PROTEIN ENGINEERING BY CHEMICAL METHODS

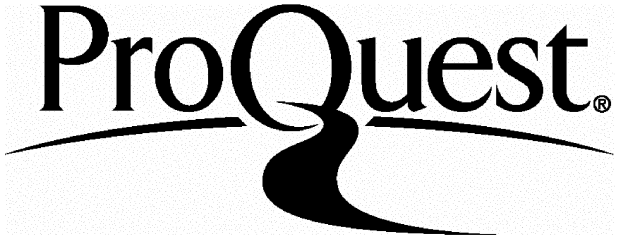
A thesis submitted in partial fulfilment of the requirements of the University of London for the degree of Doctor of Philosophy in the faculty of Medicine

by

Haydn Lewis Ball,
Batchelor of Pharmacy (Hons),
School of Pharmacy (London).

Department of Pharmaceutical Chemistry,
School of Pharmacy,
University of London,
29-39 Brunswick Square,
London, WC1

April 1993



To Glyn, Joy & Annalisa
for their love and support

ABSTRACT

Procedures and methodologies for the routine chemical synthesis and purification of proteins is reported. Optimised automated protocols of classical stepwise Merrifield solid phase peptide synthesis (SPPS) have been used to synthesise 99 residue M. tuberculosis and 97 residue E. coli Heat Shock 10kD proteins. Since peptide bond formation in stepwise SPPS is not unequivocable and leads to the generation of a family of chromatographically similar deletion and truncated peptides, a purification system that is independent of sequence length and amino acid composition has been developed. The proposed system is based on the combination of (i) an effective capping protocol after each coupling step and (ii) the addition of a removable protecting group to the N-terminus of the last amino acid, bearing either lipophilic, acidic or basic functions. Purification of the crude cleavage product has been performed on reversed-phase or ion-exchange media, depending on the probe used. After purification the probe molecule is quantitatively removed from the peptide chain through a base catalysed β-elimination reaction to yield the pure homogeneous product. The potential of these chromatographic probes is demonstrated by application to model peptides (linear and cyclic) from 17 to 104 residues in length.

An alternative to the chemical synthesis of proteins by stepwise SPPS, is the fragment condensation of fully protected peptide segments. The synthesis of a model peptide, GRF(1-44)+Gly, using solid phase fragment condensation is described. Attention is focused on the use of relatively large protected peptides (15-20 residues) with reference to (i) the most effective linker and therefore synthetic chemistry (Boc vs. Fmoc) for their production, (ii) their purification on a large scale and (iii) their coupling to a resin-bound peptide fragment. The most effective protocol is to synthesise the 15 and 17 residue fragments on Sasrin resin (Fmoc chemistry) and to purify them in large quantities using perfusion chromatography. Solid phase fragment condensation is performed in DMSO/NMP, using DCC/HOBt activation, with overall yield of 50%.

ACKNOWLEDGEMENTS

My sincere gratitude goes to Dr. Paolo Mascagni who not only introduced me to this wonderful world of peptide chemistry, but also to my wife. The high standards set by his dedication and scientific ability provided the levels of achievement to which I have tried to aspire.

I wish to thank Professor William A. Gibbons for his support and patience, to the Science and Engineering Research Council for their financial assistance and to Italfarmaco, for taking over where the School of Pharmacy left off. For their analytical intervention my thanks go to Colin James, Kevin Welham and Wilf Baldeo. My acknowledgement extends to Dr. Fernando Albericio and Professor Ernst Giralt for providing technical expertise, handles for the synthesis of protected peptides and a month in Barcelona.

CONTENTS

LIST OF FIGURES

LIST OF TABLES AND SCHEMES

ABBREVIATIONS AND NOMENCLATURE

SPPS	solid phase peptide synthesis
4-COOH-Fmoc	9-(hydroxymethyl)fluorene-4-carboxylic acid
4-COR-Fmoc	9-(hydroxymethyl)fluorene-4-carboxylate derivative
4-COR**X**-Fmoc-Gly	9-(hydroxymethyl)fluorene-4-carboxylate glycyl derivative substituted at 4 position with group R=**X**
Boc	t-butyloxycarbonyl
Fmoc	fluorenylmethoxycarbonyl
N-terminal	N^{α}-amino terminal group
MBHA	*p*-methylbenzhydrylamine
PAM	phenylacetamidomethyl
Bzl	benzyl
For	formyl
DNP	dinitrophenyl
CBZ	carbobenzoxy
HOBt	hydroxybenzotriazole
EDC	1-ethyl-3-(3-dimethylaminopropyl)carbodiimide
DCC	dicyclohexylcarbodiimide
DSC	disuccinimidyl carbonate
DIEA	diisopropylethylamine
TEA	triethylamine
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DCM	dichloromethane
MeOH	methanol
AcCN	acetonitrile
NSu	N-succinimidyl carbonate
TFA	trifluoroacetic acid
HF	hydrogen fluoride
TFMSA	trifluoromethylsulphonic acid
MPLC	medium pressure liquid chromatography
NMR	nuclear magnetic resonance
$CDCl_3$	deuterated chloroform
DMSO-d_6	deuterated dimethylsulfoxide
FAB-MS	fast atom bombardment mass spectrometry
RP-HPLC	reversed-phase high pressure liquid chromatography
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
IEF	iso-electric focusing
HCl	hydrochloride

CHAPTER ONE

PEPTIDE CHEMISTRY - AN INTRODUCTION

The idea that proteins could be best considered as chains of amino acids linked to eachother through amide bonds was first suggested by Hofmeister [1902] and Fischer [1906]. It was also Fischer [1902] who predicted that the chemical synthesis of synthetic enzymes would become a reality. The discovery of biologically active peptides and the role which they and peptide hormones play in the regulation of life processes provided the necessary impetus for the development of synthetic procedures. The pioneering work on oxytocin leading to its total synthesis by du Vigneaud *et al* [1953] and the elucidation of the structure of insulin by Sanger [1953] indicated the need for peptide chemistry as a tool of medicine to produce molecules that were difficult to obtain from natural sources.

Simple amino acids consist of two functional groups (ie. amino and carboxyl functionalities linked by a substituted alkyl carbon) both of which are involved in the formation of an amide bond. When the intended reaction is to link one amino acid to another amino acid but avoid any self-coupling the concept of selective "deactivation" or protection of specific functionalities is introduced to enable controlled bond formation. Further complication results when the side chain of the amino acid carries a functional group which also requires protection (eg. the side chain COOH of glutamic acid).

The first protecting groups examined were not suitable for peptide chemistry since they could not be removed without destruction of the amide bond that linked the amino acids together. Research was then devoted to the discovery of easily removable protecting moieties to overcome this problem. The breakthrough was made by Bergmann and Zervas [1932] with their synthesis of the benzyloxycarbonyl (Z) group. The discovery of this group was important for the further development of peptide chemistry since it could be quantitatively removed, without the need for special equipment, leaving the peptide bond and side chain functions intact. The Z group could be removed by catalytic hydrogenation at room temperature and pressure as well as by reduction with

sodium in liquid ammonia [Sifferd and du Vigneaud, 1935] and acidolysis [Ben-Ishai and Berger, 1952]. Importantly, the Z group was able to protect the chiral integrity of the amino acid, to which it was attached, during amide bond formation. The ability to cleave the protecting group with acid prompted other researchers to develop alternative groups removable under milder conditions using either acids or bases. The more significant acid labile groups to result were the tert-butyloxycarbonyl (Boc) [Carpino, 1957; Anderson and McGregor, 1957; McKay and Albertson, 1957], *o*-nitrophenylsulphenyl (Nps) [Goerdeler, and Holst, 1959; Zervas *et al*, 1963] and biphenylisopropyloxycarbonyl (Bpoc) [Sieber and Iselin, 1968] groups. Of the groups sensitive to mild organic bases, 9-fluorenylmethoxycarbonyl (Fmoc) [Carpino and Han,1970 and 1972] deserves special mention since it heralded an alternative approach to the commonly used acidolytic method of peptide synthesis, the Fmoc chemistry protocol.

Activation

To form the amide bond it is necessary to activate the carboxyl group of an amino acid:

R-COOH $\rightarrow$ R-CO-X

where X represents an electron withdrawing group.

While researchers searched for better protecting groups which were more amenable to synthetic procedures, parallel investigations were conducted in the area of carboxyl group activation. The azide method developed by Curtius [1902] is still used today, particularly in the solution phase coupling of protected peptide fragments [Meienhofer, 1979]. However the more commonly used approach for routine peptide bond formation is via the symmetrical or mixed anhydride. The advantage of the former is that a second acylation product is not formed which could generate unwanted side products. Alternatively, the activated esters (eg. HOBt) of the amino acids can be used.

The acylating reagent dicyclohexylcarbodiimide (DCC) which was introduced by Sheehan and Hess [1955] was used for the synthesis of active esters and mixed or symmetrical anhydrides for peptide bond formation. Carbodiimide-mediated activation has been widely adopted for SPPS, since when used to activate Boc protected amino acids, little racemisation is observed [Stewart and Young, 1984]. However DCC suffers some drawbacks, namely inducement of racemisation during segment condensation, especially when long reaction times are required [Barton *et al*, 1973] and the C-terminal amino acid is not a glycine or proline residue. This unwanted side-reaction occurs more often in polar solvents with a high dielectric constant (eg. DMF) and can be reduced upon the addition of mildly acidic HOBt [Konig and Geiger, 1970abc]. Another important side reaction involves the dehydration of the unprotected side chains of aspargine and glutamine [Gish *et al*, 1956].

In line with the constant desire of scientists to refine and attempt to improve existing methodologies several new activating agents which offer superior reaction kinetics over DCC were developed (eg. HBTU [Dourtoglou *et al.*, 1984], TBTU [Knorr *et al*, 1989] and BOP [Castro *et al*, 1975]). However in the case of BOP the risk of racemisation [Knorr *et al*, 1989] during fragment condensation is still a problem in addition to the generation of a highly carcinogenic byproduct. The application of BOP and BTU-reagents for the synthesis of protected peptides on highly acid-sensitive resins such as the trialkoxybenzhydryl based [Rink, 1987] and 2-chlorotrityl based [Barlos *et al*, 1989] resins have been reported to cleave the peptide prematurely when these reagents are used for coupling. Even the addition of an excess of the tertiary base diisopropylethylamine (DIEA) fails to overcome this problem [Florsheimer and Riniker, 1990].

Solid phase vs. Solution phase

The desire of the protein and peptide chemist has been to understand the structural features of proteins and how they exert their biological responses. Towards this goal the chemistry of peptide synthesis became highly developed

following an understanding of the mechanisms involved in the formation of the amide bond and knowledge of the side reactions that can occur during deprotection, activation and coupling. However the preparation of longer sequences, in solution, was found to be difficult and extremely time consuming. The need for a more eloquent approach was called for and appeared in the guise of solid phase peptide synthesis (SPPS). The technique of SPPS, from the C-terminus was first suggested by Merrifield [1963] and overcame many of the problems that were experienced with solution phase methodologies (Table 1).

	Synthetic protocol	
	Solution phase	Solid phase
Strategy	Segment/convergent	Stepwise
Side-Chain Protection	Maximal	Maximal
Solubility Problems	Frequent/serious	Occasional/minimal
Chemistry	Highly specialised	Simple/general
Automation	No	Yes
Yields	Low	High
Purity	Excellent	Acceptable
Manpower:		
Skill level/Amount	High + experience/50x	Less/1x

Table 1. Comparison of solution and solid phase peptide synthesis
[Taken from Kent and Clark-Lewis, 1985]

Apart from the advantages of simplicity, suitability for automation and a significant reduction in the time required for a synthesis that SPPS offers, there are some problems associated with the technique. Principally, the incomplete coupling during the stepwise addition of each residue results in the accumulation of sequences that lack one or more residues. Consequently, when the peptide is cleaved from the resin support the chromatographic profile is complicated due to the presence of a family of closely related peptidic impurities and can ultimately lead to poor recover of the desired sequence.

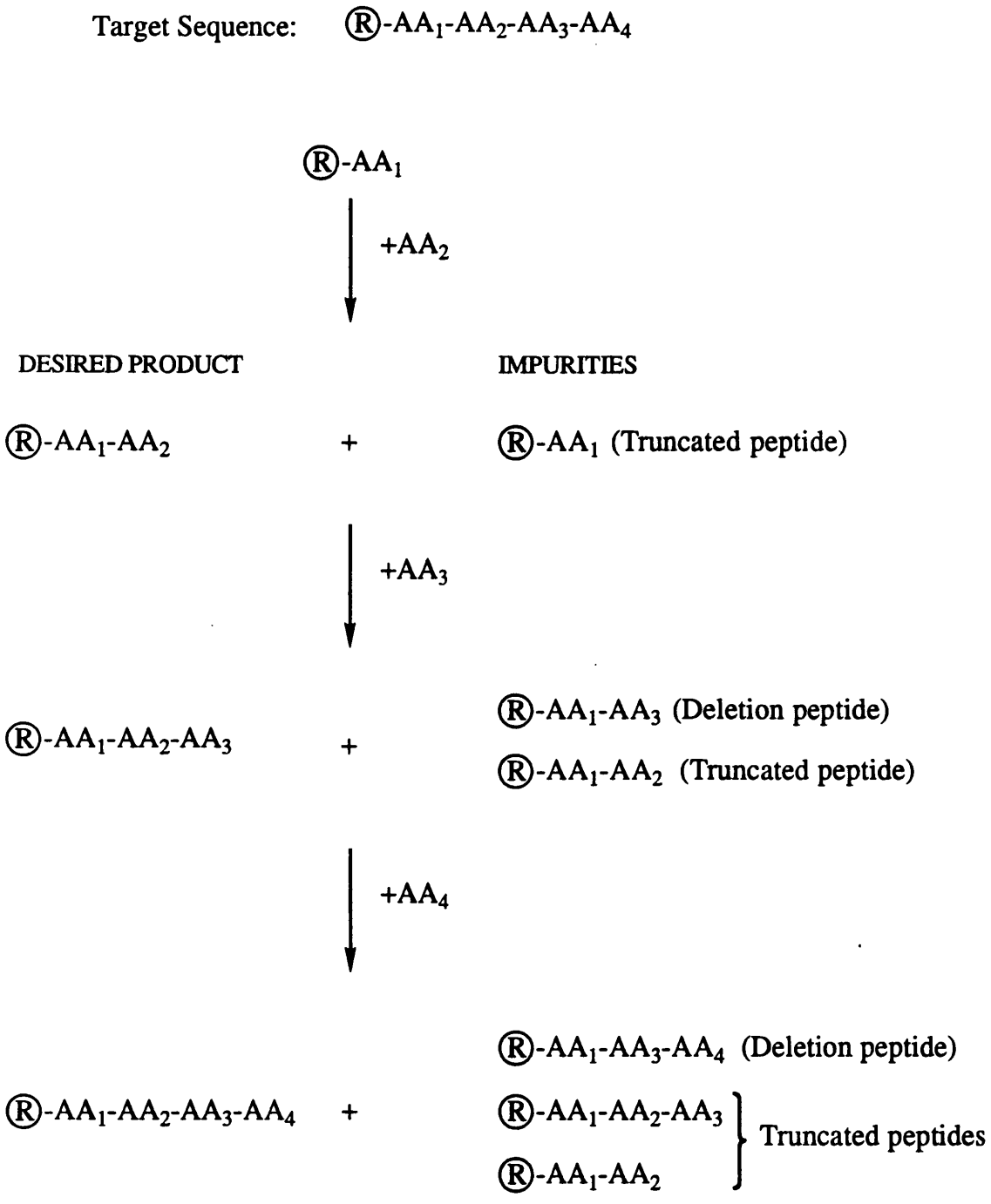
SIDE REACTIONS IN SPPS

Peptide synthesis, involving the three basic steps of activation, coupling and deprotection, would ideally result in the formation of an homogeneous product. However, the desired reaction that is based on unequivocal chemistry is reserved for only a few examples in organic synthesis. Peptide synthesis and all the reactions that this entails are usually accompanied by competing undesired side reactions. Although the chemical protocols that have been developed for SPPS are carefully conceived to minimise the occurance of side reactions, there are however a number of side reactions that can occur at each step in the synthesis. The significance of these extraneous reactions often relate to the amino acid being coupled or correlated to the particular sequence being synthesised. They are particularly problematic when large molecules are prepared, since a high number of repeating chemical steps are involved, thus resulting in the accumulation of by-products (Fig. 1). Here follows a list of the major competing side reactions leading to the generation of impurities which are of particular importance in SPPS:

Deletion peptides

Since coupling reactions generally fail to go to completion, a small but significant proportion of the free amino groups from the penultimate coupling reaction can react with the next amino acid (Fig. 1). The peptide chain that now lacks one residue can then react with the next and subsequent amino acids that follow during the stepwise synthesis. When this undesired side reaction repeats itself, to a lesser or greater extent at every amino acid coupling, the result is that a family of peptide sequences exist lacking one or more residues. Some residues are particularly susceptible to incomplete reaction due to steric hindrance generated by their bulky side chain groups (eg. isoleucine).

An alternative route to the formation of deletion sequences arises from the reactions used to derivatise resins with the first amino acid which often result in



Scheme 1. Scheme outlining generation of deletion and truncated impurities during SPPS

the formation of benzaldehyde groups on the solid support. The benzaldehyde groups are not stabilised and are released during coupling cycles forming Schiff's base imines with the neutralised free amino terminus thus preventing acylation by the incoming activated amino acid. Since the Schiff's base imines are reversible the unacylated peptide chain can be released [Kent, 1984] and available for coupling to the next amino acid which results in the generation of a family of deletion peptides. For longer peptide sequences (>40 residues) these deletion forms, lacking one or more residues, can have similar chromatographic characteristics to the target sequence making purification troublesome.

Terminated peptides

During peptide synthesis the N-terminal residue can become terminated ('capped') thus making it unable to react with the next amino acid to be coupled. One possible cause of peptide chain termination is the presence of contaminants in the commercially available Boc amino acids, usually as secondary urethane-protected compounds.

Trifluoroacetylation

Trifluoroacetylation of the free N-terminal amino group can occur due to the instability of the benzyl ester bond between the resin and the peptide chain to the TFA used to remove the Boc group [Gutte and Merrifield, 1971]. When this occurs the peptide is prematurely cleaved from the resin and replaced by a TFA benzyl ester group. The latter is mildly activated to nucleophilic attack by the neutralised amino group of the N-terminal residue after treatment with TFA, thus preventing further coupling [Kent *et al*, 1979].

The typical occurrence of chain loss, deletion, truncated and trifluoroacetylation peptide formation is 0.01, 0.05, 0.02 and 0.02% respectively, when stable resins are used [Kent and Clark-Lewis, 1985]. For short peptides where the coupling efficiency is high, the significance of 0.1% impurity at each step of chain elongation is not critical. Unfortunately for longer sequences the cumulative effect is dramatic as illustrated in Table 2.

Average	Number of residues				
Yield per step	11	21	31	51	100
96%	66%	44%	29%	13%	1.7%
99.8%	98%	96%	94%	90%	82%

Table 2. Recovery of desired sequence after stepwise SPPS. [Table taken from Kent and Clark-Lewis 1985]

Racemisation

When the acidic proton from the carboxyl group is removed and replaced by the activating group, with its electron withdrawing properties, the activity of the α-hydrogen is enhanced and therefore more susceptible to base catalysed racemisation. The possible loss of chiral purity is a particularly important consideration during the coupling of protected fragments in segment condensation and protein synthesis. The degree of risk to chiral purity depends on several factors including reaction solvent, presence or absence of tertiary bases and most importantly on the method of activation. Unfortunately, in the latter case, any increase in activation will tend to increase the acidity of the α-proton and therefore render it sensitive to abstraction by bases. The effect of the polarity of the solvent has been shown to be significant [Bodanszky and Bodanszky, 1967; Kovacs *et al*, 1975]. While polar solvents (eg. DMSO or DMF) promote racemisation, the reaction with less polar liquids (eg. acetonitrile, DCM, THF or dioxane) is not so fast. The amino acid derivatives, however, are often insoluble in the latter and coupling efficiency is reduced due to their poorer resin-swelling properties. As a result in SPPS a compromise exists where the activation procedure is conducted in DCM containing a proportion of DMF to aid dissolution. In addition, coupling reactions are conducted at high concentration to ensure high coupling rates. The urethane-type N-terminal protecting group (eg. Boc) has been found to be particularly useful in preventing the rearrangements that are undergone in those solvents that promote racemisation [Benoiton and Chen, 1981]. The addition of nucleophilic molecules eg. HOBt [Konig and Geiger, 1970abc] and N-

hydroxysuccinimide [Weygand *et al*, 1966; Wunsch and Drees, 1966] that offer acidic protons more easily abstracted than the proton from a chiral centre can also help maintain chiral integrity of the amino acid.

Loss of chiral purity can also occur through protonation of the carbonyl oxygen in the presence of strong acids (eg. HF [Sakakibara and Shimonishi, 1965]) used to remove side chain protecting groups and, in SPPS, cleavage of peptide from the solid support. The problem of racemisation at this stage in the synthesis can generally be overcome by performing the cleavage reaction at 0°C.

Cyclisations

Diketopiperazine (DKP) formation is a particular hazard as a result of the thermodynamically prefered six-membered ring structure. This undesired side reaction is accelerated by bases and is particularly problematic in SPPS when the C-terminal amino acid is a proline or glycine [Mergler *et al*, 1988] resulting in premature loss of peptide from the resin [Lukenheimer and Zahn, 1970; Gisin and Merrifield, 1972; Khosla *et al*, 1972; Rothe and Mazanek, 1974].

Peptide sequences containing Asp-Gly together are notorious for cyclisation due to β-benzyl-aspartyl protonation during treatment with strong acids [Ondetti *et al*, 1968]. The advent of β-cyclohexyl protection [Tam *et al*, 1979] for the side-chains of Asp and Glu has somewhat avoided this particular side reaction.

Alkylation

Alkylation of sensitive amino acids (eg. Tyr, Met and Trp) can occur during removal of protecting groups and cleavage of the peptide from the resin, due to the generation of alkylating species when reactive cations interact with solvent molecules. The inclusion of scavengers like dialkyl sulphides or thiols can limit this problem by 'mopping-up' the reactive species before they have the opportunity to damage the amino acid side chains.

Chain fragmentation

Strong acids like liquid HF or TFMSA are capable of causing an N→O shift. For example the acyl group, when attached to the nitrogen group of a serine residue in the peptide chain, can be encouraged to migrate to the hydroxy oxygen on the side chain [Shin *et al*, 1962; Sakakibara *et al*, 1962]. This alteration in the peptide chain backbone can be reversed, in solution at pH7-8, with sodium bicarbonate solution, although some hydrolysis can occur leading to fragmentation of the chain.

Side reactions associated with individual amino acids

Table 3 describes the major side reactions that are attributed to the chemical properties of each amino acid. The reactions leading to these unwanted products can occur at different stages during peptide synthesis (ie. during activation, coupling and final cleavage of peptide from the resin).

Amino Acid	Side Reaction	Consequence and Remedy
Glutamine	Dehydration of carboxamide to nitrile. Hydrolysis of carboxamide to acid. Ring closure of N-terminal residue to pyroglutamyl, catalysed by weak acids eg. Boc amino acids and aided by polar aprotic solvents like DMF.	As described for Asn. As described for Asn. Peptide chain termination.
Cysteine/ Cystine	Disulphide interchange in strong acids catalysed by thiols. Base catalysed β-elimination.	 Leads to racemisation.
Methionine	Oxidation of thioether to a sulphoxide. Alkylation of sulphur atom during acidolytic cleavage of protecting groups.	Reversible on addition of reducing agent. Protected with scavengers.

Amino Acid	Side Reaction	Consequence and Remedy
Lysine	ε-amino group chemically similar to the N-terminus.	Loss of ε-amino protection can be followed by ε-acylation.
Arginine	Intramolecular acylation at either N-group of guanidine during activation.	Total protection against lactam formation only offered when both N-groups are protected.
Serine/ Threonine	Alcoholic -OH susceptible to acid and base catalysed acylation. N→O acyl migration in strong acids eg. HF.	Limit time of exposure to catalysing agents. Rearrangement reversible in aqueous solution of sodium bicarbonate.
Aspartic Acid	Tendancy for β-alkyl aspartyl residues to change to amino-succinyl groups (catalysed by acids and bases). Intramolecular nucleophilic attack. Asp-Pro sequence sensitive to acid.	Ring closure. Succinimide ring can be opened with water or dilute base to give the β-aspartyl impurity. DKP formation. Peptide chain degradation.
Glutamic Acid	Friedel Crafts acylation of anisole to the γ-carboxyl, by liquid HF. Cyclisation due to activation of midchain glutamyl γ-carboxyl Cyclisation due to activation of N-terminal glutamyl γ-carboxyl	Capping of glutamyl side chains. Formation of six membered ring system leading to transpeptidation. Cyclisation into five membered pyroglutamyl ring system.
Asparagine	Dehydration of side chain carboxamide during activation. Hydrolysis of carboxamide to acid especially near Arg or Lys.	Conversion to nitrile suppressed by the addition of HOBt or reversed by hydration with strong acids eg.HF. Avoid extremes of pH and high temperatures especially with HF.

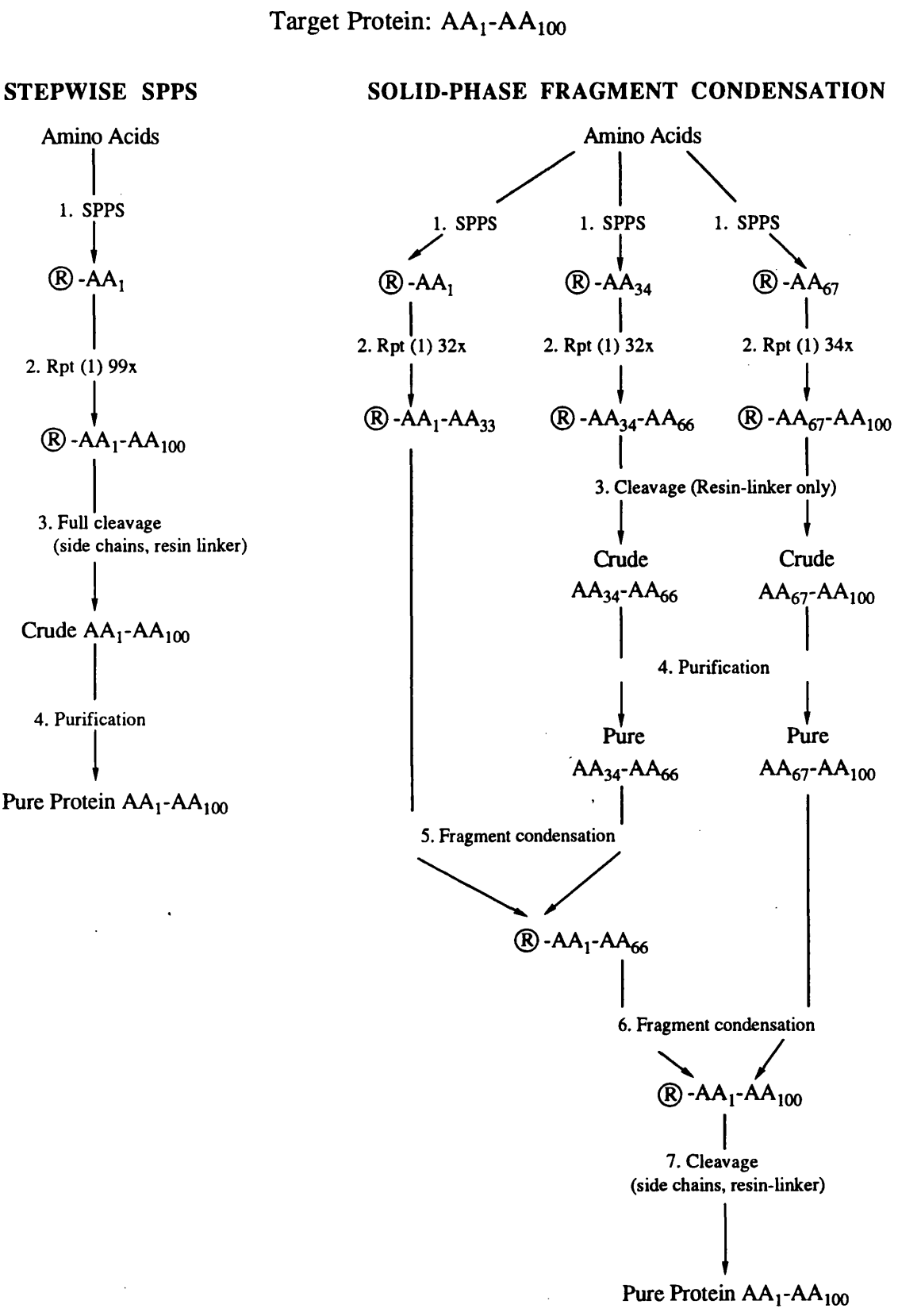
Amino Acid	Side Reaction	Consequence and Remedy
Glycine	The absence of a bulky side-chain can allow the acylated amino group to accept another acyl group. The presence of base can promote the formation of hydantoin rings where Gly is the second residue in the sequence.	Branching of peptide chain. Termination of peptide chain.
Valine/ Isoleucine	Branching side chain at β-carbon causes steric hindrance.	Poor coupling yields. O → N migration of intermediates during DCC activation.
Proline	Ring conformation generates steric hindrance. Particular geometry of imine structure. Secondary amine susceptible to N-alkylation.	Formation of N-acyldicyclohexylurea derivative. Aids DKP formation. Peptide chain terminated.
Phenylalanine	Aromatic ring possesses no activating substituents.	No important side-reactions.
Tyrosine	Phenolic hydroxyl group sensitive to electrophilic aromatic substitution.	C-benzylation. Prevented, during cleavage by addition of scavengers or protection of -OH with halogen substituted benzyl groups.
Tryptophan	Indole ring sensitive to oxidation, dimerisation and alkylation in acidic conditions.	Cleavage reaction performed at low temperatures and with scavengers. Formylation of the ring offers some protection.
Histidine	Imidazole ring slightly basic	Catalyses O-acylation of -OH containing residues or racemisation by intra-molecular proton abstraction.

Table 3. Side reactions of individual amino acids. [Compiled from Bodanszky 1984 and the references cited therein]

Even with thorough preparation and consideration of possible causes leading to byproducts, there will always be a degree of uncertainty as to the success of a synthesis due to sequence related problems rather than to difficulties with specific amino acids. Sequence dependent problems often resulting in poor coupling yields have been attributed to intermolecular β-sheet formation and occurs most usually in peptides containing 10-20 residues [Meister and Kent, 1983]. The use of DMF with its superior resin swelling properties combined with the solvating properties of the solid support itself can help to minimise this phenomenon of peptide chain aggregation. Another example of a sequence related problem can occur with Asp-Gly and Asp-Pro amide bonds which are particularly sensitive to the acidolytic conditions used to remove the N-terminal protecting group (Boc chemistry); in both cases the result is fragmentation of the peptide chain [Piszkiewicz *et al*, 1970]. In addition, Asp-Gly sequences are notorious for promoting acid catalysed β-shift side reactions [Merrifield, 1967], where the peptide backbone transfers to the side chain carboxyl group of Asp.

Coupling efficiency has been shown to be improved if the reaction is performed at elevated temperatures. Lloyd *et al* [1989] reported that the failure sequences for rat ANF(1-28) [Otteson *et al*, 1988] could be improved when the couplings were performed at 60°C. Unfortunately, the advantages of reduced coupling times had to be balanced against a significant amount of dehydrated material. The occurrence of other side reactions at elevated temperatures such as racemisation and cyclisation (eg pyroglutamate formation) were not discussed.

The addition of inorganic salts (eg. LiCl) to the coupling mixture have been shown to improve the swelling properties of some resins, particularly the more polar polystyrene and Kieselghur-based solid supports by disrupting intermolecular interactions [Thaler *et al*, 1991]. and affect the reaction dynamics. In some cases the rate of coupling was reduced but the overall addition of amino acid exceeded that achieved using conventional techniques.



Scheme 2. General scheme for the chemical synthesis of peptides

PEPTIDE AND PROTEIN SYNTHESIS

The aim of this thesis was to develop, improve and use synthetic and chromatographic techniques for the preparation of large polypeptides. To this target the work has focused on three different areas (i) the synthesis and purification of large peptides using highly optimised stepwise SPPS chemical protocols, (ii) the design, synthesis and application of chemical molecules to aid the purification of large synthetic peptides and (iii) the synthesis of large peptides using orthogonal approaches (Fig. 2).

The desire to extend the newly developed technology of SPPS quickly accelerated with the notion that proteins could be synthesised in a routine manner. Several well-known papers exist that describe early attempts to synthesise protein sequences ranging from 104 to 188 residues in length; Ribonuclease A (124 residues [Gutte and Merrifield, 1969]), Cytochrome c (104 residues [Sano and Kurihara, 1969]), Growth hormone (188 residues [Li and Yamashiro, 1970]), Ribonuclease T_1 (104 residues [Izumiya *et al*, 1972]) and Lysozyme (129 residues [Sharp *et al*, 1973]). Unfortunately, the initial optimism held by peptide chemists quickly dissipitated when normal criteria used to judge purity could not be successfully applied in the case of the synthetic proteins. Since these early days, the development of high-resolution chromatographic and electrophoretic techniques for purification and analysis, together with structural characterisation supplied by mass and nuclear magnetic resonance spectrometry have offered the means by which the integrity of the product can be judged more precisely.

The advent of fully automated and reliable SPPS instruments has potentially made the task of synthesising proteins a more attainable goal. Since the routine was removed from the activation, coupling and deprotection cycles, research was centered on improving finer details of amide bond formation.

OPTIMISATION OF SYNTHETIC PROTOCOLS

Apart from using stable amino acid-resins that are free from extraneous functionalities and controlling the purity of the Boc protected amino acids, there are other aspects of the synthetic protocol that can be improved to reduce the occurrence of impurities. The optimization of SPPS can also be extended to chemical procedures as outlined in Table 4. The consequence is that application of all the considerations mentioned can result in a coupling efficiency of >99.5% per amino acid, to produce a product, after cleavage, with a high degree of purity (even for large peptides; Table 2).

Feature	Properties	Advantages
1% copoly (S-DVB)	swollen polymer network	enhances peptide solvation/reactivity
"Pam" resin	clean, stable	minimal side reactions
Flexible reaction protocols		rapid, minimal side reactions
Coupling in DMF	good solvent	complete reaction
Quantitative monitoring	to <0.1%	real time feedback control
Boc/benzyl protected amino acids	high purity, HF-labile	cheap, well understood
HF deprotection	mild, general	universal solvent, well understood

Table 4. Optimised protocol for stepwise SPPS using Boc chemistry. [Taken from Kent and Clark-Lewis 1985]

In an effort to make the task of purifying synthetic peptides a simpler one, peptide chemists have described many techniques that involve a repetitive 'capping' procedure after each coupling reaction or the attachment of an additional molecule to the N-terminus of the desired sequence, at the end of the

synthesis. In the former case the unreacted amino groups are derivatised with a molecule to alter the chromatographic characteristics of the impurities while the latter relies on the alteration of the properties of the desired peptide.

'Capping' procedures

To prevent the unreacted N-terminal amino group from further reacting with subsequent amino acids, several acylating agents have been employed which form stable adducts with the residual amino groups thus preventing deletion peptide formation. The original and most commonly used capping reagent is acetic anhydride which is used in large excess in combination with an organic base [Merrifield, 1963; Blake and Li, 1968; Mitchell *et al*, 1978]. Other compounds have been used with equal or better success eg. N-acetylimidazole [Markley and Dorman, 1970] and isopropenyl formate [van Melick and Wolters, 1972; Wolters *et al*, 1974], although these require longer reaction times. An alternative group is 3-nitrophthalic anhydride [Wieland *et al*, 1969; Penke and Birr, 1974] which facilitates the separation of terminated impurities by ion exchange chromatography. As a consequence of this procedure a series of terminated sequences will be generated, which it is hoped will have chromatographic properties sufficiently different from the target peptide to allow separation.

PURIFICATION

For the purification of short synthetic peptides, up to about 40 residues in length, the application of conventional chromatographic techniques have proved to be extremely successful. Standard laboratory methods used to obtain the pure product are based on separations by chromatography. Many different stationary phases are available which exploit different physico-chemical properties of the peptides. Table 5 below lists some of the important chromatographic techniques, the charateristics upon which separation is achieved and some of their limitations.

Property	Technique	Advantages	Disadvatages
Size and shape	Gel filtration	Uncomplicated and straightforward	Peptide-peptide and peptide-gel interactions
Net charge and distribution of charged groups	Ion-exchange chromatography	High resolving power	Requires desalting
Isoelectric Point	Chromatofocusing	Simple operation	Large sample volume Labour intensive
Hydrophobicity	Reversed-phase chromatography	Simple operation	Loss of bioactivity
Metal binding	Immobilised metal ion affinity chromatography	Specific for His/ Trp containing peptides	Requires metal binding sequence
Affinity for ligand,antibody etc	Affinity chromatography	Highly specific	Cost

Table 5. Comparison of some chromatographic methods for the purification of proteins

Other important techniques used for the purification of peptides include counter current distribution where separation is achieved by differential partitioning between two solvents. This method can be used successfully for the

purification of non-hydrophilic peptides, especially protected peptide fragments, that could not normally be purified in large quantities due to the poor solubility in most solvents [Nyfeler *et al*, 1991].

Where repetition of one technique may not yield the homogeneous product, a successive procedure or series of procedures which utilise different separation criterias could be employed. However this approach is time consuming and requires the operator to have a good practical knowledge of different purification methods in order to optimise the series of steps. Running the same material through different procedures will also reduce the final yield of product.

As previously mentioned the reciprocal strategy involving the derivatisation of the growing peptide chain with a 'chromatographic probe' has been investigated by many researchers. The methods of purification are based on the permanent or reversible coupling of molecules to the synthesised peptides to alter their chromatographic nature. The first examples cited for SPPS are based on the further addition of one or more amino acids eg. a Lys residue [Suzuki *et al*, 1976] or a Cys-Met dipeptide [Krieger *et al*, 1976; Lindeberg *et al*, 1978] or Cysteic acid [Hubbuch *et al*, 1977] for purification on a cation-exchange column, an organomercurial affinity column or by counter-current distribution, respectively. Recovery of the desired peptide from the extended sequence was accomplished after CNBr treatment or Edman degradation. Using a similar means of removal covalently bound sulfonyl groups [Birr *et al*, 1970; Dwulet and Gurd, 1976] were used to enable separation on a mild anion exchange column. Merrifield and Bach [1978] described the use of the base-labile Fmoc group, with a sulfonyl side chain at position 2, to derivatise non-terminated peptides and so facilitate purification by ion-exchange chromatography. The Fmoc sulfonyl probe was used for the purification of neutral peptides, or peptides containing only one charged residue. In both cases the peptides were no longer than six residues and the application of this sulfonyl Fmoc derivative to the purification of larger peptides has not yet been described.

Recently Engstrom *et al* [1989] purified peptides between 25 and 40 residues in length on an immobilised metal-ion column. General applicability of this method is also lacking due to its reliance on the presence of His and/or Trp residues in the N-terminal region of the peptide.

Affinity chromatography has been extended to the purification of synthetic proteins. Lobl *et al* [1988] constructed the 153 residue IL-1 synthetic protein and went on to attach biotin to the N-terminus residue through a covalent linkage, thus enabling separation of biotinylated chains on an avidin-agarose column. Although this purification was successful biotin could not be removed from the purified protein due to its covalent linkage.

The strong affinity biotin has for avidin was also exploited by Bannow *et al* [1991]. Retrieval of the purified of 99 residue SIV protease polypeptide from the biotinylated form was achieved through self-processing, where the protein itself cleaved an Ala-Pro sequence to release the desired product.

FRAGMENT CONDENSATION

Towards Fischers original belief that proteins could be chemically synthesised work has focused on two different approaches. While one group of researchers have concentrated on optimising the chemistry involved in stepwise solid phase synthesis, others have taken the approach that smaller fully protected peptide fragments can be linked together to form the desired protein sequence. Convergent fragment condensation [Finn and Hofmann, 1976] involves synthesising short segments of the desired sequence which retain their side chain protecting groups. Since the peptide fragments are small it was envisaged that conventional chromatographic techniques could be applied to purify them to homogeneity. In the majority of cases, previous research towards the chemical synthesis of polypeptides has been performed in the solution phase [Denkewalter *et al*, 1969; Finn and Hofmann, 1976; Kenner *et al*, 1979; Yajima and Fujii, 1980] or a combination of protected peptide synthesis on a resin support followed by fragment coupling in solution [Kaiser *et al*, 1989; Kamber and Riniker, 1991; Nyfeler *et al*, 1991]. The third option is to couple purified protected fragments, that have been synthesised on a resin, directly onto a resin-bound fragment. Some encouraging results with regards to coupling yields have been obtained [Yajima *et al*, 1974; Maruyama *et al*, 1976; Barlos *et al*, 1991]. Unfortunately, many problems were encountered the most serious of which was the poor solubility of protected peptide intermediates in almost all organic solvents. The net result of the latter property was the inability to form a peptide bond [Kiyama *et al*, 1984] due to the tendancy of the extremely hydrophobic peptide chains to aggregate with eachother [Pillai and Mutter, 1981]. Other problems exist due to racemisation of the C-terminal amino acid [Fridkin and Patchornik, 1974] when this is not a glycine or proline residue. There is also a lack of highly resolving methodologies for the purification of protected peptides and the degree of experience required by the peptide chemist is often too great to be applied in most laboratories for the synthesis of polypeptides by solution phase.

The alternative technique of solid phase fragment condensation helps to overcome some of the problems encountered with solution phase methods by combining the advantages of stepwise SPPS with the ability to purify short fully protected peptide fragments. Polypeptides synthesised by fragment condensation as opposed to stepwise SPPS have also been found to have greater biological activity [Yajima *et al*, 1974]. However to be successful the technique is dependent on a rapid and efficient method of preparing and purifying the protected peptide intermediates. The 'linker' group on which the protected peptides are synthesised must offer (i) ease of derivatisation to the resin support, (ii) allow efficient stepwise synthesis of fragments, (iii) stability to the normal coupling conditions and (iv) the cleavage protocol should leave side chain and N-terminus protecting groups unaffected and be quantitative.

In recent years many linker groups have been developed some offering true orthogonality, while others rely on differential sensitivity to the reagent used for cleavage. Examples of the latter are the super acid sensitive resins used for Fmoc chemistry ie. Sasrin developed by Mergler *et al* [1988], a trialkoxy-diphenylmethylester resin [Rink, 1987] and a HMPB (hydroxymethylmethoxy-phenoxybutyric acid) linker [Kamber and Riniker, 1991], which have been used successfully to synthesis protected peptide fragments up to 20 residues in length. The peptide-resin bond and side chain protecting groups are cleavable on the basis of differential acidolysis ie. the concentration of acid (eg. TFA) required to cleave the peptide-resin bond is less than that needed to fully deprotect the amino acid side chains. However differential acidolysis is not ideal since the selectivity of the acid for the peptidyl-resin linker over side chain protecting groups is not absolute [Atherton *et al*, 1981]. This is particularly important for more acid sensitive protecting groups eg. Fmoc-Lys(Boc). Furthermore, the mild acidity of the coupling catalysts (eg. HOBt) can cause premature loss of peptide chains during chain assembly due to the high acid sensitivity of the resin to peptide linker [Rink, 1987]. Consequently the most desirable methods are based on chemical selectivity where the peptide-resin bond is broken using a

technique that does not involve the same chemical process as that used to cleave the side chain protecting groups.

In an attempt to address these problems several linker groups have been proposed for use with Boc chemistry that are cleaved by organic bases or by photolysis. Additionally, two allylic anchoring groups namely 4-bromocrotonic acid [Kunz and Dombo, 1988] and 4-trityloxy-2-but-2-enyloxyacetic acid [Guibe *et al*, 1989] employ an hydrogenation procedure (at atmospheric pressure) to cleave the peptide from the resin. Since the cleavage conditions leave all side chain and N-terminal protecting groups intact, these allylic resin linkers are compatible with both Boc and Fmoc chemistries [Kunz *et al*, 1991]. The resin-linker molecules have been used to synthesise a variety of protected peptides [Lloyd-Williams *et al*, 1991a].

One of the techniques used to develop an orthogonal approach to the removal of protected peptide chains from the solid support was photolysis [Rich and Gurwara, 1973; Rich and Gurwara, 1975; Wang, 1976; Tjoeng *et al*, 1978; Tam *et al*, 1980]. The linker used was based on a phenyl ring substituted with a nitro group at the *ortho* position, which increases the polarity of the resin and consequently improves swelling in polar solvents (eg. DMF) [Rich and Gurwara, 1975]. Unfortunately, the former property also promotes the loss of peptide chains through diketopiperizine formation [Giralt *et al*, 1981]. This problem can be overcome by incorporating the third amino acid using the method of Suzuki and Endo [1977] or by coupling the second and third amino acids as a preformed dipeptide. Cleavage of the peptide sequence from the resin was achieved by photolysis, without affecting benzyl-based side chain protecting groups. Histidine protected with dinitrophenyl group could not be used due to its instability to photolysis [Rich and Gurwara, 1975]. The generation of impurities during the cleavage procedure has been experienced and attributed to the formation of secondary byproducts of photolysis (eg. azo compounds resulting from *o*-nitrosobenzaldehyde formation [Patchornik *et al*, 1970]) and the long reaction time required [Whitney *et al*, 1984].

Several laboratories have explored resin-handles that are sensitive to mild base catalysed cleavage reagents. The Fmoc based 9-(hydroxymethyl)fluorene-4-carboxylic acid (4-COOH-Fmoc) was described by Mutter and Bellof [1984] as a novel anchor group for the synthesis of protected peptides which used acid-labile N-terminal amino protecting groups. The 4-COOH-Fmoc molecule was stable to the normal treatments of Boc chemistry. The application of the base-labile handle was demonstrated by the synthesis of a model tetrapeptide. It was reported that the loading of the first amino acid onto the 9-hydroxymethyl functionality proceeded in high yield without racemisation. The synthesis of a similar Fmoc based molecule was reported by Liu *et al* [1990]. In this case attachment of the handle to the resin support was through a carboxyl group at position 2 on the fluorene ring. Unlike the fluorene handle developed by Mutter and Bellof, an extra methylene group separated the ring from the carboxyl group, which Liu *et al* [1990] argued, rendered the linker less susceptible to premature β-elimination, due to the slight basicity of the free amino group of the growing peptide chain. Addition of mildly acidic HOBt to the coupling mixture was found to counteract this undesired effect and thereby maximise the yield of peptide. To investigate the application of both these linker molecules to solid phase synthesis of protected peptides, five residue [Mutter and Bellof, 1984] and a seven residue [Liu *et al*, 1990] peptides were constructed using Boc chemistry.

A 2-[4-(hydroxymethyl)phenylacetoxy]propionyl-resin synthesised by Whitney *et al* [1984] represents a second type of resin sensitive to base-catalysed cleavage conditions. A series of protected peptides were synthesised ranging from seven to sixteen residues in length, all with C-terminal glycine to avoid possible racemisation during subsequent fragment coupling procedures. This resin-linker was found to overcome many of the drawbacks associated with the use of the 2-bromopropionyl-resin including loss of peptide chains due to Schiff base formation between the C-terminal glycine and carbonyl group of the propionyl-resin. Diketopiperizine formation was also avoided with this base-labile resin. It was found that the two hindered, non-nucleophilic organic bases

tetramethylguanidine (TMG) and 1,8-diazabicyclo-[5.4.0]-undec-7-ene (DBU) provided the necessary conditions for cleavage. Importantly, the success of cleavage was not dependent on the length of peptide chain. Both Whitney *et al* [1984] and Liu *et al* [1990] advise caution when subjecting peptides containing aspartic or glutamic acids to organic bases (eg. piperidine) due to the possibility of cyclisation to pyroglutamic and aspartimide products. The side reaction can be avoided by using cyclohexyl as opposed to the benzyl ester protecting groups [Schon *et al*, 1983].

Other methods have been developed to generate protected peptide fragments. Tam *et al* [1977] found that when potassium cyanide was complexed with 18-Crown-6, its nucleophilicity could be enhanced to an extent that it became a suitable reagent to cleave Boc-aminoacyloxyacyl-polystyrene resins.

The principle difficulty of the fragment condensation approach has been the purification of the protected peptide intermediates on a larger scale. Since protected peptides are extremely hydrophobic in nature their solubility in many solvents is limited. Consequently, highly solubilising solvents (eg. DMF) have been used in order to maintain the material in solution and allow purification [Lloyd-Williams *et al*, 1991b]. It has also been shown that protecting the side chains of glutamine and asparagine can improve dramatically the solubility of protected fragments by disrupting the electrostatic forces that encourage aggregation [Atherton *et al*, 1990; Kamber and Riniker, 1991]. For this reason, Fmoc chemistry offers distinct advantages over the Boc method, since the trityl protecting group is maintained on the side chains throughout the synthetic and cleavage procedures. Unfortunately, the xanthanyl group used to protect the Boc derivatives of asparagine and glutamine residues is sensitive to the acidolytic conditions used to remove the N-terminal protecting group and therefore can only be maintained during coupling step.

The development of suitable methods to purify fully protected peptides has been attempted by several groups. The techniques used have included simple precipitation [Liu *et al*, 1990] and chromatographic methods based on polarity [Gabriel *et al*, 1977; Barlos *et al*, 1991], size [Kaiser *et al*, 1989],

hydrophobicity [Nyfeler *et al*, 1991] or a combination of different techniques [Pedroso *et al*, 1982]. Precipitation whilst quick and easy to perform generally tends to yield all peptidic material and hence peptidic impurities. Chromatographic separation on the basis of polarity using silica gel, especially on open-column apparatus, is generally slow and suffers from poor resolution of closely related compounds. Gel filtration on Sephadex LH20 [Nystrom and Sjovall, 1963; Meienhofer *et al*, 1971], LH-60 [Zeiger and Anfinsen, 1973] and G-50 [Galpin *et al*, 1975] have been described using chloroform, DMF and alcohols. The low efficiency of these techniques however required the use of large columns and long elution times (1-4 days). Probably the most important method of purification is based on hydrophobicity. Reversed-phase chromatographic media has been developed for MPLC systems [Lloyd-Williams *et al*, 1991b] but require large quantities of DMF, both in terms of volume and proportion in the eluting solvent to maintain the peptide in solution. However, the advantage of MPLC is that the column can handle a high loading of crude material.

The poor solubility of protected peptide fragments not only creates problems for purification but is critical to the success of fragment condensation procedures. In some respects fragment condensation onto a resin bound peptide does present certain advantages since the resin matrix offers a solubilising effect [Kent, 1985]. In many instances poor coupling yields are obtained possibly due to inter- and intramolecular interactions between the peptide chains [Albericio *et al*, 1989]. Thus it has been shown, through NMR experiments, that long peptide chains experience restricted mobility, especially rotational movement, due to the physical structure of the resins interconnecting matrix. The conventional microporous polystyrene-1%-divinylbenzene was found to possess the physical properties most suited to segment coupling through its superior swelling characteristics [Albericio *et al*, 1989]. Additionally, it has been found that the use of appropriate solvent, temperature and a sufficient excess between 1.25 and 2.5 equivs. of protected peptide helped to achieve quantitative coupling of fragments up to four residues in length. The choice of coupling reagent has

important implications with regards to the optical purity of the C-terminal residue. Consequently, whenever possible this amino acid should be a glycine or proline to avoid problems of racemisation.

The practical difficulties encountered when using protected peptides can be partially overcome by purifying the short peptide fragments in an unprotected form and then chemically ligating them with minimal protection. This approach was illustrated by the synthesis of a 92 amino acid protein [Blake *et al*, 1986]. A highly inventive method of protein synthesis was suggested by Schnolzer and Kent [1992]. They coupled two unprotected synthetic peptide segments corresponding to the two halves of the HIV-1 protease sequence, through a thioester linkage. Consequently, the resulting protein was not completely natural. An alternative technique involves enzyme-catalysed peptide bond formation and has been explored by Nakatsuka *et al* [1987] and Wallace and Campbell [1989]. The advantage arising from the use of peptidases like trypsin are that the coupling is absolutely specific, without side reactions and requires no side chain protection of the fragments to be coupled. The success of the method has been variable and possibly more could be achieved by using these specific proteases as stoichiometric reagents rather than as catalysts [Kent, 1988].

The review of research in the field of protein and peptide chemistry that has been described previously has included some quite eloquent methods for the synthesis and purification of large peptides. Unfortunately in the latter case the majority of methods are not generally applicable or require specialist equipment and knowledge. However, if protein engineering is to become at a relatively routine procedure further progress in the areas of synthesis and purification are required. This thesis it is hoped provides additional 'pieces of the jigsaw' by assimilating what has gone before and developing methodolgies (with the aid of model peptides), so that eventually, peptides of unlimited length can be chemically synthesised.

CHAPTER TWO

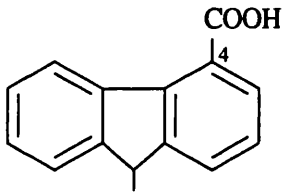
PURIFICATION OF SYNTHETIC PEPTIDES USING CHROMATOGRAPHIC PROBES BASED ON THE FMOC MOLECULE

INTRODUCTION

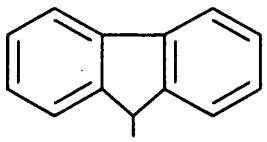
The synthesis of peptide chains on solid phase supports, using the stepwise approach suffers inherently from a proportion of incomplete coupling of the activated amino acid at each step during the synthesis, as was previously discussed in the introduction. The percentage of free amino groups that remain unreacted after a coupling step is often related to either the peptide sequence that may cause intermolecular interactions between adjacent chains or the physical properties of the N-terminus and/or incoming amino acid (ie. steric hindrance). The end result is the generation of a family shorter sequence lacking on or more amino acids which either react with subsequent residues (ie. deletion peptides) or remain terminated at the point of acylation failure (ie. truncated peptides). For relatively short sequences (less than about 40 residues) the separation of these undesired impurities can usually be performed using conventional chromatographic techniques. The purification method proposed is based on the selective incorporation of a probe that would favour separation of the desired peptide chain from closely related impurities. The use of chromatographic probes that potentiate a particular physical characteristic (ie. charge/hydrophobicity) or exploit the affinity of a particular amino acid or molecule for another group has been described. Merrifield and Bach [1978] used a reversible Fmoc group with a substituted sulfonyl moiety to enable the separtion of short peptides (up to six residues) on ion-exchange media. The application of the 2-sulf-Fmoc group to longer sequences was not demonstrated. The strong binding affinity between biotin and avidin lead to the covalent attachment of the former to the N-terminus of the synthetic 153 residue IL-1 protein [Lobl *et al.*, 1988]. Any biotinylated protein was then retained and underivatised impurities simply eluted through the immobilised avidin column. Unfortunately, due to the covalent bond linking the affinity label to the protein,

the biotinylation was irreversible. A reversible approach was suggested by Bannow *et al.*, [1991], where the synthetic protein itself (the 99 residue SIV protease) cleaved the N-terminal Ala-Pro bond. As a result of the requirement for specific amino acid residues the method did not offer a general applicability.

The procedure that has been developed was based on the addition of a molecular label to the N-terminus of the last amino acid of the peptidyl-resin. To avoid the unwanted derivatisation of shorter deletion sequences, a capping procedure following the coupling of each amino acid would be required. The key molecule (9-hydroxymethyl)fluorene-4-carboxylic acid (4-COOH-Fmoc, **1**) was synthesised from fluorene-4-carboxylic acid and possessed the chemical properties of the fluorenylmethoxycarbonyl (Fmoc, **2**) group.



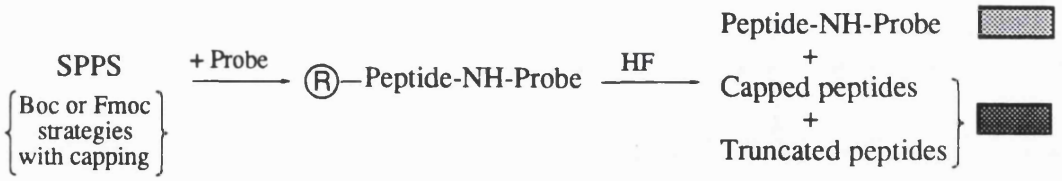
1



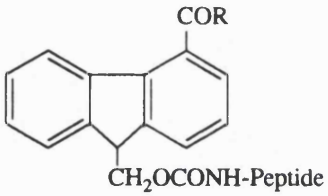
2

It was Mutter and Bellof [1984] who first suggested the 4-COOH-Fmoc group as a base-labile linker for the synthesis of protected peptide fragments. The 4-carboxyl moiety was attached to the solid support while the peptide was attached at the C-terminus to the 9-hydroxymethyl group. For the purposes of this research the carboxylic acid function of 4-COOH-Fmoc enabled the incorporation of extremely hydrophobic groups or charged residues to give molecules with enhanced chromatographic characteristics. It was envisaged that these new molecules (4-COR-Fmoc) could be coupled to peptide chains with the desired sequence, through a base-labile urethane group. Stability of the analogous 2-sulf-Fmoc group to the acidolytic cleavage conditions used to remove side chain protecting groups and cleave the peptides from the resin [Merrifield and Bach, 1978] was also believed to extend to the peptide bound

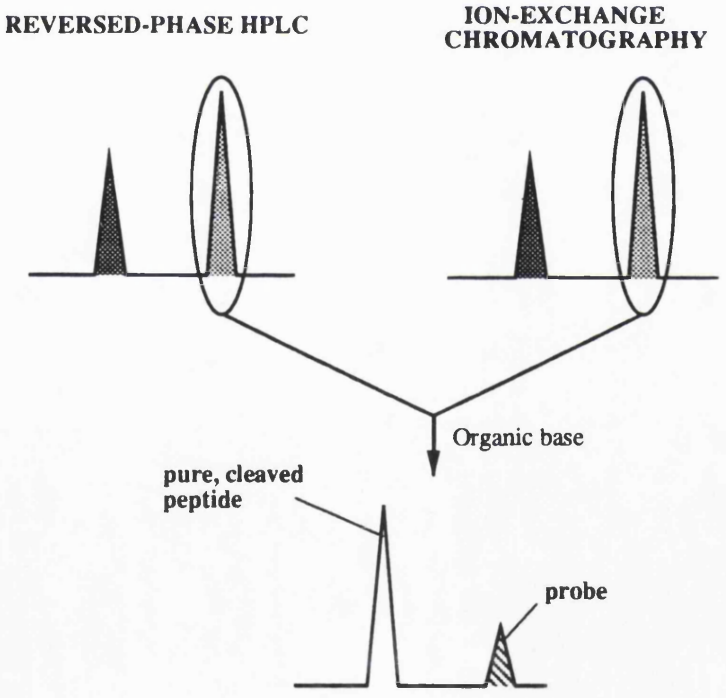
probe (4-COR-Fmoc) molecules. The susceptability of the urethane bond to β-elimination (which is the chemical property exploited for Fmoc chemistry) was used to yield the free peptide following purification. Furthermore, since the groups attached to the carboxylic function of 4-COOH-Fmoc and therefore the chromatographic characteristics could be selected to give optimal performance, purification would be independent of chain length and amino acid composition. The following characteristics of the chromatographic probes were considered during the design stage of the project. Firstly, the derivatisation of the peptidyl-resin with the chromatographic probe should be quantitative and only occur at the N-terminus of the target peptide and not shorter deletion sequences. Thus, to avoid the latter problem an effective capping protocol was performed after the addition of each amino acid. The next problem was how to make the 4-COR-Fmoc derivatised peptide sufficiently different chromatographically from underivatised sequences, especially for longer peptides (greater than about 40 residues), to allow separation. The 4-COOH-Fmoc molecule offered two functional groups, the first to enable the incorporation of groups with particular physical properties and the second to provide a link between the probe and the peptide that could be cleaved simply without affecting the peptide itself. Thus, a highly lipophilic molecule was attached to the reversible 4-COOH-Fmoc probe molecule that was designed to delay the elution of peptide chains to which it was attached. However, in certain instances it is not desirable for the peptide to be in contact with organic solvents and so the versatility of the 4-COOH-Fmoc molecule was such that charged groups, with acidic or basic characteristics, could be attached for purification on ion-exchange media. The application of the probes to a wide range of chromatographic media should permit the purification of synthetic peptides independently of their chain length or amino acid composition. The general philosophy of the proposed purification method is shown in Scheme 1.



Probes:



where R is a lipophilic or charged group



Scheme 1. General scheme showing the application of 4-COR-Fmoc chromatographic probes to the purification of synthetic peptides.

RESULTS AND DISCUSSION

SYNTHESIS AND COUPLING OF THE CHROMATOGRAPHIC PROBES

The method described in the literature [Mutter and Bellof, 1984] for the synthesis of 9-(hydroxymethyl)fluorene-4-carboxylic acid (4-COOH-Fmoc, **1**), involved an initial protection of the 4-carboxyl function with a t-butyl group (**3**:R**1**). Since the reaction required about 7 days to go to completion, alternative protecting groups were investigated as shown in figure 1.

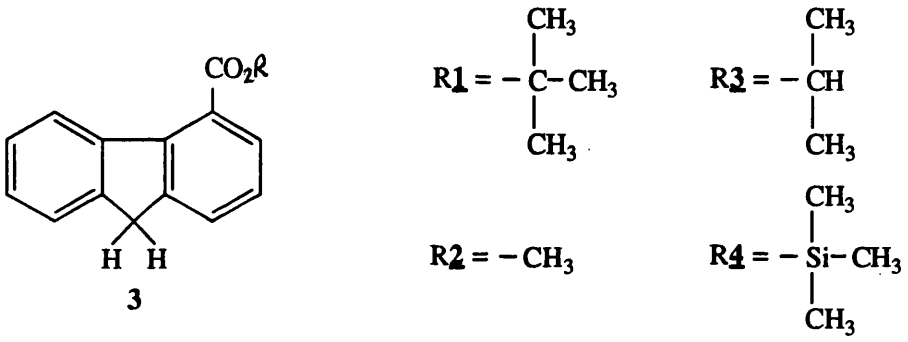


Figure 1. Protecting groups used for fluorene-4-carboxylic acid.

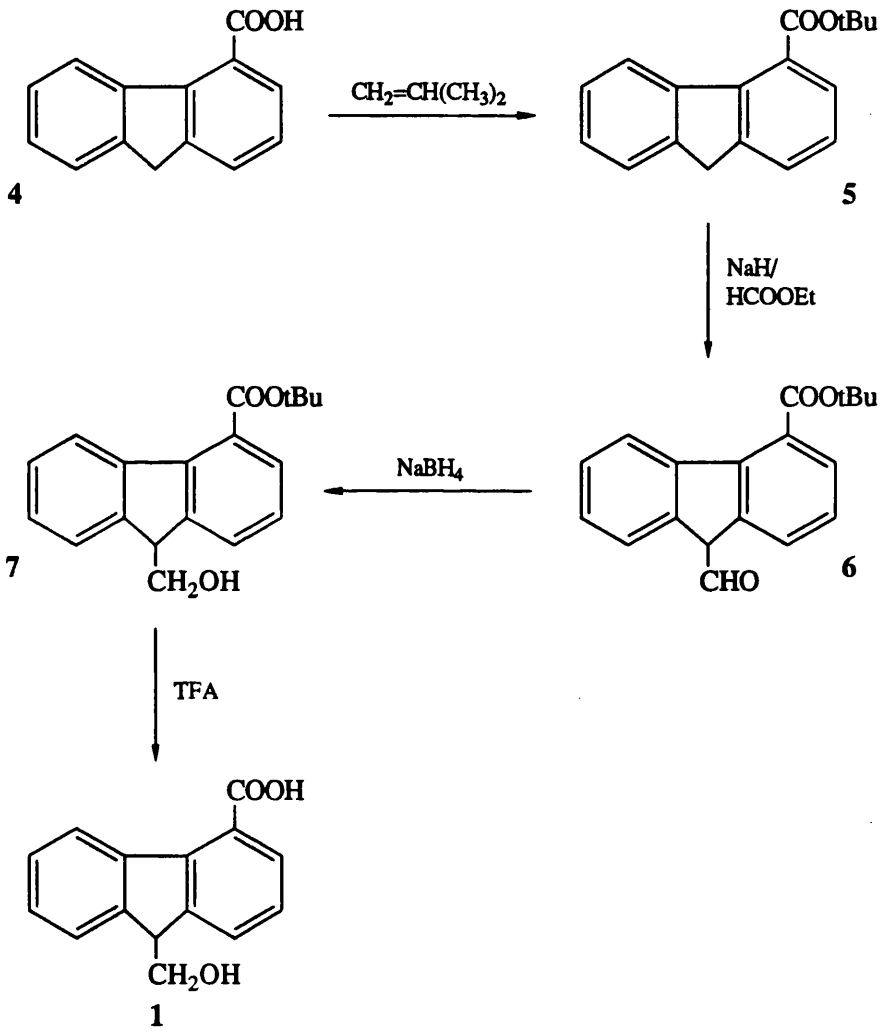
An attempt was made to react the free carboxyl compound, but was unsuccessful due to its poor solubility in the reaction solvent. To shield the carboxyl function and thereby prevent salt formation in the organic solvent, a complex of the potassium salt with 18-Crown-6 was performed. While the complexation was successful, its subsequent solubility in the organic solvent (diethyl ether) was disappointing and this approach was not persued further. Since ionic means of protecting the carboxyl group could not be developed further, direct covalent derivatisation using the groups shown in figure 1 were examined. The synthesis of the methyl ester compound (**3**:R**2**) proceeded smoothly, achieving 79% yield after recrystallisation. Unfortunately, the ester linkage was not stable to the subsequent reaction conditions, which caused

significant hydrolysis to the free acid. An alternative protecting group which was easier to prepare than the tert.butyl ester was the isopropyl ester group (3:R3). Unfortunately, while the reaction with sodium hydride proceeded to the 9-formylated intermediate, the similar reactivity of the isopropyl ester C-H group an additional formylation had occurred. The presence of the di-formylated compound was supported by NMR and FAB-MS data which gave a mass of 28 units higher than expected. It was therefore concluded that due to the instability of the isopropyl protecting group to the reaction conditions, it should be abandoned and attention focused on different groups. The final derivative synthesised was the trimethylsilane group (3:R4) which was found to be too labile for the intended applications. Despite the various attempts no improvements in time and yields were obtained, over the original method, which was therefore used for the synthesis of (3:R1).

Synthesis 9-(hydroxymethyl)fluorene-4-carboxylic acid (1)

Based on the protocol described by Mutter and Bellof [1984] and outlined in Scheme 2, fluorene-4-carboxylic acid (**4**) was treated with 2-methyl-1-propene by bubbling the gas into 10% sulphuric acid in dioxane, at 0°C, until saturated. The acidic, mildly pressurised conditions, resulted in the formation of the t-butyl ester group on the reactive 4-carboxylate group. The esterification was allowed to proceed for 7 days at room temperature and the product was then purified by flash chromatography to yield 80.8%. **5** was then treated with distilled ethyl formate and sodium hydride, under anhydrous conditions to introduce the aldehyde group and give t-butyl 9-formyl fluorene-4-carboxylate (**6**) yielding 81.6% crude product. The crude product (**6**) from the formylation reaction was reduced with sodium borohydride, in isopropanol, to t-butyl 9-(hydroxymethyl)fluorene-4-carboxylate (**7**). The latter product (**7**) was then purified by flash chromatography to yield 89.6%. The t-butyl ester protecting group was removed by mechanically shaking **7** with 50% TFA in DCM for 30 min to give **1**. The final product was purified on silica gel using the flash

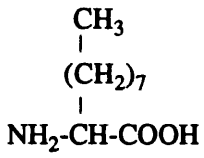
chromatography technique (yields 75.9%). The overall yield for the four step procedure to synthesise 9-(hydroxymethyl)fluorene-4-carboxylate was 45%.



Scheme 2. Synthetic procedure for preparation of key molecule (**1**).

Effect of increased lipophilicity on chromatographic properties

Once the central molecule **1**, which would provide the reversible link to the peptide chain, had been synthesised it was necessary to design suitable groups that would dictate the chromatographic charateristics of the probe. One such candidate available in the laboratory was the lipophilic molecule, α-amino-decanoic acid (**8**) [Albertson and Archer, 1945; Albertson and Tuller,1945].



8

This molecule, it was thought, would increase the retention time of any peptide to which it was attached when injected on a reversed-phase column. To measure the effects of these lipophilic molecules on the retention time of peptides, the 20 amino acid Foot-and-Mouth disease virus (FMDV Vp1 141-160) peptide was synthesised in the laboratory and then further elongated with one, two and three residues of the α-aminodecanoic acid. These fatty groups were attached to the N-terminal end of the peptide using standard activation methods. The retention times of free and derivatised peptides were then evaluated using a C_{18} reversed-phase column. The difference in retention time between each peptide was approx. 2 min (Fig. 2), thus indicating that the use of these fatty amino acids in the 4-COOH-Fmoc molecule (**1**), for the purification of peptides by reversed-phase chromatography was feasible.

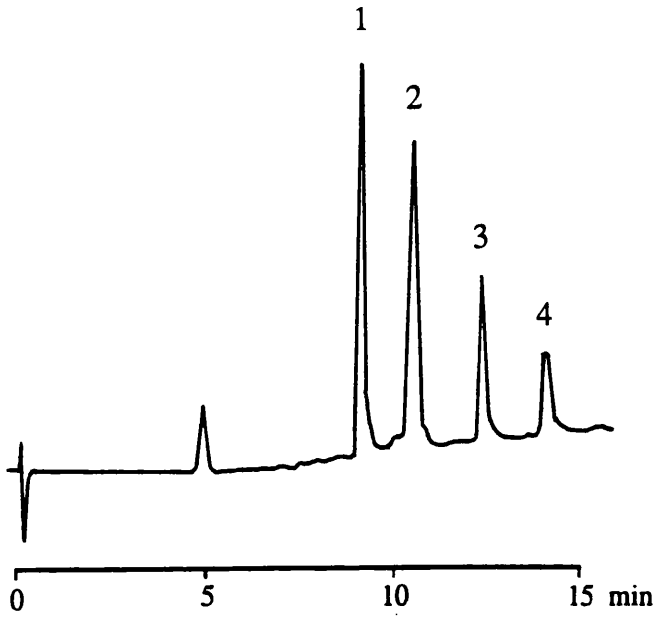
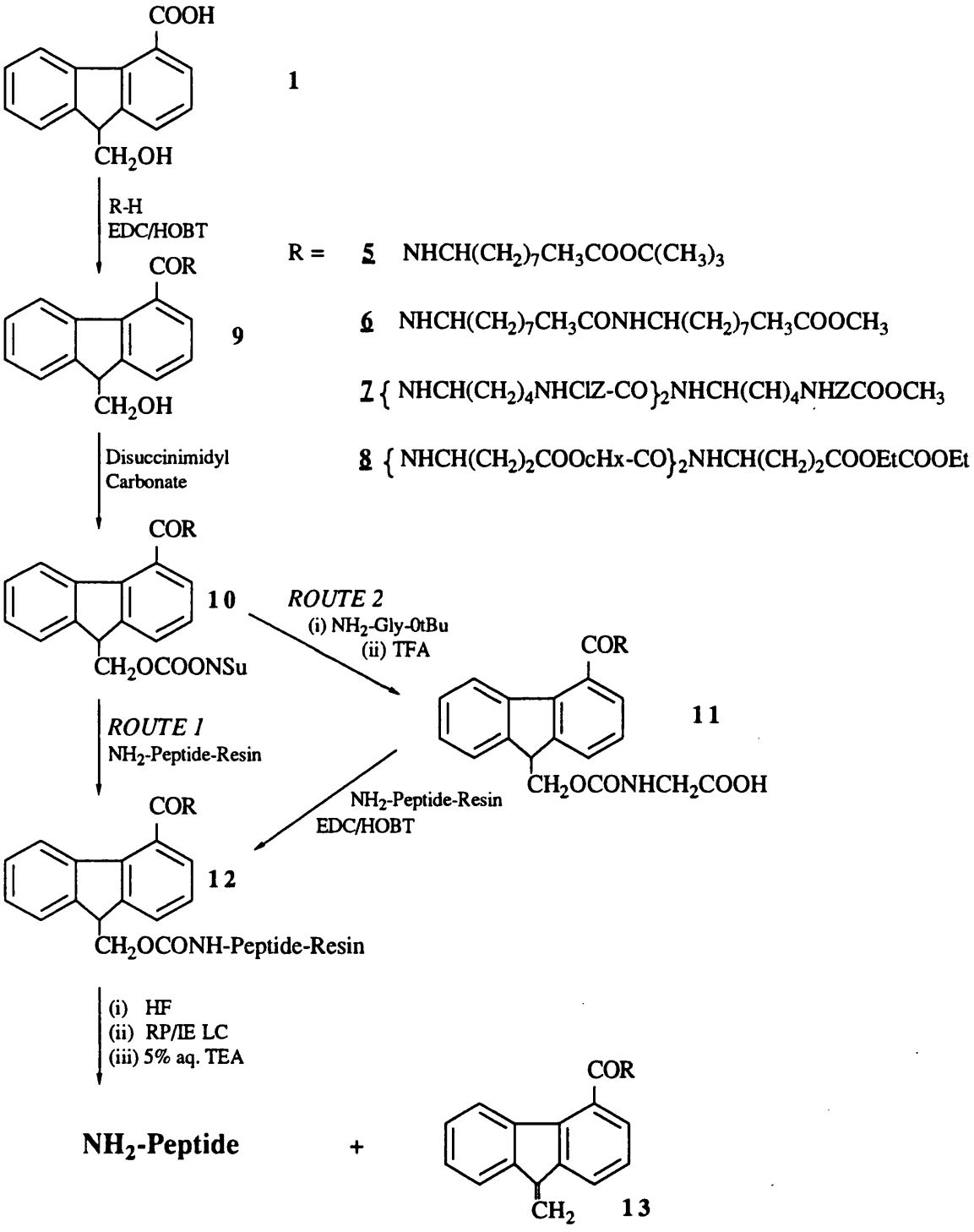


Figure 2. C_{18} RP-HPLC of FMDV Vp1 141-160 (1) derivatised with one (2), two (3) and three (4) α-aminodecanoic acid residues (System 1).

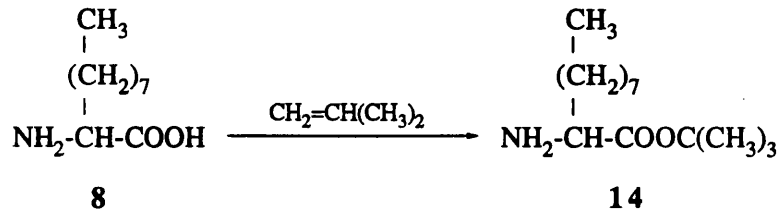
Synthesis of lipophilic 4-COR-Fmoc probes

As expected, by increasing the number of hydrophobic groups and therefore the hydrophobicity of a molecule, the result was an increase in the retention time on reversed-phase media. Hence two derivatives of 4-COOH-Fmoc (**1**) were synthesised with the incorporation of one (D,L) α-aminodecanoic acid group (**9**:R**5**) and two (D,L) α-aminodecanoic acid units (**9**:R**6**), as a dipeptide, as shown in Scheme 3. Racemic fatty amino acid was used since for the purpose of demonstrating that separation of 4-COR-Fmoc derivatised peptide from underivatised peptide could be achieved, the use of optically pure material was not necessary. Consequently, due to the presence of diastereoisomers more peaks were to be expected on the reversed-phase chromatogram.



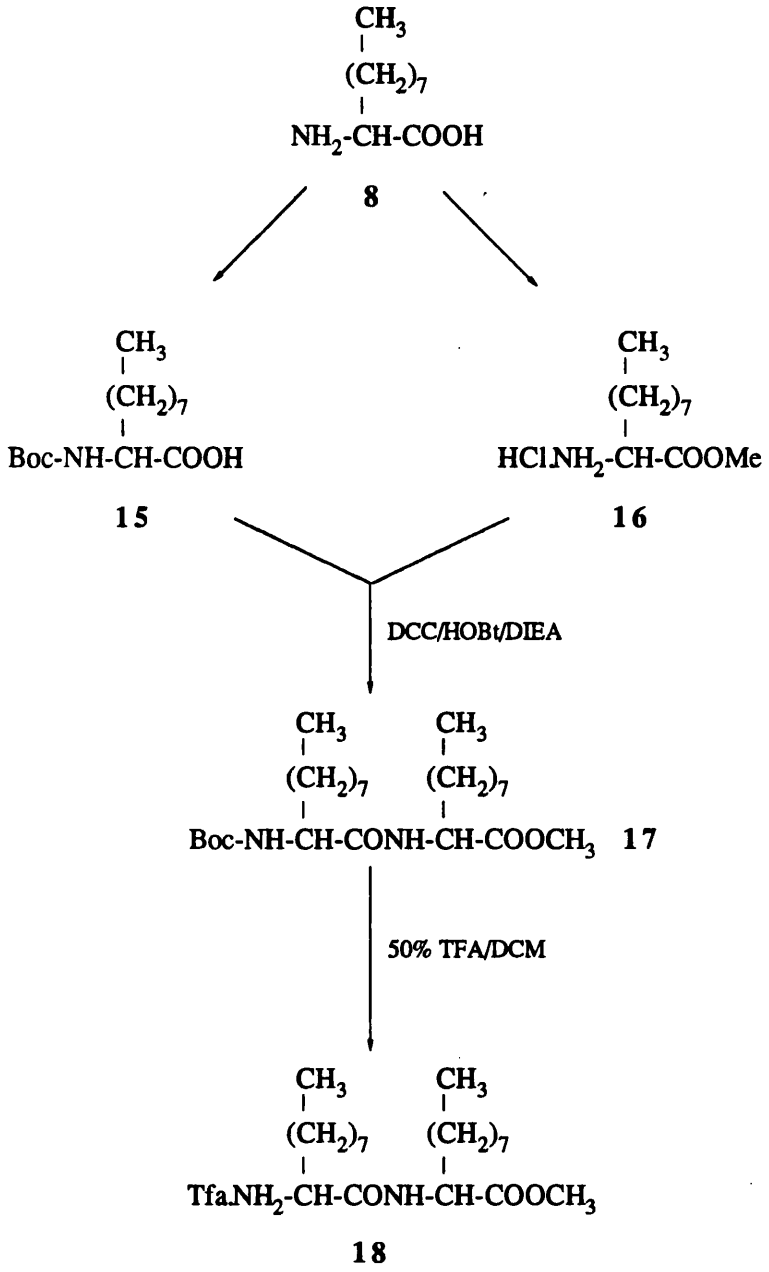
Scheme 3. Outline of the synthesis of chromatographic probes 4-COR(**5**-**8**)-Fmoc, their incorporation onto the peptidyl-resin and removal to yield the purified peptide.

A



8 14

B



Scheme 4. Outline of synthesis of 'fatty' amino acid derivatives

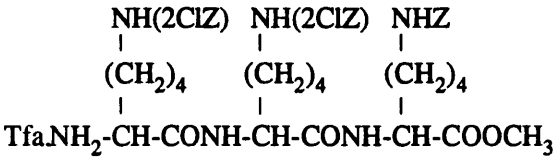
To synthesise the simpliest of the two derivatives 4-COR-Fmoc (**9**:R**5**), it was first necessary to protect the C-terminal group of **8** as a t-butyl ester, using the reaction conditions described for the esterification of fluorene-4-carboxylic acid (Scheme 4A). The coupling of the monomeric (D,L) α-aminodecanoic acid t-butyl ester (**14**) to 4-COOH-Fmoc (**1**) was performed using of 1.1 equivs. of 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and HOBt. Purification was achieved by flash chromatography on silica gel with a yield of 61.5%.

Solution phase peptide synthesis methodologies were used to couple two (D,L) α-aminodecanoic acid molecules together, which could then be linked to 4-COOH-Fmoc (**1**) to form the second hydrophobic probe molecule. However, unlike the mono-fatty acid derivative, the C-terminal group and appropriate N-terminal protection was achieved using the methyl ester and Boc groups, respectively, as shown in Scheme 4B. The dipeptide (D,L) α-aminodecanoic acid methyl ester was then coupled to 4-COOH-Fmoc (**1**), using 1.1 equivs. each of HOBt and EDC, 1 equiv. of DIEA (Scheme 3) and the product (**9**:R**6**) purified by flash chromatography to give a solid 'glassy' material (yield; 68.3%).

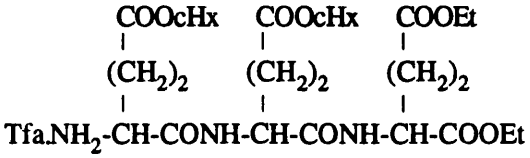
Synthesis of charged 4-COR-Fmoc probes

So far the discussion has focused on chromatographic probes that are designed to be used on reversed-phase media. However, in theory, the 4-carboxyl group on 4-COOH-Fmoc (**1**) offered the opportunity to introduce molecules with different physical properties (eg. ionic groups). To examine the versatility of this purification scheme two 4-COR-Fmoc derivatives were prepared, one possessing positively charged side chains (ie. lysine) and the other negative charges (ie. glutamic acid), for use on ion-exchange chromatographic media. The idea was that upon addition of these 4-COR-Fmoc probes to the target peptide, it would be possible to alter its overall net charge and thus facilitate separation from underivatised impurities.

The ion-exchange probes developed for the extension of the purification system to alternative chromatographic media were based on the addition of the charged amino acids Lys and Glu to the 4-COOH-Fmoc moiety (**1**). As representative molecules of this class, tripeptides of lysine (**19**) and glutamic acid (**20**), each made up of three residues, were synthesised.



19



20

The tripeptides were synthesised by solution phase peptide synthetic methods. The first derivative was the tri-lysine analogue which was protected at the C-terminus as the methyl ester. The side chain protecting groups were carbobenzoxy (CBZ) and 2-chloro CBZ. The first two amino acids were coupled together using 1.1 equivs. each of HOBt and DCC and 1 equiv. of DIEA. Further elongation to the tripeptide was performed in the same fashion and following purification by flash chromatography on silica gel, gave an overall yield of 67.7%.

The tripeptide of glutamic acid was synthesised in the same way as described for the trilysine derivative. The carboxyl terminus was protected as an ethyl ester and side chain protection was accomplished using an ethyl and cyclohexyl molecules. The overall yield after purification by flash chromatography was 43.7%. Since both carboxyl groups on the C-terminal glutamic acid residue were protected as ethyl esters and anhydrous HF is unable to hydrolyse the ester

linkage, the result was that in this example two less negative charges would be present on the 4-COR-Fmoc molecule (**9**:R**8**).

Coupling the trilysine peptide to the 4-carboxyl group of 4-COOH-Fmoc was mediated using 1.1 equivs. HOBt and EDC and 1 equiv. DIEA. Purification of the desired product, 4-COR-Fmoc (**9**:R**7**) was by recrystallisation from hot acetonitrile with a recovery of 70%. The same conditions as previously described yielded the triglutamic acid chromatographic probe 4-COR-Fmoc (**9**:R**8**). To purify the latter molecule flash chromatography was used successfully to obtain 81% of the required product.

Synthesis of 4-COR-Fmoc succinimidyl carbonates

Various methods exist for activating the alcohol group of the Fmoc molecule and thus enable their introduction onto the N-terminus of the peptidyl-resin. These include the synthesis of the corresponding acid chloride using the extremely toxic mustard gas phosgene or its less harmful analogues, di- or tri-phosgene [Eckert and Forster, 1987]. A more satisfactory route to the generation of the base-labile urethane which was reported to proceed in high yield was via the intermediate N-hydroxysuccinimidyl carbonate (4-COR-Fmoc-NSu). In the case of 4-COR-Fmoc (**9**:R**5**) the N-hydroxysuccinimidyl carbonate was prepared following a modified method described by Sauer and Morris [1987]. This involved treating 4-COR-Fmoc (**9**:R**5**) with 2 equivs. of disuccinimidyl carbonate (DSC) and 1 equiv. of pyridine in acetonitrile at r.t. for 7h. The activated compound thus formed was found to be stable in aqueous solvents and was purified by flash chromatography to homogeneity as determined by thin-layer chromatography, ^{1}H-NMR and FAB-MS. The yields of active carbonate (**10**:R**5**) obtained after purification by flash chromatography was 97%. The other three probes were activated in the same way and purified by flash chromatography to yield 67% (**10**:R**6**) and 96% (**10**:R**8**) or recrystallised from hot absolute ethanol, 74% (**10**:R**7**).

Incorporation of 4-COR-Fmoc probes onto peptidyl-resin

For the incorporation of the 4-COR-Fmoc probes onto the peptidyl-resin two alternative routes were investigated; (i) direct attachment of the active 4-COR-Fmoc-NSu carbonate to the free amino group of the resin bound peptide (Route 1; Scheme 3) and (ii) pre-derivatisation with a single amino acid for which Gly was chosen as a representative residue (Route 2; Scheme 3).

Synthesis of 4-COR-Fmoc-Gly intermediates

With reference to the latter approach initial attempts were made to react free glycine with the succininidyl carbonate compound (**10**:R**5**) in the presence of 1 equiv. of sodium carbonate in aqueous solution, according to a method described in the literature [Bodanszky and Bodanszky, 1984]. The reaction was allowed to continue at r.t. for 2h 20 min and then acidified to pH3 with 0.1M HCl, to convert any unreacted glycine to the corresponding hydrochloride. The latter was then removed by partitioning between chloroform, which contained the 4-COR-Fmoc derivative and water. After drying over anhydrous sodium sulphate, the chloroform layer was evaporated down and the residue purified on a HPLC silica agel column, to yield just 21% of the final product (**11**:R**5**). The failure was attributed to poor solubility of **10**:R**5** in the reaction solvents used (ie. acetone and water).

Since free glycine was not soluble in organic solvents it was necessary to use a derivative with a protected carboxyl group (eg. HCl salt of Gly-OtBu). Thus, 1 equiv. of the activated 4-COR-Fmoc-NSu compound (**10**:R**6**) was dissolved in DCM, together with a small quantity of DMF to aid dissolution. To the reaction vessel was added an equimolar quantity of HOBt and 1.5 equivs. each of Cl^- NH_3^+-Gly-OtBu and DIEA, based on the method described by Paquet [1982]. The mildly acidic HOBt was found effective at preventing premature β-elimination of the desired product, which was found to occur in the presence of basic DMF.

The reaction was followed by TLC and found to be complete after 60 min. The reaction mixture was washed with 0.1M HCl and water, then purified by flash chromatography to give the desired product with a yield of 72%. To enable the glycyl derivative to react with the free amino terminus of the peptide, the t-butyl protecting group was removed using 50% TFA in DCM and the cleaved product (**11**:R**6**) passed through a silica gel column for purification (yield 77%). The same procedure was used for the tri-lysine derivative of 4-COR-Fmoc-NSu (**10**:R**7**) to give the corresponding glycine derivative (**11**:R**7**) with a yield of 72%, after purification.

PURIFICATION OF SYNTHETIC PEPTIDES ON THE BASIS OF LIPOPHILICITY

PURIFICATION OF A SEVENTEEN RESIDUE PEPTIDE

Derivatisation of peptidyl-resin

The effects on the retention time caused by the addition of 4-COR-Fmoc probe (**11**:R**5**), were first examined using the FMDV Vp1 141-157 sequence [Bittle *et al*, 1982].

Gly-Ser-Gly-Val-Arg-Gly-Asp-Ser-Gly-Ser-Leu-Ala-Leu-Arg-Val-Ala-Arg-NH_2

The peptide was assembled on an MBHA resin using an automated synthesiser and standard Boc chemistry. No capping procedure was performed at this stage. Quantitative ninhydrin analysis [Sarin *et al*, 1981] indicated an average coupling efficiency of 98.6% (Fig. 3). The Boc group from Ser^{142} was removed, manually, with 50% TFA and the resulting deprotected peptidyl-resin then treated for 90 min with a DMF solution containing 4-COR-Fmoc derivative (**11**:R**5**), HOBt and DCC in equimolar quantities.

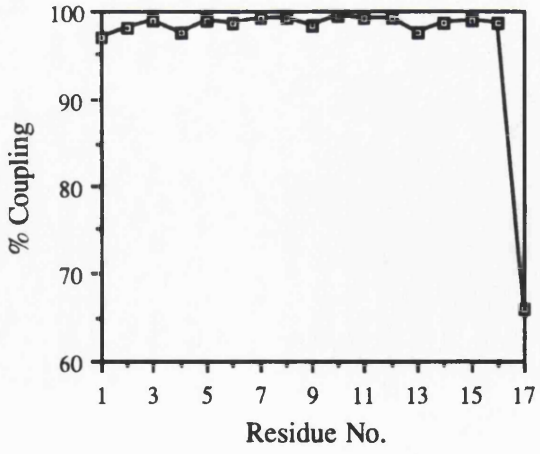
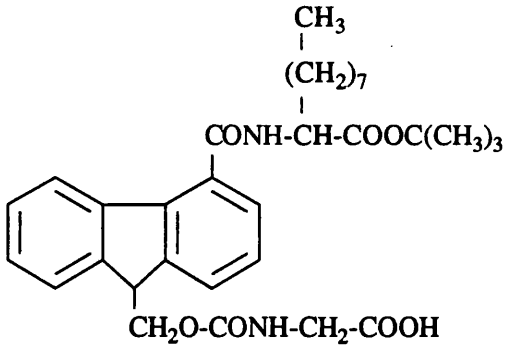


Figure 3. Ninhydrin analysis for synthesis of FMDV 141-157 peptide. (C-terminus = Residue 1)



11:R5

In order to generate a mixture of derivatised and underivatised peptides, which would demonstrate the separation by RP-HPLC, less than 1 equiv. of 4-COR-Fmoc (**11**:R5) was used, thus resulting in the 65% incorporation of residue 17 ie. Gly^{141} (Fig. 3), as determined by ninhydrin analysis. The incorporation of the 4-COR-Fmoc probe, was followed spectrophotometrically by removing an aliquot (3-4mg) of peptidyl-resin at 30 min intervals; this was washed on a sintered glass filter with 50% MeOH/DCM and then dried under vacuum. The peptidyl-resin was accurately weighed, treated with 20% piperidine in DMF for 15 min to release the 4-COR5-Fmoc derivative from the peptide and the absorbance of the supernatant measured at 270nm. Coupling was considered complete (usually 90 min-2h) when two successive readings were identical. After coupling the derivatised peptidyl-resin was washed with DMF, DCM, dried from 50% MeOH/DCM to give 175mg (47μmol) of peptidyl-resin and then subjected to Low TFMSA/High HF treatment [Kent, personal communication]. Although previous work had shown that the urethane bond between the fluorene ring and the peptide chain was stable to strong acidolytic treatment [Merrifield and Bach, 1978] the stability of the entire 4-COR5-Fmoc probe molecule had yet to be demonstrated. Interestingly, the Fmoc group (**2**) which had been attached to the N-terminus of a peptide synthesised using Boc chemistry [Grode *et al*, 1992] was subject to electrophilic aromatic substitution. The impurities were identified and were shown to arise from the migration of amino acid protecting groups. The two cleavage procedures that constitute the LowTFMSA/HighHF procedure were applied as described in the experimental

section and the crude peptide product precipitated with dry diethyl ether. Following filtration the crude product was dissolved in 5% acetic acid solution and lyophilised to give a yield of 93mg (94%) crude material.

Purification of 4-COR-Fmoc derivatised peptide

The crude derivatised peptide was readily soluble in water and injected on C_{18} reversed-phase media. The chromatogram obtained showed peaks that eluted in two distinct regions along the acetonitrile gradient (Fig. 4).

To show that these corresponded to derivatised and underivatised peptides, 30mg (14 μmol) of crude material was purified by semipreparative C_{18} reversed-phase column. The earlier eluting peak labelled 1 (Fig. 4) was collected (6.5mg; 22%) and shown by FAB-MS ($[M+H]^+$: expected 1583; found 1584) to be the underivatised FMDV 142-157 peptide fragment. Thus the cluster of smaller peaks on either side of the main peak labelled 1 probably represented truncated and underivatised deletion sequences. The second major product which appeared to be represented by two closely eluting peaks labelled 2 (Fig. 4) were then collected (7.5mg; 25%). Due to their increased retention time on the reversed-phase media (indicating a product with higher lipophilicity) they were expected to represent the racemic 4-COR**5**-Fmoc derivatised peptides. Thus FAB-MS ($[M+H]^+$: expected 2092; found 2093) and amino acid analysis (Table 1) confirmed that this material was indeed the desired 4-COR**5**-Fmoc derivatised product. The 4-COR**5**-Fmoc derivatised peptide generated two peaks (labelled 2; Fig. 4) due to the use of racemic (D,L) α-aminodecanoic acid whose diastereoisomers were separated on the reversed-phase chromatographic media. Furthermore, the result obtained confirmed, as envisaged, that the HF cleavage conditions to which the 4-COR-Fmoc derivatised peptidyl-resin were subjected to were not detrimental to the chromatographic probe.

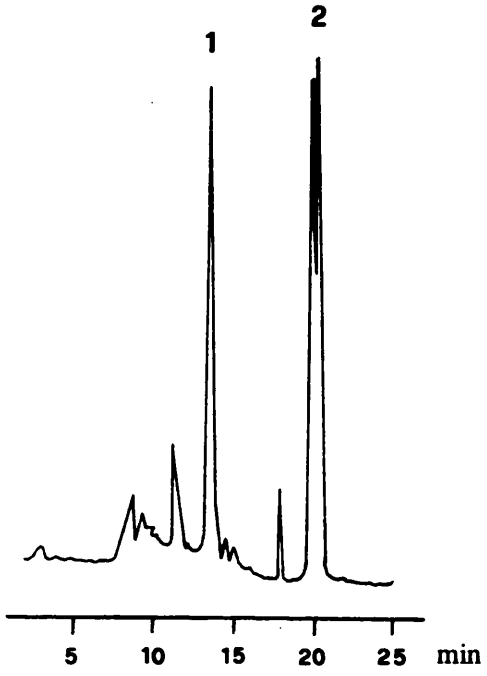


Figure 4. C_{18} RP-HPLC of crude HF cleavage material containing 4-COR5-Fmoc derivatised FMDV 141-157 peptide. Peak (1) is the underivatised peptide and the two peaks labelled (2) are the diastereoisomeric forms of the 4-COR5-Fmoc derivatised peptide (System 2).

The relatively short gradient (0-100% acetonitrile in 30 min) was capable of achieving a separation of 6 min between the two major products. The cluster of smaller peaks that eluted for several minutes both before, after and possibly directly beneath the underivatised peptide with the correct sequence (labelled 1; Fig. 4) emphasised the difficulties that can exist during conventional purification procedures. Thus, the increase in retention time by 6 min which resulted from the addition of the hydrophobic 4-COR-Fmoc probe (**11**:R5) was sufficient to pull the attached peptide away from the contaminating peptidic impurities thereby avoiding the collection of co-eluting impurities. The significance of these results were that, like for the FMDV peptide substituted with one, two and three α-aminodecanoic acid groups as described above, the addition of this

hydrophobic probe drastically changed the chromatographic characteristics of the derivatised peptide, with the difference that the lipophilic group could be removed in the case of the 4-COR$\underline{5}$-Fmoc treated material to yield the purified free peptide.

Cleavage of 4-COR-Fmoc derivative from peptide

Since the group that linked the chromatographic probe to the peptide was a urethane moiety, the same base catalysed β-elimination reaction used for the fluorenylmethyloxycarbonyl (Fmoc) SPPS strategy, could also be used to free the purified peptide.

Studies into the sensitivity of the urethane group to various organic bases in different solvents have been performed. These have included (i) 50% piperidine in DCM for 30 min [Chang and Meienhofer, 1978] (ii) 20% piperidine in DMF for 10 min [Atherton *et al*, 1978a; Atherton *et al*, 1978b], (iii) 10% morpholine in DCM or DMF for 20 min [Merrifield and Bach, 1978]. In general the use of DMF with its superior resin-swelling properties offered better cleavage kinetics than the less polar solvents (eg. DCM), when combined with 10-20% piperidine.

All the cleavage mixtures that have been mentioned above rely on the use of organic solvents (eg. DMF or DCM) which are ideal when side chain protected resin-bound peptides are being treated. However, the treatment of free derivatised peptide with organic solvents would not be desirable, particularly if the product was biologically active in addition to the poor solubility of larger peptide sequences. Consequently, aqueous based systems were investigated. Merrifield and Bach [1978] suggested the use of dilute aqueous solutions of TEA, sodium carbonate or sodium hydroxide. Various attempts were made on a small aliquots of purified 4-COR$\underline{5}$-Fmoc derivatised peptide, which indicated that a 30 min treatment with 5% aqueous TEA was sufficient to achieve quantitative cleavage of the chromatographic probe. The optimised procedure was then performed on a larger scale.

The cleavage mixture containing the free FMDV 141-157 peptide was injected on a C_{18} RP-HPLC column to give the chromatogram shown in figure 5. The peak labelled 2 was collected and shown by FAB-MS ($[M+H]^+$: expected 1657; found 1656) and amino acid analysis (Table 1) to correspond to the free FMDV 141-157 sequence. The base (TEA) eluted with the solvent front which facilitated its removal from the product.

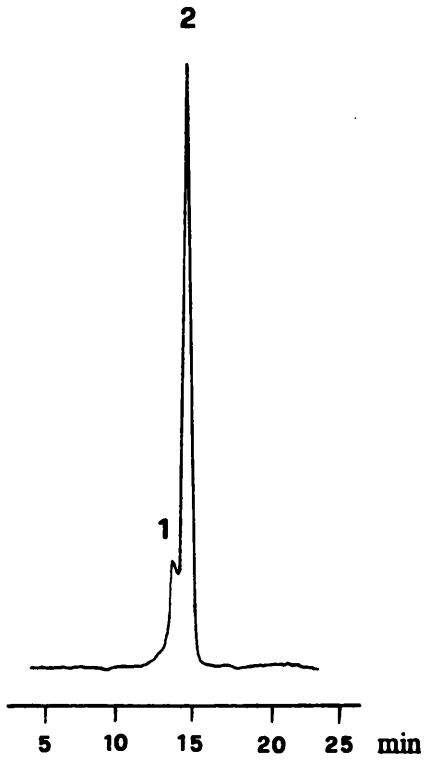


Figure 5. Free peptide (2) after treatment with 5% TEA. Peak (1) represents the deletion sequence (System 2).

Residue	FMDV Vp1 142-157 + 4-COR5 -Fmoc-Gly		FREE PEPTIDE	HIV-1 gp120 fragment + 4-COR6 -Fmoc-Gly		FREE PEPTIDE	HIV-1 gp120 fragment + 4-COR7 -Fmoc-Gly		FREE PEPTIDE	cyclic FMDV fragment + 4-COR8 -Fmoc		FREE PEPTIDE
Asp/Asn	1.3	(1)	1.1	2.9	(3)	3.6	3.2	(3)	3.1	1.1	(1)	1.0
Thr	--		--	0.8	(1)	1.2	0.9	(1)	1.0	--		--
Ser	3.0	(3)	2.9	1.0	(1)	1.1	1.0	(1)	1.0	1.9	(2)	2.0
Glu/Gln	--		--	4.5	(5)	4.9	4.6	(5)	4.7	4.3[x]	(1)	1.2
Pro	--		--	3.3	(4)	3.9	3.7	(4)	3.8	2.2	(2)	1.9
Gly	4.3	(4)	3.8	4.5	(5)[x]	4.6	4.7	(5)[x]	4.8	4.6	(4)	3.9
Ala	2.0	(2)	2.0	--		--	--		--	2.0	(2)	2.0
Val	1.9	(2)	1.7	--		--	--		--	1.7	(2)	2.4
Ile	--		--	0.8	(1)	1.0	0.9	(1)	0.9	--		--
Leu	2.1	(2)	2.1	1.0	(1)	1.0	1.0	(1)	1.0	2.1	(2)	2.0
Phe	--		--	--		--	--		--	1.0	(1)	0.9
Lys	--		--	--		--	2.4[x]		--	--		--
Arg	3.3	(3)	3.4	4.9	(5)	5.5	5.5	(5)	5.2	2.9	(3)	2.8

[x] Additional amino acids from 4-COR-Fmoc probes.

Table 1. Amino acid analysis of purified 4-COR-Fmoc derivatised peptides and free peptides.

The important observation was that the increased retention time induced by the hydrophobic chromatographic probe was lost upon its removal from the peptide (ie. retention time reduced from 20 min to 14 min). An integration of the peak which corresponded to the free FMDV peptide indicated that a purity of 95% had been achieved. An impurity (labelled 1; Fig. 5) which made up the remaining 5% was collected and shown by FAB-MS to correspond to a deletion sequence consisting of the target peptide, less Ser. The explanation as to why the impurity was also derivatised with 4-COR-Fmoc probe (**11**:R**5**) lay in the fact that no capping procedure was performed during the synthesis. Consequently, any free amino terminii, even those associated with deletion peptides, would also be derivatised, as was demonstrated.

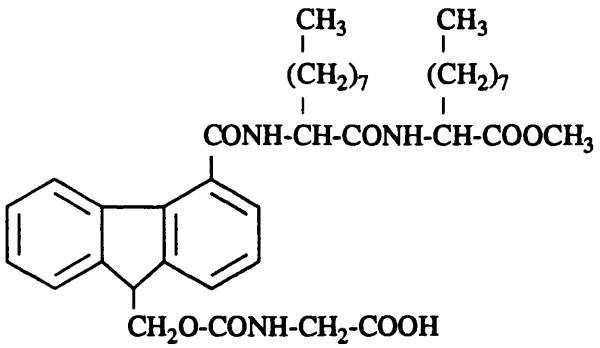
Capping procedures

Since the coupling of an amino acid is rarely quantitative it is desirable to prevent the residual free amino groups from further reacting and thereby forming deletion sequences. Thus, "capping" involves the formation of stable adducts, by reacting these free amino groups with acylating agents, to replace a potential family of deletion peptides with one of terminated sequences. During the course of these investigations different 'capping' procedures were assessed on the automated peptide synthesiser. The first protocol involved reacting the peptidyl-resin with DCM solutions containing 20% acetic anhydride and 2.5% DIEA, for 6 min. The success of the capping reaction to remove unreacted free amino groups was determined by performing a quantitative ninhydrin test. Although the method described above gave the expected results (ie. 100% acylation of free amino groups) further attempts were made to reduce the time involved in the reaction. Due to its superior resin-swelling properties, DMF was used to replace DCM as the solvent for acylation. The final capping protocol adopted involved a prewash with 15% DIEA in DMF, followed by a 2.5 min treatment with 18% acetic anhydride and 10% DIEA. The run cycles on the

automated synthesizer were modified accordingly to deliver the required reagents for capping, after the addition of each amino acid.

PURIFICATION OF A TWENTY-FIVE RESIDUE PEPTIDE

The successful derivatisation, cleavage and purification of the model 17 residue FMDV peptide fragment using the lipophilic reversible chromatographic probe (**11**:R**5**) now led to the examination of the more hydrophobic 4-COR-Fmoc-Gly molecule (**11**:R**6**) for application to longer peptide sequences.



11:R**6**

Furthermore, from the experience acquired with the 17 residue FMDV fragment, it was necessary to include a capping procedure after each coupling step to avoid derivatising deletion sequences. Thus, a twenty-five residue peptide was synthesized representing the HIV-1 gp120 loop sequence [Robson *et al*, 1986], using Boc chemistry.

Leu-Pro-Thr-Pro-Arg-Gly-Pro-Asp-Arg-Pro-Glu-Gly-Ile-
Glu-Glu-Glu-Gly-Gly-Glu-Arg-Asp-Arg-Asp-Arg-Ser-amide

The peptide was constructed on an MBHA resin and the capping procedure described above was applied during the stepwise synthesis.

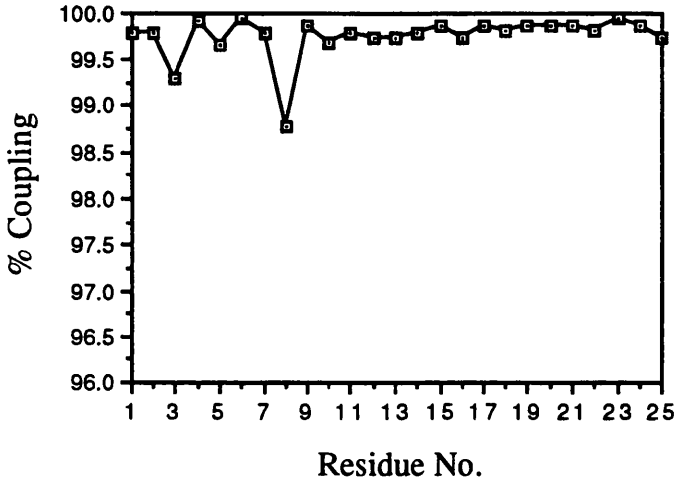


Figure 6. Ninhydrin analysis for synthesis of HIV-1 gp120 loop sequence. (C-terminus = Residue 1)

After completion of the synthesis and Boc removal, the free amino terminal of the peptidyl-resin was reacted with 2 equivs. of 4-COR-Fmoc-Gly derivative (**11**:R**6**) as the HOBt ester of the pre-formed glycine compound. UV monitoring of the coupling efficiency, as measured by the release of the 4-COR-Fmoc-Gly group into solution, indicated that the reaction was complete in 90 min.

150mg (25.5μmol) 4-COR**6**-Fmoc-Gly derivatised peptidyl-resin was then cleaved using the High HF procedure, the crude peptide precipitated with diethyl ether, dissolved in 5% acetic acid and finally lyophilised to a fluffy white powder with a yield of 72 mg (81%) crude product. A small quantity of crude cleavage product was injected onto an analytical C_8 reversed-phase column to produce the chromatogram shown in Figure 7A. The material that eluted, appeared in two distinct regions of the chromatogram, a cluster of low intensity peaks between 22 and 40 min (labelled 1, Fig. 7A) followed by two large peaks at 46 min (labelled 2, Fig. 7A). The profile was similiar to that observed for the

4-COR5-Fmoc derivatised FMDV peptide, where the two later eluting peaks were found to be the diastereomeric forms of the chromatographic probe.

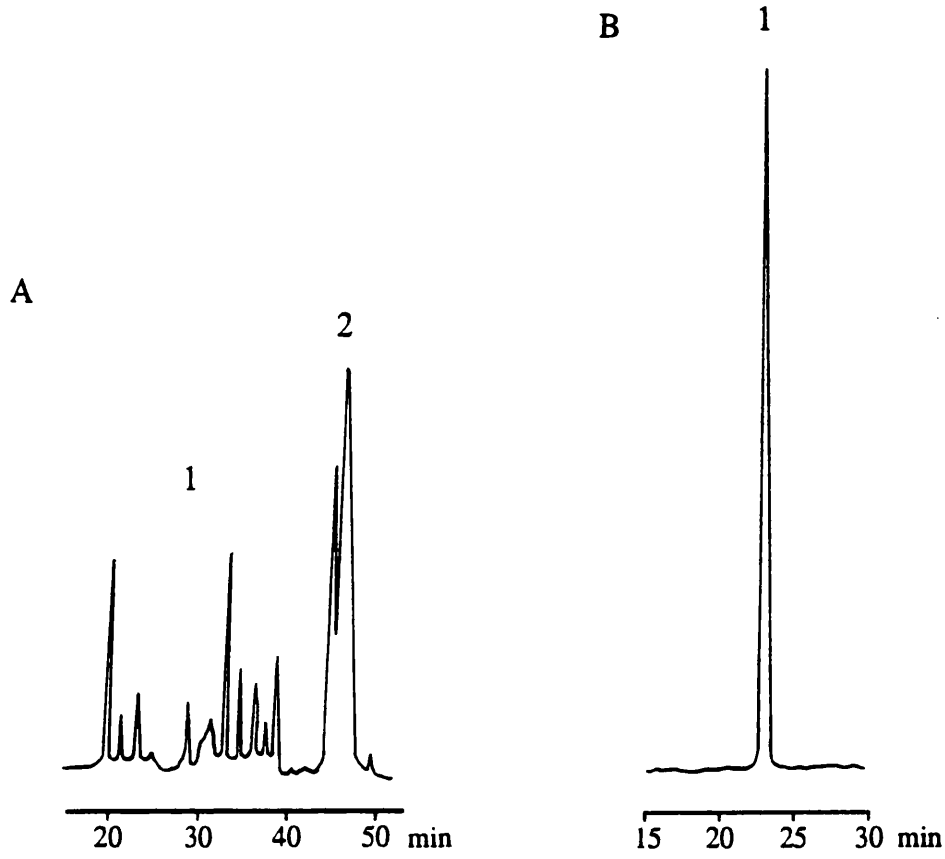


Figure 7. C_8 RP-HPLC of HIV ENV gp120 fragment (System 3).
A) Crude 4-COR6-Fmoc-Gly derivatised peptide (2) after HF cleavage.
B) Free peptide (1) after treatment with 5% aqueous TEA.

The two peaks which eluted at 46 min were collected and characterised by FAB-MS ($[M+H]^+$: expected 3476; found 3476) and amino acid analysis (Table1) to confirm that the diastereomeric forms of the 4-COR6-Fmoc-Gly derivatised HIV-ENV peptide had been isolated. As further corroborating evidence, when the purified material was treated with 5% aqueous TEA the two peaks that originally eluted at 46 min became one peak at 23 min, with no trace of unwanted chains, in accordance with the loss of the hydrophobic chromatographic probe (Fig. 7B). The single peak was collected and analysis by

FAB-MS ($[M+H]^+$: expected 2877; found 2878) and amino acid analysis (Table 1) confirmed that the target peptide had been purified. Thus the low intensity peaks in the crude cleavage mixture, which appeared between 22 and 26 min were probably terminated sequences.

As demonstrated with the 17 residue FMDV peptide, the hydrophobic probe (**11**:R**5**) was sufficiently lipophilic to retain the attached peptide long enough for all underivatised material to elute off the reversed-phase column and therefore facilitate purification of the target peptide. Furthermore, the acetylation of unreacted amino groups after each step of solid phase synthesis prevented the unwanted derivatisation of deletion chains, as was indicated by the presence of a single peak for the free peptide after treatment with base.

PURIFICATION OF FORTY-SEVEN RESIDUE PEPTIDE

To determine the potential of the purification system the next application was to a 47 residue sequence. The peptide representing the 39-86 fragment of the HIV-1 TAT protein [Sodrosky *et al*, 1985] was synthesized on a pre-derivatised PAM resin using Boc SPPS protocols [Kent and Clark-Lewis, 1985] and a capping procedure.

Thr-Lys-Ala-Leu-Gly-Ile-Ser-Thr-Gly-Arg-Lys-Lys-Arg-Arg-Gln-Arg-
Arg-Arg-Pro-Pro-Gln-Gly-Ser-Gln-Thr-His-Gln-Val-Ser-Leu-Ser-Lys-
Gln-Pro-Thr-Ser-Gln-Pro-Arg-Gly-Asp-Pro-Thr-Gly-Pro-Lys-Glu-COOH

The peptidyl-resin had been stored in DMF at -70°C for 12 months prior to N-terminal deprotection and treatment with the 4-COR-Fmoc-Gly (**11**:R**6**). Consequently, the possibility existed that the N-terminus could have undergone rearrangements [Kent and Clark-Lewis, 1985] and therefore give rise to a proportion of terminated impurities.

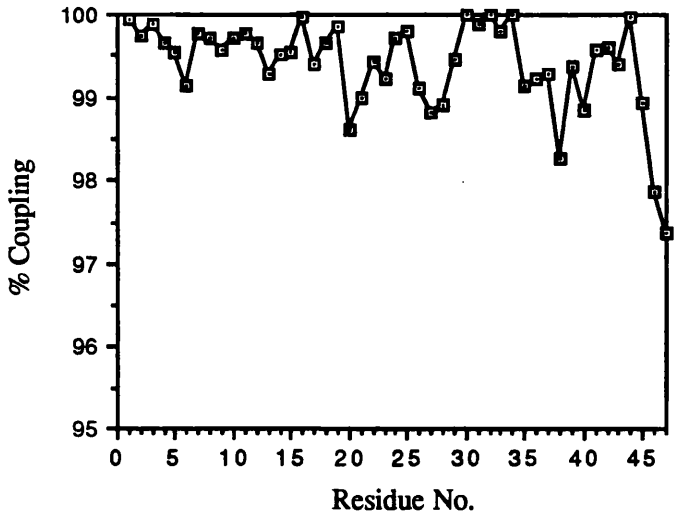


Figure 8. Ninhydrin analysis for synthesis of HIV-1 TAT 39-86 fragment. (C-terminus = Residue 1)

The Boc group was removed with 50% TFA in DCM and derivatised with 4-COR6-Fmoc-Gly (**11**:R6), as the preformed glycine compound, in four-fold excess (Scheme 3). Coupling was complete after 60 min as determined by the UV monitoring technique. Normally, the protecting groups His(DNP) and Trp(CHO) are removed prior to HF cleavage. However, in this case the His residue was left protected since the basic conditions required for the thiolysis of the His-DNP bond (ie. β-mercaptoethanol/DIEA in DMF) would have prematurely cleaved the urethane bond linking the chromatographic probe to the peptidyl-resin. 71mg (7μmol) of 4-COR6-Fmoc-Gly derivatised TAT 39-86 peptidyl-resin was cleaved using the lowTFMSA/highHF procedure. The initial low TFMSA part involved the addition of the scavengers, 1,4-butanedithiol, *p*-cresol and dimethyl sulphide in TFA to the peptidyl-resin, followed by the addition of TFMSA dropwise at 0°C. After 2h at 0°C the partially deprotected peptidyl-resin was filtered and washed with diethyl ether. The high HF part of the reaction required the use of the scavengers 1,4-butanedithiol and *p*-cresol, followed by the addition of anhydrous HF, at 0°C. The reaction was continued

for 1 h. After distilling off the HF and precipitating the cleaved peptide with diethyl ether the crude material was dissolved in 10% acetic acid and lyophilised to give 42.3mg (95%) of crude product. An SDS-PAGE (Fig. 9) of the crude cleavage product showed a band with a molecular weight approx. 9Kd.

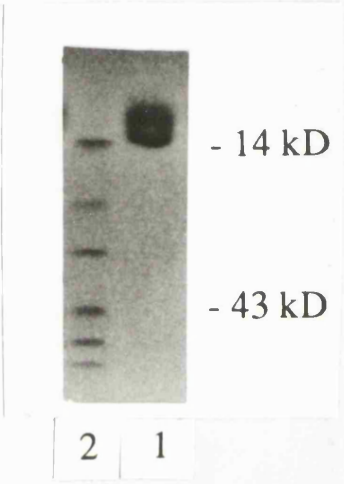


Figure 9. SDS-PAGE of crude 4-COR6-Fmoc-Gly derivatised TAT 39-86 peptide (lane 1) and molecular weight markers (lane 2).

The lack of small molecular weight bands confirmed that no degradation of the treated polypeptide had occured during acidolytic cleavage. The discrepancy between the predicted molecular weight of about 6 KDa and the observed migration corresponding to 9 KDa, on the SDS-PAGE gel has been reported [Sodrosky *et al*, 1985] and was probably caused by the highly basic nature of the peptide. Chromatographic analysis was then performed with the column equilibriated at 37°C. The crude cleavage product was injected onto a C_4 reversed-phase media (Fig. 10) to give (i) a broad peak (**1**) at approx. 19 min., which resembled that of the underivatised TAT fragment and (ii) a second less intense peak (**2**) with a longer retention time (ie. 30 min.) and probably

represented the desired 4-COR6-Fmoc-Gly derivatised product. The eluting peaks were collected and reinjected on an analytical C_4 reversed-phase column and monitored at 220nm (Fig. 11).

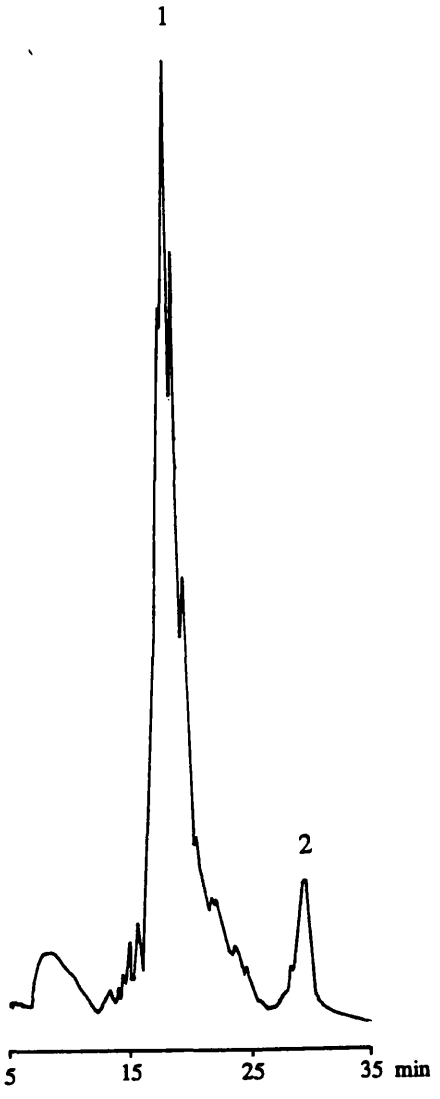


Figure 10. C_4 RP-HPLC of crude 4-COR6-Fmoc-Gly derivatised 39-86 TAT fragment (1). (System 4)

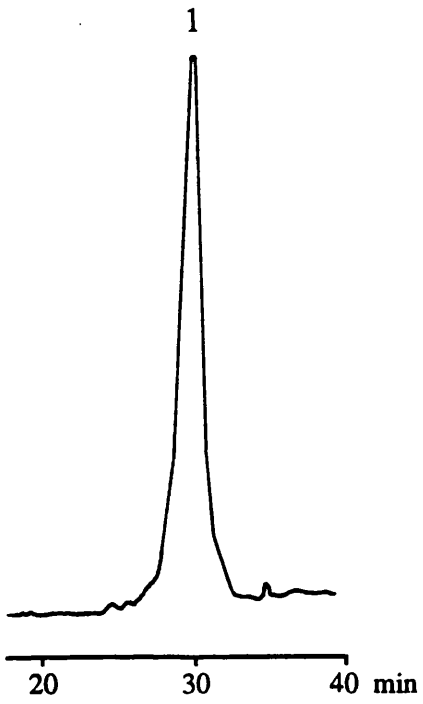


Figure 11. C_4 RP-HPLC of purified 4-COR6-Fmoc-Gly derivatised 39-86 TAT fragment (1). (System 4)

To show that the later eluting peak (labelled 2; Fig. 10) was the 4-COR6-Fmoc-Gly derivatised TAT fragment, an aliquot of purified material was treated with 5% aqueous TEA. The cleavage solution was reinjected on the reversed-phase media and gave a single peak (labelled 1; Fig. 12A) at 21 min. The observed retention time was in close accordance with that recorded for the supposed underivatised material (labelled 1; Fig. 10). To provide confirmation that the polypeptide purified using the chromatographic probe 4-COR6-Fmoc-Gly was correct, a sample of untreated 39-86 TAT fragment that had been purified by convention chromatographic means was injected onto the C_4 reversed-phase column. The elution time of 21 min (Fig. 12B) corresponded to that obtained using the 4-COR6-Fmoc-Gly purification system. Furthermore, SDS-PAGE analysis (Fig. 13) of the free Gly-39-86 peptide following base treatment gave a band that was approx 9kD, confirming that the small later eluting material was the desired product. Interestingly comparison with the

crude product (Lane 2, Fig. 13) showed a tighter band, thus indicating that the family of closely related peptides (truncated and deletion peptides) had probably been removed.

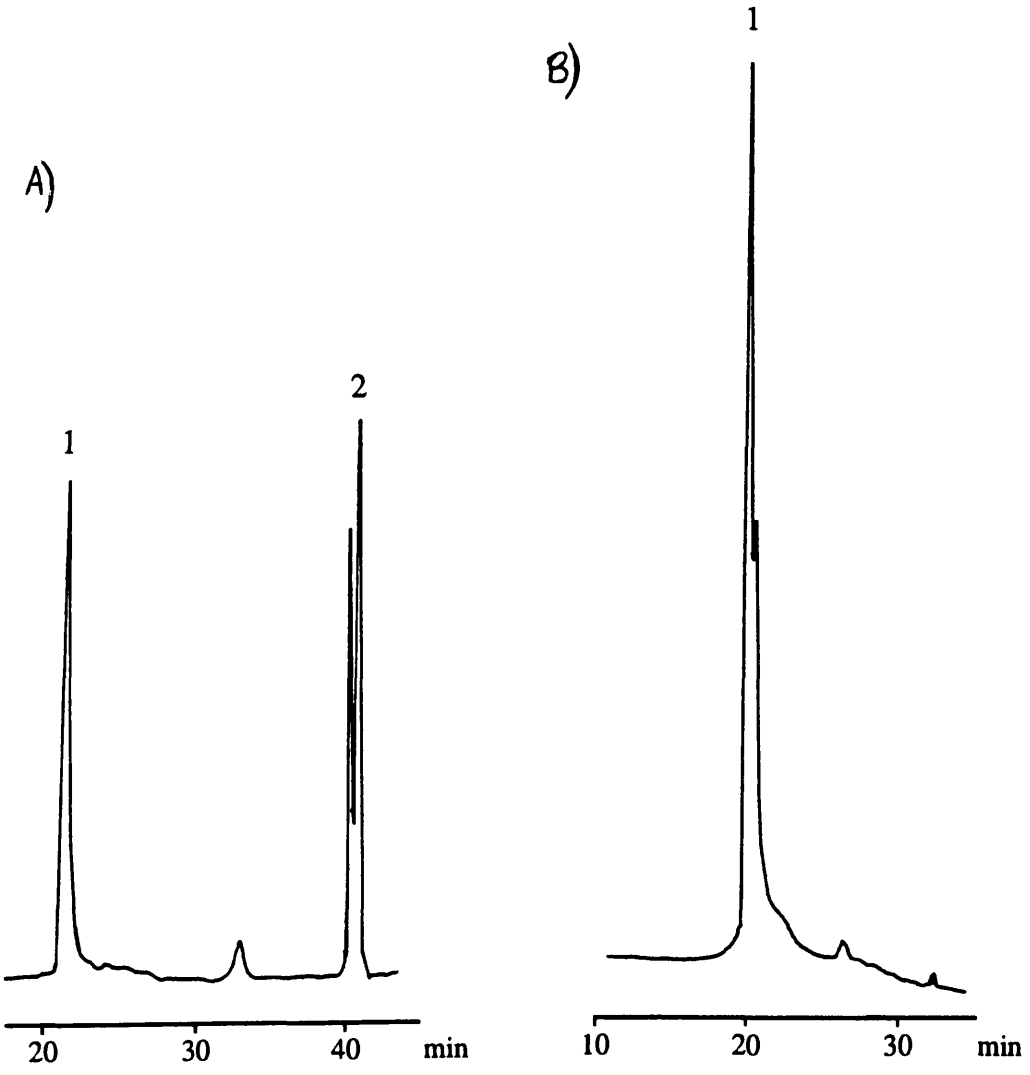


Figure 12. C_4 RP-HPLC of 39-86 TAT fragment. (System 4)
A) 4-COR6-Fmoc-Gly derivatised peptide following treatment with 5% aqueous TEA. Peak 1 represents underivatised peptide. Peak 2 = 13(R=6)
B) Reference sample of purified 39-86 TAT peptide.

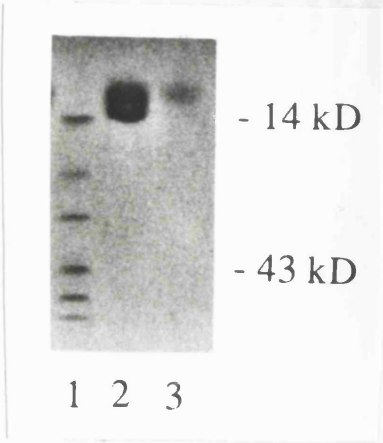


Figure 13. SDS-PAGE of 39-86 TAT fragment.
Lane 1 = protein markers (14-66kD).
Lane 2 = Crude 4-COR6-Fmoc-Gly derivatised polypeptide.
Lane 3 = Purified 4-COR6-Fmoc-Gly derivatised polypeptide.

The previous experiments served to demonstrate (i) that it was possible to derivatise larger peptides, (ii) that separation of derivatised from underivatised material was achievable and (iii) that the purified polypeptide was cleaner than crude product as judged by SDS gel electrophoresis. Unfortunately, the yields were very poor, due to either scarce reactivity of the bulky 4-COR6-Fmoc-Gly group or availability of the peptidyl-resin N-terminus for reaction after several months in storage. The former explanation was probably unlikely since probe (**11**:R6) had reacted well with the HIV-ENV gp120 peptide fragment.

PURIFICATION OF A ONE HUNDRED AND FOUR RESIDUE PROTEIN

The ultimate aim of this piece of research was to establish a method of purification that offered a simple, one-step approach irrespective of the amino acid sequence or length. In order to determine the full potential of the 4-COR-Fmoc-Gly derivatives with regards to the latter property, a peptidyl-resin comprising 104 residues, fragment 270-373 of the HIV-1 p24 GAG protein [Peptide sequence reference; Mascagni *et al*, 1990], was derivatised with 4-COR-

Fmoc-Gly probe (**11**:R**6**). The polypeptide had been assembled on a PAM resin, Boc chemistry [Kent and Clark-Lewis, 1985], without a capping procedure.

Leu-Asn-Lys-Ile-Val-Arg-Met-Tyr-Ser-Pro-Thr-Ser-Ile-Leu-Asp-Ile-Arg-Gln-GlyPro-Lys-Glu-Pro-Phe-Arg-Asp-Tyr-Val-Asp-Arg-Phe-Tyr-Lys-Thr-Leu-Arg-Ala-Glu-Gln-Ala-Ser-Gln-Glu-Val-Lys-Asn-Trp-Met-Thr-Glu-Thr-Leu-Leu-Val-Gln-Asn-Ala-Asn-Pro-Asp-Ala-Lys-Thr-Ile-Leu-Lys-Ala-Leu-Gly-Pro-Ala-Ala-Thr-Leu-Glu-Glu-Met-Met-Thr-Ala-Ala-Gln-Gly-Val-Gly-Gly-Pro-Gly-His-Lys-Ala-Arg-Val-Leu-Ala-Glu-Ala-Met-Ser-Gln-Val-Thr-Asn-Ser-COOH

In order to observe the difference in retention time generated by the addition of the chromatographic probe (**11**:R**6**) onto the 104 residue protein, it was necessary to have a proportion of underivatised material. Since the protein-resin had been prepared many months in advance of this experiment, it was believed that a sufficiently high number of truncated chains existed for the required comparison to be made (Kent and Clark-Lewis, 1985).

To attach the 4-COR**6**-Fmoc-Gly probe, 300mg (18μmol) of the 104-mer peptidyl-resin was allowed to swell in DMF for several hours, washed well with DCM and then treated with 50% TFA in DCM to remove the Boc protecting group. The deprotected resin-bound 104-mer was treated with a DMF/DCM solution containing an estimated two-fold excess of 4-COR**6**-Fmoc-Gly, HOBt and EDC (1:1:1). After 90 min reaction time, no further addition of 4-COR-Fmoc-Gly derivative (**11**:R**6**) was observed as determined by UV and ninhydrin analysis. The peptidyl-resin was washed and dried from MeOH/DCM (1:1), in preparation for the LowTFMSA/HighHF reaction. As described for the 4-COR**6**-Fmoc-Gly derivatised TAT fragment, His(DNP) and Trp(For) residues were not deprotected prior to acidolytic cleavage because the basic conditions required would also have removed the chromatographic probe. It was envisaged that cleavage of these groups could be performed after reversed-phase

purification, although this would not be the best approach due to the possible instability of the protein to organic solvent. It would be more suitable to use protecting groups that were acid labile (eg. His(Cbz)) or were more acid sensitive than Trp(For). The acidolytic treatment of the derivatised was conducted as previously described for FMDV Vp1 141-157 and TAT peptides, to be precipitated by diethyl ether, redissolved in 5% acetic acid and finally lyophilised to 198mg (87%) of a pale yellow powder.

Previous discussion for the application for the reversed-phase purification of 17 to 49 residue peptides has demonstrated that the 4-COR6-Fmoc-Gly derivatised material could be separated from underivatised peptide by up to 26 min. Thus in the case of the 104 residue protein it was necessary to observe a peak or group of peaks with a greater retention time than the corresponding underivatised analogues. From the RP-HPLC of the crude polypeptide on a C_4 column two broad peaks labelled 1 and 2 (Fig. 14) where obtained with retention times of 49 and 53 min, respectively.

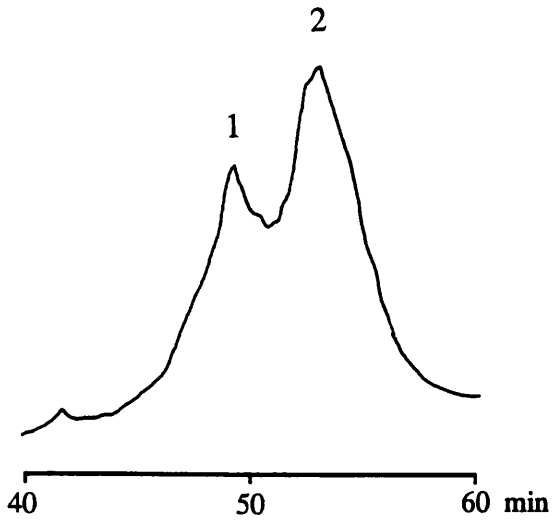


Figure 14. C_4 RP-HPLC of crude 4-COR6-Fmoc derivatised 104-mer. (System 5)

The difference in retention time of 4 min could be increased to 7 min between the two main peaks (labelled 1 and 2) using a C_8 HPLC column, under the same eluting conditions (Fig. 15A). On C_{18} reversed-phase medium the time separating the two principle peaks (labelled 1 and 2) was also 7 min but gave almost baseline resolution of the eluting material (Fig. 15B). It was then necessary to determine which species were represented by the two observed groups of peaks. As a reference point 170mg (9μmol) of untreated 104 residue p24 GAG fragment was cleaved using the same conditions (Low TFMSA/High HF), with the exception that the His and Trp were deprotected, as used for the derivatised protein to yield after lyophilisation 112mg (99%) crude protein. The crude underivatised material was loaded onto the C_{18} reversed-phase column and eluted as a broad peak at 48 min and therefore similar characteristics to peak labelled 1 (Fig 15B). The two peaks labelled 1 and 2 (Fig. 15B) were then isolated and the later eluting material reinjected on C_{18} reversed-phase column (Fig. 15C).

To prove that the second peak (labelled 2, Fig. 15B) was 4-COR6-Fmoc-Gly derivatised 104-mer, an aliquot of the later eluting peak, after purification (Fig. 15C), was treated with dilute aqueous TEA. After 15 min, a second peak appeared on the RP-HPLC chromatogram with the same retention time as the underivatised polypeptide. An additional 15 min reaction resulted in complete conversion of the peak at 55 min to the peak at 48 min, which was consistant with the loss of the hydrophobic probe 4-COR6-Fmoc (Fig. 15D). Further evidence in support of the later eluting peak being 4-COR6-Fmoc-Gly derivatised 104-mer was provided by SDS-PAGE electrophoresis (Fig. 16), which gave a band corresponding to the expected average molecular weight (ie. approx. 11kD). Amino acid composition as determined by analysis (Table 2) of the polypeptide represented by peak labelled 2 (Fig. 15B) indicated that this was also 104-mer. Thus it was concluded that the second peak corresponded to the 4-COR6-Fmoc-Gly derivative of the 104 residue p24 GAG 270-373 fragment.

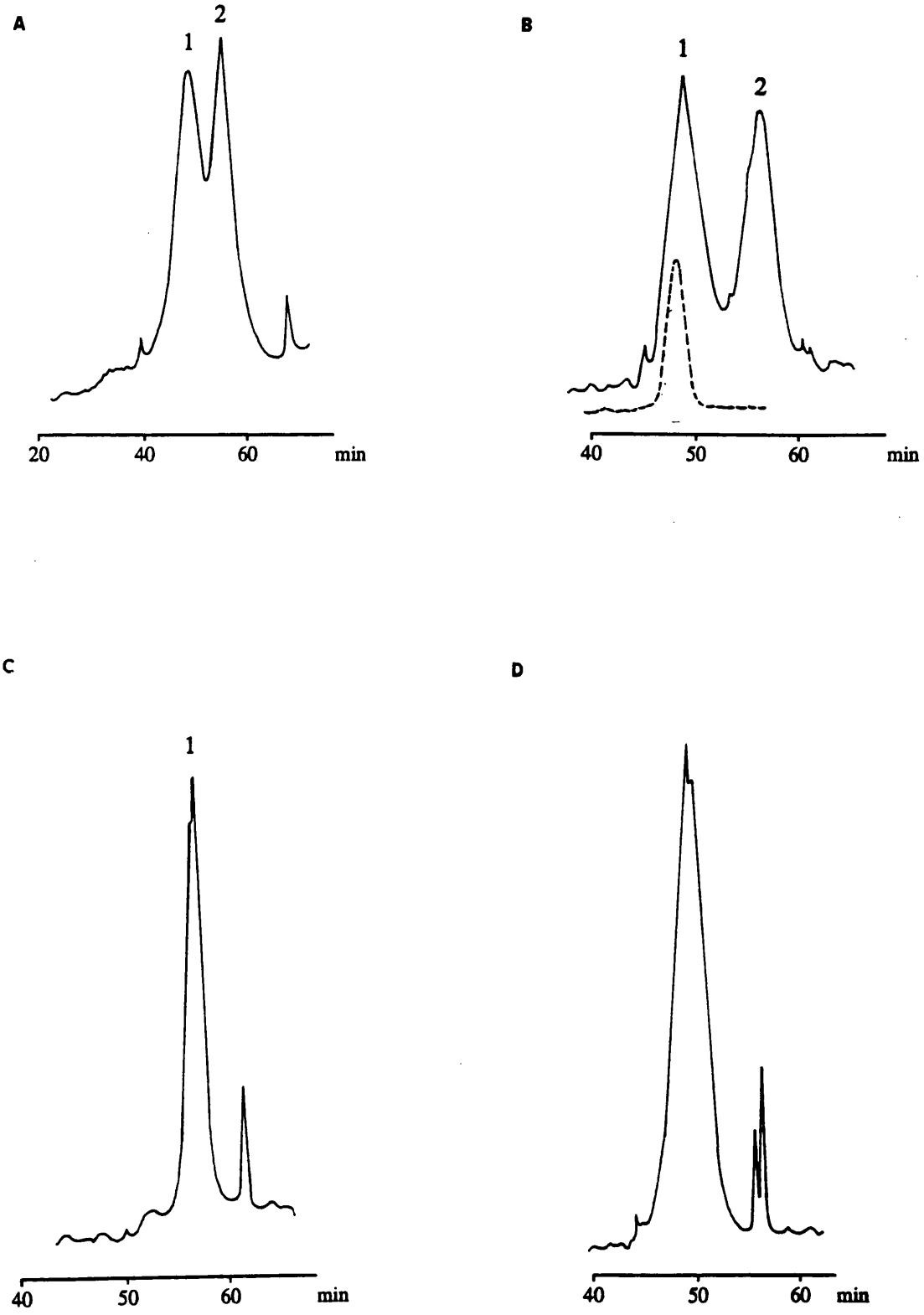


Figure 15. RP-HPLC of HIV-1 p24 GAG 270-373.
A) Crude material after HF cleavage on C_8 media. (System 6)
B) Crude 4-COR6-Fmoc-Gly material on C_{18} media. (System 7)
(------ Underivatised crude 104-mer)
C) Purified 4-COR6-Fmoc-Gly derivatised p24 fragment. (System 7)
D) Free 104-mer after treatment with 5% aqueous TEA. (System 7)

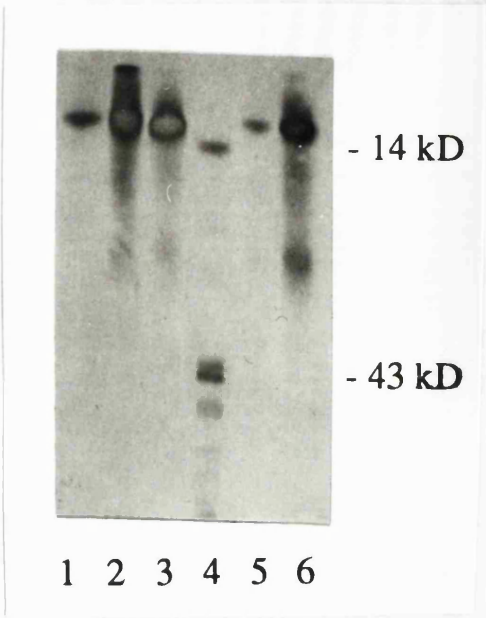


Figure 16. SDS-PAGE of purified material from RP-HPLC (Fig. 15B).
Lane 1, 2 and 3 represent increasing concentrations of underivatised 104-mer.
Lane 4 shows low molecular weight markers.
Lane 5 and 6 represent low and high concentrations of 4-COR6-Fmoc-Gly derivatised 104-mer.

Once the experimental data had confirmed that the later eluting peak was in fact 4-COR6-Fmoc-Gly derivatised 104-mer, it was then necessary to determine the homogeneity of the separated polypeptide. Thus, 40mg (3μmol) of crude cleavage product was purified on C_{18} reversed-phase media and a series of fractions were taken through both the early and late eluting peaks. Amino acid analysis was then performed on each of the combined fractions from the two peaks. The results (Table 2) showed that the poorest expected to found ratios were for samples taken from the early eluting peak (ie. underivatised). This was not unexpected and confirmed that it represented the family of terminated sequences. The amino acid composition of the fractions taken through the later eluting peak were in general closer to the theoretical values. In particular the fraction corresponding to the central portion of the later eluting peak showed an amino acid composition that was identical (within ±10%) to the expected results. Furthermore, the amino acid analysis obtained for the central fraction indicated

a product that had similar homogeneity to underivatised 104-mer that had previously been purified using a combination of conventional techniques (ie. dialysis, RP-HPLC, size exclusion and ion-exchange chromatography; [Mascagni *et al*, 1990]), as shown in Table 2. The purity of the material taken from the central part of the later eluting peak was further proven by IEF gel electrophoresis (Fig. 17). This indicated that the peptide focused at pH5.6 and was judged about 90% homogeneous. The appearance of an additional band at approx. pH9.3 was attributed to a different conformational form of the protein as described in Appendix II.

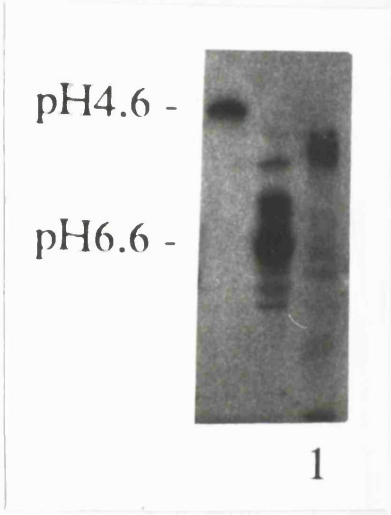


Figure 17. IEF gel of 4-COR6-Fmoc-Gly derivatised 104-mer (lane 1),

In conclusion, the separation obtained with the 104-mer using chromatographic probe (**11**:R6) demonstrated the applicability of this purification scheme to synthetic polypeptides ranging from 17 to 104 residues in length. The degree of homogeneity obtained, in the case of the larger peptides, as determined by amino acid analysis and SDS gel electrophoresis, after just one purification step by RP-HPLC was similar if not better than material that had undergone four conventional purification steps.

Residue	Expected	C-18 RP HPLC Purified 104mer		
		No Fmoc probe RT=49 mins	Fmoc probe attached RT=57 mins	104mer purified by standard four-step scheme
Asp/Asn	9	6.3	8.5	9.5
Thr	8	6.6	7.6	8.6
Ser	5	3.3[x]	4.7	4.9
Glu/Gln	13	11.1	13.0	13.5
Pro	6	4.6	6.1	6.5
Gly	6	6.0	7.1[y]	6.2
Ala	12	12.0	12.3	12.7
Val	7	5.0	6.8	6.9
Met	5	4.0	5.3	5.6
Ile	4	1.6	3.3	2.8
Leu	9	7.1	9.0	8.9
Tyr	3	1.4	2.7	2.5
Phe	2	1.1	1.8	1.8
Lys	7	4.9	7.0	7.1
His	1	--[z]	--[z]	0.9
Arg	6	3.7	6.6	6.1

[x] The hydrolyses was performed in conditions that did not prevent the loss of Ser and Thr residures.

[y] Extra Gly residue was introduced with the Fmoc probe.

[z] The polypeptide contained a $N_{(Im)}$-DNP-His residue which was not deprotected during hydrolysis.

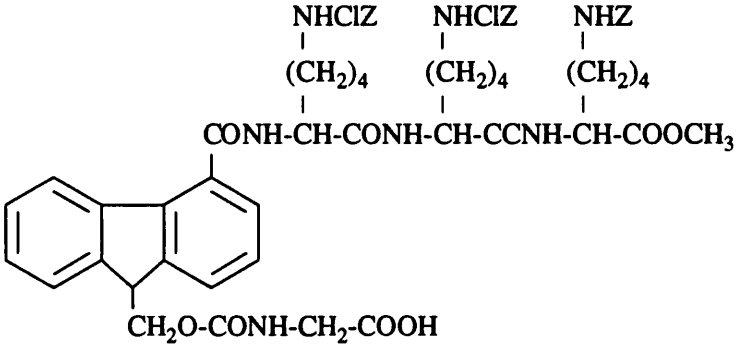
Table 2. Amino acid analyses of p24 270-373 purified with and without the use of 4-COR6-Fmoc-Gly probe.

PURIFICATION OF SYNTHETIC PEPTIDES ON THE BASIS OF CHARGE

PURIFICATION OF A TWENTY-FIVE RESIDUE PEPTIDE

In the previous sections the application of the hydrophobic 4-COR-Fmoc-Gly probes (**11**:R5 and 6) to the purification of peptides by RP-HPLC has been discussed. The results have indicated that with the more hydrophobic probe (**11**:R6), separation of peptides up to 104 residues is possible. To avoid the use of organic solvents necessary for RP-HPLC, which may be deleterious to the conformational integrity of proteins, an alternative system based on aqueous solvents would be advantageous. Ion-exchange chromatography enabling separation of material on the basis of charge would provide a useful extension to the application of 4-COR-Fmoc probes for the purification of synthetic peptides. Thus, 9-(hydroxymethyl)fluorene-4-carboxylic acid (**1**) was derivatised with amino acids with charged side chains.

An aliquot of HIV-1 ENV gp120 loop peptidyl-resin (sequence p.66) was deprotected at the N-terminus with 50% TFA in DCM. The resulting free amino group was then derivatised using 2 equivs. of the preformed glycine analogue 4-COR-Fmoc-Gly (**11**:R7) ion-exchange probe in accordance with route 2 (Scheme 3). Coupling was mediated by the HOBt/DCC activating combination (two fold excess of each) and completed after 60 min.



11:R7

172mg (29μmol) of treated peptidyl-resin was cleaved using the High HF procedure. The crude material that precipitated on addition of diethyl ether was filtered off and dissolved in 10% acetic acid. A white powder was recovered after lyophilisation weighing 112mg (110%). The higher than expected mass was probably due to the presence of scavengers in the crude product. The effect of incorporating the positively charged 4-COR$\underline{Z}$-Fmoc-Gly probe was to alter the net charge from two negative in the free peptide to zero in the derivatised peptide analogue. Consequently it would be possible to achieve separation by anionic exchange chromatography because the negatively charged impurities would be retained longer on the column. Thus, a sample of the crude cleavage product was purified using a strong anionic exchanger on an FPLC instrument. The buffer system used was bis-Tris HCl at pH6.5 with 1M NaCl in the eluting solvent. Using this buffer system a salt concentration of 0.15M NaCl was required to displace the underivatised impurities (peak 2; Fig. 18A) as opposed to 0.05M for the 4-COR$\underline{Z}$-Fmoc-Gly derivatised peptide (peak 1; Fig. 18A).

Even though the addition of probe (**11**:R$\underline{Z}$) provided the peptide with an overall net charge of zero, a possible explanation for the slight retention on the column could be that the pI of the molecule was close to 6.5 or that other matrix-peptide interactions were responsible (ie. not charge-charge). The peak which eluted at 0.05M salt was collected and desalted on a C_4 reversed-phase column and found by FAB-MS ($[M+H]^+$: expected 3543; found 3545) and amino acid analysis (Table 1) to be the desired 4-COR$\underline{Z}$-Fmoc-Gly derivatised peptide. An analytical quantity of the desalted 4-COR$\underline{Z}$-Fmoc-Gly derivatised peptide was then reinjected on the C_4 reversed-phase media (Fig. 18B) and eluted at 28 min. A separate aliquot of purified 4-COR$\underline{Z}$-Fmoc-Gly derivatised peptide was then treated with 5% aqueous TEA and after 30 min injected on the reversed-phase column. The cleavage product eluted as a single peak with a shorter retention time at 24 min (Fig. 18C), due to the loss of the 4-COR$\underline{Z}$-Fmoc probe. The homogeneity of the peptide was then confirmed following isolation of the

pure cleaved product using FAB-MS ($[M+H]^+$: expected 2877; found 2878) and amino acid analysis (Table 1).

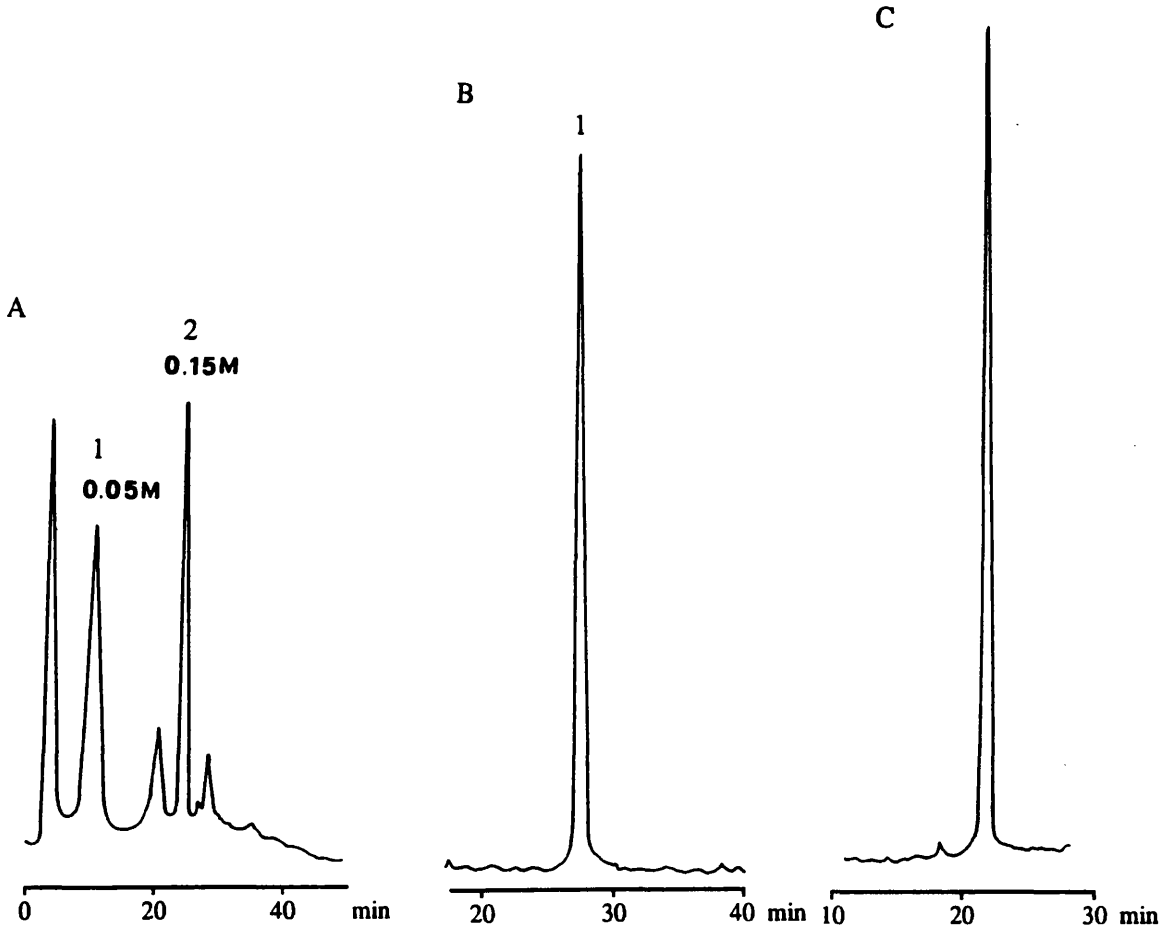


Figure 18. Chromatographic analysis of 4-CORZ-Fmoc-Gly derivatised HIV-1 ENV gp120 peptide.

A) Anionic exchange FPLC of crude material after HF cleavage. (System 8)

B) C_4 RP-HPLC of purified 4-CORZ-Fmoc-Gly derivatised peptide. (System 5)

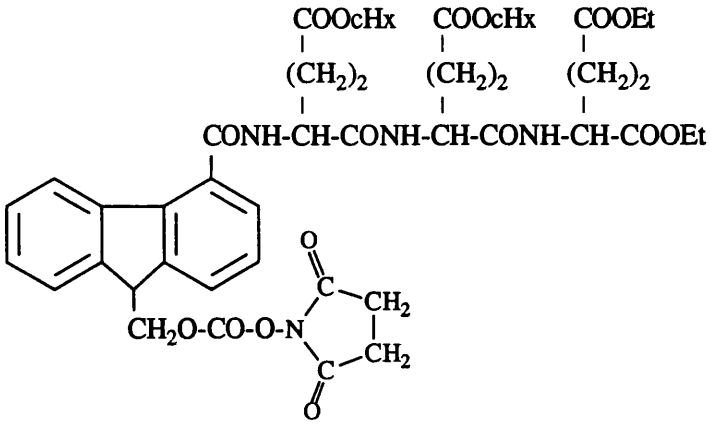
C) C_4 RP-HPLC of free peptide after treatment with 5% TEA. (System 5)

PURIFICATION OF A CYCLIC TWENTY-FIVE RESIDUE PEPTIDE

In the previous investigations the 4-COR-Fmoc derivatives were incorporated onto the N-terminus of the peptidyl-resins as preformed amino acid derivatives. The procedure for introducing the chromatographic probes onto the peptidyl-resin could be much simplier if activated analogues were used (eg. HOBt or succinimidyl) indicated in Scheme 3 (route 1). The suitability of this approach was examined by derivatising a twenty-five residue FMDV peptide [Bittle *et al*, 1982] with the negatively charged 4-COR-Fmoc-NSu probe (**10**:R**8**). The peptide had previously been synthesised using Boc chemistry, with capping.

NH_2-Cys-ACA-Gly-Ser-Gly-Val-Arg-Gly-Asp-Phe-Gly-Ser-Leu
S
S
CO_2H-Cys-ACA-C(Acm)-Pro-Leu-Gln-Arg-Ala-Val-Arg-Pro-Ala

The sequence included two Cys residues which were added to enable cyclisation.



10:R**8**

The use of the ion-exchange probes to a cyclic peptide would add a further dimension to the generality of the proposed purification system. It should be noted that even though three glutamic acid residues were incorporated into probe (**10**:R**8**) only two were charged since the C-terminal Glu was protected at both carboxyl groups as ethyl esters. Methyl and ethyl esters are not hydrolysed by anhydrous HF [Roeske, 1981] and would remain in place on the chromatographic probe. The overall net charge on the free peptide was +2 and upon addition of the 4-COR-Fmoc-NSu derivative (**10**:R**8**) would be changed to -1. Altering the polarity of the material would therefore enable separation from underivatised impurities on an anionic exchange column. The N-terminal protection on the peptidyl-resin was removed with 50% TFA in DCM and then treated with a four-fold excess of 4-COR-Fmoc-NSu (**10**:R**8**). All the previous 4-COR-Fmoc probes had been used as the prederivatised Gly compounds (route 2; Scheme 3). However, after β-elimination Gly would still be attached to the N-terminus of the purified peptide which, while appropriate for FMDV 141-157, with its N-terminal Gly residue, would require synthesising different probes to cover all the available amino acids. Thus the application of the succinimidyl ester was an obvious advantage. A small quantity of the tertiary amine DIEA was added to catalyse the coupling reaction. The incorporation of the 4-COR-Fmoc derivative (**10**:R**8**) was monitored spectrophotometrically using the method described above and found to be complete in 90 min. 129mg (26μmol) of 4-COR**8**-Fmoc derivatised peptide was cleaved using the High HF procedure The crude cleaved peptide following HF treatment was then dissolved in 50ml of 10% acetic acid, the solution diluted with water to 200ml and the pH increased to 6-7 to allow cyclisation by air oxidation at room temperature for 48h [Stewart and Young, 1984]. The usual pH of 7.8 was not used due to the possibility of premature loss of the 4-COR-Fmoc probe from the peptide.

The lyophilised, crude cyclised material was then purified by anion exchange FPLC (Fig. 19A), since the addition of the 4-COR**8**-Fmoc probe had altered the

net charge of the FMDV peptide from two positive to one negative. Separation was performed on Mono Q media using 20mM phosphate buffer at pH6.5 and a linear elution gradient (0.5M NaCl) over 35 min (0-40%). The positively charged underivatised peptidic material and uncharged impurities eluted with the solvent front while the desired derivatised cyclic peptide was retained and eluted at a salt concentration of 0.1M NaCl. The peak labelled 2 (Fig. 19A) was found by FAB-MS to represent the desired material ($[M+H]^+$: expected 3354; found 3356).

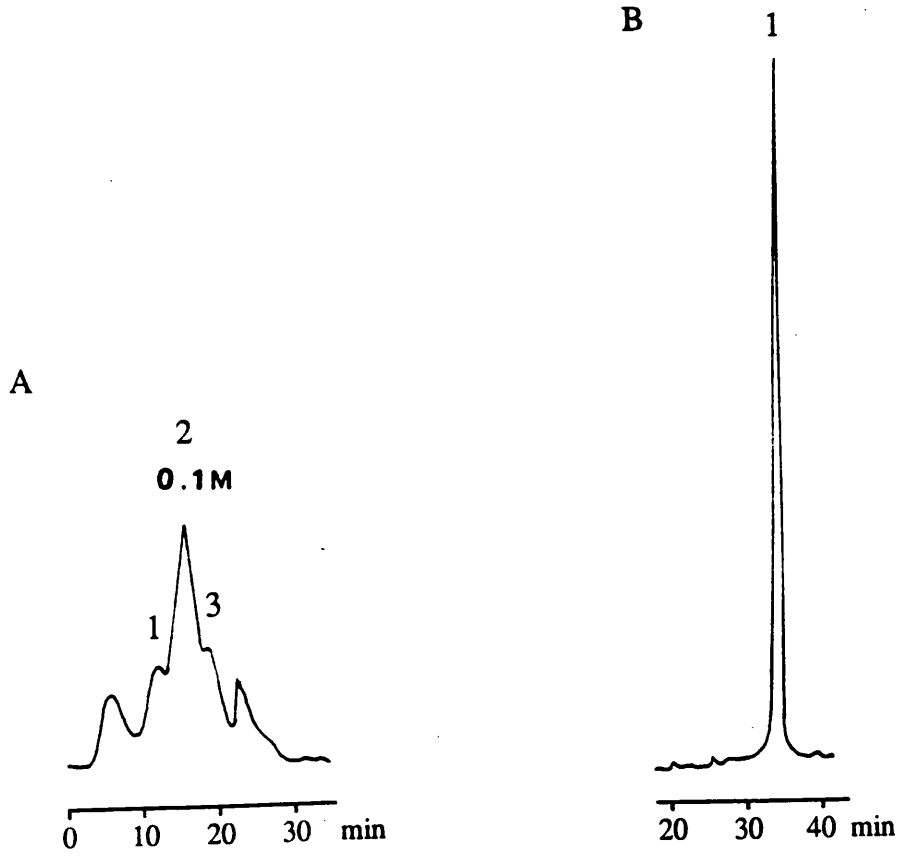


Figure 19. Chromatographic analysis of 4-COR8-Fmoc derivatised cyclic FMDV peptide.

A) Anionic exchange FPLC of crude material after HF cleavage. (System 9)

B) C_4 RP-HPLC of free cyclic peptide after treatment with 5% TEA (System 5)

The two flanking peaks (labelled 1 and 3; Fig. 19A), were collected and analysed by FAB-MS. The results indicated that the left-hand shoulder labelled 1 in figure 19A was underivatised peptide ($[M+H]^+$: found 2645) and the right-hand shoulder (labelled 3) was either a dimeric form or the uncyclised 4-COR$\underline{8}$-Fmoc derivatised peptide ($[M+H]^+$: expected 3354; found 3352). The reversed-phase profile of the free peptide (Fig. 19B), after the β-elimination of the 4-COR$\underline{8}$-Fmoc group with mild base (5% aqueous TEA), indicated the presence of an impurity (approx. 5%) as a shoulder on the main peak. This impurity probably originated from the overlapping of the peaks 1 and 2 on the ion-exchange media. The pure form of the free peptide had the expected molecular weight by FAB-MS ($[M+H]^+$: expected 2645; found 2645) and amino acid analysis (Table 1).

CONCLUSIONS

While synthetic methodologies have made significant advances over the years, simple and quick techniques that are capable of separating closely related peptide impurities have not been equally developed. The aim of this project was to design a molecule that could be used to 'label' the desired sequence and that when attached would drastically change the chromatographic characteristics of the derivatised peptide and thereby facilitate its separation from underivatised impurities. The coupling of the chromatographic probe had to be simple and once incorporated be stable to the cleavage conditions. Furthermore, after purification of the derivatised peptide, retrieval of the pure peptide from the derivatised analogue must be possible. The fluorene molecule, linked to the N-terminus of the peptide chain through a base-labile urethane group, offered these properties of acid stability and reversibility. The properties possessed by the 9-(hydroxymethyl)fluorene group were then extended to its derivative, 9-(hydroxymethyl)fluorene-4-carboxylic acid (4-COOH-Fmoc, **1**). Apparently, the sensitivity described for the Fmoc group (**2**) to electrophilic substitution during HF cleavage (Grode *et al*, 1992) was not encountered with the 4-COR-Fmoc probes due to the electron-withdrawing effects of the 4-carboxyl function. The carboxyl functionality at position 4 provided the means by which various chemical groups could be attached. Previous work in the laboratory on the covalent attachment of one, two and three α-aminodecanoic acid groups to a 20 residue peptide and the resulting increase in retention time on RP-HPLC media, suggested that the similar effect could be achieved through the 4-COOH-Fmoc molecule. Thus, two lipophilic 4-COR-Fmoc-Gly probes (**11**:R**5** and **6**) were synthesised to allow separation of derivatised peptides from underivatised impurites on the basis of hydrophobicity. To avoid using the organic solvents necessary in reversed-phase chromatography, which could disrupt the biological activity of the synthetic peptides, alternative 4-COR-Fmoc probes (**11**:R**7** and **10**:R**8**) were synthesised. The latter possessed positive and negative charged side

chains (trilysine and triglutamic acid, respectively) and were designed to alter the overall charge of the attached peptide, for purification on ion-exchange media. To attach the chromatographic probes to the peptidyl-resin, two different methods were used. The first and more complicated involved the pre-derivatisation of the probes with an amino acid, for which Gly was chosen as representative of group of compounds. Simple amide bond formation was then required to couple the 4-COR-Fmoc-Gly probe to the last residue of the peptide chain. The alternative route required the use of the activated succinimidyl carbonate compound 4-COR-Fmoc-NSu, which upon addition to the free N-terminus of the peptidyl-resin, generated the base-labile urethane linkage directly. To release the derivatised peptide from the resin and cleave the side chain protecting groups a treatment with anhydrous HF was necessary. The extreme acidolytic conditions were found not to be detrimental to the 4-COR-Fmoc probes.

Chromatographic analysis of the crude derivatised peptides, ranging from 17 to 104 residues then followed. The use of less hydrophobic analogue 4-COR-Fmoc compound (**11**:R**5**) for the purification of the 17 residue FMDV fragment served to show that (i) a separation of 6 min could be achieved between derivatised and underivatised peptide and (ii) the undesired coupling of **11**:R**5** to a deletion sequence indicated the importance of a capping protocol during peptide synthesis. The application of the more hydrophobic 4-COR-Fmoc probe (**11**:R**6**) to a longer 25 residue HIV-1 gp120 fragment demonstrated that on reversed-phase media a separation of 26 min could be achieved between closely related peptide impurities and the desired sequence which possessed the chromatographic probe. The significant increase in retention time for the relatively short 25 amino acid peptide with 4-COR**6**-Fmoc-Gly attached lead to an extension of the purification method towards one of the original aims of the research ie. to develop a means to purify large synthetic polypeptides. Thus, the attachment of 4-COR-Fmoc-Gly probe (**11**:R**6**) to the TAT 39-86 fragment and

the 104 residue p24 fragment produced a separation of approx. 10 min and 7 min on reversed-phase media, respectively. In the case of the 104-mer amino acid analysis of the material purified with the aid of the chromatographic probe indicated that its degree of homogeneity compared favourably with that of the same polypeptide purified by a conventional four-step procedure.

The flexibility of the purification procedure was proven by its application not only to reversed-phase but also to ion-exchange media. The idea was that by incorporating charged 4-COR-Fmoc molecules onto the peptide of interest it would be possible to alter the ionic properties of the derivatised peptides and hence its chromatographic characteristics. Thus, by adding 4-COR-Fmoc-Gly (**11**:R7) to HIV-1 gp120 loop peptide the overall net charge was altered from +2 in the underivatised form to zero. The effect was to cause the underivatised impurities to bind to a cationic ion exchange media while the desired product with no overall charge eluted shortly after the solvent front. The use of negatively charged tri-glutamic acid 4-COR-Fmoc-NSu probe (**10**:R8), besides showing the feasibility of introducing the probes as succinimidyl carbonates, also demonstrated that the method was not restricted to linear sequences (ie. 25 residue cyclic FMDV fragment).

β-elimination of the 4-COR-Fmoc group to release the free purified peptide was accomplished in aqueous phase. Quantitative cleavage of the chromatographic probe, in the aqueous system, was shown to occur in 30 min. This was demonstrated by a time dependent reduction in the peak corresponding to the 4-COR6-Fmoc-Gly derivatised p24 GAG 270-373 protein and the appearance of another peak with a shorter retention time due to the free protein, when followed on reversed-phase media. The integrity the purified 4-COR6-Fmoc-Gly derivatised and free peptides were determined by FAB-MS, amino acid analysis and also, SDS-PAGE and IEF gel electrophoresis. All the methods used indicated that the expected 4-COR-Fmoc derivatives had been separated and that following base treatment the correct free peptide was obtained.

CHAPTER THREE

PEPTIDE SYNTHESIS BY FRAGMENT CONDENSATION

INTRODUCTION

Since the advent of solid phase peptide synthesis (SPPS) back in the early 60's, a considerable amount of effort has been directed towards the chemical synthesis of proteins. As has been described in the introduction, the use of SPPS has been limited to molecules of about 150 residues, due to the exponential accumulation of closely related impurities, which after 150 cycles represent a large proportion of the crude product and make purification to homogeneity difficult if not impossible. Furthermore, current methods for determining sequence homology are generally unable to offer definitive analysis of the total protein sequence ie. the presence of a sequence lacking one amino acid would be difficult to detect. Hence fragment condensation may not only offer a route to improving SPPS for short proteins, but may also open an avenue towards the construction of sequences longer than 150 residues. Fragment condensation involves constructing short segments of the desired sequence which retain their side-chain protecting groups. These fully protected peptide fragments are then purified before being coupled together to give the complete peptide chain. The ability to fully characterise these shorter fragments would give far greater confidence to the homology of the final product than if the same protein had been synthesised in a stepwise fashion.

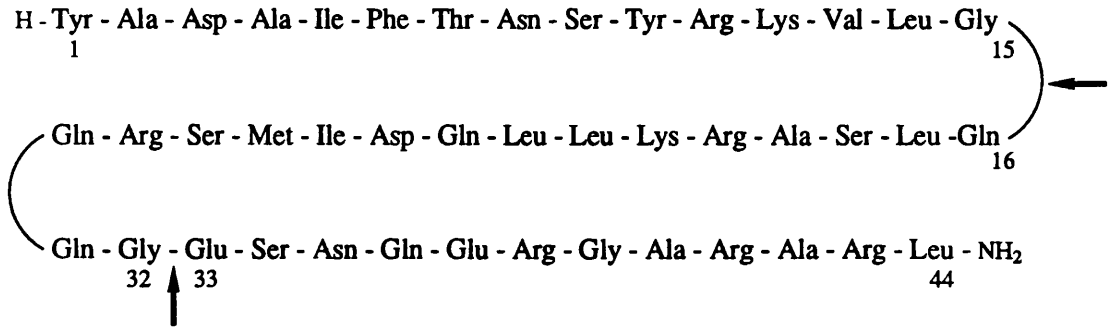
Although solution phase segment coupling has been well documented [Finn and Hofmann, 1976], a solid phase approach where the purified and fully characterised protected peptides are coupled to the resin bound C-terminal fragment is less well reported. The advantages of the latter approach are those that apply equally well to stepwise SPPS ie. simplification of repeated procedures and ease of removal of excess reagents. The resin also serves to protect the carboxyl terminus thus preventing self-activation and coupling. In the solution phase the reactive carboxylic acid requires a protecting group (eg. t-butyl; [Kamber and Riniker, 1991]) that requires removal at a later stage.

The major limitations of the fragment condensation approach often relate to the extreme hydrophobicity of the protected peptide intermediates, a characteristic attributed to the amino acid side chain protecting groups. The net result are products with poor solubility in many solvents which limits their purification on many chromatographic systems and makes it difficult to obtain a sufficiently concentrated solution for coupling. The aim of the research undertaken was to examine ways of improving methods for the synthesis of proteins by fragment condensation. However, before embarking on the route towards this goal, it was necessary to evaluate the many techniques present in the literature for the synthesis, purification, and eventual coupling of protected peptide fragments. The difficulties encountered during the course of this study, in part due to initial inexperience, but mainly as a result of the nature of the research and of the model peptide chosen, would give the impression that the suggested improvements to be instigated, following each experiment, were not applied strictly in a flowing manner. Consequently, the discussion contained in this chapter reflects the many necessary attempts and changes that were made as more experience was gained.

In general, the synthesis, as described in the introduction, involves the coupling of short protected segments (<9 residues in length) onto a short resin bound fragment. Consequently, this would necessitate a large number of manipulations in order to construct a protein of say more than 150 residues. The division of the protein sequence into larger segments would thus be more desirable in that this would reduce the number of manipulations.

RESULTS AND DISCUSSION

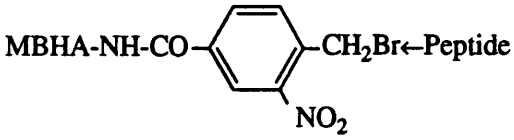
To gain experience in handling large protected segments of approx. 15-20 residues and to determine which methods were most successful, in our hands, for synthesising, purifying and coupling such peptides, a model peptide, human Growth Releasing Hormone (1-44) was chosen. The 44 residue sequence was cut at Gly residues to divide the peptide conveniently into three fragments; 1-15, 16-32 and the C-terminal 33-44.



The use of C-terminal Gly or Pro residues overcame the serious risk of racemisation during fragment coupling due to oxazalone formation. Unfortunately, the latter amino acid and to a lesser extent the former can facilitate diketopiperazine (DKP) formation [Pedroso *et al*, 1986; Mergler *et al*, 1988], resulting in the loss of peptide chains from the resin. This unwanted side reaction could be significantly reduced if the third amino acid was coupled manually using the method described by Suzuki and Endo [1977]. It was decided initially to work on the solid phase fragment condensation technique due to its simplicity. This required stepwise SPPS of the resin bound C-terminal peptide and the synthesis and purification of the other protected peptide intermediates. To determine which methodologies were most appropriate the following rationale was adopted; (i) the two protected peptide fragments (1-15 and 16-32) were synthesised on four different handles (Fig. 1), which were cleavable under a variety of conditions, using Boc and Fmoc chemistries, (ii) the crude protected fragments were purified on various chromatographic media and the best method in terms of yield, purity and handling of the material was determined, (iii) the C-terminal (33-44) fragment was synthesised using Fmoc and Boc chemistries to evaluate the effect of the different side chain protecting groups on segment condensation and (iv) a variety of activating agents/solvent combinations were investigated to find which gave the highest incorporation yields of protected peptide.

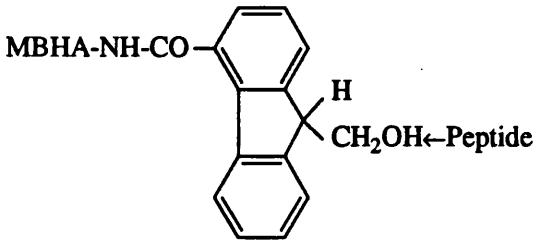
RESIN-HANDLES FOR BOC CHEMISTRY

Photolabile Handle

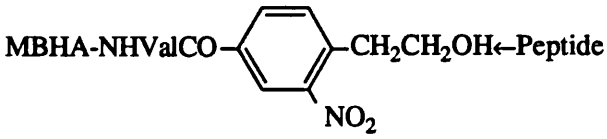


21

Base-labile Handles



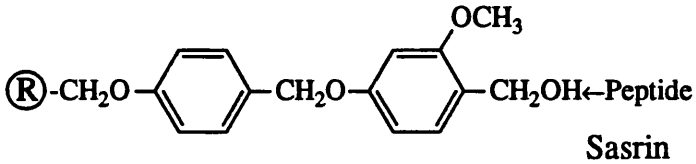
22



23

RESIN-HANDLE FOR FMOC CHEMISTRY

Acid-labile Handle



Sasrin

24

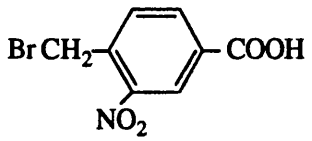
Figure 1. Structures of resin-handles used for synthesis of protected peptides

BOC CHEMISTRY

SYNTHESIS OF PROTECTED PEPTIDES USING BOC CHEMISTRY

Synthesis on photolabile resin-handle (21)

The first resin handle to be investigated was the photolabile *o*-nitrobenzyl linker [Rich and Gurwara, 1975; Giralt *et al*, 1982].



25

The initial step involved coupling 4-bromomethyl-3-nitrobenzoic acid (**25**) as a symmetrical anhydride to polystyrene based MBHA resin. 2 equivs. of the former were activated with 1 equiv. of DCC in DCM. Ninhydrin analysis indicated that quantitative incorporation had been achieved after 90 min. The attachment of the first amino acid through the bromo group was accomplished by reacting Boc-Gly-OH as the caesium salt in dry DMF [Giralt *et al*, 1982]. The caesium salt was formed by neutralising a 25% ethanol solution of the Boc protected amino acid with caesium carbonate. The solvent was reduced on a rotary evaporator and the residual liquid dried by lyophilisation. A three fold excess of Boc-glycine-O^-Cs^+ was added to a suspension of **21** in DMF which had been pre-heated in a water bath set at 40°C. The reaction mixture was stirred mechanically without the use of a magnetic stirring bar that may have damaged the polystyrene matrix. Stirring was continued for 18h and the extent of amino acid incorporation onto the photolabile handle was determined by the method described by Gisin [1972] and amino acid analysis. Both methods of analysis indicated an incorporation of approximately 33% with respect to the theoretical value derived from the resin substitution (Table 1).

Handle	% incorporation of 1st amino acid	% peptide on resin		% peptide remaining on resin after cleavage	
		1-15	16-32	1-15	16-32
21	33%	89%	91%	30%	24%
22	45% 0.194mmol/g	95%	94%	T=30min nil	T=30min nil
23	99% 0.502mmol/g	94%	83%	T=30min 95%[$] T=6h 17%	T=30min 58%

Table 1. Experimental analysis of resin handles for Boc chemistry

Residue	Expected		Found					
			1-15-®			16-32-®		
	1-15	16-32	**21**	**22**	**23**	**21**	**22**	**23**
Tyr	2	-	1.5	2.0	1.9	-	-	-
Ala	2	1	2.3	2.1	2.0	0.9	1.4	0.9
Asp/Asn	2	1	2.0	2.0	2.1	1.0	1.3	1.0
Ile	1	1	0.8	0.8	0.8	1.4	0.5	0.7
Phe	1	-	0.9	1.0	1.0	-	-	-
Arg	1	2	0.9	0.9	0.9	1.1	0.9	2.0
Lys	1	1	1.2	1.0	1.0	1.0	1.1	1.0
Val	1	-	1.2	0.8	2.2[#]	-	-	1.1[#]
Leu	1	3	0.9	1.1	1.1	2.2	2.2	3.2
Gly	1	1	2.1	1.1	1.1	1.4	0.1	1.1
Glu/Gln	-	4	-	-	-	4.3	3.8	3.8
Met	-	1	-	-	-	1.0	0.6	0.3

Table 2. Amino acid analysis of peptidyl-resins following synthesis of the protected peptides by Boc strategy. (# One residue of Val from Internal Standard)

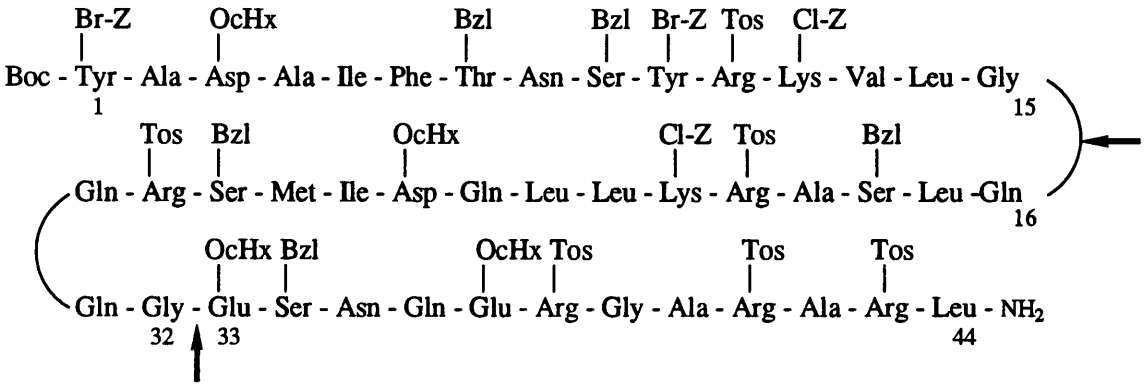
In an attempt to improve the incorporation of Boc-Gly-OH onto the resin, the reaction was repeated several times without improvement. Possible causes of the poor incorporation of amino acid are (i) the inclusion of water from either reagents, eg Boc-glycine-O^-Cs^+, or solvent, resulting in the premature hydroxylation of the **21**, (ii) incomplete reactivity with the bromine group, (iii) the existence of a competing reaction resulting in the termination of the reacting groups on the resin or (iv) a poor batch of MBHA resin. In answer to point (iv), a ninhydrin analysis of the free amino groups was performed with a larger dilution factor to compensate for the high absorbance [Sarin *et al*, 1981]. The resin was found to possess 91.2% free amino groups which allowing for the error, confirmed that the problem was not due to the starting resin.

The DIEA salt was used as an alternative to the caesium salt [Giralt *et al*, 1982]. The coupling reaction was performed in freshly distilled, dry ethyl acetate and refluxed. After 18, 46 and 54 hours samples were taken and analysed using the standard ninhydrin test. An improvement was observed with a maximum incorporation of 44%. Adding further quantities of Boc-glycine-O^- $DIEA^+$ salt did not improve the yields, suggesting that handle **21** had been terminated. Further evidence for this conclusion was offered by an analysis of bromine content on the resin before and after derivatisation. The presence of bromine on the derivatised resin was investigated using a quantitative analysis which involved standard Lassaigne, followed by titration of the free bromine against silver nitrate [Vogel, 1948]. Using this approach the resin was found negative for the presence of bromine.

A third approach to achieve derivatisation of the resin involved the use of a crown ether which is known to complex and solubilise inorganic salts of amino acids [Mascagni *et al*, 1987]. The experiment involved reacting 1 equiv. Boc-Gly-O^-K^+ with an excess of 18-Crown-6 in different solvents ie. DMF, ethyl acetate and acetonitrile, for 3h. The amounts of amino acid incorporated as determined by the ninhydrin method were 33, 40.5 and 37%, respectively. These results did not improve on further reaction or on addition of more reagents. Since it was not possible to increase the yields above about 40%, it was

decided to use the resin which had 33% incorporation of the first amino acid. However, to ensure that any remaining reactive groups on the resin support were made unreactive, the derivatised resin with the first amino acid was shaken with 5% acetic anhydride in DMF for 10 min.

The photolabile resin handle derivatised with the first amino acid (ie. Boc-Gly-OH) was then used for synthesising the 15 and 17 residue peptides by the stepwise Boc strategy.



However, due to the possibility of diketopiperazine formation induced by the *o*-nitrobenzyl group, it was necessary to attach the third amino acid manually using the method described by Suzuki and Endo [1977]. The Boc group of the second amino acid was deprotected with 4N HCl in dioxane and 2.5 equivs. of the third amino acid was introduced as the N-methylmorpholine salt together with 2.5 equivs. DCC. The percentage coupling for the third amino acid for both fragments was complete as determined by quantitative ninhydrin analysis. The peptidyl-resin was then returned to the automated synthesiser to attach the remainder of the residues. On completion of the synthesis amino acid analysis was performed on the peptidyl-resin which indicated a reasonable found to expected amino acid composition (Table 2) and that no premature loss of peptide from the photolabile handle had been experienced during couplings (ie. 89% (1-15) and 91% (16-32), based on the substitution value of the resin), as shown in Table 1.

The cleavage of the two peptidyl-resins was performed by the peptide group in Barcelona under the supervision of Dr F. Albericio. To detach the fully protected peptides from the support it was necessary to suspend the resin bound peptide in DCM/TFE (7:3) and irradiate the stirred solution with ultraviolet light (λ=350nm) for 24h. After evaporating the solvent the yield of crude material was 71% (1-15) and 76% (16-32), of the theoretical peptide content on the resin.

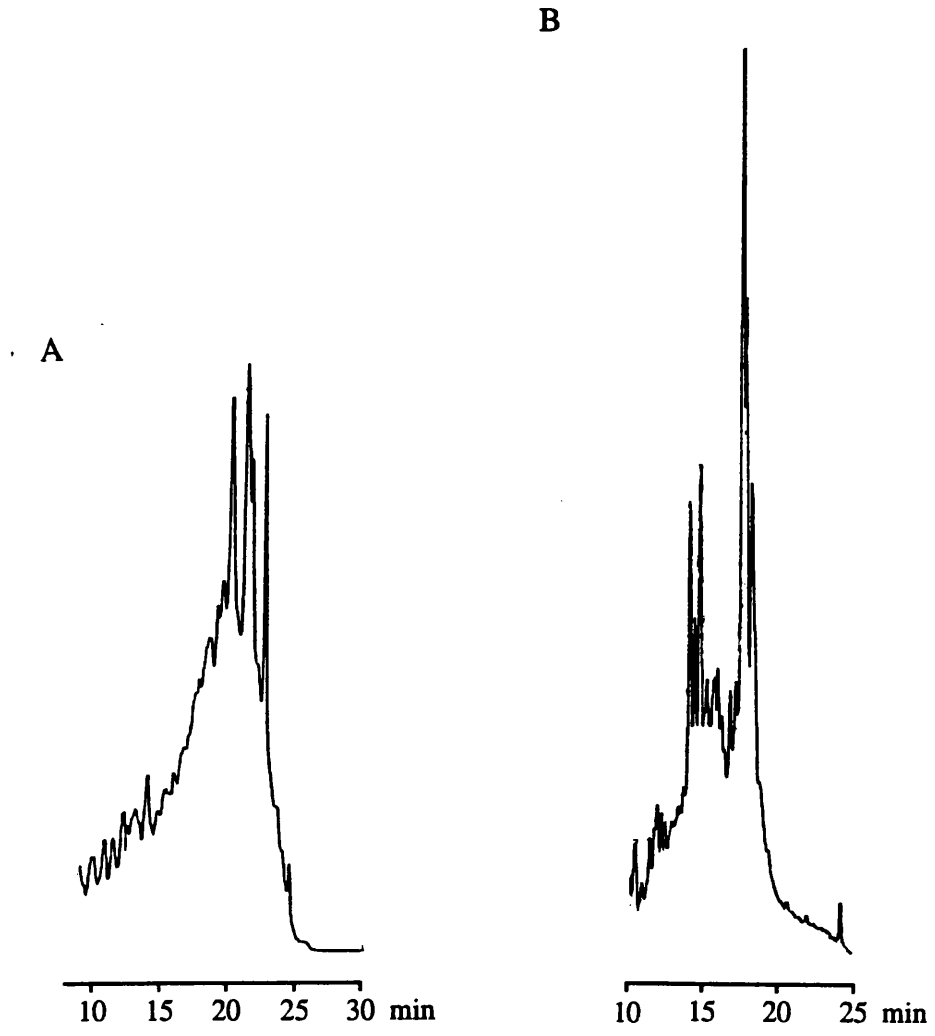
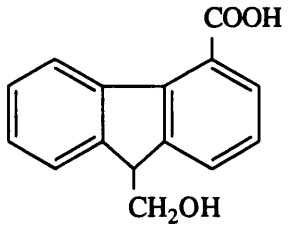


Figure 2. C_{18} RP-HPLC of crude protected peptide fragments from handle **21**. (System 10)
A) Sequence 1-15
B) Sequence 16-32

The crude products for both protected fragments were analysed on C_{18} reversed-phase media (Fig. 2). Both crude cleavage products gave complicated mixtures of peaks, making purification of the desired peptides almost impossible. Furthermore, the glassy crude products were poorly soluble in most organic solvents and only soluble in DMF with heating. The complexity of the crude cleavage product after photolysis necessitated the consideration of alternative resin handles for the synthesis of the protected peptide fragments.

Synthesis on base-labile resin-handle (22)

Base-labile handles offer true orthogonality since the side chain and N-terminal protecting groups are sensitive to acidolytic conditions, while the peptidyl-resin ester linkage requires an organic base to release the product. The first example of this category to be examined was the MBHA resin derivatised with 9-(hydroxymethyl)fluorene-4-carboxylic acid (**1**).



1

This molecule had been previously described by Mutter and Bellof [1984] for application to the synthesis of small protected peptides. To couple **1** to the solid support, a two fold excess was activated with DCC in 50% DMF/DCM. Using these reaction conditions Mutter and Bellof [1984] reported no difficulties with the derivatisation of the resin (eg. self-reaction through the alcohol and carboxylic acid groups to form esters). The DMF was required due to the poor solubility of 9-(hydroxymethyl)fluorene-4-carboxylic acid (**1**) in DCM. The resulting DCU byproduct was filtered off and the supernatant transferred to a suspension of MBHA resin in 50% DMF/DCM. After 6h, ninhydrin analysis indicated that 90% incorporation had been attained, at which point the reaction was stopped. The newly synthesised resin was derivatised with the first amino

acid, Boc-Gly-OH by reacting the resin with a ten fold excess of amino acid and DCC [Bodansky, 1984; Rink, 1987], together with a catalytic quantity of DMAP (0.2 equivs.).

To determine the degree of incorporation of the first amino acid an aliquot of the derivatised resin was treated with 50% TFA/DCM, the quantitative ninhydrin analysis was performed. This indicated that 65% glycine, as a percentage of the total theoretical incorporation had been achieved. To check the reproduceability of these absorbance readings, a quantity of Boc-Gly-PAM was similarly deprotected and treated with ninhydrin. The calculated presence of amino acid was 61.4%, almost 39% less than expected, in accordance with the inaccuracy of the ninhydrin technique for high substitution values of free amino groups. The degree of incorporation of the first amino acid after 2h reaction time, as indicated by amino acid analysis was in fact 45% (Table 1). To prevent the formation of deletion peptides the unreacted hydroxyl groups were capped using acetic anhydride and pyridine in DCM (1:1:2) for 30 min [Liu *et al*, 1990].

The two peptide fragments were synthesised using Boc chemistry in a stepwise fashion, using the same amino acid derivatives as described for synthesis on resin-handle **21**. At the end of the syntheses amino acid analysis was performed on the peptidyl-resins, showing that 95% (1-15) and 94% (16-32) peptide were present on the solid supports (Table 1). These results indicated that no significant loss of peptide from the resin had occured during capping and subsequent synthetic procedures.

In the original aricle of Mutter and Bellof [1984], an investigation was made into the sensitivity of the ester group to bases. It was reported that very mild alkaline conditions were required to cleave a tetrapeptide from the solid support (eg. 20% piperidine in DMF for 30 min). Application of the same cleavage conditions for the GRF fragments indicated, by amino acid analysis, that no peptide remained on the resin after 30 min (Table 1).

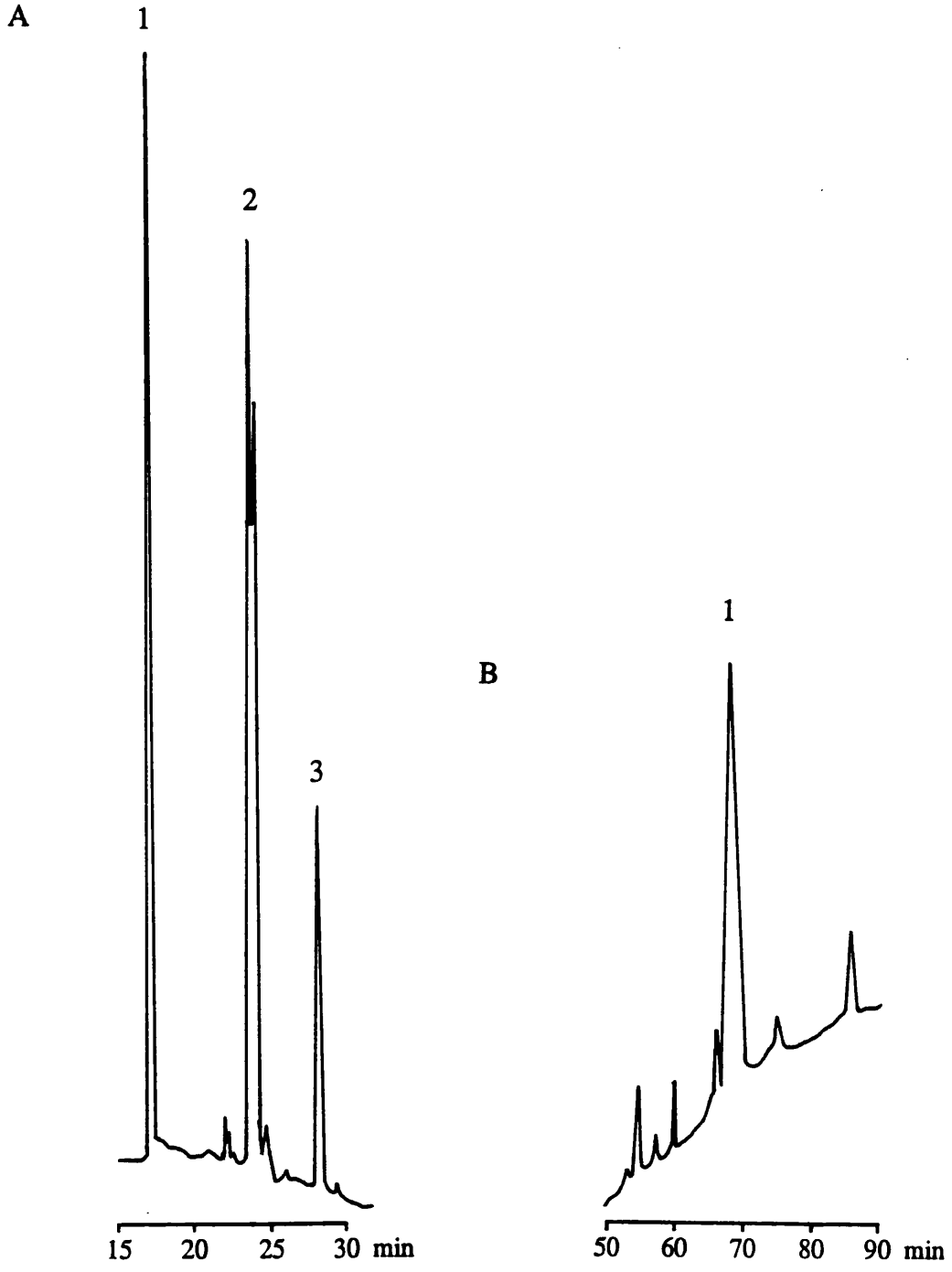


Figure 3. Crude protected peptide synthesised on handle **22**
A) C_{18} RP-HPLC of fragment 1-15. (System 11)
B) C_8 RP-HPLC of fragment 16-32. (System 12)

Residue	Expected		1-15			16-32	
	1-15	16-32	**22**	**22***	**23**	**22**	**23**
Tyr	2	-	1.9	2.1	1.8	-	-
Ala	2	1	2.0	2.0	2.0	1.1	0.9
Asp/Asn	2	1	2.1	2.2	2.0	1.9	0.9
Ile	1	1	1.0	0.8	0.8	0.9	1.0
Phe	1	-	0.9	0.9	0.9	-	-
Arg	1	2	1.1	0.6	1.1	2.2	2.4
Lys	1	1	1.1	1.0	1.0	1.0	1.0
Val	1	-	1.0	0.8	0.9	-	-
Leu	1	3	1.2	1.0	1.0	3.1	3.4
Gly	1	1	1.4	1.1	1.1	1.1	0.9
Glu/Gln	-	4	-	-	-	4.0	4.2
Met	-	1	-	-	-	0.7	0.8

Table 3. Amino acid analysis of fragments 1-15 and 16-32 from **22** and **23**, purified by analytical RP-HPLC. **22*** represents analysis of peak 2 (Fig.3A)

Following the 30 min base treatment the crude cleavage products were analysed by RP-HPLC. Fragment 16-32 gave one major peak (labelled 1, Fig. 3B) which was collected and shown by amino acid analysis (Table 3) and FAB-MS ($[M+H]^+$ Expected: 2812; Found; 2812) to correspond to the desired peptide. The RP-HPLC profile of crude 1-15 fragment, after a 30 min treatment with base (Fig. 3A), was more complicated with the appearance of two peaks at 17 and 28 min (labelled 1 and 3, respectively in Fig. 3A) together with two closely eluting peaks (labelled 2, Fig. 3A). Amino acid analysis of peak 1 showed it to be non-peptidic, while peaks 2 and 3 gave the correct amino acid compositions (labelled **22*** and **22** in Table 3, respectively). FAB-MS analyses of peak 2 ($[M+H]^+$ Expected: 2829; Found; 2322) and peak 3 ($[M+H]^+$ Expected: 2829; Found; 2427) did not, however, give the correct molecular ion that corresponded to the expected sequence. Since the correct amino acid composition was obtained the results would suggest some instability of one or more side chain protecting groups to the base, which may or may not have been related to the sequence of fragment 1-15. If degradation was a problem, a

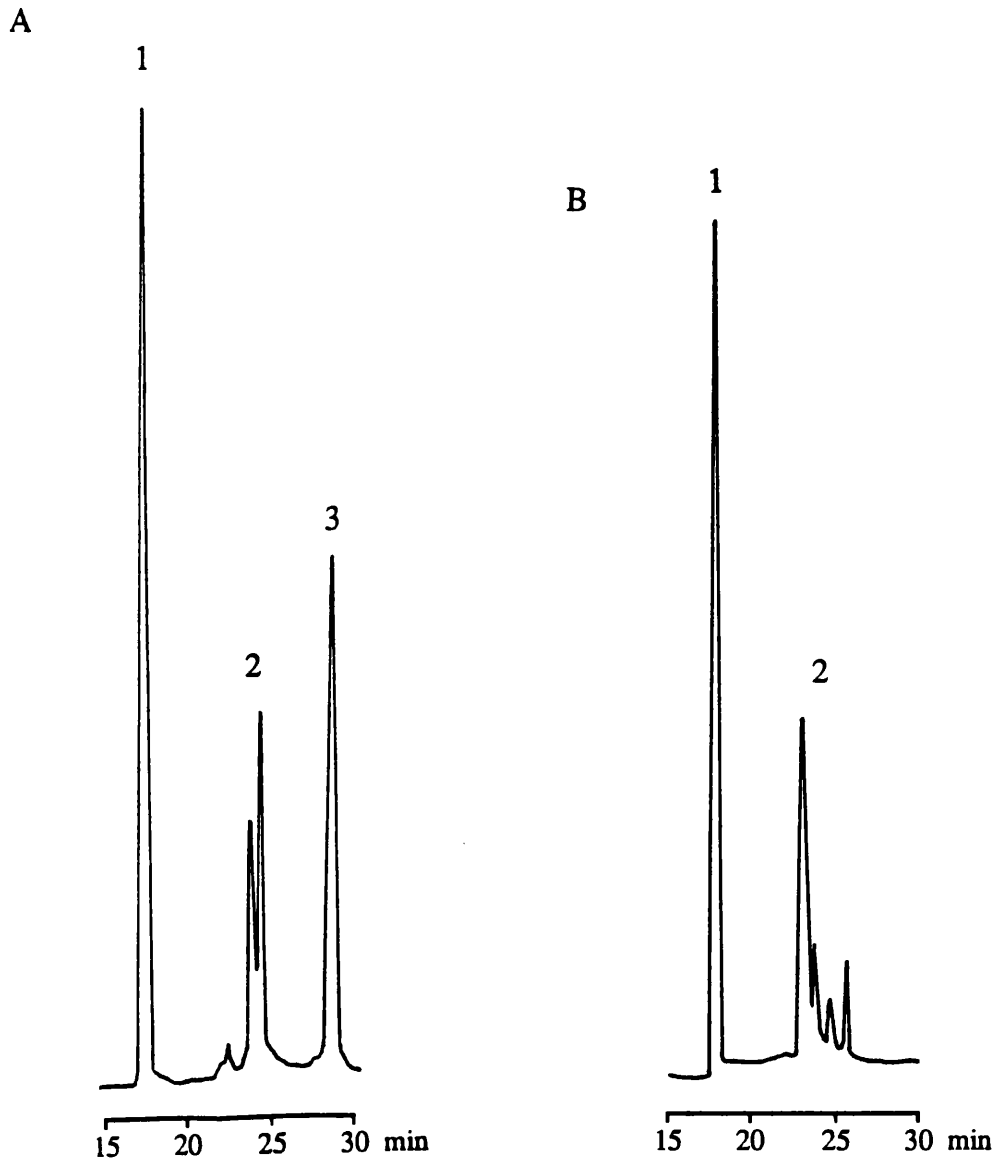


Figure 4. C_{18} RP-HPLC of crude fragment 1-15 synthesised on handle **22** in 20% piperidine/DMF. (System 11)
A) T=5 min
B) T=90 min

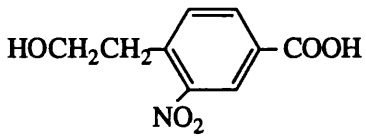
shorter base treatment (ie. 5 min) might enable the recovery of the desired product. Unfortunately, as seen in Figure 4, the same pattern of peaks was obtained although the proportion of the peaks labelled 2 was less. After a 90 min treatment (Fig. 4B) all peak 3 was lost. Thus, there appeared to be two levels of degradation, one that was almost instantaneous and a second, slower

reaction, which eventually converted all peak 3 to the peaks labelled 2 (Fig. 3A), probably due to sensitivity of other Bzl-based protecting groups to piperidine.

Sensitivity of benzyl-based chemical groups to base has been reported. For example, Schon *et al* [1983] demonstrated that the Bzl ester protecting group for Asp/Glu residues was unstable in the presence of piperidine and that Kuroda *et al* [1992] stated that the Br-Z protecting group for Tyr was affected by base.

In summary, the synthesis of fragment 16-32 on handle **22** proceeded smoothly and following treatment with the secondary amine, piperidine, gave the desired product as determined by amino acid analysis and FAB-MS. However, the protected peptide 1-15 while synthesised successfully, did not give the expected product following removal from the resin using base. Thus, to determine whether the observed degradation could be eliminated, a second base-labile handle (**23**) was examined.

Synthesis on base-labile resin-handle (23)



26

For this synthesis an internal standard, Val, was used to provide a convenient means of determining the substitution value of the resin. MBHA resin was reacted with a DCM solution containing a four fold excess of Boc-Val-OH and 0.5 equivs. with respect to the amino acid of the activating agent DCC. Quantitative coupling was achieved, as indicated by the ninhydrin test, after 30 min. The N-terminal protecting group was removed with 50% TFA/DCM and the resin neutralised with 5% DIEA/DCM. 1.5 equivs. of 4-(hydroxyethyl)-5-nitrobenzoic acid (**26**) was activated as the HOBt ester using equimolar quantities of DCC and HOBt. The solution containing the active ester was added to a suspension of the NH_2-Val-MBHA resin. The reaction was allowed to

proceed for 60 min after which time ninhydrin analysis showed complete coupling had been achieved. To the resin-handle which had been swelling in 50% DMF/DCM, Boc-Gly-OH was added in ten fold excess together with an equivalent amount of DCC and a catalytic quantity (0.2 equivs.) of DMAP. After 2h the reaction was interrupted and amino acid analysis indicated that 99% coupling of the first amino acid had been achieved to give a resin with a substitution value of 0.5mmol/g (Table 1). From this point of view the 4-(hydroxyethyl)-5-nitrobenzyl-MBHA handle (**23**) gave better results compared to the 9-(hydroxymethyl)fluorene-4-carboxyl group (**22**).

The synthesis of the two protected peptide fragments was performed with the same amino acid derivatives and coupling protocols as used for handle (**21**). At the end of the synthesis an amino acid analysis of the peptidyl-resin indicated that 94% fragment 1-15 and 83% (fragment 16-32) of the theoretical yield was present (Table 1). Thus, no loss of peptide had occured through premature cleavage of the peptide-resin linker or through diketopiperizine formation.

While the introduction of the first amino acid was more successful for **23**, its cleavage kinetics were less favorable than in the case of the previous base-labile handle (**22**). Amino acid analysis of the residual peptidyl-resin following treatment with 20% piperidine in DMF for 30 min indicated that 95% (fragment 1-15) and 58% fragment 16-32 still remained attached to the resin (Table 1). After 6h treatment the percentage of peptide still left on the resin was 17% and 12%, respectively.

RP-HPLC analysis of the 1-15 crude cleavage product gave two main peaks at 17 and 28 min (labelled 1 and 2, Fig. 5A), which corresponded to the peaks labelled 1 and 3 (Fig. 3A) from handle **22**. The main difference however was that the two peaks labelled 2 (Fig. 3A) were not present after handle **23** was treated for 30 min with base.

Figure 5. C_{18} RP-HPLC of crude protected peptides from handle (**23**) (System 11)
A) Fragment 1-15
B) Fragment 16-32

Amino acid analysis of the two peaks from handle **23** confirmed that peak 1 (Fig. 5A) was non-peptidic, while peak 2 possessed the correct amino acid content (Table 3). However, the latter did not give a mass corresponding to the expected molecular ion ($[M+H]^+$ Expected: 2829; Found: 2404). In fact the difference recorded (425 mass units) was the same as that obtained for peak labelled 3 (Fig. 3A) less 23 (possible Na^+) from the synthesis of fragment 1-15 on handle **22**. The molecular ion obtained was therefore sequence 1-15 with the loss of Br-Z (mass=214) from its two tyrosine residues, due to piperidine treatment, as stated by Kuroda *et al*, [1992]. Furthermore, the additional degradation, which resulted in the appearance of several peaks at approx. 23 min (labelled 2, Fig. 3A), observed previously also occured with handle **23**, albeit at a slower rate (data not shown). Therefore, the degradation of 1-15 appeared to be related to an instability of certain benzyl-based protecting groups to piperidine and not to the handle, which however appeared to influence the rate of further degradation.

Fragment 16-32 was also injected on reversed-phase media (Fig. 5B) and gave the desired product as determined by FAB-MS ($[M+H]^+$ Expected: 2812; Found: 2813) and amino acid analysis (Table 3). Furthermore, the sharper peak observed compared to that obtained from the previous synthesis on handle **22** would appear to reflect an increased purity.

The following summarises the results obtained for the Boc synthesis of protected peptide fragments. The photolabile handle (**21**) was used successfully to synthesise the two GRF protected peptide fragments, although poor incorporation of the first amino acid was experienced. A potential problem reported with the use of handle (**21**) was diketopiperazine formation [Barany and Albericio, 1985; Giralt *et al*, 1981], which was not surprising due to the electron-withdrawing effect of the 3-nitro substituent and the consequent increased susceptibility of the resin-peptide ester linkage toward nucleophiles (eg. free amino group following N-terminal deprotection). This unwanted side reaction was overcome by applying the method reported by Suzuki and Endo

[1977], where the third amino acid was coupled manually. The principle reasons why resin handle **21** was not thought suitable (based on the experience gained with the protected peptide GRF fragments 1-15 and 16-32) for the routine synthesis of protected peptides were (i) the presence of many degradation products as indicated by the complexity of the HPLC analysis compared with same peptide when synthesised on other resin-linkers and (ii) the need for special equipment to perform the photolytic cleavage.

The 9-(hydroxymethyl)fluorene-4-carboxylic acid resin handle (**22**) did not require additional manipulation to avoid the DKP side reaction, but suffered from a poorer derivatisation of the first amino acid; only 45%. The synthesis of the 15 and 17 residue protected peptides proceeded well. However, cleavage of the peptide from the resin support with base was found to be problematic possibly due to sensitivity of the certain benzyl based protecting groups to base. A repeat synthesis of (1-15) on resin-handle (**23**), followed by piperidine treatment also yielded degradation products which were not observed for fragment (16-32). Furthermore, it was observed that the poor cleavage kinetics of the resin-peptide bond in handle (**23**), which required longer base treatment, lead to additional degradation.

PURIFICATION OF BOC SYNTHESISED PROTECTED PEPTIDES

Although the correct protected fragment (1-15) could not be obtained at this stage, a preliminary evaluation of chromatographic methods could be performed on fragment (16-32) from both base-labile handles. As previously described, silica gel as a separation media was found not to offer the resolving power needed for these large protected fragments and therefore reversed-phase chromatography was performed.

Purification by reversed-phase MPLC

Previous work towards the purification of short protected peptides on RP media has focused on a medium pressure system (RP-MPLC) [Lloyd-Williams *et al*, 1991]. The solvent systems used were based on DMF, which helped maintain the protected fragments in solution. The advantages offered by MPLC were that large columns (eg. C_8 Lichroprep column (440x37mm)) could be used to purify large quantities of material. The major disadvantages of using DMF were the necessity to extensively purify large volumes of solvent to remove amines and that it was not possible to monitor the eluting peaks at 220nm due to absorbance by DMF. Peptides that did not contain aromatic residues, for example GRF fragment (16-32) could not be monitored at 254 or 270nm. Therefore it was necessary to analyse the fractions individually after purification, by analytical RP-HPLC using a standard water/acetonitrile based system.

To determine the effectivness of MPLC, fragment (16-32) (from a synthesis on handle (**23**)) was dissolved in DMF and loaded onto the C_8 Lichroprep column (440x37mm), equilibrated in solvent A (DMF/water/propionic acid; 6:4:0.5). The eluting solvent (DMF/AcCN/propionic acid; 6:4:0.5) was then applied in gradient fashion. 5ml fractions were collected and analysed by RP-HPLC, using a standard water/acetonitrile based system. The fractions containing the desired product were combined, the DMF removed under vacuum and the residual solvent lyophilised to a glassy, yellow solid.

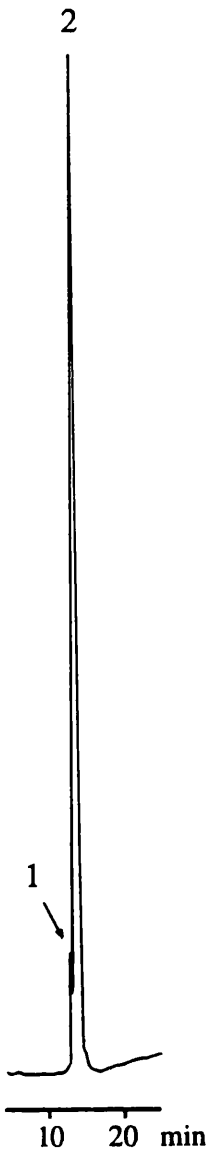


Figure 6. C_8 RP-HPLC of fragment 16-32 from handle **23**, purified by RP-MPLC. (System 13)

Although this method was able to purify fragment 16-32, which appeared as two peaks due to oxidised and reduced forms of Met (labelled 1 and 2, Fig. 6), the yield of purified material was only 10% (Table 4). An additional problem was the lack of solubility of the protected peptide in its dry, glassy, purified form in solvents (eg. DMF, DMSO, NMP etc) generally used for the coupling reactions.

Residue	Expected 16-32	Reversed-phase MPLC (% recovery)	Reversed-phase HPLC (% recovery)
Ala	1	1.3	0.9
Asp	1	1.2	0.9
Ile	1	1.4	1.0
Arg	2	1.7	2.4
Lys	1	1.0	1.0
Leu	3	3.0	3.4
Gly	1	1.3	0.9
Glu/Gln	4	3.9	4.2
Met	1	0.5	0.8
		(10%)	*(4%)*

Table 4. Amino acid analysis and yields for fragment 16-32 after purification

Purification by reversed-phase HPLC

Since in separate experiments it had been noticed that lyophilisation of water/AcCN solutions of purified 16-32 peptide fragment (from separation of analytical quantities of material), yielded a fluffy white material, which proved to be more soluble in the above mentioned solvents, it was decided to focus efforts on refining the aqueous based reversed-phase method. Furthermore, RP-HPLC should offer increased resolution over silica-based chromatography due to the peptides extremely hydrophobic nature. However application of RP-HPLC on a larger scale meant adapting the chromatographic conditions due to the poor solubility of protected peptide fragment (16-32) in solvent mixtures containing high proportions of water. After several attempts an example of the conditions chosen for semi-prep RP-HPLC were as follow. 100mg (22μmol) of fragment 16-32 peptidyl-resin which had been synthesised on **23** was cleaved using 20% piperidine in DMF and loaded onto the semi-prep C_8 RP column (100x10mm), previously conditioned in 40% AcCN. Fractions through the two main peaks were collected (labelled 1 and 2 in Fig. 7) and analysed by analytical RP-HPLC. Those containing the same material were combined and lyophilised

to a white powder. The two peaks were found to be the oxidised and reduced forms of fragment 16-32. The main product labelled 2 gave a yield of only 2.5mg (4%; Table 4). Further investigations to locate the remaining product showed that a large proportion of protected peptide had precipitated on the column despite using the initial concentration of 40% AcCN. To improve the yields, successive purifications were carried out using the same water/AcCN system to which was added 10% DMF. Despite these precautions the yield of product remained almost equal to that obtained previously. Furthermore, rather than a white powdery product, the same clear, poorly soluble glass-like material discussed previously, was obtained after lyophilisation.

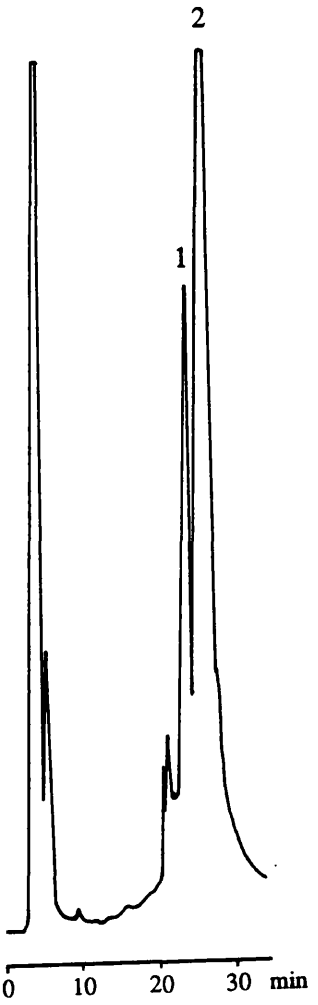


Figure 7. C_8 RP-HPLC of crude protected fragment 16-32 synthesised on handle (**23**) showing oxidised and reduced (labelled 1 and 2) forms. (System 14)

An alternative means of reducing loss of material through precipitation would be to load less protected peptide. Thus, the semi-preparative HPLC column was re-equilibriated with the aqueous based solvents and the loading was reduced to about a quarter (6μmol) of the original quantity. Using this approach precipitation was reduced and it was possible to obtain a powder product with yield of peptide of 1.5mg (10%), based on the peptide content of the resin. Following repeated injection, collection and analysis procedures 14mg (5μmol) fragment 16-32 (Fig. 8) from (**23**) was eventually collected. This material was then used to perform preliminary fragment condensation experiments.

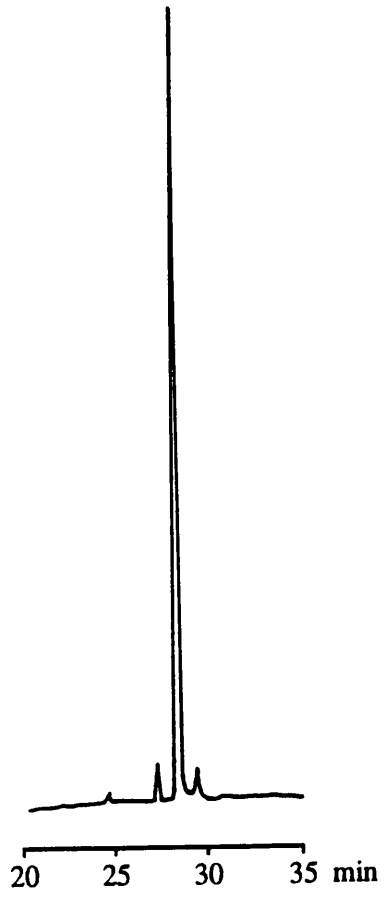


Figure 8. C_4 RP-HPLC of purified protected fragment 16-32 synthesised on handle (**23**). (System 11)

SOLID PHASE FRAGMENT CONDENSATION OF BOC SYNTHESISED PROTECTED PEPTIDES

The coupling of protected peptide fragments to resin-bound fragments has represented a critical step in the orthogonal approach to the synthesis of polypeptides. The important factors to consider are (i) concentration of nucleophile [Albericio *et al*, 1989], (ii) coupling reagent [Florsheimer and Riniker, 1990], (iii) reaction solvent and (iv) coupling time. The literature suggests that with increasing chain length, the efficiency of coupling protected peptides to a resin-bound peptide correspondingly decreases [Kaiser *et al*, 1989]. The explanation for this observation is that the movement of molecules within the polystyrene resin matrix is only due to translational movement and not a combination of translational and rotational effects, caused by the length of the peptide chain. A molecule that occupies a smaller space (eg. a single amino acid) is able to move with the aid of both effects and its ability to do so is reflected in superior coupling kinetics.

Although the rate of acylation reactions, like those described here, are highly dependent on the reactant concentrations, the limited solubility of the protected fragments prevented the use of the high concentrations that would provide the quantitative coupling desired. Grandas *et al* [1989] suggested that 1.6-2.8 equivs. of the peptide to be coupled be used depending on the length of the resin-bound peptide and to obtain quantitative coupling it was occasionally necessary to repeat a coupling.

Since the research was aimed at developing solid phase fragment condensation it was necessary to synthesise the C-terminal sequence 33-44, using stepwise Boc chemistry. The peptide was synthesised on pre-derivatised polystyrene based PAM resin which possessed the superior swelling properties [Giralt *et al*, 1984; Albericio *et al*, 1989] necessary for fragment condensation. Quantitative ninhydrin analysis indicated an average coupling yield of 99.6%. At the end of the synthesis the N-terminal Boc protecting group was left attached and the peptidyl-resin stored at -70°C in DMF.

14mg (5μmol) of the protected peptide 16-32, which represented 1.5 equivs. were dissolved in 200μl DMF together with 1.5 equivs. each of HOBt and benzotriazole-1-yl-oxy-tris-(dimethylamino)-phosphonium hexafluorophosphate (BOP). The final concentration of protected peptide in the reaction mixture was 0.025M. The solution was left to activate for 30 min at 0°C. The N-terminal Boc group was removed from the 33-44 peptidyl-resin with 50% TFA/DCM, then neutralised with a 5% DIEA in DCM and dried from DCM. 11mg (~3μmol) of peptidyl-resin was then rapidly weighed out and added to the reaction vessel containing the activated protected peptide and 1.8μl (10μmol) DIEA. After 24h mixing, the peptidyl-resin was washed with DMF, dried from DCM/MeOH (1:1) and a small aliquot taken for amino acid analysis. Taking Lys as a reference the calculated incorporation of fragment 16-32 was found to be 28% (Table 5). In an attempt to improve the coupling a second aliquot of peptide (1.5 equivs.) was reacted with fresh, deprotected peptidyl-resin, using the same conditions as described above, with the exception of coupling agent which was changed to DCC. However, also in this case the coupling yields (20%) were similar to that previously obtained. Since poor coupling did not appear to be due to the activating agent, then either the reactivity of protected fragment (16-32) was intrinsically poor or intermolecular β-sheet formation of the resin-bound chains was making the free amino groups inaccessible.

To understand whether intermolecular interactions, if they existed, were responsible for the poor yields a further experiment was performed in which under the same reaction conditions a single amino acid (ie. Boc-Gly) was coupled. Ninhydrin analysis showed that after 24h quantitative coupling had been achieved and therefore β-sheet formation of resin-bound peptide was either not occuring or if it did, was not responsible for the poor coupling yields.

Once it was established that the resin-bound peptide was capable of freely reacting, at least with Boc-Gly, attention was focussed on the 16-32 protected peptide and in particular on whether the latter aggregated in solution.

Fragments Coupled	Coupling Reagents	Time	% coupling as determined by AAA
16-32 → 33-44-® t-Boc 1.5eq	BOP/HOBt/DIEA 1eq 1eq 3eq	24h	28.1%
	DCC/HOBt/DIEA 1eq 1eq 3eq	48h	20.4%
1-15 → 16-44-® t-Boc 2eq	DCC/HSu/DIEA 1eq 1eq cat	24h	3.7%
		48h	12.9%
	BOP/HOBt/NMM 1eq 1eq 3eq	72h	20.2%

Table 5. Fragment condensation using Boc/Bzl protected peptides

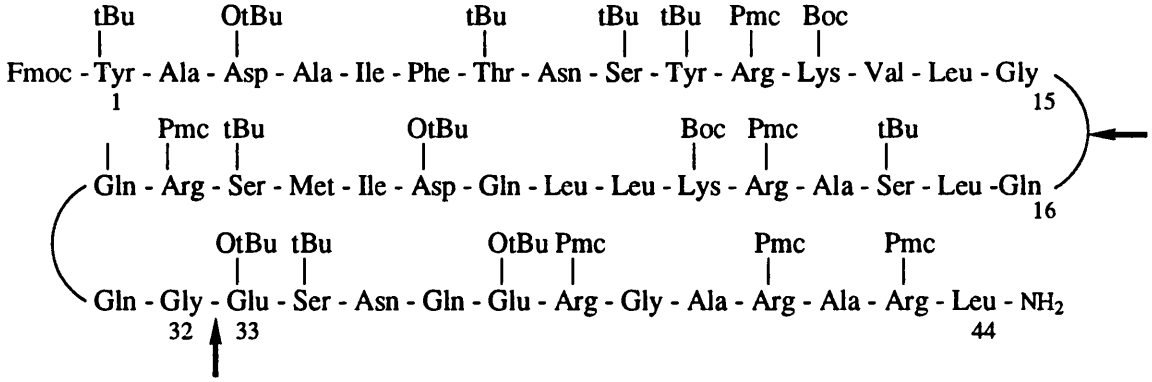
Hydrophobic protected peptide fragments are renouned for their tendancy to aggregate, which is often related to their sequence and length [Kaiser *et al*, 1989] and results in poor solubility. The situation is further aggravated by free asparagine and glutamine residues, whose side chains promote intermolecular aggregation (ie. β-sheet formation) [Atherton *et al*, 1990, Kamber and Riniker, 1991]. Consequently, the question was asked that if Asn/Gln amino acids were protected, would the reduced aggregation be reflected in improved solubility and thereby ease protected peptide purification and coupling?

The requirement for side chain protection on Asn/Gln residues was that it must be retained during the synthetic procedure and after cleavage from the solid support. Unfortunately, the only readily available protection for Boc amino acids is the xanthanyl group which, while maintained for the activation and coupling step (to avoid dehydration) is subsequently partly removed upon treatment with TFA, during the next cycle of the synthesis. Furthermore, since attempts to synthesise fragment 1-15 using Boc chemistry on the base-labile handles was unsuccessful, an alternative synthetic strategy was investigated, ie Fmoc chemistry.

FMOC CHEMISTRY

SYNTHESIS OF PROTECTED PEPTIDES USING FMOC CHEMISTRY

The cleavage conditions required for the generation of protected peptides from Fmoc resins enabled a new range of amino acid derivatives to be used. One of the advantages offered by Fmoc chemistry was a protecting group for Asn/Gln (ie. trityl) [Kamber and Riniker, 1991] that can be maintained throughout synthetic and cleavage procedures. However before attempting the syntheses of the GRF fragments using Asn(Trt)/Gln(Trt), it was first necessary to investigate whether or not the problems experienced with the Boc peptides (ie. solubility) could be overcome by merely changing to the alternative synthetic strategy and the consequent use of different protecting groups. Thus, GRF (1-15) was resynthesised on Sasrin resin (**24**) [Mergler *et al*, 1988].



Synthesis on acid sensitive Sasrin resin

The synthesis of fragment 1-15 was performed using HBTU/NMP chemistry. Amino acid analysis was performed on the peptidyl-resin (Table 6) which indicated that the synthesis had been successful. Qualitatively, the found to expected ratio were much in aggreement (except Tyr which had a lower than expected value), but only 63% peptide was found on the resin. The lower than

Residue	Expected			1-15-®		16-32-®	33-44+G-®
	1-15	16-32	33-44+G	24	24*	24*	24*
Tyr	2	-	-	0.9	1.8	-	-
Ala	2	1	2	2.0	2.0	0.9	2.0
Asp/Asn	2	1	1	1.9	2.1	1.0	1.0
Ile	1	1	-	0.7	0.9	0.8	-
Phe	1	-	-	0.9	0.9	-	-
Arg	1	2	3	1.1	1.0	2.3	2.9
Lys	1	1	-	1.1	1.1	1.0	-
Val	1	-	-	1.0	1.0	-	-
Leu	1	3	1	1.0	1.0	3.0	1.0
Gly	1	1	2	1.0	1.1	1.0	1.9
Glu/Gln	-	4	3	-	-	3.9	2.9
Met	-	1	-	-	-	0.5	-

Table 6. Amino acid analysis of peptidyl-resins following synthesis of the protected peptides by Fmoc strategy. * Asn/Gln(Trt)

expected peptide content could have resulted from premature cleavage of the peptide-resin bond by acidic components in the reaction mixture (eg. HOBt and BF_3 from HBTU). Loss of peptide has been reported for Rink resin (an alternative resin for the synthesis of protected peptides), due to sensitivity to mildly acidic HOBt and was overcome with the addition of 1% DIEA to the HOBt stock solution [Rink, 1987]. Synthesis of fragment 16-32 was not conducted at this stage until some preliminary information became available as to the behaviour of a protected peptide (ie. fragment 1-15), in terms of premature cleavage from the resin, the chromatographic properties of the material and its solubility.

The protected peptide 1-15 fragment was cleaved with two treatments of 1% TFA in DCM. The cleavage solution was neutralised with 200µl DMF and the DCM finally evaporated off. The resulting crude product was also poorly soluble when attempts were made to dissolve it in organic solvents (eg. DMF, NMP and DMSO). The poor solubility properties observed were similar to those that had been synthesised using the Boc approach. Amino acid analysis of the resin following cleavage indicated that only 3% peptide still remained

attached to the solid support. The DMF solution containing the protected peptide was then analysed on reversed-phase media (Fig. 9) and used with the aim of developing a batch purification procedure.

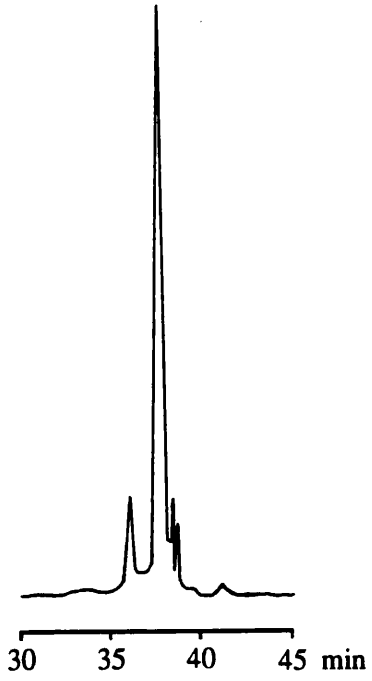


Figure 9. C_4 RP-HPLC of crude protected fragment 1-15 synthesised on handle (**24**). (System 12).

PURIFICATION OF FMOC SYNTHESISED PROTECTED PEPTIDE FRAGMENTS

Gel filtration

As yet it had not been possible to purify large quantities of material without losing resolution or solubility after lyophilisation. The latter seemed to be particularly severe when the material was in a glassy form. Thus to improve both the ability to perform batch purification on a large scale and solubility of the final product, gel filtration chromatography [Pedroso *et al*, 1982; Giralt *et al*, 1982; Kneib-Cordonier *et al*, 1990] on a solvent mixture containing MeOH/DMF was used. Gel filtration permitted the quick and inexpensive preparation of columns for high loading, DMF solubilised the crude peptide and MeOH, it was hoped, would improve the physical properties of the final product. Although the resolving power of gel filtration is limited it was felt that due to the relatively high purity of the crude peptide, its purification by gel filtration could be accomplished. In particular it was felt that if the method were successful, its routine application on a laboratory scale would be feasible as long as the crude products were already highly homogeneous.

100mg fragment 1-15 attached to Sasrin resin was cleaved with 1% TFA/DCM and the combined supernatant reduced in volume on a rotary evaporator, to leave the DMF solution containing the cleaved protected peptide. The LH20 gel filtration media was loaded into a column as a slurry in DMF/MeOH (1:1) and then washed with 5 column volumes of solvent. The DMF solution containing the protected peptide was loaded onto the top of the column and positive pressure (8 psi) was applied. The eluting solvent, DMF/MeOH (1:1) was added at which point a gel developed at the top of the column, probably due to insolubility of the crude protected peptide. The purification was continued and fractions of approx. 3ml were collected, their absorbance at 280nm recorded and those containing similar material combined. The fractions containing the highest concentration of material eluted in approx. 45ml solvent, however high absorbance readings continued to be recorded in the following 45ml of solvent. This observation was consistent with the protected

peptide precipitating on the column and gradually redissolving on the addition of further solvent. The combined fractions were lyophilised to a poorly soluble solid and the yield was calculated to be 34%, based on the peptide content of the resin. An aliquot of the protected peptide thus purified was dissolved in DMF and injected on analytical reversed-phase media to give a tailing peak centered at 37 min (Fig. 10). The retention time was the same as that obtained for the crude peptide shown in figure 9, thus indicating that purification had been achieved. However, the apparent incompatibility between the solvent and the protected peptide, as indicated by the gelling on the column, the observed elution characteristics and the physical properties of the final product prevented the batch purification of fragment 1-15 by gel filtration and therefore further analysis and characterisation was not performed.

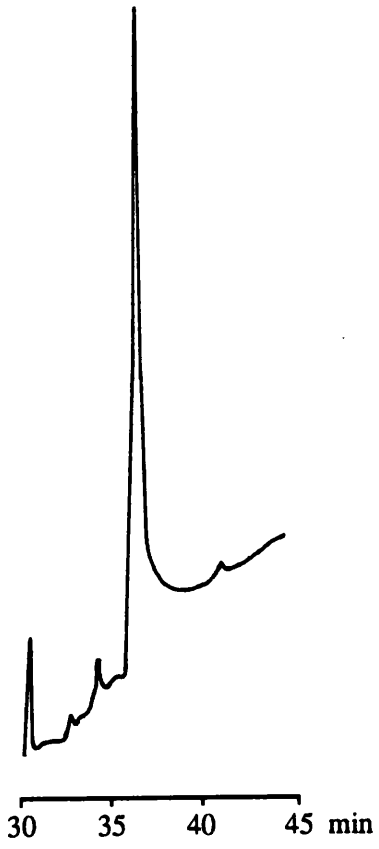


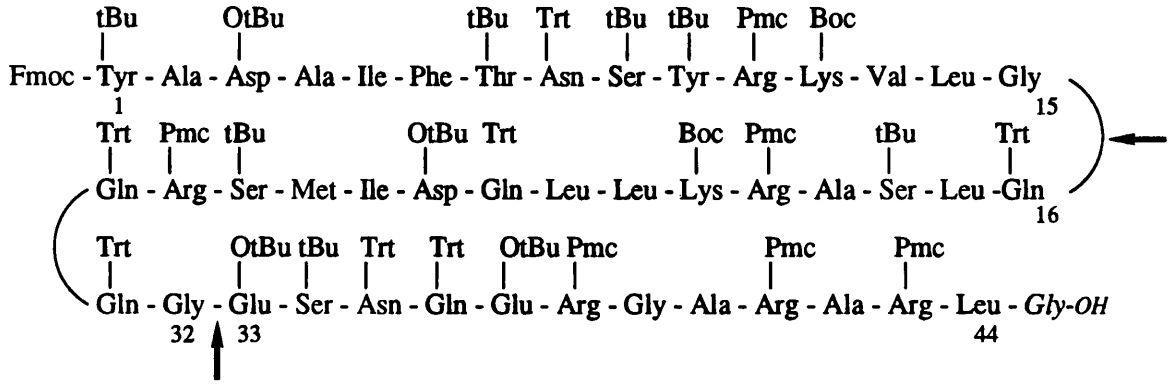
Figure. 10. C_8 RP-HPLC of protected fragment 1-15 synthesised on handle (**24**) and purified by gel filtration. (System 15)

The previous synthesises of fragments 1-15 and 16-32 using either Boc or Fmoc chemistry had yielded poorly soluble products. Some dissolution was achieved at low concentrations in DMF, however this was restricted to protected peptides in their crude state. The addition of more polar solvents (eg. MeOH, water etc) immediately caused the protected peptide to gel with the consequent effect that the yields from purification procedures were low (eg. fragment 16-32 was recovered with 4% yield from RP-HPLC, as described on page 115). The results obtained appeared to indicate that a physico-chemical interaction was present which influenced the physical characteristics of the protected peptides. For example, intermolecular interactions (eg. β-sheet formation) which have been reported for peptides containing unprotected Asn/Gln residues could cause aggregation, resulting in poor solubility. Since the GRF(1-44) sequence possesses eight such amino acids, their contribution to the insolubility of fragments 1-15 and 16-32 could be significant. Thus the use of the trityl group for the protection of Asn/Gln residues and Fmoc chemistry were evaluated in combination with Sasrin resin.

SYNTHESIS OF PEPTIDE FRAGMENTS USING FMOC CHEMISTRY (ASN/GLN PROTECTED)

Synthesis on acid-sensitive Sasrin resin

The peptide sequences 1-15 and 16-32 were synthesised in a similar manner as described previously, using HBTU/HOBt chemistry. A complete protection scheme was used as indicated in the sequence shown below.



At the end of the syntheses, amino acid analysis of the peptidyl-resins showed good accordance with the expected amino acid composition and that 67% (fragment 1-15) and 86% (fragment 16-32) were obtained based on the theoretical amount of peptide present (Table 6).

Small aliquots of the two protected peptide fragments 1-15 and 16-32 were cleaved from the solid support, using 1% TFA/DCM and injected on an analytical RP-HPLC column (Fig. 11A and 11B). The single peaks observed suggested that the crude peptide possessed good homogeneity. Furthermore, in accordance with previous work with peptides containing trityl protected Asn/Gln residues, solubility of the crude peptides were found to be far greater in several organic solvents independent of the sequence or amino acid composition. Both GRF fragments 1-15 and 16-32 dissolved readily in DMF, DCM and even AcCN containing up to 50% water. The ability to achieve high concentrations in the solvents used for fragment condensation was extremely important if the coupling efficiency was to be maximised. In addition, since the solubility of the crude peptides thus synthesised seemed no longer to be a

problem, several different procedures were examined in an attempt to purify them on a relatively larger scale.

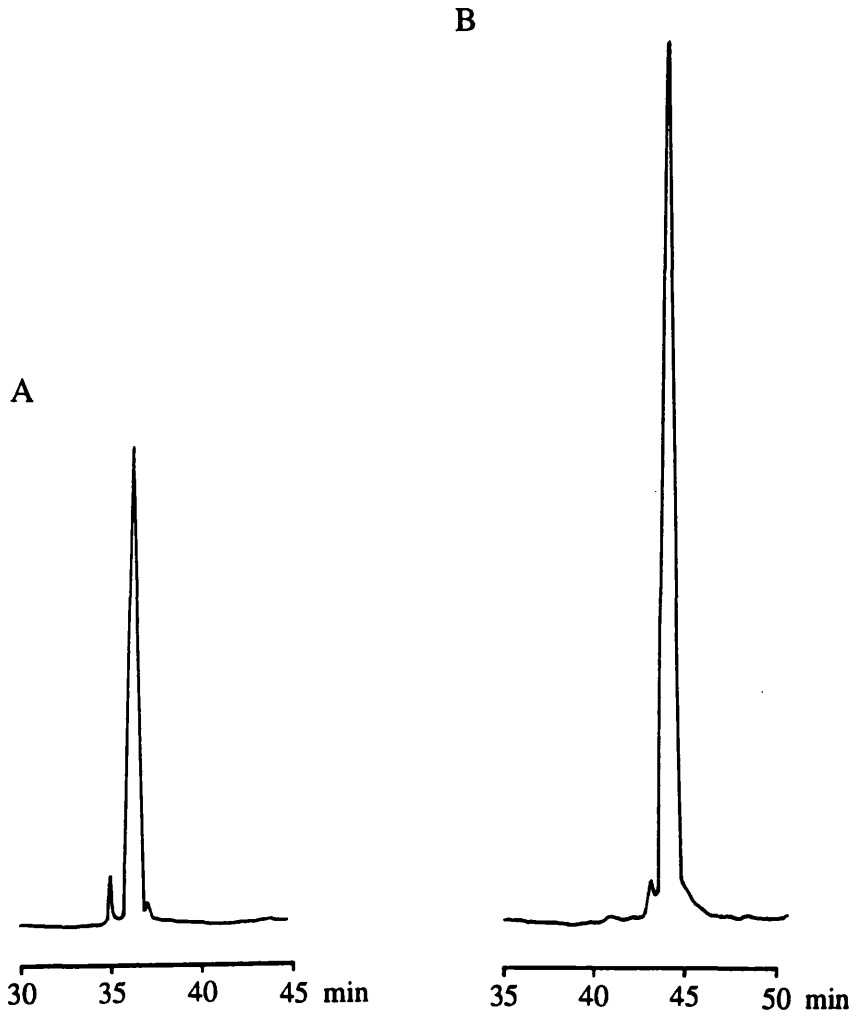


Figure 11. C_8 RP-HPLC of crude protected fragments: Asn/Gln protected (System 15)
A) Fragment 1-15
B) Fragment 16-32

PURIFICATION OF FMOC SYNTHESISED PEPTIDE FRAGMENTS (ASN/GLN PROTECTED)

Purification by solvent extraction

The first method investigated, was a simple solvent extraction technique. 90mg peptidyl-resin (fragment 16-32) was treated with 1% TFA in DCM and the organic solvent was shaken with water and brine, before drying over anhydrous sodium sulphate. The DCM was removed under vacuum to give a yield of 47mg (73%), based on the theoretical peptidyl-resin substitution, of a glassy pale yellow product. Fragment 1-15 was treated in a similar manner and yielded 42mg (66%) material after removal of the DCM solvent. Amino acid analysis of both products corresponded well with the expected values (Table 7). Mass spectroscopy of the fragments 1-15 (Expected: 2828; Found: M^++Na: 2851) and 16-32 (Expected: 3962; Found: M^++Na: 3986) showed that the correct material had been synthesised and purified. In the case of fragment 1-15 the mass analysis result proved for the first time that, using Fmoc chemistry, the correct protected fragment had been synthesised and recovered with all protecting groups intact following removal from the resin support.

Solubility, at high concentration in various solvents (eg DMF, DCM, NMP and DMSO) was easily achieved with no apparent gelling that had afflicted previous peptides synthesised without protection on Asn/Gln residues. Analytical quantities of fragments 1-15 and 16-32, following solvent extraction were injected onto reversed-phase media (Fig 12A and 12B). The resulting chromatograms indicated a good degree of purity. While solvent extraction did not represent a true chromatographic method of purification, it was useful for preparing large quantities of highly soluble material from homogeneous crude peptides.

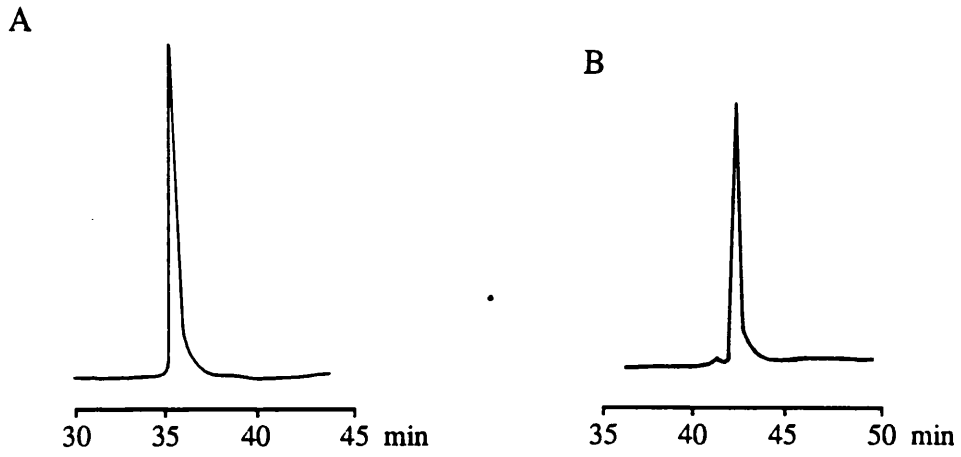


Figure 12. C_8 RP-HPLC of protected peptides after solvent extraction. (System 15)
A) Fragment 1-15
B) Fragment 16-32

Residue	Expected		Fragment 1-15	Fragment 16-32		
	1-15	16-32	Solvent Extraction *(% recovery)*	Solvent Extraction	RP-HPLC *(% recovery)*	Poros
Tyr	2	-	2.0	-	-	-
Ala	2	1	2.0	0.9	0.9	1.0
Asp	2	1	2.0	0.9	0.9	1.0
Ile	1	1	0.8	1.0	1.0	0.7
Phe	1	-	1.0	-	-	-
Arg	1	2	1.0	2.4	2.4	1.7
Lys	1	1	1.0	1.0	1.0	1.0
Val	1	-	1.0	-	-	-
Leu	1	3	1.0	3.4	3.4	2.9
Gly	1	1	1.0	0.9	0.9	1.0
Glu	-	4	-	4.2	4.2	3.9
Met	-	1	-	0.8	0.8	0.4
			(66%)	*(73%)*	*(21%)*	*(53%)*

Table 7. Amino acid analysis of purified protected peptides from Sasrin resin, containing Asn/Gln(Trt) residues.

Purification by reversed-phase HPLC

The second method investigated was semi-preparative RP-HPLC using a water/acetonitrile based solvent system. It was envisaged that the improved solubility imparted to peptide fragments as a result of protecting the Asn/Gln side chains would allow the method to be used for batch purification. Thus 50mg each of fragment 16-32 (9.6μmol) peptidyl-resins was cleaved with 1% TFA/DCM. The acid solution containing the protected peptide was neutralised with DMF and the volatile solvent removed under vacuum. The residual DMF solution was loaded onto the semi-preparative reversed-phase column (100x10mm). Fractions were collected, combined and lyophilised to dryness yielding 8mg (21%) of pure fragment 16-32. Amino acid (Table 7) and mass spectrometry analysis (Expected: 3962; Found: M^++Na: 3986) confirmed that the correct material had been purified. It therefore appeared that the use of Asn/Gln derivatives with protected side chains had improved the solubility of the protected peptide and thus resulted in a higher recovery (ie. from 4% to 21%).

Although conventional RP-HPLC described here was useful for purifying a 17 residue fully protected peptide to give a highly soluble product, the same problems that were described for RP-HPLC purification of Boc synthesised peptides (ie. precipitation) also applied to the Fmoc protected peptides, albeit to a lesser extent. The precipitation of material on the column was confirmed when a blank run was performed immediately following the purification. The resulting chromatogram (data not shown) was identical to that obtained for the previous run, although less intense. Thus the amount of material that could be purified in one run was restricted to the limited loading capacity of the semipreparative column. Since large quantities of peptide were required for coupling experiments repeated injections were necessary and conventional RP-HPLC required relatively long gradient times to achieve resolution. If the quantity of material that could be loaded was restricted the answer would be to reduce the gradient time, so that for a given period a larger amount of material could be processed. Perfusion chromatography was thus an obvious method to

investigate with its reasonable loading capacity and fast gradient times due to high flow rates.

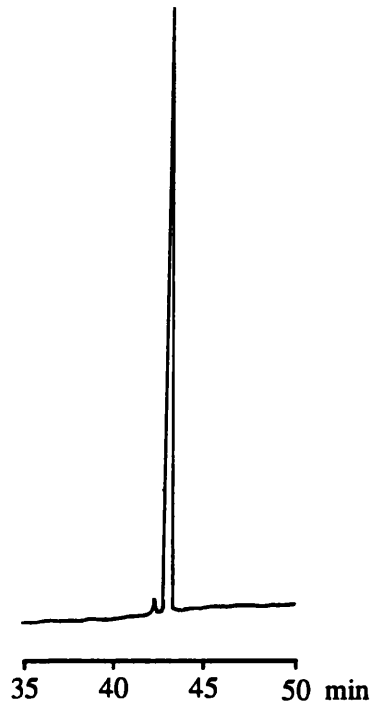


Figure 13. C_8 RP-HPLC of protected peptide fragment 16-32 after purification on conventional reversed-phase column. (System 15)

Purification by perfusion chromatography

The basis of separation (ie. hydrophobic interaction with the stationary phase) is the same as conventional RP-HPLC, but unlike conventional "diffusion-based" chromatographic media, where molecules arrive at the inner surface of the pores by the relatively slow process of molecular diffusion, the perfusion material combines large flow-through pores for mass transport of molecules with smaller "diffusive pores" possessing short pathlengths for increased resolution [Afeyan *et al*, 1991]. These properties could be important for relatively bulky molecules, like protected peptides, that may have restrictive diffusion characteristics resulting in poor resolution on conventional reversed-

phase media. Furthermore, the application of perfusion chromatography with its extremely high flow rates was thought to reduce the risk of precipitation on the column, by reducing the opportunity for intermolecular peptide-peptide interaction in favour of peptide-matrix interaction. In addition the shorter gradient times required to give the same degree of resolution obtained on a conventional matrix would enable more material to be prepared in a given time.

90mg of fragment 16-32 peptidyl-Sasrin resin containing a theoretical amount of 68mg peptide was treated with several aliquots of 1% TFA in DCM which were then dropped onto 200µl DMF and the DCM finally evaporated off under vacuum. The residual DMF solution containing the protected peptide was injected directly onto the semi-preparative perfusion column (100x7.5mm). A flow rate of 6ml/min was used with the application of an elution gradient from 50% to 100% AcCN in 20 min. The fractions collected were analysed on a conventional reversed-phase column, combined and lyophilised to a white powder with a yield of 36mg (53%). The material was 98% pure from HPLC (Fig. 14) and gave the expected amino acid analysis (Table 7), indicating that the expected homogeneity had been achieved and that the correct material had been isolated.

It had been noticed previously that following purification on RP-HPLC media using TFA buffered solvents, the lyophilised product still contained traces of acid. Therefore, since traces of acid could affect the fragment condensation reaction, it was necessary to either (i) extract the acid in water or (ii) use solvents that were not buffered. The latter approach was adopted with the consequent loss of some resolution.

A blank run following the purification demonstrated that no precipitation had occured unlike on the ordinary reversed-phase media. Furthermore, like the material recovered from the standard reversed-phase system the white powder product obtained had a similar solubility profile to the protected peptides purified by the standard reversed-phase methodology, with the major difference that in the latter case only 21% recovery was achieved, as opposed to 53% from perfusion chromatography.

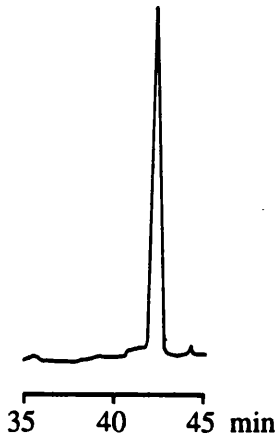


Figure 14. C_8 RP-HPLC of protected peptide fragment 16-32 after purification on reversed-phase perfusion media. (System 15)

SOLID PHASE FRAGMENT CONDENSATION OF FMOC SYNTHESISED PEPTIDES (ASN/GLN PROTECTED)

Batch purification of sufficient quantities of fragments 1-15 and 16-32 was achieved by the perfusion chromatography technique. The next step was hence to attempt fragment condensation using the fully protected GRF segments. In the previous fragment coupling experiment using Boc/Bzl protected peptides, only approx. 20% incorporation was achieved. The poor coupling was attributed to the physical properties of the protected peptide rather than β-sheet formation between peptide chains of the resin-bound 33-44 fragment, since a single amino acid was coupled without difficulty. The same results were not necessarily true when the side chain groups of the C-terminal fragment are protected with t-butyl type groups. Therefore, sequence 33-44 was resynthesised using Fmoc chemistry with trityl protecting groups on the side chains of Asn and Gln. Furthermore, a doubt still remained concerning the sensitivity of some Boc amino acid side chain protecting groups to piperidine, which would be required to remove the Fmoc group from the N-terminus of the middle fragment (16-32), before coupling the 1-15 sequence. Therefore it was felt prudent to avoid possible complications and to synthesise the C-terminal resin-bound fragment by stepwise SPPS, using Fmoc chemistry.

Stepwise SPPS of C-terminal fragment 33-44+Gly

The synthesis of fragment 33-44 was performed on Sasrin resin which was pre-derivatised with Fmoc-Gly. HBTU/NMP chemistry protocol, using single coupling cycles, except for Arg which was double coupled. No loss of peptide through premature cleavage was found as determined by amino acid analysis (91% based on the theoretical amount of peptide present).

Coupling an amino acid to GRF(33-44)+Gly peptidyl-resin

The first step involved coupling a single amino acid (eg. Fmoc-Lys(Boc)-OH) to prove, as in the coupling of Boc-Gly-OH to 33-44-PAM resin, that the free amino group was available for coupling following N-terminal deprotection. 15mg (substitution value 0.27mmol/g; 4μmol) of (33-44)+Gly peptidyl-resin was treated with 20% piperidine in DMF. A two fold excess of activated Fmoc-Lys(Boc)-OH was prepared using 1 equiv. each of DCC/HOBt in DMSO/NMP. DMSO was shown by Barlos *et al* [1990] to be a good solubilising solvent for protected peptide fragments. Due to the highly hygroscopic nature of DMSO/NMP it was necessary to dry the solvent over molecular sieves prior to use. The solution containing the activated amino acid, had a final concentration of 0.07M. After 24h a sample of resin was removed and the extent of Lys incorporation determined by amino acid analysis. The coupling yield was found to be 96% and therefore the amino group of the peptidyl-resin was free to react and not hindered by, for instance, β-sheet formation.

Coupling protected peptide to resin-bound amino acid

To evaluate the feasibility of activating the larger protected fragments and their ability to acylate, a second experiment was conducted which involved coupling fragment 16-32 to a resin with a single amino acid attached (ie. Fmoc-Gly-Sasrin). The N-terminal protecting group was removed with 20% piperidine in DMF, dried from DCM/MeOH (1:1) and then quickly weighed to avoid any unwanted rearrangements at the free amino group. 7.3mg (4.74μmol) deprotected resin was transferred to a 300μl micro reaction vessel and allowed to swell in 50μl DMSO/NMP. 2 equivs. of purified fragment 16-32 (37mg, 9.3μmol) was activated with 2 equivs. each of DCC and HOBt in 50μl DMSO/NMP. The solution was added to the peptidyl-resin and left to react at room temperature with occasional vortexing. The final concentration of fragment 16-32 was 0.07M. After 72h an aliquot of resin was removed for amino acid analysis. The analysis results indicated that 44% incorporation of fragment 16-32 had occured. To improve the fragment condensation an

additional 1 equiv. DCC/HOBt was added, but did not result in further coupling after 24h. The resin was washed with DMSO/NMP to remove excess reagents and a further 2 equivs. of protected fragment 16-32, DCC and HOBt were added. After another 72h, amino acid analysis showed that the extent of coupling had increased to 79%.

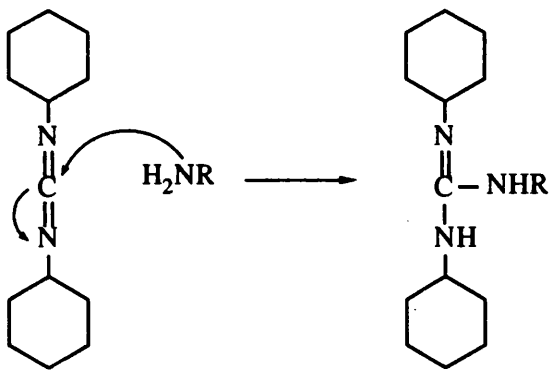
Coupling protected peptide to GRF(33-44)+Gly peptidyl-resin

Once it was established that both the resin-bound peptide and protected peptide were able to react, in the conditions described above, the third experiment involved the fragment condensation of purified fragment 16-32 onto a resin bound peptide, (33-44)+Gly-Sasrin. The reaction conditions were identical to those described for coupling to Gly-Sasrin, using a two fold excess of protected peptide, 2 equivs. DCC/HOBt in DMSO/NMP and at a concentration of 0.07M. After 72h amino acid analysis showed that 38% incorporation had been achieved and was increased to 63% after recoupling with fresh reagents.

The three experiments suggested that reactivity of activated protected fragment 16-32 was generally independent of chain length and mainly a matter of kinetics. Degradation of the protected peptide active ester, which may have occured due to the extended reaction time, may explain why more complete coupling was achieved with the coupling of Fmoc-Lys(Boc)-OH and that when fresh activated peptide was added further coupling was observed. The slightly worse coupling yields observed when coupling one peptide fragment to another, compared to a single amino acid, possibly could have been due to steric hindrance generated by the N-terminal residue (Glu^{33}(OtBu)).

SOLID PHASE FRAGMENT CONDENSATION OF GRF (1-44)+GLY AND ITS PURIFICATION

Once an efficient protocol had been established for fragment condensation (ie. activation with DCC/HOBt in DMSO/NMP at a concentration of 0.07M), the synthesis of the target peptide (1-44)+Gly was attempted. Since the coupling of fragment 16-32 to fragment (33-44)+Gly-Sasrin had apparently reached maximum incorporation of 63% after 6 days, then the same resin was used for the final coupling of fragment 1-15. However, before performing the last fragment condensation an aliquot of the resin was removed and fully cleaved with 95%TFA/5% water plus scavengers, to discover the true peptide composition on the resin. The crude RP-HPLC chromatogram (Fig. 15) showed four major products which were collected and analysed for amino acid composition (Table 8) and mass. The first peak labelled 1 (Fig. 15; 27%) corresponded to the underivatised peptide 33-44+Gly (FAB-MS $[M+H]^+$ Expected 1442; Found 1444). The second peak labelled 2 (13%) had the same amino acid composition as the first, less one Glu residue, but had a mass that was 206 higher than expected ($[M+H]^+$ Expected 1442; Found 1651). The analysis results for peak 2 were interpreted as representing a side product due to the reaction of the N-terminus with DCC to form a guanidine-type product. The formation of this condensation product has been described in the literature and its extent depends on the N-terminal amino acid [Barany and Merrifield, 1979].



Peaks labelled 3 (FAB-MS $[M+H]^+$ Expected 3396; Found 3399) and 4 (FAB-MS $[M+H]^+$ Expected 3396; Found 3398) were found to be the desired product containing oxidised and reduced Met and had a combined proportion of 60%, which corresponded well with the result from amino acid analysis (ie. 63%).

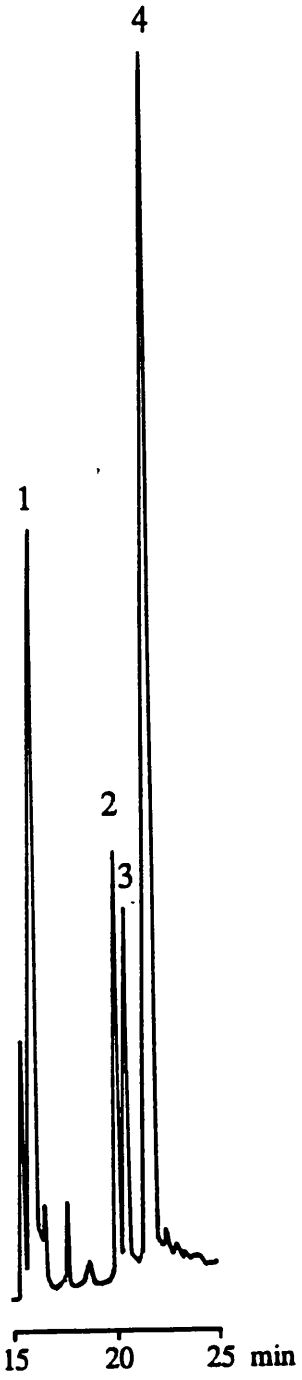


Figure 15. C_{18} RP-HPLC of crude fully deprotected GRF(16-44)+Gly. (System 16)

Before resuming the synthesis by adding the final fragment (1-15) and to prevent possible derivatisation to any unreacted amino groups from the previous coupling, a capping procedure was performed. This involved treating the 8mg of 16-44+Gly peptidyl-resin with 5% acetic anhydride in DMF, for 5 min, with 1% DIEA added to neutralise any acetic acid that could prematurely cleave the acid-sensitive Sasrin resin. After capping, the N-terminal Fmoc protection was removed with 20% piperidine/DMF in preparation for the incorporation of the final 15 residue fragment.

Following the same procedure as described for fragment 16-32, 2 equivs. of purified fragment 1-15 was activated with 2 equivs. each of DCC/HOBt in DMSO/NMP to give a final concentration in the reaction vessel of 0.07M. The reaction was allowed to proceed for 3 days with occasional vortexing, at which point an aliquot was removed for amino acid analysis. The result obtained indicated that 45% coupling of 1-15 to (16-44)+Gly-Sasrin resin had been achieved. The reagents were then removed, the peptidyl-resin washed and a further 2 equivs. fragment 1-15, HOBt and DCC in dry DMSO/NMP added to the reaction vessel. After an additional 3 days the amino acid analysis indicated that the incorporation of N-terminal fragment had increased to 68%. However, the latter result did in fact represent a coupling of 100% because 37% was the percentage of 33-44+Gly peptidyl-resin that previously had not reacted with fragment 16-32 and had been shown to be acetylated with acetic anhydride or guanidated by DCC. Given the margin of error for amino acid analysis of resin-bound material, the actual coupling of the last series of residues (1-15), in the sequence of GRF(1-44)+Gly, was hence estimated to be quantitative.

A full cleavage reaction was performed, using 95% TFA in the presence of scavengers. When injected on C_{18} RP-HPLC media a series of peaks eluted (labelled 1 to 6, Fig. 16). Each peak was collected and analysed for amino acid content and by FAB-MS with the exception of the cluster of peaks labelled 3 and 4 which were collected together.

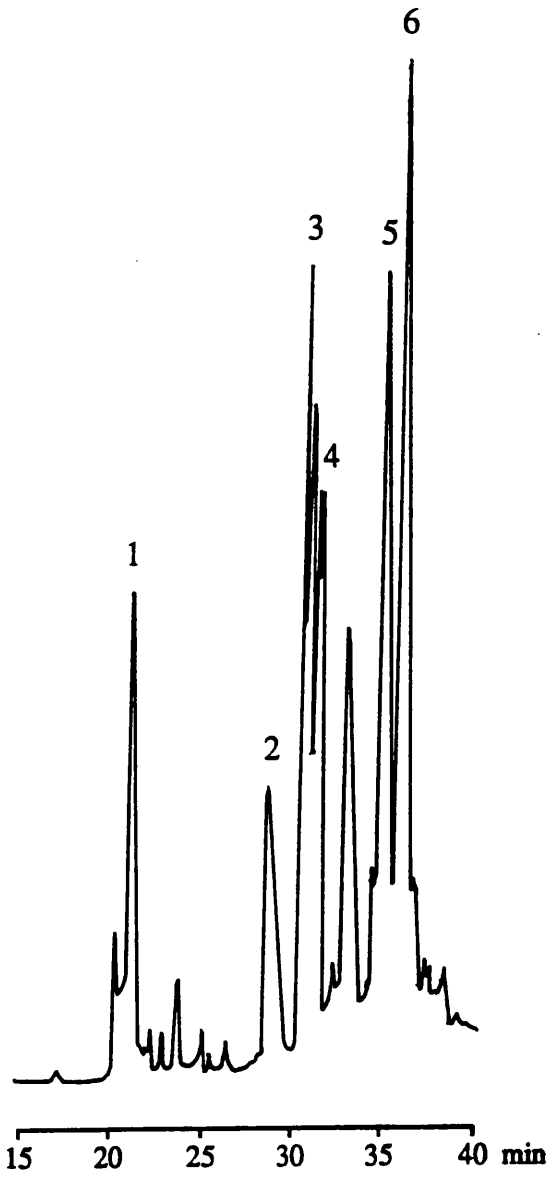


Figure 16. C_{18} RP-HPLC of crude fully deprotected GRF(1-44)+Gly. (System 7)

On the basis of the analyses performed and FAB-MS data, the crude mixture was made up of the following components: The peak labelled 1 (16%) (Fig. 16) was found to be the acetylated C-terminal fragment 33-44+Gly ($[M+H]^+$: Expected 1485; Found 1486). The peak labelled 2 (12%), which was also observed in the previous reaction, probably represented the guanidated side

product that would be unable to react further ($[M+H]^+$:Expected 1648; Found 1650). The peaks labelled 3 ($[M+H]^+$:Expected 3397; Found 3400) and 4 ($[M+H]^+$:Expected 3397; Found 3399), which together accounted for 22% were found to contain the intermediate sequence, 16-44+Gly as the oxidised and reduced forms of Met^{27}. Finally, peaks 5 ($[M+H]^+$:Expected 5097; Found 5115) and 6 ($[M+H]^+$:Expected 5097; Found 5099), together accounting for 50%, were found to be the full 1-44+Gly peptide (Met, oxidised and reduced).

Peak 6 was collected and found to be a single component when reinjected on an analytical RP-HPLC column (Fig. 17), with the correct amino acid composition (Table 8).

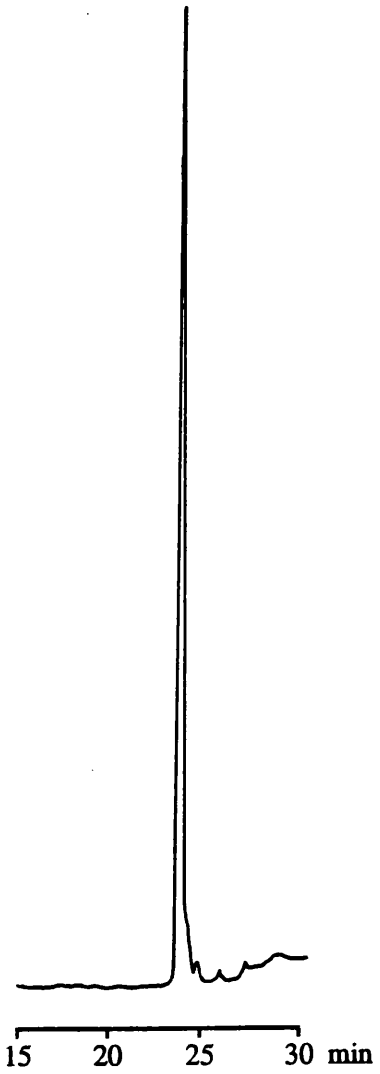


Figure 17. C_{18} RP-HPLC of purified GRF(1-44)+Gly. (System 7)

Residue	Expected			Peptide 16-44+G (Fig. 15)				Peptide 1-44+G (Fig. 16)	
				Peak No.				Peak No.	
	33-44+G	16-44+G	1-44+G	1	2	3	4	1	6
Tyr	-	-	2	-	-	-	-	-	1.8
Ala	2	3	5	2.0	2.0	3.0	3.0	2.0	4.8
Asp/Asn	1	2	4	1.0	1.0	2.0	2.0	1.2	4.0
Ile	-	1	2	-	-	0.9	0.8	-	1.8
Phe	-	-	1	-	-	-	-	-	1.0
Arg	3	5	6	2.8	3.0	4.6	4.3	3.4	5.9
Lys	-	1	2	-	-	1.0	1.0	-	2.0
Val	-	-	1	-	-	-	-	-	1.0
Leu	1	4	5	1.0	1.0	4.1	4.2	1.0	5.0
Gly	2	3	4	2.0	2.1	3.1	3.1	2.3	4.4
Glu/Gln	3	7	7	3.0	2.1	7.1	7.1	3.1	7.2
Met	-	1	1	-	-	0.7	0.4	-	1.0

Table 8. Amino acid analysis of fully deprotected peptides after fragment condensation

CONCLUSIONS

Synthesis of proteins by the solid phase fragment condensation approach, using short protected peptide fragments (up to approx. 10 residues in length) have been performed [Yajima *et al*, 1974; Barlos *et al*, 1991]. The major difficulty commonly encountered when attempts are made to use larger peptide sequences was poor solubility of protected peptide intermediates. The consequence of these poor dissolution charateristics was that purification using conventional chromatographic techniques was usually low yielding and that the high concentrations required for coupling could not be achieved.

The research described here demonstrated that longer protected peptide sequences could be (i) synthesised successfully, (ii) purified using chromatographic separation methods and finally (iii) coupled to a resin bound peptide. Furthermore, a purification method was described that allowed large protected peptides to be purified in larger quantities and yielded a product that possessed high solubility in many organic solvents. However, to arrive at a method that could be applied to the synthesis, cleavage, purification and coupling of large protected peptide fragments, a prossess of trial and elimination, combined with the inevitable learning process was necessary.

Several different resin-handles were tested. The photolabile nitrobenzyl resin handle for application to Boc chemistry proved to be problematic, both because of poor derivatisation with the first amino acid and a tendancy to generate DKP byproducts at the coupling of the third amino acid unless manually incorporated. Furthermore, the long cleavage times, poor solubility of the crude product and a complicated reversed-phase HPLC due to the presence of many impurities, meant that alternative resin handles were examined.

The major difference between the base-labile handles **22** and **23** was the extent of derivatisation with the first amino acid (**22**:45% and **23**:99%) and the cleavage kinetics when the peptidyl-resin was treated with base. The former handle quantitatively released the protected peptide after a 30 min treatment with base, while the latter still retained 17% peptide after 6h reaction. The synthesis of fragment 16-32 on handles **22** and **23** yielded the correct material.

Unfortunately, fragment 1-15 was not obtained from either handle due to a two stage degradation process which appeared to be due to an incompatability of some benzyl-based protecting groups with piperidine. The primary degradation product formed almost immediately while the secondary degradation products were formed on continued exposure to piperidine and the rate of degradation seemed to be related to the handle used (**22** > **23**).

The next step was purify the large protected peptides to homogeneity. Since success had been reported with a variety of chromatographic procedures, each was evaluated using the GRF 1-15 and 16-32 protected peptides. While not all purification methods were used to purify every synthetic peptide, the applicability of a given technique for the purification of hydrophobic peptide fragments was judged on the basis of loading, recovery, purity and physical properties of the purified material (Table 9).

Various chromatographic methods were used to purify the Boc/Bzl based protected peptides. Purification on the basis of polarity, using silica gel was found to be a poorly resolving technique, which yielded a product with poor solubility. Likewise, gel filtration suffered from precipitation on the column and therefore gave poor yields. A more exacting means of purification, especially of these extremely hydrophobic species, was on reversed-phase media. High capacity, MPLC purification of Boc synthesised fragment 16-32 gave better recovery and purity than previously obtained, but also a product that was sparingly soluble in organic solvents. Semi-preparative RP-HPLC using a water/AcCN based solvent system was low yielding due to precipitation on the column, but offered the advantage that the white powder product obtained was reasonably soluble in a variety of solvents (eg. DMF, NMP and DMSO).

Preliminary solid phase fragment condensation experiments using Boc/Bzl based protected peptides were low yielding with a maximum coupling of only 28%. Almost quantitative coupling of a single amino acid, under the same reaction conditions showed that the free amino group on the peptidyl-resin was available for reaction and therefore the poor reactivity was due to the protected peptide. These observations combined with a tendancy for the protected peptides

Method	Solvent	Loading	Separation Time	Recovery	Purity	Solubility of product
Solvent Extraction	DCM	N/A	Good	Good	Poor	Excellent
Gel Filtration	DMF	Low	Good	Poor	Poor	Poor
RP-MPLC (C_8)	DMF/H_2O	High	Poor	Poor	Excellent	Poor
RP-HPLC (C_8)	H_2O/AcCN	Good	Good	Poor	Excellent	Excellent
Perfusion HPLC	H_2O/AcCN	High	Excellent	Good	Excellent	Excellent

Table 9. Table comparing different purification techniques for protected peptides

to 'gel' upon standing in solution, possibly due to intermolecular interactions between unprotected Asn/Gln residues lead to the use of Fmoc chemistry.

Consequently, the Fmoc strategy was examined using the Sasrin acid-sensitive resin. The pre-derivatised Sasrin resin yielded fragment 1-15 in good yield and purity as determined by analytical RP-HPLC, but as previously reported, solubility in most solvents used for purification and coupling was minimal. Previous reports indicating that improved solubility was achieved using protected Asn/Gln residues [Atherton *et al*, 1990, Kamber and Riniker, 1991] were corroborated by their use for the synthesis of 1-15 (one Asn residue) and 16-32 (four Gln residues). Thus, Asn/Gln(Trityl) containing protected peptides were synthesised on Sarin resin and were found to be largely soluble in most solvents.

The increased solubility of the Asn/Gln protected peptides allowed various methods to be used for batch purification of large quantities of material. Thus, solvent extraction, while not capable of separating peptidic impurities, was useful in cases where the crude protected peptide was highly homogeneous. The yield of purified material from conventional RP-HPLC was higher when protected Asn/Gln residues were used (ie. 21% from 4% recovery) but still suffered from precipitation on the column. The third method was perfusion chromatography which like the conventional RP column had a loading that was dependent on the column dimensions. The advantage however, apart from the improved recovery (ie. 53% from 21%) was that shorter gradient times (without loss of resolution) could be applied, resulting in a higher turnover of crude material and yielding a highly soluble product.

The experience gained during the course of these studies suggested that the most suitable products for solid phase fragment condensation could be obtained using Fmoc chemistry combined with a complete protection scheme (ie. Asn/Gln(Trt)), since the solubility profile was far superior when these groups were prevented from interacting.

A variety of coupling conditions were tried and the best combination was found to be DCC/HOBt. Fragment condensation of the highly soluble (0.07M in

DMSO/NMP) fully protected 17 residue 16-32 peptide onto resin-bound 33-44+Gly was achieved to a yield of 60% (from RP-HPLC). Unfortunately, a competing side reaction which involved the irreversible guanidation of the free N-terminus by DCC was detected (13%). The extent of DCC capping appeared to be related to the N-terminal amino acid present, since the coupling of fragment 1-15 to 16-44+Gly peptidyl-resin was almost quantitative with no significant guanidated 16-44+Gly product. The final yield of the desired peptide sequence 1-44+Gly was achieved with an overall yield of 50% (from RP-HPLC) which would indicate that, although many problems were encountered, protein synthesis through fragment condensation on a resin support represents a route worth exploring further.

CHAPTER FOUR

PROTEIN SYNTHESIS- STEPWISE SOLID PHASE APPROACH

INTRODUCTION

The solid phase approach to the synthesis of peptides introduced by Merrifield [1963], combined with improvements in the chemistry and automation, gave the peptide chemist the opportunity to chemically synthesise proteins. Today, almost twenty years later, the stepwise synthesis of highly homogeneous peptides upto approx. 40 residues has become a relatively routine procedure. Unfortunately, attempts to chemically synthesise proteins by extending the peptide chain further (ie. continuing the repeating cycle of N-terminal deprotection, followed by coupling of the next amino acid) were usually difficult and low yielding. The main reason being that the proportion of the desired sequence, in the crude mixture, was often small with respect to its related impurties. One of the earliest examples of stepwise solid phase protein synthesis were conducted by Gutte and Merrifield [1969] who synthesised the 124 residue Ribonuclease A protein. A pure, homogeneous product was not obtained and the low specific activity (13-24%) observed when compared to natural material was attributed to the presence of related impurities (ie. deletion and truncated sequences). This was followed by the synthesis of the 104 residue cytochrome c (2% activity; Sano and Kurihara, [1969]), growth hormone (188 residues (5-10% activity; Li and Yamashiro, [1970]) and lysozyme (129 residues with 9-25% activity; Sharp *et al*, [1973]). In every case the purification of the target sequence was complicated due to the presence of many closely related impurities. In general the impurities obtained were intrinsic to the stepwise approach which is not based on unequivocal reactions. Furthermore, these early syntheses used amino acid derivatives and resins that were not always stable when subjected repeatedly to the harsh conditions required at every stage of the synthesis. The diversity and development of high resolution purification techniques (eg. reversed-phase media that possess varying degrees of hydrophobicity, ion-exchange media with strong and weak binding interactions, affinity methods and separation based on the proteins isoelectric point) that have

become available in recent years, especially when used in combination, have resulted in some success in obtaining the desired sequence from a complicated mixture.

Theoretically, the coupling of pure, fully characterised peptide fragments offered the most logical means of synthesising proteins, given the inherent difficulties associated with the stepwise approach. Several well executed syntheses have been performed, both in solution [Kuroda *et al*, 1992] and on the solid phase [Barlos *et al*, 1991]. The latter represents the most elegant system where reagents can be added and removed simply and efficiently. The major difficulty faced which is independent of the technique used, is poor solubility of short protected peptide fragments (5-10 residues in length) in most organic solvents [Albericio *et al*, 1989] and almost complete insolubility of larger fragments (10-20 residues in length). As a result of the extreme hydrophobicity of protected peptides, development of convenient methods for batch purification (10's of mgs) and the inability to attain the high concentrations of reagents needed for efficient coupling have limited the use of fragment condensation for protein synthesis to a few examples.

Recent research on the enzyme-mediated ligation of fully deprotected peptides [Blake *et al*, 1986] overcame the problems of insolubility of fully protected fragments. The application of this technology is restricted to the experience acquired by a limited number of laboratories. Alternatively, the formation of a thioester linkage between two unprotected peptide fragments [Schnolzer and Kent, 1992] provided an unique method of synthesising proteins which overcame the inherent difficulties of protected peptide fragment condensation.

The research discussed in this chapter, with regards to protein synthesis, is restricted to advancement of stepwise solid phase peptide synthesis.

RESULTS AND DISCUSSION

Synthesis of proteins by stepwise solid phase method

Stepwise protein synthesis and peptide synthesis are in effect one and the same thing, where the former represents an increase in the number of repeating cycles. In reality however, success in protein synthesis is governed by additional factors which in many cases are less significant in the synthesis of shorter peptide chains. Firstly, the accumulation of deletion peptides, which occurs exponentially according to the equation X^y (where X = average yield per step and y = number of residues), results in a crude product that is difficult to purify. For example, in the case of a protein of 100 residues synthesised with an average coupling efficiency of 99.8%, the proportion of deletion sequences in the crude product would be approx. 18%. However, if the average incorporation was 96%, the proportion of deletion peptides would increase to 98.3%. Secondly, the extended periods of contact with harsh chemicals can lead to unwanted reactions on amino acid side chains. Thirdly, the 3D structure and the pathways that lead to a biologically active conformation are generally more complicated for proteins compared to shorter peptides. Thus, the programs used for automated peptide synthesis on the ABI 430A were modified extensively to reduce the risk of modifications during chain assembly. The programs were provided by Dr S. B. H. Kent and were further modified and their final form was as described below:

OPTIMISATION OF SYNTHETIC PROTOCOLS

Resin and amino acid derivatives

Syntheses were performed on stable phenylacetamidomethyl (PAM) resins that are not susceptable to premature cleavage caused by TFA [Gutte and Merrifield, 1971] or the neutralised N-terminal amino group after deprotection [Kent *et al*, 1979]. The PAM resins were obtained prederivatised with the first amino acid to avoid generating resin-bound benzaldehyde groups during functionalisation. If present they can react with free N-terminal amino groups to

form Schiff's base imines, which due to their reversible nature would give rise to a population of deletion peptides. The amino acid derivatives used were of high purity, without the contamination of secondary urethane-protected derivatives which are stable to the TFA used to remove the Boc group [Kent and Merrifield, 1983]. The amino acid side chain protecting groups (Table 1) were chosen for their resistance to acid and activating reagents.

Amino acid	Side chain protecting group	Remarks
Alanine	tosyl	
Arginine		
Asparagine	cyclohexyl	greater stability
Aspartic acid	Acm	cleaved post-HF
Cysteine		
Glutamine	cyclohexyl	greater stability
Glutamic acid		
Glycine	dinitrophenyl	reduces side reactions
Histidine		
Isoleucine		
Leucine	Cl-Z	
Lysine		
Methionine		
Phenylalanine		
Proline	benzyl	
Serine	benzyl	
Threonine	formyl	
Tryptophan	Br-Z	
Tyrosine		
Valine		

Table 1. Amino acid derivatives used for protein synthesis

Activation

The majority of the Boc-protected amino acids were activated as symmetrical anhydrides, which were prepared immediately before use, upon the addition of 0.5 equivs. DCC. The advantage of this means of coupling is that peptide chains could not be prematurely terminated by DCC-activation of any residual TFA which could be present at the coupling step [Stewart and Young, 1984].

Furthermore, symmetrical anhydrides have also been found to offer superior rates of reaction in many difficult sequences. To minimise the risk of incomplete coupling, a second coupling reaction was performed.

The exception to anhydride coupling were the residues arginine, asparagine and glutamine. Arginine, when protected on the side chain with tosyl, has been found to undergo a side reaction, whereby more than one residue of arginine may be added to the peptide chain in a single coupling step [Stewart and Young, 1984]. In the case of asparagine and glutamine DCC can cause dehydration of the side chain [Stewart and Young, 1984] and therefore these residues and arginine were coupled as the HOBt activated esters. The disadvantage of the active ester approach is that due to its low reactivity, it was necessary to give longer coupling times. The second addition of arginine, during a double coupling cycle, was as the symmetrical anhydride due to its better reaction kinetics. Since the incorporation was almost complete in the first half, the risk of undesired insertion reactions were negligible.

Coupling

The cartridges used for the peptide synthesiser each contained 2mmol of amino acid derivative. Using a synthetic protocol which was based on a 0.5mmol scale, two equivs. of symmetrical anhydride would be delivered to the deprotected peptidyl-resin. However, in the case of the HOBt activated ester, 4 equivs. were delivered. The additional activated intermediate helped to overcome the slower reaction times.

Deprotection and neutralisation

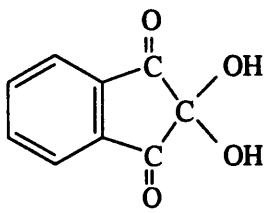
The N-terminal amino protecting group, Boc, was removed using a 5 min treatment with 100% TFA. The resin was then washed with DMF and neutralised with DIEA. Following a final washing procedure the peptidyl-resin was ready to receive the activated amino acid derivative.

Capping

A capping procedure was performed to prevent the formation of deletion sequences, following the addition of each amino acid. This involved a prewash with 15% DIEA in DMF, followed by a 2.5 min treatment with 18% acetic anhydride and 10% DIEA in DMF, as described on p.65.

Monitoring coupling efficiency

In an ideal situation it should be possible to determine the extent of amino acid incorporation, at the time of coupling, using a non-destructive assay procedure. Unfortunately, the technology required for on-line monitoring and 'intelligent' manipulation of the synthetic protocol was not available on the peptide synthesiser used. In Boc chemistry, using the 'batch' synthesis procedure, only destructive tests are available for routine determination of amino acid incorporation and were conducted post-synthesis. The test used involved reacting an aliquot of peptidyl-resin after the completion of each amino acid coupling with ninhydrin, which when reacted with secondary amines, generated a blue colour [Sarin *et al*, 1981]. The intensity of the colouration, which could be measured on a spectrophotometer at 570nm, was directly proportional to the number of free amino groups. The absorbance reading, combined with the weight of peptidyl-resin and theoretical number of moles of free amino groups on the resin, could be inserted in the equation shown in figure 1, to determine the percentage incorporation of amino acid at a particular point in the sequence. The method is sensitive and found to be accurate down to less than 0.1% residual amine (>99.9% coupling) with a reproducibility of ±0.05% [Kent, 1988]. The conditions are also sufficiently vigorous (ie. 5 min at 100°C) to reveal amine that was unavailable for coupling (eg. due to inter- or intramolecular aggregation or Schiff's base formation; Kent, [1984]).



Ninhydrin

$$\text{amine } (\mu\text{mol/g}) = \frac{(\text{Abs sample} - \text{Abs blank}) \times 3\times10^{6}}{\text{Extinction coefficient} \times \text{sample weight (mg)}}$$

where extinction coefficient = 15000 $M^{-1}cm^{-1}$

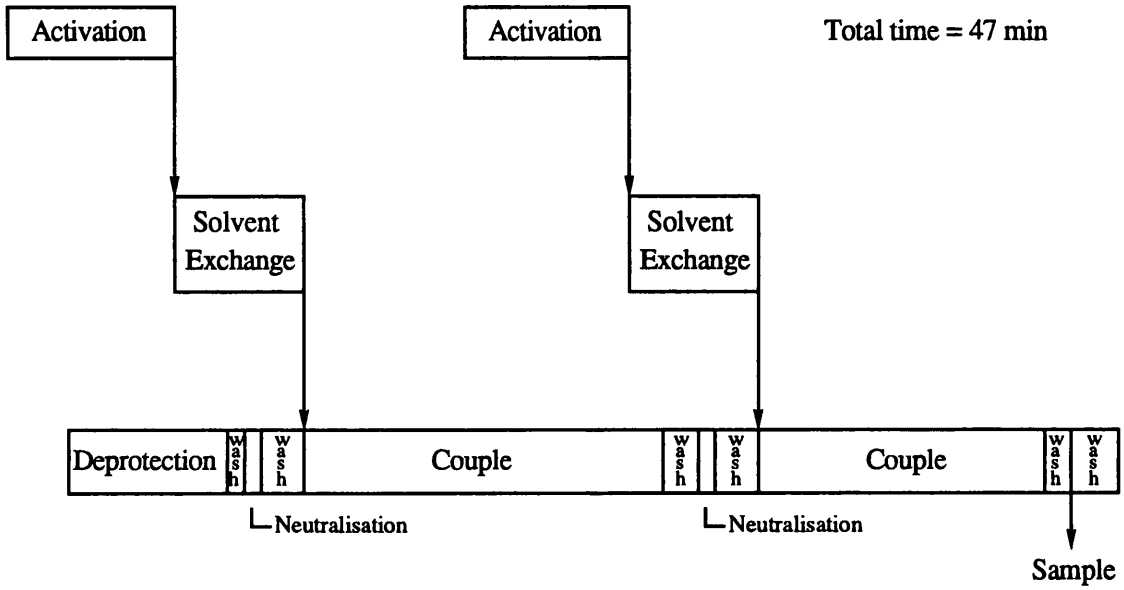
$$\%\ \text{Coupling} = \left\{ 1 - \frac{\text{amine } (\mu\text{mol/g})}{10^{3} \times \text{substitution (mmol/g)}} \right\} \times 100$$

Figure 1. Equation for calculating amino acid incorporation

Automation

All the aspects described above were incorporated into a computer program which was used to control an ABI 430A automated peptide synthesiser. The instrument essentially consisted of three parts, each operating independently of one another and performing separate functions. The first part of the instrument was the activation vessel where DCC was delivered, followed by the dissolved amino acid derivative. After activation the solution was transferred to the concentrator, leaving behind the precipitated DCU byproduct. The concentrator besides reducing the volume of solvent and replacing DCM with DMF, also served as the activating vessel for the preparation of HOBt active esters. The activated amino acid derivative (symmetrical anhydride or active ester) was then transferred to the reaction vessel containing the freshly N-terminal amino deprotected peptidyl-resin. To ensure efficient mixing of the reagents the reaction vessel was agitated using a mechanism that caused the contents to vortex and circulate constantly. Scheme 1 below illustrates operation of the peptide synthesiser with respect to time showing overlapping function of the activator, concentator and reaction vessel. Using these optimised procedures a double coupling cycle could be accomplished in approx. 45 min, compared to 105 min for the standard procedure. Thus, a protein (eg. 100 residues) could be synthesised in approx. 3 days.

During the course of a synthesis (approx. every 25 residues) it was necessary to remove a portion of the resin, firstly, to reduce the bulk of the peptidyl-resin and thereby ensure continued efficient mixing of the reagents. Secondly, the peptide fragments could be sequenced to provide a qualitative indication of the success of the synthesis, more accurate than that obtained from the ninhydrin analysis.



Scheme 1. General scheme of double coupling cycles performed by the automated peptide synthesiser. Taken from Kent and Clark-Lewis [1985]

Cleavage

Short peptides with uncomplicated sequences (eg. not containing Trp) were usually cleaved from the resin support using a High HF procedure only. However, chemically synthesised proteins which possessed a more varied mixture of amino acids and therefore protecting groups are subject, in theory, to a higher occurance of incomplete cleavage and side reactions induced by HF, resulting in a further complicated crude cleavage product. Consequently, a modification of the Low/High HF technique [Tam *et al*, 1983] was used for the cleavage of proteins. The Low TFMSA/High HF protocol [Kent, 1988] offered the advantages that TFMSA could be more easily removed that HF and that

mechanistically the low TFMSA part of the reaction removed the majority of benzyl-based protecting groups through an SN_2, as opposed to an SN_1 mechanism, which in theory should reduce side reactions because reactive carbonium and nitronium ions are generated via the former route.

Before acidiolytic cleavage was undertaken, pretreatment of the peptidyl-resin for His(DNP), N-terminal Boc and Trp(Formyl) was necessary, according to Kent [1988]. The dinitrophenyl protecting group for His is generally cleaved by thiolysis and the reaction can be performed on either the peptidyl-resin or on the free protein after HF cleavage. For reasons of convenience, the former approach was usually adopted, whereby the Boc protected peptidyl-resin was treated with 20% β-mercaptoethanol, 10% DIEA in DMF for 1h. The Boc group was then removed with a 30 min treatment in 50% TFA/DCM. The formyl group for Trp is sensitive to acidiolytic treatment, but to ensure complete removal, especially when incorporated in the sequence of a protein, a reaction with ethanolamine at 0°C in DMF was necessary. After washing and drying the peptidyl-resin was then ready for low TFMSA treatment. To the peptidyl-resin were added scavengers (ie. 1,4-butanedithiol, *p*-cresol, dimethyl sulphide and TFA) which was then cooled to 0°C before adding TFMSA dropwise. After 2h the partially deprotected peptidyl-resin and any prematurely cleaved protein was precipitated with diethyl ether, filtered from the cleavage mixture and dried under vacuum in preparation for the final part of the cleavage. Since the majority of the benzyl-based protecting groups had been removed in the previous reaction with TFMSA, fewer groups were left which could generate reactive species when cleaved by HF and therefore the fully deprotected protein should be less contaminated by impurities resulting from side chain modification. To ensure that side reactions were minimised during the HF reaction, scavengers were added (ie. 500µl 1,4-butanedithiol and 500µl *p*-cresol). After 1h at 0°C the HF was distilled off and the cleaved protein was precipitated with dry diethyl ether. The protein and residual resin was filtered and the former dissolved in 10% acetic acid. The acetic acid solution was lyophilised to give a crude product ready for purification.

E. COLI 10kD HEAT SHOCK PROTEIN

SYNTHESIS

Two syntheses of the 97 residue 10kD Heat Shock protein of E. coli were performed on a 0.5mmol scale using the automated peptide synthesiser and the protocols described previously. The difference between the two syntheses was that one involved capping following the addition of each amino acid, according to the protocol described above.

Met-Asn-Ile-Arg-Pro-Leu-His-Asp-Arg-Val-Ile-Val-Lys-Arg-Lys-Glu-Val-Glu-Thr-Lys-Ser-Ala-Gly-Gly-Ile-Val-Leu-Thr-Gly-Ser-Ala-Ala-Ala-Lys-Ser-Thr-Arg-Gly-Glu-Val-Leu-Ala-Val-Gly-Asn-Gly-Arg-Ile-*Leu-Glu-Asn-Gly-Glu-Val-Lys-Pro-Leu-Asp-Val-Lys-Val-Gly-Asp-Ile-Val-Ile-Phe-Asn-Asp-Gly-Tyr-Gly-Val-*Lys-Ser-Glu-Lys-Ile-Asp-Asn-Glu-Glu-Val-Leu-Ile-Met-Ser-Glu-Ser-Asp-Ile-Leu-Ala-Ile-Val-Glu-Ala-OH

0.63g PAM resin derivatised with the first amino acid (Boc-Ala; substitution 0.8mmol/g) was loaded into the reaction vessel. Samples of peptidyl-resin (approx. one third) were removed from the reaction vessel at amino acids marked with *. The progress of the syntheses were determined using spectrophotometric analysis of the reaction between ninhydrin and the unreacted N-terminal amino groups after every double coupling cycle or, in the case of the second synthesis, after every double coupling cycle but before the capping procedure (Fig. 2). The average coupling was 99.35% (uncapped) and 99.71% (capped). As can be observed from the general trend (Fig. 2), the synthesis which involved capping gave better amino acid incorporation (except for Ile^{7} and Leu^{49}), especially between residues 45 to 70.

Cleavage of 1-97 peptidyl-resin

The DNP from His and the N-terminal amino protecting group were removed from 275mg 1-97 peptidyl-resin as described previously. After drying under vacuum the peptidyl-resin weighed 260mg and was then treated using the

low TFMSA/High HF protocol. The crude protein was dissolved in 10% acetic acid and lyophilised; 54mg (approx. 34%) crude protein was thus recovered.

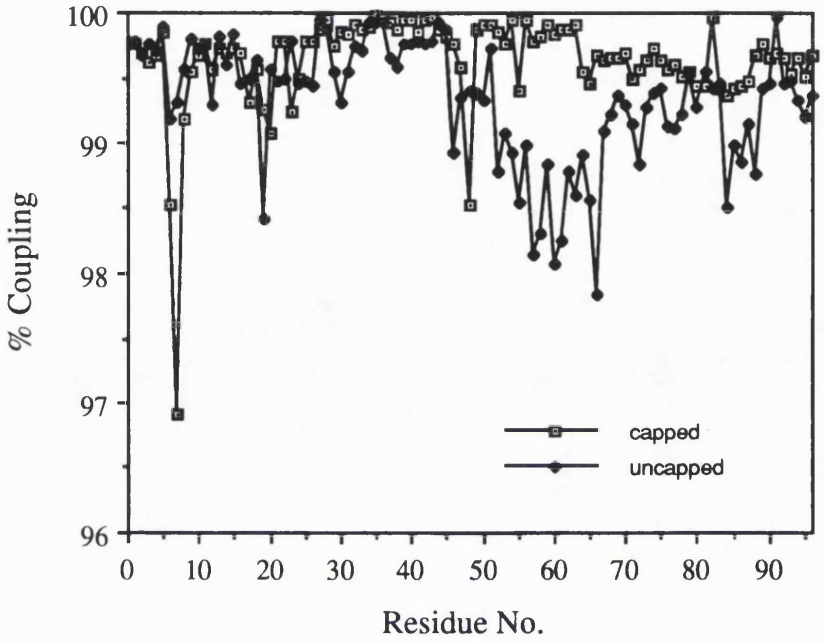


Figure 2. Ninhydrin analysis for synthesis of E. coli 10kD HS protein
C-terminus = Residue 1

PURIFICATION AND ANALYSIS

The crude HF product was run on SDS-PAGE and gave a band that corresponded to a molecular weight of approx. 16kD (Lane 1; Fig. 3). The higher than expected value was also observed for recombinant 10kD protein [Chandrasekhar *et al*, 1986]. This confirmed that a protein with the correct approximate molecular weight had been obtained.

One of the aims of this piece of work was to establish whether molecules like the E. coli 10kD and which are generally termed chaperones, due to their ability to help other proteins to find their biologically active 3D structure [Ellis and van der Vies, 1991], are capable of spontaneously finding their active structure or whether they need the assistance of other chaperones.

Since both natural and recombinant 10kD protein from E. coli assemble into an heptameric structure, the first assay performed was size exclusion chromatography. The crude synthetic protein was run on a size exclusion column in non-denaturing buffer (Fig. 4) where four distinct peaks were observed. These were found to correspond to various oligomers. In particular it was found that one of the four aggregates had a molecular weight of approx. 70kD consistant with an heptameric structure. Further work relevant to the biological activity of this molecule was conducted in the laboratory and the results described in the paper entitled 'Chemical Synthesis Of 10kD Chaperonin' (Publications, p.238). With respect to the chemical characterisation of the protein no further work was carried out at this stage.

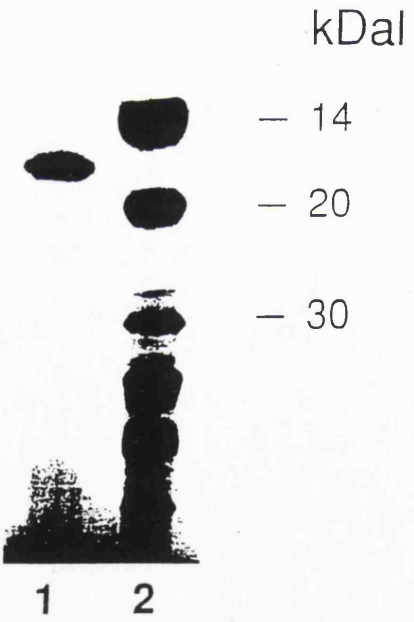


Figure 3. SDS-PAGE of crude synthetic E. coli 10 kD HS protein (lane 1) and low molecular weight markers (lane 2)

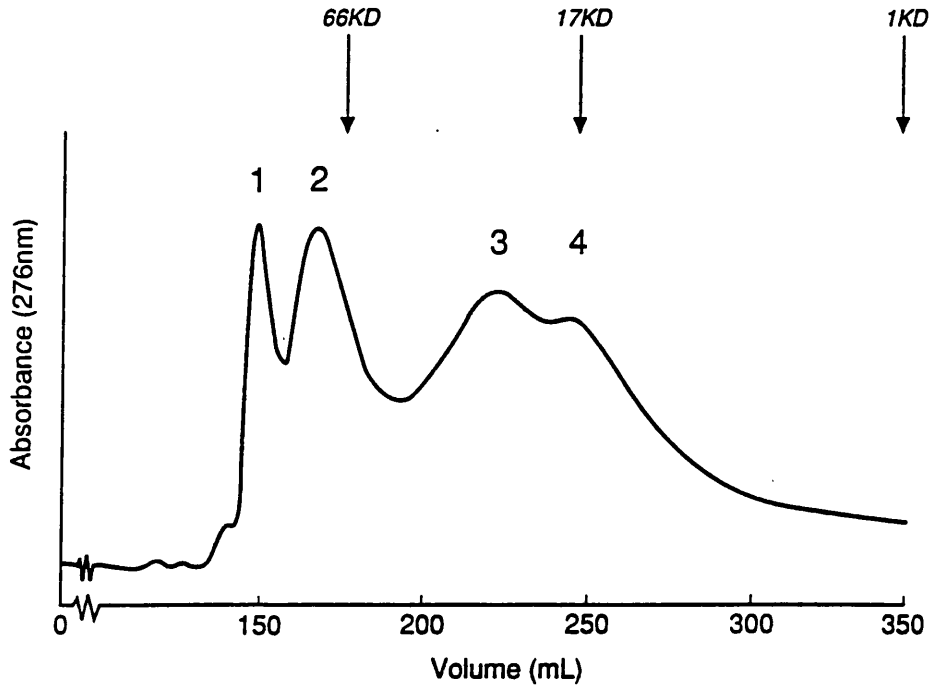


Figure 4. Size exclusion chromatogram of crude E. coli 10 kD HS protein (System 17)

M. TUBERCULOSIS 10KD HEAT SHOCK PROTEIN

SYNTHESIS

The chemical synthesis of the 10kD Heat Shock protein of M. tuberculosis was performed on the automated peptide synthesiser using the same protocol described for the E. coli protein. No capping procedure was used however for this synthesis.

Ala-Lys-Val-Asn-Ile-Lys-Pro-Leu-Glu-Asp-Lys-Ile-Leu-Val-Gln-Ala-Asn-Glu-Ala-Glu-Thr-Thr-Thr-Ala-Ser-*Gly-Leu-Val-Ile-Pro-Asp-Thr-Ala-Lys-Glu-Lys-Pro-Gln-Glu-Gly-Thr-Val-Val-Ala-Val-Gly-Pro-Gly-Arg-Trp-*Asp-Glu-Asp-Gly-Glu-Lys-Arg-Ile-Pro-Leu-Asp-Val-Ala-Glu-Gly-Asp-Thr-Val-Ile-Tyr-Ser-Lys-Tyr-Gly-Gly-Thr-Glu-Ile-Lys-Tyr-Asn-Gly-Glu-Glu-Tyr-Leu-Ile-Leu-Ser-Ala-Arg-Asp-Val-Leu-Ala-Val-Val-Ser-Lys-OH

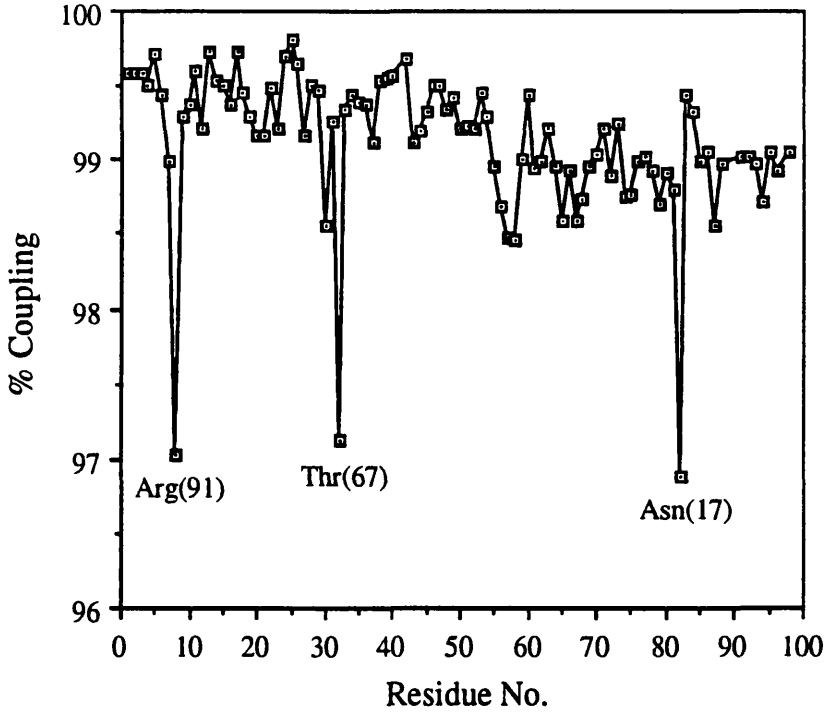


Figure 5. Ninhydrin analysis for synthesis of M. tuberculosis 10kD HS protein
C-terminus = Residue 99

0.77g of prederivatised PAM resin (Boc-Lys; substitution 0.65mmol/g) was used for the synthesis of the protein. At two points (*) during the course of the synthesis, samples of resin were removed from the reaction vessel. The progress of the synthesis was determined using spectrophotometric analysis of the reaction between ninhydrin and the unreacted N-terminal amino groups after every double coupling cycle (Fig. 5). The average coupling was 99.12%.

Cleavage

The low TFMSA/high HF protocol described for the cleavage of the E. coli 10kD Heat Shock Protein was applied to 300mg peptidyl-resin. However, due to the poor solubility of the crude peptide in 10% acetic acid solution, 20mmol ammonium acetate, pH 7.8, was used. Thus, after lyophilisation 260mg (theoretical yield 185mg) crude protein was recovered. The higher than expected yield was probably due to incomplete removal of the ammonium acetate salt and/or residual water molecules. An amino acid analysis of the crude protein was performed and indicated a good found to expected ratio (Table 2).

Residue	Expected	Peptidyl-resin	Crude	Pure
Tyr	4	5.0	5.2	4.5
Ala	9	8.9	8.8	8.3
Asp/Asn	10	10.1	10.0	9.4
Ile	7	6.3	6.4	5.9
Arg	3	3.1	3.2	3.1
Lys	9	8.6	8.4	8.1
Val	11	11.6	10.6	10.0
Leu	7	7.0	7.4	6.9
Gly	9	9.7	10.1	9.5
Glu/Gln	13	13.0	13.3	12.3
Thr	7	5.6	5.9	5.6
Ser	4	4.1	4.1	4.0
Pro	5	4.8	4.7	4.4

Table 2. Amino acid analyses for synthetic M. tuberculosis 10kD HS protein

PURIFICATION AND ANALYSIS

The crude protein (Fig. 6A) was then purified using a three step protocol (ie. dialysis, ion exchange and reversed-phase chromatography). Each procedure separated the desired protein sequence from smaller molecular weight impurities, deletion and truncated proteins using different physical characteristics.

The first procedure which was aimed at removing small molecular weight molecules (eg. scavengers and small peptides) from the crude HF cleavage product involved dialysis. 150mg of crude protein was dissloved in 75ml 20mM ammonium acetate pH7.8 buffer to give a 2mg/ml solution and placed in a dialysis bag with Mwt. cutoff 8000. The protein solution was first dialysed against 20mM ammonium acetate pH7.8 buffer for a total of 6h with two further buffer changes and stirred at room temperature. The ammonium acetate buffer was then followed by two 90 min treatments with water. The contents of the dialysis bag were removed and lyophilised to give 120mg crude protein.

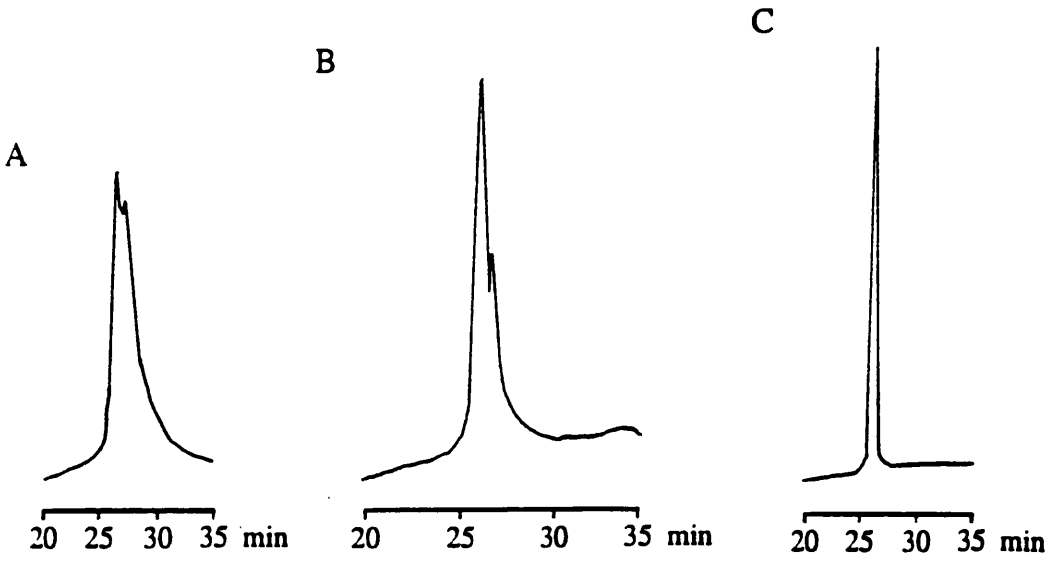


Figure 6. C_{18} RP-HPLC (System 18) of
A) Crude HF cleavage product
B) Semi-purified 10kD protein
C) Purified M. tuberculosis 10kD following 3 stage purification scheme

Ion exchange chromatography was then applied to the dialysed protein. Since the pI of the protein was 4.4 anionic ion-exchange medium was chosen, equilibriated with 20mM TrisHCl pH7.8. The eluting buffer contained 0.5M NaCl. 30mg dialysed crude protein was injected on the semi-preparative MONO Q column and a gradient consisting of 30 min at 0% salt, followed by the application a salt gradient from 0% to 100% (0.5M NaCl) in 130 min was applied. The flow was 2ml/min and detection was at 220 and 274nm. The product was found to elute after 120 min which corresponded to 0.35M NaCl. Fractions were taken through the main peak and checked by analytical RP-HPLC. Those fractions containing the same material were combined, desalted on RP-HPLC media and lyophilised. An analytical quantity of the semi-purified protein was injected on a RP-HPLC and gave a cleaner chromatogram, compared to crude material, as shown in Figure 6B.

The final stage of the purification was reversed-phase chromatography which also served to desalt the product from ion exchange chromatography. 60mg of semi-purified protein containing salt was loaded onto a semi-preparative RP column. A 290 min gradient was applied, from 20-100% acetonitrile, with a flow of 2.5ml/min. Fractions were taken through the main peak, which eluted at approx. 45% acetonitrile and analysed by RP-HPLC. Fractions that contained the same material were pooled and lyophilised to yield 3mg (10%, based on crude dialysed material) purified M. tuberculosis 10kD Heat Shock Protein (Fig. 6C). Amino acid analysis (Table 2), FAB-MS/Electrospray MS ($[M+H]^+$: Expected 10666; Found $[M+Na]^+$ 10684±10) and protein sequencing (Table 3) confirmed that the correct sequence had been isolated.

SEQUENCES 99-75 & 74-50					SEQUENCE 49-25		
Expected	Found / pmol		Found / pmol		Expected	Found / pmol	
A	A	8668			D	D	259
K	K	7720			E	E	435
V	V	7510			D	D	240
N	N	5629			G	G	467
I	I	6406			E	E	413
K	K	6375			K	K	829
P	P	4433			R	R	111
L	L	4271			I	I	659
E	E	3067			P	P	526
D	D	2573			L	L	558
K	K	3133			D	D	173
I	I	2678			V	V	437
L	L	2718			A	A	426
V	V	3160			E	E	161
Q	Q	1546			G	G	307
A	A	2102			D	D	136
N	N	1279			T	T	89
E	E	1511			V	V	198
A	A	1960			I	I	226
E	E	1439			Y	Y	200
T	T	707			S	S	36
T	T	859			K	K	102
T	T	739			Y	Y	120
A	A	1114			G	G	105
S	S	125			G	G	72
G			G	789			
L			L	3390			
V			V	2676			
I			I	1406			
P			P	1225			
D			A	341			
T			T	413			
A			A	793			
K			K	997			
E			E	257			
K			K	787			
P			P	648			
Q			Q	193			
E			E	328			
G			G	307			
T			T	58			
V			V	694			
V			V	765			
A			A	214			
V			V	571			
G			G	199			
P			P	232			
G			G	284			
R			R	77			
W			W	26			

Table 3. Sequence analysis of peptide fragments from M. tuberculosis 10kD Heat Shock Protein

CHAPTER FIVE

EXPERIMENTAL

MATERIALS AND METHODS

Boc, methyl ester and t-butyl ester protected amino acids were purchased from Novobiochem (Nottingham, UK). Phenylacetamidomethyl (PAM) and *p*-methylbenzhydrylamine (MBHA) resins and peptide synthesis coupling reagents were supplied by Applied Biosystems (Warrington, UK). Apparatus for HF cleavage was supplied by The Peptide Institute, Japan. Dichloromethane (BDH; Poole, UK) and dimethylformamide (Rathburn; Walkerburn, UK) were analytical and peptide synthesis grade respectively. Acetic anhydride (BDH) was redistilled, bp 138°C, prior to use. Peptides were synthesised on an ABI 430A peptide synthesizer, using Boc chemistry. Peptides synthesised using the Fmoc strategy were prepared on an ABI 431A. Fluorene-4-carboxylic acid was obtained from Lancaster Synthesis. All other reagents and solvents used to make the Fmoc probes were analytical reagent grade. Thin-layer chromatograms were run on precoated 0.25 mm silica gel $G_{254/356}$ plates from Merck. The solvent systems used for both TLC analysis and flash chromatography were as follow: A, chloroform/MeOH 7:3; B, chloroform/AcOEt 9:1; C, chloroform/MeOH 9:1; D, chloroform/AcOEt 1:1; E, chloroform/AcOEt 8:2; F, chloroform/MeOH 19:1; G, chloroform/MeOH 6:4; H, cyclohexane/EtOAc 1:1; I, chloroform/ MeOH/EtOAc 8:1:1; J, chloroform/MeOH 1:1; K, chloroform/AcOEt 6:4; L, chloroform/MeOH 8:2; M, petroleum ether/AcOEt 2:1; M, petroleum ether/AcOEt 3:1. Spots were visualized under the short wavelength UV lamp or, where appropriate, after spraying with a 0.2% solution of ninhydrin in acetone.

Nuclear Magnetic Resonance Spectrometry (NMR)

^{1}H-NMR for analysis of organic molecules was performed on were either Varian (300MHz) or Bruker (500MHz) instruments. Samples were dissolved in either deuterated chloroform ($CDCl_3$) or dimethylsulphoxide (DMSO-d_6), depending on solubility. Resonance readings were taken at r.t.

LIQUID CHROMATOGRAPHY

High pressure liquid chromatography

The HPLC instruments used were either Applied Biosystems 150A or Shimadzu LC10 systems. The former was a dual piston, two pump, gradient system, connected to a single wavelength detector and dedicated integrator. The latter was also a gradient instrument with the differences that the pumps were single piston, with a dual wavelength detector, controlled temperature chamber and autosampling ability for sample analysis. In general the water/acetonitrile solvents were buffered with 0.045% and 0.036%TFA, respectively. The exception was the purification of protected peptide fragments by perfusion chromatography where the solvents were used unbuffered. Detection was at 220nm (amide bonds), 254nm (aromatic), 270nm (Fmoc) and 301nm (Fmoc).

The reversed-phase columns for analysis, purification and desalting were either obtained from either Vydac or Brown-Lee. The perfusion chromatography columns were obtained from Perseptive Biosystems.

FPLC

Low-pressure chromatography was performed on a Pharmacia FPLC instrument, which consisted of a gradient programmer, two single piston pumps and single wavelength detector. For ion-exchange chromatography, analytical (50x5mm) and semi-preparative (160x10mm) anionic (Mono Q) or cationic (Mono S) columns, obtained from Pharmacia were used. For size exclusion separation a Sephacryl S-100 semi-preparative (160x10mm) column was used.

Flash Chromatography

The flash chromatography apparatus from Rhone-Poulenc consisted of essentially two main parts. The lower piece formed the column was 30cm in length with an internal diameter of 2.5cm. The top of the apparatus formed the solvent reservoir and was attached to a cylinder of nitrogen. The separating medium was silica gel 60 and was loaded into the column as a slurry, using the eluting solvent to a depth of 20cm. At the top of the silica gel bed was placed a

disc of filter paper to prevent disturbance. The sample was placed at the top of the column in a minimum volume of liquid (usually 2-3ml) and allowed to enter the bed under gravity. 200-300ml of solvent was then placed in the reservoir and a positive pressure applied (10psi). 25ml fractions were collected and analysed by TLC to locate those that contained purified material. Similar fractions were combined and evaporated on a rotary evaporator.

Amino acid analysis

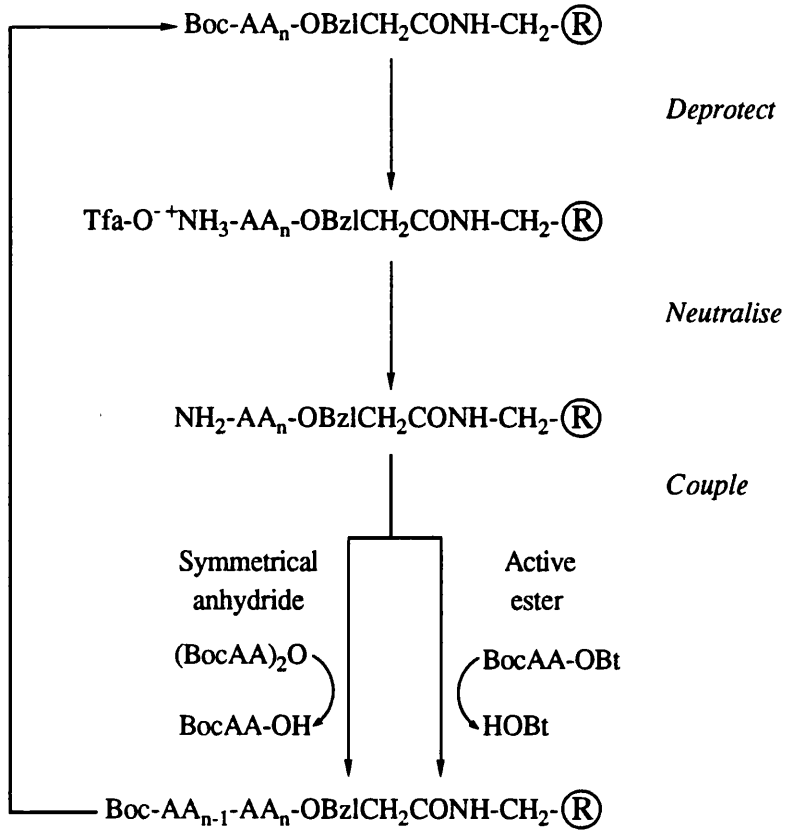
Amino acid analyses were performed on a Beckman System Gold analyser. The peptide or peptidyl-resin sample was hydrolysed in 1ml of a 1:1 mixture of hydrochloric acid/propionic acid (Sigma) at 110°C for 21h. The acid solution was evaporated to dryness, redissolved in buffer, which contained β-alanine (internal standard) and then filtered in preparation for analysis. The complete analysis system was fully automated and involved taking a 25µl sample and separating the individual amino acid components on an ion-exchange column (150x3mm), using a flow rate of 0.3ml/min. The amino acids were then derivatised with ninhydrin, at 130°C and detected at 500nm. The computer then calculated the relative percentage of each component following normalisation.

Amino acid side chain-protecting groups

Amino acid	N-alpha Boc	N-alpha Fmoc
Asp, Glu	OcHx	Ot-But
Ser, Thr	Bzl	t-But
Lys	2ClZ	Boc
Tyr	BrZ	t-But
Cys	Acm	not used
His	DNP	Trityl
Arg	Tos	Pmc
Trp	Formyl	not used
Asn, Gln	unprotected	Trityl

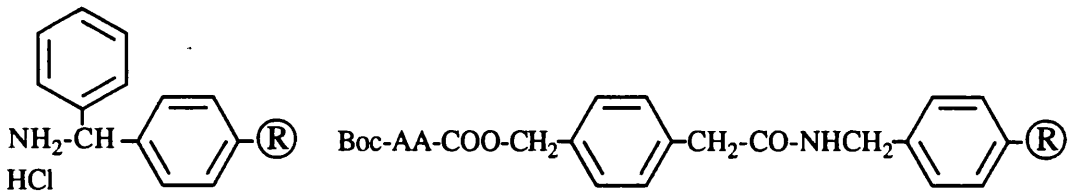
Table 1. Table of amino acid side chain protecting groups

GENERAL PEPTIDE SYNTHESIS USING BOC CHEMISTRY



Scheme 1. General scheme for stepwise peptide synthesis using Boc chemistry

Peptide syntheses were performed on an automated peptide synthesiser ABI 430A, which was designed for batch synthesis. Every step required to build a peptide chain (ie. deprotection, activation and coupling), shown in Scheme 1, were carefully coordinated to achieve high overall yields. Polystyrene based MBHA.HCl resin (**27**), required a loading step to introduce the first residue, after neutralising the HCl salt and yielded a C-terminal amide. Alternatively, "PAM" resins (**28**), pre-derivatised with the first amino acid were used and gave a peptide with a free carboxyl C-terminus. The cross-linked polymeric resin support existed as a suspension copolymer of styrene-divinylbenzene (S-DVB), in the form of beads with an approximate diameter of 50 microns. The substitution value of such resins was usually of the order 0.4-0.8mmol/g.



27 28

Deprotection

Acidolytic deprotection of the N-terminal Boc group was accomplished using 100% TFA. Following a DMF wash the trifluoroacetate salt was neutralised with the tertiary amine DIEA. The resin was then washed again with DMF in readiness for receiving the activated amino acid.

Amino acid activation

All the amino acids with the exception of arginine, glutamine and asparagine were activated as symmetrical anhydrides, using DCC. A four fold excess of amino acid (2mmol) was used for a large scale (0.5mmol) synthesis, which when activated with 1mmol DCC, produced a two fold excess of symmetrical anhydride. Activation was performed in DCM, with the exception of Lys, His, Leu, Trp and Tyr to which was added 20% DMF to aid dissolution. The three other amino acids were coupled as HOBt esters in four fold excess. The arginine derivative was dissolved in a mixture of DCM/DMF in the ratio 1:10, while glutamine and asparagine were dissolved in 100% DMF. If each amino acid was double coupled, the second delivery of glutamine and asparagine (dissolved in DMF) would also be as the HOBt active ester, but in the case of arginine (dissolved in 50% DCM/DMF), the symmetrical anhydride was prepared.

During the activation of amino acids as symmetrical anhydrides, the volume of liquid was reduced by evaporating off the volatile DCM solvent. At the same time the part of the DCM was replaced with DMF. The activated amino acid was then delivered to the N-terminal amino deprotected peptidyl-resin. The residual DCU byproduct was dissolved in one of two ways, depending on the type of activation (ie. symmetrical anhydride: MeOH; HOBt active ester: DMF), since HOBt activation was performed in the concentrator vessel.

Coupling and determination of coupling efficiency

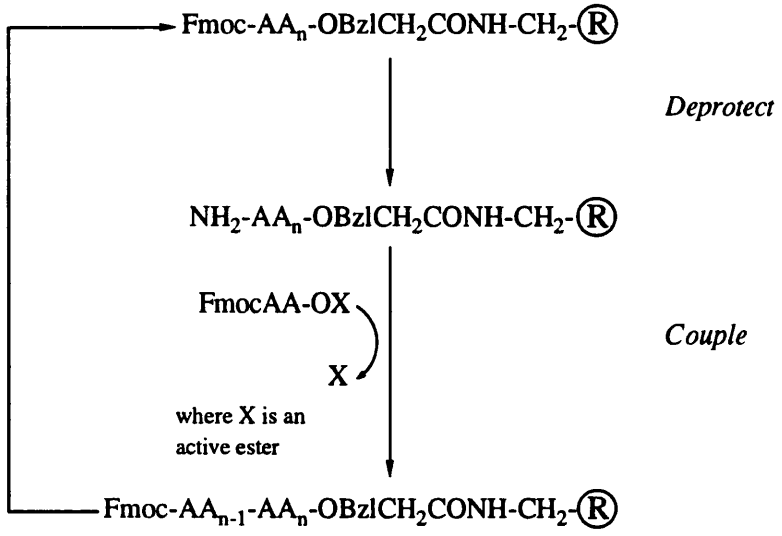
Coupling of each activated amino acid was performed in solvent containing a high proportion of the polar solvent DMF, which is known to give high levels of incorporation. Efficient mixing of reagents in the reaction vessel was achieved by means of a vortex mixer. The length of time allowed for each coupling was dependent on the type of activation. For the symmetrical anhydrides 10.3 min was allowed, while slightly longer (ie. 17.3 min) was given for the slower reacting HOBt esters.

Following the incorporation of each amino acid, an aliquot of peptidyl-resin was automatically removed from the reaction vessel and delivered to a fraction collector, together with DMF and a small quantity of DIEA. Routinely, a quantitative ninhydrin reaction [Sarin *et al*, 1981] was performed on the separated peptidyl-resin to determine the amount of residual uncoupled amine, as an indication of coupling efficiency. The dried resin sample was treated with three reagents, phenol in ethanol, potassium cyanide in pyridine and ninhydrin. After heating at 100°C for 5 min the absorption at 570nm was recorded. The absorption value related to the proportion of unreacted amino groups and by using the equation shown on page 154, the coupling efficiency could be calculated. The results obtained could then be plotted to visualise the difficult sequences.

Capping

In some instances, a capping cycle was performed following the addition of each amino acid. An aliquot of peptidyl-resin was removed before any capping was performed for ninhydrin analysis. The capping procedure itself involved a prewash with 15% DIEA in DMF, followed by a 2.5 min treatment with 18% acetic anhydride and 10% DIEA. The peptidyl-resin was finally washed with DMF.

GENERAL PEPTIDE SYNTHESIS USING FMOC CHEMISTRY



Scheme 2. General scheme for stepwise peptide synthesis using Fmoc chemistry

The alternative peptide synthesis strategy was based on the base-labile Fmoc group. These syntheses were performed on an ABI431A using Fastmoc (HBTU) protocols, following the basic scheme shown in Figure 2, on a 0.25mmol scale. HBTU when used in conjunction with Fmoc/NMP chemistry is reported to be faster and more complete than carbodiimide-mediated reactions, resulting in shorter cycle times and increased coupling efficiency [Knorr *et al*, 1989].

The resin support used was a copolymer of styrene-divinylbenzene (S-DVB), with an additional group that enabled cleavage of a protected peptide using 1% TFA in DCM. If 95% TFA was used, the fully deprotected peptide could be obtained. The structure of the Sasrin resin (**24**) is shown on page 97 and was prederivatised with the first amino acid, Fmoc-Gly-OH (substitution 0.65mmol/g).

Deprotection

The N-terminal amino Fmoc group was removed with two 3 min treatments with a 20% mixture of the secondary amine piperidine in NMP. The deprotected peptidyl-resin was then washed with NMP.

Amino acid activation

Dissolution and activation of the amino acid derivative was performed in the cartridge. At the start of the cycle, 2.2ml NMP and 2.2ml 0.45M HBTU/HOBt in DMF were delivered and activation was allowed to continue for 6 min. The solution was then transferred directly to the reaction vessel containing the N-terminal amino deprotected peptidyl-resin.

Coupling

As soon as the activated amino acid derivative was delivered to the reaction vessel, followed by the addition of 0.3ml DIEA. A period of 28 min was programmed for coupling during which time the reaction vessel was vortexed to ensure efficient mixing of the components.

SYNTHESIS OF CHROMATOGRAPHIC PROBES

Synthesis of 9-(hydroxymethyl)fluorene-4-carboxylic acid (1)

7 was prepared according to Mutter and Bellof [1984]. All intermediates (**5-7**) were purified as described below and shown to be pure by TLC, FAB-MS and ^{1}H-NMR. Scheme 2 (p.48) illustrates the synthetic procedure.

10g (47.6mmol) fluorene-4-carboxylic acid (**4**) was dissolved in 200ml *p*-dioxane containing 10% conc. sulphuric acid and saturated with isobutylene. After 6 days stirring at r.t. the reaction mixture was neutralised and the product (**5**) separated into chloroform. The organic layer was shaken with water, 3% sodium hydrogen carbonate ($NaHCO_3$), 0.1M HCl and finally water. The chloroform layer was dried over sodium sulphate (Na_2SO_4) and evaporated to dryness. Purification of **5** was achieved using flash chromatography (M) to give 10.2g (80.8%). 4.73g (17.8mmol) compound **5** was dissolved in 30ml diethyl ether containing 5 equivs. sodium hydride and distilled ethyl formate, then refluxed for 6.5h. The reaction mixture was then poured onto cold water and the aqueous layer collected. The latter was acidified with glacial acetic acid and the oily product extracted with diethyl ether. Water soluble impurities were removed by washing the diethyl ether layer with water, 3% $NaHCO_3$ and finally water. The organic layer was dried over Na_2SO_4, evaporated to dryness to yield 4.32g (81.6%) crude product (**6**). The crude material was then reduced without further purification. 4.11g (14mmol) compound (**6**) was dissolved in isopropyl alcohol to which was added 1.5 equiv. sodium borohydride. After stirring at room temperature for 3.5h the reaction mixture was poured onto water, extracted into diethyl ether and treated as described for **6**. Intermediate **7** was then purified by flash chromatography (N) to yield 3.71g (89.6%). 3.7g of **7** was deprotected by treatment with 30ml 50% TFA in DCM for 30 mins and then purified by flash chromatography (A). The solvent was removed on a rotory evaporator to yield 2.85g (75.9%) of pale yellow solid (**1**): R_f (A) 0.4. ^{1}H-NMR (DMSO-d_6): 3.8-4.1 (m, 3H, H-C(9), CH_2OH); 7.2-7.8 (m, 6H, aryl H); 8.3 (d, 1H, H-C(3)). MS: 240 ($[M+H]^+$); 223 ($[M+H]^+$ - OH); 210 ($[M+H]^+$ - CH_2O); 165 ($[M+H]^+$ - CO_2).

Synthesis of (D,L) α-aminodecanoic acid (8)

2.5g sodium metal (0.11M) was weighed after removal of oxidised material. The sodium was dissolved over 1h in 85ml abs. EtOH which had been cooled to 0°C. To the solution was added 23.7g 1-bromooctane (0.11M) and 24.3g diethyl acetoamidomalonate (0.13M) with stirring. The reaction mixture was refluxed overnight. The solution was allowed to stand and after attaining room temperature (r.t.) was poured onto 160ml ice/water. Stirring was continued at 0°C for 45 mins. The precipitated solid was filtered off and washed with aliquots of ice-cold water. The residue was allowed to dry on the filter under vacuum. The pale brown solid was dissolved in 180ml 36% HCl in a 1 litre round-bottomed flask. Refluxing was performed overnight. The acidic solution was diluted with 500ml water/EtOH (3:1), refluxed for 5 mins and hot filtered through cotton wool to remove any oily residues. The solution was cooled and neutralised with a quantity of liq. ammonia to give pH 7, at which point light tan precipitate appeared. The solid was filtered on paper and washed with water and finally dried under vacuum over phosphorus pentoxide. The (D,L) α-aminodecanoic acid (**8**) was crystallised from hot glacial acetic acid to yield 15.8g (77.3%) white crystals.

Synthesis of DL-α-aminodecanoic acid derivative (14)

2g (10.7mmol) DL-α-aminodecanoic acid (**8**) was dissolved in 33ml of dioxane/sulphuric acid (10:1), sealed with a rubber septum and cooled to 0°C. Isobutylene gas was bubbled into the solution through a 10cm long syringe needle until the volume had approximately doubled. The temperature was increased to ambient and stirring continued for 4 days. The excess isobutylene gas was allowed to escape through a purge needle before pouring the reaction mixture onto 120ml ice-cold 50:50 mixture of diethyl ether and 0.25M NaOH. An emulsion formed in the organic layer which separated in a separatory funnel and was collected. The aqueous layer was washed with a further two 20ml aliquots of diethyl ether. The organic layers were combined and evaporated down to a white oil. The oil was taken up into 5ml of water and the pH raised to

9 by adding a quantity of 0.25M NaOH. The product was separated into chloroform which was collected and dried over Na_2SO_4. 1 equiv. (1.1ml) 36% HCl, based on the complete conversion to t-butyl ester, was added to the solution. The solvent was evaporated to an oil and the protected amino acid recrystallised from hot petroleum ether 60:40, as the HCl salt, with a yield of 2.38g (90%). R_f (G) 0.8. ^{1}H-NMR ($CDCl_3$): 1.5 (s, 9H, $-OC(CH_3)_3$). MS: 244 ($[M+H]^+$); 188 ($[M+H]^+ - C(CH_3)_3$); 142 ($[M+H]^+ - COOC(CH_3)_3$).

Synthesis of Boc-DL-α-aminodecanoic acid (15)

10g (53.5mmol) (DL)-α-aminodecanoic acid (**8**) was suspended in 150ml of tert. BuOH/water (2:3). The pH was adjusted to approx. 13 using 8M NaOH. Whilst stirring at r.t. 17.3g (80.4mmol) ditertiary-butyl-dicarbonate dissolved in 100ml tert. butanol was added slowly. The pH of the solution was maintained at 12-13 with the 8M NaOH. Stirring was continued for 2h with frequent monitoring of the pH. The reaction mixture was poured onto 100ml water and the pH reduced to 3 with citric acid. The aqueous solution was washed with 5x50ml EtOAc to extract the Boc protected amino acid The organic phases were combined and shaken with 2x50ml brine. The EtOAc solution was dried over Na_2SO_4 and evaporated to dryness on a rotary evaporator. The residue was taken up in 200ml AcCN and left o.n. at 4°C. The precipitated solid was filtered off and the AcCN solution evaporated to an oil. A further 10ml AcCN was triturated into the oil at 0°C and then left at 4°C for 2h. A yield of 10.3g (66.8%) of a sticky off-white solid was obtained. R_f (C) 0.8. ^{1}H-NMR ($CDCl_3$): 1.5 (s, H-C(9), Boc); 5.0 (m, 1H, αC-H). MS: 287 ($[M+H]^+$); 186 ($[M+H]^+$ - Boc).

Synthesis of DL-α-aminodecanoic acid, methyl ester (16)

7.8ml (0.11mmol) of freshly distilled thionyl chloride was added dropwise to 60ml MeOH, cooled to 0°C, over a period of 5 mins. 4g (21.4mmol) DL-α-aminodecanoic acid (**8**) was added and the reaction mixture refluxed for 22h. After cooling to r.t. the solvent was evaporated off and the methyl ester

recovered by recrysallisation from $MeOH/Et_2O$. The semi-purified product was recrystallised a second time from $MeOH/Et_2O$ to yield 3.53g (69%). R_f (C) 0.5. ^{1}H-NMR ($CDCl_3$): 3.75 (s, 3H, $-OCH_3$). MS: 202 ($[M+H]^+$); 185 ($[M+H]^+ - CH_3$); 142 ($[M+H]^+ - COOCH_3$).

Synthesis of di-DL-α-aminodecanoic acid derivative (18)

2.42g (8.42mmol) of Boc-DL-α-aminodecanoic acid (**15**) was reacted with 2g (8.42mmol) DL-α-aminodecanoic acid, methyl ester (**16**) to form the di-peptide using standard solution phase conditions. The Boc protected di-peptide (**17**) was purified by flash chromatography (B) to yield 2.74g (69%). The Boc protected dipeptide (2.74g, 5.83mmol) was treated with 50% TFA in DCM for 30 mins at r.t. to remove the Boc group. After purification by flash chromatography (C) a yield of 2.72g (97%) was obtained for compound **18**. R_f (C) 0.4 and 0.6 (diastereoisomers). ^{1}H-NMR ($CDCl_3$): 4.05 and 4.5 (m, 2H, αC-H); 1.3-1.9 (m, 28H, alkyl H); 0.9 (t, 6H, alkyl $-CH_3$). MS: 370 ($[M+H]^+$).

Synthesis of tri-lysine derivative (19)

Boc-L-Lys(2Cl-Z)-OH (1.88g, 4.53mmol) was coupled to $HCl.NH_2$-L-Lys(Z)-OMe (1.5g, 4.53mmol) using standard solution phase reactions. Similarly the tri-peptide was synthesized and purified by flash chromatography (C) to yield the pure tri-peptide (**19**), with an overall yield of 3.15g (67.7%). R_f (C) 0.35. ^{1}H-NMR ($CDCl_3$): 1.3-1.8 (m, 24H, alkyl H); 3.75 (s, 3H, $-OCH_3$); 4.1-4.5 (m, 3H, α-H); 5.0-5.2 (m, 6H, CH_2-(Z)); 7.2-7.35 (m, 13H, aryl H (Z)) MS: 887 ($[M+H]^+$).

Synthesis of tri-glutamic acid derivative (20)

The solution phase chemistry used for the synthesis of the tri-lysine derivative was utilised for the production of an oligo-glutamic acid molecule. After purification by flash chromatography (D), the overall yield of compound **20** was 43.7%. R_f (C) 0.45. ^{1}H-NMR ($CDCl_3$): 1.1-2.8 (m, 22H, alkyl H); 4.3-4.7 (m, 3H, α-H); 5.1 (s, 2H, $-CH_2$ (Z)); 7.3-7.35 (m, 5H, aryl H (Bzl)). MS: 634 ($[M+H]^+$).

Synthesis of hydrophobic 4-COR5-Fmoc derivative, 9:R5

600mg (2.5mmol) compound **1** was suspended in 15ml DCM. The addition of 435µl (2.5mmol) DIEA, 696mg (2.5mmol) DL-α-aminodecanoic acid, t-butyl ester (**14**) and 383mg (2.5mmol) HOBt caused **1** to dissolve. The reaction mixture was cooled to 0°C and whilst stirring 525mg (2.74mmol) EDC was added. Stirring was continued for 1h at 0°C and at r.t. overnight. The reaction mixture was diluted with 20ml DCM and shaken with water, 0.1M HCl, 3% $NaHCO_3$, water, dried over Na_2SO_4 and evaporated *in vacuo* to a pale yellow oil. Purification was achieved on silica gel using solvent system H. Fractions of pure material were combined and evaporated to yield 715mg (61.5%) of a pale yellow solid (**9**:R5). R_f (H) 0.5. ^{1}H-NMR ($CDCl_3$): 1.5 (s, 9H, $-OC(CH_3)_3$); 1.3-1.4 (m, 14H, alkyl H); 4.0-4.1 (m, 3H, H-C(9), $-CH_2OH$); 4.8 (m, 1H, α-H); 7.2-7.6 (m, 6H, aryl H); 8.0 (d, 1H, H-C(3)). MS: 467 ($[M+H]^+$).

Synthesis of hydrophobic 4-COR6-Fmoc derivative, 9:R6

1.5g (6.25mmol) 9-(hydroxymethyl)fluorene-4-carboxylate (**1**) was dissolved in 6ml of a DCM:DMF mixture (1:1) and cooled to 0°C. HOBt hydrate (0.96g, 6.25mmol) and EDC (1.31g, 6.88mmol) were added to the reaction mixture, followed by the dropwise addition of 3ml of a DCM solution containing **18** (3g, 6.25mmol) and DIEA (1.09ml, 6.25mmol). The reaction was left stirring at 0°C for 1h then for 3h at room temperature. At the end of the reaction the urea derivative and any other impurities were extracted into water, 0.1M HCl, 3% $NaHCO_3$, water, dried over Na_2SO_4 and evaporated *in vacuo* to a yellow sticky solid. The crude product was purified by flash chromatography (D) to give 2.53g (68%). R_f (D) 0.5 and 0.6 (diastereoisomers). ^{1}H-NMR ($CDCl_3$): 0.9-2.0 (m, 34H, alkyl H); 3.6 (d, 3H, $-OCH_3$); 3.8-3.95 (m, 3H, H-C(9), $-CH_2OH$); 4.5 and 4.9 (m, 2H, α-H); 7.2-7.6 (m, 6H, aryl H); 7.95 (d, 1H, H-C(3)). MS: 592 ($[M+H]^+$).

Synthesis of positively charged 4-COR7-Fmoc derivative, 9:R7

9-(hydroxymethyl)fluorene-4-carboxylate (**1**) (0.80g, 3.35mmol) was dissolved in 5ml of a DCM:DMF mixture (2:3) and cooled to 0°C. The HOBt (0.51g, 3.35mmol) mediated coupling reaction using EDC (0.70g, 3.69mmol) to the tri-lysine derivative **19** (3.35g, 3.35mmol) was performed in the presence of DIEA (0.58ml, 3.35mmol), for 1h at 0°C and then 3h at room temperature. Impurities and the urea derivative were extracted into water, 0.1M HCl, 3% $NaHCO_3$, water, dried over Na_2SO_4 and evaporated *in vacuo* to a yellow sticky solid. The pure 4-COR-Fmoc derivative was purified by crystallization from hot AcCN; 2.6g (70%). R_f (C) 0.6. ^{1}H-NMR (DMSO-d_6): 1.3-2.4 (m, 24H, alkyl H); 3.6 (s, 3H, -OCH_3); 3.8 (m, 2H, -CH_2OH); 4.0 (t, 1H, H-C(9)); 4.1-4.5 (m, 3H, α-H); 5.0-5.05 (m, 6H, CH_2-(Z and 2Cl-Z)); 7.2-7.3 (m, 13H, aryl H (Z and 2Cl-Z)); 7.4-7.8 (m, 6H, aryl H); 8.0 (d, 1H, H-C(3)). MS: 1110 ($[M+H]^+$).

Synthesis of negatively charged 4-COR8-Fmoc derivative, 9:R8

The coupling reaction to the 9-(hydroxymethyl)fluorene-4-carboxylate group (**1**) was performed using the same conditions described for **9**:R7. After flash chromatography (D) 81% of purified **9**:R8 was obtained. R_f (D) 0.25. ^{1}H-NMR (DMSO-d_6): 1.1-2.4 (m, 22H, alkyl H); 3.8 (m, 2H, -CH_2OH); 4.0-4.1 (m, 5H, -OCH_2 -H-C(9)); 4.2-4.6 (m, 3H, α-H); 4.65 (m, 1H, -OCH (cHx)); 5.05 (s, 2H, -CH_2 (Bzl)); 7.2-7.9 (m, 11H, aryl H (Bzl and fluorene)); 7.95 (d, 1H, H-C(3)). MS: 856 ($[M+H]^+$).

Synthesis of 4-COR5-Fmoc succinimidyl carbonate, 10:R5

530mg (1.14mmol) 4-COR-Fmoc derivative **9**:R5 was dissolved in 15ml acetonitrile, that had previously been dried over molecular sieves. To the reaction vessel was added 1 equiv. dry pyridine (92µl, 1.14mmol) and 2 equivs. DSC (583mg, 2.28mmol), then left to stir at r.t. for 7h. At the end of the reaction the AcCN was evaporated *in vacuo*, the residue dissolved in chloroform and washed with H_2O, 0.1M HCl, H_2O, dried over Na_2SO_4 and evaporated *in vacuo* to a pale yellow solid, 670mg (97%) of crude material. Compound **10**:R5

was used without further purification. R_f (I) 0.65. ^{1}H-NMR ($CDCl_3$): 2.8 (m, 4H, CH_2(ONSu)). MS: 607 ($[M+H]^+$).

Synthesis of 4-COR6-Fmoc succinimidyl carbonate, 10:R6

The same procedure as described for **10**:R5 was used to synthesise **10**:R6. 1.25g (2.11mmol) **9**:R6 was treated with 1 equiv. dry pyridine (160µl, 2.11mmol) and 1.08g (4.22mmol) DSC. 1.66g (107%) crude material was purified by flash chromatography on silica gel (E) to give 1.03g (67%) of the succinimidyl carbonate. R_f (E) 0.4 and 0.25 (diastereoisomers). ^{1}H-NMR ($CDCl_3$): 2.8 (m, 4H, CH_2(ONSu)). MS: 733 ($[M+H]^+$).

Synthesis of 4-COR7-Fmoc succinimidyl carbonate, 10:R7

The same procedure as described for **10**:R5 was used to synthesise **10**:R7. 500mg(0.45mmol) **9**:R7 was treated with 1 equiv. dry pyridine (34µl, 0.45mmol) and 2 equivs. DSC (230mg 0.9mmol). 632g (112%) crude material was purified by crystallisation from absolute ethanol to give 418mg (74%) of the succinimidyl carbonate. R_f (C) 0.6. ^{1}H-NMR (DMSO-d_6): 2.6 (s, 4H, CH_2(ONSu)). MS: 1250 ($[M+H]^+$).

Synthesis of 4-COR8-Fmoc succinimidyl carbonate, 10:R8

The same procedure as described for **10**:R5 was used to synthesise **10**:R8. 260mg (0.3mmol) **9**:R8 was treated with 1 equiv. dry pyridine (23µl, 0.3mmol) and 2 equivs. DSC (155mg 0.61mmol). After purification by flash chromatography (K), 290mg (96%) was obtained. R_f (F) 0.65. ^{1}H-NMR (DMSO-d_6): 2.6 (s, 4H, CH_2(ONSu)). MS: 995 ($[M+H]^+$).

Synthesis of 4-COR5-Fmoc-glycine derivative, 11:R5

0.67g (1.11mmol) **10**:R5 was dissolved in 5ml acetone. 91mg (1.22mmol) glycine was dissolved 5ml water containing 124mg (1.22mmol) sodium carbonate. The solution containing glycine was added dropwise to the **10**:R5. A precipitate developed which dissolved after 5 min. Stirring was continued at r.t.

for 2h 20 mins. A further 10ml water was added and the solution acidified to pH3 with 0.1M HCl. The product was extracted into 3 x 20ml portions of chloroform, which were combined, washed with water, dried over Na_2SO_4.and evaporated to an oil. The product was purified on a HPLC silica gel column, using a hexane/isopropanol gradient (10% iPrOH to 50% iPuOH in 30 min) to yield 167mg (21.3%). R_f (J) 0.5. ^{1}H-NMR (DMSO-d_6): 3.2 (m, 2H, CH_2(Gly)). MS: 567 ($[M+H]^+$).

Synthesis of 4-COR6-Fmoc-glycine derivative, 11:R6

0.99g (1.36mmol) **10**:R6 was dissolved in 3ml DCM and 0.5-1ml DMF to aid dissolution. 1 equiv. of HOBt hydrate (207mg, 1.36mmol) was then added, followed by 1.1 equiv. t-butyl-glycine hydrochloride (285mg, 1.49mmol) and 1.1 equiv. DIEA (260µl, 1.49mmol). After 1h stirring at room temperature any impurities were removed by extraction into water and the glycyl derivative purified by flash chromatography (E) to give 0.72g (72%) product. R_f (E) 0.5 and 0.55 (diastereoisomers). ^{1}H-NMR (DMSO-d_6): 1.4 (s, 9H, -OtBu); 3.65 (d, 2H, CH_2(Gly)). MS: 749 ($[M+H]^+$). The t-butyl group was removed using 50% TFA in DCM at r.t. for 30 min. After evaporation of the solvent, purification was achieved by flash chromatography (chloroform/AcOEt 19:1) leaving a pale yellow solid (**11**:R6), 0.73g (77%). R_f (C) 0.1. ^{1}H-NMR (DMSO-d_6): 3.65 (m, 2H, CH_2(Gly)). MS: 693 ($[M+H]^+$).

Synthesis of 4-COR7-Fmoc-glycine derivative, 11:R7

1.2g (0.96mmol) **10**:R7 was treated as described for the synthesis of **11**:R6, using 147mg (0.96mmol) HOBt hydrate 202mg (1.06mmol) t-butylglycine and 184µl (1.06mmol) DIEA. Purification was achieved using flash chromatography (C) to give 1.11g (91%). R_f (C) 0.5. ^{1}H-NMR (DMSO-d_6): 1.4 (s, 9H, -OtBu); 3.65 (s, 2H, CH_2(Gly)). MS: 1266 ($[M+H]^+$). 1.1g (0.87mmol) of the t-butyl protected product was treated with 50% TFA in DCM at r.t. for 30 min. The final product was purified by flash chromatography (J) to yield 0.76g (72%) of the correct material **11**:R7. R_f (L) 0.5. ^{1}H-NMR (DMSO-d_6): 3.6 (s, 2H, CH_2(Gly)). MS: 1211 ($[M+H]^+$).

COUPLING CHROMATOGRAPHIC PROBES TO PEPTIDYL-RESIN

The peptidyl-resin was prepared for derivatisation by treating with 50% TFA in DCM for 30 mins to allow N-terminal deprotection. The TFA salt was neutralised by shaking for 2x2 mins with 5% DIEA in DCM. The deprotected peptidyl-resin was transferred to a modified ABI reaction vessel, sealed at one end to prevent loss of material. Then, depending on the route of introduction (Scheme 3, p.51), one of the following procedures was performed:

Reaction with 4-COR-Fmoc succinimidyl carbonate derivatives (Route 1)

To the swollen, freshly deprotected peptidyl-resin was added 4 equivs. of the 4-COR-Fmoc succinimidyl carbonate, based on the substitution value upon the addition of the last amino acid in the sequence. The active carbonate was added dissolved in sufficient DCM to give a final molarity of 0.03M. The reaction mixture was vortexed and aliquots of resin were removed at half hourly intervals, to determine the degree of incorporation of the 4-COR-Fmoc derivative onto the peptide. The sample of peptidyl-resin was washed with DCM and dried from MeOH/DCM (1:1). The dried resin was weighed and treated with 20% piperidine in DMF for 15 mins at r.t. The concentration and therefore the degree of incorporation was obtained from the absorbance reading at 270nm, due to the released Fmoc moiety in the supernatant. When two consecutive readings corresponded to eachother the reaction was terminated. The derivatised peptidyl-resin was washed well with DCM and dried from MeOH/DCM (1:1).

Reaction with glycine derivatised 4-COR-Fmoc groups, 11:R5-7 (Route 2)

In the case where the preformed amino acid 4-COR-Fmoc-Gly derivative had been synthesised, the following procedure was performed. The freshly N-terminal amino deprotected peptidyl-resin was left swollen in DCM. Coupling of the 4-COR-Gly-Fmoc compounds (2 equivs. based on the substitution value

of the last residue) was achieved using 2.2 equivs. DCC, 2 equivs. HOBt and 2 equivs. DIEA in DCM. To the reaction vessel containing the resin was added the activated 4-COR-Fmoc-Gly derivative and then vortexed until the coupling had gone to completion (approx. 1-2.5h), as described above. The derivatised peptidyl-resin was then transferred to a sintered glass filter and washed with DMF, DCM and finally dried from MeOH/DCM(1:1).

REMOVAL OF SIDE CHAIN PROTECTING GROUPS AND CLEAVAGE OF PEPTIDE FROM RESIN

Removal of DNP from His (where appropriate)

The peptidyl-resin was treated with a solution containing 20% β-mercapto-ethanol and 10% DIEA in DMF for 30 mins. The resin was washed with DCM and treatment repeated for a further 30 mins. The resin was finally washed with DCM.

Removal of N-terminal Boc protecting group

The peptidyl-resin was swollen in DCM and treated with 50% TFA in DCM twice, for 1 min then 30 mins. The resin was washed and neutralised with 5% DIEA in DCM. Following a DCM wash the peptide-resin was dried from MeOH/DCM (1:1).

Removal of formyl from Trp (where appropriate)

The resin was transferred to a 50 ml round-bottomed flask and cooled to 0°C. The resin was suspended in 15ml DMF, 1ml ethanolamine and 1ml water. The mixture was stirred gently for 5 mins. The resin was filtered off and washed with DCM, TFA and dried from MeOH/DCM (1:1).

High HF only

The cleavage conditions used for short peptides or peptides with simple sequences consisted of a High HF protocol only. A maximum of 500mg of peptidyl-resin, following pre-treatment for Boc, His (DNP) or Trp (formyl) as necessary, was added 800µl *p*-cresol and 200µl *p*-thiocresol. 10ml HF was distilled over and the cleavage reaction conducted at 0°C for 1h. The work-up of the free peptide was as described for the Low TFMSA/High HF protocol.

LOW TFMSA/HIGH HF CLEAVAGE PROTOCOL

Low TFMSA

The peptidyl resin was treated as described above depending on the residues present in the sequence and transferred to a teflon reaction vessel. For 100mg of resin the following scavengers were added: 20µl 1,4-butanedithiol (fresh); 80µl *p*-cresol and 300µl dimethyl sulphide (fresh). The resin was cooled to 0°C to which was added 500µl TFA. Whilst stirring 100µl TFMSA was slowly added dropwise. Stirring at 0°C was continued for 2h. After 2h the resin was filtered and washed with diethyl ether, then DCM and dried under vacuum. The partially deprotected resin was then ready for the High HF reaction.

High HF

The scavengers used for the two stage cleavage protocol were 500µl each of 1,4-butanedithiol and *p*-cresol. After distilling 10ml of HF into the reaction vessel, magnetic stirring of the peptidyl resin was continued at 0°C for 1h. The HF was distilled off and the fully cleaved peptide precipitated with dry Et_2O. The resin and peptide were collected on a glass filter (pore size no. 3). After several washes with Et_2O to remove scavengers, the peptide was dissolved in 3ml aliquots of 10% AcOH. If the peptide was particularly hydrophobic or dissolved preferentially in a basic solution then a little AcCN or a lyophilisable salt (eg. ammonium acetate) respectively, was used. The crude peptide was recovered from solution by lyophilisation.

Cleavage of His (DNP) from free peptide:

The peptide was dissolved in a solution of 6M guanidine HCl, 50mmol Tris acetate at pH 8.5, at a concentration of 10-20mg peptide/ml. β-mercaptoethanol to a concentration of 20% v/v was added and the reaction mixture heated to 30°C for 4h. The solution was diluted with water and transferred to a dialysis bag (Mwt cutoff=8000) and dialysed against 10mmol AcOH at 4°C. The dialysis was continued for 3 days with regular change of the dialysis solvent.

FRAGMENT CONDENSATION

Synthesis of protected peptides on photolabile handle (21)

0.72g (2.76mmol) bromomethyl-3-nitrobenzoic acid (**25**) was dissolved in 5ml DCM. To the solution was added 5.52ml (2.76mmol) of a 0.5M solution of DCC. The reaction mixture was cooled to 4°C and left stirring for 1h. The temperature was then raised to r.t. and the precipitated DCU by-product filtered off. 1.39g (1mmol) of MBHA.HCl resin (substitution value; 0.72mmol/g) was treated with 30%TFA in DCM for 15 min, washed with four aliquots of DCM and then neutralised with 5% DIEA in DCM. The activated solution of **25** was then added to the resin and shaken for 1h. Quantitative ninhydrin analysis on an aliquot of resin indicated that complete derivatisation had occured. The yield of derivatised MBHA resin (**22**) was 1.54g (96%).

0.88g (5mmol) Boc-Gly-OH was dissolved in 20ml ethanol/water (3:1). The pH of the solution was adjusted to pH7 with a 1.5M solution of caesium carbonate. The volatile solvent was evaporated off on a rotary evaporator and the remaining solvent removed by lyophilisation. The resulting material was then left o.n., under vacuum over phosphorus pentoxide. 1.54g (1mmol) of resin-handle (**21**) was then suspended in 30ml DMF, to which was added 0.60g (1.95mmol) of the Boc-Gly-O^-Cs^+ salt. The reaction mixture was heated to 40°C and mechanically stirred for 18h. The resin was then filtered from the reaction mixture and washed with DMF, followed by DMF/water (4:1), DMF, DCM and then finally dried from DCM/MeOH (1:1) to give 1.63g of derivatised resin-handle. Amino acid analysis of the resin showed that 33% coupling of Boc-Gly had been achieved. Any reactive groups remaining were then capped using 5% acetic anhydride in DMF and treating for 10 min.

The synthesis of fragments 1-15 and 16-32 were then performed on the automated peptide synthesiser as described previously. The exception was the third amino acid, which was incorporated manually using the method described by Suzuki and Endo [1977]. The N-terminal protected peptidyl-resin, containing the first two amino acids, was removed from the reaction vessel and washed

with DCM and dioxane. The peptidyl-resin was then treated with 4N HCl in dioxane for 0.5 min, followed by a second 30 min treatment. The resin was washed with dioxane and DCM. To the peptidyl-resin was added a DCM solution containing 2.5 equivs. DCC, followed by a 3 min reaction with 2.5 equivs. of the third amino acid, as the N-methylmorpholine salt dissolved in DCM. The peptidyl-resin was then washed with DCM, DMF and returned to the peptide synthesiser reaction vessel to complete the synthesis.

Synthesis of protected peptides on base-labile handle (22)

1g (0.62mmol) of MBHA.HCl resin (substitution; 0.62mmol/g) was shaken with 50% TFA in DCM for 15 min and then neutralised with 5% DIEA in DCM. 0.3g (1.24mmol) compound **1** was dissolved in 4ml DCM/DMF (1:1) and activated with 2.73ml (1.37mmol) of a 0.5M solution of DCC. The reaction was allowed to proceed for 1h at 0°C, at which point the precipitated DCU was filtered from the solution. The activated form of **1** was then transferred to the suspension of MBHA resin. After 6h vortexing, quantitative ninhydrin analysis showed that 90% incorporation had occured, yielding 0.98g of derivatised resin.

0.2g (1.15mmol) Boc-Gly-OH was dissolved in 3ml DCM to which was added 2.51ml (1.26mmol) 0.5M DCC solution and 5.6mg (46μmol) dimethylamino-pyridine (DMAP). 0.45g (0.23mmol) of derivatised MBHA resin was swollen in 2ml DMF and then added to the reaction mixture. The coupling reaction was conducted for 2h when amino acid analysis indicated that 45% incorporation of Boc-Gly had been achieved, 0.31g. Any unreacted groups were capped with a solution containing acetic anhydride/pyridine/DCM (1:1:2) for 30 min [Liu *et al*, 1990].

Synthesis of the protected peptides was then performed on the automated peptide synthesiser using the Boc chemistry protocols described previously.

Synthesis of protected peptides on base-labile handle (23)

0.63g (0.5mmol) of MBHA.HCl resin (substitution; 0.81mmol/g) was shaken with 50% TFA in DCM for 15 min and then neutralised with 5% DIEA in DCM. An internal standard, Val was then incorporated into peptidyl-resin. 0.43g (2mmol) Boc-Val-OH was dissolved in 3ml DCM and activated with 4.4ml (2.2mmol) of 0.5M DCC. The DCU was filtered off and the activated amino acid transferred to the MBHA resin. After 1h, the incorporation was complete as determined by quantitative ninhydrin analysis. The resin was washed with DCM/MeOH (1:1) and dried to yield 0.72g. The N-terminal Boc group was removed with a 30 min treatment with 50% TFA in DCM and then neutralised with 5% DIEA in DCM. 0.14g (0.68mmol) 4-(hydroxyethyl)-5-nitrobenzoic acid (**26**) was activated with 0.14g (0.68mmol) DCC and 0.11g (0.68mmol) HOBt. Coupling was complete after 2h.

0.7g (4mmol) Boc-Gly-OH was then activated with 0.82g (4mmol) DCC and the coupling catalysed with 9.8mg (80μmol) DMAP. The activated amino acid was added to the derivatised MBHA resin and left to react for 2h. The amino acid resin was washed with DCM and dried from DCM/MeOH (1:1) to yield 0.96g. Amino acid analysis indicated that 99% coupling had been achieved.

Solid phase fragment condensation

Fragment condensation was performed in 300μl micro-reaction vessels. The protected peptide was dissolved in sufficient solvent together with activating agents to give a concentration of approx. 0.07M and activation was allowed to occur for 30 min. The N-terminal protecting group was removed from the peptidyl-resin with either 50% TFA in DCM or 20% piperidine in DMF, depending on whether Boc or Fmoc chemistry had been used for the synthesis. In the case, of the former a neutralisation step was included with 10% DIEA in DCM. The resins were washed, dried from DCM and then immediately transferred to the reaction vessel containing activated protected peptide. The reaction vessel was left at r.t. with occasional vortexing. The coupling was followed by removing an aliquot of peptidyl-resin and analysing the amino acid content. When necessary a second coupling was performed by washing the

reagents from the resin on a sintered glass filter and then adding the resin to freshly activated protected peptide. The resin was finally washed and dried.

Cleavage of photolabile handle (21)

The photolysis reaction was performed in the laboratory of Dr. F. Albericio, University of Barcelona. The cleavage involved irradiating the peptidyl-resin, suspended in DCM:trifluoroethanol (7:3), with ultraviolet light (λ=350nm) for 24h. The solvent was then removed on a rotary evaporator.

Cleavage of base-labile handles (22) and (23)

For 100mg of peptidyl-resin, base cleavage of the protected peptide was achieved using 1ml 20% piperidine in DMF for 15 min. The piperidine/DMF solvent containing the protected peptide was collected and the peptidyl-resin treated with another aliquot of 20% piperidine in DMF. The combined solution was then reduced in volume on a rotary evaporator.

Cleavage of Sasrin resin

(a) protected peptide

100mg of peptidyl-resin was swollen in DCM, drained and then treated with 5ml 1% TFA in DCM for 10 min. The solution was filtered off and neutralised with 0.3ml DMF. The procedure was repeated three times and the DCM solvent, containing the protected peptide combined. The DCM solution was then either shaken with aqueous liquid for solvent extraction of removed on a rotary evaporator for chromatographic purification.

(b) full cleavage

Before deprotection, the N-terminal Fmoc protecting group was removed with 20% piperidine in DMF. For quantities of peptidyl-resin ≤100mg the following proportions were used. Scavengers: 0.05g phenol, 50µl thioanisole, 25µl ethanedithiol and 50µl water. 1ml of TFA was then added and the mixture stirred at r.t. for 3.5h. The residual resin was filtered off using glass wool and the fully deprotected peptide precipitated with diethyl ether. The precipitate was separated by centrifugation and washed twice with further diethyl ether. Finally, the peptide was dissolved in 10% acetic acid solution and lyophilised.

APPENDIX I

CHROMATOGRAPHIC CONDITIONS

Eluent A: water/0.045% TFA
Eluent B: acetonitrile/0.036% TFA
Flow: 1 ml/min
Detection: 220nm
(unless otherwise stated)

System 1

Eluent A: water/0.1% TFA
Eluent B: 60% acetonitrile/0.1% TFA
Detection: 218nm
Column: Vydac C_{18} RP (150x4.6mm)
Gradient: 0% to 55%B in 15 min

System 2

Eluent A: water/0.1% TFA
Eluent B: acetonitrile/0.1% TFA
Column: Vydac C_{18} RP (150x4.6mm)
Gradient: 0% to 100%B in 30 min

System 3

Eluent A: water/0.1% TFA
Eluent B: acetonitrile/0.1% TFA
Column: Brown-Lee C_8 RP (30x4.6mm)
Gradient: 0%B for 5 min, then 0 to 100% in 80 min

System 4

Eluent A: water/0.1% TFA
Eluent B: acetonitrile/0.1% TFA
Detection: 216nm
Column: Vydac C_4 RP (150x4.6mm)
Gradient: 0% to 100%B in 40 min

System 5

Column: Vydac C_4 RP (150x4.6mm)
Gradient: 0% to 100%B in 80 min

System 6
Column: Brown-Lee C_8 RP (30x4.6mm)
Gradient: 0%B for 5 min, then 0 to 100%B in 80 min

System 7
Column: Vydac C_{18} RP (150x4.6mm)
Gradient: 0%B for 5 min, then 0 to 100%B in 80 min

System 8
Buffer A: 20mM phosphate pH6.5
Buffer B: 20mM phosphate pH6.5 + 0.5M NaCl
Detection: 263nm
Column: Pharmacia Mono Q (50x5mm)
Gradient: 0 to 40%B in 30 min

System 9
Buffer A: 20mM phosphate pH6.5
Buffer B: 20mM phosphate pH6.5 + 0.5M NaCl
Detection: 263nm
Column: Pharmacia Mono Q (50x5mm)
Gradient: 0 to 40%B in 35 min

System 10
Column: Vydac C_{18} RP (150x4.6mm)
Gradient: 20% to 100%B in 30 min

System 11
Column: Vydac C_{18} RP (150x4.6mm)
Gradient: 20% to 80%B in 30 min

System 12
Column: Brown-Lee C_8 RP (30x4.6mm)
Gradient: 0% to 100%B in 80 min

System 13
Column: Brown-Lee C_8 RP (30x4.6mm)
Gradient: 40%B for 5 min, then 40 to 100%B in 15 min

System 14
Column: Brown-Lee C_8 RP (30x4.6mm)
Flow: 5ml/min
Gradient: 40%B for 5 min, then 40 to 100%B in 58 min

System 15
Column: Brown-Lee C_8 RP (30x4.6mm)
Gradient: 30%B for 2 min, then 30 to 100%B in 45 min

System 16
Column: Vydac C_{18} RP (150x4.6mm)
Gradient: 0%B for 5 min, then 0 to 100%B in 40 min

System 17
Column: Pharmacia Sephacryl S-100
Solvent: 100mM Tris-HCl, 10mM $MgCl_2$, pH 7.8
Flow: 2.5ml/min
Detection: 276nm

System 18
Column: Merck C_{18} RP-100 (100x4.6mm)
Gradient: 0 to 50%B in 30 min

APPENDIX II

The following is a draft copy of a paper in preparation for publication:

THE STRUCTURAL AND AGGREGATION PROPERTIES OF THE SYNTHETIC C-TERMINAL HALF (104-MER) POLYPEPTIDE FROM HIV p24gag RESEMBLES THOSE OF FULL-LENGTH PROTEIN

H.L. Ball, E. Chan, E. Hallakova, W.A. Gibbons, A. Coates and P. Mascagni

INTRODUCTION

The chemical synthesis of the C-terminal half sequence ([Ala^{330}, Ala^{350}, 270-373) of the HIV p24gag protein (hereafter referred to as p24) and its immunological properties have been reported (ref).

p24 is believed to be self-associated in the HIV core (ref) although it is not known yet whether self-association initiates with the p24 precursor (p55) or whether it takes place after p55 processing. Assembly of p24 is however believed to be necessary for virion maturation. Recently (refs) a possible homology between p24 and picornaviral VP2 coat protein has been suggested. According to these proposals p24 has the eight-stranded anti-parallel β barrel typical of other viral capsid proteins. These predictions contrast with those which assigned to p24 a predominantly α-helix secondary structure (ref) and the secondary structure composition of recombinant protein calculated by Circular Dichroism (CD) spectroscopy (ref). In this latter study the CD spectrum of SIV p27 was assigned to an a+b structure containing 44% alpha helix and 56% beta sheet (ref).

The secondary structure of p24 calculated from CD spectroscopy is likely to be that of a dimeric form of the protein. Thus self-association which occurs in vivo during virus particle assembly has been demonstrated to occur in-vitro by analytical affinity chromatography (ref). The results show that in solution containing up to ≤400μg/ml p24 dimerizes, the dissociation constant being between $1\text{-}4\text{x}10^{-5}$M at pH 7. However at concentrations (80mg/ml) close to the

estimated concentration of p24 in the virion, self-association seems to proceed predominantly to dimer with some population of trimers or tetramers (ref).

In view of the above background the characterisation of the conformational properties of synthetic 104-mer and its ability (if any) to self-associate was carried out. Naturally during the course of the study the question arose as to whether molecules like the 104-mer can fold into structures contained in the full-length protein. Thus in the discussion that follows attempts will be made to correlate the structural findings for the synthetic 104-mer with the properties of p24.

RESULTS AND DISCUSSION

The first step in this study was to purify synthetic 104-mer using the same protocol described earlier (ref). The salient points of this purification scheme which bear on the structure of the synthetic polypeptide are discussed in the following section.

Purification of synthetic 104-mer (method A)

After cleavage from the resin crude polypeptide was dissolved in 6M guanidine HCl, 50mM Tris acetate 20% 2-mercaptoethanol, pH 8.5 and stirred at room temperature for 2 hrs to remove residual His side-chain protecting groups. The solution was then dialysed against 150mM ammonium acetate for 2 days at 4°C to remove smaller impurities. The content of the dialysis tubing was then lyophilized and subjected to semi-preparative RP-HPLC. A broad peak eluting between 70 and 80% acetonitrile was collected the solvent reduced by evaporation in vacuum and the resulting solution loaded onto a preparative size exclusion column. The main fraction, shown by SDS-PAGE to contain synthetic polypeptide with the expected molecular weight (data not shown), was lyophilized and further purified by preparative cationic-exchange chromatography. The use of the latter was suggested by (i) isoelectrofocusing gel electrophoresis of crude 104-mer which revealed two main components focusing at approximately 9.3 and 5.6 respectively (Fig. 1; lane 1) and (ii) the theoretical pI of 104-mer which was calculated to be 8.85.

Thus a solution containing approximately 10 mg of lyophilised material per ml of buffer was loaded onto the column and the material which eluted at 0.13M collected (Fig. 2A). SDS-PAGE (Fig. 3A) indicated that the latter had the expected molecular weight and was a single component when re-injected on analytical ion-exchange chromatography (Fig. 2B). Further confirmation of its homogeneity was obtained by isoelectrofocusing electrophoresis which indicated that the polypeptide focused as a single band with a pI of 9.3 (Fig. 3B).

Although the material obtained from this purification was a single component by ion-exchange analysis (Fig. 4A) and electrophoresis analysis, during subsequent experiments it was noticed that its chromatographic profile would slowly change as a function of time. In particular when a solution containing purified 104-mer and stored at -20°C for several days was subjected to ion-exchange chromatography a new component eluting at the beginning of the salt gradient was observed in the chromatogram (Fig. 4B). Furthermore, the concentration of this new component increased upon dilution of the solution stored at -20°C indicating an equilibrium between either two differently folded proteins or different aggregates of the same species (Fig. 4C). Repetition of these dilution experiments using size-exclusion chromatography confirmed the second of the two hypotheses since the two ion-exchange components eluted as proteins of molecular weight 53kD and 14kD consistent with tetrameric and monomeric species. Interestingly a third minor component corresponding to a dimeric form of the 104-mer was found by further dilution of monomer isolated from ion-exchange (data not shown). Collectively, these results indicated that:

1) Purification of synthetic 104-mer by the procedure described above had yielded a tetrameric species whose pI was 9.3.

2) Disaggregation to monomers could be achieved by dilution of tetramer solutions. The monomeric species thus obtained seemed to have a pI intermediate between 9.3 and 8 where the latter was the pH of the buffer solution used for ion-exchange chromatography.

3) Dilution of purified monomer solutions yielded a new species whose molecular weight was estimated by size exclusion chromatography to be that of dimer.

While the existence of higher aggregates of the 104-mer was not totally unexpected given the relatively high concentration of protein used during purification (see above) and the tendency shown by p24 to aggregate both in-vivo and in-vitro, the formation of a dimer in dilute monomeric solutions seemed to be consistent with the formation of a new folded state. Furthermore these observations raised the question as to whether the component focusing at pH 5.6 in the IEF gel of crude 104-mer was indeed an impurity as initially concluded.

In light of these observations a new purification by ion-exchange chromatography was carried out. This time the concentration of crude 104-mer was maintained as low as ≤500μg/ml. As expected this removed completely the tetramer. However, the decreased concentration of the latter did not result in the expected increase of monomer concentration. Thus the majority of 104-mer eluted prior to the application of the gradient consistent with the existence of a negatively charged species (results not shown). Only a small amount of monomer was recovered.

Purification of synthetic 104-mer (method B)

The results from the above purification had indicated that different aggregated forms of the peptide existed as an equilibrium in solution and that their relative concentrations were dependent upon the solution conditions. To investigate therefore the nature of the various forms of the peptide and to separate them according to their pI, a second purification protocol was used. This was based on the preparative IEF technique in free solution using the Rotofor cell (ref).

30mgs of crude polypeptide were dissolved in a 6mM NaCl solution (40ml) and the pH adjusted to 7.3 with 0.1N NaOH. After an o.n. standing at 4°C the solution was transferred to the Rotofor cell, ampholytes added and focusing allowed to take place for 5.1 hrs at 4°C. At the end of the run twenty fractions

covering the entire pH gradient (3-10) were collected and an aliquot of each analysed by RP-HPLC. In Figure 5A a diagram of fraction concentration vs pH indicates that the synthetic material existed in solution in three main forms characterized by pIs of 5.6, 8 and 9.3 respectively. This correlated well with the results from the ion-exchange experiments. Thus the material with a pI of 9.3 was the aggregated peptide while the less intense peak at pH 8 was assigned to the monomeric form (M) detected above and whose pI had been estimated to be between 7 and 9.3. Finally, the material focusing at 5.6 corresponded to the peptide unretained by the cationic-exchange medium. The latter had the expected amino acid composition for the 104-mer (data not shown) and the RP-HPLC retention time was the same as that obtained for the various preparations of 104-mer (data not shown). To further characterize this new component, fraction 8 (see Fig. 6A) from the rotofor purification was subjected to anionic-exchange chromatography. In Figure 6A the synthetic polypeptide eluted as a broad and jagged peak at a salt concentration of about 0.1M. Although the results from this chromatography confirmed the existence of a negatively charged folded species, the complexity of the fplc profile was unexpected. Possible explanations to these observations were:

1) fraction 8 contained a mixture of proteins having closely related sequences

2) fraction 8 contained a mixture of proteins having the same sequence but closely related conformations.

Although impurities corresponding to closely related sequences are normally generated during the course of chemical synthesis and their presence is expected in the crude material, it was felt that this was no longer the case after the protein had been purified using the highly selective IEF technique. Therefore to prove the existence of multiple conformations of the same protein ion-exchange chromatography was carried out under a variety of different conditions. These were designed so to alter the conformational equilibria that might be present in solution and included temperature variations and the use of MeOH as co-solvent for ion-exchange chromatography.

Although lowering the temperature to 4°C had important effects on the shape of the chromatographic peak by reducing its width (Fig. 6B), the most dramatic effects were observed upon addition of MeOH. This induced the progressive narrowing of the chromatographic peak and its collapse into a single sharp line for a MeOH concentration of 20% (Fig. 6C). This was isolated and shown to remain a single species when reinjected on the same chromatographic column (Fig. 6D).

That the chromatographic results were to be ascribed to conformational changes was confirmed by CD spectroscopy. Thus the spectrum of fraction 8 in 20% MeOH resembled that of a structure with an appreciable alpha helix content which contrasted with a mainly random coil conformation obtained in aqueous solution (Fig. 8 and see below for discussion).

These results together with those described in the previous section indicated that depending upon the experimental conditions it was possible to isolate three main forms of the synthetic material and distinguishable for their size and pI: a tetramer (T) with pI=9.3 its protomeric form (M) with a pI of 8 and finally a third species which had a pI of 5.6. A purified form of the latter, subjected to size exclusion chromatography, was shown to be the dimer detected in the dilution experiments described in the previous section.

Conformational analysis of purified synthetic 104-mer

To characterize at a structural level the three different forms of 104-mer discussed above, their secondary structure composition was measured using CD spectroscopy and compared to that obtained using secondary structure prediction algorithms. However the first question to answer before addressing the problem of the structure of the synthetic 104-mer and its relationship with p24 was how the folding of the former was influenced by the replacement of the two Cys residues at positions 330 and 350 in the original sequence of p24 with two Ala residues.

The ala-cys replacement

Using chemical synthesis it was shown that a Cys residue can be substituted for the iso-steric Aib amino acid without affecting the correct folding and activity of proteins (refs). However, in the case of the molecule described here two Ala residues were used in place of the two Cys present in p24. Thus in order to assess the effects of this replacement on the folding of the 104-mer a new molecule was synthesised containing the two Cys residues at 330 and 350 (ref to a paper to be published). The sequence covered in this new synthesis was 270-363. The shortening of the original 104-mer sequence by ten residues was suggested by findings that had become available during this study. Thus according to published observations the cleavage of p24 from its p55 precursor takes place between residues 363 and 364 (ref) ten residues down-stream from the initial hypothesis that had guided the synthesis of the 104-mer.

After chain assembly performed using the same chemical protocol applied to the synthesis of the 104-mer, crude 94-mer was dissolved in 6M Guanidine.HCl and extensively dialysed against 10mM Tris.HCl pH 7.8 buffer containing oxidised and reduced glutathione. Folding was allowed to proceed at 4°C for 48 hr. The content of the dialysis tubing was then lyophilised and purified by Rotofor isoelectrofocusing. Figure 5B shows the distribution of 94-mer through the 20 fractions thus collected. Two main products were obtained which corresponded to pI 5-6 and approx. pI 8. These were isolated and by the aid of the NTTB reagent shown not to contain free Cys residues. These results correlated well with those obtained with the 104-mer and indicated that the Cys-Ala replacement had not affected the folding of 104-mer. The only difference noticed between the two molecules was the extent of aggregation to tetramers which appeared to be negligible in the case of 94-mer.

Secondary structure prediction

Having determined that the replacement of Cys with Ala residues had not affected the folding of the synthetic 104-mer its secondary structure was initially studied using prediction algorithms. To this aim seven sequences corresponding to the 270-363 region of the p24 proteins of HIV and related

viruses were extracted from the SWISS-PROT database (ref) and then aligned using a multiple sequence alignment algorithm scoring on sequence homology only (ref). The proteins in the database (Table 1) thus obtained were about 35% homologous to each other. The alignment was then divided into shorter segments where the secondary structure elements were considered separately. The decision on where the breaking point of a secondary structure element takes place was based upon the identification within the alignment of gaps and/or insertions which generally identify loop regions. Although the latter often have defined secondary structures, in the following discussion they were always assigned to coil regions.

The assignment of secondary structure to each segment was based on three different methods, Chou-Fasman (ref), GOR (ref) and Momany (ref). The helical wheel was used to determine whether the section could be an amphiphilic helix. Furthermore assignments were also made based on sequence homologies with fragments found in proteins with known 3D structure. When two or more of these prediction methods were in agreement then the structure was used.
Finally the secondary structures thus assigned were modelled using the Quanta program (ref) and refined by at least 1000 steps of steepest descendent minimization.

The results of this prediction process, summarized in Table 2 and Figure 7, are discussed hereafter.

The sequence 95-104 not present in the p24 proteins considered was assigned to coil consistent with its being at the C-terminus of the synthetic protein. Similarly we assigned to coil the 1-7 region although in the longer p24 this is likely to be a β-strand.

The 8-15 was predicted by all three methods to be a β-strand. Since 70% of its residues are hydrophobic it is likely to be an interior strand. The 16-31 region is the most conserved region among the 7 proteins compared (60% homology with 3 conserved residues). The 16-22 sequence (IRQGPKE) was predicted to be an amphiphilic helix with Q and E on the same side and located at the surface of the protein. The Pro in the middle provided a slight change of

direction in the helix while the Pro residue at position 23 provided a turn followed by another helix (DYVDRF). The 32-35 sequence was assigned to a coil preceding a long amphiphilic surface helix (residues 36 to 50). Also, the 46-50 region shares homology with chymotrypsin where the equivalent sequence NWMVQQ is an helical region. The 51-69 was also another conserved region among the seven gag proteins. The first 4 residues TLLV were assigned to a β-strand while the QNANPD region to a coil consistent with the NMR structure of the synthetic (NANP)n polypeptide model of Plasmodium falciparum circumsporozoite protein (ref). The 61-66 sequence was an helix with both K residues on the same side. The alanine rich, GP containing 67-72 region was an interior coil followed by another conserved region, 73-81. The latter was predicted to be a surface amphiphilic helix. Interestingly a synthetic peptide corresponding to this region has been found to prevent the ability of the p55 gag polyprotein to form virus-like particles (ref). The 82-88 region contains several G and P residues consistent with its being a coil and turn region. Finally the 89-94 sequence was assigned to a β-strand.

Circular Dichroism studies

To experimentally determine the secondary structure composition of the three forms of 104-mer CD spectroscopy was used. The first folded state studied was that corresponding to the dimeric form (D).

Figure 8A shows the CD spectrum of an aqueous solution of D purified by the Rotofor technique. The spectrum, characterised by a negative band at about 200nm, was similar to those that have been observed for charged polypeptides presumed to be in the so-called random coil conformation (ref). However, in the spectrum of these polypeptides the negative shoulder at about 220nm in the spectrum of D is replaced by a positive CD contribution at about the same wavelength. Furthermore, a reinterpretation of the CD random coil conformation of poly(L-Lys) recently published (ref) assigns the CD spectrum of the polypeptide at pH 7.6 to that of a conformation aith an appreciable left-handed helix content. Whether the latter applied to the spectrum of D could not be concluded at this stage. However the chromatographic experiments described

above and the need to interpret the discrepancy between the calculated and found pI seemed to rule out a random coil conformation for D.

This conclusion was confirmed by MeOH into buffer titration CD experiments. Thus the plot of e^{-220} versus percentage of MeOH showed a plateau for MeOH concentrations larger than about 30% (Fig. 9); this together with an isosbestic point at about 208nm was consistent with an equilibrium between at least two states. The first state was represented by the spectrum at higher concentration of MeOH (Fig. 8C) and characterised by a predominately alpha-helical structure and a second state, that in aqueous solution. The ion-exchange experiments described above when combined with the observation that the pI of D was 4 units less than the calculated value from amino acid composition seemed to suggest that this second state was characterized as containing a significant amount of secondary structure and a compact form but lacking a well-defined tertiary structure. Whether the secondary structure elements were mainly the left-handed helix or a combination of the two could not be decided on the basis of these data.

The existence of defined elements of secondary structure in D was further confirmed when dialysis of 104-mer was carried out at pH 5.5, close to the pI of D. Thus, as expected some precipitate was formed during dialysis which was isolated, resuspended in buffer and quickly analyzed by CD spectroscopy. The spectrum thus obtained was very similar to that of D in buffer/MeOH mixtures (Fig. 8C) although a shift of the minimum to approx. 204nm indicated that the protein was rapidly generating the less compact conformation described above.

Having established that the spectrum in MeOH was representative of the folded state of D the secondary structure composition of the latter was calculated from the CD spectrum in 50% MeOH using the variable selection method described by Manavalan and Johnson. The results of this analysis indicated that the structure of D is of a a+b type with 43 residues (39%) being in helical conformation and about 10 residues (10%) in beta conformation. The remaining residues were turns (24) and coil (29) (Table 3). These results

correlated well with the secondary structure prediction described above (Table 3).

The CD structures of M and T

When an aliquot of T, directly obtained from ion-exchange purification, or a diluted solution of T which had been demonstrated to yield M were subjected to CD, the spectra in buffer pH 7 were identical to eachother and similar to that of the D form with a negative band at 200nm and a negative shoulder at about 220nm (Fig. 10A). However unlike D where the existence of well-defined secondary structure elements had been demonstrated there were no indications as to the presence of the latter in the structure of T. Alternatively, a specific aggregation of M to tetramers made it unlikely the possibility of a random coil for the conformation of T and M. Therefore, the assignment of CD the spectra of the latter to the above discussed left-handed helix seemed to be realistic.

Addition of MeOH to either M or T induced changes identical to those observed for D (Fig. 10B and 10C). These results together with those from the size exclusion dilution experiments described above indicated the existence of an equilibrium between T and M and between M and D. As to the interpretation of the CD spectra of T and M in buffer, their assignment to a conformation containing a substantial fraction of left-handed helix was not only consistent with the literature but also substantiated the conclusions inferred from these published observations, ie. that the left-handed helix might represent a common intermediate in the folding and unfolding of peptides and proteins. Thus the 104-mer in state M evolved either to T in concentrated solutions or to the folded state D in diluted solutions.

CONCLUSIONS

The chromatography experiments described above have highlighted the existence of three different forms of the synthetic 104-mer. These differed from one another not only for their size, monomer, dimer and tetramer, but also for their pI, which were found to be 8, 5.6 and 9.4 respectively. The existence of these three forms did not seem to be a direct consequence of the replacement of the two Cys residues at 330 and 350 with alanine. Thus a synthetic analogue of

the 104-mer which contained the two naturally occurring Cys gave during the IEF purification in solution two products having pI similar to those of M and D. Furthermore separate ion-exchange experiments carried out at higher peptide concentrations indicated the existence of a third form of synthetic 94-mer strongly retained by the cationic exchange medium (results not shown). This was in agreement with the aggregation to tetramers observed for the 104-mer.

CD spectroscopy studies indicated that the D form was, in aqueous solutions, a compact molecule lacking however a well defined tertiary structure. This conclusion was demonstrated by MeOH into buffer titration studies and consistent with the pI of D, about four units less than the expected value calculated from its amino acid composition. Thus one possible explanation for this discrepancy would be that some of the ionizing groups (the positively charged residues) expected to be exposed on the surface were buried inside the globular protein. In this assumption, the increased hydrophobicity of the solution, following the addition of MeOH, would result in the stabilisation of the globular state rather than favour its unfolding.

In 50% MeOH the CD structure of D was calculated to be that of a predominantly α-helix molecule in agreement with secondary structure predictions. By contrast the CD structures of M and T were interpreted as containing an appreciable amount of left-handed helix.

Dilution experiments as well as CD spectroscopy studies indicated that an equilibrium existed between M, D and T. This equilibrium could be perturbed by varying the peptide concentration or by the addition of MeOH according to the scheme in Fig. 11. Given the tendency shown by M to aggregate to tetramers and/or fold into D, M could represent an early folding intermediate of the 104-mer. Early folding intermediates generally termed Molten Globule are characterised by possessing a relatively hydrophobic character which may therefore favour aggregation. A corollary to these conclusions was that in the case of synthetic 104-mer, folding and/or aggregation would stem from a common conformational state in the left-handed helix structure.

Of the three forms of 104-mer detected in these experiments, the D form seemed the most likely one to resemble the folded state of the corresponding region in full-length p24. Several pieces of evidence were in agreement with this conclusion:

1. The pI of full-length p24 is 6.25.
2. Full-length p24 has been shown to exist in solution as a dimer. The D form of 104-mer is also a dimer.
3. The CD spectra of full-length p24 (Fig. 8D) and that of D are those of an a+b protein.

Additionally, although very little is known about the properties of the tertramer of p24 obtained at a high protein concentration, also synthetic 104-mer tends to form tetrameric structures at high concentration.

In light of the above observations it seems reasonable to conclude that the conformational properties of the synthetic protein are maintained in full-length p24. A corollary to this conclusion is that the aggregation motif which leads to the dimerization of p24 is contained in the C-terminus half of the protein. This conclusion is consistent with a recent observation that a synthetic peptide (PAATLEEMMTA) corresponding to the 339-349 region of p24 interferes with the self-association process of p55/p24 leading to the assembly of infectious viral particles (Ref).

A further corollary arising from having established a correlation between the structural properties of p24 and its synthetic C-terminal half is that studies aiming at interfering with viral assembly through inhibition of p24 self-association could benefit from the use of chemically synthesised proteins. Thus, the latter are obtained in relative large amounts (tens of mgs) using synthetic methodologies which are both fast, reliable and when compared to the cost of recombinant p24, inexpensive.

Furthermore, the solution structure of the C-terminus half could be more easily assigned by high-field NMR than that of the much larger p24 protein thus providing additional information on the structure activity relationship of the gag protein.

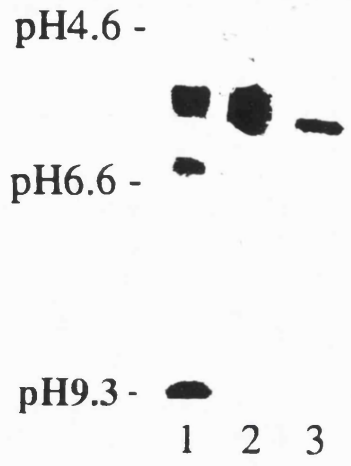


Figure 1. IEF gel of p24 gag protein
Lane 1 = Crude synthetic 104-mer
Lane 2 = Purified dimer(D) of synthetic 104-mer
Lane 3 = Recombinant full-length p24

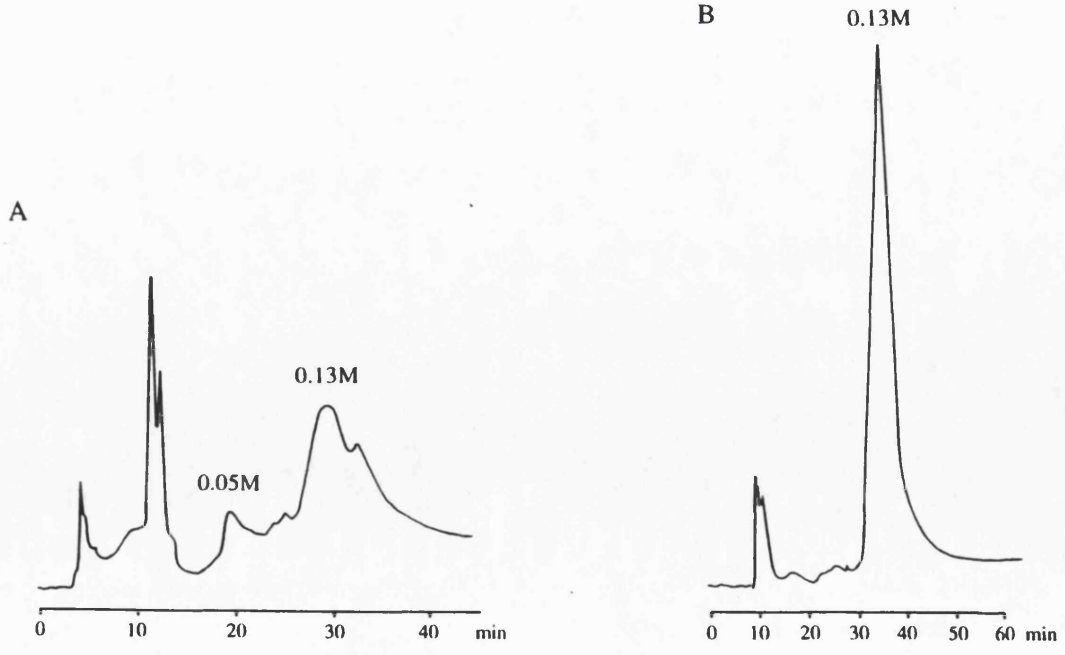


Figure 2. Cationic-exchange chromatography
A) Semi-preparative purification of crude 104-mer
B) Purified 104-mer

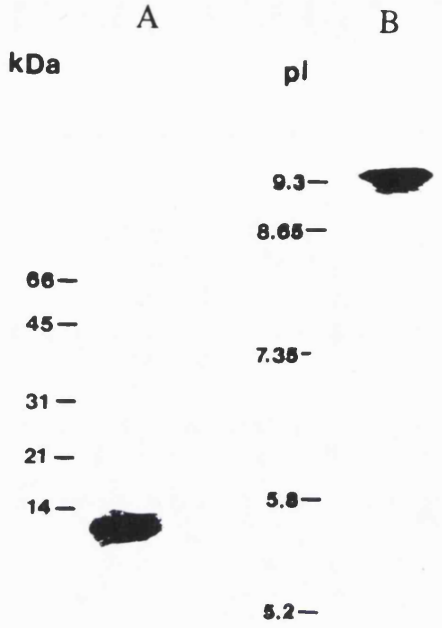


Figure 3. Gel electrophoresis of synthetic 104-mer after purification on cationic exchange media.
A) SDS-PAGE
B) IEF

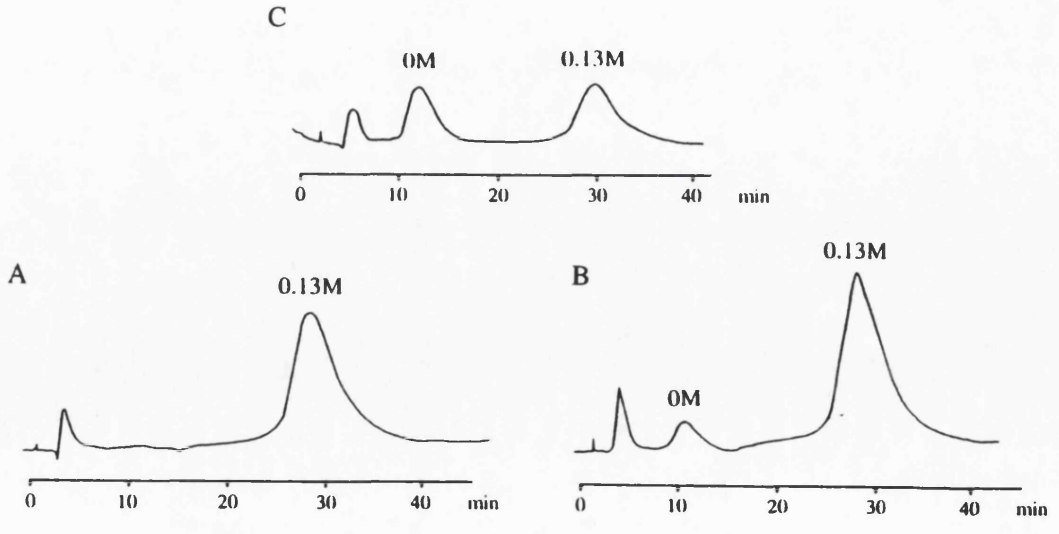


Figure 4. Cationic-exchange chromatography
A) Purified 104-mer after storage at -20°C for 3 days
B) A) after dilution

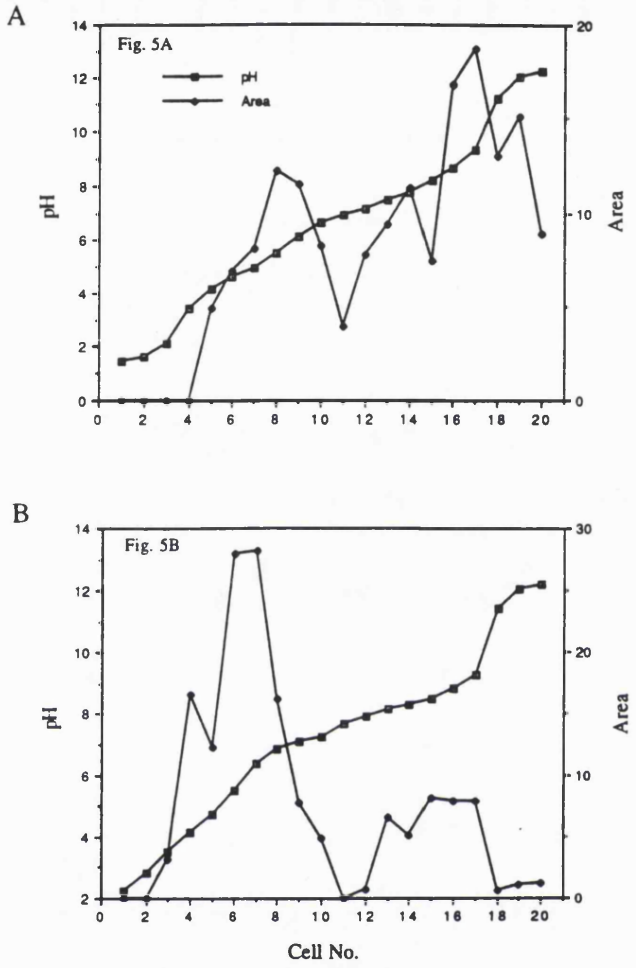


Figure 5. Purification of synthetic protein by IEF on Rotofor system
A) 104-mer
B) 94-mer

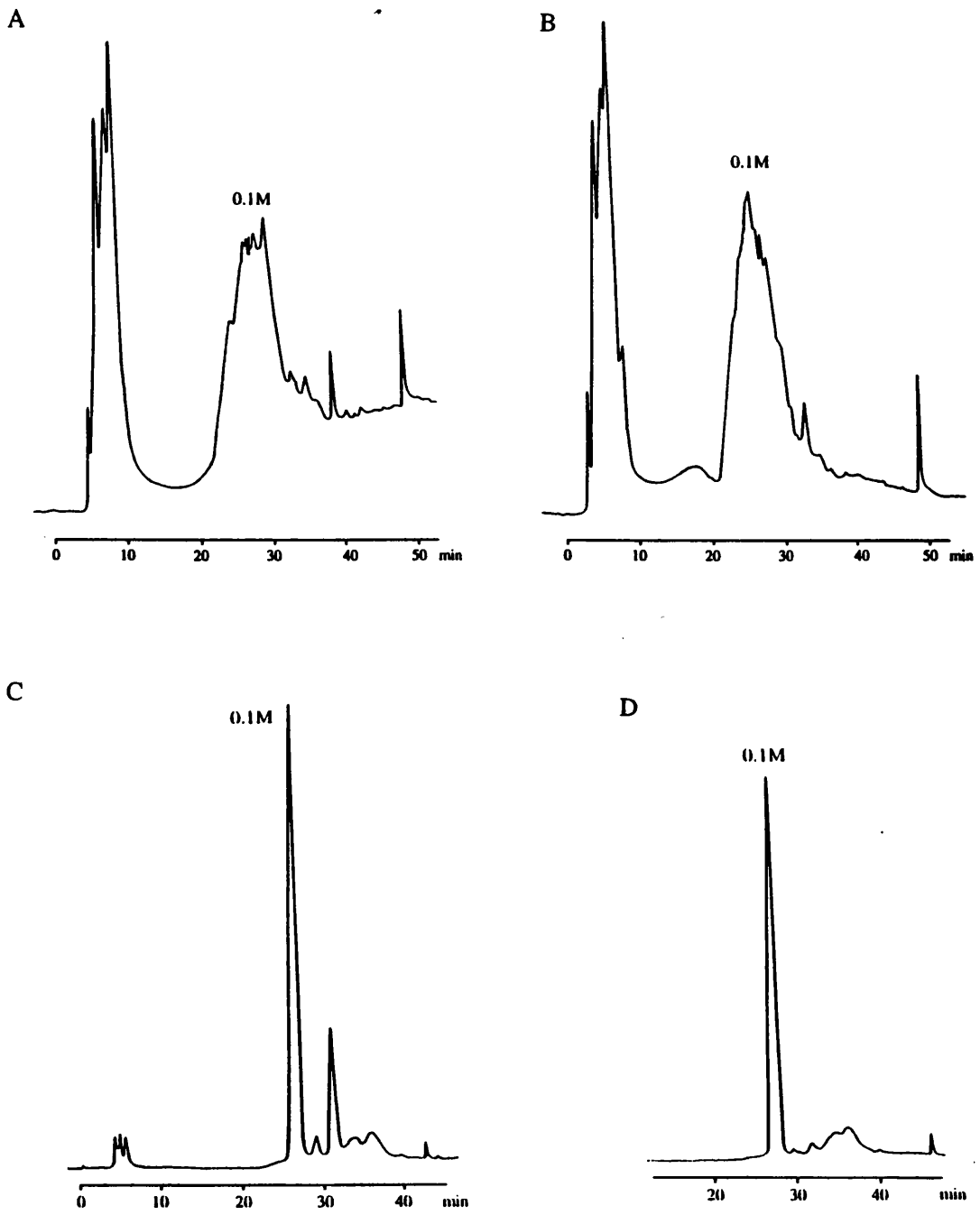


Figure 6. Anionic exchange chromatography of 104-mer

A) Cell 8 from Rotofor purification

B) As (A) at 4°C

C) As (A) plus 20% MeOH

D) Dimer after anion exchange purification in 20% MeOH

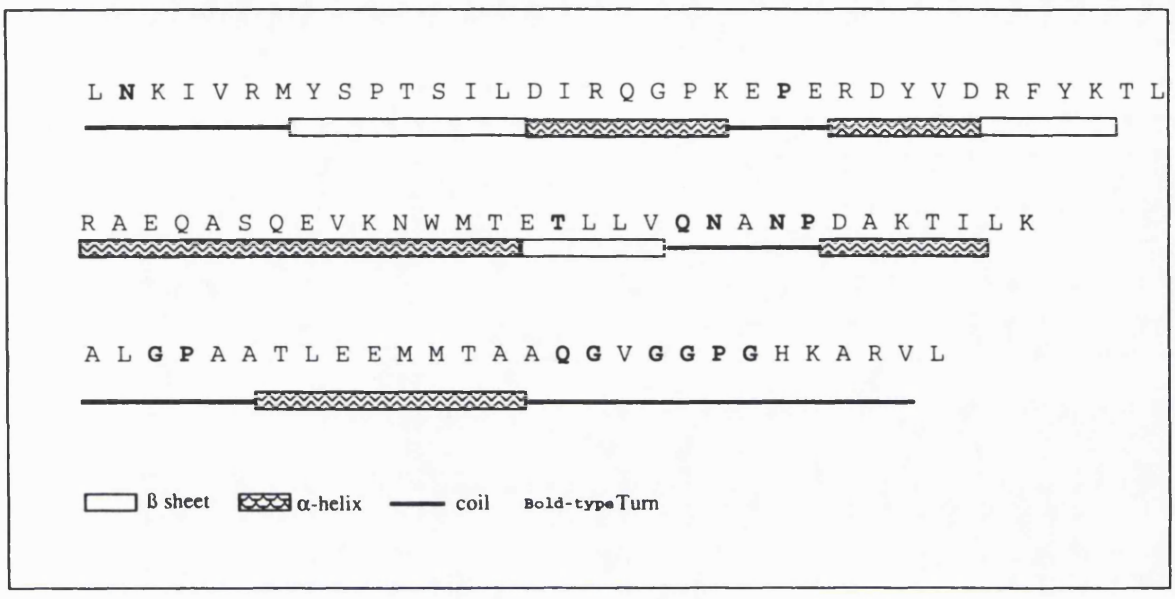


Figure 7. Secondary structure prediction for 104-mer

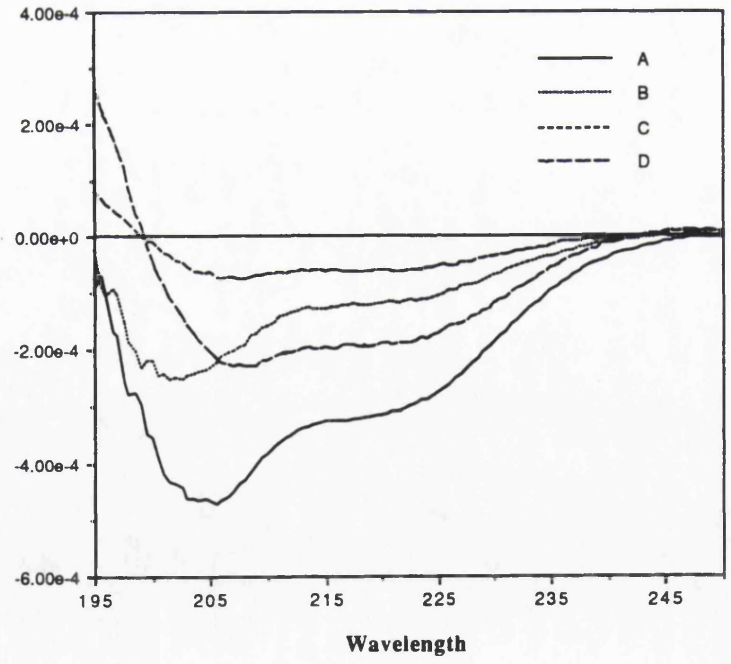


Figure 8. Circular dichroism analysis

A) Crude synthetic 104-mer in 20mM MOPS, pH 7
B) Crude synthetic 104-mer in water
C) Crude synthetic 104-mer in water + 50% MeOH
D) Recombinant full-length p24 in 20mM MOPS, pH 7

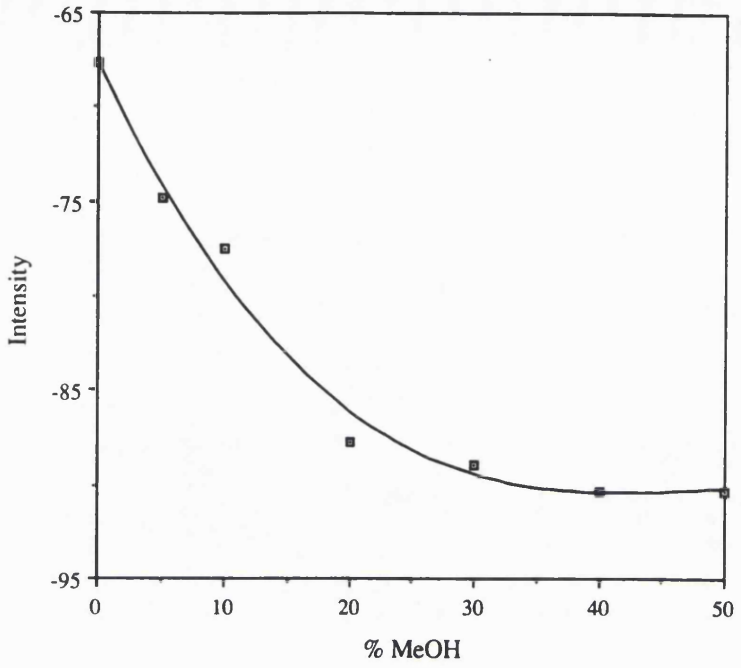


Figure 9. Results of MeOH into buffer titration experiments for synthetic 104-mer

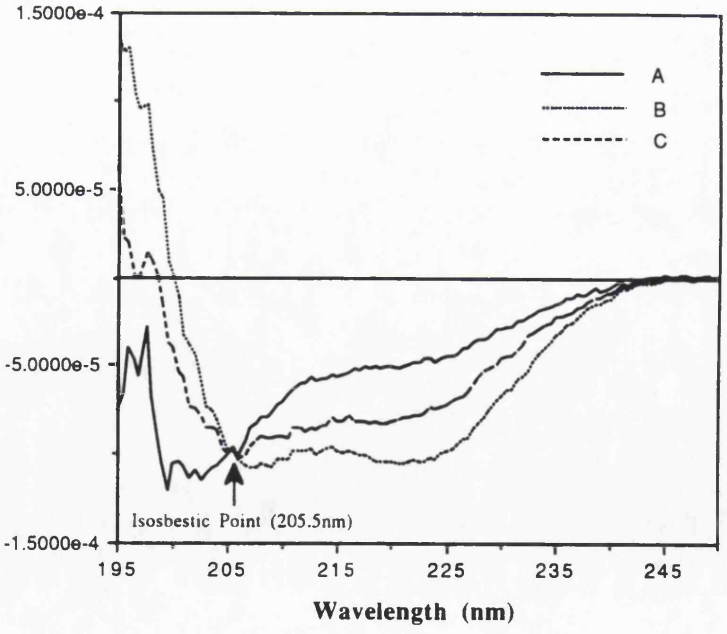


Figure 10. Circular dichroism studies of tetrameric (T) form of 104-mer
A) T in 20mM MOPS, pH 7
B) T in 20mM MOPS + 20% MeOH, pH 7
C) T in 20mM MOPS + 50% MeOH, pH 7

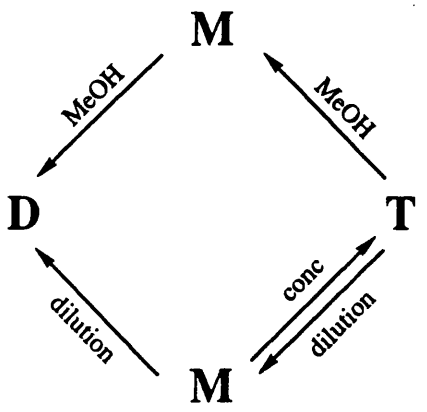


Figure 11. Scheme illustrating the proposed equilibrium state of 104-mer conformation forms

```
          1                  10          20          30
hiv       ln       kivr m    ysptsi ld irqgpkepfrdyvdrfyktlr
vlv       lqwvitalrsvrhms    hrpgnp ml vkqknteyedfiarlleaid
eiv       ms       egikvmi   gkp ka qn irqgakepypefydrllsqik
siv       l         pvk      gdpgas ltgvkqgpdepfadfvhrlittag
olv       l        rsvrhms    hrpgnp ml ikqknseyedfiarlleaid
biv       lq       aame vaqakhatpgp in ihqgpkepytdfinrlvaale
blv       ll        vlq l         qpwst ivqgpaessvefvnrlqisl

          40             50            60           70           80
hiv       ae qasqev kn wmtetll    vq nanpd aktil  kalgp   aatleemmta
vlv       ae pvtdpi kt ylkvtls    yt nastd cqkqmd rtlgtrvqqatveekmqa
eiv       se ghpqei sk fltdtlt    iq nanee crnam  rhlrp   edtleekmya
siv       rifgsaeag vd yvkq la    ye nanpa cqaai  rp  pr kktdltgyirl
olv       te pvtdpi kt ylkvtls    ft nastd cqkqmd rvlgtrvqqasveekmqa
biv       gm aapett ke yllqhls    id haned cqsil  rplgp   ntpmekkiea
blv       ad nlpdgvlrnplltp lvmqmlt esvskfcrgeasgrg gak  tagl

                 90
hiv       aqgvggpghkarvl
vlv       crdvgsegfkmqll
eiv       crdigttkqkmmll
siv       csdig psyqqgl
olv       crdvgsegfkmqll
biv       crvvgsqkskmqfl
blv        rtigpprmkqpall
```

Table 1. Sequence alignments for 104-mer

AA	C-F	GOR	Momany	Helical wheel	Final assignment
1-7	B	B	B		C
8-15	B	C	B		B
16-22	H	C	H	A	AH
26-30	H	H	H		H
31-35	B	C	B		B
36-50	H	H	A	A	AH
51-55	B	B	B		B
56-60	C	C	C		C
61-66	H	B	H	A	AH
67-72	C	H	C		C
73-81	H	H	H	A	AH
82-88	C	C	C		C
89-94	B	B	B		C

H: helix; B: ß-strand; C: coil; A: amphiphilic; AH: amphiphilic helix

Table 2. Secondary structure prediction of p24 94-mer by various methods

	Helix	↑↓β	↑↑β	turn	coil	total
in %	39	7	3	23	28	100
in AA	41	7	3	24	29	104
pred.AA	43	14		17	30	104

Table 3. Secondary structures from CD and prediction

PUBLICATIONS

Int. J. Peptide Protein Res. **40**, 1992, 370–379

Purification of synthetic peptides using reversible chromatographic probes based on the Fmoc molecule

H.L. BALL and P. MASCAGNI

Italfarmaco Research Centre, Milan, Italy

A rapid, versatile, reversible procedure for purifying synthetic peptides has been developed based on the specific incorporation of 4-carboxylate Fmoc derivatives onto the terminal amino acid of peptidyl-resins. The acid stable 4-COR-Fmoc derivatives were synthesised with a variety of chemical groups thus altering the chromatographic properties of the "target" peptides and permitting their convenient purification, either by reversed-phase HPLC or ion exchange chromatography. The assembly of the peptides involved a capping step to prevent the formation of deletion forms. The 4-COR-Fmoc derivatives were incorporated either as preformed amino acid conjugates or as activated succinimidyl esters. After HF cleavage and purification the 4-COR-Fmoc probes were quantitatively removed with organic bases. The efficiency of the technique was demonstrated by the purification of small to large sized peptides, including a cyclic analogue.

Key words: capping; fluorenylmethoxycarbonyl; peptide synthesis; purification

The chemical protocols for automated synthesis of peptides by either t-Boc or Fmoc strategies have been sufficiently developed to allow both routine synthesis of small to medium sized peptides and the synthesis of proteins up to about 100 residues. However, these protocols, although highly developed and sophisticated, are generally unable to prevent those side reactions which can occur during stepwise chain assembly and lead to the formation of deletion and truncated sequences. The separation of the latter from the "target" peptide using conventional chromatographic or electrophoretic techniques is particularly troublesome when the peptide chain length is greater than approximately 40 residues. Recent papers have described methods of purification based on the covalent or transitory coupling of molecules to synthetic peptides to alter their chromatographic nature. Merrifield & Bach (1) described the use of a reversible 2-Sulf-Fmoc group, with a sulfonyl side-chain at position 2, to derivatise non-terminated peptides and so facilitate purification by ion-exchange chromatography. The Fmoc sulfonyl probe was used for the purification of neutral peptides, or peptides containing only one charged residue. In both cases the peptides were no longer than about 20 residues and the application of this sulfonyl Fmoc derivative to the purification of larger peptides has not yet been described.

In an effort to purify the 153 residue IL-1 synthetic protein, Lobl *et al.* (2) covalently attached biotin to the peptide chain and separated the biotinylated chains on an avidin-agarose column. Although this purification was successful, biotin could not be removed from the purified protein due to its covalent linkage. The biotinylation of a 99-residue SIV protease synthetic protein was also performed by Bannow *et al.* (3). In this case reversal of the derivatisation was accomplished by the synthetic enzyme itself, which cleaved the biotinylated *N*-terminus Ala-Pro sequence. Recently Engstrom *et al.* (4) purified peptides between 25 and 40 residues in length on an immobilised metal-ion column. The method relies on the presence of His and/or Trp residues in the *N*-terminal region of the peptide and consequently does not offer general applicability. An overview of the use of various affinity labels to remove terminated peptides has been discussed by Barany & Merrifield (5).

The selective purification procedure that we have de-

Abbreviations follow the recommendations of the IUPAC-IUB Commission on Biochemical Nomenclature (*Biochem. J.* **126**, 773–780 (1972)). Additional abbreviations: DCC, dicyclohexylcarbodiimide; EDC, 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide; HOBt, hydroxybenzotriazole; DMF, dimethylformamide; Fmoc, fluorenylmethoxycarbonyl; TEA, triethylamine; TFA, trifluoroacetic acid; HF, hydrogen fluoride; t-Boc, *tert.*-butyloxycarbonyl; PAM, phenylacetamido methyl; MBHA, *p*-methylbenzhydrylamine; FAB-MS, fast atom bombardment-mass spectrometry.

veloped is based on the combination of a capping protocol (to prevent the formation of deletion peptides) during peptide chain assembly and on the addition, to the last residue of the sequence, of a reversible protecting group which bears either lipophilic, acidic or basic functions. These protecting groups are based on the 4-carboxylate fluorenylmethoxycarbonyl (4-COOH-Fmoc) molecule and have been designed with the following criteria:

1) The introduction of extremely hydrophobic groups or charged residues onto the 4-COOH-Fmoc group generates new molecules with enhanced chromatographic features (4-COR-Fmoc). If the properties of these molecules are transferred to a synthetic peptide through derivatisation, then the separation of the latter from its deletion and truncated forms by either reversed-phase HPLC or ion-exchange chromatography should be possible.

2) The urethane bond between the Fmoc group and peptide is stable to acids (6). Thus 4-COR-Fmoc protection should be maintained during the acid treatments aimed at deprotecting synthetic peptides and removing them from the solid support.

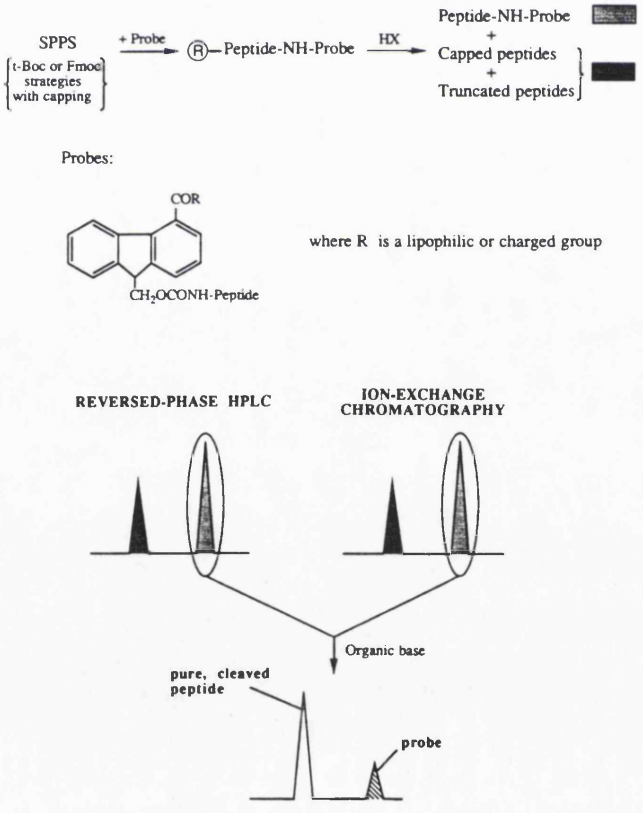
3) The instability of the 4-COR-Fmoc urethane linkage to mild organic bases should enable its quantitative cleavage from the purified peptide and subsequent retrieval of the latter in its free form.

4) The derivatisation of the peptide at the level of the last residue should permit the use of this purification-scheme for peptides synthesised with either the t-Boc or Fmoc chemical strategies.

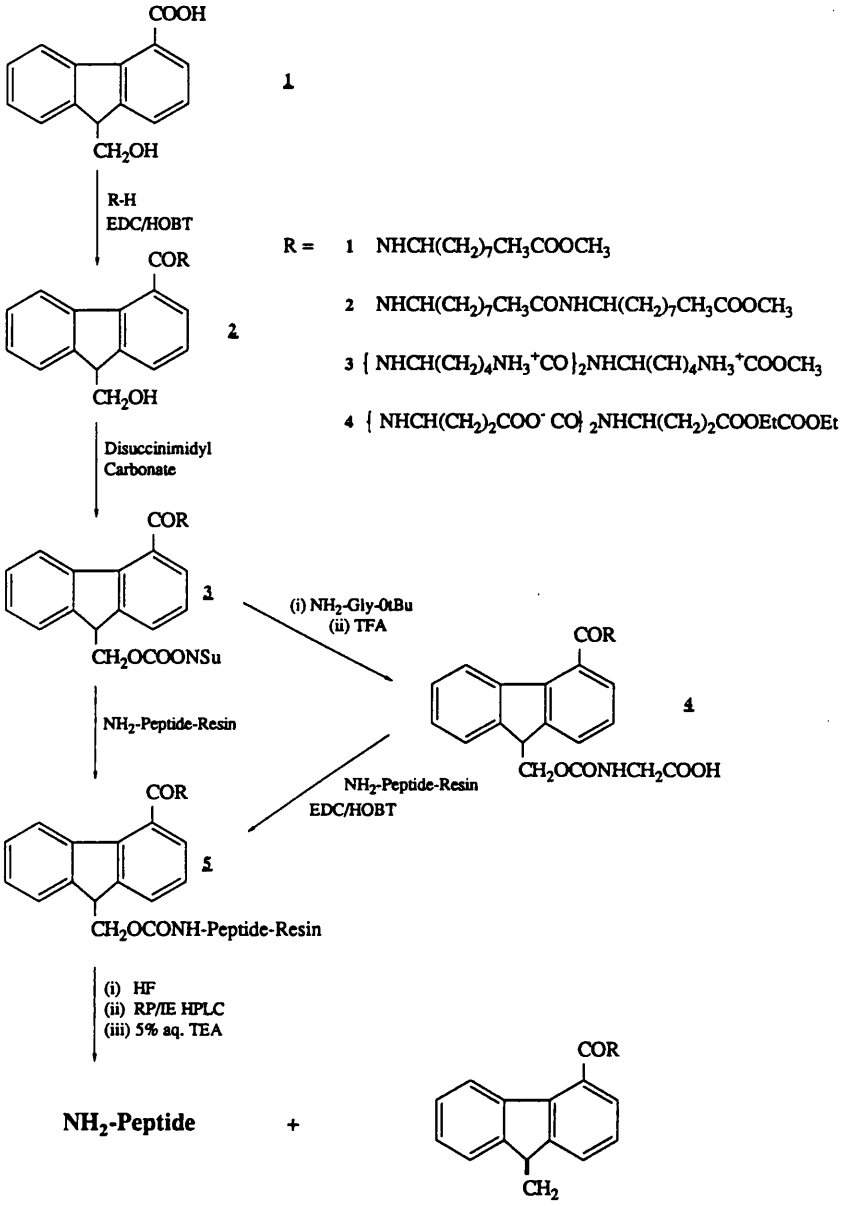
5) Finally the application of the 4-COR-Fmoc probes to a wide range of chromatographic media should permit the purification of synthetic peptides independently of their chain length or amino acid composition.

A pictorial representation of the proposed method is shown in Scheme 1.

The 4-COOH-Fmoc **1** was first described by Mutter and co-workers (7) as a handle for SPPS; in our case we used the carboxylic function in position 4 to introduce either extremely hydrophobic groups like (D,L) α-aminodecanoic acid or charged, e.g. Lys or Glu. The



SCHEME 1
Graphical representation illustrating the use of reversible 4-COR-Fmoc-probes for the purification of synthetic peptides/proteins.



SCHEME 2
Synthesis of 4-COR-Fmoc derivatised peptides.

new 4-COR-Fmoc derivatives thus obtained were coupled to the peptidyl-resin either as a preformed amino acid-probe molecules or directly to the last amino acid of the sequence via an active ester (e.g. NSu), as shown in Scheme 2. In the experiments described below we used the 4-COR-Fmoc-Gly probe as the representative of the former class of 4-COR-Fmoc derivatised amino acids. After acidolytic cleavage, the derivatised peptide was separated from unwanted sequences using either reversed-phase HPLC or ion-exchange chromatography and deprotected by treatment with an organic base.

To illustrate this purification method we have used

TABLE 1

Amino acid analysis of 4-COR-Fmoc derivatised and free peptides after purification

Residue	FMDV Vp1 142-157			HIV-1 gp120 fragment			HIV-1 gp120 fragment			Cyclic FMDV fragment		
	+ 4-COR1 -Fmoc		Free peptide	+ 4-COR2 -Fmoc		Free peptide	+ 4-COR3 -Fmoc		Free peptide	+ 4-COR4 -Fmoc		Free peptide
Asp/Asn	1.3	(1)	1.1	2.9	(3)	3.6	3.2	(3)	3.1	1.1	(1)	1.0
Thr	–		–	0.8	(1)	1.2	0.9	(1)	1.0	–		–
Ser	3.0	(3)	2.9	1.0	(1)	1.1	1.0	(1)	1.0	1.9	(2)	2.0
Glu/Gln	–		–	4.5	(5)	4.9	4.6	(5)	4.7	4.3[a]	(1)	1.2
Pro	–		–	3.3	(4)	3.9	3.7	(4)	3.8	2.2	(2)	1.9
Gly	4.3	(4)[a]	3.8	4.5	(5)[a]	4.6	4.7	(5)[a]	4.8	4.6	(4)	3.9
Ala	2.0	(2)	2.0	–		–	–		–	2.0	(2)	2.0
Val	1.9	(2)	1.7	–		–	–		–	1.7	(2)	2.4
Ile	–		–	0.8	(1)	1.0	0.9	(1)	0.9	–		–
Leu	2.1	(2)	2.1	1.0	(1)	1.0	1.0	(1)	1.0	2.1	(2)	2.0
Phe	–		–	–		–	–		–	1.0	(1)	0.9
Lys	–		–	–		–	2.4[a]		–	–		–
Arg	3.3	(3)	3.4	4.9	(5)	5.5	5.5	(5)	5.2	2.9	(3)	2.8

[a] Additional amino acids from 4-COR-Fmoc probes.

peptides of various lengths (17 to 104 residues) and of different chemical characteristics.

The first peptide investigated was the 142–157 sequence from the FMDV Vp1 capsid protein (8)*. The peptide was assembled on an MBHA resin using standard t-Boc chemistry. No capping procedure was performed at this stage. Quantitative ninhydrin analysis (9) indicated an average coupling efficiency of 98.6%. To introduce the 4-COR-Fmoc probe the t-Boc group was removed from the *N*-terminal residue of the peptidyl-resin with TFA and the resulting free amino group treated for 90 min with a DMF solution containing the 4-COR-Fmoc derivative **4** (R = **1**), hydroxybenzotriazole (HOBt) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) in equimolar quantities. In order to generate a mixture of derivatised and underivatised peptides, which would demonstrate their separation by reversed-phase HPLC, less than 1 equiv. of **4** (R = **1**) was used. The degree of incorporation of the 4-COR-Fmoc "handle" was monitored by measuring, at 270 nm, the release of the 4-COR-Fmoc group from base treated aliquots of peptidyl-resin. When two consecutive readings were identical the reaction was terminated. After washing with DMF, DCM and drying from 50% MeOH/DCM, the peptidyl-resin was cleaved using HF and scavengers. Under these conditions it was anticipated, that the acid stable 4-COR-Fmoc derivative would be retained on the peptide chain. Reversed-phase HPLC of the crude product confirmed this conclusion (Fig. 1a). The two major peaks labeled 1 and 2 in Fig. 1a were isolated and shown by FAB-MS* and amino acid analysis (Table 1) to correspond to free and 4-COR-Fmoc derivatised peptide, respectively. In the case of the late eluting material more than one peak can be seen in the HPLC chromatogram due to the presence of

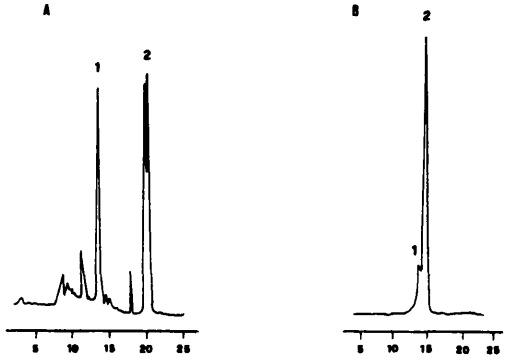


FIGURE 1

Reversed-phase HPLC of 4-COR1-Fmoc derivatised FMDV 142–157 linear peptide. A) Crude material after HF cleavage. B) Free peptide after treatment with 5% aqueous Et_3N; 1 represents a deletion peptide impurity (see text). Chromatographic analysis was performed on a Vydac C_{18} analytical column (0.46 × 15 cm) using a linear acetonitrile gradient over 30 min (0–100%, 0.1% TFA) at 218 nm.

* *Foot and Mouth Disease Virus (FMDV) Vp.1 142–157*
S-G-V-R-G-D-S-G-S-L-A-L-R-V-A-R-CONH$_2$
FAB-MS for 4-COR1-Fmoc-peptide: $[M+H]^+$: expected 2092; found 2093
FAB-MS for free peptide + Gly: $[M+H]^+$: expected 1657; found 1656

racemic α-aminodecanoic acid compound. The underivatised and derivatised peptides eluted at 40% and 59% acetonitrile respectively, clearly demonstrating that the addition of one α-amino-decanoic acid group to the 4-COOH-Fmoc molecule drastically changed the chromatographic characteristics of the peptide bearing this probe.

Once the feasibility of this purification approach had been established, the next step involved the determination of optimum conditions for the removal of the 4-COR-Fmoc group from the purified derivatised peptide and the retrieval of pure underivatised material. Studies into the sensitivity of the urethane group to various organic bases in different solvents have been performed (1). The use of dilute (5%) aqueous Et_3N gave satisfactory results with quantitative cleavage being achieved in 30 min and no side reactions observed. HPLC analysis of the free peptide obtained with this cleavage protocol gave the expected peptide in 95% yield (Fig. 1b). However, an impurity of about 5% corresponding to a deletion peptide ($[M+H]^+$ 1570; expected sequence less Ser^{142}) was seen in the HPLC chromatogram (Fig. 1b). This was not unexpected, given the ninhydrin values obtained during chain assembly and confirmed the need for a capping step during peptide synthesis, for the proposed purification method to be completely effective.

Accordingly, during the synthesis of the 25 residue HIV-1 gp120 peptide (10)**, a capping procedure (5% acetic anhydride in DMF) was performed after each coupling step. Applying the methods described above, probe **4** (R = **2**; twofold excess) was incorporated quantitatively after a reaction time of 90 min. Since this 4-COR-Fmoc derivative possesses two α-amino-decanoic acid groups it was expected that the difference in retention times, between derivatised and underivatised peptide, would be greater than that obtained in the previous example when the mono "fatty" compound **4** (R = **1**) had been used. Following HF cleavage, analysis of crude lyophilised product by reversed-phase HPLC showed an intense peak at 46 min (Fig. 2a). Isolation and characterisation of this material indicated that it was the 4-COR-Fmoc **4** (R = **2**) derivatised peptide. Crude underivatised peptide used as a control eluted at 20 min, thus indicating an increase of 26 min in the retention time of the 4-COR-Fmoc modified material.

As described in the previous case, treatment with 5% aqueous Et_3N for 30 min resulted in the quantitative removal of the 4-COR-Fmoc probe and the formation of a single peak eluting at 20 min (Fig. 2b). FAB-MS** and amino acid analysis (Table 1) confirmed this to be the free peptide with the expected sequence.

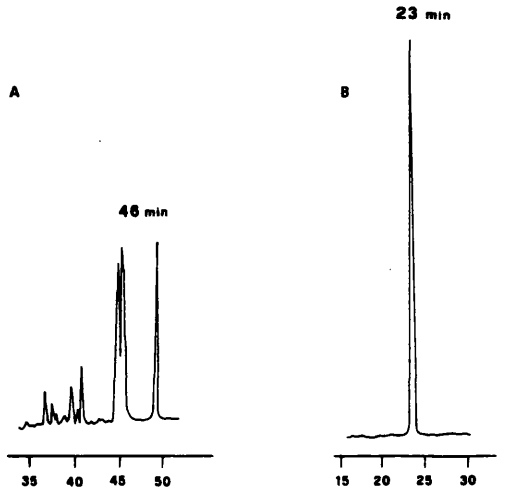


FIGURE 2

Reversed-phase HPLC of 4-COR2-Fmoc derivatised HIV-1 ENV gp120. A) Crude material after HF cleavage. B) Free peptide after treatment with 5% aqueous Et_3N. Chromatographic analysis was performed on a C_8 Brown-Lee analytical column (0.21 x 3 cm) using a linear acetonitrile gradient over 80 min (0–100%, 0.1% TFA) at 218 nm.

In conclusion these preliminary experiments had indicated that by using the lipophilic 4-COR-Fmoc probes, small to medium sized peptides could be rendered sufficiently hydrophobic to enable their separation from closely related impurities by HPLC. Furthermore the results obtained with **4** (R = **2**) suggested that it could be applied to the purification of peptides larger than those described above.

To test this hypothesis, a peptidyl-resin comprising 104 residues (11)† previously synthesised in our laboratory, was derivatised with **4** (R = **2**). The polypeptide had been assembled using highly optimised t-Boc chemistry (13) which, however, did not involve capping unreacted amino groups after each coupling step. Although a capping procedure is necessary for a successful application of this purification method, it was argued that the crude polypeptide contained a sufficiently high number of truncated impurities to test the validity of our approach.

** *HIV-1 gp120*
L-P-T-P-R-G-P-D-R-P-E-G-I-E-E-E-G-G-E-R-D-R-D-R-S-CONH$_2$
FAB-MS for 4-COR2-Fmoc-peptide: $[M+H]^+$: expected 3476; found 3476
FAB-MS for 4-COR3-Fmoc-peptide: $[M+H]^+$: expected 3543; found 3545
FAB-MS for free peptide + Gly: $[M+H]^+$: expected 2877; found 2878

† *HIV-1 p24 GAG Fragment 270–373*
L-N-K-I-V-R-M-Y-S-P-T-S-I-L-D-I-R-Q-G-P-K-E-P-F-R-D-Y-V-D-R-F-Y-K-T-L-R-A-E-Q-A-S-Q-E-V-K-N-W-M-T-E-T-L-L-V-Q-N-A-N-P-D-A-K-T-I-L-K-A-L-G-P-A-A-T-L-E-E-M-M-T-A-A-Q-G-V-G-G-P-G-H-K-A-R-V-L-A-E-A-M-S-Q-V-T-N-S-COOH

To attach the 4-COR-Fmoc probe **4** (R = **2**), an aliquot of the 104mer peptide-resin was allowed to swell in DMF for several hours and then treated with TFA to remove the t-Boc protecting group. The deprotected resin-bound 104mer was treated with a DMF/DCM solution containing an estimated twofold excess of **4** (R = **2**), HOBt and EDC (1:1:1). After 90 min reaction time, no further addition of 4-COR-Fmoc derivative was observed as determined by qualitative UV analysis. The peptide-resin was washed, dried and cleaved using a low TFMSA/high HF procedure. Reversed-phase HPLC of the crude polypeptide on C_4 media showed the presence of two large peaks centered at 46 and 51 min (data not shown). The difference in retention time of 5 min was increased to 7 min when using C_8 reversed-phase media (Fig. 3a). On C_{18} media the time separating derivatised and underivatised polypeptide was about 8 min, with almost baseline resolution of the eluting peaks (Fig. 3b). A control run using untreated 104mer indicated that this had an identical retention time as the peak eluting at 49 min in the chromatogram of Fig. 3b. Thus it was concluded that the peak at 57 min corresponded to the 4-COR-Fmoc derivative. This conclusion was supported by the evidence accumulated after isolation and characterisation of the material eluting with a retention time of 57 min. Thus when the latter was treated with 5% aqueous Et_3N for 15 min two peaks on a C_{18} column, at 49 and 57 min were obtained; after a further 15 min of reaction the 57-min peak converted completely into the 49-min peak (Fig. 3d), consistent with the removal of the 4-COR-Fmoc probe. Furthermore SDS-PAGE electrophoresis of the same material gave a band with the expected average molecular weight (approximately 11 kd, data not shown) and amino acid analysis was consistent with the expected amino acid composition (Table 2). Significantly, amino acid analysis of the peak which eluted at 49 min gave a more inferior found to expected amino acid ratio (Table 2). Furthermore, a consistency between expected and found amino acid composition had been obtained earlier (12), for underivatised 104mer, but only after a four-step purification protocol which involved dialysis, reversed-phase HPLC, size-exclusion and ion-exchange chromatography (Table 2).

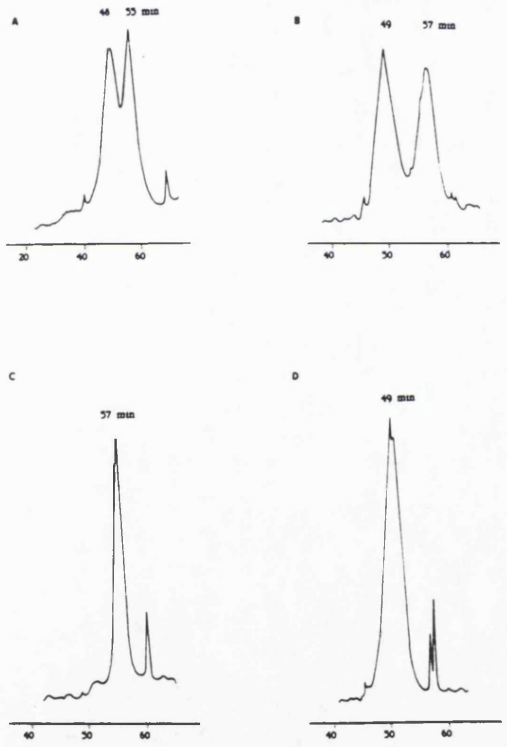


FIGURE 3

Reversed-phase HPLC of 4-COR2-Fmoc derivatised HIV-1 p24 GAG 270–373. A) Crude material after HF cleavage on a C_8 Brown-Lee analytical column (0.21 × 3 cm). B) As (A) but applied to a C_{18} Vydac analytical column (0.46 × 15 cm). Eluting conditions as described in Fig. 2. C) Isolated 4-COR2-Fmoc derivatised polypeptide. D) Free polypeptide after treatment with 5% aqueous Et_3N. Chromatographic analyses for (C) and (D) were performed on a Vydac C_{18} analytical column (0.46 × 15 cm) using conditions described in Fig. 2.

The experiments described above indicated that by changing the hydrophobic properties of a synthetic peptide with one of the 4-COR-Fmoc probes (R = **1** or **2**) separation of closely related impurities by means of reversed-phase HPLC was possible. Furthermore, by using the more lipophilic probe (R = **2**) we have shown that this separation was applicable to a large polypeptide and that the latter was obtained in a highly homogeneous form after just one chromatographic step.

To extend the application of this method to other chromatographic media we synthesised 4-COR-Fmoc probes containing charged amino acids and used them for the purification of synthetic peptides by ion-exchange chromatography.

The first example of this approach was provided by the 25 residue HIV-1 peptide described above. This was reacted with the 4-COR-Fmoc derivative **4** (R = **3**) which reduced the net charge from two negative in the free peptide to zero in its derivatised form. Separation of the latter was therefore possible by anionic exchange chromatography. Thus, when the crude mixture was applied to a MONO-Q resin (Fig. 4a), pure 4-COR-Fmoc derivatised peptide was found in the fraction eluting with 0.05 M NaCl, while the peptidic impurities were retained and finally eluted with 0.15 M NaCl. As found in the previous examples, treatment of isolated 4-COR-Fmoc derivatised peptide (Fig. 4b) with 5% aqueous Et_3N gave a pure product as shown by reversed-phase HPLC (Fig. 4c), FAB-MS** and amino acid analysis (Table 1).

TABLE 2

Comparison of the amino acid analysis of 104 residue HIV-1 p24 GAG 270-373 peptide purified with Fmoc probe 4 (R = 2) against the product purified using conventional techniques

Residue	Expected	Crude underivatised 104mer	C-18 RP HPLC Purified 104mer		104mer purified by standard four-step scheme
			No Fmoc probe RT = 48 min	Fmoc probe attached RT = 55 min	
Asp/Asn	9	7.1	6.3	8.5	9.5
Thr	8	3.8[a]	6.6	7.6	8.6
Ser	5	0.7[a]	3.3[a]	4.7	4.9
Glu/Gln	13	11.4	11.1	13.0	13.5
Pro	6	5.4	4.6	6.1	6.5
Gly	6	6.0	6.0	7.1[b]	6.2
Ala	12	12.0	12.0	12.3	12.7
Val	7	6.7	5.0	6.8	6.9
Met	5	4.3	4.0	5.3	5.6
Ile	4	2.5	1.6	3.3	2.8
Leu	9	7.6	7.1	9.0	8.9
Tyr	3	2.0	1.4	2.7	2.5
Phe	2	1.4	1.1	1.8	1.8
Lys	7	5.3	4.9	7.0	7.1
His	1	–[c]	–[c]	–[c]	0.9
Arg	6	4.0	3.7	6.6	6.1

[a] The hydrolyses was performed in conditions that did not prevent the loss of Ser and Thr residues.
[b] Extra Gly residue was introduced with the 4-COR2-Fmoc probe.
[c] The polypeptide contained a $N_{(Im)}$-DNP-His residue which was not deprotected during hydrolysis.

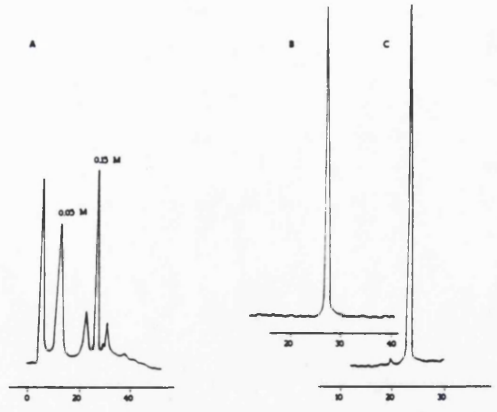


FIGURE 4

Chromatographic analysis of 4-COR3-Fmoc derivatised HIV-1 ENV gp120. A) Anionic exchange FPLC of crude material after HF cleavage. Analyses were performed on Pharmacia Mono Q medium using 20 mmol phosphate buffer at pH 6.5 and a linear elution (0.5 M NaCl) gradient over 30 min (0–40%) at 263 nm. B) Reversed-phase HPLC of purified 4-COR3-Fmoc derivatised peptide. C) Reversed-phase HPLC of free peptide after treatment with 5% aqueous Et_3N. Chromatographic analysis was performed on a Vydac C_4 analytical column (0.46 × 15 cm) using conditions described in Fig. 2.

A second example of ion-exchange purification was provided by a 25 residue cyclic analogue of the FMDV Vp1 peptide (9)††. The net charge of +2 changed to −1 by adding probe **3** (R = **4**) to the free amino terminus of the linear peptide. Unlike the previous cases the 4-COR-Fmoc molecule was introduced as the succinimidyl ester; this active ester approach was used to test the feasibility of direct derivatisation of the peptidyl-resins without the need of preparing 4-COR-Fmoc-amino acid constructs. Quantitative coupling was obtained in 90 min thus justifying this approach.

After HF cleavage the linear peptide was cyclised at pH 6 in dilute solution, then subjected to ion-exchange chromatography. Analysis of the material thus obtained indicated that the main fraction eluting with 0.1 M NaCl (Fig. 5a) was the 4-COR-Fmoc-cyclic peptide. The flanking peaks labelled 1 and 2 in Fig. 5a corresponded

†† *FMDV (cyclic) peptide*

C-ACA-G-S-G-V-R-G-D-F-G-S-L-A-P-R-V-A-R-Q-L-P-C(Acm)-
S-S
ACA-C-COOH

ACA = δ-amino caprioic acid

FAB-MS for 4-COR4-Fmoc-peptide: $[M+H]^+$: expected 3354; found 3356

FAB-MS for free peptide: $[M+H]^+$: expected 2645; found 2645

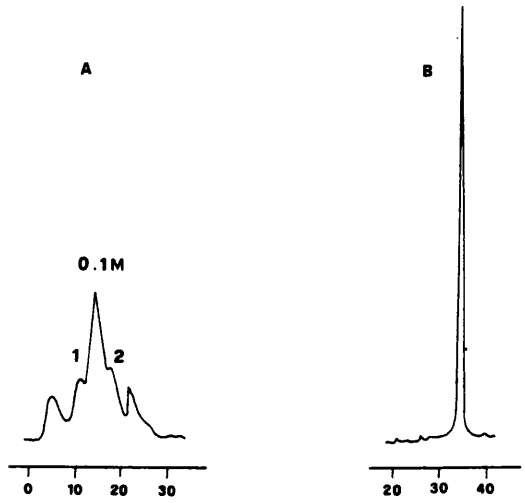


FIGURE 5

Chromatographic analysis of 4-COR4-Fmoc derivatised cyclic FMDV Vp1. A) Anionic exchange FPLC of crude material after HF cleavage. Analyses were performed on Pharmacia Mono Q medium using 20 mmol phosphate buffer at pH 6.5 and a linear elution (0.5 M NaCl) gradient over 35 min (0–40%) at 263 nm. B) Reversed-phase HPLC of free cyclic peptide after treatment with 5% Et_3N. Chromatographic analysis was performed on a Vydac C_4 analytical column (0.46 x 15 cm) using conditions described in Fig. 2.

to underivatised peptide and either the 4-COR-Fmoc derivatised dimeric form or derivatised uncyclised peptide, respectively, as indicated by FAB-MS (data not shown). Sample desalting by reversed-phase HPLC column, followed by treatment with base, gave free cyclised peptide (Fig. 5b) with the expected molecular weight[++] and amino acid analysis (Table 1).

CONCLUSIONS

New 4-COR-Fmoc derivatives containing either extremely hydrophobic groups or charged residues have been synthesised. These new molecules possess enhanced chromatographic characteristics which were transferred to a synthetic peptide following derivatisation of the peptidyl-resin. Separation of the derivatised peptide from its truncated forms was then possible by either reversed-phase HPLC or ion-exchange chromatography. When this purification scheme was applied in conjunction with a capping procedure during peptide chain assembly, highly homogeneous material was obtained after one purification step.

Successful application of this scheme has been shown with small to medium sized peptides of various amino acid composition. However, since peptides of up to about 40 residues can usually be purified using conventional chromatographic techniques, we have used one of the new 4-COR-Fmoc probes for the purification of a chemically synthesised protein of 104 residues. Although the latter had been assembled without the use of a capping protocol, it contained an elevated proportion of truncated forms so that its purification using the hydrophobic 4-COR-Fmoc derivative (**4** (R = **2**)) afforded highly homogeneous material after just one reversed-phase HPLC procedure. This was indicated not only by amino acid analysis but also by ion-exchange FPLC and the readiness with which this material adopted a folded structure (13).

EXPERIMENTAL PROCEDURES

Synthesis of peptides

The HIV-1 gp120 and FMDV Vp1 fragments were synthesised on an ABI 430A automated peptide synthesiser, using standard t-Boc protocols. In the case of the former, a capping protocol was performed after the addition of each amino acid using 5% acetic anhydride in DMF for 3 min.

The di- and tri- peptides (R = **2**, **3** and **4**) used to derivatise the 4-COOH-Fmoc **1** group were synthesised using conventional t-Boc solution phase methods. The *C*-terminal protecting groups were methyl (R = **1**, **2** and **3**) and ethyl (R = **4**) esters and consequently would not be removed by anhydrous HF. The side chain protecting groups were 2-chlorobenzoxy (R = **3**) and benzyl/cyclohexyl (R = **4**). One residue used in the synthesis of triglutamic acid contained an ethyl ester side chain protecting group.

Synthesis of 9-(hydroxymethyl)fluorene-4-carboxylate (1)

1 was prepared according to Mutter & Bellof (6). The overall yield for the four step synthesis was 44.8% and the final product was purified by flash chromatography using DCM/MeOH (7:3): Rf (DCM/MeOH 7:3) 0.4. ^{1}H-NMR (DMSO-d_6): 3.8–4.1 (*m*, 3H, H—C(9), CH_2OH); 7.2–7.8 (m, 6H, aryl H); 8.3 (d, 1H, H—C(3)). MS: 240 (M^+); 223 (M^+-OH); 210 (M^+-CH_2O); 165 ($M^+-CH_2O-CO_2$).

Synthesis of derivatised 9-(hydroxymethyl)fluorene-4-carboxylate compound 2 (R = 2)

1.5 g (6.25 mmol) **1** was dissolved in 6 mL DCM/DMF (1:1) and cooled to 0°. To the solution was added 0.96 g HOBt hydrate (6.25 mmol) and 1.31 g (6.88 mmol) 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide. After 30 min stirring to allow for activation, 3 g (6.25 mmol) of the TFA salt of di-α-(D,L)aminodecanoic acid methyl ester was dissolved in 3 mL DCM and added with 1.1 mL (6.25 mmol) DIEA to the reaction mixture. The reaction was left stirring at 0° for 60 min then at room temperature for 3 h. At the end of the reaction the urea derivative and any other impurities were extracted by successive washings with water, 0.1 M HCl, 3% $NaHCO_3$, water, dried over anhydrous Na_2SO_4 and evaporated *in vacuo*. The crude product

was purified by flash chromatography to give 2.52 g of product (68.2%, MS: 592 (M^+)).

The other 4-COR-Fmoc probes, **2** (R = **1**, **3** and **4**) were synthesised using the same protocol to yield 61.5% (MS: 423 (M^+)), 70% (MS: 1110 (M^+)) and 81% (MS: 856 (M^+)), respectively.

Synthesis of 9-(succinimidylhydroxymethyl)fluorene-4-carboxylate compound 3 (R = 2)

1.07 g (1.82 mmol) **2** (R = **2**) was dissolved in 10 mL acetonitrile. To the solution was added 138 μL (1.82 mmol) dry pyridine and 0.93 g (3.64 mmol) disuccinimidyl carbonate. The reaction mixture was stirred at room temperature for 8.5 h. The acetonitrile was removed on a rotary evaporator and the residue taken up in 20 mL DCM. The organic phase was shaken with water, 0.1 M HCl, water and then dried over anhydrous Na_2SO_4. 1.1 g pure product (82%, MS: 733 (M^+)) was obtained after flash chromatography on silica gel.

The active succinimidyl esters **3** (R = **1**, **3** and **4**) of the other compounds were made in a similar fashion to yield 85% (MS: 564 (M^+)), 82% (MS: 1250 (M^+)) and 96% (MS: 995 (M^+)), respectively.

Synthesis of 9-(glycylmethoxycarbonyl)fluorene-4-carboxylate 4 (R = 2)

1.0 g (1.36 mmol) **3** (R = **2**) was dissolved in 3 mL DCM, to which was added 0.21 g (1.36 mmol) HOBt hydrate. 0.29 g (1.49 mmol) Gly-OtBu in 5 mL DCM was added dropwise to the solution containing **3** (R = **2**) together with 260 μL (1.49 mmol) DIEA. The reaction mixture was stirred at room temperature for 1 h. After the reaction had gone to completion 20 mL DCM was added and the organic layer shaken with 20 mL aliquots of water, 0.1 M HCl, water and finally dried over anhydrous Na_2SO_4. The solvent was removed on a rotary evaporator and the product purified by flash chromatography to yield 0.72 g (71%; MS: 748 (M^+)).

4 (R = **2**) was obtained by treating purified t-butyl ester with 50% TFA in DCM for 30 min. The solvent was removed under vacuum and the free carboxyl compound purified by elution through a flash chromatography column to give an overall yield for the two step procedure of 0.66 g (64%; MS: 692 (M^+)).

The succinimidyl esters **3** (R = **1** and **3**) were also converted to **4** (R = **1** and **3**) in a similar way to give the following yields: **4** (R = **1**) 59% (MS: 511 (M^+)) and **4** (R = **3**) 65% (MS: 1211 (M^+)).

Attachment of the 4-COR-Fmoc probes to the peptidyl-resin

Coupling of 4 (R = 2) to peptidyl-resin. The peptidyl-resin was treated with 50% TFA in DCM for 30 min to remove the *N*-terminal protecting group and then neutralised with 5% DIEA in DCM. A twofold excess of **4** (R = **2**) was converted to the corresponding HOBt ester following the addition of 1 equiv. each of HOBt and DCC in 3 mL DCM. After allowing 30 min for activation the precipitated DCU was filtered off and the supernatant transferred to the reaction vessel containing the peptidyl-resin. The reaction mixture was vortexed and aliquots of resin were removed at half hourly intervals to allow assessment of coupling efficiency. The degree of incorporation of **4** (R = **2**) was determined by treating the peptidyl-resin with 20% piperidine in DMF for 15 min and measuring the absorbance of the released Fmoc group at 270 nm. When two consecutive results were identical the reaction was stopped by filtering the derivatised resin, washing with DMF and drying from 50% MeOH in DCM. The coupling time was found to range between 60 and 90 min.

4 (R = **1** and **3**) were coupled in the same way as HOBt activated esters.

Coupling of 3 (R = 4) to peptidyl-resin. A twofold excess **3** (R = **4**) was dissolved in 3 mL DCM and added to the *N*-terminally deprotected and neutralised peptidyl-resin, previously swollen in 1 mL DMF. The progress of the coupling was monitored according to the procedure detailed above. At the end of the reaction the derivatised peptidyl-resin was washed with DMF and dried from 50% MeOH in DCM in preparation for treatment with anhydrous HF. The coupling time was also found to be between 60 and 90 min.

Cleavage of 4-COR-Fmoc-peptidyl-resin

The 4-COR-Fmoc derivatised and parent FMDV and HIV-1 gp120 peptides were cleaved using a High HF only protocol. The cleavage mixture used for the High HF reaction was 800 μL *p*-cresol, 200 μL *p*-thiocresol and 10 mL HF. The reaction mixture was left stirring at 0° for 1 h. The fully deprotected peptide was precipitated with diethyl ether and finally dissolved in 10% acetic acid before lyophilisation.

In the case of the 104mer a Low TFMSA/High HF procedure was adopted. The Low TFMSA reaction was performed at 0° for 2 h, in a glass round-bottomed flask containing 20 μL 1,4-butanedithiol (fresh), 80 μL *p*-cresol and 300 μL dimethylsulphide (fresh), 500 μL TFA and 100 μL TFMSA. The partially cleaved peptidyl-resin was washed and dried with diethyl ether before performing the High HF reaction, as described above.

ACKNOWLEDGEMENTS

H.L. Ball was supported by an SERC grant (no. 87301535). Special thanks are due to Applied Biosystems for supporting this research and performing the amino acid analysis.

REFERENCES

1. Merrifield, R.B. & Bach, A.E. (1978) *J. Org. Chem.* **43**, 4808
2. Lobl, T.J., Deibel, M.R. & Yen, A.W. (1988) *Anal. Biochem.* **170**, 502

3. Bannow, C.A., Tomasselli, A.G., Zurcher-Neely, H.A., Smith, C.W. & Heinrikson, R.L. (1991) 12th American Peptide Symposium, Boston, Abstracts, p. 460
4. Engstrom, A., Bennich, H. & Lindemberg, G. (1989) 11th American Peptide Symposium, San Diego, Abstracts, p. 110
5. Barany, G. & Merrifield, R.B. (1980) in *Peptides. Analysis, Synthesis, Biology* (Gross, E. & Meienhofer, E., eds.), Vol. 2, pp. 163–165, Academic Press, London
6. Carpino, L.A. & Han, G.Y. (1972) *J. Org. Chem.* **37**, 3404–3409
7. Mutter, M. & Bellof, D. (1984) *Helv. Chim. Acta* **67**, 2009–2016
8. Bittle, J.L., Houghten, R.A., Alexander, H., Shinnick, T.M., Sutcliffe, J.G., Lerner, R.A., Rowlands, D.J. & Brown, F. (1982) *Nature* **298**, 30
9. Sarin, V.K., Kent, S.B.H., Tam, J.P. & Merrifield, R.B. (1981) *Anal. Biochem.* **117**, 147–157
10. Robson, B., Fishleigh, R.V. & Morrison, C.A. (1986) *Nature* **325**, 395
11. Mascagni, P., Sia, D.Y., Coates, A.R.M. & Gibbons, W.A. (1990) *Tetrahedron Lett.* **31**, 4637
12. Kent, S.B.H. & Clark-Lewis, I. (1985) in *Synthetic Peptides in Biology and Medicine* (Alitalo, K., Partanen, D. & Vaheri, A., eds.), Elsevier Science Publishers, Amsterdam
13. Ball, H.L., Coates, A.R.M., Gibbons, W.A. & Mascagni, P., Manuscript in preparation

Address:

P. Mascagni
Italfarmaco Research Centre
Via Dei Lavoratori 54
Cinisello Balsamo 20092
Milan
Italy

Peptides 1990
E. Giralt and D. Andreu (Eds.)

Selective purification of large synthetic peptides using removable chromatographic probes

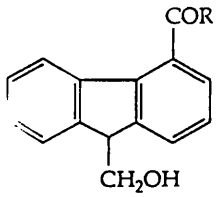
H.L. Ball[a], S.B.H. Kent[b] and P. Mascagni[a,*]
[a]The School of Pharmacy, University of London, London WC1N 1AX, U.K.
[b]Bond University, Brisbane, Australia

Introduction

The principal impurities arising from stepwise solid phase peptide synthesis (SPPS) are truncated and deletion peptides. The purification of peptides larger than 40 residues by conventional chromatographic and electrophoretic methods becomes a difficult task. Purification methods addressing this problem have been published [1–4]. However, a general purification scheme whose applicability is independent of the peptide sequence and length has not been demonstrated.

Results and Discussion

The selective purification system developed is based on the combination of (a) a capping protocol after each coupling step during SPPS and, (b) the addition of a removable protecting group to the terminal amino acid, bearing either lipophilic, acidic or basic functionalities (Fig. 1). The acid-stable 9-(hydroxymethyl)fluorene-4-carboxylate group (**1a**) [5] can be derivatised at the 4-position with chromatographic probes and then attached to the 'target' peptide through a base-labile urethane link.



1b R= -NHCH[$(CH_2)_7CH_3$]CONHCH[$(CH_2)_7CH_3$]$COOCH_3$

1c R= -{NHCH[$(CH_2)_4NH_3^+$]CO}$_2$NHCH[$(CH_2)_4NH_3^+$]$COOCH_3$

1d R= -{NHCH[$(CH_2)_2COO^-$]CO}$_2$NHCH[$(CH_2)_2COO^-$]$COOCH_3$

Fig. 1. **1a**

To illustrate this purification method we have synthesised peptides of various lengths (25–104 residues) and different chemical characteristics (Fig. 2). The twenty-five residue HIV-1 gp120 was synthesised using *t*-Boc chemistry and capped with 5% acetic anhydride in DMF after each coupling step. The peptidyl resin was derivatised with **1b** as the preformed Gly compound and cleaved with

*To whom correspondence should be addressed.

A. HIV-1 gp120 ENVELOPE PEPTIDE

L P T P R G P D R P E G I E E E G G E R D R D R S-$CONH_2$

B. FMDV CYCLIC PEPTIDE

C-Nle-G S G V R G D F G S L A P R V A R Q L P C(Acm)-Nle-Ahx-C-COOH

Fig. 2. Amino acid sequences.

HF. Purification of the target peptide was achieved by HPLC using a C-8 RP column. A separation between derivatised and underivatised peptides of 26 min resulted from the addition of the lipophilic Fmoc group (Table 1). Treatment of the derivatised peptide with 5% aqueous Et_3N for 30 min yielded pure HIV-1 gp120 peptide, as determined by FAB-MS and amino acid analysis (Table 1).

To show the general applicability of this purification scheme HIV-1 gp120 peptide was also derivatised with an ion exchange probe **1c**, which changed the net charge from −2 to zero. Following purification by anionic exchange FPLC, FAB-MS and amino acid analysis confirmed that the derivatised and underivatised peptides, after treatment with 5% aqueous Et_3N for 30 min, were the expected products (Table 1).

A third Fmoc probe **1d** was attached to a twenty-five amino acid FMDV peptide (Fig. 2) as the active succinimidyl ester, changing the net charge from positive to negative. After HF cleavage the peptide was cyclised, through two Cys residues at the N and C termini, by air oxidation in dilute solution. Purification by anionic exchange chromatography followed by removal of the Fmoc probe with organic base, yielded the cyclic FMDV peptide (Table 1) and a small impurity. The impurity probably resulted from a slight overlapping on the ion exchange column due to non-optimised conditions.

The 104-residue HIV-1 p24 GAG (270-373) fragment was synthesised using *t*-Boc chemistry, without capping and stored at −70°C before derivatisation with the preformed Gly analogue of **1b**. Separation between derivatised and underiva-

Table 1 *Comparison of HPLC and FPLC retention times and fast atom bombardment mass spectrometry (FAB-MS) of Fmoc derivatised and underivatised peptides*

Peptide	HPLC and FPLC analysis			FAB-MS analysis[b]	
	Retention time (min)				
	C-4 RP	C-8 RP	IE	Expected	Found
A (+1b)	23 (51)	20 (46)	–	2877 (3476)	2878 (3476)
A (+1c)	23 (27)	–	4[a] (10)	2877 (3543)	2877 (3545)
B (+1d)	–	33 (34)	4[a] (10)	2645 (3354)	2645 (3356)
HIV-1 (+ 1b)	48 (54)	48 (58)	–	–	–

Values in parentheses refer to the derivatised peptides, [a]Solvent front; [b]Mass of molecular ion, M^++H.

tised peptides was 6 min on a C-4 RP column and could be increased to 10 min on a C-8 column (Table 1), thus enabling effective purification. The molecular weight of the peptide, with and without the Fmoc probe, was assessed by SDS-PAGE and in both cases found to correspond to the expected masses.

Conclusions

The purification of peptides ranging from 25 to 104 residues using the proposed scheme has been successful. Reversed phase HPLC using α-aminodecanoic acid groups (**1b**) offers simpler and more general applicability and could be extended to Fmoc peptide chemistry. However, the versatility of this system allows for the use of aqueous chromatographic conditions (**1c**, **1d**), when organic solvents may have an adverse effect on the peptides stability or activity.

Acknowledgements

We wish to thank SERC for supporting H.L. Ball's Ph.D research. Special thanks go to Applied Biosystems for supplying the Model 151A HPLC system, Brownlee C-8 columns and performing the amino acid analysis.

References

1. Merrifield, R.B. and Bach, A.E., J. Org. Chem., 43(1978)4808.
2. Lobl, T.J., Deibel, M.R. and Yen A.W., Anal. Biochem., 170(1988)502.
3. Engstrom, A., Bennich, H. and Lindemberg, G., 11th American Peptide Symposium, San Diego, 1989, Abstracts p. 110.
4. Barany, G. and Merrifield, R.B., Peptides: Analysis, Synthesis, Biology, Vol. 2, In Gross, E. and Meienhofer, E. (Eds.), Academic Press, London, 1980, pp. 163–165.
5. Mutter, M. and Bellof, D., Helv. Chim. Acta., 67(1984)2009.

Affinity methods for purifying large synthetic peptides

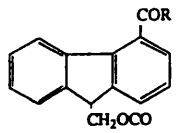
H. Ball[a], C. Grecian[a], S.B.H. Kent[b] and P. Mascagni[a]
[a]School of Pharmacy, 29–39 Brunswick Square, London WC1N 1AX, U.K.
[b]California Institute of Technology, Pasadena, CA 91125, U.S.A.

Introduction

The principal impurities arising from the stepwise nature of peptide synthesis are truncated peptides and deletion peptides. For peptides larger than 40 residues, the purification of the target peptide from these impurities, using conventional chromatographic and electrophoretic methods, becomes a difficult task.

Results and Discussion

We have devised a purification scheme based on the acid-stable Fmoc-4-carboxylic acid **1a** which improves on previous work involving either attachment of a Sulph-Fmoc group to the peptide [1], or a covalent biotin-peptide derivative [2]. **1a** can be derivatized with one of the groups shown in Fig. 1, and introduced via the urethane moiety onto the last amino acid of the sequence being synthesized. After HF cleavage, the purification of the peptide, thus derivatized, should be carried out using either reverse phase, ion-exchange or affinity chromatography. If this purification scheme is combined with capping procedures after each coupling step of the synthesis, then the purified peptide should also be free of deletion peptides.



1a

1b R1 = $NH_2CH(CH_2)_7CH_3COOtBu$

1c R2 = $NH_2CH(CH_2)_7CH_3CONHCH(CH_2)_7CH_3COOCH_3$

1d R3 = NH_2-(2ClCBZ)Lys-(2ClCBZ)Lys-(CBZ)Lys-OCH_3

Fig. 1. Fmoc derivatives used in this study.

To probe our hypothesis, we synthesized FMDV VP1-143-160 using the Boc-amino acid chemistry, and introducing the last residue (Gly) as the Fmoc derivative **1b**. After HF cleavage, the crude peptide was eluted through a C-18 reverse phase column. Figure 2 shows the HPLC chromatogram of the FMDV derivative; the peptide containing the Fmoc derivative eluted 8 min later than the underivatized peptide, thus confirming our hypothesis.

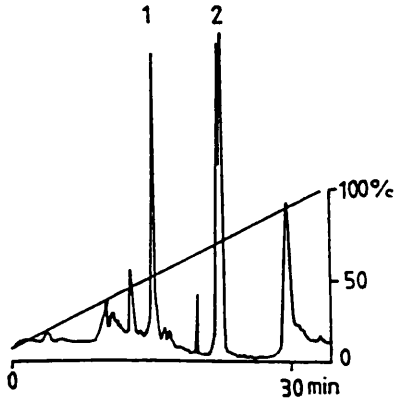


Fig. 2. C-18 RPHPLC of the FMDV(143–160) derivative with **1b**. *1 = underivatized peptide; 2 = derivatized peptide.*

Next we synthesized the 39–86 HIV-1 TAT fragment using the Boc strategy. However, this time, capping with 20% acetic anhydride in DCM containing 5% DIEA was performed after each coupling. The resin-peptide was stored in DMF at -70°C, and subsequently treated with TFA to remove the Boc group. Only 20% of the peptide chains were deprotected, as indicated by a quantitative ninhydrin test conducted after two TFA treatments.

Using a manual shaker, the Gly derivative of **1c** was then added as the HOBt active ester. The incorporation was monitored by taking small aliquots of peptide-resin and treating them with 20% piperidine in DMF. Readings of fluorene UV absorbance at 270 nm were then taken. The reaction was stopped when two successive such readings were found to be identical. After HF cleavage, the crude peptide was eluted on a C-4 column with a AcCN linear gradient. The derivatized peptide eluted at 68% AcCN, which contrasted with 33% for the uncoupled peptide. RPHPLC of 15 mg of crude material allowed the separation of these two peptides. Their molecular weight was assessed by SDS-PAGE and, in both cases, found to correspond to the expected masses.

Conclusions

These experiments have shown that the proposed scheme can be successfully applied to the purification of peptides of about 50 residues. The application of this procedure to 100-residue peptides using **1d** is currently under investigation.

References

1. Merrifield, R.B. and Bach, A.E., J. Org. Chem., 43 (1978) 4808.
2. Lobl, T.J., Deibel, M.R. and Yen, A.W., Anal. Biochem., 170 (1988) 502.

Peptides 1990
E. Giralt and D. Andreu (Eds.)

Chemical synthesis and biological characterisation of the *E. coli* 10 kDa chaperonin

M.S.L. Lim[a], H.L. Ball[a], M. Tonolo[a], A.R.M. Coates[b] and P. Mascagni[a,*]
[a]*The School of Pharmacy, University of London, London WC1N 1AX, U.K.*
[b]*St George's Hospital Medical School, London SW17 0RE, U.K.*

Introduction

Chaperonins (cpns), or heat-shock proteins, are important antigens in certain infectious diseases [1], may help trigger arthritis [2], and are involved in protein folding [2,3]. Biosynthesis of cpns increases after stress, such as exposure to heat or toxic chemicals [4], indicating a protective role for cpns; these proteins may also be a link in the immune response [5]. Cpns have been highly conserved during evolution and are ubiquitous in nature [4].

Two cpns from *Escherichia coli*, groEL (60 kDa) and groES (10 kDa), behave as molecular chaperones by assisting in oligomeric protein assembly [3]. We synthesised the 97-residue groES by automated solid phase methodologies, comparing the results from two independent chain assemblies which differed by an additional capping step, and also evaluating the coupling efficiencies with DDC and DIC. The fragments containing 24, 37, 59 and 79 residues were also obtained.

We report here the ability of the synthetic groES to (a) fold into the active multimer, and (b) restore, with recombinant groEL, the activity of denatured ribulose bisphosphate carboxylase (Rubisco) [3].

Results and Discussion

The groES protein was synthesised according to the sequence from the *E. coli* gene [6], on an ABI 430A using modified Merrifield chemistry [7]. The coupling efficiencies (Fig. 1a), as monitored by quantitive ninhydrin assays, were in most cases much higher with capping than without; the average yield increased from 99.4% to 99.7% with capping when DCC was used. DIC was found to be less effective as an activating agent than DCC (Fig. 1a).

After low TFMSA-high HF cleavage of groES, the crude material was divided into three portions which were taken up in (1) 5% HOAC, (2) phosphate buffered saline (PBS), and (3) 6 M urea; each was dialysed against 50 mM Tris-HCl at pH 7.6 then analysed by size exclusion chromatography. In solutions (1) and (3), GroES and the 19-97 fragment aggregated spontaneously, whereas solution (2) contained only the monomeric form.

*To whom correspondence should be addressed.

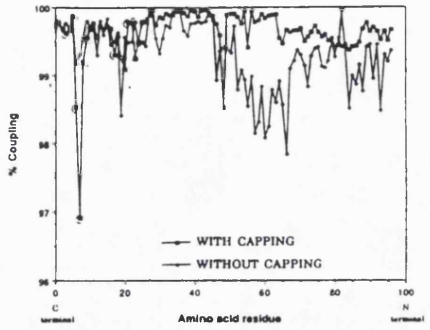


Fig. 1a. Percentage yield of each coupling step; encircled residues were coupled with DIC activation.

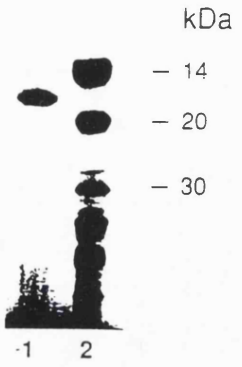


Fig. 1b. PhastGel Homogenous 20% SDS-PAGE of groES chaperonin.

The relative concentrations of the aggregates of full groES were: dimer (30%), tetramer (40%), hexamer/heptamer (20%) and higher aggregates (10%). SDS-PAGE of each fraction of groES each gave one main band at about 16 kDa (Fig. 1b), as reported in the literature for natural groES [4].

The folded groES and its 19–97 fragment were tested for activity in the Ribulose reconstitution assay [3]. Rubisco was denatured (Figs. 2a,b) and its ability to

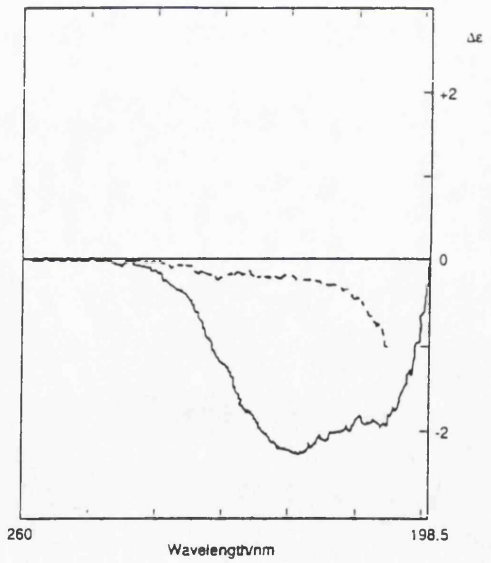


Fig. 2a. CD spectra of (i) natural, folded Rubisco (—) and (ii) Rubisco denatured in 6 M guanidine-HCl (---).

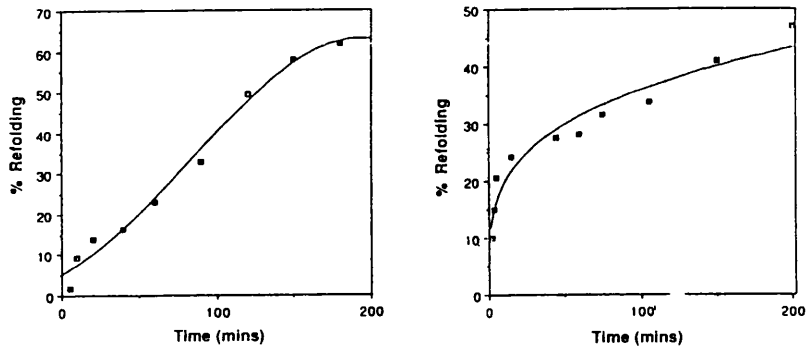


Fig. 2b. Activity profiles of Rubisco refolded by groES, groEL and Mg-ATP after denaturation by (i) 6 M guanidine-HCl and (ii) 0.1 M glycine-HCl.

incorporate ^{14}C-bicarbonate into ribulose bisphosphate under the assay conditions [3] was restored by incubation with the synthetic groES but not with the 19-97 fragment (Fig. 2b).

Acknowledgements

We thank Prof. G. Lorimer (E.I. du Pont de Nemours, U.S.A) and Dr A.A. Gatenby (Birkbeck College, University of London) for samples of recombinant Rubisco and groEL respectively. We also thank the Maplethorpe Foundation (MSLL), the Science and Engineering Research Council (HLB) and Applied Biosystems, U.K.

References

1. Baird, P.N., Hall, L.M.C. and Coates, A.R.M., J. Gen. Microbiol., 135(1989)931.
2. Res, P.C.M., Schaar, C.G., Breedveld, F.C., Eder, W. van, Embden, J.D.A. van, Cohen, I.R. and Vries, R.R.P. de, Lancet, (1988)478.
3. Goloubinoff, P., Christeller, J.T., Gatenby, A.A. and Lorimer, G.H., Nature, 342(1989)884.
4. Ellis, R.J., Sem. Cell Biol., 1(1990)1.
5. Lamb, J.R., Bal, V., Mendez-Samperio, P., Mehlert, A., So, A., Rothbard, J., Jindal, S., Young, R.A. and Young, D.B., Int. Immunol., 1(1989)191.
6. Hemmingsen, S.M., Woolford, C., Vies, S.M. van der, Tilly, K., Dennis, D.T., Georgopoulos, C.P., Hendrix, R.W. and Ellis, R.J., Nature, 333(1988)330.
7. Kent, S.B.H., Annu. Rev. Biochem., 57(1988)957.

Volume 286, number 1,2, 201–203 FEBS 09963 July 1991

ADONIS 001457939100670R

Chemical synthesis of 10 kDa chaperonin

Biological activity suggests chaperonins do not require other molecular chaperones

Paolo Mascagni[1,*], Mirta Tonolo[1], Haydn Ball[1,*], Marguerita Lim[1], R. John Ellis[2] and Anthony Coates[3]

[1]*Department of Pharmaceutical Chemistry, The School of Pharmacy, 29/39 Brunswick Square, London, WC1N 1AX, UK,* [2]*Department of Biological Sciences, University of Warwick, Coventry, CV4 7AL, UK* and [3]*Department of Medical Microbiology, St George's Hospital Medical School, Cranmer Terrace, London, SW17 ORE, UK*

Received 6 April 1991

Molecular chaperones are required for the correct folding and assembly of certain other polypeptides. It is not known whether molecular chaperones themselves require other chaperones to become functional. A 97-amino acid chaperone, the chaperonin 10 protein was chemically synthesised so that during synthesis and purification there was no contact of the chaperone with any other protein. The purified, synthetic chaperonin 10 protein formed oligomeric structures spontaneously and was biologically active as a chaperonin. This is the first description of a chemically synthesised chaperonin, and suggests that no other chaperones are required for correct folding, polymerisation and biological activity of this chaperone.

Chaperonin 10; GroES; Heat-shock protein; Chemical synthesis; Protein folding; Rubisco

1. INTRODUCTION

Molecular chaperones, are a family of ubiquitous cellular proteins whose function is to prevent incorrect interactions between the surfaces of other molecules; there is evidence that chaperones are required for the correct folding and assembly of certain other polypeptides [1,2]. This raises the question as to whether molecular chaperones themselves require other chaperones to become functional. In the yeast mitochondrion functional pre-existing chaperone, chaperonin 60 (cpn 60) [1,2] is needed so that new chaperonin 60 can be assembled from imported subunits [3]. In contrast, *Escherichia coli* cpn 60 can self assemble [4] in the presence of Mg-ATP and is stimulated by chaperonin 10 (cpn 10) [5]. So far, the chaperones which have been studied have either been purified from whole cell extracts or have been made by recombinant DNA techniques. In both cases, the opportunity is present for the chaperone protein to transiently interact with other chaperones within the cell during its folding and/or assembly. We have addressed this question by chemically synthesising a chaperone so that during the synthesis and purification there was no contact of the chaperone with any other protein. A 97-amino acid chaperone, the cpn 10 or 10 kDa GroES [5] protein of *E. coli* was synthesised.

**Present address:* Peptide Chemistry Department, Italfarmaco Research Center, Via Lavoratori 54, Cinesell Balsama 20092, Milan, Italy

Correspondence address: Professor Anthony R.M. Coates, Department of Medical Microbiology, St George's Hospital Medical School, Cranmer Terrace, London, SW17 0RE, UK

Abbreviations: cpn 10, chaperonin 10; SDS-PAGE, sodium dodecyl sulphate-polyacrylamide gel electrophoresis; cpn 60, chaperonin 60; BSA, bovine serum albumin; TFA, trifluoroacetic acid; HF, hydrogen fluoride; DNP, dinitrophenyl; DMF, dimethylformamide; TFA, trifluoracetic acid; HF, hydrogen fluoride

2. MATERIALS AND METHODS

2.1. *Chemical synthesis of E. coli chaperonin 10*

The synthesis of the polypeptide was carried out in a stepwise fashion using a 430A ABI synthesizer and the chemical scheme based on the combination of N(α)-BOC, benzyl-based side chain-protected amino acids. Other side chain-protecting groups included: tosyl for Arg, dinitrophenyl (DNP) for His, Cl- and Br-carbobenzoxy for Lys and Tyr, respectively. As the solid support we used the phenylacetamidomethyl resin [6] containing 0.34 mmol of Ala. All amino acids were double coupled using the chemical protocols developed by Kent [7] and the efficiency of each coupling step monitored by the quantitative ninhydrin test [8]. The average incorporation thus calculated was 99.35% [9]. After the completion of chain synthesis, the DNP group from the His side chain was removed by treating an aliquot of peptide resin twice with a dimethylformamide (DMF) solution containing 2-mercaptoethanol (20% v/v) and the t-BOC group from the last residue with 100% TFA.

To complete the side chain deprotection and remove the peptide from the resin the low-high HF procedure was used [10]; scavengers utilised were dimethylsulphide, *p*-cresol and *p*-thiocresol (6:0, 8:0.2 ml/0.5 g resin-peptide) for the low HF cleavage and *p*-cresol and *p*-thiocresol (0.8:0.2) during the high HF step.

Separation of the crude, diethylether-precipitated cpn 10 from the peptide-resin mixture proved difficult owing to the poor solubility of the crude polypeptide. However, after several washings with 5% acetic acid and lyophilisation of the combined solutions, 40 mg of crude polypeptide were obtained from 200 mg of resin-peptide material.

For the refolding assay 15 mg of lyophilised material was dissolved in 5 ml of 6 M urea (50 mM Tris buffer, pH 7.5) and dialysed at 4°C for 24 h against 4 liters of the same Tris buffer solution.

After dialysis the solution was tested for its protein content using the Bio-Rad Protein Assay reagent, concentrated by partial lyophilisation to the desired protein concentration (see below) and used without any further purification for the refolding assay, and for the FPLC and SDS-PAGE analysis.

2.2. *Size-exclusion chromatography of synthetic chaperonin 10*

The separation was performed with a Pharmacia Sephacryl S-100 column using a Pharmacia FPLC apparatus. The solvent (100 mM Tris-HCl, 10 mM $MgCl_2$, pH 7.8) was applied at a 2.5 ml/min rate; monitoring was with a UV recorder at 276 nm.

2.3. *SDS-PAGE*

SDS-PAGE analysis was performed on the Pharmacia Phast System with 20% gradient gels.

2.4. *Refolding of recombinant rubisco protein using synthetic* E. coli *cpn 10 and recombinant cpn 10*

Rubisco was denatured in 6 M guanidine-HCl and refolded using the protocol described in [11]. Briefly, reconstitution was initiated by diluting 21 μl denatured rubisco into 1659 μl solution containing 50 mM Tris-HCl buffer, pH 8.0, 3 mM ATP, 7 mM $MgCl_2$, 20 mM glucose, 4 μM cpn 60 and 6 μM cpn 10 at 23°C. At various times the reconstitution reaction was quenched with hexokinase. Reaction mixtures were supplemented with 50 mM [^{14}C]$NaHCO_3$ (300 dpm·nmol^{-1}) and allowed to activate for a further 5 min. Ribulose biphosphate (1 mM) was added and rubisco activity measured as the formation of acid-stable ^{14}C over 20 min.

3. RESULTS AND DISCUSSION

The purified, synthetic cpn 10 formed oligomeric structures which are also found in recombinant GroES [5]. The formation of oligomers was demonstrated by size-exclusion FPLC in non-denaturing buffer. We found that there were peaks at about 200 kDa (numbered 1 in Fig. 1), 90 kDa (peak 2), 30–40 kDa (peak 3) and 15–30 kDa (peak 4). The 90 kDa peak probably corresponds to an 80 kDa oligomer described previously [5]. The 30–40 kDa peak may represent a

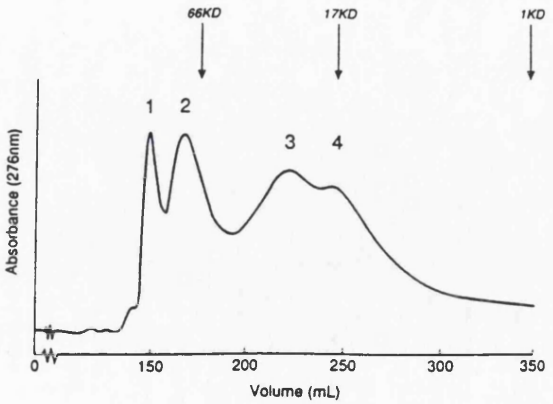


Fig. 1. Size-exclusion chromatography of synthetic chaperonin 10. Peaks 1, 2, 3 and 4 eluted at 150, 170, 225 and 245 ml, respectively. In a calibration run, BSA (mol. wt. 66000) eluted at 175 ml and cytochrome *c* (mol. wt. 13000) at 250 ml.

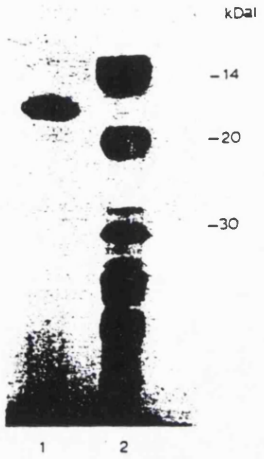


Fig. 2. SDS-PAGE of synthetic cpn 10. (Track 1) Synthetic cpn 10 from dialysis. (Track 2) markers.

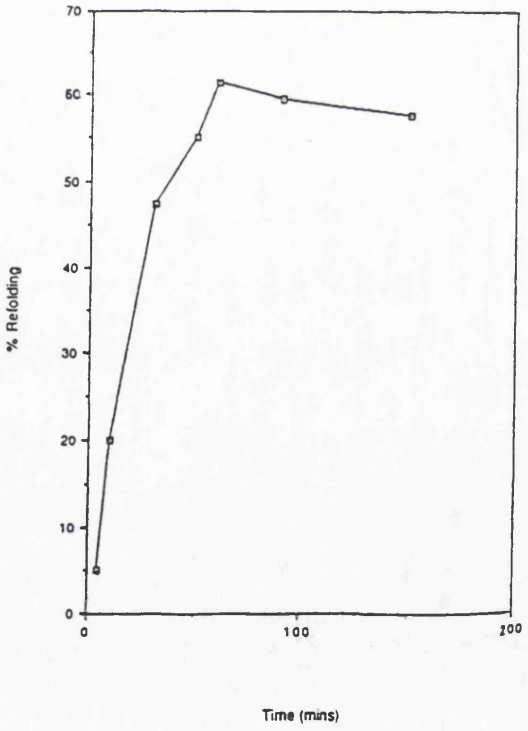


Fig. 3. Refolding of recombinant rubisco protein using whole, unfractionated synthetic cpn 10 and recombinant cpn 60 (rubisco binding protein). Rubisco refolding is expressed as a percentage of the enzymic activity of an equal quantity of non-denatured native enzyme.

dimer or a trimer. On SDS-PAGE (Fig. 2), protein from the 30–40 kDa size-exclusion peak migrated as a single band at about 15 kDa, consistent with the previous observations with recombinant cpn 10 [5]. SDS-PAGE of the 90 kDa and 200 kDa peaks in the size-exclusion chromatography also give a single band at 15 kDa (data not shown).

We then tested the synthetic protein for biological activity as a chaperonin. Recently Goloubinoff and colleagues [11,12] have described an in vitro assay for cpn activity. Denatured recombinant rubisco protein from *Rhodospirillum rubrum* is biologically inactive, but in the presence of recombinant *E. coli* cpn 10 and the associated cpn 60, as well as Mg-ATP and potassium ions, rubisco assembles into a dimer and regains its enzymic activity. We repeated this assay with recombinant reagents kindly supplied to us by A.A. Gatenby and G. Lorimer (Dupont de Nemours & Co., Wilmington, USA) except that we used chemically synthesised rather than recombinant cpn 10. The addition of dialysed, complete synthetic cpn 10 to denatured rubisco in the presence of recombinant cpn 60, Mg-ATP and potassium ions yielded a rubisco which was enzymically active as shown by the fixation of radioactive carbon dioxide (Fig. 3). Denatured rubisco incubated either alone or in the presence of either cpn 60 or cpn 10 was completely inactive in this assay. The synthetic cpn 10 showed comparable specific activity to that published for the recombinant cpn 10 [11]. Denaturation of synthetic cpn 10 with urea, followed by dilution of the urea, also resulted in functionally active cpn 10. To our knowledge this is the first description of the chemical synthesis of a biologically active chaperonin.

These experiments suggest that in the case of the cpn 10, no other chaperones are required for correct folding, polymerisation and biological activity. In other words, unlike rubisco, the cpn 10 does not need chaperones to reconstruct its activity in an in vitro system. Although *E. coli* cpn 60 can self-assemble [4] it is not yet known whether this renatured cpn 60 is active in the rubisco assay. It is unlikely that the addition of cpn 60 in the rubisco assay causes the unfolded cpn 10 to fold up correctly because the cpn 10 which was added was already in oligomeric form. These observations suggest that the cpn 10 molecular chaperone falls into the category of proteins whose folding and assembly into functional structures occurs with sufficient probability to dispense with the services of chaperones [2]. Whether this is true for all chaperones remains to be determined.

Acknowledgements: We thank A.A. Gatenby and G.H. Lorimer (Dupont de Nemours & Co., Wilmington, USA) for the supply of rubisco protein and rubisco binding protein.

REFERENCES

[1] Ellis, R.J. and Hemmingsen, S.M. (1989) Trends Biochem. Sci. 14, 339–342.
[2] Ellis, R.J. (1990) Sem. Cell. Biol. 1, 1–72.
[3] Cheng, M.Y., Hartl, F.-U. and Horwich, A.L. (1990) Nature 348, 455–458.
[4] Lissin, N.M., Venyaminov, S. Yu and Girshovich, A.S. (1990) Nature 348, 339–342.
[5] Chandrasekhar, G.N., Tilly, K., Woolford, C., Hendrix, R.W. and Georgopoulos, C. (1986) J. Biol. Chem. 261, 12414–12419.
[6] Mitchell, A.R., Erickson, B.W., Ryabtsev, M.N., Hodges, R.S. and Merrifield, R.B. (1976) J. Am. Chem. Soc. 98, 7357.
[7] Kent, S.B.H. and Clark-Lewis, I. (1985) Synthetic Peptides in Biology and Medicine (K. Alitalo, D. Partanen and A. Vaheri eds.) Elsevier, Amsterdam.
[8] Sarin, V.K., Kent, S.B.H., Tam, J.P. and Merrifield, R.B. (1981) Anal. Biochem. 117, 147.
[9] This value does not take into account a background colour which accumulated during the synthesis. This was estimated to be about 0.5% by testing the amino acids which follow the Pro residues.
[10] Tam, J.P., Heath, W.F. and Merrifield, R.B. (1983) J. Am. Chem. Soc. 105, 2845.
[11] Goloubinoff, P., Christeller, J.T., Gatenby, A.A. and Lorimer, G.H. (1989) Nature 342, 884–889.
[12] Viitanen, P.V., Lubben, T.H., Reed, J., Goloubinoff, P., O'Keefe, D.P. et al. (1990) Biochemistry 29, 5665–5671.

REFERENCES

Albericio, F., Pons, M., Pedroso, E. and Giralt, E., (1989), *J. Org. Chem.* **54** 360-366

Albertson, N. F. and Archer, S., (1945), *J. Am. Chem. Soc.*, **67**, 308

Albertson, N. F. and Tullor, B. F., (1945), *J. Am. Chem. Soc.*, **67**, 502

Anderson, G. W. and McGregor, A. C., (1957), *J. Am. Chem. Soc.* **79**, 6180

Atherton, E., Fox, H., Harkiss, D., Logan, C. J., Sheppard, R. C. and Williams, B. J., (1978a), *J. Chem. Soc. Chem. Commun.* 537-539

Atherton, E., Fox, H., Harkiss, D. and Sheppard, R. C., (1978b), *J. Chem. Soc. Chem. Commun.* 539-540

Atherton, E., Brown, E., Priestley, G., Sheppard, R. C. and Williams, B. J., (1981), *Proc. Seventh American Peptide Symposium,* Pierce Chemical Co., Rockford, p.163-174

Atherton, E., Cameron, L. R., Cammish, L. E., Dryland, A., Goddard, P., Priestley, G. P., Richards, J. D., Sheppard, R. C., Wade, J. D. and Williams, B. J., (1990), *Innovation and Perspectives in Solid Phase Synthesis*, Oxford, UK, Ed. Epton, R., SPCC, Birmingham, p.15-25

Bannow, C. A., Tomasselli, A. G., Zurcher-Neely, H. A., Smith, C. W. and Heinrikson, R. L., (1991), *Proc. Twelfth American Peptide Symposium*, Boston, USA, Eds. Smith, J. A. and Rivier, J. E., ESCOM, p.738-739

Barany, G. and Albericio, F., (1985), *J. Am. Chem. Soc.* **107**, 4936- 4942

Barlos, K., Dimitrios, G., Gatos, D., Kallitsis, J., Papaphotiu, G., Sotirin, P., Wenqin, Y. and Schafer, W., (1989), *Tet. Lett.*, **30**, 3943

Barlos, K., Gatos, D. and Schafer, W., (1990), *Proc. Twenty-First European Peptide Symposium*, Barcelona, Spain, Eds. Giralt, E. and Andreu, D., ESCOM, p.247-248

Barlos, K., Gatos, D. and Schafer, W., (1991), *Angew. Chem. Acta. Int. Ed. Engl.*, **30**(5), 590-593

Barton, M. A., Lemieux, R. U. and Savoie, J. Y., (1973), *J. Am. Chem. Soc.*, **95**, 4501-4506

Ben-Ishai, D. and Berger, A., (1952), *J. Org. Chem.*, **17**, 1564

Benoiton, N. L. and Chen, F. M. F., (1981), *Canad. J. Chem.*, **59**, 384

Bergmann, M. and Zervas, L., (1932), *Ber. Dtsch. Chim. Ges.*, **65**, 1192-1201

Birr, C., Reitel, C. and Wieland, T., (1970), *Angew. Chem. Int. Ed. Engl.*, **9**, 731

Blake, J. and Li, C. H., (1968), *J. Am. Chem. Soc.*, **90**, 5882-5884

Blake, J., Yamashiro, D., Ramashara, K. and Li, C. H., (1986), *Int. J. Peptide Protein Res.*, **28**, 468-476

Bodanszky, M. and Bodanszky, A., (1967), *Chem. Commun.*, 591-593

Bodanszky, D., (1984), *Principles of Peptide Chemistry*, Springer-Verlag, Berlin, p.224

Carpino, L. A., (1957), *J. Am. Chem. Soc.*, **79**(98), 4427

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

Carpino, L. A. and Han, G.-Y., (1972), *J. Org. Chem.* **37**, 3404-3409

Castro, B., Domoy, J. R., Evin, G. and Selve, C., (1975), *Tet. Lett.*, **14** 1219-1222

Chandrasekhar, G. N., Tilly, K., Woolford, C., Hendrix, R. W. and Georgopoulos, C., (1986), *J. Biol. Chem.*, **261**, 12414-12419

Curtius, T., (1902), *Ber. Dtsch. Chim. Ges.*, **35**, 3226-3233

Denkewalter, R. G., Veber, D. F., Holly, F. W. and Hirschmann, R., (1969), *J. Am. Chem. Soc.*, **91**, 502

Dourtoglou, V., Gross, B., Lambropoulou, V. and Zioudrou, C., (1984), *Synthesis* , 572-574

du Vigneaud, V., Ressler, C., Swan, J. M., Roberts, C. W., Katsoyannis, P. G. and Gordon, S., (1953), *J. Am. Chem. Soc.*, **75**, 4879

Dwulet, F. E. and Gurd, F. R., (1976), *Anal. Biochem.*, **76**, 530-538

Eckert, H. and Forster, B., (1987), *Angew. Chem. Int. Ed. Engl.*, **26**(9), 894-895

Ellis, J. E. and van der Vies, (1991), *Ann. Rev. Biochem.*, **60**, 321-347

Engstrom, A., Bennich, H. and Lindemberg, G., (1989), Poster Presentation at Eleventh American Peptide Symposium, San Diego, USA, Abstracts p.110

Finn, F. M. and Hofmann, K. (1976), *The Proteins*, Ed. Neurath, H. and Hill, R., 3rd Edn vol 2, Academic Press, New York, p.105-253

Fischer, E., (1902), *Ber. Dtsch. Chem. Ges.*, 1095

Fischer, E., (1906), *Ber. Dtsch. Chim. Ges.*, **39**, 530-610

Florsheimer, A. and Riniker, B., (1990), *Proc. Twenty-First European Peptide Symposium*, Barcelona, Spain, Eds. Giralt, E. and Andreu, D., ESCOM, p.131-133

Fridkin, M. and Patchornik, A., (1974), *Ann. Rev. Biochem.*, **43**, 419-443

Gabriel, T. F., Jimenez, M. H., Felix, A. M., Michalewsky, J. and Meienhofer, J., (1977), *Int. J. Peptide Protein Res.*, **9**, 129-136

Galpin, I. J., Kenner, G. W., Ohlsen, S. R. and Ramage, R., (1975), *J. Chromatography*, **129**, 287-293

Giralt, E., Eritja, R. and Pedroso, E. (1981) *Tet. Lett.* **22**, 3779-3782

Giralt, E., Albericio, F., Pedroso, E., Granier, C. and van Rietschoten, J. (1982), *Tetrahedron*, **38**(9), 1193-1201

Giralt, E., Rizo, J. and Pedroso, E., (1984), *Tetrahedron*, **40**, 4141-4152

Gish, D. T., Katsoyannis, P. G., Hess, G. P. and Stedman, R. J., (1956), *J. Am. Chem. Soc.*,**78**, 5954

Gisin, B. F. and Merrifield, R. B., (1972), *J. Am. Chem. Soc.*, **94**, 3102-3106

Gisin, B. F., (1972), *Ann. Chim. Acta.*, **58**, 248

Goerdeler, J. and Holst, A., (1959), *Angew. Chem.*, **71**, 775

Grandas, A., Albericio, F., Josa, J., Giralt, E., Pedroso, E., Sabatier, J. and van Rietschoten, J., (1989) *Tetrahedron*, **45**(14) 4637-4648

Grode, S. H., Strother, D. S., Runge T. A. and Dobrowolski, P. J., (1992), *Int. J. Peptide Protein Res.*, **40**, 538-545

Guibe, F., Dangles, O., Balavoine, G. and Loffet, A., (1989), *Tet. Lett.*, **30**, 2641-2644

Gutte, B. and Merrifield, R. B., (1969), *J. Am. Chem. Soc.*, **91**, 1348

Gutte, B. and Merrifield, R. B., (1971), *J. Biol. Chem.*, **246**, 1922

Hofmeister, F., (1902), *Ergeb. Physiol. Biol. Chem. Exp. Pharmacol.*, 759

Hubbuch, A., Dahno, W. and Zahn, H., (1977), *Proc. Fifth American Peptide Symposium*, Eds. Goodman, M. and Meienhofer, J., p.540-542

Izumiya, N., Waki, M., Kato, T., Ono, M., Aoyagi, N. and Mitsuyasu, N., (1972), *Chemistry and Biology of Peptides*, Ed. Meienhofer, J., Ann Arbor Sci. Publ., Ann Arbor, USA, p.269

Kaiser, E. T., Mihara, H., Laforet, G. A., Kelly, J. W., Walters, L., Findeis, M. A. and Sasaki, T., (1989), Science, **243**, 187-192

Kamber, B. and Riniker, B., (1991), *Proc. Twelfth American Peptide Symposium*, Boston, USA, Ed. Smith, J. A. and Rivier, J. E., ESCOM, pp. 525-526

Kenner, G. W., Ramage, R. and Shepherd, R. C., (1979), *Tetrahedron* **35**, 2767

Kent, S. B. H., Mitchell, A. R., Engelhard, M. and Merrifield, R. B., (1979), *Proc. Natl. Acad. Sci.*, **76**, 2180

Kent, S. B. H., (1984), *Peptides: Structure and Function* , Eds. Hruby, V. J. and Rich, D. H., Pierce Chemical Co., p.99-102

Kent, S. B. H. and Clark-Lewis, I, (1985), *Synthetic Peptides in Biology and Medicine*, Eds. Alitalo, K., Partanen, D. and Vaheri, A., Elsevier Science Publ., Amsterdam, p.29-57

Kent, S. B. H., (1988), *Ann. Rev. Biochem.*, **57**, 957-989

Khosla, M. C., Smeby, R.R. and Bumpus, F. M., (1972), *J. Amer. Chem. Soc.*, **74**, 4721

Kiyama, S., Fujii, N., Yajima, H., Moriga, M. and Takagi, A., (1984), *Int. J. Peptide Protein Res.*, **23**, 174-186

Kneib-Cordonier, N., Albericio, F. and Barany, G., (1990), *Int. J. Peptide Protein Res.*, **35**, 527-538

Knorr, R., Trzeciak, A., Bannwarth, W. and Gillessen, D., (1989), *Tet. Lett.*, **30**, (15), 1927-1930

Konig, W. and Geiger, R., (1970a), *Chem. Ber.*, **103**, 788-798

Konig, W. and Geiger, R., (1970b), *Chem. Ber.*, **103**, 2024-2034

Konig, W. and Geiger, R., (1970c), *Chem. Ber.*, **103**, 2034-2040

Kovacs, J., Cover, R., Jham, G., Hsieh, Y. and Kalas, T., (1975), *Peptides: Chemistry, Structure and Biology*, Eds. Walter, R. and Meienhofer, J., Ann Arbor Science Publ., Ann Arbor, p.317-324

Krieger, D. E., Erickson, B. W. and Merrifield, R. B., (1976) *Proc. Natl. Acad. Sci. USA*, **73**, 3160-3164

Kunz, H. and Dombo, B., (1988), *Angew. Chem. Int. Ed. Engl.*, **27**, 710-713

Kunz, H., Kosch, W. and Marz, J., (1991), *Proc. Twelfth American Peptide Symposium*, Boston, USA, Eds. Smith, J. A. and Rivier, J. E., ESCOM, p.502-504

Kuroda, H., Chen, Y-N., Kimura, T. and Sakakibara, S., (1992), *Int. J. Peptide Protein Res.*, **40**, 294-299

Li, C. H. and Yamashiro, D., (1970), *J. Am. Chem. Soc.*, **92**, 7608

Lindeberg, G., Tengborn, J., Bennich, H. and Ragnarsson, U., (1978), *J. Chromatography*, **156**, 366-369

Liu, Y.-Z., Ding, S.-H., Chu, J.-Y. and Felix, A. M., (1990), *Int. J. Peptide Protein Res.*, **35**, 95-98

Lloyd-Williams, P., Jou, G., Albericio, F. and Giralt, E., (1991a), *Tet. Lett.*, **32**(33), 4207-4210

Lloyd-Williams, P., Gairi, M., Albericio, F. and Giralt, E., (1991b), *Tetrahedron*, **47**(47), 9867-9880

Lloyd, D. H., Petrie, G. M., Noble, R. L. and Tam, J. P., (1989), *Proc. Eleventh American Peptide Symposium*, San Diego, USA, Eds. Smith, J. A. and Rivier, J. E., ESCOM, p.909-910

Lobl, T. J., Deibel, M. R. and Yen, A. W., (1988), *Anal. Biochem.*, **170**, 502

Lukenheimer, W. and Zahn, H., (1970), *Liebigs Ann. Chem.*, **740**, 1

Markley, M. D. and Dorman, L. C., (1970), *Tet. Lett.*, 1787-1790

Maruyama, H., Matsueda, R., Kitazawa, E., Takahagi, H. and Mukaiyama, T., (1976), *Bull. Chem. Soc. Jpn.*, **49**, 2259

Mascagni, P., Hyde, C. B., Charalambous, M. A. and Welham, K. J., (1987), *J. Chem.. Soc. Perkin Trans. II*, 323-327

McKay, F. C. and Albertson, N. F., (1957), *J. Am. Chem. Soc.*, **79**, 468

Meienhofer, J., Cotton, R. and Atherton, E., (1971), *J. Org. Chem.*, **36**, 3746-3748

Meienhofer, J., (1979), "The Azide Method in Peptide Synthesis" in *The Peptides* Vol 1, Eds. Meienhofer, J. and Gross, E., Academic Press, p.197-239

Meister, S. M. and Kent, S. B. H., (1980), *Proc. Sixteenth European Peptide Symposium*, p103-106

Mergler, M., Nyfeler, R., Tanner, R., Gosteli, J. and Grogg, P., (1988), *Tet. Lett.* **29**(32), 4009-4012

Merrifield, R. B., (1963), *J. Am. Chem. Soc.*, **85**, 2149-2154

Merrifield, R. B., (1963), *Recent Prog. Hormone Res.*, **23**, 451-482

Merrifield, R. B. and Bach, A. E., (1978), *J. Org. Chem.*, **43**, 4808-4816

Mitchell, A. R., Kent, S. B. H., Engelhard, M. and Merrifield, R. B., (1978), *J. Org. Chem.*, **43**, 2845-2852

Mutter, M. and Bellof, D., (1984), *Helv. Chim. Acta.* **67**, 2009-2016

Nakatsuka, T., Sasaki, T. and Kaiser, E. T., (1987), *J. Am. Chem. Soc.*, **109**, 3808-3810

Nyfeler, R., Wixmerten, U., Seidel, C. and Mergler, M., (1991), *Proc. Twelfth American Peptide Symposium*, Boston, USA, Eds. Smith, J. A. and Rivier, J.E. ESCOM, p.661-663

Nystrom, E. and Sjovall, J., (1963), *J. Chromatography*, 208-212

Ondetti, M. A., Deer, A., Sheehan, J. T., Pluscec, J. and Kocy, O., (1968), *Biochemistry*, **7**, 4069

Otteson, K. M., McCurdy, S. M., Harrison, J. L., Farnam, W. E. and Pierce, L. G., (1988), Poster Presentation at Thirteenth Ann. Lorne Conference on Protein Structure and Function.

Paquet, A., (1982), *Can. J. Chem.*, **60**, 976

Patchornik, A., Amit, B. and Woodward, R. B., (1970), *J. Am. Chem. Soc.*, **92**, 6333

Pedroso, E., Grandas, A., de las Heras, X., Eritja, R. and Giralt, E., (1986), *Tet. Lett.* **27**(6), 743-746

Pedroso, E., Grandas, A., Saralegui, M. A., Giralt, E., Granier, C. and van Rietschoten, J., (1982), *Tetrahedron*, **38**(9), 1183-1192

Penke, B. and Birr, C., (1974), *Justus Liebigs Ann. Chem.*, 1999-2002

Pillai, V. N. and Mutter, M., (1981), *Acc. Chem. Res.*, **14**, 122-130

Piszkiewicz, D., Landon, M. and Smith, E. L., (1970), *Biochem. Biophys. Res. Commun.*, **40**, 1173

Preston, P. N. and Tennant, G., (1972) *Chem. Rev.*, **72** 627

Rich, D. H. and Gurwara, S. K., (1973), *J. Chem. Soc. Chem. Commun.*, 617

Rich, D. H. and Gurwara, S. K. (1975), *J. Am. Chem. Soc.*, **97**, 1575-1579

Rink, H., (1987), *Tet. Lett.*, **28**(33), 3787-3790

Rothe, M. and Mazanek, J., (1974), *Liebigs Ann. Chem.*, 439

Sakakibara, S., Shin, K. H. and Hess, G. P., (1962), *J. Am. Chem. Soc.*, **84**, 4921

Sakakibara, S. and Shimonishi, Y., (1965), *Bull. Chem. Soc. Jpn.*, **38**, 1412

Sanger, F., (1953). *Nature*, **171**, 1025

Sano, S. and Kurihara, M., (1969), *Hoppe Seyler's Z. Physiol. Chem.*, **350**, 1183

Sarin, V. K., Kent, S. B. H., Tam, J. P. and Merrifield, R. B., (1981), *Anal. Biochem.*, **117**, 147-157

Schnolzer, M. and Kent, S. B. H., (1992), *Science*, **256**(5054), 221-225

Schon, I., Colombo, R. and Csehi, A., (1983), *J. Chem. Soc. Chem. Commun.*, 505-507

Sharp, J., Robinson, A. B. and Kamen, M. D., (1973), *J. Am. Chem. Soc.*, **95**, 6097

Sheehan, J. C. and Hess, G. P., (1955), *J. Am. Chem. Soc.*, **77**, 1067-1068

Shin, K. H., Sakakibara, S., Schneider, W. and Hess, G. P., (1962), *Biochem. Biophys. Res. Commun.*, **8**, 288

Sieber, P. and Iselin, B., (1968), *Helv. Chim. Acta.* **51**, 614-622

Sifferd, R. H. and du Vigneaud, V., (1935), *J. Biol. Chem.*, **108**, 753

Stewart, J. M. and Young, J. D., (1984), *Solid Phase Peptide Synthesis,* 2nd Ed., Pierce Chem. Co., Rockford, Illinois, p.10

Suzuki, K., Nitta, K. and Sasaki, Y., (1976), *Chem. Pharm. Bull.*, **24**, 3025-3033

Suzuki, K. and Endo, N., (1977), *Chem. Pharm. Bull.* **23**, 1143-1146

Tam, J. P., Cunningham-Rundles, W. F., Erickson, B. W. and Merrifield, R. B., (1977), *Tet. Lett.*, **46**, 4001-4004

Tam, J. P., Wong, T. W., Rieman, M.W., Tjoeng, F. S. and Merrifield, R. B., (1979), *Tet. Lett.*, 4033

Tam, J. P., Tjoeng, F. S. and Merrifield, R. B., (1980), *J. Am. Chem. Soc.*, **102**, 6117

Thaler, A., Seebach, D. and Cardinaux, (1991), *Helv. Chem. Acta.*, **74**(1), 617-627 and **74**(2), 628-643

Tjoeng, F. S., Tong, E. K. and Hodges, R. S., (1978), *J. Org. Chem.*, **43**, 4190

van Melik, J. E. and Wolters, E. T., (1972), *Synthetic Commun.*, **2**, 83-86

Wallace, C. J. A. and Campbell, L. A., (1989), *Proc. Eleventh American Peptide Symposium*, San Diego, USA, Eds. Smith, J. A. and Rivier, J. E., ESCOM, p.1043-1045

Wang, S. S., (1976), *J. Org. Chem.*, **41**, 3258

Weygand, F., Hoffmann, D. and Wunsch, E., (1966), *Naturforsch*, **216**, 246

Whitney, D. B., Tam, J. P. and Merrifield, R. B., (1984), *Tetrahedron*, **40**(21), 4237-4244

Wieland, T., Birr, C. and Flor, F., (1969), *Angew. Chem, Int. Ed. Engl.*, **8**, 764-765

Wolters, E. T., Tesser, G. I. and Nivard, R. J., (1974), *J. Org. Chem.*, **39**, 3388-3392

Wunsch, E. and Drees, F., (1966), *Chim. Ber.*, **99**, 110

Yajima, H., Kiso, Y., Okada, Y. and Watanabe, H., (1974), *J. Chem. Soc. Chem. Commun.*, 106-107

Yajima, H. and Fujii, N., (1980), *Chem.. Commun.*, 115

Zeiger, A. R. and Anfinsen, C. B., (1973), *J. Am. Chem. Soc.*, **95**, 880-886

Zervas, L., Borovas, D. and Grazis, E., (1963), *J. Am. Chem. Soc.*, **85**, 3660