Synthesis of Peptides by the Solid Phase Method

Peter L. Roach

Doctor of Philosophy

University of Edinburgh

1990

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 submitted, in whole or in part, for any degree at this or any other university.

University of Edinburgh

October 1990

Acknowledgements.

I would like to thank Professor R. Ramage for his patience, guidance and encouragment throughout the course of this work.

I wish to thank Mr. K. Shaw and Mr. B. Whigham for their technical assistance. I would like to express my thanks to those involved in the departmental analytical and spectroscopic services for their efficient work. I also thank Laporte Fine Organics for provision of a research grant.

I would like to thank Drs. C. Bladon, T. Brown, J. Parkinson and D. W. Thomas and Mr. D. Camp for proof reading and helpful discussions. Finally, I would like to thank my family, friends and colleagues who helped to make my time in Edinburgh so enjoyable.

To my parents, Peter and Lucille
and my sister, Sherry.

ABSTRACT

An investigation into the use of a novel N^{α} protecting group in peptide synthesis is described. The protecting group, 2,2-*bis*(4'-nitrophenyl)ethoxycarbonyl (Bnpeoc), has been introduced into the full range of amino acids, including those requiring appropriate side chain protection.

The Bnpeoc amino acid derivatives have been successfully applied to the solid phase synthesis of a number of peptides from ubiquitin. In addition, the peptides were assembled using a wide range of coupling reagents and conditions.

The chemical syntheses of phage λ Cro protein and peptides from its DNA binding region have also been investigated, using both the novel N^{α}-Bnpeoc protecting group and the established N^{α}-fluorenylmethoxcarbonyl (Fmoc) protecting group.

N.m.r. experiments have indicated that residues corresponding to the DNA recognition α-helix of the Cro protein may adopt a similar conformation in a peptide from that region.

Contents

Abbreviations.

Bnpeoc	2, 2-*bis*(4'-Nitrophenyl)ethoxycarbonyl
Boc	Butyloxycarbonyl
BOP	Benzotriazole-oxy-tris-(dimethylamino)-phosphonium hexafluorophosphate
Bpoc	Biphenylpropyloxycarbonyl
Cbz, Z	Benzyloxycarbonyl
CM	Carboxymethyl
COSY	Correlated spectroscopy
Cro	Control of Repressor and Other proteins
DBN	Diazabicyclo-(4.3.0)-non-5-ene
Dbs	Dibenzosuberanyl
Dbse	Dibenzosuberenyl
DBU	Diazabicyclo-(5.4.0)-undec-7-ene
DCM	Dichloromethane
DIC	Diisopropylcarbodiimide
DMAP	N,N-Dimethylaminopyridine
DMF	N,N-Dimethylformamide
DMSO	Dimethylsulphoxide
EDT	1,2-Ethanedithiol
ex.	Exchangable with solvent
F.A.B.	Fast atom bombardment
Fmoc	Fluorenylmethoxycarbonyl
HBTU	2-(1-Hydroxy-1,2,3-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HFIP	Hexafluoroisopropanol
h.p.l.c.	High performance liquid chromatography
hrms	High resolution mass spectrometry
HOBt	1-Hydroxy-1,2,3-benzotriazole
HTH	Helix-turn-helix
Mbh	4,4'-Dimethoxybenzhydryl
Mtr	Methoxytrimethylbenzene sulphonyl
NMM	N-Methylmorpholine
nOe	Nuclear Overhauser enhancement
NOESY	Nuclear Overhauser Enhancement Spectroscopy
PAM	Phenylacetamidomethyl

Pmc	2,2,5,7,8-Pentamethylchroman-6-sulphonyl
Sasrin	Super acid sensitive resin
SDS PAGE	Sodium dodecyl sulphate poylacrylamide gel electrophoresis
SPPS	Solid phase peptide synthesis
TBTU	2-(1-Hydroxy-1,2,3-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate
TEA	Triethylamine
TFA	Trifluoroacetic acid
THF	Tetrahydrofuran
TOCSY	Total correlated spectroscopy
Trt	Trityl
UV	Ultra-violet
Cha	Cyclohexylamine
Dcha	Dicyclohexylamine.

Glossary.

Conventionally genes are indicated by an italic abbreviation (e.g. *Cro, cI, Antp*). The protein products of these genes are indicated by the use of Roman letters (e.g. Cro, Iac, Antp).

Bacteriophage	A virus which infects bacteria.
Lysogeny	The incorporation of viral DNA into the host bacterial chromosome
Lytic Cycle	The reproduction of many viral particles until the bacterial host cell bursts.
Operator	An area of DNA to which one or more repressor proteins may bind in order to prevent transcription (and hence expression) of a gene.
Operator Site	The small region(s) of DNA within an operator to which the repressor binds.
Promoter	A region of DNA to which RNA polymerase binds to initiate transcription.
Repressor	A protein which binds to DNA in such a way as to reduce the expression of a gene

1. Introduction

1.1 Solid Phase Peptide Synthesis

1.1.1 Historical perspective.

The diverse range of functional groups found in peptides and proteins, often arrayed in complex structures, makes them ideally suited to the enormous range of biological functions which they perform. Yet it is this diversity of functional groups that has challenged the organic chemist with many interesting problems since the turn of the century.

Theodor Curtius (1857-1928) was perhaps the first to synthesise a peptide derivative, hippurylglycine[1], but it was the pioneering work of Emil Fisher (1852-1919) which led to the isolation of the first free peptide[2], glycylglycine, in 1901. Some six years later, Fisher was able to report[3] the preparation of a sixteen residue peptide. All this was achieved at a time when the nature of peptides and particularly proteins was still in some doubt.

In his original work, Fisher had used ethylchloroformate to protect the amino function as a urethane. Urethanes remain the principle method of amine protection during peptide synthesis to this day. However, his efforts to remove this protection were unsuccessful. The next major step forward in peptide synthesis was to be the development of a reversible amino protecting group, by Bergmann and Zervas[4] in 1932. They described the preparation of a number of N^{α}-benzyloxycarbonyl amino acid derivatives which could readily be deprotected by hydrogenolysis with palladium on charcoal as a catalyst. This has been applied to a wide range of solution syntheses with great success[5,6].

Much of the early work had been directed towards the synthesis of proteins, but it was the emerging nature of the peptide hormones and antibiotics that provided organic chemists in the late 1940s and 1950s with viable synthetic targets. The isolation, structure and total synthesis of oxytocin, a lactogenic nonapeptide amide hormone, by du Vigneaud *et al.* [5] showed the great potential of the solution phase synthesis of peptides. However, the isolation of increasing numbers of peptide hormones and the increasing complexity of the sequences began to indicate the limitations of solution synthesis: the syntheses were both laborious and time consuming.

In 1963, with these problems clearly in mind, Merrifield introduced[7] solid phase peptide synthesis. Within a year, he reported[8] the synthesis of bradykinin, a nonapeptide amide, which was prepared in an overall yield of 32% in only eight days. This was much faster than was possible by solution methods. Merrifield began by attaching an N^{α}-protected amino acid derivative to a polystyrene based solid support, through the carboxylic acid group, then deprotecting the N terminus and coupling the next N^{α}protected amino acid to it. He then extended the chain, one residue at a time, in a stepwise process of activation, coupling and deprotection. The concept revolutionised peptide synthesis and had a profound effect on the development of methods to synthesise oligonucleotides.

1.1.2 The Strategy of Solid Phase Peptide Synthesis.

The solution phase synthesis of peptides containing more than a few residues is often slow and painstaking work. The solid phase approach has a number of advantages:

1). It avoided the loses normally associated with the isolation and purification of all the intermediates in solution phase synthesis.

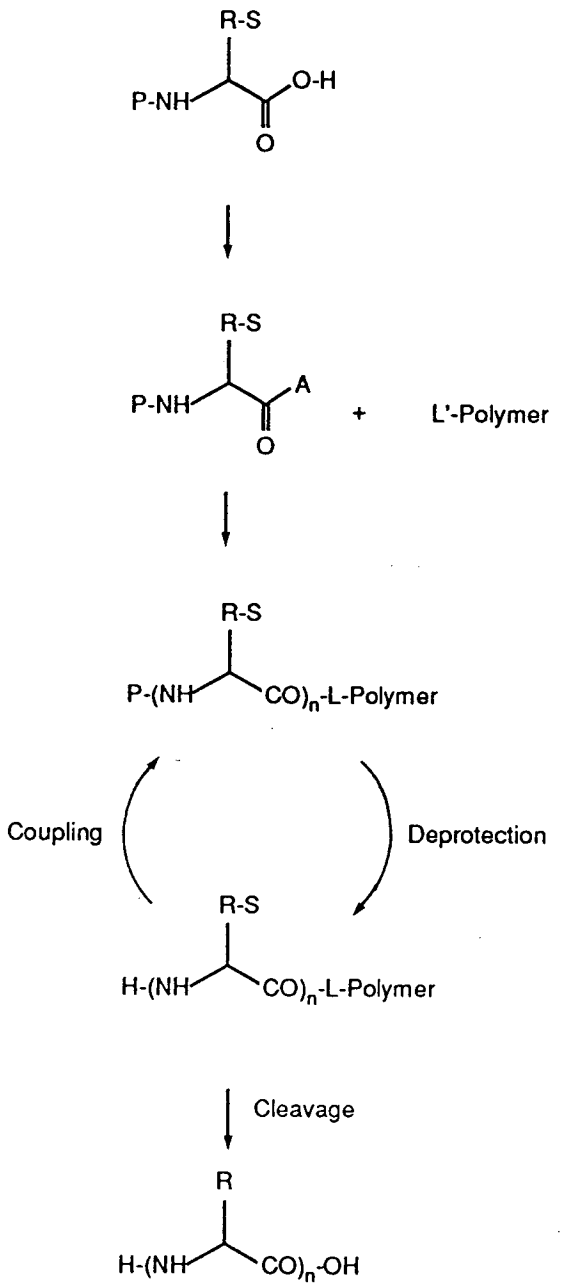
2). Individual reactions can be forced to completion by the use of large excesses of the reagent. These can be easily removed by washing the peptidyl resin.

3). The repetitive synthetic cycle lends itself readily to automation; modern automated synthesisers assemble peptides much more rapidly, often adding one residue every three or four hours.

4). The use of a single vessel reduces mechanical losses which typically occur with the repeated transfer of material.

The strategy[9] is outlined in scheme 1 and descibed below. The insoluble support is a synthetic polymer or 'resin' which contains reactive groups (L'). The C-terminal N^{α}-protected residue is firstly activated by some appropriate chemistry (A) and then coupled to the polymer. The N^{α} protecting group is chosen so that its subsequent removal does not interfere with either the link to the polymer (L) or the semi-permanant protecting groups (S) on functional groups of the amino acid side chains (R). Once the N^{α} protecting group has been removed, the next activated N^{α}-protected amino acid derivative is coupled. The cycle of deprotection and coupling is then repeated, increasing the number of residues (n) until the peptide chain is complete. The side chain protecting groups (S) may then be removed and the peptide resin link (L) cleaved. These are often chosen so that they are cleaved by the same reagent, but these steps may be carried out sequentially if the nature of the peptide resin link and the side chain protection permit.

There have been two major applications of this strategy. The first was the 'graded acidolysis' method which was used originally by Merrifield[8]. The side chain protection and peptide resin link were cleaved by strong acids (such as HF and trifluoromethanesulphonic acid) and for the N-terminal protection the Boc group was used and cleaved by TFA. The other principle method is the orthogonal strategy developed by Sheppard[10,11]. In this method the N-terminal protection is the Fmoc group[12] which is cleaved with piperidine in DMF. This does not effect the side chain protecting groups and peptide resin link which are both labile to milder acid (TFA). The concept of orthogonality, in which the conditions used to cleave each of more than one set of protecting groups does not affect the others, is of fundamental importance in this approach.



Scheme 1. The strategy of solid phase peptide synthesis.

1.1.3 The Solid Support

The solid support has a number of requirements[13]:

1). Reactive sites that provide an attachment point for the peptide, from which it can later be removed. This must be stable to the conditions of synthesis.

2). The support must allow rapid and unhindered contact between the growing peptide chain and the reagents. This is usually achieved by swelling the polymer matrix in an appropriate solvent.

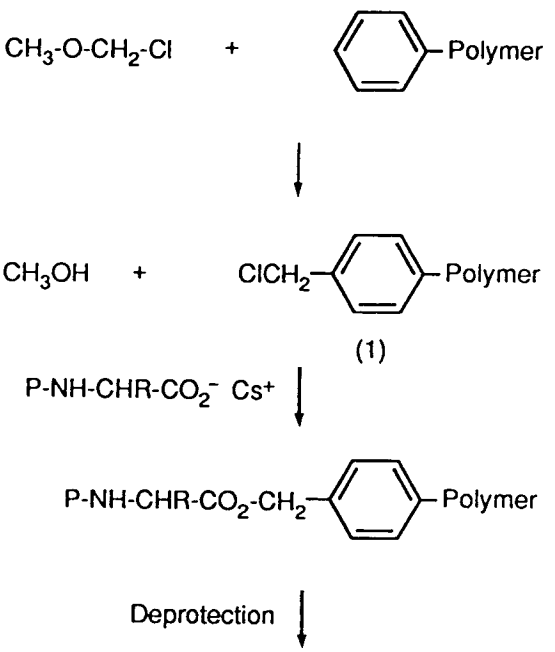
3). The polymer must be composed of particles that allow ready manipulation and rapid filtration of liquids.

The support selected by Merrifield was gelatinous polystyrene beads made of a co-polymer obtained from 2% divinylbenzene in styrene[7]. More cross linking (e.g. 8 or 16% divinylbenzene)[7] lead to rigid, less permeable beads, whilst less cross linking (e.g. 0.5% divinylbenzene)[14] leads to soft beads, for which filtration can no longer be used as a method of separation.

In solution synthesis, if the growing peptide chain is not adequately solvated it may aggregate and this often results in an insoluble precipitate. On the solid phase support, such aggregation is more difficult, but can still occur and causes the internal collapse of the resin. This led Sheppard[15] to develop the concept of a polar matrix and solvent. The excellent solvating properties of DMF were well known from solution phase syntheses and made it an excellent solvent for solid phase synthesis. For a solid phase he chose polyamide resins, which swell very well in DMF and have a 'peptide-like' character[16,17]. The use of polyamide resin as the solid support and DMF as the solvent has been successfully applied in commercially available automated synthesisers. Both polystyrene and polyamide supports have seen widespread use.

1.1.4 The peptide resin linkage.

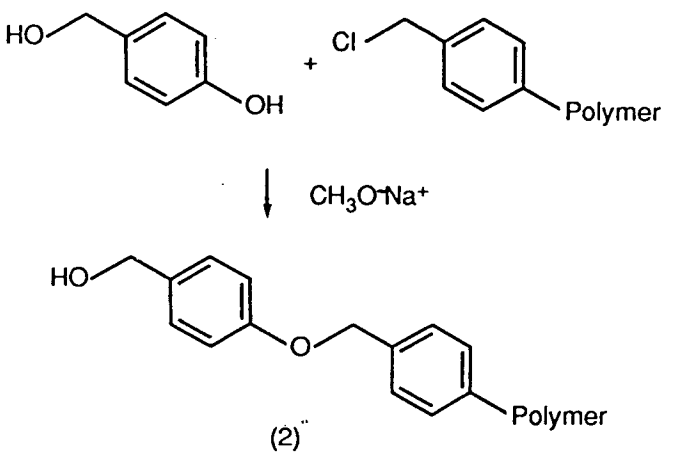
The most common method for attaching the peptide to the resin is through a benzyl ester. The earliest methods used by Merrifield involved the chloromethylation of the resin, as shown in scheme 2. The chloromethylation of the resin produces an arylchloromethyl derivative (1), which, upon reaction with the triethylammonium or caesium salt of the C-terminal amino acid derivative, gives a benzyl ester linkage.



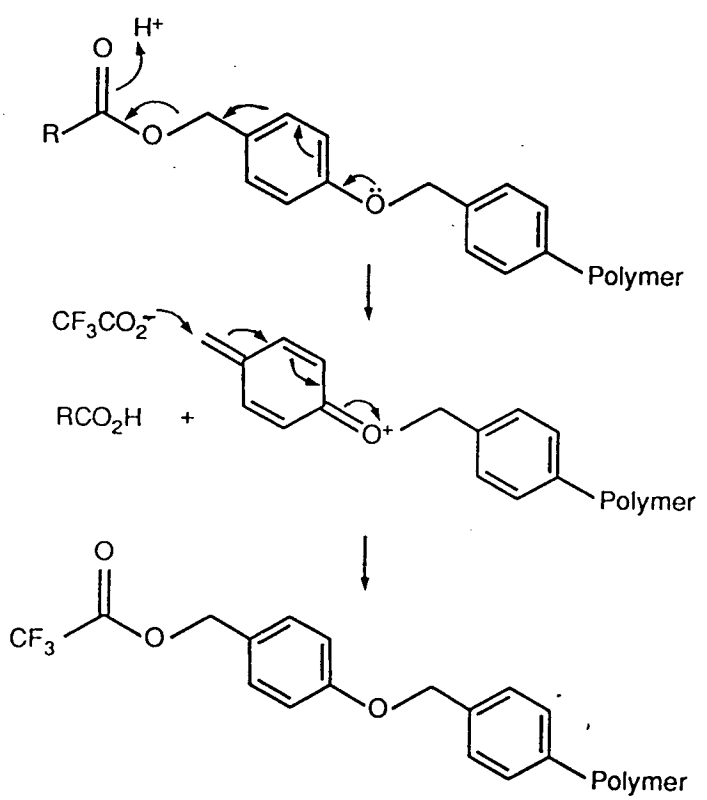
Scheme 2. Attachment of the first amino acid to the resin.

The N-terminal group used by Merrifield in the bradykinin synthesis[8] was *tert.*butyl-oxycarbonyl, which was removed by treating the peptidyl resin with 50% trifluoroacetic acid in dichloromethane. The benzyl ester was not entirely stable to these conditions and this led to a steady 'bleeding' of about 1.5% of the peptide from the resin per cycle. For bradykinin, with only nine cycles, this was acceptable, but as more ambitious syntheses were undertaken, this became increasingly significant. An electron withdrawing group introduced *para* to the benzyl ester should increase its acid stability and this was the principle behind the design of the phenylacet-amidomethyl (Pam) resin, developed by Merrifield *et al*.[18,19]

In an effort to circumvent the vigorous conditions required for cleavage and deprotection used in the Merrifield method, the very acid labile Bpoc[20] and the base labile Fmoc[12] amino terminal protecting groups were introduced. This required a linker of greater acid lability. The *p*-alkoxybenzyl alcohol resin (2) was originally developed by Wang[21] (scheme 3) for the synthesis of protected fragments. It has been widely adopted for use in the Fmoc strategy of solid phase peptide synthesis and the mechanism for its cleavage by trifluroacetic acid and water (95:5) is shown in scheme 4.

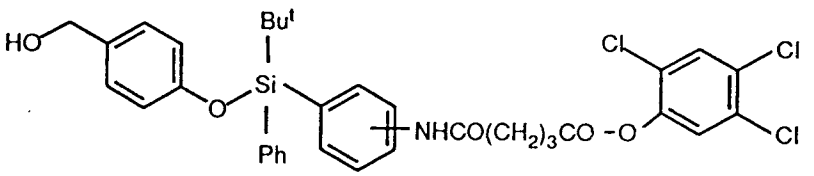


Scheme 3. The synthesis of the Wang linker[21].

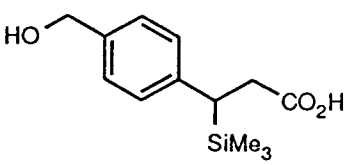


Scheme 4. The cleavage of the Wang linker by acid.

A further development has been the extension of orthogonality into a third dimension with the introduction of fluoride labile linkers. The linkers developed by Barany[22](3) and Ramage[23](4) rely on the attack on a silicon centre by fluoride and both linkers allow the cleavage of the protected peptide from the resin under neutral conditions (e.g. using tetrabutylammonium fluoride in DMF).

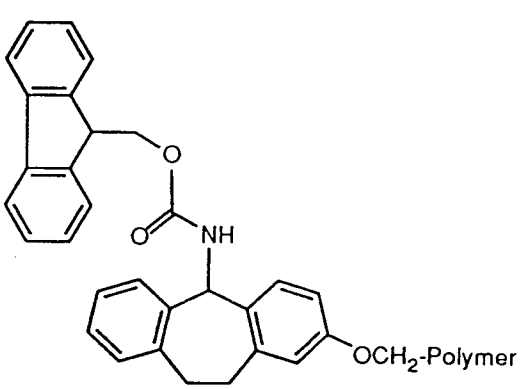


(3)



(4)

The preparation of C-terminal peptide amides requires the use of a different linker. Recently a number of examples of these have been developed for use in the Fmoc strategy[24,25,26]. The dibenzosuberanyl derivative (5) has been developed in this research group and applied to the synthesis of peptide amides[27].



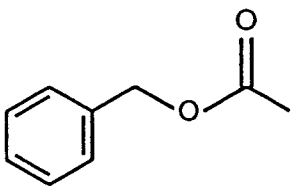
(5)

Figure 1. Linker for the preparation of peptide amides using the Fmoc strategy.

1.1.5 N^{α} protecting groups.

The N^{α} protecting groups which are most widely used in peptide synthesis are urethane derivatives, in which the urethane ester moiety is first cleaved, with subsequent loss of carbon dioxide from the resulting carbamate, giving the unblocked amine.

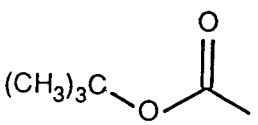
a). The benzyloxycarbonyl (Z or Cbz) group.



(6)

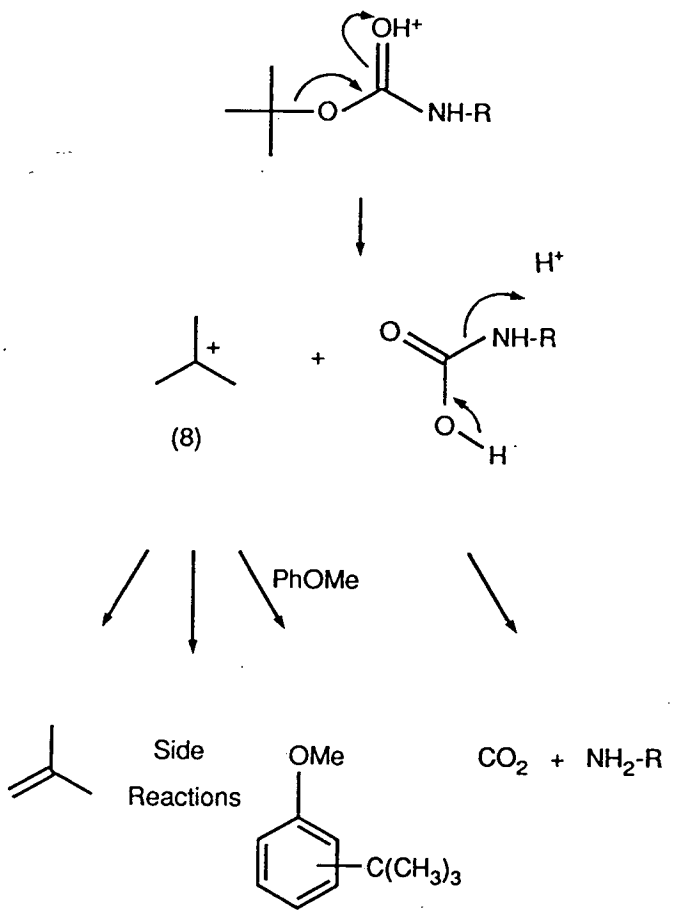
The first reversible N^{α} protecting group was the benzyloxycarbonyl or Z protecting group (6), which was introduced by Bergmann and Zervas[4] in 1932. The principle methods of cleavage are hydrogenolysis with palladium on charcoal as a catalyst, or acid catalysed solvolysis using HBr in acetic acid. A wide variety of substitutions on the aromatic ring have been investigated[28] in order to vary the acid lability of the protecting group.

b). The tert. butyloxycarbonyl (Boc) group.



(7)

This protecting group (7) was an essential element of the Merrifield 'graded acidolysis' strategy[8] and is still the most popular acid labile N^{α} protecting group used for solid phase peptide synthesis. Cleavage is usually achieved with a solution of trifluoroacetic acid in dichloromethane (1:1). The mechanism of cleavage, shown in scheme 5, includes the generation of a carbonium ion (8), which is a very reactive alkylating species and can cause problems with any sulphur containing residues (Met and Cys) and some of the aromatic amino acids (Tyr and Trp). The use of cation scavengers such as anisole or dimethylsulphide helps to minimise this problem.

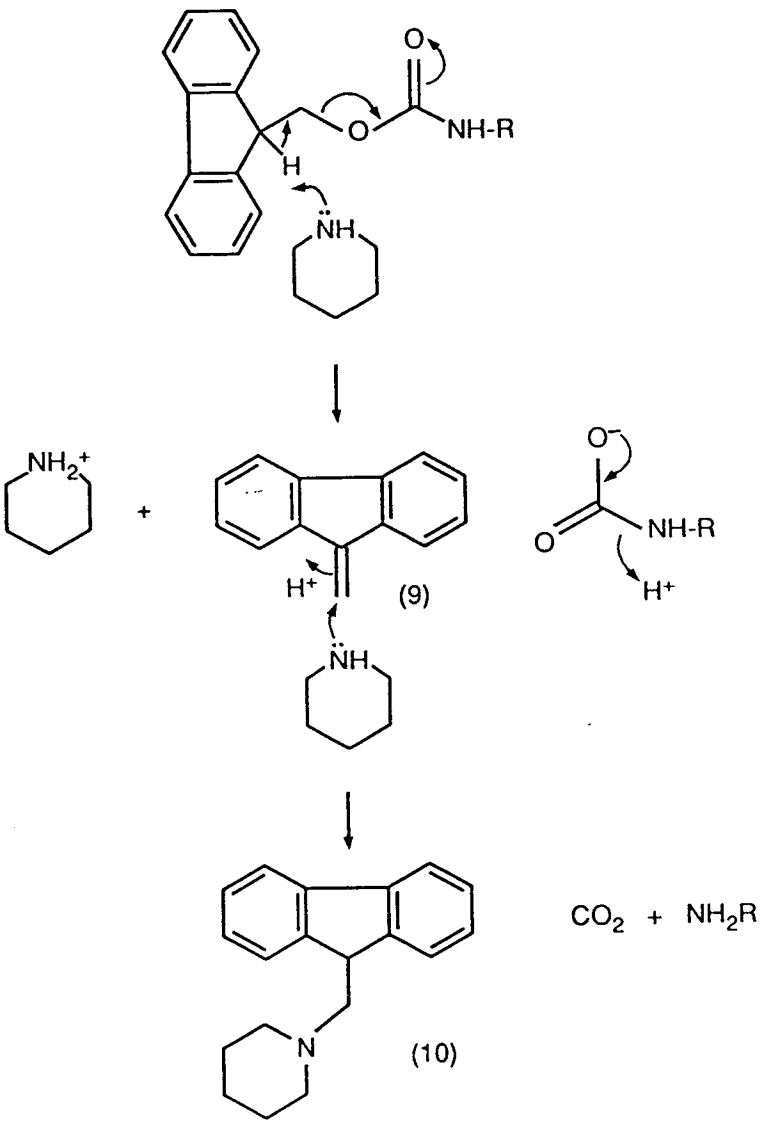


Scheme 5. The cleavage of the Boc group by acid.

c). β-Eliminating Protecting Groups.

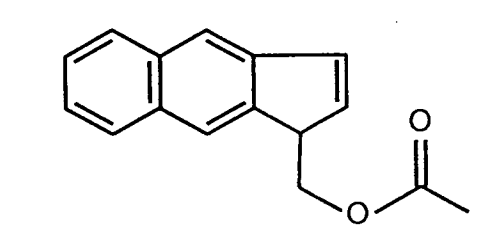
Base labile protecting groups have been used in natural product synthesis for many years[29]. In solid phase peptide synthesis, the most widely used group has been the 9-fluorenylmethoxycarbonyl (Fmoc) group, which was introduced by Carpino[12] and applied to peptide synthesis by Sheppard[10, 11].

The incorporation of a base labile N^{α} protecting group allows the linker and side chain protection to be cleaved using rather milder acid conditions than those required by Merrifield[8]. The cleavage of the Fmoc group is usually carried out with 20% (v/v) piperidine in DMF which allows the olefin (9), which is initially produced by β-elimination, to be trapped as its piperidine adduct (10), as shown in scheme 6.

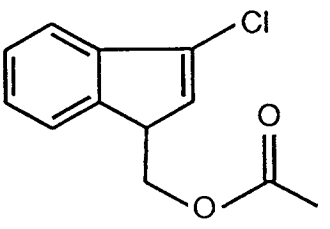


Scheme 6. The cleavage of the Fmoc protecting group by piperidine.

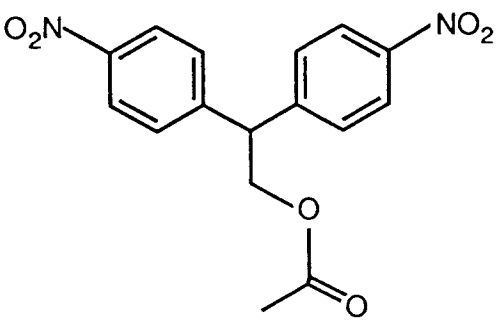
Recently, Carpino *et al.* [30] have introduced the 2-chloro-3-indenylmethoxycarbonyl (Climoc) (11) and benz[*f*]inden-3-ylmethoxycarbonyl (Bimoc) (12) groups, (shown in figure 2). These were applied to peptide synthesis, but using an adapted solid phase approach. The peptide remained in solution but the active ester required for coupling and the base used for deprotection were both attached to solid supports. The Climoc group was not completely stable in a DMF solution, but a tetrapeptide, Climoc-Gly-Gly-Phe-Leu-OBzl was assembled in dichloromethane from H-Leu-OBzl in 74% yield. With the Bimoc group the sequence Fmoc-Tyr(OBzl)-Gly-Gly-Phe-Leu-OBzl was assembled, this time in DMF, with a yield of 74%.



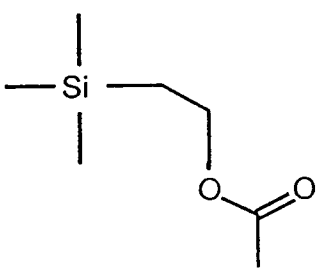
(11)



(12)



(13)

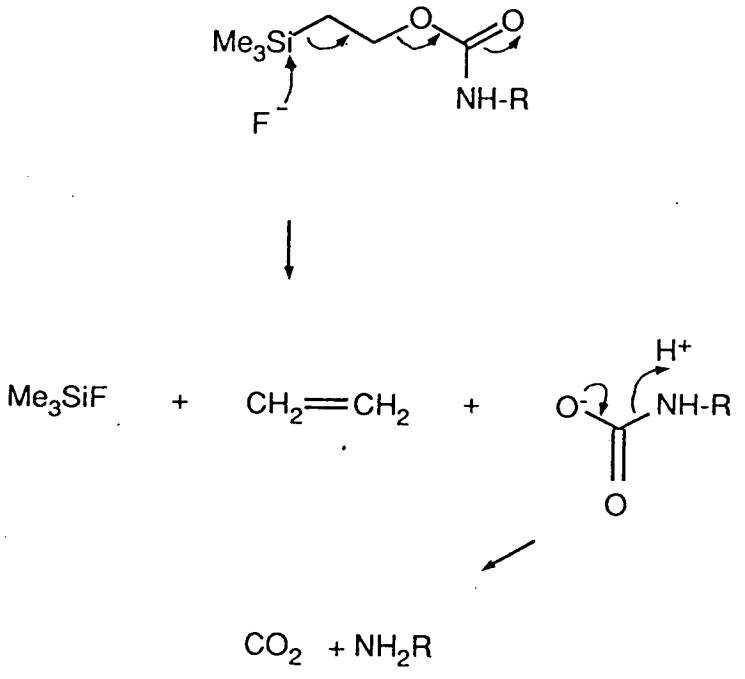


(14)

Figure 2. The structures of β-eliminating protecting groups: Bimoc (11), Climoc (12) and Bnpeoc (13) and Teoc (14) protecting groups.

Another recently introduced base labile protecting group is the 2,2-*bis*(4'-nitrophenyl)ethoxycarbonyl (Bnpeoc) protecting group (13)[31,32]. This group can be cleaved in 20% piperidine in DMF, in a similar fashion to Fmoc and has been used to assemble two small test peptides, H-Leu-Ile-Phe-Ala-Gly-OH and H-Thr-Leu-Ser-Ile-Gly-OH[33].

The β-(trimethylsilyl)ethoxycarbonyl (Teoc) protecting group (14) developed by Carpino *et al.*[34] is an elegant development of the β-eliminating amino protecting group and has been applied to peptide synthesis[35]. It is cleaved by fluoride, as shown in scheme 7, with β-elimination of the silyl moiety. It can also be cleaved by anhydrous TFA.



Scheme 7. The cleavage of the Teoc (14) protecting group by fluoride.

1.1.6 Side chain protection.

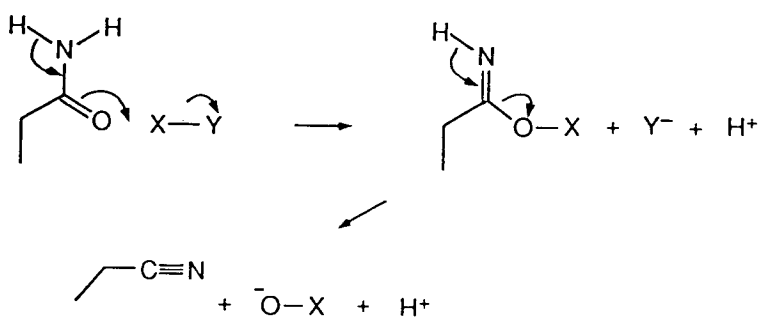
The use of Boc as the N^{α} protecting group in the Merrifield strategy precludes the use of any side chain protection which is cleaved to a significant extent by 50% TFA in dichloromethane. Most protecting groups used with this strategy are based on the benzyl group and the derivatives typically used are shown in table 1.

In the Fmoc strategy, the more acid labile *tert*.butyl derivatives can be used for most side chain protecting groups, although some amino acids require special consideration. Arginine can be protected with the pentamethylchroman sulphonyl protecting group[36] to provide acid lability which is comparable to Bu^t esters and ethers. This has been applied to a number of successful syntheses[37].

Residue	Boc strategy	Fmoc strategy
Asp, Glu	OBzl, OcHxl	Bu^t
Ser, Thr	Bzl	Bu^t
Lys	2Cl-Z	Boc
His	π-Bom, τ-Dnp	π-Bum, τ-Trt
Arg	Tos, NO_2, Mts	Pmc
Tyr	BrZ	Bu^t
Trp	Formyl	None

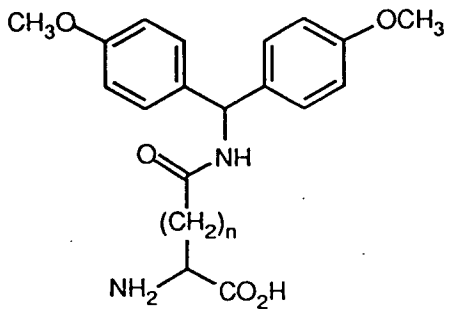
Table 1. Side chain protecting groups that are often used for solid phase peptide synthesis in the Boc and Fmoc strategies (adapted from ref. 9).

The carboxamido group of asparagine and glutamine have often been left unprotected during peptide synthesis, but they can undergo dehydration during activation and coupling to give β-cyanoalanine and γ-cyanobutyrine respectively. The formation of these derivatives has been observed with several coupling reagents, including DCCI, phosphoryl chloride[38] and during the use of phosgene to generate N-carboxyanhydrides[39]. The mechanism for this process is not obvious[39,40] but a hypothetical mechanism is shown in scheme 8.

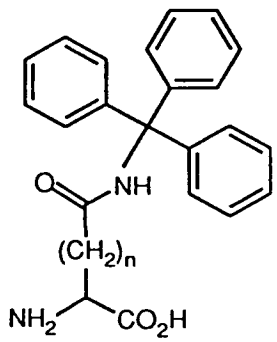


Scheme 8. A mechanism for the dehydration of Asn and Gln residues during activation and coupling.

The formation of these derivatives can be prevented by the use of acid labile alkyl protecting groups on the carboxamido group. The N-*bis*(4-methoxy)methyl (Mbh)[41] (15) and N-trityl (Trt)[42] (16) derivatives, shown in figure 3, have been used successfully. However, these large protecting groups reduce the solubility of the derivatives in some solvents and also may slow the coupling reaction through steric hinderance.



(15)



(16)

Figure 3. The protection of asparagine (n=1) and glutamine (n=2) with Mbh (15) and Trt (16).

The protection of the sulphydryl group of cysteine residues is necessary as it is a potent nucleophile and can be acylated to form a thioester or be oxidised, even by air, to a disulphide. Several protecting groups have been used for sulphydryl protection and the groups chosen depend on whether the final target molecule has a free thiol or one or more disulphide bridges. In the case of multiple disulphide bridges, the selective formation of these depends on the orthogonality of protection used for pairs of residues.

1.1.7 Coupling methods.

The coupling of amino acids is the most frequent point of failure in solid phase peptide synthesis. An ideal coupling reagent has several important features:

1). It must be easy to handle and manipulate on an automated synthesiser.

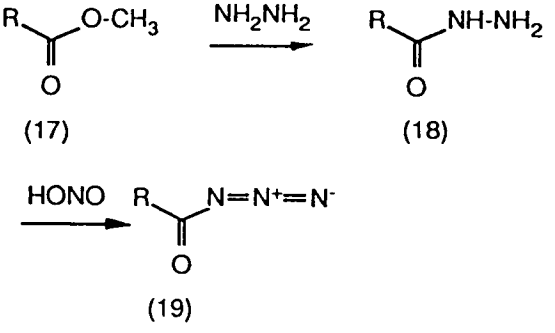
2) It must not allow racemisation to occur.

3). It must give rapid and efficient coupling.

Some of these requirements are, to some extent, mutually exclusive. For example, reagents which couple most rapidly are often air or moisture sensitive and therefore difficult to handle. Some of the methods of coupling amino acids in peptide synthesis are discussed below.

a). The Azide Method.

Developed by Curtius[43], the azide coupling method has been widely and successfully used in the solution phase synthesis of peptides[6]. As can be seen in scheme 9, acyl hydrazides were originally prepared by treating an alkyl ester (17) with hydrazine and then treating the resultant hydrazide(18) with nitrous acid to give the acyl azide (19). These methods have been improved by the introduction of alkyl nitrites to convert the hydrazide to an azide.



(17) (18) (19)

Scheme 9. The preparation of acyl azides.

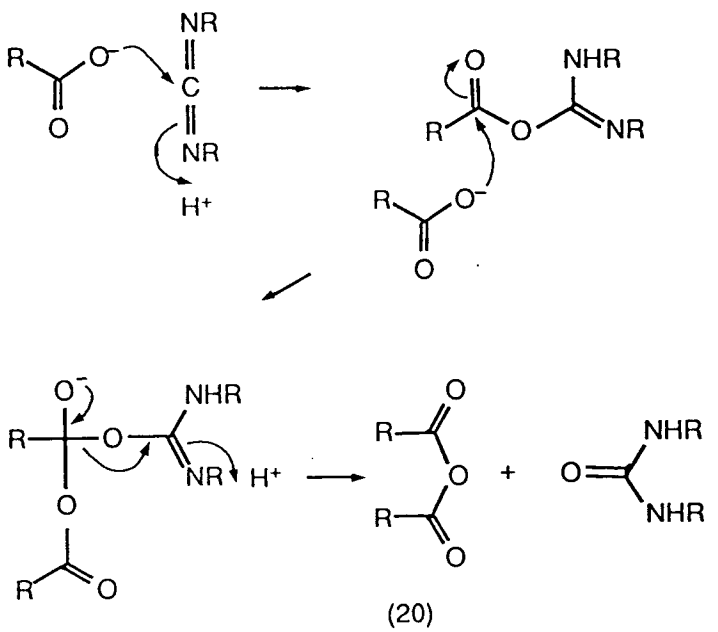
The main problem with this coupling method is the side reaction of acyl azides, the Curtius rearrangment. This side reaction is promoted by heating and by the presence of Lewis or protic acids. Azide couplings are therefore usually performed at low temperatures and in high concentration, immediately after preparing the azide.

b). Acid Chlorides

Acid chlorides have a long history of use in solution phase synthesis of peptides, but have been little used in solid phase work. One reason for this is undoubtedly the tendency of the protected amino acid chlorides to form azlactones, with the concomitant risk of racemisation during activation and coupling. However, Carpino *et al.* [44] have shown the long term stability and superior reactivity of Fmoc amino acid chlorides and suggested that problems of racemisation with Fmoc protected amino acid chlorides are remarkably low (less than 0.1% loss of chirality).

c). Symmetrical anhydrides.

These are often prepared using carbodiimides[45], as shown in scheme 10. Originally dicyclohexylcarbodiimide was used, but the superior solubility of the reaction by-products from diisopropylcarbodiimide has made this increasingly popular during automated solid phase peptide synthesis.



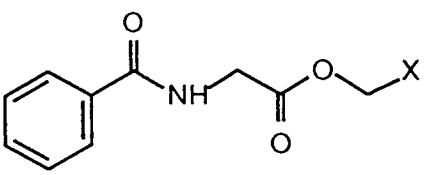
(20)

Scheme 10. The formation of a symmetrical anhydride (20) using a dialkylcarbodiimide.

While these reagents have enjoyed considerable success in both the Boc[46,47] and Fmoc[48,49] stratagies, there are a number of problems associated with their use. Dichloromethane, which provides near optimal conditions of activation[50,51] is not a particularly good solvent for some of the Fmoc amino acids, or for some of the newly formed symmetrical anhydrides[52]. These derivatives are sometimes slow to couple, particularly when the anhydride is prepared from an amino acid derivative with a sterically bulky side chain and is to be coupled to a hindered amino component (e.g. Leu, Ile, or Val residues). The highly activated O-acyl urea which is initially formed can undergo rearrangement to the unreative N-acyl urea and cannot then couple. Finally, the symmetrical anhydride method wastes more than half of the (often expensive) protected amino acid derivative.

d). Active ester methods.

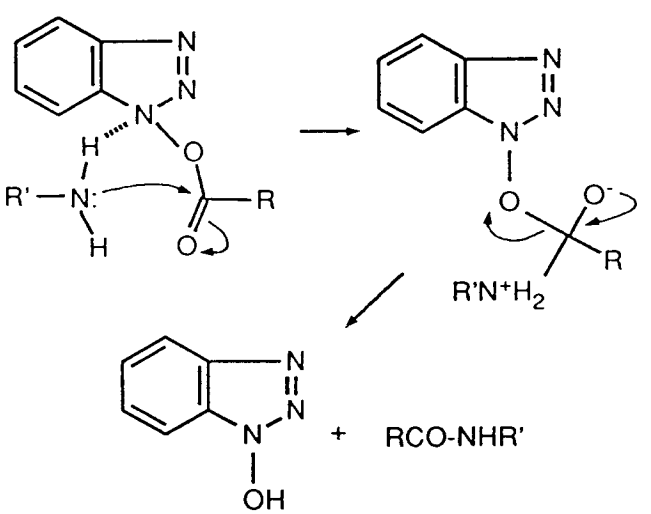
The formation of a peptide bond by the aminolysis of an ester dates from the work of Schwyzer *et al.* [53] who investigated the reaction of modified hippuric acid methyl esters (21).



(21) X = electron withdrawing group

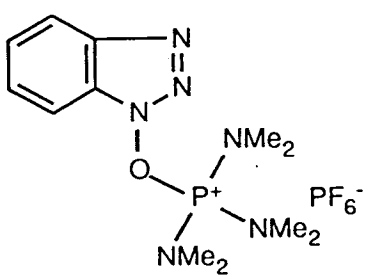
The development of aryl esters for peptide synthesis, such a *p*-nitrophenyl esters, was applied successfully to the solution phase synthesis of oxytocin[54] in 1959. Since then, a wide variety of alcohol components, for the preparation of active esters, have been examined[55] and applied with varying success.

In a somewhat similar vein are the O-acyl derivatives of substituted hydroxylamines[56].The superior reactivity of these derivatives is thought to be caused by the anchimeric assistance provided by the nitrogen atom. An example, using the popular 1-hydroxybenzotriazole ester, is shown in scheme 11.

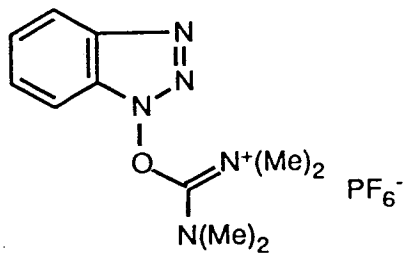


Scheme 11. The formation of a peptide bond by an HOBt ester

These compounds can be used either as an additive in a catalytic concentration in combination with another method of activation, or in stoichiometric quantities. A number of methods have been used to prepare these active esters. The use of dialkylcarbodimides is a common method. More recently they have been prepared from benzotriazolyloxy-tris(dimethyl)amino-phosphonium hexafluorophoshate (BOP reagent) (22) of Castro *et al.*[57] which has been successfully applied to peptide synthesis[58]. Knorr *et al.*[59] have used the somewhat similar 2-(1-hydroxy-1,2,3-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) (23).



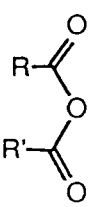
(22)



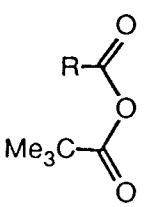
(23)

e). Mixed Anhydrides.

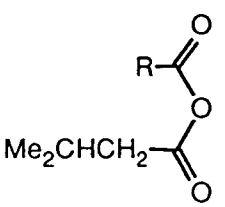
The use of mixed or unsymmetrical anhydrides (24) has been the subject of active research for many years[60,61].



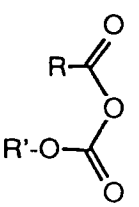
(24)



(25)



(26)

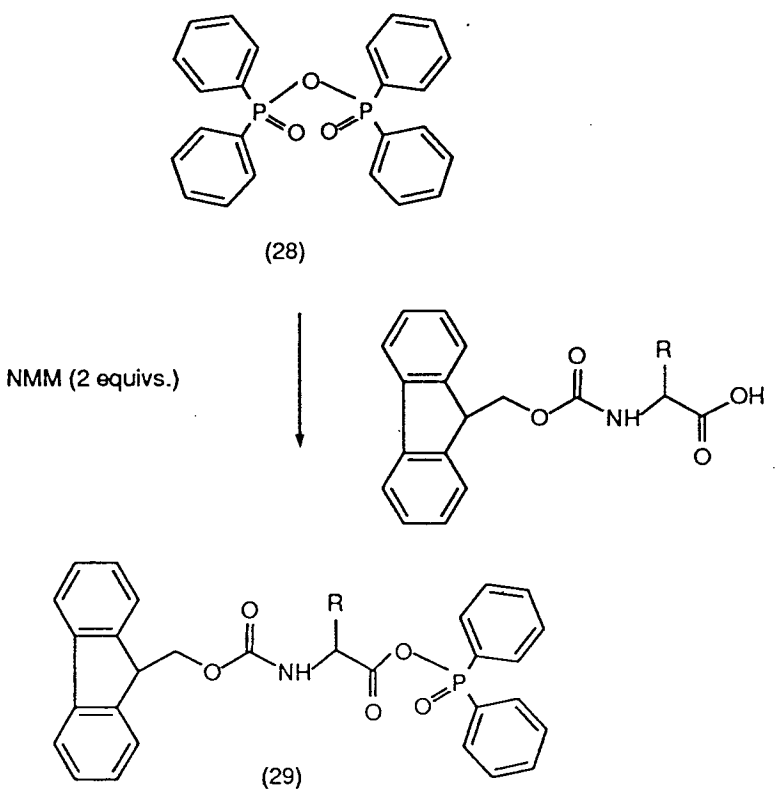


(27)

The principle disadvantage of such anhydrides is the presence of two electrophilic centres which can (and do) give a mixture of products. By varying the nature of R' it is possible to effect the electron density at the carbonyl centre and favour nucleophilic attack at one or the other position. Thus trimethylacetic anhydrides (where R=Me_3C) (25) which donate strongly and isovaleryl (26) mixed anhydrides, give only traces of the unwanted peptide. This method was used successfully in the solution phase synthesis of oxytocin by du Vigneaud *et al.*[5] Replacement of the carboxylic acid moiety with half-esters of carbonic acid[62,63] (27) greatly reduces the electrophilicity of the adjacent carbonyl. This is caused by the electron donating effect of the neighbouring oxygen

Two problems associated with the use of mixed anhydrides are the disproportionation to symmetrical anhydrides and the possibilty of racemisation. These are minimised by use of short activation times at low temperatures.

A further development has been the replacement of the carboxylate group with a phosphonate in a mixed anhydride. These do not seem to undergo attack by the incoming amino group in the phosphorus moiety and this is due, in part, to its excellent leaving group properties. Many phosphoric and phosphinic acid derivatives have been investigated as coupling reagents in peptide synthesis. Recent work within this research group had lead to the development of diphenylphosphinic anhydride (28) which can be used to prepare mixed anhydrides (29) as shown in scheme 12. This has been used to prepare a number of peptides by both solution and solid phase methods[33].



Scheme 12. Preparation of mixed anhydrides from diphenylphosphinic anhydride.

This brief overview of solid phase peptide synthesis indicates the wide variety of options that are available in the nature of the solid support, in coupling methods and in protection strategy. By choosing the correct options it is possible to extend peptide synthesis into the successful synthesis of proteins[9,49,50].

1.2 The use of the helix-turn-helix motif for sequence specific DNA binding.

This topic has recently been reviewed in depth by Harrison and Aggarwal[64]: only the salient points are presented here. The control of gene expression is of central importance in biology. At any given time, each cell uses only a subset of its genes to direct the synthesis of other molecules. This process of controlled gene expression is usually regulated at the level of transcription and an essential part of this control is the binding of regulatory proteins to the recognition sequences of the appropriate genes. The recognition of specific DNA sequences is achieved through discrete DNA binding domains which are generally small: in the region of 40 to 100 amino acid residues. Proteins and peptides of this size are increasingly becoming accessible by solid phase synthesis[9,47] and are of a suitable size for ^{1}H n.m.r. stucture determination[65,66,67].

1.2.1 Recognition by the helix-turn-helix-motif.

The recognition of specific DNA sequences by many proteins bearing the helix-turn-helix (HTH) motif is now well characterised. Many such proteins exhibit a low affinity binding to non-operator sequences, but a much higher affinity (roughly 10^5 fold) affinity for their particular operator sequence from the DNA. There are several reasons for this:

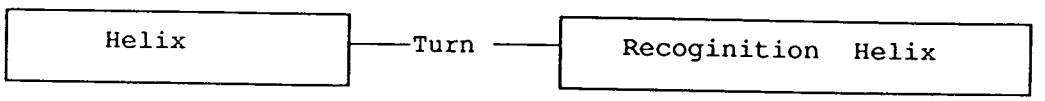
1). The edges of base pairs in B-DNA project hydrogen bond doners and acceptors into the major groove in a sequence specific manner[68]. These provide protein to base pair hydrogen bonds that are very important in the sequence specific binding.

2). The second HTH helix lies in the major groove of B-DNA and the specificity of the interaction between this helix and the DNA has been demonstrated in an experiment which swapped the recognition helices of two DNA binding proteins. The DNA binding sequence specificity of the two proteins had also been exchanged[69,70]. However, this does not work in every case, indicating that other interactions are also important.

3). The local conformation of B-DNA shows considerable variability. In addition, many helix-turn-helix bearing proteins cause considerable bending of the DNA binding site. It is entirely feasible that a protein could 'indirectly' recognise DNA through the variations in the position of the phosphate and sugar groups that occur upon binding.

1.2.2 The helix-turn-helix sequence.

The helix-turn-helix motif may be described as a sequence of about twenty residue. A selection of these is shown in table 2.



λ-Cro	16	Gln-Thr-Lys-Thr-Ala-Asp-Lys-Leu-Gly-Val-Tyr-Gln-Ser-Ala-Ile-Asn-Lys-Ala-Ile-His	35
λ-Rep	33	Gln-Glu-Ser-Val-Ala-Asp-Lys-Met-Gly-Met-Gly-Gln-Ser-Gly-Val-Gly-Ala-Leu-Phe-Asn	52
434 Cro	17	Gln-Thr-Glu-Leu-Ala-Thr-Lys-Ala-Gly-Val-Lys-Gln-Gln-Ser-Ile-Gln-Leu-Ile-Glu-Ala	36
434 Rep	17	Gln-Ala-Glu-Leu-Ala-Gln-Lys-Val-Gly-Thr-Thr-Gln-Gln-Ser-Ile-Glu-Gln-Leu-Glu-Asn	36
CAP	169	Arg-Gln-Glu-Ile-Gly-Glu-Ile-Val-Gly-Cys-Ser-Arg-Glu-Thr-Val-Gly-Arg-Ile-Leu-Lys	108
Trp Rep	68	Gln-Arg-Glu-Leu-Lys-Asn-Glu-Leu-Gly-Ala-Gly-Ile-Ala-Thr-Ile-Thr-Arg-Gly-Ser-Asn	87
Lac Rep	6	Leu-Tyr-Asp-Val-Ala-Arg-Leu-Ala-Gly-Val-Ser-Tyr-Gln-Thr-Val-Ser-Arg-Val-Val-Asn	25
Antp	31	Arg-Ile-Glu-Ile-Ala-His-Ala-Leu-Cys-Leu-Thr-Glu-Arg-Gln-Ile-Lys-Ile-Trp-Phe-Gln	50
λ cII	26	Thr-Glu-Lys-Thr-Ala-Glu-Ala-Val-Gly-Val-Asp-Lys-Ser-Gln-Ile-Ser-Arg-Trp-Lys-Arg	45

Table 2. Sequences of several helix-turn-helix motifs, compiled from refs. 64, 69 and 71. The sequences come from phage λ (cII, rep and cro); from phage 434 (434 rep and cro); from the *E. coli* catabolite gene activator protein (CAP); from the tryptophan operon repressor (trp rep); from the lactose operon repressor (lac rep) and from the *Antennapedia* gene product (Antp) from *Drosophila*..

Residues 1-7 form the first HTH helix and residues 12-20 form the second. Residue 9 is usually Gly because of the backbone flexibility required in the bend region. In addition, a number of hydrophobic residues at positions 4, 8, 10, 16 and 18 are also conserved. The second helix is sometimes refered to as the 'recognition' helix as it is this region which slots into the major groove at the operator site, as shown in figure 4. The region of DNA that a protein binds to is called an operator. Within an operator there may be several operator sites to which the protein may bind.

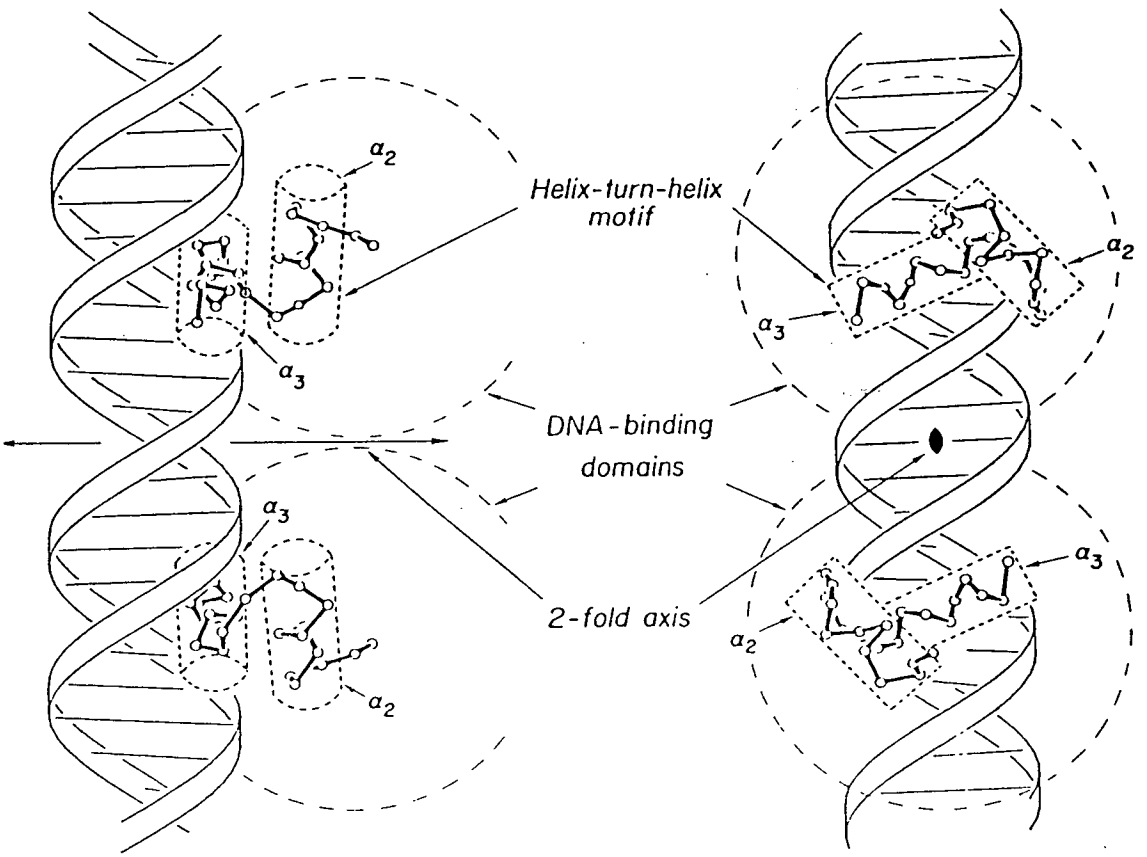


Figure 4. The binding of phage λ cro protein dimer to DNA. Helices 2 and 3 form the helix-turn-helix motif and helix 3, which fits into the major groove, is often refered to as the recognition helix. The Cro protein dimer shows a twofold axis of symmetry which is coincident with an imperfect axis of symmetry in the operator site. Adapted from ref. 71.

1.2.3 Determination of the structure of helix-turn-helix proteins by n.m.r

The first requirement in determining protein and peptide structures by n.m.r is a sequential assignment of all the resonances in the complex ^{1}H n.m.r. spectrum. With the wide range of ^{1}H homonuclear methods that have been developed it is now possible to obtain this information for proteins of up to 100 residues[65,72,73]. Once this hase been achieved, then structural constraints, derived principally from 2D nuclear Overhauser enhancement spectroscopy, are then used to generate a structure using distance geometry calculations and restrained molecular dynamics.

The lac repressor headpiece.

The lac repressor protein DNA-binding domain (or 'headpiece') is a globular domain of 51 to 59 amino acid residues from the lac repressor. ^{1}H n.m.r. has been used to determine the conformation of the 51 residue fragment[74] which confirmed the presence of the helix-turn-helix motif. Subsequent n.m.r. studies of the 56 residue fragment complexed to a 14 base pair half-operator site of DNA[75] have provided a low resolution view of the way in which the protein domain interacts with DNA. The structure of this domain is shown in figure 5.

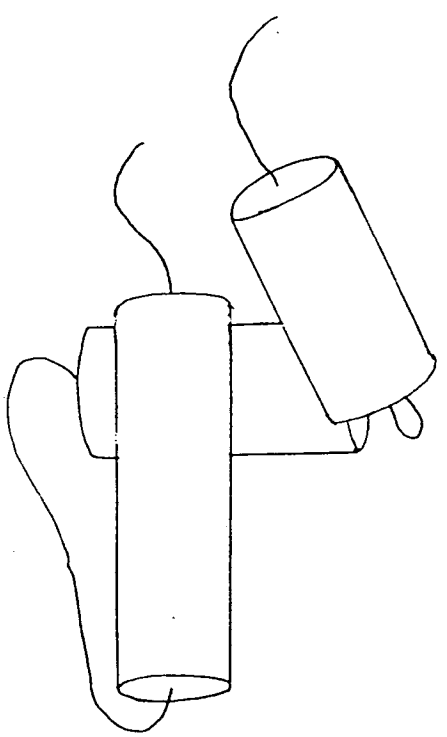


Figure 5. The structure of the lac repressor headpiece.

The Antp *homeodomain*

A number of *Drosophila* developmental gene products and mammalian transcription factors contain a 60 amino acid residue region called the homeodomain[76,77]. The homeodomain region of the *Antennapedia* (*Antp*) gene product from *Drosophila*, which has 68 residues has been studied by 2D n.m.r and a complete sequential assiginment carried out[88]. The use of nOe data[89] has allowed the determination of the structure, which has been shown to include a helix-turn-helix motif very similar to that found in many prokaryotic proteins, despite having substantial differences in other regions of the molecule.

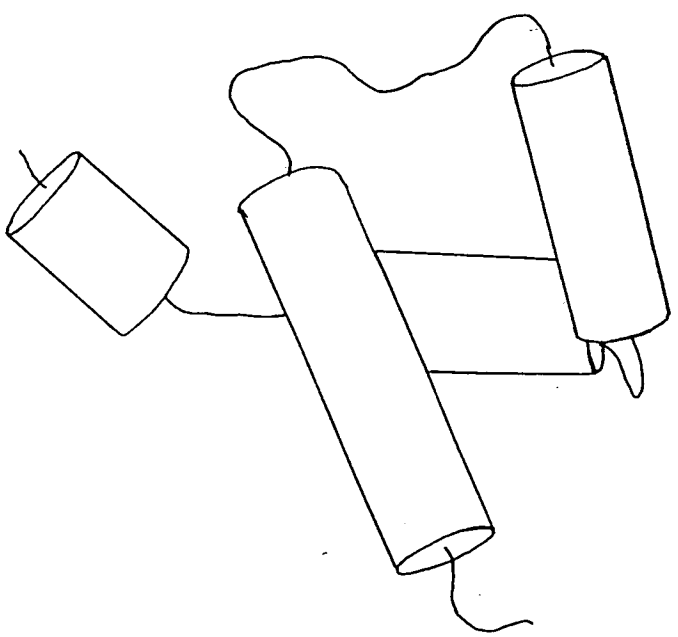


Figure 6. The structure of the Antp homeodomain, maintaining the same orientation of the helix-turn-helix motif as used in figure 5.

1.2.4 The phage λ Cro protein

The bacteriophage λ is a virus which infects the bacterium *E.coli*. One of the proteins coded for by its DNA is the Cro protein, which is involved in controlling the molecular switch that decides between two developmental pathways for the virus. It can undergo lysogeny, in which the quiescent viral DNA is incorporated into (and replicated with) the host cell DNA. Alternatively, it can undergo the lytic pathway, in which the virus takes control of the host cells protein synthesis

and DNA replication machinery in order to produce a large number of viral progeny, until the bacterium bursts open. The progeny then spread and infect other bacteria. The molecular biology of this switch and its control are complex (but lucidly reviewed in a monograph by Ptashne[69]).

The choice of developmental pathway is determined by the balance of two viral DNA binding proteins called the λ repressor and λ Cro. These both compete for binding sites on a small region of the viral DNA called the right operator (O_R). This small but busy region of viral DNA (shown in figure 7A) contains two sites to which RNA polymerase can bind, called promoters (P_{RM} and P_R). Transcription from P_R produces the mRNA that codes for Cro protein. Transcription along the other strand, in the opposite direction, produces the mRNA that codes for repressor.

The region also contains three similar (but not identical) regions of DNA which both the repressor and Cro protein bind to. They bind to these operator sites (O_R1, O_R2 and O_R3) with differing affinities and the binding of the proteins blocks the promoters so that RNA polymerase cannot bind to that promoter.

In the first case (shown in figure 7B), if the Cro protein has bound to its two preferred operator sites, OR_3 and OR_2, then only the P_{RM} promoter is blocked and RNA polymerase only has space to bind at the P_R promoter. This leads to the trancription of more *Cro* mRNA and to the transcription of other proteins downstream from *Cro* which are required for DNA replication and other lytic growth functions.

The other state (shown in figure 7C) occurs when the repressor protein is bound to its two preferred sites O_R1 and O_R2. In this case, P_R is blocked, stopping the production of Cro and the other lytic proteins, but P_{RM} is open to allow RNA polymerase to bind to the P_{RM} promoter and hence produce more repressor. However, the binding of RNA polymerase to the P_{RM} is not very strong and interactions with the repressor at O_R2 help the polymerase to bind. This state occurs when the virus has incorporated its DNA into the host cell chromosome and is called lysogeny.

When certain environmental factors endanger a lysogenic bacterial host cell (one which is carrying the virus), the bacterium produces a protein called RecA which acts as a protease on the repressor. This causes the concentration of the repressor to fall and eventually exposes the P_R promoter. Once Cro production is underway, it binds tightly to O_R3 and O_R2, stopping the production of repressor and sending the virus into the lytic phase of its life cycle.

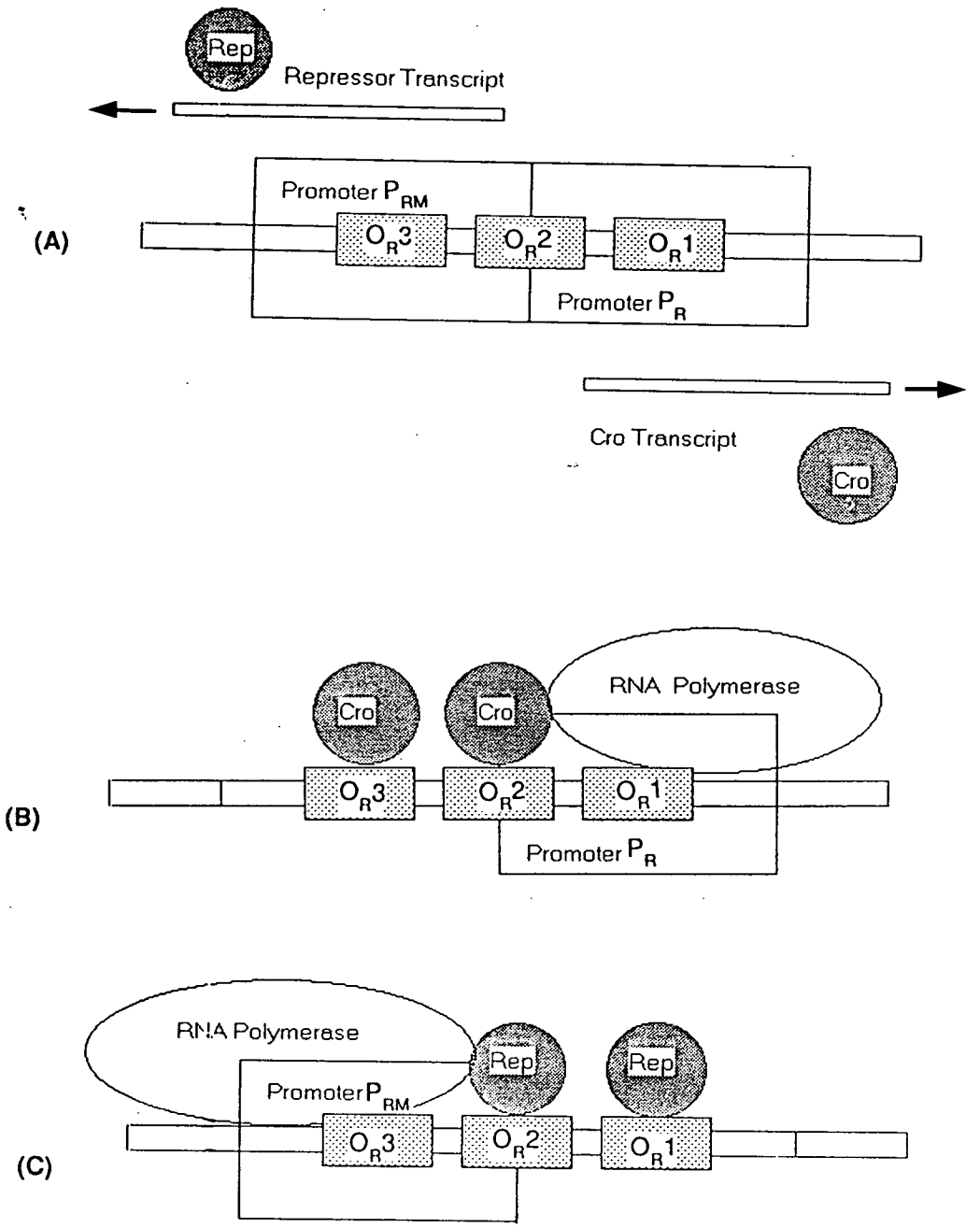


Figure 7. A schematic representation of the molecular switch which controls the development of phage λ. (A) show the arrangement on the viral DNA of the promoters (P_R and P_{RM}), to which RNA polymerase (shown as a large oval in 7B and 7C) can bind. It also shows the operator sites (O_R1, O_R2 and O_R3, all shaded) to which the repressor and Cro proteins (shown as filled circles) can bind. 7B shows the arrangent during lytic growth and 7C shows the arrangment during the lysogenic phase.

Cro has only 66 residues with a molecular weight of 7351 Da per monomer and has been well characterised[80,81,82,83]. Cro folds into a single globular domain and the Cro monomers have a high affinity for each other and form a stable dimer (figure 4), even at the low concentrations typically found in infected cells. It has been crystallised and the structure has been determined[94]. This has shown that Cro has a very simple structure: three strands of antiparallel β-sheet (residues 2-6, 39-45 and 48-55) and three α-helices (residues 7-14, 15-23 and 27-36).

```
1                           15
M E Q R I T L K A Y A M R F G

16                          30
Q T K T A K A L G V Y Q S A I

31                          45
N K A I H A G A L I P L T I A

46                          60
A D G S V T A E E V K P F P S

61      66
N K K T T A
```

Figure 8. The sequence of phage λ Cro.

The Cro structure was used to propose the molecular basis for DNA-protein recognition[85], as shown in figure 3. A complex between λ cro and a 17 base pair duplex of DNA containing the O_R3 sequence was subsequently crystallised and its structure studied[86]. This has revealed that although the main substance of the model described above is correct, there are considerable changes in the positions of the two subunits of the dimer relative to each other.

The protein-DNA interactions have also been studied by protection and base substitution experiments. These have provided an indication of which specific hydrogen bonds and other interactions are occurring (figure 9).

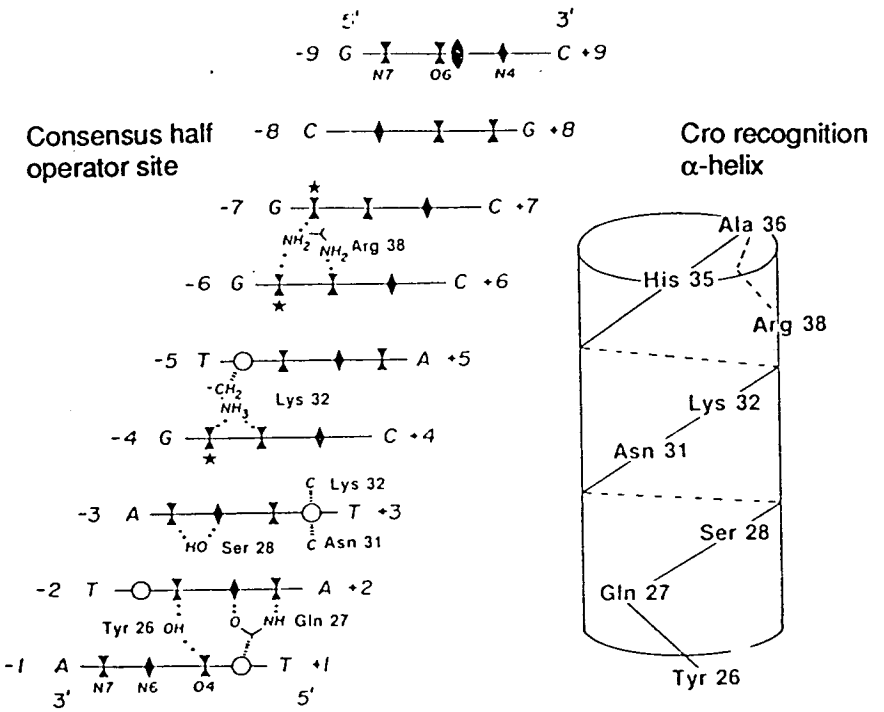
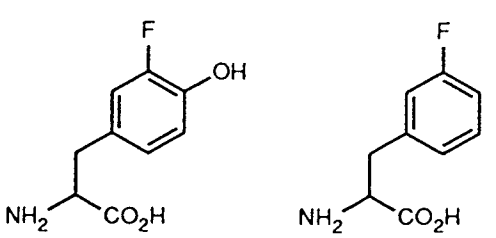


Figure 9. The sequence specific interactions of the Cro protein with DNA taken from ref. 87. Symbols: ⧗ hydrogen bond acceptor, ♦ hydrogen bond donor, ○ Me group of thymine, ★ N-7 of guanine, ···· and ıllıllı are presumed hydrogen bonds.

1.2.5 N.m.r. studies of Cro

Cro had originally been studied by 1D techniques[88] but these experiments yielded little information except that Cro was globular[89]. The Cro dimer, with a molecular weight of 14.7 kDa, was a viable target for sequential assignment. However, the complex of Cro with a synthetic DNA sequence from the operator site (with a molecular weight of about 25 kDa) has a long correlation time and, accordingly, broad ^{1}H n.m.r. resonances.

Cro contains a number of well spaced of aromatic residues (Tyr 10, 26 and 51 and Phe 14, 41 and 58). By feeding *E. Coli* cells overproducing Cro with 3-fluorotyrosine (30) or 3-fluorophenylalanine (31), Ardnt *et al*,[96] were able to prepare Cro labelled with a ^{19}F n.m.r probe in these positions.



(30) (31)

They then used this to study the binding of Cro to DNA by following the effect on the ^{19}F spectrum of titrating in various quantities of non-specific DNA fragments. The shifting of only one tyrosine and none of the phenylalanine resonances may indicate that a single tyrosine is in direct contact with the DNA and this consistant with the model proposed by Ohlendorf *et al*.[85](shown in figure 4).

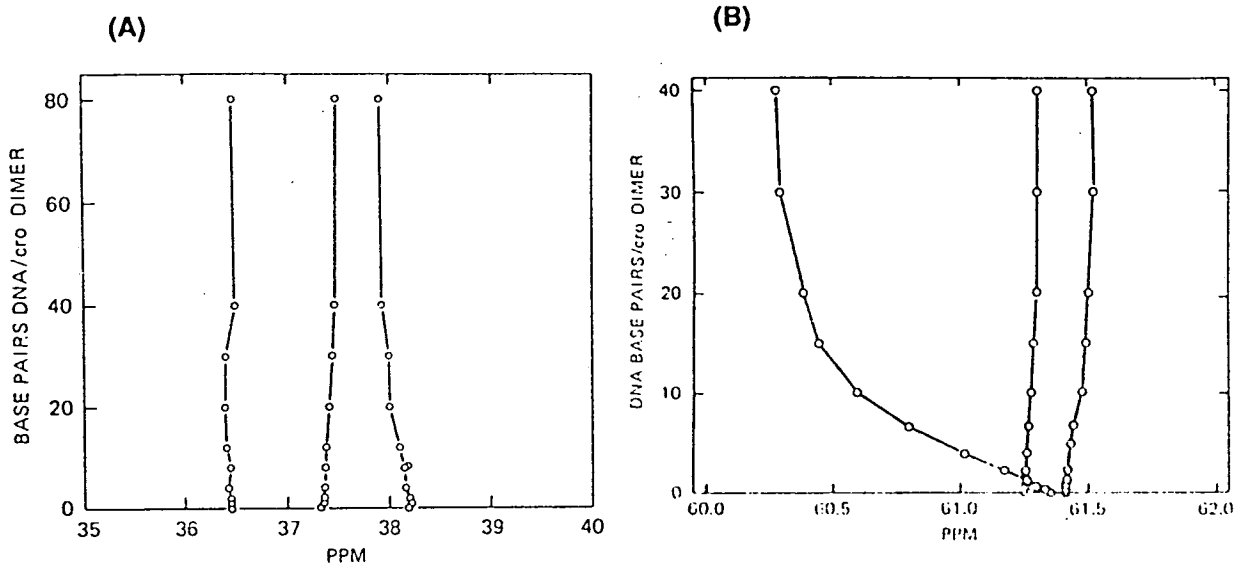
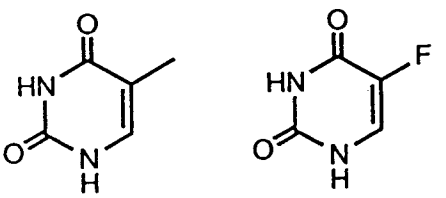


Figure 10. The effect of titrating ^{19}F labelled Cro with non-specific DNA fragments. (A) shows the effect on the resonances of the 3-fluorotyrosine substituted Cro and (B) the effect on resonances of the 3-fluorophenylalanine substituted Cro.

A nearly complete sequential assignment of the Cro ^{1}H n.m.r spectrum has been achieved[91] and and the binding of Cro protein to a 17 bp DNA sequence from O_R3 examined using ^{15}N labelling experiments[92]. The ^{15}N experiments allowed the C-terminal residues (which are not well resolved in the crystal structure) to be observed and they were shown to interact with the DNA.

Metzler and Lu[93] chose to exchange each of the thymines (32) in a 17 base pair duplex containing the O_R3 site with 5-fluorouracil (33). They then measured the solvent accessibility of each of these in the presence and absence of Cro. These results indicated that the binding of Cro to the DNA bends the operator site and the the two halves of the Cro dimer are not bound symmetrically.



(32) (33)

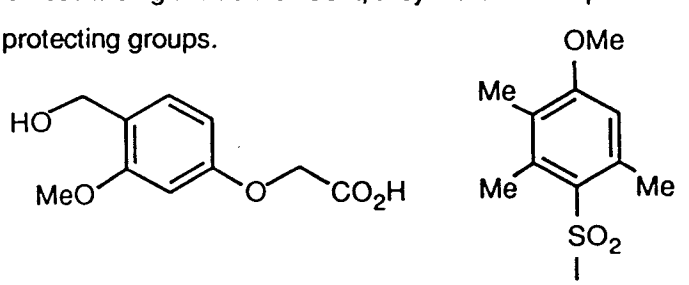
Figure 11. The structures of thymine (32) and 5'-fluorouracil (33).

1.2.6 Synthetic Studies on Cro

The first studies on peptides from the Cro protein were carried out before the crystal structure was established. Gutte *et al.*[94] prepared a 22 residue fragment from the amino terminus of Cro. They also prepared the same sequence except with an additional Cys residue on the carboxy terminal. This allowed them to dimerise the peptide and this was found to bind preferentially to λ DNA and to polyG.

Mayer *et al.* [95] have synthesised a peptide corresponding to the recognition helix (residues 26-39) and studied its folding using 1D n.m.r., using hexafluoroisopropanol (HFIP)/ water mixtures to induce some secondary structure. They also studied the addition of a synthetic DNA duplex to the peptide in HFIP/water (1:4) and showed changes occured in the 1D ^{1}H n.m.r. spectrum. Whilst these results are interesting and indicate DNA:peptide interactions are occurring, more conclusive evidence would come from a re-examination of the system using 2D n.m.r. techniques.

Recently Sheppard *et al*. have published[96] the results of some extensive studies into the synthesis of Cro by fragment condensation methods (figure 12). By using the very acid labile linker (34) it was possible to assemble protected peptides and cleave them from the solid support with 1% TFA in dichloromethane. As they point out, this was not strictly an orthogonal strategy (c.f. the fluoride labile linkers of Barany[22] and of Ramage[23]), but by carrying out the cleavage as a flow process and neutralising the acid effluent, they were able to practically eliminate the loss of Bu^t side chain protecting groups.



(34) (35)

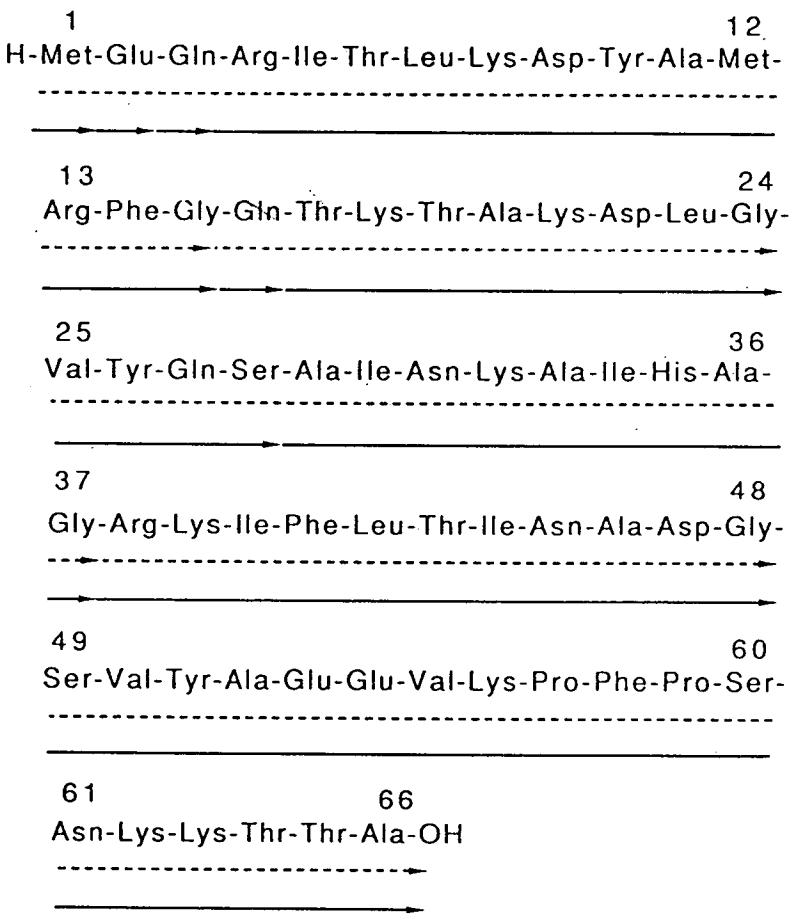


Figure 12. The sequence of λ Cro, showing the planned (dashed lines) and actual (solid lines) protected fragments synthesised for the assembly of Cro. Taken from ref. 96.

One of the principle problems that they encountered was the insolubility of the fully protected peptide fragments that were produced. This remains an obstacle to the preparation of extended protein sequences by fragment condensation. Unfortunately, experiments with a purified sample of the natural λ Cro showed that it was intrinsically unstable to trifluoroacetic acid and was substantially degraded after 2 hours of exposure to the acid. The arginine protecting groups which had been used, Mtr (35), required protracted treatment with TFA to be completely removed and the chemical synthesis of pure Cro has not yet been accomplished.

2. Discussion

2.1 The preparation of Bnpeoc derivatives

2.1.1 The design of a novel base labile protecting group

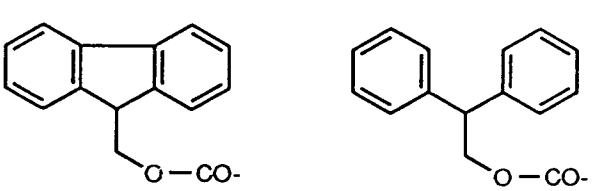
The base labile 9-fluorenylmethoxycarbonyl protecting group has proved very successful in solid phase peptide synthesis[48,49,50]. The concept of a truly orthogonal protecting strategy such as the Fmoc/Bu^t system is both intellectually appealing and practically convienient.

The removal of the Fmoc group is usually achieved with 20% (v/v) piperidine in DMF and the initial product, dibenzofulvene, is then trapped as its piperidine adduct (scheme 6, page 11). However there are a number of problems associated with the use of Fmoc amino acid derivatives in solid phase peptide synthesis. A number of the Fmoc amino acids or their symmetrical anhydrides are not soluble in dichloromethane[52], which is thought to be the best solvent for activation[50,51] and has been used routinely during solid phase peptide synthesis in the Boc/Bzl strategy. Futhermore, the Fmoc amino acids are between two and five times more expensive as their Boc equivalents.

When considering potential novel protecting groups for solid phase peptide synthesis, several important criteria must be considered:

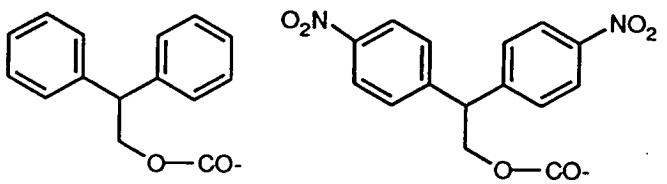
1. The protected derivatives must have a good shelf life, preferably at room temperature.
2. The protecting group must be readily prepared and introduced.
3. The deprotection conditions used must be almost quantitative and selective under mild conditions.
4. The by-products of deprotection should be easily removed.
5. The reagent should not be toxic.

Ramage and Florence[31] chose to develop an alternative to the Fmoc group using the following approach. The removal of the bond between the two phenyl rings of Fmoc gives diphenylethoxycarbonyl (36).



(36)

The diphenylethoxycarbonyl group is a very poor β-eliminating group because of the reduced acidity of the β hydrogen. In order to increase this, the introduction of an electron withdrawing nitro group in the *para* position was considered.



Consideration of the pK_a's of the compounds shown in figure 13 suggests that, providing ease of deprotection is related solely to the acidity of the β hydrogen, then the 2,2*bis*(4'-nitrophenyl)ethoxycarbonyl protecting group should be labile to weaker bases than is Fmoc, since the α leaving group is identical. This protecting group was independently developed, but inadequately characterised, by König *et al.*[97] Five amino acid derivatives were prepared by them, but only one, Bnpeoc-Phe-OH, was obtained as a crystalline solid. They therefore stopped at this point.

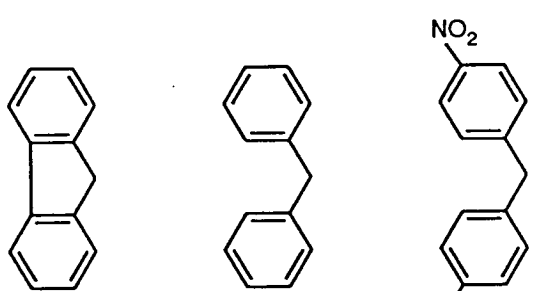
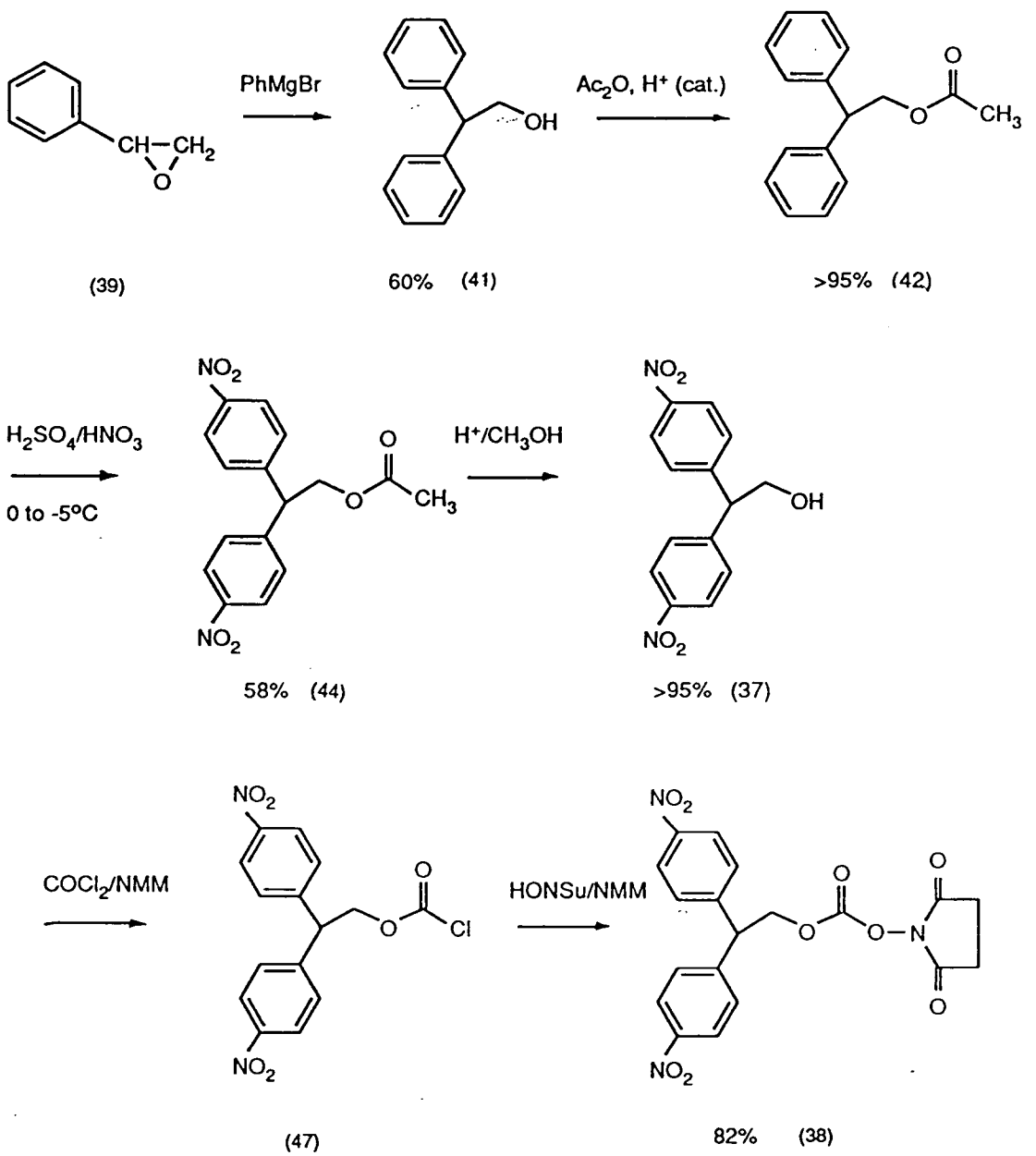


Figure 13. The relationship between the acidity of the methine protons of three diarylmethanes.

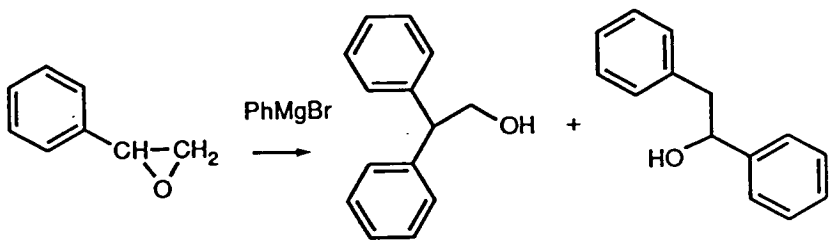
2.1.2 The preparation of 2,2-*bis*(4'-nitrophenyl)ethoxycarbonyl N-hydroxysuccinimide.

Florence[31] developed two routes to the preparation of the key intermediate, 2,2-*bis*(4'-nitrophenyl)ethanol (37) and the more viable route is shown in scheme 13. Several modifications of this route have been studied. The synthesis of large quantities (1-2 kg) of 2,2-*bis*(4'-nitrophenyl)ethoxycarbonyl-N-hydroxysuccinimide (BnpeocONSu) (38) was carried out using this route, without modification but substantially scaled up.



Scheme 13. The synthesis of Bnpeoc-ONSu (38).

The first step in the route involves the reaction of styrene oxide (39) with phenyl magnesium bromide. The Grignard reagent can attack either of the carbon atoms in the epoxide ring leading to two possible products (scheme 14).



(39) (41) (40)

Scheme 14. The reaction of phenylmagnesium bromide and styrene oxide.

Kharash and Clapp[98] showed that the product of attack at the least hindered site, 1,2 diphenylethanol (40), was obtained when a solution of the Grignard reagent was added to the styrene oxide solution. However, addition of the styrene oxide to the Grignard gave the 2,2-diphenylethanol (41). If this were the result of similar kinetics for attack at both carbon atoms (which might be predicted for a highly reactive nucleophile) then a 1:1 mixture of the isomers would be expected. However, yields in excess of 70% of the 2,2 diphenylethanol can be obtained. A possible explanation of this is the binding of a Lewis acid to the oxygen atom as shown in figure 14.

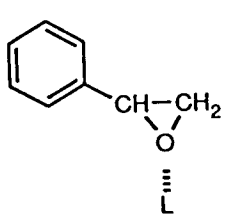
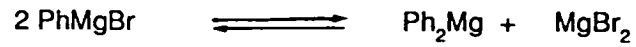


Figure 14. The binding of a Lewis acid (L) to styrene oxide.

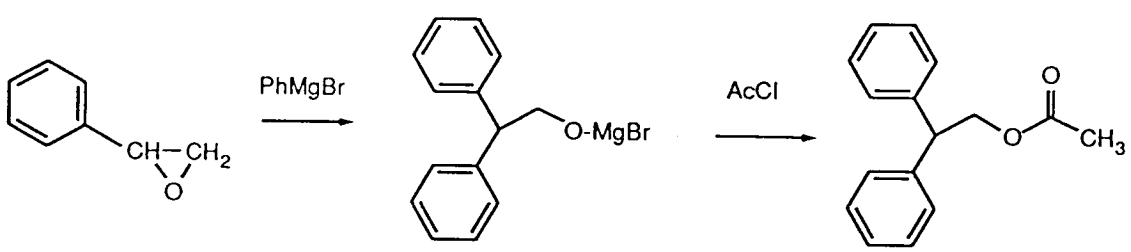
The polarisation of the carbon oxygen bond leads to a partial carbocation that can be stabilised by the phenyl group. Therefore the carbon atom proximal to the phenyl group is more susceptible to attack, by suitable nucleophiles, leading in our case to the required alcohol (41). Grignard reagents are known to be equilibrium mixtures as shown in scheme 15 and it seems likely that the magnesium ions are acting as the Lewis acid in this case.



Scheme 15.

In practice, the yield of the reaction proved variable (between 40 and 60%) and the lower yields were experienced in the larger scale (3.0 mol) preparations. Attempts to distil (under high vacuum) more of the product from the liquors that remained after crystallisation gave no more of the required product. The main fraction isolated from this distillation was biphenyl.

The 2,2-diphenylethanol (41) was readily acetylated, in quantitative yield, with acetic anhydride in the presence of a catalytic quantity of conc. sulphuric acid. In an attempt to reduce the number of steps in the synthesis, a direct route of synthesising 2,2-diphenylethylacetate (42) was attempted as shown in scheme 16 . Instead of isolating the product of the reaction between the Grignard reagent and styrene oxide, a slight molar excess of acetyl chloride was added to the magnesium 2,2-diphenylethoxide (43) solution. Unfortunately a mixture of products, including both the 2,2-diphenylethanol (41) and its acetate (43) were obtained. Given the excellent yield of the acid catalysed acetylation reaction, this work was not carried any further.



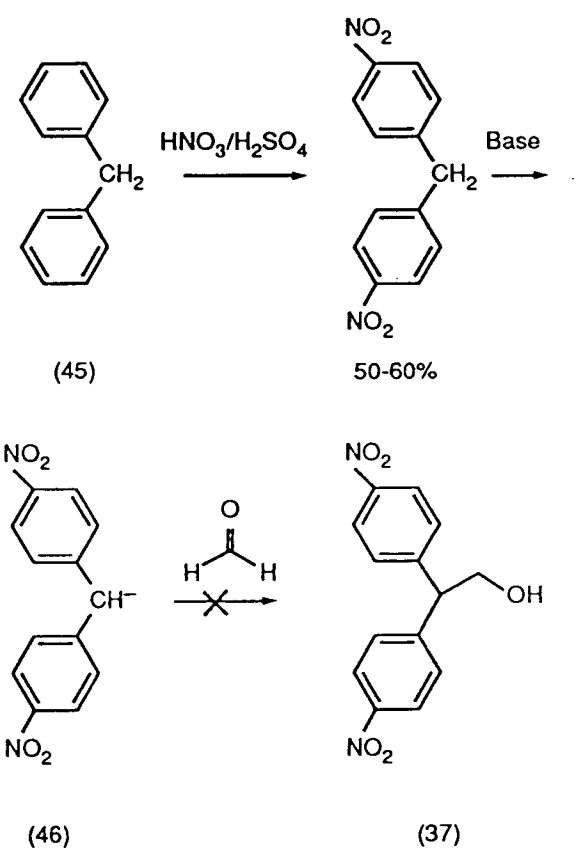
(43) (42)

Scheme 16. The direct synthesis of 2,2-diphenylethylacetate (42).

The nitration of the 2,2-diphenylethylacetate (42) using a mixture of conc. sulphuric and conc. nitric acid proceeded with an acceptable yield of 58%. The temperature of the reaction mixture was found to be critical: if the temperature fell below -5°C the rate of reaction became slow, but if allowed above 0°C the yield was reduced, probably through side reactions such as nitration at the *ortho* position. The reaction was exothermic and careful precautions (such an ice/ salt bath and vigorous stirring) had to be taken to ensure that the heat generated was quickly dissipated and the temperature thereby kept constant.

The acid catalysed methanolysis of the 2,2-*bis*(4'-nitrophenyl)ethylacetate (44) to the 2,2-*bis*(4'-nitrophenyl)ethanol (37) proceeded very cleanly and was essentially quantitative.

Recently Gray[33] has investigated an alternative route to 2,2-*bis*(4'-nitrophenyl)ethanol shown in scheme 17. Nitration of the diphenylmethane (45) proceeded in acceptable yield (50-60%) and the anion (46) was readily formed, as evidenced by the intense blue colour which appeared immediately on adding the base. Unfortunately, the expected reaction of this anion with formaldehyde did not occur because of the highly delocalised nature of the anion (46), which renders it a poor nucleophile.

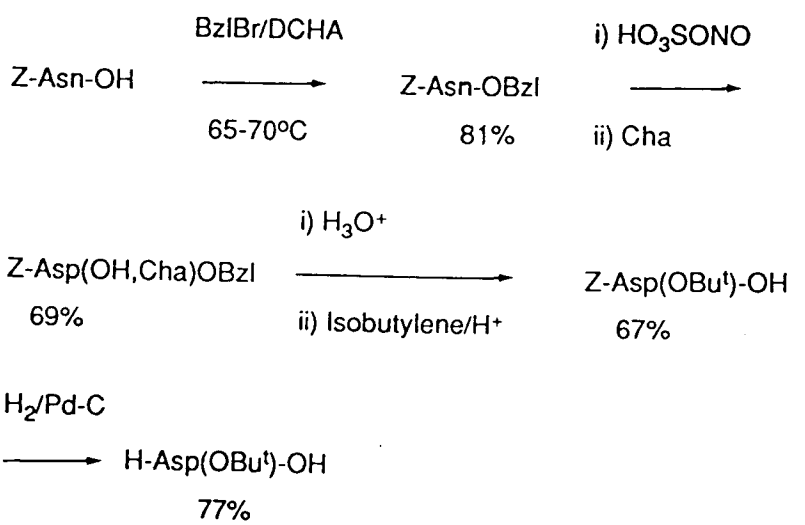


Scheme 17. An alternative route to 2,2-*bis*(4'-nitrophenyl)ethanol (37).

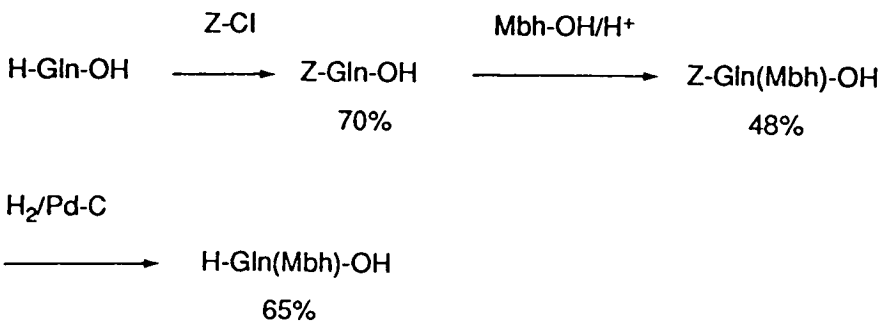
Treatment of the 2,2-*bis*(4'-nitrophenyl)ethanol (37) with phosgene (used as a toluene solution) and N-methylmorpholine afforded 2,2-*bis*(4'-nitrophenyl)ethylchloroformate (47), which could be isolated as a white crystalline solid (with identical physical characterisitics to those decribed by Florence[31]). However, the chloroformate (47) was usually not isolated. The N-methylmorpholine hydrochloride salt which had precipitated out during the reaction was removed by filtration and then the crude chloroformate solution was converted directly to the N-hydroxysuccinimidyl carbonate (38) by the addition of a further equivalent each of N-methylmorpholine and N-hydroxysuccinimide. The solution was again filtered to remove the precipitated N-methylmorpholine hydrochloride salt. Subsequent removal of the solvent *in vacuo* gave a yellow solid, which after washing with cold actetone and petrol, gave Bnpeoc-ONSu as a white solid in 82% yield. The overall yield from styrene oxide was 26%.

2.1.3 The preparation of Bnpeoc protected amino acids.

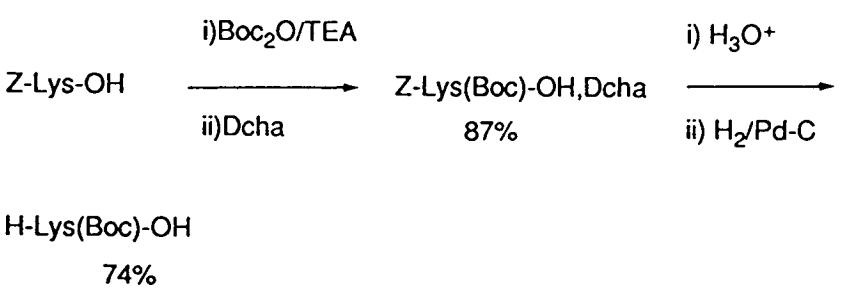
Many of the simpler N^{α}-Bnpeoc protected amino acid derivatives, which do not require side chain protection, have been synthesised and described by Florence[31]. In addition, Bnpeoc-Ser(Bu^t)-OH and Bnpeoc-Thr(Bu^t)-OH have been prepared by Gray[33]. Many of the expensive side chain protected derivatives were synthesised from the parent amino acid. These were prepared by literature methods (or adaptations of them) and are described and referenced in chapter 3 (experimental). The routes used are outlined in schemes 18 to 22.



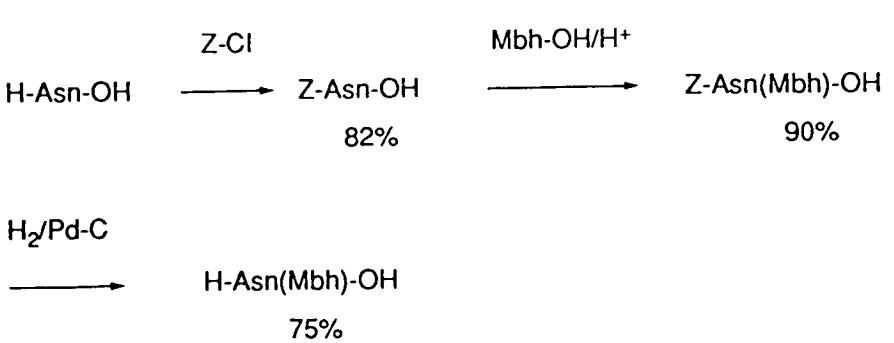
Scheme 18. The preparation of H-Asp(OBu^t)-OH



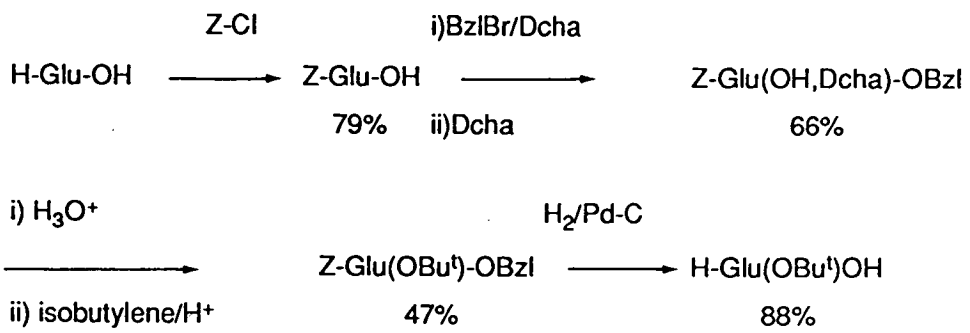
Scheme 19. The preparation of H-Gln(Mbh)-OH



Scheme 20. The preparation of H-Lys(Boc)-OH



Scheme 21. The preparation of H-Asn(Mbh)-OH



Scheme 22. The preparation of H-Glu(OBu^t)-OH

The Bnpeoc amino acids were all prepared by an adaptation of the method of Rich[99]. The amino acid or side chain protected derivative was dissolved or suspended in a mixture of dioxan and water (1:1) and 1.5 equivs. of triethylamine was added. The solid Bnpeoc-ONSu (38) was then added and the reaction followed by tlc until it had reached completion (4-16 hours). The Bnpeoc protected amino acid was isolated by acidifying the medium and extracting the derivative into an organic phase. The amino acid derivatives that readily crystallised were purified by recrystallisation, but the derivatives which proved difficult to crystallise were purified by conversion to the cyclohexylamine (Cha) or dicyclohexylamine (Dcha) salts. These salts were recrystallised and were then converted back to the pure free acids in near quantitative yields by partitioning the salt between an organic solvent and aqueous potassium hydrogen sulphate solution (2.0 M). The free acids were obtained either as a foam or as a precipitate from chloroform/petroleum ether. This process was used to prepare a suitable quantity of each amino acid derivative for solid phase peptide synthesis (usually between 5 and 30 grams). Some of the amino acid derivatives that have been prepared are listed in table 3.

Name	M.p. (°C)	$[\alpha]_D^{RT}$ (Degrees)	Yield (%)	Notes
Bnpeoc-Ala-OH	147-148	+0.8 (1.0 DMF)	90	A
Bnpeoc-Arg(Pmc)-OH	120-140	+3.6 (1.0 $CHCl_3$)		B, C, D
Bnpeoc-Arg(Pmc)-OH,Cha	136-137		79	B
Bnpeoc-Asn-OH	180-182	+0.5 (1.0 DMF)	83	A, G
Bnpeoc-Asn(Mbh)-OH	114-115	-2.1 (1.0 DMF)	87	
Bnpeoc-Asp(OBu^t)-OH	75-79			C, D
Bnpeoc-Asp(OBu^t)-OH,Dcha	160-164	+17.4 (1.0 $CHCl_3$)	69	
Bnpeoc-Cys(SBu^t)-OH,Cha	175-178		75	H
Bnpeoc-Gln-OH	189-190	-3.9 (1.0 DMF)	94	A
Bnpeoc-Gln(Mbh)-OH	97-107	+5.8 (1.0 DMF)	62	
Bnpeoc-Glu(OBu^t)-OH				A, D
Bnpeoc-Glu(OBu^t)-OH,Dcha	129-132	+5.3 (1.0 DMF)	88	A
Bnpeoc-Gly-OH	156-159		90	A
Bnpeoc-His(Trt)-OH	134-138	+4.4 (1.0 DMF)	85	
Bnpeoc-Ile-OH				A, C, E
Bnpeoc-Ile-OH, Dcha	158-159	+7.7 (1.0 DMF)	92	A
Bnpeoc-Leu-OH	68-70	+4.9 (1.0 DMF)	92	A
Bnpeoc-Lys(Boc)-OH	76-79	+7.6 (1.0 $CHCl_3$)		C, D
Bnpeoc-Lys(Boc)-OH,Dcha	146-149	+2.9 (1.0 DMF)	93	
Bnpeoc-Met-OH		+8.1 (1.0 DMF)		E
Bnpeoc-Met-OH,Dcha	186-189		81	A
Bnpeoc-Phe-OH	77-78	-16.5 (1.0 DMF)	93	A
Bnpeoc-Pro-OH	74-79	-18.7 (1.0 DMF)	91	A,G
Bnpeoc-Ser(Bu^t)-OH		+13.7 (1.0MeOH)		C, E, F
Bnpeoc-Ser(Bu^t)-OH,Cha	182-185	+11.0 (1.0 DMF)	93	
Bnpeoc-Thr(Bu^t)-OH		+9.7 (1.0 MeOH)		C,E,F
Bnpeoc-Thr(Bu^t)-OH,Cha	153-156	+11.0 (1.0 DMF)	55	
Bnpeoc-Trp-OH	94-97	-24.0 (1.0 DMF)	91	A,G
Bnpeoc-Tyr(Bu^t)-OH	73-76	-20.0 (0.8 DMF)	75	
Bnpeoc-Val-OH	142-145	+7.5 (1.0 DMF)	89	A,G

Table 3. Bnpeoc amino acid derivatives.

Notes

The optical rotations were all recorded at room temperature (approx. 25°C). *c* and the solvent are given in brackets.

A: Reference 31

B: Reference 36

C: For those derivatives isolated as their Dcha and Cha salts, the yield for the conversion to the free amino acid is nearly quantitative.

D: Isolated as a powder. When analysis indicated these were pure and free of solvents the optical rotation and m.p. were recorded.

E: Isolated as a foam.

F: Reference 33

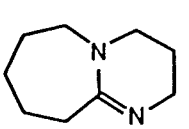
G: Reference 100

H. Reference 101

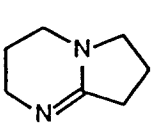
2.1.4 The deprotection of Bnpeoc derivatives

Experiments carried out by Florence[31] have indicated that the Bnpeoc group is very stable to acids and no modification is detected in TFA over 24 hours or in 3N HCl in methanol over 18 hours. It is not cleaved by pyridine or 2,6-lutidine. Not surprisingly, catalytic hydrogenation (for two hours) gave two major products (in a mixture with the starting material) which, after separation by silica chromatography, were shown to be the products of reduction of one or both nitro groups: the 2(4-nitrophenyl)-2(4-aminophenyl)ethyl and the 2,2-*bis*(4'-aminophenyl)ethyl derivatives.

Gray[33] has undertaken a comprehensive study of the stability of Fmoc- and Bnpeoc-Ser-OBzl in several solvents (DMF, DMA and NMP). The results have suggested that both of the protecting groups are cleaved by basic impurities present in these solvents. The rate of cleavage is dependent on the purity of the solvent, but the Bnpeoc derivative is more sensitive. However, the Bnpeoc amino acids, in which the carboxylic acid functionality is unprotected, are not affected by these solvents over 44 hours. This is probably because the carboxylic acid is protonating any basic impurities in the solvents.

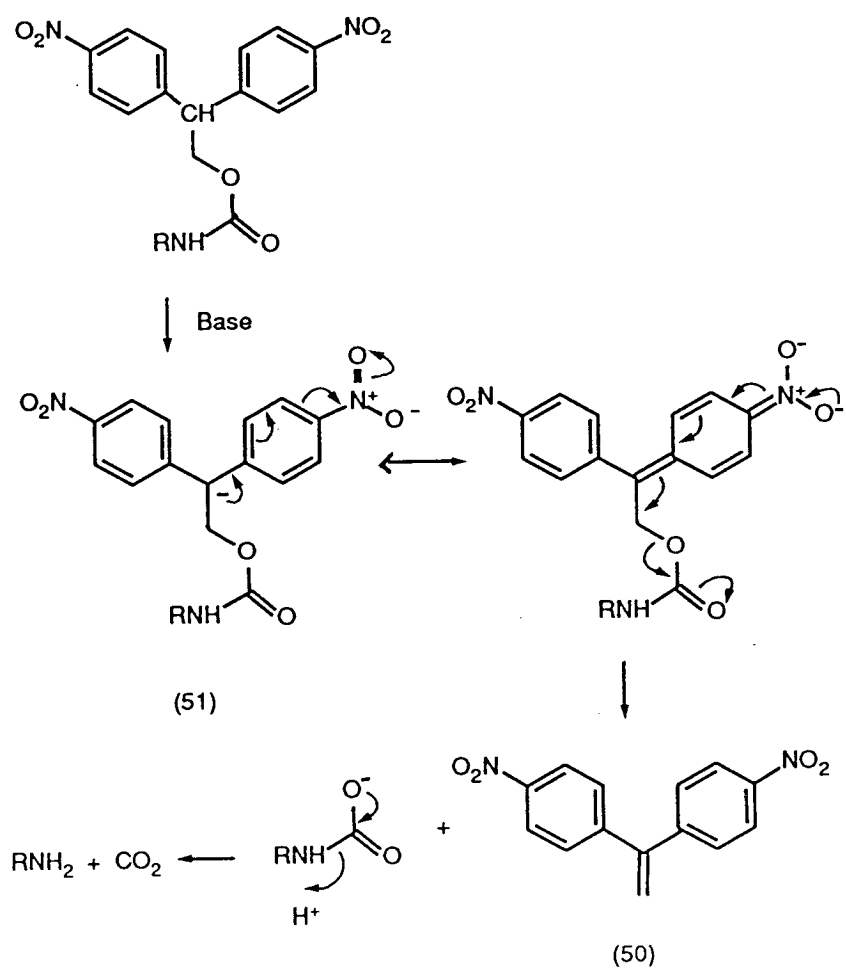


DBU (48)



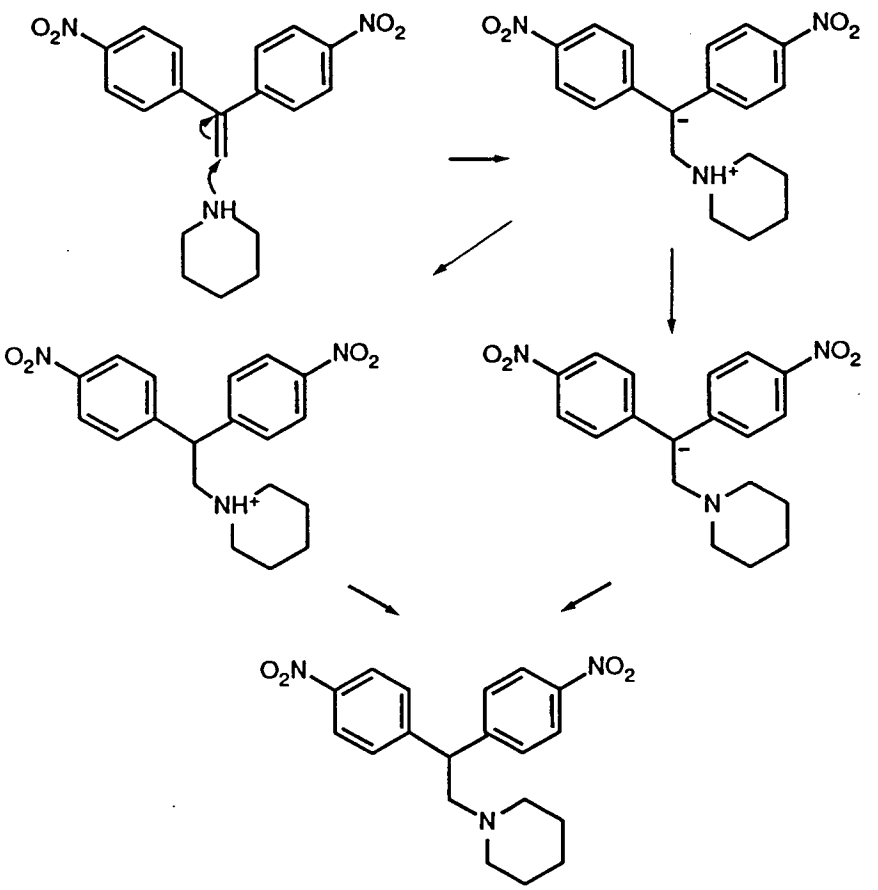
DBN (49)

The Bnpeoc group may be cleaved by the addition of strong tertiary bases such as diazabicyclo-(5.4.0)-undec-7-ene (DBU) (48) and diazabicyclo-(4.3.0)-non-5-ene (DBN) (49) with rapid β-elimination of 1,1-*bis*(4'nitrophenyl)ethene (50) (scheme 23). A transient blue colour is observed in the reaction vessel, characteristic of anions of the type (51) which fades as the reaction reaches completion. This process occurs in the presence of equimolar quantities of base, even if these are buffered with acetic acid[31].



Scheme 23. The cleavage of the Bnpeoc group by DBN or DBU.

The Fmoc group is usually cleaved during solid phase peptide synthesis using a solution of 20% piperidine (v/v) in DMF and results in the formation of the piperidine adduct. The analogous reaction occurs for the Bnpeoc group (scheme 24).



(52)

Scheme 24. The reaction of 1,1-*bis*(4'-nitrophenyl)ethene (50) with piperidine to give the piperidine adduct (52).

Addition of piperidine to a solution of the olefin in DMF results in a strong blue colour (indicating the formation of an anion) which fades as this reaction reaches completion. It remains unclear whether the quaternary ammonium centre loses its proton before or after the delocalised carbanion gains one (and both routes are shown in scheme 25). UV spectroscopic studies of the conversion of the olefin ($\lambda_{max.}$ 304 nm) to the piperidine adduct ($\lambda_{max.}$ 270 nm) have indicated[33] that the addition reaction occurs very quickly (essentially complete within three minutes) at the concentrations typically used during automated solid phase peptide synthesis.

2.2 Solid phase peptide synthesis

In the this section and in the experimental chapter, the yield quoted is calculated from the resin loading. This may not truly represent the success of the synthesis because during chromatographic separations the fractions were selected on a basis of purity and not yield. The objective of the work described in this section was to use the Bnpeoc protected amino acids in conjunction with a variety of other reagents to prepare peptides from two proteins, Ubiquitin and phage λ Cro.

2.2.1 Monitoring the progress of peptide synthesis.

The most frequent point of failure in solid phase peptide synthesis is the coupling reaction and any indication of where a poor coupling occurs can be used to optimise the synthesis. An on-line direct method of measuring of the coupling reaction would be ideal but until recently[102,103] it has proved elusive. However the piperidine adduct of dibenzofulvene is an excellent chromophore and its release from the peptidyl resin during the N-terminal deprotection step can be monitored at 314 nm and the automated synthesiser has been adapted to measure this[104]. During each cycle, after removing the Fmoc group with 20% piperidine in DMF, the effluent from the reaction vessel is passed through a UV cell, monitoring at 315 nm and the absorbance displayed on a chart recorder (figure 15).

Because the deprotection reaction is usually very efficient (with respect to the coupling reaction) the integrated peak area can give a rough indication (+/- 5%) of the efficiency of the previous coupling reaction. This system has been further adapted to allow the monitoring of syntheses carried out using the Bnpeoc protected amino acid derivatives by changing the monitoring wavelength, firstly to 365 nm and subsequently to 405 nm.

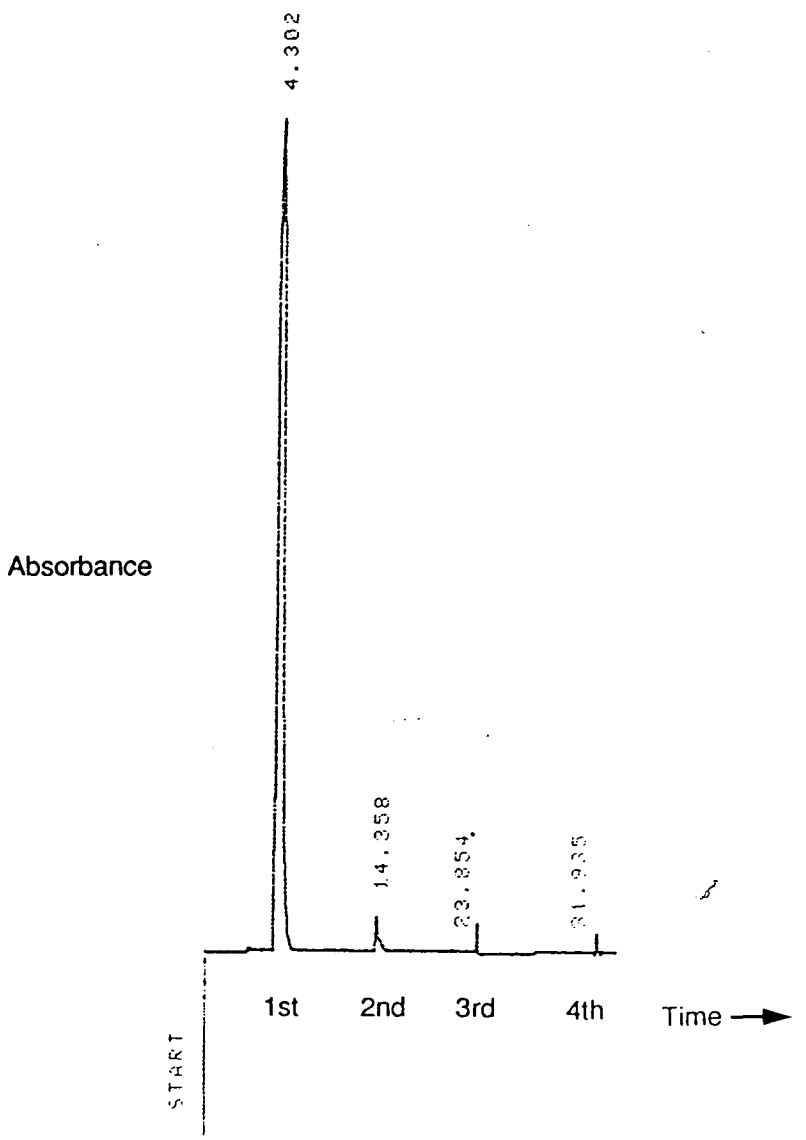


Figure 15. The cleavage of the Fmoc group, monitored at 315 nm. During each cycle the peptidyl resin is washed and vortexed four times with 20% (v/v) piperidine in DMF (for 3, 3, 3 and 1 minutes respectively). After each wash (labelled 1st to 4th) the effluent from the reaction vessel is passed through a UV absorbance detector.

2.2.2 Preparation of the loaded resins.

Target peptides with C-terminal amides were synthesised using the dibenzosuberanyl derivative (5)[27] (figure 1, page 8) as a linker. This is cleaved under conditions compatible with the final deprotection used in the Fmoc/Bu^t strategy. For preparation of peptides with C-terminal carboxylic acids, the first residue was coupled to the *p*-alkoxybenzylalcohol resin outwith the automated peptide synthesiser. Two methods of coupling were used. In the first case, Fmoc-Gly-OH was coupled successfully (35%) using DIC to prepare the symmetrical anhydride in DMF. DMAP was also required (0.2 equivs) and a final loading of 0.32 mmol/g was obtained. However this method required the use of a six fold excess of the amino acid derivative and such large excesses are routinely required to form the benzyl ester.

The second amino acid derivative to be loaded was Bnpeoc-Gln(Mbh)-OH, which had to be synthesised from its parent amino acid. Such a wasteful method of loading was unacceptable and an alternative method was used. The diphenylphosphinic mixed anhydride (scheme 12, page 20) was prepared from diphenylphosphinic anhydride (prepared by the method of Gray[33]) and then coupled to the resin in the presence of DMAP (0.1 mmol) in an excellent yield (92%).

Comparing the loading of Fmoc-Gly-OH and Bnpeoc-Gln(Mbh)-OH, it is clear that, despite the bulky side chain protecting group and the use of fewer equivalents of amino acid, the diphenylphosphinic anhydride activation resulted in a more efficient coupling reaction.

It is desirable that only the most unhindered fraction of all the available functionalised sites within the resin matrix are acylated, so that there is a large volume for the peptide chain to grow into, particularly during the synthesis of longer sequences. However, this incomplete acylation of the resin leaves available the remaining sites for acylation during subsequent couplings. In order to stop this occuring the loaded resins were each capped using benzoyl chloride, by the method described by Wang[21].

2.2.3 Peptides from Ubiquitin.

The small protein called Ubiquitin has 76 amino acids (figure 16) and has been the target of a substantial synthetic effort.[105] A large number of peptides from this protein have been synthesised[33,36], with the objective of understanding how each of the structural elements contribute to the stable tertiary structure. Some of the peptides were synthesised as C-terminal amides and with the N-terminus acetylated, since the charges of the terminal carboxylate and ammonium ions (which are not present in the natural protein sequence) might distort the structure of the peptide.

```
   1
H-Met-Gln-Ile-Phe-Val-Lys-Thr-Leu-Thr-Gly-

   11
   Lys-Thr-Ile-Thr-Leu-Glu-Val-Glu-Pro-Ser-

   21
   Asp-Thr-Ile-Glu-Asn-Val-Lys-Ala-Lys-Ile-

   31
   Gln-Asp-Lys-Glu-Gly-Ile-Pro-Pro-Asp-Gln-

   41
   Gln-Arg-Leu-Ile-Phe-Ala-Gly-Lys-Gln-Leu-

   51
   Glu-Asp-Gly-Arg-Thr-Leu-Ser-Asp-Tyr-Asn-

   61
   Ile-Gln-Lys-Glu-Ser-Thr-Leu-His-Leu-Val-

   71
   Leu-Arg-Leu-Arg-Gly-Gly-OH
```

Figure 16. The sequence of Ubiquitin.

Three small β-turn regions from the molecule (residues 6-12, 44-49 and 50-56) and a region spanning two of the turns (residues 36-53) provided an excellent opportunity to test both the Bnpeoc group and the diphenylphosphinic anhydride coupling methods.

Ubiquitin residues (36-53). (53)

H-Ile-Pro-Pro-Asp-Gln-Gln-Arg-Leu-Ile-Phe-Ala-Gly-Lys-Gln-Leu-Glu-Asp-Gly-OH

This synthesis was carried out to test the efficiency of the diphenylphosphinic anhydride coupling reagent[33] in conjunction with the Fmoc/Bu^t protection strategy[10,11]. The Fmoc amino acids were double coupled using two equivalents of the mixed anhydride. The diphenylphosphinic anhydride was loaded on the synthesiser using the small cartridges that are used to deliver the Fmoc amino acid derivatives. Glycine was only single coupled with four equivalents of the mixed anhydride. After cleavage and deprotection the crude peptide (figure 17) was purified by preparative h.p.l.c. in a yield of 45%.

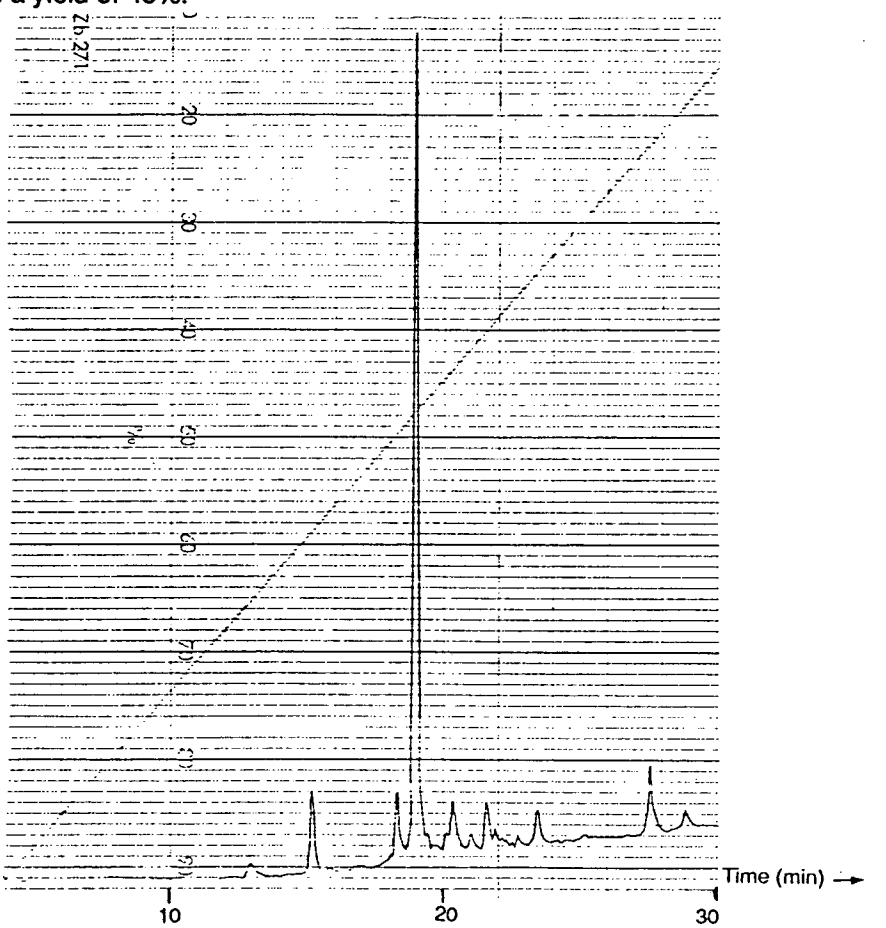


Figure 17. H.p.l.c. trace of the crude product from the synthesis of Ubiquitin residues (36-53).

Initial n.m.r. studies of the peptide in 90% H_2O/10% D_2O resulted in the identification of all the amino acid spin systems. However, these studies also indicated that the peptide was probably not adopting a stable secondary structure and the n.m.r. studies were stopped at this point.

Ubiquitin residues (44-49) (54)

H-Ile-Phe-Ala-Gly-Lys-Gln-OH

The peptide was synthesised using diphenylphosphinic anhydride as the coupling reagent and Bnpeoc protected amino acid derivatives.. Each residue was double coupled with two equivalents of mixed anhydride (except Gly which was single coupled with four equivalents). DMF was used as the solvent throughout the synthesis. The UV monitoring of the synthesis indicated that the last coupling reaction (Ile onto Phe) had not gone to completion. The poor coupling of β-branched amino acid derivatives has been previously observed[9] and so this might have been predicted. Furthermore, after cleavage and deprotection, a major and minor product were observed by h.p.l.c. (figure 18). These were each purified and both mass spectrometry and amino acid analysis indicated that the material from the minor peak was the capped (acetylated) derivative produced after the failure of the last coupling reaction.

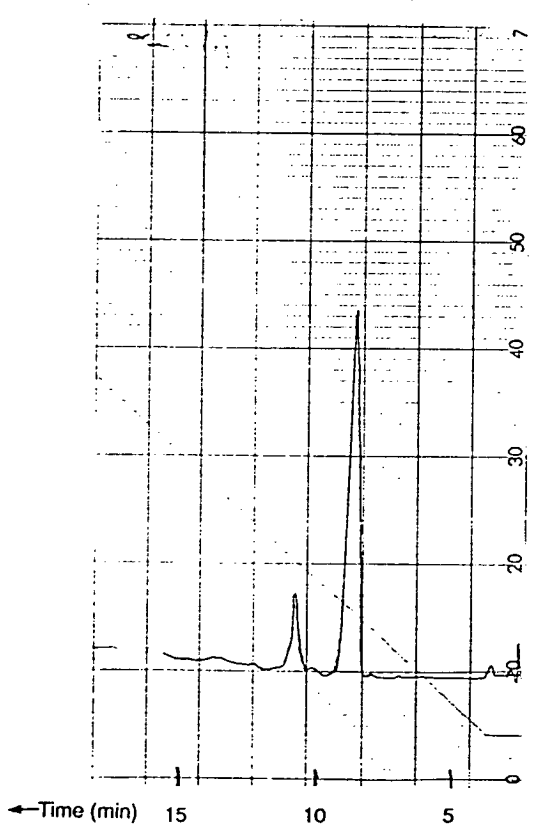


Figure 18. Analytical h.p.l.c. of crude hexapeptide (54).

Ubiquitin residues (6-12) amide (55)*and Acetyl Ubiquitin (6-12) amide.* (56)
(Ac- or H-)-Lys-Thr-Leu-Thr-Gly-Lys-Thr-NH_2

The synthesis was carried out on a 0.5 mmol scale using Bnpeoc amino acids and single two equivalent couplings with the BOP reagent[57,58] in dichloromethane. The synthesis proceeded with high efficiency (as judged by the UV monitoring of deprotection) and once the peptide had been assembled the last Bnpeoc group was removed (with 20% piperidine in DMF) and half the resin was reacted with acetic anhydride. The acetylated peptide was then cleaved from the resin and the side chains deprotected, purified by reverse phase h.p.l.c. and isolated in 50% yield. The remaining resin was cleaved and deprotected, then purified in a similar manner to give the purified peptide in 53% yield. During this synthesis the Bnpeoc amino acids (1.0 mmol) all dissolved readily in DCM (4 ml).

Ubiquitin residues (50-56) amide (57)*and Acetyl Ubiquitin (50-56) amide* (58).
(Ac- or H-) Leu-Glu-Asp-Gly-Arg-Thr-Leu-NH_2

The synthesis was carried out on a 0.5 mmol scale, double coupling with diphenylphosphinic anhydride activation[33] to give two equivalents of the mixed anhydride, except for glycine which was coupled only once with four equivalents of mixed anhydride. After the peptide had been assembled and the last Bnpeoc group had been removed, the peptide was divided into two portions. One half was deprotected and cleaved from the resin directly and the other half was acetylated after which it was cleaved and deprotected.

The crude heptapeptide (57) (isolated in 47% yield) was purified by reverse phase h.p.l.c. and obtained pure in the disappointing yield of 14%. This separation proved difficult because of the hydrophilic nature of the peptide. The acetylated derivative (58) (figure 19) was isolated in a similar manner, but with a slightly better yield of 22%. An impurity was also isolated in 8.5% yield and comparison of the mass spectra, amino acid analyses and ^1H n.m.r. indicated that the impurity was a *des*-Leu derivative. There are two Leu residues, but providing that the capping reaction used after each coupling is effective, this would seem to indicate that the coupling of the final Leu onto the preceeding Glu was incomplete.

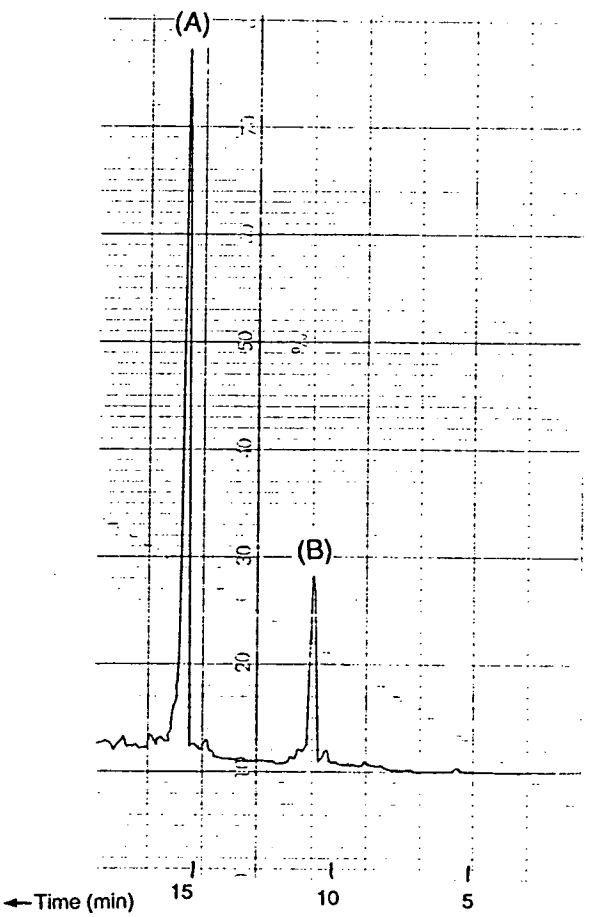


Figure 19. Analytical h.p.l.c. of crude heptapeptide (58). (A) is the required peptide and (B) is the impurity which lacks a Leu residue.

2.2.4 Synthesis of Peptides from the phage λ Cro protein.

The peptides which were selected from the Cro protein were all taken from the DNA binding region (the Cro sequence is shown in figure 12, page 32). In future n.m.r. studies it may be possible to study the interaction (if any) of these peptides with synthetic DNA operator site sequences.

The recognition helix Ac- Cro residues (25-40)-NH_2 (59).

Ac-Val-Tyr-Gln-Ser-Ala-Ile-Asn-Lys-Ala-His-Ala-Gly-Arg-Lys-Ile-NH_2

This peptide sequence contains the recognition helix of the Cro protein which binds to the major groove of the DNA operator site[69]. The three residue C-terminal extension,-Arg-Lys-Ile-, is thought to provide counterions to the negatively charged phosphate groups in the DNA sugar-phosphate backbone.

The synthesis was carried out manually in a 'bubbler' reaction vessel[106], on a 0.5 mmol scale with TBTU activation[59,107]. Single two equivalent couplings with Bnpeoc amino acids were monitored by the Kaiser test[108] and recoupled using BOP activation when required. The side chain functionality of the Gln was protected by Mbh[41].

The initial cleavage and deprotection of an aliquot the assembled peptide gave a mixture of products. The reaction was repeated and a small sample was withdrawn every half hour and subjected to analytical h.p.l.c. After two hours (figure 24A) a complex mixture was observed, but after a further hour (figure 24B) a major peak developed. Mass spectrometry of the material showed two major peaks near the molecular ion (m/z=1810), at m/z=1907 (MH^++97) which remains unidentified, and at m/z=2037 (MH^++227). This second peak is consistent with incomplete removal or scavenging of the Mbh protecting group. These ions were both reduced in intensity in the mass spectrum of material isolated after 3 hours of cleavage and deprotection. A larger batch of the material was then cleaved and deprotected and the crude peptide purified by cation exchange chromatography. The fractions containing the required material were pooled and further purified by reverse phase h.p.l.c. A poor yield (8.6%) was obtained as expected and reflects the problems of deprotection and the need for two chromatographic separations. The use of alternative carbocation scavengers (such a 1,2-ethanedithiol or phenol) might have reduced the problems at the deprotection stage and hence improve the yield.

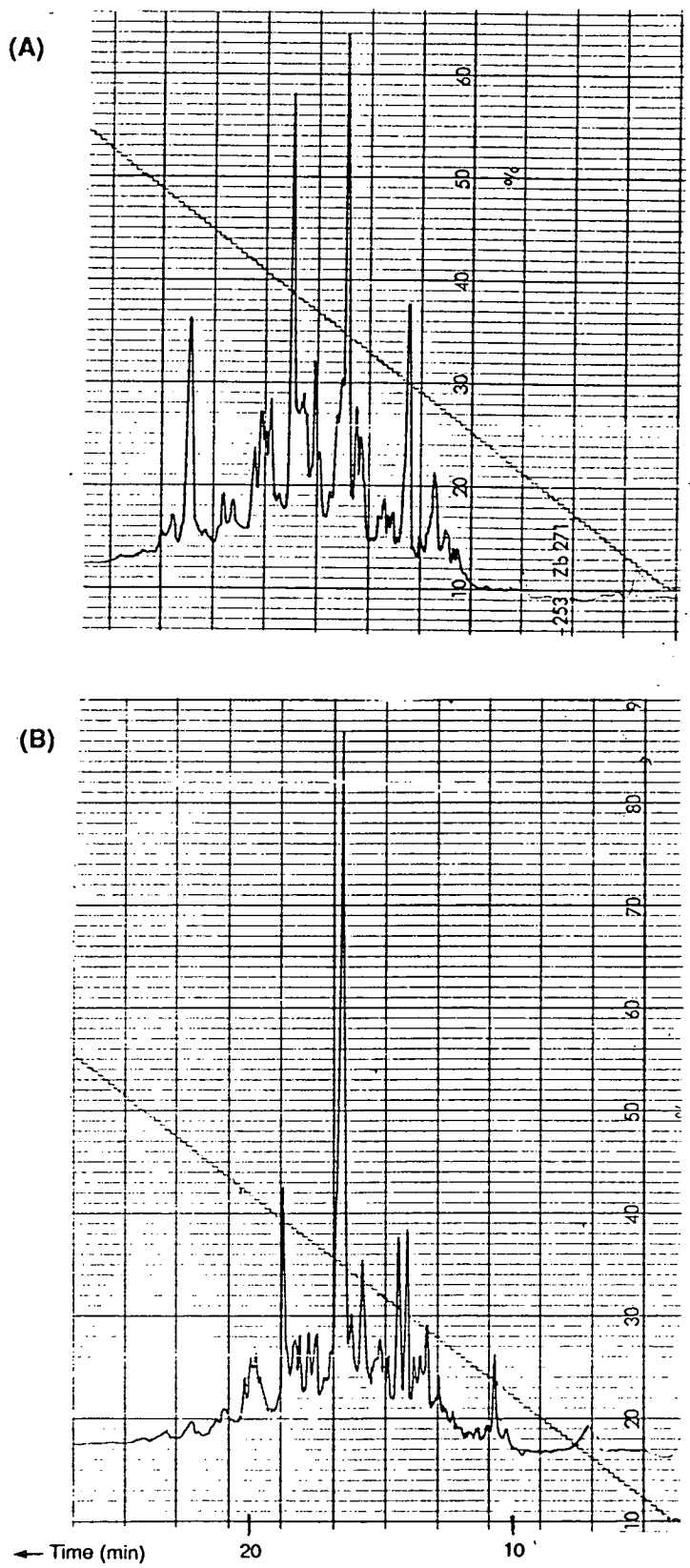
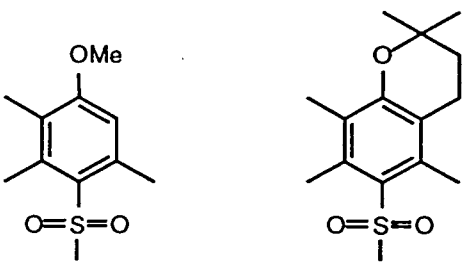


Figure 20. Analytical h.p.l.c. traces of crude (59) after 2 hours (A) and after 3 hours (B) cleavage and deprotection.

Attempted synthesis of the Cro protein.

The synthesis of the entire Cro sequence (66 residues, figure 12, page 32) was likely to be a demanding synthetic problem. The natural Cro protein was thought to be unstable to extended treatment with TFA[109] although nothing was known about the modification or degradation that occured. The problems experienced by Atherton *et al.*[96] were due, in a large part, to the conditions required (6 or more hours in TFA and scavangers) to remove the arginine protecting group then in use, Mtr (35).



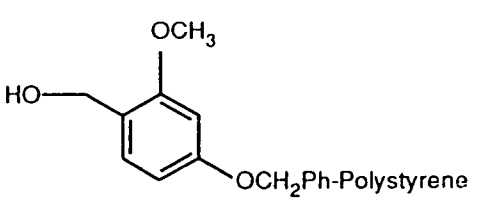
(35) (60)

Figure 21. The Mtr (35) and Pmc (60) arginine protecting groups.

The 2,2,5,7,8 pentamethylchroman-6-sulphonyl (Pmc) protecting group (figure 21) has subsequently been developed and has been shown to be cleaved at a rate (in under two hours) compatible with the Fmoc/Bu^t strategy[37]. This protecting group was thought more likely to allow the successful deprotection of the peptide, without an unacceptable degree of modification.

Experience obtained during the synthesis of Ubiquitin[105] had shown that the UV monitoring of the synthesis was useful in optimising the conditions used during the synthesis, particularly of longer (over approx. 25 residues) sequences.

To be set against these improvments in synthetic methods, the natural Cro protein is not commercially available, so it could not be used as a marker during chromatographic separations. Consequently even if the synthesis proceeded smoothly, the methods used to characterise the protein would be critical.



(61)

The synthesis was carried out on the 'Sasrin' linker (61), a linker which is very acid labile because of the methoxy group which is *ortho* to the benzyl ester. This allows the peptide to be cleaved from the resin under very mild conditions (1% TFA in DCM). The synthesis was carried out on a 0.25 mmol scale with Fmoc protected amino acid derivatives. Residues 66 to 29 were triple coupled (two equivalents) with TBTU activation[107]. However, the solubility of this reagent proved unreliable (varying from batch to batch) and this may have been responsible for several poor couplings. For the final part of the synthesis (residues 28 to 1) each amino acid was double coupled, firstly as the symmetrical anhydride (formed with DIC, two equivalents) and then with the HOBt ester (again formed with DIC, two equivalents).

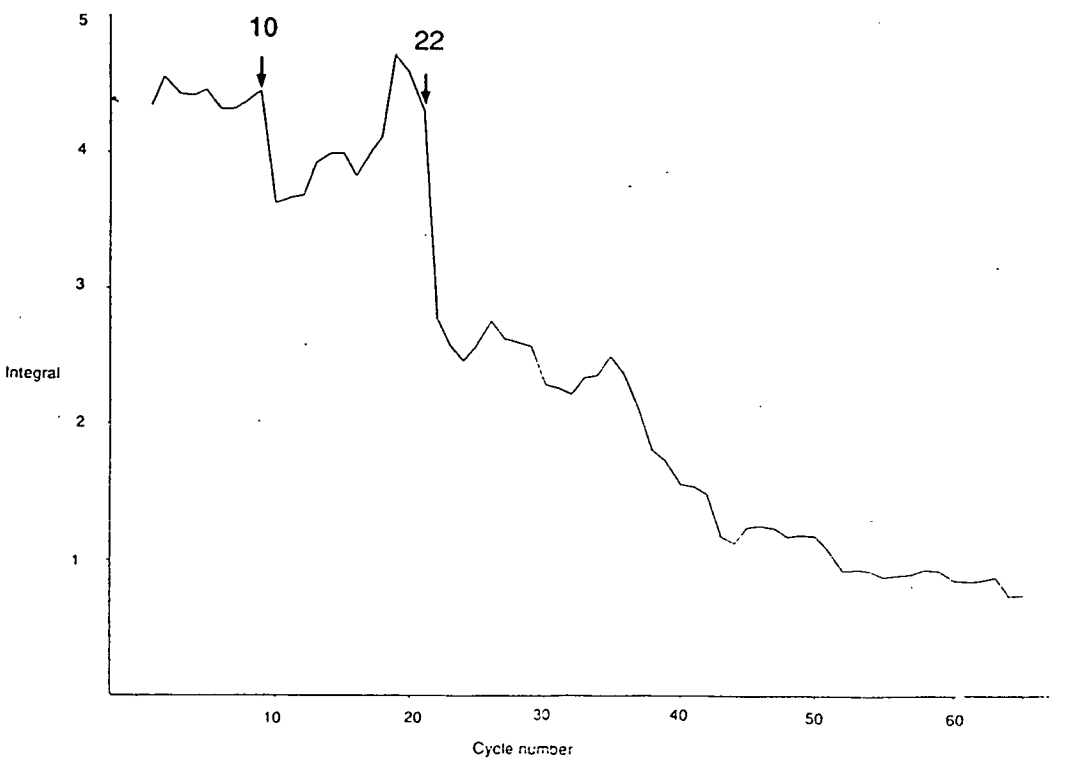


Figure 22. Plot of the UV monitoring peak area against cycle number.

The integral obtained from the UV monitoring of the synthesis was plotted against cycle number (figure 22). There were several incomplete couplings in the synthesis. The region between 8 and 16 residues has been reported as consistently difficult[9] and on cycle 10 (marked in figure 22, Pro coupled onto Phe) a substantial drop in the peak area occured. The recovery after this drop is perhaps even more worrying, since it suggests that the capping reaction did not block all the remaining amine sites left after the incomplete coupling. The deletion peptides from such a coupling failure are likely to be very similar to, and therefore difficult to separate from, the complete sequence.

At cycle 22 (Asn coupled onto Ala) a particularly poor coupling was observed. The reason for this remains unclear; it could be a result of the TBTU not dissolving or because of secondary structure formation in the growing peptide chain. The final peak area after the assembly had been completed was about 16% of the initial value. Although this is an inefficient assembly of the Cro sequence, it does serve to illustrate the usefulness of the monitoring method for indicating problems during a synthesis. Any subsequent syntheses can be optimised, with extended coupling times and recoupling of the residues that have been shown to be slow to couple. A slight change of solvent (such as a mixture of dioxan and DMF) and/or the adding of 'chaotropic' reagents[110] to disrupt secondary structure in the peptidyl resin are all possible methods for improving these couplings.

Since the Cro protein is unstable to TFA[109], only a small aliquot of the resin was cleaved and deprotected; samples were withdrawn from the mixture every 20 minutes and the crude peptide was rapidly isolated from each of these and then analysed by h.p.l.c. The optimal conditons [h.p.l.c. shown in figure 23, (A)] were obtained after 1 hour 40 minutes.

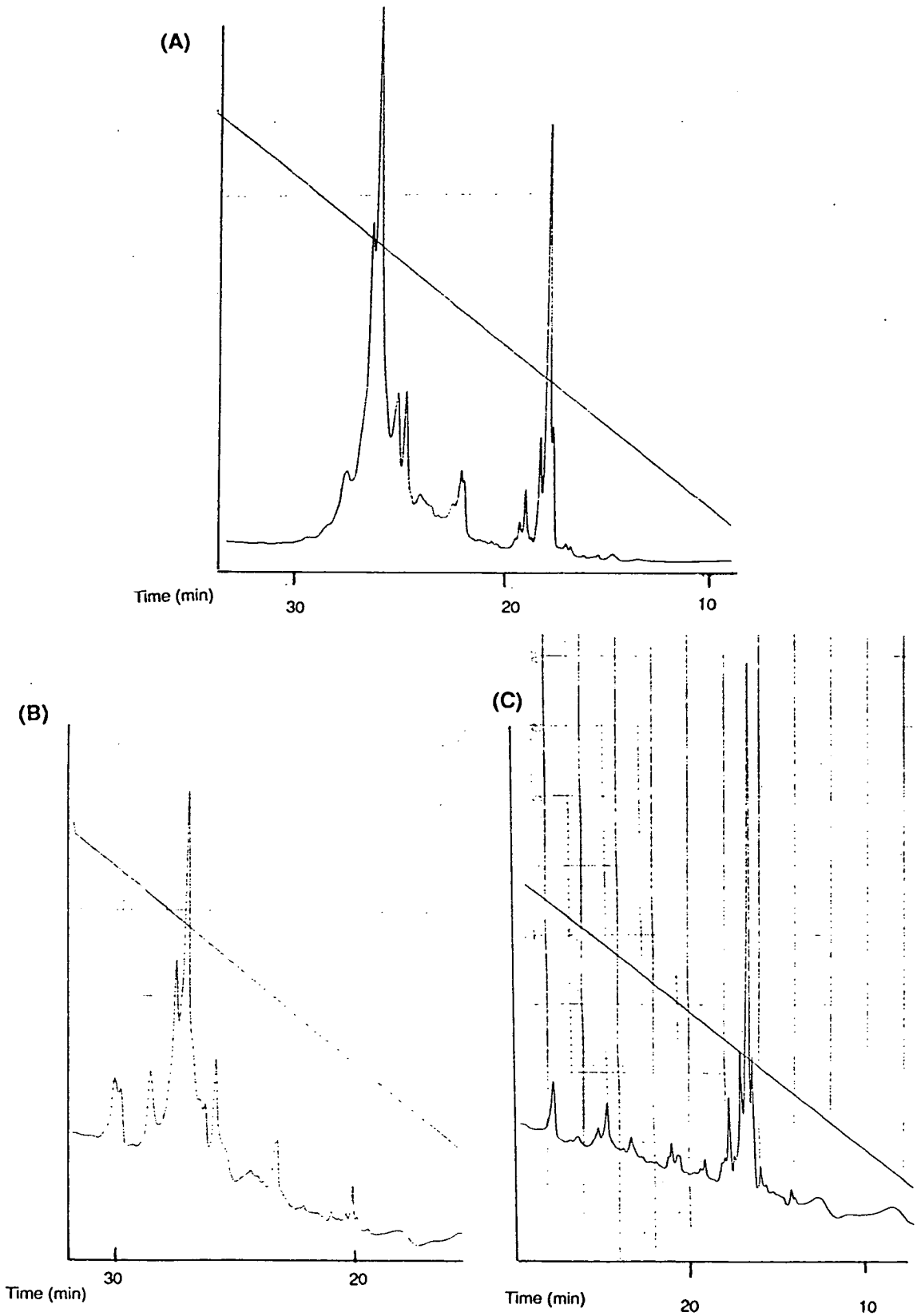


Figure 23. The analytical reverse phase h.p.l.c. of: (A) a sample of crude Cro protein after 1 hour 40 minutes; (B) the first eluted fraction from gel filtration chromatography; (C) the second eluted fraction from gel filtation chromatography.

Experiments were also carried out in which the balance of scavengers in the reaction mixture was varied. The ideal mixture was found to be: TFA/EDT/anisole/water in the ratio 10:1:1:1. A larger batch of resin was then cleaved from the resin and deprotected. The crude peptide was isolated by precipitation with diethylether. If the peptide was dried in this form and stored, it subsequently proved less soluble in aqueous media. It was therefore advantagous to redissolve the peptide in aqueous buffer immediately after its isolation. It was then applied to a Sephadex G50 column which allowed the separation of two major peptidyl fractions [figure 23, (B) and (C)]. Mass spectrometry and amino acid analysis indicated that the second eluted fraction (C) was the capped peptide produced by the incomplete coupling of cycle 22, identified with the UV monitoring method.

The other fraction was analysed by SDS polyacrylamide gel eletrophoresis. The major component had a molecular weight of approx. 6.2 kDa (as measured against standards of known molecular weight), which was compatible with the Cro protein (7.3 kDa). DNA binding proteins often have phosphate binding sites and a number have been purified by using phosphocellulose affinity chromatography. Phosphocellulose has been used for the purification of the Cro protein from natural sources[81], but gave very poor seperation of impurites from the synthetic Cro. Cation exchange chromatography using CM-Sephadex proved far more successful (figure 24) and the fractions containing the required material were pooled.

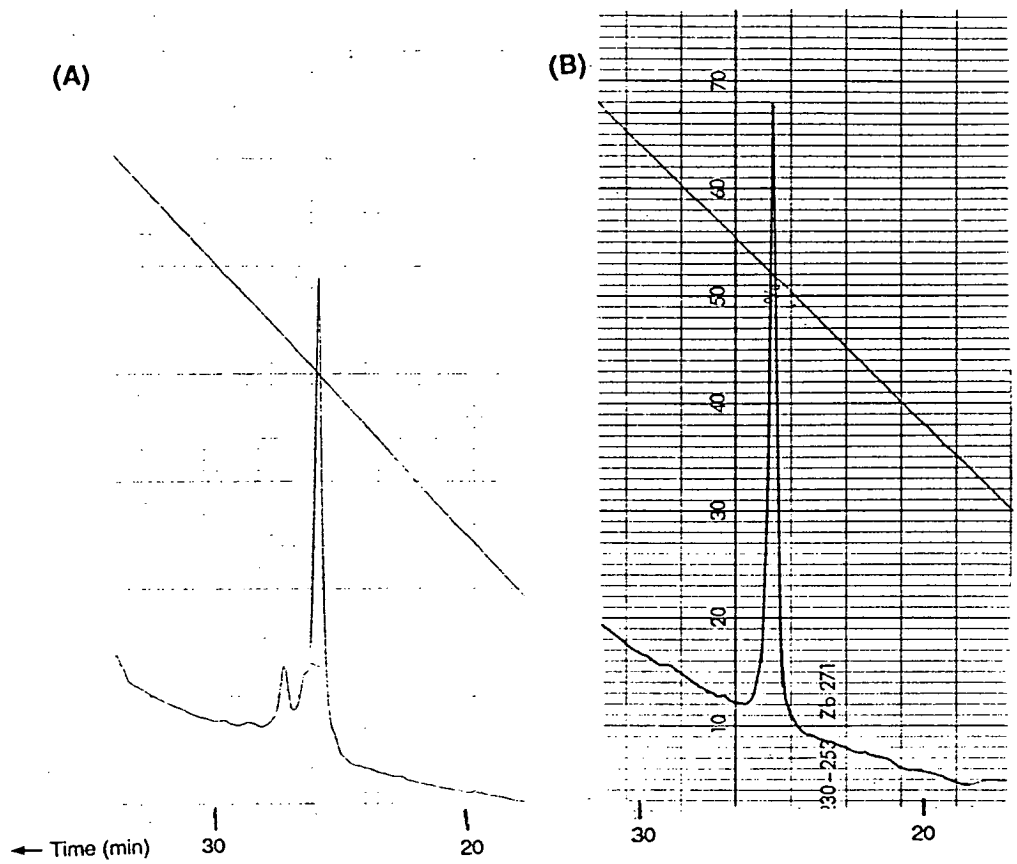


Figure 24. The major fraction after (A) cation exchange chromatography and (B) after reverse phase h.p.l.c.

An aliquot of this material was further purified by microbore h.p.l.c. to a single peak and the sequence of the first 10 residues was confirmed by automated Edman degradation[111]. Sequencing was stopped at this point until further characterisation of the purified material had been carried out. The remainder of the material was purified by preparative reverse phase h.p.l.c. and studied by ^{252}Cf plasma desorbtion mass spectrometry[112]. The mass spectrum showed a broad peak centred at 7366 (+/- 7) Da (Cro requires 7351 Da). This suggested that the sample contained a number of similar pepides and subsequent analytical h.p.l.c. indicated that the material had indeed undergone some modification since it had been purified (a period of about four weeks). This might be due to residual TFA in the h.p.l.c. buffers, or to the lyophilisation of the peptide from these buffers.

At this point the work was stopped (as the peptidyl resin from this synthesis had all been used). However, several useful points have been demonstrated:

1) The chemical synthesis of small proteins can be followed and subsequently optimised using the UV monitoring of the deprotection reaction.

2) A number of different methods of characterisation are useful in determining the nature and purity of an isolated peptide: this is particularly true when a sample of the natural material is not available.

3) A number of purification methods must be applied sequentially to obtain the peptide in a pure state.

4) Great care must be excercised when handling synthetic proteins. In particular, the use of lyophilisation to concentrate the protein or the use of preparative h.p.l.c. may have been the cause of some of the problems encountered in this purification.

Although the pure material isolated could not be well characterised it seems likely from the amino acid analysis, sequencing data and mass spectrometry, that it was either the Cro protein or a slightly modified analogue of it. In either case, these results suggest that after optimising the synthesis and improving the purification procedures, the synthesis of the Cro protein should be possible.

Synthesis of Acetyl Cro residues (15-37) amide (62)

Ac-Gly-Gln-Thr-Lys-Thr-Ala-Lys-Asp-Leu-Gly-Val-Tyr-Gln-Ser-Ala-Ile-Asn-Lys-Ala-Ile-His-Ala-Gly-NH_2

This peptide forms the helix-turn-helix motif of the Cro protein. As well as being of interest for structural studies, it was chosen as a test peptide for the comparison of synthetic methods. The same batch of resin was used for three syntheses, carried out using the methods shown in table 4.

Synthesis	Protection	Activation	UV monitor	Yield (%)
i	Fmoc	HOBt/DIC	314 nm	61
ii	Fmoc	TBTU	314 nm	31
iii	Bnpeoc	HOBt/DIC	405 nm	37

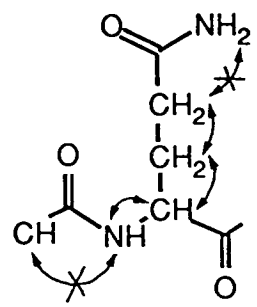
Table 4. Yield is estimated from the integrated peak areas observed using the UV monitoring of deprotection. The estimated error in these values is +/- 5%

The comparison of the TBTU and HOBt/DIC activation methods, with single five equivalent couplings, shows that the DIC/HOBt was clearly the better method. When a comparison is made using the Bnpeoc and the Fmoc groups, the Fmoc synthesis gives a clearly better yield. These results are confirmed by a comparison of the crude peptides by analytical h.p.l.c. (not shown). The Bnpeoc synthesis gave a large number of minor products, which accounts for the poor overall yield. It may be that the Bnpeoc group is slightly, (but in this case, critically) more sensitive to the basic impurities present in the DMF. Certainly, the syntheses of small Ubiquitin fragments (55) and (56) using Bnpeoc amino acids, which were carried out in DCM, gave crude products which contained far fewer impurities.

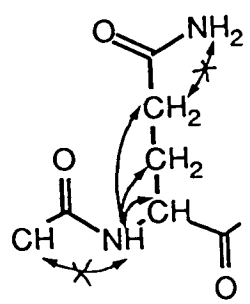
Considering the excellent solubility properties of the Bnpeoc amino acid derivatives in less polar solvents, such as DCM and dioxan, it would be interesting to carry out a comparative synthesis in one of these solvents. Since HOBt is not very soluble in these solvents, an alternative coupling reagent (for example HBTU or BOP) could be used. A direct comparison with Fmoc under these conditions would unfortunately be difficult because of the poor solubility of some Fmoc derivatives in these solvents.

The crude peptide from synthesis (i) was purified by h.p.l.c. in 38% yield. This material was used to study the secondary structure of the peptide by ^{1}H n.m.r. From the crystal structure of the Cro protein it is known that the first helix provides hydrophobic interactions to the recognition helix. In the case of the isolated peptide, such hydrophobic interactions might be mutually stabilising for both helices. In order to study the structure of a peptide or protein by n.m.r. a complete sequential assignment must be carried out. The two dimensional homonuclear techniques were used for this, in a strategy adapted[113] from that developed by Wüthrich[72].

(A) *Correlated Spectroscopy* (COSY)



(B) *Total Correlated Spectrosopy* (TOCSY)



(C) *Nuclear Overhauser Enhancement Spectroscopy* (NOESY)

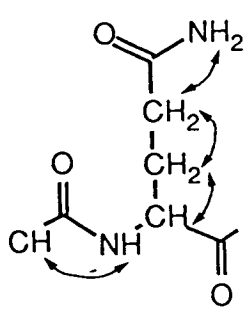


Figure 25. The interactions that can be observed in a glutamine residue by three two dimensional ^{1}H n.m.r experiments.

In a COSY spectrum, the observed cross peaks are usually for two (geminal) or three bond couplings. Such spectra cannot show couplings across the peptide bond, but instead show the isolated spin system of each amino acid. The TOCSY spectrum has similar constraints, except that it can show longer distance couplings provided that the intermediate carbon or nitrogen atoms are bonded to a hydrogen atom.

NOESY spectra show through space interactions and so this technique can show the interaction of the αCH of residue i and the nearby αNH of residue i+1, across the intervening peptide bond. This therefore allows the spin systems of the amino acids which have been identified by COSY and TOCSY to be linked together and hence sequentially assigned.

The NOESY spectrum can also provide structural constraints, which in conjunction with restrained molecular dynamics can produce a high resolution structure of the protein or peptide.

Initial studies of the peptide in aqueous solution suggested that the peptide had no stable secondary structure. The use of methanol (which is known to promote secondary structure formation) produced spectra which indicated the formation of secondary structure and therefore a more detailed analysis was carried out.

The spin systems of each residue were usually identified from the TOCSY spectrum and then cross referenced to the COSY spectrum. The TOCSY spectrum was particularly useful in assigning the γ, δ and ε protons in extended amino acid side chains. However, the side chains of the Ile, Val and Leu residues were more easily assigned from the less crowded COSY spectrum. At this stage the Ser, Asp, Asn, His, Phe and Tyr systems were just classified as AMX.

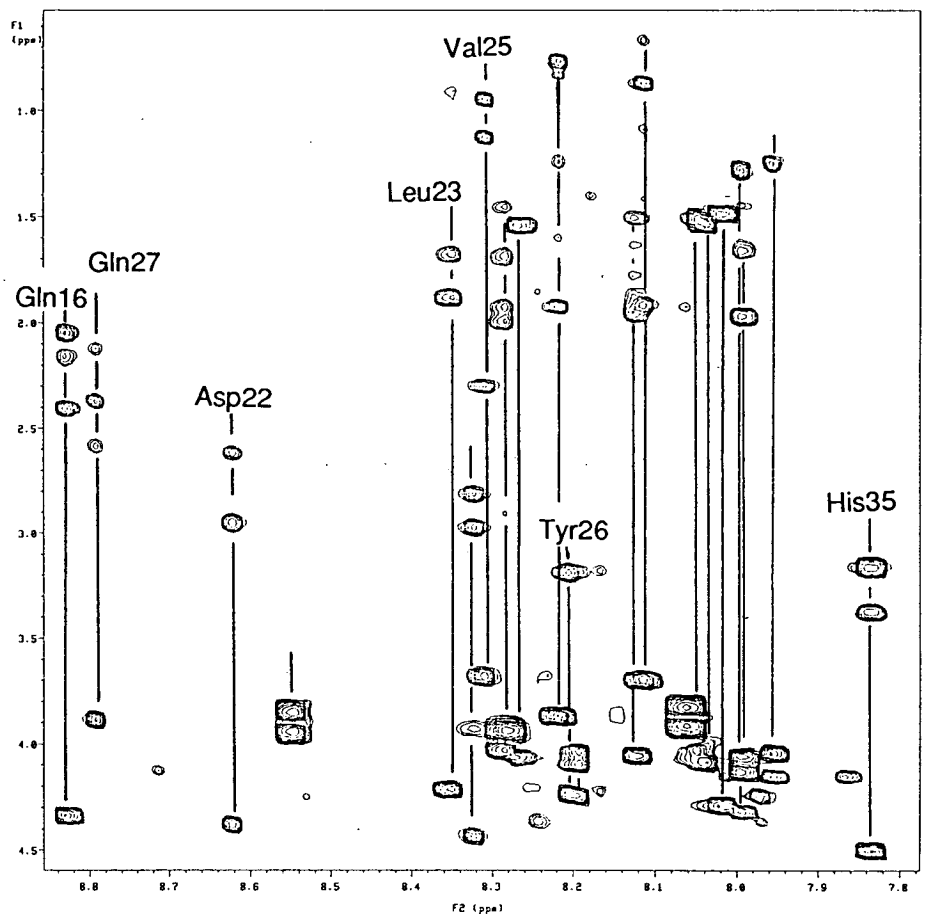


Figure 26. A region of the TOCSY spectrum of (62) showing the couplings to amide protons.

The sequential assignment was carried out using the same region as shown in figure 27 but from the NOESY spectrum. Initially pairs of residues showing clear cross peaks were linked, and then the sequence extended along the chain in both directions. In regions where the αCH_i to NH_{i+1} cross peaks overlapped other cross peaks, other interactions, such as βCH_i to NH_{i+1} cross peak was used. These cross peaks were also used to check other assignments wherever possible. This provided a complete sequential assignment, a part of which is shown in figure 27.

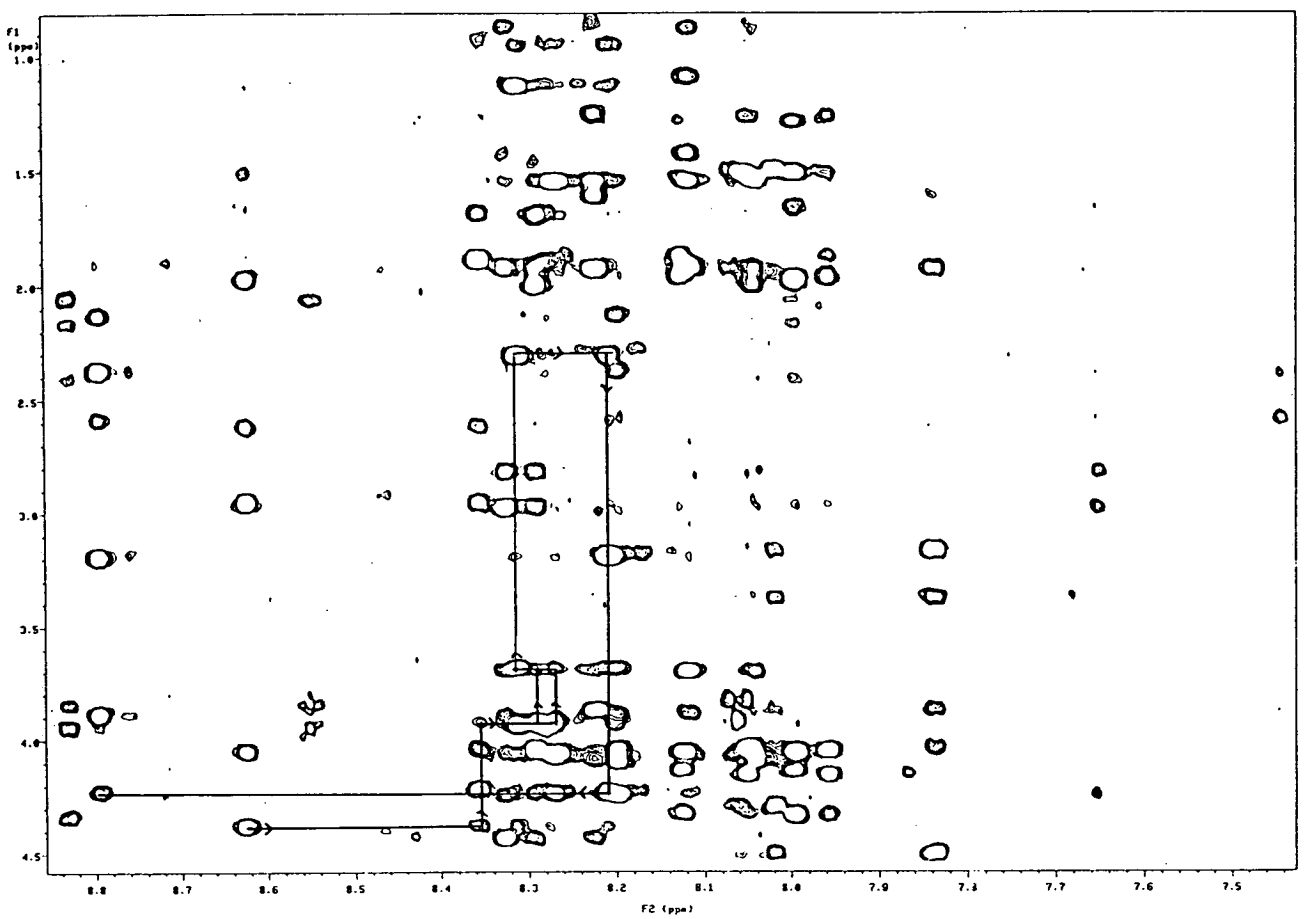


Figure 27. Part of the sequential assignment of peptide (62) using a region of the NOESY spectrum.

The last requirements were the assignment of the side chain amides of Gln 16 and 27 (via the NOESY crosspeak of the γCH_2 and the δNH_2), the side chain amide of Asn 31 (via a similar NOESY cross peak between the βCH_2 and the γNH_2) and the C-terminal amide (assigned by elimination of all the other peaks). The 3,5-H and 2,6-H aromatic resonances of Tyr were differentiated by the NOESY cross peak from the Tyr βCH_2 to the 2 and 6-H. The 2-H of His showed a NOESY cross peak to the βCH_2 and a TOCSY cross peak to the 4-H.

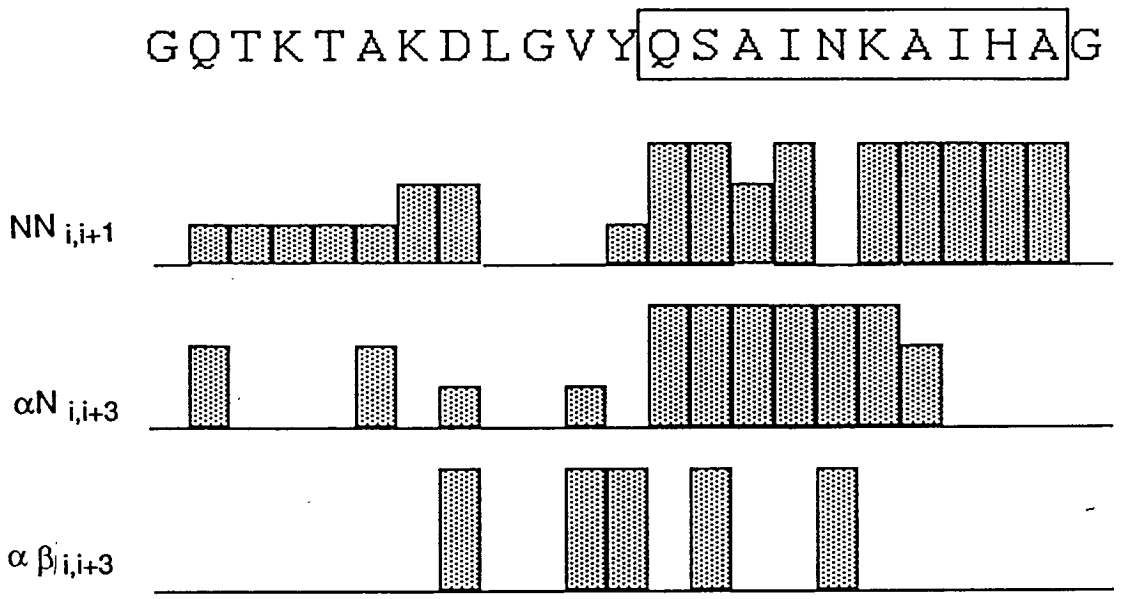


Figure 28. NOESY cross peaks indicative of an α-helix for the peptide (62). Cross peaks between sequential NH's ($NN_{i,i+1}$), between the αCH of residue i and either the NH of residue i+3 ($\alpha N_{i,i+3}$) or βCH of residue i+3 ($\alpha\beta_{i,i+3}$) are shown. They were classified by eye as strong, medium or weak. The proposed α-helix is boxed.

Once the sequential assignment had been carried out, some of the cross peaks to amide resonances were identified. Each type of secondary structure has a characteristic pattern of associated NOESY cross peaks. The observed pattern (figure 28) is consistant with the formation of an α-helix between residues 27 and 36, which is required for the recognition helix of the Cro protein. Many of the peaks for residues 15 to 25 overlap in the NOESY spectrum and it is difficult to determine the secondary structure for this region. The residues that would be expected to form the 'turn' show a few weak NOESY cross peaks but no clear structure has yet been determined.

The one dimensional ^1H n.m.r. spectrum was also studied in mixtures of water and methanol (a region of these spectra are shown in figure 29). Of particular interest is the γ methyl resonances of Val 25 (marked with the arrow in figure 29). The position in water was unambiguously assigned from a TOCSY experiment in water. The substantial shift of this resonance (from 0.85 to 1.14) on changing the solvent from water to methanol could be caused by a change in the relative position of the Val γMe to the aromatic ring of Tyr 26 (with its associated ring current). Alternatively, the formation of the α-helix would result in the generation of a helix dipole with the Val positioned at the positive end. Interestingly, the peak shape of the Val 25 γMe resonance becomes broader in both 10 and 30% H_2O in MeOH. It is sharp in both methanol and water and in all the other mixtures.

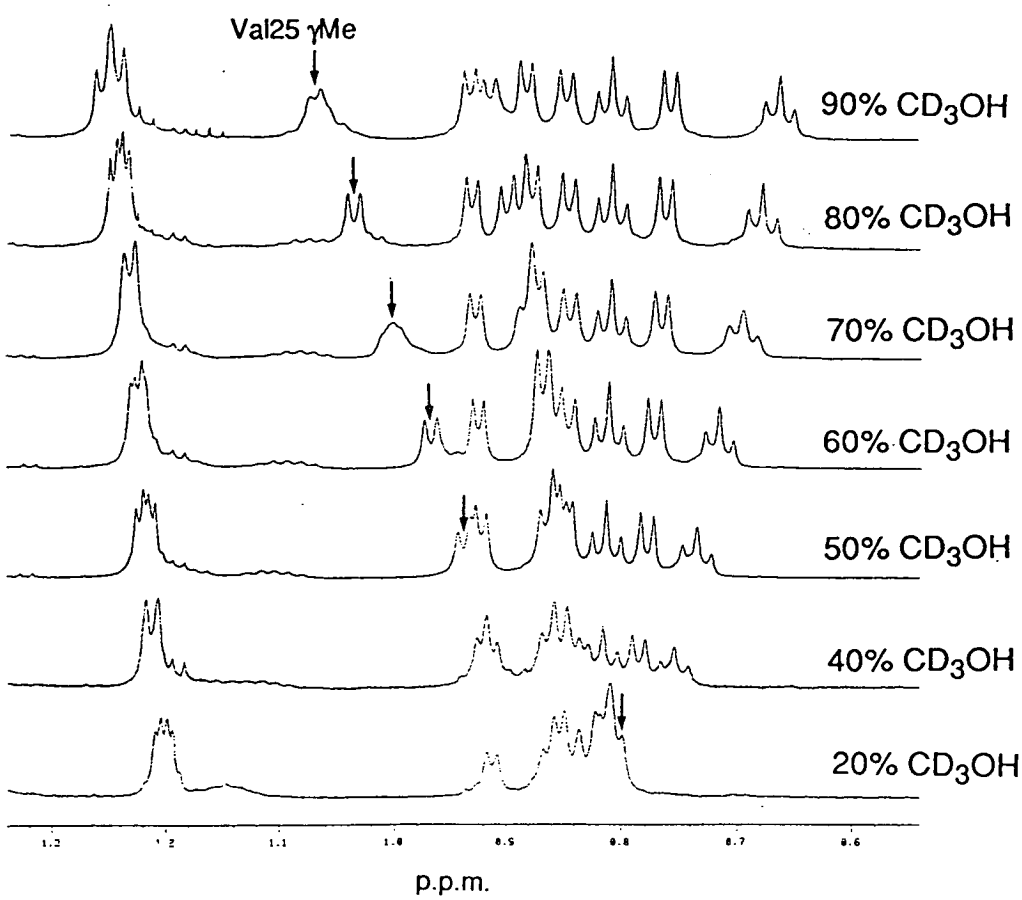


Figure 29. The 1D ^1H n.m.r. spectrum of peptide (62) in different proportions of water and CD_3OH.

Whilst these results are interesting, there are many more n.m.r. experiments that could be carried out. Of particular interest would be studies of peptide (62) bound to a synthetic DNA operator site, since this would provide a relatively simple model of the binding of the helix-turn-helix motif to DNA, which might be of use in predicting the properties of proteins other than Cro that carry this motif.

3. Experimental

3.1 Notes

All amino acids used were purchased from the SAS group of companies; protected derivatives were purchased from Bachem or Novabiochem and were used as supplied. Melting points were taken in open capillaries on an electrically heated Buchi 510 melting point apparatus, or on microscope slides on an electrically heated Reichart 7905 melting point apparatus and are uncorrected. Thin layer chromatography (tlc) was carried out on plastic sheets coated with silica gel 60GF-254 (Merck 5735) in the following systems:

(A) 1/5, ethyl acetate/light petroleum (b.p. 40-60°C)
(B) 1/4, ethyl acetate/light petroleum (b.p. 40-60°C)
(C) 30/70, ethyl acetate/light petroleum (b.p. 40-60°C)
(D) 40/60, ethyl acetate/light petroleum (b.p. 40-60°C)
(E) 18/2/1, chloroform/methanol/acetic acid.
(F) 16/4/1, chloroform/methanol/acetic acid.
(G) 19/1, ethyl acetate/acetic acid
(H) 2/1/1/1, ethyl acetate/n-butanol/acetic acid/water
(J) 3/1/1, n-butanol/acetic acid/water
(K) 1/1/1, cyclohexane/ethyl acetate/methanol

Visualisation of the compounds was achieved by using the following methods as appropriate: iodine vapour, ultra-violet absorbtion at 254 nm, Mary's reagent [4,4'-*bis*(dimethylamino)diphenylcarbinol], acidified potassium permanganate and ninhydrin. H.p.l.c was carried out using Applied Biosystems equipment (2 x 1406A solvent delivery systems, a 1480A injector mixer and a 1783A detector controller), or its equivalent. The gradients that were used are summarised in table 5. The sample was injected, followed by the isocratic and then the gradient phases.The elution of the sample was monitored at 214 or 229 nm.

Gradient	Flow ml/min	Isocratic %A	Time (Min)	Gradient %A (1)	(2)	Time (Min)	Column	Size mm
A	1.0	10	4	10	90	25	ABI RP18	220 x 4.6
B	3.6	30	5	30	50	20	ABI RP18	100 x 10
C	1.0	0	4	0	100	30	ABI RP18	100 x 4.6
D	4.0	5	4	5	15	15	ABI RP18	250 x 10
E	1.0	5	4	5	25	20	ABI RP18	220 x 4.6
F	5.0	5	4	5	35	20	ABI RP18	250 x 10
G	1.0	5	4	5	35	20	ABI RP18	220 x 4.6
H	1.0	5	4	5	40	30	Capital ODS-SB5	250 x 10
I	3.0	5	4	5	15	15	ABI RP8	250 x 10
J	4.0	10	4	10	50	20	ABI RP18	250 x 10
K	3.5	5	5	5	40	30	ABI RP18	250 x 10
L	1.0	5	5	5	50	20	Capital ODS-SB5	250 x 4.5
M	1.0	5	4	5	20	20	ABI RP8	220 x 4.6

Table 5. The conditions used for reverse phase h.p.l.c. Buffer A was 0.1% TFA in water. Buffer B was 0.1% TFA in acetonitrile. (1) gives the percentage of A at the start of the gradient and (2) gives the percentage of A at the end of the gradient.

Amino acid analyses were carried out on either a LKB 4150 alpha or a LKB 4151A alpha plus amino acid analyser, following hydrolysis in constant boiling hydrochloric acid at 110°C for an appropriate length of time (24-72 hours). Infra-red spectra were recorded on a Perkin Elmer 781 spectrophotometer in the solvent stated or as a nujol mull, using polystyrene (1603 cm^{-1}) as a standard. Ultra-violet spectra were recorded on a Varian Cary 210 spectrophotometer. Wavelengths are quoted in nanometers and extinction coefficients are quoted in $dm^3\ mol^{-1}\ cm^{-1}$. Optical rotations were measured using an Optical Activity AA1000 polarimeter. Optical rotation measurements of known compounds were performed under identical conditions (concentration and solvent) to those stated in the literature, unless otherwise noted. Mass spectra (F.A.B.) were measured on a Kratos MS 50TC spectrometer. Proton nuclear magnetic resonance spectra were recorded on Bruker WP80 (80MHz), WP200 (200MHz) or WH360 (360MHz) instruments using tetramethylsilane as the standard (δ=0.00), or on the Varian VXR600S (600MHz). Spectra recorded on this instrument were referenced either to H_2O (at 4.8 p.p.m. at 25°C) or to CD_2HOH (at 3.315 p.p.m. in CD_3OH). Elemental analyses were carried out on a Carlo Erba model 1106 elemental analyser. The following solvents were distilled before use and were dried using the reagents given in parentheses: chloroform (phosphorus pentoxide), dichloromethane (calcium hydride), diethyl ether (sodium wire), ethyl acetate, tetrahydrofuran (benzophenone/sodium), ethanol (magnesium ethoxide). Light petroleum (b.p. 40-60°C) refers to the fraction which boils between 40° and 60°C.

3.2 Preparation of side chain protected amino acid derivatives.

N^{α}-Benzyloxycarbonylasparagine
Z-Asn-OH

The *title compound* was prepared by the method of Boissonas *et al.*[114]. Benzylchloroformate (137 ml, 0.97 mol) was added to a cooled (ice bath) stirred solution of asparagine monohydrate (132 g, 0.8 mol) dissolved in $NaHCO_3$ (1.0 M, 1.5 l) over a period of two hours. This was allowed to warm to room temperature and stirred for 16 hours. The pH of the reaction mixture was tested (pH 7-8). The solution was then extracted with diethyl ether (2 x 500 ml) to remove excess benzylchloroformate, then adjusted to pH 3 with conc. hydrochloric acid (approx. 100 ml). The solution was left to stand at 4°C for 24 hours to allow the product to crystallise, after which it was collected by filtration, washed with water (2 x 300 ml) and dried *in vacuo* at 40°C to give the *title compound* (175.5 g, 82%) as a white solid: m.p. 157-159°C (lit.,[114] 163°C); $[\alpha]_D^{22}$ +8.7°(*c* 1.0, glacial acetic acid) (lit.,[4] +9.6°).

N^{α}-Benzyloxycarbonylasparagine benzyl ester
Z-Asn-OBzl

The *title compound* was prepared by the general method of Nefkens and Nirvard[115,116].
Z-Asn-OH (70.0 g, 0.26 mol) was dissolved in DMF (350 ml) and dicyclohexylamine (54 ml, 0.27 mol) was added. The stirred mixture was warmed to 65-70°C and benzylbromide (31 ml, 0.26 mol) was added. After 30 min the dicyclohexylammonium bromide which had precipitated was removed by filtration and the mixture poured onto crushed ice (1.2 l). The white precipitate which formed was filtered and dissolved in ethyl acetate. The solution was washed with $NaHCO_3$(aq.) (800 ml) and water (3 x 500 ml) and then dried ($MgSO_4$). The solvent was reduced *in vacuo* to a volume of about 200 ml and the product was left to crystallise for 24 hours. The solution was then filtered and the *title compound* dried *in vacuo* : (64 g, 81%); m.p. 127-128°C (lit.,[115] 131-132°C); $[\alpha]_D^{22}$ -14.1° (*c* 2.5 in DMF) (lit.,[115] -12.9°).

N -Benzyloxycabonyl-α-benzyl aspartic acid cyclohexylammonium salt
Z-Asp(OH,Cha)OBzl

The *title compound* was prepared by the general method of Nefkens and Nirvard[116].
Z-Asn-OBzl (40.0 g, 112 mmol) was dissolved in glacial acetic acid (550 ml) and stirred. HO_3SONO (21.4 g , 168 mmol) was then added in small portions (3-4 g) at intervals of five min. The solution was maintained at less than 25°C throughout, then stirred for one hour. Ethyl acetate (200 ml) was then added and the solution was boiled for three min. The solvent was reduced *in vacuo* to a small volume and then the residue was poured into water (1200 ml). This suspension was extracted with ethyl acetate (3 x 300 ml) and the combined organic phase washed with water (3 x 500 ml) and brine (500 ml), then dried ($MgSO_4$) and the solvent reduced *in vacuo* to a small volume (about 100 ml). Cyclohexylamine (13.75 ml, 1.0 equivs.) was then added and the *title compound* (35.24 g, 69%) precipitated out by the addition of diethyl ether in two crops as a white solid: m.p. 126-129°C; tlc -E, R_f 0.6

N -Benzyloxycarbonyl aspartic acid α-benzyl ester
Z-Asp-OBzl

Z-Asp(OH,Cha)-OBzl (30 g, 66 mmol) was suspended in $KHSO_4$ (2.0 M, 500 ml) and ethyl acetate (500 ml). After stirring for two hours, the ethyl acetate was separated and the aqueous phase was extracted with ethyl acetate (2 x 250 ml). The combined organic phase was then washed with water (2 x 500 ml) and with brine (1 x 300 ml). The ethyl acetate was then removed *in vacuo* and the residual oil was crystallised from diethyl ether/light petroleum (b.p. 40-60°C) to give the *title compound* (21.2 g, 90%) as a white solid: m.p. 82-83°C (lit.,[116] 84-85.5°C), tlc -E, R_f 0.6; δ_H(80 MHz; $CDCl_3$) 2.96 (2H, t, βCH_2), 4.67 (1H, m, αCH), 5.10, 5.15 (2 x 2H, 2 x s, Bzl and Z CH_2's), 5.90 (1H, d, αNH), 7.29, 7.31 (2 x 5H, Z and Bzl Ar-H), 9.86 (1H, br, βCO_2H).

N -Benzyloxycarbonyl aspartic acid β-*tert*.butyl benzyl ester
Z-Asp(OBu^t)-OBzl

The *title compound* was prepared by an adaptation of the method of Schwyzer and Dietrich[117].
Z-Asp-OBzl (30 g, 84 mmol) was dissolved in dichloromethane (300 ml) and the solution was poured into a thick walled flask, then cooled in an acetone/dry ice bath. Condensed isobutylene

(100 ml), conc. sulphuric acid (0.5 ml) and CuCl (50 mg) were then added. The vessel was closed securely and left to react for 14 days at room temperature, with occasional turning to mix the reagents. The flask was then cooled in an acetone/dry ice bath, opened and allowed to warm to room temperature, with evaporation of the excess isobutylene. The solution was then washed with Na_2CO_3 (1.0 M, 3 x 100 ml) and then the organic layer was concentrated *in vacuo*. Ethyl acetate (500 ml) was then added and the organic layer washed with water (3 x 400 ml) and brine (1 x 250 ml), dried ($MgSO_4$) and the solvent removed *in vacuo* to leave an oil (23.2 g, 67%). A small sample was crystallised from ethyl acetate/light petroleum (b.p. 40-60°C) as a solid: m.p. 42-45°C (lit.,[117] 45-47.5°C); δ_H(80 MHz; $CDCl_3$) 1.36 (9H, s, 3 x CH_3, Bu^t), 2.82 (2H, t, Asp βCH_2), 4.62 (1H, m, Asp αCH), 5.11 and 5.16 (2 x 2H, 2 x s, Z, Bzl CH_2), 5.71 (1H, d, Asp αNH), 7.31, 7.32 (2 x 5H, 2 x s, Bzl, Z Ar-H's).

Aspartic acid β-*tert*.butyl ester
H-Asp(OBu^t)-OH

Z-Asp(OBu^t)OBzl (22.35 g, 52 mmol) was dissolved in methanol (100 ml), cooled to 0°C and stirred under an atmosphere of nitrogen. Palladium on charcoal catalyst (2.2 g, 10% w/w) was then added in small portions. The mixture was then stirred under an atmosphere of hydrogen for 24 hours. The solution was then filtered through a pad of pre-washed 'Celite' to remove the catalyst and concentrated *in vacuo* to a volume of 50 ml. The addition of diethyl ether (100 ml) produced a white gel which was filtered and dried *in vacuo* at 40°C to give the *title compound* (7.63 g, 77%) as a white powder: m.p. 183-184°C (lit.,[26] 189-190°C); $[\alpha]_D^{22}$ +9.9° (*c* 1.02 in 90%AcOH) (lit.,[114] +8.5°)

N^{α}-Benzyloxycarbonyl-N^{γ}-[*bis*(4-methoxyphenyl)methyl]asparagine
Z-Asn(Mbh)-OH

The *title compound* was prepared by the method of König and Geiger[41]. Z-Asn-OH (30 g, 113 mmol) and *bis*(4-methoxyphenyl)methanol (27.5 g, 113 mmol) were suspended in glacial acetic acid (350 ml) and conc. sulphuric acid (3 drops) was added. The mixture was stirred for 24 hours, then poured into water (1000 ml) and extracted with ethyl acetate (4 x 400 ml). The combined organic phase was washed with water (3 x 1000 ml) and brine (1 x 500 ml), then dried ($MgSO_4$) and the solvent removed *in vacuo* to leave a white solid. Recrystallisation from tetrahydrofuran/ light petroleum (b.p. 40-60°C) gave the *title compound* (50.1 g, 90%) as a white solid: m.p. 166-171°C (lit.,[41] 176-180°C).

N^{γ}-[*bis*(4-methoxyphenyl)methyl]asparagine
H-Asn(Mbh)-OH

The *title compound* was prepared by the method of König and Geiger[41]. Z-Asn(Mbh)-OH (10.0 g, 20.4 mmol) was suspended in glacial acetic acid (100 ml) and stirred under an atmosphere of nitrogen. 10% Palladium on charcoal catalyst (5% w/w) was added in small portions. The suspension was then stirred under an atmosphere of hydrogen for 24 hours. The catalyst was then removed by filtering the solution through a pad of prewashed 'Celite'. The solvent was then removed *in vacuo* and the product triturated with sodium acetate (1.0 M, 100 ml). The solidified product was filtered, washed with water and dried to give a pale grey solid (5.5 g, 75%); m.p. 215-220°C (lit.,[41] 215-217°C); tlc -E, R_f 0.23, -J, R_f 0.47.

N^{α}-Benzyloxycarbonylglutamine
Z-Gln-OH

The *title compound* was prepared by the method of Boissonas *et al.*[114]. Glutamine (50 g, 0.34 mol) was dissolved in $NaHCO_3$ (1.0 M, 1.0 l) and cooled to 0°C with stirring. Benzylchloroformate (50 ml, 0.35 mol) was added dropwise over 30 min and the solution was then allowed to warm to room temperature. The solution was stirred for 24 hours, then extracted with diethyl ether (2 x 500 ml) and acidfied to pH 1.0 with conc. hydrochloric acid. The resulting precipitate was filtered, washed with cold water, dried *in vacuo* at 40°C for 24 hours to give the *title compound* (66.9 g, 70%) as a white solid which was used without further purification: m.p. 127-128°C (lit.,[114]135°C).

N^{α}-Benzyloxycarbonyl-N^{δ}-[*bis*(4-methoxyphenyl)methyl]glutamine
Z-Gln(Mbh)-OH

The *title compound* was prepared by the method of König and Geiger[41]. Z-Gln-OH (56 g, 0.2 mol) was suspended in glacial acetic acid (500 ml) and solid *bis*(4-methoxyphenyl)methanol (49 g, 0.2 mol) was added. Conc. sulphuric acid (0.5 ml) was then added dropwise. The mixture was stirred for 20 hours and then poured into water (1.5 l) and extracted with ethyl acetate (3 x 500 ml). The combined organic phase was washed with water (3 x 700 ml) and brine (1 x 500 ml), dried (Na_2SO_4) and the solvent removed *in vacuo*. The residue, a colourless oil, was then crystallised

from THF/light petroleum (b.p. 40-60°C) to give the *title compound* (48.9 g, 48%) as a white solid: m.p. 111-115°C (lit.,[41] 117-120°C); $[\alpha]_D^{22}$ +2.4° (*c* 2.0 in DMF) (lit.,[41] +2.43°); δ_H [80 MHz; $(CD_3)_2SO$] 2.51 (2H, m, Gln βCH_2), 2.60 (2H, m, Gln γCH_2), 3.69, 3.71 (2 x 3H, 2 x s, 2 x OCH_3, Mbh), 4.30 (1H, m, Gln αCH), 5.03 (2H, s, Z CH_2), 6.0 (1H, d, Gln αNH), 6.75-8.25 (8H, m, Mbh Ar-H), 7.33 (5H, s, Z Ar-H), 8.60 (1H, d, Gln γNH).

N^{δ}-[*bis*(4-Methoxyphenyl)methyl]glutamine
H-Gln(Mbh)-OH

The *title compound* was prepared by the method of König and Geiger[41]. Z-Gln(Mbh)-OH (25 g, 49.4 mmol) was dissolved in a mixture of glacial acetic acid (150 ml) and methanol (150 ml) and the solution was cooled to 0°C. 10% Palladium on charcoal catalyst (1.25 g, 5% w/w) was added under an atmosphere of nitrogen. The mixture was then stirred for 16 hours under an atmosphere of hydrogen at room temperature. The catalyst was then removed by filtration through a pad of prewashed 'Celite'. Diethyl ether (200 ml) was added to the solution to precipitate the *title compound* (11.9 g, 65%) as a grey solid: m.p. 200-202°C (lit.,[41] 205-206°C); $[\alpha]_D^{22}$ +9.2° (*c* 2.0 in glacial acetic acid) (lit.,[41] +7.5°).

N -(Benzyloxycarbonyl)glutamic acid
Z-Glu-OH

The *title compound* was prepared by the method of Goldschmidt and Jutz[118]. Glutamic acid (73.4 g, 0.5 mol) was dissolved in a mixture of $NaHCO_3$ (1.0 M, 500 ml) and NaOH (4.0 M, 200 ml) and the stirred solution was cooled to -5°C. Benzylchloroformate (125 ml, 150.4 g, 0.88 mol) was added dropwise over 30 min and the solution was then allowed to warm to room temperature and stirred for 16 hours. The solution was washed with diethyl ether (2 x 200 ml) and adjusted to pH 2 with conc. hydrochloric acid. The precipitate which formed was filtered, washed well with water and dried *in vacuo* at 40°C for 24 hours., The crude product was recrystallised from ethyl acetate/ light petroleum (b.p. 40-60°C) to give the *title compound* (110.3 g, 79%) as a white solid: m.p. 117-120°C (lit.,[118] 120-121°C); tlc -J, R_f 0.75.

Benzyl-N -(benzyloxycarbonyl)glutamic acid dicyclohexylammonium salt
Z-Glu(OH,Dcha)-OBzl

The *title compound* was prepared by the method of Nefkens and Nirvard[116]. Z-Glu-OH (80.2 g, 285 mmol) was dissolved in DMF (100 ml) and the solution stirred at -5°C. Triethylamine (42 ml, 302 mmol) was added, followed by benzylbromide (40 ml, 336 mmol) and the solution was allowed to warm to room temperature. The mixture was stirred for 16 hours, then poured into water (500 ml) and the oil extracted with ethyl acetate (3 x 400 ml). The combined organic phase was washed with water (3 x 500 ml) and brine (500 ml), then dried ($MgSO_4$) and concentrated *in vacuo* to a volume of about 500 ml. This solution was then cooled to 0°C and dicyclohexylamine (63.5 ml, 318 mmol) was added. Portions of diethyl ether were added to keep the suspension mobile while the product precipitated. The product was filtered, washed in diethyl ether (2 x 100 ml) and recrystallised from ethanol to afford the *title compound* (104 g, 66%) as a white solid: m.p. 160-163°C (lit.,[116] 162-164°C).

Benzyl-N -(benzyloxycarbonyl)glutamate
Z-Glu-OBzl

Z-Glu(OH,Dcha)-OBzl (60 g, 108.5 mmol) was suspended in citric acid (aq.) (20% w/v, 1000 ml) and extracted with ethyl acetate (3 x 400 ml). The combined organic phase was washed with water (3 x 300 ml) and brine (500 ml), dried ($MgSO_4$) and the solvent removed *in vacuo* to give a yellow oil. The *title compound* (32.3 g, 81%) was crystallised from ethyl acetate/light petroleum (b.p. 40-60°C) as a white solid: m.p. 95-97°C (lit.,[119] m.p. 96.5-98.5°C); $[\alpha]_D^{22}$ -11.7 (*c* 3.15 in AcOH) (lit.,[119] -9.9°); δ_H (80 MHz; $CDCl_3$) 1.87-2.20 (2H, m, Glu βCH_2), 2.2-2.5 (2H, m, γCH_2), 4.43 (1H, m, Glu αCH), 5.08, 5.10 (2 x 2H, 2 x s, Z and Bzl CH_2's), 5.53 (1H, d, Glu αNH), 7.31 and 7.32 (2 x 5H, 2 x s, Z and Bzl Ar-H), 9.10 (1H, br, Glu CO_2H).

Benzyl-N -benzyloxycarbonyl-γ-*tert*.butylglutamate
Z-Glu(OBu^t)-OBzl

The *title compound* was prepared by the method of Morley[119]. Z-Glu-OBzl (30.0 g, 81 mmol) was dissolved in dichloromethane (200 ml) and the solution poured into a thick walled flask and cooled in a dry ice/acetone bath. Condensed isobutylene (100 ml) was then added, followed by conc. sulphuric acid (0.5 ml) and CuCl (50 mg). The vessel was closed and set aside for two weeks. The thick walled flask was then cooled in a dry ice/acetone bath, opened and allowed to warm

to room temperature, with evaporation of the excess isobutylene. The organic layer was then washed with Na_2CO_3 (1.0 M, 100 ml x 3) and the solvent removed *in vacuo*. The residual oil was then redissolved in ethyl acetate (500 ml), washed with water (3 x 500 ml) and brine (1 x 500 ml) and dried (Na_2SO_4) and the solvent removed *in vacuo*. The resultant oil solidified upon trituration with diethyl ether/light petroleum (b.p. 40-60°C) (200 ml, 2:1). This solid was then recrystallised from ethyl acetate/light petroleum (b.p. 40-60°C) to give the *title compound* (16.9 g, 47%) as a white solid: m.p. 46-48°C (lit.,[119] 49-51.5°C), tlc -K, R_f 0.8.

Glutamic acid γ-*tert*.butyl ester
H-Glu(OBu^t)-OH

Z-Glu(OBu^t)-OBzl (16.1 g, 38.8 mmol) was dissolved in DMF (100 ml) and cooled to 0°C. 10% Palladium on charcoal catalyst (1.5 g, 10% w/w) was added in small portions under an atmosphere of nitrogen. The mixture was then stirred under an atmosphere of hydrogen for 72 hours. The catalyst was then removed by filtration through a pad of pre-washed 'Celite' and the filtrate was concentrated to 15 ml *in vacuo*. The addition of diethyl ether (100 ml) precipitated the *title compound* (6.93 g, 88%) as a white solid: m.p. 180-181°C (lit., [119]190-190°C).

N^{α}-Benzyloxycarbonyl-N^{ε}-(*tert*.butoxycarbonyl)lysine dicyclohexylammonium salt
Z-Lys(Boc)-OH,Dcha

To a stirred solution of Z-Lys-OH[121] (28 g, 100 mmol) in water (100 ml) at 0°C was added solid $NaHCO_3$ (12.5 g, 1.5 equivs.). A solution of di-*tert*.-butyldicarbonate (26 g, 1.2 equivs.) in dioxane (150 ml) was then added dropwise over 10 min and the solution was diluted with dioxan (60 ml) and water (60 ml). The solution was then stirred at room temperature for 16 hours. The resulting solution was then diluted with water (400 ml) and acidified to pH 3 with $KHSO_4$ (2.0 M) and extracted with diethyl ether (3 x 600 ml). The combined organic phase was washed with water (3 x 700 ml) and brine (1 x 700 ml) and dried ($MgSO_4$). The solvent was removed *in vacuo* and the residual oil was dissolved in ethyl acetate (300 ml). Dicyclohexylamine (21.9 ml, 1.0 equivs.) was added dropwise over 20 min and the precipitate which formed was filtered, washed with ethyl acetate and dried *in vacuo* to leave the *title compound* (46.1 g, 87%) as a white solid: m.p. 155-156°C (lit.,[122]156-157°C); $[\alpha]_D^{22}$ +7.1° (*c* 1.0 in EtOH) (lit.,[122] +7.82°); δ_H (80 MHz; $CDCl_3$) 0.9-2.1 (35H, br m, 3 x CH_3, Bu^t, 10 x CH_2, Dcha, Lys β, γ, δ CH_2), 2.8-3.1 (4H, m, 2 x CH, Dcha, Lys εCH_2), 4.0 (1H, q, Lys αCH), 4.62 (1H, br, NH, Dcha), 5.05 (2H, s, Z CH_2), 5.71 (1H, d, Lys αNH), 7.29 (5H, s, Z Ar-H), 7.25-8.6 (1H, br, Lys CO_2H).

N^{ε}-(*tert*.Butyloxycarbonyl)lysine

H-Lys(Boc)-OH

The *title compound* was prepared by the method of Scott *et al.*[123] Z-Lys(Boc)-OH,Dcha (9.07 g, 16.1 mmol) was partitioned between diethyl ether (400 ml) and $KHSO_4$ (2.0 M, 400 ml). The organic phase was separated and then washed with $KHSO_4$ (2.0 M, 2 x 100 ml), dried ($MgSO_4$) and the solvent removed *in vacuo* leaving a sticky foam. ^1H n.m.r was used to confirm the removal of all of the Dcha and the foam was then dissolved in MeOH (250 ml) and water (50 ml) and cooled to 0°C. 10% Palladium on charcoal catalyst (10% w/w) was added under an atmosphere of nitrogen and the suspension was then stirred under an atmosphere of hydrogen for three days. The resulting suspension was diluted with water (600 ml) and filtered through a pad of prewashed 'Celite' to remove the catalyst. The solvents were then removed *in vacuo* to leave a white solid which was dried *in vacuo* at 40°C for 24 hours (6.67 g, 74%); m.p. 259-260°C (dec.) (lit.,[124] 237-255°C); $[\alpha]_D^{22}$ +7.2° (*c* 1.0 in 2N NH_4OH) [lit.,[124] $[\alpha]_D^{22}$ +4.4° (*c* 0.88 in 2N NH_4OH)].

3.3 The preparation of 2,2-*bis*(4'-nitrophenyl)ethyl N-succinimidyl carbonate[125]

2,2-Diphenylethanol (41)

This compound was prepared by the method of Kharash and Clapp[98].
A single crystal of iodine was added to a suspension of magnesium (7.92 g, 0.33 mol) in dry diethyl ether (300 ml) under an atmosphere of argon, then bromobenzene (53.4 g, 0.34 mol) was added dropwise at a rate sufficient to maintain the reaction at reflux. When the addition was complete the mixture was heated under reflux for 30 min with stirring to produce the Grignard reagent, phenyl magnesium bromide.
The stirred Grignard reagent was then cooled to between 0 and 5°C in an ice bath and a solution of styrene oxide (37.2 g, 0.31 mol) in dry diethyl ether (150 ml) was added dropwise over 30 min. The solution was stirred for a further 30 min after which the solution was poured onto a mixture of sulphuric acid (2.0 M, 300 ml) and crushed ice (300 ml). The organic phase was separated and the aqueous phase extracted with diethyl ether (2 x 300 ml). The combined organic phase was washed with sodium bicarbonate solution (2.0 M, 2 x 500 ml), water (3 x 500 ml) and brine (1 x 500ml), then dried ($MgSO_4$) and the solvent removed *in vacuo*. The *title compound* was then crystallised from light petroleum (b.p. 40-60°C) as a white solid (114.85g, 58%): m.p. 60-61°C (lit.,[98] 56°C); (Found: C, 84.8; H, 7.1. Calc. for $C_{14}H_{14}O$: C, 84.8; H, 7.1%); tlc -A, R_f 0.61.

2,2-Diphenylethyl acetate (42)

2,2-Diphenylethanol (41), (5.0 g, 25 mmol) was added in small portions as a solid to acetic anhydride (2.57 g, 25 mmol) and then sulphuric acid (98%, 1 drop) was added and the mixture was stirred. The solid dissolved and the solution became warm. After 30 min the products were poured onto crushed ice (150 ml) and then extracted with ethyl acetate (2 x 100 ml). The combined organic phase was washed with sodium bicarbonate solution (2.0 M; 2 x 100 ml), water (2 x 100 ml) and brine (1 x 100 ml), then dried ($MgSO_4$) and the solvent removed *in vacuo* to afford the *title compound* as a yellow solid; recrystallised from diethyl ether (5.19 g, 85%): m.p. 54-55°C (lit.,[31] m.p. 54-55°C); tlc -B, R_f 0.31.

2,2-*bis*(4-Nitrophenyl)ethyl acetate (44)

A mixture of sulphuric (98%; 70 ml) and nitric (63%: 70 ml) acids were cooled, with vigorous stirring, to between -5 and 0°C in an ice/salt bath. 2,2-diphenylethyl acetate (54 g, 0.164 mol) was added in small portions to the mixture, over a period of one and a half hours, whilst the temperature was carefully maintained between -5 and 0°C by periodic additions of salt to the ice/salt bath. The mixture was then stirred for a further two hours and then poured onto crushed ice (500 ml) and extracted with ethyl acetate (3 x 400 ml). The combined organic phase was washed with sodium bicarbonate solution (2.0 M) until the pH of the aqueous layer was neutral. The organic phase was then washed with water (3 x 500 ml) and brine (1 x 500 ml) and dried ($MgSO_4$). The ethyl acetate was removed *in vacuo*, leaving a yellow oil which was crystallised from diethyl ether to afford the *title compound* (36.39 g, 49.1%) as a pale yellow solid: m.p. 110-111°C (lit.,[31] 107-108°C); (Found C, 58.0; H, 4.2; N, 8.4; Calc. for $C_{16}H_{14}N_2O_6$: C, 58.2; H, 4.4; N, 8.4%); tlc -C, R_f 0.31.

2,2-*bis*(4-Nitrophenyl)ethanol (37)

2,2-*bis*(4-Nitrophenyl)ethyl acetate (141.3 g, 0.43 mol) was dissolved in methanol (500 ml) and conc. hydrochloric acid (16.9 ml) was added dropwise. The mixture was then heated under reflux for five hours. The mixture was then poured onto crushed ice (500 ml) and the solid product was extracted into ethyl acetate (250 ml). The organic phase was separated and the aqueous phase extracted with ethyl acetate (2 x 250 ml). The combined organic phase was then washed with water (2 x 500 ml) and brine (1 x 500 ml), dried ($MgSO_4$) and the solvent removed *in vacuo*. The product was crystallised from chloroform/diethyl ether to afford the *title compound* (122.05 g, 99%) as a white solid: m.p. 104-105°C (lit.,[31] 109-110°C); (Found C, 58.4; H, 4.2; N, 9.7. Calc. for $C_{14}H_{12}O_5N_2$: C, 58.5; H, 4.2; N, 9.7%); tlc -D, R_f 0.25

2, 2-*bis*(4-Nitrophenyl)ethyl N-succinimidyl carbonate (31)

2,2-*bis*(4-Nitrophenyl)ethanol (64.77 g, 0.225 mol) was dissolved in dichloromethane (800 ml) and a solution of phosgene in toluene (1.93 M, 176 ml, 1.5 equivs.) was added. N-methylmorpholine (25 ml) was added dropwise to this mixture over a period of one hour, with stirring, under an atmosphere of dry nitrogen and then the N-methylmorpholine hydrochloride salt which had precipitated was removed by filtration. The solvents were removed *in vacuo* and the residue was redissolved in dry 1,4 dioxan (300 ml). The solution was stirred under an atmosphere of dry

nitrogen and N-hydroxysuccinimide (29.35 g, 0.255 mol, 1.13 equivs.) was added in small portions. To this mixture was added N-methylmorpholine (25 ml); more hydrochloride salt was then observed to precipitate. The solution was stirred for 16 hours and then the N-methylmorpholine hydrochloride was removed by filtration and the dioxan removed *in vacuo*, to leave a pale yellow crystalline solid which was recrystallised from acetone/light petroleum (b.p. 40-60°C) to afford the *title compound* (79.15 g, 82%) as a white solid: m.p. 183-186°C (lit.,[31] 173-174°C); (Found: C, 53.4; H, 3.5; N, 9.85. Calc. for $C_{19}H_{15}N_3O_9$: C, 53.3; H, 3.5; N, 9.8%); tlc -E, R_f 0.65.

3.4 Preparation of N^{α}-[2,2-bis (4'-nitrophenyl)ethoxycarbonyl] amino acid derivatives.

Two routes were used for the preparation of the Bnpeoc amino acid derivatives:

Route 1: is an adaptation of the general method of Shute and Rich[99] and was used to prepare compounds which could be readily crystallised. Bnpeoc-Tyr(Bu^t)-OH is given in full as an example.

Route 2: was used for the Bnpeoc derivatives which could not be readily crystallised, in which case they were isolated as their crystalline Cha or Dcha salts. The free acid was then liberated to give a foam or an amorphous powder. Bnpeoc-Lys(Boc)-OH and its Dcha salt are given in full as examples.

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxycarbonyl]-(O-*tert*.butyl)-L-tyrosine

Bnpeoc-Tyr(Bu^t)-OH

The *title compound* was obtained by route 1 as follows: a solution of triethylamine (0.42 ml, 3 mmol, 1.5 equivs) in dioxane (10 ml) was added to a stirred suspension of H-Tyr(Bu^t)-OH. $0.75H_2O$ (502 mg, 2 mmol) in water (10 ml). Bnpeoc-ONSu (860 mg, 2 mmol) was then added and the mixture stirred for 16 hours at room temperature. The resultant solution was then diluted with water (20 ml) and the pH adjusted to 3.0 by the addition of $KHSO_4$ (2.0 M). The solution was extracted with ethyl acetate (3 x 100 ml) and the combined organic phase washed with water (3 x 300 ml) and brine (1 x 200 ml), dried ($MgSO_4$) and the solvent removed *in vacuo* to give a clear oil which was crystallised from diethyl ether/light petroleum (b.p. 40-60°C) to give the *title compound* (823 mg, 75%) as a white solid: m.p. 73-76°C; (Found C, 60.8; H, 5.3; N, 7.6. $C_{28}H_{29}N_3O_9$ requires C, 61.0; H, 5.3; N, 7.6%); tlc -E, R_f 0.65; h.p.l.c. -A, R_t 24.2 min; $[\alpha]_D^{22}$ -20.0° (*c* 0.8 in DMF); $\nu_{max.}$(CH_2Cl_2) 3420 (NH), 2985 (CH_2 and CH_3), 1760 (sh, carboxylic acid), 1720 (s, urethane), 1610 (phenyl), 1525 (NO_2), 1505, 1390, 1370 (Bu^t), 1350 (NO_2), 1160, 1110, 1040, 860, 830; $\lambda_{max.}$(MeOH) 276 (ε 25 000); δ_H (80 MHz; $CDCl_3$) 1.30 (9H, s, 3 x CH_3, Bu^t), 3.02 (2H, m, βCH_2, Tyr), 4.4-4.8 (4H, m, $CHCH_2$, Bnpeoc and αCH, Tyr), 5.07 (1H, d, αNH,Tyr), 6.75-7.05 (4H, m, Ar-H, Tyr), 7.35 (4H, d, Ar-H, Bnpeoc), 8.15 (4H, d, Ar-H, Bnpeoc); δ_C (50 MHz; $CDCl_3$) 28.65 (3 x CH_3, Bu^t), 36.71 (βCH_2, Tyr), 49.65 (αCH, Tyr), 54.52 (CH, Bnpeoc), 65.86 (CH_2, Bnpeoc), 78.65 ($\underline{C}(CH_3)_3$), 123.93, 124.13, 129.09, 129.53 (Ar CH's, Bnpeoc and Tyr), 130.09, 146.70, 147.07, 154.36 (quaternary C's, Bnpeoc and Tyr), 155.22 (urethane), 175.52 (carboxylic

acid); m/z: 551 (MH^+, 0.6%), 536 (1.2), 496 (M^+-Bu^t, 12.1), 450 (4.8), 271 (13), 225 (20), 180 (25), 136 (38), 107 (100), 89 (23), 77 (24), 58 (100); hrms: found 550.18257. $C_{28}H_{28}O_9N_3$ requires 550.18254.

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxycarbonyl)]-N^{γ}-[*bis*(4'-methoxyphenyl)methyl]asparagine
Bnpeoc-Asn(Mbh)-OH

The *title compound* was prepared by route 1 and was crystallised from dichloromethane/diethyl ether as a white solid (87%); m.p. 114-115°C; (Found: C, 60.7; H, 4.8; N, 8.3. $C_{34}H_{34}N_4O_{11}$ requires C, 60.7; H, 5.05; N, 8.3%); tlc -E, R_f 0.68; h.p.l.c. -A, R_f 24.8 min; $[\alpha]_D^{22}$ -2.1° (*c* 1.0 in DMF); $\nu_{max.}(CH_2Cl_2)$ 3420 (NH), 2840 (OCH_3, Mbh), 1750 (sh, carboxylic acid), 1725 (s, urethane), 1685 (amide), 1610 (phenyl), 1525 (NO_2), 1515 (amide), 1350 (NO_2), 1160, 1035, 860 (*p*. disubst. phenyl), 835; $\lambda_{max.}(CH_2Cl_2)$ 277 (ε 20 500); δ_H(200 MHz; $CDCl_3$) 2.83 (2H, q, βCH_2, Asn), 3.72 (3H, s, CH_3O, Mbh), 3.74 (3H, s, CH_3O, Mbh), 4.44 (1H, m, αCH, Asn), 4.5-4.8 (3H, br m, $CHCH_2$, Bnpeoc), 6.00 (1H, d, CH, Mbh), 6.08 (1H, d, ex, αNH, Asn), 6.64 (1H, d, βCH_2, Asn), 6.78-6.82 (4H, m, Ar-H, Mbh), 7.02-7.08 (4H, m, Ar-H, Mbh), 7.34 (4H, d, *J* 8.78 Hz, Ar-H, Bnpeoc), 8.12 (4H, d, *J* 8.71 Hz, Ar-H, Bnpeoc); δ_C (50 MHz; $CDCl_3$) 32.1 (βCH_2, Asn), 49.4 (αCH, Asn), 53.4 (CH, Bnpeoc), 55.1 (CH_3O x 2, Mbh), 56.0 (CH, Mbh), 65.9 (CH_2, Bnpeoc), 113.8, 123.9, 128.3, 129.1 (aromatic CH's, Bnpeoc and Mbh), 133.2, 146.8, 158.6 (quaternary C's, Bnpeoc and Mbh), 155.8 (urethane), 172.1 (amide), 173.8 (CO_2H); m/z: 673 (MH^+, 0.7%), 671 (1.0), 656 (1.4), 268 (7.0), 255 (10), 242 (50), 227 (100), 213 (19); hrms: found 673.21453. $C_{34}H_{35}N_4O_{11}$requires 673.21456.

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxycarbonyl]-N^{ε}-(*tert*.butyloxycarbonyl)lysine dicyclohexylammonium salt
Bnpeoc-Lys(Boc)-OH,Dcha

The *title compound* was prepared by route 2 as follows: triethylamine (0.76 ml, 5.5 mmol, 1.5 equivs.) in dioxane (10 ml) was added to a stirred suspension of H-Lys(Boc)-OH (0.9 g, 3.65 mmol) in water (10 ml). To this mixture was added solid Bnpeoc-ONSu (1.56 g, 3.65 mmol) and the mixture was stirred for 16 hours at room temperature. The solution was then diluted with water (30 ml) and the pH adjusted to 3.0 with $KHSO_4$ (2.0 M). This suspension was then extracted with ethyl acetate (3 x 150 ml) and the combined organic phase washed with water (3 x 250 ml) and brine (1 x 250 ml), dried ($MgSO_4$) and the solvent removed *in vacuo* to give a pale green gum.

This was redissolved in dichloromethane (20 ml) and cooled to 0°C. Dicyclohexylamine (1.19 ml, 1 equiv.) was then added and the solution stirred for 15 min, after which diethyl ether (100 ml) was added to afford the *title compound* (4.12 g, 93%) as a white solid: m.p. 146-149°C(dec.); (Found C, 61.2; H, 7.4; N, 9.4. $C_{38}H_{55}N_5O_{10}$ requires C, 61.2; H, 7.4; N, 9.4%); $[\alpha]_D^{22}$ +2.9° (1.0 in DMF); tlc -E, R_f 0.55; h.p.l.c. -A, R_t 24.8 min; $\nu_{max.}$ (CH_2Cl_2) 3440 (NH), 2970 (CH_2 and CH_3), 2840, 1720 (urethanes), 1635, 1610, 1600 (phenyl), 1525 (NO_2), 1395, 1365 (Boc Bu^t), 1350 (NO_2); $\lambda_{max.}$ (MeOH) 275 (ε 21 300); δ_H (80 MHz; $CDCl_3$) 0.8 - 2.2 (37H, br m, CH_2 x 3, Lys, 10 x CH_2, Dcha, 2 x CH, Dcha, 3 x CH_3, Boc), 3.0 (2H, m, εCH_2, Lys), 3.85 (1H, m, αCH, Lys), 4.55 (3H, m, $CHCH_2$, Bnpeoc), 5.65 (1H, αNH, Lys), 7.40 (4H, d, Ar-H, Bnpeoc), 8.10 (4H, d, Ar-H, Bnpeoc); δ_C (50MHz; $CDCl_3$) 22.3 (CH_2, Lys), 24.53 (2 x CH_2, Dcha), 25.00 (CH_2, Lys), 28.26 (CH_3 x 3, Boc), 29.23 (CH_2, Dcha), 32.99 (CH_2, Lys), 40.42 (CH_2, Lys), 49.79 (αCH, Lys), 52.44 (CH, Dcha), 55.52 (CH, Bnpeoc), 65.19 (CH_2, Bnpeoc), 78.84 ($\underline{C}(CH_3)_3$, Boc), 123.85, 129.40 (Bnpeoc aromatic CH's), 146.92, 147.12 (Bnpeoc quaternary C's), 154.96, 155.78 (urethanes), 178.65 (carboxylate); m/z: 743 (MH^+, 0.3%), 741 (0.76), 560 (0.21), 504 (0.4), 460 (2.1), 306 (3.1), 182 (100), 154 (23); hrms: found 742.40266. $C_{38}H_{56}N_5O_{10}$ requires 742.40269.

N^{α}-[*bis*(4'-Nitrophenyl)ethoxycarbonyl)-N^{ε}-(*tert*. butyloxycarbonyl)lysine
Bnpeoc-Lys(Boc)-OH

Bnpeoc-Lys(Boc)-OH,Dcha (4.30g, 6.75mmol) was suspended in $KHSO_4$ (2.0 M, 500 ml) and extracted with ethyl acetate (3 x 300 ml). The combined organic phase was washed with water (3 x 500 ml) and brine (1 x 500 ml) and then dried ($MgSO_4$). The solvent was removed *in vacuo* and the residual oil redissolved in dichloromethane and evaporated *in vacuo* to give the *title compound* (3.36 g, 89%) as a yellow foam. A small quantitiy of the material was precipitated from chloroform/light petroleum (b.p. 40-60°C) as a white amorphous powder: m.p. 76-79°C; (Found C, 55.8; H, 5.8; N, 9.9. $C_{26}H_{32}N_4O_{11}$ requires: C, 55.6; H, 6.1; N, 10.0%); tlc -E, R_f 0.55; h.p.l.c. -A, R_t 24.8 min; $[\alpha]_D^{22}$ +7.6° (*c* 1.0 in $CHCl_3$); $\nu_{max.}$ (CH_2Cl_2) 3440 (NH), 2960 (CH_2 and CH_3), 1720 (s, urethanes and CO_2H), 1605, 1600 (phenyl), 1520 (NO_2), 1350, 1240, 1170, 1080, 870 (*p*.-disubst. phenyl); $\lambda_{max.}$ ($CHCl_3$) 275 (ε 19 500); δ_H (80 MHz; $CDCl_3$) 1.5 (15H, m, 3 x CH_3, Boc, β, γand δ CH_2, Lys), 3.05 (2H, m, εCH_2, Lys), 4.25 (1H, m, αCH, Lys), 4.63 (3H, m, $CHCH_2$, Bnpeoc), 5.53 (1H, d, αNH, Lys), 7.37 (4H, d, Ar-H, Bnpeoc), 8.15 (5H, m, εNH, Lys and Ar-H, Bnpeoc); δ_C (50MHz; $CDCl_3$) 22.06 (CH_2 x 2, Lys), 28.04 (CH_3 x 3, Boc), 31.24 (CH_2, Lys), 39.71 (CH_2, Lys), 49.48 (αCH, Lys), 53.47 (CH, Bnpeoc), 65.66 (CH_2, Bnpeoc), 79.30 ($\underline{C}(CH_3)_3$, Boc), 123.77, 129.04 (Bnpeoc aromatic CH's), 146.81 (Bnpeoc quaternary C's), 156.19 (urethanes), 175.52

(carboxylic acid); m/z: 561 (MH^+, 0.3%), 506 (MH^+-Bu^t, 2), 254 (10), 178 (8), 128 (17), 92 (17), 85 (25), 58 (100); hrms: found 561.21966. $C_{25}H_{33}N_4O_{10}$ requires 561.21965.

N^{α}-[2,2-*bis* (4'-Nitrophenyl)ethoxycarbonyl]alanine
Bnpeoc-Ala-OH

Bnpeoc-Ala-OH prepared by route 1 and crystallised from chloroform/diethyl ether to give the *title compound* (90%); m.p. 147-148°C (lit.,[31] 148-149.5°C); $[\alpha]_D^{22}$ +0.8° (*c* 1.0 in DMF) (lit.,[31] $[\alpha]_D^{22}$+0.7°); tlc -E, R_f 0.60; δ_H [$(CD_3)_2O$; 200 MHz] 1.36 (3H, d, βCH_3), 4.20 (1H, q, αCH), 4.70 - 4.82 (3H, m, Bnpeoc $CHCH_2$), 7.70, 8.22 (2 x 4H, 2 x d, Bnpeoc Ar-H); m/z: 404 (MH^+, 10%), 257 (40), 62 (100).

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxycarbonyl] N^{G}-(2,2,5,7,8-pentamethylchroman-6-sulphonyl)-arginine cyclohexylamine salt
Bnpeoc-Arg(Pmc)-OH,Cha

The *title compound* was prepared by route 2 and crystallised from dichloromethane /diethyl ether as a pale yellow solid (79%); m.p. 136-137°C; (Found C, 57.8; H, 6.73; N, 11.6. $C_{41}H_{55}N_7O_{11}S$ requires C, 57.7; H, 6.5; N, 11.6%); tlc -E, R_f 0.4; $\nu_{max.}$(CH_2Cl_2) 3440 (NH), 1720 (urethane), 1610 (phenyl), 1525 (NO_2), 1350 (NO_2); $\lambda_{max.}$(MeOH) 435 (1200), 320 (13 600), 282 (53 000), 256 (48 000); δ_H (80 MHz; $CDCl_3$) 1.0-1.9 (22H, br m, Pmc 2 x CH_3, Arg β and γ CH_2's, Cha 5 x CH_2), 2.03 (3H, s, Pmc CH_3), 2.55 (2H, m, Pmc CH_2), 3.05 (3H, m, δCH_2, Arg and CH, Cha), 3.89 (1H, m, Arg αCH, 4.5 - 4.7 (3H, m, $CHCH_2$, Bnpeoc), 6.0 - 7.0 (7H, br m, Arg αNH, guanadino NH's, Cha NH_3^+), 7.32 (4H, d, Ar-H, Bnpeoc), 8.10 (4H, d, ArH, Bnpeoc); m/z: 754 [Bnpeoc-Arg(Pmc)-OH, 7.6%], 366 (2.9), 326 (2.7), 219 (18), 203 (36), 147 (67), 100 ($ChaH^+$, 100).

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxycarbonyl] N^{G}-(2,2,5,7,8-pentamethylchroman-6-sulphonyl)-arginine
Bnpeoc-Arg(Pmc)-OH

The *title compound* was prepared by route 2 and precipitated from chloroform with light petroleum (b.p. 40-60°C) as a pale yellow solid: [75% from H-Arg(Pmc)-OH]; m.p. 120-140°C (lit.,[36] *ca.* 140°C); (Found C, 55.3; H, 5.4; N, 10.8. Calc. for $C_{35}H_{42}N_6O_{11}S$: C, 55.7; H, 5.6; N, 11.1%);

$[\alpha]_D^{22}$+3.6° (*c* 1.0 in $CHCl_3$); m/z 755 (MH^+, 9%), 499 (5), 219 (20), 203 (50), 147 (100); hrms: found 755.27104. $C_{35}H_{43}N_6O_{11}S$ (MH^+) requires 755.27103.

N^{α}-[2,2-*bis* (4'-Nitrophenyl)ethoxycarbonyl]-β-(*tert*.butyl)aspartic acid dicyclohexylammmonium salt

Bnpeoc-Asp(OBu^t)-OH,Dcha

The *title compound* was prepared by route 2. The crude product formed a gel from dichloromethane/diethyl ether which was dried *in vacuo* to give a white solid which was then recrystallised from dichloromethane/diethyl ether (69%); m.p. 160-164°C; (Found C, 61.2; H, 7.0; N, 8.2. $C_{35}H_{48}N_4O_{10}$ requires C, 61.4; H, 7.1; N, 8.2%); tlc -E, R_f 0.66; h.p.l.c. -A, R_t 23.6 min; $[\alpha]_D^{22}$ +17.8° (*c*1.0 in $CHCl_3$); $\nu_{max.}$ 3440 (NH), 2830, 2730, 1725 (urethane), 1640, 1605, 1600 (phenyl and carboxylate), 1525 (NO_2), 1395, 1365 (*tert.* butyl), 1350 (NO_2), 1150, 1080, 870, 830; $\lambda_{max.}$(MeOH)274 (ε 14 600); δ_H (80 MHz; $CDCl_3$) 0.9-1.2 (20H, br m, 10 x CH_2, Dcha), 1.38 (9H, s, 3 x CH_3, *tert.* butyl), 2.73 (2H, m, 2 x CH, Dcha), 2.85 (2H, m, βCH_2, Asp), 4.05 (1H, m, αCH, Asp), 4.85 (3H, m, $CHCH_2$, Bnpeoc), 5.83 (1H, d, αNH, Asp), 7.38 (4H, d, Ar-H, Bnpeoc), 8.15 (4H, d, Ar-H, Bnpeoc); δ_C (50 MHz; $CDCl_3$) 24.39, 29.92 (CH_2's, Dcha), 27.64 (3 x CH_3, *tert.* butyl), 37.63 (βCH_2, Asp), 49.26 (αCH, Asp), 51.24 (2 x CH, Dcha), 53.03 (CH, Bnpeoc), 65.60 (CH_2, Bnpeoc), 80.95 [$\underline{C}(CH_3)_3$], 123.77, 129.50 (aromatic CH's, Bnpeoc), 146.81, 146.90 (quaternary C's, Bnpeoc), 155.16 (urethane), 170.24 ($\underline{C}O_2Bu^t$), 174.43 (CO_2H); m/z: 685 (MH^+, 1.2%), 670 (1.3), 491 (21), 426 (100), 399 (56), 288 (37), 271 (39), 255 (40); hrms: found 685.34482. $C_{35}H_{49}N_4O_{10}$ (MH^+) requires 685.34484.

N^{α}[2,2-*bis* (4'-Nitrophenyl)ethoxycarbonyl]-β-(*tert*.butyl)aspartic acid

Bnpeoc-Asp(OBu^t)-OH

The *title compound* was prepared by route 2 and precipitated as white solid from chloroform/light petroleum (b.p. 40-60°C) (95%); m.p. 75-79°C; (Found C, 57.9; H, 6.50; N, 7.55. $C_{23}H_{25}N_3O_{10}$ requires C, 54.9; H, 5.01; N, 8.35%); tlc -E, R_f 0.66; h.p.l.c. -A, R_t23.6 min; $\nu_{max.}$(CH_2Cl_2) 3440 (NH), 2960 (CH_2, CH_3), 2860 (CH), 1725 (urethane), 1610, 1600 (phenyl), 1525 (s, NO_2), 1390, 1375 (*tert.* butyl), 1350 (NO_2), 1230, 1150, 1070, 870, 830; $\lambda_{max.}$(MeOH) 277 (ε 9 700); δ_H (80 MHz; $CDCl_3$) 1.36 (9H, *tert.* butyl), 2.75 (2H, m, βCH_2, Asp), 4.32 (1H, m, αCH, Asp), 4.63 (3H, $CHCH_2$, Bnpeoc), 5.73 (1H, d, αNH, Asp), 7.37, 8.15 (4H x 2, d x 2, Ar-H, Bnpeoc); δ_C (50 MHz; $CDCl_3$) 27.66 (CH_3 x 3, Bu^t), 37.09 (βCH_2, Asp), 49.47 (αCH, Asp), 52.53 (CH, Bnpeoc), 65.92

(CH_2, Bnpeoc), 82.05 ($\underline{C}(CH_3)_3$), 123.82, 129.04 (Ar CH's, Bnpeoc), 146.72, 146.87 (quaternary C's, Bnpeoc), 155.45 (urethane), 169.96 ($\underline{C}O_2Bu^t$, Asp), 174.86 ($\underline{C}O_2H$, Asp); m/z: 502 (0.17%), 448 (1.67), 426 (1.34), 271 (8), 186 (16), 182 (100), 138 (5), 134 (7); hrms: found 504.16177: $C_{23}H_{26}N_3O_{10}$ (MH^+) requires 504.16180.

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxycarbonyl]glutamine
Bnpeoc-Gln-OH

The *title compound* was prepared by route 1 and crystallised from acetone/light petroleum (b.p. 40-60°C) as a white solid (94%): m.p. 189-190°C (lit.,[31] 185-187°C); (Found C, 52.4; H, 4.5; N, 11.9. Calc. for $C_{20}H_{20}N_3O_9$ C, 52.2; H, 4.4; N, 12.2%); tlc -F, R_f 0.6; δ_C (50 MHz; [$(CD_3)SO$] 26.5 (βCH_2, Gln), 31.3 (γCH_2, Gln), 48.9 (αCH, Gln), 53.46 (CH, Bnpeoc), 65.05 (CH_2, Bnpeoc), 123.6, 129.5 (Bnpeoc Ar CH's), 146.4, 148.1 (Bnpeoc quaternary C's), 155.77 (urethane), 182.4 (carboxylic acid).

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxycarbonyl]-N^{γ}-[*bis*(4'-methoxyphenyl)methyl]glutamine
Bnpeoc-Gln(Mbh)-OH

The *title compound* was prepared by route 1 and was crystallised from THF/light petroleum (b.p. 40-60°C) as a white solid (62%); m.p. 97-107°C; (Found C, 60.5; H, 5.1; N, 7.9. $C_{35}H_{34}N_4O_{11}$ requires C, 61.2; H, 5.0; N, 8.2%); tlc -E, R_f 0.66; h.p.l.c. -A R_t 24.9 min; $[\alpha]_D^{22}$ +5.8° (*c* 1.0 in DMF); $\nu_{max.}$ (CH_2Cl_2) 3420 (NH), 2740 (w, OCH_3, Mbh), 1740 (urethane and CO_2H), 1670 (amide), 1610 (phenyl), 1525 (NO_2 and amide), 1350 (NO_2); $\lambda_{max.}$ (MeOH) 277 (ε 23 000); δ_H (80 MHz; $CDCl_3$) 2.00 (2H, m, βCH_2, Gln), 2.25 (2H, m, γCH_2, Gln), 3.71 (6H, s, 2 x OCH_3, Mbh), 4.20 (1H, m, αCH, Gln), 5.5 (3H, br m, $CHCH_2$, Bnpeoc), 5.99 (2H, m, αNH, Gln and CH, Mbh), 6.74 (4H, d, J_{AB} 8.7 Hz, Mbh Ar-H), 7.05 (4H, d, J_{AB} 8.7 Hz, Mbh Ar-H), 7.30 (4H, d, J_{AB} 7.7 Hz, Bnpeoc Ar-H), 8.10 (4H, d, J_{AB} 7.7 Hz); δ_C (50 MHz; $CDCl_3$) 27.72 (Gln βCH_2), 32.08 (Gln γCH_2), 49.45 (Gln αCH), 53.40 (Bnpeoc CH), 55.04 (2 x OCH_3, Mbh), 55.96 (Mbh CH), 65.93 (Bnpeoc CH_2), 118.00, 127.88, 128.29, 129.05 (aromatic CH's, Bnpeoc and Mbh), 133.25 (Mbh Ar 4-C), 146.87 (Bnpeoc Ar 1-C and 4-C), 155.74 (urethane), 158.64 (Mbh Ar 4-C), 172.64 (amide), 173.8 (carboxylic acid); m/z: 686 (MH^+, 73%), 669 (100), 578 (MH^+-PhOMe), 506 (42); hrms: found 687.23023. $C_{35}H_{34}N_5O_{11}$ requires 687.23021.

N^{α}-[2,2-*bis* (4'-Nitrophenyl)ethoxycarbonyl]-γ-(*tert*.butyl)glutamic acid dicyclohexylammonium salt

Bnpeoc-Glu(OBu^t)-OH,Dcha

The *title compound* was prepared by route 2. The crude product formed a gel from chloroform/ diethyl ether which was dried *in vacuo* to give a white solid which was then recrystallised from dichloromethane/diethyl ether (88%); m.p. 129-132°C (lit.,[31] 124.5-126°C); (Found C, 61.5; H, 7.25; N, 7.9. Calc. for $C_{36}H_{50}N_4O_{10}$: C, 61.9; H, 7.2; N, 8.05); $[\alpha]_D^{22}$+5.3° (*c* 1.0 in DMF) (lit.,[31] $[\alpha]_D^{22}$+5.3°); tlc -E, R_f 0.61, -G, R_f 0.41; $\nu_{max.}$ (CH_2Cl_2) 3420 (NH), 2920 (CH_3 and CH_2), 2870 (CH), 1725 (urethane), 1640, 1620, 1595 (phenyl and carboxylate), 1525 (NO_2), 1490, 1350 (NO_2), 1160, 860, 830; m/z: 699 (MH^+, 1.7%), 670 (1.4), 491 (25), 462 (30), 426 (100), 254 (58); hrms: found 699.36050. Calc. for $C_{36}H_{51}N_4O_{10}$(MH^+) requires 699.36049.

N^{α}-[2,2-bis(4'-Nitrophenyl)ethoxycarbonyl]-γ-(*tert*.butyl)glutamic acid

Bnpeoc-Glu(OBu^t)-OH

The *title compound* was prepared by route 2 and precipitated from chloroform/light petroleum (b.p. 40-60°C) as a white solid (82% from H-Glu(OBu^t)-OH); (Found C, 55.2; H, 5.25; N, 8.1. Calc. for $C_{24}H_{27}N_3O_{10}$: C, 55.7; H, 5.2; N, 8.1); $\lambda_{max.}$ (MeOH) 275 (18 900); δ_H(80 MHz; $CDCl_3$) 1.40 (9H, s, *tert*.butyl), 2.10 (2H, m, Glu βCH_2), 2.25 (2H, m, Glu γCH_2), 4.29 (1H, m, Glu αCH), 4.5-5.0 (3H, m, $CHCH_2$, Bnpeoc), 5.5 (1H, d, αNH, Glu), 7.35 (4H, d, Bnpeoc Ar-H), 8.16 (4H, d, Bnpeoc Ar-H); m/z: 462 (40%), 416 (20), 271 (60), 254 (100), 225 (60), 179 (67); hrms: found 518.17745. $C_{24}H_{28}N_3O_{11}$ requires 518.17745.

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxycarbonyl]-N^{τ}-trityl-histidine

Bnpeoc-His(Trt)-OH

The *title compound* was prepared by route 1 and was then precipitated from chloroform/ diethyl ether as a white powder (85%); m.p. 134-138°C; (Found C, 65.6; H, 4.5; N, 9.6; $C_{40}H_{33}N_5O_8.H_2O$ requires C, 65.8; H, 4.8; N, 9.6%); tlc -E, R_f 0.65; $[\alpha]_D^{22}$+4.4° (*c* 1.0 in DMF); $\nu_{max.}$ (CH_2Cl_2) 3420 (NH), 1760 (sh, carboxylic acid), 1720 (urethane), 1610 (phenyl), 1525 (NO_2), 1495, 1350 (NO_2), 1130, 865, 850, 830; $\lambda_{max.}$ (MeOH) 274 (ε 23 900); δ_H(80 MHz; $CDCl_3$) 3.1 (2H, m, βCH_2, His), 4.1-4.7 (4H, m, $CHCH_2$, Bnpeoc and αCH, His), 6.05 (1H, d, αNH, His), 6.66 (1H, s), 6.9-7.5 (19H, m, Ar-H, Trt and Bnpeoc), 7.88 (1H, s), 8.10 (4H, d, Ar-H, Bnpeoc), 9.75 (1H, br s, carboxylic acid);

δ_C (50 MHz; $CDCl_3$) 28.71 (βCH_2, His), 49.58 (αCH, His), 53.73 (CH, Bnpeoc), 65.66 (CH_2, Bnpeoc), 120.7-137.3 (9 peaks, trityl, His imidazole and Bnpeoc CH's), 146.75, 146.92 (Bnpeoc quaternary C's), 154.99 (urethane), 173.12 (carboxylic acid); m/z: 712 (MH^+, 2.5%), 470 (MH^+-Trt, 2.3), 243 (Trt^+, 100), 165 (52), 136 (47), 77 (30); hrms: found 712.24071. $C_{40}H_{34}N_5O_8$ requires 712.24072.

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxycarbonyl]isoleucine dicyclohexylammonium salt
Bnpeoc-Ile-OH,Dcha

The *title compound* was prepared by route 2 and crystallised from dichloromethane/diethyl ether as a white solid (92%); m.p. 158-159°C (lit.,[31] 160-162°C); $[\alpha]_D^{22}$ +7.7° (*c* 1.0 in DMF) (lit.,[31] $[\alpha]_D^{22}$ +7.7); tlc -E, R_f 0.50; $\lambda_{max.}$ (MeOH) 276 (ε 20 300).

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxycarbonyl]isoleucine
Bnpeoc-Ile-OH

The *title compound* was prepared by route 2 and precipitated from chloroform/light petroleum (b.p. 40-60°C) as a white amorphous powder (85% from H-Leu-OH); m.p. 59-62°C; tlc -E, R_f 0.5.

N^{α}-[2,2-*bis* (4' Nitrophenyl)ethoxycarbonyl]leucine
Bnpeoc-Leu-OH

The *title compound* was prepared by route 1 and crystallised from chloroform/diethyl ether (92%); m.p. 68-70°C (lit.,[31] 67-71°C); $[\alpha]_D^{22}$ +4.9° (*c* 1.0 in DMF) (lit.,[31] $[\alpha]_D^{27}$+4.1°); δ_H (80 MHz; $CDCl_3$) 0.87 (6H, d, 2 x δCH_3, Leu), 1.58 (3H, m, βCH_2 and γCH, Leu), 4.30 (1H, m, αCH, Leu), 7.38 (4H, d, Ar-H, Bnpeoc), 8.13 (4H, d, Ar-H, Bnpeoc); m/z: 446 (MH^+, 2%), 271 (20), 254 (20), 225 (27), 178 (30), 86 (37), 42 (90), 28 (100).

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxcarbonyl]methionine dicyclohexylammonium salt
Bnpeoc-Met-OH,Dcha

The *title compound* was prepared by route 2 and crystallised from methanol/diethyl ether as a pale yellow solid (81%); m.p. 186-189°C (lit.,[31] 179-180°C); (Found C, 59.3; H, 6.8; N, 8.7. Calc. for $C_{32}H_{44}N_4O_8S$: C, 59.6; H, 6.8; N, 8.7%); tlc -F, R_f 0.60; $\nu_{max.}$ (CH_2Cl_2) 3440 (NH), 2840 (CH_2 and CH_3), 2860 (CH), 1720 (urethane), 1640, 1610, 1600 (phenyl and carboxylate), 1525 (NO_2), 1350

(NO_2), 860, 830; $\lambda_{max.}$ (MeOH) 278 (ε 17 700); δ_H(360 MHz, $CDCl_3$) 1.0-1.9 (20H, 10 x CH_2, Dcha), 2.0 (5H, βCH_2, and CH_3, Met), 2.43 (2H, m, γCH_2, Met), 2.73 (2H, m, CH x 2, Dcha), 4.07 (1H, m, αCH, Met), 4.5-4.7 (3H, m, $CHCH_2$, Bnpeoc), 5.67 (1H, d, αNH, Met), 7.38 (4H, d, Ar-H, Bnpeoc), 8.15 (4H, d, Ar-H, Bnpeoc); m/z: 399 (10), 182 (100), 98 (30), 83 (43), 56 (70), 42 (62).

N^{α}-[2,2-*bis* (4'-Nitrophenyl)ethoxycarbonyl]methionine
Bnpeoc-Met-OH

The *title compound* was prepared by route 2 and was obtained as a foam from dichloromethane (78% from methionine); tlc -F, R_f 0.6; δ_H[80 MHz; $(CD_3)_2CO$] 2.0 (obsc. by solvent, εCH_3 and γCH_2, Met), 2.5 (2H, m, βCH_2, Met), 4.30 (1H, m, αCH, Met), 4.77 (3H, m, $CHCH_2$, Bnpeoc), 7.70 (4H, d, Ar-H, Bnpeoc), 8.22 (4H, d, Ar-H, Bnpeoc); m/z: 464 (11%), 257 (30), 217 (100), 199 (40), 181 (90).

N^{α}-[2,2-*bis*(4' Nitrophenyl)ethoxycarbonyl)]phenylalanine
Bnpeoc-Phe

The *title compound* was prepared by route 1 and was recrystallised from diethyl ether to afford a white crystalline solid (93%); m.p. 77-80°C (lit.,[31] 79-80°C); (Found: C, 57.8; H, 4.5; N, 8.3. $C_{24}H_{21}N_3O_8.H_2O$ requires C, 57.9; H, 4.6; N, 8.4%); $[\alpha]_D^{22}$ -16.5° (*c* 1.0 in DMF) (lit.,[31] $[\alpha]_D^{27}$-15.7°); δ_C (50 MHz; $CDCl_3$) 37.24 (βCH_2, Phe), 49.52 (αCH, Phe), 54.41 (CH, Bnpeoc), 65.73 (CH_2, Bnpeoc), 123.56, 127.11, 128.47, 129.01 (Ar CH's, Bnpeoc and Phe), 135,23 (Ar quaternary C, Phe), 146.64, 146.90 (Ar quaternary C's, Bnpeoc), 155.18 (urethane), 175.79 (carboxylic acid).

Comparison of the solubility properties of some Bnpeoc and Fmoc amino acid derivatives.

The solvent (0.5 ml) was vigorously stirred in a small (0.5 ml) reaction vessel (from a Wheaton Micro-Kit). An accurately weighed sample of the solid material (e.g. Bnpeoc amino acid) was added to the solution in small portions until either 0.25 mmol had been added or an insoluble precipitate had formed. The remaining sample was then weighed and the difference used to calculate the solubility. These are shown in table 6.

	Solubility in mmol/ml		
Reagent	DCM	Dioxane/DCM mixtures	Dioxane
Fmoc-Phe-OH		<0.14	
Bnpeoc-Phe-OH	0.57		
Fmoc-Gln(Dbse)-OH	0.01		
Bnpeoc-Gln(Dbse)-OH	0.40		
Bnpeoc-Asn(Mbh)-OH	0.21		
Bnpeoc-Asn(Dbse)-OH	0.18		
Bnpeoc-Gly-OH	0.06	0.14 (1:3)	>0.6
		0.29 (1:1)	
Bnpeoc-Ala-OH	0.05	0.50 (3:1)	
Bnpeoc-Arg(Pmc)-O	>0.40		
Bnpeoc-Lys(Boc)-O	>0.55		
Bnpeoc-Gly-OH	>0.5*		

Table 6. The solubility of some Bnpeoc amino acid derivatives.

Notes

*: in this case the DCM contained 0.5 mmol/ml BOP reagent and 2 equivs. of DIEA

The figures in parentheses indicates the ratio of the two solvents (e.g. dioxane:DCM)

3.5 Solid Phase Peptide Synthesis

General Notes

The preparation of all the peptides described in this section were carried out using an Applied Biosystems 430A automated solid phase synthesiser. All the solvents used were either freshly distilled (DCM) or used as commercially supplied by Applied Biosystems, Rathburn, or Aldrich (DMF).

The C-terminal residue of those peptides with a C-terminal carboxylic acid (as opposed to a C-terminal amide) were coupled to the *p*-alkoxybenzyl alcohol resin outwith the synthesiser. The extent of this coupling was determined by deprotecting a small resin sample of known weight and measuring the absorbance of the olefin piperidine adduct produced using a UV spectrophotometer.

The following semi-permanant side chain protecting groups were used throughout: Bu^t ethers (Ser, Thr, Tyr); Bu^t esters (Glu, Asp), Pmc (Arg), Trityl (His), Boc (Lys). Asn and Gln residues were used with no side chain protection unless otherwise stated.

Synthesis Cycle

The syntheses were achieved using the following common procedures:

1. Capping. Acetic anhydride (0.5 M in DMF, 3 ml) and pyridine (0.5 M in DMF, 3 ml) vortexed with the resin (6 min).
2. Washing. (5 x 1 min, DMF).
3. Removal of the N^{α}-protecting group. With 20% piperidine in DMF. The effluent from the reaction vessel was passed through a UV cell before being sent to waste. This allowed the efficiency of the peptide chain assembly to be monitored (30 min).
4. Washing. With the coupling solvent (either DCM or DMF, 4 x 1.5 min).
5. Activation. Carried out in the activator vessel.
6. Coupling. Carried out in the reaction vessel.

A variety of procedures were used for the activation and coupling of each residue and these are described in detail below. In double couple cycles steps 4, 5 and 6 are repeated at this stage.

7. Washing. With DMF (x 5, DMF, 8 min).

The completed peptidyl resins were stored at -20°C, swollen in DMF with the last N^{α}–protecting group uncleaved. This was removed, when required, by washing the resin in 20% piperidine in DMF [(10 ml for 5 min) x 3] and washed with DMF (3 x 30 ml).

Activation and coupling procedures

With diphenylphosphinic anhydride[33]

During activation the N^{α}-protected amino acid (1.0 mmol), diphenylphosphinic anhydride (1.0 mmol) and NMM (1.0 mmol) in the coupling solvent (9 ml total volume) were mixed with a stream of nitrogen in the 'activator vessel'. During the coupling reaction, in which the activated amino acid derivative was vortexed with the peptidyl resin in the 'reaction vessel', NMM (0.5 mmol) and 2,6 lutidene (2.0 mmol) in the coupling solvent (3 ml total volume) were added.

With the BOP reagent [56,58]

During activation the N^{α}-protected amino acid derivative (1.0 mmol), BOP reagent (1.0 mmol) and diisopropylethylamine (1.5 mmol) in the coupling solvent (9 ml total volume) were mixed in the 'activator vessel' (25 seconds) and then transferred to the 'reaction vessel' and vortexed with the peptidyl resin.

With TBTU [59]

During activation the N^{α}-protected amino acid derivative (1.0 mmol), TBTU (1.0 mmol), HOBt (1.0 mmol) and diisopropylethylamine (1.5 mmol) in the coupling solvent (9 ml total volume) were mixed in the activator vessel (25 seconds) and then transferred to the 'reaction vessel' and vortexed with the peptidyl resin.

With diisopropylcarbodiimide

During activation the N^{α}-protected amino acid derivative (2.0 mmol) and diisopropylcarbodiimide (1.0 mmol) in the coupling solvent (9 ml total volume) were mixed in the 'activator vessel' by a stream of nitrogen gas for 15 min. The mixture was then transfrered to the 'reaction vessel' and vortexed with the peptidyl resin.

With diisopropylcarbodiimide/1-hydroxybenzotriazole

During activation the N^{α}-protected amino acid derivative (1 mmol), diisopropylcarbodimide (1 mmol) and HOBt (1 mmol) in the coupling solvent (9 ml total volume) were mixed in the 'activator vessel' by a stream of nitrogen for 30 min. The mixture was then transfered to the 'reaction vessel' and vortexed with the peptidyl resin.

Loading the C-terminal amino acid residue onto *p*-alkoxybenzyl alcohol resin.

N^{α}-(9-Fluorenylmethoxycarbonyl)glycyl-*p*-alkoxybenzyl alcohol resin) (63)
Fmoc-Gly-(O-$CH_2C_6H_4$-OR)

Diisopropylcarbodiimide (0.48 ml, 3.05 mmol) was added to a stirred solution of Fmoc-Gly-OH (1.78 g, 6.0 mmol) in DMF (10 ml) at room temperature. After 10 min this solution was added to *p*-alkoxybenzyl alcohol resin (Bachem, 1.05 mmol/g) (1.43 g, 1.5 mmol), then DMAP (37 mg, 0.3 mmol, 0.2 equivs.) was added and the mixture was agitated in an ultrasonic bath at room temperature for 3 hours. The resin was then filtered, washed with DCM (3 x 30 ml), DMF (3 x 30 ml) and diethyl ether (3 x 30 ml), and dried *in vacuo*. The coupling yield was found to be 34% (0.32 mmol/g) by UV deprotection study[126]. The resin was then capped (see below).

N^{α}-(9-Fluorenylmethoxycarbonyl)alaninyl-2-methoxy-4-alkoxybenzyl alcohol resin (64)
Fmoc-Ala-[O-CH_2-(2-OCH_3)Ph(1,4)-OR]

Fmoc-Ala-OH.H_2O (1.32 g, 4 mmol) was dissolved in DMF (10 ml). To this was added diisopropylcarbodiimide (0.31 ml, 2 mmol) and DMAP (12.2 mg, 0.1 mmol). The reaction mixture was stirred at room temperature for 15 min and then added to 2-methoxy-4-alkoxybenzyl alcohol resin[127,128] (Bachem, 0.945 mmol/g, 1.0 g). The mixture was then agitated in an ultrasonic bath for two hours, after which the resin was filtered, washed with DMF (3 x 10 ml) and diethyl ether (3 x 10 ml) and dried *in vacuo* . The coupling yield was found to be 75% (0.62 mmol/g) in a UV deprotection study[126] and the resin was then capped (see below).

N^{α}-[2,2-*bis*(4'-Nitrophenyl)ethoxycarbonyl]-N^{γ}-[*bis*(4'-methoxyphenyl)methyl]glutamyl *p* -alkoxybenzyl alcohol resin (65)
Bnpeoc-Gln(Mbh)-(O-$CH_2C_6H_4$-OR)

Bnpeoc-Gln(Mbh)-OH (1.38 g, 2 mmol) was dissolved in DMF and diphenylphosphinic anhydride (840 mg, 2 mmol) was added and the solution stirred for 10 min. To this mixture was added N-methylmorpholine (0.218 ml, 2 mmol), DMAP (12 mg, 0.1 mmol) and *p*-alkoxybenzyl alcohol resin (Novabiochem, 0.79 mmol/g) (650 mg, 0.51 mmol). The mixture was agitated in an ultrasonic bath at room temperature for two hours. The resin was then removed by filtration, washed in DMF (3

x 30 ml), DCM (3 x 30 ml) and diethyl ether (3 x 30 ml) and then dried *in vacuo*. The coupling yield was found to be 92% (0.49 mmol/g) in a UV deprotection study[126] and the resin was then capped (see below).

Capping the unreacted benzyl alcohol groups of the loaded resins.

This procedure[21] was carried out on each batch of resin functionalised with a benzyl alcohol type linker after the first amino acid had been coupled to it. As an example: the functionalised resin, Fmoc-Gly-(O-$CH_2C_6H_4$OR) (1.5 g; 0.32 mmol/g) was suspended in dichloromethane (30 ml) and cooled to 0°C with gentle stirring. Pyridine (0.5 ml) and benzoyl chloride (0.5 ml) were added. The resin was then agitated in an ultrasonic bath for 12 min at room temperature. The resin was then filtered, washed with dichloromethane (3 x 30 ml) and diethyl ether (3 x 30 ml). The resin was then dried *in vacuo* and stored at -20°C until required.

Ubiquitin H-(residues 36-53)-OH (53).

H-Ile-Pro-Pro-Asp-Gln-Gln-Arg-Leu-Ile-Phe-Ala-Gly-Lys-Gln-Leu-Glu-Asp-Gly-OH

The synthesis was achieved using the functionalised resin (63), Fmoc-Gly-(O-$CH_2C_6H_4$-OR), on a 0.5 mmol scale employing N^{α}-Fmoc protected amino acid derivatives and diphenylphosphinic anhydride activation. Couplings were carried out in DMF, using a double couple cycle (2 x 2 equivs. of mixed anhydride) except for glycine which was incorporated using a single couple cycle (1 x 4 equivs. mixed anhydride).

The peptide was cleaved from the resin support with simultaneous removal of the side chain protecting groups by suspending an aliquot of the peptidyl resin in a mixture of TFA, water and thioanisole (10:1:1; 12 ml). The suspended resin was agitated in an ultrasonic bath for two hours, after which the resin was removed by filtration, washed with TFA (3 x 5 ml) and the combined filtrates were concentrated *in vacuo* to give a yellow oil. This oil was solidified upon adding diethyl ether (30 ml), to give the crude peptide (107.3 mg, 82%) as a white powder. An aliquot of this material (80.5 mg) was purified by preparative h.p.l.c. (gradient B) and the h.p.l.c. solvents were removed by lyophilisation to afford the *title compound* as a white solid (43.8 mg, 45%): h.p.l.c. -C, R_t 15.6 min; δ_H(600 MHz; 90% H_2O/10% D_2O) tabulated below; m/z: 2026 (MH^+, 17%), 1645 (5.2), 1233 (26), 1007 (29), 788 (42), 732 (100); hrms: found 2026.08026, $C_{90}H_{146}N_{25}O_{28}$ requires 2026.08019; amino acid analysis: Asp_2 1.95, Glu_4 3.96, Pro_2 1.60, Gly_2 2.09, Ala_1 1.02, Ile_2 1.78, Leu_2 2.08, Phe_1 0.98, Lys_1 1.09, Arg_1 0.97.

Spin system	αNH	αCH	βCH	Others
Ala	8.29	4.31	1.39	
AMX 1	8.53	4.60	2.85	
AMX 2	8.46	4.33	2.82, 2.93	
AMX 3	8.30	4.66	3.03, 3.16	
AM(PT)X 1	8.45	2.02, 2.11	2.57	
AM(PT)X 2	8.36	2.02, 2.14	2.46	
AM(PT)X 3	8.26	2.06, 2.12	2.38	
AM(PT)X 4	8.25	1.98, 2.19	2.38	
Gly 1	8.12	3.92		
Gly 2	7.95	3.95		
Ile 1		4.17	1.79	
Ile 2	7.97	4.10	1.79	γCH_2 1.40, 1.14 $\delta,\gamma CH_3$ 0.81
Leu 1	8.41	4.31	1.67	γCH 1.52, 2 x δCH_3 0.93
Leu 2	8.32	4.35	1.64	γCH 1.64, 2 x δCH_3 0.93
Lys	8.14	4.29	1.76, 1.85	γCH_2 1.45, δCH_2 1.43, εCH_2 3.01, NH_3^+ 7.56
Pro 1		4.43	2.11, 1.94	γCH_2 2.05, δCH_2 3.84, 3.62
Pro 2		4.46	2.08, 1.97	γCH_2 2.08, δCH_2 3.87, 3.74
Arg	8.21	4.28	1.82	γCH_2 1.64, δCH_2 3.25 δNH, 7.24

Table 7. ^1H spin systems assigned to the Ubiquitin fragment H-(residues 36-53)-OH. AM(PT)X systems are Glu and Gln, AMX systems are Phe and Asp[72] .

Lysinylthreoninylleucinylthreoninylglycyllysinylthreonine amide (55)

H-Lys-Thr-Leu-Thr-Gly-Lys-Thr-NH_2

The synthesis was carried out using the functionalised resin Fmoc-NH-Dbs(5,2)-OCH_2Ph-R (5) (0.74 mmol/g, 0.66 g) on a 0.5 mmol scale employing N^{α}-Bnpeoc amino acid derivatives and the BOP reagent[57,58] (Penninsula) for activation. Couplings were carried out in dichloromethane using single couplings (1 x 2 equivs. activated amino acid derivative).

The peptide was cleaved from the resin support with simultaneous removal of the side chain protecting groups by suspending an aliquot of the peptidyl resin (20% by weight) in a mixture of TFA, water and thioanisole (10:1:1, 12 ml). The suspended resin was agitated in an ultrasonic bath for two hours, after which the resin was removed by filtration, washed with TFA (3 x 5 ml) and the combined filtrates concentrated *in vacuo* to give a colourless oil. The crude peptide was precipitated as a white powder (72 mg, 96.5%) by the addition of diethyl ether (30 ml). An aliquot of this material (40 mg) was purified by preparative h.p.l.c. (gradient H) and the h.p.l.c. solvents were removed by lyophilisation to give the *title compound* (22 mg, 53%) as a white hygroscopic solid: h.p.l.c. -L, R_t15.8 min; $[\alpha]_D^{22}$-35.3° (*c* 0.91 in H_2O); δ_H (360 MHz; 90%H_2O/10%D_2O) 0.92, 0.98 (2 x 3H, 2 x d, 2 x Leu δCH_3), 1.25 (9H, m, 3 x Thr γCH_3), 1.48 (3H, m, Leu γCH and βCH_2), 1.68-1.85 (10H, m, 2 x Lys βCH, 2 x Lys γCH_2, 2 x Lys δCH_2), 1.95 (2H, m, 2 x Lys βCH), 3.03 (4H, m, 2 x Lys εCH_2), 4.0-4.6 (10H, m, Gly CH_2, 5 x αCH, 3 x Thr βCH), 7.17, 7.61 (2 x 1H, $CONH_2$), 8.18, 8.20, 8.31 (3 x 1H, 3 x d, 3 x αNH), 8.43 (1H, t, Gly αNH), 8.66, 8.74 (2 x 1H, 2 x d, 2 x αNH); m/z 747 (MH^+, 100%,), 728 (5), 702 (7), 434 (5). 374 (22), 246 (25); hrms: found 747.47284. $C_{32}H_{63}N_{10}O_{10}$ (MH^+) requires 747.47283; amino acid analysis: Thr_3 2.87, Gly_1 1.00, Leu_1 0.95, Lys_2 2.10.

Acetyllysinylthreoninylleucinylthreoninylglycinyllysinylthreonine amide (56)

Ac-Lys-Thr-Leu-Thr-Gly-Lys-Thr-NH_2

Peptidyl resin (20% by weight) from the synthesis of (55) was suspended in 20% piperidine in DMF [(10 ml for 5 min) x 3] and washed in DMF (5 x 10 ml). The N-terminal amino group was then reacted with acetic anhydride (0.5 M in DMF, 10 ml) and pyridine (0.5 M in DMF, 10 ml) for 12 min and then washed with DMF (5 x 10 ml).

The peptide was then cleaved from the resin support and the side chain protecting groups removed simultaneously by suspending the peptidyl resin in a mixture of TFA, water and thioanisole (10:1:1, 12 ml). The suspended resin was then agitated in an ultrasonic bath for two

hours. The resin was then filtered, washed in TFA (3 x 5 ml) and the combined filtrates concentrated *in vacuo* to leave a colourless oil. The addition of diethyl ether (30 ml) precipitated the crude peptide as a white powder (70 mg, 89%). An aliquot of this material (30 mg) was purified by preparative h.p.l.c. (gradient H, R_t 10.0 min) and the solvents removed by lyophilisation to afford the *title compound* (17.0 mg, 50%) as a white hygroscopic solid: h.p.l.c. -L, R_t10.1 min; $[\alpha]_D^{22}$-50.7°(*c* 0.52 in H_2O); δ_H (360 MHz; 90H_2O%/10% D_2O) 0.92, 0.98 (2 x 3H, 2 x d, 2 x Leu δCH_3), 1.10 (9H, m, 3 x Thr γCH_3), 1.4-1.6 (3H, Leu βCH_2 and γCH), 1.65-1.95 (12H, m, 2 x Lys β, γ and δCH_2), 2.09 (3H, s, acetyl CH_3), 2.99-3.12 (4H, m, 2 x Lys εCH_2), 4.0-4.5 (11H, m, 6 x αCH, 3 x Thr βCH, 1 x Gly CH_2), 7.07, 7.63 (1H x 2, s x 2, $CONH_2$), 7.57 (6H, br, 2 x Lys εNH_3^+), 8.13, 8.21, 8.24, 8.28, 8.34 (5 x αNH), 8.42 (2H, m, Gly NH, 1 x αNH); m/z: 812 (MNa^+, 60%) 790 (MH^+, 100), 385 (10), 357 (10), 272 (95), 254 (54), 215 (58); hrms: found 789.48342. $C_{34}H_{65}N_{10}O_{11}$(MH^+) requires 789.48339.

Leucylglutamylaspartylglycylargininylthreoninylleucine amide (57)

H-Leu-Glu-Asp-Gly-Arg-Thr-Leu-NH_2

The synthesis was carried out using the functionalised resin Fmoc-NH-Dbs(5,2)-OCH_2Ph-R (5) (0.74 mmol/g; 0.66 g) on a 0.5 mmol scale employing N^{α}-Bnpeoc protected amino acid derivatives and diphenylphosphinic anhydride activation. Couplings were carried out in DMF using a double couple cycle (2 x 2 equivs. of mixed anhydride) except for glycine which was incorporated using a single couple cycle (1 x 4 equivs. of mixed anhydride).

The peptide was cleaved from the resin support and the side chain protecting groups removed simultaneously by suspending the peptidyl resin (160 mg) in a mixture of TFA, water and thioanisole (10:1:1; 12 ml). The suspended peptidyl resin was agitated in an ultrasonic bath for two hours, after which time the resin beads had taken on a deep red colour. The resin was then filtered, washed with TFA (3 x 10 ml) and the solvent removed from the combined filtrates *in vacuo*. Addition of diethyl ether (30 ml) precipitated the crude peptide as a white solid (32mg, 47%; amino acid analysis: Asp_1 0.99, Thr_1 1.00, Glu_1 1.13, Gly_1 1.03, Leu_2 1.84, Arg_1 1.00).

This material was purified by preparative h.p.l.c. (gradient D) (R_t 11.0 min) and the h.p.l.c. solvents removed by lyophilisation to afford the *title compound* (9.8 mg, 14.%) as a white solid: h.p.l.c. -E, R_t 11.0 min; $[\alpha]_D^{22}$-28.4° (*c* 0.37 in H_2O); δ_H(600 MHz; D_2O) 0.93 (3H, d, Leu δCH_3), 0.98 (9H, m, 3 x Leu δCH_3), 1.25 (3H, d, Thr γCH_3), 1.6-1.95 (10H, m, 2 x Leu βCH_2, 2 x Leu γCH, Arg βCH_2, and Arg γCH_2), 2.06, 2.18 (2 x 1H, 2 x m, Glu βCH_2), 2.52 (2H, m, Glu γCH_2), 2.95 (2H, m, Asp βCH_2), 3.25 (2H, t, Arg δCH_2), 4.0 (2H, d, Gly CH_2), 4.09 (1H, m, αCH), 4.38 (2H, m, 2 x αCH), 4.24 (1H, m, Thr βCH), 4.47, 4.49, 4.74 (3 x 1H, 3 x m, 3 x αCH); m/z: 802 (MH^+,

100%), 757 (5), 399 (25), 329 (25), 165 (50); hrms: found 802.44225. $C_{33}H_{59}N_{11}O_{12}$ (MH^+) requires 802.44226; amino acid analysis: Asp_1 1.02, Thr_1 0.97, Glu_1 1.03, Leu_2 2.03, Arg_1 0.93.

N-Acetylleucylglutamylaspartylglycylargininylthreoninylleucine amide (58).
Ac-Leu-Glu-Asp-Gly-Arg-Thr-Leu-NH_2

Peptidyl resin from the synthesis of (57) was suspended in 20% piperidine in DMF [(10 ml for 5 min) x 3] and washed in DMF (5 x 10 ml). The N-terminal amino group was then reacted with acetic anhydride (0.5 M in DMF; 10ml) and pyridine (0.5 M in DMF; 10 ml) for 12 min and the resin was then washed with DMF (3 x 30 ml).

The peptide was then cleaved from the resin support and the side chain protecting groups removed simultaneously by suspending the peptidyl resin in a mixture of TFA, water and thioanisole (10:1:1; 12 ml). The resin was stirred at room temperature for two hours after which the resin was filtered, washed with TFA (3 x 10 ml) and the solvent was then removed from the combined filtrates *in vacuo*. Addition of diethyl ether (30 ml) precipitated the crude peptide as a white solid (44 mg, 66%): amino acid analysis: Asp_1 0.98, Thr_1 0.99, Glu_1 1.08, Gly_1 1.01, Leu_2 1.93, Arg_1 1.01. This material was purified by preparative h.p.l.c. (gradient F) and two fractions were collected (R_t 15.0 and 20.0 min). The h.p.l.c. solvents were removed by lyophilisation to yield two white powders: the first material eluted was an impurity (5.7 mg, 8.5%) : δ_H(600 MHz; D_2O) 0.93 (3H, d, Leu δCH_3), 0.97 (3H, d, Leu δCH_3), 1.25 (3H, d, Thr γCH_3), 1.60 (1H, m, Leu γCH), 1.70 (4H, m, Leu βCH_2 and Arg γCH_2), 1.78, 1.87 (2 x 1H, 2 x m, Arg βCH_2), 2.03, 2.18 (2 x 1H, 2 x m, Glu βCH_2), 2.08 (3H, s, acetyl CH_3), 2.51 (2H, m, Glu γCH_2), 2.93 (2H, m, Asp βCH_2), 3.25 (2H, t, Arg δCH_2), 4.00 (2H, m, Gly CH_2), 4.30 (2H, m, 2 x αCH), 4.37 (1H, d, Thr βCH), 4.43, 4.46, 4.76 (3 x 1H, 3 x m, 3 x αCH); m/z, 731 (100%), 687 (5), 424 (15), 329 (20); amino acid analysis: Asp_1 0.98, Thr_1 0.99, Glu_1 1.08, Gly_1 1.01, Leu_2 1.04, Arg_1 1.01; h.p.l.c. R_t -G 15.2 min) and the second material eluted was the *title compound* (14.9 mg, 22.3%): h.p.l.c -G, R_t 20.2 min; δ_H (600 MHz; H_2O) tabulated below; m/z 844 (MH^+, 100%), 826 (10), 799 (11), 399 (7), 242 (8); hrms: found 844.45285. $C_{35}H_{62}N_{11}O_{13}$ requires 844.45282; amino acid analysis Asp_1 0.92, Thr_1 0.93, Glu_1 1.10, Gly_1 1.00, Leu_2 2.05, Arg_1 0.95.

Amino acid	αNH	αCH	βCH	Others
Arg	8.04	4.44	1.81, 1.90	δCH_2 3.23 δNH 7.21
Asp	8.49	4.82	2.88, 2.97	
Glu	8.51	4.42	2.01, 2.15	γCH_2 2.48
Gly	8.42	3.96		
Leu A	8.26	4.31	1.56	δCH_3 0.89, 0.93
Leu B	8.33	4.35	1.64	δCH_3 0.89, 0.93
Thr	8.25	4.35	4.21	γCH_3 1.22
Ac				CH_3 2.05
$CONH_2$	7.09, 7.59			

Table 8. ^1H resonances assigned from COSY experiments (20 mM, 90% H_2O/10% D_2O)

Isoleucinylphenylalaninylalaninylglycyllysinylglutamine (Ubiquitin residues 44-49) (54)
H-Ile-Phe-Ala-Gly-Lys-Gln-OH

The synthesis was acheived using the functionalised resin Bnpeoc-Gln(Mbh)-(O-$CH_2C_6H_4$-OR) (65), on a 0.5 mmol scale employing N^{α}-Bnpeoc protected amino acid derivatives and diphenylphosphinic anhydride activation[33]. Couplings were carried out in DMF using a double couple cycle (2 x 2 equivs. of mixed anhydride) except for glycine which was single coupled (1 x 4 equivs. of mixed anhydride).

The peptide was removed from the resin support with simultaneous removal of the side chain protecting groups by suspending the peptidyl resin (20% by weight) in a mixture of TFA, water and thioanisole (10:1:1, 12 ml). The suspended resin was agitated in an ultrasonic bath for two hours, after which the resin was removed by filtration and washed with TFA (3 x 10 ml). The combined filtrates were then concentrated *in vacuo* to give a pale yellow oil. Addition of diethyl ether (30 ml) precipitated the crude peptide as a white powder (67 mg, overweight). An aliquot of this material (40 mg) was purified by preparative h.p.l.c. (gradient J) and the h.p.l.c. solvents removed by lyophilisation to afford the *title compound* (22.1 mg, 59%) as a white solid: h.p.l.c. - M, R_t 8.5 min; δ_H (600 MHz, 90% H_2O/10% D_2O) tabulated below; m/z: 664 (MH^+), 9%), 332 (3),

261 (5), 120 (20), 90 (30), 62 (55), 48 (82), 32 (100); hrms: found 663.38295. $C_{31}H_{51}N_8O_{14}$ requires 663.38296; amino acid analysis: Glu_1 1.05, Gly_1 1.03, Ala_1 0.97, Ile_1 0.95, Phe_1 0.99, Lys_1 1.07.

Amino Acid	αNH	αCH	βCH	Others
Ile		3.86	1.99, 1.48	δCH_3 0.93, γCH_3 0.99 γCH_2 1.21
Phe	8.72	4.71	3.11	2,4,6-H 7.33, 3,5-H, 7.39
Ala	8.43	4.31	1.36	
Gly	7.59	3.89		
Lys	8.19	4.38	1.79, 1.89	γCH_2 1.48, δCH_2 1.71 εCH_2 3.02
Gln	8.39	4.31	2.19, 1.99	γCH_2 2.36

Table 9. ^1H resonances of the hexapeptide (54) assigned from COSY and ROESY experiments (600 MHz; 90%H_2O/10%D_2O).

Acetyl Cro (residues 25-40) amide (59).

Ac-Val-Tyr-Gln-Ser-Ala-Ile-Asn-Lys-Ala-Ile-His-Ala-Gly-Arg-Lys-Ile-NH_2

The *title compound* was prepared manually using a 'bubbler' reaction vessel, described by Atherton and Sheppard[106], using the functionalised resin Fmoc-NH-Dbs(5,2)-OCH_2Ph-R (0.71 mmol/g; 700 mg) (5) on a 0.5 mmol scale. Activation of the N^{α}-Bnpeoc amino acid derivatives was carried out with TBTU[59,107], with single couplings (1 x 2 equivs. of the activated amino acid derivative) except where a requirement for recoupling was indicated by the Kaiser test[108]. The following cycles required recoupling: 1 (Ile), 2 (Lys), 10 (Asn) and 15 (Tyr). All these second couplings were carried out with the BOP reagent[57,58]. The side chain amides of Asn and Gln residues were protected as their N-[*bis*-(4'-methoxyphenyl)methyl] derivatives [i.e. Bnpeoc-Asn(Mbh)-OH and Bnpeoc-Gln(Mbh)-OH].

A study of the deprotection of this peptide was undertaken, in which an aliquot of the peptidyl resin (30 mg) was suspended in a mixture of TFA, water, thioanisole and anisole (10:1:1:1, 13

ml) and agitated in an ultrasonic bath for four hours. Samples of the reaction mixture were removed at hourly intervals, filtered, the solvent was removed *in vacuo* and diethyl ether (10 ml) was added to precipitate the crude peptide (approx. 3 mg). These were studied by analytical h.p.l.c. and mass spectrometry to determine the optimum time for deprotection which was three and a half hours. A larger batch of the peptidyl resin (250 mg) was then deprotected and the peptide cleaved from the resin support under the optimised conditions, then purified as detailed previously, to yield the crude peptide (105 mg, 75%). An aliquot (80 mg) of this material was purified by cation exchange chromatography on a CM-Sephadex column (2.5 x 25 cm) eluting with ammonium acetate buffer (aq., 50 mM, pH 8.0) with gradient elution up to 700 mM. The fractions containing the major product were pooled and the solvents removed by lyophilisation to leave the partially purified peptide (29 mg, 27%). This material was further purified by preparative h.p.l.c. and the h.p.l.c. solvents were removed by lyophilisation to leave the *title compound* (9.2 mg, 8.6%) as a white solid: δ_H (360 MHz; D_2O) 0.87-1.00 (24H, m, Val δCH_3 x 2, Ile δCH_3 x 3, Ile γCH_3 x 3), 1.23 (3H, m, 3 x Ile γCH), 1.42-1.54 (15H, m, Ala CH_3 x 2, Ile βCH x 3, Ile γCH x 3), 1.70-2.10 (19H, m, 2 x Lys β,γ and δ CH_2, Val βCH, Gln βCH_2, Arg β and γ CH_2), 2.08 (3H, s, Acetyl CH_3), 2.8-3.4 (12H, m, 2 x Lys εCH_2, Arg δCH_2, Asn, His and Tyr βCH_2), 3.70-4.60 (17H, m, 15 x αCH, Ser βCH_2), 4.60 (1H, m, αCH), 6.88 (2H, d, Tyr Ar-H), 7.18 (2H, d, Tyr Ar-H), 7.38 (1H, d), 8.70 (1H, d); m/z 1810.9 (100%), 1646.8 (64), 905.4 (91), 822.9 (48), 764.1 (27); hrms: found 1810.04507. $C_{81}H_{137}N_{26}O_{21}$ requires 1810.04509; amino acid analysis: Asp_1 1.07, Ser_1 0.91, Glu_1 1.01, Gly_1 1.12, Ala_3 3.32, Val_1 0.91, Ile_3 2.56, Tyr_1 0.91, His_1 0.96, Lys_2 2.14, Arg_1 1.06.

N–Acetyl Cro (residues 15-37) amide (60)

Ac-Gly-Gln-Thr-Lys-Thr-Ala-Lys-Asp-Leu-Gly-Val-Tyr-Gln-Ser-Ala-
Ile-Asn-Lys-Ala-Ile-His-Ala-Gly-NH_2

Three comparative syntheses of this peptide were carried out and the differences are shown in table 10 below. Care was taken to ensure that the results were strictly comparable (for example, the same batch of functionalised resin was used for all three syntheses).The syntheses were carried out on a 0.2 mmol scale using the functionalised resin Fmoc-NH-Dbs(5,2)-OCH_2Ph-R (0.71 mmol/g; 280 mg) (5) with the N^{α}-protected amino acid derivatives and coupling procedures indicated. Couplings were carried out in DMF using a single couple cycle (1 x 5 equivs. activated amino acid). The removal of the N^{α}-protecting groups were monitored using a UV absorbance detector (ABI 757 absorbance detector) at 314 (Fmoc) and 405 nm (Bnpeoc) respectively. A portion of each of the peptidyl resins were suspended in a mixture of TFA, water , anisole and 1,2-ethanedithiol (10:1:1:1; 13 ml) and agitated in an ultrasonic bath for two hours. The resin was then

removed by filtration, washed with TFA (3 x 5 ml) and then the combined filtrates were concentrated *in vacuo*. The crude peptide was then precipitated by the addition of diethyl ether (30 ml). An analytical h.p.l.c. trace of each sample was recorded and a small aliquot of the most successful synthesis (synthesis i) was purified by preparative h.p.l.c. (gradient J) and the h.p.l.c. solvents were removed by lyophilisation to leave the *title compound* (22 mg, 38%) as a white solid: δ_H (600 MHz; CD_3OH) tabulated below; m/z 2412.8 (100%), 2227.7 (19), 1371.5 (10), 1206.9 (25), 979.7 (30); hrms: found 2413.30316. $C_{105}H_{175}N_{32}O_{33}$ (MH^+) requires 2413.30319: amino acid analysis: Asp_2 2.09, Thr_2 2.00, Ser_1 0.91, Glu_2 2.04, Gly_3 3.05, Ala_4 4.36, Val_1 0.98, Ile_2 1.82, Leu_1 1.01, Tyr_1 0.82, His_1 0.83, Lys_3 3.07 (mean deviation/residue=0.056).

Synthesis	Protection	Activation	UV monitor	Yield (%)
i	Fmoc	HOBt/DIC	314 nm	61
ii	Fmoc	TBTU	314 nm	31
iii	Bnpeoc	HOBt/DIC	405 nm	37

Table 10. Yield is estimated from the peak areas observed with the UV monitored cleavage of the N^{α}-protecting group by 20% piperidine. The estimated error in these values is +/-5%

Amino acid	αNH	αCH	βCH	Others
Ac				CH_3 2.04
Gly15	8.55	3.85, 3.92		
Gln16	8.83	4.35	2.05, 2.15	γCH_2 2.41, δNH_2 7.06, 7.36
Thr17	7.99	4.04	4.30	γCH_3 1.29
Lys18	8.18	4.05	1.87, 1.95	γCH_2 1.50, δCH_2 1.64, 1.77, εCH_2 2.96
Thr19	7.96	4.04	4.14	γCH_2 1.27
Ala20	8.05	4.02	1.52	
Lys21	7.99	4.05	1.97	γCH_2 1.67, δCH_2 1.48 εCH_2 2.88
Asp22	8.62	4.66	2.61, 2.95	
Leu23	8.36	4.21	1.68	γCH 1.88, δCH_3 0.93, 0.96
Gly24	8.28	3.94		
Val25	8.31	3.68	2.36	δCH_3 0.96, 1.14
Tyr26	8.21	4.23	3.18	3,5-H 7.01, 2,6-H, 6.66
Gln27	8.79	3.88	2.12, 2.37	γCH_2 2.59, δNH_2 7.41, 6.72
Ser28	8.20	4.23	4.03, 4.08	
Ala29	8.27	4.06	1.54	
Ile30	8.12	3.68	1.92	γCH_2 1.44, 1.10 γCH_3 0.69, δCH_3 0.89
Asn31	8.32	4.43	2.96, 2.81	γNH_2 7.61, 6.92
Lys32	8.28	4.03	1.93,1.99	γCH_2 1.47, δCH_2 1.69 εCH_2 2.88
Ala33	8.04	4.28	1.49	
Ile34	8.22	3.85	1.92	γCH_2 1.60, δCH_3 0.83 γCH_3 0.77
His35	7.84	4.48	3.37,3.16	4-H 7.31, 2H 8.44
Ala36	8.02	4.29	1.50	
Gly37	8.08	3.82, 3.90		Amide 6.81, 7.54

Table 11. Assignment of the proton n.m.r. resonances for (62) in CD_3OH, taken from COSY, TOCSY and NOESY experiments.

Attempted synthesis of the Cro protein.

The synthesis was carried out on a 0.25 mmol scale using the functionalised resin N^{α}-(9-fluorenylmethoxycarbonyl)-alaninyl-2-methoxy-4-alkoxybenzyl alcohol resin (64)[127,128] (403 mg; 0.62 mmol/g). Residues 29 to 66 were triple coupled using TBTU (3 x 2 equivs. activated amino acid) and residues 1 to 28 were double coupled with a symmetrical anhydride coupling (1 x 2 equivs) followed by a DIC/HOBt coupling (1 x 2 equivs.).

The peptide was cleaved from the resin with simultaneous removal of the side chain protecting groups by suspending the peptidyl resin (20% by weight) in a mixture of TFA, 1,2-ethanedithiol, water and anisole (10:1:1:1; 13 ml). The suspended resin was agitated in an ultrasonic bath for 2 hours and the resin was then removed by filtration, washed with TFA and the combined filtrates were concentrated *in vacuo* to leave a pale yellow oil. Addition of diethyl ether (30 ml) precipitated the crude product as a white powder (259 mg, 88%) (table 12, column 2). A number of purification protocols were investigated and the success of the method assessed by amino acid analysis, SDS gel electrophoresis and analytical h.p.l.c. The crude peptide (120 mg) was applied to a Sephadex G50 (superfine grade) column (10 x 800 mm), previously equilibrated with ammonium acetate buffer (10 mM, pH 7.2). The required fractions were pooled and lyophilised in two major fractions (total 71.1 mg, 59%). The second (lowest molecular weight) fraction (15.5 mg, 12.9%) contained a major impurity: m/z: 2280 (100%), 1941 (56), 1835 (50); amino acid analysis: table 12, column 1.

The other fraction (amino acid analysis: table 12, column 3) was further purified by cation exchange chromatography on a CM-Sephadex column (2.5 x 25 cm) eluting with buffer A (TrisHCl, 10 mM, NaCl, 100 mM, pH 7.9) for one column volume then gradient elution with buffer B (TrisHCl, 10mM, NaCl 1.0M, pH 7.9). The fractions were analysed by analytical h.p.l.c. and those containing the major component were pooled, lyophilised and desalted (Sephadex G15, 2 x 100 cm column) (3.1 mg, 2.5%): (amino acid analysis: table 12, column 4). The peptide was further purified by preparative h.p.l.c. (gradient J) then the solvents were removed by lyophilisation to give a white solid (1.1 mg): ^{1}H n.m.r (NaH_2PO_4, 10 mM, pH 6.8) indicated the material was different from the material isolated from natural sources[88,89,91]; m/z: (pdms) found 7365.7 (broad peak); required (Cro MH^+) 7352; amino acid analysis: table 12, column 5.

Expected		1	2	3	4	5
Thr	6	2.09	4.75	5.06	5.83	5.33
Ser	3	1.62	2.91	2.96	3.32	2.51
Glu	6	2.33	5.80	5.66	7.19	6.00
Pro	2	1.71	4.60	2.54	2.25	2.05
Gly	4	1.25	3.08	3.99	4.18	4.85
Ala	8	3.68	8.16	9.06	8.33	8.37
Val	3	1.97	5.09	3.56	3.88	3.82
Met	2	0.00	2.06	0.48	0.20	1.82
Ile	5	0.82	3.80	5.17	4.41	5.00
Leu	3	0.33	1.71	2.37	3.15	2.94
Tyr	3	1.01	2.51	3.17	2.78	2.83
Phe	3	1.52	3.00	3.15	2.87	2.98
His	1	0.10	0.84	1.20	1.03	1.03
Lys	8	3.46	7.68	7.56	7.93	7.54
Arg	3	0.00	2.06	1.96	3.00	2.80
Mean deviation/residue			0.18	0.13	0.11	0.07

Table 12. Amino acid analyses taken throughout the purification. See the main text for details. Mean deviation per residue is the average difference between observed and expected values for all the residues. The analyser has an error of 3% (0.03)

References.

1 T. Curtius, *J. Prakt. Chem.*, 1882, **26**, 145.

2 E. Fisher and E. Fourneau, *Ber. Dtsch. Chem. Ges.*, 1901, **34**, 2868.

3 E. Fisher, *Ber. Dtsch. Chem. Ges.*, 1907, **40**, 1754.

4 M. Bergmann and L. Zervas, *Ber. Dtsch. Chem. Ges.*, 1932, **65**, 1192.

5 V. du Vigneaud, C. Ressler, J. M. Swan, C. W. Roberts and P. G. Katsoyannis, *J. Am. Chem. Soc.*, 1954, **76**, 3115.

6 R. A. Boissonnas, St. Guttmann and P. -A. Jaquenoud, *Helv. Chim. Acta*, 1960, **43**, 1349.

7 R. B. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149.

8 R. B. Merrifield, *J. Am. Chem. Soc.*, 1964, **86**, 304.

9 S. B. H. Kent, *Annu. Rev. Biochem.*, 1988, **57**, 957.

10 E. Atherton and R.C. Sheppard in 'The Peptides, Analysis, Synthesis, Biology', ed. S. Udenfriend and J. Meinhofer, Academic Press, New York, 1987, p. 1.

11 E. Atherton, C. J. Logan and R. C. Sheppard, *J. Chem. Soc., Perkin Trans. 1*, 1981, 538.

12 L. A. Carpino and G. Y. Han, *J. Am. Chem. Soc.*. 1970, **92**, 5748.

13 W. A. Erickson and R. B. Merrifield, in 'The Proteins', ed. H. Neurath and R. L. Hill, Academic Press, New York, 1976, p. 3.

14 C. Birr and W. Lochinger, *Synthesis*, 1971, 319.

15 E. Atherton, D. L. J. Clive and R. C. Sheppard, *J. Am. Chem. Soc.*, 1975, **97**, 6584.

16 R. Arshady, E. Atherton, D. L. J. Clive and R. C. Sheppard, *J. Chem. Soc., Perkin Trans. 1*, 1981, 529.

17 R. Arshady, E. Atherton, M. J. Gait, K. Lee and R. C. Sheppard, *J. Chem. Soc., Chem. Commun.*, 1979, 423.

18 A. R. Mitchell, B. W. Erickson, M. N. Ryabtser, R. S. Hodges and R. B. Merrifield, *J. Am. Chem. Soc.*, 1976, **98**, 7357.

19 A. R. Mitchell, S. B. H. Kent, M. Engelherd and R. B. Merrifield, *J. Org. Chem.*, 1978, **43**, 2845.

20 P. Sieber and B. Iselin, *Helv. Chim. Acta*, 1968, **51**, 614.

21 S. -S. Wang, *J. Am. Chem. Soc.*, 1973, **95**, 1328.

22 D. G. Mullen and G. Barany, *J. Org. Chem.*, 1988, **53**, 5240.

23 R. Ramage, C. A. Barron, S. Bielecki and D. W. Thomas, *Tetrahedron. Lett.*, 1987, **28**, 4105.

24 P. Sieber, *Tetrahedron. Lett.*, 1987, **28**, 2107.

25 G. Breipohl, J. Knolle and W. Stuber, *Tetrahedron Lett.*, 1987, **28**, 5651.

26 B. Penke and J. Rivier, *J. Org. Chem.*, 1987, **52**, 1197.

27 C. McInnes, personal communication.

28 M. Bodanszky, 'The Principles of Peptide Synthesis', Springer-Verlag, Heidelberg, 1984, p. 91.

29 E. Fisher, *Ber. Dtsch. Chem. Ges.*, 1914, **47**, 196.

30 L. A. Carpino, B. J. Cohen, Y. -Z. Lin, K. E. Stephens and S. A. Triolo, *J. Org. Chem.*, 1990, **55**, 251.

31 M. R. Florence, Ph. D. Thesis, University of Edinburgh, 1987.

32 R. Ramage, S. Bielecki and M. R. Florence, Proceedings of the 19th European Peptide Symposium, Porto Carras, Greece, 1986, p. 135.

33 T. Gray, Ph. D. Thesis, University of Edinburgh, 1989.

34 L. A. Carpino, J. -H. Tsao, H. Ringsdorf, E. Fell and G. Hettrich, *J. Chem. Soc., Chem. Commun.*, 1978, 358.

35 E. Wunch, L. Moroder and O. Keller, *Hoppe-Seyler's Z. Physiol. Chem.*, 1981, **362**, 1289.

36 J. Green, Ph. D. Thesis, University of Edinburgh, 1987.

37 J. Green, O. M. Ogunjobi, R. Ramage and A. S. J. Stewart, *Tetrahedron. Lett.*, 1988, **29**, 4341.

38 B. Liberek. *Chem. Ind. (London)*, 1961, 987.

39 R. Wilchek, S. Ariely and A. Patchornik, *J. Org. Chem.* , 1968, **33**, 1258.

40 M. Bodanszky, 'The Principles of Peptide Synthesis', Springer-Verlag, Heidelberg, 1984, p. 190.

41 W. König and R. Geiger, *Chem. Ber.*, 1970, **103**, 2041.

42 P. Sieber and B. Riniker, in 'Innovation and Perspectives in Solid Phase Synthesis', ed. R. Epton, SPCC (UK) Ltd., Birmingham, 1990, 577.

43 T. Curtius, *Ber. Dtsch. Chem. Ges.*, 1902, **35**, 3226.

44 L. A. Carpino, B. J. Cohen, K. E. Stephens, S. Y. Sadat-Aalaee, J. -H. Tien and D. C. Langridge, *J. Org. Chem.*, 1986, **51**, 3732.

45 J. C. Sheehan and G. P. Hess, *J. Am. Chem. Soc.*, 1955, **77**, 1067.

46 I. Clark-Lewis, R. Aerbersold, H. Ziltner, J. W. Schrader, L. E. Hood and S. B. H Kent, *Science*, 1986, **231**, 134.

47 K. Nokihara and T. Semba, *J. Am. Chem. Soc.*, 1988, **110**, 7847.

48 E. Brown, R. C. Sheppard and B. J. Williams, *J. Chem. Soc., Perkin Trans.*, 1983, 75.

49 C. -R. Wu, V. C. Stevens, G. W. Tregear and J. D. Wade, *J. Chem. Soc., Perkin Trans. 1*, 1989, 81.

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

51 D. Hudson, *J. Org. Chem.*, 1988, **53**, 617.

52 M. Calquist, *Acta Chem. Scand. Ser. B*, 1987, **41**, 494.

53 R. Schwyzer, B. Iselin and M. Feurer, *Helv. Chim. Acta*, 1955, **38**, 69.

54 M. Bodanszky and V. du Vigneaud, *J. Am. Chem. Soc.*, 1959, **81**, 5688.

55 M. Bodanszky, 'The Principles of Peptide Synthesis', Springer-Verlag, Heidelberg, 1984, p. 33.

56 M. Bodanszky, 'The Principles of Peptide Synthesis', Springer-Verlag, Heidelberg, 1984, p. 35.

57 B. Castro, J. R. Dormoy, G. Evin and C. Selve, *Tetrahedron. Lett.*, 1975, **14**, 1219.

58 D. Le-Nguyen, A. Heitz and B. Castro, *J. Chem. Soc., Perkin Trans. 1*, 1987, 1915.

59 R. Knorr, A. Trzeciak, W. Bannwarth and D. Gillessen, *Tetrahedron. Lett.*, 1989, **30**, 1927.

60 T. Wieland, W. Kern and R. Sehring, *Liebigs Ann. Chem.*, 1950, **569**, 117.

61 T. Wieland and R. Sehring, *Liebigs Ann. Chem.*, 1950, **569**, 122.

62 J. R. Vaughan, *J. Am. Chem. Soc.*, 1951, **73**, 3547.

63 J. R. Vaughan and R. L. Osato, *J. Am. Chem. Soc.*, 1952, **74**, 676.

64 S. C. Harrison and A. K. Aggarwal, *Annu. Rev. Biochem.*, 1990, **59**, 933.

65 P. E. Wright, *Trends Biochem. Sci.*, 1989, **14**, 255.

66 G. Parraga, S. J. Horvath, A. Eisen, W. E. Taylor, L. Hood, E. T. Young and R. E. Klevit, *Science*, 1988, **241**, 1489.

67 T. C. Oas, L. P. McIntosh, E. K. O'Shea, F. W. Dahlquist and P. S. Kim, *Biochemistry*, 1990, **29**, 2891

68 N. C. Seeman, J. M. Rosenberg and A. Rich, *Proc. Natl. Acad. Sci. USA.*, 1976, **73**, 804

69 M. Ptashne, 'A Genetic Switch', Blackwell Scientific Publications, Oxford, 1980.

70 R. P. Wharton and M. Ptashne, *Trends Biochem. Sci.*, 1986, **11**, 71.

71 Y. Takeda, D. H. Ohlendorf, W. F. Anderson and B. W. Matthews, *Science*, 1983, **221**, 1020

72 K. Wüthrich, 'NMR of Proteins and Nucleic Acids', John Wiley and Sons, New York, 1986.

73 K. Wüthrich, *Science*, 1989, **234**, 45.

74 R. Kaptein, E. R. P. Zuiderweg, R. M. Scheek, R. Boelens and W. F. van Gunstern, *J. Mol. Biol.*, ,1985, **182**, 179.

75 R. Boelens, R. M. Scheek, J. H. van Boom and R. Kaptein, *J. Mol. Biol.*, 1987, **193**, 213.

76 W. McGinnis, M. S. Levine, E. Hafen, A. Kuroiwa and W. J. Gehring, *Nature (London)*, 1984, **308**, 428.

77 M. P. Scott and A. J. Weiner, *Proc. Natl. Acad. Sci. USA*, 1984, **81**, 4115

78 G. Otting, Y. Q. Qian, M. Miller, M. Affoller, W. Gehring and K. Wüthrich, *EMBO. J.*, 1988, **7**, 4305.

79 Y. Q. Qian, M. Billeter, G. Otting, M. Mutter, W. J. Gehring and K. Wüthrich, *Cell*, 1989, **59**, 573.

80 T. M. Roberts, H. Shimatake, C. Brady and M. Rosenberg,*Nature (London)*, 1977, **270**, 274.

81 Y. Takeda, A. Folkmanis and H. Echols, *J. Biol. Chem.*, 1977, **252**, 6177.

82 M. W. Hsiang, R. D. Cole, Y. Takeda and H. Echols, *Nature (London)*, 1977, **270**, 275.

83 A. A. Pakula and R. T. Sauer, *Proteins: Structure, Function and Genetics*, 1989, **5**, 202.

84 W. F. Anderson, D. H. Ohlendorf, Y. Takeda and B. W. Matthews,*Nature (London)*, 1981, **290**, 274.

85 D. H. Ohlendorf, W. F. Anderson, R. G. Fisher, Y. Takeda and B. W. Matthews,*Nature (London)*, 1982, **298**, 719

86 R. G. Brennan and B. W. Matthews,*Trends Biochem. Sci*, 1989, **14**, 286.

87 Y. Takeda, A. Sarai and V. M. Rivera, *Proc. Natl. Acad. Sci.,USA*, 1989, **86**, 439.

88 M. P. Kirpichnikov, A. V. Kurochkin and K. G. Skryabin, *FEBS Lett.*, 1982, **150**, 407.

89 A. V. Kurochkin and M. P. Kirpichnikov, *FEBS Lett.*, 1982, **150**, 411.

90 K. T. Ardnt, F. Boshelli, J. Cook, Y. Takeda, E. Tecza and P. Lu, *J. Biol. Chem.*, 1983, **258**, 4177.

91 P. L. Weber, D. E. Wemmer and B. R. Reid, *Biochemistry*, 1985, **24**, 4553.

92 P. Leighton and P. Lu, *Biochemistry*, 1987, **26**, 7262.

93 W. J. Metzler and P. Lu, *J. Mol. Biol*, 1989, **205**, 149.

94 B. Gutte, S. Schindler, F. Standar and E. Wittschieber, 1980, *Biochem. Biophys. Res. Commun.*, 1980, **95**, 1071

95 R. Mayer, G. Lancelot and C. Hélène, *FEBS Lett.*, 1983, **153**, 339.

96 E. Atherton, L. R. Cameron, L. E. Cammish, A. Dryland, P. Goddard, G. B. Priestly, J. D. Richards, R. C. Sheppard and B. J. Williams, in 'Innovation and Perspectives in Solid Phase Synthesis', ed. R. Epton, SPCC (UK) Ltd, Birmingham, 1990, 11.

97 W. König, D. Büchner, R. Knüttel, K. Lindner and A. Volk, Proceedings of the Akabori Conference, Grainau-Eibsee, Bavaria, 1985, p. 32.

98 M. S. Kharash and H. G. Clapp, *J. Org. Chem.*, 1939, **3**, 355

99 R. E. Schute and D. H. Rich, *Synthesis*, 1987, 346.

100 Prepared on a 10 gram scale by G. Raphy.

101 Prepared by M. Pitkeathly.

102 M. Flegel and R. C. Sheppard, *J. Chem. Soc., Chem. Commun.*, 1990, 536.

103 S. A. Salisbury, E. J. Tremeer, J. W. Davies and D. E. I. A. Owen, *J. Chem. Soc., Chem. Commun.*, 1990, 538.

104 K. Shaw, personal communication.

105 R. Ramage, J. Green and O. M. Ogunjobi, *Tetrahedron Lett.*, 1989, **30**, 2149.

106 E. Atherton and R. C. Sheppard, in 'Solid Phase Peptide Synthesis, A Practicle Approach', IRL Press, Oxford, 1989

107 TBTU was prepared by T. Muir.

108 E. Geiger, R. L. Colescoft, C. D. Bossinger and D. L. Cook, *Anal. Biochem.*, 1970, **34**, 595.

109 E. Atherton, presented at 'Peptide Synthesis; solid phase, enzymatic and recombiant approaches', London, May 1989.

110 J. M. Stewart and W. A. Klis, in 'Innovation and Perspectives in Solid Phase Synthesis', ed. R. Epton, SPCC (UK) Ltd, Birmingham, 1990, p. 1.

111 Automated Edman degradation was carried out by the Welmet Sequencing facility, University of Edinburgh.

112 Plasma desorbtion mass spectrometry was carried out by Applied Biosystems Ltd.

113 W. C. Chan, L. -L. Lian, B. W. Bycroft and G. C. K. Roberts, *J. Chem. Soc., Perkin Trans. 1*, 1990, 2359

114 R. A. Boissonas, St. Guttman, P. -A. Jaquenoud and J. P. Waller, *Helv. Chim. Acta,* 1955, **38**, 1491.

115 G. H. L. Nefkens and R. J. F. Nirvard, *Recl. Trav. Chim. Pays-Bas,* 1965, **84**, 1315.

116 G. H. L. Nefkens and R. J. F. Nirvard, *Recl. Trav. Chim. Pays-Bas,* 1964, **83**, 199.

117 R. Schwyzer and H. Dietrich, *Helv. Chim. Acta,* 1961, **44**, 2003.

118 S. Goldschmidt and C. Jutz, *Chem. Ber.*, 1953, **86**, 1116.

119 J. S. Morley, *J. Chem. Soc. C.*, 1967, 2410.

120 R. Roeske, *J. Org. Chem.*, 1963, **28**, 1251.

121 Prepared by O. Ogunjobi.

122 E. Wunch and A. Zwick, *Chem. Ber.*, 1964, **97**, 2497.

123 J. W. Scott, D. Parker and D. R. Parrish, *Synth. Commun.*, 1981, **11**, 303.

124 R. Schwyzer and H. Rittel, *Helv. Chim. Acta*, 1961, **44**, 159.

125 The work described in this section was carried out with G. Raphy.

126 R. Holden, Ph. D. Thesis, University of Edinburgh, 1989.

127 M. Mergler, R. Tanner, J. Gosteli and P. Grogg, *Tetrahedron Lett.*,1988, **29**, 4005.

128 M. Mergler, R. Nyfeler, R. Tanner, J. Gosteli and P. Grogg, *Tetrahedron Lett.*, 1988, **29**, 4009

Courses Attended.

Organic Research Seminars (Various speakers).

NMR Spectroscopy (Dr. I. H. Sadler, University of Edinburgh).

Biochemistry Dept. Seminars (Various speakers).

Solid Phase Peptide Synthesis Symposium, Oxford (Various speakers).

Molecular Biology Dept. Seminars (Various speakers).

Medicinal Chemistry (Professor R. Baker, MS&D).

Peptide Synthesis: Problems and Progress, Exeter (Various speakers).