This thesis is submitted in part fulfilment of the requirement for the degree of Doctor of Philosophy at the University of Edinburgh. Unless otherwise stated, the work is original, and has not been previously submitted, in whole or in part, for any degree at this, or any other university.
"Time held me green and dying
Though I sang in my chains like the sea"

*Extract from 'Fern Hill' by Dylan Thomas.*
To my Granda
Abstract

The following work documents three studies undertaken using solid phase synthesis techniques.

Interaction of the zinc metallopeptidase, Endothelin Converting Enzyme-1 (ECE-1), with its natural peptidic substrate, big endothelin-1 (big ET-1), has been investigated via an SAR study using solid phase peptide synthesis (SPPS). Truncated forms of the substrate had been previously reported to inhibit ECE-1, this was confirmed however the big ET-1 analogues were shown to be alternative substrates for the enzyme. A short study of the substrate specificity of ECE-1 was carried out.

The synthesis of vast libraries of peptides using combinatorial synthesis has been used to accelerate the drug discovery process. Purification of these mixtures has not been previously attempted. 17-Tetrabenzo[a,c,g,i]fluorenylmethoxycarbonyl (Tbfmoc), developed for use with single peptides and proteins, has been used to achieve facile purification of five peptide libraries synthesised using SPPS. The methodology was fully optimised for the efficient separation of the desired library members from all impurities by exploiting the affinity of Tbfmoc for carbon.

A potential small molecule inhibitor of the zinc metalloenzyme, Farnesyl transferase (FTase), was designed. The efficient solution phase synthesis of this novel structure is reported. Its adaptation to solid phase synthesis is described, with the view to using multiple parallel synthesis techniques to synthesise a range of analogues.
Acknowledgements

My sincere thanks to Professor Ramage for granting me the opportunity to undertake my Post-graduate research at the University of Edinburgh, under his supervision. His encouragement, optimism and ideas have been a continual source of inspiration.

I would also like to express my gratitude to Parke Davis for providing the financial support for the work, also to Kay Ahn and her group for the biological testing of my compounds. I extend a particular thank you to Daniele Leonard, on whose personal contact and input I have greatly depended.

I must also acknowledge the Chemistry Department analytical services for their prompt service and ever amicable reception, together with the assistance provided by Kevin Shaw and Brian Whigham, to all of whom I am greatly indebted. Thanks also to my friends and colleagues in lab. 29.

Finally, my greatest debt of gratitude is to my family, especially to my parents, whose unceasing support, devotion and love will sustain me always.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ac</td>
<td>acyl</td>
</tr>
<tr>
<td>Acm</td>
<td>acetamidomethyl</td>
</tr>
<tr>
<td>Ar</td>
<td>aryl</td>
</tr>
<tr>
<td>BET</td>
<td>big endothelin</td>
</tr>
<tr>
<td>big ET</td>
<td>big endothelin</td>
</tr>
<tr>
<td>Boc</td>
<td>t-butyloxycarbonyl</td>
</tr>
<tr>
<td>Bn</td>
<td>benzyl</td>
</tr>
<tr>
<td>Bu</td>
<td>butyl</td>
</tr>
<tr>
<td>Bum</td>
<td>t-butoxymethyl</td>
</tr>
<tr>
<td>Cbz</td>
<td>benzylxycarbonyl</td>
</tr>
<tr>
<td>CNS</td>
<td>central nervous system</td>
</tr>
<tr>
<td>COS</td>
<td>combinatorial organic synthesis</td>
</tr>
<tr>
<td>CTF</td>
<td>C-terminal fragment</td>
</tr>
<tr>
<td>d</td>
<td>doublet (nmr)</td>
</tr>
<tr>
<td>Da</td>
<td>Daltons</td>
</tr>
<tr>
<td>DCC</td>
<td>dicyclodicyclohexylcarbodiimide</td>
</tr>
<tr>
<td>DCM</td>
<td>dichloromethane</td>
</tr>
<tr>
<td>DEAD</td>
<td>diethyl azodicarboxylate</td>
</tr>
<tr>
<td>DIC</td>
<td>N, N'-diisopropylcarbodiimide</td>
</tr>
<tr>
<td>DIEA</td>
<td>N, N-diisopropylethylamine</td>
</tr>
<tr>
<td>DMAP</td>
<td>dimethylaminopyridine</td>
</tr>
<tr>
<td>DME</td>
<td>dimethoxyethane</td>
</tr>
<tr>
<td>DMF</td>
<td>N, N- dimethylformamide</td>
</tr>
<tr>
<td>DMS</td>
<td>dimethylsulphide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulphoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>DVB</td>
<td>divinylbenzene</td>
</tr>
<tr>
<td>ECE</td>
<td>Endothelin Converting Enzyme</td>
</tr>
<tr>
<td>ECGC</td>
<td>electron capture gas chromatography</td>
</tr>
<tr>
<td>EDT</td>
<td>ethane-1, 2-dithiol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>EDRF</td>
<td>Endothelium derived relaxing factor</td>
</tr>
<tr>
<td>EI</td>
<td>electron impact</td>
</tr>
<tr>
<td>ELISA</td>
<td>Enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>eq</td>
<td>equivalents</td>
</tr>
<tr>
<td>ES</td>
<td>electrospray</td>
</tr>
<tr>
<td>Et</td>
<td>ethyl</td>
</tr>
<tr>
<td>ET</td>
<td>endothelin</td>
</tr>
<tr>
<td>FAB</td>
<td>fast atom bombardment</td>
</tr>
</tbody>
</table>
PPA  polyphosphoric acid
ppm  parts per million
Pr   propyl
PS   polystyrene
q    quartet (nmr)
q    quaternary
RB   round bottomed (flask)
Rf   retention factor
RF   radio frequency
RNA  ribonucleic acid
RP   reverse phase
rt   room temperature
s    singlet (nmr)
SAR  structure activity relationship
solECE-1 soluble recombinant endothelin converting enzyme-1
SPOS solid phase organic synthesis
SPPS solid phase peptide synthesis
t    triplet (nmr)
t    tertiary
TBDMS t-butyldimethylsilane
Tbfmoc 17-tetrabenzo[a,c,g,i]fluorenlymethoxycarbonyl
'tBu t-butyl
TFA  trifluoroacetic acid
THF  tetrahydrofuran
TIS  triisopropylsilane
TLC  thin layer chromatography
TMG  1,1,3,3-tetramethylguanidine
Trityl triphenylmethyl
UV   ultraviolet
Z    benzloxycarbonyl
### Amino Acid Abbreviations

**SIDE CHAIN FUNCTION**

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Abbreviation</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>alanine</td>
<td>-CH$_3$</td>
<td>Ala A</td>
</tr>
<tr>
<td>asparagine</td>
<td>-CH$_2$CONH$_2$</td>
<td>Asn N</td>
</tr>
<tr>
<td>cysteine</td>
<td>-CH$_2$SH</td>
<td>Cys C</td>
</tr>
<tr>
<td>glutamine</td>
<td>CH$_2$CH$_2$CONH$_2$</td>
<td>Gln Q</td>
</tr>
<tr>
<td>glycine</td>
<td>-H</td>
<td>Gly G</td>
</tr>
<tr>
<td>isoleucine</td>
<td>-CH(CH$_3$)CH$_2$CH$_3$</td>
<td>Ile I</td>
</tr>
<tr>
<td>leucine</td>
<td>-CH$_2$CH(CH$_3$)$_2$</td>
<td>Leu L</td>
</tr>
<tr>
<td>methionine</td>
<td>-CH$_2$CH$_2$SCH$_3$</td>
<td>Met M</td>
</tr>
<tr>
<td>phenylalanine</td>
<td>-CH$_2$Ph</td>
<td>Phe F</td>
</tr>
<tr>
<td>proline*</td>
<td></td>
<td>Pro P</td>
</tr>
<tr>
<td>serine</td>
<td>-CH$_2$OH</td>
<td>Ser S</td>
</tr>
<tr>
<td>threonine</td>
<td>-CH(OH)CH$_3$</td>
<td>Thr T</td>
</tr>
<tr>
<td>tryptophan</td>
<td></td>
<td>Trp W</td>
</tr>
<tr>
<td>tyrosine</td>
<td></td>
<td>Tyr Y</td>
</tr>
<tr>
<td>valine</td>
<td>-CH(CH$_3$)$_2$</td>
<td>Val V</td>
</tr>
<tr>
<td>aspartic acid</td>
<td>-CH$_2$COOH</td>
<td>Asp D</td>
</tr>
<tr>
<td>glutamic acid</td>
<td>-CH$_2$CH$_2$COOH</td>
<td>Glu E</td>
</tr>
<tr>
<td>arginine</td>
<td>-CH$_2$CH$_2$CH$_2$NHC(=NH)NH$_2$</td>
<td>Arg R</td>
</tr>
<tr>
<td>histidine</td>
<td></td>
<td>His H</td>
</tr>
<tr>
<td>lysine</td>
<td>-CH$_2$CH$_2$CH$_2$CH$_2$NH$_2$</td>
<td>Lys K</td>
</tr>
<tr>
<td>α-aminoisobutyric acid*</td>
<td>-C(CH$_3$)$_2$</td>
<td>Aib -</td>
</tr>
<tr>
<td>2-aminoisobutyric acid</td>
<td>-CH$_2$CH$_3$</td>
<td>Abu -</td>
</tr>
<tr>
<td>penicillamine</td>
<td>-C(SH)(CH$_3$)$_2$</td>
<td>Pen -</td>
</tr>
</tbody>
</table>

* Complete structure of Pro is given.
Table of Contents

Abstract .................................................. I
Acknowledgements ....................................... ii
Abbreviations ........................................... iii
Amino acid abbreviations ............................. vi
Table of Contents ......................................... vii

Part I: Chemical Peptide Synthesis

Chapter 1: Introduction to Solid Phase Peptide Synthesis ......................................... 1
  1.1 Overview ........................................ 1
  1.2 Chemical Peptide Synthesis .................. 3
  1.3 Solid Phase Peptide Synthesis ............. 4
    1.3.1 Amide Bond Formation .................. 7
    1.3.2 Orthogonal Protection .................. 10
    1.3.3 Nature of the Solid Support and Linker Group 13
  1.4 References ..................................... 19

Chapter 2: Investigation of the Endothelin Converting Enzyme through its Natural Substrate 22
  2.1 Introduction .................................... 22
    2.1.1 Discovery of the Endothelins .......... 22
    2.1.2 Biosynthesis ............................. 24
    2.1.3 Endothelin Converting Enzyme (ECE) .... 26
    2.1.4 Endothelin Receptors .................. 29
    2.1.5 Pathophysiology ......................... 31
  2.2 Structure Activity Relationship Study of Big ET-1 ......................................... 34
Part II: Heterocyclic Synthesis

Chapter 4: Synthesis of a Potential Farnesyl Transferase Inhibitor

4.1 Introduction

4.1.1 Protein Prenylation

4.1.2 Ras Proteins

4.1.3 Protein Farnesyltransferase

4.1.4 Protein FTase Inhibitors

4.4 Design and Synthesis of a Potential FTase Inhibitor

4.2.1 Design of a Novel Structure

4.2.2 Initial Route to the Target

4.2.3 Modification or the Synthetic Strategy

4.2.4 Synthesis of Precursor B

4.2.5 Alternative Route to 2-Fluoro-4-hydroxybenzoic Acid

4.2.6 Alternative Synthetic Strategy

4.2.7 Solid Phase Synthesis of the Target Structure

4.3 Biological Testing and Conclusions

4.4 References

Chapter 5: Experimental

Part I: Solid Phase Peptide Synthesis

Part II: Heterocyclic Synthesis

Appendices

Courses Attended

Presentations and Awards
Part I: Solid Phase Peptide Synthesis

Chapter 1

Introduction to Chemical Peptide Synthesis

1.1 Overview

Peptides and proteins are highly versatile biological components of enormously varying size, form and function. They are employed in vivo to execute an immense array of biological functions, from mediating biological response, as enzymes and hormones, to cellular construction. In addition, peptides are conversely responsible for a whole range of pathological diseases but can also be of great therapeutic value. Primary metabolites are important probes in the study of biological systems, enabling us to understand at a molecular level the highly complex and critical mechanisms fundamental to life.

These biological species are constructed through the repetitive condensation of α-amino acids to produce polymers of enormously varying length, containing a trans amide backbone. Peptides have a relatively low molecular weight compared with proteins; while the distinction between the two is not absolute, polymers of 50 amino acids or less are generally considered peptides; between 50 and 80 are sometimes referred to polypeptides; while those of greater than 80, proteins.

A more definite criterion for distinction is based on the overall biopolymer structure. Large polypeptides with definite three-dimensional structure, or tertiary structure, are normally classified as proteins. This structure, unique to each protein, is
essential for protein function. It is achieved through the interplay of two other structural properties of the protein: primary structure, the sequence of \( \alpha \)-amino acids in the peptide chain, and secondary structure, local conformation within the chain in blocks of 4-20 residues. All peptides and proteins possess primary structure by definition, and may possess secondary structure; however, only proteins then fold into a distinct tertiary structure, often bringing residues close together that are far apart in the primary structure. The tertiary structure is thermodynamically favoured and is maintained via interatomic forces such as hydrogen bonding, hydrophobic and charge interactions. Some proteins also possess quaternary structure, existing when the final, active form of the 'protein' consists of more than one polypeptide chain interwoven into a larger complex, interacting through non-covalent forces.

Nature has designed a highly sophisticated, fast and efficient method for protein synthesis. DNA is the permanent record of all biological peptide and protein sequences, contained within the cell nucleus; this information is transcribed into a nucleic acid template, messenger RNA, which moves to the ribosome protein in the cytoplasm, which translates information into the required protein species. The majority of biological peptides originate from much larger protein precursors, through enzyme mediated post-translational cleavage and modification of the polymer chain.

Recent developments in recombinant DNA and peptide synthesis have provided powerful tools with which to study many biological processes; however,
they have limitations: recombinant peptide synthesis is restricted to the 20 naturally occurring amino acids; there is minimum control over protein expression and isolation of the desired protein sequence from the fermentation broth can be difficult. Therefore an alternative method would be beneficial.

1.2 Chemical Peptide Synthesis

One alternative approach is chemical peptide synthesis. This technique is not restricted to the 20 naturally occurring amino acids, and is easily adaptable to primary structure variation via specific site alteration, with complete control over the peptide construction. This flexibility allows the incorporation of non-genetically encoded material into the polypeptide, such as artificial amino acids or NMR probe nuclei at a specific site. These modifications provide information on how individual residues contribute to the overall activity and conformation of the peptide or protein, and furnishes an insight into the complex mechanisms associated with these species. It also opens up the possibility of designing totally artificial structures for therapeutic use.

The first reported chemical synthesis of an amide bond was made in 1901 by Emil Fischer. He synthesised the dipeptide glycylglycine from two glycine monomers, see fig. 1.2.

![Figure 1.2: Dipeptide with highlighted amide bond.](image)

The synthetic construction of the amide bond through the condensation of two $\alpha$-amino acids is hampered by two fundamental factors. The zwitterion nature of the amino acids means they will not spontaneously react, but must be activated in some way, most obviously at the carboxyl terminus. In addition, all non-participating functional groups, including reactive side chains (see page vi), must be protected to prevent their participation in the reaction.
Several solution phase strategies were developed in the early part of the 20th century to facilitate chemical peptide synthesis. These methods were laborious and time consuming, with a large number of repetitive steps required for the synthesis of even very small peptides. Each synthetic step was required to go almost to completion in order to ensure the final product could be achieved. Using solution phase synthesis meant that each of the intermediates required purification, hence reducing the yield and making the synthesis of longer peptides impractical. These restrictions limited the application of solution phase methodology, hence progress in the area of synthetic peptide synthesis was slow.

1.3 Solid Phase Peptide Synthesis.

In 1963 Bruce Merrifield offered his solution. His idea was to attach one starting material to a functionalised, otherwise inert, solid support; to add excess of a second starting material in solution; then wash away that excess, leaving the isolated product attached to the solid support. This technique has two important advantages over solution phase strategies:

- The reaction could be forced to completion by the addition of a large excess of the reagent, which could later be easily removed without affecting the product.
- Purification and characterisation of intermediates was not required, as they remained attached to the solid support, hence were isolated by simple filtration.

These advantages made the synthesis faster and the yield higher. The concept is now employed in much oligonucleotide and oligosaccharide synthesis, with much traditional heterocyclic chemistry also being adapted to utilise the methodology. Merrifield’s original protocol for solid phase peptide synthesis (SPPS) involved the stepwise elongation from the C-terminus, see fig. 1.3.
Part I: Introduction to SPPS

His synthesis heralded a new era in synthetic peptide chemistry, and forms the basis of all solid phase synthesis performed today.

In biological systems, peptides are synthesised from the \( N \)- to the \( C \)-terminal, in SPPS, however, the peptides are grown nearly exclusively from the \( C \)- to the \( N \)-terminus. The necessity for this was recognised early, when it was realised that synthesising from \( N \) to \( C \) could result in the loss of stereochemical integrity through oxazolone formation, see fig. 1.4. On activation of the carboxylic acid (see later), it is possible that the growing peptide chain could undergo intramolecular cyclisation, instead of reaction with the next amino acid.

**Figure 1.4:** Intramolecular oxazolone formation when SPPS is performed from \( N \)- to \( C \)-terminal.

The basic principle of solid phase peptide synthesis involves the attachment of the \( C \)-terminal amino acid to a solid support (resin) via the carboxyl function, which effectively acts as a \( C^\alpha \) protecting group; after removal of the \( N^\alpha \)-protecting
group, the coupling reaction is carried out as in solution phase strategies, however, the system allows the use of a large excess of the required pre-activated amino acid species. The excess reagent and soluble impurities are then washed away, leaving the dipeptide product attached to the resin. Any of the original resin-bound amino acid that failed to react is capped using acetic anhydride, to prevent further involvement. The process is then repeated with the third and subsequent amino acids, until the desired sequence is achieved, see fig 1.5.

![Diagram of the general protocol for stepwise solid phase peptide synthesis.](image)

**Figure 1.5:** General protocol for stepwise solid phase peptide synthesis.
Part I: Introduction to SPPS

The highly repetitive nature of stepwise synthetic peptide synthesis makes it amenable to automation. Merrifield was quick to realise this, and published details of his automated peptide synthesiser in 1965. He synthesised the nonapeptide bradykinin, in greater yield and less time than was possible by manual synthesis. Proteins of over 100 residues, impractical by manual techniques, are now routinely synthesised by automated syntheses.

An alternative strategy to stepwise SPPS, which is gaining popularity as methodology advances, is fragment coupling or convergent peptide synthesis. The protein is constructed as a number of smaller fragments, using conventional stepwise methodology, which are then linked together to produce the final target sequence. This technique has been used to great effect for a number of proteins, achieving the synthesis of much longer sequences than is possible via its stepwise counterpart.

As has been described, solid phase peptide synthesis developed rapidly from Merrifield's original pioneering work, and is now well established as a highly efficient, versatile method for peptide and protein synthesis. New methodology continues to advance the technology. Several comprehensive reviews have been published, and so this discussion will be limited to three key considerations.

1.3.1. Amide Bond Formation

As described previously, peptide bond formation is achieved by the activation of the carboxyl group, making it susceptible to nucleophilic attack by the α-amine function of the growing peptide chain. The activated species must be capable of achieving rapid, quantitative amide formation, under mild conditions, without compromising the stereochemical integrity of the growing peptide chain. Although the activating substituent is required to be a good leaving group, if it is too labile it may promote racemisation at the chiral centre.

The most commonly used method of carboxylate activation in other areas of organic chemistry is as the acid chloride, as used by Fischer in his early work. Unfortunately, it is of limited use in peptide chemistry due to its lability and tendency to cause racemisation. Use of the acyl azide, pioneered by Curtius, also dates back to
the beginning of the century.\textsuperscript{13} This method is also highly activating, but with negligible racemisation; however, it is disadvantaged due to a tendency to rearrange to the primary amide and the isocyanate via the Curtius rearrangement, see Fig. 1.6.

\begin{center}
\includegraphics[width=\textwidth]{curtius_rearrangement.png}
\end{center}

\textbf{Figure 1.6: Curtius rearrangement of the acid azide.}

Carboxylate activation via the anhydride has achieved greater success in peptide synthesis. The symmetric anhydride has been constructed using both dicyclohexylcarbodiimide (DCC)\textsuperscript{14} and diisopropylcarbodiimide (DIC),\textsuperscript{15} and offers rapid acylation of amines. DIC is the preferred reagent for SPPS, since the \textit{N-N'\textsuperscript{-}}diisopropylurea by-product is fully soluble. The highly reactive \textit{\alpha\textsuperscript{-}acyl-isourea} intermediate can unfortunately cause racemisation in some amino acids, namely His. Also, where the amino acid possesses an amide side chain function, nitrile formation is possible.\textsuperscript{16} Two equivalents of amino acid are required for formation of the symmetric anhydride; although the method is effective, a major drawback is that it is highly wasteful, using only half of the expensive protected amino acid species in the peptide.

\begin{center}
\includegraphics[width=\textwidth]{anhydride_activation.png}
\end{center}

\textbf{Figure 1.7: Amino acid activation as the symmetrical anhydride using DIC.}
Alternatively, a mixed anhydride species may be used. These also offer rapid acylation of amines, and have the advantage of requiring only one equivalent of the amino acid. As there are two possible sites for nucleophilic attack in such a system, the amine must be encouraged to attack at the correct site by either decreasing the reactivity of the activating portion, or by increasing its steric bulk. Several strategies have been developed to achieve this, but unfortunately such methods tend to be more suited to solution phase peptide synthesis, as they generally require low temperature to prevent disproportionation to the symmetric anhydride.

The most popular method of activation used in SPPS is via the active ester. DCC or DIC is used to form the highly reactive O-acyl substituted urea, which is then reacted with some form of hydroxyl function to form an active ester. This activated species is capable of rapid acylation of amines, but is more stable than the anhydride, hence racemisation is suppressed. The first active esters employed were ortho- and para-nitrophenyl esters, subsequently a range of alcohols have been developed. One species which has shown exceptional properties is N-hydroxybenzotriazole, see fig. 1.8. Originally designed to suppress racemisation, its potential as an effective coupling agent was soon recognised. A number of alternative N-hydroxyl compounds have been designed based on this structure, including ethyl 1-hydroxy-1,2,3-triazole-4-carboxylate, developed by Ramage et al. This compound offers increased coupling efficiency and hence reduced by-product formation, whilst retaining negligible racemisation.

![Diagram of N-Hydroxybenzotriazole (HOBt) and Ethyl 1-hydroxy-1,2,3-triazole-4-carboxylate (HOCt)](image)

**Figure 1.8:** N-Hydroxyl compounds used in the formation of active esters.
A number of different methods are employed for the effective chemical synthesis of peptides and proteins, and ongoing research continues to develop alternative improved strategies for amide bond formation, comprehensive reviews of which are available.\textsuperscript{12}

1.3.2 Orthogonal Protection.

In order to achieve a single unique product from the condensation of two $\alpha$-amino acids, all non-participating functional groups must be protected. This includes all non-participating $\alpha$-amino and $\alpha$-carboxyl groups, plus all reactive side chains. In SPPS, the $\alpha$-carboxyl group is tethered to the solid support which effectively acts as a protecting group.

If a peptide of more than two amino acids is desired, it is necessary to implement a strategy whereby the $N^\alpha$ amino protection can be cleaved selectively, leaving the other protecting groups intact for the entire synthesis; this is termed orthogonal protection. Merrifield incorporated orthogonality into his 1963 tetrapeptide synthesis by using an acid labile $N^\alpha$ protecting group with a base labile resin, see fig. 1.3.

A prerequisite for all protecting strategies is that the protecting group be stable to the reaction conditions to which it is exposed during the assembly phase, but capable of quantitative deprotection when required, under specific, mild conditions. Failure to completely remove the $N^\alpha$ protecting group could lead to the build up of deletion peptides, i.e. sequences with one or more of the amino acid residues missing. The sequences retaining $N^\alpha$ protection are unable to couple with the next amino acid, however, subsequent deprotection steps may liberate the $N^\alpha$ amine allowing the depleted sequence to participate in succeeding chain elongation. Deletions are a particular problem in peptide synthesis, as they result in both a reduced yield and difficult purification.
Part I: Introduction to SPPS

[Diagram of amino acid racemisation with use of N\textsuperscript{a} acyl protection]

**Figure 1.9:** Amino acid racemisation with use of N\textsuperscript{a} acyl protection.

The most obvious choice of N\textsuperscript{a} protection was via an acyl group; however, this form of protection has a tendency to form the intramolecular oxazolone during carboxylate activation, see fig. 1.9. Hence the main class of N\textsuperscript{a} protecting groups employed in peptide synthesis are the urethanes. These do not have this tendency, and offer stable protection under the required conditions. On cleavage of the urethane group, the resulting carbamic acid spontaneously decarboxylates to give the free α-amino terminus, see fig. 1.10.

**Figure 1.10:** Removal of the urethane protecting group.

The first N\textsuperscript{a} urethane protecting group was developed in 1932\textsuperscript{21}. The benzyloxycarbonyl (Z or Cbz) group, used by Merrifield, was cleaved either via strong acid, for example HBr in HOAc, or by hydrogenation.

**Figure 1.11:** Acid labile N\textsuperscript{a} protecting groups.
The demand for milder deprotection conditions, which did not damage the polypeptide product, led to the development of the t-butyloxycarbonyl (Boc) group which could be cleaved under milder acidic conditions. Subsequent developments gave rise to other more acid labile protecting groups. Consideration of the various deprotection requirements of these groups enabled the peptide chemist to design suitable orthogonal protection schemes, based on alternative acid labilities. The most commonly used strategy was that of $N^\alpha$-Boc with side chain Z group protection.

The development of the 9-fluorenylmethoxycarbonyl (Fmoc) protecting group offered the first truly orthogonal protecting strategy. It was developed due to concern that the repeated TFA acidolysis used in Boc chemistry, and the strong acid conditions (HF) required for cleavage from the resin, could have a deleterious effect on the growing peptide. By contrast, Fmoc cleavage was performed under mild basic conditions, using the secondary amine piperidine.

![Diagram of Fmoc cleavage](image)

**Figure 1.12: Cleavage of the Fmoc protecting group.**

The fulvene-piperidine adduct produced during Fmoc cleavage is UV active (isosbestic point at 302nm), offering a method for monitoring the coupling efficiency of each step by measuring the absorbance of the deprotection effluent (see Chapter 2). Under this strategy, the side chains are typically protected with highly acid labile groups, which remain completely intact during repeated base cleavage of the $N^\alpha$.
Part I: Introduction to IjFJFZ5 protection. Taking this in conjunction with more acid labile resins (see later), the requirement for strong acid treatment was eliminated.

Choice of side chain protection will depend on the choice of the transient α-amino protecting group. A multitude of side chain protecting groups have been developed to be compatible with the Fmoc strategy; these have been comprehensively reviewed elsewhere. Researchers continue to develop new, innovative protecting groups to replace those with less than ideal properties. Table 1.1 gives a summary of the most common groups currently employed for use with the N\(^\alpha\) Fmoc protection strategy; the majority of functionalities are efficiently protected using trimethyl phenyl (Trt), t-butyl ethers or t-butyl esters, however, the side chains of Arg, Cys and His require special consideration.

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Side Chain PG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg</td>
<td>Pmc(^{25})/Pbf(^{26})</td>
</tr>
<tr>
<td>Asn/Gln</td>
<td>Trt</td>
</tr>
<tr>
<td>Asp/Glu</td>
<td>(^1)Bu esters</td>
</tr>
<tr>
<td>Cys</td>
<td>Acm(^{27})/Trt/S(^4)Bu(^{28})</td>
</tr>
<tr>
<td>His</td>
<td>(\tau)-Trt(^{29})/(\pi)-Bum(^{30})</td>
</tr>
<tr>
<td>Lys</td>
<td>Boc</td>
</tr>
<tr>
<td>Ser/Thr/Tyr</td>
<td>(^1)Bu ethers</td>
</tr>
<tr>
<td>Trp</td>
<td>Boc</td>
</tr>
</tbody>
</table>

Table 1.1: Summary of the most common side chain protecting groups used with N\(^\alpha\) Fmoc protection.

1.3.3 Nature of the Solid Support and Linker Group.

The basic criteria for a solid support (resin) are that it be insoluble; contain a functionality to which the first reactant can be attached; be completely inert to all reagents/solvents to which it will be exposed during the synthesis; and finally, on completion of the synthesis, give quantitative release of the product under mild conditions which will not degrade the product.

Merrifield performed his synthesis on a co-polymer of styrene and divinylbenzene (DVB), functionalised with chloromethyl groups. Cross-linking gave an extended rigid structure, whilst chloromethylation afforded reactive sites through
which the C-terminal amino acid could be attached via an ester linkage. Cleavage occurred through an alkyl-oxygen fission, see fig. 1.13.

![Cleavage of Peptide from Merrifield resin.](image)

**Figure 1.13: Cleavage of Peptide from Merrifield resin.**

The resin is produced in the form of small spherical beads of typical diameter ~75-200μm. Over 99% of the functional sites are located within the spherical structure, it is therefore vital that the resin swells to allow maximum penetration of the reagents into the structure to reach the active sites. Merrifields' polystyrene based resin offered favourable swelling properties in aprotic, non-polar solvents suitable for much organic and peptide synthesis.

An alternative polymeric support, a polyamide resin, was developed by Sheppard, by copolymerisation of dimethylacrylamide with ethylene bisacrylamide, and functionalisation with acryloylsarcosine methyl ester. The more polar support was designed to be more compatible with the polarity of the growing peptide chain, aiding solvation of the peptide and preventing the fall in yield with increased peptide length witnessed with the polystyrene based resins. These solid supports are now available as co-polymers with polyethylene glycol (PEGA resins). They have demonstrated improved swelling properties in aqueous solvents, however, the polyamide based resins have not superseded their polystyrene based counterparts, which remain the most popular choice of polymeric support.

The equivalent polystyrene/polyethylene glycol (PEG) co-polymer (Tentagel™) was also developed on this same principle, to counteract the hydrophobic nature of the polystyrene. These resins also offered the advantage of separating the growing peptide from the bulky polymeric support, thus reducing steric hindrance. The long flexible linker group provided greater mobility, creating a
more ‘solution-like’ environment. This has proved favourable for many solid phase organic reactions (SPOS). However, Tentagel™ has a comparatively low level of functionalisation, which has limited its application.

The development of side chain protecting groups susceptible to cleavage using TFA (see earlier), encouraged the development of resins cleavable under similarly mild conditions. This was the stimulus for the development of the ‘linker’ system, in which a linker or spacer group was introduced between the polymeric support and the growing peptide, making it possible to control the lability of the peptide-resin link. The vast majority of resins used today are based on Merrifield's DVB cross-linked polystyrene (PS) matrix, but with the chloromethylene group derivatised with alternative linker groups.

At an early stage, Merrifield modified his original resin to counteract the depletion of the growing peptide chain from the solid support observed on repeated treatment with strong acid (TFA). This gave rise to PAM resin, see fig. 1.14, still a popular resin for use with Boc chemistry today.33

![Ester bond formed with peptide is cleaved in strong acid (HF)](image)

**Figure 1.14:** PAM resin, for use with Boc chemistry.

The first real breakthrough in resin technology was made by Wang.34 He designed a linker system attached to the original PS matrix, for use with the Fmoc orthogonal protecting strategy. It has remained one of the most popular and widely used resins to date.

![Ester bond is cleaved in mild acid (TFA), however is stable to mild base (i.e. piperidine).](image)

**Figure 1.15:** p-Benzylxoybenzyl alcohol (Wang) resin.
Although it was synthesised from the Merrifield resin by substitution of the chloromethylene group, it did not require the use of strong acid conditions to afford the final peptide cleavage. This could be achieved using TFA, by virtue of the p-oxygen which allowed resonance stabilisation of the carbocation intermediate.

Figure 1.16: Resonance stabilisation of the Wang resin carbocation.

Resin technology is a major area of research for solid phase synthesis, and an enormous array of linkers attached to the PS polymeric support are now commercially available. Virtually all early work with solid phase technology concentrated on peptide synthesis, hence the majority of resins were designed to immobilize carboxylic acids, Table 1.2 gives five of the resins used for this purpose.

<table>
<thead>
<tr>
<th>Resin Type</th>
<th>Cleavage Conditions</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-Chlorotrityl resin;</td>
<td>10% AcOH</td>
<td>2-Chlorotrityl resin, cleaved with 10% AcOH, enabling the preparation of side chain protected peptides.</td>
</tr>
<tr>
<td>Sieber Amide resin;</td>
<td>1% TFA/DCM</td>
<td>Sieber Amide resin, cleaved with 1% TFA/DCM giving the peptide amide product.</td>
</tr>
<tr>
<td>Rink acid (X=OH);</td>
<td>AcOH/DCM</td>
<td>Rink acid (X=OH), cleaved with AcOH/DCM to give the acid.</td>
</tr>
<tr>
<td>Rink amide (X=NHFmoc);</td>
<td>AcOH/DCM</td>
<td>Rink amide (X=NHFmoc), cleaved with AcOH/DCM to give the amide.</td>
</tr>
<tr>
<td>4-Hydroxymethylbenzoic acid</td>
<td></td>
<td>4-Hydroxymethylbenzoic acid (HMBA) resin; extremely versatile resin capable of producing numerous C-terminal derivatised peptides (see later).</td>
</tr>
</tbody>
</table>

Table 1.2: Five resins designed for immobilizing carboxylic acids.
In addition to being extremely acid labile, the 2-chlorotrityl resin prevents racemisation of the C-terminal amino acid, which can be a problem with other resins. This is especially beneficial for the synthesis of peptides with C-terminal His or Cys. The shear bulk of the resin also suppresses diketopiperazine formation by C-terminal Pro, which occurs readily with the use of benzyl ester linkages more susceptible to nucleophilic attack.

Many biologically active peptides have non-acid C-terminal functionalities. Such modifications are therefore of potential interest in drug applications, in peptidomimetics, and also offer a site for further chemical modification. Many resins have therefore been purposely designed to liberate derivatised carboxyl terminal peptides.

A frequently encountered C-terminal modification in biological proteins is the C-terminal amide, consequently many resins have been developed to achieve this. Sieber and Rink amide resins (shown in table 1.2), are early examples of resins enabling this modification. They incorporate nitrogen into the linker, allowing attachment of the C-terminal acid group through an amide linkage. Subsequent cleavage of the completed product liberates an amide functionality, as opposed to the free carboxylic acid.

The hydroxymethylbenzoic acid or HMBA resin is one of the most versatile peptide resins. The completed peptide can be cleaved under a wide variety of conditions, and changing the nucleophile alters the product (see Table 1.3). Hence this enables a diverse range of C-terminal modifications to be achieved using the same resin.
<table>
<thead>
<tr>
<th>Cleavage Conditions</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>NH$_3$/MeOH</td>
<td>Peptide carboxamide</td>
</tr>
<tr>
<td>NH$_2$NH$_2$/dioxane</td>
<td>Peptide hydrazide</td>
</tr>
<tr>
<td>aq. NaOH</td>
<td>Peptide acid</td>
</tr>
<tr>
<td>MeOH/Et$_3$N</td>
<td>Peptide methyl ester</td>
</tr>
<tr>
<td>NaBH$_4$/EtOH</td>
<td>Peptide alcohol</td>
</tr>
</tbody>
</table>

Table 1.3: *Effect of changing the nucleophile when cleaving the HMBA resin.*

The increased interest in the application of solid phase technology to organic small molecule synthesis has seen the development of linkers capable of immobilizing functional groups other than carboxylic acids, many of these are now commercially available.$^{35(b,c)}$ Attachment of the starting material to the resin is usually achieved through some polar functionality, which is recovered as the same or a slightly modified group. Strong polar groups are often undesirable in drug candidates, hence the development of ‘traceless’ linkers which enable compounds to be synthesised with no obvious linker attachment point.$^{40}$ Other popular resins include the safety-catch$^{41}$ and photolabile types.$^{42}$ Alternatively, another approach sees the product synthesised in solution, with the reagents introduced immobilized on resin.

There are an immense array of resins now commercially available, and researchers continue to develop new linker systems capable of achieving ever more sophisticated chemical manipulations on solid phase. Thus, the chemist may select the one most suited to their specific requirements.

Since Merrifield's pioneering work, solid phase synthesis has developed into an important, diverse area of organic chemistry, including both synthetic peptide and small molecule synthesis. The basic principle has been applied to achieve all manner of on-resin chemical manipulations. In the following two chapters, SPPS is used to synthesise a variety of peptides and peptide libraries (*see later*) in a stepwise manner.
1.4 References


Chapter 2

Investigation of the Endothelin Converting Enzyme through its Natural Substrate.

2.1 Introduction

2.1.1 Discovery of the Endothelins

The critical role of the endothelium layer in the control of vascular tone, through release of vasoactive substances, was first demonstrated through its secretion of a vasodilatory agent termed Endothelin Derived Relaxing Factor, later identified as nitric oxide. The observation that the endothelium could also cause vasoconstriction was made in 1981. Hickey et al demonstrated the presence of an endothelium-derived vasoconstrictor, subsequently identified to be of peptidic nature. It was Yanagisawa et al, in 1988, who first isolated and characterised this vasoconstrictor from cultured porcine aortic endothelial cells, which they named endothelin (ET). It was shown to act on the vascular smooth muscle cells causing a powerful and long lasting vasopressor effect. It has been recognised as the most potent vasoconstrictor known to date, with an activity ten times more potent than angiotensin II.

Its structure was determined to be a twenty one amino acid peptide, containing two intramolecular disulphide bonds across the cysteines in positions 1-15 and 3-11. The N-terminal domain was shown to determine its affinity for the receptor, while the C-terminal domain contains the receptor binding site. The parent cDNA was identified and cloned. Three distinct genes were identified, encoding three closely related peptides. The original endothelin, re-named endothelin-1 (ET-1),
was determined to be one of a family of closely related peptides, with endothelin-2 (ET-2) and endothelin-3 (ET-3) differing from ET-1 by only two and six residues, respectively. The endothelins also show considerable sequence homology with that of the sarafotoxins, a family of vasoactive peptides found in the venom of the snake *Atractaspis engaddenis.*10 Both families of peptides activate the same receptors in the brain and heart.11

**Figure 2.1:** Structural similarities between the endothelin family members.

ET-1 was first isolated from the endothelium, subsequent studies have isolated it in non-vascular tissues such as lung, pancreas, kidney, spleen and brain.12 ET-2 and ET-3 have been detected in other tissues such as the intestine and adrenal gland. ET-3 is relatively abundant in neuronal tissues and is thought to be the neural form of the endothelins.13

In addition to being potent vasoconstrictors, the endothelin peptides have been shown to possess mitogenic activity and may participate in vascular modelling and regulation of cell proliferation.14 They have been detected to different degrees in
Investigation of ECE

a wide variety of both vascular and non-vascular tissues, each invoking slightly different biological responses. Increased levels of endothelin have also been detected in a number of cardiovascular disease states.\textsuperscript{15} It is this that has led to the widespread interest in the role of the endothelins and the potential benefits that inhibiting the action of these peptides may have in the treatment of cardiovascular diseases.

ET-1 was the first of the endothelins to be isolated and identified. It is not only the most abundant, but is also the most potent vasoconstrictor of the three. Hence the majority of research into the endothelins has focused on the ET-1 peptide and unless otherwise stated, the overview of research reported here refers to ET-1.

2.1.2 Biosynthesis

Endothelial cells do not contain dense secretory granules in which to store the ET for later release. Instead, under normal conditions, ET is secreted continuously from endothelial cells through the constitutive secretory pathway and is regulated at the level of peptide synthesis. The parent gene can be switched on by a variety of physiological stimuli including: adrenaline; thrombin; vasopressin, angiotensin II; insulin; cytokines; growth factors and physical stimuli.\textsuperscript{19} ET production can also be reduced by the presence of various chemical factors including nitric oxide and prostacyclin.\textsuperscript{16,17} The most important physiological factor regulating the production and release of ET may be blood flow.\textsuperscript{18}

As mentioned previously, the endothelins are derived from three distinct genes which encode for three precursor proteins, known as preproendothelins (prepro-ET). All are of highly conserved sequences of approximately 200 amino acids. Contained within this initial species, sandwiched between pairs of basic residues which serve as recognition sites for an endopeptidase, is a sequence of 38 or 39 amino acids (species dependant) named big endothelin (big ET). This species is almost devoid of vasoconstriction activity, less than 1% of that of the corresponding ET.\textsuperscript{19} To achieve the fully active mature ET species, a further unique processing event takes place, catalysed by a previously unknown enzyme termed Endothelin Converting Enzyme (ECE).\textsuperscript{7} This species specifically cleaves the Trp\textsuperscript{21}-Val\textsuperscript{22} bond.
in big ET-1 and big ET-2, and the corresponding Trp\textsuperscript{21}-Ile\textsuperscript{22} bond in big ET-3, to give the mature ET and the C-terminal fragment (CTF). Fig. 2.2 illustrates the biosynthetic pathway established for endothelin.

**Figure 2.2:** Biosynthetic pathway of ET-1.

ECE is a zinc metallopeptidase characteristically inhibited by phosphoramidon, see later.\textsuperscript{20} Two ECE isoenzymes have been identified, distinguishable by their localization, optimal pH and their sensitivity to phosphoramidon. They are located both on the surface of vascular smooth muscle
cells and the endothelial cells, where the majority of conversion is suspected to take place as a post secretory event.

After ET is released from the endothelium it is secreted into the basolateral compartment where it acts in a paracrine manner as a local hormone on the surface of the endothelial cells and the vascular smooth muscle. It is not released into the plasma as a circulating hormone, in fact, under normal physiological conditions the plasma concentration is insufficient to achieve vasoconstriction. ET has a very short half life in blood because it bonds quickly with the tissue where an Endothelin Degrader Enzyme (EDE) is present, which metabolizes it rapidly.

2.1.3 Endothelin Converting Enzyme (ECE)

Identification, isolation and cloning of the physiologically relevant ECE took six years from the first isolation of ET-1. The first suggestion that the final step in the biosynthesis of the endothelins was achieved via a unique catalytic event was made in 1988 by Yanagisawa. It was not until 1993, that the species responsible was finally established, despite considerable, wide ranging research.

The search for the relevant ECE embraced all classes of proteases including a chymotrypsin type, aspartyl, serine and thiol proteases. However, all were superseded by the discovery that the conversion was sensitive to phosphoramidon, a potent inhibitor of many zinc metallopeptidases, see fig. 2.3. This halted other lines of investigation, and subsequent work focused on ECE as a phosphoramidon sensitive, zinc metalloprotease. Since phosphoramidon did not readily enter the cell, it was proposed that ECE was an ectoenzyme and that the final stage in the ET processing occurred as a post-secretory event on the surface of the cell.

![Figure 2.3: Structure of phosphoramidon.](image_url)
The zinc metalloprotease, neutral endopeptidase-24.11 (E-24.11 or NEP) has been shown to exhibit similar properties to ECE and has consequently been used as a model for ECE (see later). The two enzymes show a 37% sequence homology, with 10 of the 14 Cys residues conserved between the two and evidence suggesting similar tertiary structures. Both proteases are inhibited by phosphoramidon, however, unlike NEP, ECE is not sensitive to other common metalloprotease inhibitors such as thiorpan or the ACE inhibitor, captopril. NEP cleaves peptide bonds on the amino side of hydrophobic residues, hence it cleaves the Trp-Val bond of big ET. However, it will also degrade ET at other hydrophobic residues, unlike ECE it is not specific to that one site.

Isolation of ECE was hampered firstly because no rich tissue source of ECE was found for use as starting material, and secondly, since ECE was thought to be so specific to the natural big ET substrate, attempts to develop a rapid, sensitive assay to test for the enzyme have been unsuccessful. In 1993 the isolation of ECE from rat lung and porcine aortic endothelium was simultaneously reported by Takahashi et al and Ohnaka et al, who purified the enzyme to homogeneity. The corresponding gene was identified and cloned in 1994.

The enzyme was identified as a membrane bound, zinc metalloprotease with an optimal activity at neutral pH. It was shown to possess a preference for big ET-1 over other isoforms of bigET. ECE mRNA has been located in a wide variety of tissues including lung, pancreas, placenta, adrenal gland, ovary and testis. With the enzyme itself localised in the extracellular face of the endothelial cell membrane in a variety of tissues including the aorta, lung, kidney, liver, heart and some secretory cells. It was not however detected in the thyroid gland or brain. Evidence supports the existence of another non-selective form of ECE located on the surface of smooth muscle cells, enabling big ET released from the endothelial cells to be converted to the mature ET directly on the smooth muscle surface.

In 1995, a second form of ECE was cloned, this was named ECE-2 (with the ECE originally isolated re-named ECE-1). It was found to posses an overall 59% amino acid sequence homology with ECE-1, but was shown to have an acidic pH optimum of 5.5, in contrast with the neutral pH optimum for ECE-1, and presents a
200-fold greater sensitivity to phosphoramidon. Like ECE-1, it is selective for big ET-1 over both big ET-2 and big ET-3. This second ECE is located within the endothelial cells, acting as an intracellular enzyme, cleaving endogenously synthesised bigET-1 to ET-1, probably at the trans-Golgi network where the vesicular fluid is acidified. The significance of the contributions from both intra and extracellular sources to the final ET concentration is unclear, as is the relevant ECE involved in the major biosynthetic route.

Two isoforms of ECE-1 with distinct N-terminal tails have been found to exist, categorised as ECE-1α and ECE-1β. Both are encoded by the same gene by way of alternative splicing of the mRNA transcripts. No significant differences in location, specific activities or substrate specificities have yet been observed between the two isoforms.

Both ECE-1 and ECE-2 are type II integral membrane proteins with a short N-terminal cytoplasmic tail, a transmembrane hydrophobic domain and a large extracellular domain containing the catalytic site and a typical zinc-binding motif HExXXH (His-Glu-X-X-His, where X is a hydrophobic amino acid). Both isoforms show significant sequence homology, particularly in the C-terminal region, to NEP and the human Kell blood group protein. It is also evident from molecular estimations that both enzymes are highly glycosylated, with a predicted ten N-glycosylation sites. It was first suggested in 1995 that ECE-1 existed as a disulphide linked dimer when it was observed that ECE appeared as a 130kDa protein under reducing conditions, but as ~300kDa under non-reducing conditions. Site directed mutagenesis established that ECE-1 forms a dimer through its C412 residue. This dimeric structure is preferential for effective conversion of bigET to ET. The C412 residue is conserved in ECE-2, strongly suggesting that it also exists as a disulphide linked dimer.

Both ECE-1 (ECE-1α and ECE-1β) and ECE-2 have been shown to cleave bigET-1 more efficiently than either bigET-2 or bigET-3, hence it is predicted that there may well be additional members of the ECE family to be discovered which would preferentially convert bigET-2 and/or bigET-3, especially in the brain.
Site directed mutagenesis studies are building up information on the important structural features and residues essential for ECE activity.\textsuperscript{1,36} However, results from different sources are in some cases contradictory and the conclusions presented are based on limited facts. Detailed cell biology of the processing of the endothelin precursors, especially the site of action of ECE in the final stage, remains controversial.

Evidence continues to support the hypothesis that the final processing of bigET-1 to ET-1 occurs as post-secretory processing event, predominantly at the endothelin surface on the extracellular face of the plasma membrane. Not all data support this location. When ECE-1 cDNA is expressed at high levels, intracellular accumulation of ET-1 is observed, particularly in the Golgi complex.\textsuperscript{29} Some groups suggest that the intercellular conversion of bigET-1 to ET-1 (probably by ECE-2) is the major route. Further analysis of the endothelin processing is required to resolve these discrepancies, amongst others, and to evaluate the relevant roles and locations of the different ECE isoforms.

2.1.4 Endothelin Receptors

After release from the endothelial cells, the mature ET peptide rapidly achieves its physiological action through attachment to specific endothelin receptors. Endothelin receptors exist both on the surface of vascular smooth muscle and the endothelium. In 1990, two receptors were isolated and cloned, termed ETA and ETB.\textsuperscript{37} Characterisation showed both to consist of seven transmembrane segments of 20-27 amino acids. Their structures possess a sequential and topographic similarity to the G protein-coupled superfamily.

The ETA receptor exists on the surface of vascular smooth muscle. It shows a differential affinity for the three isoforms of endothelin in the order ET-1>ET-2>>ET-3, the ETB receptor however shows no selectivity.\textsuperscript{38} In 1993, the existence of two ETB receptor subtypes was confirmed, named ETB\textsubscript{1} and ETB\textsubscript{2}.\textsuperscript{39} ETB\textsubscript{1} receptors are located on the surface of the endothelium and ETB\textsubscript{2} on the vascular smooth muscle surface.
The endothelins control vascular tone by both causing vasoconstriction and dilation. This is achieved by stimulation of the different receptors. Stimulation of the ET<sub>A</sub> and ET<sub>B2</sub> receptors on the vascular smooth muscle causes vasoconstriction, while stimulation of ET<sub>B1</sub> on the endothelium causes vasodilation via release of relaxing factors. In 1994, a third type of ET receptor was cloned, termed ET<sub>C</sub>, it has been shown to display greater affinity for ET-3 over ET-1. Located on the endothelium surface, stimulation results in vasodilation by release of NO. Fig. 2.3 gives a summary of the ET interactions.

**Figure 2.4:** Endothelin, other endothelin-derived vasoactive factors and their interactions.

Research employing ET-1, the most potent and widespread of the endothelins, revealed that under physiologically normal conditions, the level of endothelin produced in the endothelium is very low. At this concentration the hormone acts locally on the receptors in the endothelium itself, ET<sub>B1</sub> and ET<sub>C</sub>. These cause the release of other endothelium derived vasoactive factors, nitric oxide synthase and cyclooxygenase 1, which can pass both into the adjacent smooth muscle causing vasodilation and into the lumen inhibiting platelet aggregation. It is now generally accepted that ET-1, under normal physiological conditions, is a vasodilator acting on the endothelium ET<sub>B1</sub> receptors.

Under certain physiological stimuli, ET levels increase and the peptide becomes capable of passing through the endothelial cells into the smooth muscle, so
acting directly on it via stimulation of the ET$_{A}$ and ET$_{B2}$ receptors, causing vasoconstriction through a heightened level of intracellular Ca$^{2+}$. Furthermore, under pathophysiological conditions ET levels can become so high that it can pass into the lumen and act as a circulatory hormone. Increased plasma levels of ET have been reported in a variety of disease states.

In addition to its effect on vascular tone, stimulation of the different endothelin receptors in different tissues causes a variety of responses. The documented mitogenic activity of ET-1 promotes proliferation of vascular endothelial cells, vascular smooth muscle and fibroblasts, mediating cardiac growth and hypertrophy. Recently it has been shown, through targeted gene disruption of mice, that the endothelin peptides may also have a morphogenic role in a variety of developmental events. Disruption of the ET genes caused severe craniofacial abnormalities. Undoubtedly, further biological processes in which the endothelins play an essential role will be revealed.

2.1.5 Pathophysiology

Increased ET levels have been indicated in a wide variety of disease states.

<table>
<thead>
<tr>
<th>Cardiovascular diseases:</th>
<th>Renal disease:</th>
</tr>
</thead>
<tbody>
<tr>
<td>myocardial ischemia</td>
<td>acute/chronic renal failure</td>
</tr>
<tr>
<td>congestive heart failure</td>
<td></td>
</tr>
<tr>
<td>arrhythmia</td>
<td></td>
</tr>
<tr>
<td>unstable angina</td>
<td></td>
</tr>
<tr>
<td>hypertension</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bronchoconstriction:</th>
<th>Vascular disorders:</th>
</tr>
</thead>
<tbody>
<tr>
<td>pulmonary hypertension</td>
<td>atherosclerosis</td>
</tr>
<tr>
<td>asthma</td>
<td>other vascular disorders,</td>
</tr>
<tr>
<td></td>
<td>e.g. Buergers disease, Takayasu’s arteries</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Neuronal action:</th>
<th>Other:</th>
</tr>
</thead>
<tbody>
<tr>
<td>cerebral vasospasm</td>
<td>endotoxic shock, septicemia</td>
</tr>
<tr>
<td>subarachnoid hemorrhage</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Endocrine:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>pre-eclampsia</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: Disease states in which the endothelins have been implicated.
Investigation of ECE

Their role in a number of cardiovascular, renal and central nervous system diseases is well documented, including hypertension and heart failure, myocardial infarction, acute renal failure, neointimal thickening, asthma, global and focal cerebral ischemia and vasospasm, including subarachnoid haemorrhage and Raynauld’s disease. In addition to causing sustained vasoconstriction, ET stimulates the generation of a cocktail of neurotransmitters and hormones including: renin; angiotensin II; aldosterone and adrenaline.

The majority of ET synthesised in the endothelium is released towards the underlying smooth muscle hence does not necessarily result in increased plasma levels, however elevated circulatory ET-1 levels have been reported in a variety of pathogenic conditions. In pulmonary hypertension and hypertension developed during pregnancy, heightened plasma levels have been recorded due to ET-1 over production. Increased levels are seen during acute myocardial infarction and unstable angina, where ET receptor antagonists have been reported to reduce myocardial infarction size. ET-1 levels are also significantly increased in the cerebrospinal fluid of patients after subarachnoid haemorrhage causing cerebral vasoconstriction. Other disease states in which endothelin may play a role continue to be identified.

The increasing experimental and clinical evidence that overproduction of the endothelins (especially ET-1) may play a significant pathogenic role in such a wide variety of disease states, has attracted significant attention to the possibility that inhibition of the endothelin activity could be a useful therapeutic tool for use in cardiovascular medicine, particularly in conditions in which these peptides are clearly implicated. Research has focused on inhibiting the normal endothelin processing pathway at two points, either at the ET receptor using an antagonist, or by inhibition of ECE-1, see fig. 2.5.
Due to the relatively early identification and cloning of the endothelin receptors, the development of potential endothelin antagonists has been widely pursued. Researchers have embraced all manner of chemical candidates, this work has been comprehensively reviewed elsewhere.23,51 One of the most promising to have emerged (currently in clinical trials) is RO47-0203, known as bosentan, see fig 2.7.52

Studies have shown that systemic administration of bosentan to patients with severe chronic heart failure produced systemic, pulmonary and peripheral venous vasodilation and improved cardiac performance. Several potential antagonist candidates have now reached clinical trials, with early results suggesting a beneficial therapeutic use combined with ACE inhibitors in the treatment of chronic heart failure.
ECE-1 proved more elusive and was not isolated until 1993. This restricted the development of ECE-1 inhibitors, hence research in this area lags behind that of the endothelin antagonists. ECE-1 offers an ideal candidate for blocking the endothelin pathway since it appears highly specific to its natural substrate big ET, cleavage of which is essential for full ET activity.\textsuperscript{53} Despite problems obtaining the pure enzyme, much research has been carried out with the view to developing viable ECE-1 inhibitors, again this has been comprehensively reviewed elsewhere.\textsuperscript{54}

### 2.2 Structure Activity Relationship Study of Big ET-1.

#### 2.2.1 Background

Although much progress in our appreciation of ECE-1 has been made since its isolation, there is still a great deficit in our fundamental knowledge of the enzyme. Significant understanding of the enzyme active site and critical residue interactions can be gained by Structure Activity Relationship (SAR) studies of the natural substrate. Such studies are widely used to probe the conformational and biochemical requirements of ligand/receptor interaction. The major substrate of ECE-1 is big ET-1 (BET-1), a 38 amino acid peptide containing the major structural features: two disulphide bridges between residues 1-15 and 3-11; α-helix: residues 8-12; β-sheet: residues 27-30 and 32-35; β-turns: residues 4-7 and 22-26.\textsuperscript{55}

![Figure 2.7: Big endothelin-1 (bigET-1).](image-url)
As discussed earlier, BET-1 is cleaved by ECE-1 exclusively at the Trp$^{21}$-Val$^{22}$ bond to give the mature ET-1 peptide and CTF.$^{56}$

Several groups have performed limited investigations into the effect of truncating the BET-1 substrate on enzyme recognition. In 1991, it was reported that BET-1[17-26] was not recognised by ECE-1, suggesting that residues within the substrate, distant from the site of cleavage, may be essential for enzyme recognition.$^{57}$ Okada et al. established, through specific site mutations, that residues 27-34 within the C-terminal region were important for enzyme recognition, though in what capacity was not understood.$^{58}$ They later proposed BET-1[19-34] as the minimum sequence to be recognised by ECE-1. The N-terminal disulphide loop appeared to hinder conversion to the mature species, as BET[16-37] showed a 2 fold faster rate of hydrolysis than the natural substrate, while BET[18-34, Phe$^{22}$] showed a 12 fold increase.$^{59}$

A limited study by Fabbrini et al. demonstrated that substitution of residues bordering the scissile bond was tolerated, with the mutants being hydrolysed by ECE-1, indicating that the conformation rather than sequence of the substrate was important for enzyme recognition.$^{60}$ Tentative suggestions were made that, while the P1 residue (Trp) was critical, the P1' residue could tolerate substitution, but only by a hydrophobic amino acid.$^{61}$

The fact that ECE-1 recognised a relatively long carboxyl-terminal sequence in addition to the processing site made the enzyme unique. Heightened interest in the big ET substrate was triggered when Morita et al. reported that the enzyme did not cleave truncated analogues with D-amino acids at positions 20, 21, 22 & 23, and that one of the analogues, BET[16-38, D-Val 22], showed potent competitive inhibitory activity, comparable to that of phosphoramidon (Ki 2.5x10$^{-5}$M).$^{62}$ These studies were very limited, to determine definitively the requirements for enzyme recognition, further research exploring sequence changes and truncation of the substrate would be necessary.
2.2.2 Preliminary SAR Investigation of ECE-1.

A systematic SAR investigation of big ET-1 analogues was undertaken to elucidate the minimum peptide sequence required for enzyme recognition and the residues critical to binding. The inhibitory nature of these species provided a handle with which to compare and evaluate each analogue. It also offered the possibility of designing a potential ECE inhibitor drug candidate.

A suitable peptidic ECE inhibitor candidate would possess a minimal IC$_{50}$ value, offset against the shortest chain length for enzyme recognition. Longer peptides are more prone to in vivo enzymatic cleavage, the shorter the sequence the more easily it can be protected against hydrolysis, making it potentially orally active and more favourable pharmacokinetic properties. In addition, shorter sequences offer better templates on which to base small molecules, which are preferred drug candidates. An inhibitor must be able to attach to the active site of the enzyme, preventing the natural substrate from binding, but must not be cleaved by the enzyme in any way.

The project was initially taken up by Jiang, whose investigation overlapped with the subsequent findings made here. A preliminary discussion of this study should therefore be made.
<table>
<thead>
<tr>
<th>PEPTIDE</th>
<th>% INHIBITION</th>
<th>IC$_{50}$(µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Big ET-1(16-38)</td>
<td>92.1@1µM</td>
<td>0.038</td>
</tr>
<tr>
<td>2 Big ET-1(19-38)</td>
<td>39@200µM</td>
<td>12.5</td>
</tr>
<tr>
<td>3 Big ET-1(19-34)</td>
<td></td>
<td>2.5</td>
</tr>
<tr>
<td>4 Big ET-1(19-34, Leu$^{34}$)</td>
<td></td>
<td>7.0</td>
</tr>
<tr>
<td>5 Big ET-1(19-34, Leu$^{32,34}$)</td>
<td>70.1@200µM</td>
<td></td>
</tr>
<tr>
<td>6 Big ET-1(27-34, Leu$^{34}$)</td>
<td>0@200µM</td>
<td></td>
</tr>
<tr>
<td>7 Big ET-1(19-38, Pen(Acm)$^{22}$)</td>
<td>12.1@200µM</td>
<td>1.0</td>
</tr>
<tr>
<td>8 Big ET-1(19-38, Pen$^{22}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9 Big ET-1(19-34, Cys(Acm)$^{22}$)</td>
<td>46.3@100µM</td>
<td>17.3</td>
</tr>
<tr>
<td>10 Big ET-1(19-34, Cys$^{22}$)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Big ET-1(19-34, Pen(Acm)$^{22}$)</td>
<td>43.4@100µM</td>
<td></td>
</tr>
<tr>
<td>12 Big ET-1(19-34, Pen$^{22}$)</td>
<td>60.5@100µM</td>
<td></td>
</tr>
<tr>
<td>13 Big ET-1(16-34, Cys(Acm)$^{22}$)</td>
<td></td>
<td>0.33</td>
</tr>
<tr>
<td>14 Big ET-1(16-34, Cys$^{22}$)</td>
<td></td>
<td>0.040</td>
</tr>
<tr>
<td>15 Big ET-1(16-34, Pen(Acm)$^{22}$)</td>
<td></td>
<td>5.5</td>
</tr>
<tr>
<td>16 Big ET-1(16-34, Pen$^{22}$)</td>
<td></td>
<td>10.2</td>
</tr>
<tr>
<td>17 Big ET-1(Cys(Acm)$^{1,3,11,15}$, Trpφ(CH$_2$NH)Val)</td>
<td></td>
<td>13.5</td>
</tr>
<tr>
<td>18 Big ET-1(Cys$^{1,3,11,15}$, Trpφ(CH$_2$NH)Val)</td>
<td>5.3@10µM</td>
<td>32.2</td>
</tr>
<tr>
<td>19 Big ET-1(12-38, Cys(Acm)$^{15}$, Trpφ(CH$_2$NH)Val)</td>
<td></td>
<td>13.2</td>
</tr>
<tr>
<td>20 Big ET-1(19-38, Trpφ(CH$_2$NH)Val)</td>
<td>0@100µM</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.2: SAR Results from study by Jiang.

From the % inhibition and IC$_{50}$ results it could be confirmed that truncation from both the N and C termini was possible whilst retaining enzyme recognition. The BET-1(16-38) truncate demonstrated a highly favourable IC$_{50}$, suggesting it was efficiently recognised by ECE-1. Both BET-1(19-38) and BET-1(19-34) were recognised by the enzyme, in concordance with the previously reported results, however the truncation was accompanied by a large increase in the IC$_{50}$. Truncation of the C-terminal from 38 to 34 did not have such a dramatic effect on the IC$_{50}$ as cutting the N-terminal from 16 to 19, suggesting that residues 16 to 18 may be more important for enzyme recognition.
Peptides 4 and 5, with Leu substituted for Gly at residues 32 and 34, showed a greater IC\textsubscript{50} value than BET-1(19-38) suggesting that residues 35 to 38 were not essential for enzyme recognition. BET-1(27-34, Leu\textsuperscript{34}) showed no inhibitory activity, whether it was in fact recognised by the enzyme was not established.

Previous studies had shown that some substitutions of residues 21 and 22, flanking the natural cleavage site of BET-1, were tolerated. Sulphur containing amino acids are frequently found to give interesting results in enzyme inhibitory studies, hence a number of analogues containing both protected and unprotected Cys and Pen at residue 22 were synthesised. All SH protected species showed decreased inhibitory activity compared with the unprotected analogues, indicating that the free SH was important. Comparison of the free Pen with the corresponding free Cys analogues, showed that both the size of the residue and the free SH group were significant. The BET-1(16-34, Cys\textsuperscript{22}) analogue showed the greatest inhibitory activity, and presented the most likely potential ECE-1 inhibitor candidate.

Jiang also made a series of reduced isostere truncates to alter the conformation of the peptide at the scissile bond, with the natural amide bond mimicked by the reduced peptide bond ϕ(CH\textsubscript{2}NH). The large increase in IC\textsubscript{50} suggested the peptides were no longer efficiently recognised by the enzyme.

The main points of interest from this study were that BET-1 could be truncated to BET-1(19-34) whilst retaining enzyme recognition, in concordance with the results from Okada et al, however residues 16 to 18 appear to play an important role. BET-1(16-38) displayed a favourable IC\textsubscript{50} value. Substitution of residues 32 and 34 were tolerated, hence further C-terminal truncation maybe possible without jeopardising recognition. Substitution of residue 22 with sulphur containing amino acids revealed a potential ECE-1 inhibitor candidate, BET-1(16-34, Cys\textsuperscript{22}).
2.2.3 SPPS of Truncated Big ET-1 Analогues.

All peptides in this chapter were made singly using automated SPPS on an Applied Biosystems 430A peptide synthesiser, employing the $N^\alpha$-Fmoc orthogonal protecting strategy (see chapter 1). Where necessary, the amino acid side-chains were protected using standard acid labile protecting groups.

All peptides, except BET-1(16-27) and BET-1(16-27, Cys$^{22}$), employed Wang resin as the polymeric support. The C-terminal amino acid was manually coupled to the functional sites of the resin via the symmetrical anhydride. The level of loading was quantitatively checked by deprotection of the $N^\alpha$-Fmoc, using piperidine and the UV absorbance of the deprotection solution measured at 302nm (corresponding to the chromophoric piperidine-fulvene adduct). The two remaining peptides were synthesised on 2-chlorotrityl chloride resin as they possessed C-terminal His which is prone to racemisation on the Wang resin.

The degree of amino acid loading onto both resins was determined by the reaction time. Low levels of loading are often desirable when synthesising large peptides as steric interactions between growing peptide chains can be detrimental to the synthesis; however, all the peptides contained here were relatively short, hence amino acid loadings of 65-80% of the total functionalised sites on the resin (achieved through 2.5-3 hours sonication) were used.

Each synthesis was performed on a 0.25mmol scale, via a stepwise strategy from the C to N terminal, using the standard protocol detailed in chapter 5. Activation of the protected amino acid was achieved using DIC and HOCt, a novel coupling agent designed and synthesised by Ramage et al. (see chapter 1), to form the actived ester. This was then introduced to the deprotected N-terminus of the resin bound peptide and coupling allowed to proceed. It is essential that near quantitative coupling be achieved for each residue in order to ensure the synthesis goes to completion. Previously, using conventional DIC and DIC/HOBT coupling methods, this was achieved by coupling each amino acid twice, known as double coupling. However, the use of HOCt made this unnecessary, as a single coupling step for each
residue was normally sufficient to achieve the required > 90% coupling efficiency. This has made the peptide synthesis both faster and more economical.

The $N^\alpha$-Fmoc was cleaved from the resin-bound peptide between residues, in preparation for the next amino acid, using 20% piperidine in 1:1 DMF/dioxane. The chromophoric nature of the piperidine-fulvene adduct provides a non-invasive means of monitoring the coupling efficiency of each residue, hence estimate the potential yield of the completed peptide. An aliquot of the deprotection solution is automatically transferred from the synthesiser to an on line UV detector, set at 302nm, connected to an integrator. This allows the size of the deprotection peak to be compared with that of the previous amino acid, enabling the relative coupling efficiency of each step to be monitored. A deprotection profile can be constructed for the peptide synthesis, *fig. 2.10* gives the profile for the BET-1(16-38, Leu$^{22}$) peptide, typical of all the peptides constructed herein.

![Deprotection profile for big ET-1(16-38, Leu$^{22}$).](image)

**Figure 2.10:** Deprotection profile for big ET-1(16-38, Leu$^{22}$).

After completion of the synthesis, the $N^\alpha$-Fmoc of the final $N$-terminal amino acid was left intact allowing the quantitative calculation of that amino acid, hence the completed peptide, attached to the resin.

The Fmoc group was cleaved manually using piperidine, prior to acidolytic cleavage of the peptide from the resin using 95% TFA with a cocktail of scavengers. This treatment cleaved both the peptide from the resin and removed all side-chain
investigation of ECE protecting groups, giving the fully functionalised, free peptide. The scavengers are carbocation sequestering agents, their presence in the cleavage mixture prevents electrophilic attack of the amino acid side chains by the protecting groups liberated as carbocations, a consequence of the acidolytic cleavage. The scavengers selected for the cleavage depend on the amino acids (and protecting groups) in the sequence, a general guide is given in Table 2.3.

<table>
<thead>
<tr>
<th>Choice of Scavenger</th>
<th>Amino Acid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>All residues</td>
</tr>
<tr>
<td>Ethanedithiol (EDT)</td>
<td>All 'Butyl and Trityl protected amino acids, esp. Trp, Met and Cys(Trt)</td>
</tr>
<tr>
<td>Thioanisole</td>
<td>Met, Arg(Pmc)</td>
</tr>
<tr>
<td>Triisopropylsilane (TIS)</td>
<td>Substitute for EDT, particularly for Arg(Pmc) and Trp(Boc)</td>
</tr>
<tr>
<td>Phenol</td>
<td>Tyr, Trp</td>
</tr>
</tbody>
</table>

Table 2.3: Use of Scavengers in Acidolytic Cleavage.

The optimum cleavage time was established for each peptide as 2-2.5 hours by trial cleavages with small quantities of resin. The crude product was isolated from the TFA and scavengers by trituration with diethyl ether, then re-dissolved in a suitable solvent system to a concentration of ~5mg/ml. The majority of peptides completely dissolved in 1:2 acetonitrile/water.

Purification was achieved via preparative RP HPLC. The analytical RP HPLC profile of the crude material provided guidelines by which to set the preparative conditions. The exact gradient profile was individually tailored to each peptide to afford maximum separation, the general gradient profile is detailed in chapter 5.

Analytical RP HPLC confirmed the purified material to be single peak, and its composition was analysed via MALDI ToF mass spectroscopy and amino acid analysis (see Appendices I and II). Each peptide was then sent to Parke Davis Pharmaceuticals, Ann Arbor, Michigan, where its IC₅₀ value was measured against cloned soluble ECE-1 (solECE-1). This value offered a simple method for comparison and evaluation of each analogue.
2.2.4 SAR Study of Truncated BigET-1 Analogues.

<table>
<thead>
<tr>
<th>PEPTIDE</th>
<th>%INHIBITION</th>
<th>IC$_{50}$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  Big ET-1(16-38, D-Val$_{22}$)</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>2  Big ET-1(16-34)</td>
<td>0.085</td>
<td></td>
</tr>
<tr>
<td>3  Big ET-1(17-34)</td>
<td>0.57(±0.057)</td>
<td></td>
</tr>
<tr>
<td>4  Big ET-1(18-34)</td>
<td>37%@50µM</td>
<td></td>
</tr>
<tr>
<td>5  Big ET-1(16-32)</td>
<td>0.14</td>
<td></td>
</tr>
<tr>
<td>6  Big ET-1(16-27)</td>
<td>0.37</td>
<td></td>
</tr>
<tr>
<td>7  Big ET-1(16-38, Cys$_{22}$)</td>
<td>0.059</td>
<td></td>
</tr>
<tr>
<td>8  Big ET-1(16-32, Cys$_{22}$)</td>
<td>0.15</td>
<td></td>
</tr>
<tr>
<td>9  Big ET-1(16-32, Cys$_{21}$)</td>
<td>&gt;1</td>
<td></td>
</tr>
<tr>
<td>10 Big ET-1(16-32, Cys$_{21,22}$)</td>
<td>&gt;1</td>
<td></td>
</tr>
<tr>
<td>11 Big ET-1(16-27, Cys$_{22}$)</td>
<td>1.07</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.4: Inhibition results for the SAR study of Big ET-1 analogues.

Initially BET-1(16-38, D-Val$_{22}$) was synthesised for comparison with that same peptide made by Morita et al. The IC$_{50}$ was determined to be 2µM, this was in concordance with that previously reported.

In continuation of the study made by Jiang, the effect of further truncation of the BET-1(16-38) analogue was more closely studied. Truncation of the N-terminal to residue 34 caused a two fold increase in the IC$_{50}$; however this figure is comparable with that of BET-1(16-38) when the associated error is taken into consideration. Subsequent reduction from the C-terminal to residues 32 and 27 caused a greater increase in the IC$_{50}$, suggesting enzyme recognition may have been more significantly weakened by the removal of C-terminal amino acids after residue 34.

Truncation of the N-terminal, peptides 3 and 4, saw a more dramatic increase in the IC$_{50}$ suggesting that residue 16, His, was important for ECE-1 recognition. A potential ECE-1 inhibitor based on the BET-1 substrate should therefore contain this residue to enable efficient enzyme recognition, whilst truncation of the C-terminal to residue 27 may be tolerated.
Jiang’s investigation of the effect of sulphur containing amino acids was expanded by substitution of the favoured unprotected Cys, see earlier, for the residues flanking the natural cleavage site in a number of truncated analogues. Substitution of Cys into position 22 (P1’) of BET-1(16-32) and BET-1(16-38) had virtually no effect on the IC₅₀, while substitution into the (16-27) analogue caused a significant increase. This was in contrast with the previous investigation of BET-1(16-34, Cys²²), which had been shown to posses two fold greater activity (IC₅₀=0.040) than the corresponding BET-1(16-34). Substitution of Trp for Cys at position 21 (P1) was not tolerated (peptides 9 and 10).

From this investigation it can be said that although BET-1 can be truncated from the N-terminal to residue 19 and still retain some enzyme recognition, residue 16 is essential for strong binding of the peptide to ECE-1. Truncation of the C-terminal to residue 27 is also tolerated. The dramatic fall off in activity associated with the N-terminal truncation was not as obvious; however, as the residues 38 to 27 were systematically removed, the IC₅₀ was seen to increase in an exponential-like manner, allowing for the conclusion that residues 28 to 34 are also significant to enzyme recognition. From this study it could be extrapolated that BET-1(19-27) would be recognised by the enzyme, however, the Km value for BET-1(16-27) was determined (Parke Davis, Michigan) to be significantly greater than that of BET-1(16-38) (see later), indicating that it bound more weakly to ECE-1.

A potential peptidic enzyme inhibitor candidate should posses the optimal balance between IC₅₀ value, strength of binding and minimum chain length. Of the analogues investigated, the three most promising candidates were:

<table>
<thead>
<tr>
<th>IC₅₀ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big ET-1(16-38)</td>
</tr>
<tr>
<td>Big ET-1(16-34)</td>
</tr>
<tr>
<td>Big ET-1(16-34, Cys²²)</td>
</tr>
</tbody>
</table>

All showed strong enzyme binding characteristics and worthy of further investigation.
2.2.5 Substrate Testing of Potential ECE-1 Inhibitor Leads.

The truncated analogues BET-1[16-34] and BET-1[16-38] were shown to bind to ECE-1, with IC₅₀ values comparable with ECE-1 inhibitors reported by other groups. Both were considered worthy of more extensive subsite investigation, by systematically substituting key residues with alternative amino acids to optimise inhibitory activity, using the powerful technique of combinatorial chemistry (see chapter 3).

Combinatorial chemistry is the term given to the principle whereby, instead of synthesising single compounds, multiple (in some cases many millions of) compounds are generated simultaneously. This technique has gained popularity, particularly in the pharmaceutical industry where recent advances in high throughput biological screening has placed a great demand on the supply of novel compounds for testing.

The BET-1[16-34], being four residues shorter, was the preferred candidate for combinatorial exploration, and it had already been demonstrated that substitution of Val²² by Cys halved the IC₅₀ value. However, prior to launching further investigations, it was necessary to establish whether or not these potential lead candidates acted as substrates for ECE.

The theoretical post cleavage C-terminal peptide fragments of the BET-1[16-34] and BET-1[16-34, Cys²²] analogues (BET-1[22-34] & BET-1[22-34, Cys²²]) were synthesised and sent to Parke Davis, where they were used as references when analysing the products from the peptide interaction with ECE-1 by HPLC.

Both analogues showed cleavage by ECE-1. BET-1[16-34] was rapidly hydrolysed exclusively at the Trp²¹-Val²² bond. Substitution of the Val²² by Cys slowed the rate of cleavage of the 21-22 bond, and mass spectrometric analysis revealed the formation of several products. In addition to the predicted [16-21] and [22-34] products, fragments [20-34], [21-34], [20-28] and [21-28] were detected due to the cleavage of the Ile¹⁹-Ile²⁰, Ile²⁰-Trp²¹ and Val²⁸-Val²⁹ bonds. The fragments
Investigation of ECE

got on to form di-sulphide bridged species; [22-34] formed a di-sulphide linkage with the free substrate, while either [20-34] bridged to [21-28] or [21-28] bridged to [20-34] (both possessing the same mass).

It appeared that, contrary to what had been previously thought in earlier research of ECE, the enzyme was capable of cleaving several bonds in the BET-1 peptide, however only at a much slower rate than that of the Trp\textsuperscript{21}-Val\textsuperscript{22}. In the BET-1[16-34] analogue, the Trp\textsuperscript{21}-Val\textsuperscript{22} hydrolysis occurred rapidly, allowing insufficient time for the relatively slow cleavages at the other sites to occur at a detectable rate. Once the substrate was cleaved at the Trp\textsuperscript{21}-Val\textsuperscript{22}, ECE-1 was incapable of cleaving the resulting smaller peptides. However, because the rate of hydrolysis of the Trp\textsuperscript{21}-Cys\textsuperscript{22} bond was much slower, this allowed cleavage at the alternative sites to occur at a comparable rate before the substrate was hydrolysed into shorter peptides which behave as poor ECE-1 substrates.

2.2.6 Conclusions

The truncated BET-1 analogues, being alternative substrates for the enzyme, were unfortunately unsuitable as potential ECE-1 inhibitors, bringing this line of investigation to a close. The SAR study did however furnish us with considerable information into the enzyme/substrate interaction; residue 16 (His) was critical for enzyme recognition, the minimum sequence which demonstrated strong binding was BET-1[16-34], limited substitution at P1' position was also tolerated. Hence these findings concur with previous reports that the 27-34 sequence C-terminal sequence is important for ECE-1 recognition.

The revelation that the truncated big ET-1 analogues were cleaved both at the 21-22 bond and alternative sites along their length, is in contrast to the previously held belief that ECE-1 hydrolyses the Trp\textsuperscript{21}-Val\textsuperscript{22} bond in its natural BET-1 substrate exclusively. This observation opened up an alternative line of investigation.
2.3 Subsite Specificity Study of ECE-1.

2.3.1 Introduction to Kinetic Study of Substituted BigET-1(16-38) Analogues.

A recent substrate specificity study by Parke Davis, Ann Arbor, revealed that ECE-1 was capable of metabolising a number of biologically active peptides differing in both size and structure, including: neurotensin; substance P; angiotensin I and oxidised insulin B chain, as efficiently as its known in vivo substrate big ET-1.\textsuperscript{64} This study was further to that made by Turner \textit{et al.},\textsuperscript{65} who observed that ECE-1 was capable of hydrolysing bradykinin. This was in contrast to previous assertions that ECE-1 cleaved the big ET peptides exclusively.

The study revealed ECE-1 to have a substrate specificity similar to that of neprilysin (NEP or neutral endopeptidase 24.11), a mammalian cell-surface zinc metallopeptidase of the same enzyme family as ECE-1, preferring to cleave the substrate at the amino side of hydrophobic residues. NEP had been used previously as a model for ECE-1 due to their chemical and physical similarities, both are phosphoramidon-sensitive zinc metalloproteases with significant amino acid sequence and structural homology. However a number of differences in their substrate specificity and sensitivity to inhibitors had been observed; NEP was known to metabolise a broad range of substrates, while ECE-1 was thought to be specific to the endothelin precursors.

The results of the study by Parke Davis opened up the possibility that ECE-1 is potentially involved in the metabolism of a broad range of biologically active peptides distinct from the endothelins, both at the cell surface and in the secretory pathway. The revelation that all the truncated big ET-1 analogues from section 2.2.4 were also alternative substrates for ECE-1, presented a new area for investigation of these species. As discussed previously, the analogues were cleaved by ECE-1 at varying rates (i.e. with different $k_{cat}$ values). It was also observed that if the cleavage at the 21-22 site was hindered, as in big ET-1(16-34, Cys\textsuperscript{22}), ECE-1 was able to cleave the peptide at a number of other sites along the backbone at a compatible rate.
When the 21-22 bond was rapidly hydrolysed, further cleavage of the resulting peptide fragments was not seen, presumably the shorter peptide fragments were poor substrates for ECE-1. The substrate specificity study by Parke Davis had observed that ECE-1 had been incapable of metabolizing substrates smaller than hexapeptides, also that the enzyme preferred to cleave amide bonds with a hydrophobic P1' residue but that a variety of residues were tolerated at the P1 position.

A kinetic study of the P1 and P1' positions, flanking the cleavage site of the big ET-1(16-38) truncate, with systematic sub-site variation was undertaken to elucidate the sub-site specificity of ECE-1. Such information is useful in the design of both enzyme substrates and inhibitors.

2.3.2 *Michaelis-Menten Enzyme Kinetics.*

The relationship between the rate of catalysis and substrate concentration takes the form of a hyperbola for the majority of enzymes, as shown in Fig 2.11.

![Figure 2.11: Relationship between rate of enzyme catalysis versus substrate concentration.](image)

At low substrate concentrations, for a fixed concentration of enzyme (E), the rate is first order (linearly proportional), but at high substrate (S) concentrations the rate becomes independent of that concentration and is therefore zero order.
In 1913 Leonor Michaelis and Maud Menten proposed a simple model to account for these characteristics, based on the assumption that the enzyme forms a specific intermediate complex with the substrate prior to product formation. This model is illustrated by the scheme:

\[
\begin{align*}
&E + S \xrightleftharpoons[k_{-1}]{k_1} ES & \text{FAST STEP} \quad (1) \\
&ES \xrightarrow{k_2} \text{Products} + E & \text{SLOW STEP} \quad (2)
\end{align*}
\]

where \( E \) = enzyme  \\
\( S \) = substrate  \\
\( ES \) = enzyme-substrate complex

At high concentrations of substrate all the enzyme active sites are occupied, the enzyme will exist solely as the enzyme-substrate complex and the reaction velocity will reach maximum. Then, the rate of decomposition of the complex to the product(s) will determine the overall speed of the process, i.e. is the rate determining step. Hence,

\[
\text{Rate}_{\text{forward}} (v_f) = k_2[ES] \quad (3)
\]

Solving this rate equation gives the relationship:

\[
\frac{v_f}{V_{\text{max}}} = \frac{[S]}{K_m + [S]} \quad \text{MICHAELIS-MENTEN Equation}
\]

- \( V_{\text{max}} \) = Maximum velocity, achieved when enzyme is fully saturated with substrate.  
- \( [S] \) = Concentration of substrate  
- \( K_m \) = Michaelis Constant

where \( K_m = \frac{k_{-1}}{k_1} \), based on an equilibrium approach  

or, \( K_m = \frac{k_1 + k_2}{k_1} \), based on a steady-state approach  

The steady state approach is often considered the more realistic as the equilibrium method assumes a true equilibrium to be achieved in the first step. The \( K_m \) value is routinely determined when characterising an enzyme as it is a reproducible constant, unique to that enzyme. It is dependant on both the specific substrate and environmental conditions such as pH, temperature and ionic strength.
The Km value has two meanings; firstly it is equal to the substrate concentration at which the reaction rate is half its maximum value, that is, the substrate concentration when half the enzyme active sites are filled, see fig. 2.11.

\[ \text{Km} = [S] \text{ at } 0.5V_{\text{max}} \tag{6} \]

Hence the lower the Km value, the more compatible the substrate is with the enzyme active site, and the lower the concentration required to saturate the enzyme.

Secondly, the Km value is related to the rate constants of the individual steps of the catalytic scheme, in fact it is equal to the dissociation constant of the ES complex, see equations 4 and 5. This relationship is only true for the equilibrium approach if \( k_1 \) is much greater than \( k_2 \) (\( k_1 \gg k_2 \)), that is, when the dissociation of the ES complex to E and S is much more rapid than the formation of E + products, see steps 1 and 2. When this is true Km is a measure of the strength of the ES complex (how tightly the substrate binds); a high Km indicates weak binding, while a low Km indicates strong binding of the substrate to the enzyme active site.

The turnover number, \( k_{\text{cat}} \), of an enzyme is the number of substrate molecules converted into product by an enzyme molecule in a unit time, when the enzyme is fully saturated with substrate, i.e. at \( V_{\text{max}} \) when \( [S] \gg \text{Km} \).

Then, \( k_{\text{cat}} = k_2 \tag{7} \)

Hence, \( V_{\text{max}} = k_{\text{cat}}[E_{\text{total}}] \), \tag{8} \]

since \( [ES] = [E_{\text{total}}] \)

It is a reflection of how fast the enzyme will work when saturated with substrate, and is dependant on the amount of enzyme present. The turnover numbers for most enzymes with their physiological substrates are in the range 1 to \( 10^4 \) per second.

Under normal physiological conditions however, most enzymes are not saturated with substrate. When \( [S] \ll \text{Km} \) and most of the enzyme active sites are unoccupied, the enzymatic rate is much less than \( V_{\text{max}} \). Under these conditions the enzyme velocity, that is, the rate of forward reaction, \( v_f \), depends on the value of \( k_{\text{cat}}/\text{Km} \) and on \( [S] \). The \( k_{\text{cat}}/\text{Km} \) value (specificity constant) is independent of the
amount of enzyme present and is comparable against other enzymes, making it a very useful parameter. The higher the $k_{\text{cat}}/K_m$ ratio, the better the substrate.

The ultimate limit on the value of $k_{\text{cat}}/K_m$ is set by the rate of diffusion controlled encounter between the enzyme and its substrate, hence the maximum $k_{\text{cat}}/K_m$ is between $10^8$ and $10^9 \text{m}^{-1}\text{s}^{-1}$. If the $k_{\text{cat}}/K_m$ ratio of an enzyme lies within this range, it is said to have obtained kinetic perfection and its catalytic velocity is restricted only by the rate at which it encounters substrate in solution. The $k_{\text{cat}}/K_m$ ratio holds for enzymes with more complex reaction pathways than the simple two step scenario given here; therefore it is a commonly used parameter for comparison in enzyme kinetic studies.

Although the Michaelis-Menten model does not hold for all enzymes, the big ET-1(16-38)-ECE-1 interaction was found to fit the Michaelis-Menten kinetics, see fig. 2.12.

![Figure 2.12: Plot of reaction velocity verses [Big ET-1(16-38)].](image)

In the subsequent subsite specificity study of big ET-1(16-38) analogues, the initial velocity data for each substrate was plotted as a function of concentration to fit the Michaelis-Menten equation, using KaleidaGraph Software (Synergy Software, Reading, PA) to obtain $K_m$ and $V_{\text{max}}$ values. Turnover numbers, $k_{\text{cat}}$, were calculated from the expression $k_{\text{cat}} = V_{\text{max}}/\left[E\right]$, using a subunit molecular mass of $1.2 \times 10^5 \text{Da}$ for solECE-1 (recombinant soluble ECE-1).
2.3.3 Subsite Specificity Study of ECE-1.

In order to evaluate the specific effects of changes at the P1 and P1' positions, it was necessary to use a series of peptides identical in structure apart from the side chains at the P1 and P1'. The Km values of a number of previously submitted big ET-1 truncates were measured, see Table 2.5, to determine the most suitable candidate for the subsite specificity study.

<table>
<thead>
<tr>
<th>TRUNCATE</th>
<th>Km(µM)</th>
<th>k&lt;sub&gt;cat&lt;/sub&gt;(sec&lt;sup&gt;-1&lt;/sup&gt;)</th>
<th>k&lt;sub&gt;cat&lt;/sub&gt;/Km(M&lt;sup&gt;-1&lt;/sup&gt;s&lt;sup&gt;-1&lt;/sup&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big ET-1(16-38)</td>
<td>13.8</td>
<td>3.3</td>
<td>2.4x10&lt;sup&gt;5&lt;/sup&gt;</td>
</tr>
<tr>
<td>Big ET-1(16-34)</td>
<td>42.1</td>
<td>6.7</td>
<td>1.5x10&lt;sup&gt;5&lt;/sup&gt;</td>
</tr>
<tr>
<td>Big ET-1(16-32)</td>
<td>&gt;400</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Big ET-1(16-27)</td>
<td>&gt;400</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2.5: Km and k<sub>cat</sub> data for previously submitted Big ET-1 analogues.

Measurement of the Km value of big ET-1(16-32) and big ET-1(16-27) was hampered by poor solubility making these peptides unsuitable for use in the kinetic study. The big ET-1(16-38) analogue was selected as the most suitable substrate since it possessed the most appropriate (lowest) Km value for convenient measurement in the HPLC assay, it bound strongly to the enzyme, was efficiently cleaved and was shown to be the most soluble of the truncates.

A series of 12 big ET-1(16-38) analogues were individually synthesised by SPPS, using the protocols described previously in section 2.2.3, with the 6 amino acids Leu, Phe, Pro, Arg, Glu and Ala substituted at residues P1 and P1' flanking the natural cleavage site 21-22. Each peptide, purified to single peak by RP HPLC, was sent to Parke Davis, Michigan, where it was assayed against solECE-1.

The hydrolysis of each was characterised by the determination of the kinetic constants Km and k<sub>cat</sub>. The initial velocity (V<sub>0</sub>) data for the rate of cleavage at the labile bond, 21-22, was determined for each analogue by the appearance of the C-terminal (22-38) product by HPLC analysis. This was plotted as a function of substrate concentration to fit the Michaelis-Menten equation. The data were analysed
Investigation of ECE

by KaleidaGraph (Synergy Software, Reading, PA) to determine the kinetic parameters displayed in Table 2.6.

<table>
<thead>
<tr>
<th>PEPTIDE</th>
<th>Km (µM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$k_{cat}/K_m$ (M$^{-1}$s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big ET-1</td>
<td>2</td>
<td>0.052</td>
<td>2.5x10$^4$</td>
</tr>
<tr>
<td>Big ET-1(16-38)</td>
<td>13.8</td>
<td>3.3</td>
<td>2.4x10$^5$</td>
</tr>
<tr>
<td>Big ET-1(16-38, Phe$^{21}$)</td>
<td>7.1</td>
<td>1.83</td>
<td>2.6x10$^5$</td>
</tr>
<tr>
<td>Big ET-1(16-38, Leu$^{21}$)</td>
<td>70</td>
<td>6.1</td>
<td>8.7x10$^4$</td>
</tr>
<tr>
<td>Big ET-1(16-38, Ala$^{21}$)</td>
<td>4.2</td>
<td>0.63</td>
<td>1.5x10$^5$</td>
</tr>
<tr>
<td>Big ET-1(16-38, Pro$^{21}$)</td>
<td>57</td>
<td>5.2</td>
<td>9.1x10$^4$</td>
</tr>
<tr>
<td>Big ET-1(16-38, Glu$^{21}$)</td>
<td>190</td>
<td>6.9</td>
<td>3.6x10$^4$</td>
</tr>
<tr>
<td>Big ET-1(16-38, Arg$^{21}$)</td>
<td>360</td>
<td>15.4</td>
<td>4.3x10$^4$</td>
</tr>
<tr>
<td>Big ET-1(16-38, Phe$^{22}$)</td>
<td>18.2</td>
<td>15.4</td>
<td>8.5x10$^5$</td>
</tr>
<tr>
<td>Big ET-1(16-38, Leu$^{22}$)</td>
<td>29.6</td>
<td>15.36</td>
<td>5.2x10$^5$</td>
</tr>
<tr>
<td>Big ET-1(16-38, Ala$^{22}$)</td>
<td>131</td>
<td>26.3</td>
<td>2.0x10$^5$</td>
</tr>
<tr>
<td>Big ET-1(16-38, Pro$^{22}$)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Big ET-1(16-38, Glu$^{22}$)</td>
<td>134</td>
<td>7.9</td>
<td>5.7x10$^4$</td>
</tr>
<tr>
<td>Big ET-1(16-38, Arg$^{22}$)</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2.6: Kinetic data from systematic subsite specificity study of ECE-1.

All the 16-38 analogues were substrates for ECE-1, except big ET-1(16-38, Pro$^{22}$) and big ET-1(16-38, Arg$^{22}$) which were not hydrolysed at a detectable rate. None of the peptides in this study, including big ET-1(16-38), were found to bind to ECE-1 as effectively as big ET-1, indicated by the high $K_m$ values, making them comparatively poor substrates. However, higher $k_{cat}/K_m$ values were seen for the majority of analogues, indicating a higher rate of hydrolysis of the labile bond. This observation is consistent with previous reports suggesting the C-terminal loop of big ET-1 may hinder conversion to the mature ET-1 species in some way.

It was predicted that ECE-1 would preferentially cleave the peptide chain directly prior to a hydrophobic residue, while a number of different residues were thought to be tolerated at position 21 (P1). Hence analogues with a hydrophobic residue at position 22 (P1') should be more readily cleaved.

It was found that all analogues with a variable 21 residue were primarily cleaved at the X$^{21}$-Val$^{22}$ bond, except for the Arg$^{21}$ which was also cleaved at the
Investigation of ECE

Asp^{18}-Ile^{19} site. Small amounts of (19-38) product were detected after prolonged incubation with ECE-1 for the other peptides. All obeyed Michaelis Menten kinetics, except big ET-1 (16-38, Ala^{21}), although big ET-1 (16-38, Arg^{21}) required prolonged incubation. From these results it is apparent that varying the P1 residue does not have a dramatic effect on the rate of hydrolysis by ECE-1 as the k_{cat}/K_m values were all comparable with big ET-1(16-38). However, the charged side chains of hydrophilic Glu and Arg made relatively poor substrates, with significantly higher K_m values for those peptides. The corresponding rate of hydrolysis was slow allowing, in the case of big ET-1(16-38, Arg^{21}), cleavage at the alternative 18-19 site.

Big ET-1(16-38, Ala^{21}) showed strong binding to ECE-1, with a K_m value only two fold greater than big ET-1 itself; however, its rate of hydrolysis was ten fold greater due to its larger turnover number. It obeyed Michaelis-Menten kinetics up to 30μM, but was shown to inhibit ECE-1 at higher concentration (IC_{50}>35μM).

It was predicted that the P1' position would be more specific. Although the majority of analogues were still recognised as substrates, substitution of the P1' site with the hydrophilic residues Pro and Arg, led to compounds neither cleaved by ECE-1 nor that were inhibitors of the enzyme. Since no detectable hydrolysis at the labile bond was observed, the kinetic study was not performed for either. Big ET-1(16-38, Glu^{22}) was found to be an ECE-1 substrate, although the big ET-1(16-38, Pro^{22}) analogue, containing the less hydrophilic Pro residue, was not. This phenomenon has been seen for other peptidases on substituting Pro at the P1' site of the substrate. It is thought that it must cause a distinct conformational change at the cleavage site. Phe and Leu are more hydrophobic than the natural Trp, although they did not bind as strongly as big ET-1(16-38), they were the most rapidly hydrolysed of all the analogues.
2.3.4 Summary and Conclusions

As predicted, substitution at the P1 position was tolerated with $k_{cat}/K_m$ values remaining fairly consistent, although the large $K_m$ values indicated weak binding. All obeyed Michaelis-Menten kinetics, except big ET-1 (16-38, Ala$^{21}$), although big ET-1 (16-38, Arg$^{21}$) required prolonged incubation. This is concordant with the predicted results, that the P1 site was not specific. Big ET-1 (16-38, Ala$^{21}$) obeyed Michaelis-Menten kinetics up to 30μM concentration, above this it acted as an ECE inhibitor, although no explanation has been offered.

Substitution at the P1' position also appeared to be tolerated, except for substitution with Arg and Pro. This was unexpected, one explanation is that other residues in the peptide sequence are recognised by the enzyme, enabling binding of the substrate despite substitution of the P1' residue.

These findings support assertions that ECE-1 is not wholly specific to the cleavage of big ET to give the mature ET, but is capable of cleaving a variety of substrates.
2.4 References


Chapter 3

A Novel Purification Protocol for Soluble Combinatorial Peptide Libraries

3.1 Introduction to Combinatorial Chemistry

It has been estimated that a new drug requires the preparation and evaluation of approximately 10,000 compounds over a period of twelve years at a cost of ~£224 million. These figures, together with ever increasing competitiveness, have forced the pharmaceutical industry to adopt new technologies which promise more efficient and economic approaches to drug discovery.

Advances made in the 1980's, such as the development of robotics and the miniaturisation of in vitro testing methods, have enabled biologists to screen thousands of compounds against a large number of biological targets in a matter of days. This high-throughput screening technology revolutionised the pharmaceutical industry's ability to evaluate novel compounds for desirable biological activity, which had previously been a bottle neck in the drug discovery process. The rate limiting step was then defined by the ability of the chemist to provide a continuous supply of novel molecules for testing, the philosophy being that the greater the number of compounds screened, the greater the probability of identifying one with desirable activity.

Traditional sources of molecules include the internal chemical portfolio of the company, together with commercial and university collections; or alternatively from
natural sources such as fermentation broths, plant and marine animal extracts. However, these sources are exhaustive and represent only a small fraction of the potential diversity possible. Over the past decade a new thinking has emerged to revolutionise traditional chemical synthesis, bringing it in line with the high-throughput screening technologies.

The term combinatorial chemistry refers to the simultaneous synthesis of multiple (in some cases many millions of) compounds, all chemically and structurally distinct. Whereas traditionally each compound is synthesised and purified one at a time before submission for biological testing; combinatorial chemistry aims to rapidly produce a diverse range of permutations of a given molecule, without necessarily maintaining total control of the absolute chemical nature of each species within the mixture. This, in theory, should accelerate the drug discovery process at the ‘lead generation’ and ‘lead optimization’ stages. These vast mixtures, or libraries, are created within a similar time frame to that traditionally taken by the chemist to create single compounds, and opens the door to a degree of chemical diversity previously inaccessible.

The extent of diversity achievable through combinatorial chemistry is dependant on the number of locations within the molecule selected to introduce diversity at, and the number of different functionalities or building blocks introduced at those positions. When considering a library of peptides, the number of members of that library is determined by the number of residues into which diversity is introduced and the number of amino acids introduced at that residue according to the relationship \( N = b^x \), where \( N \) = number of compounds in library, \( b \) = number of building blocks and \( x \) = number of synthetic steps. Hence a library of hexapeptides, varied at each residue with each of the 20 naturally occurring amino acids, would contain \( 20^5 = 3.2 \) million members. This is further illustrated in Table 3.1.
Purification of Peptide Libraries

<table>
<thead>
<tr>
<th>Number of Steps at which diversity is applied</th>
<th>Number of building blocks at each step</th>
<th>Theoretical size of library</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>400</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>8000</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>160000</td>
</tr>
<tr>
<td>5</td>
<td>20</td>
<td>3200000</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td>5</td>
<td>15625</td>
</tr>
<tr>
<td>12</td>
<td>5</td>
<td>244140625</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>3135000000</td>
</tr>
</tbody>
</table>

Table 3.1: Size of library achieved depends upon both the number of building blocks used and the number of steps at which diversity is introduced.

The sudden explosion of interest in combinatorial synthesis was accompanied by the development of numerous technologies for the simple, efficient synthesis and deconvolution of ever increasing numbers of libraries containing the maximum possible diversity.

Combinatorial chemistry is synonymous with solid phase synthesis. Only using this methodology can such vast numbers of compounds be practically created. This is due to the well documented advantages of the technique: purification of intermediates is unnecessary, isolation of the resin-bound intermediates is achieved through simple filtration; reactions may be driven to completion using excess reagents and catalysts, which are easily removed from the mixture after the reaction; solid phase techniques are amenable to automation.

Perhaps due to this, the combinatorial principle was initially used to create vast libraries of peptides using the well established SPPS methodology, although recent years has seen a revival of interest in Solid Phase Organic Synthesis (SPOS) as the potential benefit of small molecule libraries has been realised. Advances in immunology, the human genome project and the molecular cloning of numerous bioactive receptors, enzymes, hormones etc. placed a huge demand on the supply of synthetic peptides for SAR exploration studies, which also threw combinatorial peptide synthesis into the limelight.
3.1.1 Combinatorial Peptide Libraries.

It was in 1984 that Mario Geysen first developed a technique to synthesize multiple peptides simultaneously, known as the 'mimotope strategy'. He synthesised 96 different peptides concurrently on a resin support covalently linked to the spherical pinhead surface of polyacrylic acid-grafted polyethylene pins arrayed in a 96 microtitre plate format. By exposing each pin to a different activated amino acid, it was possible to synthesize only one peptide per pin and to know the structure of the individual peptides on each pin head. Using multiple plates it was possible to synthesize several thousand peptides simultaneously with specific spatial location.

The biological activity of these peptides was tested, with the peptides still attached to the pins, using an ELISA binding assay against a biological target of pharmaceutical interest. Those pins carrying peptides which successfully interacted with the biological target could be visually identified, thus allowing manual separation and sequencing to reveal their identity.

The methodology was initially met with much scepticism, but pressure on the pharmaceutical industry to produce more compounds forced the synthetic chemistry community to acknowledge the potential benefit of Geysen's concept. Modification of the original Multipin technology has improved and broadened its application; in particular the use of linker groups attached to the resin support, allowing cleavage of the peptides into solution, enabling competitive binding experiments to be run. This powerful methodology has been applied to the synthesis of a broad array of libraries.

Houghten, another early pioneer of combinatorial chemistry, devised an ingenious method to rapidly produce large numbers of peptides simultaneously, using already established SPPS methodology, by physically dividing the resin into a number of individual porous polypropylene bags, known as the 'tea-bag' method. Each bag was separately exposed to different individual activated amino acids, then the bags combined for all common steps such as washing and deprotection; in this way specific, individual peptides were synthesised in each. Labelling each bag with its synthetic history enabled the incredibly simple identification of any peptides...
Purification of Peptide Libraries displaying desirable biological activity. The size of the library was limited only by the number of bags used.

The most obvious way to synthesise multiple peptides simultaneously might be to couple mixtures of amino acids to a single resin support, so that each bead in theory supports all the library members. This method has been used with marked success; however, there is little control over the final composition of the library as certain sequences would be favoured over others due to competing coupling reactions, hence equimolar representation is not achieved. This is essential to obtain meaningful biological results. The problem is circumvented by dividing the resin into a number of equal aliquots, according to the number of different amino acids to be coupled, then pooling the resin for all common steps. This technique was devised by Furka et al. and is illustrated in fig. 3.1. Each of the sublibraries could either be screened separately or recombined into one large library.

![Diagram of Furka's Split-Pool Synthesis]

Figure 3.1: Furka's Split-Pool Synthesis.

Libraries of up to approximately $10^6$ members are obtainable using these techniques; however libraries of $\sim 10^3$-$10^4$ members have been reported to be of
Purification of Peptide Libraries

optimal size. Both the Multipin and Tea-bag methods have been used commercially for this purpose.

One of the most sophisticated technologies which has emerged for the solid phase synthesis of peptide libraries is that of spatially discrete arrays, created on a flat surface or plate, employing photolithographic methods. The technique uses a photolabile N\textsuperscript{α}-amino protecting group which can be photochemically removed from specifically defined portions of the plate surface. Therefore only those sequences which were previously deprotected will react with the activated amino acid to which the entire surface is subsequently exposed. The pattern of illumination (or masking against illumination) and the sequence of reactants define the products and their location.

The number of compounds possible is limited only by the number of synthetic sites that can be addressed with the available masking strategy. It is possible to prepare up to 400 different peptides on a 10mm x 10mm surface (with a spatial resolution of 0.5mm x 0.5mm), and, from a practical point of view, this technology could be used to prepare much larger libraries (~10\textsuperscript{5}) than the multiple peptide synthesis strategies discussed previously.

Of vital importance in combinatorial chemistry is the ability to identify unambiguously the library components showing desirable biological activity, that is the ability to deconvolute the library. On-resin analysis of peptides or oligonucleotides is relatively simple, if a one-bead-one-product protocol has been employed, using reversible, colourmetric biological analysis to isolate the bead, followed by an appropriate microsequencing technique to identify the species. On-resin colourmetric assays are also applicable to spatially addressable libraries including the Multipin, Tea-Bag and Photolithographic techniques, where the location and identity of each product is recorded throughout the synthesis.

However, this strategy is not suitable if there are multiple peptides on any given bead, or the library has been cleaved from the solid support to be tested in solution. If the library is deemed to possess favourable activity, the component or components cannot be immediately isolated and identified; the active members are determined by an iterative process of sub-library synthesis and re-analysis. This
method, originally devised by Houghton et al., though time consuming, laborious and wasteful, is often unavoidable.

Direct microsequencing analysis is limited to oligomeric libraries, since such sensitive methods of absolute identification are not possible when the library consists of small organic molecules. An alternative strategy, to avoid iterative testing and re-synthesis, is to attach a unique identifier tag or code to the individual resin beads directly after each building block, using a split/pool technique. Each product can be deduced by the corresponding unique sequence of tags which give the complete chemical history of that bead.

The tags must: have mutually compatible chemistries with the desired libraries; be stable to all reagents used in the synthesis; be readily cleaved under specific, mild conditions, independently of the product; be capable of very high sensitivity detection, providing unambiguous identification. A variety of tags have been employed to date including nucleotides for encoding peptides, peptides for encoding non-peptides or non-sequencable products and photolabile aromatics (see fig. 3.2) again for encoding both peptides and small organic molecules.

Non-chemical methods of library tagging have also been reported, the most successful of which uses radio frequency (RF). The resin is divided into an appropriate number of porous micro-reactors each containing an RF encodable chip. A Split/Mix synthetic strategy is employed, as described previously, but prior to re-

Figure 3.2: Photolabile aromatics used to 'tag' peptide libraries.
Purification of Peptide Libraries

combining each sub-section is subjected to a unique RF pulse, hence each RF chip has a record of the complete synthetic history of that reactor. This principle has been exploited in one of the most commercially successful combinatorial technologies, developed primarily for small molecule library synthesis, known as IRORITM.

The high sensitivity of many of the encoding strategies makes unambiguous identification of minute quantities of material possible, enabling libraries of up to $10^8$-$10^9$ members to be rapidly deconvoluted. The technology is also amenable to small molecule synthesis and has facilitated the screening of combinatorial small molecule libraries.

3.1.2 Combinatorial Organic Synthesis

Synthetic peptide chemistry has the advantage over recombinant techniques in that it is not restricted to the 20 natural amino acids genetically encoded for, but can incorporate any form of unnatural amino acid or amino acid mimic that it is possible to create, thus greatly expanding the potential diversity achievable. Peptides are of central importance in nature, where they are used to control, catalyse and modulate all biological responses. Random screening and SAR investigation of peptides and proteins provides essential information into highly complex biological processes and often furnishes the researcher with a template on which to design novel pharmacophores.

Unfortunately, peptides are frequently unsuitable as drug candidates due to their poor bioavailability and unfavourable pharmacokinetic and pharmacodynamic properties, being rapidly broken down by peptidase enzymes in vivo. The vast majority of pharmaceutical drugs available are small organic molecules. To truly explore the enormous potential of combinatorial chemistry, the unlimited reaction types as well as building blocks must be exploited.

The development of Combinatorial Organic Synthesis (COS) was initially hampered by the limited solid phase methodology for organic transformations. It was Leznoff, in the 1970's, who first realised the advantages of performing organic reactions on a solid support. In its most simple sense, the polymeric resin could be
considered an insoluble protecting group and Leznoff reported its use as such in the synthesis of several small organic molecules, in particular for use with symmetrical bifunctional compounds.\textsuperscript{11}

Solid phase organic synthesis (SPOS) fell into disrepute due to the limitations of the methodology compared with solution phase synthesis: the range of compatible solvents was limited to those which enabled resin swelling; reaction temperatures were relatively low to avoid resin breakdown; reaction times were longer and analysis of resin-bound compound was problematic. Recently, however, SPOS has witnessed a major revival with the enormous potential of combinatorial chemistry. Current literature continuously reports the adaptation of known organic transformations to solid phase, making this one of the most rapidly expanding areas of organic chemistry.\textsuperscript{12}

To date, small molecule libraries have been based on organic molecule templates of known biological interest. The libraries have been built up around the multifunctional basic skeleton of these proven pharmocophores by attachment of a range of building blocks with complimentary functionalities. Early heterocyclic combinatorial libraries included benzodiazapines,\textsuperscript{13} β-mercapto ketones\textsuperscript{14} and quinolone antibiotics.\textsuperscript{15}

Many of the technologies developed for peptide library synthesis have been adapted for COS, in particular the encoding strategies which facilitate deconvolution of the libraries. In addition, COS has seen the development of a multitude of new technologies specifically designed for that purpose.

\textbf{3.1.3 Concluding Remarks}

Combinatorial chemistry has rapidly established itself as an important area of synthetic organic chemistry with a vast amount of literature published, scientific meetings and even companies dedicated to the advancement of the methodology. A comprehensive survey of the different methodologies available was outwith the scope of this report, however there have been several reviews published.\textsuperscript{16}
It is perhaps too early to expect to see a new drug on the market which originated from a combinatorial library; however there have been sufficient discoveries made showing drug candidate potential, to persuade the pharmaceutical industry to adapt their drug discovery programs to include this enticing methodology.\textsuperscript{17}

Combinatorial chemistry has revolutionized the drug discovery process through the vast chemical diversity achievable using this technology. The scale of research being performed and the technological advancements being made, mean that this area of synthetic organic synthesis is set to expand much further.

3.2 Purification of Soluble Peptide Libraries.

3.2.1 Aim of the Project.

The importance of studying the interaction of peptides with biological targets of interest has already been established. Peptides have proven useful both as models on which to base small molecule leads and as potential drug candidates in their own right. The production of vast peptide libraries is therefore an important application of combinatorial chemistry.

Deconvolution and evaluation of these complicated mixtures to establish which, if any, of the species posses desirable activity is not a trivial exercise; observed biological activity might be the cumulative effect of a number of the predicted library members or even due to some unpredicted impurity(ies) present. Iterative deconvolution techniques and sub-library resynthesis may fail to identify these ‘false hits’. If the library is tested on-resin using a one-product-one-bead strategy, the active peptides can be visually observed, \textit{via} an appropriate assay, and the relevant beads manually removed for microsequencing. The problem of false hits is particularly significant when the library is cleaved from the solid support and tested in solution.
Screening in solution can be advantageous as the peptides are likely to behave differently tethered to a resin compared with when completely free. Solution screening may therefore represent a more accurate evaluation of the libraries' biological activity. In addition, resin-bound libraries are limited to screening against soluble receptors, such as antibodies or enzymes.

As with all chemical synthesis, peptide synthesis results in the accumulation of impurities and by-products. In peptide synthesis much of this takes the form of chemically similar truncated peptides (truncates); shorter lengths of the desired peptide sequence which, after $N^\alpha$ deprotection, failed to couple with the next activated amino acid, hence to prevent their further participation in the synthesis were 'capped' with acetic anhydride. Due to their similar chemical nature they can be very difficult to eliminate from the desired sequence(s). When producing peptide libraries the problem of accumulating truncated peptides is obviously amplified.

Limited efforts have been made to eliminate these species from the libraries, although due to their similar nature they are likely to interfere during library screening, either by masking the activity of genuine library members or they themselves resulting in 'false hits'. The highly sensitive screening techniques required as a result of the relatively low concentration of each library member present, makes it important to eliminate any additional complicating factors so that meaningful results might be obtained from the library.

The acetylated $N$-terminal of the truncates, in principle, offers a handle with which to separate these species from the desired peptides with free $N$-termini. Ramage et al designed a pendant moiety, tetrabenzo[a,c,g,i]fluorenylmethoxycarbonyl (Tbfmoc), to exploit this distinction.\textsuperscript{18} The aim of this project was to optimize the Tbfmoc purification protocol of chemically synthesised peptides and adapt it to the facile purification of soluble peptide libraries.
3.3 Tetrabenzo[a,c,g,i]fluorenyl methoxycarbonyl.

![Diagram of Tetrabenzo[a,c,g,i]fluorenyl methoxycarbonyl]

**Figure 3.3:** Tetrabenzo[a,c,g,i]fluorenyl methoxycarbonyl (Tbfmoc) moiety attached to the peptide through its $N^\alpha$ amino acid.

Tbfmoc was based on the Fmoc protecting group, *see section 1.3.2.* It is attached to the $N^\alpha$-amino terminus of the completed peptide sequence, *via* the chloroformate, after removal of the $N^\alpha$ protecting group but prior to side chain deprotection and acidolytic cleavage from the solid support. The truncated peptides, having acetylated $N$-termini, will not attach the Tbfmoc thereby introducing a definite distinction between the peptide and truncates. The moiety imparts a number of useful properties onto the peptide to which it is attached, these have been exploited to facilitate the separation of the desired sequence from the truncates and other impurities accumulated during SPPS, *see table 3.2.*

<table>
<thead>
<tr>
<th>Property</th>
<th>Purification Advantage</th>
</tr>
</thead>
<tbody>
<tr>
<td>UV absorbance maximum at 364nm.</td>
<td>Can easily differentiate between Tbfmoc and non-Tbfmoc material.</td>
</tr>
<tr>
<td>Imparts different solubility characteristics on peptide.</td>
<td>Tbfmoc peptides can elute later on RP HPLC than non-Tbfmoc impurities.</td>
</tr>
<tr>
<td>Extremely high affinity for carbon</td>
<td>Tbfmoc-peptides adsorb, truncates can be washed away.</td>
</tr>
</tbody>
</table>

**Table 3.2:** Properties imparted on the peptide by the Tbfmoc.
The extended conjugation system of the Tbfmoc species furnishes it with a characteristic absorbance at 364nm, enabling any peptides possessing this moiety to be easily identified by UV or RP HPLC. Tbfmoc also imparts different solubility characteristics on the peptide, increasing the hydrophobicity of the species. This can result in the peptide having a longer retention time on RP HPLC, so separating it from material which would normally co-elute. Tbfmoc, being an essentially planar, aromatic molecule, has been shown to possess a high affinity for carbon. This has enabled the desired peptide sequence to be pulled out of the crude synthesis material via an affinity chromatography technique; whereby all desired sequences are adsorbed onto the insoluble carbon, via the Tbfmoc, leaving all truncates and other impurities in solution. The carbon can then be washed, removing all remaining impurities, before cleaving the desired peptides back into solution.\textsuperscript{19}

Tbfmoc has been used previously to purify single peptides and proteins, the following work describes its application to the purification of five randomly selected soluble peptide libraries.

3.2.3 Synthesis of Combinatorial Peptide Libraries

Five randomly selected libraries were synthesised via automated SPPS using the Fmoc orthogonal protecting strategy on Wang resin, see fig 3.4.
The first two libraries each consisted of a mixture of three truncated analogues of the 21 amino acid peptide hormone Endothelin-1 (see chapter 2) in cooperation with a separate investigation on going at that time. The third and forth libraries were both based on a truncated segment of the leumorphin protein, selected due to its ease of synthesis. Three amino acid diversity was introduced at three sites in library 3, resulting in 27 unique peptides. Purification of this library to its predicted 27 members only, demonstrated the ability of Tbfmoc to pull out the relevant sequences from larger mixtures. The final library consisted of nine analogues (including the natural species) of the 40 amino acid β-Amyloid protein, involved in the pathology of Alzheimer’s disease. The successful isolation of this library demonstrated the adaptability of Tbfmoc to the purification of small protein libraries.

Each of the libraries was synthesised on an Applied biosystems 430A peptide synthesiser on a 0.25mmol scale using the general protocol outlined in Chapter 5, together with the most simple method of multiple peptide synthesis to achieve the
Purification of Peptide Libraries

diversity. This entailed the cartridges corresponding to the sites of diversity being loaded with equimolar quantities (1/3 mmol each) of the three amino acids to be introduced at that position, such that the total amino acid content of the cartridge remained at 1mmol. This enabled the libraries to be synthesised completely on the automated synthesiser without adapting the existing apparatus in anyway.

Although this method could result in a non-equimolar concentration of each library member, it was selected due to its rapid, facile nature and yielded libraries of sufficient quality to which the Tbfmoc purification protocol could be applied. Amino acid analysis results of the libraries did however indicate a relatively good equimolar presence of each product, see Appendices III-VII.

If the library was to be submitted for biological testing then equimolar quantities of each library member is essential to ensure equal representation in the assay. Since this is not guaranteed using the above method, a split/pool strategy was adopted in the synthesis of library 4. This was accomplished by dividing the resin into three equal portions at the positions of diversity and manually coupling each separately with the appropriate activated amino acid in the sonic bath, using similar conditions to those employed by the synthesiser. The portions were then re-combined for filtration and washing, before the resin was returned to the automated synthesiser which performed all common steps either to the next position of diversity, where the process was repeated, or to the end of the synthesis.

Non-optimum conditions were applied to each synthesis using DIC/HOBT with single coupling throughout, except for library 5 (see section 3.7) in order to demonstrate the effectiveness of the Tbfmoc application in achieving pure peptide libraries. On completion of the synthesis the Nα-Fmoc was cleaved from the resin-bound library of peptides and Tbfmoc loaded via the chloroformate, see fig. 3.5. The crude libraries were all cleaved from the solid support and the functionalized side chains deprotected using a standard acidolytic cleavage protocol.
AcOHN—peptide — peptide—NHFmoc

20% piperidine
in dioxane

AcOHN—peptide — peptide—NH₂

Cl

TbfmocCl/DCM

AcOHN—peptide — peptide—NHTbfmoc

TFA/scavengers

HO—peptide—NHTbfmoc

+ HO—peptide—NHOAc (Truncated peptides with acetylated N-terminal)

Figure 3.5: Loading of Tbfmoc onto the peptide libraries via the chloroformate.

3.2.4 Initial Study Using the Original Tbfmoc Purification Protocol.

Library 1 was solubilized in 1:1 6M guanidine hydrochloride/isopropanol (1:1 6M GuHCl/iPrOH). RP HPLC of the crude material showed the Tbfmoc group to have caused increased retention of the three desired peptides, such that they were sufficiently separated from the other material to allow the use of simple preparative RP HPLC to achieve purification. All material with a corresponding 364 nm absorption was collected and analysed by analytical RP HPLC, MALDI ToF mass spectroscopy and amino acid analysis, which confirmed the presence and identity of the three peptides (see Appendix III).

The RP HPLC trace of the crude second library was more complicated than that of the first, with the 364nm material overlapping with unidentified foreign material, such that simple preparative HPLC was unsuitable. Instead, for this and all subsequent libraries, Tbfmoc’s affinity for carbon was exploited. Libraries two and three used porous graphitised carbon (PGC) to affect purification.
The crude material, isolated from resin cleavage, was adsorbed onto PGC from the 1:1 6M GuHCl/iPrOH solution by vortexing, leaving all non-Tbfmoc material, including truncates, in solution. A polar solvent system was required to achieve effective adsorption of the Tbfmoc-peptides from solution and a 1:1 6M GuHCl/iPrOH system was chosen due to the poor solubility of the Tbfmoc-peptide libraries. Complete adsorption of 364nm material was confirmed by HPLC, the suspension centrifuged and the supernatant discarded. The PGC was washed repeatedly with the solvent system, using the vortex/centrifuge technique, until a flat baseline was attained on RP HPLC, indicating that all the truncates and impurities had been washed away. The libraries were then cleaved from the Tbfmoc moiety using 10% piperidine in the solvent system, leaving the Tbfmoc group still adsorbed on the carbon. The piperidine supernatant, containing the library, was neutralised with glacial acetic acid and desalted using RP HPLC.

Lyophilisation resulted in high quality, white peptide which was analysed by MALDI ToF mass spectroscopy and amino acid analysis, thus confirming the presence of each of the library members only. Fig. 3.6 shows the MALDI ToF mass spectrum of library 3 after carbon purification, this clearly illustrates the diversity within the library. Of the 27 members, 10 were of degenerate mass giving 17 mass variants (given in the table).

<table>
<thead>
<tr>
<th>Predicted Distribution</th>
<th>Peak/Value</th>
<th>Predicted distribution</th>
<th>Peak/Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1285</td>
<td>A/1287.9</td>
<td>1361</td>
<td>E/1360.9</td>
</tr>
<tr>
<td>1299</td>
<td></td>
<td>1363</td>
<td>F/1364.3</td>
</tr>
<tr>
<td>1303</td>
<td>B/1303</td>
<td>1375</td>
<td></td>
</tr>
<tr>
<td>1313</td>
<td></td>
<td>1379</td>
<td>G/1379.6</td>
</tr>
<tr>
<td>1317</td>
<td>C/1319.4</td>
<td>1393</td>
<td>H/1398.3</td>
</tr>
<tr>
<td>1331</td>
<td></td>
<td>1407</td>
<td></td>
</tr>
<tr>
<td>1345</td>
<td>D/1334.3</td>
<td>1411</td>
<td>I/1413.4</td>
</tr>
<tr>
<td>1349</td>
<td></td>
<td>1425</td>
<td>J/1425.4</td>
</tr>
</tbody>
</table>

Figure 3.6: MALDI ToF mass spectroscopy results for Library 3.
3.2.5 Optimisation of the Tbfmoc Purification Protocol

It was considered that improvements could be made to the Tbfmoc purification protocol. Having cleaved the library from the Tbfmoc, leaving the moiety adsorbed onto the carbon, the PGC is effectively spent. PGC being a very expensive form of carbon, meant that it was desirable to find a cheaper alternative to use in the purification protocol. This was found in ordinary acid washed, decolourizing, chromatographic charcoal. In addition, the guanidine hydrochloride solvent system was not ideal, being a charged species it tended to stick to the peptides making it difficult to eliminate. The purification protocol was thus optimised with the exclusion of this medium.

The Tbfmoc strategy was revised by subjecting each of the nine members of Library 4, synthesised individually, to a number of modified procedures. Under the conditions described previously, in section 3.2.4, only 50% of the peptides were recovered as clean, white material, the rest being obtained as viscous gums after purification. This was possibly due to poor de-salting of the peptide as a result of the guanidine. A trial cleavage of the peptides without Tbfmoc showed the crude material to be highly soluble in 1:1 acetonitrile/water. The protocol was therefore modified; the Tbfmoc library was adsorbed onto charcoal from 1:1 6M GuHCl/iPrOH, then, after washing twice with that system, the charcoal was washed with 1:1 acetonitrile/water and the peptide cleaved off the Tbfmoc using 10% piperidine in 1:1 acetonitrile/water. De-salting was achieved using preparative RP HPLC, holding the gradient at 2 % acetonitrile until all the salt had been eluted.

Unfortunately these modifications proved insufficient, the 6M guanidine hydrochloride was very persistent, therefore it was desirable to avoid the use of this solvent altogether. The 1:1 6M GuHCl/iPrOH system had been selected due to the strongly hydrophobic nature of the peptides when attached to the Tbfmoc. In order to eliminate the use of this system entirely, the peptides were adsorbed onto the charcoal directly from the TFA cleavage mixture without prior isolation of the crude product, making the Tbfmoc purification step an extension of the acidolytic cleavage.
The chromatographic charcoal was found to cause a persistent brown discolouration of the peptides. Although the carbon was pre-washed with each system to which it would be exposed, this was insufficient to prevent the charcoal from imparting these coloured impurities to the peptides. They were introduced during the 10% piperidine cleavage step, resulting in discolouration of the piperidine supernatant and a large number of fines. It was found, however, that after repeated washing with fresh 10% piperidine in 1:1 acetonitrile/water, the supernatant eventually remained colourless. The charcoal could then be rinsed with the neutral solvent system and finally with the TFA/scavenger cocktail, then used immediately to adsorb the Tbfmoc-peptides. It was found that the washed charcoal could not be dried or stored for later use.

Having revised the purification strategy, the vortex/centrifuge method was directly compared with a dry flash approach. In both, equimolar quantities of a single resin-bound member of library 4 (DPNGYSGELFDA) with Nα-Tbfmoc was used, with approximately the same weight of carbon (~0.5 g). The peptide was loaded straight from the TFA cleavage solution onto the carbon in both cases.

The vortex/centrifuge method was performed as described previously. In the dry-flash method, the charcoal was loaded into a small column with a sintered glass frit at the bottom (a layer of cotton wool and Celite was necessary to prevent the charcoal fines from passing through the sinter). The column was fitted to a water pump vacuum filter apparatus, and the TFA cleavage solution of the peptide poured through the column three times until all the material with a corresponding 364 nm absorbance had been adsorbed from the solution. The carbon was washed with 1:1 acetonitrile/water, before cleaving the peptide using 10% piperidine. The charcoal was allowed to remain in contact with the piperidine by removing the vacuum for several minutes, and by re-pouring the solution repeatedly through the column. Finally the carbon was rinsed with 1:1 acetonitrile/water.

In both cases, the resulting piperidine solution was acidified with glacial acetic acid, the acetonitrile removed in vacuo and the solution subjected to preparative RP HPLC. The single peak, eluted at ~33% acetonitrile, was collected and freeze dried. Although both methods resulted in clean, white product, the
vortex/centrifuge method resulted in a yield of 83.8% pure peptide, while the dry flash method 62.5%. Hence it was concluded that the vortex/centrifuge method was the better strategy for Tbfmoc purification.

The efficiency of charcoal versus PGC was also assessed, again using the Tbfmoc-DPNGYSGELFDA peptide, using the vortex/centrifuge method with equal amounts of each form of carbon. The PGC was found to give a 53.7% yield compared with the 83.8% yield for charcoal, making charcoal the preferred form of carbon for the Tbfmoc purification strategy.

\[
\text{Figure 3.7: Optimised Purification protocol.}
\]
Library 4 was synthesised using a split/pool strategy, see section 3.1.1, giving a final coupling efficiency of 46%, hence the library contained a significant amount of impurities. When it was subjected to the optimised purification strategy (method 2a, see chapter 5) an 82% yield of clean, white material was obtained. The power of the Tbfmoc/carbon purification strategy is illustrated in fig. 3.8, which compares the analytical RP HPLC of the crude Library 4, with that of the library after it had been subjected to the optimised purification strategy.

\[\text{Crude Library 4 with Tbfmoc} \quad \text{(cross-hatched area shows the peptidic material containing Tbfmoc)}\]

**Figure 3.8:** RP HPLC traces of Library 4 before and after Tbfmoc carbon purification. With dual wavelength monitoring at 214nm and 364nm

The purified library trace gave a pattern of seven separate peaks, all were eluted between 31% and 38% acetonitrile so were easily separable from any salts present. The composition of the library was confirmed by MALDI ToF mass spectrometry and amino acid analysis, see Appendix VI.

### 3.2.6 Synthesis and Purification of Library 5.

The final library of β-amyloid analogues, Library 5, was constructed in order to demonstrate the ability of Tbfmoc to effectively purify libraries of small proteins. The library was synthesised, as described previously, with the amino acid cartridges at the two positions of diversity being loaded with multiple amino acids. The positions of diversity were randomly selected. The residues incorporated were chosen in order to achieve the maximum mass variation.
Initially diversity was introduced at residues 13 and 29, see fig. 3.5, with the incorporation of His, Arg & Ser and Gly, Lys & Tyr respectively. Being a longer peptide, the improved coupling agent, HOCt, was employed as opposed to HOBt. Unfortunately the synthesis failed, even when it was repeated with double coupling.

The positions of diversity were revised and instead of replacing Gly(29) with much larger amino acids, Lys(28), a bulky amino acid, was chosen as the site of diversity and the less bulky Ala and Arg residues incorporated at this position. It was hoped that this would prevent the large drop in coupling efficiency observed at this residue during the previous two attempts at the library synthesis. HOCt double coupling was employed throughout, with extended double coupling times at the sites of diversity. The synthesis was performed on a 0.2 mmol scale, as opposed to 0.25, meaning five equivalents of amino acid were available for each coupling instead of four. The final quantitative Fmoc loading test indicated only a final 23% coupling efficiency.

The library was cleaved as described previously, using standard acidolytic techniques. A trial cleavage of the library without Tbfmoc revealed that the crude material dissolved in 3:2 acetonitrile/water, hence this was selected as the solvent system for purification. The 'dry flash' column method (method 2b, chapter 5) was adopted for the purification, as opposed to the vortex/centrifuge technique; although the yield was not as good using this method, it was suspected that exposing the charcoal to less disruptive conditions might reduce the release of the unidentified, coloured impurities.

Preparative HPLC of the piperidine cleavage solution showed a distinct series of sharp peaks, which were collected together and lyophilised giving clean, white peptide material (60% yield). Electrospray mass spectroscopy and amino acid analysis results gave good correlation with the predicted results, see Appendix VII.
3.3 Summary and Conclusions

Tbfmoc has been successfully used to facilitate the separation and purification of five soluble peptide libraries from truncated peptide fragments and other impurities which accumulate during SPPS. The affinity of Tbfmoc for carbon, both PGC and charcoal, has been exploited in the purification of four of the five libraries with potentially better recovery from the charcoal than PGC.

Adsorption of the peptide directly from the TFA cleavage solution eliminated the problem of insolubility of the Tbfmoc-peptides and hence the necessity of using non-ideal solvent systems. For each library the exact purification conditions were optimised, with an appropriate solvent system selected by establishing the solubility of the crude library without Tbfmoc.

The instability of the charcoal under basic conditions can be overcome by thoroughly pre-washing the charcoal, however, the charcoal must be used immediately and cannot be stored dry. Comparison between the vortex/centrifuge and 'dry-flash' column methods of charcoal purification suggest that, although the yield from the latter is slightly lower, disturbing the charcoal less during the purification protocol results in less brown discolourisation and hence a higher quality of resulting product.

The versatility of the Tbfmoc purification strategy has been demonstrated with its effective purification of both Library 5, and much larger peptides and proteins at Edinburgh, which accumulate increased numbers of impurities during their synthesis. The Tbfmoc principle could therefore be adapted to achieve facile purification of small protein libraries.

The methodology affords rapid, facile separation of peptide libraries from their chemically similar impurities, which could potentially interfere with the library screening. The method is adaptable and simple, resulting in high quality peptide libraries containing the predicted members only, whilst retaining the essence of the combinatorial approach.
3.4 References

Purification of Peptide Libraries

Part II: Heterocyclic Synthesis

Chapter 4

Synthesis of a Potential Farnesyl Transferase Inhibitor

4.1 Introduction

4.1.1 Discovery of Protein Prenylation

The prenylation of proteins in vivo is an important post-translational modification, whereby either a 15 carbon (farnesyl) or a 20 carbon (geranylgeranyl) isoprenoid group is covalently attached to the carboxy-terminal of a protein, via a thioether linkage to a cysteine residue.

The first prenylated protein, discovered in 1978, was a mating factor from the fungus Rhodospiridium Toruloids. The undecapeptide, named Rhodotorucine A, was shown to possess a carboxy-terminal S-farnesyl-L-cysteine methyl ester. Other species of fungus were subsequently observed to produce similarly farnesylated proteins.

In 1980, a study into the effect of compactin (an inhibitor of the isoprenoid biosynthesis) on mammalian cell growth and morphology revealed the presence of prenylated proteins in mammalian cells. It was shown that high concentrations of compactin inhibited the cell development of cultured fibroblasts. This effect was reversed by addition of mevalonic acid, suggesting that this, or some isoprenoid metabolite derived from it, was required for cell proliferation, see fig. 4.1. Insertion of radiolabelled mevalonic acid into compactin treated cells resulted in incorporation
of the label into cellular proteins. The first mammalian protein shown to undergo prenylation was the nuclear membrane-associated protein Lamin B in 1988.

\[
\text{Acetoacetyl CoA + Acetyl} = \text{CoA} + \text{H}_2\text{O}
\]

\[
\begin{align*}
\text{Acetoacetyl CoA} & \quad \text{Compactin} \\
\text{Mevalonate} & \quad \text{Mevalonate}
\end{align*}
\]

\[
\begin{align*}
\text{H}_3\text{C} & - \text{O} - \text{OP} - \text{OP} - \text{O} - \\
\text{CH}_2 & - \\
\text{H}_3\text{C} & - \text{O} - \text{OP} - \text{OP} - \text{O} - \\
\text{OP}_2\text{O}_5^3^- & \\
\text{Farnesyl pyrophosphate} & \\
\text{Geranylgeranyl pyrophosphate} & \\
\end{align*}
\]

**Figure 4.1: Biosynthetic pathway of mevalonic acid and its metabolites.**

There are two types of prenylation which can occur in mammalian cells. Farnesylation refers to the attachment of the 15-carbon farnesyl group and
geranylgeranylation refers to the 20-carbon geranylgeranyl group. The latter modification is responsible for the prenylation of 80-90% of all known prenylated proteins.

All known farnesylated and geranylgeranylated proteins are characterised by the carboxyl-terminal sequence CAAX, where C is cysteine, A is any aliphatic amino acid and X is variable. The identity of X is a major factor in determining whether the protein is farnesylated or geranylgeranylated; when X is methionine, serine, glutamine or alanine the protein is farnesylated, but when X is leucine or phenylalanine, geranylgeranylation occurs. The CAAX motif is therefore critical to protein prenylation. This criterion opened up the investigation of other proteins containing the C-terminal CAAX motif, and has revealed a wide variety of protein families which undergo prenylation. A list of known prenylated proteins has been recently published.

Of particular interest was the discovery that the Ras family of proteins, key cellular signalling molecules, were farnesylated. The Ras proteins are plasma membrane bound GTP-binding proteins, involved in the regulation of cell replication and transformation. Oncogenic versions of the Ras proteins have been shown to undergo similar post-translational modification with subsequent membrane localisation. These species can lead to uncontrolled cell growth and have been implicated in a wide variety of proliferative diseases, ~25% of all human cancers are associated with oncogenic mutants of Ras in tumour cells.

Farnesylation of these oncogenic variants of Ras and subsequent membrane localisation is imperative to the transformation of the cell to a tumorigenic state. This revelation has triggered widespread research into possible methods to disarm these mutated Ras proteins.
4.1.2 Ras Proteins

Three functional Ras genes have been identified, cloned and sequenced in humans, H-Ras-1, K-Ras-2 and N-Ras.\(^{14}\) They encode for three highly related proteins known as p21. Two other genes, H-Ras-2 and K-Ras-1, have also recently been identified in rats and humans.\(^ {15}\) All are members of the Ras superfamily of GTPases which also includes Rho/Rac, Rab, Ran and ARF (ADP-ribosyl factor).\(^ {16(a)}\) The Ras p21 proteins are either of 188 or 189 amino acids; the first 85 residues are completely conserved, the subsequent 80 posses 85\% homology and the remaining are variable, except for the terminal CAAX motif present in all.

After construction in the cystol, the p21 Ras proteins undergo a series of post translational modifications. These are common to the majority of known prenylated proteins, enabling their localisation into the inner plasma cell membrane, where they trigger a cellular signalling cascade.\(^ {16}\) Fig. 4.2 illustrates these post translational modifications for the Ras proteins.

\[\text{Figure 4.2: Post-translational modification of the Ras } p21 \text{ proteins.}\]
The initial step is the prenylation of the cysteine residue in the CAAX motif with farnesyl pyrophosphate (FPP) via a dedicated farnesyltransferase enzyme, see later, to give a thioether linkage. Subsequent modification involves the proteolytic cleavage of the three C-terminal amino acids via a specific membrane-bound endoprotease. This is followed by methylation of the S-prenyl-cysteine by a membrane-bound methyl transferase (using S-adenosyl-L-methionine as the methyl donor), to give the C-terminal methyl ester characteristic of the majority of prenylated proteins.

The farnesylation step has been shown to be obligatory to membrane localisation and hence Ras p21 activity; however the subsequent two steps, although shown to be important for the proper targeting of Ras to the plasma membrane and for full Ras function, have been shown to be less critical.

K-Ras possesses a stretch of basic residues in its sequence, which are thought to increase the binding affinity of the protein. H-Ras and N-Ras lack in this sequence, but are acylated with palmitic acid (palmitoylation) at Cys residues upstream of the farnesylated Cys (Cys for N-Ras, Cys and Cys for H-Ras) prior to membrane localisation. This additional modification has been shown to greatly potentiate H-Ras function.

Figure 4.3: Palmitoylation of the Ras proteins.
The prenyl group serves as a hydrophobic anchor to bind the protein to the cell membrane. One study has suggested that the farnesyl group inserts into the hydrophobic interior of the membrane, rather than binding to a well-defined pocket on a protein receptor. Once localised in the membrane, the Ras p21 proteins are capable of binding GDP/GTP and possess intrinsic GTPase activity.

The protein normally exists in the inactive GDP bound state, however interaction with a growth signal protein of the tyrosine kinase family, causes the protein to bind GTP, altering its conformation. This change in conformation allows the active form to interact with cytosolic serine/threonine kinases, the best characterised being Raf-1. Raf-1 is recruited to the plasma membrane where activation occurs through an unknown mechanism, triggering an intracellular signalling cascade through the phosphorylation of MAP kinase kinase (MEK), which in turn phosphorylates the MAP (mitogen-activated protein) kinases. These species are then translocated to the nucleus, where they phosphorylate and consequently activate multiple transcription factors, including c-jun and c-fos, leading to cell proliferation and differentiation. This simplified explanation of the signalling pathway is illustrated in fig. 4.4.

Figure 4.4: Intracellular signalling cascade.
Other effector molecules have been shown to be involved in alternative Ras signalling pathways, investigation of which is continuing.\textsuperscript{25}

The Ras protein acts as a switch, which is turned on or off depending on the guanine nucleotide attached to the protein. It regulates the on/off state of the signalling pathway by cycling between the GTP-bound active form and the GDP-bound inactive form, depending on stimulus by the growth signal protein. In normal Ras, the active form is rapidly deactivated by the GAP NF1 enzyme, which dephosphorylates the bound GTP, returning the Ras to its original inactive conformation and hence shutting down the signalling pathway.\textsuperscript{26}

As mentioned previously, mutated forms of Ras proteins are found in a wide variety of human tumour types, see Table 5.1. The most commonly found mutations occur at positions 12, 13, 61 and 63, particularly at 12.\textsuperscript{27}

<table>
<thead>
<tr>
<th>Tumour Type</th>
<th>Predominant Mutant Ras Gene</th>
<th>Occurrence of Ras Gene (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>pancreas</td>
<td>K-Ras</td>
<td>80-90</td>
</tr>
<tr>
<td>colon</td>
<td>K-Ras</td>
<td>30-60</td>
</tr>
<tr>
<td>small intestine</td>
<td>H-Ras</td>
<td>31</td>
</tr>
<tr>
<td>lung</td>
<td>K-Ras</td>
<td>27-60</td>
</tr>
<tr>
<td>prostate</td>
<td>H- and K-Ras</td>
<td>0-25</td>
</tr>
<tr>
<td>liver</td>
<td>K- and N-Ras</td>
<td>12-26</td>
</tr>
<tr>
<td>skin</td>
<td>H-, K- and N-Ras</td>
<td>0-46</td>
</tr>
<tr>
<td>thyroid</td>
<td>H-, K- and N-Ras</td>
<td>0-60</td>
</tr>
<tr>
<td>leukemic</td>
<td>K- and N-Ras</td>
<td>6-40</td>
</tr>
</tbody>
</table>

Table 4.1: Occurrence of Ras proteins in human tumours.\textsuperscript{1(o)}

K-Ras is the most prevalent of the proteins found in oncogenic form in human tumours. Mutations disrupt the GTPase activity of the Ras by locking the protein in its active GTP-bound conformation, causing the signalling pathway to be permanently "on". This leads to the continuous production of growth promoting transcription factors and hence oncogenicity.
It is therefore desirable to disarm these Ras oncogens and so halt the signalling pathway. The most favourable point for interference in the pathway is thought to be the farnesylation step, possibly by inhibition of the dedicated farnesyl transferase enzyme. The farnesyl group is essential for the location of Ras to the inner membrane, such localisation is obligatory both for normal Ras function and also to the activity of oncogenic forms of the protein.

4.1.3 Protein Farnesyltransferase

There are three known prenyl transferase enzymes, all αβ heterodimers consisting of two subunits: farnesyltransferase (FTase),\(^ {28}\) geranylgeranyltransferase type I (GGTase I)\(^ {29}\) and geranylgeranyltransferase type II (GGTase II)\(^ {30}\). The former is responsible for the farnesylation of biological proteins, while the latter two are responsible for their geranylgeranylation. Both FTase and GGTase I possess identical α-subunits (48 kDa),\(^ {31}\) while the β-subunits are distinct (46 kDa and 43 kDa respectively). GGTase II, involved with the prenylation of proteins characterised by the C-C or C-X-C motif, has distinct α and β subunits. All have been identified, cloned and characterised. Since this study is interested specifically in the farnesylation of Ras proteins, the discussion is limited to the FTase enzyme.

FTase is a zinc metalloenzyme, containing a single bound zinc ion.\(^ {32}\) It is responsible for the transfer of a farnesyl group from FPP to the cysteine SH of the CAAX motif of the relevant protein; the exact mechanism for which has not yet been precisely determined. One group has proposed an electrophilic mechanism for yeast FTase.\(^ {33}\) A nucleophilic mechanism has been proposed for the mammalian enzyme, in which the zinc ion activates the cysteine thiol of the peptide substrate as a thiolate anion for nucleophilic attack on the isoprenoid substrate.\(^ {34}\)
Several studies support the catalytic role of the zinc, and one has shown the metal ion to be co-ordinated by the thiol of the CAAX Cys. The rate limiting step for the process has been shown to be the dissociation of the farnesylated protein product from the enzyme.

The crystal structure of this enzyme was published in 1997; within a highly helical, tightly bound structure, two potential binding sites were shown to exist. The β-subunit forms an α-α barrel structure, or deep pocket, containing a hydrophobic core. Modelling studies have shown that a farnesyl pyrophosphate (FPP) molecule fits exactly into this cavity. Placing the terminal carbon of the isoprenoid at the bottom of the cleft allows the diphosphate moiety to directly interact with the catalytic zinc at the enzyme active site. This is concordant with previous studies, which suggested that the FPP bound exclusively to the β-subunit.

The exact depth of the cavity explains the enzyme discrimination between the FPP and GGPP substrates. Although the GGPP could also bind in this cavity, the five additional carbons places the diphosphate group out of range of the active site and hence FTase is unable to transfer this substrate to the Ras protein.

The second binding site is near the subunit interface, and intersects the putative isoprenoid binding pocket at the enzyme active site (marked by the bound zinc ion between the two subunits). This site takes the form of a hydrophilic cleft running nearly orthogonally to the isoprenoid binding pocket. Evidence suggests that the protein substrate binds in this cleft through the final nine residues at the C-
terminus.\textsuperscript{37} This would bring the Cys of the CAAX moiety exactly into register with both the zinc ion and the FPP $\alpha$-phosphate. Site-directed mutagenesis studies have begun to identify residues critical to substrate binding, specificity and the catalytic activity of the FTase.\textsuperscript{39}

4.1.4 Protein FTase Inhibitors

As previously described, inhibition of FTase would prevent the membrane localisation of oncogenic versions of Ras, and hence may offer a valuable method for the treatment of a number of forms of cancer, many of which are resistant to current treatments. This possibility has been the focus of much research across the pharmaceutical industry over the past decade. The enormous variety of potential inhibitors that have emerged from these studies has been extensively reviewed.\textsuperscript{40}

The Ras proteins were among the first shown to be farnesylated, the FTase enzyme appearing therefore to be exclusive to Ras. This presented an ideal candidate for inhibition. Some 20 other proteins have subsequently been shown to be farnesylated \textit{via} the same enzyme, however so far this not been shown to be a problem, with some reports suggesting that the inhibition may even be specific to malignant cells.\textsuperscript{41}

Many of the potential inhibitors to emerge have shown great promise as novel anticancer agents, showing in some cases almost complete tumour regression apparently without any toxic side effects.\textsuperscript{42} Remarkably, some FTase inhibitors have been shown to be active against cancer cells which do not contain mutated Ras proteins. This suggests that other FTase substrates may play a role in the pathogenesis of proliferative diseases.\textsuperscript{42,43}

Early in the investigation of the Ras proteins, the CAAX motif of the natural \textit{H}-Ras p21 substrate (CVIL) was found to be a Ras-competitive inhibitor of FTase,\textsuperscript{44} hence the design of many of the early inhibitors was based on this. Having explored peptidic analogues, non-peptide CAAX mimics emerged with a view to improving the pharmacokinetic properties of these species. Another key modification was the replacement of the thiol group with a more suitable candidate, due to concern about
the possible long term adverse effects of these functionalities. These investigations have produced some of the most potent compounds to date, and provided much of the information on the biological consequences of FTase inhibition.

FTase inhibitors have also been designed based on the farnesyl moiety, offering competitive inhibition with the FPP substrate as opposed to the Ras substrate. This class of inhibitors is less popular, since FPP is the substrate for several enzymes other than FTase, hence leading to side effects. However, Parke Davis have developed an FPP competitive inhibitor which has not shown such toxicological problems. On the whole, however, this class of FTase inhibitors has generally been less potent, and has not demonstrated any in vivo activity.

An alternative approach to specific design in the drug discovery process is through the high throughput screening of chemical libraries. This technique has been responsible for a number of novel, potent FTase inhibitors, several of which are now in clinical trials. The mode of enzyme inhibition, for all these candidates, is via competitive inhibition with the Ras protein substrate. Preliminary clinical results show these compounds not only to provide potent in vivo activity, but also to be remarkably non-toxic.

![FTase inhibitor leads from (a) Janssen and (b) Schering-Plough.](image)

The structures of the two most clinically advanced compounds are shown in fig. 4.6. The Janssen compound ((+) isomer) is currently in Phase II trials; it is administered orally, with acceptable toxicity allowing further dose increases, and a mean terminal half-life in vivo of 13-14 hours. The Schering-Plough compound is also administered orally, and is currently between Phase I and Phase II trials. Merck
has a potential FTase inhibitor in Phase I trials but its structure has not yet been disclosed, nor has the structure of a reported Bristol-Myers-Squibb lead.

Numerous other lead compounds have been reported, showing the continued level of interest in FTase inhibitors. Indeed, preliminary data from these first compounds to enter clinical trials have shown remarkable promise, with suggestion that they operate in a more complicated manner than simply preventing Ras farnesylation alone. Why these compounds should be so non-toxic has also created an on-going debate.

The results suggest that a range of human cancers might respond to FTase inhibitor therapy. In addition, synergy between FTase inhibitors and other anti-cancer treatments currently available may further extend their therapeutic potential.

The alternative forms of Ras (H-, N- and K-Ras) are differently distributed throughout the various forms of cancer, K-Ras oncogenes being the most predominant. Other distinctions between the proteins are emerging, this raises the question as to whether the FTase inhibitors should be specific to each form. In addition, recent discoveries have shown that when the farnesylation of K-Ras is prevented by specific FTase inhibitors, the protein can be geranygeranylated allowing it to function. This 'back-up' system may account for the lack of toxicity observed for these inhibitors, however it also raises the question of whether FTase specific inhibitors are sufficient or whether comprehensive FTase/GGTase inhibitors are required.

Although the exact mechanism for tumour regression under FTase inhibition is not understood, it is hoped that the recent x-ray structure of the enzyme will aid understanding of this, and also elucidate fundamental aspects of protein prenylation and the consequences of suppressing this process. Ultimately, it should aid the design of improved FTase inhibitors.
4.2 Design and Synthesis of a Potential FTase Inhibitor.

The aim of the project was to design and synthesise a novel heterocyclic structure which may offer potential FTase inhibitor activity, by modelling the target around other known inhibitors. The investigation would be complimentary to the ongoing work at Parke Davis Pharmaceuticals, Michigan.

4.2.1 Design of a Novel Structure

The ongoing Ras program at Parke Davis was based around a structure isolated from compound library screening, see fig. 4.7.

![Figure 4.7: FTase inhibitor isolated from mass screening at Parke Davis.](image)

An extensive SAR study based on this molecule was being carried out within Parke Davis, with the view to improving its *in vitro* and cellular potency against the enzyme. Work had concentrated solely on substitution at positions 2 and 6; early indications favoured the retention of the imidazole species at position 6, but with substitution at position 2 with a thiophene group, creating the structure given in fig. 4.8(a).

![Figure 4.8: (a) Favoured structure from early Parke Davis study, (b) Janssen Pharmaceuticals patented lead structure.](image)
A novel structure was designed based on the Parke Davis findings, with incorporation of key features from the Janssen study which had culminated in the highly promising drug candidate R115777. Both these compounds act by competing with Ras protein substrate binding. The tetralone skeleton was replaced with a quinolinone structure. Since the 2-quinolinone structure was comprehensively protected by Janssen Patents, the 4-quinolinone skeleton was substituted, retaining the N-methyl group. In addition, aromatic bulk at the imidazole was mimicked by the incorporation of a p-benzoic acid group, see fig. 4.9.

![Proposed target structure](image)

Figure 4.9: Proposed target structure.

The benzoic acid modification would also serve as a site for solid support attachment, with view to the solid phase synthesis of the target.

The structure opened up a new line of investigation, complimentary to that at Parke Davis. It also offered the possibility of using combinatorial chemistry to perform an SAR study around the molecule, with a number of potential sites for diversity (X, Y, Z).
4.2.2 Initial Route to the Target.

![Initial Synthetic Route](image)

**Figure 4.10:** Initially proposed synthetic strategy.

The initial synthetic route was designed around Hay’s synthesis of Ciprofloxacin, which contained the 4-quinolinone skeleton. The advantage of this
was that the strategy had been fully optimised for both solution phase and solid phase synthesis, therefore the conditions could be directly applied to the synthesis of the target at the common steps (steps 1, 4, 5 & 6). Retrosynthesis of the target through this route reduced the structure to two precursors, marked A and B, see fig. 4.10.

Neither of these were commercially available; however, precursor A was a known compound, prepared by Grey et al to its unprotected form via a three step synthesis from the relatively cheap starting material 3-fluorophenol.53

![Synthesis reaction](image)

**Figure 4.11: Synthesis of 3-fluoro-4-hydroxybenzoic acid by Grey et al.**

As this method was time consuming, a one step para-carboxylation of the 3-fluorophenol using a Kolbe-Schmitt type procedure was preferable.54 The conditions required for para-carboxylation, as opposed to ortho, could not be met (50 atm CO, 240°C); fortunately, alternative conditions to achieve this had been reported by Sasson and Radinsky,55 based on previous work by Reimer and Tiemann.56 They reported the selective para-carboxylation of phenol using carbon tetrachloride (CCl₄) in strong base, with a copper powder catalyst. 3-Fluorophenol was subjected to these conditions, giving the 2-fluoro-4-hydroxybenzoic acid product exclusively, in 24% yield. The copper was replenished twice during the reaction, as it was deactivated with time.

Sasson and Radinsky proposed a mechanism for the process; they suggested the formation of the trichloromethyl cation on the surface of the copper catalyst. Preference for para-carboxylation over ortho was determined by the concentration of aqueous sodium hydroxide, with an initial 50% w/w NaOH solution in 50% molar excess giving the highest selectivity. They suggested that the concentrated basic solution induced a degree of geometrical organisation at the copper-aqueous
interface, forcing the phenolate anion to approach the trichloromethyl cation on the catalyst surface preferentially through the para position.

Prior to formation of the β-keto ester, via the acid chloride (fig. 4.10, step 1), the phenolic function of the 3-fluoro-4-hydroxybenzoic acid required protection. Initially, the silicon based protecting group tert-butyldimethylsilyl (TBDMS)\textsuperscript{57} was investigated for the purpose. Unfortunately, the TBDMS ester was formed together with the ether on reaction with the TBDMSCl.

Literature precedence existed for the direct conversion of the TBDMS ester to the acid chloride, using oxalyl chloride in the presence of a catalytic amount of DMF.\textsuperscript{58} Unfortunately, repeated attempts using the cited conditions failed to achieve the desired manipulation.

![Figure 4.12: Direct conversion of the TBDMS ester to the acid chloride.](image)

Benzyl protection was then attempted, by reaction with benzyl bromide.\textsuperscript{59} Again, the benzyl ester formed in addition to the ether, however the ester could be selectively removed via mild base hydrolysis to give the precursor A in an overall ~15% yield for the three steps, as illustrated in fig. 4.13.

![Figure 4.13: Formation of intermediate A.](image)

The benzoic acid was then converted to the acid chloride using thionyl chloride, and the product used directly in the formation of the β-keto ester precursor.
with potassium ethyl malonate, see fig. 4.14, as in the synthesis reported by Wemple.  

\[
\text{EtO-}\text{O-K}^+ 
\begin{array}{c}
\text{1. MgCl}_2, \text{Et}_3\text{N}, \text{CH}_3\text{CN, rt, 2.5hrs} \\
\text{F} \text{Cl} \\
\text{2. Cl}, \text{rt, 16hrs} \\
\text{EtO\text{O-Bn}} \\
\end{array}
\rightarrow
\text{EtO-}\text{O-Bn}
\]

**Figure 4.14:** Formation of the β-keto ester via the acid chloride.

This methodology had originally been developed as a route into the β-(fluoraryl)-β-oxo-ester precursors for the synthesis of quinoline antibacterial agents. Optimum conditions were established using potassium ethyl malonate with mild base, anhydrous magnesium chloride and triethyl amine to give a stable enolate intermediate.

Selective cleavage of the benzyl ether via hydrogenation at atmospheric pressure gave 3 (47% yield for the combined steps); however the subsequent Mitsunobu reaction (see fig. 4.10) required the second precursor B, the synthesis of which had been hampered (see later).

### 4.2.3 Initial Modification of the Synthetic Strategy

In the absence of B, the synthetic strategy required modification. Phenol had been reported to open the epoxide function of styrene oxide under basic conditions in a clean, high yielding reaction. Attack was favoured at the more hindered position, giving a mixture of both the primary and secondary alcohols in a 4:1 ratio, see fig. 4.15.
Figure 4.15: Opening of styrene oxide with phenol.

This manipulation could be applied to the synthesis of the target structure by using the phenolic side chain of intermediate 3 to open 4-styrene oxide benzoic acid, with substitution of the resulting alcohol group with imidazole at a later step in the synthesis, see fig. 4.16.

To achieve the proposed target structure, the epoxide was required to open at the least hindered position, giving the secondary alcohol, the disfavoured isomer. It was considered, however, to be an acceptable deviation from the originally intended structure, and of equal potential interest to use the major isomer of this reaction.

The target molecule was also modified by the removal of the benzoic acid group, see fig. 4.10. This functionality was originally included as a site for attachment to a conventional solid support, however, it was deemed unnecessary, as this could be achieved through the β-keto ester side chain, marked in fig. 4.16, used in the synthesis of Ciprofloxacin. Elimination of this group meant that styrene oxide could be substituted for 4-styrene oxide benzoic acid, synthesised from the highly expensive 4-vinyl benzoic acid (see section 4.2.4). These modifications resulted in the modified target structure 17.
Optimum conditions for the epoxide opening were established using the phenol/styrene oxide model, before applying these to 3. The ring opened exclusively at the more hindered position, giving the product 15 in sufficient quantity for identification purposes; however, there was insufficient material for further steps to be attempted towards the target isomer and supply of 3 was limited.

**Figure 4.16: Revised synthetic strategy to the target isomer.**
4.2.4 Synthesis of Precursor B.

For the synthesis of B, a possible strategy was to start from the commercially available 4-vinyl benzoic acid. Initially the benzoic group was protected as the methyl ester, for the duration of the target synthesis (if the reaction had been performed on solid phase, the resin would have been attached at that point); the epoxide was then formed using MCPBA (overall 70.5% yield).

The epoxide was required to open from the more hindered side, using imidazole as a nucleophile, to produce B. An epoxide would normally be expected to open under nucleophilic attack in neutral or basic conditions via a SN2 mechanism at the least sterically hindered position. Under acidic conditions, it exists as the protonated species and may open by either an SN1 or SN2 mechanism, however, for both, substitution usually occurs at the most hindered carbon. When a phenyl group is attached to the epoxide, the effect of the aromatic ring makes that position more electrophilic, and hence this would be the anticipated site of nucleophilic attack, even in basic conditions. This is illustrated by the previously described reaction between styrene oxide and phenol in aqueous NaOH.

Imidazole is a weaker nucleophile than the phenolate ion, and requires activation to give epoxide opening at the most hindered site. Unfortunately, although exposed to a variety of reagents, see fig. 4.17, suitable conditions for the epoxide opening from either end were not established.

The previous concession made that an isomer of the originally proposed target was an acceptable deviation (see fig. 4.16), allowed the methodology put forward by Sala et al to be utilized. The method enabled the synthesis of 20 in good
yield (74%), by the reaction of imidazole with styrene oxide in DMF, shown in fig. 4.18.

\[ \text{Styrene Oxide} + \text{Imidazole} \rightarrow \text{Product} \]

(Alternative isomer to B)

![Figure 4.18: Opening of styrene oxide with imidazole.](image)

This would then enable the Mitsunobu reaction in the initial synthetic strategy (fig. 4.10) to be performed between 3 and 20, with subsequent completion of 17. Unfortunately, supply of the CCl₄ starting material from commercial manufacturers was delayed, hence an alternative method for the synthesis of the p-hydroxybenzoic acid precursor of A was required before the Mitsunobu reaction could be attempted.

4.2.5 Alternative Route to 2-Fluoro-4-hydroxybenzoic Acid.

Toluene is readily oxidized to benzoic acid under a variety of conditions; this presented the possibility of using such a strategy to oxidise 3-fluoro-4-methyl phenol to the required benzoic acid 10. This substituted aromatic starting material was not commercially available, however its isomer 4-fluoro-3-methyl phenol was, giving the consequent modified target shown in fig. 4.19.

![Figure 4.19: Consequent target structure resulting from the starting material 4-fluoro-3-methyl phenol.](image)

Three possible methods were investigated: KMnO₄, KMnO₄/NaOH and PbO₂. Although all successfully oxidised toluene to benzoic acid, the first offered...
the cleanest, most high yielding route. This method was therefore applied to the 4-fluoro-3-methyl phenol, but without success.

The method could only be used to achieve the benzoic acid of this system if the phenolic side chain was first protected as the acetate, then the methyl transformed into the benzyl bromide using NBS (fig. 4.20).  

\[
\begin{align*}
\text{Acetic Anhydride} & \rightarrow F_{\text{OH}} \rightarrow \text{NBS} \rightarrow F_{\text{BrH}_2C} \rightarrow \text{KMnO}_4 \rightarrow HOOC_{\text{OH}}
\end{align*}
\]

**Figure 4.20:** Transformation of 4-fluoro-3-methylphenol to 2-fluoro-5-hydroxybenzoic acid.

This method was low yielding and indirect, offering no advantage over the three step strategy used by Grey et al., see fig. 4.11.

### 4.2.6 Alternative Synthetic Strategy

The difficulty in synthesising A made any route requiring this precursor impractical. Fortunately, an extensive investigation of the literature presented an alternative, much shorter route to the target 17, through the initial synthesis of 7-hydroxy-2,3-dihydroquinolin-4-one *via* a three step process. The two pendant groups were attached, as in the initial route, at positions 2 and 6 after formation of the heterocyclic skeleton.
Figure 4.21: Solution phase synthesis of the target isomer 17.

It was Clemo and Perkin in 1924 who first observed that certain β-anilinopropionic acids underwent ring closure, as their N-p-toluenesulfonyl (tosyl) derivatives, to give the dihydro-4-quinolone species. In 1946, Elderfield and Johnston simultaneously investigated this transformation, and reported the synthesis of a series of β-anilinopropionic acids and their subsequent cyclisations. Later, Koo published the cyclisation of the β-anilinopropionic acids to the corresponding 4-hydroxyquinolone species via a simple one step treatment with polyphosphoric acid (PPA).

Synthesis of the 7-hydroxy-2,3-dihydro-1H-quinolin-4-one analogue of 28 (containing the free secondary amine species) had been reported by Kavrakova in 1990. The communication reported a PPA induced cyclisation, as given by Koo,
although no experimental details were given. Hence, PPA cyclisation of 3-(3-
hydroxy-phenylamino)-propionic acid, synthesised from commercially available 3-
aminophenol and acrylic acid, was initially performed prior to N-methyl protection. However, the 2,4-dihydroquinolone product was not isolated from the work-up after the cyclisation step.

The literature suggested that protection of the amine was beneficial, therefore methylation of the 3-aminophenol with methyl iodide was made the first step. This afforded a mixture of the mono- and dimethylated species as a colourless oil (combined yield 86%). Silica chromatography afforded the product separate from the starting material, but unfortunately only partially resolved from the dimethylated species. Reduced pressure distillation also failed to achieve separation, although the dimethylated species was identified as a solid, while the monomethylated product an oil. This was not considered a significant problem however, as the dimethylated species would not participate in the subsequent steps, and could be eliminated later in the synthesis.

Reaction of 25 with acrylic acid afforded intermediate 27 as a red oil. TLC showed all the monomethylated species to have been consumed, and the crude product to consist of three well separated components. Unfortunately, the product proved to be unstable to silica chromatography, and hence PPA cyclisation was performed on the crude material.

Thus, 27 was combined with a ten fold excess of PPA, and the viscous mixture stirred manually at 100°C until the colourless paste had turned a dark red. On work-up a canary yellow solid was obtained, which was purified to single spot by silica chromatography, and identified as the cyclised product 28. The overall yield for the three steps was ~10%.

Elderfield attributed the characteristic canary yellow colour of the 4-
hydroxyquinolinones to the existence of these substances as a series of tautomers, see fig. 4.22. The quinoid chromophore exists in one of the latter two tautomers which are zwitterion in nature.
Optimum conditions for the Mitsunobu reaction were then established, using 20 and phenol as a model system (fig. 4.23).\textsuperscript{72}

Unfortunately, the product 30 was inseparable from the triphenylphosphine oxide (PO Ph\textsubscript{3}) by-product by either chromatography or washing; the use of tributylphosphine did not circumvent this problem. This was also true for the product 29.

The small quantity of 29 obtained was insufficient to attempt other means of separation, however contamination would not interfere with the biological testing of the compound. Hence 29 and 30, both contaminated with PO Ph\textsubscript{3}, together with 28, were submitted for testing by Parke Davis, Michigan (see later).

The final step to complete the target synthesis (Aldol reaction with 2-thiophenecarboxaldehyde) was not performed, as later results from the ongoing FTase inhibitor study of the tetralone analogues at Parke Davis indicated that the presence of the thiophene group was unnecessary, and in some cases actually resulted in an increase in the IC\textsubscript{50} value.\textsuperscript{46}

The ultimate aim of the project was to use combinatorial chemistry to investigate the SAR. To achieve this, it was necessary to adapt the solution phase synthesis to solid phase. This was also prompted by the failure to separate the Mitsunobu product from the PO Ph\textsubscript{3} by-product. The oxide remained entirely soluble.
under the reaction conditions, making it simple to remove under solid phase conditions.

4.2.7 Solid Phase Synthesis of the Target Structure 29.

The solution phase strategy was highly amenable to solid phase chemistry. In addition, the 2,4-dihydroquinolinone structure offered a number of potential sites at which diversity could be easily introduced onto the basic skeleton for combinatorial exploration.

Initially, REM resin, was proposed as the solid support since the linker group, which would be cleaved with the final product, was equivalent to acrylic acid. Solution phase investigation of the 3-aminophenol/benzyl acrylate model demonstrated that the N-methylation could be performed after the Michael addition of the aniline, hence coupling of the 3-hydroxyaniline to the REM resin was performed first. Glacial acetic acid was required to promote the reaction in the absence of the free carboxylic acid function of the acrylic acid. This meant that the N-alkylation step would be performed on the solid support, enabling diversity to be introduced at the N, prior to cleavage and cyclisation to give the target structure, see fig. 4.24.

Figure 4.24: Proposed solid phase synthesis on REM resin.
It was realised that Wang resin could mimic the REM resin by attachment of acrylic acid directly onto the linker, using the standard peptide coupling system of DIC/DMAP. The expense of REM resin made Wang the preferred solid support. Having loaded the acrylic acid, the subsequent procedures could be carried out as in the solution phase strategy, see fig. 4.25.

Figure 4.25: Solid phase synthesis on Wang resin.

On-resin analysis of each intermediate proved challenging. Although estimation of the extent of loading could not be made, IR spectroscopy of 32 and 33 showed firstly the appearance of the C=O (1721.8 cm\(^{-1}\)) and C=C (1654.8 cm\(^{-1}\)) stretches, after loading the acrylic acid onto the resin, followed by increase of the former (1731.9 cm\(^{-1}\)) and disappearance of the latter as the conjugated system was lost on addition of 3-hydroxyaniline.

Elemental analysis of 33 was also used to confirm the presence of N. The result suggested over 100% loading, however the value was subject to a relatively large error margin. \(^{13}\)C gel phase nmr was also performed on each of the
intermediates, but the resulting spectra did not provide useful information. Intermediate 34 differed from 33 by a single methyl group, and on-resin analysis techniques (including IR spectroscopy) could not detect this small change, hence a sample of 34 was cleaved from the resin to confirm the completion of this step.

Ideally, it was hoped to achieve cleavage and cyclisation in one step; however, PPA failed to liberate the product from the solid support, possibly as it was too viscous to penetrate the resin matrix. Hence, 30% TFA in DCM was used to cleave the product, then the solvent was then removed in vacuo and the residue treated with PPA as before. Under this procedure, 34 was successfully cleaved from the resin and cyclised to give 28.

This study demonstrated the successful adaptation of the target synthesis to solid phase. The basic 2,3-dihydroquinolinone skeleton was achieved in reasonable yield, with construction of the precursors on resin. Although the two step cleavage and cyclisation protocol was not ideal, particularly for combinatorial library production, it was simple and highly effective. Alternatively, another resin could be employed which cleaved the ester linkage giving an activated carboxylic acid, which could then spontaneously cyclise without the use of PPA. Diversity could be easily introduced, both at the N-alkylation step and in the Mitsunobu reaction.
4.3 Biological Testing and Conclusions.

The efficient solution phase synthesis of the novel target structure 29 was achieved and the reaction conditions shown to be viable to solid phase synthesis, although combinatorial exploration was not pursued. Three compounds were submitted for biological testing of their FTase inhibitor activity, the results of which are displayed in Table 4.2.

<table>
<thead>
<tr>
<th>Compound</th>
<th>IC₅₀ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>7-Hydroxy-1-methyl-2,3-dihydro-1H-quinolin-4-one 28</td>
<td>&gt;30</td>
</tr>
<tr>
<td>1-(2-Phenoxy-2-phenyl-ethyl)-1H-imidazole 30</td>
<td>4.2</td>
</tr>
<tr>
<td>7-(2-Imidazole-1-yl-1-phenyl-ethoxy)-1-methyl-2,3-dihydro-1H-quinolin-4-one 29</td>
<td>2.6</td>
</tr>
</tbody>
</table>

Table 4.2: IC₅₀ values of the compounds screened for their FTase inhibitor potential.

Although target molecule 29 did not possess exceptional FTase inhibitor activity, it was comparable to PD 405123 (structure first identified from mass screening). The basic dihydroquinolin-4-one skeleton did not show significant activity; interestingly, however, 30 showed inhibitor activity comparable to that of the target. This simpler structure would be worthy of further investigation, and offers an ideal candidate for combinatorial exploration.
4.4 References


40. (a) C. Omer and N. Kohl, *TiPS* 1997, 18, 437; (b) S. Sebti and A. Hamilton, *Pharmacol. Ther.* 1997, 74, 103; (c) G. Kelloff et al., *Cancer Epidemiology,*


48. (a) E. Bernhard, G. Kao, A. Cox et al., Cancer Res. 1996, 56, 1727; (b) E. Bernhard, W. McKenna, A. Hamilton, S. Sebti, Y. Qian, J. Wu and R. Muschel, Cancer Res. 1998, 58, 1754.


Chapter 5

Experimental


5.1 General notes.

All Fmoc protected amino acids and functionalised polystyrene resins were purchased either from Novabiochem or Bachem and, unless otherwise stated, the amino acids were of the L-configuration. All other required reagents were purchased from either Acros or Aldrich and were checked and purified, when necessary, before use. All solvents were purchased from Acros, unless otherwise stated; and the following were dried by refluxing over the reagents given in parenthesis when required: dichloromethane (calcium hydride), tetrahydrofuran (sodium wire/benzophenone) and acetonitrile (calcium hydride). The dimethylformamide and 1,4-dioxane used were peptide synthesis grade, supplied by Rathburn Chemicals. Melting points were determined in open capillaries using a Buchi 510 melting point apparatus. Ultra violet-visible (UV-vis) spectra were recorded on a Perkin Elmer UV/Vis Lambda 11 spectrometer in the solvent indicated. Mass spectra were recorded on a Perspective Biosystems Voyager Workstation Matrix assisted laser desorption ionisation, time of flight (MALDI ToF) mass spectrometer (337nm UV laser, 1.2m flight tube) using either 3,5-dimethyloxy-4-hydroxy cinnamic acid or α-cyano-4-hydroxy cinnamic acid as the ionisation promoter. Electrospray ionisation mass spectroscopy was performed on a Micromass Platform II instrument with Mass Lynx 2.3 Build 5 software.
5.2 Notes on Solid Phase Peptide Synthesis (SPPS)

All peptides described were synthesised on an Applied Biosystems 430A automated peptide synthesiser, equipped with an Applied Biosystems 757 absorbance detector linked to a Hewlett Packard HP3396A integrator for online monitoring of the deprotection solution.

The Fmoc orthogonal protecting strategy was employed in the synthesis of all peptides and peptide libraries. Unless otherwise stated, standard acid-labile side chain protection groups were used for all amino acids containing reactive side chain functionalities, see chapter 1. p-Benzyloxybenzyl alcohol (Wang) resin was employed as the solid support for all peptides unless the C-terminal amino acid was His, Cys or Pro, when 2-chlorotrityl resin was used. Both resins required the manual loading of the C-terminal amino acid onto the resin, see sections 5.6 and 5.8.

Unless otherwise stated, each synthesis was performed on a 0.25mmol scale and was run continuously until the completed sequence was attained by the sequential, stepwise addition of each protected amino acid from the C- to N-terminus. The protocol was modified for the synthesis of the peptide libraries, see note 5.10.

Single coupling was used throughout (except in exceptional cases). The amino acid was activated as the HOCl active ester using DIC in 1:1 DMF/1,4-dioxane as the activating agent for all single peptides, and the HOBt active ester in the synthesis of peptide libraries. 4 equivalents, 1mmol, of amino acid was used per coupling, loaded into the cartridge corresponding to that residue. If the residue was His, 8 equivalents (2mmol) of HOBt was included in the cartridge to prevent racemisation. Following the coupling step, any unreacted N\textsuperscript{α} amino groups were capped using acetic anhydride to prevent their further participation in the synthesis.
The standard run cycle for the addition of one amino acid residue is detailed below:

1. Capping of any free $N^\alpha$ amino groups: 0.5M acetic anhydride in 1:1 DMF/1,4-dioxane with 0.125M DIEA and 0.2% HOBt.
2. Washing: 5 times with 1:1 DMF/1,4-dioxane.
3. Deprotection of the $N^\alpha$ Fmoc protection by treating twice (4 min and 1.5 min) with 20% piperidine in 1:1 DMF/1,4-dioxane. Resin drained and washed 4 times with DMF/1,4-dioxane between deprotections.
4. Washing: 8 times with 1:1 DMF/1,4-dioxane.
5. Activation of the protected amino acid as the active ester, prior to addition to the resin, using 1:1 0.25M DIC in 1,4-dioxane/0.25M HOBT or HOCl in DMF.
6. Transfer of the activated species to the reaction vessel, coupling allowed to proceed in 1:1 DMF/1,4-dioxane for 30 mins.
7. Washing: 6 times with 1:1 DMF/1,4-dioxane.

After each $N^\alpha$ deprotection the relative coupling efficiency was monitored by UV at 302nm by the transfer of a 2ml aliquot of the effluent, containing the chromophoric piperidine-fulvene adduct, in a continuous flow mode from the synthesiser to the UV detector.

5.3 Notes on Reverse Phase High Pressure Liquid Chromatography (RP HPLC)

All RP HPLC was performed on either an ABI 151 or Gilson instrument (305 pump and programmer, 306 pump, 805 solvent delivery system, 811C mixer and a 118 UV-vis detector). Unless otherwise stated, all analytical HPLC was performed on a reverse phase Aquapore (4.6x100mm, 20μm) or Vydac (4.6x250mm, 20μm) C18 column employing a flow rate of 1 ml/min, monitoring at wavelength 214 nm. When Tbfmoc was
incorporated into the peptide, dual wavelength monitoring was used at wavelengths 214 and 364 nm. Semi-preparative HPLC was accomplished on an aquapore C18 reverse phase column (10x250mm, 20μm) using a flow rate of 5 ml/min with monitoring at 214 nm. The programmed solvent polarity gradient system was tailored to each individual peptide’s requirements, however the standard protocol, used in the majority of both analytical and preparative cases, is detailed below. The peptides were dissolved in a suitable solvent system, to a maximum concentration of 5mg/ml, for injection onto the column.

<table>
<thead>
<tr>
<th>TIME (mins)</th>
<th>%B</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>32</td>
<td>90</td>
</tr>
<tr>
<td>34</td>
<td>10</td>
</tr>
</tbody>
</table>

where solvent A = H2O, 0.1% TFA
solvent B = CH3CN, 0.1% TFA

5.4 Notes on Amino Acid Analysis

Peptide (1mg) was accurately weighed into a carius tube and 6M hydrochloric acid (containing 5mg/ml Na2SO3) added (4cm depth). The tube was connected to a vacuum trolley pump via a manifold. It was cooled in liquid nitrogen until solidified, then evacuated for 5 minutes. The vacuum was closed, the tube gently heated to melt any remaining solid, allowed to stand for 10 minutes, then the process repeated a further 2 times. Having de-gassed the solution it was re-frozen, the vacuum opened and the tube sealed using a glass blowing torch. It was placed in an oven at 110°C for the required time (standard for small peptides 20 hours), after which it was allowed to cool, opened and the contents quantitatively transferred to a round bottomed flask. The solvent was lyophilised to remove the HCl. The residue was washed twice with a small volume of water, again removing the solvent, then diluted with pH 2.2 sodium citrate loading buffer. This was quantitatively transferred to a 2ml standard flask and made up to the mark with additional buffer. The sample was filtered through a 0.45 micron filter membrane before submission for amino acid analysis. Analysis was performed on a
5.5 Synthesis of ethyl-1-hydroxyl-1,2,3-triazole-4-carboxylate (HOCt).

Step 1: Preparation of ethyl diazoacetate
Ethyl glycinate hydrochloride (210g, 1.5 mol) in water (300ml) was mixed with dichloromethane (500ml) under a nitrogen blanket. The mixture was cooled to -15°C, and cold sodium nitrite (124.5g, 1.8 mol) in water (200ml) added in one portion with stirring. The temperature was lowered to -20°C and a 5% aqueous sulphuric acid solution at 0°C added dropwise, with continuous stirring, keeping the temperature below -15°C. After complete addition of the acid solution, the temperature was allowed to rise to 0°C. The reaction mixture was transferred to a cooled separating funnel and the organic layer allowed to run into 5% sodium carbonate solution (400ml) at 0°C. The aqueous layer was extracted once with DCM (400ml) and this was combined with the DCM/sodium carbonate mixture, which was stirred until no further gas evolved. The layers were separated, and the aqueous extracted once with DCM (400ml). The combined organic layers were dried (magnesium sulphate) filtered and the solvent evaporated giving a yellow liquid which was used directly in step 3.

Step 2: Preparation of the Vilsmeier reagent.
Freshly distilled thionyl chloride (53ml, 0.72mol) was added dropwise, with stirring, to DMF (56ml, 0.72mol) under nitrogen, at room temperature. After complete addition, the mixture was warmed to 40°C and stirred for 2 hours until a thick oil was obtained. The
Experimental

oil was then evaporated in vacuo until an off white solid was obtained (2 hours). The product was used immediately in step 3.

**Step 3: Preparation of the iminium chloride.**
The Vilsmeier reagent was dissolved in dry chloroform (400ml) and cooled in an ice/salt bath under nitrogen. Ethyl diazoacetate was then added dropwise, keeping the temperature below 5°C. Addition complete, the reaction was stirred at room temperature for 30 minutes, then evaporated to give an oily residue. Addition of ether to the residue caused the product to precipitate out as a yellow solid, which was filtered off under nitrogen with suction and washed with ether. The sticky pale yellow solid was stored overnight in a desiccator and then used in step 4.

**Step 4: Preparation of diazo-oxime.**
NH₂OH.HCl (44g, 0.63mol) was dissolved in a minimum amount of water and cooled in an ice/salt bath. Sodium carbonate (33.6g, 0.32mol) was added slowly with stirring to achieve pH 7.5. The product from step 3 (130g, 0.63mol) was added, with stirring, and the diazo-oxime product precipitated out as a yellow solid. After several minutes, the precipitate was collected by filtration and washed with a small volume of ice cold water.

**Step 5: Preparation of ethyl 1-hydroxyl-1,2,3-triazole-4-carboxylate (HOCl).**
The product from step 4 was dissolved in chloroform (150ml) and dried over magnesium sulphate, before filtration and addition of concentrated acetic acid (0.5ml). The solution was stood at room temperature in light until cyclisation was complete (several days). The solvent was removed in vacuo, giving an off white solid. Ethyl acetate was added, the solid isolated by filtration, washed with ethyl acetate and dried in vacuo yielding the product as fine, white crystals. The filtrate was concentrated and stood at 4°C, forcing more of the product to crystallise out. This was filtered off and the process repeated until no further crystals were formed. Overall yield 21.84g, 10%.
Experimental

m.p. 98-100°C. \(^1\)Hnmr(CDC\(_3\), 200MHz, \(\delta/\text{ppm}\)) 1.34(3H, t), 4.38(2H, q), 8.06(1H, s), 12.5(1H, s). \(^{13}\)C\(^{1}\)Hnmr(CDC\(_3\), 62MHz, \(\delta/\text{ppm}\)) 159.2 (qC), 136.5 (qC), 121.7 (CH), 61.7 (CH\(_2\)), 13.9 (CH\(_3\)).

5.6 General protocol for diisopropylcarbodiimide coupling of the C-terminal amino acid onto \(p\)-benzyloxybenzyl alcohol (Wang) resin.

The Fmoc-amino acid (5.5 mmol, 6 equivalents relative to resin loading) was activated as the symmetric anhydride by treatment with diisopropylcarbodiimide (DIC) (450 \(\mu\)l, 2.9 mmol, 3 equivalents) in DMF (5 ml). The mixture was sonicated for 10 minutes, before being added to Wang resin (1.2 g, 0.96 mmol), previously swollen in DMF (5 ml). 4-Dimethylaminopyridine (DMAP) (10 mg, 0.09 mmol,) was added and the mixture sonicated for 2 hours. The resin was then filtered and washed consecutively with DMF and 1,4-dioxane.

Determination of resin loading efficiency

A sample of the resin was dried \textit{in vacuo}, after washing with diethyl ether. The dry resin (3-4 mg) was accurately weighed into a 10 ml volumetric flask and made up to the mark with 20\% (v:v) piperidine in DMF, sonicated for 10 minutes and the UV absorbance of the supernatant measured at 302 nm. The loading of the resin was calculated using the relationship:

\[
\text{Resin Loading (mmol/g)} = \frac{10 \times \text{Absorbance}}{9 \times \text{weight of resin (mg)}}
\]

This method was used to determine both the initial manual Fmoc-amino acid loading to the resin and the final Fmoc-amino acid coupling to the growing peptide sequence which gave the overall loading of peptide on to the resin.
5.7 General protocol for acidolytic cleavage from Wang resin

The N\textsuperscript{\textalpha} Fmoc protecting group was removed by sonication of the resin in 20% piperidine in DMF for 20 min, filtering and washing consecutively with DMF, 1,4-dioxane, DCM and diethyl ether.

The dry resin-bound peptide was placed in a round bottomed flask. A cocktail of scavengers (up to 1 ml volume) was added to the resin and the mixture stirred for 30 min. The choice of scavengers is dictated by the amino acids (hence protecting groups) in the sequence; the standard mixture, employed in the cleavage of all the peptides detailed here, contained ethanedithiol (EDT) (0.35 ml), water (0.4 ml), triisopropylsilane (TIS) (0.1 ml), thioanisole (0.15 ml) and phenol (0.2 g). To this was added trifluoroacetic acid (TFA) (10 ml), and the mixture stirred under nitrogen for 2 hours at room temperature. The cleaved resin was then filtered and washed with TFA (1 x 2 ml), then cold diethyl ether (20 ml) poured into the crude peptide/TFA filtrate, causing the peptide to precipitate out. This suspension was stood in ice for 15 mins, before centrifugation and removal of the supernatant liquid. The peptide was washed a further 3 times with fresh cold ether, the supernatant being removed after centrifugation.

5.8 General protocol for the coupling of the C-terminal amino acid onto 2-chlorotritylchloride resin.\textsuperscript{2}

The amino acid (dried overnight over NaOH) (1.17 mmol) was dissolved in dry DCM (10 ml). Dry DMF (catalytic amount) was added until the amino acid had completely dissolved, before adding the solution to 2-chlorotritylchloride resin, (dried overnight over NaOH) (1 g at 1.17 mmol/g). One third of the total DIEA to be used (0.98 mmol) was added, and the reaction stirred for 5 mins, before addition of the remaining two thirds DIEA in a 1:1 mixture with dry DCM. The reaction was stirred under nitrogen for 1 hour at room temperature. Excess HPLC grade methanol (0.8 ml) was added, and the
reaction stirred for 10 mins, before filtering off the resin and washing consecutively with: DCM (3x10ml), DMF (2x10ml), isopropanol (2x10ml), DMF (2x10ml), isopropanol (2x10ml), methanol (2x10ml) and diethyl ether (2x10ml). The resin was then dried (NaOH).
The loading level of the amino acid on the resin was determined as described above for Wang resin.

5.9 Acidolytic cleavage of resin bound-peptide from 2-chlorotrityl resin.

The dry resin-bound peptide was stirred at room temperature in a solution of 2:2:6 acetic acid/trifluoroethanol/DCM (20ml/g of resin) for 2 hours. The resin was filtered and washed with the cleavage mixture (3x5ml), hexane was added (5x vol. of solvent) and the solvents removed in vacuo. The residue was washed a further 3 times with hexane, each time evaporating to dryness.

5.10 Notes on the synthesis of the peptide libraries.

All peptide libraries were synthesised on a 0.25mmol scale (except Library 5, 0.2 mmol scale) on an Applied Biosystems 430A automated peptide synthesiser, using the Fmoc orthogonal protecting strategy. Non-optimum conditions were applied, using HOBt/DIC single coupling for the synthesis of Libraries 1-4, to give peptide libraries containing a significant proportion of impurities and truncates.

The synthesis of Library 5 was found to fail under these conditions, hence HOCt active ester coupling was employed, with double coupling at all residues (except Gly, where an extended coupling time was applied). At the three sites of diversity the resin was removed from the synthesiser and coupled in the sonic bath, allowing extended coupling
time, in addition to being double coupled. Under these conditions Library 5 was successfully constructed with an overall 23% final coupling efficiency.

Diversity was introduced into the libraries in one of two ways:

1. Using simple multiple peptide synthesis, where, at the sites of diversity, the cartridge corresponding to that residue was loaded with 1mmol equivalent of different amino acids. All the libraries in this work used 3 amino acids at each site of diversity, hence the corresponding cartridges were loaded with 1/3mmol each of the 3 selected amino acids. This technique was used in the synthesis of Libraries 1, 2, 3 and 5.

2. Library 4 was produced using a split/pool strategy; whereby, at the sites of diversity the resin was removed from the automated peptide synthesiser (after deprotection of the N\textsuperscript{\textalpha} Fmoc group), divided into 3 equal portions and the 3 selected amino acids coupled separately for 30 min. in a sonic bath. After coupling, the 3 portions of resin were re-combined, filtered and washed, then returned to the automated synthesiser which continued the synthesis either to the next site of diversity, where the process was repeated, or to the end of the peptide sequence.

After completion of the synthesis the N-terminal Fmoc was removed from the libraries, Tbfmoc was loaded (see below) and the libraries cleaved from the resin using standard methods. Libraries 2-5 were then subjected to an affinity purification protocol using carbon, see section 5.12.
5.11 General protocol for the loading of tetrabenzo[a,c,g,i]fluorenyl-17-methoxy carbonyl (Tbfmoc) onto the \( N^\alpha \) of the resin-bound peptide.

The \( N^\alpha \)-Fmoc was removed by sonicating the resin bound peptide in 10% piperidine in 1,4-dioxane for 20 minutes. After washing thoroughly with 1,4-dioxane and DCM, the resin (1 g) was swollen in DCM (10ml), and TbfmocCl (3 equivalents) dissolved in DCM (10ml) added, together with DIEA (3 equivalents). The mixture was sonicated, in the absence of light, for 3 hours at room temperature. The resin was then filtered and washed with copious DCM and diethyl ether, and dried \textit{in vacuo}.

**Determination of resin loading efficiency**

A known weight of dry resin (3-6mg) was weighed into a 10ml volumetric flask and made up to the mark with 10% piperidine in 1,4-dioxane. After sonicating for 10 minutes, the UV absorption of the solution was measured at 364 nm. The loading of the Tbfmoc onto the peptide was calculated using the relationship:

\[
\text{Tbfmoc loading (mmol/g) = \frac{\text{Absorbance} \times 0.613}{\text{Weight of resin (mg)}}}
\]

5.12 Carbon purification of Tbfmoc-peptides and peptide libraries.

\textit{Using porous graphitised carbon (PGC)}

**Method 1: Used for the purification of Libraries 2 and 3.**

The crude peptide material with \( N^\alpha \) Tbfmoc (200mg), obtained from acidolytic cleavage, was dissolved in 1:1 6M guanidine hydrochloride/isopropanol (10 ml). PGC (150mg) was added, and the suspension vortexed (10 mins), until all material with a corresponding 364 nm absorbance had been adsorbed, as assessed by HPLC after
centrifugation. The supernatant was then discarded, and the PGC washed repeatedly with 1:1 6M GuHCl/iPrOH (4x20 ml) using the vortex/centrifuge technique. On obtaining a flat baseline on RP HPLC, the peptide was cleaved from the Tbfmoc (which remained adsorbed onto the PGC) by vortexing for 15 minutes in 10% piperidine in 1:1 GuHCl/iPrOH (10ml). This was repeated, then the carbon rinsed twice with neutral solvent, the supernatants combined and neutralised with glacial acetic acid to pH 5.5. The isopropanol was removed, then the mixture desalted by preparative RP HPLC. The HPLC conditions were as detailed previously, except the gradient was held at 10% acetonitrile until all the guanidine and salt had been eluted. The relevant peptide peaks were collected and lyophilised.

**Using activated charcoal**

The activated decolourising charcoal was prepared by washing repeatedly using a vortex/centrifuge technique with 10% piperidine in a suitable polar solvent system, until the supernatant ran clear (6x25ml). The charcoal was then rinsed with that same solvent system without piperidine (4x25ml), followed by TFA (10ml) for immediate use.

The solvent system was selected by cleaving a small portion of resin, carrying the peptide(s) without Tbfmoc, and establishing a suitable solvent to solubilize the crude material. 1:1 Acetonitrile/water was found to solubilize Library 4, while 3:2 acetonitrile/water was demonstrated to effectively dissolve crude Library 5 without Tbfmoc.
Method 2a: Optimised procedure used for the purification of Library 4.

Using the vortex/centrifuge technique, the library (200mg of peptide) was adsorbed onto the pre-washed charcoal (100mg) straight from the TFA cleavage solution, without isolating the crude material. The charcoal was washed with 1:1 acetonitrile/water (4x25ml) until a flat baseline was obtained on RP HPLC. The library was then cleaved using 10% piperidine in 1:1 acetonitrile/water (2x10ml). The carbon was washed twice further with 1:1 acetonitrile/water (2x10ml) to remove any remaining peptidic material, the supernatants combined and neutralised with glacial acetic acid. The acetonitrile was removed by rotary evaporation and the residual peptide solution de-salted by preparative RP HPLC, holding the gradient at 2% acetonitrile until the salt had been eluted. The relevant peaks were collected and lyophilised.

Method 2b: Optimised procedure used in the purification of Library 5.

Pre-washed charcoal was loaded into a small sintered glass column containing a layer of celite. The column was then operated by a dry flash column chromatography technique. The \(N^\alpha\)-Tbtf-moc-Library 5 was adsorbed onto the charcoal straight from the TFA cleavage by passing the TFA solution through the column 3 times. After checking the supernatant by HPLC (all 364nm material adsorbed from solution), the charcoal was washed with 3:2 acetonitrile/water (100ml) (flat baseline on RP HPLC), then the peptides cleaved with 10% piperidine in acetonitrile/water (2x20ml). The piperidine was allowed to remain in contact with the charcoal for several minutes by removing the vacuum and passing the solution repeatedly through the column. The charcoal was further washed with 3:2 acetonitrile/water (2x20ml), the washings combined with the piperidine cleavages, neutralised with glacial acetic acid and the acetonitrile removed. The peptide solution was then de-salted by preparative RP HPLC, holding the gradient at 2% acetonitrile until the salt had eluted, the relevant peaks collected and lyophilised.
5.13 *In vitro* testing of truncated big ET-1 analogues. (Performed by Parke Davis)3

A mixture containing bigET-1 (0.1μM), Hepes-KOH (pH 7.0) (100mM), NaCl (50mM), pepstatin A (50μM), leupeptin (100μM), phenylmethylsulphonyl fluoride (200μM), an indicated concentration of BigET-1 analogue (as the inhibitor, DMSO for control), and the membrane fraction of CHO/human ECE-1 cells. The final concentration of DMSO was 1.5%. After incubation for 1 hr at 37°C, the reaction was stopped by adding EDTA to give a final concentration of 10mM. This final mixture was then directly analysed for mature ET-1 by enzyme-linked immunosorbant assay.

The analogues were judged to be substrates if the hydrolysis could be completely inhibited by a specific ECE-1 inhibitor, as determined by RP HPLC analysis (*Ref. Ahn et al. BBRC 1998, 243, 184*).

5.14 Kinetic Testing of the big ET-1(16-38) analogues. (Performed by Parke Davis)3

The assays were performed as described in section 5.13. They were carried out for 1 hour and then analysed by RP HPLC (C18 column) using a linear gradient from 0-50% acetonitrile. Rates were determined by appearance of the [22-38] product. A standard plot was used to relate moles of [22-38] to peak area of the peptide. Initial rate data was analysed by Kaleidagraph (Synergy Software, Reading, PA), which fitted to the Michaelis-Menten equation hyperbola (*see chapter 2*).
Part 2: Heterocyclic Investigations into a Potential Ras FTase Inhibitor.

5.15 Techniques and instrumentation

Thin layer chromatography (TLC) was run on aluminium sheets pre-coated with silica gel (Kieselgel 60 F$_{254}$) from Merck. The compounds were visualised by ultra-violet absorption at 254nm and potassium permanganate or 0.2% bromocresol in ethanol, in the case of acid functionalities. Wet flash chromatography was performed using silica gel 60 (230-400 mesh) (from BDH). Fourier Transform Infra Red (FTIR) Spectroscopy was performed on a Bio-Rad FTS-7 spectrometer, either in the solvent indicated or prepared as a KBr disc. Proton NMR spectra were recorded on either a Gemini Varian 200, Brucker WP 200 (200 MHz) or Brucker AC 250 (250 MHz) instrument, in the deuterated solvent indicated, using tetramethylsilane (TMS) as an external standard. Carbon and fluorine NMR were recorded on either of the latter instruments. High and low resolution mass spectra were measured either by fast atom bombardment (FAB) on a Kratos MS50TC machine, or by electron impact (EI) on a Kratos 902MS. CHN elemental analysis was carried out on a Carlo Erba 1106 or Perkin Elmer 2400 instrument. All further techniques and instrumentation employed in the analysis of the work has been detailed previously. All reactions were performed under anhydrous conditions.
5.16 Preparation of 2-fluoro, 4-hydroxyl benzoic acid (10).^4

![Chemical Structure]

3-Fluorophenol (0.094mol, 1 equiv.), sodium hydroxide (0.846mol, 9 equiv.), water (33.84ml) and copper powder (2.2mmol, 0.023 equiv.) were combined in a round bottom flask and stirred mechanically. The slurry was heated to 80°C, and carbon tetrachloride (0.097mol, 1.03 equiv.) added gradually over 5 hours at such a rate as to keep the reaction at gentle reflux. Additional copper powder (0.088g) was added after 2 and 4 hours. The reaction was allowed to stir at 80°C overnight. The resulting slurry was diluted with water and filtered, before being cooled on an ice bath and acidified (conc. hydrochloric acid), causing the product to precipitate out as a pale orange solid. After recrystallising from water three times, with decolourising charcoal, (10) was obtained as a white crystalline solid (3.52g, 0.02mol, 24%).

**Rf** 0.80 (diethyl ether). **m.p.** 203-204°C (Lit. 204-205.5°C). **FTIR** $\nu_{\text{max}}$/cm$^{-1}$(nujol) 1520, 1593 (Ar); 1686 (C=O); 3469 (OH). $\lambda_{\text{max}}$/nm (methanol, $\varepsilon$/dm$^3$/mol$^1$/cm$^{-1}$) 246(13884). $^1$Hnmr(DMSO, 200MHz, $\delta$/ppm) 10.73(1H, s, OH); 7.88-7.82(1H, m, ArH); 6.84-6.72(2H, m, ArH). $^{13}$C{$^1$H}nmr (DMSO, 62MHz), $\delta$/ppm 171.44 (COOH); 166.41 (d, J=251.1Hz, ArCF); 163.33 (d, J=13.9Hz, q ArC o to F); 132.94 (d, J=11.2Hz, ArH p to F); 110.15 (d, J=2.7Hz, q ArCOH); 107.18 (d, J=22.3Hz, ArCH m to F); 103.95 (d, J=24.2Hz, ArCH o to F). $^{19}$Fnmr (DMSO), $\delta$/ppm -104.26 (ArF). **HRMS** (EI, M$^+$/z) C$_7$H$_5$FO$_3$ (M$^+$) requires 156.0223, found 156.0222.
5.17 Preparation of 1\textit{butyldimethylsilyl} (2-fluoro-4-t-butyldimethylsilyloxy) benzoate (9).\textsuperscript{5}

\begin{center}
\includegraphics[width=0.5\textwidth]{product_diagram.png}
\end{center}

Under an N\textsubscript{2} atmosphere, 2-fluoro, 4-hydroxy benzoic acid (6.41mmol, 1 equiv.) was dissolved in dry THF (12.8ml). Freshly distilled triethylamine (19.23mmol, 3 equiv.) was added, followed by t-butyldimethylsilyl chloride (12.82mmol, 2 equiv.) solution in THF (3ml), and the reaction was stirred at room temperature for two hours. The solvent was then removed, and the residue taken up in diethyl ether and washed with water, sat. NaHCO\textsubscript{3} solution, water and brine. The organic layer was dried (Na\textsubscript{2}SO\textsubscript{4}) and the solvent removed to give the product as an oil. Purification was achieved via Kugelrohr distillation under trolley pump vacuum, giving a colourless oil (1.297g, 3.4mmol, 53\%).

\textbf{Rf} 0.92 (diethyl ether). \textbf{FTIR} $\nu_{\text{max}}$/cm\textsuperscript{-1}(DCM) 1571, 1614; 1696 (C=O); 2886, 2931 (CH). $\lambda_{\text{max}}$/nm (methanol, $c$/dm\textsuperscript{3}mol\textsuperscript{-1}cm\textsuperscript{-1}) 238(12375). $^{1}$\textbf{Hnrmr(CDC1\textsubscript{3}, 200MHz, $d$/ppm)} 7.66-7.57 (1H, m, ArH); 6.44-6.30 (2H, m, ArH); 0.77 (9H, s, $^{1}$Bu); 0.74 (9H, s, $^{1}$Bu); 0.13 (6H, s, 2xCH\textsubscript{3}); 0.00 (6H, s, 2xCH\textsubscript{3}). $^{13}$\textbf{C$^{1}$H$^{1}$nrmr (CDC1\textsubscript{3}, 62MHz), $d$/ppm} 169.32 (qC=O), 169.24 (qArC), 164.41 (d, J=262.4Hz, ArC-F), 162.61 (qArC), 162.54 (2xqC), 134.3 (d, J=1.7Hz, ArCH $p$ to F), 116.42 (d, J=2.6Hz, ArCH $m$ to F), 109.04 (d, J=23.4Hz, ArCH $o$ to F), 26.00 (4xCH\textsubscript{3}), 25.91 (6x CH\textsubscript{3}). \textbf{HRMS (FAB, M$^{+}$/z)} C\textsubscript{19}H\textsubscript{33}FO\textsubscript{3}Si\textsubscript{2}(MH$^{+}$) requires 385.1953, found 385.1968.
5.18 Preparation of benzyl (2-fluoro-4-benzyloxy)benzoate (11).\(^6\)

![Structure of benzyl (2-fluoro-4-benzyloxy)benzoate (11)](image)

2-Fluoro, 4-hydroxy benzoic acid (0.10g, 6.41mmol, 1 equiv.) was dissolved in DMF (12ml), under N\(_2\). To this was added potassium carbonate (1.97g, 12.82mmol, 2 equiv.), followed by benzyl bromide (1.65g, 9.62 mmol, 1.5 equiv.), and the reaction was stirred at 60°C overnight. The resulting suspension was filtered, the solvent removed by evaporation and the residue re-dissolved in ethyl acetate. After washing with 1M HCl (3x10ml) and brine (2x10ml), the organic layer was dried (MgSO\(_4\)) and the solvent removed to give the product. Re-crystallisation from hexane gave (11) as a white crystalline solid (1.10g, 3.27mmol, 51%).

**Rf** 0.83 (1:1 diethyl ether/hexane). **FTIR** \(\nu_{\text{max}}/\text{cm}^{-1}\) (DCM) 1276.9 (C-O); 1616.9 (CH); 1715.7 (C=O); 3063.2 (ArCH). \(\lambda_{\text{max}}/\text{nm}\) (methanol, \(\varepsilon/\text{dm}^3\text{mol}^{-1}\text{cm}^{-1}\)) 252 (13230).

**\(^1\)Hnmr** (CDCl\(_3\), 200MHz, \(\delta/\text{ppm}\)) 7.97-7.88 (1H, m, ArH); 7.48-7.32 (10H, m, ArH); 6.82-6.68 (2H, m, ArH); 5.35 (2H, s, CH\(_2\)); 5.10 (2H, s, CH\(_2\)). **\(^{13}\)C\(^{\text{1}}\)Hnmr** (CDCl\(_3\), 62MHz), \(\delta/\text{ppm}\) 163.8 (qArCOOBn), 163.7 (qArCOBn), 163.6 (d, \(J=11.7\text{Hz}\), qArC \(o\) to F), 163.4 (d, \(J=260.7\text{Hz}\), qArC-F), 135.9 (qArC), 135.5 (qArC), 133.4 (d, \(J=1.8\text{Hz}\), ArCH \(p\) to F), 128.6 (2xArCH), 128.4 (2xArCH), 128.3 (ArCH), 128.0 (ArCH), 127.9 (2xArCH), 127.4 (2xArCH), 110.8 (d, \(J=2.6\text{Hz}\), ArCH \(m\) to F), 103.1 (d, \(J=26.0\text{Hz}\), ArCH \(o\) to F), 70.4 (CH\(_2\)), 66.4 (CH\(_2\)). **HRMS** (FAB, \(M^+/z\)) \(\text{C}_{21}\text{H}_{17}\text{FO}_3\) (MH\(^+\)) requires 337.1162, found 337.1150.
5.19 Preparation of 4-benzyloxy-2-fluorobenzoic acid (A).

(11) (0.94mmol, 1 equiv.) was suspended in 2:1 2M sodium hydroxide/1,4-dioxane (3ml). After stirring overnight at room temperature, the solvent was removed and the residue diluted in water. After extracting twice with diethyl ether, the aqueous layer was cooled in an ice bath and acidified with conc. hydrochloric acid, causing (A) to precipitate out as a white solid (0.15g, 0.63mmol, 67%).

Rf 0.82 (diethyl ether). FTIR ν max/cm⁻¹ (DCM) 1619.2; 1704.8 (C=O); 3414.4 (OH).
λ max/nm (methanol, ε/dm³mol⁻¹cm⁻¹) 241 (4182).

¹H nmr (CDCl₃, 200MHz, δ/ppm) 8.01–7.93 (1H, m, ArH); 7.43–7.38 (5H, m, ArH); 6.85–6.70 (2H, m, ArH); 5.11 (2H, s, CH₂).

¹³C {¹H} nmr (CDCl₃, 62MHz), δ/ppm 164.8 (qC=O), 164.8 (qArCOBn), 163.3 (d, J=11.6Hz, qArC), 162.9 (d, J=257.3, ArC-F), 136.2 (qArC), 133.5 (d, J=2.4Hz, ArCH p to F), 128.7 (2xArCH), 128.4 (ArCH), 128.1 (2xArCH), 111.5 (d, J=2.6Hz, ArCH m to F), 103.4 (d, J=26.0Hz, ArCH o to F), 70.2 (CH₂).

HRMS (FAB, M⁺/z) C₁₄H₁₁FO₃ (MH⁺) requires 247.0692; found 247.0771.
5.20 Preparation of ethyl 3-keto-3-(4-benzyloxy-2-fluorobenzene) propanoate via the Acid Chloride (2).\(^7\)

Under \(N_2\) atmosphere, 4-benzyloxy-2-fluorobenzoic acid (0.78 mmol) was heated in thionyl chloride (5 ml) for 1.5 hours. The excess thionyl chloride was then removed by distillation. (FTIR \(v_{\text{max}}/\text{cm}^{-1}\text{(liquid film)}\) 1769 (C=O))

Under \(N_2\) atmosphere, potassium ethyl malonate (1.6 mmol, 2.06 equiv.) in dry MeCN (1.5 ml) was cooled to 10-15°C, and to this was added Et\(_3\)N (1.56 mmol, 2 equiv.), followed by MgCl\(_2\) (1.95 mmol, 2.5 equiv.). Stirring was then continued at room temperature for 2.5 hours. The resulting slurry was re-cooled to 0°C, and the acid chloride in dry MeCN (1.5 ml) added dropwise over 20 mins, followed by additional Et\(_3\)N (0.17 mmol, 0.22 equiv.). The mixture was allowed to stir overnight at room temperature, after which it was concentrated and taken up in toluene (10 ml) twice. Further toluene was added (10 ml), and the mixture stirred and cooled to 10-15°C. Aqueous HCl (13%, 3 ml) was added cautiously, keeping the temperature below 25°C. The aqueous layer was separated and the organic layer washed with 13% aq HCl (2x5 ml), followed by H\(_2\)O (2x5 ml), then concentrated under vacuum to give (2) as a white solid (145.4 mg, 0.46 mmol, 59%).

Rf 0.80 (diethyl ether). FTIR \(v_{\text{max}}/\text{cm}^{-1}\text{(DCM)}\) 1616.3 (C=C), 1691.5 (C=O), 1728.8 (ester C=O), 2996.4 (CH), 3063.9 (ArCH). \(\lambda_{\text{max}}/\text{nm}\) (methanol, \(\varepsilon/\text{dm}^3\text{mol}^{-1}\text{cm}^{-1}\)) 246 (26062). \(^1\)Hnmr(CDCl\(_3\), 200 MHz, \(\delta/\text{ppm}\)) 7.99-7.92 (1H, m, ArH); 7.42-7.34 (5H, m, ArH); 6.83-6.69 (2H, m, ArH); 5.10 (2H, s, CH\(_2\)Ph); 4.24 (2H, dd, J=0.06, 0.03 Hz, CH\(_2\)Me); 3.44 (2H, s, -OC-CH\(_2\)-CO-); 1.30 (3H, t, \(^3\)J=0.03 Hz, CH\(_3\)). \(^{13}\)C\(^{\text{1H}}\)nmr (CDCl\(_3\), 62 MHz), \(\delta/\text{ppm}\) 171.06 (q ArC-OBn); 168.96 (d, J=3.7 Hz, q C=O); 167.04 (q
C=O); 164.42 (d, J=11.8Hz, q ArC o to F); 164.05 (d, J=262.1Hz, q ArC-F); 135.32 (q ArC); 134.02 (d, J=2.2Hz, ArCH p to F); 128.67 (2xArCH); 128.37 (ArCH); 127.41 (ArCH); 110.98 (d, J=2.9, ArCH m to F); 103.20 (d, J=25.7, ArCH o to F); 70.49 (CH2); 61.93 (CH2); 40.55 (CH2); 13.90 (CH3). $^{19}$Fnmr (CDCl$_3$), δ/ppm -104.84 (ArF). HRMS (FAB, M$^+$/z) C$_{18}$H$_{17}$FO$_4$ (MH$^+$) requires 317.1111, found 317.1125.

5.21 Preparation of ethyl 3-keto-3-(2-fluoro-4-hydroxybenzene) propanoate (3).

![Structural formula](image.png)

Ethyl 3-keto-3-(4-benzyloxy-2-fluorobenzene) propanoate (2) (128.8mg, 0.408mmol) was dissolved in methanol (5ml) and a catalytic amount of 10% palladium on charcoal added. The vessel was charged with H$_2$ three times, then the suspension allowed to stir under a hydrogen atmosphere for 2.5 hours. The suspension was then filtered and the solvent removed in vacuo to yield (3) (58mg, 0.26mmol, 63.4%).

Rf 0.7 (diethyl ether). FTIR $\nu_{\text{max}}$/cm$^{-1}$(DCM) 1687.4 (C=O), 1735.1 (ester C=O), 2984.1 (CH), 3041.9 (ArCH). $\lambda_{\text{max}}$/nm (methanol, $\epsilon$/dm$^3$mol$^{-1}$cm$^{-1}$) 244 (9176).

$^1$Hnmr(CDCl$_3$, 200MHz, δ/ppm) 7.99-7.92 (1H, m, ArH); 6.83-6.69 (2H, m, ArH); 4.24 (2H, dd, J=0.06, 0.03Hz, CH$_2$Me); 3.44 (2H, s, -OC-CH$_2$-CO-); 1.30 (3H, t, $^3$J=0.03Hz, CH$_3$). HRMS (FAB, M$^+$/z) C$_{11}$H$_{11}$FO$_4$ (MH$^+$) requires 227.0641, found 227.2043.

140
5.2.2 Synthesis of 2-phenoxy-2-phenyl ethanol (13) and 1-phenyl-2-phenoxy ethanol (14).

Phenol (2g, 21.3mmol) was dissolved in aqueous sodium hydroxide (0.85g, 21.3mmol in 2ml H₂O). The solution was heated to 70°C and styrene oxide (2.43ml, 21.3mmol) added dropwise over 3 min. The reaction was stirred for 4.5 hours, then poured onto ice causing a precipitate to form. After ensuring the solution was basic, the insoluble material was extracted with diethyl ether (2x20ml). The organic layers were combined, dried (sodium sulphate) and the solvent removed giving the product as a mixture of two isomers (13) and (14) (~4:1) (2.4g, 11.2mmol, 53%). Partial separation was achieved by flash column chromatography (1:1 diethyl ether/hexane).

(13) Rf 0.51 (1:1 diethyl ether/hexane). FTIR v max/cm⁻¹ (DCM) 1491.2, 1594.9 (Ph); 2359.4, 2335.2; 3061.9 (ArCH); 3590.6 (OH). λ max/nm (methanol, ε/dm³mol⁻¹cm⁻¹) 265 (1818), 271 (1817), 278 (1413). ¹Hnmr(CDCℓ₃, 200MHz, δ/ppm) 7.41-7.17 (7H-1, m, ArH), 6.95-6.87 (3H-1, m, ArH), 5.29 (1H, dd, J=8.04, 3.7Hz, CH), 3.88 (2H, sddd, CH₂), 2.54 (1H, dd, J=9.25, 4.11Hz, OH). ¹³C{¹H}nmr (CDCℓ₃, 62MHz), δ/ppm 157.58 (qArC), 137.64 (qArC), 129.25 (2xArCH), 128.57 (2xArCH), 127.98 (ArCH), 126.11 (2xArCH), 121.04 (ArCH), 115.73 (2xArCH), 80.90 (CH), 67.37 (CH₂). HRMS (FAB, M⁺/z) C₁₄H₁₄O₂ (MH⁺) requires 215.0994, found 215.1002. Elemental Analysis C₁₄H₁₄O₂ requires 78.48% C, 6.59% H; found 77.94% C, 6.54% H.

(14) Rf 0.48 (1:1 diethyl ether/hexane). FTIR v max/cm⁻¹ (DCM) 1276.9 (C-O); 1490.8, 1594.1; 2921.0 (CH); 3051.4, 3073.5 (ArH); 3590.5 (OH). λ max/nm (methanol, ε/dm³mol⁻¹cm⁻¹) 264 (1664), 271 (1902), 278 (1427). ¹Hnmr(CDCℓ₃, 200MHz, δ/ppm)
7.34-6.92 (10H, m, ArH), 4.07 (1H, dt, J=8.8, 3J=2.9, CH), 4.0 (2H, m, CH₂), 3.0 (1H, d, J=3.1Hz OH). **HRMS (EI, M⁺/z)** C₁₄H₁₄O₂ (M⁺) requires 214.0994 , found 214.0990. **Elemental Analysis** C₁₄H₁₄O₂ requires 78.48% C, 6.59% H; found 78.12% C, 6.38% H.

(3) (0.03g, 0.13mmol) was used to open styrene oxide under the conditions detailed above. The product was shown to consist solely the isomer (15) (crude yield ~25%). **Rf** 0.74 (1:1 hexane/ethyl acetate). **FTIR** \( \nu_{max}/cm^{-1}(DCM) \) 1690.4 (C=O); 1745.2 (C=O ester); 2990.3 (CH); 3064.8, 3081.7 (ArH). **¹Hnmr(CDCl₃, 200MHz, δ/ppm)** 7.99-7.91 (1H, m, ArH), 7.47-7.11 (5H, m, ArH), 6.91-6.59 (2H, m, ArH), 6.32 (1H, dd, J=7.91, 2.9Hz, CH), 4.21 (2H, dd, J=0.05, 0.03Hz, CH₂), 4.01 (2H, m, CH₂OH), 3.46 (2H, s, CH₂), 2.26 (1H, dd, J= 9.15, 4.02Hz, OH), 1.26 (3H, t, 3J=0.03Hz, CH₃). **HRMS (FAB, M⁺/z)** C₁₉H₁₉O₅F (MH⁺) requires 346.1216; found 346.1222.

### 5.23 Preparation of methyl 4-vinyl benzoate (18).

4-Vinylbenzoic acid (2g, 13.5mmol), methanol (15ml) and concentrated sulphuric acid (catalytic amount) were combined, and the mixture heated at reflux for 4 hours. The excess methanol was then removed, the residue dissolved in DCM (15ml) and extracted with water(20ml). The aqueous layer was then extracted twice with DCM (2x10ml), the organic layers combined and washed with saturated sodium bicarbonate until no further gas was evolved, then washed once with water. The organic extract was dried (MgSO₄) and the DCM removed, giving (18) as a low melting, crystalline solid (1.9g, 11.7mmol, 87%).
Rf 0.62 (1:1 diethyl ether/hexane). FTIR $\nu_{\text{max/cm}}^{-1}(\text{DCM})$ 1289.5 (C-O), 1607.9 (C=C), 1720.6 (C=O), 3063.4 (ArH). $\lambda_{\text{max/nm}}$ (DCM, $\varepsilon$/dm$^3$mol$^{-1}$cm$^{-1}$) 241 (11448).

$^1$Hnmr(CDC$l_3$, 200MHz, $\delta$/ppm) 7.94-7.90 (2H, m, ArH); 7.41-7.37 (2H, m, ArH); 6.68 (1H, dd, J=17.6, 11.0Hz, vinyl H); 5.79 (1H, d, J=17.6Hz, vinyl H); 5.31 (1H, d J=11.0Hz, vinyl H); 3.84 (3H, s, CH$_3$).

$^{13}$C$^1$Hnmr (CDC$l_3$, 62MHz), $\delta$/ppm 166.51 (C=O), 130.05 (qArC), 129.98 (qArC), 129.82 (2xArCH), 129.42 (2xArCH), 126.69 (vinyl CH), 52.33 (vinyl CH$_2$), 52.03 (CH$_3$). HRMS (EI, M$^+$/z) C$_{10}$H$_{10}$O$_2$ (M$^+$) requires 162.0681; found 162.0681.

5.24 Preparation of methyl (4-styrene oxide)benzoate (19). 9

Methyl 4-vinylbenzoate (0.56g, 3.5mmol) was added to a solution of metachloroperbenzoic acid (MCPBA) (0.59g, 3.5mmol) in chloroform (10ml). The reaction was stirred at 0°C for 24 hours, then shaken with 10% NaOH and washed with water. The chloroform solution was dried (MgSO$_4$), and the solvent removed giving (19) as a white solid (0.50g, 2.84mmol, 81%).

Rf 0.64 (1:1 diethyl ether/hexane). FTIR $\nu_{\text{max/cm}}^{-1}(\text{DCM})$ 1722 (C=O); 2304; 2953, 2987 (CH); 3053 (ArCH). $\lambda_{\text{max/nm}}$ (DCM, $\varepsilon$/dm$^3$mol$^{-1}$cm$^{-1}$) 240 (14351).

$^1$Hnmr(CDC$l_3$, 200MHz, $\delta$/ppm) 7.96-7.90 (2H, m, ArH); 7.41-7.19 (2H, m, ArH); 3.84 (4H, s, CH$_3$ + epoxide H); 3.10 (1H, dd, J=5.5, 4.2Hz, epoxide H); 2.71 (1H, dd, J=5.5, 2.6Hz, epoxide H). $^{13}$C$^1$Hnmr(DMSO, 62MHz), $\delta$/ppm 166.65 (C=O), 129.93 (qArC), 129.83 (qArC), 129.67 (2xArCH), 125.27 (2xArCH), 52.02 (CH$_3$), 51.80 (epoxide CH), 51.32 (epoxide CH$_2$). HRMS (EI, M$^+$/z) C$_{10}$H$_{10}$O$_3$ (M$^+$) requires 178.0630, found 178.0629.
5.25 Preparation of 2-imidazol-1-yl-1-phenyl-ethanol (20).\(^\text{10}\)

![Chemical Structure](image)

Imidazole (2g, 0.029mol) and freshly distilled styrene oxide (3.4ml, 0.029mol) in DMF (20ml), were heated at 90°C with stirring under nitrogen for 50 hours. After cooling to room temperature, the mixture was diluted with water (300ml) and extracted with diethyl ether (2x200ml). The organic layers were combined, dried (Na\(_2\)SO\(_4\)) and the solvent evaporated, affording the crude product, which was recrystallised from diethyl ether to give a beige crystalline solid. Yield 75%.

**Rf** 0.69 (1:1 DCM/methanol). **m.p.** 146°C (lit. °C). **FTIR** \(\nu_{\text{max}}/\text{cm}^{-1}\) (bromofom) 877.6 (C=O); 1389.4 (OH bending); 1444.9;1709.2 (C=O);3386.9 (N-H). \(\lambda_{\text{max}}/\text{nm}\) (methanol, \(\epsilon/\text{dm}^3\text{mol}^{-1}\text{cm}^{-1}\)) 252 (193), 257 (217), 263 (169). \(^1\text{H}n\text{mr}(\text{DMSO}, 200\text{MHz}, \delta/\text{ppm})\)

7.57(1H, s, ImCH); 7.43-7.35(5H, m, ArH); 7.15(1H, t, \(^3\)J=1.23Hz, ImCH); 6.98(1H, t, \(^3\)J=1.06Hz, ImCH); 5.00(1H, m, CH); 4.26(2H, m, CH\(_2\)). \(^{13}\text{C}\{^1\text{H}\}n\text{mr}(\text{DMSO}, 62\text{MHz}), \delta/\text{ppm}\) 141.07(qArC); 137.10(CH); 127.54(2xCH); 127.00(CH); 126.45(CH); 125.18(2xCH); 199.59(CH); 72.30(CH); 53.38(CH\(_2\)). **HRMS**(FAB, \(M^+/z\)) C\(_{11}\)H\(_{12}\)N\(_2\)O (MH\(^+\)) requires 189.0950, found 189.0958. **Elemental Analysis** C\(_{11}\)H\(_{12}\)N\(_2\)O requires 70.19% C, 6.43% H, 14.88% N; found 70.11% C, 6.26% H, 14.85% N.
5.26 Oxidation of toluene.

Method 1

Water (9.2ml), toluene (2.5g, 0.0275mol) and KMnO₄ (4.59g, 0.029mol, 1.06 equiv.) were heated at gentle reflux until the colour had disappeared (1h.). A second portion of KMnO₄ (4.59g, 0.029mol, 1.06 equiv.) was added with water (2ml), and the heating continued until the purple colour was destroyed (2.5h.). The reaction was allowed to cool slightly, and the brown manganese dioxide precipitate filtered off and washed with hot water. The filtrate was concentrated under reduced pressure and acidified to pH3 with conc. HCl, causing the benzoic acid product to crash out as a white solid. It was recrystallised from ethanol, with hot filtration to remove the salts (yield 86%).

m.p. 121-122°C (lit. 122.6-123.1°C). MS(FAB, M⁺/z) C₇H₆O₂ (MH⁺) requires 122, found 122.

Method 2

Toluene (2.5g, 0.088mol) and KMnO₄ (13.33g, 0.264mol, 3.07 equiv.) were added to a solution of NaOH (22g, 1.76mol, 20 equiv.) in water (330ml). The mixture was heated under reflux at 100°C with stirring for 15mins, then the slurry allowed to cool and ethanol (5ml) added, to react with the excess KMnO₄. The manganese dioxide was filtered off with suction through a plug of celite, giving a colourless solution. The filtrate was allowed to cool in an ice bath and carefully acidified with conc. HCl causing the benzoic acid to precipitate out. It was recrystallised from ethanol, with hot filtration to remove the salts (yield 23%).

m.p. 118-120.5°C (lit. 122.6-123.1°C). MS(FAB, M⁺/z) C₇H₆O₂ (MH⁺) requires 122, found 122.
**Method 3\textsuperscript{13}**

A RB flask, fitted with a mechanical stirrer, was charged with KOH pellets (25.1g, 0.46mol) and water (5.2ml). When the slurry had cooled, toluene (2.5g, 0.0275mol) was added and stirring commenced. Lead dioxide (25.12g, 0.105mol) was added and the mixture heated \textit{via} a graphite bath on a hot plate. The temperature was maintained at 240\textdegree C, and after several minutes the lumpy brown mass rapidly formed a bright orange melt. After a further 15 minutes at this temperature, the melt was briefly heated to 250\textdegree C before being allowed to cool slightly. After a few minutes the liquid was poured into a large beaker, forming a thin film on the walls of the beaker. When the material had cooled to room temperature, water (100ml) was poured into the beaker and the contents stirred for 1h. The suspension was filtered under suction and the solids washed with water, the filtrate partially neutralized with conc. HCl (15ml) and sufficient sodium sulfide added to precipitate out the lead. The suspension was heated to coagulate the lead sulphide, then allowed to cool slightly and filtered under suction. The filtrate was placed in an ice bath and acidified with conc. HCl (15ml). The product precipitated out as a cream solid which was filtered off (crude yield 81%). The quality of the product achieved was poor and would require considerable purification, hence the KMnO\textsubscript{4} oxidation method \textsuperscript{1} was preferred.

\textbf{m.p.} 112-115\textdegree C (lit. 122.6-123.1\textdegree C). \textbf{MS(FAB, M\textsuperscript{+}/z)} C\textsubscript{7}H\textsubscript{6}O\textsubscript{2} (MH\textsuperscript{+}) requires 122, found 122.
5.27 Preparation of acetic acid 3-fluoro-4-methyl-phenyl ester (22).^{14}

To 3-methyl-4-fluorophenol (5.0g, 39.6 mmol) in dry THF (150ml) at room temperature was added triethylamine (6.18ml, 44.6mmol, 1.126 equiv), forming a yellow solution. Acetic anhydride (8.56ml, 90.58mmol, 2.29 equiv.) and DMAP (10mol%) were stirred in dry THF (15ml) for 15 minutes, before slow addition to the phenolate via a cannula. Stirring was continued for 90 minutes, then the solution quenched with 1M HCl and the organic products extracted with ether (2x20ml). The combined organic layers were washed with sodium bicarbonate solution and dried (Na₂SO₄). The solvent was then evaporated, giving the product as a yellow liquid (6.01g, 90.4%).

\[
\text{Rf} \quad 0.80 \ (1:1 \ \text{diethyl ether/hexane}) \quad \text{FTIR} \quad \nu_{\text{max}}/\text{cm}^{-1}(\text{DCM}) \quad 1195.5, \ 1370.3, \ 1494.9, \ 1756.3 \ (\text{C}=\text{O}), \ \lambda_{\text{max}}/\text{nm} \quad \text{(methanol, } \varepsilon/\text{dm}^3\text{mol}^{-1}\text{cm}^{-1}) \quad 266 \ (1581), \ 272 \ (1482).
\]

\[
\text{\textsuperscript{1}Hnmr(CDCI₃, 200MHz, } \delta/\text{ppm} \quad 7.04-6.82 \ (3\text{H, m, ArH}), \ 2.27 \ (3\text{H, s, CH}_3), \ 2.26 \ (3\text{H, s, CH}_3).
\]

\[
\text{\textsuperscript{13}C{'H}nmr(CDCI₃, 62MHz), } \delta/\text{ppm} \quad 169.28 \ (q\text{C}=\text{O}), \ 158.51 \ (d, \ J=243.0, \ \text{ArC-F}), \ 145.91 \ (d, \ J=2.74Hz, \ q\text{ArC p to F}), \ 188.62 \ (d, \ J=19.4Hz, \ q\text{ArC o to F}), \ 123.98 \ (d, \ J=5.5Hz, \ \text{ArCH m to F}), \ 119.82 \ (d, \ J=8.8Hz, \ \text{ArCH m to F}), \ 115.21 \ (d, \ J=24.9Hz, \ \text{ArCH o to F}), \ 20.59 \ (\text{CH}_3), \ 14.23 \ (d, \ J=3.1Hz, \ \text{ArCH}_3), \ \text{\textsuperscript{19}Fnmr (CDCI₃, } \delta/\text{ppm} - 121.53 \ (\text{ArF}). \ \text{HRMS(FAB, } M^{+}/z) \quad \text{C}_9\text{H}_9\text{O}_2\text{F (MH}^{+}) \quad \text{requires} \ 169.0587; \ \text{found} \ 169.0601.
\]
5.28 Preparation of acetic acid 4-bromomethyl-3-fluoro-phenyl ester (23).\textsuperscript{15}

Acetic acid 3-fluoro-4-methyl-phenyl ester (2.0g, 11.9mmol), N-bromosuccinimide (2.12g, 11.9mmol) and a catalytic amount of azobisisobutyronitrile were combined in DCM (20m1) and heated at reflux for 1 hour under irradiation with a 500W bulb. The solution was then filtered, and the filtrate washed with aqueous sodium carbonate, dried (Na\textsubscript{2}SO\textsubscript{4}) and the solvent evaporated giving the crude product as a yellow oil (2.18g, 75.0%).

Rf 0.72 (1:1 diethyl ether/hexane). FTIR v\textsubscript{max}/cm\textsuperscript{-1}(DCM) 1196.5 (C-O), 1370.4, 1497.9, 1761.5 (C=O), 2985.0 (CH). λ\textsubscript{max}/nm (methanol, ε/dm\textsuperscript{3}mol\textsuperscript{-1}cm\textsuperscript{-1}) 272 (1647).

\textsuperscript{1}Hnmr(CDC\textsubscript{3}, 200MHz, δ/ppm) 7.16-7.01 (3H, m, ArH), 4.47 (2H, s, CH\textsubscript{2}), 2.28 (3H, s, CH\textsubscript{3}). \textsuperscript{13}C\textsuperscript{(1}H)nmr(CDC\textsubscript{3}, 62MHz), δ/ppm 169.41 (qCO), 157.75 (d, J=248.3Hz, ArC-F), 146.24 (d, J=3.0Hz, ArC p to F), 125.86 (d, J=16.5Hz, qArC o to F), 123.87 (d, J=3.3Hz, ArCH m to F), 123.28 (d, J=8.5Hz, ArCH m to F), 116.29 (d, J=15.2Hz, ArCH o to F), 24.75 (d, J=4.2Hz, CH\textsubscript{2}), 20.76 (CH\textsubscript{3}). \textsuperscript{19}Fnmr (CDCl\textsubscript{3}), δ/ppm -120.85 (ArF).

HRMS(FAB, M\textsuperscript{+}/z) C\textsubscript{9}H\textsubscript{8}O\textsubscript{2}BrF (MH\textsuperscript{+}) requires 246.9692; found 246.9770.

5.29 Preparation of 4-acetoxy-2-fluoro-benzoic acid (24).\textsuperscript{11}

Acetic acid 4-bromomethyl-3-fluoro-phenyl ester (1.13g, 4.05mmol), potassium permanganate (0.68g, 4.28mmol, 1.06 equiv.) and water (20ml) were stirred at reflux until the purple colour had disappeared (1 hour). A second portion of potassium permanganate (1.06 equiv.) was added with water (4ml), and heating continued until the colour was again destroyed (2.5 hours). The reaction mixture was allowed to cool slightly, then the precipitate filtered off and washed thoroughly with hot water. The
filtrate was concentrated, then acidified to pH3 with conc. HCl. The solution was evaporated to dryness, and the solid residue re-dissolved in methanol. The remaining solid (KCl) was filtered off, and the filtrate evaporated to dryness, giving the product (crude yield 0.1g, 13%).

Rf 0.2 (1:1 DCM/methanol). FTIR ν max/cm⁻¹(bromoform) 1455.8, 1622.2, 1704.4 (C=O). λ max/nm (methanol, ε/dm³mol⁻¹cm⁻¹) 290 (2860).¹Hnmr(CD₃OD, 200MHz, δ/ppm) 7.25-7.08 (3H, m, ArH), 2.29 (3H, s, CH₃). HRMS(FAB, M⁺/z) C₉H₇O₄F (MH⁺) requires 199.0328; found 199.0313.

5.30 Preparation of 3-methylaminophenol (25).¹⁶

![Chemical Structure](image)

3-Aminophenol (8.05g, 74mmol) and methyl iodide (10.2g, 73mmol) were heated in DMF (40ml) with potassium carbonate (10g, 0.74mmol) for 2 hours at 100°C under nitrogen. The reaction mixture was then evaporated, the residue diluted with water (20ml) and extracted with diethyl ether (2x20ml). The organic extracts were combined, washed with sat. sodium bicarbonate solution and water, dried (Na₂SO₄) and the solvent evaporated. Partial purification was achieved by Kugelrohr distillation (b.pt. 144°C at 0.09 mbar). Unfortunately, the monomethylated product was contaminated with the dimethylated species, hampering analysis of the product (combined yield 86%).

Rf 0.79 (dimethylated (26)), 0.70 (monomethylated (25)) (diethyl ether). FTIR ν max/cm⁻¹(DCM) 1152.4, 1334.3, 1514.3, 1620.9, 2971.9 (CH), 3053.7 (ArH), 3584.4 (OH). λ max/nm (methanol) 250, 293 (unable to calculate ε due to unknown weight).¹Hnmr(CDC¹³, 200MHz, δ/ppm) 7.10-7.00(m, ArH); 6.40-6.08(m, ArH); 5.31(s, ArNH); 2.87(s, CH₃); 2.74(s,CH₃). HRMS(EI, M⁺/z) (25) C₇H₉NO (M⁺) requires 123.0684, found 123.0684; (26) C₈H₁₁NO (M⁺) requires 137.0841, found 137.0842.
The same reaction conditions were employed for the methylation of (34) on Wang resin, using 4 equivalents of methyl iodide and base. After 3 hours reaction the resin was filtered off and washed consecutively with water, DMF, DCM and diethyl ether. Confirmation of the reaction was made by cleavage and cyclisation of the product, see section 5.37.

5.31 Preparation of 3-[(3-hydroxyphenyl)methylamino]-propionic acid (27).\(^{17}\)

A solution of crude 3-methylaminophenol (25) (12.5g, 0.1mol) and acrylic acid (7.2g, 0.1mol) in dry THF (50ml) was stirred with heating at reflux for 16 hours. The THF was then removed in vacuo, yielding a red oil (crude yield 89%), which was used directly in the next step. Again, analysis was hampered by difficulty in purification.

\(R_f\) 0.57 (diethyl ether). FTIR \(\nu_{\max}/\text{cm}^{-1}(\text{bromoform})\) 1168.3, 1238.4, 1613.9, 1714.4 (C=O). \(\lambda_{\max}/\text{nm (methanol)}\) 257, 296 (unable to calculate \(\epsilon\) due to unknown weight).

\(^1\)Hnmr(CD\(_3\)OD, 200MHz, \(\delta/\text{ppm}\)) 7.03-6.95(m, ArH); 6.39-6.12(m, ArH); 3.60(t, \(J=7.25\text{Hz}, \text{CH}_2\)); 2.87(s, CH\(_3\)); 2.85(s, CH\(_3\)); 2.50(t, \(J=7.25\text{Hz}, \text{CH}_2\)). HRMS(EI, \(M^+/	ext{z}\)) \(C_{10}H_{13}O_3N\) requires 195.0895, found 195.0897.
5.32 Preparation of 7-hydroxy-1-methyl-2,3-dihydro-1H-quinolin-4-one (28).\textsuperscript{18}

![Chemical Structure](image)

3-[(3-Hydroxy-phenyl)-methyl-amino]-propionic acid (27) (2.02g, 0.011mol) and polyphosphoric acid (44.3g, 0.11mol, 10 equiv.) were combined with manual stirring. The viscous mixture was gradually heated, then the temperature maintained at 100°C until a deep red colour appeared (10 minutes). The reaction was then allowed to cool slightly before being poured into a beaker of ice/water and stirred for an hour. The product was extracted into diethyl ether (3x 50ml), the organic extracts combined and dried (Na\textsubscript{2}SO\textsubscript{4}). Removal of the solvents under reduced pressure afforded a viscous dark yellow oil. Purification was achieved by silica chromatography (ethyl acetate/hexane, 3:2) giving the product as a yellow solid (overall yield for 3 steps 25-28 10.4%).

\textbf{Rf} 0.49 (diethyl ether). \textbf{FTIR} \nu_{max}/cm\textsuperscript{-1}(DCM) 1587.3, 1624.5, 1665.5, 1730.5 (C=O), 2068.1, 2222.1, 2234.3, 2674.8. \lambda_{max}/nm (methanol, \varepsilon/dm\textsuperscript{3}mol\textsuperscript{-1}cm\textsuperscript{-1}) 236 (17307), 253 (21043), 284 (10227), 379 (4917). \textbf{\textsuperscript{1}Hnmr}(CD\textsubscript{3}OD, 200MHz, \delta/ppm) 7.68(1H, d, J=8.79Hz, ArH); 6.19(1H, dd, J=8.79, 2.20Hz, ArH); 6.11(1H, d, J=2.20Hz, ArH); 3.41(2H, t, J=7.33Hz, CH\textsubscript{2}); 2.93(3H, s, CH\textsubscript{3}); 2.60(2H, t, J=7.32Hz, CH\textsubscript{2}). \textbf{\textsuperscript{13}C{\textsubscript{\textsuperscript{1}}H}nmr}(CDCl\textsubscript{3}, 62MHz), \delta/ppm 194.09 (qC=O), 165.20 (qArC), 156.03 (qArC), 137.17 (qArC), 130.37 (ArCH), 106.65 (ArCH), 98.28 (ArCH), 51.32 (CH\textsubscript{2}), 38.40 (CH\textsubscript{3}), 37.70 (CH\textsubscript{2}). \textbf{HRMS}(EI, M\textsuperscript{+}/z) C\textsubscript{10}H\textsubscript{11}NO\textsubscript{2} (MH\textsuperscript{+}) requires 177.0790, found 177.0792. \textbf{Elemental Analysis} C\textsubscript{10}H\textsubscript{11}NO\textsubscript{2} requires 67.78% C, 6.26% H, 7.90% N; found 67.57% C, 6.04% H, 7.62% N.
5.33 Preparation of 7-(2-imidazole-1-yl-1-phenyl-ethoxy)-1-methyl-2,3-dihydro-1H-quinolin-4-one (29). The alcohol (20) (0.076g, 0.41mmol) and triphenylphosphine (0.161g, 0.615 mmol) were combined in THF (2ml), under an N₂ atmosphere, and cooled in an ice bath. A solution of the 2,4-dihydroquinolone (28) (0.072g, 0.41mmol) and DEAD (97µl, 0.615mmol) in THF, was added dropwise over 1 hour. The resulting solution was allowed to stir at room temperature for 2 days. The THF was then removed in vacuo, and the residue dissolved in ethyl acetate. This was washed with water, sat. sodium bicarbonate solution and brine, dried (MgSO₄) and the solvent evaporated. Partial separation of the product was achieved by silica chromatography (5% methanol in DCM), unfortunately the product was co-eluted with triphenylphosphine oxide ((O)PPh₃). This made determination of yield and analysis of the product difficult. Rf 0.20 (2% methanol in DCM). FTIR νmax/cm⁻¹ (DCM) 1120.4; 1610.0, 1665.2 (conjugated C=C); 2069.3; 2541.9; 2676.7. λmax/nm (methanol) 273, 279, 370 (unable to calculate ε due to unknown weight). ¹Hnmr(CD₃OD, 200MHz, δ/ppm) 7.70-6.92 (9-H, m, Im+ArH); 6.26 (1H, dd, J=8.72, 2.30Hz, ArH); 6.11 (1H, d, J=2.28Hz, ArH); ~3.38 (m, CH₂); 2.78 (3H, s, CH₃); ~2.52 (2H, m, CH₂). ¹³C{¹H}nmr(CD₃OD, 62MHz), δ/ppm 194.05 (qC=O), 164.16 (qArC), 155.28 (qArC), 138.26 (qArC), 137.94 (qArC), 52.84 (CH₂), 51.17 (CH₂), 38.31 (CH₃), 37.61 (CH₂). Contamination with Ph₃PO made complete assignment difficult, several possibilities existed for the remaining CH groups. HRMS(FAB, M⁺/z) C₂₁H₂₁N₃O₂ (MH⁺) requires 348.1634, found 348.1648.
Conditions for the Mitsunobu reaction described above were optimised by the reaction of styrene oxide with (20), to give (30). Again, the product co-eluted with POPh₃, making analysis difficult.

\[ \text{Rf 0.54 (DCM). FTIR } \nu_{\text{max/ cm}^{-1}}(\text{DCM}) \ 1511.9, 1657.3, 26.65.7, 3105.6, 3372.9. \]

\[ ^1\text{Hnmr}(\text{CD}_3\text{OD, 200MHz, } \delta/\text{ppm}) \ 7.57-6.89 \ (13\text{H, m, ArH}), 5.14 \ (1\text{H, m, CH}), 4.46 \ (2\text{H, m, CH}_2). \]

\[ ^{13}\text{C}^1\text{H}nmr \ (\text{CD}_3\text{OD, 62MHz}, \delta/\text{ppm}) \ 171.90(\text{ArC-OH}); \]

5.34 Preparation of benzyl 3-(3-hydroxy-phenylamino)-propanoate.²⁰

3-Aminophenol (10g, 0.09mol), benzyl acrylate (14.6g, 0.09mol) and glacial acetic acid (5ml) were combined in THF (20ml) under nitrogen. The solution was heated at reflux for 2 days until a deep red colour was observed. The THF was removed and the residue dissolved in diethyl ether. This was washed consecutively with water, 5% sodium bicarbonate solution and water, dried (MgSO₄) and the solvent removed giving a solid residue. This was recrystallised twice from toluene, giving the product as a white crystalline solid (52%).

\[ \text{Rf 0.82 (diethyl ether). FTIR } \nu_{\text{max/ cm}^{-1}}(\text{DCM}) \ 1258.1, 1513.6, 1621.5, 1732.2 \ (\text{C} = \text{O}), \]

\[ 3421.8 \ (\text{OH}). \lambda_{\text{max/ nm}}(\text{methanol, } \varepsilon/\text{dm}^3\text{mol}^{-1}\text{cm}^{-1}) \ 208 \ (33423), 248 \ (8431), 290 \ (2409). \]

\[ ^1\text{Hnmr}(\text{CD}_3\text{OD, 200MHz, } \delta/\text{ppm}) \ 7.43-7.38(5\text{H, m, ArH}); \]

\[ 7.04-6.97(1\text{H, m, ArH}); \]

\[ 6.25-6.20(3\text{H, m, ArH}); \]

\[ 5.20(2\text{H, s, CH}_2); \]

\[ 3.45(2\text{H, t, } ^3\text{J}=6.77\text{Hz, CH}_2); \]

\[ 2.71(2\text{H, t, } ^3\text{J}=6.77\text{Hz, CH}_2). \]

\[ ^{13}\text{C}^1\text{H}nmr \ (\text{CD}_3\text{OD, 62MHz}, \delta/\text{ppm}) \ 171.90(\text{ArC-OH}); \]
157.38(qArC-N); 149.06(qC=O); 135.63(qArC); 128.98(ArCH); 127.61(2xArCH); 127.25(3xArCH); 104.41(ArCH); 103.72(ArCH); 99.16(ArCH); 65.41(CH2); 38.72(CH2); 33.10(CH2). \text{HRMS(FAB, M^+}/z) \text{C}_{16}H_{17}NO_3 (MH^+) \text{requires} 272.1208, \text{found} 272.1219. \text{Elemental Analysis} \text{C}_{16}H_{17}NO_3 \text{requires} 70.83\% \text{C}, 6.32\% \text{H}, 5.16\% \text{N}; \text{found} 70.58\% \text{C}, 6.06\% \text{H}, 5.13\% \text{N}.

These same conditions were used to load 3-aminophenol onto both REM (31) and Wang resin (33), using a 10 fold excess of reagent. The minimum volume of THF was used to swell the resin, in a 1:1 solution with acetic acid. The resin was then filtered off, and washed with DCM and diethyl ether.

\text{FTIR \nu_{max}/cm^{-1}(DCM)} (31) 1726.1(C=O), (33) 1731.9(C=O). \text{Elemental Analysis (33)} \%N 0.35\pm0.3.

5.35 Attachment of acrylic Acid to Wang Resin (32).

Wang resin (1.0g at 0.56mmol/g, 0.56mmol) was swollen in DMF (5ml) and a catalytic amount of DMAP added. DIC (0.106g, 1.68mmol) was added to acrylic acid (0.242g, 3.36mmol) in DMF (5ml), and the solution sonicated for 15mins. This mixture was then poured into the Wang resin suspension and sonicated for 3 hours. The resin was then filtered off, and washed consecutively with DMF, DCM and diethyl ether.

\text{FTIR \nu_{max}/cm^{-1}(DCM)} 1654.8(C=C), 1721.8(C=O).
5.36 Cleavage and cyclisation of 7-hydroxy-1-methyl-2,3-dihydro-1H-quinolin-4-one (28) from Wang resin.\textsuperscript{18}

The dry resin (1g, presuming 0.56mmol/g) was stirred in 30% TFA in DCM (10ml) for 40 minutes at room temperature. The resin was then removed, and the solvent evaporated \textit{in vacuo}. The residue was treated with PPA (1g), and the mixture heated with stirring until the temperature reached 100°C. The temperature was maintained at 100°C until a deep red colour appeared (10 minutes), then the reaction allowed to cool slightly before being added to ice/water and stirred for 1 hour. The product was extracted with diethyl ether (3x 20ml), the organic extracts combined and dried (Na$_2$SO$_4$), and the solvent removed under reduced pressure, affording a yellow oil (crude yield 0.04g, 40%).

The analysis was concordant with that of the solution phase synthesis product.

\textbf{Rf} 0.50 (diethyl ether). \textbf{FTIR $\nu$ max/cm$^{-1}$ (DCM)} 1589.4, 1669.5, 1732.8 (C=O), 2070.4, 2676.1. $\lambda_{max}$/nm (methanol, $\varepsilon$/dm$^3$mol$^{-1}$cm$^{-1}$) 233 (17707), 261 (19887), 280 (10107), 382 (4942). $^1$Hnmr(CD$_3$OD, 200MHz, $\delta$/ppm) 7.68(1H, d, $J=8.79$Hz, ArH); 6.19(1H, dd, $J=8.79$, 2.20Hz, ArH); 6.11(1H, d, $J=2.20$Hz, ArH); 3.41(2H, t, $^3J=7.33$Hz, CH$_2$); 2.93(3H, s, CH$_3$); 2.60(2H, t, $^3J=7.32$Hz, CH$_2$). \textbf{HRMS(EI, $M^+$/z)} C$_{10}$H$_{11}$NO$_2$ (MH$^+$) requires 177.0790, found 177.0793.
References

Appendices

Appendix I: MALDI ToF mass spectroscopy results for truncated big ET-1 analogues.

<table>
<thead>
<tr>
<th>PEPTIDE</th>
<th>PREDICTED VALUE</th>
<th>MASS DETERMINED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big ET-1[16-34]</td>
<td>2158.11</td>
<td>2160.73</td>
</tr>
<tr>
<td>Big ET-1[17-34]</td>
<td>2022.07</td>
<td>2023.57</td>
</tr>
<tr>
<td>Big ET-1[18-34]</td>
<td>1910.18</td>
<td>1908.86</td>
</tr>
<tr>
<td>Big ET-1[16-27]</td>
<td>1473.75</td>
<td>1475.50</td>
</tr>
<tr>
<td>Big ET-1[16-38, Cys22]</td>
<td>2590.29</td>
<td>2591.55</td>
</tr>
<tr>
<td>Big ET-1[16-32, Cys21]</td>
<td>1905.93</td>
<td>1907.21</td>
</tr>
<tr>
<td>Big ET-1[16-32, Cys21,22]</td>
<td>1906.86</td>
<td>1908.46</td>
</tr>
<tr>
<td>Big ET-1[16-27, Cys22]</td>
<td>1477.69</td>
<td>1476.81</td>
</tr>
<tr>
<td>Big ET-1[16-38, D-Val22]</td>
<td>2586.35</td>
<td>2586.48</td>
</tr>
</tbody>
</table>

MALDI ToF mass spectroscopy results from SAR investigation of ECE-1.
<table>
<thead>
<tr>
<th>PEPTIDE</th>
<th>PREDICTED VALUE</th>
<th>MASS DETERMINED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big ET-1[16-38, Phe^{21}]</td>
<td>2547.32</td>
<td>2546.53</td>
</tr>
<tr>
<td>Big ET-1[16-38, Leu^{21}]</td>
<td>2523.33</td>
<td>2521.56</td>
</tr>
<tr>
<td>Big ET-1[16-38, Ala^{21}]</td>
<td>2471.30</td>
<td>2470.38</td>
</tr>
<tr>
<td>Big ET-1[16-38, Pro^{21}]</td>
<td>2497.30</td>
<td>2495.35</td>
</tr>
<tr>
<td>Big ET-1[16-38, Glu^{21}]</td>
<td>2529.29</td>
<td>2541.04</td>
</tr>
<tr>
<td>Big ET-1[16-38, Arg^{21}]</td>
<td>2556.35</td>
<td>2554.68</td>
</tr>
<tr>
<td>Big ET-1[16-38, Phe^{22}]</td>
<td>2634.33</td>
<td>2637.33</td>
</tr>
<tr>
<td>Big ET-1[16-38, Leu^{22}]</td>
<td>2600.34</td>
<td>2602.18</td>
</tr>
<tr>
<td>Big ET-1[16-38, Ala^{22}]</td>
<td>2555.31</td>
<td>2551.85</td>
</tr>
<tr>
<td>Big ET-1[16-38, Pro^{22}]</td>
<td>2584.31</td>
<td>2588.4</td>
</tr>
<tr>
<td>Big ET-1[16-38, Glu^{22}]</td>
<td>2616.30</td>
<td>2618.36</td>
</tr>
<tr>
<td>Big ET-1[16-38, Arg^{22}]</td>
<td>2643.36</td>
<td>2645.27</td>
</tr>
</tbody>
</table>

*MALDI ToF mass spectroscopy results from subsite specificity study of ECE-1*
Appendix II: Amino acid analysis results for truncated big ET-1 analogues.

<table>
<thead>
<tr>
<th>PEPTIDE</th>
<th>His</th>
<th>Leu</th>
<th>Asp</th>
<th>Ile</th>
<th>Val</th>
<th>Thr</th>
<th>Pro</th>
<th>Glu</th>
<th>Tyr</th>
<th>Gly</th>
<th>Ser</th>
<th>Arg</th>
<th>Cys</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big ET-1[16-34]</td>
<td>2/2.0</td>
<td>2/1.8</td>
<td>2/1.9</td>
<td>2/1.8</td>
<td>3/2.5</td>
<td>1/0.8</td>
<td>2/1.8</td>
<td>1/0.9</td>
<td>1/1.0</td>
<td>2/2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Big ET-1[17-34]</td>
<td>1/1.1</td>
<td>2/1.9</td>
<td>2/1.8</td>
<td>2/1.7</td>
<td>3/2.6</td>
<td>1/0.9</td>
<td>2/1.6</td>
<td>1/0.8</td>
<td>1/1.1</td>
<td>2/2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Big ET-1[18-34]</td>
<td>1/0.9</td>
<td>1/1.0</td>
<td>2/1.9</td>
<td>2/1.7</td>
<td>3/2.5</td>
<td>1/0.9</td>
<td>2/1.6</td>
<td>1/1.1</td>
<td>1/1.0</td>
<td>2/2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Big ET-1[16-32]</td>
<td>2/1.9</td>
<td>1/1.0</td>
<td>2/2.1</td>
<td>2/1.8</td>
<td>3/2.8</td>
<td>1/0.7</td>
<td>2/2.4</td>
<td>1/0.8</td>
<td>1/0.9</td>
<td>1/1.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Big ET-1[16-27]</td>
<td>2/1.8</td>
<td>1/0.8</td>
<td>2/1.9</td>
<td>2/1.9</td>
<td>1/0.8</td>
<td>1/1.0</td>
<td>1/0.9</td>
<td>1/0.9</td>
<td>1/1.0</td>
<td>2/2.0</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Big ET-1[16-38,Cys]</td>
<td>2/2.0</td>
<td>2/2.2</td>
<td>2/2.1</td>
<td>2/1.8</td>
<td>2/1.9</td>
<td>1/0.8</td>
<td>3/2.6</td>
<td>1/0.9</td>
<td>1/0.9</td>
<td>2/2.0</td>
<td>2/1.7</td>
<td>1/1.1</td>
<td>1/0.8</td>
</tr>
<tr>
<td>Big ET-1[16-32,Cys]</td>
<td>2/2.2</td>
<td>1/0.9</td>
<td>2/2.2</td>
<td>2/1.7</td>
<td>2/1.8</td>
<td>1/0.9</td>
<td>2/2.2</td>
<td>1/0.8</td>
<td>1/1.1</td>
<td>1/1.0</td>
<td>-</td>
<td>-</td>
<td>1/0.7</td>
</tr>
<tr>
<td>Big ET-1[16-32,Cys2]</td>
<td>2/2.0</td>
<td>1/1.0</td>
<td>2/2.2</td>
<td>2/1.8</td>
<td>3/2.7</td>
<td>1/0.9</td>
<td>2/1.8</td>
<td>1/0.8</td>
<td>1/1.1</td>
<td>1/1.0</td>
<td>-</td>
<td>-</td>
<td>1/1.8</td>
</tr>
<tr>
<td>Big ET-1[16-32, Cys21]</td>
<td>2/2.2</td>
<td>1/0.8</td>
<td>2/2.4</td>
<td>2/1.6</td>
<td>2/1.8</td>
<td>1/1.0</td>
<td>2/1.7</td>
<td>1/0.9</td>
<td>1/0.8</td>
<td>1/1.0</td>
<td>-</td>
<td>-</td>
<td>2/1.6</td>
</tr>
<tr>
<td>Big ET-1[16-27,Cys]</td>
<td>2/2.1</td>
<td>1/1.0</td>
<td>2/1.7</td>
<td>2/1.7</td>
<td>-</td>
<td>1/0.8</td>
<td>1/0.7</td>
<td>1/0.6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1/0.8</td>
</tr>
<tr>
<td>Big ET-1[16-38,D-Val]</td>
<td>2/1.9</td>
<td>2/2.1</td>
<td>2/1.9</td>
<td>2/1.8</td>
<td>3/2.6</td>
<td>1/0.9</td>
<td>3/2.8</td>
<td>1/1.0</td>
<td>1/0.9</td>
<td>2/2.0</td>
<td>2/1.8</td>
<td>1/1.1</td>
<td>-</td>
</tr>
</tbody>
</table>

Amino acid analysis results from SAR investigation of ECE-1.
### Amino Acid Analysis Results from Subsite Specificity Study of ECE-1

| Peptide            | His | Leu | Asp | Ile | Val | Thr | Pro | Glu | Tyr | Gly | Ser | Arg | Phe | Ala |
|--------------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| BigET-1[16-38, Phe²] | 2/2.1 | 2/2.0 | 2/2.1 | 2/1.6 | 3/2.7 | 1/0.8 | 3/3.2 | 1/0.8 | 1/0.8 | 2/2.0 | 2/1.8 | 1/1.0 | 1/1.0 | -   |
| BigET-1[16-38, Leu²] | 2/2.0 | 3/3.1 | 2/2.0 | 2/1.8 | 3/2.9 | 1/0.7 | 3/2.7 | 1/1.0 | 1/0.7 | 2/2.0 | 2/1.7 | 1/1.1 | -   | -   |
| BigET-1[16-38, Ala²] | 2/2.2 | 2/2.1 | 2/2.1 | 2/1.5 | 3/2.6 | 1/0.9 | 3/3.1 | 1/1.1 | 1/0.9 | 2/2.0 | 2/1.7 | 1/1.0 | 1/1.0 | -   |
| BigET-1[16-38, Pro²] | 2/2.0 | 2/1.7 | 2/2.3 | 2/1.6 | 3/2.8 | 1/0.8 | 4/4.1 | 1/0.8 | 1/0.9 | 2/2.0 | 2/1.8 | 1/0.7 | -   | -   |
| BigET-1[16-38, Glu²] | 2/2.1 | 2/1.9 | 2/1.8 | 2/1.7 | 3/2.8 | 1/0.7 | 3/3.2 | 2/1.8 | 1/0.7 | 2/2.0 | 2/1.8 | 1/1.2 | -   | -   |
| BigET-1[16-38, Arg²] | 2/2.0 | 2/2.2 | 2/2.1 | 2/1.7 | 3/2.7 | 1/0.9 | 3/2.8 | 1/0.7 | 1/0.8 | 2/2.0 | 2/1.8 | 2/2.0 | -   | -   |
| BigET-1[16-38, Phe²²] | 2/2.2 | 2/1.7 | 2/2.0 | 2/1.8 | 2/1.6 | 1/0.6 | 3/3.2 | 1/0.8 | 1/0.8 | 2/2.0 | 2/1.6 | 1/1.1 | 1/0.9 | -   |
| BigET-1[16-38, Leu²²] | 2/2.0 | 3/2.9 | 2/2.0 | 2/1.5 | 2/1.9 | 1/0.7 | 3/3.3 | 1/0.9 | 1/0.9 | 2/2.0 | 2/1.7 | 1/0.9 | -   | -   |
| BigET-1[16-38, Ala²²] | 2/2.0 | 2/1.8 | 2/2.1 | 2/1.6 | 2/1.8 | 1/0.8 | 3/3.0 | 1/1.0 | 1/0.8 | 2/2.0 | 2/1.6 | 1/1.1 | 1/1.0 | -   |
| BigET-1[16-38, Pro²²] | 2/2.1 | 2/1.9 | 2/2.3 | 2/1.8 | 2/1.9 | 1/0.7 | 4/4.1 | 1/1.1 | 1/0.9 | 2/2.0 | 2/1.7 | 1/1.0 | -   | -   |
| BigET-1[16-38, Glu²²] | 2/2.0 | 2/1.0 | 2/2.0 | 2/1.9 | 2/1.6 | 1/0.8 | 3/2.6 | 2/2.0 | 1/0.9 | 2/2.0 | 2/1.6 | 1/0.7 | -   | -   |
| BigET-1[16-38, Arg²²] | 2/2.1 | 2/1.7 | 2/2.2 | 2/1.8 | 2/1.8 | 1/0.9 | 3/2.7 | 1/1.0 | 1/0.9 | 2/2.0 | 2/1.7 | 2/1.8 | -   | -   |

*Amino acid analysis results from subsite specificity study of ECE-1.*
Appendix III: Mass spectroscopy and amino acid analysis results for Library 1.

<table>
<thead>
<tr>
<th>PREDICTED MASS DISTRIBUTION</th>
<th>DETERMINED MASS DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1306.61</td>
<td>1308.27</td>
</tr>
<tr>
<td>1290.62</td>
<td>1291.42</td>
</tr>
<tr>
<td>1329.63</td>
<td>1331.04</td>
</tr>
</tbody>
</table>

MALDI ToF mass spectroscopy results for Library 1.

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>PREDICTED RATIO</th>
<th>ACTUAL RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>FIRST   SECOND</td>
</tr>
<tr>
<td>Asp</td>
<td>3</td>
<td>3.1      3.1</td>
</tr>
<tr>
<td>Val</td>
<td>3</td>
<td>1.7      2.8</td>
</tr>
<tr>
<td>Ile</td>
<td>6</td>
<td>5.3      4.9</td>
</tr>
<tr>
<td>Leu</td>
<td>3</td>
<td>3.4      3.1</td>
</tr>
<tr>
<td>Tyr</td>
<td>4</td>
<td>1.9      1.8</td>
</tr>
<tr>
<td>Phe</td>
<td>1</td>
<td>1.0      0.9</td>
</tr>
<tr>
<td>His</td>
<td>3</td>
<td>2.9      2.9</td>
</tr>
</tbody>
</table>

Amino acid analysis results for Library 1.
Appendix IV: Mass spectroscopy and amino acid analysis results for Library 2.

<table>
<thead>
<tr>
<th>PREDICTED MASS DISTRIBUTION</th>
<th>DETERMINED MASS DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1732.79</td>
<td>1733.12</td>
</tr>
<tr>
<td>1748.78</td>
<td>1749.37</td>
</tr>
<tr>
<td>1771.80</td>
<td>1772.14</td>
</tr>
</tbody>
</table>

MALDI ToF mass spectroscopy results for Library 2.

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>PREDICTED RATIO</th>
<th>ACTUAL RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>6</td>
<td>7.2</td>
</tr>
<tr>
<td>Glu</td>
<td>3</td>
<td>3.2</td>
</tr>
<tr>
<td>Val</td>
<td>3</td>
<td>3.4</td>
</tr>
<tr>
<td>Ile</td>
<td>6</td>
<td>4.6</td>
</tr>
<tr>
<td>Leu</td>
<td>3</td>
<td>2.6</td>
</tr>
<tr>
<td>Tyr</td>
<td>1</td>
<td>1.1</td>
</tr>
<tr>
<td>Phe</td>
<td>4</td>
<td>4.2</td>
</tr>
<tr>
<td>His</td>
<td>3</td>
<td>2.5</td>
</tr>
<tr>
<td>Lys</td>
<td>3</td>
<td>2.9</td>
</tr>
</tbody>
</table>

Amino acid analysis results for Library 2.
Appendix V: Mass spectroscopy and amino acid analysis results for Library 3.

<table>
<thead>
<tr>
<th>PREDICTED MASS DISTRIBUTION</th>
<th>DETERMINED MASS DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1285</td>
<td>1287.9</td>
</tr>
<tr>
<td>1299</td>
<td></td>
</tr>
<tr>
<td>1303</td>
<td>1303</td>
</tr>
<tr>
<td>1313</td>
<td></td>
</tr>
<tr>
<td>1317</td>
<td>1319.4</td>
</tr>
<tr>
<td>1331</td>
<td></td>
</tr>
<tr>
<td>1335</td>
<td>1334.3</td>
</tr>
<tr>
<td>1345</td>
<td></td>
</tr>
<tr>
<td>1349</td>
<td></td>
</tr>
<tr>
<td>1361</td>
<td>1360.9</td>
</tr>
<tr>
<td>1363</td>
<td>1364.3</td>
</tr>
<tr>
<td>1375</td>
<td></td>
</tr>
<tr>
<td>1379</td>
<td>1379.6</td>
</tr>
<tr>
<td>1393</td>
<td>1398.3</td>
</tr>
<tr>
<td>1407</td>
<td></td>
</tr>
<tr>
<td>1411</td>
<td>1413.4</td>
</tr>
<tr>
<td>1425</td>
<td>1425.4</td>
</tr>
</tbody>
</table>

MALDI ToF mass spectroscopy results for Library 3.

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>PREDICTED RATIO</th>
<th>RATIO DETERMINED</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FIRST</td>
<td>SECOND</td>
</tr>
<tr>
<td>Asp</td>
<td>11</td>
<td>12.2</td>
</tr>
<tr>
<td>Thr</td>
<td>1</td>
<td>1.4</td>
</tr>
<tr>
<td>Ser</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>Glu</td>
<td>2</td>
<td>1.8</td>
</tr>
<tr>
<td>Pro</td>
<td>1</td>
<td>0.0</td>
</tr>
<tr>
<td>Gly</td>
<td>3</td>
<td>2.9</td>
</tr>
<tr>
<td>Ala</td>
<td>6</td>
<td>5.9</td>
</tr>
<tr>
<td>Leu</td>
<td>3</td>
<td>2.6</td>
</tr>
<tr>
<td>Tyr</td>
<td>4</td>
<td>3.8</td>
</tr>
<tr>
<td>Phe</td>
<td>4</td>
<td>4.3</td>
</tr>
</tbody>
</table>

Amino acid analysis results for Library 3.
Appendix VI: Mass spectroscopy and amino acid analysis results for Library 4.

<table>
<thead>
<tr>
<th>PREDICTED MASS DISTRIBUTION</th>
<th>DETERMINED MASS DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>1270.52</td>
<td>1269.99</td>
</tr>
<tr>
<td>1284.54</td>
<td>1285.68</td>
</tr>
<tr>
<td>1298.56</td>
<td>1299.69</td>
</tr>
<tr>
<td>1312.94</td>
<td>1312.98</td>
</tr>
</tbody>
</table>

MALDI ToF mass spectroscopy results for Library 4.

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>PREDICTED RATIO</th>
<th>DETERMINED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>9</td>
<td>8.5</td>
</tr>
<tr>
<td>Ser</td>
<td>3</td>
<td>2.9</td>
</tr>
<tr>
<td>Glu</td>
<td>3</td>
<td>3.0</td>
</tr>
<tr>
<td>Gly</td>
<td>4</td>
<td>3.5</td>
</tr>
<tr>
<td>Ala</td>
<td>4</td>
<td>3.7</td>
</tr>
<tr>
<td>Abu</td>
<td>1</td>
<td>0.7</td>
</tr>
<tr>
<td>Val</td>
<td>1</td>
<td>1.0</td>
</tr>
<tr>
<td>Ile</td>
<td>1</td>
<td>0.9</td>
</tr>
<tr>
<td>Leu</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>Tyr</td>
<td>3</td>
<td>2.4</td>
</tr>
<tr>
<td>Phe</td>
<td>3</td>
<td>2.7</td>
</tr>
<tr>
<td>Pro</td>
<td>3</td>
<td>1.5</td>
</tr>
</tbody>
</table>

Amino acid analysis results for Library 4.
Appendix VII: Mass spectroscopy and amino acid analysis results for Library 5.

<table>
<thead>
<tr>
<th>PREDICTED MASS DISTRIBUTION</th>
<th>OBSERVED MASS DISTRIBUTION</th>
</tr>
</thead>
<tbody>
<tr>
<td>4224</td>
<td>4224</td>
</tr>
<tr>
<td>4274</td>
<td>4272</td>
</tr>
<tr>
<td>4281</td>
<td>4284</td>
</tr>
<tr>
<td>4293</td>
<td>4296</td>
</tr>
<tr>
<td>4309</td>
<td>4310</td>
</tr>
<tr>
<td>4331</td>
<td>4331</td>
</tr>
<tr>
<td>4350</td>
<td>4353</td>
</tr>
<tr>
<td>4359</td>
<td>4359</td>
</tr>
<tr>
<td>4378</td>
<td>4377</td>
</tr>
</tbody>
</table>

Electrospray mass spectroscopy results for Library 5.

<table>
<thead>
<tr>
<th>AMINO ACID</th>
<th>PREDICTED RATIO</th>
<th>DETERMINED RATIO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>36</td>
<td>33.6</td>
</tr>
<tr>
<td>Ala</td>
<td>30</td>
<td>26.7</td>
</tr>
<tr>
<td>Glu</td>
<td>36</td>
<td>37.8</td>
</tr>
<tr>
<td>Phe</td>
<td>27</td>
<td>24.7</td>
</tr>
<tr>
<td>Arg</td>
<td>15</td>
<td>13.7</td>
</tr>
<tr>
<td>His</td>
<td>21</td>
<td>22.8</td>
</tr>
<tr>
<td>Ser</td>
<td>21</td>
<td>19.7</td>
</tr>
<tr>
<td>Gly</td>
<td>54</td>
<td>50.8</td>
</tr>
<tr>
<td>Tyr</td>
<td>9</td>
<td>8.6</td>
</tr>
<tr>
<td>Val</td>
<td>57</td>
<td>51.9</td>
</tr>
<tr>
<td>Lys</td>
<td>9</td>
<td>8.7</td>
</tr>
<tr>
<td>Leu</td>
<td>18</td>
<td>19.9</td>
</tr>
<tr>
<td>Ile</td>
<td>18</td>
<td>16.4</td>
</tr>
<tr>
<td>Met</td>
<td>9</td>
<td>8.7</td>
</tr>
</tbody>
</table>

Amino acid analysis results for Library 5.
Courses Attended


RSC Perkin Division, Scottish Meeting, University of St. Andrews, 1998; University of Strathclyde, 1997; University of Edinburgh, 1996; University of Glasgow, 1995.

SCI Graduate Symposium, University of Edinburgh, 1998; University of St. Andrews, 1996.


European Peptide Symposium, University of Edinburgh, September 1996.
Presentations and Awards

