-NH-C_{12}
3	Benzoyl-L-valine-L-phenylalanylamido-decane	BZ-V-P-NH-C_{10}
4	Benzoyl-L-valine-L-phenylalanylamido-dodecane	BZ-V-P-NH-C_{12}
5	Benzyloxycarbonyl-L-valine-L-tyrosylamido-decane	Z-V-T-NH-C_{10}
6	Benzyloxycarbonyl-L-valine-L-tyrosylamido-dodecane	Z-V-T-NH-C_{12}
7	Benzyloxycarbonyl-L-valine-L-phenylalanylamido-decane	Z-V-P-NH-C_{10}
8	Benzyloxycarbonyl-L-valine-L-phenylalanylamido-dodecane	Z-V-P-NH-C_{12}
9	Benzyloxycarbonyl-L-valine-L-valine-L-tyrosylamido-decane	Z-V-V-T-NH-C_{10}
10	Benzyloxycarbonyl-L-valine-L-valine-L-tyrosylamido-dodecane	Z-V-V-T-NH-C_{12}
11	Benzyloxycarbonyl-L-valine-L-valine-L-phenylalanylamido-decane	Z-V-V-P-NH-C_{10}
12	Tertiary butyloxycarbonyl-L-valine-L-tyrosylamido-decane	BOC-V-T-NH-C_{10}
13	Tertiary butyloxycarbonyl-L-valine-L-tyrosylamido-dodecane	BOC-V-T-NH-C_{12}

(a) The abbreviations in the short formulae are represented as follows:

BZ = benzoyl BOC = tertiarybutyloxycarbonyl Z = benzyloxycarbonyl

T = L-tyrosine P = L-phenylalanine V = L-valine

C_{10} = $-(CH_2)_9CH_3$ C_{12} = $-(CH_2)_{11}CH_3$

1.6 Objectives of this study

In summary, the aims of this investigation are:

1. To synthesize and test the compounds with verifying the structure of synthetic compounds.
2. To test percentage inhibition against the protease (HLE and/or Cathepsin G) include other proteolytic enzymes in the body (trypsin and chymotrypsin) in order to ascertain their mode of actions.
3. To synthesize and test compounds with various hydrocarbon chains, protecting groups and being di- or tri-peptide in order to study the effect of them.
4. To determine the type of inhibition of these compounds with respect to specific enzymes.