The Synthesis and Conformational Analysis of Biological Molecules

Kenneth G. Boyd

A thesis submitted for the degree of Doctor of Philosophy.

University of Edinburgh July 1992

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

University of Edinburgh

July 1992

To my parents and my wife.

Acknowledgements

I would like to thank Dr. R.L. Baxter for his guidance during the course of this work. I would also like to thank Dr. A. Hayman for helpful discussion and advice.

Sincere thanks are due to Mr. K. Shaw for technical assistance with peptide synthesis, the academic and technical staff responsible for the aquisition of nmr spectra and all other members of the departments technical staff for their support.

I wish to thank the Science and Engineering Research Council for the provision of the research grant.

Abstract

This thesis discusses the synthesis and structural analysis of molecules pertinent to current problems in the areas of biosynthesis, structure activity relationships and cell biology.

Chapter one investigates the structure of an N-terminal blocked derivative of methionine^{5}enkephalin, in DMSO solution, using single and two dimensional nmr techniques. Analysis of the data using standard methods indicates the formation of a type I' beta turn centered around the two glycine residues.

Chapter two deals with the development and optimisation of a synthesis of D,L tyrosine which is suitable for the incorporation of isotopic labels to give enriched species such as [2-^{13}C, ^{15}N]tyrosine which is a potentially useful biosynthetic probe. Possible uses of the probe are discussed and three routes are described. The most successful route is based on a modification of the Strecker reaction.

Chapter three describes the synthesis, purification and conformational analysis of a fifteen residue peptide corresponding to the signal sequence of the hemagglutinin protein of the influenza virus A/WSN/33, the portion of the polypeptide responsible for its targeting to the cell surface. Circular dichroism studies of this N-terminal peptide in various solvents indicate that the peptide forms a predominantly β-sheet structure in aqueous solution and in alcoholic solutions of low hydrophobicity, while preferentially adopting a helical structure in more lipophilic solvent systems. The relevence of this to targeting to the cell surface is discussed.

Abbreviations

Ac	acetyl
AIR	aminoriboimidazoleribonucleotide
b	broad
Boc	t-butoxycarbonyl
Bu	butyl
CD	circular dichroism
cDNA	complementary DNA
COSY	correlated spectroscopy
d	doublet
DIC	N,N'-diisopropylcarbodiimide
DMAP	N,N-dimethylaminopyridine
DMF	dimethylformamide
dmso	dimethyl sulphoxide
Dpp	diphenylphosphinyl
E. coli	*Escherichia coli*
EI	electron impact
ER	endoplasmic reticulum
EtOAc	ethyl acetate
FAB	fast atom bombardment
Fmoc	9-fluorenylmethoxycarbonyl

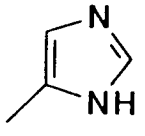
GF	Greenfield and Fasman
GPI	guinea-pig ileum
HA	hemagglutinin
HFIP	1,1,1,3,3,3-hexafluoroisopropanol
HOBt	1-hydroxybenzotriazole
HPLC	high performance liquid chromatography
LAH	lithium aluminium hydride
m	multiplet
Me	methyl
mRNA	messenger RNA
MS	mass spectrometry
NA	neuraminidase
nmr	nuclear magnetic resonance
nOe	nuclear Overhauser effect
NOESY	nOe spectroscopy
POMC	pro-opiomelanocortin
q	quartet
RNA	ribonucleic acid
ROESY	rotating frame NOESY

s	singlet
SDS	sodium dodecylsulphate
SPPS	solid phase peptide synthesis
SRP	signal recognition particle
SPRP	signal recognition particle receptor
t	triplet
t-Bu	tertiary butyl
TFA	trifluoroacetic acid
TFE	trifluoroethanol
TGN	trans-reticular network
TOCSY	total correlation spectroscopy
TRNOE	transfered NOE
UV	ultra violet
vRNA	viral RNA

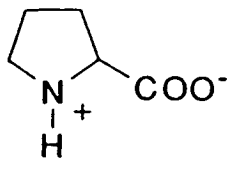
Genetically Coded Amino Acids

Amino Acid	Side Chain	3-Letter Code
Alanine	CH_3	Ala
Arginine	$(CH_2)_3NHC(NH)NH_2$	Arg
Asparagine	CH_2CONH_2	Asn
Aspartic Acid	CH_2COOH	Asp
Cysteine	CH_2SH	Cys
Glutamic Acid	$(CH_2)_2COOH$	Glu
Glutamine	$(CH_2)_2CONH_2$	Gln
Glycine	H	Gly
Histidine	$CH_2(im)$[1]	His
Isoleucine	$CH(CH_3)C_2H_5$	Ile
Leucine	$CH_2CH(CH_3)_2$	Leu
Lysine	$(CH_2)_4NH_2$	Lys
Methionine	$(CH_2)_2SCH_3$	Met
Phenylalanine	$CH_2C_6H_5$	Phe
Proline[2]		Pro
Serine	CH_2OH	Ser
Threonine	$CH(OH)CH_3$	Thr
Tryptophan	$CH_2(ind)$[3]	Trp
Tyrosine	$CH_2C_6H_4OH$	Tyr
Valine	$CH(CH_3)_2$	Val

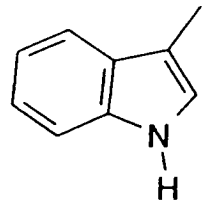
Notes:



1. The imidazolyl group of histidine



2. Proline



3. The Indole Group of tryptophan

Contents

CHAPTER 1

1. Enkephalins

1.1 Introduction

The opiates have long held a particular fascination for scientists and laymen alike. This reflects the unique therapeutic value of these agents, the mystery surrounding their mode of action and the devastating effects wrought by opiate abuse. The narcotic analgesics produce a wide spectrum of pharmacological effects, which may vary from species to species, and present a bewildering picture to the observer. For many years the uncertainties surrounding the mechanism of opiate action attracted a great deal of interest but it was not until the early 1970's that the opiate receptor became a defined pharmacological entity[1].

An added impetus to narcotic research was given by two separate but related developments in *in vitro* analysis. Firstly, the development of the guinea-pig ileum (GPI) as a model for opiate action[2] and the discovery of the mouse vas deferens as a second model of opiate action[3]. Secondly, it became possible to demonstrate directly the existence of pharmacologically relevant, opiate receptor binding sites in the brain and gut[4,5,6]. It was therefore possible to define opiate action in strict pharmacological terms at both the cellular and molecular level.

In view of the discovery of the morphine receptor it was proposed that morphine (1) might be analogous to other plant alkaloids like nicotine (2) or muscimol (3) which mimic the action of endogenous neurochemicals such as acetylcholine (4) and γ-aminobutyric acid (5)[7,8]. This triggered off the search for an endogenous ligand for the opiate receptor. Evidence for such a compound was obtained when an unknown factor isolated from pig brain showed pharmacological properties similar

H.Tyr.Gly.Gly.Phe.Met.Thr.Ser.Glu.Lys.Ser.Gln.Thr.Pro.Leu.Val.Thr.OH

(8)

H.Tyr.Gly.Gly.Phe.Met.Thr.Ser.Glu.Lys.Ser.Gln.
Thr.Pro.Leu.Val.Thr.Leu.Phe.Lys.Asn.Ala.Ile.
Ile.Lys.Asn.Ala.His.Lys.Lys.Gly.Gln.OH

(9)

H.Tyr.Gly.Gly.Phe.Met.Thr.Ser.Glu.Lys.
Ser.Gln.Thr.Pro.Leu.Val.Thr.Leu.OH

(10)

H.Tyr.Gly.Gly.Phe.Leu.Arg.Arg.Ile.Arg.
Pro.Lys.Leu.Lys.Trp.Asp.Asn.Gln.OH

(11)

to morphine[9,10]. The unknown factor was soon characterised as a mixture of low molecular weight peptides[11] and mass spectra of the acylated and methylated derivatives were initially used to sequence the peptides showing them to be Leu^5-enkephalin (6) and Met^5-enkephalin (7)[12]. Independent confirmation of these sequences was later obtained by the isolation of (6) and (7) from calf brain[13].

H.Tyr.Gly.Gly.Phe.Leu.OH (6)

H.Tyr.Gly.Gly.Phe.Met.OH (7)

With further study it soon became apparent that the concept of one opiate receptor and one or two ligands was an oversimplification. This was made evident with the discovery of the 91 amino acid polypeptide β-lipotropin[14,15], which in itself was inactive, and a number of peptides corresponding to portions of its C-terminal region namely α,β and γ endorphin, (8), (9) and (10), which did possesses opiate like properties. This was followed by the identification of dynorphin (11)[16] as the active factor of a pituitary extract. While (11) possess opiate receptor activity its properties are significantly different from the enkephalins.

1.2 Biogenesis

The explosion of interest in the endogenous opioid peptides in the mid to late seventies resulted in the extensive characterisation of the opioid peptides, thus providing a valuable model system for the investigation of general features such as biosynthesis.

Peptide hormones are in general derived from the enzymic cleavage of larger, and generally inactive, prohormones which are synthesised on

membrane bound ribosomes. The cleavage of the signal sequence on entry to the endoplasmic reticulum yields the prohormone which is sequentially cleaved, yielding one or more biologically active peptides that are available for secretion. In the case of cellularly secreted hormones, cleavage appears to follow an ordered pattern and initially involves proteolytic attack at paired basic amino acid sequences. Attack at the carboxy terminus of the basic amino acids leaves the active peptide with a basic amino acid on the C-terminus which is subsequently removed by a carboxypeptidase B like enzyme to give the active peptide[17].

It was originally thought that β-lipotropin was a precursor of Met^5-enkephalin as its N-terminal sequence corresponds to the primary structure Met^5-enkephalin[18]. This is however highly unlikely; apart from not explaining the origin of Leu^5-enkephalin, the Thr-Ser sequence which links the Met^5-enkephalin sequence to the rest of β-lipotropin would be a very unusual cleavage site for the production of the pentapeptide. It should also be noted that β-endorphin and the enkephalins are not found in the same organs. The enkephalins have been found in the posterior lobe of the pituitary whereas β-endorphins have been found only in the anterior and intermediate lobes of the pituitary[19], and tissues such as the striatum and adrenal medulla have been found to contain the enkephalins but no β-endorphin[20,21]. However for conclusive evidence that β-lipotropin is not involved in the biosynthetic pathway of enkephalins it would be necessary to isolate a precursor containing the enkephalins' sequences.

The work undertaken in attempting to elucidate the origin of the opioid peptides has been the subject of various review articles[22,23]. It is now clear that there are three separate families of endogenous peptide opiate receptor ligands: enkephalins, endorphins and dynorphins. They

are derived from three different prohormones: proenkephalin pro-opiomelanocortin and prodynorphin, which are coded by mRNA's derived from three separate genes, fig 1.1.

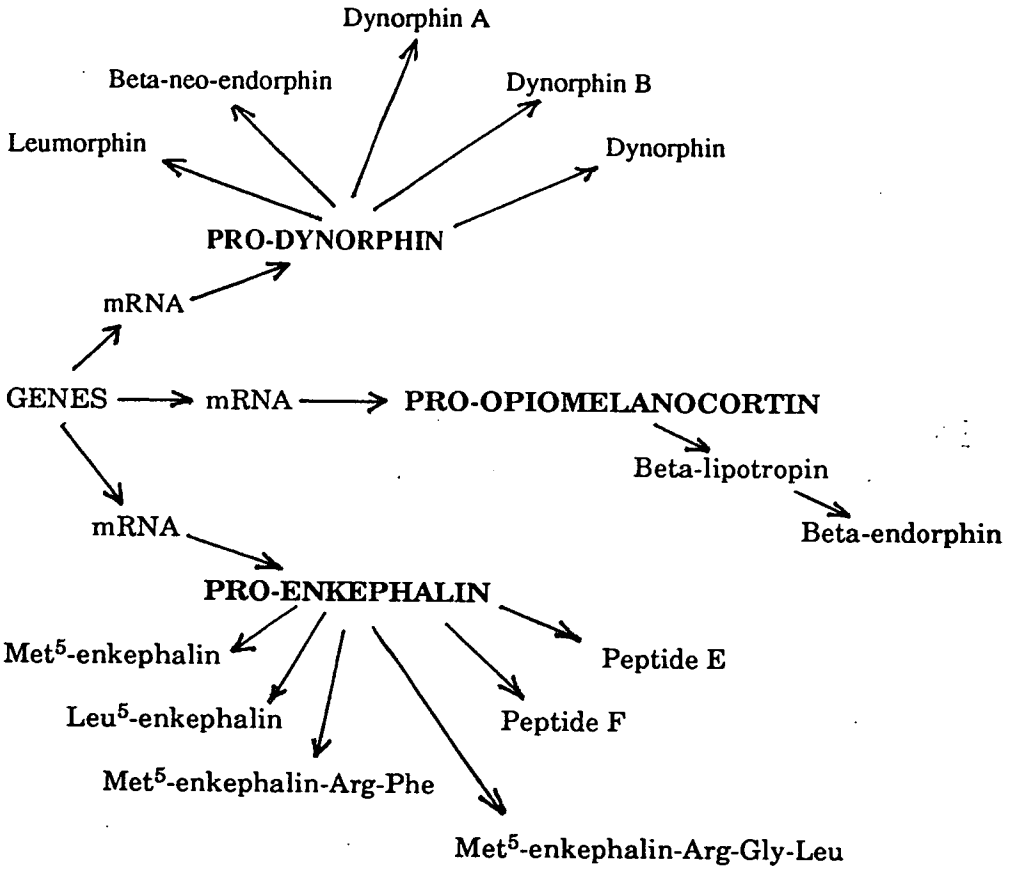


Fig. 1.1 Opioid peptides and their precursors

The precursor of both natural enkephalins known as proenkephalin A, contains the amino acid sequences of both Met^5 and Leu^5-enkephalin in a fixed ratio of six to one respectively, although some of the Met^5-enkephalin sequences incorporate several more amino acids to produce larger, C-terminally extended, opioids. Pro-

opiomelanocortin(POMC) is the common precursor of the opioid β-endorphin and although this contains a Met^5-enkephalin sequence at its amino terminus no Met^5-enkephalin is derived from POMC. Prodynorphin, also known as proenkephalin B, contains Leu^5-enkephalin sequences but no Met^5-enkephalin sequences. The opioid peptides derived from this precursor include the dynorphins, β-neodynorphin and leumorphin, all of which contain the Leu^5-enkephalin sequence at their amino terminus. There is evidence to suggest that Leu^5-enkephalin is also derived from proenkephalin B[24]. This is feasible in light of the fact that in the dynorphins the Leu^5-enkephalin sequence is followed by two arginine residues, which would facilitate enzymatic cleavage, and the distribution of the dynorphins and enkephalins is often contiguous.

The identification of proenkephalin and prodynorphin would appear to have solved the problem of the biosynthesis of brain enkephalins. Moreover the existence of Leu^5-enkephalin in two separate precursors would neatly explain how the ratio of Leu^5 and Met^5-enkephalin can vary in different regions of the brain.

1.3 Opioid Receptors

The concept that there are several types of opioid receptors was originally derived from the observation that in man nalorphine(12) had a dual action, antagonising the analgesic effect of morphine and also acting as an analgesic in its own right. It has been suggested that the analgesic effect of nalorphine is mediated by a receptor which is different from the morphine, μ-receptor, and has been referred to as the κ-receptor.

When the enkephalins were first discovered it was assumed that they adopted a conformation analogous to that of morphine, with a rigid

structure, however it was noted that some alkaloid opiates could not displace [^{3}HTyr^{1},Met^{5}]enkephalin (13) from its receptor efficiently and the enkephalins could not compete for alkaloid binding sites[25]. While it has been proposed that this is due to different modes of binding to the same receptor or conformational changes at the receptor induced on the enkephalins[26], later work has suggested that these binding phenomena are due to the existence of multiple classes of opiate receptor with differing affinities for various opiates[27,28].

From the results of various bioassays and radioreceptor assay studies the alkaloid opiates have been found to be more potent in the GPI test than the enkephalins whereas the reverse is true for the vas deferens assay[29]. These observations indicated that the opioid peptides interact with two types of receptor : the μ-receptor, with which morphine preferentially interacts, appeared to predominate in the guinea-pig ileum, whereas a δ-receptor, which is different from the μ or κ receptor and with which the enkephalins interact preferentially, appeared to predominant in the mouse vas deferens[30].

Work done on the selectivity of the opiate peptides towards the three receptor subtypes have shown the enkephalins to have selectivity for the δ-receptor. Most of the products of prodynorphin interact with κ-receptors while all three families of opioid peptides have affinity for the μ-receptor[23,31].

The three classes of opiate binding sites are now firmly recognised. Although there is some recent evidence for the existence of subdivisions of the κ and μ sites[23], no firm evidence for further subdivisions of receptor types has been shown.

1.4 Structure Activity Studies

In order to design super-agonists a great deal of structure activity work has been done on the enkephalins, with the aim of identifying the essential functionalities required for receptor site binding. Early work showed that the blocking of the C-terminal either by dansylation or amidation resulted in analogs with equal potency[32]. Blocking of the N-terminus generally resulted in analogues with less activity[33,34] but some alkylated and acylated analogs were found to be active[35] as were analogs with additional amino-acid residues attached to the N-terminus[35]. Replacement of L-Tyr^1 with D-Tyr or with other amino acids[36,37] results in a dramatic loss of activity as does derivitisation of the phenolic hydroxyl[32,38]. These results suggest that an intact L-tyrosine moiety in position one is essential for activity.

Early work demonstrated that the enkephalins were subject to rapid deactivation when exposed to tissue homogenates and purified enzyme preparations such as carboxypeptidase-A and leucine aminopeptidase. In view of these results it was not surprising that in *in vivo* studies, even after central administration, the agonist activity of the enkephalins were at best transient. Examination of the degradation products by thin layer chromatography showed the deactivation of the enkephalins to take place principally through cleavage of the Tyr^1-Gly^2 amide bond yielding the inactive tetrapeptides (14) and (15)[39].

H.Gly.Gly.Phe.Met.OH (14)

H.Gly.Gly.Phe.Leu.OH (15)

In order to render the amide bond inaccessible to cleavage Gly^2 was replaced with D-alanine to give the analog (16)[39], which was equipotent to Met^5-enkephalin in receptor binding assays but was not degraded. It also showed analgesic activity when administered intracereboventrically. Further work on analogues substituted at the two position showed that replacement of Gly^2 with D-Met or D-Ser produced even more potent analogues[40]. However analogues with L-amino acids other than glycine in position two had diminished activity[41,42].

Position three was initially shown to be less tolerant towards N- or C- substitutions and replacement of Gly^3 with with alanine resulted in complete loss of activity[43]. The presence of an aromatic functionality at position four was found to be essential for activity, though introduction of a nitro functionality in the para position of Phe^4 gave a compound with very high potency in the GPI assay[44].

H.Tyr.DAla.Gly.Phe.Leu.NH_2

(16)

A wide variety of manipulations have been reported at the C-terminus. It was found that an aliphatic residue in position five and an unmodified carboxy group at the C-terminus were required for efficient binding to the δ-receptor[45]. On the other hand, interaction with the μ-receptor is not appreciably effected by structural modifications at the C-terminal and omission of the fifth residue produces analogs with enhanced activity in the GPI assay[44,46].

Thus by the late seventies the minimum structural requirements for interaction with the opiate receptor were perceived as encompassing

the elements of an intact L-Tyr^1 moiety and the aromatic nucleus of the Phe^4 separated by an appropriate dipeptide spacer. However several interesting modifications of the enkephalin backbone performed in order to enhance the stability of the peptides towards enzymatic degradation showed that replacement of the Tyr-Gly amide bond with a CH_2-NH linkage, producing an analogue (17) did not result in a drop in activity[47]. A similar replacement of the amide bond using an olefinic linkage produced an analogue (18) which had three times the potency of Leu^5-enkephalin methyl ester in the GPI assay[48]. However replacement of the Gly-Gly amide bond with a CH_2-NH bond results in complete loss of activity[47]. This showed that a peptide bond between Tyr^1 and Gly^2 was not a necessary constituent for opioid activity.

H.Tyr.Pro.Phe.Gly.Pro.Ile.OH

(19)

H.Tyr.DAla.Phe.Gly.Tyr.Pro.Ser.NH_2

(20)

H.Tyr.D-Met(o).NHCH$_2$CH$_2$CH$_2$Ph

(21)

H.Tyr.Pro.Phe.Pro.NH_2

(22)

The early view that a dipeptide spacer was essential for opiate activity was challenged with the discovery that β-casomorphin (19)[49] and dermorphin (20)[50] both had opiate activity while having only a single amino acid residue between the two essential tyrosine and phenylalanine residues. With this in mind several analogs were synthesised including (21)[51] and (22)[52] which were at least as potent as Met^5-enkephalin demonstrating that of the Phe^4 residue only the aromatic ring is necessary for interaction with the opiate receptor and an amide bond two carbon atoms distant from the phenyl ring is unnecessary, as is an amino acid in position three.

Recent attempts to synthesise biologically active enkephalin analogs with minimal structure produced analogs (23)[53] and (24)[54]. (23) efficiently inhibits morphine activity and the Leu^5-enkephalinamide analogue (24) is reported to be equipotent with the parent amide in the GPI assay. It is now generally accepted that the minimal structure neccessary for interaction with the opiate receptor is a basic nitrogen, a hydroxylated phenyl ring and a second phenyl ring fixed at appropriate points in space.fig. 1.2.

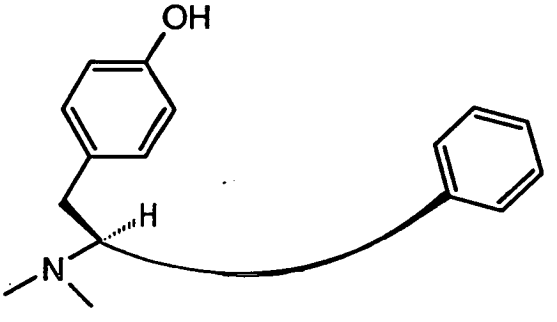


Fig. 1.2. Minimal structural requirements for opiate activity

1.5 Conformation

The most attractive rationalisation for the pharmacological activity of the enkephalins and their analogues is that their receptor binding is similar to that of the morphine alkaloids. If this is the case it would be expected that the conformation adopted by the peptides at the receptor site would be topologically similar to that of the rigid morphine skeleton, and models of the enkephalins have been constructed on the basis of structural comparisons with the alkaloid opiates[55,56]. As a result a number of studies have been made of the conformational analysis of the endogenous opioid peptides in the search for information about their receptor-bound conformations.

The problem is a difficult one because of the inherent flexibility of small peptide molecules many conformational states are possible, especially in solution. The situation is further complicated through the existance of various opioid receptor sub-classes that have different structural requirements. The enkephalins having significant affinity for both the μ and δ receptors and may bind to the two sub-classes by assuming two distinct conformations. Three main approaches have been used in the quest to determine the structure of the enkephalins, these being, crystal structure determination, nmr spectroscopy and theoretical energy calculations, each of which has been the subject of a number of reviews.

1.5.1 Theoretical Energy Calculations

Mathematical analysis of the enkephalins have suggested a number of low energy conformations corresponding to different secondary structures. These include a type II' β bend centered around the Gly^3 and

Phe^4 residues, first proposed by Isogai *et al* in an early investigation[57], which is stabilised by a hydrogen bond between the tyrosine hydroxyl and the carbonyl group of the Gly^3 residue. Similar structures have been proposed by other workers who have found a number of minimal energy conformations characterised by a turn around the Gly^3 and Phe^4 residues, as well as low energy structures corresponding to turns around the Gly^2 and Gly^3 residues[58,59].

Other studies using the Monte-Carlo method to obtain statistical samples of conformations of Leu^5-enkephalin suggest a predominance of a local folded form in the zwitterionic state and mainly extended conformations in the cationic form[60]. Thus from theoretical calculations the enkephalins would appear to be able to access a number of low energy states and may well be present as a mixture of a number of different conformations.

1.5.2 Crystal Structure Determination

Crystal structure determination is generally accepted as being the definitive method of structure elucidation for rigid molecules. However in the case of conformationally flexible molecules the validity of crystallography as a definitive tool may be questioned for two reasons : firstly the conformation exhibited in the solid state may be dominated by intermolecular forces which are not present in dilute solution in biological fluids, and secondly interaction with a receptor site may stabilise conformations different from those imposed by crystal packing restraints.

The first X-ray structure of Leu^5-enkephalin, published by Smith and Griffin[60] demonstrated the presence of a type I β turn centered on the two glycine residues and stabilised by two intramolecular hydrogen bonds,

one between the amino group of Tyr^1 and the carbonyl oxygen of Phe^4 and the second between the amide proton of Phe^4 and and the Tyr^1 carbonyl. A similar structure containing a type I' β bend was also found for the Leu^5-enkephalin analog [(4'Br)Phe^4,Leu^5]enkephalin[62] in the solid state. However, in marked contrast to the folded conformations observed in the above studies, a separate study showed crystals of Leu^5-enkephalin to contain molecules with extended backbones in the form of an antiparallel β-sheet[63].

Only one crystal structure has been published for a Met^5-enkephalin derivative that being the brominated analog — [(4'Br)Phe^4,Met^5]enkephalin[64]. This was found to adopt a dimeric structure in the form of an antiparallel β-pleated sheet.

The occurence of two entirely different types of secondary structure in the solid state is indicative of the flexibility of the enkephalin molecules and is consistent with the the results of the theoretical studies which postulated that the low energy conformers of the enkephalins would encompass both extended and folded forms.

1.5.3 Nuclear Magnetic Resonance Studies in Solution

In the quest to determine the solution structure of the enkephalins proton nuclear magnetic resonance(nmr) has been the most extensively used technique, and the results of numerous studies have appeared in the literature[65,66]. In common with the crystal structure determinations and with theoretical energy calculations the structures proposed fall into two distinct groups : those characterised by a β-bend and those characterised by an anti-parallel β-sheet.

The most frequently postulated structures are those characterised by β-turns and these themselves fall into three distict classes. The first, a type I β turn centered on the Gly^2 and Gly^3 residues has been advanced for both the natural enkephalins end their truncation peptide H.Tyr.Gly.Gly.Phe.OH in both aqueous and dmso solutions[67,68,69]. A second class of structures, again characterised by a type I β turn in this case around the Gly^2 and Gly^3 residues, has been suggested for both Met^5 and Leu^5 enkephalin in dmso solution[70,71]. A third class of structure incorporating a different type of β-turn has recently been reported by two groups. Milon *et al* recently carried out an elegant study on Leu^5-enkephalin and its active analogs H.Tyr.(DAla).Gly.Phe.Leu.OH and its corresponding N-terminal amide in a phospholipid environment[72]. Using transferred nuclear Overhauser effects(TRNOE) to observe dipolar couplings evidence was produced which was indicative of a monomeric structure involving a typeII' β bend centered on the Gly^3 and Phe^4 residues accompanied by a γ turn around the amino acid in the two position. This was supported by the evidence for a type II β bend postulated for Leu^5 enkephalin in dmso solution on the basis of ROE measurements[73].

A structure characterised by an antiparallel β sheet has been proposed to exist for both natural enkephalins in high concentrations in dmso solutions but not in aqueous solution[71,74].

Nuclei other than hydrogen have been used to a lesser extent in investigations of enkephalin secondary structure. Experiments carried out using ^{13}C nmr favour a type I β turn around the Gly^3 and Phe^4 residues on the basis of torsion angles calculated from coupling constant data[75,76]. A study using the ^{15}N enriched tetrapeptide H.Tyr.Gly.Gly.Phe.OH in dmso

concluded that the peptide was only present in a folded conformation when in the zwitterionic form but adopted extended conformations at high and low pH[77]. ^{17}O nmr has been used in an attempt to elucidate the differences in the hydrogen bonding at Gly^2 and Gly^3, however the study was limited to a fully protected Leu^5-enkephalin and failed to produce any evidence for hydrogen bonding involving the carbonyl oxygen of these two residues[78].

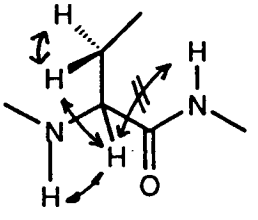
1.6 Discussion

In view of the conformational flexibility inherent in the structure of the enkephalins the crystal structures determined for these molecules, dominated as they are by intermolecular forces, may be of little relevance to the conformation adopted on binding to membrane bound receptors. Earlier nmr studies on Met^5-enkephalin have been on the whole based on single dimensional spectra which were in some instances obtained in solutions of high concentration, a condition which it is now acknowledged can give rise to dimerisation. Because of the conflicting evidence in the literature we decided to reinvestigate the solution structure of Met^5-enkephalin in dilute solution using 2D nmr techniques.

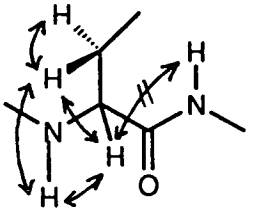
In this work we chose to study Dpp.Met^5-enkephalin an N-terminal blocked derivative of Met^5-enkephalin. This derivative was chosen for two reasons. Firstly functionalisation of the N terminal amine, with the diphenylphosphinyl group[79], inhibits the formation of dimers by denying amine carboxylate ionic interactions. Secondly, by employing a bulky hydrophobic group we hoped to stabilise a single conformation. As discussed previously, derivatisation of the amino terminus does not necessarily lead to loss of activity.

1.6.1 2D NMR

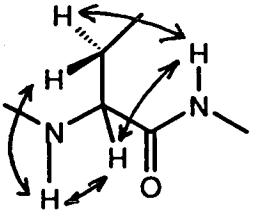
The large number of protons present even in relatively small peptides leads to very complicated nmr spectra. In order to assign the data obtained it is necessary to use two dimensional techniques which both resolve the information and provide additional aids to its assignment. Examples of the typical information which can be obtained using 2D nmr spectroscopy are shown in fig. 1.3.



COSY



TOCSY



ROESY/NOESY

Fig 1.3 Information available from 2D nmr

The method of choice for obtaining evidence of secondary structure in peptides or proteins in solution is the detection of the nuclear Overhauser effect (nOe). Such measurements allows one to pick out short

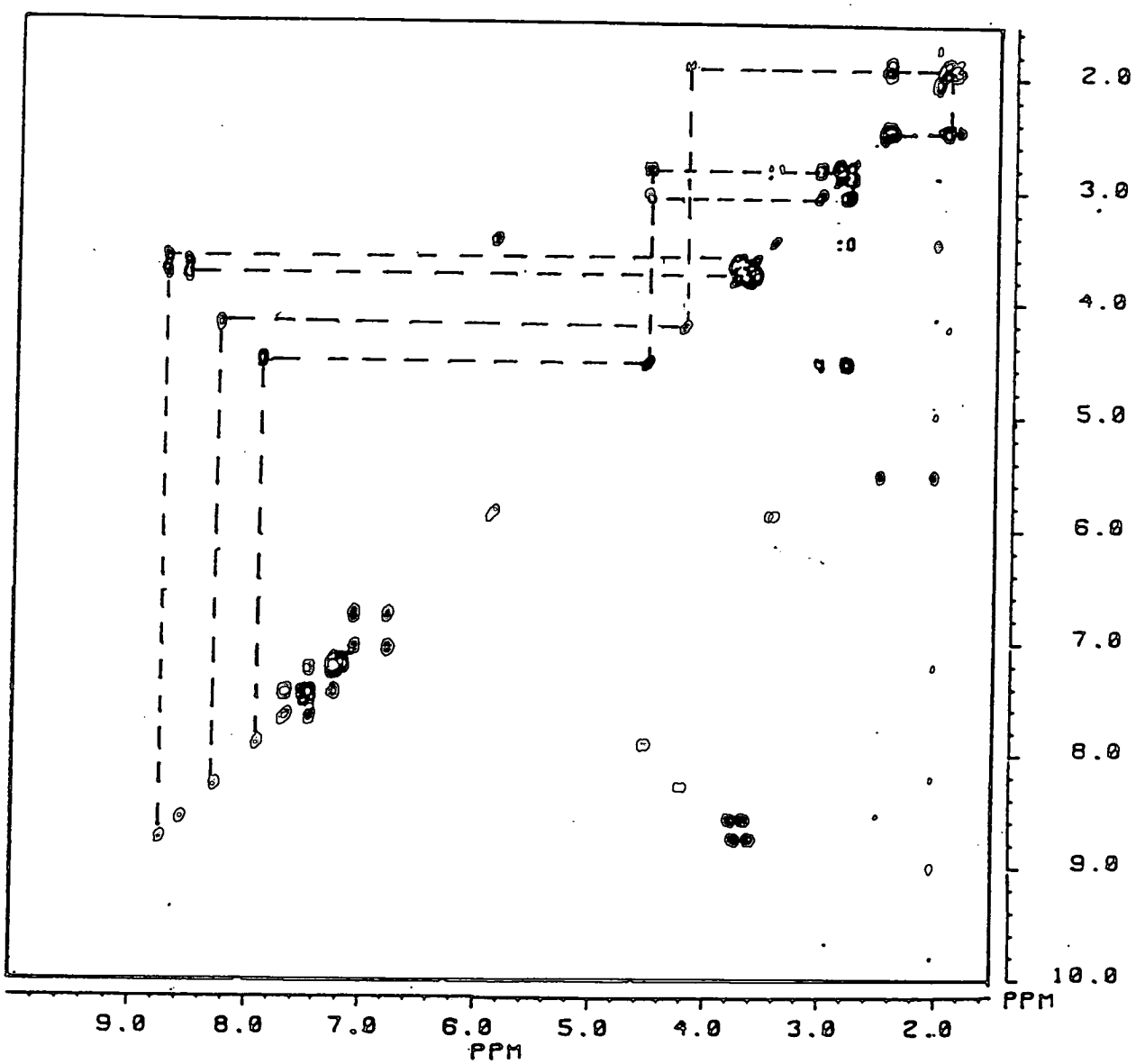


Fig. 1.4 2D COSY spectra of Dpp.Met^{5}enkephakin in dmso d_6 at 273K.

(less then 3.5-4.0Å) interproton distances. Unfortunately for small peptides in the commonly used frequency range, 300-600MHz, the nOe values are frequently close to zero[80]. However a method for measuring nOe's in the rotating frame, the rotating frame nuclear Overhauser effect(ROE)[80,81] gives ROE values which are always positive even for small peptides and is thus invaluable for the qualitative evaluation of through space proton-proton interactions in molecules such as the enkephalins. We therefore chose to use the two dimensional ROESY[81] procedure for the study of Dpp.Met^5-enkephalin.

1.6.2 Assignment of the Proton Spectra

Assignment of the proton chemical shifts was performed using 2D COSY and ROESY experiments using the methods described by Wuthrich[82]. The assignment protocol for the Gly^3, Phe^4 and Met^5 residues shown in figure 1.4 illustrate the identification of the spin systems corresponding to the individual amino acids.

The sequential assignment of the Gly^2 and Gly^3 resonances follow from consideration of the cross peaks observed on the ROESY spectrum. The two glycine residues show characteristic NH triplets at 8.56 and 8.73 ppm. The higher field amide NH resonance shows a cross peak with the CαH of Tyr^1 at 3.14 ppm and no such connectivity is observed for the Gly NH at 8.73 ppm. This unambiguously assigns the NH at 8.56 ppm to Gly^2, and consequently the one at 8.73 ppm to Gly^3, thus discriminating between the two spin systems.

The chemical shifts and coupling constants are tabulated (tables 1.1 and 1.2).

Residue	Hα	Hβ	NH	Other
Tyr^1	3.41	2.90,2.88	5.87	7.06,6.78
Gly^2	3.71,3.80	---	8.56	
Gly^3	3.65,3.75	---	8.73	
Phe^4	4.54	3.03,2.81	7.87	
Met^5	4.21	1.98	8.26	2.45,2.03

Table 1.1 Proton chemical shifts(ppm) for Dpp.Met^5-enkaphalin in dmso-d_6 at 28°C

Residue	$^2J_{\alpha_1\alpha_2}$	$^3J_{\alpha\beta}$	$^3J_{\beta_1\beta_2}$	$^3J_{NHCH\alpha}$
Tyr^1	---	a	a	8.2
Gly^2	-17.2	---	---	5.8,5.8
Gly^3	-17.3	---	---	5.8,5.8
Phe^4	---	4.7,9.0	13.7	7.7
Met^5	---	3.9,9.5	a	8.5

Table 1.2 Coupling constants(Hz) for Dpp.Met^5-enkephalin

a=not obtainable from spectrum

1.6.3 Identificaton of a β Turn

With regard to the application of nmr data to the identification of secondary structural features in peptides there are two principle criteria for the selection of "useful" proton-proton distances. In the first case they

	Tyr^1	Gly^2	Gly^3	Phe^4	Met^5
$\delta_{N,N}(i,i+2)$	✗	✗	✗	–	–
$\delta_{N,N}(i,i+1)$	✓	✓	✓	✗	–
$\delta_{\alpha,N}(i,i+1)$	✓	✓	✓	✓	✓
$\delta_{\alpha,N}(i,i)$	✓	✓	✓	✓	✓
$\delta_{\beta,N}(i,i)$	✓	–	–	✓	✓

Table 1.3 Important Dipolar couplings in the ROESY spectrum of Dpp.Met^5 enkephalin.

IRRADIATED SIGNAL	RESPONSE
Tyr^1NH	Dpp aromatics
Tyr^1CαH	Gly^2NH
Gly^2NH	Gly^3NH Tyr^1NH
Gly^3NH	Gly^2NH Phe^4NH
Phe^4NH	Phe^4CαH Gly^2NH Gly^3NH Phe aromatics
Phe^4CαH	Phe^4NH Met^5NH* Phe^4 aromatics
Phe^4β	Phe^4 aromatics
Met^5NH	PheCαH* MetCαH

Table 1.4 n.O.e results. All responses were less than 1% except for * which was 3%.

should be significantly short so as to give rise to observable interactions in ROESY or NOESY experiments, ie shorter than 4Å. The second and most important criterion is that the proton-proton interactions should be characteristic for different secondary structures.

A summary of the important dipolar couplings found in the ROESY experiment on DppMet^5-enkephalin is shown in table 1.3. The integrity of the dipolar couplings was confirmed by a series of nOe difference experiments which are summarised in table 1.4. In this case the ROE's with greatest diagnostic value are those between the NH's of Gly^2 and Gly^3 and between the NH's of Gly^3 and Phe^4 fig. 1.5.

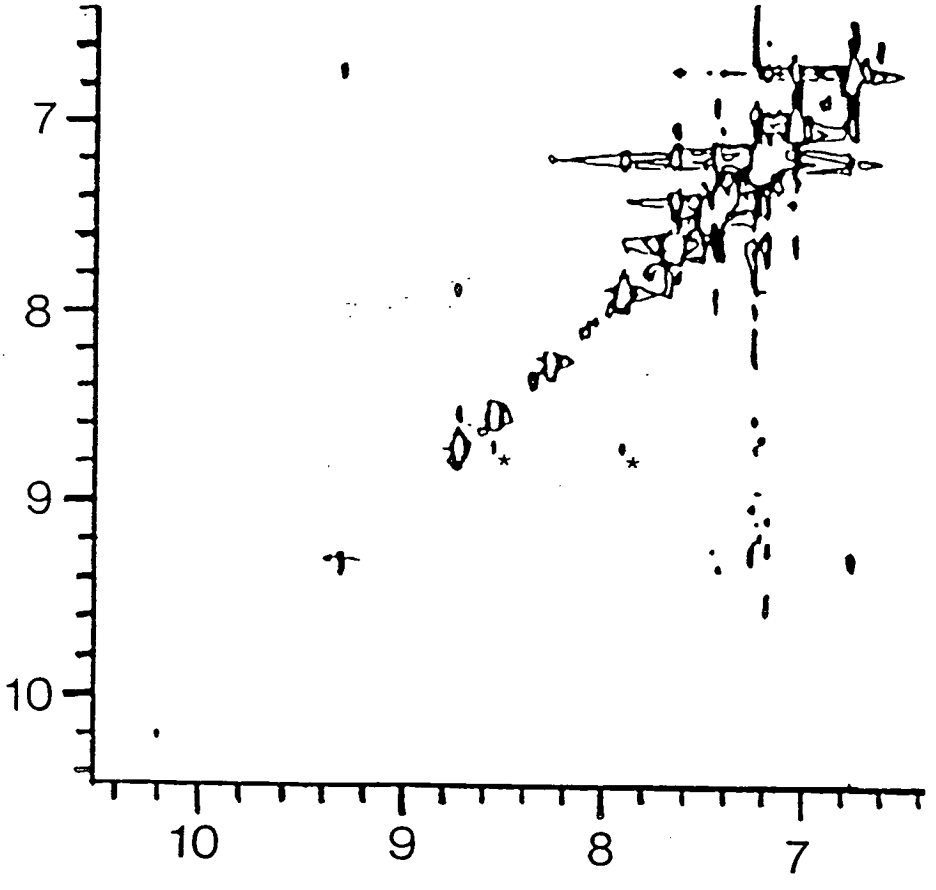


Fig. 1.5 Fingerprint region of the ROESY spectrum of Dpp.Met^5-enkephalin indicating the amide-amide cross peaks

These are indicative of a local folded β turn structure and can be used to discriminate between the various types of β turn[82]. For turns of type I the shortest proton-proton distances are those between NH^i and NH^{i+1} and between NH^{i+1} and NH^{i+2} which are of the order of <2.5Å. For turns of type II the shortest distances are those between $C\alpha H^i$ and NH^{i+1} and NH^{i+1} and NH^{i+2}, fig. 1.6.

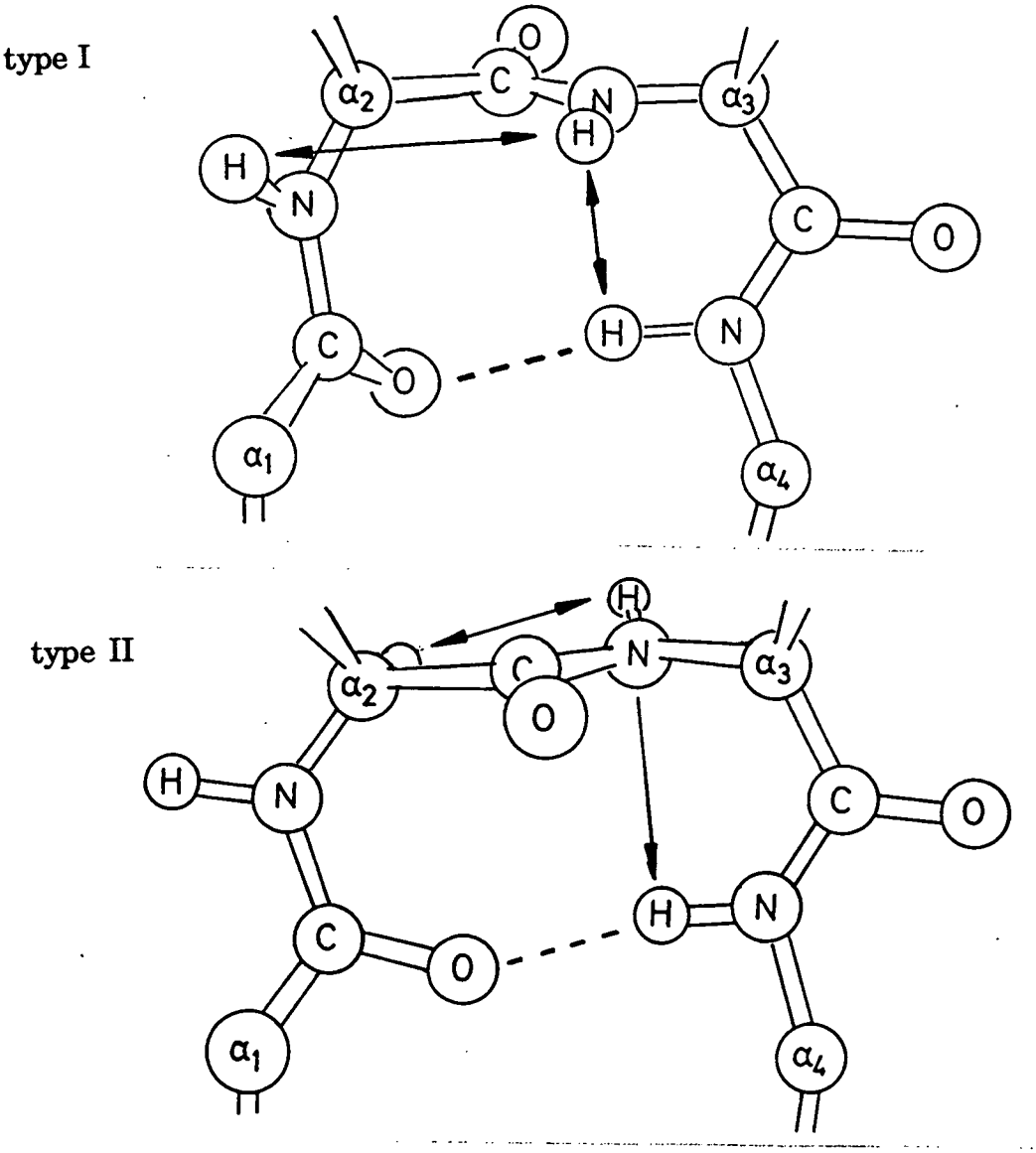


Fig. 1.6 Standard type I and type II turns showing the ^1H-^1H contacts of interest for the identification of β-turns

The distinction between the pair in this case relies on the presence (type I) or the absence (type II) of a dipolar coupling between NH^i and NH^{i+1}, since in a type II structure the distance between these two protons increases to 4.5Å which is insufficient to give rise to a strong ROE interaction. Thus the presence of a ROE interaction between the NH's of Gly^2 and Gly^3 differentiates between the two types, indicating that in our solvent system Dpp.Met^5-enkephalin contains a conformer with a type I or type I' β turn centered on the Gly^2 and Gly^3 residues.

1.6.4 Coupling Constants

Analysis of the coupling constants between interacting nuclei yields information of the spatial orientation of the participating atoms[83]. Owing to their ease of observation proton-proton coupling constants have a prominant position in the conformational analysis of peptides. The most useful are the vicinal NHCαH and H-CαCβ-H coupling constants which give the ϕ and χ torsion angles and the geminal H-Cα-H coupling constants which give information on the ϕ and ψ torsion angles of glycine containing peptides. The important torsion angles present in peptide fragments are shown in fig. 1.7.

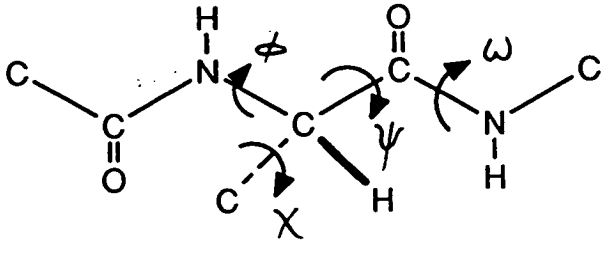


Fig. 1.7 A peptide fragment showing the torsion angles ϕ, ψ, ω, and χ

The vicinal $HNC\alpha H_2$ couplings of the glycine residue give rise to an ABX spectrum. In this case the triplet splitting of the NH signal does not enable direct evaluation of the $^3J_{NHC\alpha H}$ constant[84]. However the spacing between the outer components of the amide signal is equal to the sum of the J_{AX} and J_{BX} coupling constants and this has been shown to be related to the torsion angle ϕ by the following relationship:-

$$\Sigma(^3J_{HNC\alpha H}) = -9.4\cos^2\phi - 1.1\cos\phi + 14.9$$

For the coupling constants observed this gives a number of possible values for ϕ, -65, -125, 65 and 125°. The ϕ values required for a type I and type II β turn are almost identical, ±60° for the first residue of the turn and ±90 or ±80° for the second residue of a type I and type II turn respectively[82]. While an angle of ±60° is easily accomodated by the experimentally observed couplings a value in the region of ±90° would seem to be outwith the experimentally derived torsion angles. In any event the vicinal $NHC\alpha H_2$ coupling constant would appear to be of limited value in differentiating between the two families of β-turn which differ only by 10° in the ϕ torsion angle of the second residue of the turn.

The ψ torsion angle around the carbonyl-CαH bond provides a more definitive indication of the type of turn present. For both types of β turn the ψ angle for the second residue is 0°, however for the first residue the ψ torsion angle is markedly different for each type of turn ±30° for a type I turn and ±120° for a type II turn[82].

The geminal H-Cα-H coupling constant observed in nmr spectra of glycyl containing peptides is dependant on the ψ angle but is independant of the torsion angle ϕ.

In the spectra of Dpp.Met^5-enkephalin the geminal couplings of the methylene protons of the two glycine residues are 17.2 and 17.3 Hz for

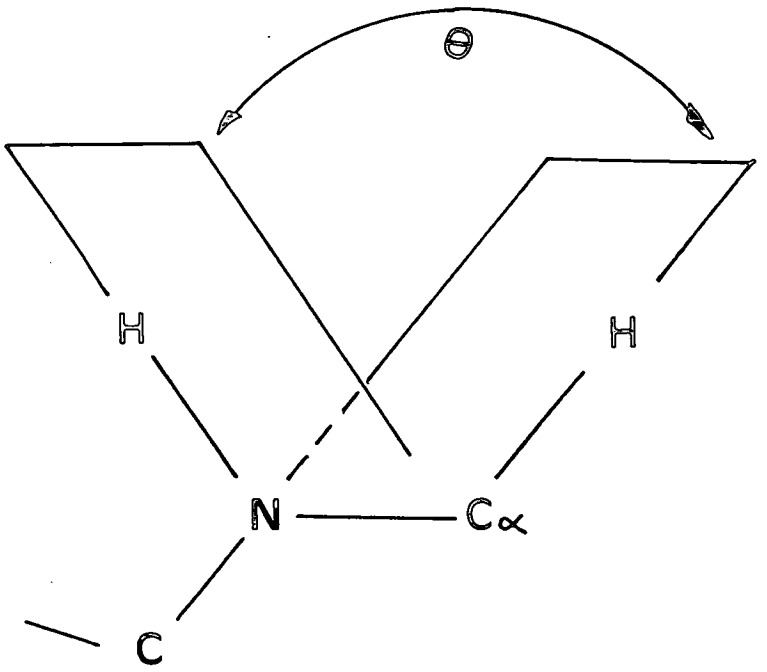


Fig. 1.8 The dihedral angle θ

Residue	Φ	Ψ	θ
Tyr^1	---	---	0,153
Gly^2	±60,±125	0,180	---
Gly^3	±60,±125	0,180	---
Phe^4	---	---	0,150
Met^5	---	---	0,155

Table 1.5 Torsion angles.

Gly^2 and Gly^3 respectively and are apparently negative in common with those of other tetrahedral carbon atoms[85].

The most widely used model for the dependance of the glycine geminal coupling constant on the ψ torsion angle is that described by Cookson[86]. According to this model geminal couplings of -17.2 and -17.3 Hz correspond to angles of 0° or 180°(±40°) for the principle torsion angle ψ.

To accomodate a type II turn a ψ torsion angle of ±120° is required for the second glycine residue. This would correspond to a geminal coupling in the region of -13 to -15 Hz which is obviously excluded by the experimentally observed values. ψ torsion angles of ±30° and 0° required for a type I or I' β turn are however easily accomodated by coupling constants in the region of 17Hz. This rules out the presence of a type II or II' β turn and provides further evidence supporting the presence of a type I or I' β turn in this molecule. However it should be noted that the torsion angles determined on the basis of nmr data do not allow discrimination between a type I and type I' β turn.

The torsion angle θ, fig. 1.8, corresponding to the angles between the amide NH and the CαH bonds for the tyrosine, phenylalanine and methionine residues were calculated from their respective CαHNH coupling constants according to the method of Bystrov[84] and are reported in table, 1.5 along with the torsion angles calculated for the glycine residues.

The averaged conformational states of peptide side chains can be determined from the vicinal H-CαCβ-H proton-proton coupling constants using the modified Karplus-type equations described by Feeney[87] and Pachler[88]. These enable evaluation of the conformational states of the

amino acid side chains of peptides in terms of their distribution among the three different rotational states where χ equals -60, 60 and 180° (tg^-, gg and tg^+ respectively.

Applying the criteria described by Pachler to the Phe^4 and Met^5 CαHCβH coupling constants indicates a distribution of 77%tg and 23%gg for the Phe^4 side chain and, 75%tg and 25%gg for the Met^5 side chain. While this method does not discriminate between the two possible trans-gauche conformations, tg^+ and tg^-, Feeney's treatment does allow discrimination between these two forms although it in turn has the disadvantage of solving to two different sets of rotamer populations. However since only one set of solutions of the Feeney equation corresponds to the poulations obtained from Pachler's equations a combination of the two methods allows the complete solution of the rotamer population set.

Application of this method gives the rotamer populations 63%tg^+, 12%tg^- and 25%gg for the phenylalanine side chain and 69%tg^+, 2%tg^- and 29%gg for the methionine side chain. This indicates that both side chains favour the trans-gauche conformation.

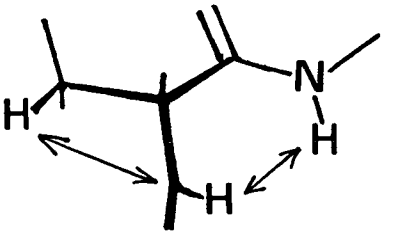
It was however impossible to calculate the rotamer populations of the tyrosine residue by this method due to couplings of the phosphorous atom to the CαH and $C\beta H_2$ signals.

1.6.5 Building a Model of $DppMet^5$-Enkephalin

The β-turn was used as a starting point for the model and was constructed using standard bond angles corresponding to type a I' β turn[82]. The tyrosine spin system shows only two interresidue ROE's one between the tyrosine amide and Gly^2 amide protons and the second the

characteristic $C\alpha H^{i},NH^{i+1}$ dipolar coupling between the tyrosine CαH and the $Gly^{2}NH$.

The H-NCα-H coupling constant of 8.2Hz corresponds to a θ angle of 0 or 180° and on the basis of modelling an angle of 0° seems more probable as an angle of 180° results in the diphenylphosphinyl group interfering with the β-turn backbone. The tyrosine side chain is most easily accomodated in a trans-gauche^{+} conformation which is in accord with the observation of an ROE between the Tyr^{1} NH and only one of the tyrosine β protons. Thus the conformation of the tyrosine residue with respect to the β turn can be defined as shown.

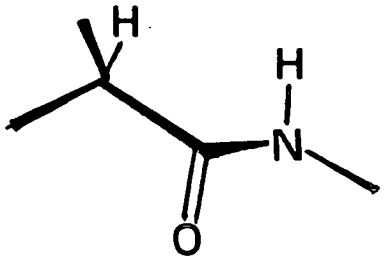


The phenylalanine residue also seems to adopt a major conformation. A NHCαH coupling constant of 7.7 Hz suggests a torsion angle θ of 10° or 150°. From modelling an angle closer to a trans arrangement would appear to be the most acceptable as this is consistent with the observed ROE between one and only one of the phenylalanine methylene protons. This is in agreement with the trans gauche^{+} conformation derived from the CαHCβH coupling constants.

The phenylalanine CαH shows a strong NOE (3.5%) with the methionine amide proton which is indicative of a parallel arrangement of the glycine CαH and the methionine NH bond. This fixes the Met relative to the Phe residue and hence to the β turn.

The methionine side chain would appear to be well removed from the rest of the structure as no inter-residue ROE's are observed for the

methionine side chain protons and the only clues to its conformation are the strong NOE linking the the Phe[4] CαH proton to Met[5] amide proton, discussed above, and the predicted tg^+ conformation of the side chain. A trans arrangement of the NHCαH bond is again favoured from modelling as this reduces the steric interactions of the side chain with the rest of the structure.



Using the above data and observations a model was constructed using the Alchemy II molecular modelling program. Figure 1.9 shows the integral β turn structure. The full structure, not including the hydrogen atoms is shown in figure 1.10. Fragments of the structure including the hydrogen atoms are shown in figures 1.11 and 1.12.

The model constructed indicates clearly the position of the bulky Dpp group below the plane of the β-turn, well removed from the rest of the structure. The tyrosine and methionine side chains extend outwards in the same plane as the β turn with the aromatic ring of the the phenylalanine residue adopting a more peripheral position.

1.6.6 Energy Minimisation

The model constructed from the nmr data was subjected to an energy minimisation procedure contained in the Alchemy II program. This produced a structure essentially the same as the one derived experimentally. Figure 1.13 shows the minimised structure minus the hydrogen atoms for clarity. The spatial arrangement of the amino acid

side chains and Dpp blocking group relative to the β turn are essentially those derived from the nmr data. Figures 1.14, 1.15 and 1.16 show the fragments of the minimised structure corresponding to those shown for the non-minimised structure.

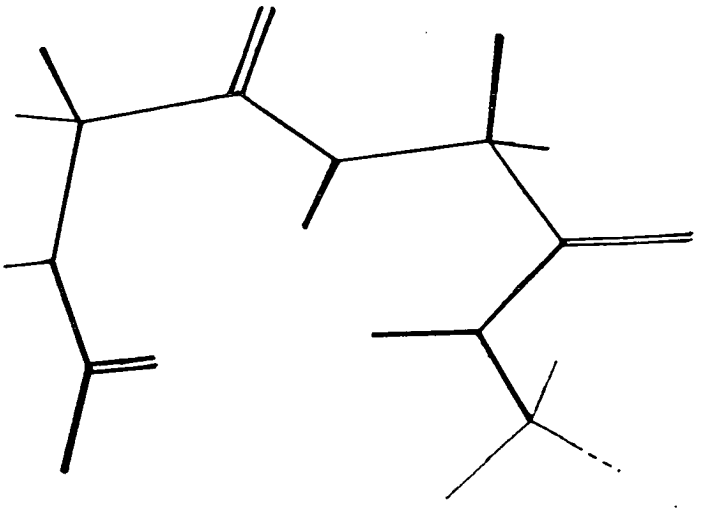


Fig. 1.9 The integral β turn of Dpp.Met^5 enkephalin

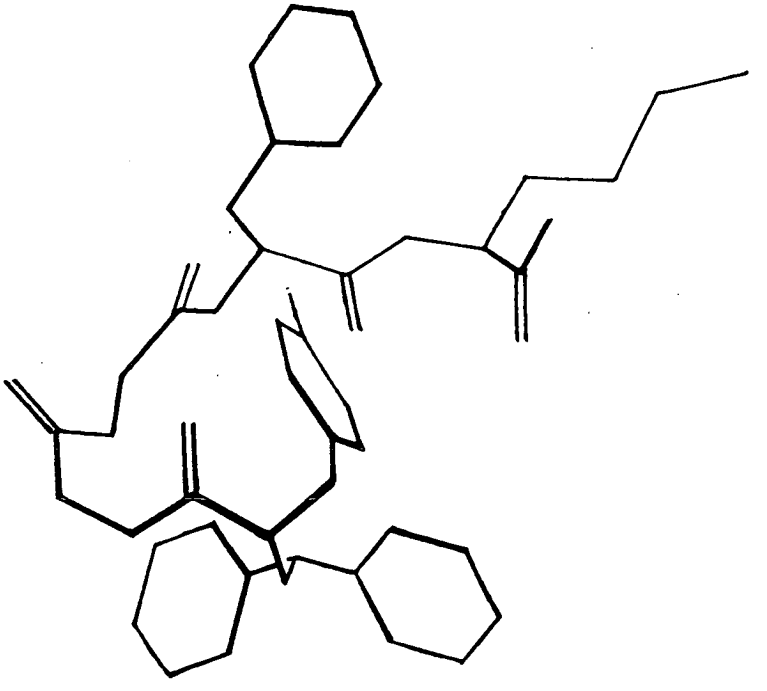


Fig. 1.10 Backbone structure of Dpp.Met^5 enkephalin

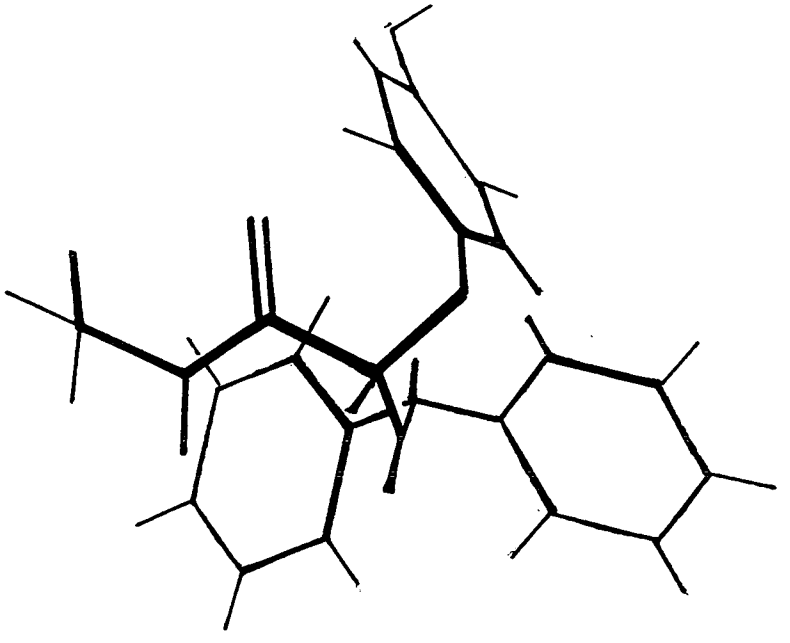


Fig. 1.11 Perspective view of the N-terminal portion of the peptide

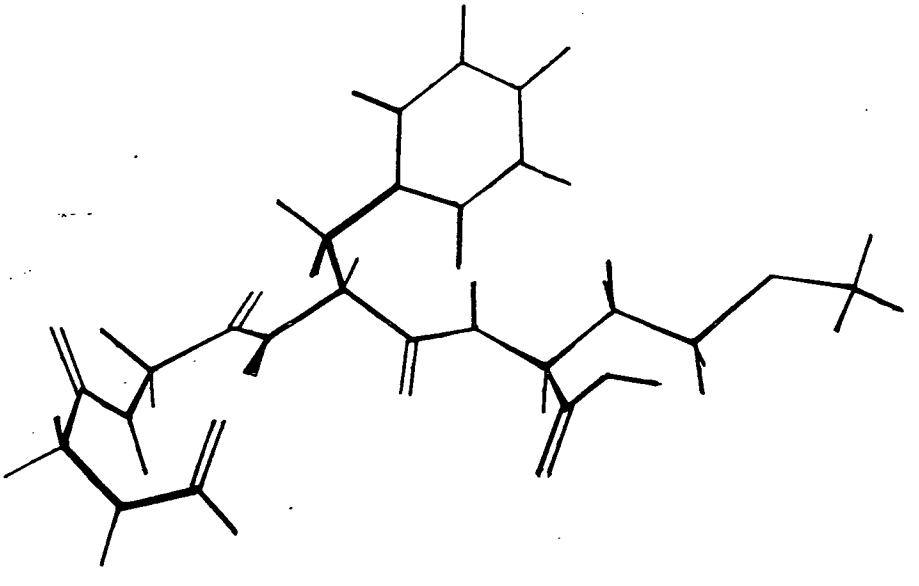


Fig 1.12 Perspective view of the C-terminal region of the peptide

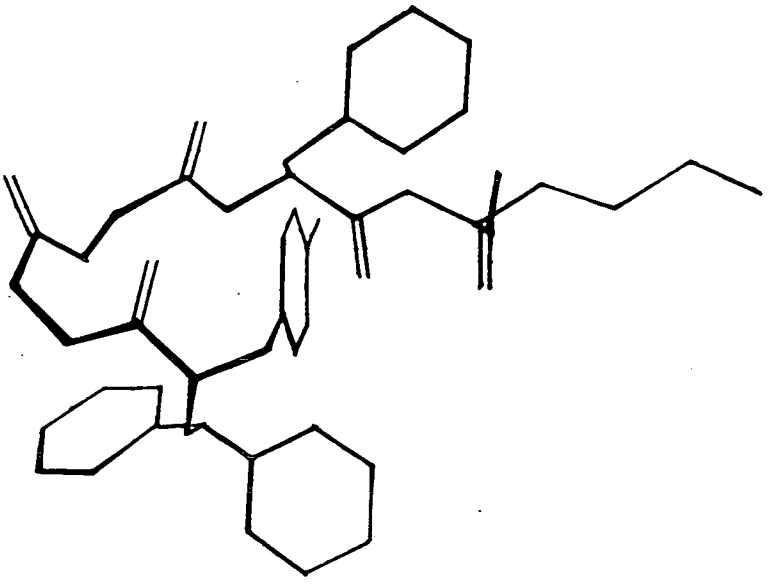


Fig. 1.13 Minimised structure

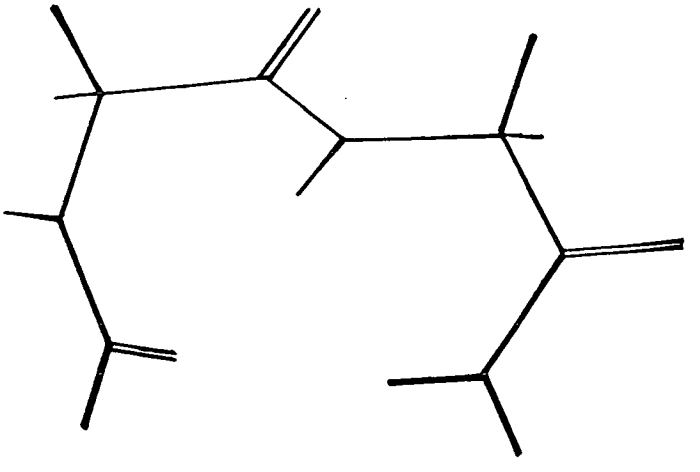


Fig 1.14 Integral β turn of the minimised structure.

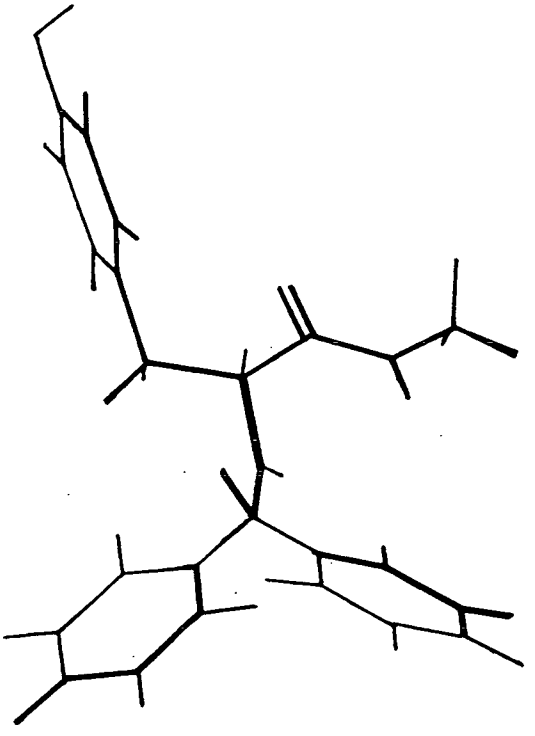


Fig. 1.15 N-terminal portion of minimised structure

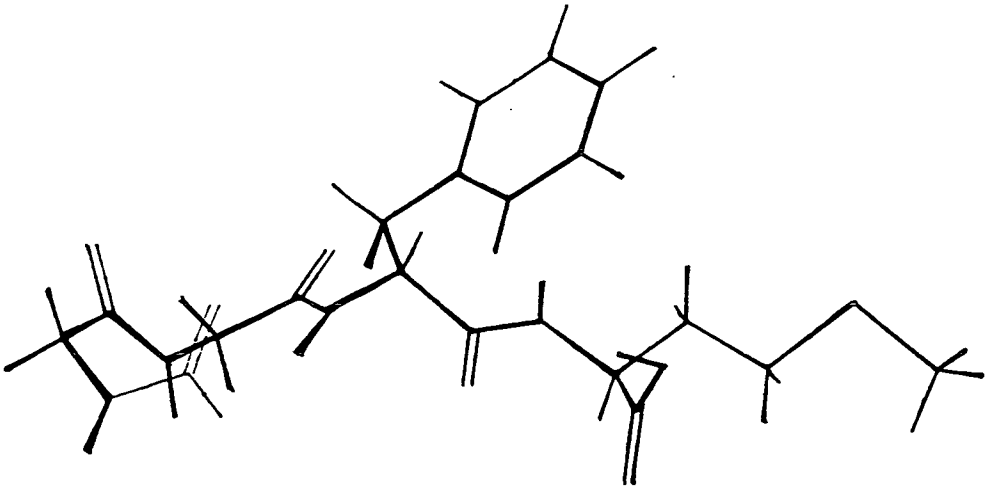


Fig. 1. 16 C-terminal portion of the minimised structure.

1.6.7 Conclusion

The initial choice of a type I' β turn in preference to a type I β turn would seem to be justified when considering the spatial requirements of the Dpp protecting group and the tyrosine side chain. While changing the β-turn from type I' to type I does not result in any conformational change at the carboxy terminus of the structure, i.e. the Phe^4 and Met^5 residues, the N-terminus is considerably disrupted, resulting in the Dpp group impinging on the carboxy terminal region of the peptide. Thus a type I' β turn structure is favoured on steric grounds.

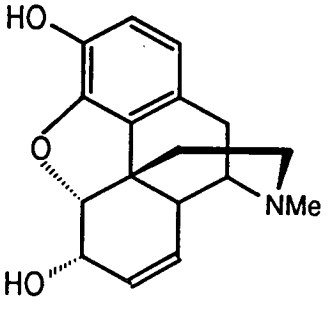
While the type I' turn described above is probably the dominant form in dmso solution the presence of other less populated folded and extended conformations cannot be ruled out. It is however clear that there exists a locally folded form containing a type I' β turn in dmso solution. It is tempting to speculate that this may be the conformation associated with binding to the opioid receptor.

CHAPTER 2

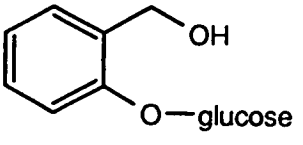
2 Tyrosine

2.1 Introduction

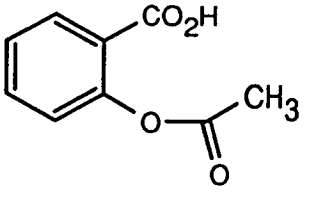
For thousands of years man has utilized the pool of natural products for uses which have included medicines, poisons and recreational aids. Examples include opium, the air dried milky exudate from unripe fruit capsules of the opium poppy, of which morphine (25) is the major active constituent. This has been used since ancient times for its marked analgesic and narcotic effects. Salicin (26) found in extracts of willow bark, a source used for centuries in a variety of healing preparations for the treatment of fevers, was partly responsible for the synthesis of a more potent analogue acetyl salicyclic acid or asprin (27). The toxic properties of hemlock have been known since the times of the ancient Greeks, being reported as the source of poison used in the execution of Socrates[89]. The plant is now known to contain the neurotoxic alkaloid coniine (28).



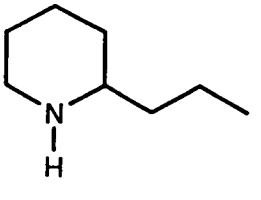
(25)



(26)



(27)



(28)

The seemingly infinite number of chemical species present in nature fascinated early chemists who could not rationally account for the production of such diverse and complex species and for many years it was thought that compounds present in living organisms contained some vital force. The realisation that natural products were formed by basically unexceptional chemical processes resulted from Wohlers "unexpected" preparation of urea, a well known constituent of urine, via the pyrolytic rearrangement of ammonium cyanate.

As the structures of an increasing number of natural products became known, attempts were made to classify them in terms of structural type. It was therefore possible to spot characteristic structural features which suggested a particular precursor. Thus followed the total synthesis of natural products via hypothetical biosynthetic pathways which provided further evidence that biosynthesis proceded via standard chemical reactions without nature resorting to some "biological trickery". In addition the hypothetical intermediates of these biomimetic pathways were often found to be present in the organism containing the natural product under investigation.

The real breakthrough in the determination of biosynthetic pathways came with the feeding of isotopically labelled precursors and the isolation of label containing natural products. Thus the direct incorporation of small molecules into natural products could be demonstrated.

2.2 Use of Labelled Precursors

Incorporation of labelled precursors is normally effected in plants by feeding solutions to the appropriate tissue, or for micro-organisms by

their addition to the culture medium during the phase of maximum production of the required metabolite. Classical feeding studies involve feeding of ^{14}C labelled precursors followed by isolation and systematic degradation of the resulting metabolite to determine the distribution of the labelling.

More recently with the advent of nmr spectroscopy non-destructive methods of locating labelled atoms have been devised based on the recognition of suitably enriched isotopes, notably ^{13}C, ^{15}N, ^{2}H and ^{17}O[90].

Nmr spectroscopy offers many advantages over other available techniques when applied to biosynthetic problems. It is the only technique that can separately detect the presence of most biosynthetically useful elements and the only non-destructive method capable of determining the locations and concentrations of isotopic labels in a metabolite. Furthermore spin-spin coupling and isotope shift effects can be used to provide direct evidence for the incorporation of intact biogenic units and for the identification of processes involving bond formation and bond cleavage. Such information can only be deduced indirectly by other methods.

The disadvantage of nmr is the quantities of labelled metabolite required for analysis, generally several milligrams, which may not always be available in some cases. This problem is not encountered when using radioisotopes due to the increased sensitivity of the detection techniques available.

The problem of analysing small quantities of stable isotope labelled metabolite can be overcome using mass spectroscopy to determine the incorporation of labelled material. Identification of the position of enrichment being possible from the composition of the fragment ions. This

is however a much more labour intensive process and one in which the signal to noise is considerably reduced[91].

2.3 ^{13}C Labelling

^{13}C labelling is the most widely employed method for probing biological pathways and many reviews exist on the subject[92,93,94]. The natural abundance of ^{13}C is 1.1% so providing the enrichment is significantly outwith the error of measurement (about 0.5 - 1.0% ^{13}C above natural abundance is usually sufficient) feeding a singly labelled precursor allows direct identification of the enriched site from the ^{13}C spectrum.

When a precursor is labelled with ^{13}C at two positions, which are spin-spin coupled, its fate can be determined by the absence or presence of this coupling in the metabolite. If coupling is preserved in the ^{13}C spectrum of the labelled metabolite it is convincing evidence that part at least of the precursor is incorporated intact. Loss of the coupling accompanied by enrichment is proof of bond scission.

Subsequently the use of amino acids enriched with ^{13}C and ^{15}N in adjacent positions as probes have been used to effect in solving biosynthetic problems. α-Amino acids enriched at the 2-position with ^{13}C and in the amino function with ^{15}N offer significant potential as they possess in essence a labelled C-N bond the integrity of which can be detected by observing the one bond coupling, $^{1}J_{CN}$, in the ^{13}C or ^{15}N nmr spectrum. This enables the characterisation of biochemical processes which can be conjectured to involve either conservation or cleavage of C-N bonds.

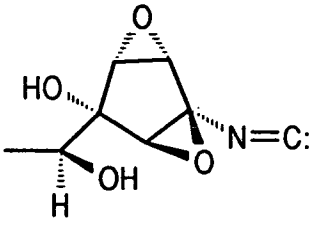
[2^{13}C,^{15}N]Glycine has been used to demonstrate that the carbon-nitrogen unit of the 2-amino-3-hydroxycyclopent-2-enone moiety of the antibiotic asukamycin(29) is derived intact from glycine[95]. Other examples of the use of [2^{13}C,^{15}N] labelled amino acids have appeared in the literature[96,97].

2.4 Tyrosine

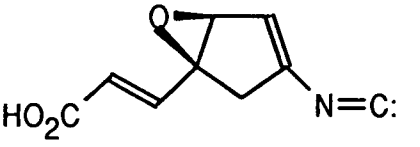
Many examples exist in the literature of biosynthetic pathways proposed to involve the incorporation of carbon-nitrogen units whose supposed origin is the 2C-N portion of tyrosine.

2.4.1 Cyclopentyl Isocyanides

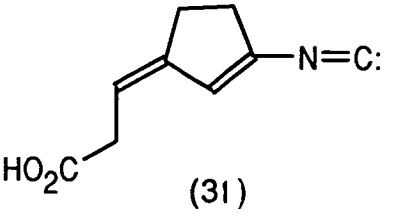
Cyclopentyl isocyanides such as trichovidrin (30), 3-(3'-isocyanocyclopent-2-enylidene)propionic acid (31), and the epoxy acid (32) are produced by strains of *Trichoderma hamatum* as major metabolites[98].



(30)

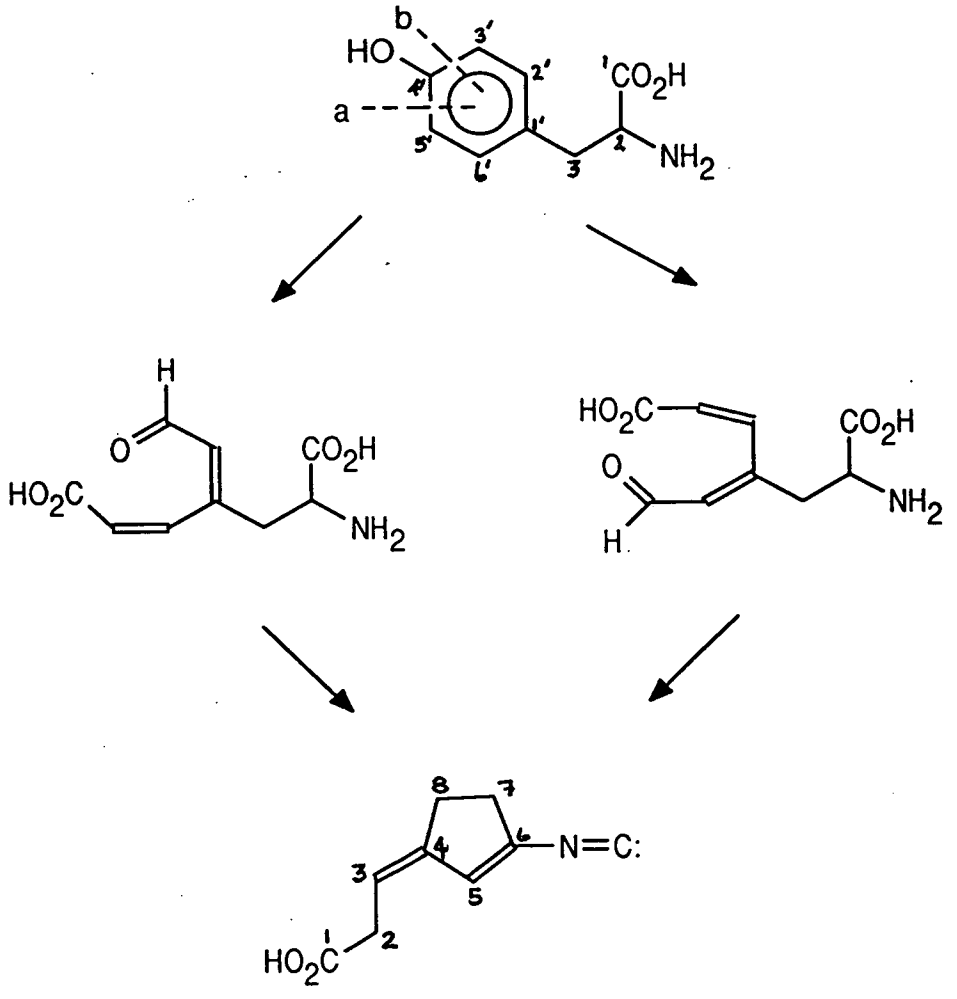


(32)



(31)

The biosynthesis of (32) has been studied extensively and the proposed biosynthetic pathway is shown in scheme 2.1. This shows the formation of (32) via the oxidative cleavage of the tyrosine aromatic ring adjacent to the phenolic function at a or b followed by a cyclisation process involving the carboxylic carbon atom and one of the two atoms originally meta to the phenolic function, depending on the initial cleavage site.



(scheme 2.1)

The proposed ring cleavage and recyclisation accompanied by loss of a single carbon atom is in agreement with the observation of a loss of

one carbon atom on feeding [U-^{14}C]tyrosine[98,99], and with the retention of only two of the three contiguous carbon atoms of [3',4',5'-$^{13}C_3$]tyrosine in the final product[99].

Feeding of [3-^{13}C] and [1-^{13}C] tyrosines produced (32) labelled in positions 5 and 7 respectively. This along with the observation that [1',2'-$^{13}C_2$]tyrosine labels both side chain and ring functions in the final product indicates formation of the cyclopentene ring from the aliphatic portion of tyrosine and two of the carbon atoms of the aromatic ring, 1' and 2' or 1' and 6' depending on the initial cleavage site[99].

Although this gives a seemingly comprehensive account of the formation of (32) the biosynthetic origin of the isonitrile moiety remains unproven. Here is a case where [2-^{13}C,^{15}N]tyrosine would be a valuable probe.

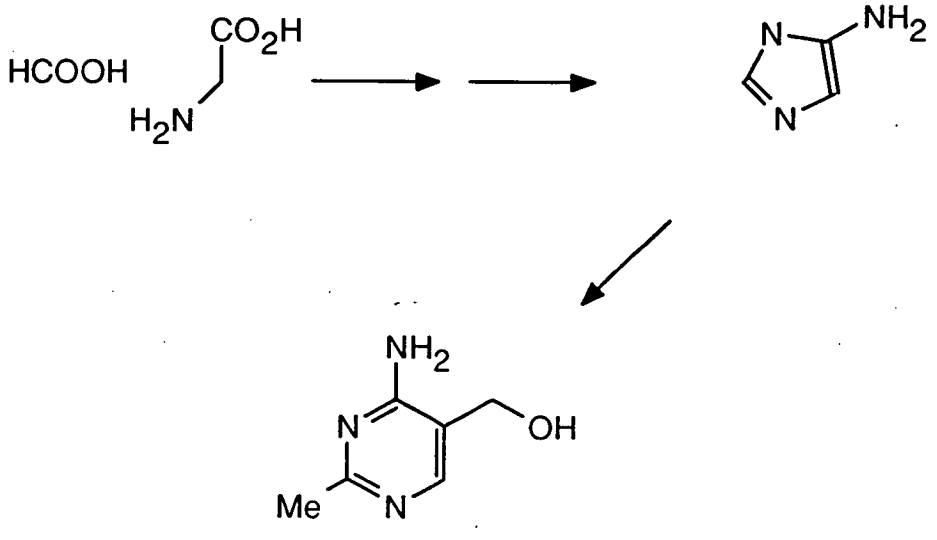
2.4.2 Thiamine

Thiamine (33), also known as vitamin B_1, is the precursor for the metabolically active thiamine pyrophosphate (34) which is the essential coenzyme involved in the enzymatic decarboxylation of α–keto acids and in the transketolase reaction.

Although much is known of the function of thiamine the biosynthesis of this compound remains largly undetermined. The enzymatic steps involved in the conversion of the pyrimidine and thiazole moietes to thiamine in yeast were elucidated some time ago[100,101], however the biosynthesis of the pyrimidine and thiazole moieties themselves are as yet undefined.

The primary precursors of the pyrimidine unit differ in eukaryotes and prokaryotes. In prokaryotes the intermediacy of 5'-

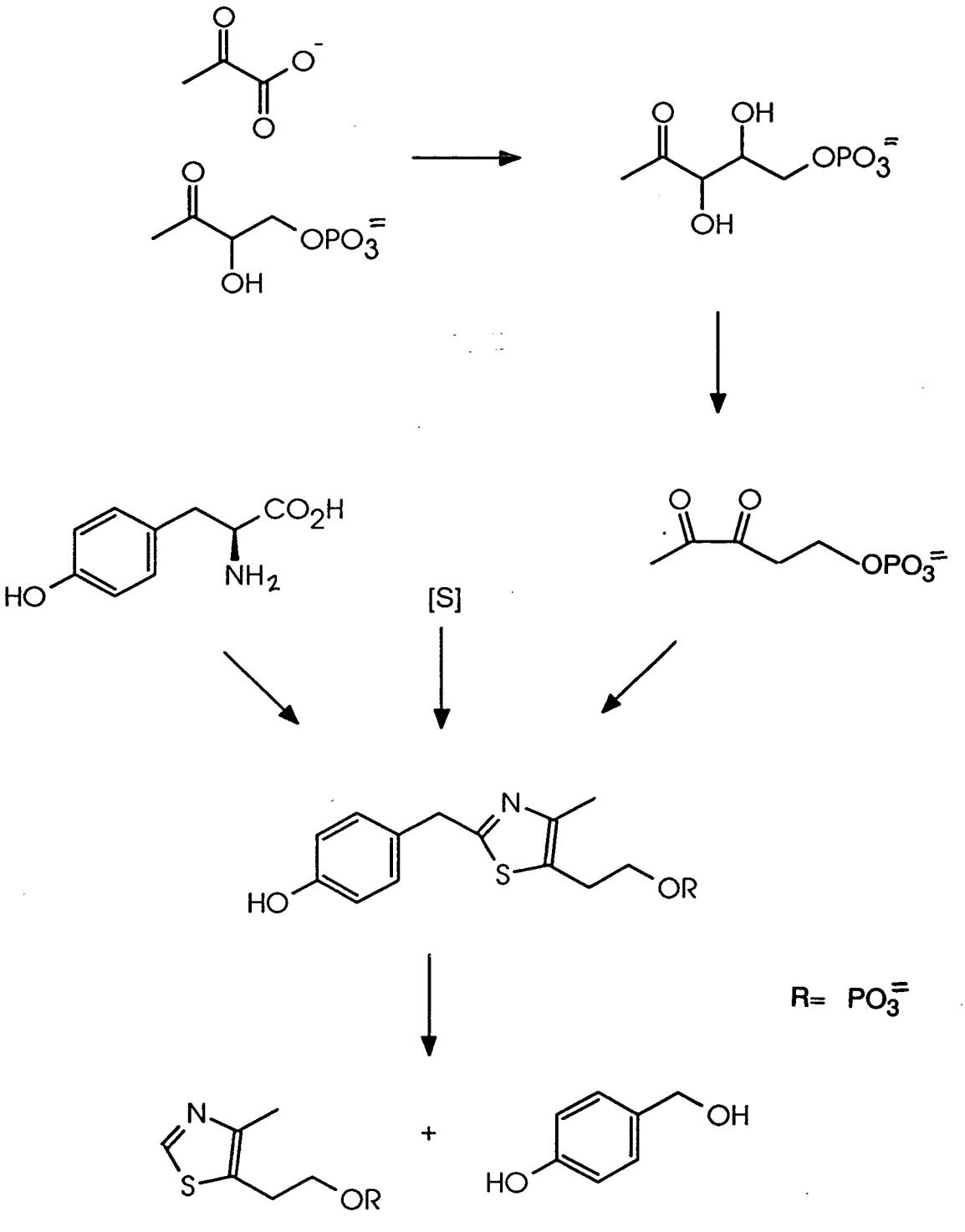
aminoriboimidazoleribonucleotide (5'-AIR) derived from formate and glycine is widely accepted, scheme 2.2. Little is known of the biosynthesis of the pyriminidine moiety in eukaryotes and it would appear that there are two competing pathways[102]. Both pathways have been shown to utilize formate and a pentose. Recently Tazuya *et al* demonstrated the incorporation of the two ring nitrogens of histidine into the N-3 and amino nitrogen at the C-4 of the pyrimidine[103] indicating that at least in the major pathway the N-1, C-2 and N-3 of histidine are incorporated into the N-3, C-4 and the amino-N at C-4 unit of the pyrimidine moiety, scheme 2.3.



Scheme 2.2

The thiazole unit originates from an amino acid, a five carbon sugar and a sulphur source. In yeast the precursors are glycine and a pentulose phosphate[104], and in bacteria they are tyrosine and a deoxypentulsoe[105]. The sulphur source is presumed to be cysteine although this has yet to be confirmed.

An outline of the proposed biosynthetic pathway for the synthesis of the thiazole moiety in *E. coli* is shown in scheme 2.4.



(scheme 2.4)

The presence of tyrosine in the scheme was shown by a series of experiments with [U-^{14}C], [2-^{14}C] and [^{15}N] tyrosine in *E. coli* and *Salmonella typhinurium* showing that the C-2 and ring nitrogen of the thiazole moiety are derived from tyrosine[106]. White identified 4-hydroxybenzyl alcohol as a metabolite of tyrosine in *E. coli* grown in the absence of thiamine or the thiazole[107].

Here again the intact incorporation of a 2C-N unit from tyrosine could be demonstrated using [2-^{13}C, ^{15}N] labelled tyrosine although the small amounts of thiamine produced in bacteria would negate the use of nmr for detection of the label. However mass spectroscopy has been used effectively to study the incorporation of stable isotopes into thiamine.

2.5 The Synthesis of Labelled Tyrosines

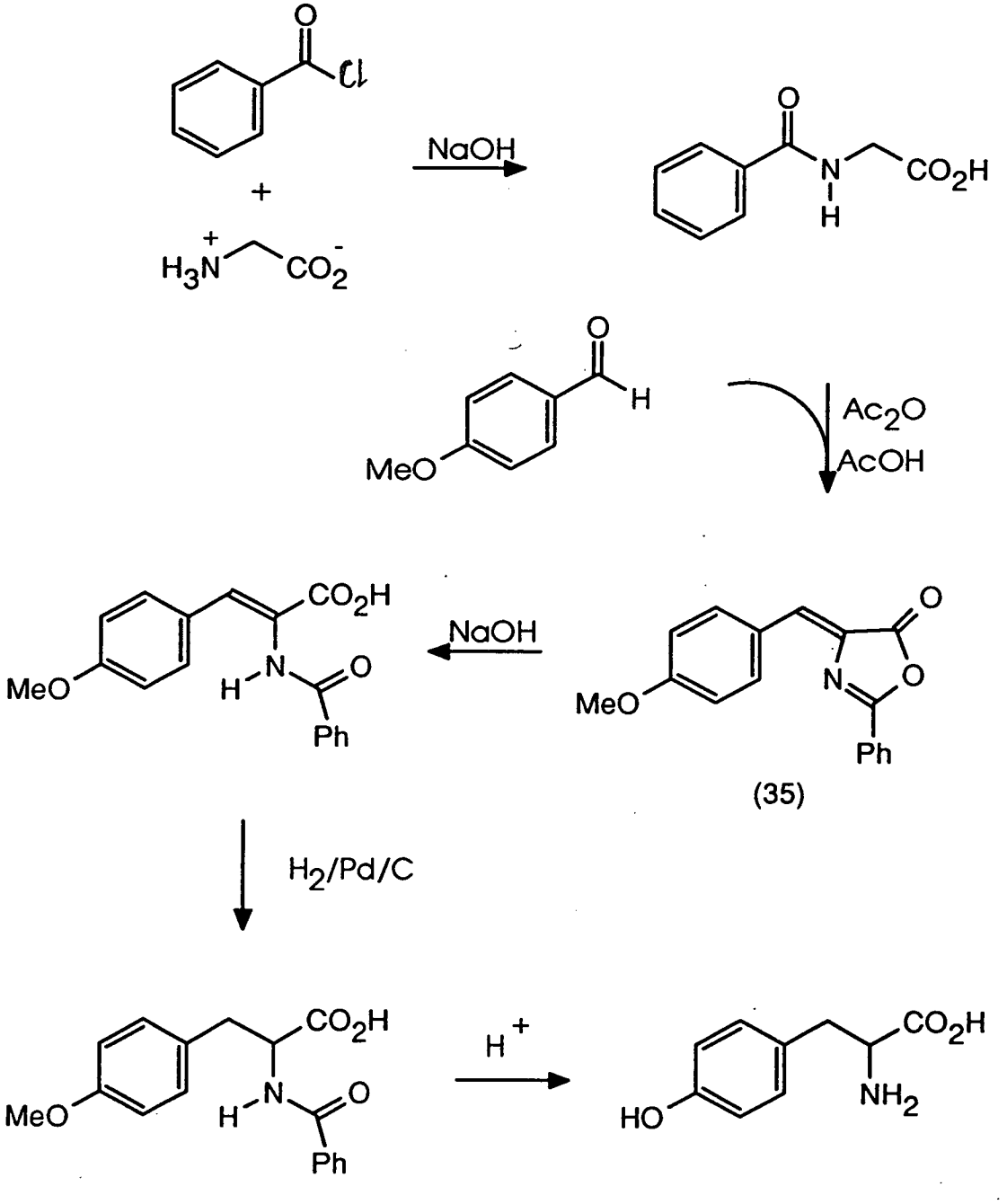
Due to the commercial unavailability of a [2-^{13}C, ^{15}N]tyrosine and its obvious usefulness as a biosynthetic probe, we decided to investigate the possibility of synthesising such a probe. Routes to many labelled tyrosines have been reported in the literature and here we review a few of the more profitable routes.

2.5.1 Carboxy Labelled Tyrosine

Many methods are available for the introduction of stable isotopes into the carboxy function of tyrosine. The two most widely reported are based on the Erlenmeyer azlactone synthesis and the Bucherer hydantoin synthesis.

Introduction of the label via Erlenmeyers' classic synthesis[108] (scheme 2.5) involves the preparation of hippuric acid from carboxy

labelled glycine followed by condensation with 4-methoxy benzaldehyde to form the oxazolone (35). Base hydrolysis of (35) yields the corresponding α-benzylamine acrylic acid and subsequent hydrogenolysis of the double bond followed by hydrolysis yields the racemic α–amino acid.



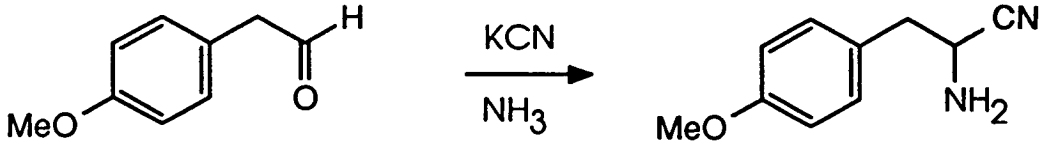
(35)

scheme 2.5

Although favoured for some time this approach has become less frequently used due to the high cost of the labelled starting material and the number of low yielding reactions involved. Yields reported for this route vary from 3-60% even for scaled up reactions[109].

The Bucherer hydantoin synthesis offers a more direct route to carboxy labelled tyrosine. This involves the reaction of the 4-methoxyphenylacetaldehyde bisulphite adduct with potassium cyanide and ammonium carbonate to produce the 4-methoxybenzylhydantoin which upon acid hydrolysis yields D,L-tyrosine (scheme 2.6). The attraction of this route lies in the label source of the the C1 atom being the inexpensive potassium cyanide and in the fact that there is only two steps to concern with. This route was first employed by Lotfield[110] to produce [1-^{14}C]-D,L-tyrosine with a reported yield of 65% from the cyanide. However later workers have been unable to reproduce his yields[111,109].

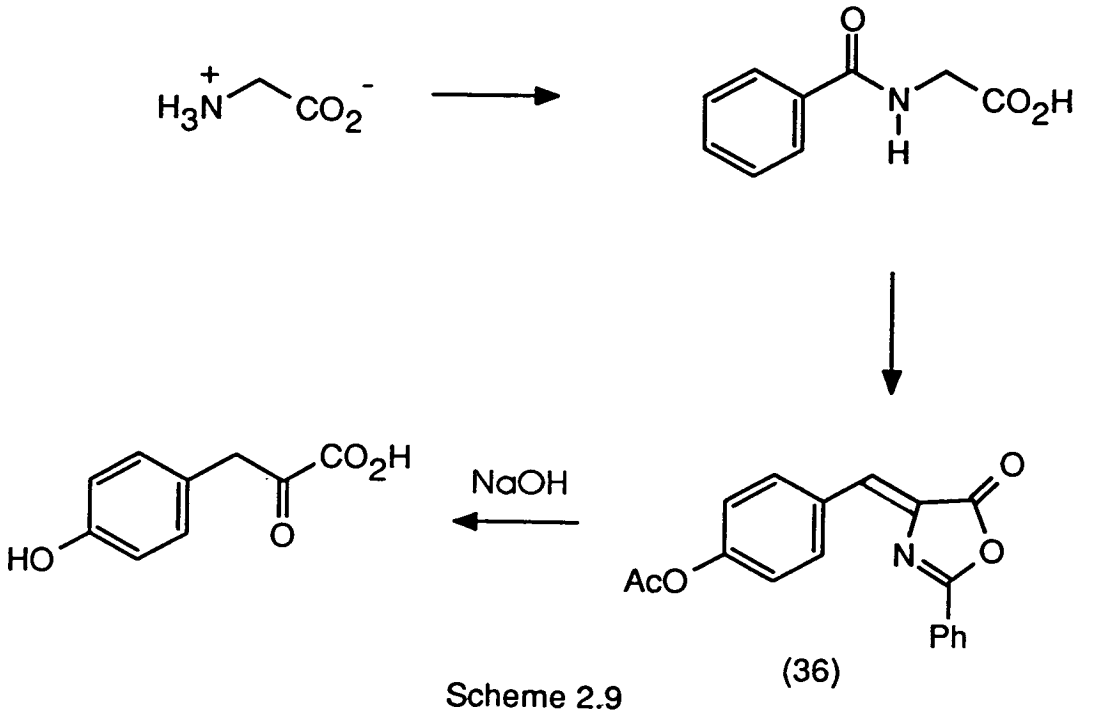
A third method of preparing carboxy labelled tyrosine has recently been reported using the Strecker route[109]. This involoves the reaction of 4-methoxyphenylacetaldehyde with potassium cyanide and aqueous ammonia to produce the corresponding α-aminonitrile, scheme 2.7. Subsequent hydrolysis of the α-amino nitrile yielded the D,L-amino acid in 45% yield based on potassium cyanide.



Scheme 2.7

2.5.2 ^{15}N Labelled Tyrosine

Two established routes exist for the synthesis of ^{15}N-labelled tyrosine. The first is the standard Erlenmeyer synthesis using ^{15}N-glycine as the label source[112]. The second is the reaction of either ^{15}N-glutamic acid or ^{15}N-aspartic acid with 4-hydroxyphenyl pyruvic acid mediated by the enzymes tyrosine amino transferase and aspartate amino transferase which gives a 55% yield of the final product with respect to the labelled precursor. This method was later improved by Baldwin *et al* who utilised ^{15}N aspartic acid, aspartate transaminase and 4-hydroxyphenylpyruvic acid to prepare ^{15}N L-tyrosine in 80% yield[113] (scheme 2.8). This approach was also used in the synthesis of [^{15}N, 2-^{13}C]tyrosine using [2-^{13}C]-4-hydroxyphenylpyruvic acid prepared via the base hydrolysis of the Erlenmeyer product (36) of [2-^{13}C]glycine (scheme 2.9) giving [2-^{13}C,^{15}N]tyrosine in a yield of 36% from [2-^{13}C]glycine[113].



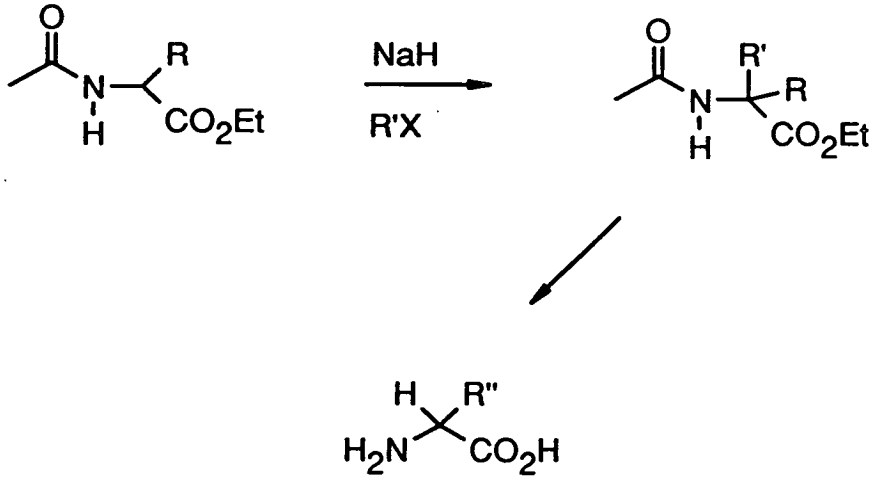
(36)

Scheme 2.9

2.5.3 Side Chain Labelled Tyrosine

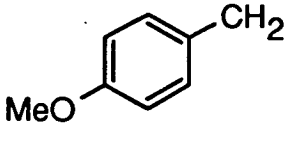
The routes reported in the literature for the synthesis of ring and 3-C labelled tyrosines include the previously mentioned Bucherer and Erlenmeyer routes. These involve the synthesis of a suitably labelled aldehyde, a phenylacetaldehyde or a benzaldehyde respectively[112,114].

A more commonly used route however is the alkylation of the sodium salt of either diethyl acetamidomalonate or ethyl acetamidocyanoacetate with an appropriately labelled 4-methoxybenzyl bromide[112,115]. Hydrolysis of the alkylation product yields the appropriately labelled tyrosine, scheme 2.10. An example is the synthesis of [3',5'-$^{13}C_2$]tyrosine performed by Hruby which gave an overall yield of 22% over ten steps[115].

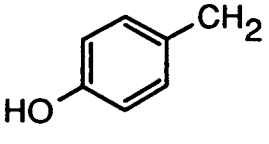


R=CN or CO_2Et

R'=



R"=



Scheme 2.10

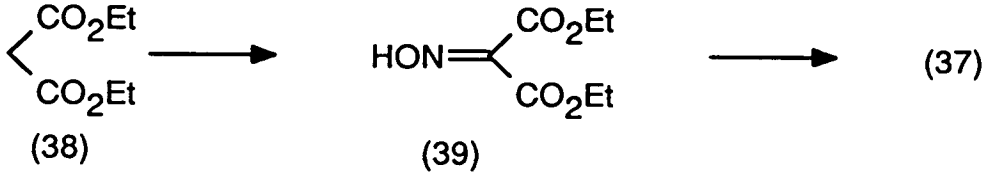
2.6 Discussion

The body of literature outlining various routes to labelled tyrosines offers a few ideas for designing a viable route to [2-^{13}C,^{15}N]tyrosine. Here we discuss three attempts at formulating a synthesis of the required product.

2.6.1 Diethylacetamidomalonate

One of the most useful and general methods for the synthesis of amino acids is the alkylation of diethyl acetamidomalonate (37) with an alkyl halide, scheme 2.11. Several applications of this procedure have been well documented in the literature[112,116,117].

In order to prepare [2-^{13}C,^{15}N]tyrosine using this methodology it is first necessary to prepare [2-^{13}C,^{15}N]diethyl acetamidomalonate. To test the viability ot this route for a labelled tyrosine synthesis (37) was prepared by the two step procedure shown in scheme 2.12[118]. Treatment of diethyl malonate (38) with sodium nitrite in the presence acetic acid produces the intermediate diethyl isonitrosomalonate (39) which was subsequently reduced and acylated to produce (37) in 65% yield.

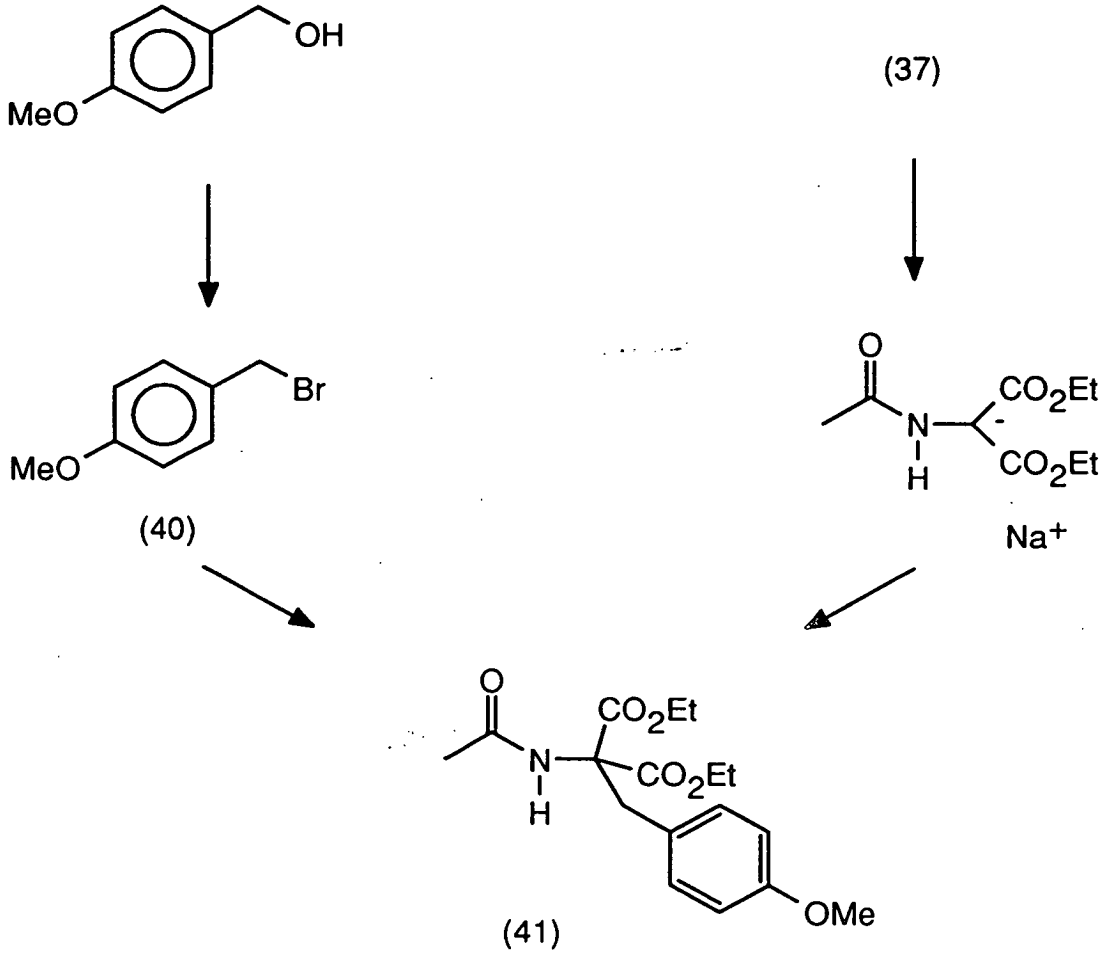


(38) (39)

scheme 2.12

Alkylation of (37) was then attempted using 4-methoxybenzyl bromide (40) prepared by the method of Baldwin *et al*[112]. Generation of the sodium salt of (37) using sodium ethoxide followed by addition of the halide (40), scheme 2.13, gave after purification (41) in only 17% yield. It

had been noted in other experiments involving the alkylation of (37) using alkylhalides that better results were obtained using sodium hydride as the base[119]. Using this method gave a much improved yield of 59%.

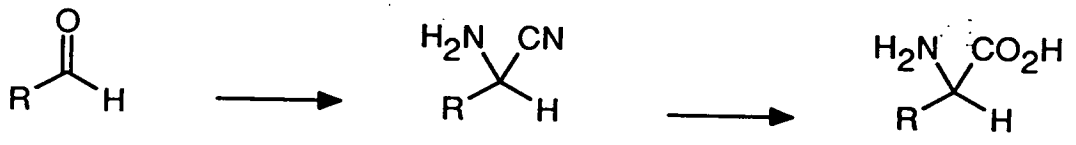


scheme 2.13

At this stage it was felt that yields of 65% and 59% for the first two steps coupled with the expensive nature of one of the starting materials, diethyl [^{13}C] malonate, did not justify this as a potential route to [2-^{13}C,^{15}N]tyrosine. This line of research was therefore discontinued at this point.

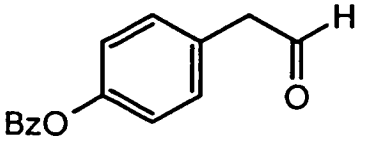
2.6.2 Strecker Route

The reaction of an aldehyde or ketone with an amine and sodium cyanide produces an α-aminonitrile which can readily be hydrolysed to give the corresponding α-amino acid (scheme 2.14). This reaction known as the Strecker reaction, has been used extensively for the synthesis of both labelled[109,120] and unlabelled[121] amino acids. While attempts to produce tyrosine by this route have not always been successful it was thought that this route may be useful for the synthesis of [2-^{13}C,^{15}N]tyrosine since the nitrogen label could be incorporated from the relativel cheap and readily available ^{15}N ammonium chloride.

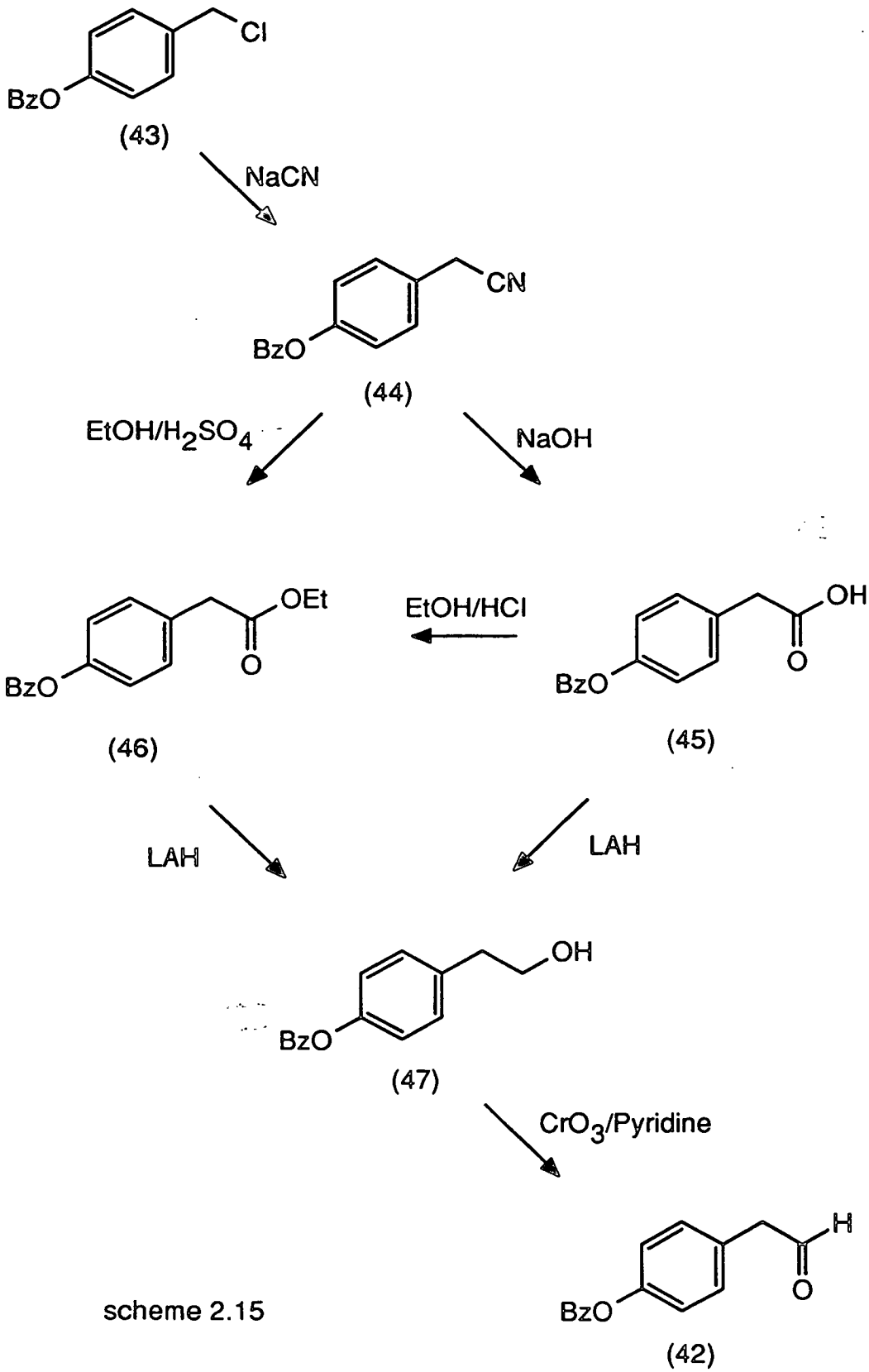


scheme 2.14

The initial target thus became the development of a satisfctory route to a suitably protected [1-^{13}C] 4-hydroxyphenylacetaldehyde. Since benzyl ethers can readily be removed by hydrogenation, hydrolysis with strong acids or sodium ammonia reduction[122] we elected to prepare 4-hydroxypenylacetaldehyde(42) as the synthon. The route adopted for the synthesis of (42) is outlined in scheme 2.15 .



(42)


Cl
BzO
(43)
NaCN
CN
BzO
(44)
EtOH/H2SO4
NaOH
OEt
O
BzO
(46)
EtOH/HCl
OH
O
BzO
(45)
LAH
LAH
OH
BzO
(47)
CrO3/Pyridine
H
O
BzO
(42)

scheme 2.15

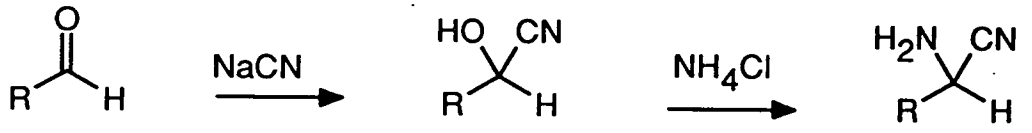
The nitrile (44) was prepared quantitatively from 4-benzyloxybenzyl chloride (43) and sodium cyanide. Although acid hydrolysis of (44) failed to provide the acid (45) without extensive decomposition of the starting material, base hydrolysis progressed smoothly to give (45) in a 73% yield.

It was originally planned to convert the acid to its ethyl ester (46) and reduce this to the corresponding alcohol (47). However conversion to the ester gave an unsatisfactory yield of 54%. Subsequent reduction of ester with lithium aluminium hydride (LAH) gave 4-benzyloxyphenethyl alcohol (47) in 73% yield; an overall yield of 40% for the two steps. Attempts to produce the ethyl ester directly from the nitrile by alcoholysis were unsuccesful. However direct reduction of the acid to the alcohol using an excess of LAH was successful giving a 73% yield of the the alcohol (47).

Oxidation of (47) using chromium trioxide and pyridine gave 4-benzyloxyphenylacetaldehyde (42) in 80% yield, 43% yield from the chloride (43). This route is therefore well suited to the preparation of [1-^{13}C]-4-benzyloxyphenylacetaldehyde, since the ^{13}C label could be easily incorpoated using ^{13}C sodium cyanide as the label source in the initial step.

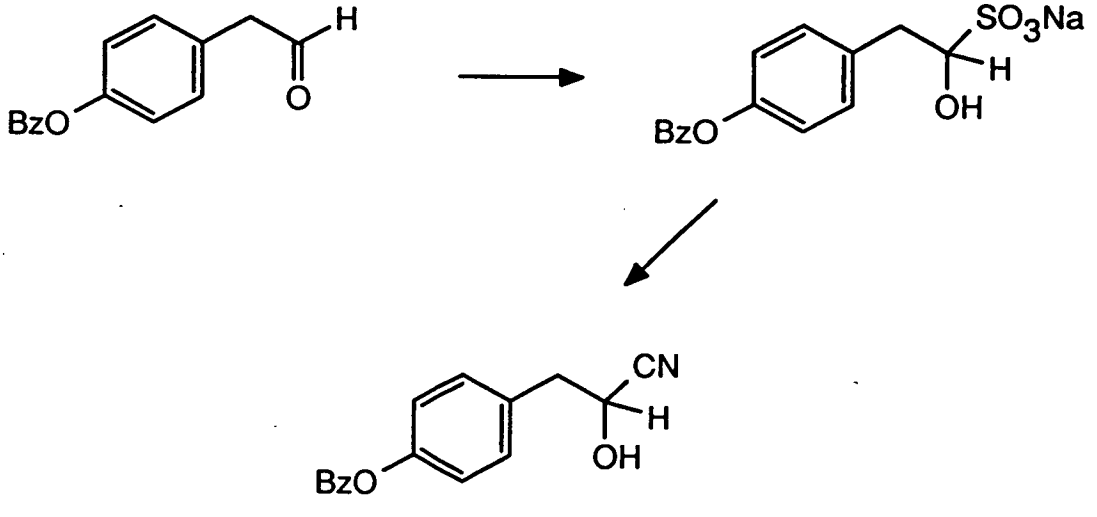
While being ideal for the preparation of [1-^{13}C] (42) this route was unsuitable for the production of the large amounts of (42) required for the optimisation of the Strecker reaction conditions. A more economical large scale route proved to be the esterification and etherification of 4-hydroxyphenylacetic acid (48) with benzyl chloride and reduction of the benzyl ester (49) with LAH to give the alcohol[123], scheme 2.16, in 53% yield from the acid (48).

It was decided to employ a modification of the Strecker procedure where the intermediate cyanohydrin is prepared and isolated before carrying out the amidation reaction (scheme 2.17). This methodology has been found to give greater yields than the classic one pot Strecker procedure[120,121].



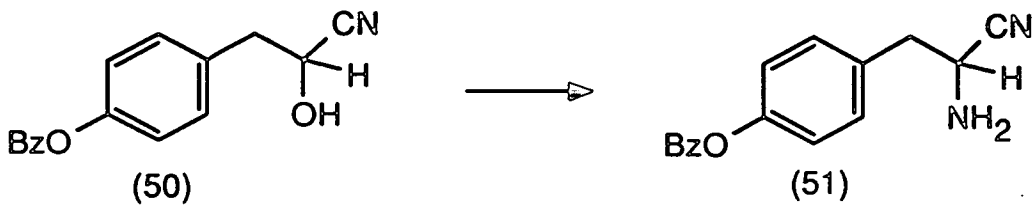
scheme 2.17

Preparation of the cyanohydrin (50) was accomplished using the method described by Pierson *et al*[121]. Thus treatment of the preformed sodium bisulphite addition product of the aldehyde with aqueous sodium cyanide gave 4-benzyloxyphenyl-α-hydroxy-propionitrile (51) in 80% yield (scheme 2.18).



(50)

scheme 2.18



(50) (51)

scheme 2.19

Temperature	Time	Yield
60°C	4 hours	40%
70°C	1 hour	41%
70°C + room temp.	30 min. + 12 hours	39%

Table 2.1

The conversion of a cyanohydrin to the corresponding α-aminonitrile has generally been effected either by passing gaseous ammonia into a solution of the cyanohydrin or by the addition of concentrated ammonia solution to an aqueous solution of the cyanohydrin[109,121]. Since neither of these procedures are practical for a ^{15}N labelled synthesis some other reagent had to be used. Ammonium chloride dissolved in aqueous sodium hydroxide provides a convenient source of ammonium hydroxide solution from the readily availible ^{15}N labelled precursor.

Simple addition of an alkaline solution of ammonium chloride to an aqueous ethanol solution of the cyanohydrin gave after purification only a 13% yield of the required product and in addition hydrolysis of the crude product with 6N hydrochloric acid provided only a very small amount of impure material. This may be due to the initial insolubility of the cyanohydrin upon addition of the ammonia solution and subsequent decomposition of the unstable aminonitrile.

Slow addition of the ammonia solution to the cyanohydrin (50) over a period of thirty minutes while maintaining the temperature at 60°C eradicated the insolubility problem and gave a 40% yield of the required product from the alcohol (47), scheme 2.19, which is higher than yields reported in the literature for similar reactions[110,124]. Attempts to optimise the reaction conditions by employing differing temperatures and reaction times did not significantly alter the yields obtained, table 2.1.

Recently the application of ultrasound for increasing the intimacy of reagents in heterogeneous reaction systems has attracted increasing interest[125]. However sonication of a suspension of the cyanohydrin in an aqueous ethanol solution of ammonium chloride and sodium hydroxide at

40°C gave only a 21% of (51). This represented a significant reduction in yield compared with that obtained using thermal agitation.

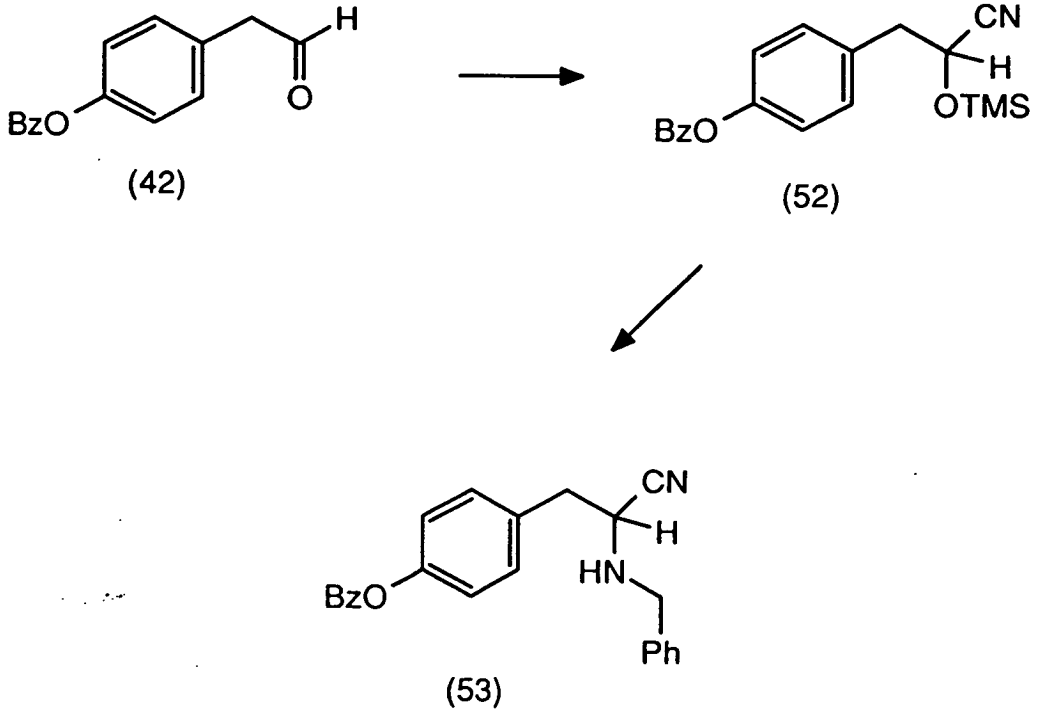
An alternative method of obtaining α-aminonitriles from aldehydes described by Hanafusa *et al* using alumina as an inorganic support with, or without, ultrasonic irradiation has been reported to give increased yields compared with standard Strecker conditions[126]. For example phenyl-α-aminopropionitrile has been succesfully prepared from benzylaldehyde, ammonium chloride and potassium cyanide in 64% yield in the abscence of ultrasonic radiation and in 90% yield ,with ultrasonic irradiation.

4-Benzyloxyphenylacetaldehyde was heated in a suspension of potassium cyanide, ammonuim chloride and alumina in acetonitrile. This however provided only a multicomponent reaction mixture from which none of the required product could be isolated. Likewise treatment of the cyanohydrin (50) with ammonium chloride and alumina in acetonitrile also failed to produce any α-aminonitrile.

Mai and Patil[127] have described a method for producing a number of α-amino nitriles from the corresponding trimethylsiloxynitriles and an amine in methanol solution : the siloxynitriles being prepared by the reaction of an aldehyde with trimethylsilyl cyanide in the presence of a catalytic amount of zinc iodide, scheme 2.20.

The necessity to use methanol as the solvent for the amination reaction prevents the use of an inorganic ammonium salt as the nitogen source. However benzylamine should be a good source of the amine function in this reaction and this in turn should be easily prepared using the Gabriel method from benzyl chloride and potassium phthalimide[128] which is commercially available in the ^{15}N labelled form.

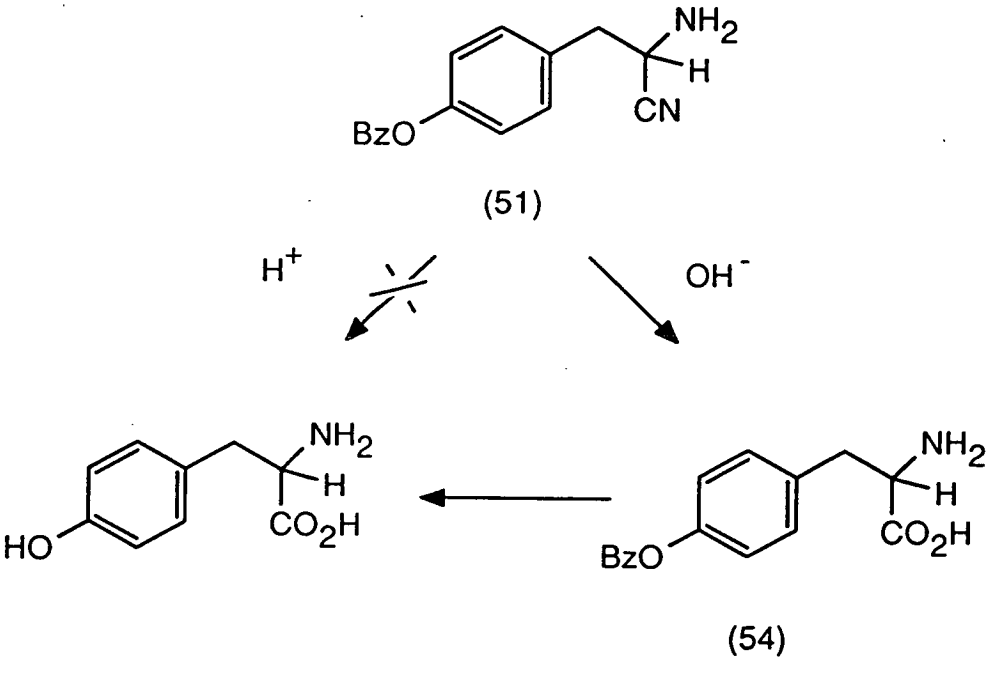
4-Benzyloxyphenylacetaldehyde (42) was converted to the trimethysilyloxy nitrile (52), a methanolic solution of benzylamine added and the solution heated under reflux. Fractionation of the reaction mixture gave a 27% yield of the 4-benzyloxyphenyl-α-aminobenzyl-propionitrile (53), scheme2.21.



scheme 2.21

Of the methods discussed here for the preparation of α-aminonitriles from aldehydes it is clear that the alumina mediated reactions and the reaction via the trimethylsilyloxy nitrile are unsuitable for a synthesis of [2-^{13}C,^{15}N]tyrosine due to the low yields of the required products. However the modified Strecker reaction gives a 42% yield of the ∝-amino-nitrile from the alcohol, over three steps, (corresponding to an average yield of 75% for each of the three steps). It was therefore decided to proceed with this route and to develop suitable conditions for the hydrolysis of (51) to D,L-tyrosine.

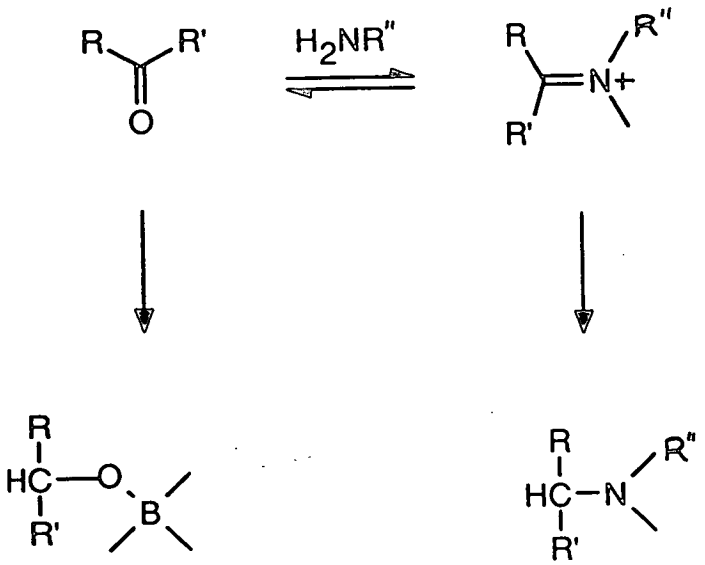
Attempts to produce tyrosine directly from the acid hydrolysis of (51) using conc. HCl, 9N HCl or 48% HBr afforded only low yields of an impure product. Base hydrolysis using potassium hydroxide in an aqueous ethanol solution however produced O-benzyl-tyrosine (54) in excellent yield, scheme 2.22. Final deprotection of (54) was acccomplished using a mixture of trifluoroacetic acid and trifluoromethanesulphonic acid at 0°C[129] to give D,L-tyrosine in 79% yield from (54), (17% yield from 4-benzyloxybenzyl chloride).



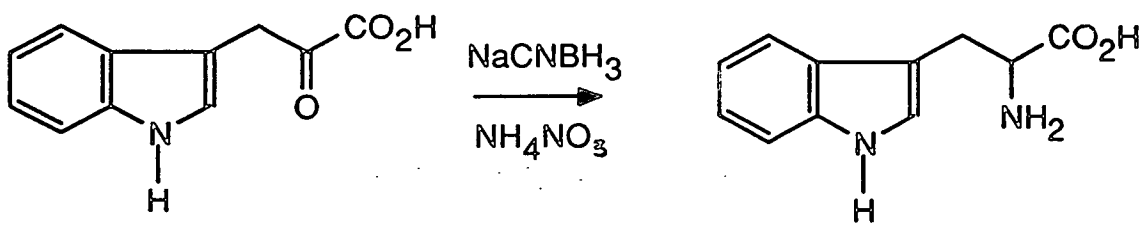
scheme 2.22

2.6.3 Reductive Amination of 4-Hydroxyphenylpyruvic Acid

The route described by Baldwin et al[99] for the synthesis of [2-^{13}C,^{15}N]tyrosine utilizing the Erlenmeyer intermediate prepared from hippuric acid and 4-acetoxybenzaldehyde offers the only literature



scheme 2.23



(55)

scheme 2.24

synthesis of this compound. The paper however offers no experimental details and although an overall yield of 36% is quoted the amount of tyrosine produced by this enzymatic method if not mentioned and it is likely that this route will only be useful in producing small amounts of labelled material. It would therefore be useful if another method of reductively aminating the pyruvic acid could be found.

Sodium cyanoborohydride is a versatile reagent that can reduce a variety of functional groups with remarkable selectivity[130]. It has been observed that sodium cyanoborohydride will reduce an iminium ion much faster than a carbonyl group (scheme 2.23).

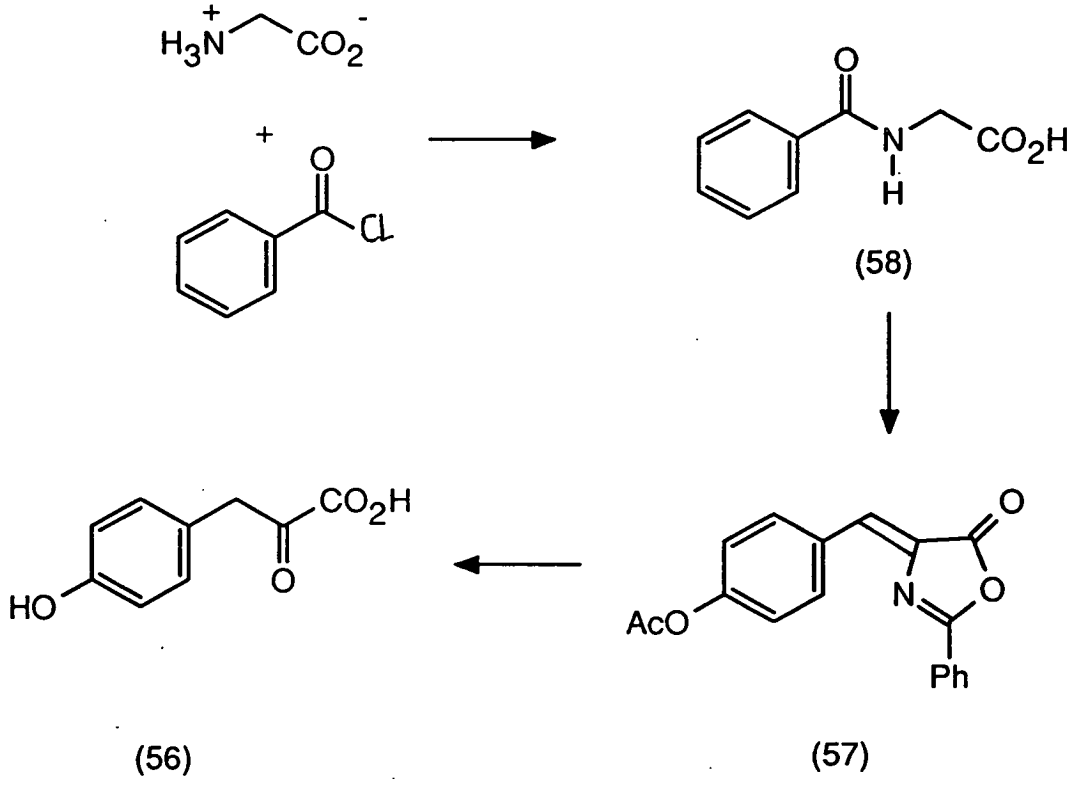
It is thus possible to reductively aminate an aldehyde or ketone by simply reacting the carbonyl compound with an amine in the presence of sodium cyanoborohydride. The optimum pH for this reaction being 3-4. This methodology has been extended to the reductive amination of substituted pyruvic acids providing a new synthesis of α-amino acids. This is claimed to be the most efficient and economical route to ^{15}N-labelled α-amino acids and has been used to synthesis ^{15}N-tryptophan (55) using ammonium nitrate as the amine source, scheme 2.24, giving a 23% yield of (55)[131].

To test the viability of this method as a route to ^{15}N-tyrosine, 4-hydroxyphenylpyruvic acid was reductively aminated, using the same procedure as described above for the tryptophan synthesis giving a 30% yield of tyrosine, scheme 2.25. The reaction was also carried out using ammonium bromide as the amine source which gave a more satisfactory yield of 90%.

In order to utilize this approach for the synthesis of [2-^{13}C,^{15}N] tyrosine it was necessary to prepare [2-^{13}C]-4-hydroxyphenylpyruvic acid

(56). Baldwin et al synthesised (56) from the oxazolone (57) via base hydrolysis. However no experimental details are available and it was therefore necessary to repeat the synthesis to establish the optimum conditions.

N-Benzoylglycine (58) was prepared from benzoylchloride and glycine in 93% yield using a method previously reported by Baldwin *et al*[98], scheme 2.26. The azlactone, 4-(4'-acetoxybenzylidene)-Δ^2-3N-oxazolin-5-one (57) was prepared from (58) under standard Erlenmeyer conditions giving (57) in 61% yield. Hydrolysis of the azlactone to produce 4-hydroxyphenylpyruvic acid however proved to be more difficult.



scheme 2.26

Unsaturated azlactones are converted to α-ketoacids by strong mineral acids or alkalies in either aqueous or alcoholic solutions. Alkaline hydrolysis has generally proven more effective[132].

Refluxing the azlactone (57) in 10% aqueous sodium hydroxide for one hour produced only the acylaminoacrylic acid (59) indicating the need to use more forcing conditions. Heating (57) in a 6.25M sodium hydroxide solution at 20°C, in a sealed reaction vessel, produced an intractible solid which could not be removed from the reaction vessel.

Attempts to produce the pyruvic acid using acid hydrolysis produced a deep purple coloured solution from which none of the required product could be isolated.

A more satisfactory procedure for the preparation of substituted pyruvic acids from azlactones has been reported[133] involving the base hydrolysis of the azlactone followed by saturation of the cooled reaction mixture with sulphur dioxide, allowing the removal of the benzoic acid bisulphite addition product from which the aketo acid is regenerated using acid hydrolysis.

The azlactone (57) was therefore heated in 10% sodium hydroxide solution at 100°C for nine hours. Saturation of the resulting solution with sulphur dioxide enabled the removal of the benzoic acid however upon heating the filtrate with hydrochloric acid to regenerate the α-keto acid a deep purple coloured solution was formed from which no product could be isolated

In summary, of the three routes investigated here the one involving the modified Strecker reaction appears to be the most satisfactory. This should enable the production of enough [2-^{13}C,^{15}N]tyrosine to carry out the biosynthetic studies discussed earlier.

Further work would have to be carried out on the third route in order to develop this into a feasible route for the production of [2-^{13}C,^{15}N]tyrosine.

CHAPTER 3

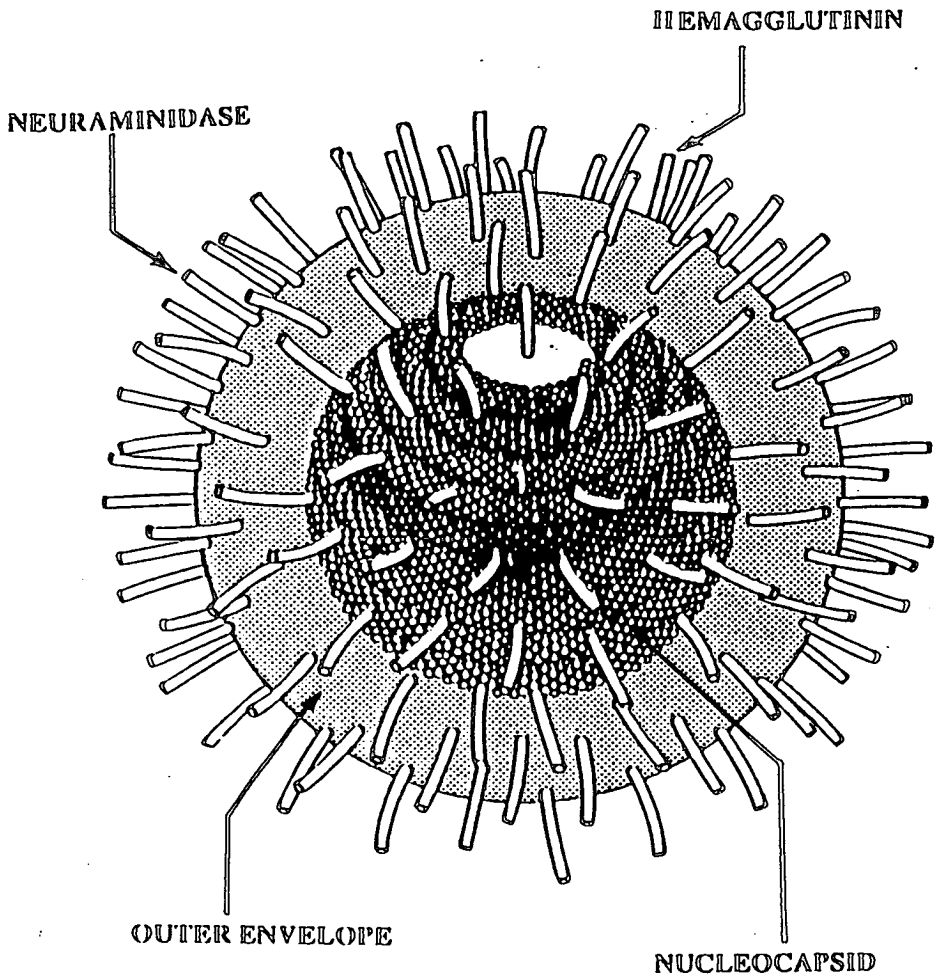


Fig. 3.1 Influenza virus particle.

3 Signal Sequence Structure

3.1 Introduction

The influenza virus is a member of the family of negative strand RNA viruses known as myxo viruses (Gk. myxa meaning mucus). These are again subdivided into two groups : the orthomyxoviruses, or influenza viruses, and the paramyxoviruses or parainfluenza viruses which include mumps, measles, rubella and Newcastle disease viruses.

A number of detailed monographs on the virology of the influenza viruses are available. The most informative of these are by Fields[134] and Krug.[135]

The influenza virus particle, or virion, has four main structural units, fig.3.1. These are:-

1. The nucleocapsid, the core particle containing the genetic information, RNA polymerases and core structural proteins.

2. The outer envelope, a lipoprotein membrane which is partially derived from the outer envelope of the host cell.

3. Hemagglutinin (HA), the major surface glycoprotein which is involved in the initial stages of virus-host cell interaction.

4. Neuraminidase (NA), the surface glycoprotein responsible for the removal of sialic acid groups from the host cell surface.

Influenza viruses are divided into three types A, B and C based on their antigenic determinants. Influenza A is the principal cause of epidemic influenza while influenza B is usually associated with a milder disease, but can cause winter epidemics. Influenza C is not thought to be a human pathogen.

The fact that the influenza virus continues to be an important human pathogen, frequently causing widespread disease and significant

loss of life is due to the extensive variation of the hemagglutinin and neuraminidase glycoproteins. Two types of variation have been noted. Antigenic drifts which result in minor antigenic changes causing isolated outbreaks known as epidemics, and antigenic shifts which result in major antigenic changes causing worldwide outbreaks known as pandemics. Antigenic drifts usually occur every one or two years while antigenic shifts occur at less frequent intervals. Both glycoproteins vary independently to produce variant viruses, but HA variation is considered more important because it is quantitatively the major surface glycoprotein and the antigen against which neutralising antibodies are directed.

Pandemics this century have been associated with variations of the HA with or without variation of the NA. The 1918-1919 pandemic was particularly severe killing at least 20 million people world wide and was associated with a swinelike influenza virus termed H1N1, where H1 and N1 designate the antigenic description of the HA and NA glycoproteins.

Since 1933, when the influenza virus was first isolated from humans two major antigenic shifts have occured; in 1957 the H2N2 subtype, Asian influenza, replaced the H1N1 subtype and in 1968 the Hong Kong, H3N2, virus appeared. The reappearance of the H1N1, subtype in 1977 caused infection only in people under the age of twenty one indicating that older people had antibodies to the virus due to exposure before 1957.

Since 1977 both H3N2 and H1N1 strains have co-circulated throughout the world with the ratio of H3 to H1 isolates changing dramatically from year to year and an underlying cyclical turnover of infection moving between the two strains[136].

3.2 Influenza Virus Proteins

The influenza virus comprises eight negative strand RNA segments, ten proteins and a plasma membrane. The plasma membrane is derived from the host cell while the genetic and protein components are synthesised in the infected host cell.

The eight RNA segments code for the ten polypeptides. Three involved in transcription of the influenza virus, two surface proteins, two matrix proteins, two non-structural proteins and a structural protein associated with the nucleocapsid. This is summarised in table 3.1.

RNA Segment	Encoded Polypeptide	Function
1	PB2	Component of transcription complex.
2	PA	Component of transcription complex.
3	HA	Major Surface Glycoprotein.
4	NP	Structural protein of the nucleocapsid.
5	NA	Surface glycoprotein.
6	M1 M2	Major component of virion underlying the lipid bilayer. Integral membrane protein.
7	NS1	Non-structural protein, function unknown.
8	NS2	Non-structural protein, function unknown.

Table 3.1

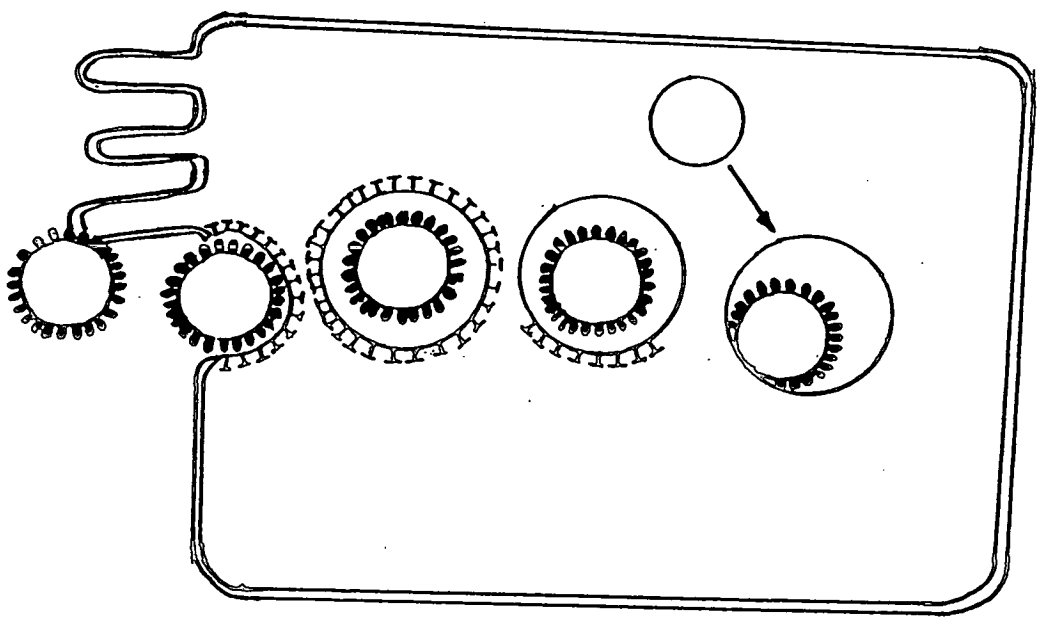


Fig. 3.2. Entry of the virus into the cell.

3.3 Replication

3.3.1 Attachment and Entry

Since viruses do not possess the required synthetic machinery for self-reproduction they must enter the host cell to replicate.

The attachment of the influenza virion to the host cell is mediated by the HA protein whose spikes interact with the sialic acid rich glycoprotein receptor of the host cell. Following binding the virion is transported into the cell *via* the endocytotic pathway. The virus particle becomes surrounded by a coated pit, a specialised region of the cell membrane where the cytoplasmic side of the membrane is coated with the protein clathrin. The coated pit then folds inwards and pinches off into the cytoplasm to form a "coated vesicle", the outer surface of which is now surrounded by clathrin. Inside the cell the vesicle loses its clathrin coat and fuses with an endosome. The endosome-vesicle complex then merges with a lysosome. The interior of the lysosome is more acidic than other regions of the cell and this acidity is thought to cause the outer viral membrane to fuse with the membrane of the lysosome. The nucleocapsid is thus freed into the cytoplasm initiating infection[135,137], figure 3.2.

3.3.2 Translation and Transcription

Once within the cell the viral genome is transported to the nucleus where replication takes place. The mechanism by which this transport takes place is not known. The RNA of the virus itself is non-infectious. The mRNA's of the virion are transcribed by the virion associated RNA transcriptase.

The transcription of the viral genome is the first event after entry

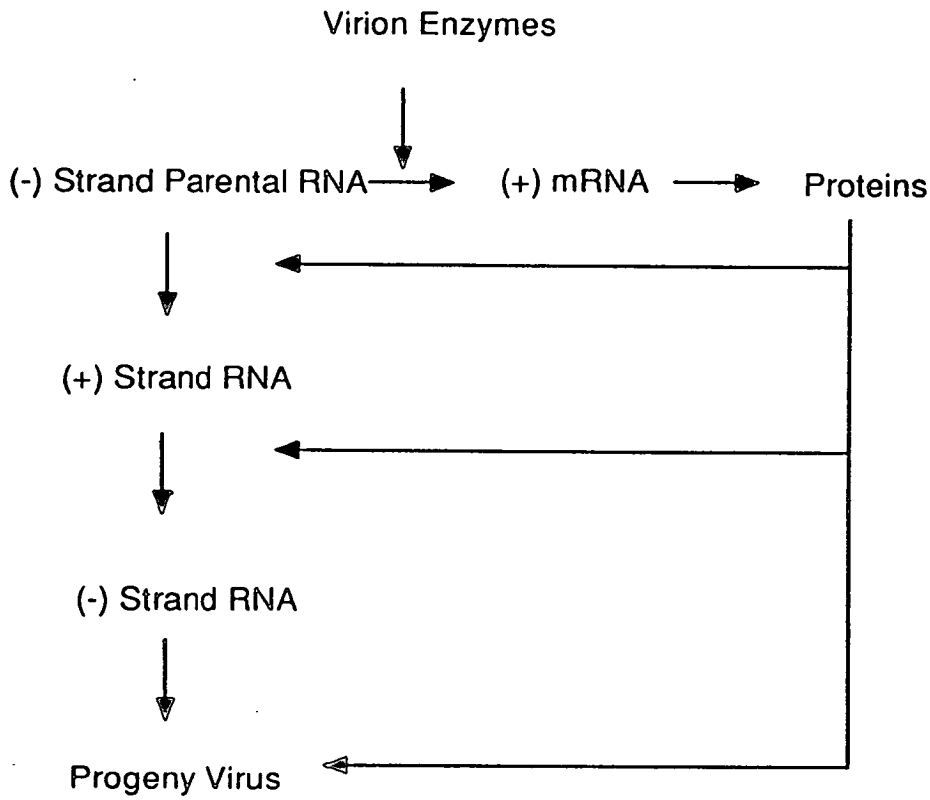


Fig. 3.3 Protein / RNA cycle.

to the cell. Replication of the viral genome then proceeds under the direction of the newly synthesised viral proteins. A full positive strand is made which serves as a template for the synthesis of negative strand genomic RNA's. Fig. 3.3 shows the translation-transcription cycle.

3.3.3 Virion Assembly and Release

The two modules that comprise the influenza virion are assembled in different locations : the nucleocapsid in the nucleus and the envelope at the cell surface membrane.

Assembly of the nucleocapsid is the least well understood of the two assembly processes, however it must involve the transport of proteins from the cytoplasm to the nucleus where the ribonucleoprotein assembly takes place.

The assembly of the viral envelope is better understood. In common with the other negative strand RNA viruses the last step in assembly of the influenza virus is associated with its egress from the cell[138].

The HA and NA glycoproteins become inserted into both the inner and outer surface of the plasma membrane. The NA glycoprotein is anchored to the membrane by an uncleaved signal-sequence near its N-terminus forming tetrameric spikes at the cell surface. The HA glycoprotein is synthesised with a cleavable N-terminal signal peptide and is anchored in the membrane by a hydrophobic domain near its C-terminus forming spike shaped trimers at the cell surface.

The M2 protein is also an integral membrane protein that is expressed at the cell surface of infected cells with an extracellular domain of 18-28 amino acids. This along with the HA and NA glycoproteins constitutes the polypeptide fraction of the virion lipoprotein membrane.

The M1 protein is thought to underly the lipid bilayer and appears to serve the important role of bridging the gap between the nucleocapsid and the carboxy-terminal cytoplasmic domains of the glycoproteins which are already anchored into the nascent virus envelope patches. The M1 protein first associates with the membrane bound proteins and budding begins when the nucleocapsid diffuses to the membrane and begins to associate with the M1 protein. As more M1 proteins are bound the membrane of the cell is forced to curve around the nucleocapsid which finaly breaks off releasing the virus which is now free to renew the process of infection. An overview of the replication cycle is shown in fig. 3.4.

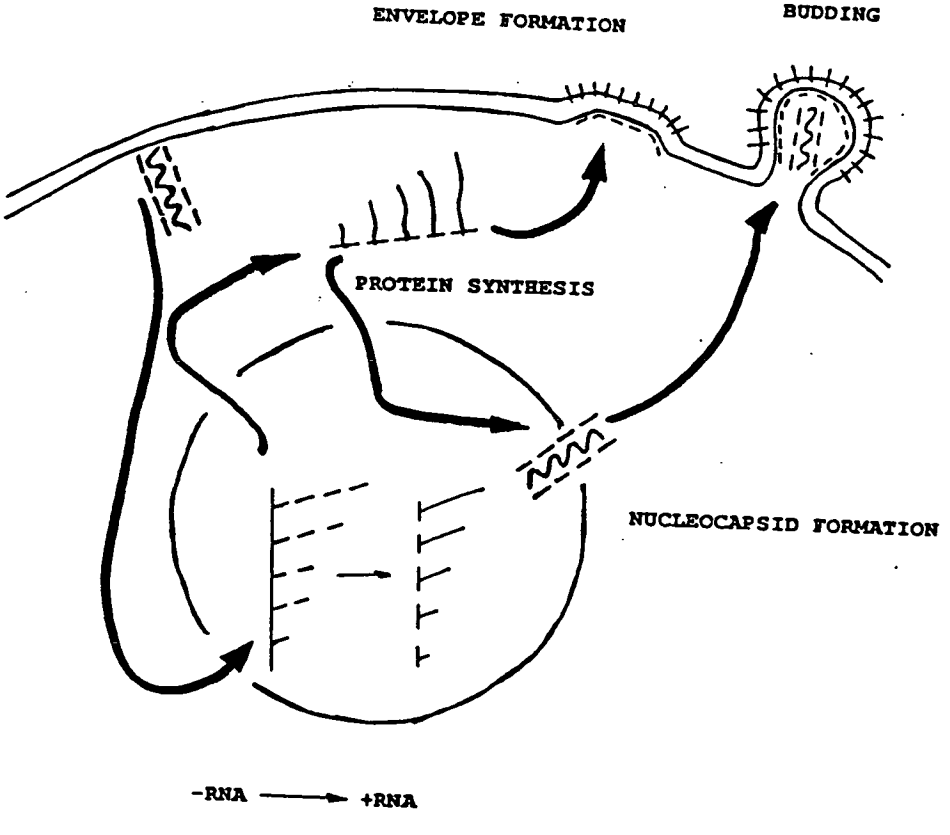


Fig. 3.4 Overview of the replication process.

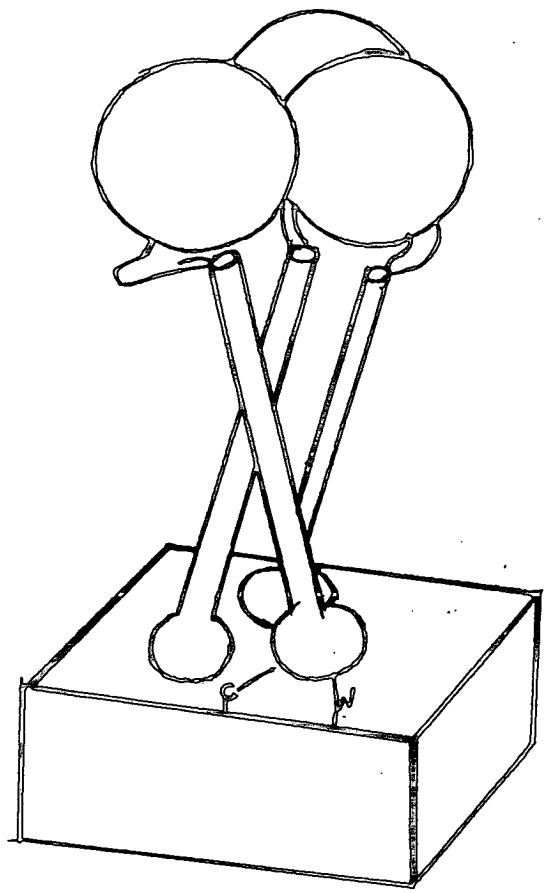


Fig. 3.5 The HA trimer.

3.4 Hemagglutinin

The hemagglutinin glycoprotein of the influenza virus is the antigen against which neutralising antibodies are directed and is probably the most widely studied components of the influenza virion.

The three dimensional structure of the HA molecule has been solved using X-ray crystallography[139] and sequencing has provided the complete primary structure of a number of HA proteins allowing the functional domains and antigenic regions to be interpreted[140]. The primary structure of the HA proteins reveal some important features.

The amino terminus commences with a signal sequence of 15-16 amino acids. The next region, approximately 550 amino acids long comprise, the HA1 subunit (approximately 330 residues) and the HA2 subunit (approximately 220 residues). The HA1 and HA2 subunits are linked by a single arginine residue which is cleaved after the formation of a stable tertiary structure. The two units remain closely associated after cleavage, held together not only by non-covalent bonds but also by a disulphide bridge linking the N-terminus of the HA1 with the HA2 subunits. The HA molecule is anchored to the cell membrane at the carboxy terminus of the HA2 subunit by a stretch of hydrophobic residues which span the lipid bilayer. The remainder of the carboxy terminus, a region of approximately ten residues, protrudes into the cytoplasm. The polypeptide is heavily glycosylated at specific asparagine residues.

The HA proteins form a glycoprotein spike on the cell surface, each spike being a trimer of HA molecules, figure 3.5. Each HA monomer forms an elongated structure consisting of a stalk capped with a large globular region at one end and a smaller globular region at the membrane end. The HA1 chain comprises the small globular region at the stalk and the entire

globular region while the HA2 chain is confined to the stalk and the smaller globular region at the membrane surface.

The attachment sites for the cellular receptors are located on the top surfaces of the the large globular regions where the folding of the HA1 chain creates small pockets to accomodate the sialic acid of the glycoprotein receptor of the host cell. The large globule also seems to contain all the sites involved in virus neutralisation.

3.5 Protein Targeting

The targeting of proteins from their point of synthesis in the cytoplasm to their required destination within the cell has attracted much interest over the past few years[141].

Eukaryotic cells, in contrast with prokaryotic cells, are subdivided into functionally distinct compartments, known as organelles, each of which is surrounded by one or more membranes. Each organelle contains its unique compliment of proteins enabling it to carry out specific functions within the cell. These proteins, generally synthesised in the cytoplasm, must be transported from this matrix into or through various membrane structures to reach their required destination. This raises one of the fundamental questions in cell biology, namely : how are proteins made in one subcellular compartment accurately and efficiently transported to their final destination in another?

Eukaryotic DNA is located mainly in the nucleus, although a small fraction is found in the mitochondria. The nuclear DNA encodes the majority of proteins within the cell.

Protein synthesis in the nucleus begins with the transcription of nuclear DNA into messenger RNA (mRNA) which, after splicing, is

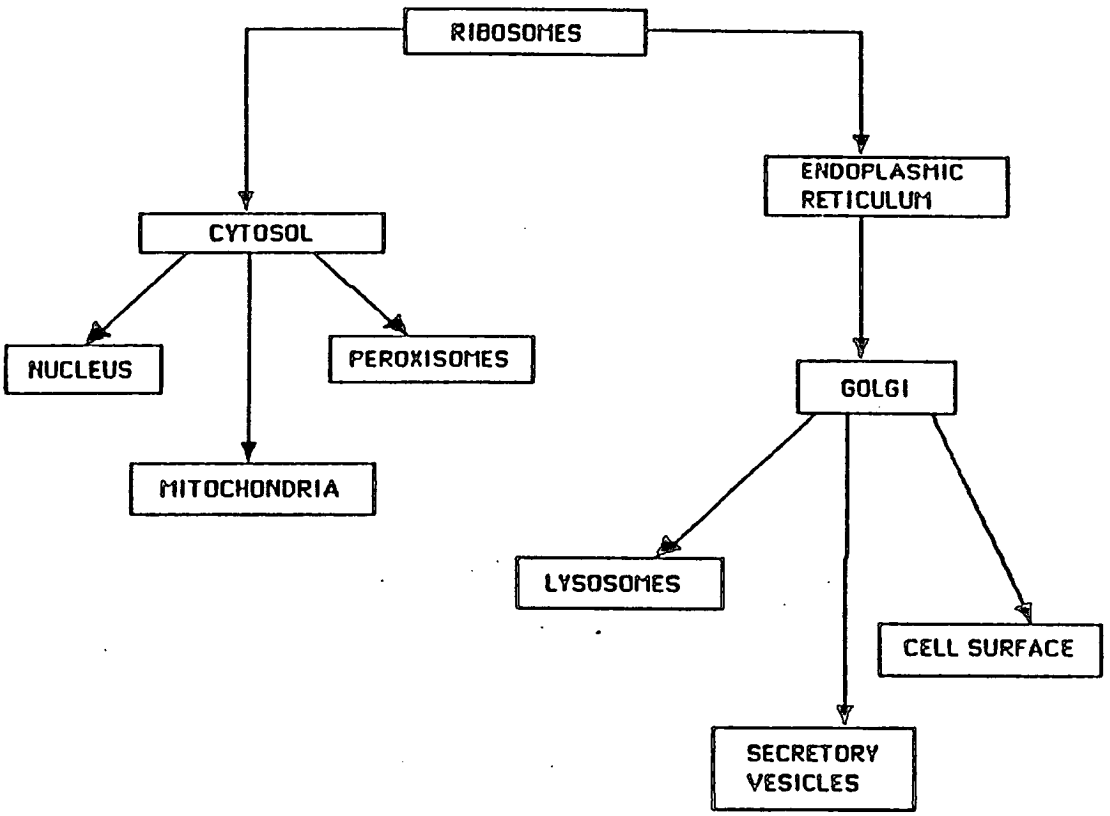


Fig. 3.6 Protein trafficing pathways.

transferred to the cytoplasm where it is translated into protein on ribosomes. Hereafter the proteins must be directed to their appropriate intracellular locations.

After synthesis on cytoplasmic ribosomes the fate of different proteins diverge : proteins bound for the endoplasmic reticulum (ER), Golgi apparatus, lysosome and extracellular locations enter the secretory or Palade pathway whereas those targetted to other intracellular organelles such as the nuclei, mitochondria or peroxisome are processed in a different manner.

Targetted proteins not destined for the Palade pathway are transported directly from the cytoplasm by mechanisms that do not necessarily appear to be tied to translation. Delivery of proteins to the remaining membranes and organelles of the cell is initiated by transfer of the nascent polypeptide across the membrane of the endoplasmic reticulum by a process tightly coupled to translation; chain elongation and translocation across the membrane appearing to occur simultaneously. From the ER, proteins can be transported by the exocytotic or Palade pathway to various intracellular organelles or to the cell surface or exterior either continuously, "constitutive secretion" or as a result of stimulation by some extracellular agent, "regulated secretion". Thus in eukaryotic cells there are variety of different pathways for protein trafficing, fig. 3.6. There must therefore be some mechanism by which proteins destined for different organelles are specifically recognised and accurately transported.

Some insight into the mechanisms by which such processes might operate was gained when in 1972 it was discovered that the immunoglobulin light chain protein was synthesised as a precursor that

was slightly larger than the secreted protein[142].

Subsequently studies of this and other secretory proteins have shown that the extra material contained within the precursor proteins of targetted polypeptides consists of a polypeptide extension to the N-terminal of the protein[143,144]. It has now been shown that the N-terminal presequences contain the information necessary for the targeting of the proteins into the secretory pathway and the term "signal peptide" has been applied to the precursor extensions.

Presequence peptides are also necessary for the targeting of endoplasmic reticular, nuclear, proxisomal, mitochondrial and chloroplast proteins of eukaryotic cells, and are also required for the secretion of secretory proteins in prokaryotic organisms. In prokaryotes the signal sequences are required to direct secretory and integral membrane proteins to and across the bacterial cytoplasmic membrane.

There has been considerable interest in these signal peptides since their discovery, and while a considerable amount of knowledge has been gained about the characteristics of these sequences there is still little evidence as to the actual mechanism by which they mediate the translocation of large water soluble proteins across the hydrophobic environment of the phospholipid bilayer core.

3.6 Signal Sequences

Viral glycoproteins are transported from their site of synthesis in the cytoplasm to the outer-membrane of the host cell utilizing the same machinery used for transport of host cell integral membrane proteins. It is now clear that a common pathway exists for the transport of proteins to the cell surface. This pathway is organised spatially and temporally as a

sequence of membrane bound compartments each containing its own compliment of processing enzymes. Movement through the exocytic pathway is vectorial, there is no evidence for a backflow of proteins. Proteins enter the pathway at the ER membrane and travel through the Golgi to reach the cell surface or secretory granules, lysosomes or vacuoles.

The signal sequences of bacterial proteins targetted towards the bacterial secretory pathway have been found to be similar to those of eukaryotes and while comparison of ER and bacterial signal sequences show a low degree of sequence homology it also shows a number of common characteristics[145-151]. Analysis has revealed that signal sequences vary in length from 15-35 residues and are exclusively N-terminal although not always cleavable. Three structurally, and possibly functionally, distinct regions have been identified as the building blocks of a secretory presequence, fig. 3.7.

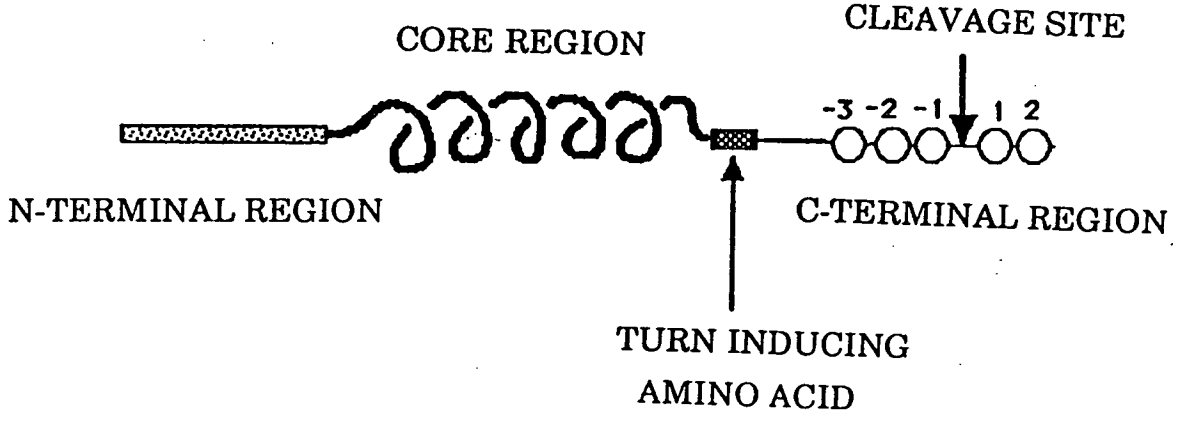


Fig. 3.7 Structurally distinct regions of secretory signal sequences

The N-terminal region can vary in length from 2-10 amino acids and usually contains at least one Lys or Arg residue and is positively charged. The second region, the core region, is essential in both pro- and eukaryotes. It consists of a stretch of nine or more residues containing hydrophobic or neutral residues and is predicted to form an α-helix. This region is usually terminated by a turn-inducing amino acid, proline or glycine. The final region, the C-terminal region precedes the signal peptidase cleavage site and consists of five to seven residues of higher polarity than those in the core region. Small apolar residues generally occupy the positions -1 and -3 (-1 being the position on the N-terminal side of the cleavage site) with position -2 usually occupied by an aromatic amino acid.

3.7 The Palade Pathway

The steps by which the nascent polypeptide is directed to and across the membrane of the ER occur rapidly and have therefore been difficult to analyse. Most of the information about this process has been derived from experiments performed *in vitro* in which translocation is at least an order of magnitude slower than *in vivo*. The precise relationship of these results to the situation *in vivo* is however unclear. However for proteins such as the HA protein which contain an N-terminal signal sequence the general sequence of events in the exocytic pathway appears to be as follows.

As the nascent polypeptide chain emerges from the ribosome the hydrophobic signal sequence is recognised by a ribonucleoprotein complex called the signal recognition particle (SRP)[152,153]. The 54 kDa component of the SRP is observed to bind to the hydrophobic portion of the signal

sequence[154,155], but only in the presence of ribosomes[153]. Binding of the SRP to the signal sequence causes a pause in translation[156,157,158]. Evidence suggests that there is only one type of SRP in a particular organism, and it follows that it must recognise different signal sequences.

The signal recognition particle-ribosome complex then interacts with the SRP receptor (SRPR) or docking protein[156,157,158], a heterodimeric complex consisting of α and β subunits. The β subunit is a peripheral membrane protein which can be anchored to the face of the ER *via* the β-subunit[162,163,164]. Sequencing of the α-subunit[162] has shown that it contains sequences similar to those in nucleic acid binding proteins and it is thought that it may bind to the 7S RNA component of the SRP. This provides a functional ribosome-membrane junction permitting translocation of the nascent peptide chain across the ER membrane. The SRP and SRPR are not constituents of the ribosome membrane junction and are likely to be recycled once they have accomplished the initial targetting[158]. Attachment of the ribosome complex can however occur in the absence of the SRP and may involve binding to a signal sequence receptor on the ER membrane[165].

The mammalian signal peptidase, responsible for cleavage of the signal sequence after translocation, has recently been purified as a complex of six polypeptides[166], two of which are glycoproteins. It is postulated that only one subunit of the isolated complex carries out signal peptide cleavage, the other subunits performing some other related functions. The amount of this protein complex present in microsomes roughly equals the amount of bound ribosomes and it can be conjectured that the signal peptidase complex may also function in ribosome binding

or in some other aspect of translocation such as forming a pore through the membrane.

It has been suggested that the movement of largely hydrophobic proteins through the lipid bilayer requires a protein pore[165]. Alternatively, it has been argued on thermodynamic grounds that a loop of polypeptide may be able to pass through the membrane without the aid of accessory proteins[167]. One possibility is that the channel through the membrane could be formed by lipids in a non-bilayer formation with their hydrophilic headgroups lining the walls. The possibility that the signal peptides themselves could form the pore has been discounted[168].

An alternative hypothesis is that a channel could be formed by a membrane protein, such as those co-purified with the mammalian signal peptidase[166], which would form the channel and fold around the secretory protein as it penetrates the membrane. While none of these putative "translocases" have been firmly identified a recent experiment carried out by Blobel *et al* using electrophysiological techniques has demonstrated the existence of a protein-conducting channel in the endoplasmic reticulum[169]. Initial results indicate that the channel closes upon dissociation of the robosome from the membrane. This suggests that the ribosome may directly interact with part of the protein conducting channel. Further purification of the components linked with the import machinery may identify the channel forming proteins.

3.8 Modification and Post-translational Processing of The HA

Oligosaccharides are added to the HA polypeptide rapidly after specific asparagine residues are exposed in the lumen of the ER[170]. Glycosylation occurs by the transfer of a branched oligosacharide of the

structure $GlcNAc_2Man_9Glc_3$ from the dolichol-lipid carrier to an asparagine of the tripeptide motif Asn.X.(Thr or Ser).

In the ER trimming of the terminal glucose residues from the mannose rich oligosaccharide substituents occurs, followed by removal of a single mannose unit. This may occur either co-translationaly or post-translationaly. Within the ER, proteins destined for transport represent only a fraction of a percent of the total protein content. The resident proteins of the ER do not cycle between the ER and the Golgi apparatus indicating the presence of a gate that functions to permit the transport of transient proteins to the Golgi while preventing the exit of the vastly more numerous ER resident proteins. The precise location of this gating function is not known although it is thought to occur before or during the sorting of the transient proteins into the transport vesicles that are responsible for moving exocytic proteins to the Golgi complex.

Trimerisation of the HA protein begins in the ER and is complete by the time the protein reaches the medial Golgi compartment. However it is not clear if trimerisation and folding of the HA is completed within the ER. The HA is then transported to the Golgi apparatus in small vesicles which move between the transitional elements of the two organelles. The protein travels through the Golgi complex where modifications to the oligosaccharide side chains take place. Upon exit from the Golgi the HA protein enters the trans-reticular network (TGN), a post Golgi compartment consisting of a tubular network of vesicles that appear to be continuous with the trans Golgi cisternae. The TGN initiates export of proteins to the cell surface in secretory granules.

3.9 Sorting of the HA to the Apical Domain of Polarised Epithelial Cells

In epithelial cells, influenza glycoproteins are sorted within the cell and are displayed only at a particular domain of the plasma membrane. Epithelia are composed of polarized cells that maintain at least two distinct specialisations of their plasma membranes, the apical and basolateral domains[171].

The glycoproteins of the influenza virus are recognised by epithelial cells as belonging to the apical class of surface molecules[172,173]. It is not clear what role, if any, these cells have in the infectious cycle of viruses, however it is worth noting that polarized epithelial cells of the upper respiratory tract are the primary site of influenza virus infection and release of viral proteins towards the apical or luminal surface favours the development of localized rather than systemic infection.

3.10 Influenza Virus HA Signal Sequences

As discussed above the signal sequence appears to contain all the information necessary for the targeting of proteins to the appropriate membrane.

Numerous mutations of HA signal sequences have been constructed in order to elucidate their functional requirements[170]. Expression of HA cDNAs from which sequences encoding the signal peptide have been deleted results in the production of HA molecules that remain in the cytoplasm, demonstrating that the HA signal sequence is necessary for translocation. This is confirmed in experiments where sequences encoding the signal peptide region of the HA protein were fused to sequences encoding for proteins that normally do not cross the ER membrane. When

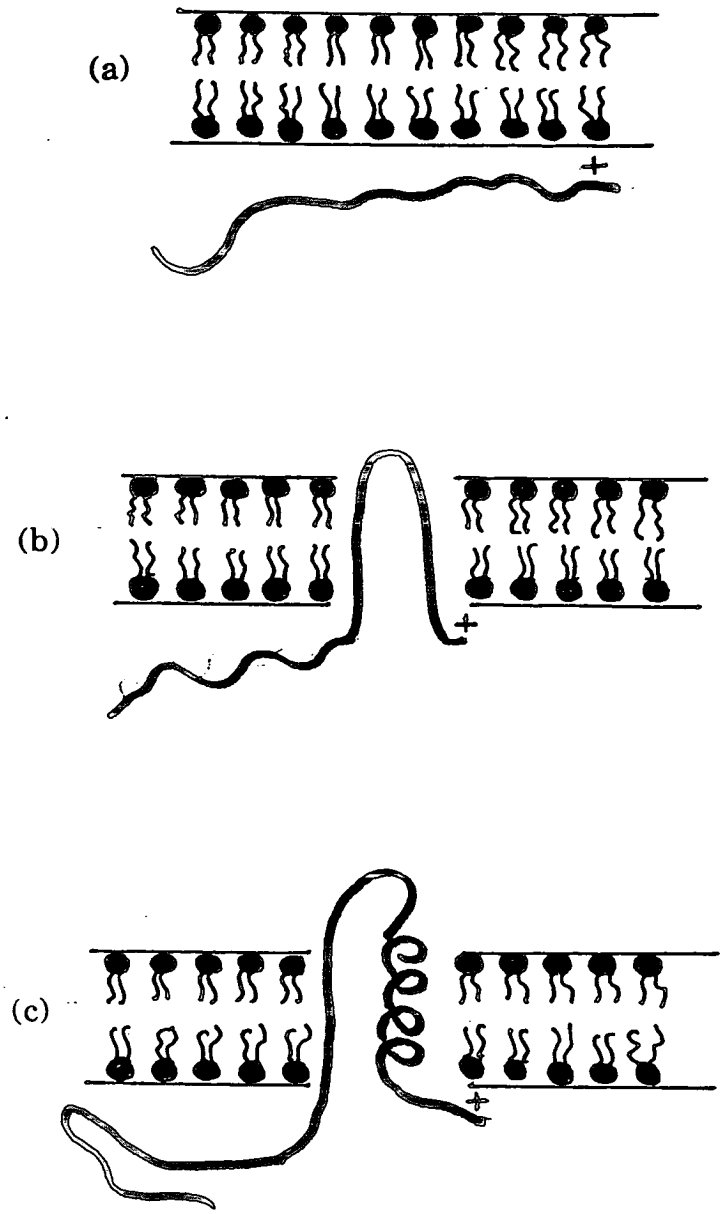


Fig. 3.8 Model for the insertion of ER signal sequences into membranes. (a) Association of the positively charged N-terminus with the lipid headgroups. (b) Insertion of the signal peptide into the membrane in a β-sheet like conformation. (c) The hydrophobic core region undergoes a conformational change to an α-helix, in the process forming a channel through which the translocation of the protein can occur.

preceded by an HA siganl sequence both SV40 large TAg and a modified aminoglycoside 3' phosphotransferase II, proteins normally located in the nucleus and cytoplasm respectively were translocated into the ER and glycosylated. It has also been shown that the HA signal can be replaced by a heterologous signal sequence corresponding to some other exported protein without affecting transport, a finding consistent with the earlier observation that there is little of no amino acid conservation between signal peptides of HA's from different virus strains.

Although the presence of the three structurally distinct regions within the signal sequences are necessary the precise structure-function relationship of specific residues are poorly understood.

Various mechanisms for the interactions of the different types of targeting sequences with membranes have been formulated and these fall broadly into two groups[174]. One view postulates the existence of specific protein assemblies within the lipid bilayer, which are responsible for facilitating the motion of proteins across the membrane during protein synthesis. The second view assumes protein insertion to be a spontaneous process that does not require specific apparatus. However the concept that a crucial role is played by the structure of the secreted protein is central to both these models.

While the structure formed by the signal peptide in a water filled protein channel could be influenced by signal peptide protein interactions, the structure of a signal peptide inserted into the lipid bilayer is expected to undergo folding in discrete stages.

A model involving the insertion of signal peptides into the lipid bilayer has been suggested[174], fig. 3.8. The positively charged N-terminus of the signal peptide interacts with the lipid headgroups. This is followed

by the insertion of the signal peptide into the membrane in a β-sheet like conformation. The hydrophobic core region of the signal peptide then undergoes a conformational change to an α-helix. As a result of the conformational change, a segment of the mature protein enters the membrane allowing translocation through the membrane to proceed.

Some clarification of the mechanism of import to the ER could be achieved if some method of determining the structure of the signal sequence during the transport process were available. While such an undertaking is not feasible using techniques currently available an approach which is experimentally feasible is the determination of the structure of the signal sequence in the type of environment encountered by the peptide during import i.e. the hydrophobic lipid bilayer. The presence or absence of defined secondary structural features under these conditions might lend credence to the validity of the various mechanisms.

3.11 Circular Dichroism

Circular dichroism (CD) measures the optical activity of asymmetric molecules in solution based on the difference between the left and right handed components of circularly polarised light.

Optical activity in peptides stems from two different forms of chirality. Asymmetry in the individual amino acids is the lesser contributor while architectural dissymmetry, for example the handedness of helices, has a major effect on the optical rotatory power of a peptide molecule. Since certain structural elements show strong characteristic spectra it is possible to draw conclusions from the optical rotation on conformation.

The far UV (amide) region of the CD spectra (170-250nm) is dominated by contributions from the peptide bonds whereas the near UV region contains contributions from aromatic regions and disulphide bridges. Structural information is therefore gained from the amide region of the spectra. The helix is the structural element showing the strongest CD spectrum. Such spectra are highly characteristic, showing a double minimum with the minima located at 200-206 and 222nm. The CD spectrum of a β-sheet is relatively weak being dependant on the length and spatial arrangement of the strands and is characterised by a single minimum. Estimates of β-sheet content in peptides and proteins from CD are considerably less reliable than those for α-helices.

3.12 Discussion

To examine the aspects of secondary structure that might be crucial for the function of HA signal peptides it was decided to synthesise the 17 residue peptide of the hemagglutinin protein from the influenza virus A/WSN/33[175] using solid phase peptide synthesis (SPPS). The primary structure of the signal peptide is shown in fig. 3.9.

H.Met[1].Lys.Ala.Lys.Leu[5].Leu.Val.Leu.Leu.
Tyr[10].Ala.Phe.Val.Ala.Thr[15].Asp.Ala.OH

Fig. 3.9 Presequence (1-17) of Hemagglutinin A/WSN/33.

3.12.1 Peptide Synthesis

The methodology of solid phase peptide synthesis (SPPS) was first introduced by Merrifield in 1962[176,177] and involves the covalent attachment of the C-terminal amino acid in a sequence to an insoluble

support. Chain elongation then proceeds using methods similar to those developed for solution phase peptide synthesis. The major advantage of this approach is that separation of the reagents can be achieved by filtration allowing the resin-bound peptide to be thoroughly washed after each step to ensure complete removal of any soluble contaminants[178].

Recently methodology has been developed for the protection of the α-amino function with base labile protecting groups allowing the utilization of side chain protecting groups and resin linkers which are cleavable under milder conditions than the very strong acid used with earlier protecting group stratgies. In these syntheses an orthogonal approach is used with transient protection of the α-amino function normally achieved using the base labile 9-fluorenylmethoxycarbonyl (Fmoc) group developed by Carpino and Han[179]. This protects the α-amino function as a carbamic acid ester, which is easily cleaved to give the amine and dibenzofulvene via a β-elimination mechanism by reaction with a base. If a secondary amine such as morpholine or piperidine is used the amine subsequently adds to the dibenzofulvene to yield a stable tertiary amine. Fig. 3.10 shows the deprotection using piperidine as the base.

The insoluble resin support used with the Fmoc protection strategy is the p-alkoxybenzylalcohol resin developed by Wang[180] from which the peptide can be cleaved using 50% trifluoroacetic acid (TFA).

3.12.2 Activation of Amino Acids

In order to convert carboxylic acids into acylating agents their hydroxyl group must be replaced by an electron-withdrawing group, thus augmenting the nucleophilic attack of the amino group. The standard

acylating agents are the symmetrical anhydride or an active ester of the amino acid.

The symmetrical anhydrides are generated *in situ* by the reaction of the amino acid with diisopropylcarbodiimide (DIC). The amino acid reacts with the DIC to form an O-acylisourea which then reacts with the carboxy component of excess unreacted amino acid to produce the anhydride.

Although the symmetrical anhydride is formed rapidly and is very reactive, there are several side reactions associated with carbodiimide activation. In addition the symmetrical anhydride approach does not make efficient use of sometimes costly protected amino acids since two equivalents of the amino acid are required to generate one equivalent of the acylating agent.

A second activation method, using active esters of amino acids, was first introduced to circumvent the side reactions experienced in the symmetrical anhydride procedure. Addition of the powerful auxiliary nucleophile, 1-hydroxybenzotriazole (HOBt), to the O-acylisourea complex results in the direct formation of the HOBt ester. The excellent acylating properties of the HOBt active esters generated in this way led to the adoption of this strategy into SPPS as an activating agent in its own right. This method of activation has the added advantage of only requiring one equivalent of the amino acid to produce the acylating agent.

The loading of the C-terminal amino acid onto the resin support is most conveniently carried out by the reaction of the resin with an excess of the symmetrical anhydride of the required amino acid generated in the presence of DIC and the base N,N-dimethylaminopyridine (DMAP).

3.12.3 Monitoring the Progress of Peptide Synthesis

The number of chemical reactions required to synthesise even a small peptide necessitates that high yields are obtained in each step. The final purity of a synthetic peptide depends on the efficiency with which each residue is incorporated into the growing peptide chain. To be sure of obtaining high yields some form of continual monitoring must be carried out.

The use of the Fmoc protecting group allows the direct spectrometric monitoring of SPPS since the UV spectra of the dibenzofulvene and dibenzofulvene-piperidine adduct generated upon deprotection of the Fmoc group are characteristic of conjugated systems, absorbing at 300nm, thus differentiating the soluble product of the deprotection reaction from the non-aromatic reagents which do not absorb at this wavelength.

In the synthesis carried out here deprotection mixtures were pumped through a UV detector and the absorbance recorded at 300nm. Comparison of the areas under the peaks recorded for each coupling cycle allowed some quantitative measure of the efficiency of each step.

3.12.4 The Synthesis and Purification of the A/WSN/33 HA Signal Sequence

This signal sequence, fig. 3.9, consists mainly of hydrophobic residues containing aliphatic side chains and as a result simplifies the synthetic problem in that only a few of the residues require to be side chain protected. The use of the Fmoc group allowed the protection of the hydroxy functions of threonine and tyrosine as t-butyl (t-Bu) ethers. The

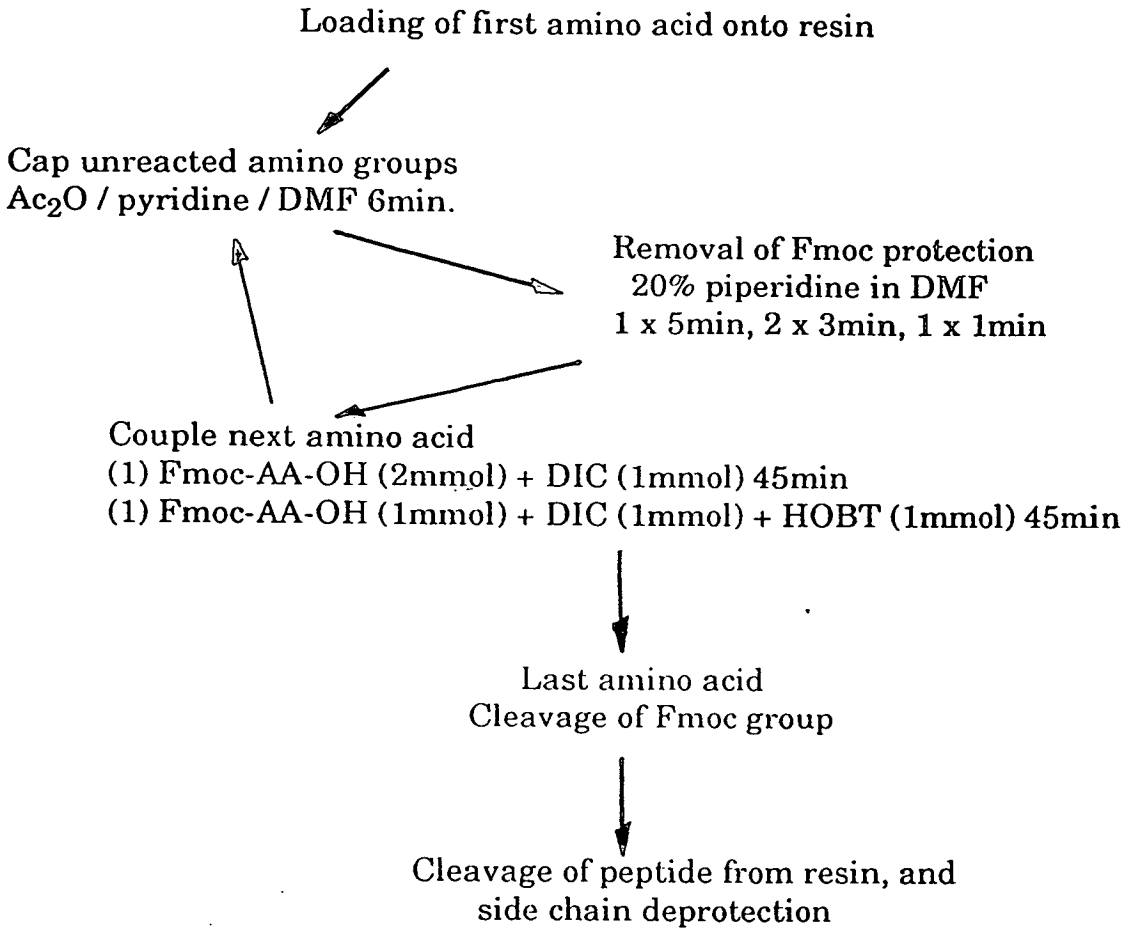


Fig. 3.11 Synthetic procedure

ω-amino group of lysine was protected with the Boc group and the aspartic acid was protected as its t-Bu ester[181].

Peptide synthesis was carried out on a 0.5mmol scale using the Applied Biosystems 430A peptide synthesiser.

The C-terminal alanine residue in the sequence was loaded onto the Wang resin as its symmetrical anhydride and the loading efficiency determined by removing and drying a sample of the resin. The dried sample was deprotected using 20% piperidine in DMF and the absorbance of the Fmoc piperidine adduct measured against a standard solution of 20% piperidine in DMF. This allowed quantification of the Fmoc-protected amino acid present on the resin.

Both the symmetrical anhydride and the HOBt active ester methods were used in tandem in a double couple cycle consisting of an initial coupling with the appropriate symmetrical anhydride followed by coupling with the HOBt active ester of the amino acid. The overall synthetic strategy is outlined in fig. 3.11.

Deprotection of the amino function was achieved by treatment of the resin bound peptide with a solution of 20% piperidine in DMF for 5 minutes followed by washing of the resin. The UV absorbance of the combined filtrates recorded at 314nm was recorded. Most of the deprotection was accomplished with the first cycle however three more short base treatments were carried out to ensure total deprotection. The resin was then thoroughly washed before the double coupling cycle of the next residue. Permanent blocking of any unreacted amine groups was achieved by the acetylation with acetic anhydride. Theoretically the final product should contain no deletion peptides, only truncated acylated peptide sequences.

In this synthesis a significant decrease in coupling efficiency was noticed on addition of the third amino acid, Thr^{15}, this was possibly due to base catalyzed ring closure to produce the 2,5-diketopiperazine[178], fig. 3.12. The coupling efficiencies for cycles 4 to 6 fluctuated by only a few percent. However in cycle 7, the addition of Tyr^{10}, the coupling efficiency appears to rise by 38%. A possibility is that deprotected Fmoc adducts from earlier deprotection reactions might have become non-covalently attached to the resin by π–π interactions and were then released from the resin by subsequent washing. The introduction of the tyrosine residue could result in displacement of these bound Fmoc groups through preferential non-covalent interaction of the tyrosine aromatic ring to the vinyl resin. The remaining coupling cycles were carried out without any major coupling failures.

Resin cleavage and side chain deprotection of the peptide was accomplished with 95% TFA using ethylmethylsulphide and ethane dithiol as scavengers. Concentration of the crude deprotection mixture and addition of ether gave a colourless precipitate which was collected by filtration and dried. The resulting solid was insoluble in acetic acid solutions and organic solvents such as methanol, methylene chloride, trifluoroethanol and acetone.

Since it is conceivable that the insolubility could be due to the presence of scavengers inhibiting penetration of solvent into the solid a second deprotected mixture was filtered into a solution of ethyl acetate : diethyl ether, 1:1. After a few minutes the peptide precipitated forming a cloudy suspension which became gelatinous on standing. The peptide was removed by centrifugation and the pellet was resuspended in the ethyl

acetate : ether solution and centrifuged. This process was repeated several times to ensure total removal of the scavengers.

In this case the peptide could be dissolved on addition of 1ml of water to a suspension of the pellet in glacial acetic acid. The volume of water was then increased to give a solution of the peptide in 30% acetic acid. Attempts to lower the acetic acid concentration by the addition of more water or methanol resulted in precipitation.

Analytical reverse phase high performance liquid chromatography (HPLC) of the crude peptide on C18-silica under standard conditions (eluting with a gradient of water(0.1% TFA) to acetonitrile(0.1%TFA)) showed no elution of the peptide from the column.

While the product gave a positive ninhydrin test both in solution and when applied to either polyacrylamide gels or paper it did not migrate either on electrofocusing polyacrylamide gels or on paper electrophoresis over a range of buffer pH's.

Paper electrophoresis of the hydrolysis product of the peptide indicated qualitatively the presence of all the constituent amino acids. The mass spectrum (positive ion FAB) of the material showed a parent ion at 1867.8, the expected mass being 1867.4. It was therefore clear that the desired product was present in the reaction mixture. It can be conjectured that the hydrophobic nature of the peptide results in the precipitation in buffered solutions.

Further attempts at reverse phase HPLC using conditions which have been used for other peptides, i.e. various concentrations of H_2O, acetonitrile with 0.1%, with various column supports were unsuccessful.

Size exclusion chromatography on Sephadex G25 eluting with 30% acetic acid proved to be a more satisfactory method of purification.

Although the peptide was eluted from the column the yields of recovered material were low.

Two major problems have been encountered in devising a reversed-phase separation of hydrophobic peptides. The first is their limited solubility in aqueous media. Secondly very hydrophobic peptides bind strongly to reversed phase matrices and require high concentrations of organic solvents for elution. Under these conditions, the peptides may not be soluble and can precipitate or may be irreversibly absorbed on the support.

The study of signal peptides and other hydrophobic peptide species have commonly been thwarted by problems associated with their purification. It would also appear that these problems are particularly pertinant to the purification of PRE A/WSN/33 HA. Several groups however have had success in purifying hydrophobic peptides using more powerful solvent systems than the standard acetonitrile-water systems. For example Halverson *et al* used the solvent system H_20;(95:5, acetonitrile : trifluoroethanol), 0.1% TFA to purify the nine residue peptide (60) on a C4 reverse phase column[182]. The addition of trifluoroethanol was claimed to be responsible for acceptable recoveries of the peptide. Formic acid in high concentration is also an extremely potent solvent for hydrophobic peptides and proteins. Hakeshoven and Dernick have previously used a solvent system of 60% formic acid with a gradient of acetonitrile to separate water insoluble polio virus structural proteins[183]. A different approach was taken by Shinnar and Kaiser who managed to purify a synthetic peptide corresponding to the leader peptide of the M13 coat protein with a gradient of 40-70% acetonitrile in 0.02M sodium phosphate buffer, pH2.7, containing 0.10M sodium perchlorate[184].

Attempts to purify PRE A/WSN/33 HA using the above protocols were uniformly unsuccessful , no peptide material being eluted from the columns. Attempts at using ion exchange chromatography were likewise frustrated by the insolubility of the peptide in aqueous solvents and its relative lack of polarity. Anion exchange chromatography eluting with 2% acetic acid gave a single peak shortly after injection, indicating that the peptide eluted in the void volume. Cation exchange chromatography gave similar results.

A useful method of purifying hydrophobic peptides has been reported by Katakai and Iizuka[185]. In this approach the peptide is repeatedly precipitated from 1,1,1,3,3,3 hexafluoroisopropanol (HFIP) with methanol. PRE A/WSN/33 HA was found to soluble in HFIP and while no precipitaion occurred upon addition of methanol, the addition of water to the HFIP solution initiated precipitation. The precipitate was collected and the process repeated several times. The dried product was found to be soluble in HFIP To date HFIP is the only solvent found which is capable of dissolving the peptide after drying.

Fig. 3.13 shows the purification procedure adopted which can be summarised as follows. Washing of the crude deprotection product with ethyl acetate-ether followed by size exclusion chromatography on Sephadex G-25 and repeated precipitation from HFIP with water. In common with Rosenblatt's description of using tlc to follow the purification of some hydrophobic peptides tlc of PRE A/WSN/33 HA was complicated by the properties of the peptide which failed to migrate from the origin and although providing no clear indication of the homogenaity of the peptide it did demonstrate the removal of migrating impurities[186]. Final purity of the peptide was confirmed by amino acid analysis.

3.13 Conformational Analysis

3.13.1 Secondary Structure Prediction

As the native conformation of peptides and proteins is coded in its amino acid sequence many efforts have been made to predict protein secondary structure from sequence data. The most widely used approach is that developed by Chou and Fasman[187] which is based on a statistical survey of a number of proteins of known secondary structure to assign helix and β-sheet potentials to each amino acid. This allows the average conformational potential of helix and β-sheet for a given sequence to be calculated. This method is based on the concept that a cluster of four residues with high helix forming potential in a group of six residues will initiate the formation of an α-helix. The helical segment will extend in both directions until a set of four residues with low helical potential, helix breakers, are encountered. Any segment with $\{P\alpha\}>1.03$ and $\{P\alpha\}>\{P\beta\}$ is predicted to be helical. $\{P\alpha\}$ and $\{P\beta\}$ being the average conformational potential for helix and β-sheet respectively. Similarly a cluster of four residues with high β-sheet forming potential, β-formers, in a stretch of five amino acids will initiate β-sheet formation. The β-sheet is then propagated in both directions as long as the sheet contains less than 30% of residues with low β-sheet forming potential. Any segment with $\{P\beta\}>1.05$ and $\{P\beta\}>\{P\alpha\}$ is predicted to be β-sheet.

Analysis of PRE A/WSN/33 HA by this method gives $\{P\alpha\}=1.18$ and $\{P\beta\}=1.10$, indicating a slight preference for helix formation. The first four residues have strong helical forming tendencies and no group of of helix breaking residues is encountered in the sequence. Thus the peptide is predicted to form an α-helix. The core region of the peptide, Leu^5 to Ala^{14},

Solution	Extremity (nm)	Molar Elipticity ($deg cm^2 mol^{-1}$)
100% HFIP	202	-9,500
10% HFIP-MeOH	215	-16,800
10% HFIP-BuOH	208 220	-22,900 -19,000
10% HFIP-TFE	207-208 215	-16,700 -15,400
50% HFIP-H_2O	214	-6,900
50% HFIP-0.6% SDS	207-208 220	-21,600 -16,900

Table 3.2

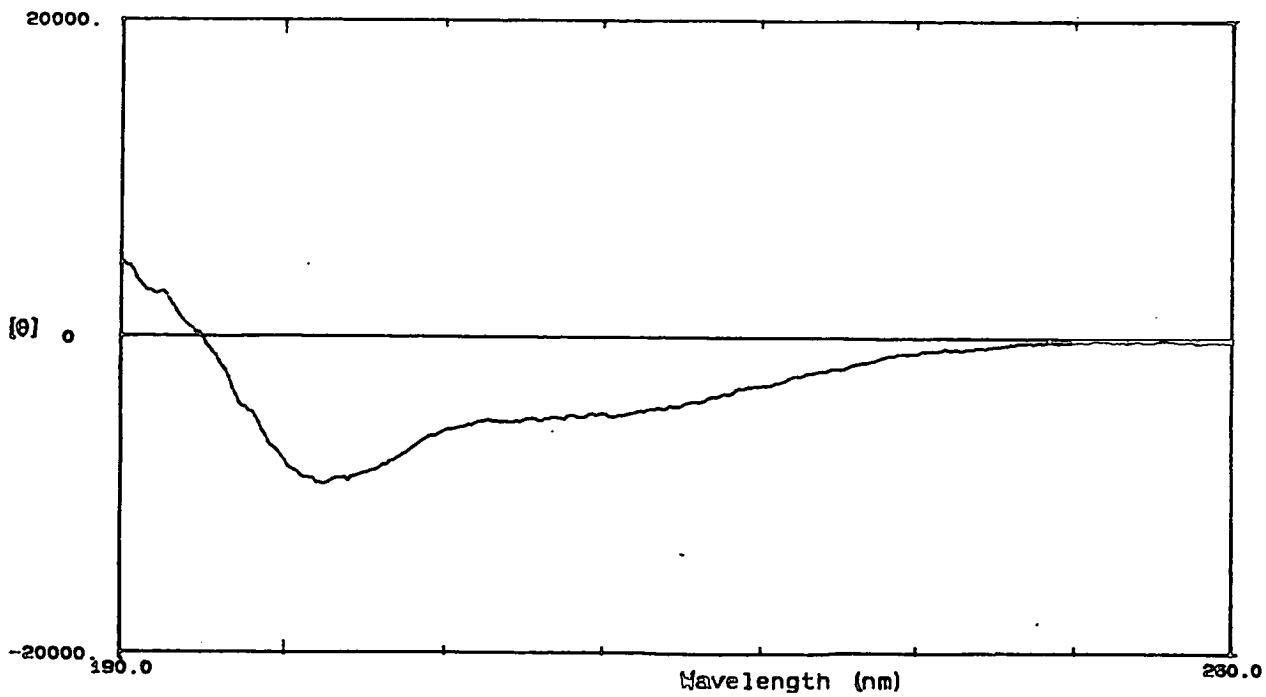


Fig. 3.14 CD spectrum of PRE A/WSN/33 HA in HFIP.

when considered separately is predicted to form a β-sheet, {Pα}=1.16 and {Pβ}=1.3.

3.13.2 Circular Dichroism

The methods available for the estimation of α-helix and β-sheet content of peptides and proteins from CD data generally fall into two categories. The first calculates the percentage helicity from the molar ellipticity of the peptide or protein at a specific wavelength[188]. This method however does not give any indication of the β-sheet content. The CONTIN program based on the method of Provencher and Glockner[189] falls into the second category and uses a linear combination of CD spectra of reference compounds with known secondary structure for the direct analysis of the CD spectrum of a peptide or protein by curve fitting. This gives both the percentages of α-helix and β-sheet and by default the proportion of random coil.

The conformation of the synthetic peptide was initially studied in aqueous and alcoholic solutions; water, methanol, n-butanol and trifluoroethanol which contained portions of HFIP. The aqueous solution was used to approximate the cytosolic medium and the alcoholic solutions to approximate the lipophilic intramembrane environment.

Fig. 3.14 shows the CD spectra of PRE A/WSN/33 HA in the very polar solvent HFIP. The peptide shows a spectrum with negative ellipticity, which is characteristic of a predominantely random structure, table 3.2. Calculations of secondary structure content, tabulated in tables 3.3 and 3.4, are concurrent with this view.

Solution	Otical Activity at 208nm	% Helix
100% HFIP	-6,700	9
10% HFIP-MeOH	-11,800	27
10% HFIP-BuOH	-22,900	65
10% HFIP-TFE	-16,700	44
50% HFIP-H_20	-6,000	7
50% HFIP-0.6% SDS	-21,600	61

Table 3.3 Percentage alpha helix as calculated by the GF method.

Solution	β-sheet	α-helix
100% HFIP	45%	9%
10% HFIP-MeOH	100%	0%
10% HFIP-BuOH	9%	91%
10% HFIP-TFE	54%	46%
50% HFIP-H_2O	75%	25%
50% HFIP-0.6% SDS	42%	58%

Table 3.4 Percentage alpha helix and beta sheet as calculated by the CONTIN program.

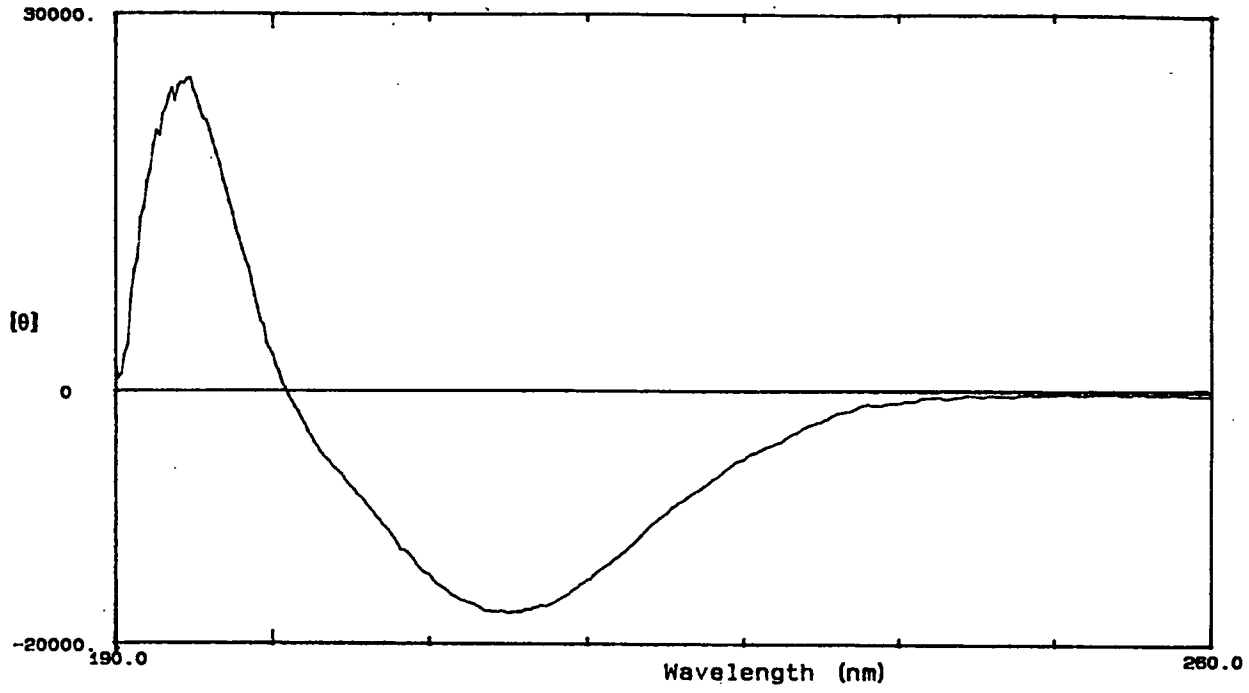


Fig. 3.15 CD spectrum of PRE A/WSN/33 HA in 10% HFIP-MeOH

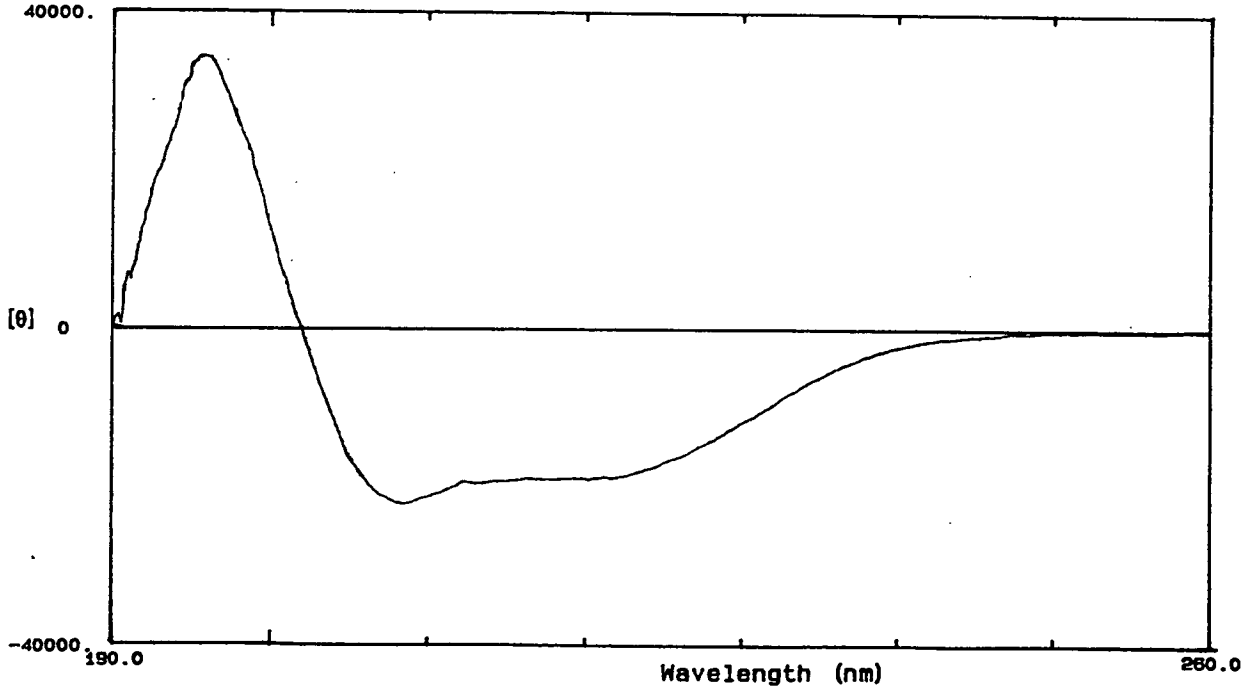


Fig. 3.16 CD spectrum of PRE A/WSN/33 HA in 10% HFIP-BuOH

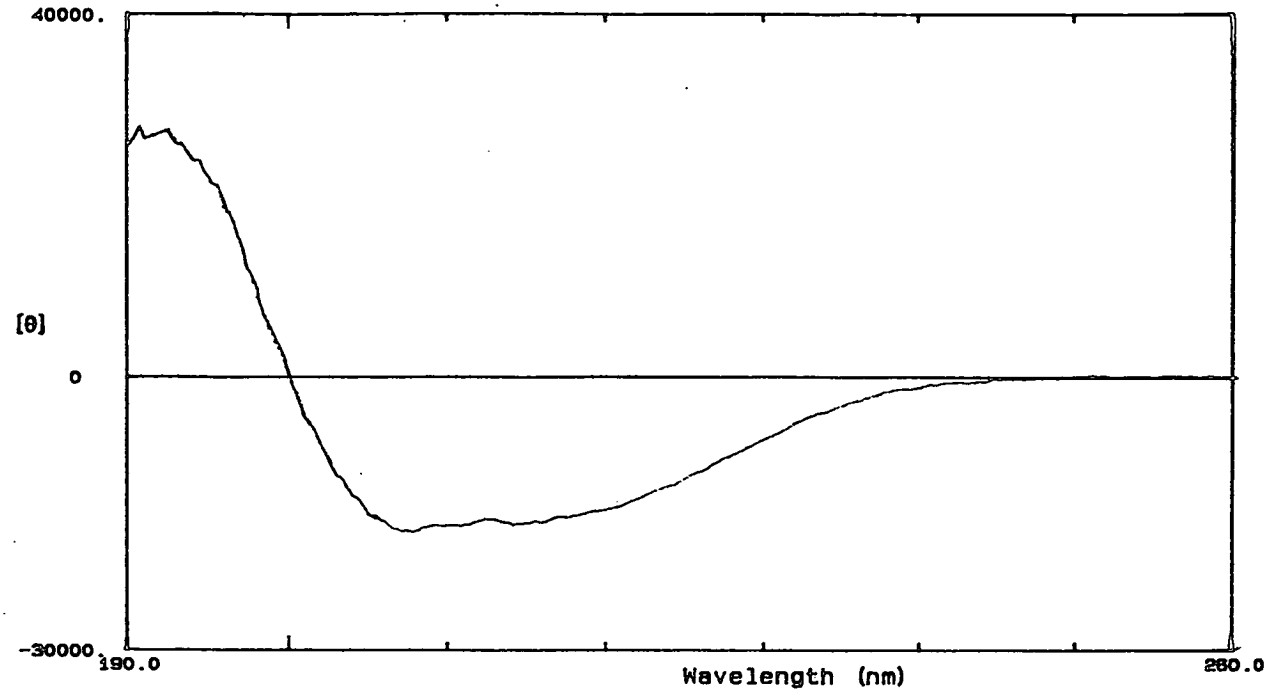


Fig. 3.17 CD spectrum of PRE A/WSN/33 HA in 10% HFIP-TFE

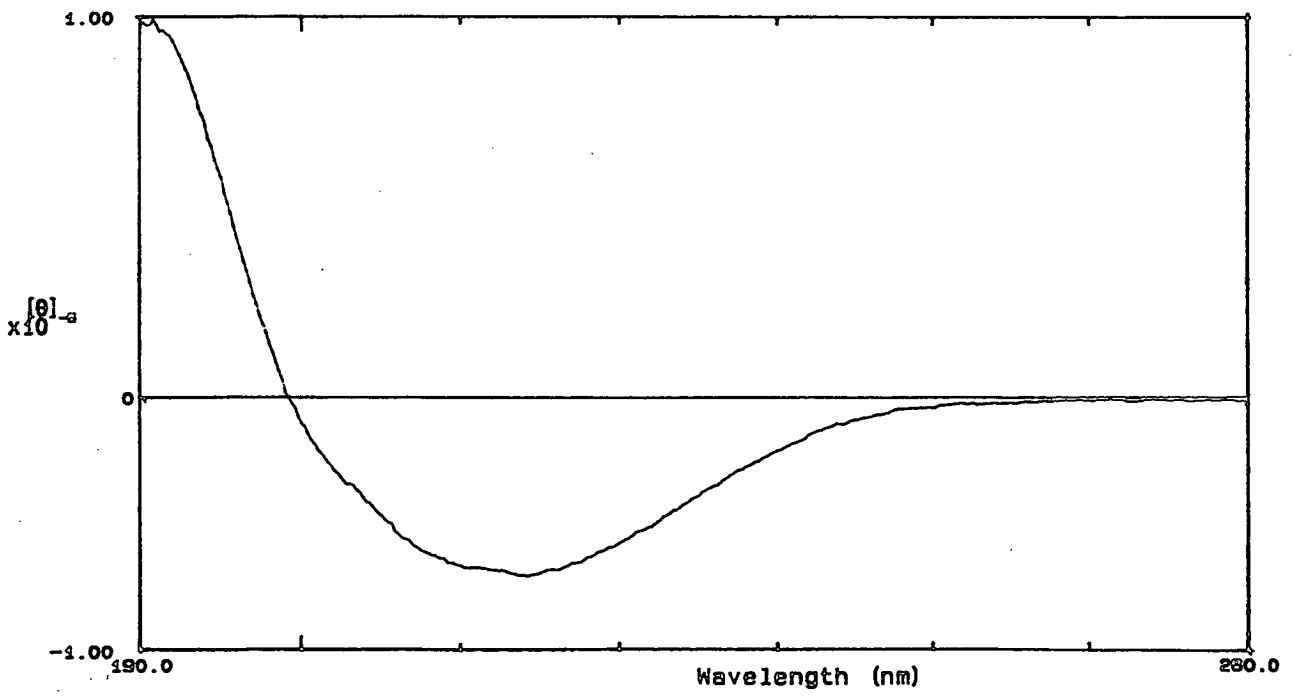


Fig. 3.18 CD spectrum of PRE A/WSN/33 HA in 50% HFIP-H_2O

A change in CD profile was observed in moving to the less polar solvent system, 10%HFIP-methanol, fig. 3.15. Here the spectrum has a single negative extreme at 215nm, $[\theta]_{215}$=-16,800degcm^{2}dmol^{-1}, indicative of a predominantely β structure, calculated as 91% β-sheet by the CONTIN program. In the more lipophilic solvent mixture, 10%HFIP-n-butanol the CD spectrum, fig. 3.16, shows two negative extremes, at 208nm and 220nm, $[\theta]_{208}$=-22,900degcm^{2}dmol^{-1} and $[\theta]_{220}$=-19,000degcm^{2}dmol^{-1}, characteristic of a predominantely helical conformation. The helical content was calculated as 65% using the Greenfield-Fasman (GF) method and 91% using the CONTIN program. The CD spectrum obtained of the peptide in 50%HFIP-water, fig. 3.17, has a single trough at 214nm, $[\theta]_{214}$=-6,900degcm^{2}dmol^{-1}. This is again characteristic of a predominantely β structure, 75% as calculated by the CONTIN program with a low helical content, (35% from the CONTIN program and 7% as calculated by the GF method). In 10%HFIP-TFE a spectrum characteristic of a predominantely helical structure is obtained, fig. 3.18, with two negative extremes one at 207-208nm and one at 215nm, $[\theta]_{208}$=-16,700degcm^{2}dmol^{-1} and $[\theta]_{215}$=-15,400degcm^{2}dmol^{-1}. Analysis using the CONTIN program showed almost equal amounts of α-helix and β-sheet (46% and 54% respectively). The GF method gave a helical content of 44%.

These results indicate the formation of a predominantely helical structure in homogeneous lipophilic media, where hydrogen bonding to solvent molecules is not possible, and a predominantely β-sheet conformation in more polar solvents.

Detergent micelles are proposed to mimic an amphiphilic surface such as a lipid bilayer. Fig. 3.19 shows the CD spectra of the PRE

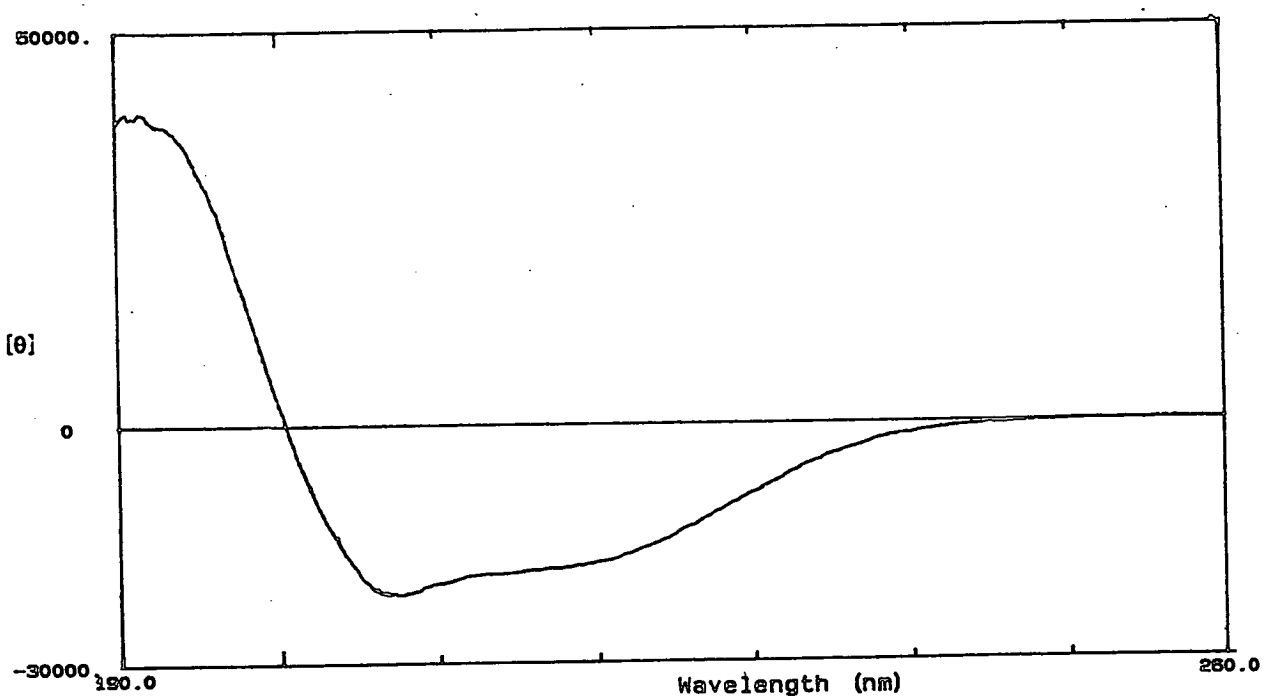


Fig. 3.19 CD spectrum of PRE A/WSN/[illegible] HA in 50% HFIP-0.6% SDS

A/WSN/33 HA peptide in 50%HFIP:aqueous sodium dodecyl sulphate (SDS) solution. This shows negative extremes at 207-208nm and 220nm, $[\theta]_{208}$=-21,600 degcm^{2}dmol^{-1} and $[\theta]_{220}$=-16,900degcm^{2}dmol^{-1}, indicating a predominance of helical structure which contrasts with the predominantely β-structure found in 50%HFIP:water. The CONTIN program gave a helical content of 58% with a β-sheet content of 42% reversing the slight predominance of β-sheet structure observed in the HFIP:H_2O solution. A slightly higher value for the helical content of this system of 61% was obtained using the GF method which gives results comparable to the helical content found for the n-butanol solution.

The series of CD studies therefore indicates an increased tendency to form an α-helical structure as the lipophilicity of the system increases whether by using more lipophilic alcohols as solvents or upon the addition of detergent micelles. These results would appear to support a model in which the transmembrane passage of the proteins is mediated in some way by the conformational transition of the hydrophobic sequences to α-helices in the lipophilic environment of the membrane.

CHAPTER 4

4 Experimental

4.1 General Methods

All starting materials were purchased from commercial sources mainly Aldrich, Lancaster and Fluka. Reagents were routinely distilled or recrystallized before use. All amino acids used for solid phase peptide synthesis were purchased from Novabiochem and used as suppplied. The following solvents were dried, using the reagents quoted in parenthesis, before distillation : diethyl ether, tetrahydrofuran (lithium aluminium hydride); acetone, methylene chloride (calcium chloride); ethanol (magnesium sulphate).

Merck Kieselgel 100 silica was used for flash column chromatography. Thin layer chromatography (tlc) was carried out on aluminium sheets precoated with Merck Kieselgel 60 silica. Solvent systems for tlc were either ethyl acetate/hexane (1:3) or n-butanol/acetic acid/water (13:2:5) unless otherwise stated. Visualisation of the components was achieved by either : iodine vapour, UV absorption at 254nm or phosphomolybdic acid spray. Compounds with free amino groups were visualized with ninhydrin spray.

High performance liquid chromatography was carried out using an Applied Biosystems system comprising 2 x 1406A solvent delivery system, a 1480A injector/mixer, and a 1783A detector/controller. The eluent was monitored at 277nm. Gel electrophoresis was carried out using the LKB Pharmacia PhastGel electrophoresis system. Amino acid analysis was carried out on an LKB 4150 alpha amino analyser following sealed tube hydrolysis in constant boiling hydrochloric acid at 100°C.

Infra red spectra were recorded on a Perkin Elmer 781 spectrophotometer using the polystyrene 1603cm^{-1} peak as a standard. Samples were tested as chloroform solutions.

Electron impact (EI) mass spectra were recorded on an AEI MS90Z or Kratos MS50TC spectrometer. Fast atom bombardment (FAB) spectra were recorded on a Kratos MS50TC spectrometer.

Proton magnetic resonance spectra were recorded on either a Jeol JNM-PMX60 (60 MHz), Bruker WH360 (360 MHz) or Varian VXR5000 (600MHz) spectrometer in the solvent indicated using tetramethylsilane (TMS) (δ 0.0) as an external standard.

Melting points were recorded on an electrically heated Reichard 7905 melting point apparatus and are uncorrected.

4.2 NMR Studies on Dpp.Met^5.Enkephalin

Nmr experiments were performed on a Varian VXR-5000 spectrometer at 600MHz or on a Bruker WH360 at 360MHZ. The experiments were carried out on a 5.2mM solution of the peptide in dmso d_6. Data sets consisting of 256 t_1 increments of 16 transients were collected for the ROESY experiment, with a mixing time of 200ms. The spectra were resolution enhanced by zero filling and sine bell multiplication.

Data sets consisting of 256 t_1 increments of 32 transients were collected for the COSY experiment and the spectrum processed as described for the ROESY spectrum.

4.3 Tyrosine Synthesis

Diethyl acetamido-4'methoxybenzyl-malonate (41)

A. Diethyl acetamidomalonate (1.10g, 5mmol), prepared by the method of Zambito and Howe[118], was added to a solution of sodium (0.11g, 5mmol) in ethanol ($30cm^3$) and the solution heated under reflux for thirty minutes. 4-Methoxybenzyl bromide (1.0g, 5mmol), freshly prepared by the method of Baldwin *et al*,[112] was added and the mixture heated under reflux for 17 hours. The reaction mixture was allowed to cool to room temperature and the solvent removed under vacuum. The residue was crystallized from aqueous ethanol before purifying by flash column chromatography on silica gel eluting with 10% ethyl acetate in toluene. Recrystallization of the product from aqueous ethanol gave the required product (0.279g, 17%). mp 97-99°C (lit., 96-97°C)[190]. δH ($CDCl_3$, 60MHz) 1.57 (6H, t, $2xO_2CH2CH_3$, J=7Hz), 2.13 (3H, s, CH_3CO_2), 3.72 (2H, s, CH_2O), 3.90 (3H, s, CH_3O), 4.40 (4H, q, $2xCO_2CH_2CH_3$, J=7Hz) 6.90 (2H, d, C_6H_4, J=8Hz), 7.08 (2H, d, C_6H_4, J=8Hz). EIMS m/z$^+$ 337 (M^+, 4), 121 (100), 43 (22), 40 (61), 32 (74%).

B. Sodium hydride (0.24g, 10mmol) was added to a solution of diethyl acetamidomalonate (2.17g, 10mmol) in THF ($25cm^3$) and the solution heated under reflux for thirty minutes. 4-Methoxybenzyl bromide (2.4g, 12mmol) was added and the mixture heated under reflux for 16 hours. The reaction mixture was allowed to cool to room temperature before filtering and evaporating to give a white solid which was purified by flash column chromatography on silica gel, eluting with 15% ethyl acetate in toluene. Recrystallization of the product from aqueous ethanol

gave the required product as white plates which were identical by nmr, tlc, mass spec and mp to an authentic sample (1.46g, 43%).

4-Benzyloxybenzyl nitrile (44)

4-Benzyloxybenzyl chloride (2.00g, 8.6mmol), sodium cyanide (0.64g, 12.8mmol) and sodium iodide (100mg) were heated under reflux in acetone (20cm^3) for eight hours with the exclusion of moisture. The reaction mixture was cooled and filtered and the solid washed with acetone (50cm^3) and toluene (50cm^3). The combined acetone extracts were evaporated to dryness and the residue taken up in toluene (50cm^3). The toluene extracts were combined before washing with water (100cm^3) and drying over sodium sulphate. Evaporation gave a yellow oil which crystallized on standing, recrystallization from ethanol gave the nitrile (1.90g, 100%). mp 66-67°C. δH ($CDCl_3$, 60MHz) 3.64 (2H, s, $ArCH_2CN$), 5.10 (2H, s, $ArCH_2O$), 6.92 (2H, d, C_6H_4, J=8Hz), 7.24 (2H, d, C_6H_4, J=8Hz), 7.36 (5H, s, ArH). EIMS m/z^+ 223 (M^+, 24), 91 (100), 65 (25). m/z(EI) 233.0994 (M^+, $C_{15}H_{13}NO$ requires 223.0997). (Found: C, 80.3; H, 5.88; N, 6.04. $C_{15}H_{13}NO$ requires C, 80.69, H, 5.87; N, 6.27%)

Attempted Synthesis of 4-Benzyloxyphenylacetic acid (45)

4-Benzyloxybenzyl nitrile (0.5g, 2.4mmol) was suspended in 10M sulphuric acid (20cm^3) and heated under reflux for fifteen minutes. After cooling to room temperature the reaction mixture was extracted with ethyl acetate (100cm^3). Tlc showed the reaction mixture to be multicomponent. No product was isolated from this reaction.

4-Benzyloxyphenylacetic acid (45)

4-Benzyloxybenzylnitrile (2.00g, 8.97mmol) was dissolved in hot ethanol ($40cm^3$). A solution of potassium hydroxide (7.5g) in water ($15cm^3$) was added and the mixture was heated under reflux for 15 hours. The reaction mixture was diluted with water ($100cm^3$) and acidified to pH1 with 6N HCl before extracting with ethyl acetate ($200cm^3$). The combined extracts were washed with 1N HCl ($200cm^3$), dried over magnesium sulphate and evaporated to dryness to give an off white solid which was recrystallized from benzene to give the acid (1.57g, 73%). mp 117-119°C. δH ($CDCl_3$, 60MHz) 3.60 (2H, s, $ArCH_2CO_2H$), 5.10 (2H, s, $ArCH_2O$), 6.97 (2H, d, C_6H_4, J=8Hz), 7.27 (2H, d, C_6H_4, J=8Hz), 7.36 (5H, s, ArH), 10.32 (1H, bs, CO_2H). EIMS m/z$^+$ 242 (M^+, 20), 91 (100), 65 (13%). m/z(EI) 242.0937 (M^+, $C_{15}H_{14}O_3$ requires 223.0942).

Attempted Synthesis of Ethyl 4-benzyloxyphenylacetic acid (46)

4-Benzyloxybenzyl nitrile (0.5g, 2.4mmol) was taken up in a mixture of ethanol ($15cm^3$) and DMF ($10cm^3$). Concentrated sulphuric acid ($10cm^3$) was added and the mixture heated under reflux for fifteen hours. The reaction mixture was allowed to cool to room temperature and diluted with ice cold water ($35cm^3$) before extracting with ethyl acetate ($100cm^3$). The organic extract was washed with water ($100cm^3$) and evaporation gave an oil which crystallized on standing. Tlc showed the product to be multicomponent. No product was isolated from this reaction.

Ethyl 4-benzyloxyphenylacetic acid (46)

4-Benzyloxyphenylacetic acid (0.5g, 2.0mmol) was dissolved in a solution of ethanol ($28.5cm^3$) containing acetylchloride ($2.5cm^3$) and the

mixture stirred at room temperature for three hours. The reaction mixture was then concentrated to approximately 15cm^3 and made basic by the addition of aqueous sodium hydroxide before extracting with ethyl acetate. Evaporation gave the ester (0.29g, 54%). δH ($CDCl_3$, 60MHz) 1.23 (3H, t, OCH_2CH_3 , J=8Hz), 3.55 (2H, s, $ArCH_2CO_2Et$), 4.17 (2H, q, OCH_2CH_3 , J=8Hz), 5.05 (2H, s, $ArCH_2O$), 6.93 (2H, d, C_6H_4, J=8Hz), 7.23 (2H, d, C_6H_4, J=8Hz), 7.42 (5H, s, ArH). EIMS m/z$^+$ 270 (M^+, 14), 91(100%).

Benzyl 4-benzyloxyphenylacetate (49)

4-Hydroxyphenylacetic acid (10.0g, 66mmol), benzyl chloride (19.8g, 156mmol), sodium iodide (10.8g), and potassium carbonate (21.6g) were stirred as a partial suspension in refluxing acetonitrile (200cm^3) for 30 hours. The reaction mixture was filtered and the residue washed with acetonitrile (100cm^3). The combined solutions were evaporated *in vacuo* and the residue recrystallized from aqueous ethanol to give the ester as a white solid (14.9g, 69%). mp 73-75°C (lit., 73-74°C)[123]. δH ($CDCl_3$, 60MHz) 3.60 (2H, s, $ArCH_2CO_2$), 5.05 (2H, s, $ArCH_20$), 5.13 (2H, s, CO_2CH_2) 6.93 (2H, d, C_6H_4, J=8Hz), 7.23 (2H, d, C_6H_4, J=8Hz), 7.33 (5H, s, ArH), 7.40 (5H, s, ArH). EIMS m/z$^+$ 332 (M^+, 14), 91 (100), 65 (4).

4-Benzyloxyphenethylalcohol (47)

A. 4-Benzyloxyphenylacetic acid (2.0g, 8mmol) in tetrahydrofuran (50cm^3) was added dropwise to a suspension of lithium aluminium hydride (1.1g, 27.5mmol) in tetrahydrofuran (50cm^3) with stirring. The resulting solution was stirred under reflux for 2 hours. The reaction mixture was allowed to cool to room temperature before adding water

($2.2cm^3$), 15% sodium hydroxide ($2.2cm^3$) and water ($3.3cm^3$). The mixture was filtered and dried over magnesium sulphate before evaporating to dryness to give an off white solid. Purification by flash column chromatography on silica gel gave a white solid which recrystallized from hexane as white plates (1.38g, 73%). m.p. 85°C (lit., 83-85°C)[123]. δH($CDCl_3$, 60MHz) , 1.43 (1H, s, OH), 2.80 (2H, t, $ArCH_2C$, J=6Hz), 3.83 (2H, t, CH_2OH, J=6Hz), 5.07 (2H, s, $ArCH_2O$), 6.93 (2H, d, C_6H_4, J=8Hz), 7.20 (2H, d, C_6H_4, J=8Hz), 7.43 (5H, s, ArH). EIMS m/z$^+$ 228 (M^+, 24), 91 (100), 65 (10). (Found: C, 78.9; H, 7.12; N, 0.00. $C_{15}H_{16}O_2$ requires C, 78.92; H, 7.06; N, 0.00%)

B. Ethyl 4-benzyloxyphenylacetate (0.40g,1.5mmol) in THF ($5cm^3$) was added dropwise to a suspension of lithium aluminium hydride (0.114g, 3.0mmol) in THF ($40cm^3$) and heated under reflux for one hour. The reaction mixture was allowed to cool to room temperature before adding water ($0.28cm^3$), 15% sodium hydroxide ($0.28cm^3$) and water ($0.69cm^3$). The suspension was filtered and the cake washed with ether. The resulting solution was dried and evaporation gave a white solid which was purified by flash column chromatography on silica gel eluting with ethyl acetate and hexane. This gave the required product as a white crystaline solid which was recrystallized from hexane as white plates (0.25g, 75%) which were identical by nmr, mp, and tlc to an authentic sample.

C. Benzyl 4-benzyloxyphenylacetate (4.50g, 13.4mmol) in THF ($50cm^3$) was added dropwise to a suspension of lithium aluminium hydride (1.20g, 31.6mmol) in THF ($100cm^3$). This was stirred under reflux

for 2 hours. The reaction mixture was allowed to cool to room temperature before adding enough saturated aqueous sodium potassium tartrate to produce a granular suspension. The solution was then filtered and the solid washed with THF. Drying over magnesium sulphate and removal of solvent gave an off white solid. Purification by flash column chromatography on silica gel and recrystallization from hexane gave the alcohol as white plates (2.31g, 75%), which were identical in m.p., tlc and nmr to an authentic sample.

4-Benzyloxyphenylacetaldehyde (42)

Chromium trioxide (1.3g, 13mmol) was added to a stirred solution of pyridine (2.3cm^3) in dry methylene chloride (40cm^3). This was stirred on an ice bath for five minutes and then at room temperature for a further ten minutes. 4-Benzyloxyphenethylalcohol (0.5g, 2.2mmol) in dichloromethane (5cm^3) was added dropwise and stirring continued for three hours. The supernatant was decanted and the residue washed with ether (200cm^3). The combined organic extracts were filtered through a celite pad and washed with 5% sodium hydroxide (200cm^3), 1N HCl (200cm^3), 5% sodium carbonate (200cm^3) and brine (200cm^3). Drying over magnesium sulphate and evaporating to dryness gave a clear oil (0.40g, 80%). δH ($CDCl_3$, 60MHz) 3.55 (2H, d, CH_2CHO, J=3Hz), 5.00 (2H, s, CH_2O), 6.92 (2H, d, C_6H_4, J=8Hz), 7.16 (2H, d, C_6H_4, J=8Hz), 7.36 (5H, s, ArH), 7.72 (1H, t, CHO). ν_{max} ($CHCl_3$) 2720 (CHO), 1718 (carbonyl), 1605cm^{-1}, (aromatic).

4-Benzyloxyphenyl-α-hydroxy-propionitrile (50)

To a stirred suspension of 4-benzyloxyphenylacetaldehyde (191mg, 0.85mmol) in water ($3cm^3$) was added a solution of sodium metabisulphite (115mg, 0.6mmol) in water ($3cm^3$) and stirring continued for one hour. A solution of sodium cyanide (115mg, 4mmol) in water ($3cm^3$) was then added dropwise with stirring and stirring continued for a further three hours. The reaction mixture was extracted with ether ($100cm^3$) and the combined ether extracts washed with brine ($100cm^3$) and sodium metabisulphite solution ($100cm^3$). Drying and evaporation gave the cyanohydrin (172mg, 80%) which was used directly in the next reaction. δH($CDCl_3$, 60MHz) , 2.10 (1H, s, OH), 3.13 (d, 2H, $ArCH_2C$, J=6Hz), 4.70 (1H, t, CH, J=6Hz), 5.07 (2H, s, $ArCH_2O$), 6.07 (2H, d, C_6H_4, J=8Hz), 7.30 (2H, d, C_6H_4, J=8Hz), 7.50 (5H, s, ArH). υ_{max} ($CHCl_3$) 3420 (OH), 2720 (CN), 1605cm^{-1} (aromatic).

4 Benzyloxyphenyl-α-amino-propionitrile (51)

A. 4-Benzyloxyphenyl-α-hydroxy-propionitrile prepared from 4-benzyloxyphenethytlalcohol (500mg, 2.2mmol) was taken up in ethanol : water ,1:1, ($30cm^3$) and treated with a solution of ammonium chloride (400mg,7.5mmol) in 2N potassium hydroxide ($1.5cm^3$) and heated at 50°C for 24 hours. The reaction mixture was allowed to cool to room temperature before evaporating in vacuo top give a yellow oil, which was resuspended in ether, filtered and dried. Subsequent evaporation gave a yellow oil which was purified by flash column chromatography on silica gel eluting with 20% ethyl acetate in hexane ($230cm^3$) and ethyl acetate ($70cm^3$) to give the aminonitrile (70mg, 13%) attemts to crystallize the product were unsuccessful. δH ($CDCl_3$, 60MHz) 1.70 (2H, bs, NH_2), 2.98

(2H, d, CH_2, J=6Hz), 3.91 (1H, t, CH, J=6Hz), 5.10 (2H, s, $ArCH_2O$), 6.95 (2H, d, C_6H_4, J=9Hz), 7.30 (2H, d, C_6H_4, J=9Hz), 7.43 (5H, s, ArH). EIMS m/z 225 (M-HCN, 17), 197 (7), 134 (51), 91(100%). υmax ($CHCl_3$) 3400 (NH_2), 2210 (CN), 1605cm^{-1} (aromatic).

B. 4-Benzyloxyphenyl-α-hydroxy-propionitrile prepared from 4-benzyloxyphenethylalcohol (180mg, 0.79mmol) was taken up in ethanol (3cm^3) and water (1cm^3) was added. The solution was slowly heated to 60°C before a solution of ammonium chloride (120mg, 2.2mmol) in 2N sodium hydroxide (1cm^3) was added dropwise *via* a syringe over a period of fifteen minutes. Heating was then continued at 60°C for a further four hours. The reaction mixture was allowed to cool to room temperature and extracted with ether (20cm^3). The aqueous layer was saturated with sodium chloride and reextracted with ether (20cm^3). The combined organic extracts were dried and evaporation gave a yellow solid, which was purified by flash column chromatography on silica gel eluting with 50% ethyl acetate in hexane (100cm^3) and ethyl acetate (100cm^3) to give a yellow solid (79mg, 40%) which was nmr and tlc identical to an authentic sample.

C. 4-Benzyloxyphenyl-α-hydroxy-propionitrile prepared from 4-benzyloxyphenethylalcohol (180mg, 0.79mmol) was taken up in ethanol (2cm^3) and the solution diluted with water (2cm^3). A solution of ammonium chloride (250mg, 4.7mmol) in 10% sodium hydroxide (1cm^3) was added dropwise to the stirred solution over a period of 30 minutes maintaing the temperature of the solution at 70°C. Heating was continued at 70°C for 1 hour before allowing the solution to cool to room

temperature. The crude reaction mixture was extracted with ether ($120cm^3$) and the combined organic extracts washed with water ($100cm^3$). Evaporation gave a yellow oil which was purified by flash column chromatography on silica gel eluting with a gradient of 0-100% ethyl acetate in hexane ($100cm^3$) and ethyl acetate ($100cm^3$) to give the aminonitrile (82mg, 41%), which was tlc and nmr identical to an authentic sample.

D. 4-Benzyloxyphenyl-α-hydroxy-propionitrile prepared from 4-benzyloxyphenethylalcohol (180mg, 0.78mmol) was taken up in ether ($4cm^3$). This was added to a solution of solution of water and ethanol, 1:1 ($4cm^3$) which was maintained at a temperature of 70°C at a rate which allowed the ether to evaoporate off slowly. A solution of ammonium chloride (55mg, 1.0mmol) in 2N sodium hydroxide ($1cm^3$) was added dropwise *via* syringe over a period of thirty minutes. The temperature was reduced to 30°C and stirring continued for a further 12 hours. The reaction mixture was allowed to cool to room temperature before evaporated to dryness and resuspending in ether. Filtration and evaporation gave a yellow solid, which was purified by flash column chromatography on silica gel eluting with 20% ethyl acetate in hexane ($150cm^3$) and ethyl acetate ($100cm^3$) to give a yellow solid (72mg, 39%) which was nmr and tlc identical to an authentic sample.

E. 4-Benzyloxyphenyl-α-hydroxy-propionitrile prepared from 4-benzyloxyphenethylalcohol (1.0g, 4.4mmol) was taken up in hot ethanol ($20cm^3$) and diluted with water ($10cm^3$). The solution was slowly heated to 70°C and a solution of ammonium chloride (700mg, 13mmol) in 10%

sodium hydroxide (2.5cm^3) added dropwise *via* syringe over a period of twenty minutes. The mixture was stirred at 70°C for a further ninety minutes before cooling to room temperature and extracted with ether (75cm^3) and the ether extract washed with water (200cm^3) before drying and evaporating to give a yellow oil, which was purified by flash column chromatography on silica gel eluting with 50% ethyl acetate in hexane (100cm^3) and ethyl acetate (100cm^3) to give a yellow solid (213mg, 27%) which was nmr and tlc identical to an authentic sample.

F. 4-Benzyloxyphenyl-α-hydroxy-propionitrile prepared from 4-benzyloxyphenethylalcohol (500mg, 2.2mmol) was taken up in ethanol (7cm^3) and the solution diluted with water (7cm^3). To the resulting suspension was added ammonium chloride (500mg, 9.3mmol) in 10% sodium hydroxide (5cm^3). This was sonicated at 40°C for 6 hours. The reaction mixture was extracted with ether (300cm^3) and evaporated to give a yellow oil which was purified by flash column chromatography on silica gel eluting with hexane (30cm^3), 50% ethyl acetate in hexane (80cm^3) and ethyl acetate (60cm^3) to give the aminonitrile (91mg, 21%), which was tlc and nmr identical to an authentic sample.

G. 4-Benzyloxyphenylacetaldehyde prepared from 4-benzyloxyphenethylalcohol (180mg, 0.78mmol) was taken up in acetonitrile (5cm^3) and added to a stirred suspension of ammonium chloride (130mg, 2.4mmol) potassium cyanide (130mg, 2.0mmol) and alumina (500mg) in acetonitrile (15cm^3). The suspension was heated, with stirring, at 50°C for 24 hours. The reaction mixture was allowed to cool to room temperature before filtering and evaporating to give a yellow oil,

which was resuspended in ether, filtered and dried. Subsequent evaporation gave a yellow oil which was found to be multicomponent by tlc. No product was isolated from this experiment.

H. 4-Benzyloxyphenyl-α-hydroxy-propionitrile prepared from 4-benzyloxyphenethylalcohol (100mg,) was taken up in acetonitrile (15cm^3). Ammonium chloride (130mg, 2.4mmol) and alumina (500mg) were added and the suspension heated at 50°C for 24 hours. The crude reaction mixture was filtered and the filtrate dried and evaporated to give a yellow oil which was multicomponent by tlc. No product was isolated from this reaction.

4 Benzyloxyphenyl-α-aminobenzyl-propionitrile (53)

4-Benzyloxyphenylacetaldehyde prepared from 4-benzyloxyphenethylalcohol (500mg, 2.2mmol) was taken up in dichloromethane (5cm^3) and zinc iodide (10mg) added. Trimethylsilylcyanide (0.30cm^3, 2.2mmol) was added to the stirred solution under argon with cooling. Stirring was then continued at room temperature for fortyfive minutes. A solution of benzylamine (0.24cm^3, 2.2mmol) in methanol (5cm^3) was added and the resulting solution heated under reflux for two hours under argon. The reaction mixture was cooled to room temperature before evaporating to dryness. The residue was taken up in ether and dried before filtering and evaporating to give a yellow oil which was purified by flash column chromatography on silica gel eluting with a gradient of 0-60% ethyl acetate in hexane (200cm^3) to give an off white solid which was crystallized from ethanol giving the aminonitrile (193mg, 27%). mp 115°C. δH ($CDCl_3$, 60MHz) 1.70 (2H, bs,

NH_2), 2.98 (2H, d, CH_2, J=6Hz), 3.91 (1H, t, CH, J=6Hz), 5.10 (2H, s, $ArCH_2O$), 6.95 (2H, d, C_6H_4, J=9Hz), 7.30 (2H, d, C_6H_4, J=9Hz), 7.39 (5H, s, ArH), 7.43 (5H, s, ArH). EIMS m\z$^+$ 315 (M-HCN, 12), 224 (10), 198 (13), 91(100%). FABMS m\z$^+$ 343 (M+H), 316, 226, 197, 91. υmax ($CHCl_3$) 3310 (NH), 2210 (CN) 1610cm^{-1} (aromatic).

Acid Hydrolysis of 4-Benzyloxyphenyl-α-amino-propionitrile (51)

A. 4-Benzyloxyphenyl-α-amino-propionitrile (74mg, 0.29mmol) was suspended in 48% hydrobromic acid (5cm^3) and the sealed tube heated at 100°C for 16 hours. The resulting dark brown solution was evaporated to dryness and taken up in water (5cm^3). The suspension was filtered through a plug of charcoal to give a clear solution which was multicomponent by tlc. Purification by cation exchange chromatography on Dowex AG50W X8 eluting with a gradient of 0-1M pyridine in water provided no pure tyrosine by tlc.

B. 4-Benzyloxyphenyl-α-amino-propionitrile (50mg, 0.20mmol) was suspended in concentrated hydrochloric acid (5cm^3). The resulting suspension was treated as above and provided no pure tyrosine by tlc.

C. 4-Benzyloxyphenyl-α-amino-propionitrile (50mg, 0.20mmol) was suspended in 6N hydrochloric acid (5cm^3) before heating and chromatographing as above. Tlc of the ninhydrin positive fractions showed no pure tyrosine.

O-Benzyl D,L-tyrosine (54)

4-Benzyloxyphenyl-α-amino-propionitrile (55mg, 0.22mmol) was taken up in ethanol (20cm^3) before adding a solution of potassium hydroxide (3.7g) in water (7.5cm^3). The solution was stirred at 100°C for five hours and cooled to room temperature. The ethanol was removed by evaporation and the aqueous neutralised with 6N HCl. The solution was filtered and filtrate applied to a column of Dowex AG50W X8 ion exchange resin and eluted with water (500cm^3) and 1N ammonium hydroxide (300cm^3). The ninhydrin positive fractions were pooled and dried to give the required product as a white solid which was tlc and nmr identical to an authentic sample (59mg, 100%).

Attempted Preparation of D,L-Tyrosine

4-Benzyloxyphenyl-α-hydroxy-propionitrile prepared from 4-benzyloxyphenethytlalcohol (180mg, 0.79mmol) was taken up in ether (3cm^3) and added to a solution of ethanol:water, 1:1, (3cm^3). This was slowly added to a heated reaction vessel at 50°C allowing the ether to evaporate. Ammonium chloride (130mg, 2.4mmol) in 2N sodium hydroxide (0.5cm^3) was added with stirring and heated at 50°C for 2 hours before cooling to room temperature and stirring for three days. The reaction mixture was evaporated to dryness and taken up in ether before filtering. The ether solution was added to 9N HCl (4cm^3) and heated gently allowing the ether to evaporate. The reaction vessel was then sealed and and heated at 110°C for 14 hours. The cooled reaction mixture was evaporated to dryness and the residue taken up in water (4cm^3). The aqueous solution was adjusted to pH7 and applied to AG50W X8 Dowex. Elution with a gradient of 0-1M pyridine in water gave ninhydrin positive

fractions which were pooled and freeze dried to give a white solid which was multicomponent by tlc.

D,L-Tyrosine

O-Benzyl-D,L-tyrosine (50mg, 0.20mmol) was treated with thioanisole (1.2cm^3), trifluoromethanesulphonic acid (0.1cm^3) and TFA (2cm^3) for 30 minutes at 0°C. The mixture was concentrated *in vacuo* to give an oil which was taken up in water (3cm^3) and neutralised with 10% NaOH dofore applying to a column of Dowex AG50W X8 and eluting with water (500cm^3) and 1N ammonium hydroxide (300cm^3). The ninhydrin positive fractions were pooled and freeze dried to give a white solid (34mg, 79%) which comigrated with an authentic sample of L-tyrosine on tlc.

4-(4'-Acetoxybenzylidene)-Δ^2-3N-oxazolin-5-one (57)

N-benzoylglycine prepared by the method of Baldwin *et al*[112] (2.21g, 12.2mmol), sodium acetate (2.21g, 26.9mmol), acetic anhydride (8.3cm^3) and 4-acetoxybenzaldehyde (4.3g, 3.68cm^3, 26.3mmol) were heated at 100C for 2.5 hours with stirring. The reaction mixture was then cooled and treated with ethanol (11cm^3). The resulting suspension was stirred for 15 minutes before standing at room temperature for 14 hours. The precipitate was collected and washed with cold ethanol (50cm^3) and hot water (50cm^3). The solid was dried under vacuum to give a yellow solid (2.28g, 61%). mp 181-182°C (lit 180-181°C)[190]. δH ($CDCl_3$) 2.16 (3H, s, CH_3CO_2), 7.00-8.16 (10H, m, ArH, CH). υmax ($CHCl_3$) 1800, 1760 (carbonyl), 1660cm^{-1} (C=N) . EIMS m\z 265 (M-AcOH, 5), 105 (100) 77(19%).

Attempted synthesis of 4-hydroxyphenylpyruvic acid

A. 4-(4'-Acetoxybenzylidene)-3N-oxazolin-5-one (500mg, 1.6mmol) was suspended in 10% NaOH ($30cm^3$) and heated at 100°C for 1.5 hours. The reaction mixture was cooled to room temperature and acidified with conc. HCl. The resulting precipitate was collected and washed with 2N HCl before drying to give 2-benzamido-3-(4'-hydroxyphenyl)-prop-2-enoic acid (270mg, 95%). mp >210°C. δH (dmso d_6) 6.63-8.00 (10H, m, ArH, CH), 9.75 (1H, s, CO_2H), 9.23 (1H, bs, NH). EIMS m\z 283 (M^+, 13), 239 (6), 105 (100), 77 (36), 44 (11), 40 (21%).

B. 4-(4'-Acetoxybenzylidene)-3N-oxazolin-5-one (500mg, 1.6mmol) was suspended in 6.25M sodium hydroxide ($1.5cm^3$) and heated at 210°C in a sealed tube for 15 minutes. The resulting solution was cooled to room temperature, the resulting glass coluld not be removed from the reaction vessel.

C. 4-(4'-Acetoxybenzylidene)-3N-oxazolin-5-one (3.0g, 9.8mmol) was suspended in a mixture of acetic acid ($130cm^3$) and 6N HCl ($37cm^3$) and heated at 100°C for 4 hours. The resulting deep purple solution was stored at 5°C overnight. No product separated on cooling.

D. 4-(4'-Acetoxybenzylidene)-3N-oxazolin-5-one (491mg, 1.6mmol) was suspended in 10% sodium hydroxide ($60cm^3$) and heated at 100°C for 9 hours. The resulting solution was cooled to room temperature and saturated with sulphur dioxide. The gelatinous solution was cooled to room temperature filtered and the filtrate acidified with conc. HCl. The

resulting solution was heated on a water bath for 4 hours giving a deep purple coloured solution. No product separtated on cooling.

D,L-Tyrosine

A. Sodiumcyanoborohydride (200mg, 3.2mmol) was added to a soltion of 4-hydroxyphenylpruvic acid (180mg, 1.0mmol) and ammonium nitrate (100mg, 1.2mmol) in methanol (15cm^3). The solution was adjusted to pH7.2 with 10% potassium hydroxide and stirred at room temperature for 60 hours. Concentrated hydrochloric acid (5cm^3) was added and stirring continued for a further hour. The reaction mixture was evaporated to dryness and taken up in water. The slution was adjusted to pH7 and applied to AG50W X8 Dowex. Elution with water (500cm^3) and 1N ammonium hydroxide (300cm^3) gave ninhydrin positive fractions which were pooled and freeze dried to give a white solid (54mg, 30%). The product was identical by tlc and nmr to an authentic sample.

B. Sodiumcyanoborohydride (200mg, 3.2mmol) was added to a solution of 4-hydroxyphenylpruvic acid (180mg, 1mmol) and ammonium bromide (200mg, 2.0mmol) in methanol (15cm^3). The solution was adjusted to pH7.2 with 10% potassium hydroxide and stirred at room temperature for 4 days. Concentrated hydrochloric acid (5cm^3) was added and stirring continued for a further hour. The reaction mixture was worked up as with the above reaction to give a white powder (162mg, 90%) which comigrated with authentic tyrosine on tlc.

4.4 Synthesis, Purification and Structure Determination of PRE A/WSN/33 HA

4.4.1 Resin Loading

N(α)-Fluorenylmethoxycarbonylalanine (1.46g, 4.6mmol) was dissolved in a minimum amount of dry DMF and diisopropylcarbodiimide (0.36cm^3, 0.3g, 2.34mmol) was added before stirring at room temperature for five minutes.

The resin (0.78mmolg^{-1}, 1.0g, 0.78mmol) was swollen in a minmum amount of DMF and a few crystals of DMAP added before sonicating for five minutes. The Fmoc-Ala-OH/DIC solution was added to the resin and the mixture sonicated for three hours. The resin was filterd and washed with dry DMF and methylene chloride before drying.

The loading of the amino acid on the resin was determined by suspending a sample of the resin in a solution of piperidine in DMF (20% v/v). The suspension was then sonicated at room temperature for 15 minutes. A UV spectrum of the solution was then obtained and the absorbance of the solution at 300nm measured. Comparison of this value against standard values obtained from the protected amino acids allowed the evaluation of the resin substitution which was found to be 82%.

4.3.2 Capping of Unreacted Resin Sites

Fmoc-peptide-resin (0.5mmol) was treated with acetic anhydride (0.10g, 1.0mmol) and pyridine (0.080g, 1.0mmol) in DMF (7cm^3). The mixture was shaken for 2.5 minutes then filtered and treated with a further identical portion of acetic anhydride and pyridine in DMF. After

shaking for a further 3.5 minutes the mixture was filtered and the resin washed thoroughly with DMF.

4.4.3 Removal of the Fmoc Group

Fmoc-peptide-resin (0.5mmol) was treated with four portions of 20% piperidine in DMF (9.0cm^3) and shaken for 5, 3, 3 and 1 minute respectively. The mixture was filtered prior to each addition and the UV absorbance of the filtrate measured at 300nm on each occasion. After the last treatment the peptide-resin was washed thoroughly with DMF.

4.4.4 Activation of Fmoc-Amino Acids

A. Fmoc-amino acid (2.0mmol) was dissolved in DMF and treated with 0.5M DIC in DMF (2.0cm^3, 1.0mmol). The solution was allowed to stand for 15 minutes.

B. Fmoc-amino acid (1.0mmol) was dissolved in DMF (4.0cm3) and treated with 0.5M DIC in DMF (2.0cm^3, 1.0mmol) and 0.5M 1-hydroxybenzotriazole (HOBt) in DMF (2.0cm^3, 1.0mmol). The solution was allowed to stend for 15 minutes.

The Fmoc-amino acids were coupled using procedure A, then recoupled using procedure B.

4.4.5 Coupling of Activated Amino Acid to the Peptide-Resin

The activated amino acid solution, from 4.4.4 A or B, was added to the peptide resin (0.5mmol). The mixture was shaken for 30 minutes then filtered and the resin washed thoroughly with DMF.

4.4.6 Automated Chain Assembly

All peptides were assembled using an Applied Biosystems 430A peptide synthesiser. The synthesis was carried out on a 0.5mmol scale steps 4.4.2 to 4.4.5 inclusive ("capping", "deprotection", "activation" and "coupling") comprise the coupling cycle used to elongate the resin-bound peptide by one residue. Once all the amino acids had been coupled, one further round of capping and deprotection gives the free N-terminus.

4.4.7 Side Chain Protection

The following Fmoc-amino acids were incorporated with the side chain protection given in parenthesis : aspartic acid, threonine, tyrosine (t-Bu); lysine (Boc). The remaining amino acids were used unprotected.

4.4.8 Deprotection

A. The resin bound peptide (100mg) was stirred as a suspension in a 1:1 mixture of ethyl methyl sulphide and ethanedithiol ($0.5cm^3$) for 10 minutes. Trifluoroacetic acid 95% ($10cm^3$) was added and stirring continued for a further 3 hours under argon. The mixture was filtered and the filtrate concentrated to approximately $1cm^3$ *in vacuo*. Addition of ice cold diethyl ether ($10cm^3$) gave an off white precipitate which was collected by filtration. The product was washed with ether ($50cm^3$) and dried. The solid was not soluble in water, glacial acetic acid, 50% acetic acid, dichloromethane or TFE.

B. A portion of the resin bound peptide (100mg) was deprotected as above. The reaction mixture was filtered into a 1:1 mixture of diethyl

ether and ethyl acetate (50cm^3). The resulting suspension was centrifuged at 2000rpm for 10 minutes and the supernatant decanted. The solid was resuspended in a further portion of the above solvent and centrifugation repeated and the supernatant again decanted. This procedure was repeated a further five times to ensure removal of scavengers. The moist peptide was directly suspended in glacial acetic acid (9cm^3) and water added to give a solution of the peptide in 90% acetic acid. Dropwise addition of water gave a solution of the peptide in 30% acetic acid. Further addition of water resulted in precipitation of the peptide.

4.4.9 Purification

The solution of the peptide in 30% acetic acid was applied to a column of G-25 Sephadex (2.5 x 30cm) and eluted with 30% acetic acid at 25cm^3 per hour. UV absorbance of the eluant was monitored at 254 and 277nm. The appropriate fractions were pooled and lyophilised to give a white solid (10mg). The product was taken up in HFIP (1cm3) and addition of water (3cm^3) resulted in precipitation of the product. The peptide was collected by centrifugation and resuspended in HFIP (1cm^3). Precipitation was repeated six times by this method. The peptide was then dissolved in HFIP (2cm^3). 30% Acetic acid (5cm^3) was added and the HFIP removed in a stream of argon. Freeze drying gave peptide as a white solid (3mg). Tlc using the solvent system pyridine:n-butanol:acetic acid:water, (10:15:3:12), showed removal of migrating impurities in during the course of the purification. FAB-MS m/z 1867.8, (C_{89},H_{147},N_{19},$O_{22}$$S_1$ requires 1867.4) Amino acid analysis : Asx_1 (0.81), Thr_1 (0.91), Ala_4 (3.51), Val_2 (2.12), Met_1 (0.88), Leu_4 (3.57), Tyr_1 (0.89), Phe_1 (0.91), Lys_2 (2.10).

4.4.10 High Performance Liquid Chromatography (HPLC)

HPLC of PRE A/WSN/33 HA was attempted using the following solvent systems on both a C-8 and C-4 reversed phase column.

A. Solvent A, acetonitrile (0.1% TFA). Solvent B, water (0.1% TFA)

(i) 10 to 100% solvent A in solvent B over 30 minutes.

(ii) 100% solvent B for 10 minutes.

B. Solvent A, acetonitrile (0.1% TFA). Solvent B, water (0.1% TFA)

(i) 50 to 100% solvent A in solvent B over 30 minutes.

(ii) 100% solvent B for 10 minutes.

C. Solvent A, (95:5 acetonitrile:TFE, (0.1% TFA)). Solvent B, water (0.1% TFA)

(i) 10% solvent A in solvent B.

(ii) 10 to 20% solvent A in solvent B over 5 minutes.

(iii) 20 to 90% solvent A in solvent B over 15 minutes.

(iv) 90% solvent A in solvent B for 5 minutes.

D. Solvent A, (95:5 acetonitrile:TFE, (0.1% TFA)). Solvent B, water.

(i) 50 to 100% solvent A in solvent B over 30 minutes.

(ii) 100% solvent B for 10 minutes.

E. Solvent A, acetonitrile. Solvent B 0.02M sodium phosphate buffer pH 2.7 containing 0.10M sodium perchlorate.

(i) 40 to 70% solvent A in solvent B over 30 minutes.

(ii) 70% solvent B for ten minutes.

F. Solvent A, acetonitrile. Solvent B 0.02M sodium phosphate buffer pH 2.7 containing 0.10M sodium perchlorate.

(i) 50 to 100% solvent A in solvent B over 30 minutes.

(ii) 100% solvent B for ten minutes.

G. Solvent A, 40% acetonitrile in 60% aqueous formic acid. Solvent B, 60% aqueous formic acid.

(i) 10 to 60% solvent A in solvent B over 30 minutes.

(ii) 60% solvent B for 5 minutes.

Ion exchange chromatography was carried out on using Whatman Partisil-10/25 SAX anion exchange and Partisil-10/25 SCX cation exchange columns with isochratic elution using 2% acetic acid and 0.1M ammonium acetate respectively.

The above systems proved totally unsuitable for the purification of PRE A/WSN/33 HA with no product being eluted from the column.

4.4.11 Circular Dichroism (CD)

Far UV CD spectra were recorded on a JASCO J600 Spectrophotometer in the region 260-190nm using a quartz cell of path length 0.1mm at 27°C The data are presented as molar ellipticities. The spectra were analysed using the method of Greenfield and Fasman[188] and by the CONTIN program[189].

A solution of the PRE A/WSN/33 HA was prepared by dissolving the peptide (1mg) in HFIP (1cm^3). A series of CD spectra were recoreded in a variety of alcoholic and aqueous solutions such that the concentration of the peptide was 0.10 mgml^{-1}. 1.20cm^3 of each of the following solutions were used for each spectrum.

A. HFIP 100%.

B. 10% HFIP in methanol.

C. 10% HFIP in butanol.

D. 10% HFIP in TFE.

E. 50% HFIP in water.

F. 50% HFIP in 0.6% SDS.

Results.

Solution	GF helix	Contin helix	Contin sheet
A	9%	9%	45%
B	27%	0%	100%
C	65%	91%	9%
D	44%	46%	54%
E	7%	25%	75%
F	61%	58%	42%

References

1. H.W. Kosterlitz, H.O.J. Collier and J.C. Villarreal, "Agonist and Antagonist Actions of Narcotic Drugs", Macmillan, London, 1972.
2. H.W. Kosterlitz and A.A. Waterfield, *Annu. Rev. Pharmacol. Toxicol.*, 1975,**15**, 29.
3. G. Henderson, J. Hughes and H.W. Kosterlitz, *Br. J. Pharmocology*, 1972, **46**, 764.
4. C.B. Pert and S.H. Snyder, *Science*, 1973, **179**, 1011.
5. E.J. Simon, J.M. Hiller and I. Edelman, *Proc. Natl. Acad. Sci. USA*, 1973, **70**, 1947.
6. L. Terenius, *Acta. Pharmacol. Toxicol.*, 1973, **32**, 317.
7. K. Beaumont, W.S. Chilton, H.I. Yamamura and S.J. Enna, *Brain Research*, 1978, **148**, 153.
8. H.I. Yamamura, S.J. Enna and M.J. Kuhar, "Neurotransmitter Receptor Binding", Raven Press, New York, 1978.
9. L. Terenius and A. Wahlstrom, *Acta Physiol. Scand.*, 1975, **94**, 74.
10. J. Hughes, *Brain Resaerch*, 1975, **88**, 295.
11. J. Hughes, *Life Science*, 1974, **16**, 1753.
12. J. Hughes, L.W. Smith, H.W. Kosterlitz, L.A. Fothergill, B.A. Morgan and H.R. Morris, *Nature*, 1975, **258**, 577.
13. R. Simatov and S.H. Snyder, *Proc. Natl. Acad. Sci. USA*, 1976, **73**, 2515
14. A.F. Bradbury, D.G. Smyth and C.R. Snell, *Biochem. Biophys. Res. Comm.*, 1976, **69**, 950.
15. N. Ling, R. Burgus and R. Guillemin, *Proc. Natl. Acad. Sci. USA*, 1976, **73**, 3942.
16. A. Goldstein, W. Fischli, L.I. Lowney, M. Hunkapiller and L. Hood,

Proc. Natl. Acad. Sci. USA, 1981, **78**, 7219.

17. L. Polgar, "Mechanisms of Protease Action", CRC Press Inc., Florida, 1990.
18. B.M. Cox, A. Goldstein and C.H. Li, *Proc. Natl. Acad. Sci. USA*, 1976, **73**, 1821.
19. J. Rossier, T.M. Vargo, S. Minnick, N. Ling, F.E. Bloom and R. Guilleman, *Proc. Natl. Acad. Sci. USA*, 1977, **76**, 5162.
20. H.-Y.T. Yang, W. Fratta, J.S. Homg, A. DiGiulio and E. Costa, *Neuropharmacology*, 1978, **17**, 433.
21. R.V. Lewis, A.S. Stern, J. Rossier, S. Stein and S. Udenfriend, *Biochem. Biophys. Res. Comm.*, 1979, **88**, 822.
22. J. Hughes, *British Med. Bull.*, 1983, **39**, 17.
23. B.J. Pleuvry, *British J. Anaesthesia*, 1991, **66**, 370.
24. H. Imura, Y. Kato, Y. Nakai, K. Nakao, I. Tanaka, H. Jingami, T. Koh, T. Yoshimasa, M. Suda, M. Sakamoto, N. Sakamoto, N. Morii, H. Takahashi, K. Tojo and A. Sugawara, *J. Endocrynology*, 1985, **107**, 147.
25. J.S. Shaw and M.J. Turnbill, *Eur. J. Pharmocology*, 1978, **49**, 313.
26. A.Z. Ronai, I. Berzeti and S. Bajusz, *Eur. J. Pharmocolgy*, 1977, **45**, 313.
27. W.R. Martin, C.G. Eades, J. Thompson, R.E. Huppler and P.E. Gilbert, *J. Pharmacol. Exp. Thr.*, 1976, **197**, 517.
28. J.A.H. Lord, A.A. Waterfield, J. Hughes and H.W. Kosterlitz, *Nature*, 1977, **267**, 495.
29. G.M. Loewe and S.K. Burt, *Proc. Natl. Acad. Sci. USA*, 1978, **75**, 7.
30. Y. Audigier, H. Marzarguil, R. Gout and J. Cross, *Eur. J. Pharmacol.*, 1980, **63**, 35.

31. S.J. Paterson, L.E. Robson and H.W. Kosterlitz, *British Med. Bull.*, 1983, **39**, 31.

32. L. Terenius, A. Wahlstrom, G. Lindberg, S. Karlson and V. Ragnarson, *Biochem. Biophys. Res. Comm.*, 1976, **71**, 175.

33. N. Ling and R. Guillemin, *Proc. Natl. Acad. Sci. USA,* 1976, **73**, 3308.

34. P.W. Schiller and J.S.C. Hilaire, *J. Med. Chem.*, 1980, **23**, 290.

35. G. Gacel, M.C. Fournie-Zaluski, E. Fellion, B.P. Roques, B. Senault, J.M. Lecomte, B. Malfroy, J.-P. Swerts and J.C. Schwartz, *Life Science,* 1979, **24**, 725.

36. D.H. Coy, A.J. Kastin, A.V. Schally, O. Morin, N.G. Caron, F. Larbie, J.M. Walker, R. Fertel, G.G. Berntson and C.A. Sandman. *Biochem. Biophys. Res. Comm.*, 1976, **73**, 632.

37. B.A. Morgan, C.F.C. Smith, A.A. Waterfield, J. Hughes and Kosterlitz, *J. Pharm. Pharmacol.*, 1976, **28**, 660.

38. J.-K. Chang, B.T.W. Fong, J. Leighton and P. Cuaterecasas, *Life Science*, 1978, **28**, 660.

39. J.M. Hambrook, B.A. Morgan, M.J. Rance and C.F.C. Smith, *Nature*, 1976, **262**, 782.

40. A.S. Dutta, J.J. Gormley, C.F. Hayward, J.S. Morley, J.S. Shaw, G.J. Stacey and M.T. Turnbill, *Life Science*, 1977, **21**, 559.

41. C.R. Beddell, R.B. Clark, G.W. Hardy, L.A. Lowe, F.B. Ubatuba, J.R. Vane, S.W. Wilkinson, K.J. Chang, P. Cuaterecasas and R.J. Miller, *Proc. Royal Soc. B*, 1977, **198**, 249.

42. N.S. Agarwal, V.J. Hruby, R. Katz, W. Klee and M. Nirenberg, *Biochem. Biophys. Res. Comm.*, 1977, **76**, 129.

43. N. Ling, S. Minick, L. Lazarus, J. Rivier and R. Guillemin,

"Peptides", Proceedings of the 5th American Peptide Symposium, ed M. Goodman and J. Meienhofer, John Wiley & Son's, New York, 1977.

44. J.V. Castell, A.N. Eberle, V.M. Kriwaczeck, A. Tunkyi, P.W. Schiller, K.Q. Do, P. Thane and R. Schwyzer, *Helv. Chim. Acta,* 1979, **62**, 525.

45. A.Z. Ronai, J.I. Szekely, I. Berzetei, E. Miglecz and S. Bujusz, *Biochem. Biophys. Res. Comm.*, 1979, **91**, 1239.

46. Y. Audigier, H. Mazaguil and J. Gros, *FEBS Lett.*, 1980, **110**, 88.

47. D. Hudson, R, Sharpe and M. Szelke, Int. *J. Pept. and Protein Research*, 1980, **15**, 122.

48. M.T. Cox, D.W. Heaton and J. Horbory, *J. Chem. Soc.,Chem. Comm.*, 1980, 799.

49. V. Brantl, H. Teschenmacher, A. Henschen and F. Lottspeich, *Hoppe-Seylers Z. Phsiol. Chem.*, 1979, **360**, 1211.

50, P.C. Montecucchi, R. de Catiglione, S. Piani, L. Gazzini and V. Erspamer, *Int. J. Pep. Protein Res.*, 1981, **17**, 275.

51. Y. Kiso, T. Miyazaki, M. Satomi, M. Inui, T. Akita, H. Moritoki, M. Takei and H. Nakamura, "Peptide Chemistry 1981", ed T. Shiori, Protein Res. Foundation, Osaka, Japan, 1981.

52. J.-K. Chang, A. Killian, E. Hazum, P. Cauteracasus and J.-K. Chang, *Science*, 1981, **212**, 75.

53. A.R. Jacobson, A.R. Gintzler and L.M. Sayre, *J. Med. Chem.*, 1989, **34**, 1708.

54. E. Roubini, R. Laufer, C. Gilon, Z. Selinger, B.P. Roques and M. Chorev, *J. Med. Chem.*, 1991, **34**, 2430.

55. A.S. Horn and J.R. Rodgers, *Nature*, 1976, **260**, 795.

56. A.F. Bradburry, D.G. Smyth and C.R. Snell, *Nature*, 1976, **260**, 165.

57. Y. Isogai, G. Nemethy and H.A. Scheraga, *Proc. Natl. Acad. Sci. USA*, 1977, **74**, 414.

58. J.L. de Coen, C. Humblet and M.H.J. Koch, *FEBS Lett*, 1977, 73, 38.

59. G.H. Loew and S.K. Burt, *Proc. Natl. Acad. Sci. USA*, 1978, **75**, 7.

60. S. Permilat and B. Maigret, Biochem. Biophys. Res. Comm., 1979, 91, 543. *J. Phys. Chem.*, 1980, **84**, 293.

61. G.D. Smith and J.F. Griffin, *Science*, 1978, **199**, 1214.

62. T. Ishida, M. Kenmotsu, Y. Mino, M. Inoue, T. Fujiwara, K. Tomita, T. Kimura and S. Sakakibara, *Biochem. J.*, 1984, **218**, 677.

63. I.L. Karle, J. Karle, D. Mastropaolo, D. Camerman and N. Camerman, *Acta Crystalogr., Sect. B*, 1983, **39**, 625.
A. Camerman, D. Mastropaolo, I. Karle, J. Karle and N. Camerman, *Nature*, 1983, **306**, 447.

64. M. Doi, T. Ishida, M. Inoue, T. Fujiwara, K. Tomita, T. Kimura and S. Sakakibara, *FEBS Lett.*, 1984, **170**, 229.

65. P.W. Schiller in "The Peptides", p220, ed. S. Udenfriend and J. Meienhofer, Academic Press Inc., London, 1984.

66. F.A. Gorin, T.M. Balasibramanian, C.D. Barry and G.R. Marshall, *J. Supramol. Structure*, 1978, **9**, 27.

67. A. Motta, D. Picone, T. Tancredi and P. Temussi, *Tetrahedron*, 1988, **44**, 975.

68. G. Gupta, M.H. Sarma, R.H. Sarma and M.M. Dhingra, *FEBS Lett.* 1986, **198**, 245.

69. M.C. Fournie-Zaluski, T. Prange, C. Pascard and B.P. Roques,

Biochem. Biophys. Res. Commun. 1977, **79**, 1199.

70. B.P. Roques, C. Garbay-Jaureguiberry, R. Oberlin, M. Anteunis and A.K. Lala, *Nature*, 1976, **262**, 778.

71. M.A. Khaled, M.M. Long, W.D. Thompson, R.J. Bradley, G.B. Brown and D.W. Urry, *Biochem. Biophys. Res. Commun.*, 1977, **76**, 224.

72. A. Milon, T. Miyazawa and T. Higashjima, *Biochemistry*, 1990, **29**, 65.

73. B. Vesterman, J. Saultis, B. Betins, E. Liepins and G.V. Nikiforvich, *Biochim. Biophys. Acta,* 1989, **998**, 204.

74. T. Miyazawa and T. Higashijima, *Biopolymers*, 1981, **20**, 1949.

75. S. Combrisson, B.P. Roques and R. Oberlin, *Tet. Lett.*, 1976, **38**, 3455.

76. C. Garbay-Jaureguiberry, B.P. Roques, R.B. Oberlin, M. Anteunis, S. Combrisson and J.Y. Lallemand, *FEBS Lett.*, 1977, **76**, 93.

77. C. Garbay-jaureguiberry, J. Baudet, D. Florentin and B.P. Roques, *FEBS Lett.*, 1980, **115**,1980.

78. E. Moret, I.P. Gerothanassis, R.N. Hunston and J. Lauterwein, *FEBS Lett.*, 1990, **262**, 173.

79. R. Ramage, D. Hopton, D. Parrot, M.J. Parrott, G.W. Kenner and G.A. Moore, *J. Chem. Soc., Perkin Trans. 1,* 1985, 461.

80. A.A. Bothner-By, R.L. Stephens, J. Lee, C.D. Warren and R.W. Jeanloz, *J. Amer. Chem. Soc.*, 1984, **106**, 811.

81. A. Bax and D.G. Davis, *J. Magn. Res.*, 1985, **63**, 207.

82. K. Wuthrich, M. Billeter and W. Braun, *J. Mol. Biol.*, 1984, **180**, 715.

83. V.F. Bystrov, *Prog. in NMR Spectroscopy*, 1976, **10**, 41.

84. V.F. Bystrov, V.T. Ivanov, S.L. Portnova, T.A. Balashova and Y.A. Ovichinnikov, *Tetrahedron*, 1973, **29**, 873.

85. M. Barfield and D.M. Grant, *J. Amer. Chem. Soc.*, 1963, **85**, 1899.

86. R.C. Cookson, J.J. Fraenkel, J. Hudec and T.A. Crabb, *Tetrahedron Suppl.*, 1966, **7**, 355.

87. J. Feeney, *J. Magn. Res.*, 1976, **21**, 473.

88. K.G.R. Pachler, *Spectrochimika Acta*, 1964, **20**, 581.

89. W.B. Ober, *New York State J. Med.*, 1977, **77**, 254.

90. W. Kemp, "NMR in Chemistry : a multinuclear approach", Macmillan, London, 1986.

91. R.L. Baxter, A.B. Hanley and H.W.S. Chan, *J. Chem. Soc., Perkin Trans. 1*, 1990, 2963.

92. A.G. McInnes and J.L.C. Wright, *Acc. Chem. Res.*, 1975, 8, 313.

93. T.J. Simpson, *Chem. Soc. Rev.*, 1975, **4**, 497.

94. M.J. Garson and J. Staunton, *Chem. Soc. Rev.*, 1979, **8**, 539.

95. A. Nakagawa, T.-S. Wu, P.J. Keller, J.P. Lee, S. Omura and H.G. Floss, *J. Chem. Soc., Chem. Comm.*, 1985, 519.

96. R.L. Baxter, E.M. Abbot, S.L. Greenwood and I.J. McFarlane, *J. Chem. Soc. Chem. Comm.*, 1985, 564.

97. C.A. Townsend and G.M. Salituro, *J. Chem. Soc. Chem. Comm.*, 1984, 1631.

98. J.E. Baldwin, A.E. Derome, L.D. Field, P.T. Gallagher, A.A. Thah and V. Thaller, *J. Chem. Soc. Chem. Comm.*, 1981, 1227.

99. J.E. Baldwin, H.S. Bansal, J. Chondrogianni, L.D. Field, A.A. Thah and V. Thaller, *Tetrahedron*, 1985, **41**, 1931.

100. R.L. Baxter and A.I. Scott in "Comprehensive Heterocyclic Chemistry", Volume 1, ed. A.R. Katritzky and C.W. Rees, Pergamon Press, Oxford, 1984.

101. D.W. Young, *Natural Prod. Rep.*, 1986, **3**, 395.

102. G. Grue-Sorense, R.L. White and I.D. Spenser, *J. Amer. Chem. Soc.*, 1986, **108**, 146.

103. K. Tazuya, K. Yamada and H. Kumaoka, *Biochim. Biophys. Acta*, 1989, **990**, 73.

104. R.L. White and I.D. Spenser, *J. Amer. Chem. Soc.*, 1982, **104**, 4934.

105. R.H. White, *Biochemistry*, 1978, **17**, 3833.

106. B. Estramareix and M. Therison, *Biochim. Biophys. Acta*, 1972, **273**, 275. E. Bellion, D.H. Kirkley and J.R. Faust, *Biochim. Biophys. Acta*, 1976, **427**, 229. R.H. White and F.B. Rudolph, *Biochim. Biophys. Acta*, 1978, **542**, 340.

108. J. Lamb and W. Robson, *Biochemical Journal*, 1931, **25**, 1231.

109. V.K.P. Unny, S. Thyagaajan and K.V. Viswanathan, Radiochem. *Radioanal. Lett.*, 1981, **47**, 367.

110. R.B. Lottfield, *J. Amer. Chem. Soc.* 1950, **72**, 2499.

111. G.R. Clemo, F.K. Duxbury and G.A. Swan, *J. Chem. Soc.*, 1952, 3464.

112. J.E. Baldwin, H.S. Bansal, J. Chondrogianni, P.T. Gallagher, A.A. Taha, A. Taylor and V. Thaller, *J. Chem. Res.(m)*, 1984, 1812.

113. J.E. Baldwin, S.C. Ng, A.J. Pratt, M.A. Russell and R.L. Dyer, *Tetrahedron Lett.*, 1987, **28**, 2303.

114. C. Halldin and B. Laangstroem, *Int. J. Radiat. Isot.*, 1984, **35**, 779.

115. V. Viswanathan and V.J. Hruby, *J. Org. Chem.*, 1979, **44**, 2892.

116. H.R. Snyder, J.F. Shekleton and C.D. Lewis, *J. Amer. Chem. Soc.*, 1945, **67**, 310.

117. A.T. Shulgin and E.M. Gal, *J. Chem. Soc.*, 1953, 1316.

118. A.J. Zambito and E.E. Howe, "Collected Organic Synthesis", p 373, ed. H.E. Baumgarten, Wiley and Sons Inc, USA, 1973.

119. D. Murphy, personal communication.

120. R.L. Baxter and E.M. Abbot, *J. Labl. Cmpds. Radiopharm.*, 1985, **22**, 1211.

121. E. Pierson, M. Giella and M. Tishler. *J. Amer. Chem. Soc.*, 1948, **70**, 1450.

122. T.W. Greene, "Protective Groups in Organic Synthesis", John Wiley and Sons, New York, 1981.

123. S.H. Hedges and R.B. Herbert, *J. Chem. Res.(m)*, 1979, 410.

124. M. Freifelder and R.B. Hasbrouck, *J. Amer. Chem. Soc.*, 1960, **82**, 696.

125. C. Petrier, J.C.S. Barlosa, C. Dupay and J.-L. Luch, *J. Org. Chem.*, 1985, **50**, 5761.

126. T. Hanafusa, J. Ichihara and T. Ashida, *Chem. Lett.*, 1987, 687.

127. K. Mai and G. Patil, *Tetrahedron Lett.* 1984, **25**, 687.

128. H.R. Ing and R.H.F. Manske, *J. Chem. Soc.*, 1926, 2348.

129. Y. Kiso, H. Isawa, K. Kitagawa and T. Akita, *Chem. Pharm. Bull.*, 1978, **26**, 2562.

130. "Selections from the Aldrichimica Acta", p.67, Aldrich Chemical Company, Wisconsin, 1984.

131. R.F. Borch, M.D. Bernstein and H.D. Durst, *J. Amer. Chem. Soc.*, 1971, **93**, 2897.

132 H.E.. Carter, *Org. Reactions*, 1946, **3**, 198.

133. E. Baltazzi, *Quart. Rev., Chem. Soc.*, 1955, **9**, 150.

134. "Virology" , volume 1, second edition, ed. B.N. Fields and D.M. Knipe, Raven Press, New York, 1990.

135. "The Influenza Viruses". ed. R.M. Krug, Plenum Press, London, 1989.

136. F.I. Smith and P. Palase in ref. 135, page 319.

137. K. Simons, H. Garoff and A. Helenius, *Scientific American*, 1982, **246**, 46.

138. D.C. Wiley and J.J. Skehel in ref. 134, page 63.

139. I.A. Wilson, J.J. Skehel and D.C. Wiley, *Nature*, 1981, **289**, 366.

140. D.W. Kingsbury in ref. 134. page 1075.

141. A. Puglsey, "Protein Targeting", Academic Press Inc., San Diego, 1989.

142. C. Milsten, G.G. Brownlee, T.M. Harrison and M.B. Mathews, *Nature*, 1972, **239**, 117.

143. G. Blobel and B. Doberstein, *J. Cell Biol.*, 1975, **67**, 835.

144. G. Blobel and B. Doberstein, *J. Cell Biol.*, 1975, **67**, 852.

145. G. von Heijne, *Eur. J. Biochem.*, 1983, **133**, 17.

146. G. von Heijne, *J. Mol. Biol.*, 1986, **184**, 99.

147. G. von Heijne, *J. Mol., Biol.*, 1986, **192**, 287.

148. M.S. Briggs and L.M. Gierash, *Adv. Protein Chem.*, 1985, **37**, 109.

149. T.A. Rapoport, *CRC Crit. Rev. Biochem.*, 1986, **41**, 73.

150. L.L. Randal, S.J.S. Hardy and J.R. Thom, *Annu. Rev. Microbiol.*, 1987, **41**, 507.

151. L.M. Gierash, *Biochemistry*, 1989, **28**, 923.

152. P. Walter and G. Blobel, *Proc. Natl. Acad. Sci. USA*, 1980, **77**, 112.

153. P. Walter and G. Blobel, *J. Cell Biol.*, 1981, **91**, 551.

154. T.V. Kurchalia, M. Wiedman, A.S. Girshovich, E.S. Bochkareva, H. Bielka and T. Papoport, *Nature*, 1986, **320**, 634.

155. V.C. Kreig, P. Walter and A.E. Johnson, *Proc. Natl. Acad. Sci. USA*, 1986, **83**, 8604.

156. P. Walter and G. Blobel, *J. Cell Biol.*, 1982, **95**, 453.

157. R. Gilmore, P. Walter and G. Blobel, *J. Cell Biol.*, 1982, **95**, 453.

158. R. Gilmor and G. Blobel, *Cell*, 1983, **35**, 677.

159. D.I. Meyer, D. Louvard and B. Dobberstein, *J. Cell Biol.*, 1982, **92**, 579.

160. R. Gilmore, G. Blobel and P. Walter, *J. Cell Biol.*, 1982, **95**, 463.

161. R. Gilmore, G. Blobel and P. Walter, *J. Cell Biol.*, 1982, **95**, 470.

162. L. Lauffer, P.D. Garcia, R.N. Harkins, C. Coussens, A. Ullrich and P. Walter, *Nature*, 1985, **318**, 334.

163. S. Tajima, L. Laufer, V.L. Rath and P. Walter, *J. Cell Biol.*, 1986, **103**, 1167.

164. M. Hortsch and D.I. Meyer, *Biochem. Biophys. Res. Commun.*, 1988, **150**, 111.

165. R. Gilmore and G. Blobel, *Cell*, 1985, **42**, 497.

166. E.A/ Evans, R. Gilmore and G. Blobel, *Proc. Natl, Acad. Sci. USA*, 1986, **83**, 581.

167. D.M. Engelman and T.A. Steitz, *Cell*, 1981, **23**, 411.

168. S.J. Singer, *Proc. Natl, Acad. Sci. USA*, 1987, **84**, 1015.

169. S. Simon and G. Blobel, *Cell*, 1991, **65**, 371.

170. M.G. Roth, M.-J. Gething and J. Sambrook in ref. 135, page 219.

171. E. Rodrriguez-Boulan and M. Pendergras, *Cell*, 1980, **20**, 45.

172. E. Rodreguez-Boulan and D.D. Sabatini, *Proc. Natl, Acad. Sci. USA*, 1978, **75**, 5071.

173. M.G. Roth, R.W. Compans, L. Giusti, A.R. Davis, D. Nayak, M.-J. Gething and J. Sambrook, *Cell*, 1983, **33**, 435.

174. M.S. Briggs and L.M. Gierash, *Adv. in Protein Chemistry*, 1985, **37**, 109.

175. A.L. Hiti, A.R. Davis and D.P. Nayak, *Virology*, 1981, **111**, 113.

176. R.B. Merrifield, *Fed. Proc. Fed. Amer. Soc. Expl. Biol.*, 1962, **21**, 412.

177. R.B. Merrifield, *J. Amer. Chem. Soc.*, 1962, *85*, 2149.

178. For an introduction to peptide synthesis see M. Bodansky, "Peptide Chemistry", Springer-Verlag, Berlin, 1988.

179. L.A. Carpino and G.Y. Han, *J. Org, Chem.*, 1972, **37**, 3404.

180. S.S. Wang, *J. Amer. Chem. Soc.*, 1973, **95**, 1328.

181. G.B. Fields and R.L. Noble, *Int. J. Pept. Protein Res.*, 1990, **35**, 161.

182. K. Haverson, P.E. Fraser, D.A. Kirschner and P.T. Lansbury Jr., *Biochemistry*, 1990, **29**, 2639.

183. J. Heukeshoven and R. Dernick, *J. Chromatography*, 1985, **326**, 91.

184. A.E. Shinnar and E.T. Kaiser, *J. Amer. Chem. Soc.*, 1984, **106**, 5006.

185. R. Katakai and Y. Iizuka, *J. Amer. Chem. Soc.*, 1984, **106**, 5715.

186. M. Rosenblatt, J.F. Harbener, G.A. Tyler, G.L. Shepard and J.T. Potts Jr., *J. Biol. Chem.*, 1979, **254**, 1414.

187. P.Y. Chou and G.D. Fasman, *TIBS*, 1977, 128.

188. N. Greenfield nad G.D. Fasman, *Biochemistry*, 1969, **8**, 1969.

189. S.W. Provenchar and J.D. Glockner, *Biochemistry*, 1981, **20**, 33.

190. O. Tsoga, M. Noguch and H. Moriyama, *Heterocycles*, 1981, **16**, 209.

Courses Attended

Organic Chemistry Research Seminars, various speakers, Department of Chemistry, University of Edinburgh, 1988-91.

Medicinal Chemistry, Prof, R. Baker and colleagues, Merck, Sharp and Dohme Ltd., Terlings Park, UK. 1988, 1990.

11th Irvine Memorial Lectures, "Asymmetry in Chemistry", various speakers, University of St. Andrews, 1988.

Advances in Organic Chemistry, various speakers, Department of Chemistry, University of Edinburgh, 1989-1991.

Industrial Chemistry, Members of ICI Grangemouth and Department of Chemical Engineering, University of Edinburgh, 1990.

Medicinal Chemistry, Members of SmithKline Beecham Pharmaceuticals Ltd., Brockham Park, UK, 1990.

2D Nuclear Magnetic Resonance, Dr's R.L. Baxter, I.H. Sadler and B. Birdsall, University of Edinburgh, 1989.

RSC Bio-organic Group, "Bio-organic Mechanisms", University of Newcastle, Newcastle, England, July, 1989.