I dedicate this 'ere Thesis to all those
who spend painful hours of failure in the lab
.....it's worth it in the end, don't lose heart!

"...when I started doing chemistry, I did it the way I fished –
for the excitement, the discovery, the adventure,
for going after the most elusive catch imaginable
in uncharted seas....The discipline, nonetheless,
is exacting: everything that can be observed
should be observed, even if it is only recalled
as the bland background from which the intriguing
bits pop out.... The goal is always finding something new,
hopefully unimagined and, better still, hitherto unimaginable."
K. Barry Sharpless, Nobel Lecture 2001

Believe nothing, no matter where you read it, or who said it,
no matter if I have said it, unless it agrees with
your own reason and your own common sense.
Gautama Buddha

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING, SCIENCE AND MATHEMATICS

SCHOOL OF CHEMISTRY

Doctor of Philosophy

Combinatorial Approaches to Peptide Receptors

by JON SHEPHERD

This thesis concerns the evaluation of molecular receptors for peptide carboxylates.

Chapter 1 outlines the scope of the work, including recent developments in the host-guest chemistry of oligopeptide carboxylates, the combinatorial paradigm, sequencing methodologies, its application to the discovery of receptors and the nature of guests of interest studied.

Chapter 2 describes the evolution of a novel method of determining a solid-phase linked peptide sequence. Limitations are discussed. The method is then successfully implemented in the synthesis of a library 'pseudo-tweezer' receptors for carboxylates and a library of authentic tweezer receptors, including the synthesis of appropriate carboxylate-binding motifs.

Chapter 3 describes the use of the libraries synthesised in solid-phase screening experiments with oligopeptide guests of biological interest. Results of these experiments are discussed, leading to the identification of receptors. Resynthesis of these receptors in both solid-phase-linked and free form is detailed and evaluation of binding is made.

Chapter 4 describes the synthesis and binding constant evaluation of a receptor for the oligopeptide of interest D-Ala-D-Ala, which had previously been identified by alternative combinatorial methods.

A perspective on the method is offered and future investigation indicated.

2005

Contents

Preface

I, Jon Shepherd, declare that

- the research described in this thesis was carried out under the supervision of Prof. J.D. Kilburn at the University of Southampton between October 2002 and October 2005;
- No part of this thesis has previously been submitted at this or any other university;
- Where the published work of others has been consulted, this is always clearly attributed;
- Where the public work of others has been quoted, this is always clearly attributed;
- I have acknowledged all main sources of help

Jon Shepherd

December 2005

Acknowledgements

Ah, whole pages all to myself to warble on about something other than Chemistry! Except that... ...

First and foremost thanks go to Jeremy for three years of encouragement, cajoling and funny looks, and not least the opportunity to learn more than I thought I could. I hope I've not been too over-the-top in all this time! Best of luck for the future, and no doubt I'll see you again.

The Kilburn Group both past and present, I salute you. Three people in particular deserve extra-special mention; Sandra; for an education in language, for teaching me appreciation of my own abilities and for how not to despair when things are 'not bloody soluble in DMF' (apologies for the faaaaaags!); Richard; for being the only person to appreciate bad puns, for understanding that things NEVER work properly and for sly digs at my incipient northern qualities (Biscuits? No, it just isn't funny anymore); and Sarah; for showing me it is possible to be both chemist and human at the same time and all the strange noises and expressions! Thankyou, thankyou, thankyou all three. To the rest of you, I am proud to have known you all. Remember to keep Süchi-Büchi clean and nice, or else.

Proof-readers are useful people. My own have been... ...

Thanks to the group of Prof. J. de Mendoza for for supplying carboxylate-binding site.

Thanks to Dr. T. Gale for preparing the dye-labelled HIV-guest. Saves a lot of time!

Thanks to John and Julie down in MS for being both extremely helpful and pleasant people. Especially for the time the TOF went loopy!

My source of free tickets: Adrian, thanks for being so generous! Get back home safe, and there's a big hug in it for you!

Many thanks to the University Buddhist Society, and in particular to Jaye and Lalitaratna, for showing me something different of life. Om Mane Padme Hum in your direction.

Thanks to all those who have put up with my domestic habits. Ailsa – I'm glad you let me live with you! I think I win our little race. Carole – It was a good two years, don't lose touch! Caterina – many lovely smells and moral support.

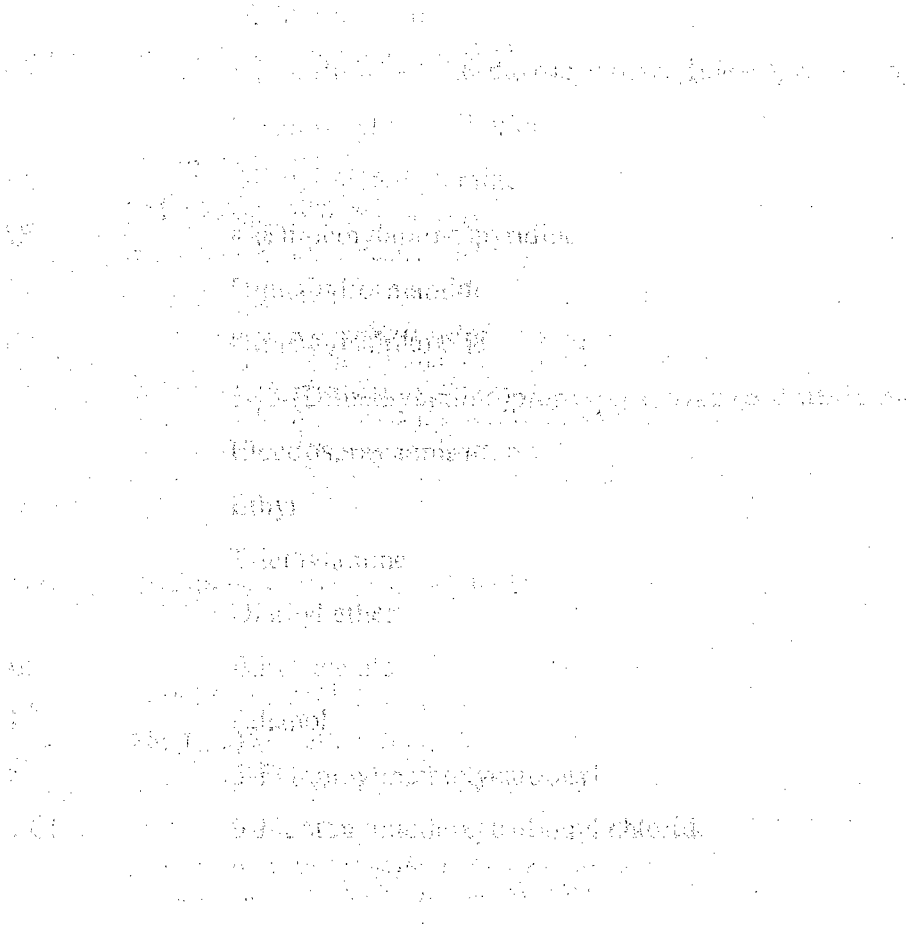
To all at the West End Little Theatre Club: you've been great, we did some good shows. I'll always have the paint on my shoes and the memory of enormous bruises! Good luck for the future.

To the family; Jason gets credit for all manner of useless information and good grace. Deb – whatever you think, I'm always there, even if crap with it. Dad, I really don't know what to say. It's difficult. I see more of you in myself every day.

And finally, with great difficultly, eternal feelings to Mum. I am forever indescribably grateful to you, for giving me the best reasons I've ever needed. Enjoy the next ride, see you again, sometime.

Jon Shepherd

December 2005



Abbreviations

AA	Amino acid
Ac	Acetyl
Ac_2O	Acetic anhydride
AcOH	Acetic acid
Bn	Benzyl
BnBr	Benzyl bromide
^tBu	*tert*-butyl
Boc	*tert*-butyloxycarbonyl
Boc_2O	Di-*tert*-butyl pyrocarbonate
CBS	Carboxylate Binding Site
Cbz	Benzyloxycarbonyl
CNBr	Cyanogen bromide
DCC	Dicyclohexylcarbodiimide
DCM	Dichloromethane
Ddpe	1-(4,4-dimethyl-2,6-dioxocyclohexylidene)phenylethyl
DIC	Diisopropylcarbodiimide
DIPEA	Diisopropylethylamine
DMAP	4-(Dimethylamino)pyridine
DMF	Dimethylformamide
DMSO	Dimethylsulphoxide
EDC	1-[3-(Dimethylamino)propyl]-3-ethylcarbodiimide hydrochloride
ES	Electrospray ionisation
Et	Ethyl
Et_3N	Triethylamine
Et_2O	Diethyl ether
EtOAc	Ethyl acetate
EtOH	Ethanol
Fmoc	9-Fluorenylmethoxycarbonyl
Fmoc-Cl	9-Fluorenylmethoxycarbonyl chloride

FT-IR	Fourier-transform Infra-Red spectroscopy
HMDS	Hexamethyldisilazane
HOBt	Hydroxybenzotriazole
HPLC	High-performance liquid chromatography
HRMS	High-resolution mass spectrometry
LRMS	Low-resolution mass spectrometry
MALDI-TOF	Matrix-assisted-Laser-desorption-Ionisation-Time-Of-Flight
Me	Methyl
MeOH	Methanol
MS	Mass Spectrometry
Mtr	4-methoxy-2,3,6-trimethylbenzenesulphonyl
NMR	Nuclear Magnetic Resonance spectrometry
Pbf	2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulphonyl
p-Br PhCOOH	*para*-bromobenzoic acid
PEG	Poly(ethylene glycol)
Ph	Phenyl
PR	HIV-1 Protease
PyBOP	Benzotriazol-1-yloxytripyrrolidinophosphonium hexafluorophosphate
RT	Reverse transcriptase
SPPS	Solid-phase peptide synthesis
TBDPS	*tert*-butyldiphenylsilyl
TFA	Trifluoroacetic acid
TMS	Trimethylsilyl
TMSBr	Trimethylsilylbromide

The common three-letter abbreviations are used for amino acids;

Ahx	γ-Aminohexanoic acid
Ala	Alanine

Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
Cys	Cysteine
Gly	Glycine
Gln	Glutamine
Glu	Glutamic acid
His	Histidine
Ile	Isoleucine
Leu	Leucine
Lys	Lysine
Met	Methionine
Phe	Phenylalanine
Pro	Proline
Ser	Serine
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine

Chapter 1

Introduction

1.1 The field of supramolecular chemistry

Supramolecular chemistry is a young and burgeoning field within chemical science, amalgamating many traditional disciplines. It addresses the chemical, physical *and* biological features of species held in union by non-covalent interactions.

In the broadest sense, supramolecular chemistry is the study of systems containing more than one molecular identity and more specifically, the 'chemistry of molecular assemblies and of the intermolecular bond'[1]. In a semantic sense, supramolecular chemistry is 'transcending molecular chemistry', turning away from the molecule as individual entity and towards molecular association and thus an interest in functionalities that confer intermolecular interaction rather than chemical reactivity.

The birth of supramolecular chemistry is traditionally recognised as Charles Pedersen's potassium-templated synthesis of crown ethers[2,3]. However, the study of molecular interaction has a much longer history, especially in medicinal chemistry. Emil Fischer formulated the basic principles of enzyme-substrate interactions in the simple analogy of "wie Schloss und Schlüssel zueinander passen müssen"[4] [must be complementary like a lock and key], which contains the idea of specificity in molecular binding, of how substrates and enzymes must interact selectively. Current studies into effective drugs rely on the study of interaction between a drug molecule and a target biomolecule involving many of the concepts developed in supramolecular chemistry, yet many practitioners would not consider themselves supramolecular chemists.

A common terminology is used within the field to describe interacting systems. Two participants in a molecular binding event are generally termed the 'host' and the 'guest'.

The host is generally regarded as the larger of the two, displaying convergent binding functionality whereas the guest is usually smaller in size and displays divergent binding functionality. These terms are flexible, since binding is a cooperative event between two species; what is termed the 'host' and what the 'guest' is an issue of perspective. Hosts are also commonly referred to as 'receptors'.

For example, epidermal growth factor receptor[5] (EGF-R) is a protein tyrosine kinase 'host' which is able to bind several different epidermal growth factor hormone 'guests'[6-8]. Such biological interactions are usually highly specific and critical for many cellular processes, including DNA-synthesis, self-assembly of protein complexes and antigen-antibody responses, all taking place in aqueous medium. A great challenge for supramolecular chemists is to mimic such highly selective processes and to generate systems which also recognise complex biological structures under physiological conditions.

The application of host-guest systems stretches beyond the biochemical world. 'Molecular machines' such as those generated by Stoddart and co-workers rely on differing recognition processes occurring under differing conditions[9,10]. For example, the molecular elevator formed by the association of tris-crown ether **1** and tripodal guest **2** can be controlled by pH, and generate modest amounts of mechanical energy (Fig 1-1).

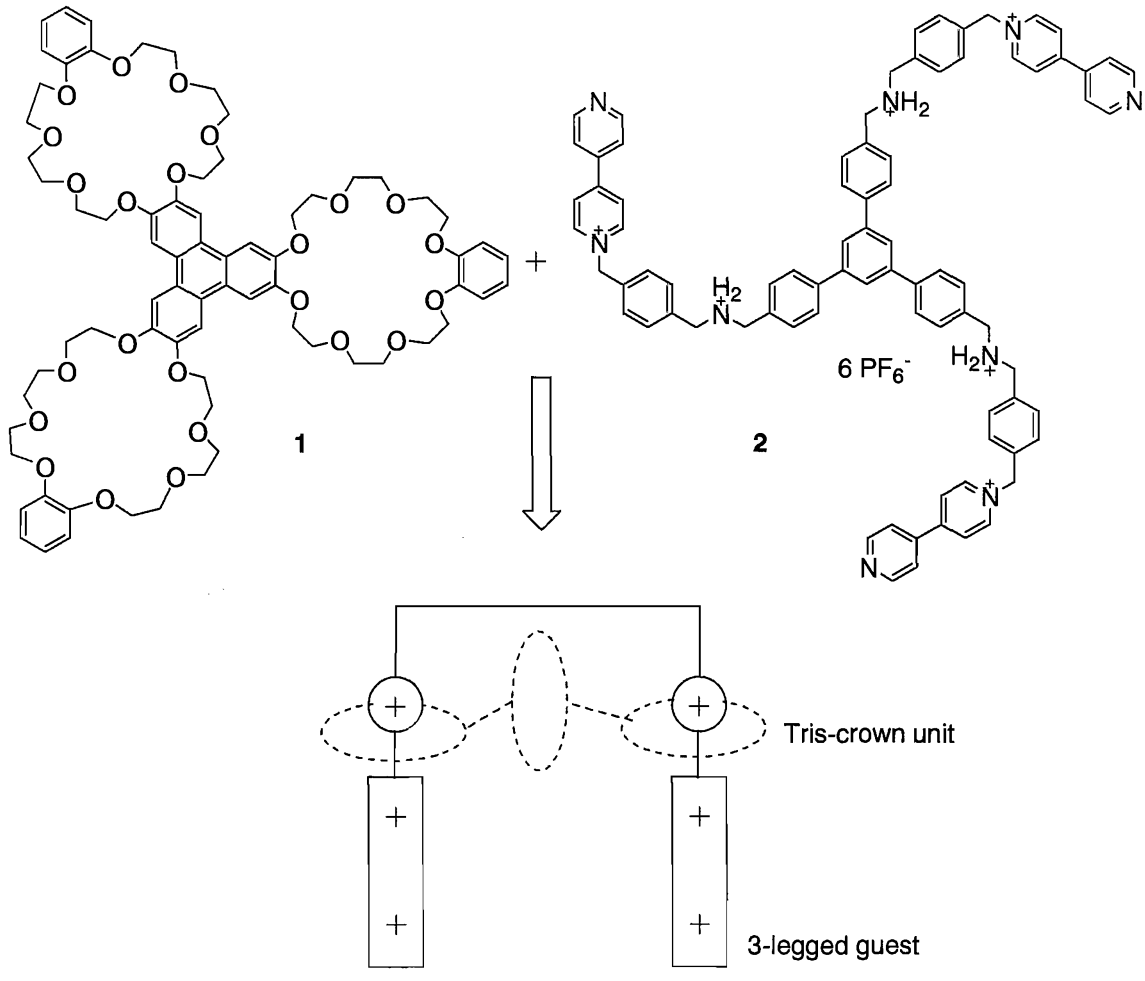


Fig 1-1 Stoddart *et al*'s molecular elevator; Association of the two components 1 and 2 into the 'molecular elevator', shown below in a simplified schematic with only two 'shafts'. 'Legs' of 2 insert into the macrocyclic rings of 1. Depending on pH (and hence the protonation state of the secondary amines in 2), the tris-crown platform occupies an 'upper' or 'lower' station

Supramolecular systems also have applications in catalysis, such as Cram *et al*'s 'synthetic-chymotrypsin' species[11-13] (Fig 1-2). A further application of supramolecular systems is in separation science, for example the enantioseparation of phenylalanine salts[14] by binding and transport across phase boundaries by hosts which are enantioselective.

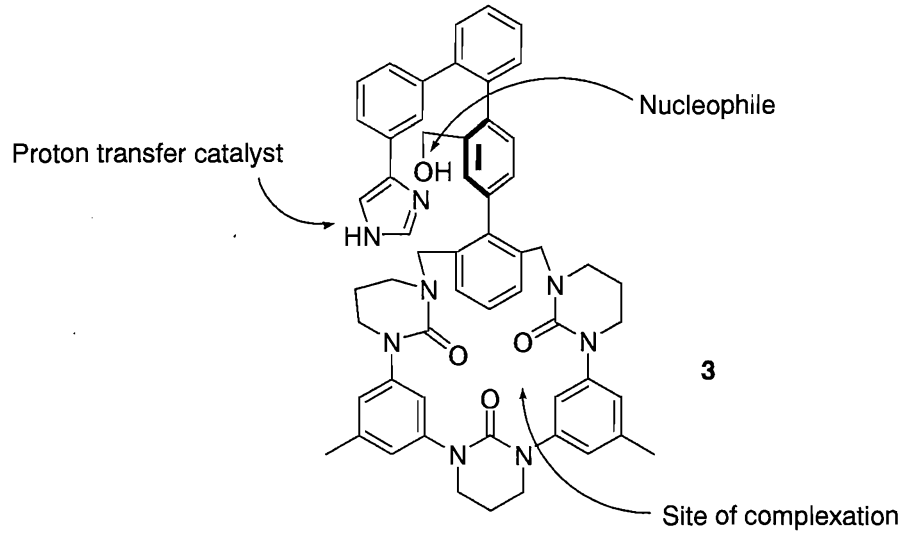


Fig 1-2 Cram *et al*'s 'synthetic chymotrypsin' 3. Displaying the same essential features as the active site of the chymotrypsin enzyme (binding site, nucleophilic OH, imidazole proton shuttle – lacks the extra carboxyl group of Asp), the species is able to catalyse a model transacylation reaction by 10^5

This brief overview indicates the broad scope of supramolecular chemistry. A large number of synthetic hosts have been prepared for the binding of a plethora of substrates for a variety of applications. In particular, there has been extensive interest in the binding of biomolecules including nucleotide bases, amino acids and oligopeptides.

1.2 Factors affecting the solution-phase binding of species

Synthetic receptors are based around the design paradigm of the eventual construction of a stable host-guest complex held together by non-covalent interactions. Such constructions are possible due to thermodynamic stabilisation gained by the system on complexation, often a synergy of several interactions.

Binding is a fluid, equilibrium-based process, with guests being able to freely associate and disassociate from hosts. Hence binding affinities are often expressed as an equilibrium constant. For example, in the simplistic binding event;

$$H + G \rightleftharpoons H.G$$

the binding constant K_a is given by;

$$K_a = \frac{[H.G]}{[H].[G]}$$

and a larger K_a represents a higher equilibrium concentration of host-guest complex and hence a more strongly binding receptor.

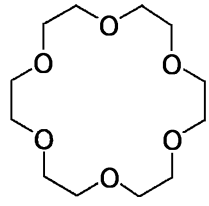
Binding constants are thus thermodynamic quantities and are related to the free energy exchanges of the binding process. Free energy change is conveniently divided into enthalpic contributions (due to the formation/destruction of intermolecular bonds) and entropic contributions, which are both discussed below. It is the combination of all such contributions which determines the ability of a host to bind a specific guest in a particular medium, hence the design of a receptor usually attempts to maximise all positive contributions. A brief discussion of some typical contributions follows.

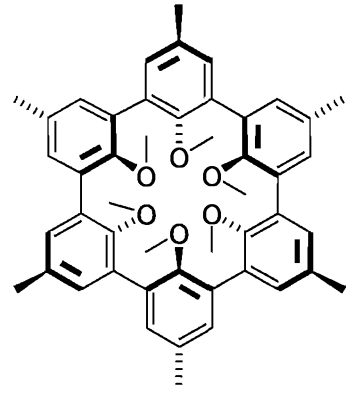
1.2.1 Conformational flexibility

Upon binding, a host-guest complex acquires a new rigidity not associated with either the free host or free guest, due to the formation of non-covalent intermolecular bonds. The geometrical demands of such bonds force the host and guest to adopt particular conformations in order to maximise the interactions. Hence both host and guest lose various degrees of freedom upon binding. A 'pre-organisation' of a host prior to a binding event is often used to denote this. The loss of flexibility is an entropic contribution and is commensurate with the transition from two independent units to a single, associated unit.

Recognition of this process has led to the design of receptors with a rigid structure. The host is then in theory perfectly defined to interact with a guest molecule, thus losing fewer degrees of freedom upon binding and in effect 'pre-paying' this entropic penalty binding at

the stage of receptor synthesis. An example is the highly rigidified spherand systems such as **5** used to bind metal cations (Fig 1-3).[15]





4: Na^+ $K_a = 2.5 \times 10^4$ **5**: Na^+ $K_a = 1.4 \times 10^{14}$

Fig 1-3 Pre-organisation of binding sites. The crown ether 4 is flexible and loses entropy on binding a metal cation, an energetic penalty. In contrast, the spherand 5 is highly rigid and creates a perfect binding site for binding certain metal cations. No energy is required to reorganise 5 into a binding conformation

However, it is rarely possible to so perfectly design a host for guests of complexity greater than spherical geometry, and host rigidity often impedes binding of guest by preventing access to sites of binding functionality. Thus a compromise is often achieved by using receptors with both rigid portions (to avoid a high entropic penalty of binding) and flexible portions which allow access by the guest and some degree of system-internal optimisation of binding geometry (rather than imposing it through rigidity).

1.2.2 Electrostatic interactions

Electrostatic interactions result from the forces between charged species, as described by Coulomb's laws. They are enthalpic in nature. The magnitude of the interaction depends on the size of the charges, the distance between them and the nature of the medium containing the charges.

Typical electrostatic interactions exist between ions and permanent dipoles. Since dipoles are vector quantities, they must be correctly aligned in order to maximise the interaction, which is attractive between opposite charges (Fig 1-4).

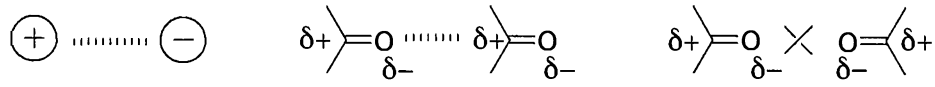
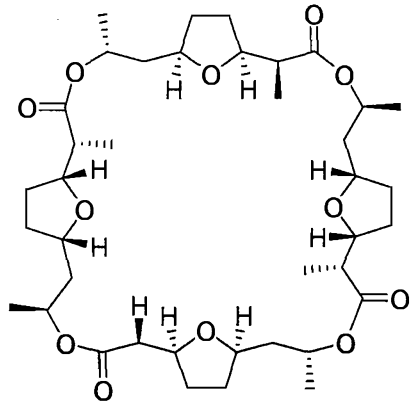


Fig 1-4 Electrostatic interactions. Attraction exists between ions of opposite charge and maximally between correctly aligned dipoles. Incorrectly aligned dipoles do not attract.

Such interactions are typified in crown ethers of Pedersen[3], which bind metal cations through ion-dipole interactions between the metal ion and the oxygen atoms of the polyether macrocycles. A correlation has been made between the cavity size of the macrocycle and the cation radius for maximum complex stability, such that a more stable complex is formed when the cation fits best inside the crown ether. A similar situation exists for the naturally occurring macrotetrolide nonactin **6** (Fig 1-5), which is a known potassium ionophore[16] and ammonium sensor[17].



6

Fig 1-5 Nonactin 6, a natural product known to effectively bind potassium and ammonium

1.2.3 Hydrogen bonds

Hydrogen bonds are ubiquitous in supramolecular complexes, and may be thought of as the basic tool in the kit of supramolecular chemistry. Such interactions are enthalpic in nature.

Hydrogen bonds are established between an electronegative heteroatom (including, but not exclusively, N, O and F) and a hydrogen atom. The heteroatom is characterised by the presence of a lone pair of electrons, which act as the hydrogen bond acceptor; the hydrogen atom is formally covalently bonded to a third atom and is regarded as the hydrogen bond donor. Examples are shown in Fig 1-6)

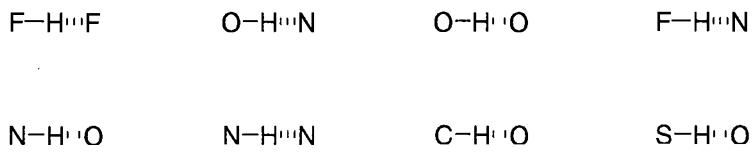


Fig 1-6 Examples of hydrogen bonds

Hydrogen bonding is independent of the charged nature of host and guest; hydrogen bonds form equally well between neutral species and between charged species. However the strength of a hydrogen bond can be reinforced if the atoms involved do carry charge. Hydrogen bonds are far weaker than standard covalent bonds (~50-100kcal/mol) but much stronger than surface-dependent van der Waals interactions (~0.05-0.5kcal/mol).[18]

Hydrogen bonds are detected in X-ray crystal structures and have a nominal length of between 1.7 and 2.0 Å, however this is not exclusive.[19] Countless examples of hydrogen bonding exist in the literature, of which a few illustrative examples are discussed.

The most famous example of hydrogen bonding between two species is given by the double-helix structure of DNA. Two strands of ribophosphate polymer backbone are held together by the complementary hydrogen bond interactions of appended nucleobases. Selective recognition is noted between the triply hydrogen-bonded guanine-cytosine pair and the doubly hydrogen-bonded adenine-thymidine pair (Fig 1-7), which gives rise many

of the information-storage properties of DNA. The selectivity arises from the matched nature of hydrogen bond donors and acceptors displayed by the nucleobases.

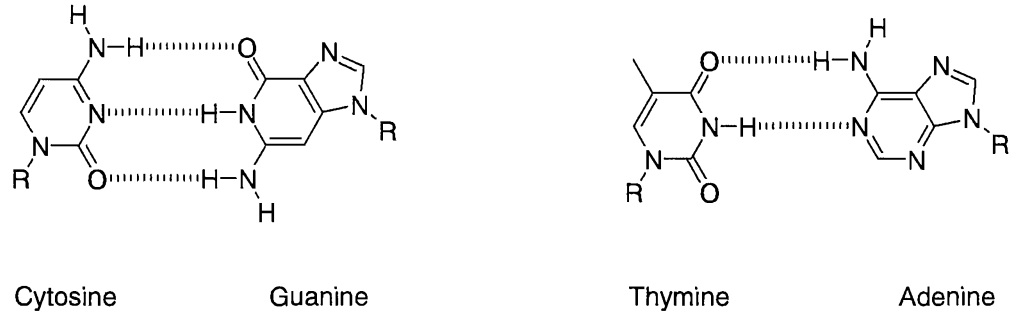


Fig 1-7 Complementary base pairs in DNA. R groups represent the point of attachment to the ribophosphate backbone. Cytosine and guanine interact through three hydrogen bonds, thymine and adenine through two. Both arrangements are donor-acceptor matched.

Hydrogen bonding in synthetic receptors is often utilised in the form of an array of donor/acceptor functionalities which match those of the guest. The three-dimensional arrangement is of critical importance, as secondary electrostatic interactions can arise, which affects both the binding strength and binding selectivity.[20]

A synthetic receptor which relies heavily on hydrogen bonding interactions is the barbiturate receptor of Hamilton *et al.*[21] This macrocyclic receptor contains six hydrogen bonding functionalities which are complementary both in nature and in arrangement to that of barbiturate (Fig 1-8). The importance of the hydrogen bonds was shown by investigating the binding of guests with fewer H-bonding functionalities (e.g. **8**, **9**), which were found to bind with a much lesser K_a.

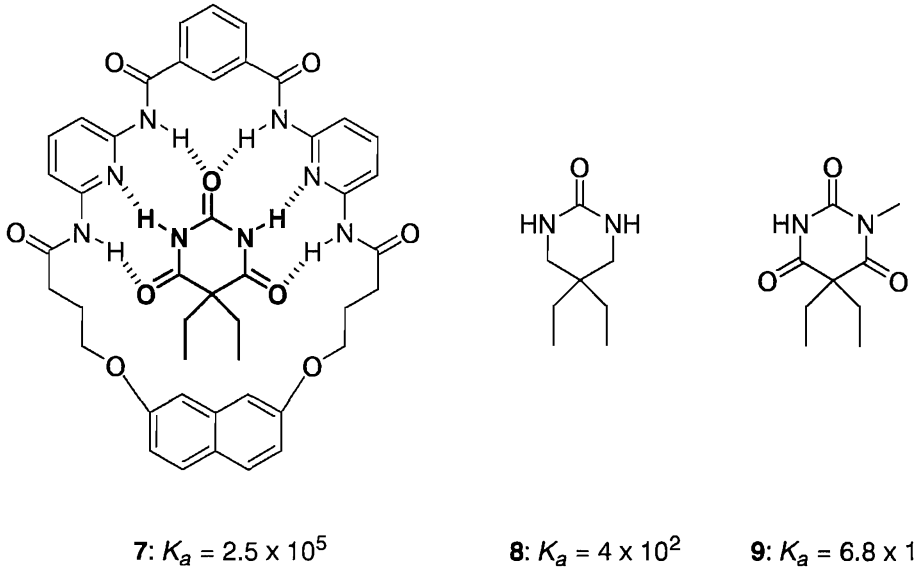


Fig 1-8 Hamilton *et al*'s barbiturate receptor. The network of six hydrogen bonds contributes to maximal binding of barbiturate. Guests lacking certain H-bond functionalities are bound less strongly

1.2.4 The effect of solvent

Binding events are usually observed under ambient conditions in a solution phase, which permits efficient access of hosts and guests to one another. Thus the nature of the solvent plays a critical role in the binding event and has great effects on complex stabilistation.[22] Hydrophobic (or more specifically solvophobic) interactions also influence a number of important processes including the formation of micellar and colloidal structures, stabilization of protein-protein complexes and protein folding.

In the first instance, there is a great entropic contribution made by solvent; both host and guest must desolvate over their mutually binding surfaces to permit intermolecular interactions to occur. This releases a number of solvent molecules to the bulk solution, which is a positive entropic contribution to binding. This is especially the case where host or guest is only weakly solvated in its unbound state, leading to the so-called 'hydrophobic effect' where a guest much prefers the environment provided by the host than bulk solution. A specific example is the binding of aromatic species by cyclodextrins[23] in aqueous solution, whereby the aromatic guest prefers the non-polar environment provided by the cavity of the cyclodextrins. Binding is accompanied by the displacement of several molecules of water.

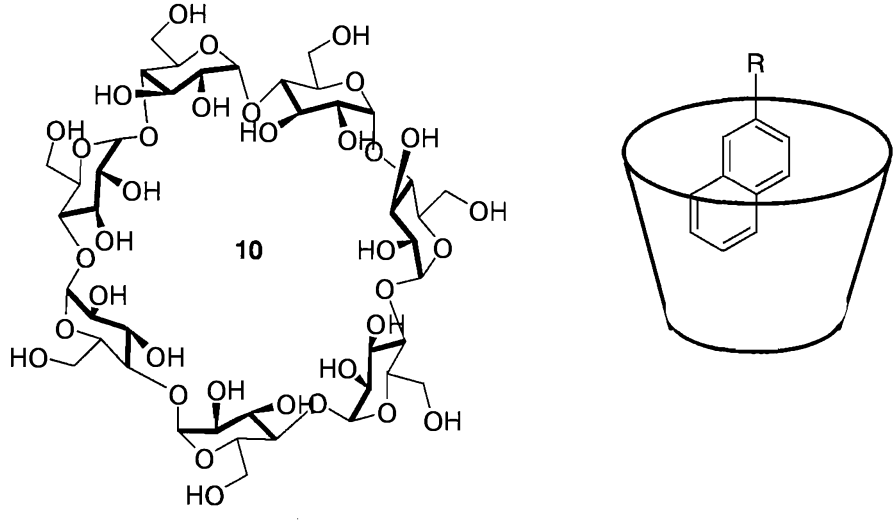


Fig 1-9 *β*-Cyclodextrin 10, an oligomer of 7 pyranoses (cyclodextrins of varying size exist). The structure forms a funnel shape displaying all the pendant hydroxyls either on the outer face or on the rims, parallel to the funnel. This creates a polar exterior coupled with a non-polar interior.

A study of complexes of cyclophane receptors displaying hydrophobic cavities with aromatic guests[24] found that binding strength decreased upon progression from aqueous systems to more organic systems, demonstrating the relationship between complexation and solvent nature.

1.3 The binding of carboxylates and oligopeptides

Carboxylates are ubiquitous in nature, commonly as amino acids but also in the form of oligopeptides and proteins. Such species have important biological functions[25], and much effort has been devoted to investigating the binding of such species. A brief introduction to the binding of carboxylates and oligopeptides is now given; more extensive information can be found in the literature.[22, 26-30]

1.3.1 Carboxylate receptors

Species used to specifically bind carboxylates are often characterised by the presence of a carboxylate-binding site (CBS). Examples of CBS include urea, thiourea and guanidinium moieties (Fig 1-10)

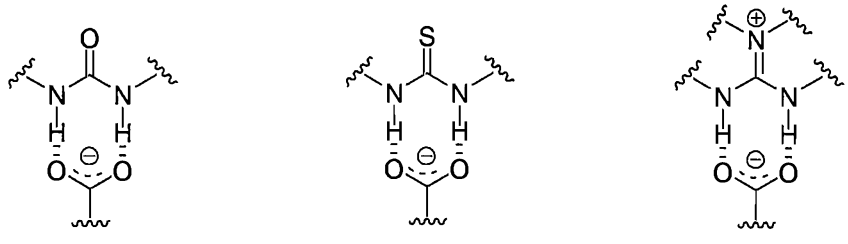
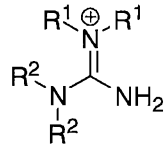


Fig 1-10 Examples of CBS. These moieties display hydrogen bond functionality matched to the carboxylate

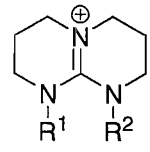
These moieties all display a bidentate hydrogen bond donor system which is matched to the bidentate hydrogen bond acceptor arrangement of the carboxylate. They are also matched in size, making these groups well-suited for use as CBS-units. In particular, the guanidinium moiety is highly appropriate as it remains protonated over a substantial pH range, thus forming not only hydrogen bonding interactions but also strong electrostatic interactions with carboxylates.[30] Nature uses the guanidinium moiety (present as the natural amino acid Arg) in order to coordinate carboxylates and other anionic groups

(present as Asp, Glu or various phosphates). The properties of natural and synthetic guanidines are amply noted in the literature.[31]

Hamilton *et al*[32] have studied a range of alkylguanidinium species as tetrafluoroborate salts by isothermal titration calorimetry with tetrabutylammonium acetate, including acyclic and bicyclic structures. Substitution of potential hydrogen bonding sites in **11** by methyl groups led to a significant reduction in binding constant (measured in DMSO). Likewise the bicyclic system **14** was found to bind acetate strongly, but the alkylated derivatives **15** and **16** gave no detectable binding.



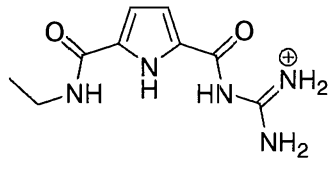
11: $R^1 = R^2 = H$, $K_a = 7.9 \times 10^3\ M^{-1}$
12: $R^1 = Me$, $R^2 = H$, $K_a = 3.4 \times 10^3\ M^{-1}$
13: $R^1 = R^2 = Me$, $K_a = 1.1 \times 10^2\ M^{-1}$



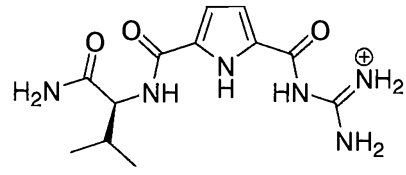
14: $R^1 = R^2 = H$, $K_a = 5.6 \times 10^3\ M^{-1}$
15: $R^1 = Me$, $R^2 = H$
16: $R^1 = R^2 = Me$

Fig 1-11 Hamilton *et al*'s model guanidinium CBSs for carboxylate binding

Binding of a carboxylate can be enhanced by incorporation of interacting moieties beyond the CBS. Schmuck[33] synthesised acylguanidinium **17**, attached to a pyrrole ring and an amide unit. This receptor was found to bind a variety of *N*-α-acetyl amino acid carboxylates in DMSO-H_2O mixtures (Fig 1-12). It was suggested that extra hydrogen bonding interactions from the pyrrole nitrogen and the amide NH account for increased binding in such polar medium. Incorporation of a Val unit into this receptor structure (**18**) conferred a degree of enantioselectivity in binding *N*-α-acetyl amino acid carboxylates.[34]



17: K_a = 770 M^{-1} (Ac-L-Ala)



18: K_a = 1610 M^{-1} (Ac-L-Ala)
K_a = 930 M^{-1} (Ac-D-Ala)

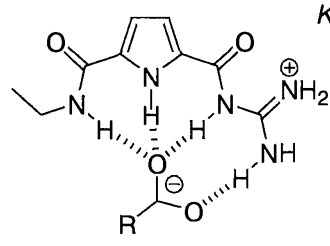
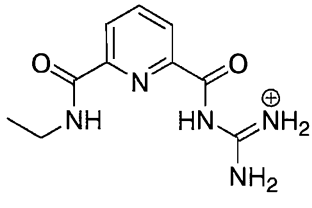
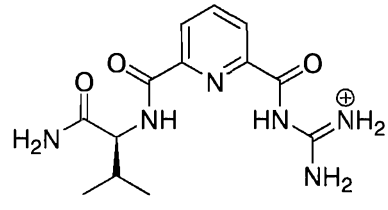


Fig 1-12 Schmuck's guanidiniopyrrole receptors for carboxylates in aqueous medium

Exchanging the pyrrole moiety for a pyridine ring gave guanidiniopyridine receptors **19** and **20**.[35] These were found to bind the same *N*-α-acetyl amino acid carboxylates with attenuated strength, and also the Val-linked receptor **20** showed no enantioselectivity. This was attributed to dipole repulsion between carboxylate and pyridine nitrogen and unfavourable steric interactions. The identical benzene system binds *N*-α-acetyl amino acid carboxylates with strength lying between the pyridine and pyrrole systems (K_a = 600 M^{-1}, Ac-L-Ala).



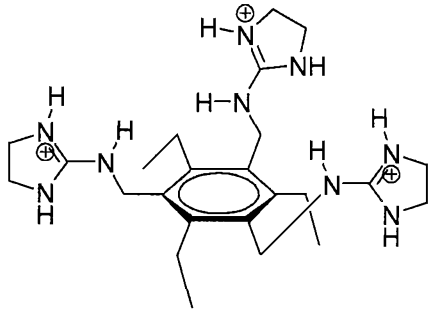
19: K_a = 460 M^{-1} (Ac-L-Ala)



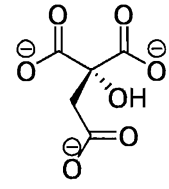
20: K_a = 360 M^{-1} (Ac-L-Ala)
K_a = 330 M^{-1} (Ac-D-Ala)

Fig 1-13 Schmuck *et al*'s guanidiniopyridine receptors

The binding of polycarboxylates has also been reported. Anslyn *et al*[36,37] synthesised trisguanidinium **21** as a chemosensor for citrate **22**, a triscarboxylate (Fig 1-14).



21

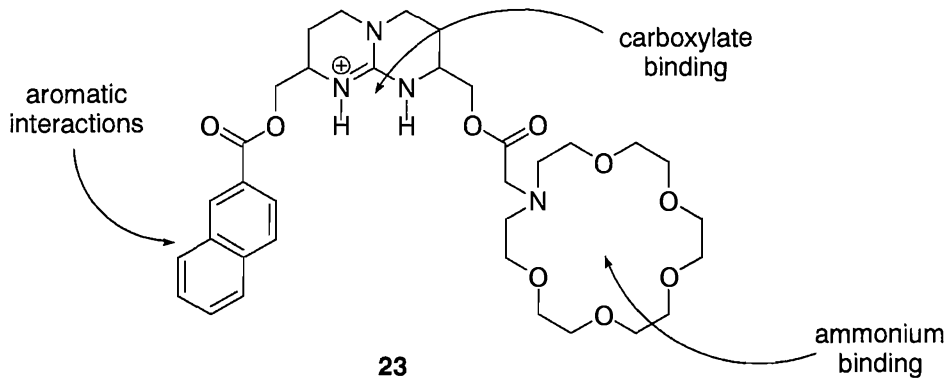


22

Fig 1-14 Anslyn *et al*'s receptor for citrate anions. The aromatic core of 21 displays three guanidinium units on the same side to bind to the three carboxylates of citrate

21 was found to be selective for citrate in water over mono- and biscarboxylates, phosphates and other salts in competition experiments.[38] The three guanidinium units are pre-organised onto the same face of the of the benzene core owing to steric effects imparted by the ethyl groups, giving a conformation displaying multiple hydrogen bond donors in positively charged moieties, perfect for binding citrate.

A noteworthy example of the use of a guanidinium in a carboxylate receptor is that of de Mendoza *et al* who devised **23** (Fig 1-15) as a receptor for L-tryptophan.[39] **23** was developed from similar structures by Schmidtchen *et al.*[30] This receptor is designed to bind the zwitterionic form of the amino acid owing to a three-point arrangement of binding motifs; the rigid bicyclic guanidinium unit for association with the carboxylate, the aza-crown macrocyclic unit to bind the ammonium (which is held in the correct region of space on one face of the bicyclic unit to bind only one enantiomer of the amino acid), and the appended naphthyl group to interact favourably with the aromatic side chain.

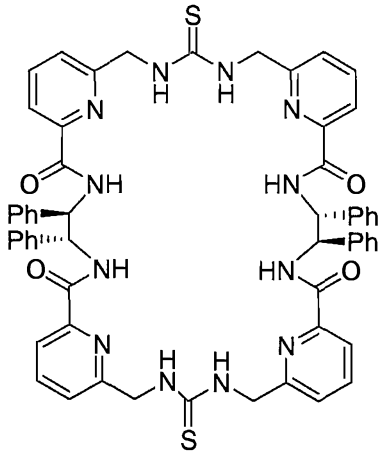


23

Fig 1-25 de Mendoza *et al*'s bicyclic guanidinium receptor for enantioselective recognition of tryptophan

Liquid-liquid extraction experiments showed that L-Trp and L-Phe were readily extracted into an organic phase by **23**, whereas L-Val was not. The corresponding D-amino acids were not extracted, illustrating the potential for highly enantioselective receptors.

Kilburn *et al* have developed the bisthiourea macrocycle **24**.[40] The presence of the pyridine rings helps to pre-organise the macrocycle for binding. **24** was found to bind *N*-protected glutamate in an enantioselective fashion and to bind guest more strongly in more polar solvent media; no detectable binding was observed in $CDCl_3$ but a binding constant of 3720 M^{-1} was measured in d_6-DMSO, which is contrary to what is usually expected.



24

Fig 1-16 Kilburn *et al*'s bisthiourea macrocycle 24 for the binding of glutamate. The pyridine nitrogens pre-organise the macrocycle into useful conformations

This anomalous behaviour was attributed to strong intramolecular hydrogen bonds which are strong in the non-polar medium but can be replaced by solvent interactions in the polar medium. Hence in polar media the macrocycle is less constrained and can readily bind guest, whereas in non-polar media a considerable energetic barrier must be overcome to break the strong internal forces.

1.3.2 Peptide receptors

Control can be exercised over cellular processes (including intercellular communication, and immune responses) through the judicious use of peptides to modulate protein function. A central challenge of supramolecular chemistry is the discovery of synthetic receptors which selectively bind such peptides, in order that their precise function can be ascertained. However, the degree of molecular flexibility inherent in even the simplest dipeptide makes *a priori* rational design of peptide receptors an extremely challenging vernture.[28] A number of peptide receptors have been reported in the literature, of which only a few choice examples are presented here. The reader is directed to the literature for further details.[28,41]

Still and co-workers[42,43] have investigated the properties of a large molecular-weight pair of receptors **25** and **26.** Both these receptors are polycyclic oligomers of a triscarboxyaryl unit and *vic*-diamines, generating a large hydrophobic pocket, and all bind certain peptides selectively in either chloroform or aqueous solution.

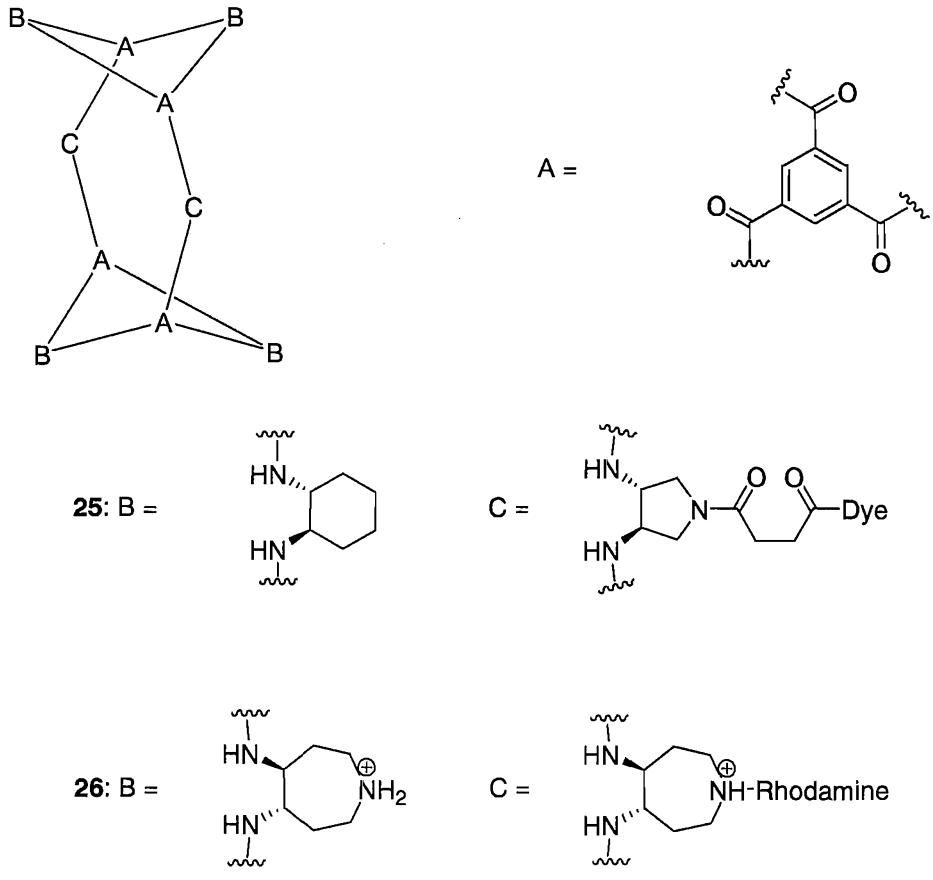
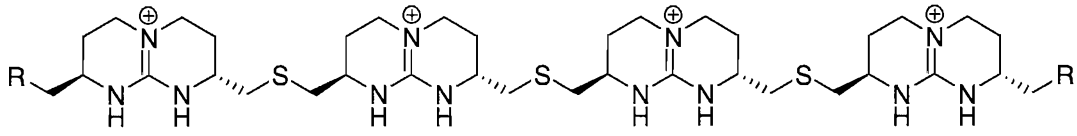


Fig 1-17 Still *et al*'s cage receptors for peptides. The indicated A,B,C units are oligomerised into the full cage structure

For side-chain protected dipeptides, **26** binds similar sequences in water to **25** in chloroform. For deprotected tripeptides, **26** binds different sequences in water to **25** in chloroform. This demonstrates that selective peptide receptors can be generated and that sequence-selective receptors for peptides in aqueous medium can potentially be derived from known sequence-selective receptors for peptides in organic medium.

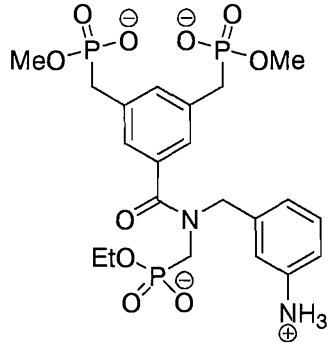
Tetraguanidinium receptor **27** (Fig 1-18), investigated by de Mendoza and co-workers[44,45], was found to strongly bind a 16-mer peptide containing four Asp residues with $K_a \sim 10^5$ M^{-1}. The peptide was stabilised in an α-helical conformation, presumably through binding of the four side-chain carboxylate residues by the four guanidinium moieties.



27: R = OTBDPS

Fig 1-18 de Mendoza *et al*'s tetraguanidinium receptor for Asp-containing peptides

Schrader *et al*[46] have reported a synthetic receptor for the tripeptide Arg-Gly-Asp, which is a known motif in several proteins which adhere to cell surfaces by interacting with integrin proteins[47]. **28** was found to bind Arg-Gly-Asp with K_a = 1300 M^{-1} in water. This binding is ascribed to the *m*-xylene bisphosphonate motif which recognises the guanidinium side-chain of Arg and the ammonium group binding to the carboxylate side-chain of Asp. pH was carefully controlled to ensure the correct protonation state of all entities.



28

Fig 1-19 Schrader *et al*'s receptor for the Arg-Gly-Asp cell adhesion peptide

Many other peptide receptors are based on the concept of the 'molecular tweezer', first described in 1978.[48] Such species consist of a 'head-group' or 'hinge', a conformationally restricted moiety that holds two or three functionalised 'side-arms' in close proximity, generating a partially enclosed binding region reminiscent of macroscopic tweezers. Both the head-group and the side-arms may be functionalised for the binding of a specific guest. In particular, the case where both guest and side-arms are peptidic, it is anticipated that β-sheet structures will form (Fig 1-20).[49]

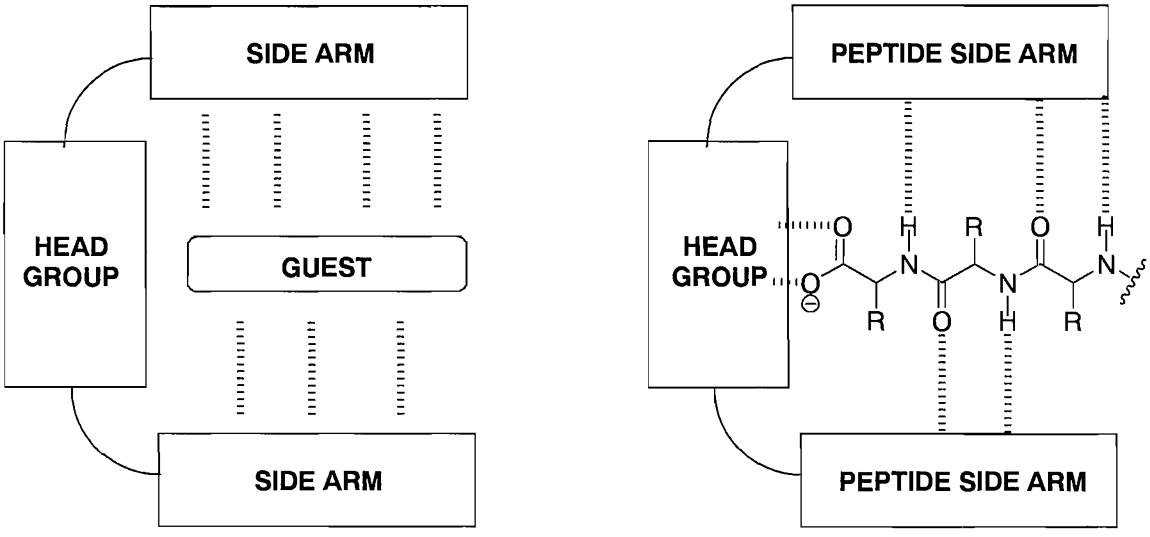
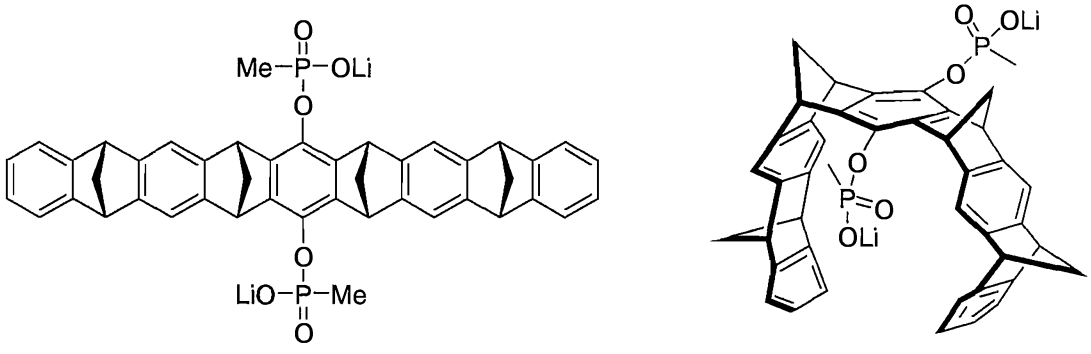


Fig 1-20 Tweezer receptors. A flexible binding cavity is formed by the combination of side-arms and head group. In the case of peptidic side-arms and guests, substantial formation of β-sheet structures may lead to strong binding

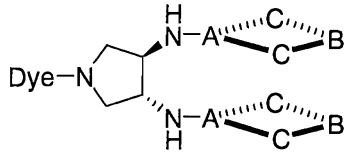
Tweezers are popular as receptor molecules owing to their great versatility. For example, Klärner *et al* have developed tweezer structures based on fused six-membered rings (Fig 1-21) as receptors for Lys and Arg in aqueous medium. It is hypothesised that the alkyl side-chain of these amino acids resides in the space enclosed by the tweezer arms, with the polar functionalities protruding into the bulk solvent, in a structure reminiscent of pseudorotaxanes.[50] Lysine-containing peptides of biological relevance were also bound in aqueous solution.



29

Fig 1-21 Klärner *et al*'s molecular tweezer for the binding of Lys. The curved structure presents a hydrophobic tunnel through which the aliphatic side-chain of Lys can thread

Returning to the binding of peptides, Still *et al* synthesised tweezer receptor **30**, based on *vic*-diamine/isophthalate oligomer side-arms. This receptor was found to bind only two tripeptide sequences in chloroform from 3375 tripeptides assayed.[51]



30

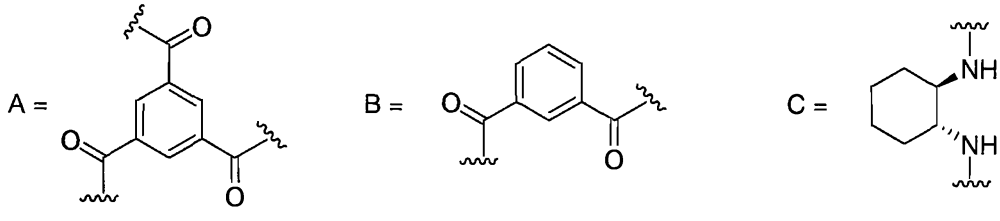


Fig 1-22 Still *et al*'s tweezer receptor for peptide binding in chloroform

Further tweezers by Still *et al*[51] showed distinct selectivity for peptides in water containing the motif L-Gln-D-Leu, and similar motifs where Gln is replaced by Asn. These results confirm the potential for selective binding even in aqueous media.

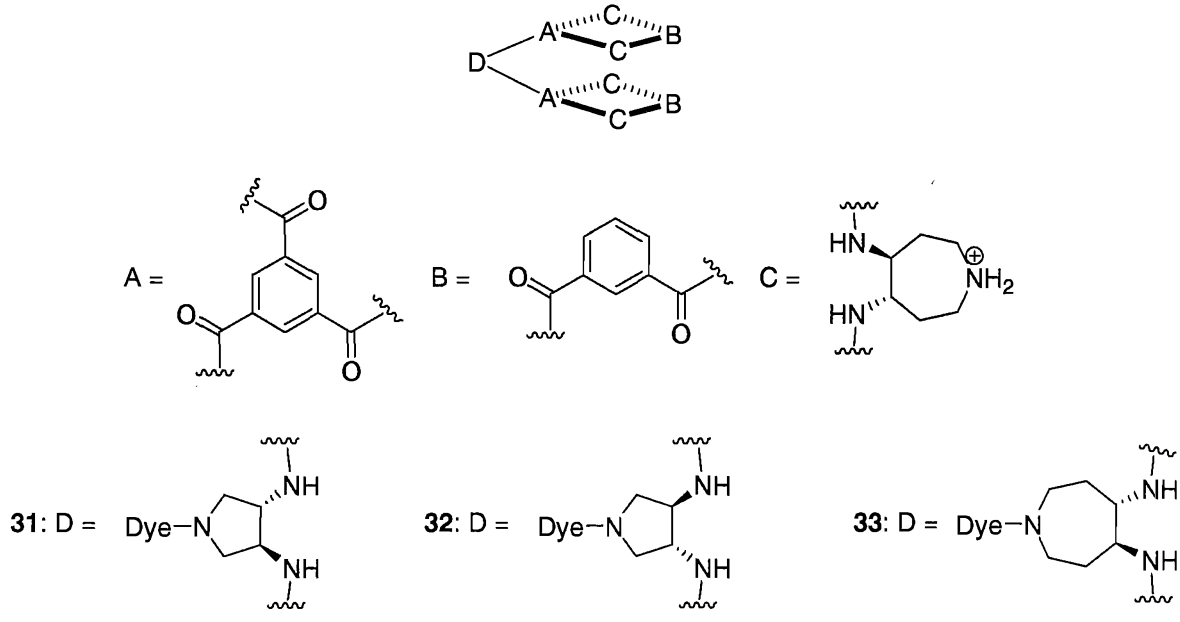
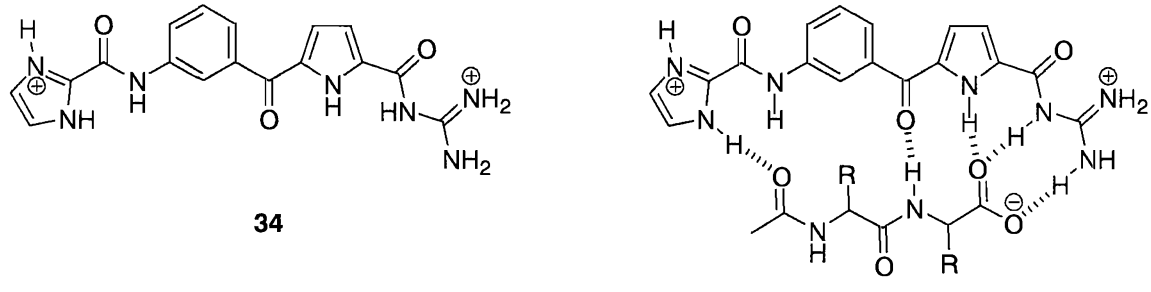


Fig 1-23 Still *et al*'s tweezer receptors for peptide binding in aqueous medium

Schmuck *et al*[52] have designed receptor **34** for dipeptides. The guanidiniopyrrole moiety already described is used as a CBS, with an extended backbone to provide further hydrogen bonding interactions and an imidazole/imidazolium moiety (depending on pH) to bind with an *N*-terminal acetyl group (Fig 1-24).



34

Fig 1-24 Schmuck *et al*'s guanidiniopyrrole receptor for dipeptides in aqueous medium

This receptor was found to bind simple dipeptides in aqueous systems with binding constants of the order 10^4 (determined by UV titration). This binding proved to be an order of magnitude larger than for the simple amino acid constituents of the dipeptides, showing the effect of the secondary hydrogen bonding interactions. Val-Val was the most strongly bound dipeptide of those tested, which was ascribed to the propensity of valine to induce β-sheet conformations[53] and the shielding of the secondary H-bonds from solvent by the bulky side-chains, thus increasing the strength of interaction.[54]

It is thus possible to generate highly selective receptors for peptides based on the tweezer model. A particular advantage of such molecules and peptide guests is that they readily lend themselves to a combinatorial treatment.

1.4 Combinatorial methods

Combinatorial chemistry is now a well-established branch of science, which has revolutionised the practices of the pharmaceutical industry. The beginnings of combinatorial chemistry lie in the advent of solid-phase synthesis of peptides by Merrifield.[5] Its paradigms are now being applied to various other fields of research beyond drug discovery.[56-58]

The main thrust of combinatorial chemistry lies in the consideration of the 'library', a diverse collection of chemical entities. The method divides into three main steps; library preparation; library screening for a desired property; determination of structures of active ('hit') library members. Five distinct methods have been identified for this process, including methods that require deconvolution processes (e.g. an iterative approach[59], positional scanning[60] and recursive deconvolution[61]), spatially-addressable methods[62] and the one-bead-one-compound (OBOC) approach.63

1.4.1 The 'split-and-mix' approach

Combinatorial library synthesis is typically exemplified by the 'split-and-mix' approach, first developed by Furka *et al*[64,65] and expanded upon by Houghten *et al*[59] and Lam *et al.*[66] The procedure may be summarised as follows; solid-support is divided into portions before independent and distinct reaction protocols are carried out on each portion. These portions are then recombined and mixed, thus randomising the distribution of the different resin-loaded entities before re-dividing and further independent reaction protocols (Fig 1-25) This process can be iterated at will, generating a library of maximally n^x members, for x iterations of split-and-mix using n independent reaction protocols.

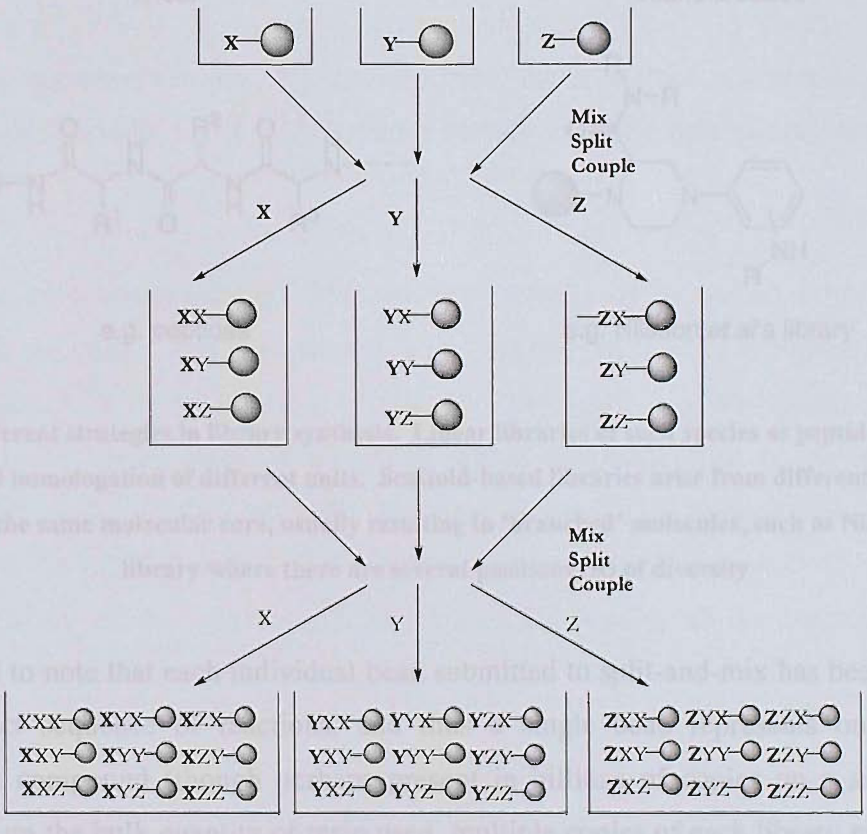


Fig 1-25 Split-and-Mix synthesis. After the first coupling, all resin is mixed and redivided. A second cycle of coupling proceeds. By continuing in this manner, all permutations of a sequence with *n* residues may be obtained over considerably fewer synthetic steps than traditional chemical synthesis.

This process is not limited to linear systems, leading to the use of 'scaffolds' which become differentially functionalised according to reactions carried out. An example is Nilsson *et al*'s library[67] derived from a 4-phenyl-2-carboxypiperzine scaffold to generate potentially bioactive compounds (Fig 1-26).

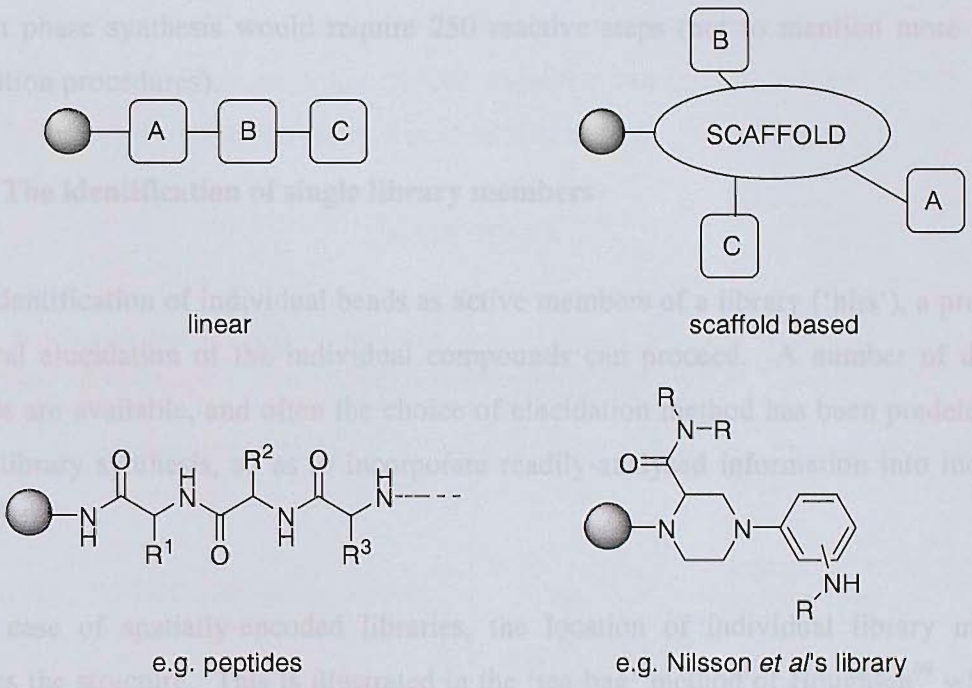


Fig 1-26 Different strategies in library synthesis. Linear libraries of such species as peptides arise from sequential homologation of different units. Scaffold-based libraries arise from different reactions applied to the same molecular core, usually resulting in 'branched' molecules, such as Nilsson *et al*'s library where there are several positions (R) of diversity

It is critical to note that each individual bead submitted to split-and-mix has been subjected to a distinct sequence of reactions, and thus a single bead represents only a single synthesised compound (though perhaps present in billions of copies on a single bead). Depending on the bulk quantity of resin used, multiple copies of each library member may be present in the bulk library; for maximum diversity, a sufficiently large number of individual resin beads must be used to allow the synthesis of each possible combination.

The power of the split-and-mix method lies in its ability to rapidly generate very large numbers of compounds, in comparison to traditional linear solution-phase synthesis. For example, using a choice of five monomers for each position (resin splits into five portions), a library of 125 trimers can be synthesised in 15 reactive steps with minimal purification procedures associated with solid-phase chemistries. To create the same 125 trimers by solution phase synthesis would require 250 reactive steps (not to mention more lengthy purification procedures).

1.4.2 The identification of single library members

Upon identification of individual beads as active members of a library ('hits'), a process of structural elucidation of the individual compounds can proceed. A number of different methods are available, and often the choice of elucidation method has been predetermined before library synthesis, so as to incorporate readily-analysed information into individual beads.[68]

In the case of spatially-encoded libraries, the location of individual library members specifies the structure. This is illustrated in the 'tea-bag' method of Houghten[69] where the synthetic history of a porous sachet of resin can be recorded on the sachet. Hence the structure of the library member contained within the sachet can be determined. A sophistication of this method has been developed as the IRORI system[70] whereby porous reactor units are labelled with a radio-frequency tag which allows both identification of synthetic history of the reactor and also automated sorting of the reactor units into appropriate batches for synthetic procedures.

Chemically-encoded libraries use chemical information supplied by bead-based entities to determine the structure of a library member. The simplest form of this encoding involves merely cleaving the library member from the solid-phase and using standard analytical methods, assisted by knowledge of the potential chemical entities represented. However, the limited material available from a single resin bead is often insufficient in both quantity

and purity to make a reliable assignation (except in the case of libraries of oligonucleotides, where the polymerase chain reaction can generate a larger quantity for analysis).

For the analysis of *N*-deprotected peptide libraries (using only α–amino acids), recourse has been made to Edman sequencing methodologies.[71] Amino acids are sequentially digested from the *N*-terminus of the peptide and digestion products characterised by HPLC at picomolar quantity. Thus an entire peptide sequence can be determined. However, this process is time-consuming and expensive to implement.

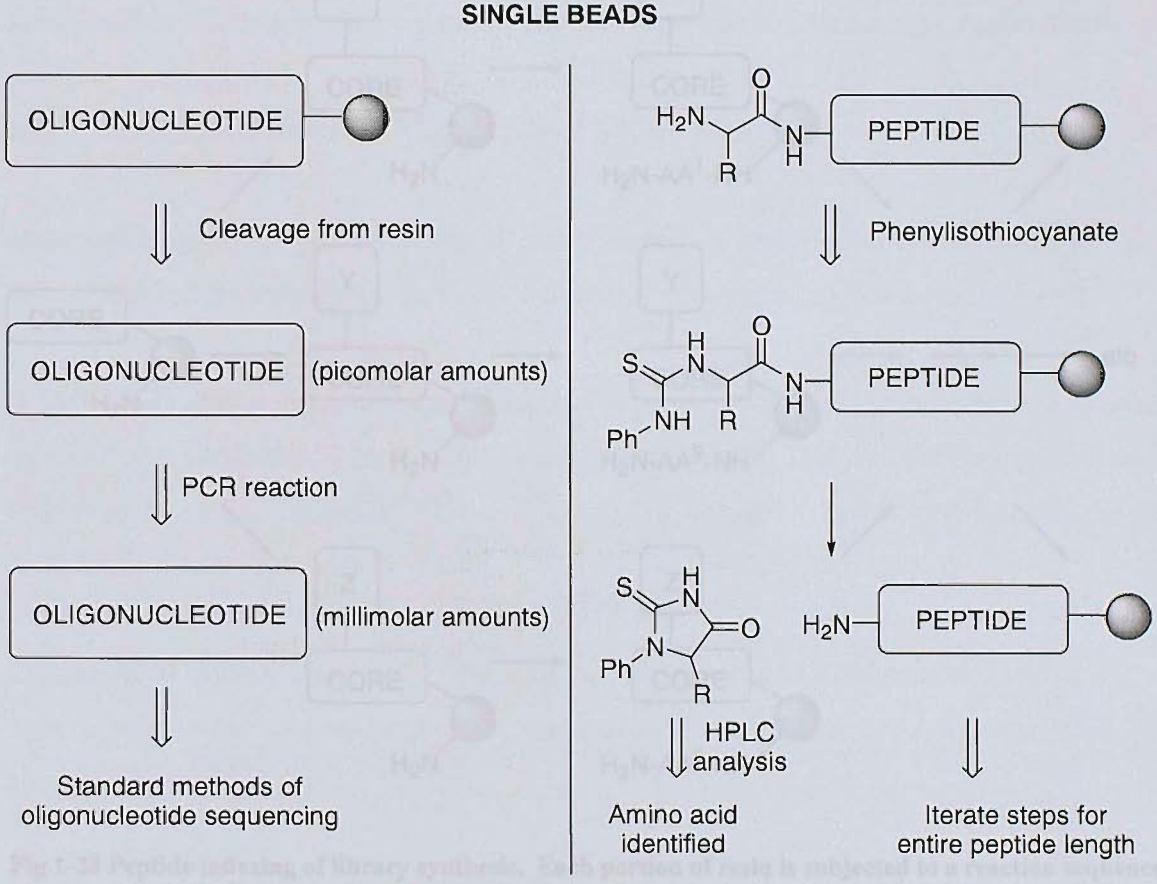


Fig 1-27 Sequencing of single-bead oligonucleotides (by amplification and standard analysis) and peptides (by Edman sequencing)

Both of these above methods can be used to index synthetic history of molecules of non-oligonucleotide/non-peptidic nature.[72-75] In tandem with the synthesis of the desired library

members, an oligonucleotide sequence or peptide sequence is synthesised. Individual nucleobases/amino acids are related to the diverse synthetic steps used to create library members (e.g. Gly codes for reaction A, Ala for reaction B etc). The synthetic history of an individual bead can thus be determined by finding the peptide/nucleotide sequence and 'translating' into the structure of a library member. This method requires that the library synthetic steps are orthogonal to the biomolecule synthesis, which limits this approach in many ways.

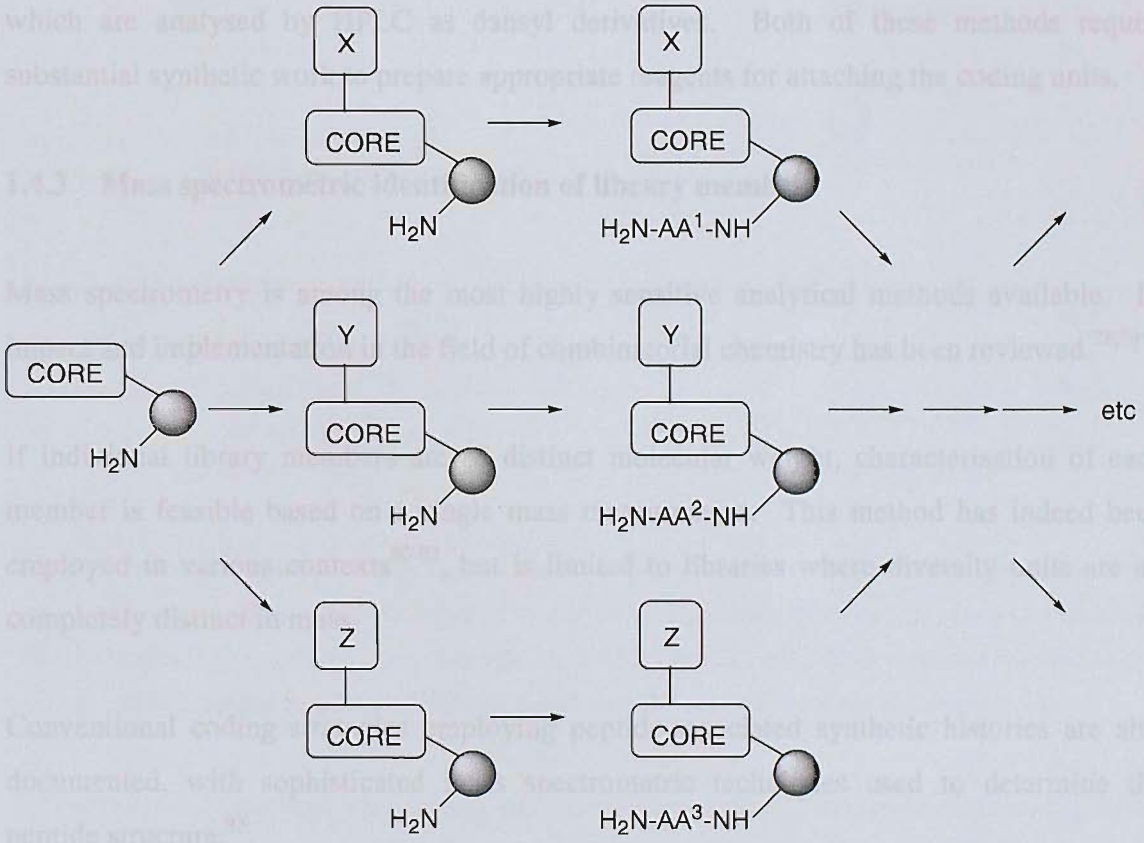


Fig 1-28 Peptide indexing of library synthesis. Each portion of resin is subjected to a reaction sequence (to install X,Y,Z) followed by coupling of a distinct amino acid ($AA^{1/2/3}$). Which reaction a bead has undergone is recorded on the bead by the presence of a specific amino acid. Continued cycles build up the library and peptide chain, which is then sequenced and 'translated' into a reaction history. Oligonucleotides can be used instead of peptides to the same effect.

The coding for synthetic steps using chemical entities is not limited to linear biomolecules. Alternative species can be used to code for synthetic steps in a 'molecular barcode' for the library member. Still *et al*[76] have used chemically inert halophenol derivatives that are directly attached to the polystyrene resin matrix which can be oxidatively cleaved and analysed in subpicomolar quantity by electron capture gas chromatography. This requires a degree of care in both attaching the appropriate coding unit and in the cleavage and analysis thereof. Gallop *et al*[77] have devised a method of coding based on secondary amines (which exist as secondary amides on the solid-support and are cleaved by hydrolysis for analysis) which are analysed by HPLC as dansyl derivatives. Both of these methods require substantial synthetic work to prepare appropriate reagents for attaching the coding units.

1.4.3 Mass spectrometric identification of library members

Mass spectrometry is among the most highly sensitive analytical methods available. Its impact and implementation in the field of combinatorial chemistry has been reviewed.[78,79]

If individual library members are of distinct molecular weight, characterisation of each member is feasible based on a single mass measurement. This method has indeed been employed in various contexts[80-82], but is limited to libraries where diversity units are all completely distinct in mass.

Conventional coding strategies employing peptide-associated synthetic histories are also documented, with sophisticated mass spectrometric techniques used to determine the peptide structure.[83]

A more useful method to determine the sequence of bead-based peptides has been described by Youngquist *et al.*[84] This method relies on the generation of 'ladder sequences' during the synthesis of the peptide on the solid phase (Fig 1-29). During synthesis of a library, a small percentage of the growing peptide chains are capped at each step by Ac-Ala to generate a series of termination products on the bead. These termination products ('ladder sequences') allow peptide sequence determination by consideration of the

mass differences between consecutive products; since the mass of each amino acid is generally distinct, a specific sequence can be simply read off from a single mass analysis. Arginine is specifically incorporated into a linker unit before split-and-mix peptide synthesis begins, to allow detection of 'ladder sequences' with great sensitivity. Further investigation by St. Hilaire *et al*[85] developed the use of matched capping groups to account for varying kinetics of different library members.

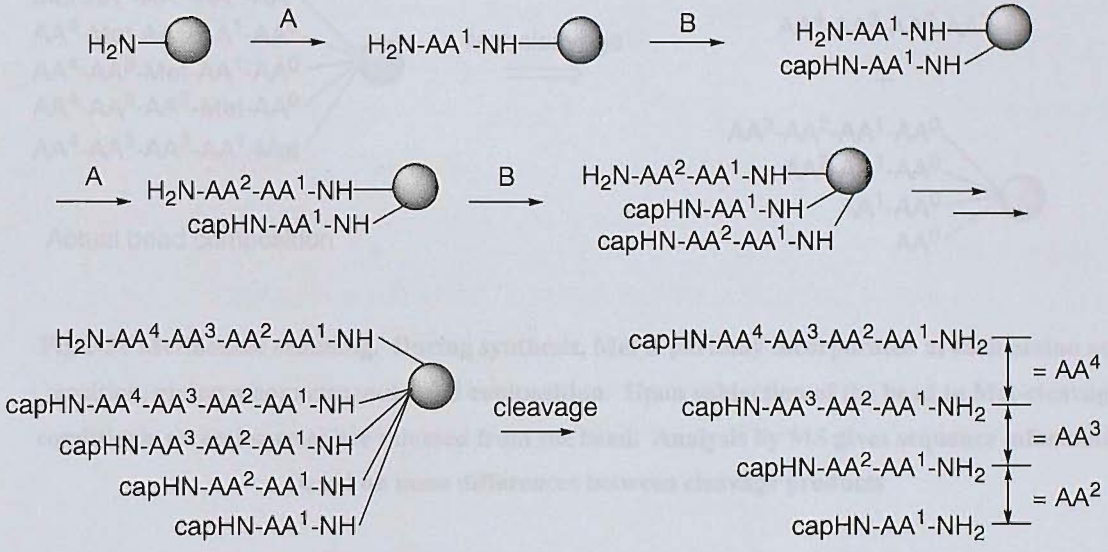


Fig 1-29 Generation of ladder sequences for rapid analysis of peptides from libraries. Two basic steps are used; amino acid coupling (A) followed by partial termination with a capping group (B), iterated for the length of the peptide. Upon cleavage, the product mixture includes terminated species which differ in mass by a single amino acid and thus the peptide sequence can be inferred

Youngquist *et al* generate the appropriate 'ladder sequences' during the course of peptide synthesis by incorporation of a capping group. Alternative methods for generating the truncated sequences have also been recorded. Bradley *et al* generate the truncated sequences for analysis by a process of 'methionine scanning'.[86] Peptide synthesis proceeds as usual, with the addition of a proportion of methionine to each coupling mixture such that Met is partially incorporated into the growing sequence at each position. To identify the peptide sequence, a single bead is subjected to CNBr-mediated cleavage of Met giving rise to cleavage products of different mass, with the mass differences again corresponding to an individual amino acid (Fig 1-30).

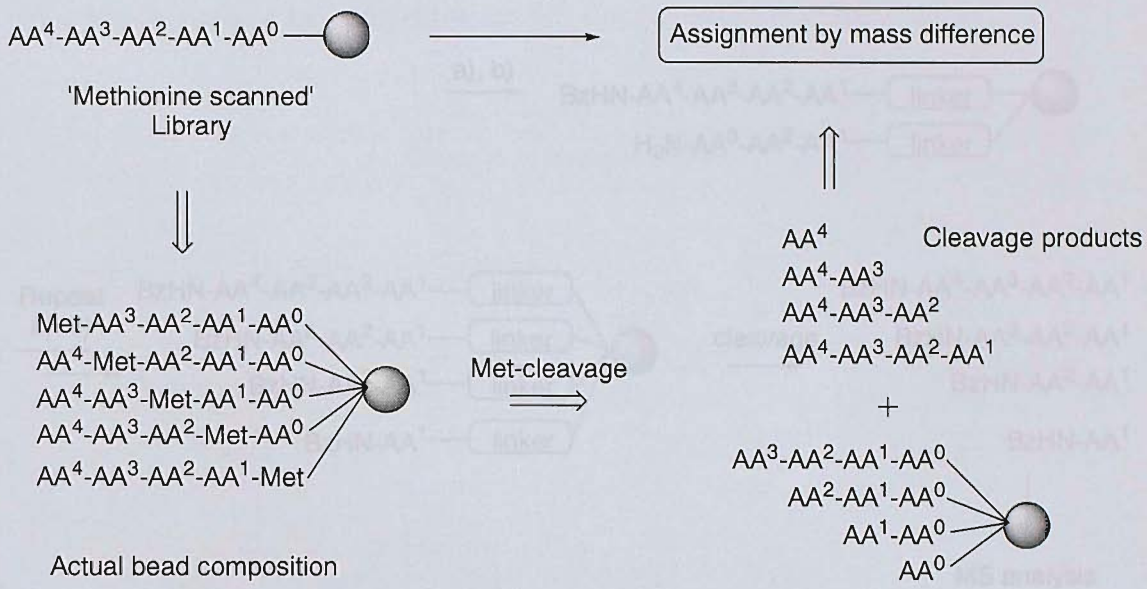
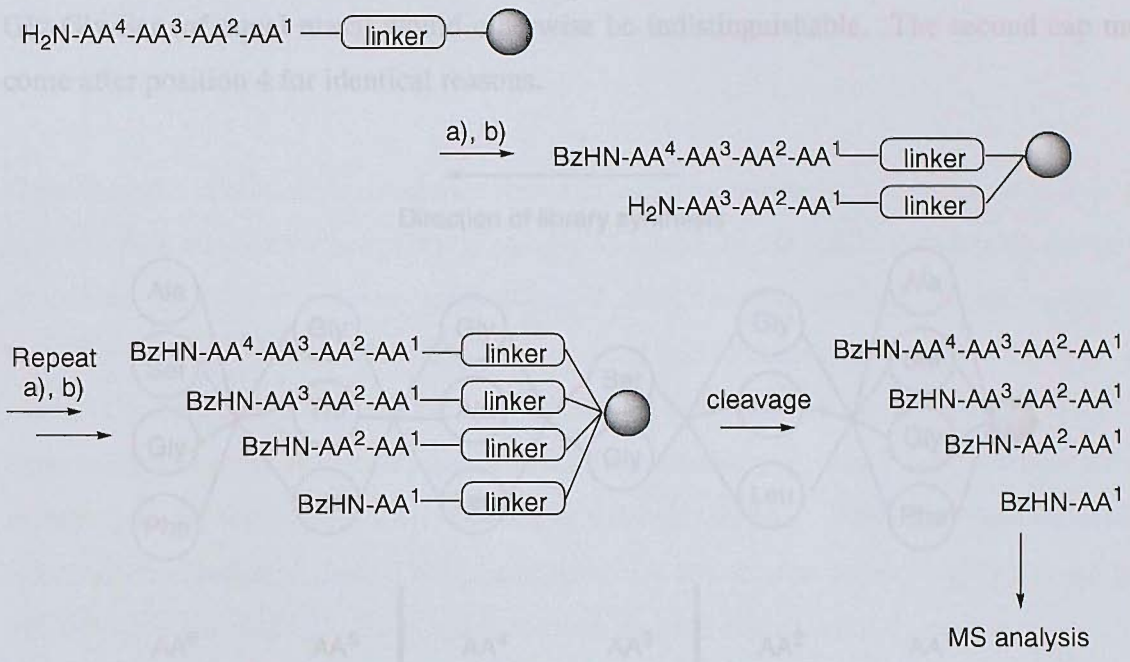


Fig 1-30 Methionine scanning. During synthesis, Met is partially incorporated at each amino acid position, giving a homogeneous bead composition. Upon subjection of the bead to Met-cleavage conditions, several species are released from the bead. Analysis by MS gives sequence information based on mass differences between cleavage products

Pei *et al*[87,88] utilise a post-screening process of Edman degradation to generate 'ladder sequences' (Fig 1-31). To prepare a bead for analysis, the full length peptide is partially capped with an inert group before Edman degradation is applied to the remaining free peptide to remove the *N*-terminal residue.



Reagents and conditions: a) PhNCS/Bz-NHS in fixed ratio, b) TFA

Fig 1-31 Post-screening Edman degradation and MS sequencing of peptides. The ratio of PhNCS to Bz-NHS allows partial degradation at each step, generating 'ladder sequences' which can be analysed by MS upon total cleavage of material

This process is iterated for the length of the peptide to be sequenced, generating the appropriate 'ladder sequences' to be analysed upon cleavage of all material from the resin. An advantage of this method is that beads present a homogeneous composition of the library member (devoid of 'ladder sequences' or chemical coding entities) for the screening process.

In order to minimise the number of 'ladder sequences' necessary for full characterisation of a peptide, Griesinger *et al* describe[89] the use of a computer algorithm ('Biblio'). Depending on library composition, it is perhaps unnecessary to cap the growing sequence after the addition of each individual residue, instead capping after the incorporation of a dipeptide unit or potentially longer units. The requirement is that each unit must still be uniquely distinguishable from every other possible unit. For example, in Fig 1-32 the first cap must come after position 2 since the potential tripeptide sequences Ser-Gly-Ser and

Gly-Gly-Ser (of equal mass) would otherwise be indistinguishable. The second cap must come after position 4 for identical reasons.

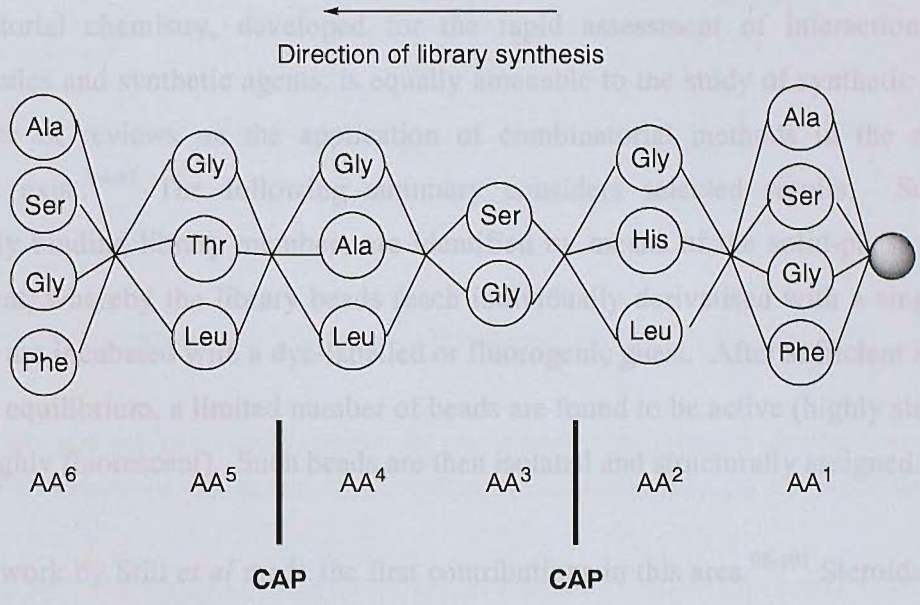


Fig 1-32 Minimal generation of ladder sequences to allow full sequencing. Depending on library composition, it is not necessary to cap after every amino acid, rather at positions in order to prevent oligopeptide units of identical mass being generated (which would be indistinguishable by MS)

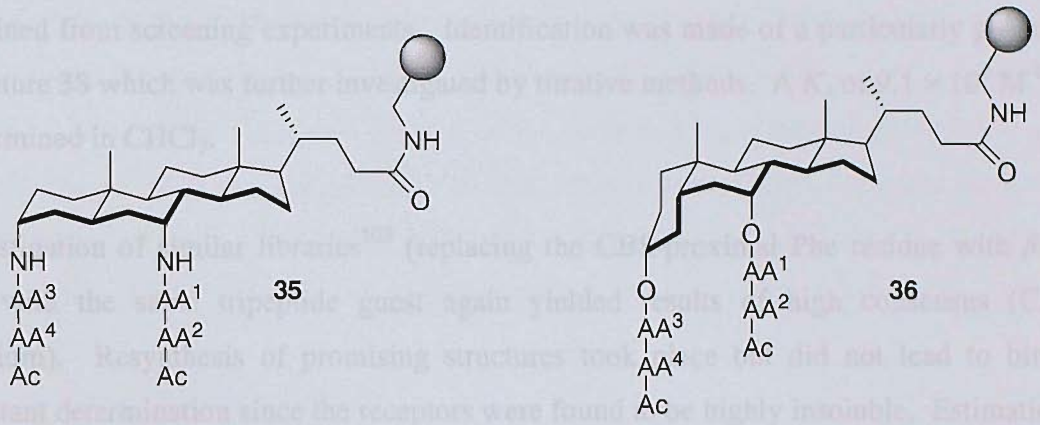
'Biblio' will perform this analysis given the details of library composition, and in fact through nomination of two or more distinct capping groups for use (rather than the same capping group for every possibility) can reduce the number or capping steps required, though this makes sequence read-out from mass spectrometric data a non-trivial exercise.

'Ladder sequence' coding by peptides has been used identify terbium(III)-binding peptides[90], small-molecule streptavidin ligands[91] and triazole-based inhibitors of a cysteine protease[92], amongst others. The methodology represents a compromise between rapid and simple determination of library members and the sophistication needed to encode libraries of considerable size. Further developments continue to appear.[66,93]

1.5 Combinatorial methods applied to the study of receptors for carboxylates and peptides

Combinatorial chemistry, developed for the rapid assessment of interaction between biomolecules and synthetic agents, is equally amenable to the study of synthetic receptors. A number of reviews on the application of combinatorial methods in the search for receptors exist.[94-97] The following summary considers selected results. Strong and selectively binding library members are identified by means of the solid-phase screening experiment, whereby the library beads (each individually derivatised with a single library member) are incubated with a dye-labelled or fluorogenic guest. After sufficient incubation to obtain equilibrium, a limited number of beads are found to be active (highly stained with dye or highly fluorescent). Such beads are then isolated and structurally assigned.

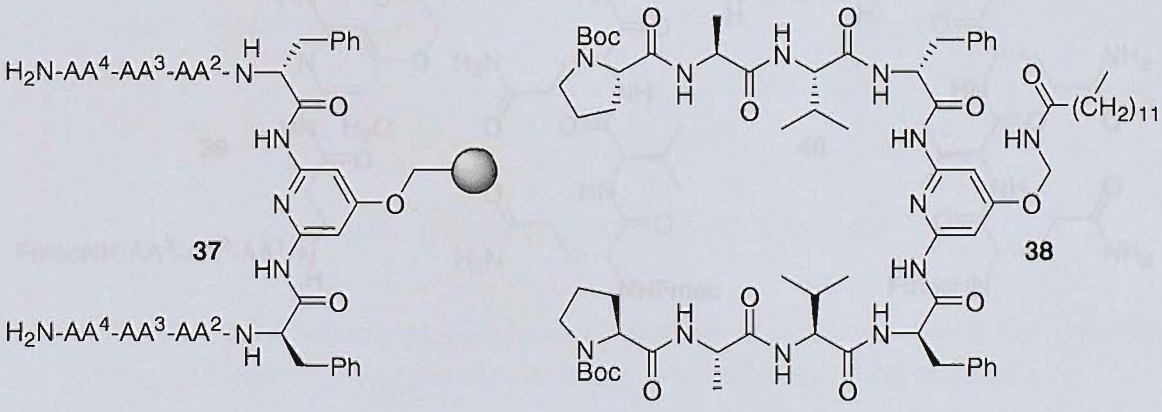
Original work by Still *et al* made the first contributions in this area.[98-101] Steroidal cores or varying conformation were attached to the solid phase and derivatised at two positions with a randomised peptide chain (Fig 1-33) Such species represent a tweezer receptor. Screening experiments of such libraries with a dye-labelled pentapeptide 'Leu-enkephalin' yielded a high consensus amongst 'hit' structures in the case of the A,B-*trans*-steroidal core **35**, showing that receptor libraries can be used to rapidly converge onto selective receptor molecules. Comparisons with results from similar libraries based on the A,B-*cis*-cholic acid core **36** show the importance of decreased receptor flexibility for selectivity of binding.



Leu-enkephalin pentapeptide: Dye-spacer-Tyr-Gly-Gly-Phe-Leu-OMe

Fig 1-33 Still *et al*'s steroidal based tweezers for the Leu-enkephalin peptide. Results from screening the *trans*-steroidal core library 35 had high consensus, results from the *cis*-core 36 less so

Kilburn *et al*[102] synthesised a tweezer library **37** based on a diamidopyridine CBS core with two identical peptide side-arms (Fig 1-34).



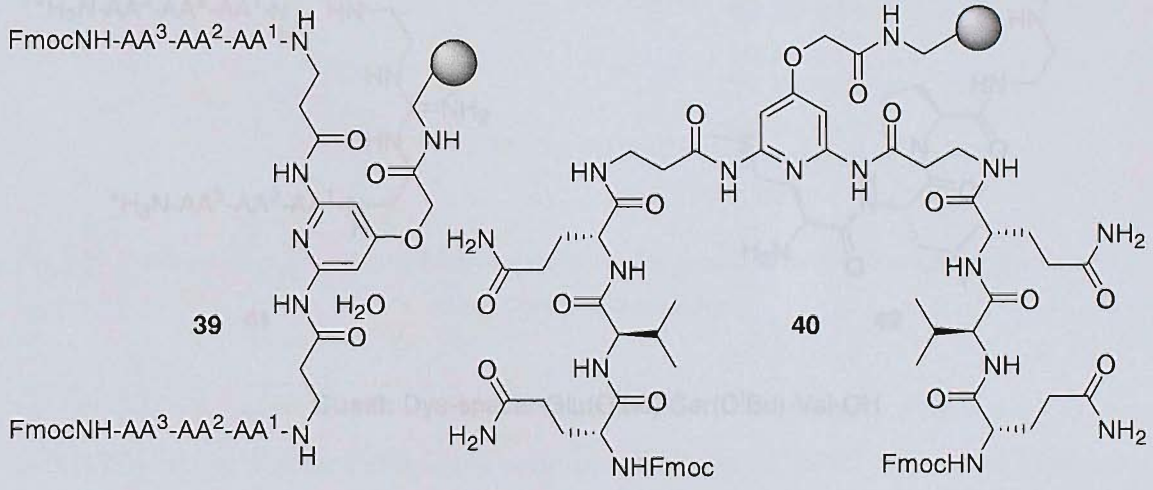
Guest: Dye-Glu(O^{t}Bu)-Ser(O^{t}Bu)-Val-OH

Fig 1-34 Kilburn *et al*'s library of diamidopyridine-based peptide receptors 37 and a specific receptor 38 derived from screening experiments which binds well to the indicated guest.

This library was screened with a variety of dye- and fluorescence labelled derivatives of the sequence Glu(O^{t}Bu)-Ser(O^{t}Bu)-Val (both side-chain protected and deprotected forms). It was noted that the nature of the chromophore label had a certain effect on the results

obtained from screening experiments. Identification was made of a particularly promising structure **38** which was further investigated by titrative methods. A K_a of 9.1×10^3 M^{-1} was determined in $CHCl_3$.

Investigation of similar libraries[103] (replacing the CBS-proximal Phe residue with β-Ala, **39**) with the same tripeptide guest again yielded results of high consensus ($CHCl_3$ medium). Resynthesis of promising structures took place but did not lead to binding constant determination since the receptors were found to be highly insoluble. Estimation of solid-phase binding constants of **40** by HPLC assay gave $K_a \sim 3 \times 10^5$ for the deprotected Glu-Ser-Val and $K_a \sim 1 \times 10^3$ for the protected guest. Estimation for the enantiomer D-Glu-D-Ser-D-Val gave $K_a \sim 9 \times 10^3$, showing a degree of enantioselective binding by this receptor.

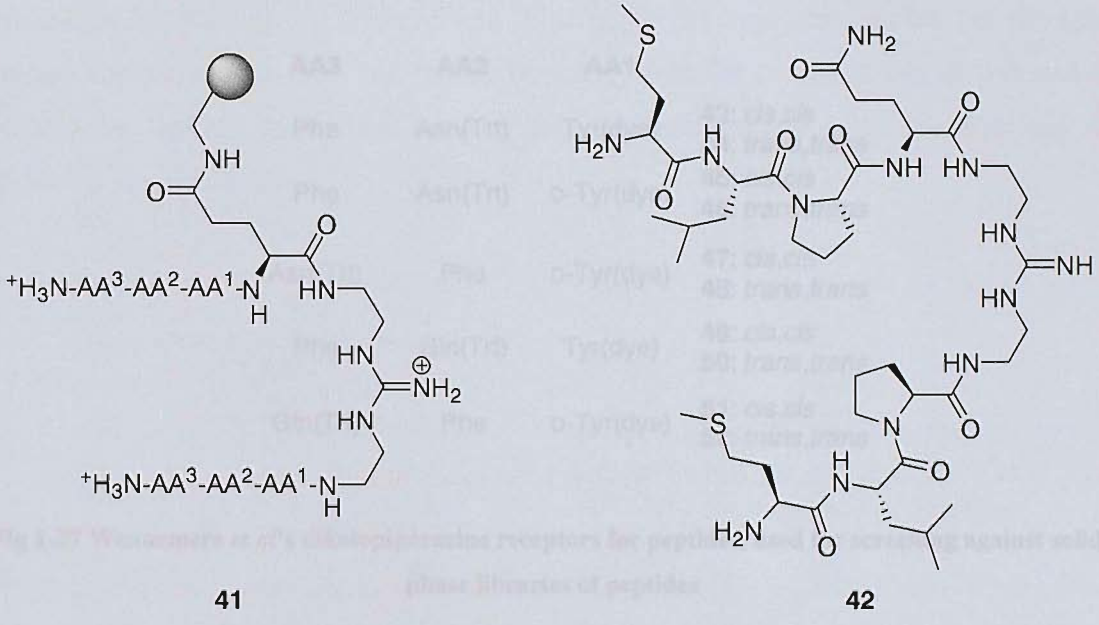


Guest: Dye-spacer-Glu-Ser-Val-OH

Fig 1-35 Kilburn *et al*'s library of more flexible diamidopyridine-based peptide receptors 39 and a specific receptor 40 derived from screening experiments which binds well to the indicated guest

A more flexible CBS based on alkylguanidiniums[104] was also used to investigate the Glu-Ser-Val guest. A symmetrical library **41** was screened against a dye-labelled derivative of this tripeptide in aqueous medium and appeared to give high selectivity, however less consensus was noted in 'hit' structures than in the screens of previous libraries in $CHCl_3$.

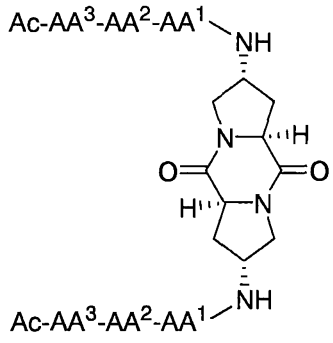
Whether this is due to the less conformationally defined nature of the receptors or to the properties of the medium is unclear. A consensus result was determined and resynthesis of **42** took place. Enantioselective binding of the protected Glu(O^{t}Bu)-Ser(O^{t}Bu)-Val guest (K_a = 8.2 × 10^4 M^{-1}) over the protected D-Glu(O^{t}Bu)-D-Ser(O^{t}Bu)-D-Val guest (K_a = 8.0 × 10^3 M^{-1}) was noted by UV titration, and the side-chain deprotected guest was found not to bind. Such results indicate the potential power of such systems for highly selective binding.



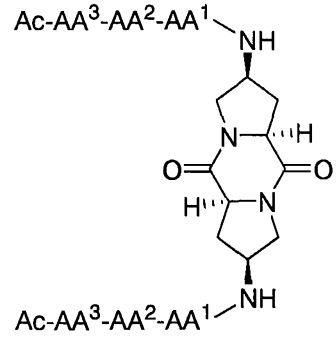
Guest: Dye-spacer-Glu(O^{t}Bu)-Ser(O^{t}Bu)-Val-OH

Fig 1-36 Kilburn *et al*'s flexible alkylguanidinium-based library of peptide receptors 41 and a specific receptor 42 identified from screening experiments which binds well to the indicated guest

Screening experiments have also been carried out in the reverse sense, to find the binding selectivity properties of a given receptor amongst a combinatorial library of guests. Wennemers *et al* have investigated[105,106] tweezer receptors based on a diketopiperazine head-group.



trans, trans-scaffold



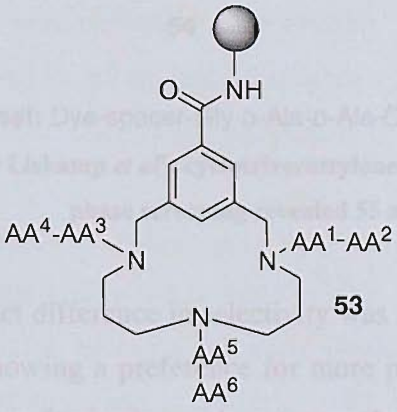
cis, cis-scaffold

AA3	AA2	AA1	
Phe	Asn(Trt)	Tyr(dye)	**43**: *cis,cis* **44**: *trans,trans*
Phe	Asn(Trt)	D-Tyr(dye)	**45**: *cis,cis* **46**: *trans,trans*
Asn(Trt)	Phe	D-Tyr(dye)	**47**: *cis,cis* **48**: *trans,trans*
Phe	Gln(Trt)	Tyr(dye)	**49**: *cis,cis* **50**: *trans,trans*
Gln(Trt)	Phe	D-Tyr(dye)	**51**: *cis,cis* **52**: *trans,trans*

Fig 1-37 Wennemers *et al*'s diketopiperazine receptors for peptides, used for screening against solid-phase libraries of peptides

Dye-labelled receptors **43-52** (on the side chain of a Tyr-residue) were found to have distinct binding preferences amongst a solid-phase library of 24389 *N*-α-acetyl tripeptides. Subtle differences in receptor structure gave rise to significantly different binding distributions. Critically, the geometry of the head-group (*cis,cis*- vs. *trans-trans*) was found to determine whether the receptors showed a great degree of selectivity or no, with the rigidly U-shaped *trans,trans*- receptors showing the best results, presumably due to enforcing a tweezer-conformation with ideal separation between peptide strands. The *cis,cis*-isomer is more linear in shape and less pre-disposed to convergent binding functionality. The methods presented exemplify a methodology used to investigate the binding properties of receptor-classes.

Recent work by Liskamp *et al*[107-109] with cyclotriveratrylene and triazacyclophane scaffolds appended with three peptide arms has discovered highly selective receptors for the motif D-Ala-D-Ala. A library of receptors **53** based on a triazacyclophane scaffold with three distinct side-arms was prepared[108] (Fig 1-38) and screened with a fluorescent-labelled D-Ala-D-Ala compound in buffered aqueous medium. High homology was obtained from 'hit' beads; The motif [AA^1 = Phe, AA^2 = Lys, AA^5 = Ile, AA^6 = Arg] was common, as were the residue combinations solely at AA^1/AA^2 or AA^5/AA^6, but AA^3/AA^4 showed less consensus. No independent binding constants were measured. The results obtained were rationalised by considering the creation of a highly hydrophobic pocket for the guest between the arms with most noticeable homology (similar to enzymatic clefts) and the potential for ion-ion interactions between the carboxylate guest and basic Lys and Arg residues in the side-arms.



53

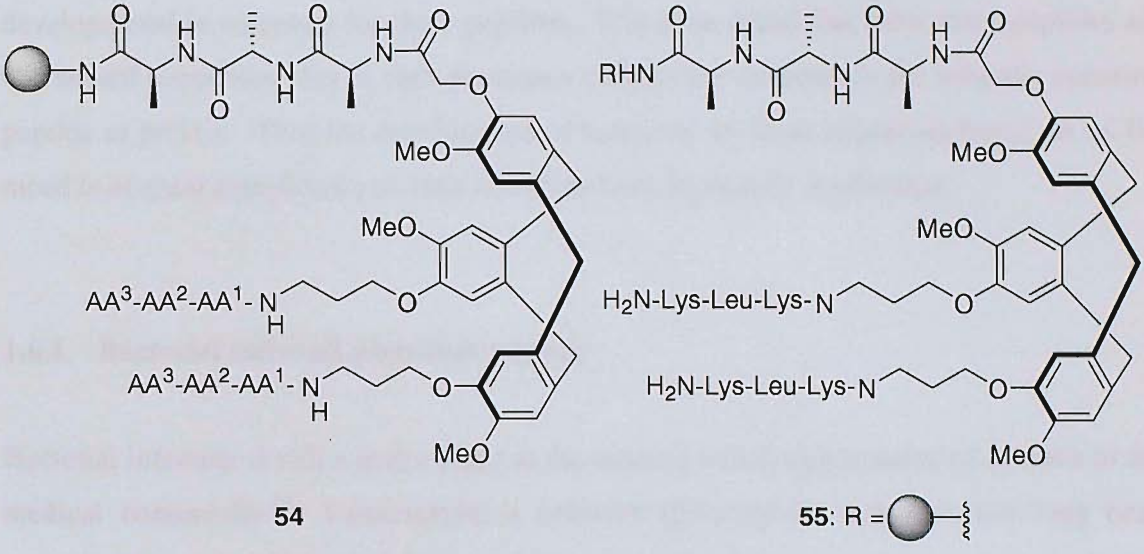
AA^1, AA^2 = Lys, Glu, Val, Phe, Gly, Ser
AA^3, AA^4 = His, Ala, Leu, Pro, Tyr, Asn
AA^5, AA^6 = Arg, Asp, Ile, Trp, Thr, Gln

Guest: Dye-spacer-Gly-D-Ala-D-Ala-OH

Fig 1-38 Liskamp *et al*'s tripodal triazacyclophane-based library 53. Each arm consists of a distinct selection of amino acids. Solid-phase screening revealed strong binding structures of the indicated guest

A library **54** based on the cyclotriveratrylene scaffold with three peptide side-arms was also prepared (Fig 1-39).[109] In this case only two of the side-arms consist of (identical) randomised sequences, the third arm being Ala-Ala-Ala for every library member.

Screening was performed with the tripeptide L-Gly-D-Ala-D-Ala and the ester analogue L-Gly-D-Ala-D-Lac labelled with a selection of chromophores (fluorogenic and dye-based).



54

Guest: Dye-spacer-Gly-D-Ala-D-Ala-OH

56: R = H

Fig 1-39 Liskamp *et al*'s cyclotriveratrylene-based library 54 for D-Ala-D-Ala and D-Ala-D-Lac. Solid-phase screening revealed 55 as a binding structure for the guest indicated

A distinct difference in selectivity was noted between the two guests, with the D-Ala-D-Lac guest showing a preference for more polar residues and the D-Ala-D-Ala guest showing a preference for basic residues (in particular Lys). There was a complete absence of Lys in the 'hit' sequences for the D-Ala-D-Lac guest. A slight difference in selectivity was noted depending on the nature of the chromophore label. The differences are all the more remarkable considering the minimal difference between the two motifs (amide NH *vs* ester O). A selected receptor was resynthesised on the solid phase (**55**) and binding was confirmed by experiments with an appropriate fluorescent guest, which also established that recognition is not due to the chromophore itself. Again, no absolute binding constant was measured. Bioassay of the non-solid phase linked receptor (**56**) showed no inhibition of bacterial growth, suggesting that the binding constant is low.

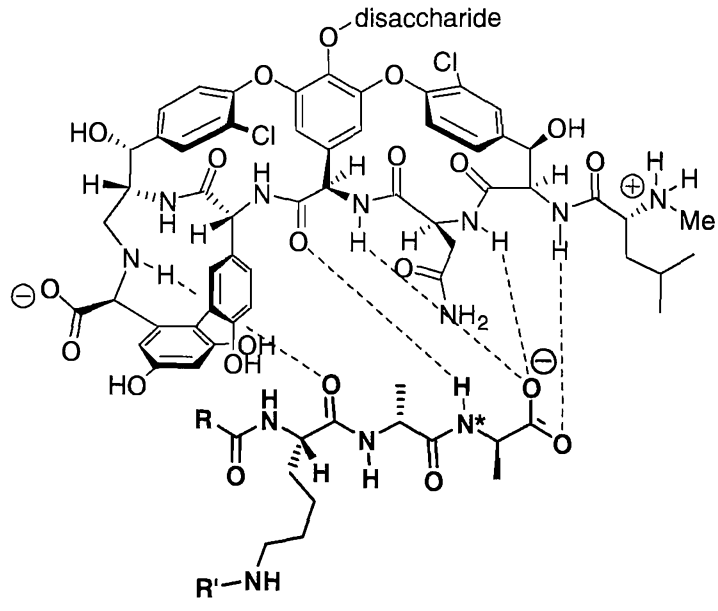
1.6 Biologically interesting oligopeptide carboxylates

The study of peptides is a major endeavour within science, including the development of molecular receptors for specific peptide sequences of biological interest.[110] Three such peptide sequences currently under study are now briefly discussed below, including any developments in receptors for these peptides. It is to be noted that these three peptides are *C*-terminal sequences, that is, each presents a carboxylate terminus in the naturally occuring peptide or protein. Thus the development of receptors for these sequences based on a CBS motif is of great significance as such receptors have immediate application.

1.6.1 Bacterial cell-wall precursor peptide

Bacterial infection is still a major killer in the modern world, and remains of concern to the medical community.[111] Vancomycin, a complex glycopeptide antibiotic has long been considered as a last line of defence against bacterial infections as for many years it was used without observation of resistance.

Vancomycin **57** is known to function through highly-specific non-covalent binding to the *C*-terminal D-Ala-D-Ala motif of the uridine diphosphate-muramyl pentapeptide that is a precursor of bacterial cell wall synthesis.[112] In many strains of bacteria, the C-terminal tripeptide sequence is conserved as Lys-D-Ala-D-Ala. The binding interaction consists of five specific hydrogen bonding interactions[113] (Fig 1-40). Due to binding, the action of transglycosidase enzymes that assemble the cell wall from the D-Ala-D-Ala fragments is prevented, thus preventing normal formation of the bacterial cell wall and causing cell lysis.



57

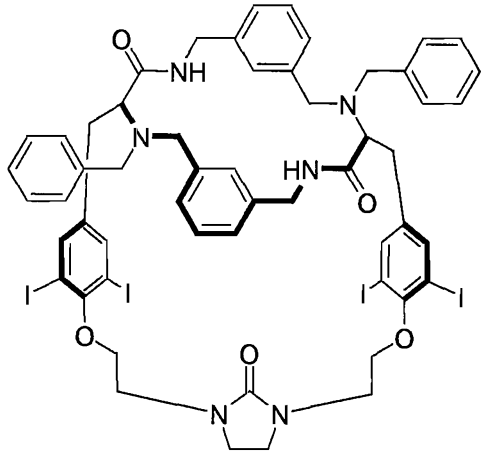
58

Fig 1-40 Vancomycin 57 and the five hydrogen bonding interactions with Lys-D-Ala-D-Ala 58. Where the NH* unit is replaced by O in Lys-D-Ala-D-Lac, only one hydrogen bond is lost but binding affinity decreases sharply

The emergence of vancomycin-resistant strains of bacteria has been traced to a mutation in the *C*-terminus of the cell-wall precursor peptide. The resulting mutant now presents the sequence Lys-D-Ala-D-Lac which is bound 1000 times less efficiently than the D-Ala-D-Ala sequence.[112] This is surprising considering the conservative nature of the change and the loss of only one hydrogen bonding interaction[114] (though it is replaced by a repulsive interaction).

Since the mode of action has been traced to the binding of a peptide, significant research has been invested in developing receptors for the D-Ala-D-Ala sequence. Vancomycin, being a glycopeptide, has proven amenable to a combinatorial approach to modifying the aglycon, peptidic portion[115] to generate receptors with significant affinity for both Ac-Lys-D-Ala-D-Ala and Ac-Lys-D-Ala-D-Lac. Such structures are however still highly complex. More simplified binding pockets[116,117] resembling the vancomycin core were found to bind carboxylic acids, suggesting a minimal CBS as a constituent of vancomycin in addition to the hydrogen bonding pattern.

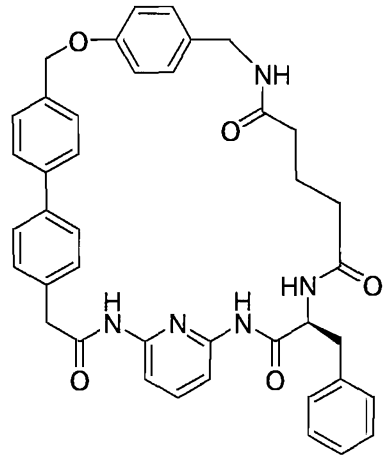
There are few non-vancomycin inspired receptors for the D-Ala-D-Ala motif reported. Still *et al* reported as early as 1990[118] a macrobicyclic structure **59** which bound the benzyl ester of Ac-D-Ala-D-Ala, but binding with the enantiomeric Ac-L-Ala-L-Ala in $CDCl_3$ was stronger.



59

Fig 1-41 Still *et al*'s macrobicyclic receptor 59 for alanylalanine in $CDCl_3$

Kilburn *et al* synthesised a simpler macrocyclic system **60**[119] based on a diamidopyridine CBS, which was found to efficiently bind various amino acids and dipeptides in $CDCl_3$ solution, including the free carboxylic acids of Cbz-L-Ala-L-Ala and Cbz-D-Ala-D-Ala. Again, the L-enantiomer was bound more strongly.



60

Fig 1-42 Kilburn *et al*'s receptor 60 for alanylalanine in $CDCl_3$. Cbz-L-Ala-L-Ala is bound more strongly than Cbz-D-Ala-D-Ala

Macrobicyclic receptors[120,121] containing the diamidopyridine unit have also been synthesised. For the flexible version with an alkyl spacer in the bridging ring **61**, binding (in $CDCl_3$) was found to be stronger with Cbz-D-Ala-D-Ala than Cbz-L-Ala-L-Ala. However the more rigid aryl-spaced version **62** was found to associate more strongly with the L-enantiomer (though association with both enantiomers were stronger than with the flexible version). The reasons for this reversal in selectivity are unclear.

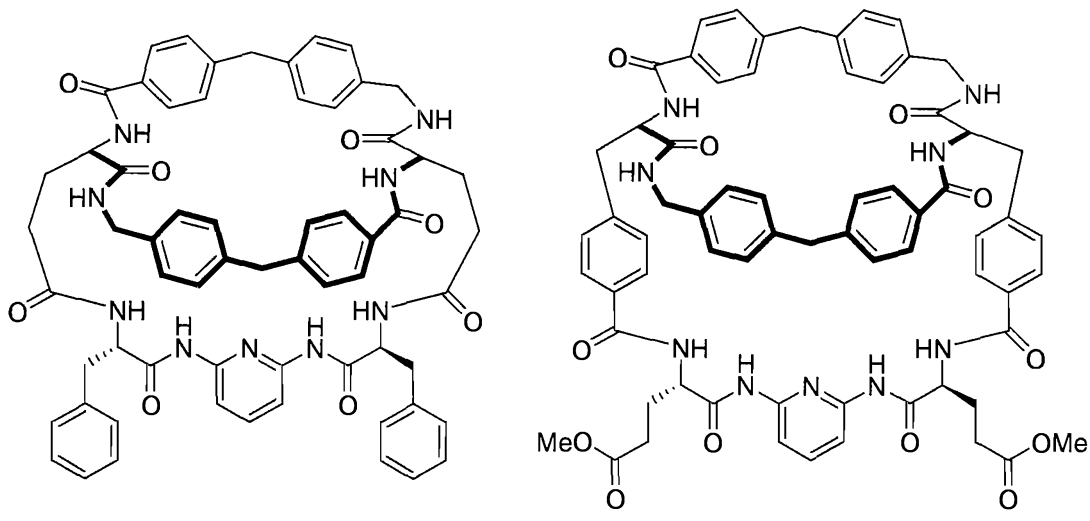
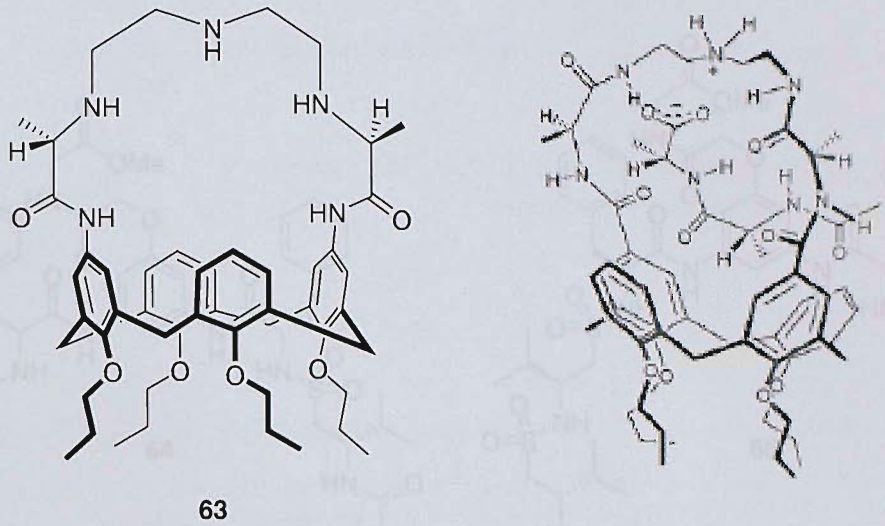


Fig 1-43 Kilburn *et al*'s macrobicyclic receptors 61,62 for alanylalanine. The rigidified receptor 62 binds more strongly to both enantiomers, but enantioselectivity is reversed between the two

Ungaro *et al*[122] have studied a macrobicyclic receptor **63** for D-Ala-D-Ala based on calixarenes (Fig 1-40). The chiral receptor was not only discovered to effectively bind D-Ala-D-Ala in $CDCl_3$ (presumably reinforced by a proton transfer from the acidic guest to the secondary amine of the host) but also to perform as inhibitor against a similar range of bacteria as vancomycin. Derivatising the secondary amine with a Boc-group or replacing it with a methylene unit completely suppressed biological activity, showing the crucial nature of this group.



63

Fig 1-44 Ungaro *et al*'s bridged calixarene receptor for alanylalanine 63 and binding model of the association between x and Ac-D-Ala-D-Ala. The peptide threads through the macrocycle and the carboxylate-ammonium ion pair accounts for much of the complex stability

A binding model was proposed based on experimental results with the guest threaded through the peptidic bridging ring with electrostatic interactions between ammonium and carboxylate and hydrogen bonding interactions between amide groups. In this case, the calixarene portion merely functions to rigidfy the arrangement of the bridging ring. However it was proposed that one of the guest methyl groups could occupy the hydrophobic space in the neck of the calixarene, especially in more aqueous media.

Tweezer receptors for D-Ala-D-Ala are less well developed. Kilburn *et al*[123] have described tweezers with a diamidopyridine CBS as head group with sulphonamidopeptide side-arms (**64**,**65**). In contrast to amides, which present both hydrogen-bond donor and acceptor functionalities (amide NH and carboxyl CO respectively), sulphonamides provide strong donor functionality but only weak acceptors. The bisulphonamide receptor **64** was found to bind both *N*-protected alanines and alanylalanines with comparable strength, however tetrasulphonamide receptor **65** was found to be markedly selective for Cbz-D-Ala-D-Ala when compared to Cbz-L-Ala-L-Ala (K_a of 2404 vs 107 M^{-1}, measured by NMR titration in $CDCl_3$).

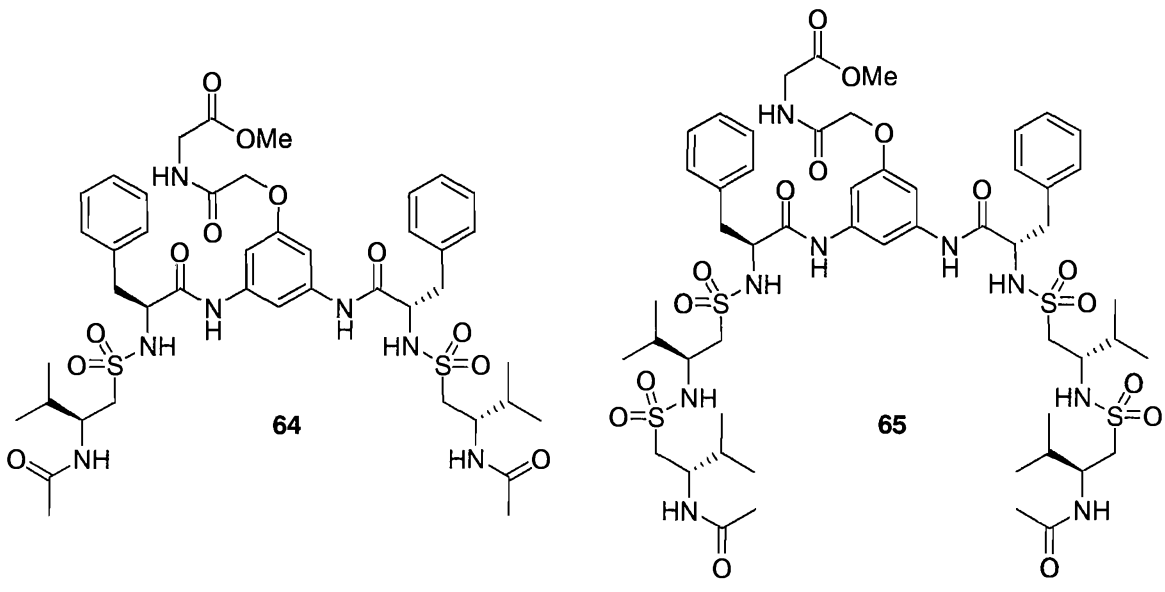
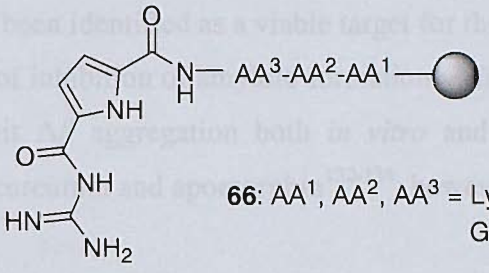


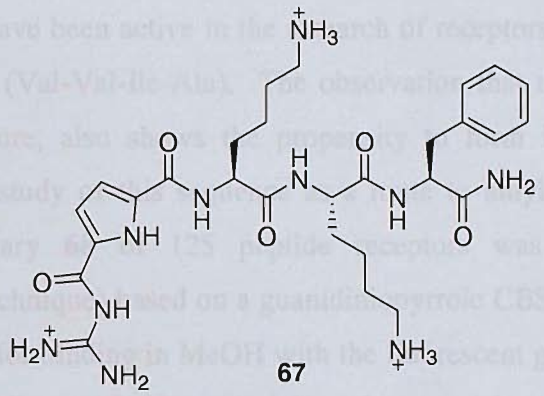
Fig 1-45 Kilburn *et al*'s sulphonamidopeptide receptors 64, 65 for alanylalanine

Liskamp *et al*'s combinatorial studies[107-109] on tweezer receptors for D-Ala-D-Ala and D-Ala-D-Lac in aqueous medium have already been described (section 1.5). Tweezers were generated which appeared to be both highly selective for either D-Ala-D-Ala or D-Ala-D-Lac and strongly binding.

Very recently, Schmuck *et al* have reported[124] a guanidiniopyrrole-based receptor library **66** for binding the tetrapeptide Glu-Lys-D-Ala-D-Ala. A selected receptor **67** from this library is reported to bind the guest in aqueous medium with a startling K_a of 1.7×10^4 M^{-1} on both solid-phase and in solution. The binding strength is attributed solely to electrostatic interactions and hydrogen bonding, with no significant contribution from hydrophobic effects (as uncharged hydrophobic library members showed only modest binding constants). For such a simplistic (though highly charged) receptor, the result is encouraging for the development of more sophisticated systems.



66: AA^1, AA^2, AA^3 = Lys(Boc), Tyr(O^{t}Bu), Ser(O^{t}Bu), Glu(OBn), Phe, Val, Leu, Trp



67

Fig 1-46 Schmuck et al's receptor library 66 for Glu-Lys-D-Ala-D-Ala and selected receptor 67 found to bind strongly in aqueous medium

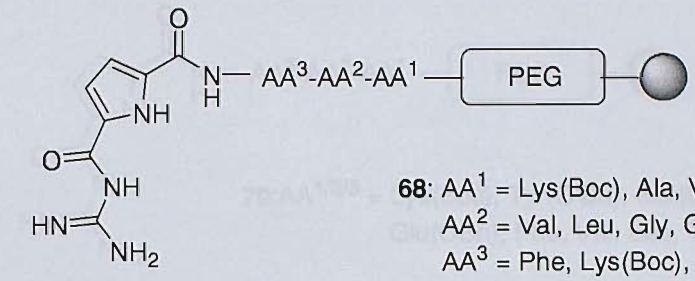
Further studies on the binding interactions between this highly important D-Ala-D-Ala sequence may continue to provide further understanding of bacterial pathology and evolve potent receptors for the sequence in aqueous medium with the potential for development into therapies.

1.6.2 Amyloid *β*-protein *C*-terminus

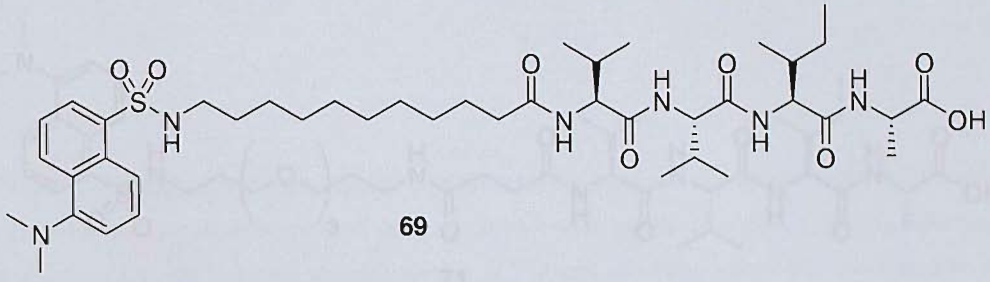
Amyloid *β*-protein (A*β*) is a relatively short protein of 39-42 amino acid residues, derived from the Alzheimer precursor protein by protein scission in the body.[125] It is known to be a major component of plaques deposited in the brain, most commonly in Alzheimer's disease patients.[126-129] It is now known that the *C*-terminus of A*β* (Val-Val-Ile-Ala) is critical for the seeding of this plaque formation[130] though whether the formation of plaques is a cause or a symptom of the degenerative process is unknown.[126]

Aβ has now been identified as a viable target for the treatment of Alzheimer's disease[131], by the process of inhibition of amyloid formation. Many inhibitors have now been discovered which inhibit Aβ aggregation both *in vitro* and in cell assays, including Congo Red, rifampicin, curcumin and apomorphin[132-134], however the mode of action of these molecules is unknown.

Schmuck *et al* have been active in the research of receptors for the *C*-terminal tetrapeptide sequence of Aβ (Val-Val-Ile-Ala). The observation that the tetrapeptide, labelled with a dansyl fluorophore, also shows the propensity to form insoluble units[135] confirms the viability of the study of this sequence as a route to amyloid precipitation inhibitors. A solid-phase library **68** of 125 peptide receptors was prepared (using the IRORI radiolabelling technique) based on a guanidiniopyrrole CBS attached to a tripeptide (Fig 1-47) and studied for binding in MeOH with the fluorescent guest **69**.[136] The receptors could be ranked according to the measured K_a values, and it was noted that small changes in receptor structure could greatly affect the binding. The most strongly binding receptors are noted.



68: AA^1 = Lys(Boc), Ala, Val, Gln(Trt), Phe
AA^2 = Val, Leu, Gly, Glu(OBn), Asp(OBn)
AA^3 = Phe, Lys(Boc), Arg(Pmc), Val, Ala

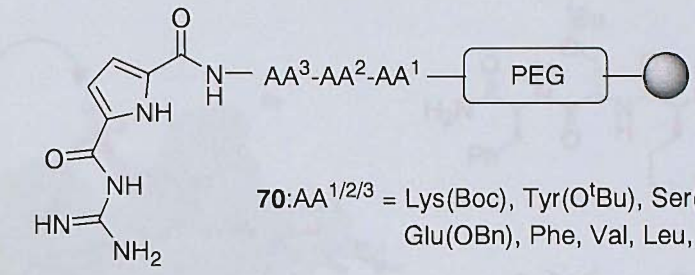


69

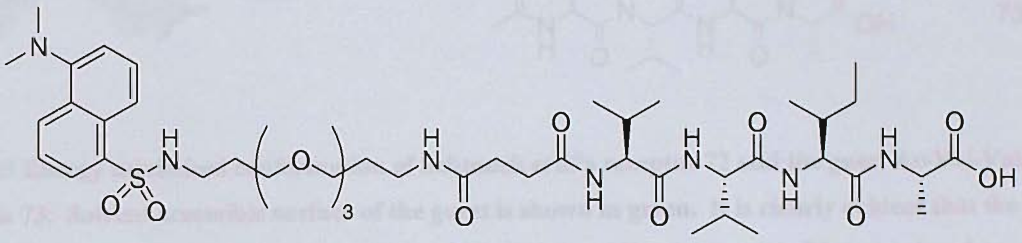
AA^1	AA^2	AA^3	K_a/M^{-1}	
Phe	Val	Val	9800	Most strongly binding library members (MeOH)
Val	Val	Val	9300	
Phe	Val	Lys(Boc)	4600	
Val	Val	Lys(Boc)	4500	

Fig 1-47 Schmuck *et al*'s library of guanidiniopyrrole receptors 68 and solid-phase binding constants with guest 69 in MeOH of the strongest binding library members

A binding model based on the formation of a β-sheet is proposed. Side-chain interactions between AA^1 and the first Val within the β-sheet complex are revealed to be of great importance. A different library **70**[135] of the same overall structure but varying in the nature of the randomised amino acids was also assayed for binding to the Val-Val-Ile-Ala sequence (appropriately *N*-terminally derivatised with PEG to increase solubility) in aqueous buffer.



70:$AA^{1/2/3}$ = Lys(Boc), Tyr(O^{t}Bu), Ser(O^{t}Bu), Glu(OBn), Phe, Val, Leu, Trp



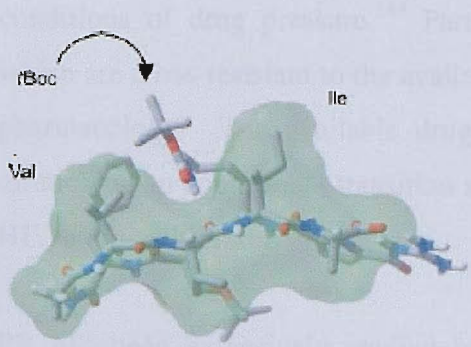
71

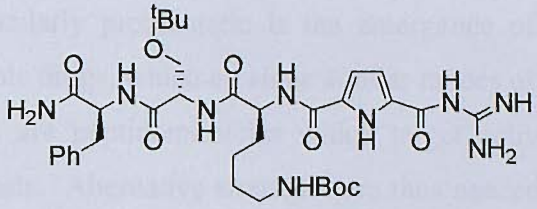
AA^1	AA^2	AA^3	K_a/M^{-1}
Lys(Boc)	Ser(O^{t}Bu)	Phe	4200
Tyr(O^{t}Bu)	Leu	Ser(O^{t}Bu)	4000
Lys(Boc)	Lys(Boc)	Glu(OBn)	3000
Lys(Boc)	Phe	Val	2900

Most strongly binding library members (aqueous buffer)

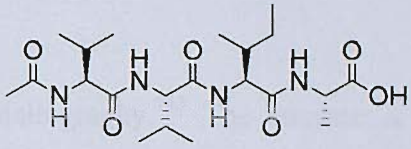
Fig 1-48 Schmuck *et al*'s library of guanidiniopyrrole receptors 70 and solid-phase binding constants with guest 71 in aqueous medium of the strongest binding library members

Binding constants were again measured and the results rationalised by the use of molecular modelling studies. In both cases the guanidinium-carboxylate ion-pair is deemed to be the major interaction, with the formation of a β-sheet structure and secondary hydrogen bonding increasing binding strength. Selectivity of binding amongst receptors is due to the nature of the side-chains, which is not immediately obvious. For example, the most strongly binding receptors in aqueous buffer were found to display AA^1 = Lys(Boc). This has been attributed to the formation of the β-sheet allowing the Boc-group to be positioned in the cleft between a Val-side chain and the Ile-side chain, thus reducing hydrophobic contact of the tetrapeptide in a favourable interaction (Fig 1-49)





72



73

Fig 1-49 Energy minimised conformation of Schmuck *et al*'s receptor 72 and the guest Ac-Val-Val-Ile-Ala 73. Solvent accessible surface of the guest is shown in green. It is clearly evident that the Boc-side chain of Lys sits between the side-chains of Val and Ile of the guest. Picture taken from Schmuck *et al*[135]

The use of these guanidiniopyrrole receptors extends beyond the theoretical study of Aβ aggregations. Certain of the receptors were shown to inhibit *in vitro* fibril formation by interaction with the C-terminus of Aβ, as measured by bioassay.[137] Thus the prospect exists for further developments of carboxylate-receptors in this area towards the generation of potential therapies for Alzheimer's disease and other neurodegenerative diseases.

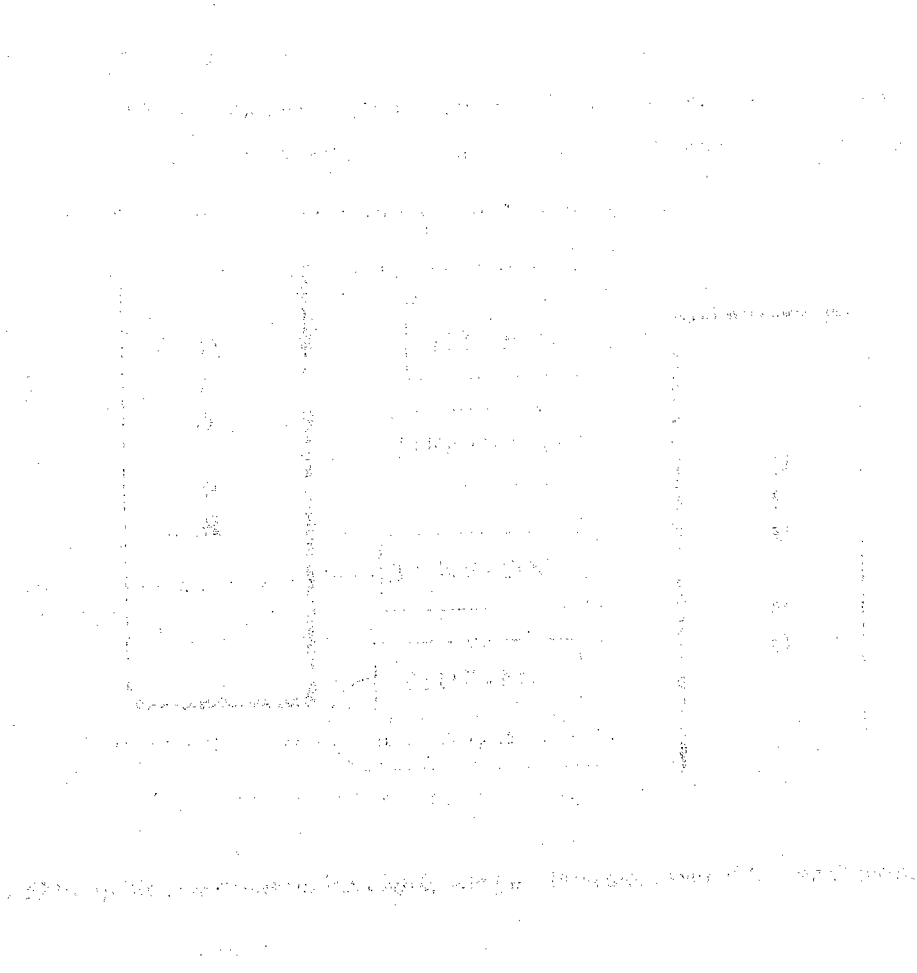
1.6.3 HIV-1 Protease *C*-terminus

The spread of the AIDS virus is of critical concern, with a 40 millions cases reported worldwide in 2001, with increasing infection rates.[138] Prophylaxis is effective in controlling the spread of the disease, but there is still a pressing need for therapies as currently no vaccination exists.

The causative agent of AIDS has been identified as HIV, a highly active retrovirus.[139] Antiretroviral drug therapies[140] are available which target the viral enzymes HIV-1 Protease (PR) and reverse transcriptase (RT), however evolution of drug resistance is becoming a problem.[141] 150 changes have been established in the PR and RT regions associated with antiretroviral resistance. Thus current antiviral therapy has led to the

conditions of drug pressure.[142] Particularly problematic is the emergence of mutations which are cross-resistant to the available drugs, which all show similar modes of action and pharmacology. The available drugs are peptidomimetics which target active sites by mimicking an appropriate transition state. Alternative strategies are thus needed to control HIV by PR inhibition.

PR has been extensively studied by crystallography.[143] The enzyme is found to self-assemble into an active homodimeric structure of two 99-amino acid monomers, forming a tube-like hydrophobic catalytic site (Figs 1-50 and 1-51). The dimer is stabilised by hydrophobic contact area and also by the interdigitation of the *N*- and *C*-terminal strands to form a β-sheet structure. This β-sheet structure is estimated to provide 75% of dimer stability.[144]



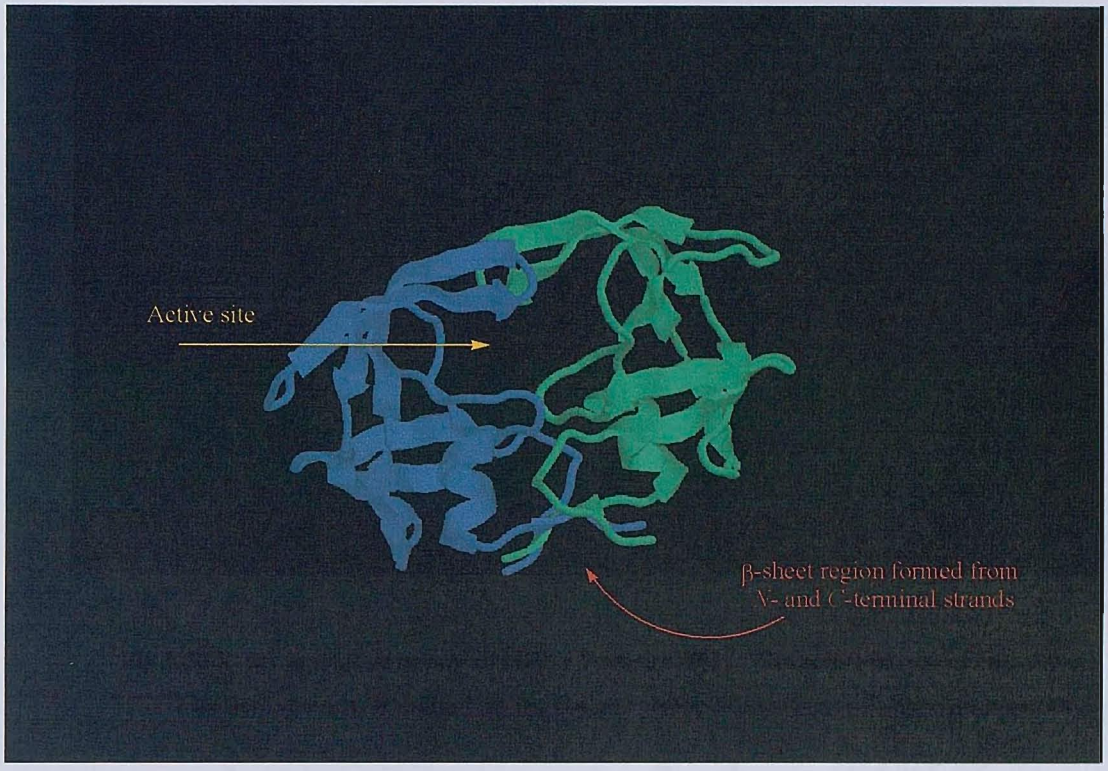


Fig 1-50 X-ray crystal structure of HIV-1 Protease (PR). The active enzyme forms from two identical strands, stabilised by the *β*-sheet region of the *N*- and *C*-terminal strands

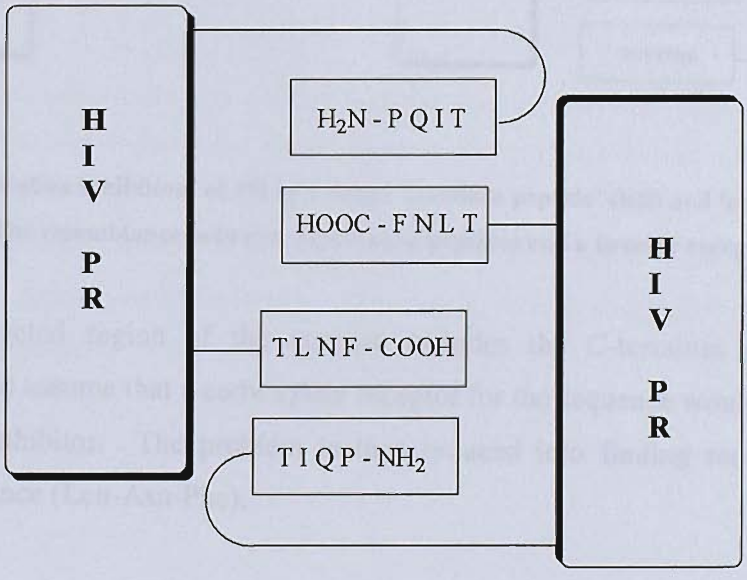


Fig 1-52 Simplified schematic of PR highlighting the interdigitation of *N*- and *C*-terminal strands

The *N*- and *C*-terminal regions of the monomeric enzyme sequences are known to be highly conserved in many strains of HIV.[145-148] Destabilisation of the dimer destroys the shape of the active site, rendering PR inactive[149,150] This evidence suggests that disrupting the β-sheet structure formed by the *N*- and *C*-terminal strands through the use of 'dimerisation inhibitors' would be an effective method of controlling HIV. A significant advantage to this mode of action is that viral mutation is unlikely to confer inhibitor resistance, as any mutation in the targeted terminal regions must be matched by a simultaneous complementary mutation in the opposite strand to maintain dimer stability. Initial studies using so-called 'interface peptides' has proven successful, generating potent inhibitors based on a consensus peptide sequence (*not* identical to the native terminal strands) and minor modifications.[151,152]

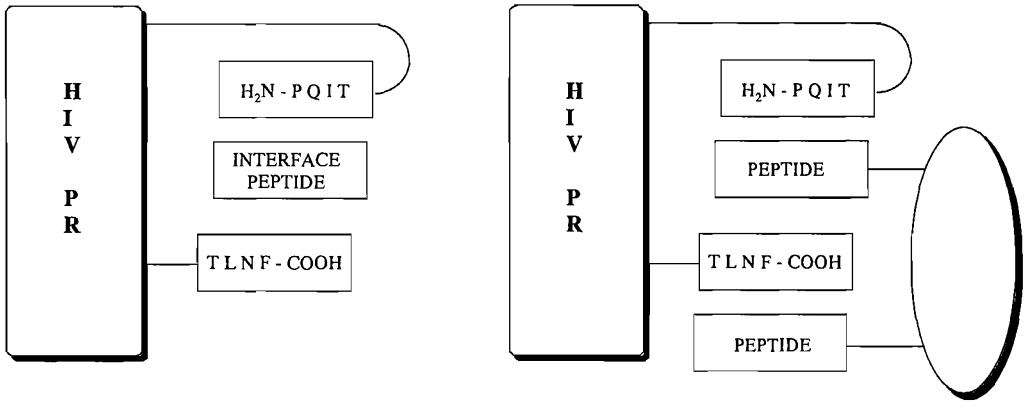
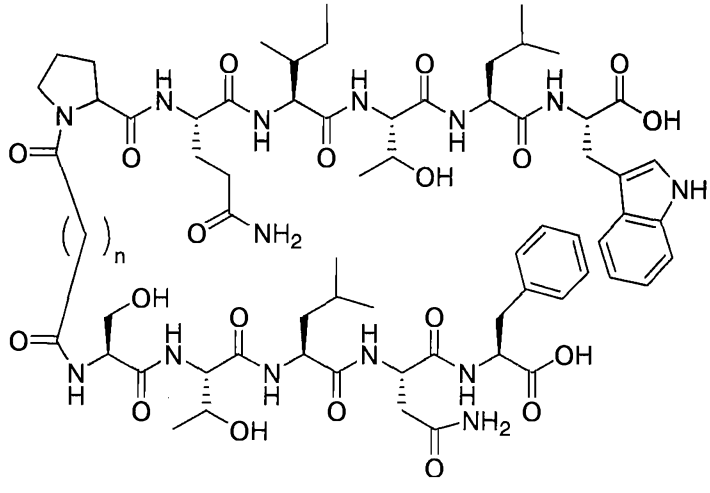


Fig 1-52'Dimerisation inhibition' of PR by a single 'interface peptide' (left) and 'cross-linked' peptides (right). The resemblance between cross-linked peptides and a tweezer receptor is obvious

Since the targeted region of the enzyme includes the *C*-terminus of PR, it is not unreasonable to assume that a carboxylate receptor for the sequence would be effective as a dimerisation inhibitor. The problem is thus reduced into finding receptors for the *C*-terminal sequence (Leu-Asn-Phe).

Dimerisation inhibitors[153,154] have already been investigated which could be considered tweezer receptors. Chmielewski *et al* have been particularly active in the field, studying

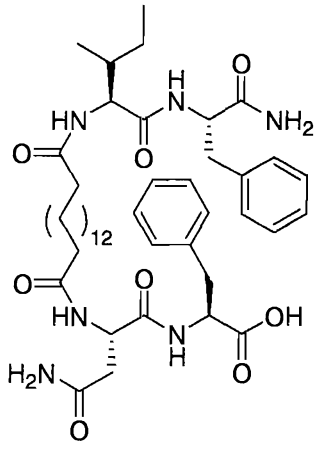
structures such as **74** (Fig 1-53) based on two peptides cross-linked by a flexible alkyl tether of varying length[155,156], though 14 methylene units was discovered to be optimal in the cases studied. This is also suggested by the 10Å strand-separation in the natural dimer.[143] An alanine-scanning approach[157] was used to determine the relative importance of each residue in the peptide side-arms. It was indicated that side-chains were more important for β-sheet formation than amide functions[158] and an investigation was carried out into which amide residues were crucial for inhibitory potency. A focused library was systematically synthesised and evaluated based on single mutations of amino acids in a previously determined optimal structure, generating highly potent analogues.[159]



74

Fig 1-53 Chmielewski *et al*'s cross-linked interfacial peptides 74 as dimerisation inhibitors of PR. Various investigations around this core structure have been undertaken

It was also shown that inhibitory potency could also be achieved using much more modest tweezers with only dipeptide side-arms (Fig 1-54), thus dramatically reducing molecular weight and complexity.[160]



75

Fig 1-54 Chmielewski *et al*'s truncated cross-linked interfacial peptides e.g. 75. Dimerisation inhibitory potency is maintained in these reduced pharmacophores

Similar cross-linked peptides have been developed based on more rigid structures. Sicsic *et al*[161] have developed aryl units symmetrically derivatised with peptides (Fig 1-55) giving more rigid receptors than the alkyl-tethered versions of Chmielewski *et al.* Bio-assay of such structures revealed highly potent inhibitors of PR dimerisation, with favoured sequences being identified. The inclusion of a pyridine-linker could be viewed as the incorporation of a primitive CBS for the carboxylate unit based on an ion-pair interaction. The dimerisation inhibitory potency in bioassay was found to decrease in the order **76** << **78** < **77**.

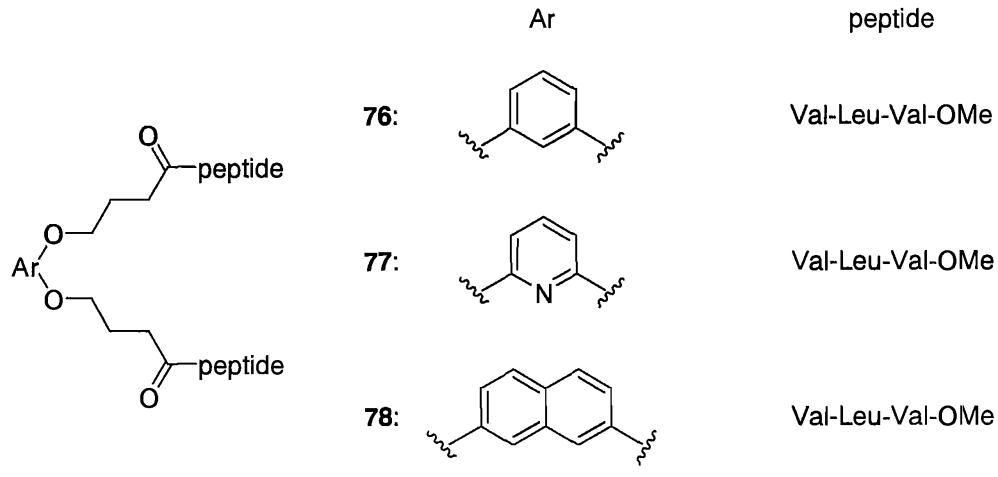
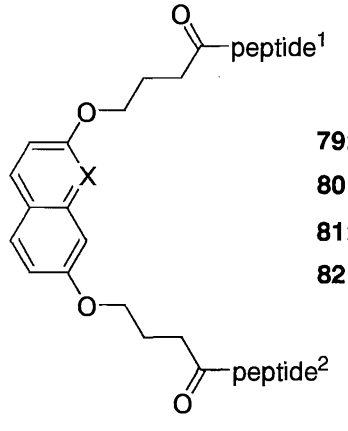


Fig 1-55 Sicsic *et al*'s rigidly tethered peptides in tweezer receptors. Various aromatic units were tested with various peptides. Only a representative selection is shown here

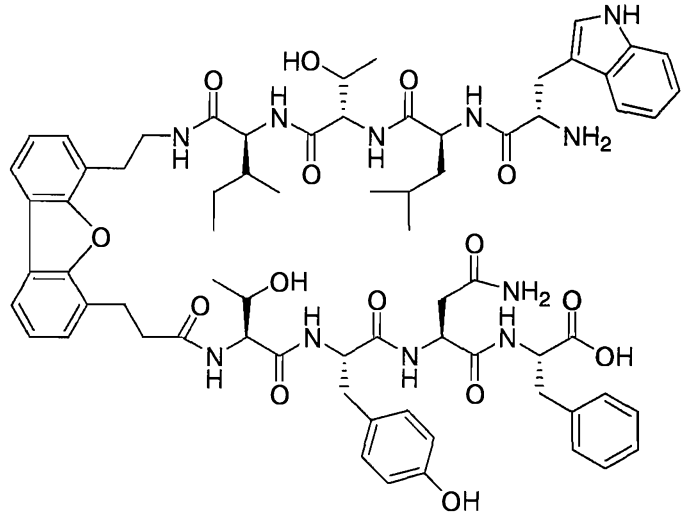
Unsymmetrical versions of the above receptors have also been developed.[162] A quinoline-based linker was utilised to incorporate the favoured ion-interaction of the pyridine-linker and the more appropriate size of the naphthalene-linker. Again, highly potent inhibitors were possible (Fig 1-56), with the receptor **82** performing best in bioassay for dimerisation inhibition.



	X	peptide^1	peptide^2	K_{id}/uM
79:	C	Val-Leu-Val-OMe	Ile-Thr-Asn-NH_2	1.50
80:	N	Val-Leu-Val-OMe	Ile-Thr-Asn-NH_2	0.25
81:	C	Val-Leu-Val-OMe	Val-Leu-Val-OMe	0.56
82:	N	Val-Leu-Val-OMe	Val-Leu-Val-OMe	0.13

Fig 1-56 Sicsic *et al*'s further development of tweezer receptors for PR. Various peptides were utilised in varying combinations; representative examples are shown

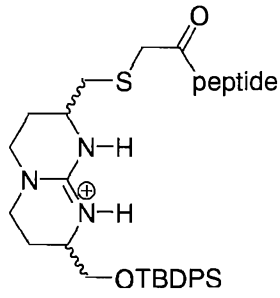
Song *et al*[163] used a 4-(2-aminoethyl)-6-dibenzofuranopropionic acid scaffold as a so called 'β-sheet nucleator', a hydrophobic cluster which was shown to nucleate a variety of peptide sequences in a β-sheet conformation.[164] Peptide side-arms were attached, however inhibitory potency was less than other receptors noted above.



83

Fig 1-57 Song *et al*'s tweezer receptor 83 with β-sheet nucleator scaffold

The most sophisticated dimerisation inhibitors to date, potentially representing true carboxylate receptors rather than peptide-based receptors, have been developed by de Mendoza *et al.*[165] Structures consisting of a bicyclic guanidinium CBS, peptides representing the *C*-terminus of PR and a hydrophobic tail were synthesised and were found to be highly potent (Fig 1-58). For example, **84**, **85**, and **86** all inhibit dimerisation of PR with respectable inhibition constants ($K_{id} \leq 0.29\mu M$) with some difference noted by the peptide sequence. The configuration of the bicyclic guanidinium was found to be inconsequential for PR inhibition; **85** and **86** were equipotent. However **87** was found to have a far weaker inhibitory power, ascribed to self-association of the guanidinium with the free carboxylate of the peptide, thus preventing efficient binding to the *C*-terminus of PR.



84: (*S*,*S*), peptide = Gly-Ile-Ser(OBn)-Tyr(OBn)-Asn-Leu-OMe

85: (*S*,*S*), peptide = Gly-Ala-Thr(OBn)-Leu-Asn-Phe-OBn

86: (*R*,*R*), peptide = Gly-Ala-Thr(OBn)-Leu-Asn-Phe-OBn

87: (*S*,*S*), peptide = Gly-Ala-Thr(OBn)-Leu-Asn-Phe-OH

Fig 1-58 de Mendoza *et al*'s dimerisation inhibitors of PR, based on carboxylate binding by the bicyclic guanidinium unit

Such structures are rigid, in contrast to the Chmielewski receptors. The potential for hydrogen bonding and electrostatic interactions provided by the guanidinium unit may contribute to inhibitory potency by increasing the binding affinity for the *C*-terminus. Further developments in the nature of the peptide side-chains are awaited.

All the above detailed receptors, based on a tweezer model, have been developed in reference to the full PR enzyme. It remains to be seen whether potent receptors for the truncated *C*-terminal sequence of PR (Leu-Asn-Phe) are also effective dimerisation inhibitors. Considering the noted modes of action of the known receptors, this seems highly probable. As an interim target, carboxylate receptors of the Leu-Asn-Phe sequence are worthy of investigation.

1.7 Scope of the current work

The current work wishes to expand on the methods of combinatorial investigation of receptors for peptides with free carboxylates in aqueous medium. The general limit of such investigations is not in the synthesis of such receptors, as most have been based on the molecular tweezer with rather simplistic head groups. Peptide side-arms are very easily introduced in solid-phase synthesis and are readily amenable to combinatorial variation. Interpretation of 'hit' structures remains a bottle-neck in such methods and the current work aims to address some of the current problems through the development of simple-to-implement procedures based on the concept of the ladder sequence.[84]

In particular, the development of receptors for biologically interesting peptides (such as those discussed above) is a goal of the current methods. Receptors currently exist for all these peptides, but with heavy limitations of either binding solely in organic medium or being of high molecular complexity. The generation and study of simpler systems with high binding affinity in aqueous medium should provide greater understanding of the biological systems and perhaps be an entry-point into synthetically-viable inhibitors.

Chapter 2

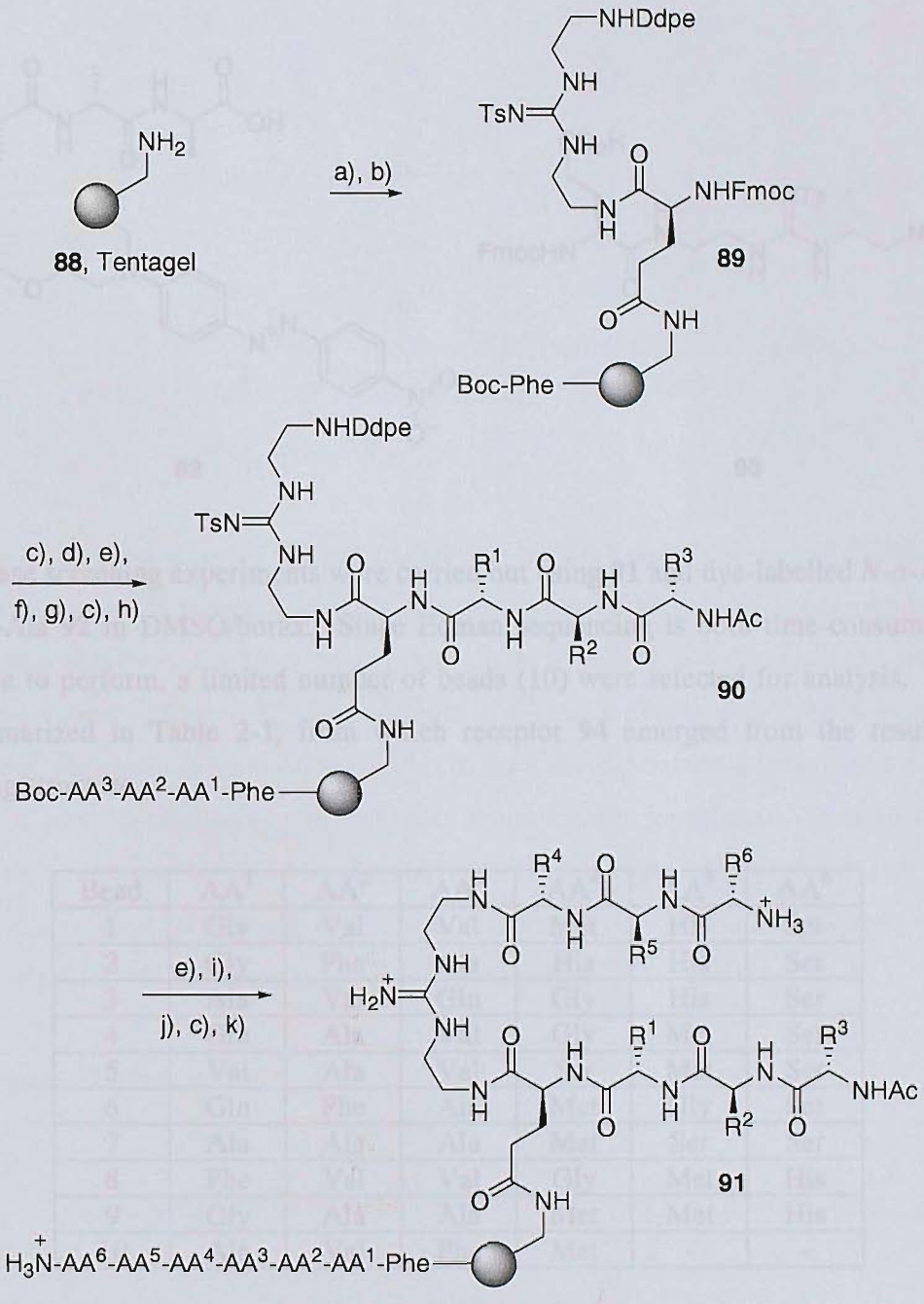
Synthesis of a receptor for a D-Ala-D-Ala guest and solid-phase binding studies

2.1 Introduction

This chapter describes the synthesis of a solid-phase-linked receptor of the tripeptide guest *N*-α-Ac-Lys-D-Ala-D-Ala, and estimation of a binding constant by solid-phase uptake assay.

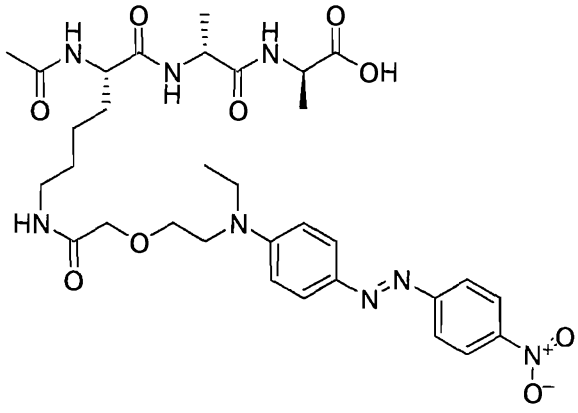
As discussed in Chapter 1, the motif Lys-D-Ala-D-Ala plays an important role in bacterial cell wall synthesis, and it is the binding of this unit by vancomycin in a complex interaction that gives this powerful antibacterial its efficacy. Vancomycin is however a sophisticated, highly functionalised molecule and interest exists in the generation of simpler systems which also bind the Lys-D-Ala-D-Ala motif. Current research has been devoted to the discovery of strongly binding receptors for this sequence and several structures have been highlighted.

Previous work within the Kilburn group[166] has involved the synthesis of libraries of tweezer receptors. Such a library **91** was synthesised in a split-and-mix protocol as detailed in Scheme 2-1. A simultaneously synthesised peptide coding strand was utilised to identify receptor structures present on individual beads by Edman sequencing, with Phe present as the last residue in each case. Owing to the need for orthogonal strategy in building the coding strand, the diversity of residues in one tweezer arm was limited.

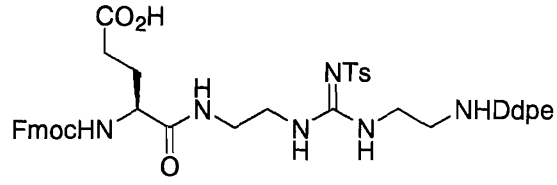


Reagents and conditions: a) 10 mol% Boc-Phe, PyBOP, HOBt, DIPEA, DMF; b) **93**, PyBOP, HOBt, DIPEA, DMF; c) 20% piperidine/DMF; d) split-and-mix Fmoc-peptide synthesis using Gly, Ala, Val, Phe, Gln; e) 30% TFA/DCM; f) split-and-mix Boc-peptide synthesis using Gly, Ala, Val, Phe, Gln; g) repeat c-f twice; h) Ac_2O, DMAP; i) 35% aqueous hydrazine; j) three-fold split-and-mix peptide synthesis using Gly, Ala, Ser, Met, His; k) liquid HF

Scheme 2-1 Kilburn *et al*'s library of tweezer receptors 91 for solid-phase screening



92

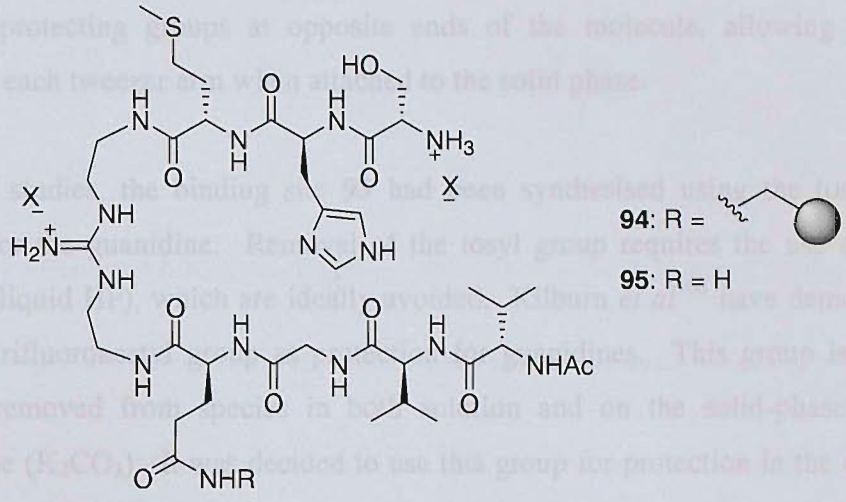


93

Solid-phase screening experiments were carried out using **91** and dye-labelled *N*-α-Ac-Lys-D-Ala-D-Ala **92** in DMSO/borax. Since Edman sequencing is both time-consuming and expensive to perform, a limited number of beads (10) were selected for analysis. Results are summarized in Table 2-1, from which receptor **94** emerged from the results as a promising structure.

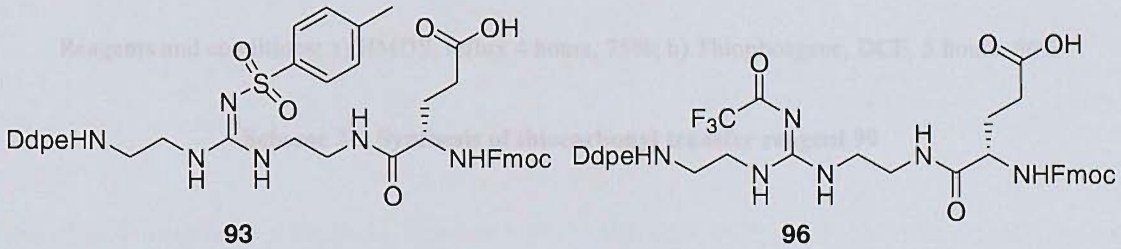
Bead	AA^1	AA^2	AA^3	AA^4	AA^5	AA^6
1	Gly	Val	Val	Met	His	Ser
2	Gly	Phe	Ala	His	His	Ser
3	Ala	Val	Gln	Gly	His	Ser
4	Gln	Ala	Val	Gly	Met	Ser
5	Val	Ala	Val	Ser	Met	Ser
6	Gln	Phe	Ala	Met	Gly	Ser
7	Ala	Ala	Ala	Met	Ser	Ser
8	Phe	Val	Val	Gly	Met	His
9	Gly	Ala	Ala	Met	Met	His
10	Ala	Val	Phe	Met	-	-

Table 2-1 Results from screening library 91 with guest 92. Sequencing was inconclusive for certain residues of bead 10, as indicated



Synthesis of isolated receptor **95** was carried out on Rink Amide resin and NMR binding studies attempted. However, owing to poor resolution of signals no meaningful binding constant was obtained. UV titration gave evidence of binding; a decrease in the absorption of the guest chromophore of **92** upon addition of **95** and a clearly identified isobestic point were noted. However, it was not possible to fit the resulting data to a presumed 1:1 binding model. 1:2 and 2:1 binding modes were also found to be inappropriate. In order to obtain an estimate of magnitude of the binding between receptor **95** and guest **92**, it was decided to perform a solid-phase binding assay using receptor **94**. Various methods are known for this measurement, including those of Still[101], Wennemers[167] and Kilburn.[103]

2.2 Synthesis of Binding Site 112



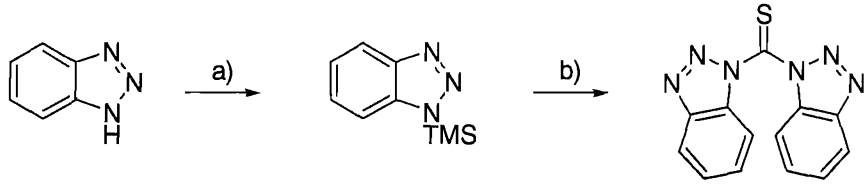
93 **96**

To obtain receptor **94** as a solid-phase linked entity, it was necessary to synthesise an appropriately protected CBS, as described below. Of particular importance are the

orthogonal protecting groups at opposite ends of the molecule, allowing independent synthesis of each tweezer arm when attached to the solid phase.

In previous studies, the binding site **93** had been synthesised using the tosyl group as protection for the guanidine. Removal of the tosyl group requires the use of hazardous conditions (liquid HF), which are ideally avoided. Kilburn *et al*[168] have demonstrated the use of the trifluoroacetyl group as protection for guanidines. This group is simply and efficiently removed from species in both solution and on the solid-phase using mild aqueous base (K_2CO_3). It was decided to use this group for protection in the current work i.e. to generate **96**.

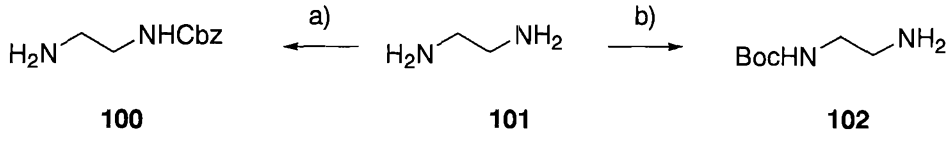
Protected guanidine **107** was synthesised according to Schemes 2-1 to 2-3. Thiocarbonyl transfer reagent **99** was synthesised according to modified literature procedure[169] (Scheme 2-1). Ethylenediamine **101** was mono-Boc-protected to give **102**. **101** was also independently reacted with benzylphenylcarbonate to give mono-Cbz-protected diamine **100** (Scheme 2-2).



97 **98** **99**

Reagents and conditions: a) HMDS, reflux 4 hours, 75%; b) Thiophosgene, DCE, 5 hours, 56%

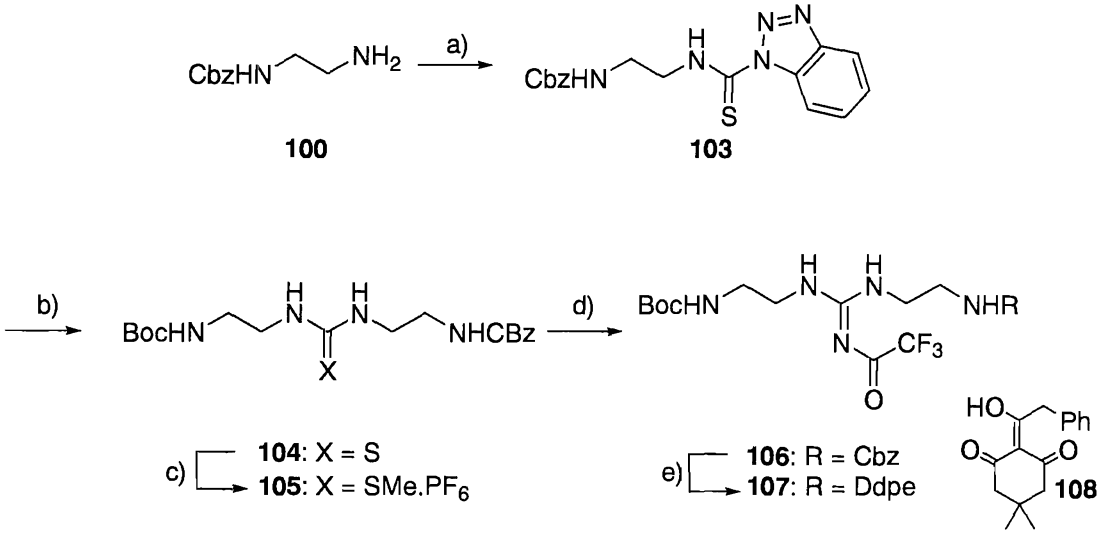
Scheme 2-1 Synthesis of thiocarbonyl transfer reagent 99



Reagents and conditions: a) Benzylphenylcarbonate, EtOH, 18 hours, 60%; b) Boc_2O, $CHCl_3$, 18 hours, 77%

Scheme 2-2 Synthesis of monoprotected ethylenediamines 101 and 102

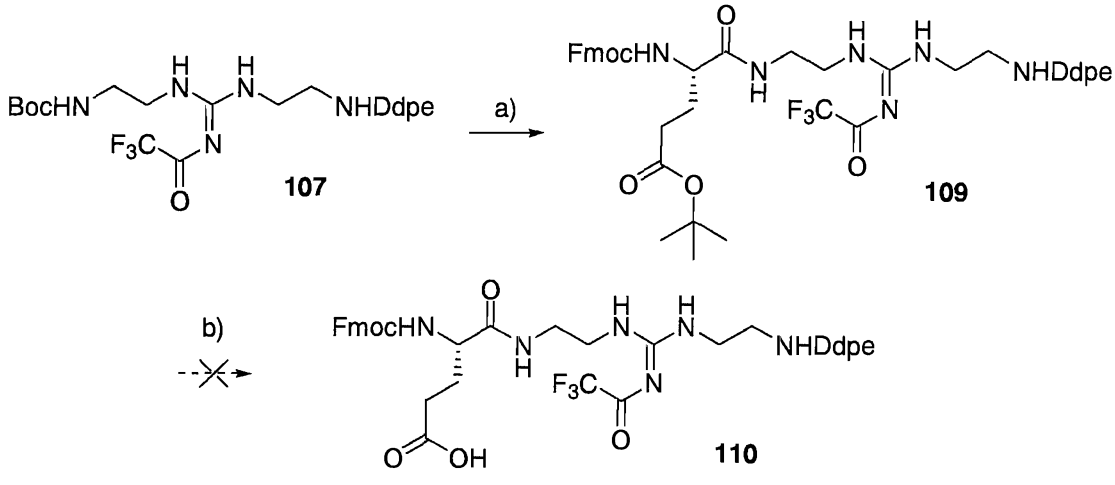
100 was reacted with **99** to give intermediate **103** which was then coupled with amine **102** to give thiourea **104**. **104** was then activated as a thiouronium salt **105** by sequential treatment with methyl iodide and ammonium hexafluorophosphate, followed by reaction with trifluoroacetamide and DBU to give protected guanidine **106**. Functional group exchange of Cbz for Ddpe led to protected guanidine **107** (Scheme 2-3). (For synthesis of **108**, see Chapter 3).



Reagents and Conditions: a) **99**, DCM, 18 hours, 74%; b) **102**, Et_3N, DCM, 18 hours, 59%; c) i) MeI, acetone, 2 hours, ii) NH_4PF_6, MeOH/DCM, 18 hours; d) F_3CCONH_2, DBU, toluene/$CHCl_3$, reflux, 18 hours, 51% for three steps; e) i) H_2, Pd/C, DMF, 6 hours, ii) **108**, DMF, 18 hours, 27% over two steps

Scheme 2-3 Synthesis of protected intermediate 107

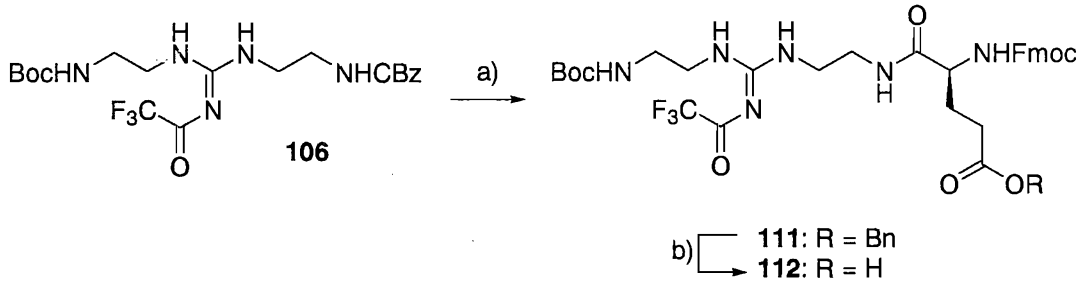
Guanidine **107** was Boc-deprotected and coupled with *N*-α-Fmoc-glutamic acid-ω-*tert*-butyl ester to give derivative **109** (Scheme 2-4). Unfortunately upon attempted deprotection of the *tert*-butyl ester using dilute TFA, a mixture of products was obtained from which it was not possible to isolate the desired acid **110**. This is partly due to the sensitivity of the trifluoroacetyl group to the presence of MeOH, precluding the possibility of chromatographic purification. Attempted recrystallisation from a range of solvents did not give pure product.



Reagents and conditions: a) i) TFA/DCM, 2 hours, ii) Fmoc-Glu(O^{t}Bu)-OH, HOBt, DIPEA, EDC, DMAP, DMF/DCM, 18 hours, 79% over two steps; b) 30% TFA/DCM, 3 hours

Scheme 2-4 Failed synthesis of protected binding site 110

An alternative protection strategy was thus devised for the synthesis. Guanidine **106** was Cbz-deprotected under hydrogenation conditions and the resulting amine directly coupled to *N*-α-Fmoc-glutamic acid-ω-benzyl ester yielding compound **111** (Scheme 2-5). It was then possible to cleave the benzyl ester under hydrogenation conditions in EtOH, which did not appear to affect the trifluoroacetyl protection. Acid **112**, thus obtained, is suitable for a solid-phase synthesis of **94** since one arm of the desired receptor contains no acid-sensitive protecting groups (Gly, Val only), and hence is compatible with a later-stage Boc-deprotection before construction of the second arm.

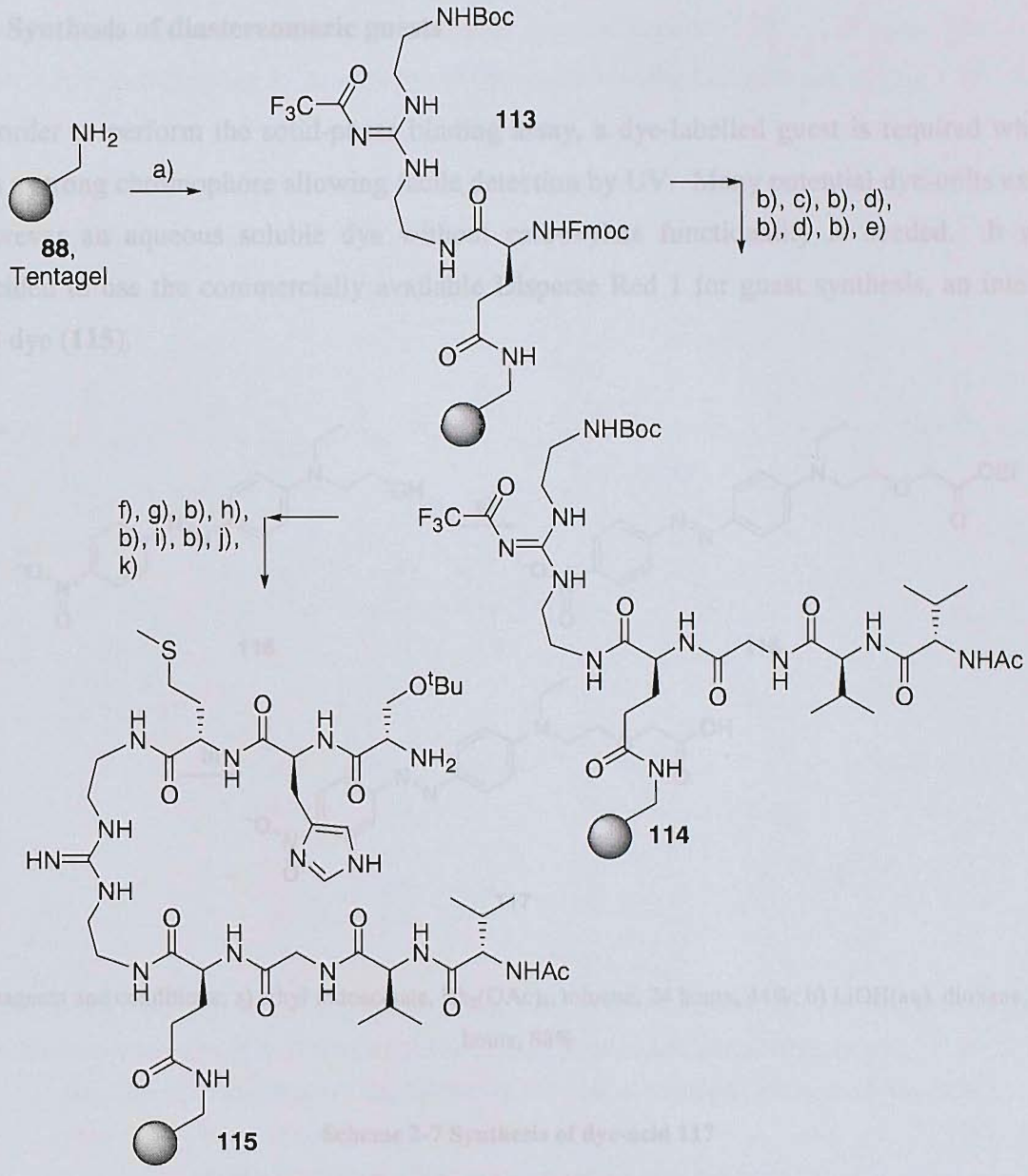


Reagents and conditions: a) i) H_2, Pd/C, DMF, 18 hours, ii) Fmoc-Glu(OBn)-OH, HOBt, EDC, DIPEA, DMAP, DMF, 18 hours, 20% over two steps; b) H_2, Pd/C, EtOH, 2 hours, 88%

Scheme 2-5 Synthesis of protected binding site 112

2.3 Synthesis of the Solid-Phase linked Receptor 94

Acid **112** was coupled to Tentagel resin (loading 0.2mmol/g) using standard conditions (DIC/HOBt, DIPEA, DMF). Standard Fmoc-SPPS methods were then used to construct the first arm of the receptor before capping with an acetyl group (**114**). A Boc-deprotection was then performed using dilute TFA and the second arm of the receptor constructed using standard Fmoc-SPPS. Side chain deprotection was carried out using 50% TFA/DCM, followed by cleavage of the trifluoroacetyl protecting-group with 0.15M K_2CO_3 in MeOH/DMF/H_2O to obtain solid-phase linked receptor **94** (Scheme 2-6). The resin was finally rinsed with Et_2O and rigourously dried *in vacuo*. All steps were monitored by the Kaiser ninhydrin test.

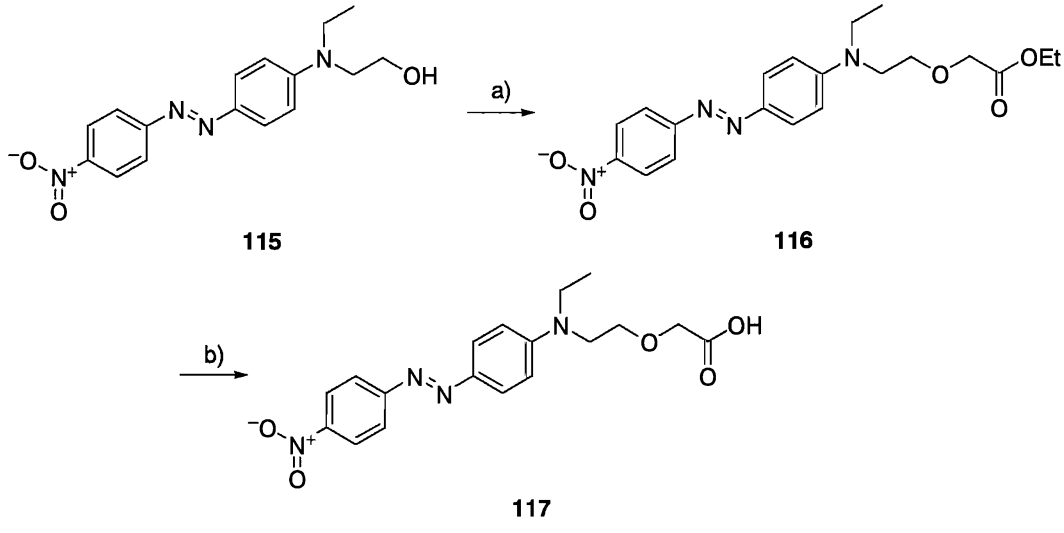


Reagents and conditions: a) **112**, DIC, HOBt, DIPEA, DMF, 18 hours; b) 20% piperidine/DMF, 20 mins; c) Fmoc-Gly-OH, DIC, HOBt, DIPEA, DMF, 3 hours; d) Fmoc-Val-OH, DIC, HOBt, DIPEA, DMF, 3 hours; e) Ac_2O, DIPEA, DMF, 3 hours; f) 20% TFA/DCM, 20 mins; g) Fmoc-Met-OH, DIC, HOBt, DIPEA, DMF, 3 hours ; h) Fmoc-His(Trt)-OH, DIC, HOBt, DIPEA, DMF, 3 hours; i) Fmoc-Ser(O^{t}Bu)-OH, DIC, HOBt, DIPEA, DMF, 3 hours; j) 95% TFA/DCM, 2 hours; k) 0.15M K_2CO_3, MeOH/DMF/H_2O

Scheme 2-6 Synthesis of solid-phase linked receptor 94

2.4 Synthesis of diastereomeric guests

In order to perform the solid-phase binding assay, a dye-labelled guest is required which has a strong chromophore allowing facile detection by UV. Many potential dye-units exist, however an aqueous soluble dye without carboxylate functionality is needed. It was decided to use the commercially available Disperse Red 1 for guest synthesis, an intense red dye (**115**).



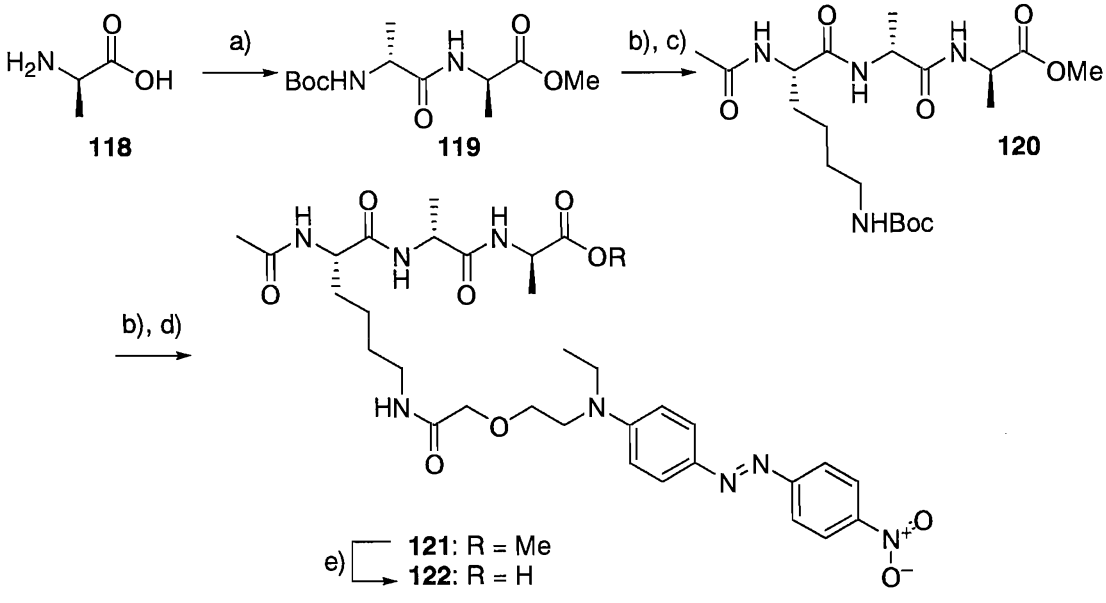
Reagents and conditions: a) ethyl iodoacetate, $Rh_2(OAc)_4$, toluene, 24 hours, 44%; b) LiOH(aq), dioxane, 18 hours, 88%

Scheme 2-7 Synthesis of dye-acid 117

First **115** was appropriately functionalised for incorporation into peptide guests as acid **117**. According to literature procedure[170], **115** was treated with ethyl iodoacetate and rhodium acetate at reflux to generate ester **116**, which was hydrolysed to give acid **117** (Scheme 2-7).

Side-chain dye-labelled *N*-α-Ac-Lys-D-Ala-D-Ala **122** was synthesised in solution according to Scheme 2-8. Conversion of D-alanine into its methyl ester followed by simple

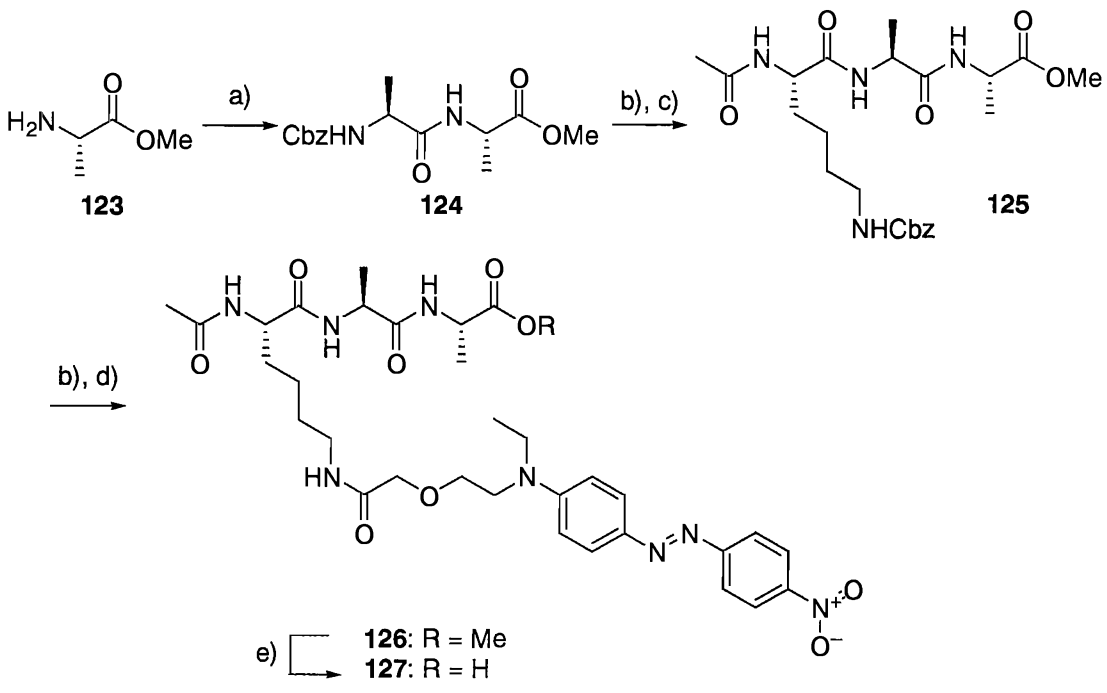
peptide coupling under standard conditions gave dipeptide **119**. **119** was then Boc-deprotected and coupled to Ac-Lys(Boc)-OH under standard conditions, giving **120**. Boc-deprotection of **120** and amide formation with dye-labelled acid **117** gave highly coloured ester **121**, which was hydrolysed in disappointing yield to give the desired dye-labelled guest **122** under basic conditions. The low yield of the final step is perhaps due to problematic isolation of the acid from aqueous conditions.



Reagents and conditions: a) i) $SOCl_2$, MeOH, reflux, 2 hrs, ii) Boc-D-Ala-OH, EDC, HOBt, DIPEA, 18 hrs, 53% over two steps; b) 20% TFA/DCM, 1.5 hrs; c) Ac-Lys(Boc)-OH, EDC, HOBt, DIPEA, 18 hrs, 27% over 2 steps; d) **117**, EDC, HOBt, DIPEA, 18 hrs, 33%; e) LiOH(aq), dioxane, 18 hours 7%

Scheme 2-8 Synthesis of dye-labelled D-Ala-D-Ala guest 122

In a similar manner, the diastereomeric guest *N*-α-Ac-Lys-L-Ala-L-Ala **127** was synthesised according to Scheme 2-9, utilising the Cbz-protecting strategy. Final hydrolysis of the methyl ester again proceeded in disappointingly low yield, most likely due to problematic isolation of the acid.



Reagents and conditions: a) Cbz-L-Ala-OH, EDC, HOBt, DIPEA, $CHCl_3$, 2 hours, 58%; b) H_2, Pd/C, MeOH, 2 hours; c) Ac-Lys(Cbz)-OH, EDC, HOBt, DIPEA, $CHCl_3$, 16 hours, 24% over two steps, d) **117**, EDC, HOBt, DIPEA, DMAP, DMF, 86% over 2 steps; e) LiOH(aq), dioxane, 4 hours, 7%

Scheme 2-9 Synthesis of dye-labelled L-Ala-L-Ala guest 127

2.5 Solid-Phase Binding Assay

Binding constants of species can be determined by a number of different methods.[171] Most often this involves a titration experiment, whereby one of the interacting partners is slowly added to the other in a solution phase, and the change in output of a particular physical property is monitored. For maximum sensitivity, the change in the physical property must be of significant magnitude, and is often related to the binding interaction. For example, NMR-titration experiments will monitor the changes in chemical shift of protons which are involved in the primary host-guest interaction, the movement of a signal in the spectrum corresponding to the changes in environment a particular proton experiences on binding (more or less deshielded). Likewise, UV-titration experiments will monitor the changes in absorbance of a chromophore associated with either the unbound host or unbound guest.

From these measured changes, numerical methods allow the computation of a binding constant based on appropriate stoichiometries of interaction.

Measurement of binding constants where one of the interacting species is immobilised on the solid-phase can be performed using the previously detailed methods[101,103,167]. In these methods, a receptor-derivatised resin is incubated with a solution of the desired guest. After incubation, and removal of the resin (by filtration) the concentration of the supernatant solution is then measured by UV or HPLC. Based on the change in concentration, it is possible to calculate the quantity of guest which has localised to the resin (and is presumably in a binding interaction with the receptor located there), and hence a 1:1 binding constant. This necessitates either an estimation of the receptor loading on the resin, or the solving of simultaneous equations which eliminate this variable (c.f. Still *et al*[101]).

In the known cases, solid-phase binding constant determination has been carried out either in organic solution or in aqueous solution using a PEG-ligated guest which is highly water-soluble. In the current case, the guests **122** and **127** are somewhat hydrophobic owing to the side-chains, and are readily absorbed from an aqueous solution into underivatised Tentagel, which represents a more hydrophobic environment than bulk solution. It is thus necessary to take account of this 'background' non-binding adsorption before a calculation of binding constants from UV data can be made.

Attempts were made to account for the non-binding adsorption using a simple model involving initial concentration of supernatant guest and mass of Tentagel resin used. However, it was not possible to fit experimental data to this model. Hence it was decided to approximate the non-binding adsorption by performing two experiments in tandem; one utilising receptor-derivatised resin and one using an equivalent mass of underivatised Tentagel. After incubation of both samples with identical solutions of dye-labelled guest, the concentration of the supernatant was determined by UV. A direct subtraction of the non-binding adsorption was then possible, giving the bound concentration and hence an estimation of the binding constant. Loading of the receptor onto Tentagel was assumed to

have proceeded efficiently, hence the loading of the underivatised Tentagel could be used for calculations. The system is too complicated for a simple elimination of this variable by solution of simultaneous equations.

A detailed account of the calculation method used to determine solid-phase binding constants is now given. Since the mass of the receptor loaded to the resin is 928.09 (for **94**) and the nominal loading of the resin used is 0.2mmol/g, the masses of resin used in the binding experiment (m_{rec} and m_{TG} for receptor-derivatised resin and Tentagel, respectively) are linked by the expression;

$$m_{rec} = 1.186 m_{TG} \qquad [1]$$

Both the sample of receptor-derivatised resin and the sample of Tentagel are then incubated with guest solution of starting concentration c_{init} and volume V. Final concentrations of the supernatant solutions are then determined by UV (c_{rec} and c_{TG} for the receptor-derivatised resin and the Tentagel, respectively). The final amounts of dye in the supernatant solution are thus given by cV and hence the amounts localised to the resin are calculated by subtraction from the starting amounts;

$$n_{final,rec} = V(c_{init} - c_{rec}) \qquad [2]$$

$$n_{final,TG} = V(c_{init} - c_{TG}) \qquad [3]$$

Since the masses of resin used are equivalent, it is possible to subtract these amounts from each other and obtain the amount *bound* by the receptor, n_{bound};

$$n_{bound} = n_{final,TG} - n_{final,rec} = V(c_{TG} - c_{rec}) \qquad [4]$$

The concentration of the host guest complex, $[HG]$ is thus given by;

$$[HG] = n_{bound} V = (c_{TG} - c_{rec}) \qquad [5]$$

and the concentration of free guest [*G*] is $c_{rec.}$ The amount of free host is calculated as the difference between the amount of receptor loaded onto the resin and the amount utilised for binding;

$$n_{freehost} = (L.m_{TG}) - n_{bound} = L.m_{TG} - V(c_{TG} - c_{rec}) \qquad [6]$$

for initial resin loading L. Hence the final concentration of host, [*H*], is given by;

$$[H] = \frac{n_{freehost}}{V} = \frac{Lm_{TG}}{V} - (c_{TG} - c_{rec}) \qquad [7]$$

A 1:1 binding constant K_a can then be calculated using the standard model;

$$K_a = \frac{[HG]}{[H].[G]} \qquad [8]$$

Note that the receptor loading must be known in order for accurate calculations to be made. In practice, an estimation was used based on the known, measurable loading of underivatised Tentagel. Note also that the calculation assumes a 1:1 binding model, instinctively appropriate yet not proven. The subtraction of 'background' non-binding adsorption by Tentagel also introduces inaccuracy, as it assumes totally homogeneous base resin which is rarely the case.

Calibration curves were made by measuring the UV absorption of the dye-chromophore of guests **122** and **127** at varying concentration in 20% DMSO/borax. Both curves gave the same relationship between concentration and absorption at 500nm. Results of the binding assay of solid-phase linked receptor **x** with guests **x** and **x** are shown in Tables 2-2 and 2-3. Errors are estimated based on a 20% error in the estimated loading of the resin.

	1	2	3

m_{rec}/ mg	5.1	3.2	2.7
m_{TG}/ mg	4.3	2.7	2.3
c/ M	1.34×10^{-4}	1.34×10^{-4}	1.34×10^{-4}
V/ ml	1.2	1.2	1.2
c_{rec}/ M	2.08×10^{-5}	1.77×10^{-5}	1.49×10^{-5}
c_{TG}/ M	4.22×10^{-5}	2.67×10^{-5}	2.30×10^{-5}
[*HG*]/ M	2.14×10^{-5}	8.97×10^{-6}	8.15×10^{-6}
[*G*]/ M	2.08×10^{-5}	1.77×10^{-5}	1.49×10^{-5}
[*H*]/ M	6.95×10^{-4}	4.41×10^{-4}	3.75×10^{-4}
K_a/ M^{-1}	1480 ± 380	1150 ± 300	1460 ± 370

Table 2-2 Results obtained in solid-phase binding assay of resin-linked receptor 94 and dye-labelled guest 122, assuming a nominal loading of 0.2mmol/g. Mean K_a = 1363 ± 185 M^{-1}

	1	2	3
m_{rec}/ mg	3.3	3.8	3.4
m_{TG}/ mg	2.8	3.2	2.9
c/ M	1.23×10^{-4}	1.23×10^{-4}	4.90×10^{-5}
V/ ml	1.2	1.2	1.2
c_{rec}/ M	1.93×10^{-5}	2.88×10^{-5}	5.21×10^{-6}
c_{TG}/ M	2.58×10^{-5}	3.14×10^{-5}	6.33×10^{-6}
[*HG*]/ M	6.51×10^{-6}	2.66×10^{-6}	1.12×10^{-6}
[*G*]/ M	1.93×10^{-5}	2.88×10^{-5}	5.21×10^{-6}
[*H*]/ M	4.60×10^{-4}	5.31×10^{-4}	4.82×10^{-4}
K_a/ M^{-1}	734 ± 185	174 ± 43	446 ± 111

Table 2-3 Results obtained in solid-phase binding assay of resin-linked receptor 94 and dye-labelled guest 127, assuming a nominal loading of 0.2mmol/g. Mean K_a = 451 ± 280 M^{-1}

It is noted that the binding is stronger with the D-Ala-D-Ala guest **122** than for L-Ala-L-Ala guest **127**, suggesting a certain diastereoselectivity by receptor **94**. However, the differing results (e.g. 734 M^{-1} vs. 174 M^{-1}, Table 2-3) obtained in assays using **127** suggest that diastereoselectivity is potentially significantly less than the order of magnitude indicated by the mean results quoted above. It is also to be noted that the calculation assumes a 1:1 binding mode, which although instinctively appropriate is in no way supported by previous UV binding studies carried out in solution[166] (where data was found not to fit 1:1 stoichiometry). It is also to be noted that the microenvironment of the resin can affect the nature of the binding event, and thus cannot be directly compared to the situation in

solution. However, it is encouraging that the values obtained suggest strong binding, confirming power of combinatorial screening in identifying receptors for particular peptide carboxylate sequences.

2.6 Conclusions

Receptor **94** was successfully synthesised on solid-phase and a binding constant with dye-labelled *N*-α-Ac-Lys-D-Ala-D-Ala **122** estimated by solid-phase UV assay. A limited diastereoselectivity is noted over *N*-α-Ac-Lys-L-Ala-L-Ala **127**. Tweezer structures appear to be viable, strongly binding receptors for the *N*-α-Ac-Lys-D-Ala-D-Ala sequence, and hence may find application in further biological studies. The power of the screening combinatorial libraries of receptors is further confirmed, indicating that development of more diverse libraries of receptor structures could reveal more potently binding receptors and/or more diastereoselective binding units; more comprehensive assessment of screening 'hit' results is however required. In particular, the ability to selectively bind peptide sequences in a highly polar medium is encouraging. Previous results (see Chapter 1) utilising non-guanidinium based binding sites have been limited to relatively non-polar media; the strong ion-pair interaction between guanidinium and carboxylate thus appears to be a prequisite for strong binding in aqueous media, and further investigation is indicated. Evaluation of receptor **94** with the known mutagenic guest *N*-α-Ac-Lys-D-Ala-D-Lac (in vancomycin-resistant bacterial strains) is also indicated.

Chapter 3

Development of a method of peptide-library sequencing and application in the construction of libraries of tweezer receptors

3.1 Introduction

This chapter describes the evolution of a method of determining the sequence of a solid-phase linked peptide or peptide-based molecule by mass spectrometry. It is based upon the 'ladder sequencing' methods of Youngquist *et al*,[84] and obviates the need for expensive and time-consuming Edman degradation procedures. The use of this methodology is then demonstrated in the construction of two libraries of receptors for carboxylates based on guanidinium binding-sites. Thus, by screening the libraries of receptors against carboxylate guests and determining the peptide sequence loaded onto a 'hit' bead, rapid and more cost-effective evaluation of promising receptors for particular oligopeptide carboxylates is envisaged.

The basic philosophy of the methodology involves the partial termination of a growing peptide sequence with a capping group after the addition of each amino acid residue, thus generating a heterogenous resin composition of not only the full peptide sequence but also the sequences lacking successively greater numbers of residues from the *N*-terminus (See Chapter 1). Hence upon cleavage of all material from the resin bead, a sequence can be interpreted from the mass differences between individual cleavage products.

Limitations of this methodology are considered. The methods are then used to build solid-phase based libraries of carboxylate-binding receptors based on the tweezer model. Such receptors consist of a single binding site (CBS) modified with randomised peptide side-arms. Randomisation arises through the application of a combinatorial split-and-mix protocol. Structural determination of a single library member is achieved by MALDI-TOF-MS analysis of the appropriate 'ladder' sequences built up on the resin during

library synthesis. The synthesis of two libraries is described, one based on a single peptide sequence capped with a CBS (the single-armed receptor library) and one authentic tweezer library with two peptide side-arms.

3.2 Development of mass spectrometric 'ladder sequencing' methods

Bradley *et al* have described a modification of Youngquist's method of 'ladder' generation[93] which involves the use of an equimolar mixture of AcOH and d_4-AcOH in tandem with PyBOP and DIPEA as the capping reagent mixture. This mixture thus theoretically generates equal amounts of sequence capped with Ac and d_3-Ac. Hence, upon cleavage and MALDI-TOF-MS analysis, relevant peaks for sequencing are readily identified by the characteristic isotopic pattern of equal intensity peaks at $[M+H]^+$ (for Ac-terminated sequence) and $[M+3+H]^+$ (for d_3-Ac-terminated sequence). Identification is also facilitated by the application of software (Masslynx v3.2, Micromass Ltd., Manchester, UK, cluster analysis function) which can identify such isotopic patterns within a limit of statistical tolerance.

Attempts to utilise this strategy proved capricious, most likely due to the sensitivity of the deuterated reagent; the method could not be used reliably in the course of this work. However, the basic principle of utilising an isotopically-differentiated mixture as the capping reagent is a useful one. Characteristic isotopic patterns occur in MS with chlorinated and brominated compounds, since these halogens occur naturally as a constant isotopic mixture. In particular, naturally occurring bromine consists of an equiabundant mixture of ^{79}Br and ^{81}Br. Hence monobrominated compounds are easily identified in MALDI-TOF-MS as the characteristic isotope pattern of $[M+H]^+$ and $[M+2+H]^+$ (molecular ion and quasi-molecular ion), which is referred to informally as a 'bromine doublet' or 'doublet' in this thesis. More complex isotope patterns are seen with polybrominated species.

It is thus feasible that a monobrominated capping reagent could be implemented in the generation of ladder sequences to the same effect as Bradley's use the of AcOH/d_4-AcOH

mixture, with similar facile identification of relevant ladder peaks from noise. It is to be noted that using such a reagent does not serve to distinguish isobaric residues (e.g. Leu from Ile), which is circumvented in other methods by using capping reagents of two distinct types (e.g. Bradley *et al*[93], by the use of either Ac/d_3-Ac or Bz/d_5-Bz for the isobaric pair).

A similar method for easy identification of relevant ladder peaks has recently been reported by Lam *et al*,[91] who incorporate bromine into a *linker* in the form of bromophenylglycine (Fig 2-1). This however means that *every* cleavage product from the bead will contain bromine, including any articles of mis-synthesis on the solid phase, potentially making MS-analysis more complex. By using the isotopic label in the capping group, as in the current work, this problem is avoided.

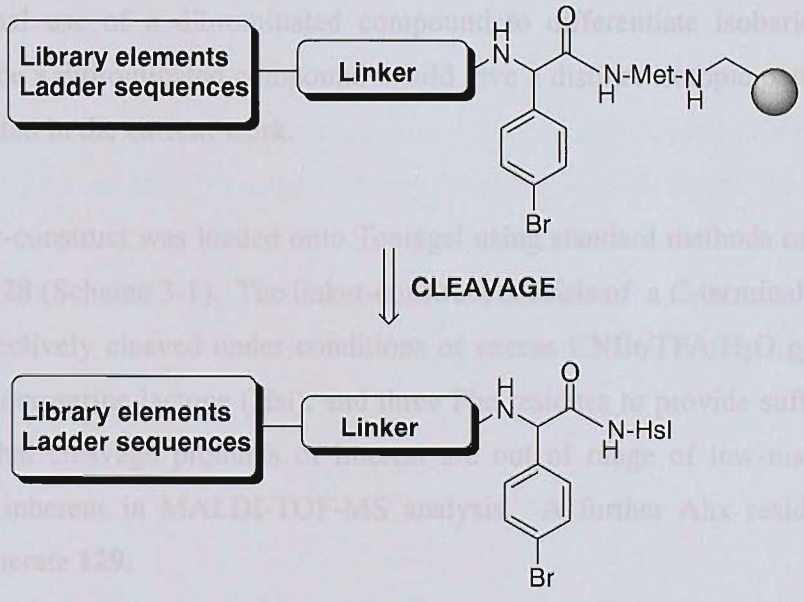


Fig 3-1 Lam *et al*'s paradigm for isotopic detection of ladder sequences Note that all species cleaved from the bead will contain bromine, including library elements *not* devoted to capping, potentially complicating MS results

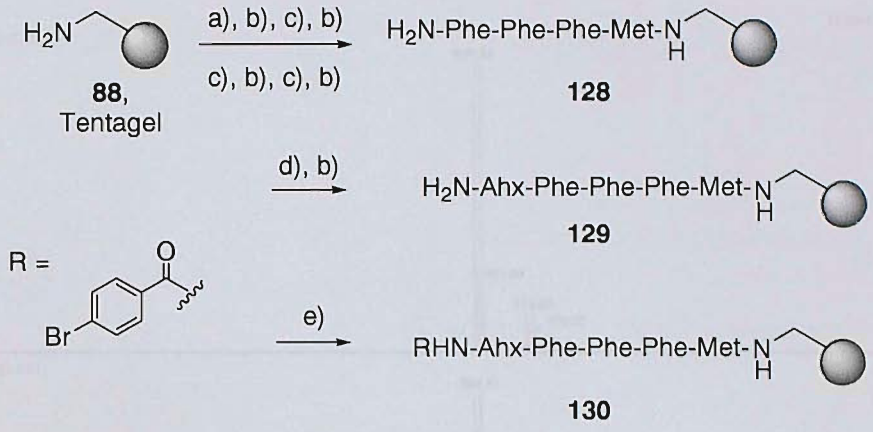
Regarding MS-analysis, MALDI-TOF is the method of choice. Owing to the nanomolar quantities of material represented by a single resin bead, and the fact that the 'ladder' sequences of interest will be present in fractions of this quantity (i.e. picomolar amounts),

a highly sensitive method is required. MALDI-TOF-MS provides this sensitivity whilst still allowing ease of analysis. ES-systems in theory should also be suited to this work, as peptides are readily protonated under ES conditions and are detectable at the amount limits indicated. However, attempted use of automated ES-MS for analysis proved fruitless, probably due to the complex nature of the mixtures released from the resin on cleavage and probable contaminants such as PEG. Reliable detection of a bromine isotope pattern was not observed for systems which were reliably analysed by MALDI-TOF, and hence ES methods were not further developed.

3.2.1 Evaluation of capping unit

para-Bromobenzoic acid (abbreviated as *p*-Br PhCOOH) was deemed a suitable capping reagent, in tandem with standard amide coupling reagents (DIC/HOBt/DIPEA in DMF). The additional use of a dibrominated compound to differentiate isobaric residues is possible, since a dibrominated compound would give a distinct isotopic pattern; this was not investigated in the current work.

First a linker-construct was loaded onto Tentagel using standard methods of Fmoc-SPPS to generate **128** (Scheme 3-1). The linker-construct consists of a *C*-terminal Met residue, which is selectively cleaved under conditions of excess CNBr/TFA/H_2O generating the *C*-terminal *homo*serine lactone (Hsl), and three Phe residues to provide sufficient 'dead' mass such that cleavage products of interest are out of range of low-mass noise and matrix ions inherent in MALDI-TOF-MS analysis. A further Ahx residue was then loaded to generate **129**.



Reagents and conditions: a) Fmoc-Met-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% piperidine/DMF, 20 mins; c) Fmoc-Phe-OH, DIC, HOBt, DIPEA, DMF, 3 hours; d) Fmoc-Ahx-OH, DIC, HOBt, DIPEA, DMF, 3 hours; e) *p*-Br PhCOOH, DIC, HOBt, DIPEA, DMF, 18 hours

Scheme 3-1 Synthesis of linker-construct 128 for investigating capping methodology and demonstration of exhaustive capping

This resin **129** was subjected to a mixture of excess *p*-Br PhCOOH and excess DIC/HOBt/DIPEA in DMF. After thorough washing and drying of the resin, single beads were picked using tweezers, individually cleaved and the cleavage mixture analysed by MALDI-TOF-MS. The presence of the expected 'doublet' for $[M+Na]^+$ at *m/z* 850/852 and the lack of a peak for the uncapped peptide confirmed the potential of this method (Fig 3-2).

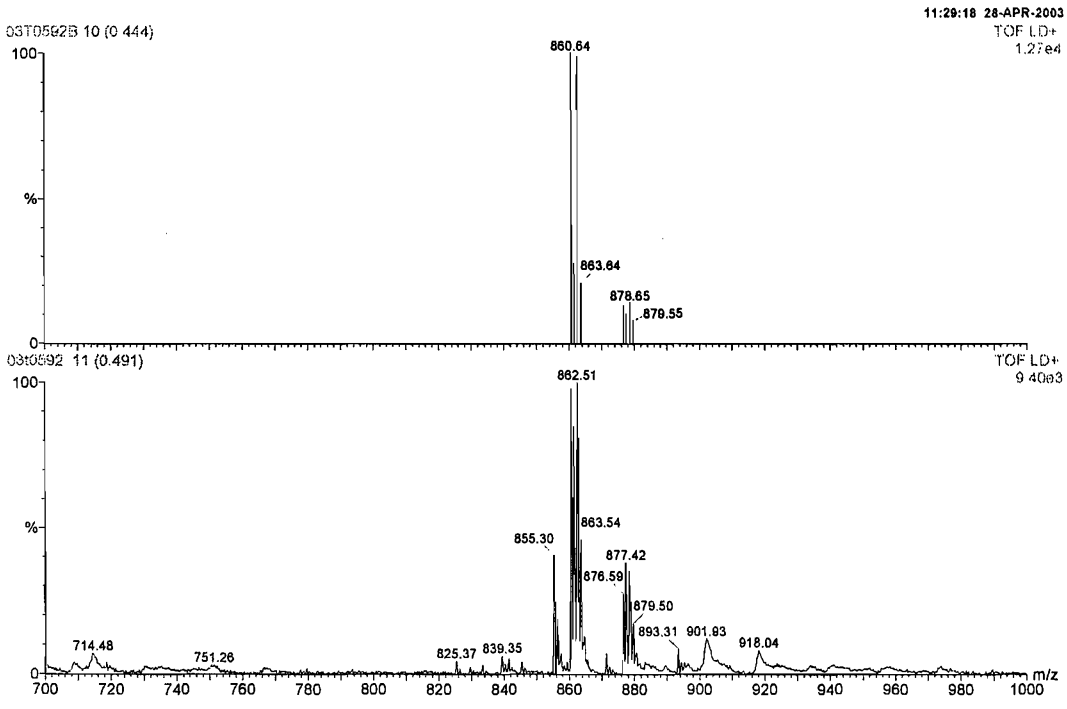
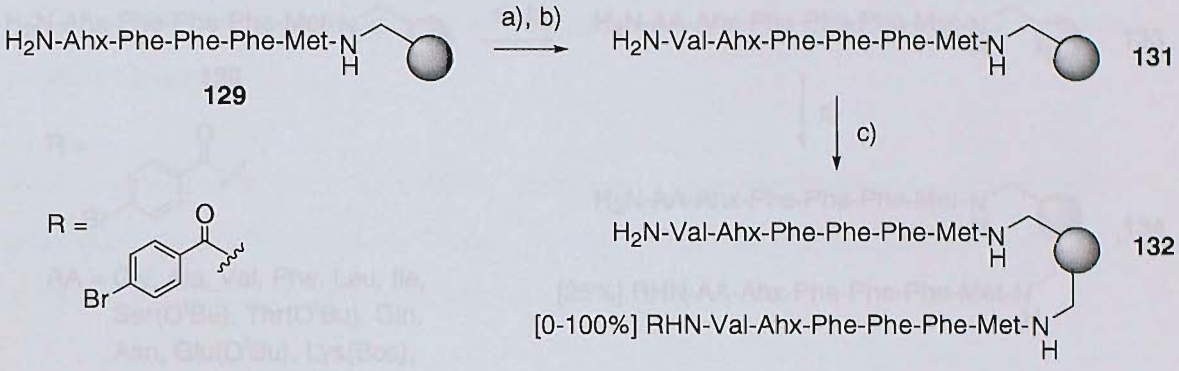


Fig 3-2 Sample MALDI-TOF-MS spectra of a single cleaved bead of 130 Lower spectrum represents raw data, upper spectrum represents software-processed data (for detection of Br-isotope patterns). The isotopic 'doublet' of the expected cleavage product is obvious, in this case as both $[M+Na]^+$ and $[M+K]^+$ adducts

Experiments were then undertaken to investigate the detection limit for capped peptide sequences. To this end, resin **131** was generated by coupling Val to resin **129** by Fmoc-SPPS (Scheme 3-2). This was to test against the most hindered amino acid, which should provide the most stringent test of reaction efficiency. Samples of resin **131** were then submitted to reaction with varying proportions of *p*-Br PhCOOH relative to the resin loading (from 0-100%) and excess DIC/HOBt/DIPEA in an 18 hour reaction. Single beads were selected from each sample, cleaved using CNBr/TFA/H_2O and the cleavage products were analysed by MALDI-TOF-MS. It was found that the capped sequence (as the $[M+Na]^+$ adduct) could be detected down to the 10%-loaded level, however the relevant 'doublet' is much more easily identified and more reliably extracted by software at higher proportions. For further experiments, it was decided to use a 25% proportion of capping group for ease of experiment read-out.

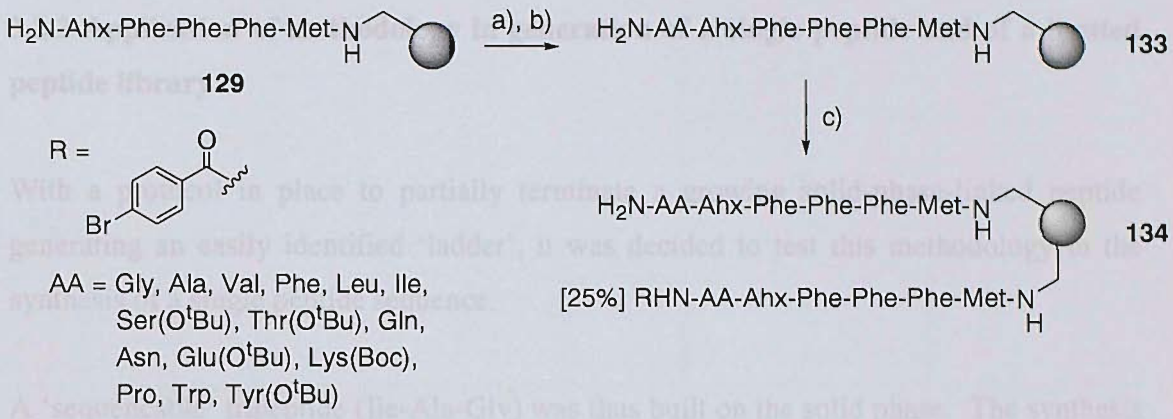


Reagents and conditions: a) Fmoc-Val-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% piperidine/DMF, 20 mins; c) *p*-Br PhCOOH (various mol%), DIC, HOBt, DIPEA, DMF, 18 hours

Scheme 3-2 Capping at various loadings of capping group

Efficiency of the capping method was then evaluated. Further samples of **131** were reacted with varying proportions of *p*-Br PhCOOH, then cleaved in bulk and the cleavage mixtures were analysed by HPLC. It appears that the capping reaction is efficient as in all cases the relative peak integration of the capped and uncapped products corresponded well to the proportion of the capping reagent used.

It was desirable to know whether each individual amino acid could be capped in an identical manner. To this end, samples of resin **129** were individually loaded with an amino acid by Fmoc-SPPS. Amino acids used were Gly, Ala, Val, Phe, Leu, Ile, Ser(O^{t}Bu), Thr(O^{t}Bu), Gln, Asn, Glu(O^{t}Bu), Lys(Boc), Pro, Trp and Tyr(O^{t}Bu). (Met and Cys are incompatible with the CNBr-mediated cleavage, His and Asp were unavailable and Arg is to be avoided in future application of this methodology in the synthesis of guanidinium-based receptors).



Reagents and conditions: a) Fmoc-AA-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% piperidine/DMF, 20mins; c) i) 25mol% *p*-Br PhCOOH, DMF, 1 hour, ii) addition of DIC, HOBt, DIPEA, DMF, 18 hours

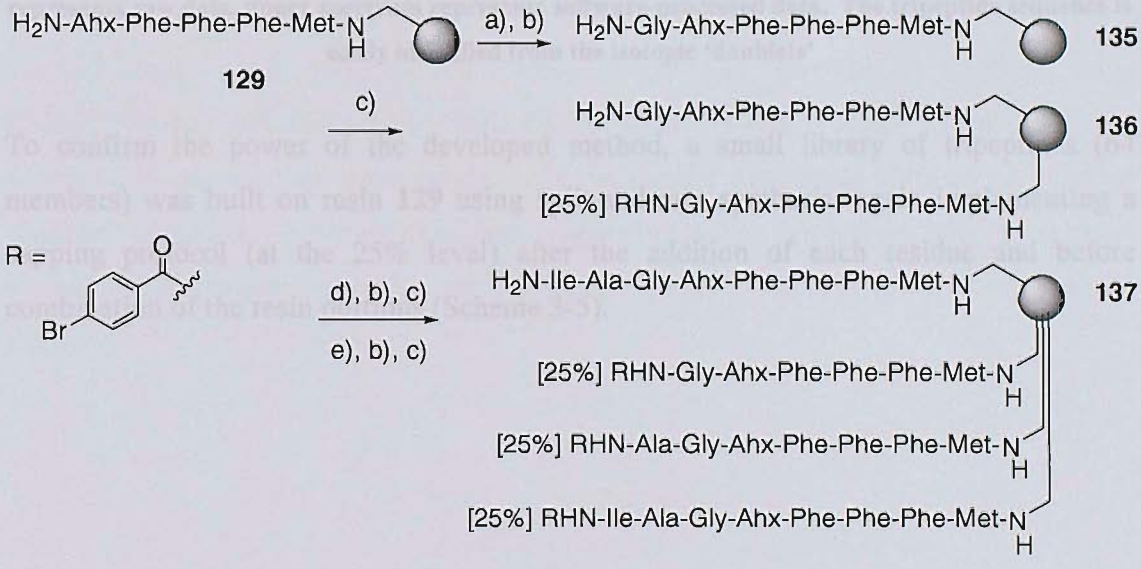
Scheme 3-3 Capping of various amino acids on the test construct 133

At this stage it was discovered that certain beads within a sample of resin were not being capped (a lack of the isotopic 'doublet' in the MS spectrum) whereas other beads were being completely capped (lack of uncapped product in MS). This was hypothesised to be due to a combination of inefficient mixing of the capping group throughout the resin sample and a highly efficient coupling reaction. To account for this, an alternative capping protocol was instigated, whereby a solution of *p*-Br PhCOOH in DMF was first allowed to pre-mix with the resin sample for at least an hour before addition of coupling reagents (DIC/HOBt/DIPEA) to allow an even distribution of the capping group throughout the resin sample (Scheme 3-3). Utilising this method, the problem of unequally distributed capping appeared to be solved; all of the tested amino acids showed detectable incorporation of the capping group at the level of a single bead (MALDI-TOF-MS analysis). It was noted that the acid-sensitive side-chains (*tert*-butyl esters, *tert*-butyl ethers, Boc groups) were fully deprotected during the cleavage protocol, except in the case of Ser(O^{t}Bu) which only partially deprotected; the reasons for this are unclear.

3.2.2 Application of methodology in generation of a single peptide and of a limited peptide library

With a protocol in place to partially terminate a growing solid-phase-linked peptide generating an easily identified 'ladder', it was decided to test this methodology in the synthesis of a single peptide sequence.

A 'sequencable' tripeptide (Ile-Ala-Gly) was thus built on the solid phase. The synthesis is illustrated in Scheme 3-4; resin construct **129** was sequentially loaded as normal for Fmoc-SPPS, but with the implementation of the capping protocol after each cycle of coupling-deprotection to generate heterogeneous resin **137**. Single bead cleavages and MALDI-TOF-MS analysis as standard clearly showed the presence of all the 'ladder' sequences and the peptide sequence could be read off from the spectrum by calculating mass differences (accounting for the different adducts in certain cases). A sample spectrum is shown in Fig 3-3).



Reagents and conditions: a) Fmoc-Gly-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% piperidine/DMF, 20 mins, c) i) *p*-Br PhCOOH (25mol%), DMF, 1 hour, ii) addition of DIC, HOBt, DIPEA, 18 hours; d) Fmoc-Ala-OH, DIC, HOBt, DIPEA, DMF, 3 hours; e) Fmoc-Ile-OH, DIC, HOBt, DIPEA, DMF, 3 hours

Scheme 3-4 Synthesis of a single, 'laddered' tripeptide 137

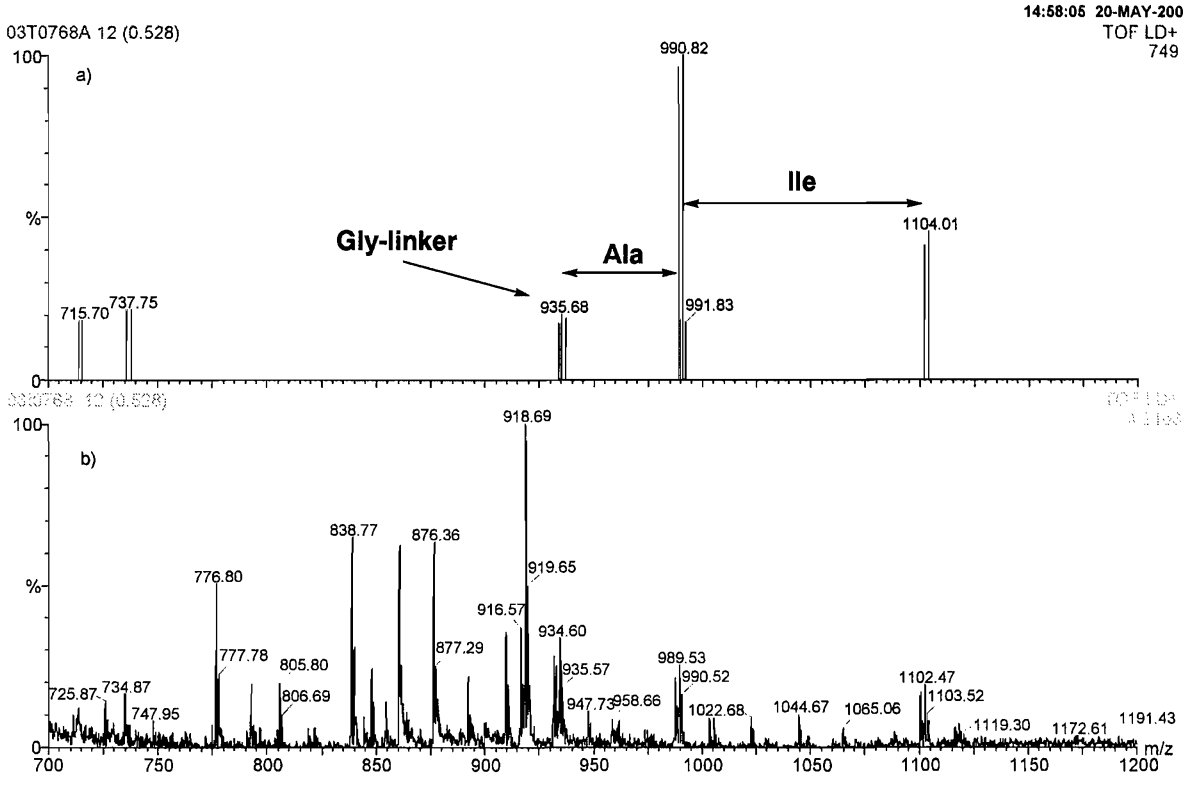
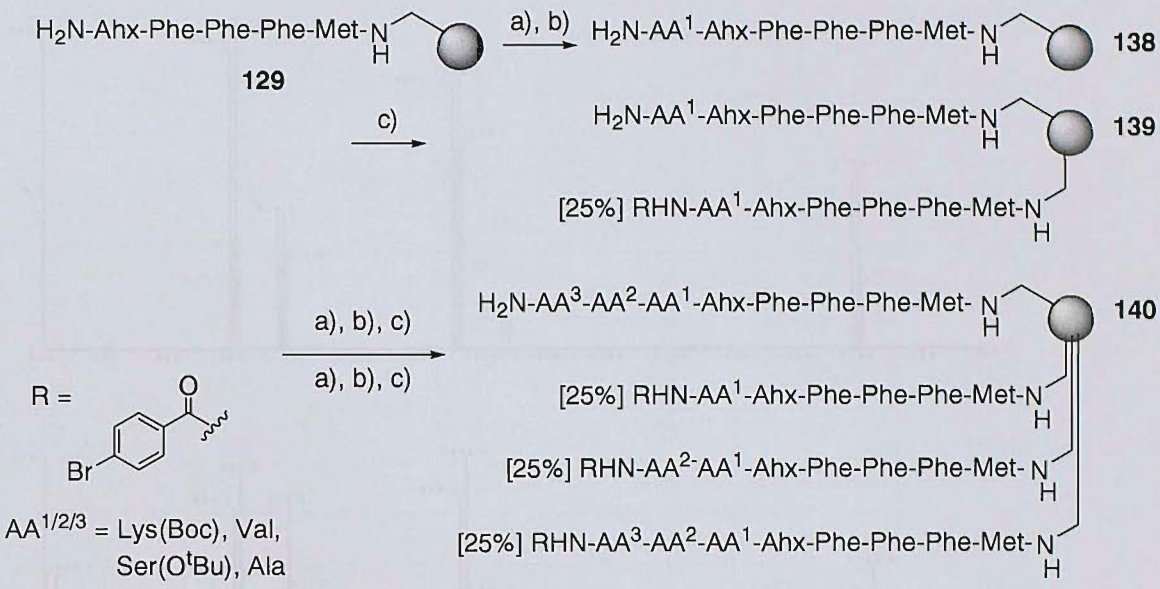


Fig 3-3 Sample MALDI-TOF-MS spectrum of a single cleaved bead of 137. Lower spectrum represents raw data, upper spectrum represents software-processed data. The tripeptide sequence is easily identified from the isotopic 'doublets'

To confirm the power of the developed method, a small library of tripeptides (64 members) was built on resin **129** using split-and-mix synthesis, again implementing a capping protocol (at the 25% level) after the addition of each residue and before combination of the resin portions (Scheme 3-5).



Reagents and conditions: a) Fmoc-AA-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% piperidine/DMF, 20 mins; c) i) *p*-Br PhCOOH (25mol%), DMF, ii) addition of DIC, HOBt, DIPEA, 18 hours

Scheme 3-5 Synthesis of a limited 'laddered' tripeptide library

Twenty single beads were selected from library **140**, cleaved as standard using CNBr/TFA/H_2O and analysed by MALDI-TOF-MS. It was possible to fully assign a sequence to eighteen of these beads. The remaining two beads showed no relevant 'doublet' signals in the mass spectrum, suggesting either imperfect resin or cleavage thereof. The resultant sequences are shown in Table 3-1. Sample spectra are shown in Fig 3-4.

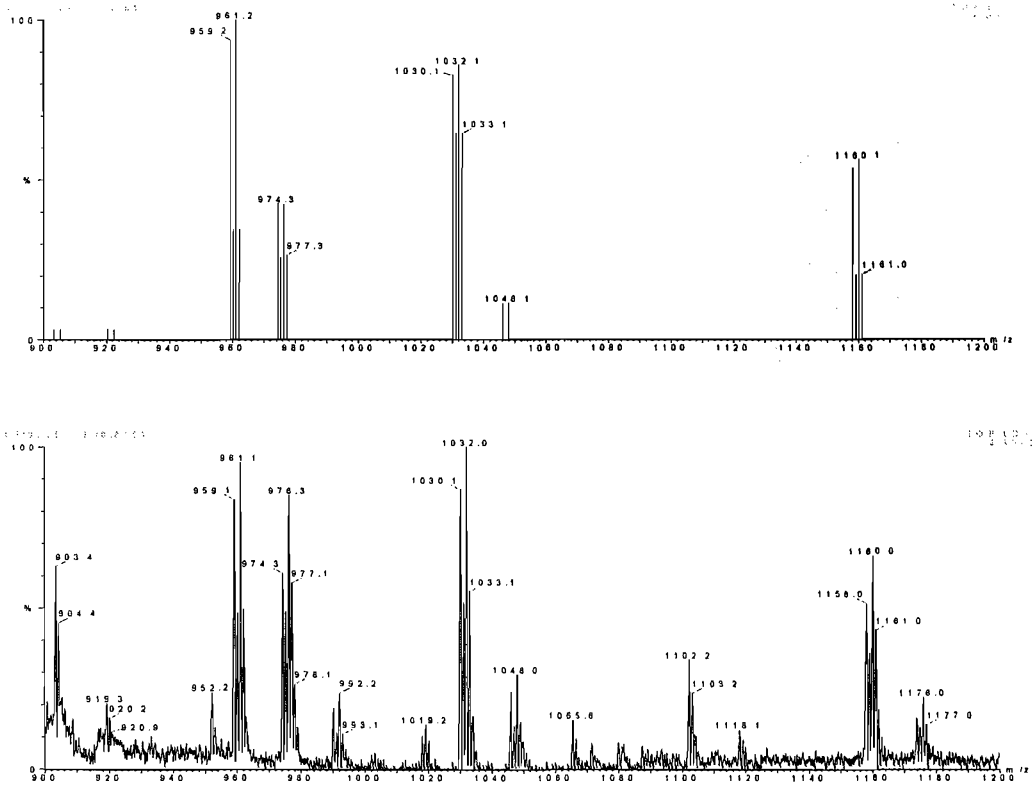


Fig 3-4 Sample MALDI-TOF-MS spectrum of a single cleaved bead of 140. Lower spectrum represents raw data, upper spectrum represents software-processed data. The tripeptide sequence is easily identified from the isotopic 'doublets' *without prior knowledge* of the sequence

Bead	AA^1	AA^2	AA^3
1	Ala	Ala	Ala
2	Ser(O^{t}Bu)	Ala	Ala
3	Lys(Boc)	Ala	Val
4	Ser(O^{t}Bu)	Val	Val
5	Val	Lys(Boc)	Ser(O^{t}Bu)
6	Lys(Boc)	Ala	Lys(Boc)
7	Lys(Boc)	Val	Ala
8	Ser(O^{t}Bu)	Ala	Lys(Boc)
9	Lys(Boc)	Val	Val
10	Lys(Boc)	Val	Ser(O^{t}Bu)
11	Ala	Lys(Boc)	Lys(Boc)
12	Ser(O^{t}Bu)	Ala	Lys(Boc)
13	Ala	Val	Val
14	Ser(O^{t}Bu)	Ala	Val
15	Ser(O^{t}Bu)	Lys(Boc)	Lys(Boc)
16	Ser(O^{t}Bu)	Ser(O^{t}Bu)	Val
17	Lys(Boc)	Ala	Val
18	Ala	Ser(O^{t}Bu)	Ser(O^{t}Bu)
19	-	-	-
20	-	-	-

Table 3-1 MS Sequencing results of random sample beads taken from library 140

3.2.3 Use of highly-ionisable linker

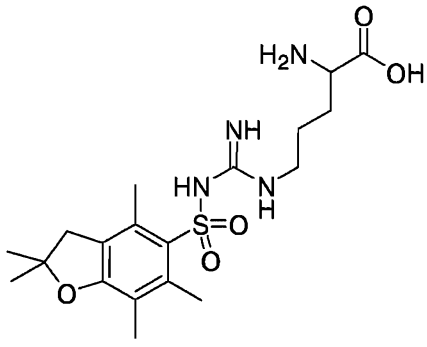
Despite the success achieved with the capping protocol to this point, its application is obviously limited. By utilising 25% of the bead composition for the generation of each 'ladder' sequence, potential diversity in library structure becomes highly limited; either a tripeptide is the limit (since it is desirable not to completely terminate a bead by blocking all free sites) or careful design avoiding isobaric oligopeptide units is required (such as by

use of the 'Biblio' algorithm[89]). Thus it was desirable to attempt to reduce the required proportion of capping group for reliable detection.

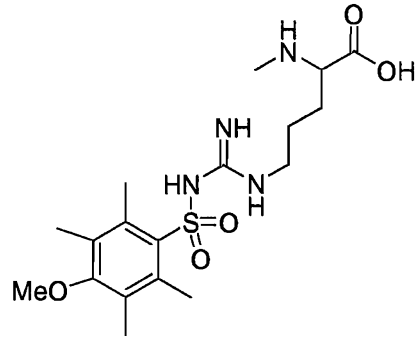
Youngquist *et al*'s original work[84] incorporated an unprotected Arg-unit into the linker. This is doubly favourable as it not only adds requisite 'dead' mass by also introduces a guanidine unit. Guanidines are highly ionisable species (pK_a of alkylguanidiniums ~14) and are detectable by MS at extremely low concentration. In this way, a lesser absolute amount of a particular 'ladder' sequence (as cleaved from a single bead) is required for detection, as the great propensity towards ionisation of the guanidine unit increases the signal to noise ratio.

Since the implementation of this methodology is to be in building libraries of receptors based on a guanidinium-derived binding site, the incorporation of a second guanidine into the linker is undesirable. Its presence would allow a higher degree of 'background' binding in a solid-phase screen, perhaps making differentiation between strongly binding beads and weakly binding beads more difficult. It is also possible that stronger binding would exist between a putative guest and the linker, hence selectivity of binding would be based on a different species than the designated library member built onto a bead.

Hence it was necessary to use a protected Arg as part of the linker, which could then be simply deprotected after screening and prior to cleavage, allowing release of the free guanidine for facile analysis by MS. Pbf[171] (Fig 3-5) is a suitable group for this function, as it is cleaved in 50% TFA/DCM, and is compatible with Fmoc-SPPS. An alternative is the more acid-stable Mtr[172], cleavable in 1M TMSBr in TFA.



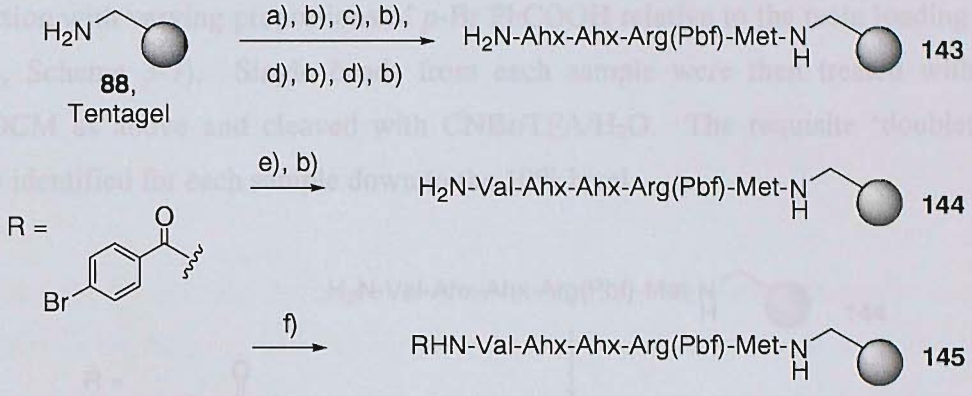
141



142

Fig 3-5 Protected Arg units for solid-phase synthesis; Arg(Pbf) 141, Arg(Mtr) 142

To test the use of an Arg(Pbf)-containing linker, the resin **143** was built by standard Fmoc-SPPS (Scheme 3-6). This linker represents that which would be used for final library synthesis, with flexible Ahx units instead of Phe units (which were also noted to cause documented problems of sequence aggregation[173] as mass loaded to the resin increased). **143** was then loaded with Val by Fmoc-SPPS to give resin **144**, which was then treated with excess *p*-Br PhCOOH and DIC/HOBt/DIPEA. After requisite washing and drying steps, single beads were selected and sequentially treated with 50% TFA/DCM (to remove Pbf) and then the standard cleavage mixture of CNBr/TFA/H_2O. MALDI-TOF-MS analysis revealed a clearly identifiable 'doublet' at the appropriate mass. (Fig 3-6)



Reagents and Conditions: a) Fmoc-Met-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% piperidine/DMF, 20 mins; c) Fmoc-Arg(Pbf)-OH, DIC, HOBt, DIPEA, DMF, 3 hours; d) Fmoc-Ahx-OH, DIC, HOBt, DIPEA, DMF, 3 hours; e) Fmoc-Val-OH, DIC, HOBt, DIPEA, DMF, 3 hours; f) i) *p*-Br PhCOOH, DMF, 1 hour ii) addition of DIC, HOBt, DIPEA, DMF, 18 hours

Scheme 3-6 Synthesis of Arg-containing-construct 143 for investigating capping methodology and demonstration of exhaustive capping

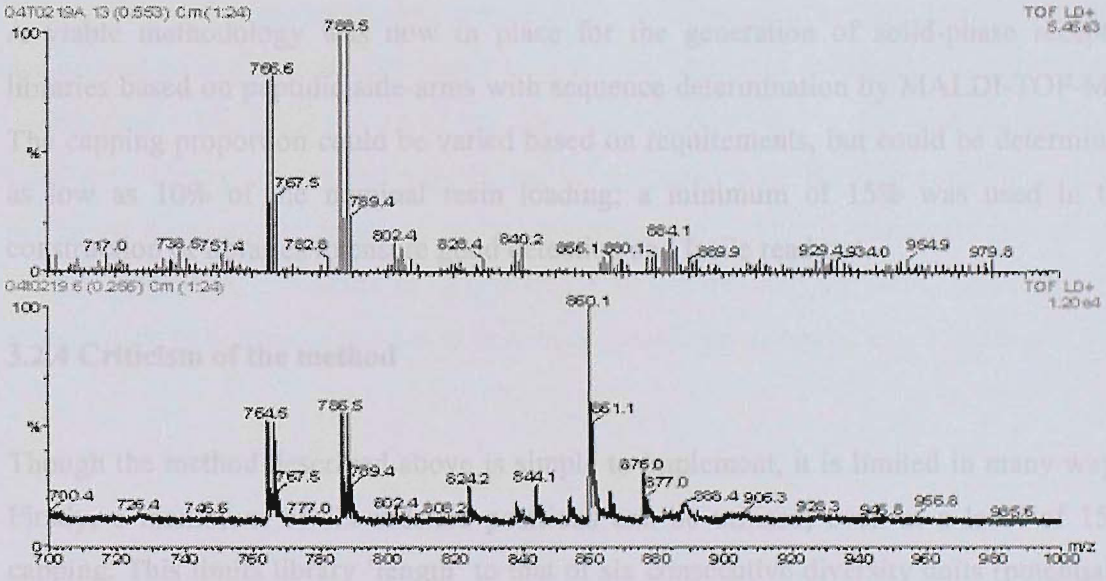
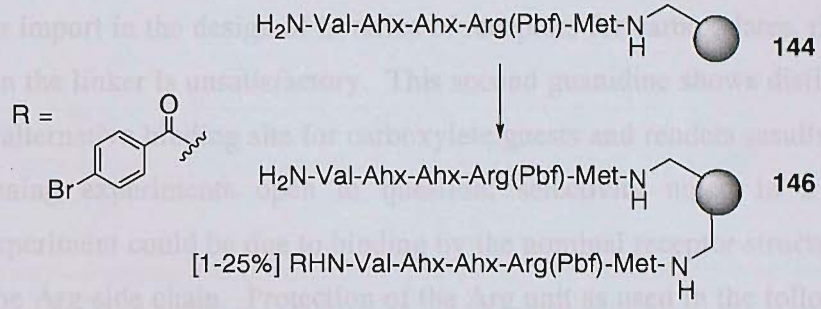


Fig 3-6 Sample MALDI-TOF-MS spectrum of a single cleaved bead of 145. Lower spectrum represents raw data, upper spectrum represents software-processed data. The isotopic 'doublets' are clearly in evidence in both spectra, as $[M+H]^+$ and $[M+Na]^+$ adducts

In a similar manner to that followed with resin **131**, samples of resin **144** were submitted to reaction with varying proportions of *p*-Br PhCOOH relative to the resin loading (from 1-25%, Scheme 3-7). Single beads from each sample were then treated with 50% TFA/DCM as above and cleaved with CNBr/TFA/H_2O. The requisite 'doublet' was clearly identified for each sample down to the 10% level.



Reagents and Conditions: a) i) *p*-Br PhCOOH (various mol%), DMF, 1 hour, ii) addition of DIC/HOBt/DIPEA 18 hours

Scheme 3-7 Capping the Arg-containing-construct at various loadings of capping group

A viable methodology was now in place for the generation of solid-phase receptor libraries based on peptidic side-arms with sequence determination by MALDI-TOF-MS. The capping proportion could be varied based on requirements, but could be determined as low as 10% of the nominal resin loading; a minimum of 15% was used in the construction of libraries to ensure good detection and facile read-out.

3.2.4 Criticism of the method

Though the method described above is simple to implement, it is limited in many ways. Firstly, a maximum of six indexed positions can be utilised, even at a level of 15% capping. This limits library 'length' to that of six consecutive diversity units (potentially longer, if an algorithm such as 'Biblio'[89] is consulted in library design). This theoretical limit cannot be achieved in practice; owing to the process of split-and-mix, the absolute quantity of available reactive sites on individual beads can only be estimated and care must be taken not to over-load the capping group at any particular point. Over-loading of

the capping group would potentially 'kill' the growing peptide sequence before library synthesis is complete by blocking all the reactive sites. This would result in a highly heterogeneous and undefined library composition. Hence the limit of six indexed positions, which would utilise $6 \times 15\% = 90\%$ of the bead composition for the generation of ladder sequences, cannot be achieved in practice.

Of particular import in the design of libraries of receptors for carboxylates, the necessary use of Arg in the linker is unsatisfactory. This second guanidine shows distinct potential to act as an alternative binding site for carboxylate guests and renders results from solid-phase screening experiments open to question; selectivity noted in a solid-phase screening experiment could be due to binding by the nominal receptor structure *or* due to binding at the Arg-side chain. Protection of the Arg unit as used in the following library syntheses is a viable solution to this problem.

3.3 Synthesis of a Single-Armed Tweezer Library 149

With a functioning methodology for generating 'ladder sequences' (and hence 'readable' libraries of beads) in place, synthesis of a simple library of receptors for carboxylates commenced. The basic design of the library is indicated in Fig 3-7, and consists of a CBS linked to a randomised peptide sequence loaded onto the Arg-containing linker necessary for sensitive detection of 'ladder' sequences. This is described as a single-armed receptor as it consists of only a single strand flanking the CBS, rather than the usual two or three strands that act to envelop a guest in an authentic tweezer receptor.

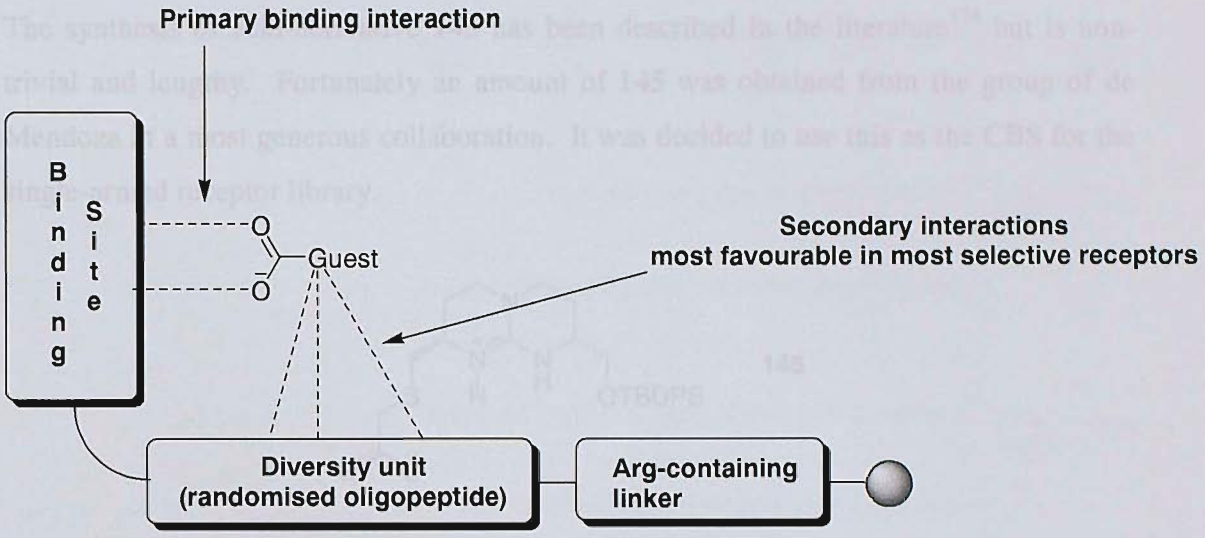


Fig 3-7 Schematic of a simplistic receptor for carboxylates. The binding site is the point of major interaction, with the diversity unit accounting for further favourable binding interactions. As the nature of the carboxylate guest varies (different peptide guests), a different diversity unit represents the most strongly binding library member.

An effective CBS was required, since it was desired to determine receptors for carboxylates in aqueous medium. One effective motif for this purpose is the bicyclic guanidinium moiety of de Mendoza *et al*[39] (Fig 3-8) which is known to bind carboxylates in aqueous systems, owing to the formation of strong ion-pair interactions. This bicyclic system is also highly conformationally defined. This is beneficial in the designing of selective receptors, since the randomised peptide portion of the receptors will be held in a defined region of space ensuring a secondary interaction between potential guests and this unit. This is particularly important for the single-armed receptor structure, where a guest will not be enveloped by side-arms on interaction with the CBS.

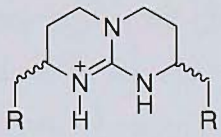
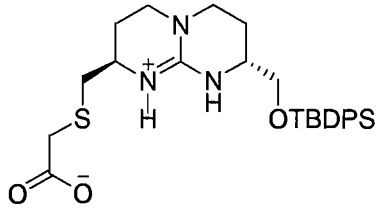


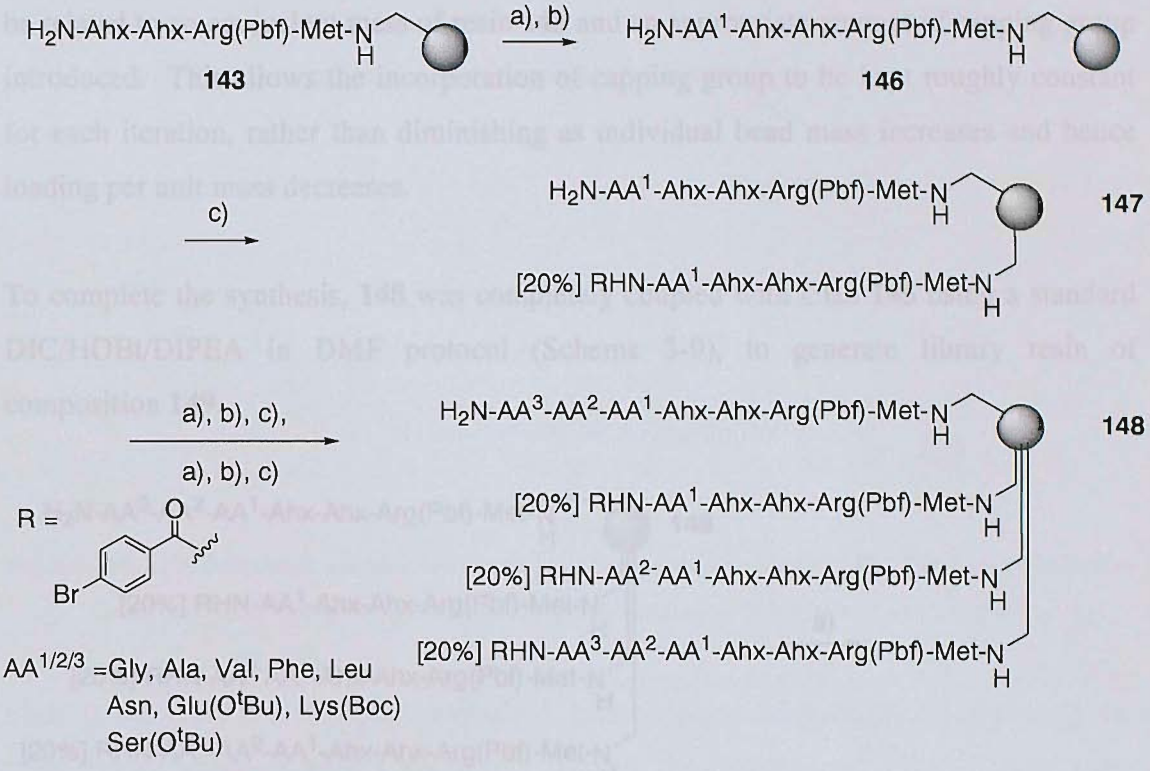
Fig 3-8 de Mendoza *et al*'s bicyclic guanidinium unit for binding carboxylates

The synthesis of acid-derivative **145** has been described in the literature[174] but is non-trivial and lengthy. Fortunately an amount of **145** was obtained from the group of de Mendoza in a most generous collaboration. It was decided to use this as the CBS for the single-armed receptor library.



145

Library synthesis was carried out as in Schemes 3-8 and 3-9. Using the construct **143** as a basis, three cycles of split-and-mix synthesis were performed as follows; in each cycle the resin was split into nine portions and each portion was individually coupled with an Fmoc-amino acid from the following list; Gly, Ala, Val, Leu, Phe, Asn, Ser(O^{t}Bu), Glu(O^{t}Bu), Lys(Boc). Fmoc-deprotection was then carried out on each portion. Both of these steps were monitored by the Kaiser ninhydrin test. Each portion of resin was then submitted to the capping protocol with *p*-Br PhCOOH (20% of nominal loading) to generate a 'ladder' sequence for that residue. The resin portions were then thoroughly mixed together by shaking whilst suspended in DMF. After three cycles, the library composition is given by **146** (Scheme 3-8).



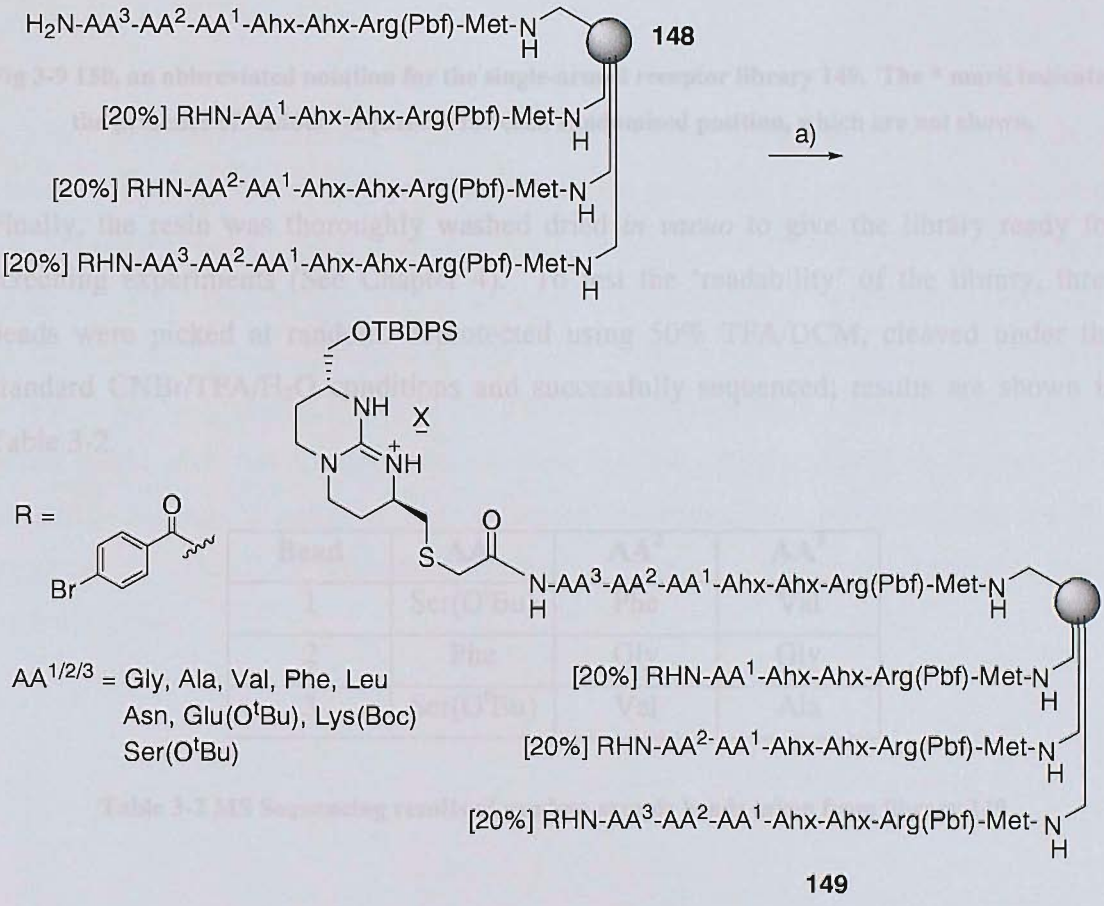
Reagents and Conditions: a) Fmoc-AA-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% piperidine/DMF, 20 mins; c) i) *p*-Br PhCOOH, DMF, 1 hours, ii) addition of DIC, HOBt, DIPEA, 18 hours

Scheme 3-8 Synthesis of a tripeptide library 148, indexed with 'ladder sequences' for MS determination of members

It should be noted that owing to the non-homogeneous nature of beads, it is not possible to define absolute NH_2-loadings at each step. This arises from the incorporation of different units during the split-and-mix procedure, from the presence of the capping group in unspecified amount and from the non-homogeneous nature of the base resin of individual beads. Hence the quantity of capping group used in each application of the capping protocol is calculated based on the average additional mass gained by a bead up to that juncture. For example, after two cycles of split-reaction-mix, the average added mass per gram of resin is equivalent to the calculated average mass of two amino acids multiplied by the measured loading of resin **143**. (The additional mass of the incorporated capping group is not accounted for). Thus a given mass of resin library can

be related to an equivalent mass of resin **143** and an appropriate amount of capping group introduced. This allows the incorporation of capping group to be kept roughly constant for each iteration, rather than diminishing as individual bead mass increases and hence loading per unit mass decreases.

To complete the synthesis, **148** was completely coupled with CBS **145** using a standard DIC/HOBt/DIPEA in DMF protocol (Scheme 3-9), to generate library resin of composition **149**.



Reagents and conditions: a) **145**, DIC, HOBt, DIPEA, DMF, 18 hours

Scheme 3-9 Addition of CBS to the '148, generating single-armed receptor library 149

149 may be usefully abbreviated as structure **150**, where the * notation indicates the presence of ladder-sequences in a non-homogenous bead (Fig 3-9).

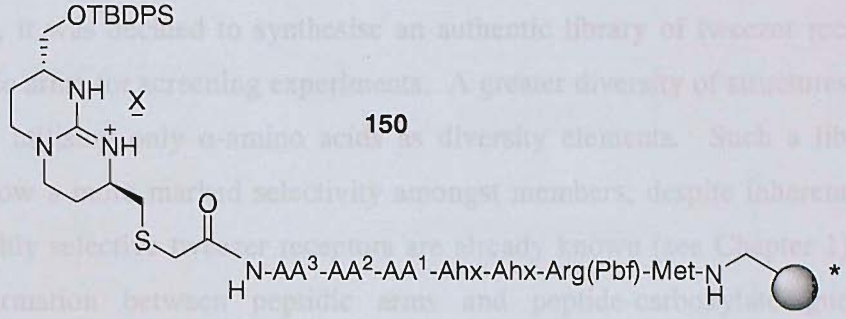


Fig 3-9 150, an abbreviated notation for the single-armed receptor library 149. The * mark indicates the presence of 'ladder' sequences for each randomised position, which are not shown.

Finally, the resin was thoroughly washed dried *in vacuo* to give the library ready for screening experiments (See Chapter 4). To test the 'readability' of the library, three beads were picked at random, deprotected using 50% TFA/DCM, cleaved under the standard CNBr/TFA/H_2O conditions and successfully sequenced; results are shown in Table 3-2.

Bead	AA^1	AA^2	AA^3
1	Ser(O^{t}Bu)	Phe	Val
2	Phe	Gly	Gly
3	Ser(O^{t}Bu)	Val	Ala

Table 3-2 MS Sequencing results of random sample beads taken from library 149

3.4 Synthesis of a Tweezer Library 184

Owing to the success observed in screening the single-armed receptor library **149** (see Chapter 4), it was decided to synthesise an authentic library of tweezer receptors with two peptidic arms for screening experiments. A greater diversity of structures is possible whilst still utilising only α-amino acids as diversity elements. Such a library would feasibly show a more marked selectivity amongst members; despite inherent flexibility, several highly selective tweezer receptors are already known (see Chapter 1). Potential β-sheet formation between peptidic arms and peptide-carboxylate guest is also anticipated.

In order to generate authentic tweezer receptors on the solid-phase, it is necessary to employ a usefully functionalised CBS. It is also necessary to consider the method of solid-phase synthesis. The primary concern is to avoid generation of 'truncated' receptor structures which contain the CBS but are missing particular diversity elements (Fig 3-10). In the context of a combinatorial library, the presence of such 'truncated' structures would imply the presence of two or more distinct receptors on an individual bead, any of which could give rise to identified selective binding associated with that bead. The diversity elements missing in the 'truncated' receptors, though present in the full structure, are effectively redundant. Screening is unlikely to reveal information about such positions.

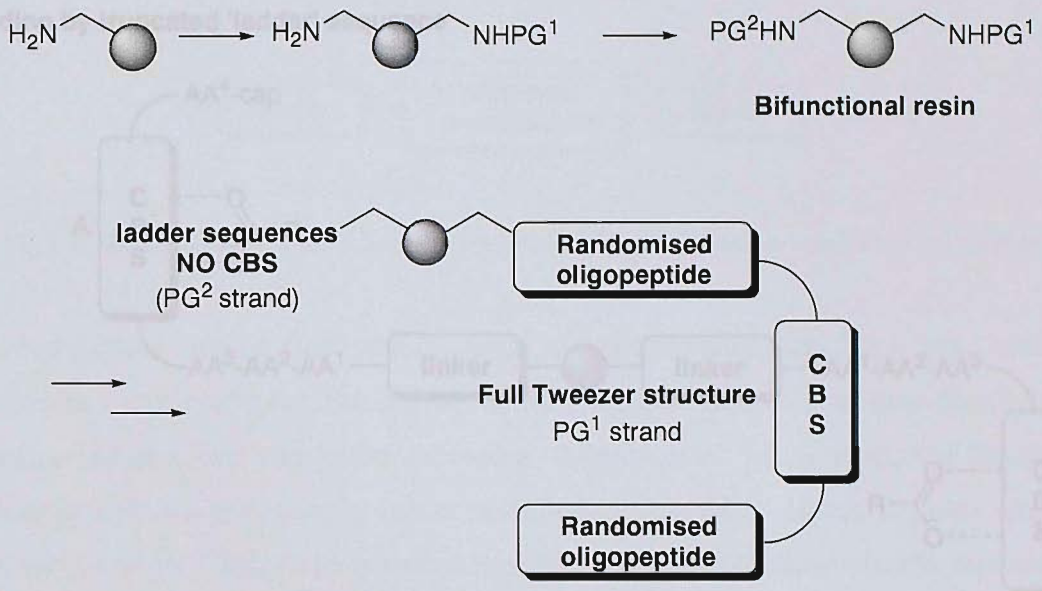


Fig 3-11 Proposed synthetic outline to a library of tweezer receptors. A "coding" strand is generated by partial derivatision of resin. 'Ladder' sequences are generated without CBS in this coding strand, full receptor structures are generated in the non-coding strand

Naturally, individual bead composition remains heterogeneous. However, since the primary interaction of any carboxylate guest will be an ion-pair interaction with a charged guanidinium binding-site, it is again conjectured that contributions from 'ladder' sequences to binding will be minimal and not strongly affect selectivity in a solid-phase screening experiment.

3.4.1 Synthesis of Binding Site

The CBS required for this library synthesis is of the general form of an extended, *N*-protected amino acid (Fig 3-12). This leads to facile incorporation of the CBS into a growing peptide strand by the usual methods of SPPS, and is amenable to standard methods of solid-phase reaction monitoring.

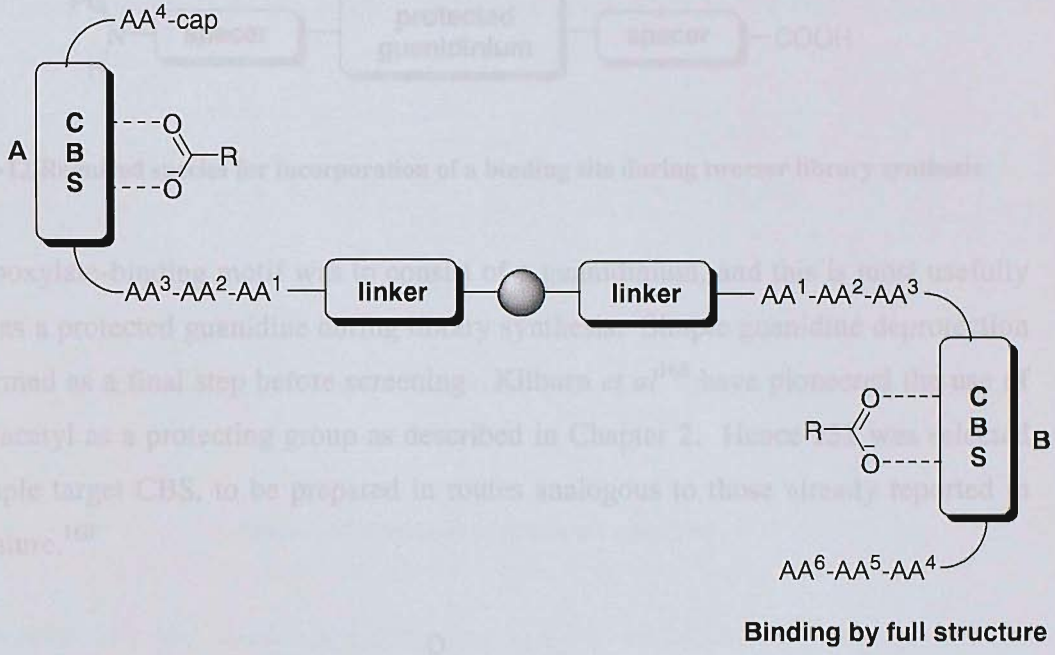
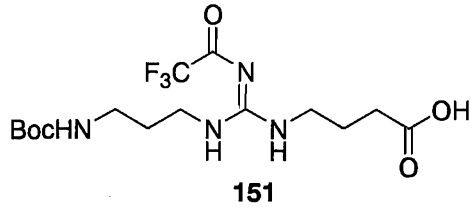


Fig 3-10 Problem of generating 'ladder' sequences in a linearly-synthesised tweezer library. If truncated sequences are present (A), the extra amino acids in the full structure (B) are effectively redundant

Unlike the single-armed receptor library **149**, where the CBS is incorporated in the final step (thus eliminating the issue), it was proposed to incorporate the CBS as an interpolation into a growing peptide sequence (as implied by Fig 3-10), which should simplify the solid-phase approach and reduce the need for orthogonal protecting groups. Hence it was inadvisable to implement capping methodology in the same simplistic manner as for the single-armed receptor library; this is guaranteed to generate truncated receptors terminated by the *p*-Br-PhCOOH group. To circumvent this problem, it was instead decided to divide the composition of the resin beads. Resin was partially derivatised with one protecting group before reaction with a second, orthogonal protecting group. This generates two distinct 'strands' (PG^1-strand and PG^2-strand). By selective deprotection, it is possible to generate the full receptor structure *with* the CBS as the sole unit in one strand. In tandem, 'ladder' sequences are generated in the other ("coding") strand *without* the presence of the CBS (Fig 3-11).

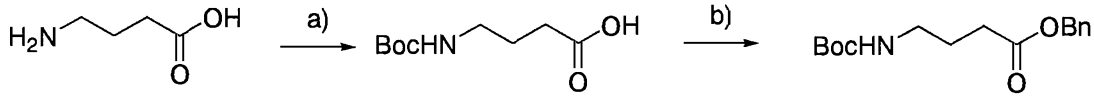
Fig 3-12 Required species for incorporation of a binding site during tweezer library synthesis

The carboxylate-binding motif was to consist of a guanidinium, and this is most usefully masked as a protected guanidine during library synthesis. Simple guanidine deprotection is performed as a final step before screening. Kilburn *et al*[168] have pioneered the use of trifluoroacetyl as a protecting group as described in Chapter 2. Hence **151** was selected as a simple target CBS, to be prepared in routes analogous to those already reported in the literature.[168]



151

First, γ-aminobutyric acid **152** was Boc-protected using a known procedure[175] to give acid **153**, and converted to the benzyl ester **154** using K_2CO_3 and benzyl bromide[176] (Scheme 3-10)

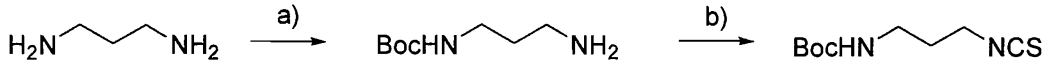


152 **153** **154**

Reagents and conditions: a) Boc_2O, Et_3N, MeOH, reflux, 3 hours, 90%; b) BnBr, K_2CO_3, wet acetone, 1 hour, 91%

Scheme 3-10 Synthesis of protected ester 154

Isothiocyanate **157** was synthesised according to a known procedure[177] via mono-Boc-protected diamine **156** (Scheme 3-11).

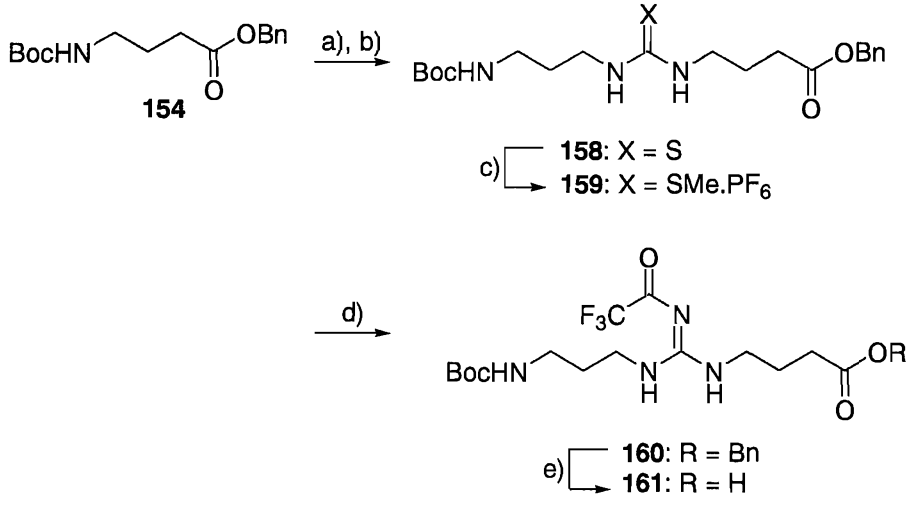


155 **156** **157**

Reagents and conditions a) 3 eq. Boc_2O, $CHCl_3$, 18 hours, 87%; b) Thiophosgene, K_2CO_3 (aq), $CHCl_3$, 18 hours, 84%

Scheme 3-11 Synthesis of isothiocyanate 157

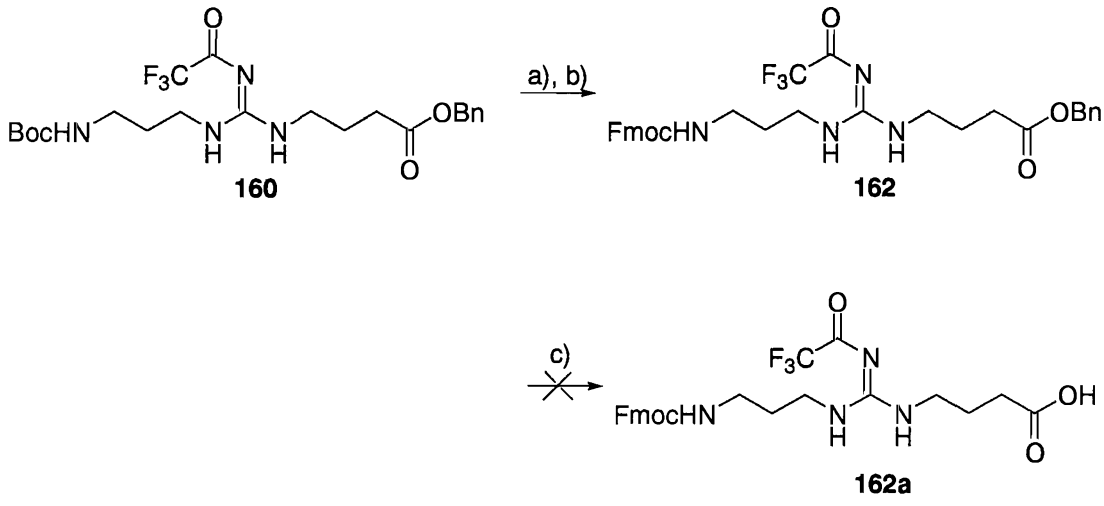
Boc-deprotection of ester **154** followed by coupling of the resulting amine salt with isothiocyanate **157** under basic conditions yielded thiourea **158** in satisfactory yield (Scheme 3-12). **158** could be converted into protected guanidine **160** using the reported three-step procedure via the thiouronium hexafluorophosphate salt **159**.[168] Final hydrogenolysis yielded the desired acid **151** in excellent yield (39% across 5-step sequence).



Reagents and conditions: a) 20% TFA/DCM, 1 hour; b) **157**, Et_3N, MeOH/DCM, 18 hours 75%; c) i) MeI, acetone, 1 hour, ii) NH_4PF_6, MeOH/DCM, 18 hours; d) F_3CCONH_2, DBU, toluene/$CHCl_3$, 1 hour, 50% over 3 steps; e) H_2, Pd/C, EtOAc, 3 hours 94%

Scheme 3-12 Synthesis of CBS acid 161

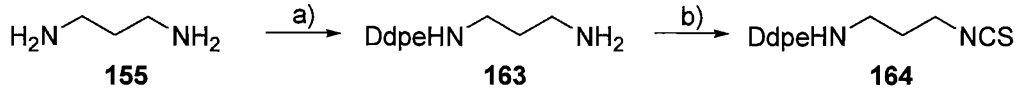
A similar route could be envisaged for the synthesis of the complementary Fmoc-protected acid **162a**. However, Fmoc is not compatible with the trifluoroacetylguanidine synthesis, owing to the use of strong base (DBU) to generate the appropriate nucleophile. It *is* possible to replace Boc with Fmoc at the stage of the guanidine ester (Scheme 3-13) to give the Fmoc compound **162** in good yield, but sadly attempted hydrogenolysis of **162** yielded a mixture of compounds from which the putative desired acid **162a** could not be satisfactorily purified.



Reagents and conditions: a) 20% TFA/DCM, 1 hour; b) Fmoc-Cl, Et_3N, DCM 1 hour, 71%; c) H_2, Pd/C, EtOAc, 2 hours

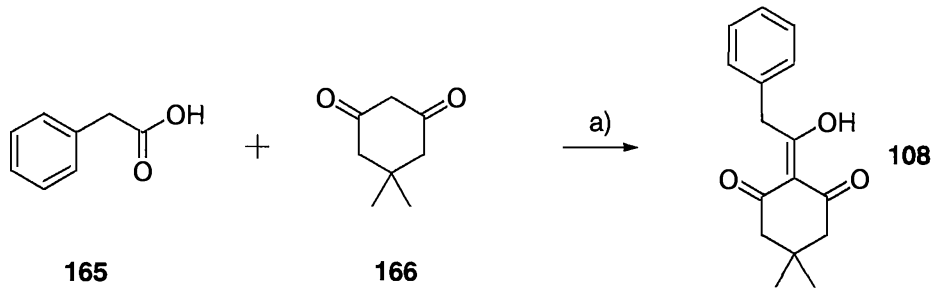
Scheme 3-13 Conversion to Fmoc-protection and failed attempted synthesis of Fmoc-protected acid 162a

An alternative to Fmoc is the Ddpe group, cleaved by hydrazine and orthogonal to Boc. CBS **170** could be synthesised in analogous manner to **151** as described below. 1,3-diaminopropane **155** was protected with Ddpe (prepared according to known procedure as in Scheme 3-15[168]) to give amine **163**. **163** was converted into the isothiocyanate **164** using standard conditions (Scheme 3-14).



Reagents and conditions: a) **108**, cat. TFA, DCM, 5 hours, quant.; b) Thiophosgene, K_2CO_3(aq), $CHCl_3$, 18 hours, 45%

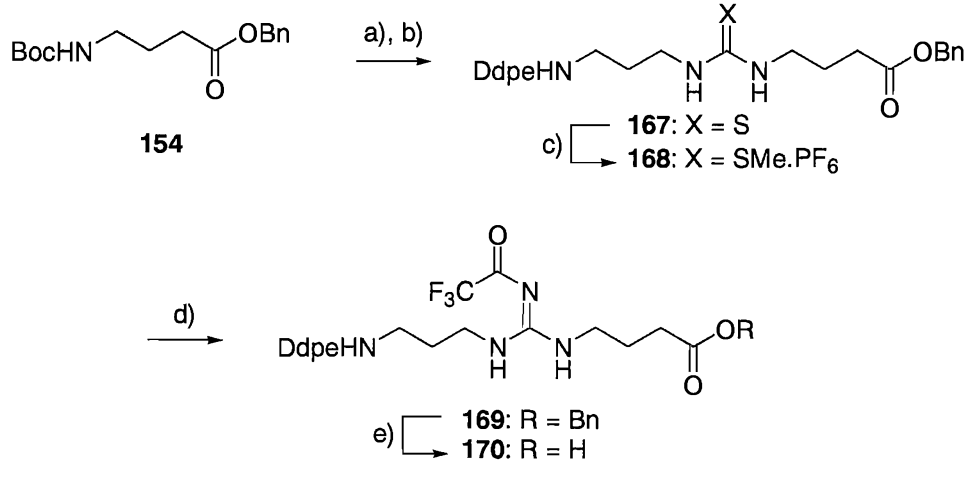
Scheme 3-14 Synthesis of isothiocyanate 164



Reagents and conditions: a) DCC, DMAP, DMF, 3 days, 42%

Scheme 3-15 Synthesis of the Ddpe protecting group 108

Ester **154** was then Boc-deprotected and the resulting amine salt was coupled with isothiocyanate **164** under basic conditions to yield thiourea **167** in good yield (Scheme 3-16). Conversion into guanidine **169** was accomplished by activating **167** as the thiouronium hexafluorophosphate **168** and reaction with trifluoroacetamide and DBU. This occurred in disappointing yield, possibly due to partial deprotection of the Ddpe group under the strongly basic conditions. Final hydrogenolysis revealed acid **170**.

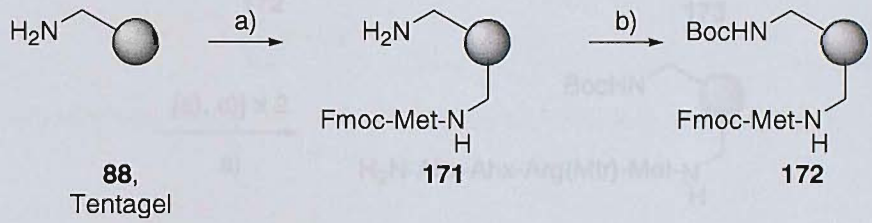


Reagents and conditions: a) 20% TFA/DCM, 1 hour; b) **164**, Et_3N, DCM, 18 hours, 84% over two steps; c) i) MeI, acetone, 3 hour, ii) NH_4PF_6, MeOH/DCM, 18 hours; d) F_3CCONH_2, DBU, toluene/$CHCl_3$, reflux, 2 hours, 32%; e) H_2, Pd/C, EtOAc, 2 hours, 98%

Scheme 3-16 Synthesis of Ddpe-protected CBS 170

3.4.2 Library Synthesis

With an appropriately protected CBS precursor unit **151** in hand, synthesis of a full tweezer library was undertaken. Firstly, high-loaded Tentagel resin (0.47mmol/g) was partially coupled (~90%) with Fmoc-Met by first shaking the resin with a solution of Fmoc-Met-OH in DMF, followed by addition of excess DIC/HOBt/DIPEA. The pre-mix is to ensure an even distribution of the amino acid throughout the resin before coupling begins, ensuring partial derivatisation at the single-bead level and not at the level of bulk resin. A quantitative Ninhydrin test[REF] gave NH_2 loading as 0.10mmol/g. The remaining free sites were reacted with excess Boc_2O to generate bifunctional resin **172** (Scheme 3-17).

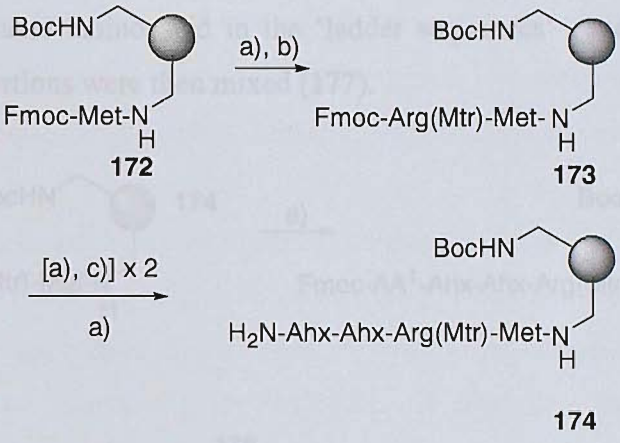


88, Tentagel **171** **172**

Reagents and conditions: a) i) Fmoc-Met-OH, DMF, 1 hour, ii) addition of DIC, HOBt, DIPEA, 18 hours; b) Boc_2O, DIPEA, DMF, 3 hours

Scheme 3-17 Partial derivatisation and protection of solid-phase in preparation for library synthesis

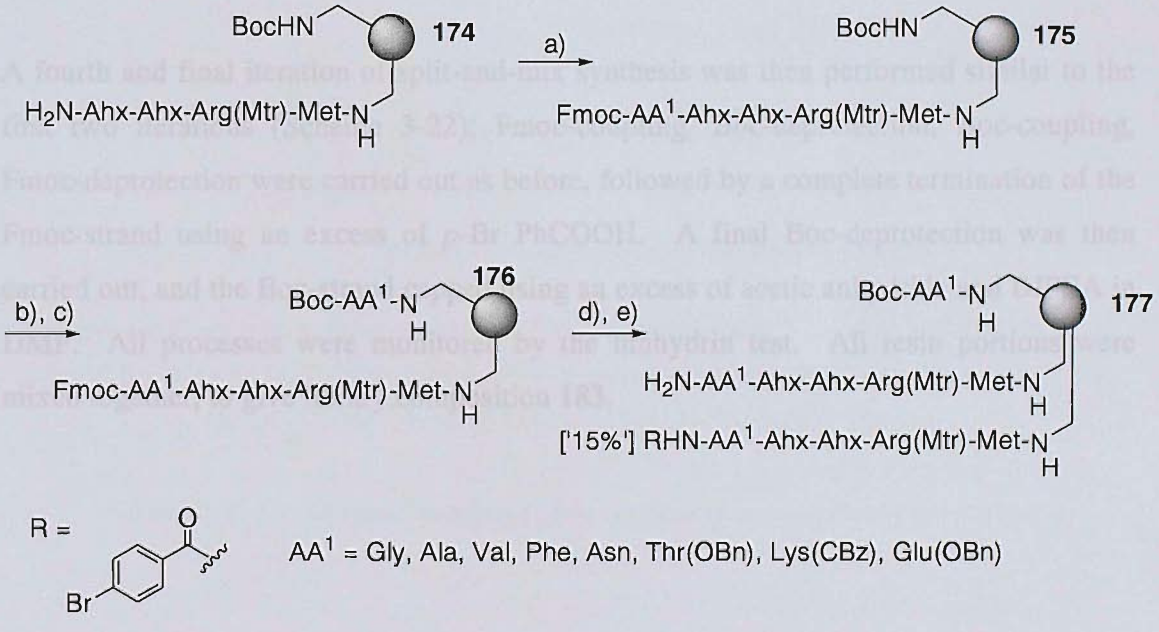
It was now necessary to incorporate the appropriate linker for generation of 'ladder' sequences. To this end, the resin was Fmoc-deprotected and standard Fmoc-SPPS was carried out to incorporate first Arg(Mtr), then two Ahx residues, giving **174** (Scheme 3-18). In this case Arg(Mtr) (cleavage in 1M TMSBr in neat TFA) is selected because it is necessary to perform several Boc-deprotections on the resin to generate the library; Mtr is much more acid-stable than Pbf. All coupling steps were monitored for completion using the Kaiser ninhydrin test. A quantitative ninhydrin test gave the NH_2-loading of **174** as 0.13mmol/g.



Reagents and conditions: a) 20% piperidine/DMF, 20 mins; b) Fmoc-Arg(Mtr)-OH, DIC, HOBt, DIPEA, DMF, 3 hours; c) Fmoc-Ahx-OH, DIC, HOBt, DIPEA, DMF, 3 hours

Scheme 3-18 Synthesis of linker for 'ladder sequence' generation in the Fmoc-strand

Split-and-mix methodology was then implemented as illustrated in Schemes 3-19 to 3-22. Firstly the bulk resin was split into eight portions, and each portion was then independently coupled to an Fmoc-amino acid (from Gly, Ala, Val, Phe, Asn, Glu(OBn), Lys(Cbz), Thr(OBn); benzyl-based protection was used owing to the need to perform many acidic Boc-deprotections), giving **175** (Scheme 3-19). Each separate portion was then Boc-deprotected using dilute TFA, followed by coupling with the same amino acid as before but with a Boc rather than Fmoc protecting group (**176**). Each sample was then Fmoc-deprotected, followed by implementation of the capping protocol (15% of nominal loading of the Fmoc-strand). In this way, the amino acid incorporated in the Boc-strand is indexed by the same amino acid in the 'ladder sequences' generated in the Fmoc-strand. The resin portions were then mixed (**177**).



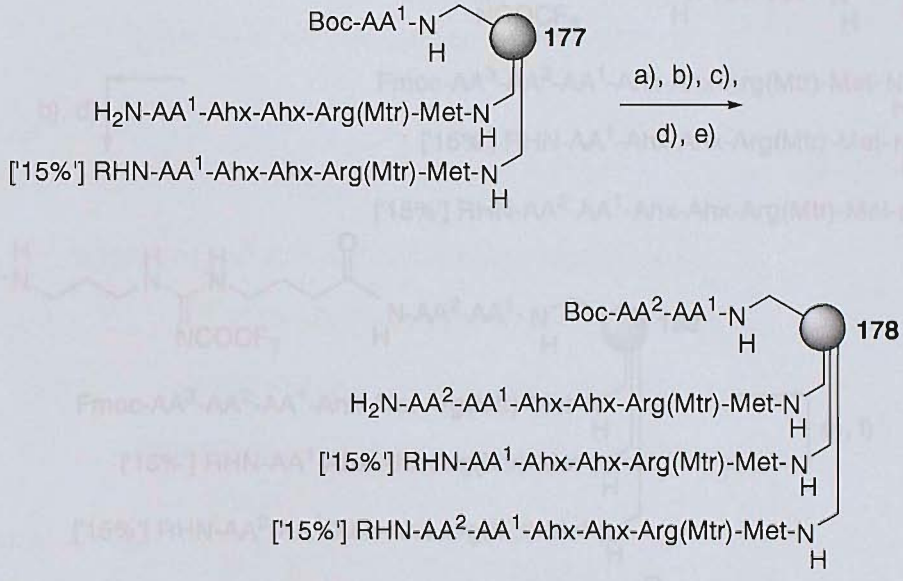
Reagents and conditions: a) Fmoc-AA-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% TFA/DCM, 20 mins; c) Boc-AA-OH, DIC, HOBt, DIPEA, DMF, 3 hours; d) 20% piperidine/DMF, 20 mins; e) i) *p*-Br PhCOOH, DMF, 1 hour, ii) addition of DIC, HOBt, DIPEA, 18 hours

Scheme 3-19 First round of split and mix synthesis

It was now desired to incorporate the CBS into the Boc-strand. This was accomplished as detailed below and shown in Scheme 3-21. The resin was split into eight portions and coupled with an Fmoc-amino acid (same diversity of amino acids as before). Boc-deprotection was then carried out using dilute TFA. However, the next step involved coupling each resin portion individually with the CBS **151** (using a standard DIC/HOBt/DIPEA in DMF, double-coupling procedure, again monitored by ninhydrin test), giving **179**. The resin portions were then again Boc-deprotected and coupled with the appropriate Boc-amino acid to match the one attached in the previous Fmoc-coupling. Fmoc-deprotection was then carried out, followed by implementation of the capping protocol (20% of original nominal loading of Fmoc-strand), and mixing of the resin portions. In this manner the CBS was incorporated into only one strand, with an appropriate ladder sequence generated. Library composition is as indicated by **181**.

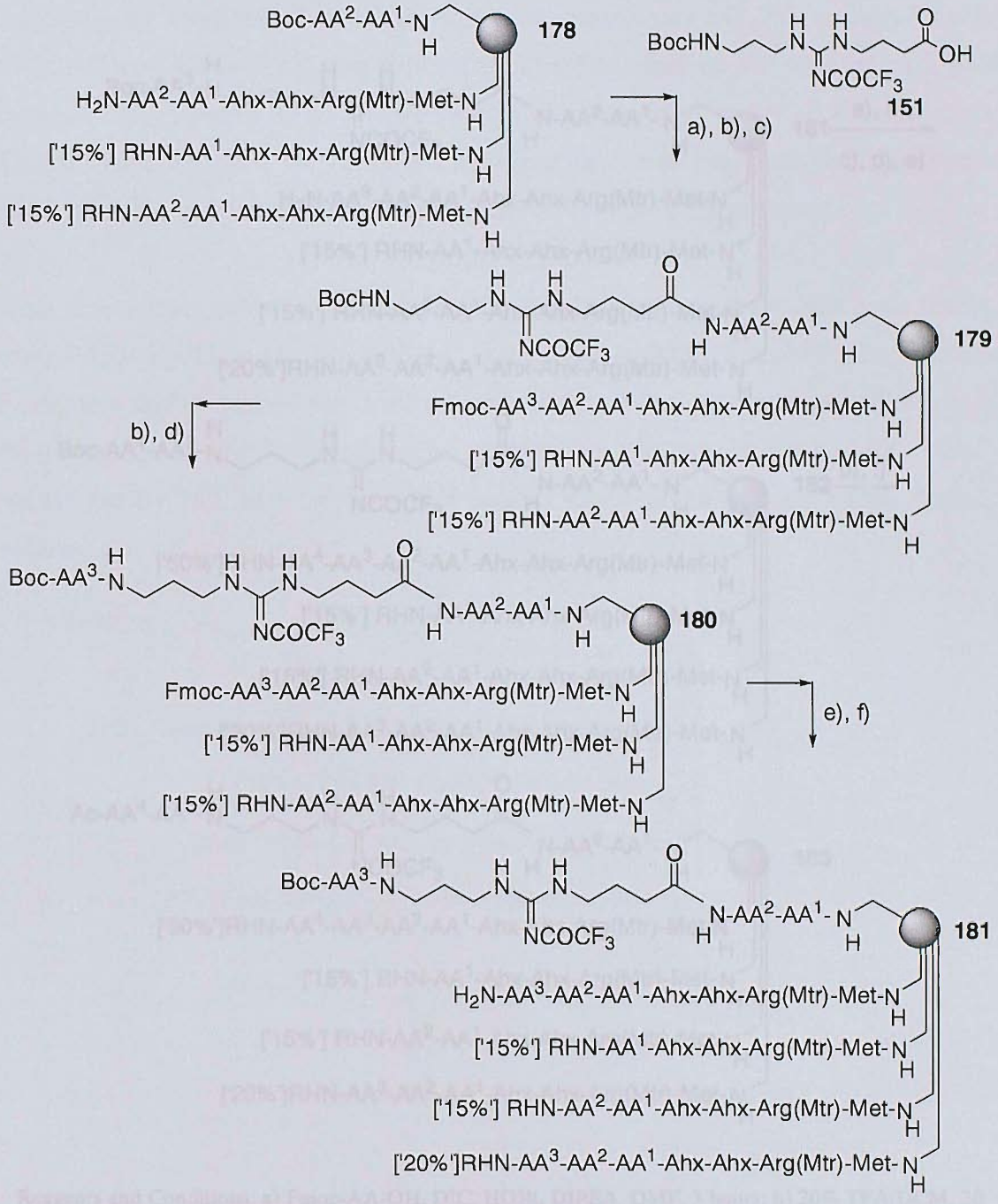
A fourth and final iteration of split-and-mix synthesis was then performed similar to the first two iterations (Scheme 3-22); Fmoc-coupling, Boc-deprotection, Boc-coupling, Fmoc-deprotection were carried out as before, followed by a complete termination of the Fmoc-strand using an excess of *p*-Br PhCOOH. A final Boc-deprotection was then carried out, and the Boc-strand capped using an excess of acetic anhydride and DIPEA in DMF. All processes were monitored by the ninhydrin test. All resin portions were mixed together, to give library composition **183**.

A second round of split-and-mix synthesis was then performed in identical fashion (Scheme 3-20); division into eight portions, Fmoc-protected coupling, Boc-deprotection, Boc-protected coupling, Fmoc-deprotection, capping protocol (15% of original nominal loading of Fmoc-strand), mixing of resin portions. The same diversity of amino acids as in the first iteration was used. Again, all steps was monitored using the Ninhydrin test. The library composition is now indicated by **178**.



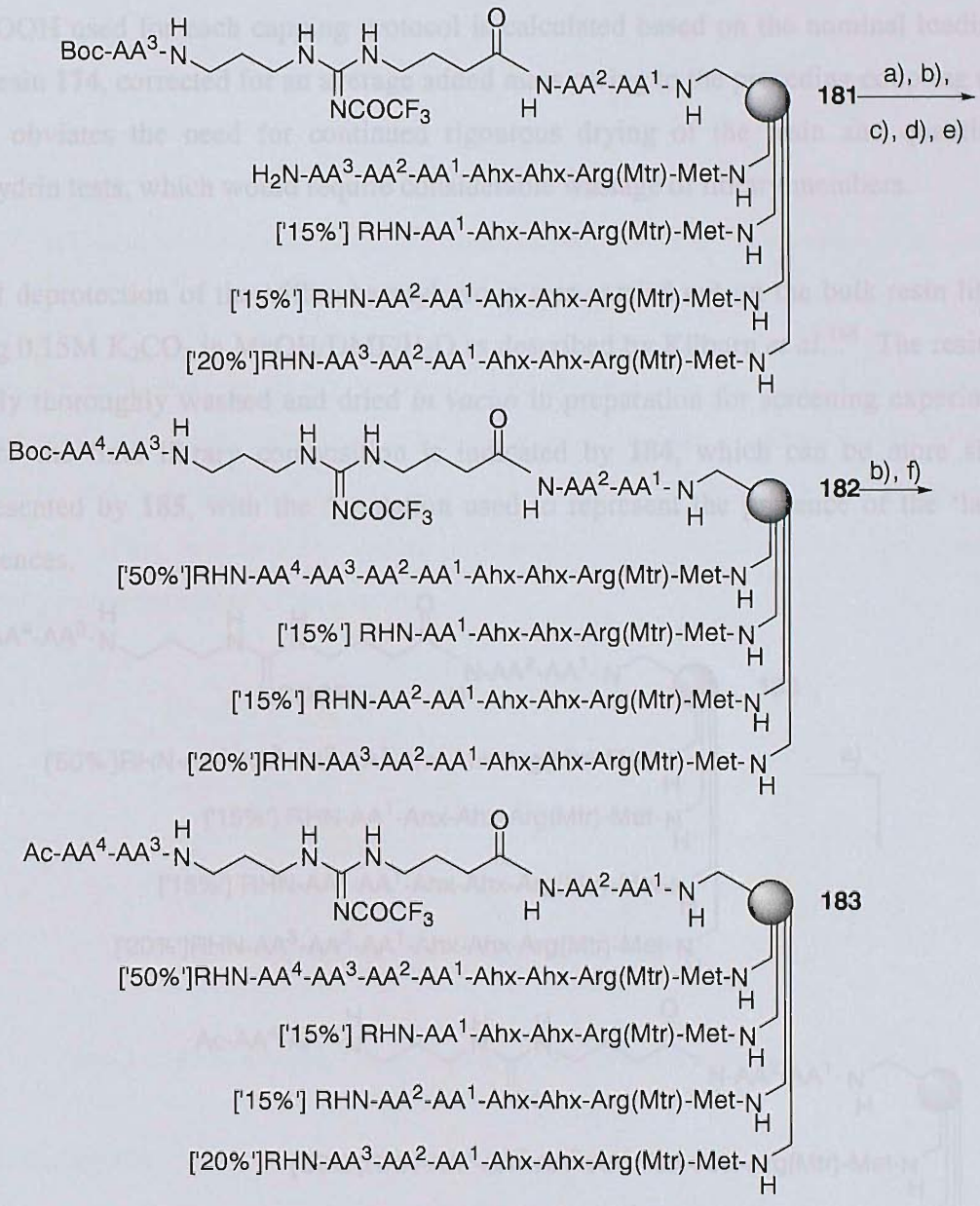
Reagents and Conditions: a) Fmoc-AA-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% TFA/DCM, 20 mins; c) Boc-AA-OH, DIC, HOBt, DIPEA, DMF, 3 hours; d) 20% piperidine/DMF, 20 mins; e) i) *p*-Br PhCOOH, DMF, 1 hour, ii) addition of DIC, HOBt, DIPEA, 18 hours

Scheme 3-20 Second round of split-and-mix synthesis in tweezer library generation



Reagents and conditions: a) Fmoc-AA-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% TFA/DCM, 20 mins; c) **151**, DIC, HOBt, DIPEA, DMF, 3 hours; d) Boc-AA-OH, DIC, HOBt, DIPEA, DMF, 3 hours; e) 20 % piperidine/DMF, 20 mins; f) i) *p*-Br PhCOOH, DMF, 1 hour, ii) addition of DIC, HOBt, DIPEA, 18 hours

Scheme 3-21 Third iteration of split-and-mix synthesis including incorporation of CBS

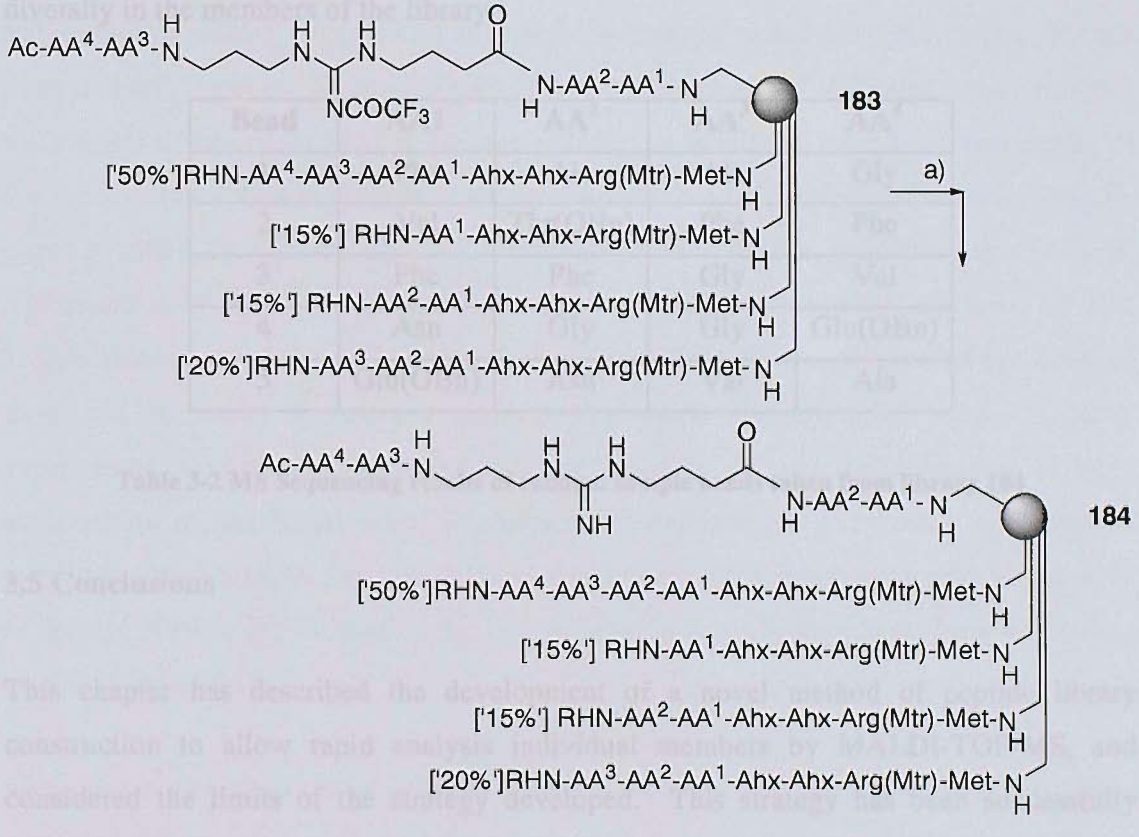


Reagents and Conditions: a) Fmoc-AA-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% TFA/DCM, 20 mins; c) Boc-AA-OH, , HOBt, DIPEA, DMF, 3 hours; d) 20% piperidine/DMF, 20 mins; e) *p*-Br PhCOOH, DIC, HOBt, DIPEA, DMF, 18 hours; f) Ac_2O, DIPEA, DMF, 3 hours

Scheme 3-22 Final round of split-and-mix synthesis in tweezer library generation

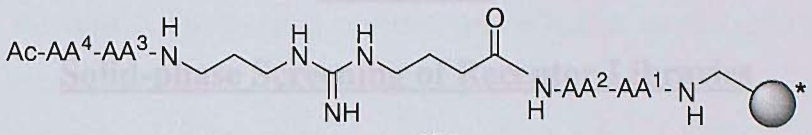
As during the synthesis of the single-armed receptor library **149**, the quantity of *p*-Br PhCOOH used for each capping protocol is calculated based on the nominal loading of the resin **174**, corrected for an average added mass owing to the preceding coupling steps. This obviates the need for continued rigourous drying of the resin and quantitative ninhydrin tests, which would require considerable wastage of library members.

Final deprotection of the trifluoroacetyl group was carried out on the bulk resin library, using 0.15M K_2CO_3 in MeOH/DMF/H_2O as described by Kilburn *et al.*[168] The resin was finally thoroughly washed and dried *in vacuo* in preparation for screening experiments. Hence the final library composition is indicated by **184**, which can be more simply represented by **185**, with the * notation used to represent the presence of the 'ladder' sequences.



Reagents and Conditions: a) 0.15M K_2CO_3, MeOH/DMF/H_2O

Scheme 3-23 Deprotection of binding site in tweezer library



185

Fig 3-13 185, an abbreviated notation for the tweezer library 184. The * mark indicates the presence of 'ladder' sequences, which are not shown.

To test the success of the library synthesis, 5 beads were selected at random, individually deprotected using 1M TMSBr in TFA, cleaved using the standard CNBr/TFA/H_2O treatment and cleavage products analysed by MALDI-TOF-MS. Results are shown in Table 3-3. This indicates successful implementation of the synthesis and a suitable diversity in the members of the library.

Bead	AA1	AA^2	AA^3	AA^4
1	Phe	Ala	Ala	Gly
2	Val	Thr(OBn)	Phe	Phe
3	Phe	Phe	Gly	Val
4	Asn	Gly	Gly	Glu(OBn)
5	Glu(OBn)	Asn	Val	Ala

Table 3-2 MS Sequencing results of random sample beads taken from library 184

3.5 Conclusions

This chapter has described the development of a novel method of peptide library construction to allow rapid analysis individual members by MALDI-TOF-MS, and considered the limits of the strategy developed. This strategy has been successfully implemented in the synthesis of two diverse libraries of carboxylate receptors with peptidic side-arms, which are to be used in solid-phase screening experiments to determine effective receptors for peptide carboxylate guests of biological significance.

Chapter 4

Solid-phase Screening of Receptor Libraries

4.1 Introduction

The screening of combinatorial libraries of ligands to determine drug candidates is a well established technique of medicinal chemistry. Likewise, strongly binding receptors of various species have been determined using libraries (see Chapter 1). It is a rapid method for probing large regions of chemical space for a variety of physical and biochemical properties.

Solid-phase screening experiments are simple in conception and implementation. For the current work, a sample of resin-bound library of receptors (of sufficient mass such that the number of beads sampled represents a statistical likelihood of at least five copies of the library) is swollen in a solvent mixture, to which a solution of dye-labelled peptide guest is added in aliquots, allowing sufficient time for equilibration between additions. Visualisation of the screen is performed under magnification with a hand lens. By eye, simple identification of heavily stained beads is possible. Further aliquots of dye-labelled guest can be added to improve distinction amongst the individual beads. Control experiments confirm that selective binding noted is not due to difference in resin composition of individual beads (by incubating dye-labelled guest with underivatised resin and noting no selective uptake), also that selective binding is not due to recognition of the dye portion (by incubating the library with the dye chromophore alone and noting no selective uptake).

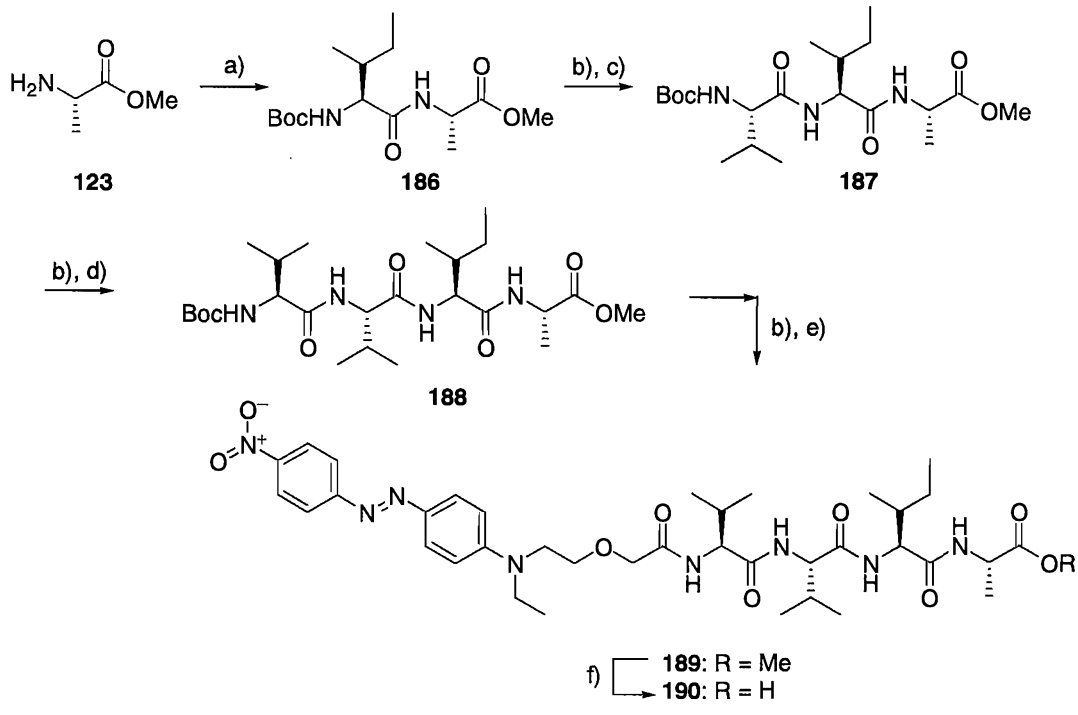
This chapter describes the identification of receptors for two peptide-carboxylate guests of biological relevance, in aqueous medium. The guests represent the *C*-termini of two proteins; an amyloid β-protein and the viral enzyme HIV-1 protease. Receptors for guests in polar media are less well-developed than those for use in organic media. Use was made of the two libraries of receptors **149** and **184** whose synthesis was described in

Chapter 3. After identification of effective receptors from solid-phase screening experiments, resynthesis and binding constant determination on the solid-phase and in solution was used to confirm the binding capability of the receptors.

4.2 Guest synthesis

Solid-phase screening requires a guest molecule with an appropriate 'output' to allow determination of which beads bind the guest most strongly. This is achieved in the current work by using a visible chromophore, allowing identification of strongly binding beads by recognition of most heavily stained beads. Such guests can also be used for UV/Vis binding studies. Many potential dye-units exist, however an aqueous soluble dye without carboxylate functionality is needed. As before (Chapter 2), it was decided to use the commercially available Disperse Red 1 for guest synthesis, an intense red dye **115**, appropriately derivatised as acid **117** (See Chapter 2).

Tetrapeptide guest **190** (Val-Val-Ile-Ala) was prepared by sequential coupling of amino acids under standard conditions to generate **188**. **188** was then coupled to dye-acid **117** using EDC and HOBt to give dye-labelled tetrapeptide ester **189**, which was then hydrolysed with mild aqueous base to give **190**. (Scheme 4-1).

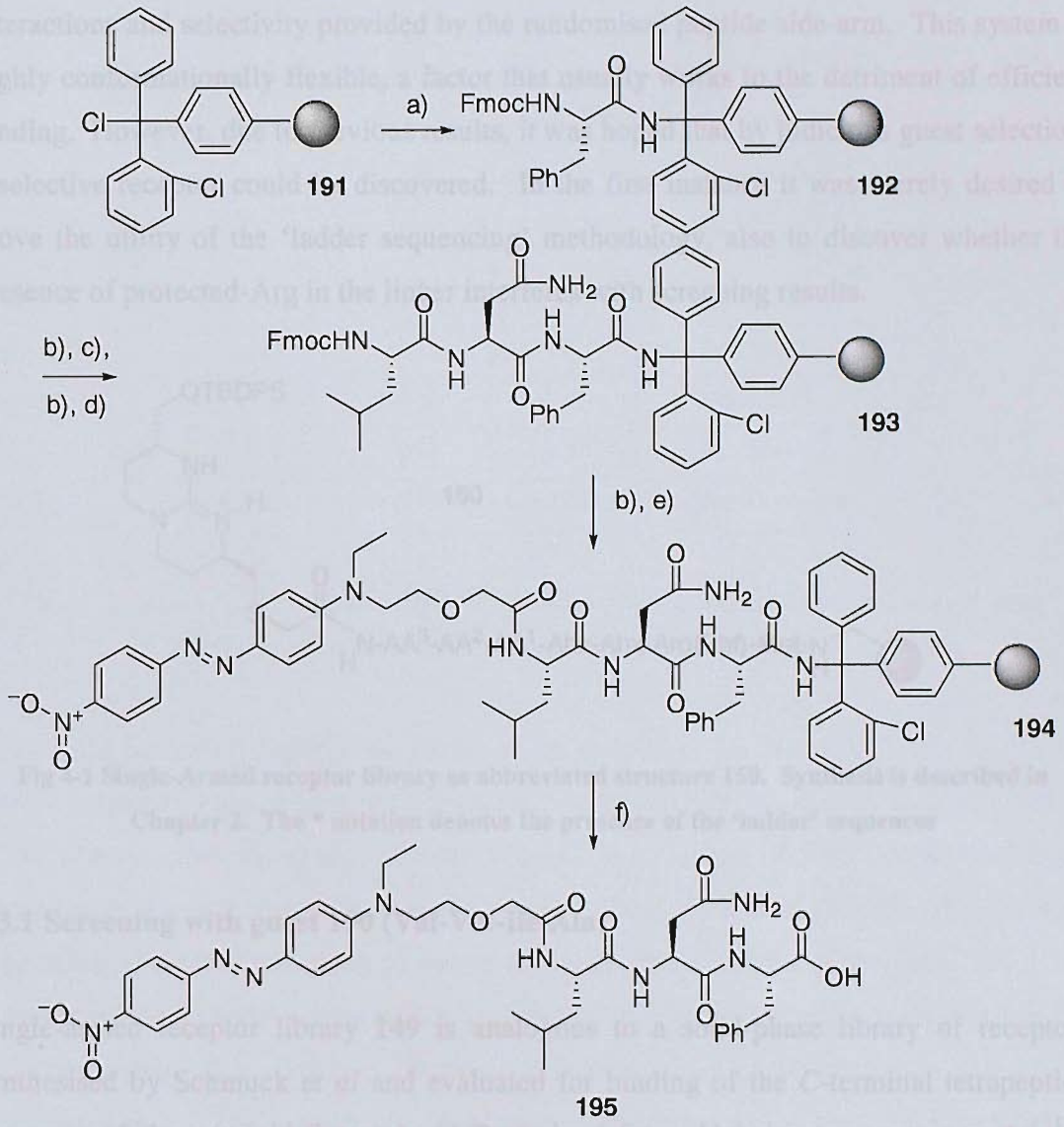


123 186 187 188

Reagents and conditions: a) Boc-Ile-OH, DIC, HOBt, DIPEA, DMAP, THF/DCM, 3.5 hours, 82%; b) 20% TFA/DCM; c) Boc-Val-OH, CDI, Et_3N, THF, 18 hours, 59%; d) Boc-Val-OH, EDC, HOBt, DIPEA, DMAP, DMF, 18 hours, 88%; e) **117**, PyBOP, HOBt, DIPEA, DMAP, THF/DCM, 18 hours, 55%; e) LiOH(aq), dioxane, 18 hours, 55%

Scheme 4-1 Synthesis of dye-labelled guest 190

Tripeptide guest **195** (Leu-Asn-Phe) was prepared by Dr. T. Gale using a solid-phase protocol on 2-chlorotrityl resin (Scheme 4-2). Fmoc-Phe was first loaded to the resin (**192**), followed by Asn and Leu in sequence using standard Fmoc-SPPS conditions, giving **193**. After deprotection, dye-acid **117** was coupled to the resin, and the final product **195** was obtained by cleavage using dilute TFA.



Reagents and conditions: a) Fmoc-Phe-OH, DIPEA, DCM, 18 hours; b) 20% piperidine/DMF, 20 mins; c) Fmoc-Asn-OH, DIC, HOBt, DIPEA, DMF, 3 hours; d) Fmoc-Leu-OH, DIC, HOBt, DIPEA, DMF, 3 hours; e) **117**, DIC, HOBt, DIPEA, DMF, 3 hours; f) 2% TFA/DCM, 2 hours

Scheme 4-2 Solid-phase synthesis of dye-labelled guest 195 (performed by Dr. T. Gale)

4.3 Single-Armed Receptor Library

Library **149** is a simplistic receptor system. The expected major interaction between library members and guest is the guanidinium-carboxylate ion pair, with secondary

interactions and selectivity provided by the randomised peptide side-arm. This system is highly conformationally flexible, a factor that usually works to the detriment of efficient binding. However, due to previous results, it was hoped that by judicious guest selection, a selective receptor could be discovered. In the first instance it was merely desired to prove the utility of the 'ladder sequencing' methodology, also to discover whether the presence of protected-Arg in the linker interferes with screening results.

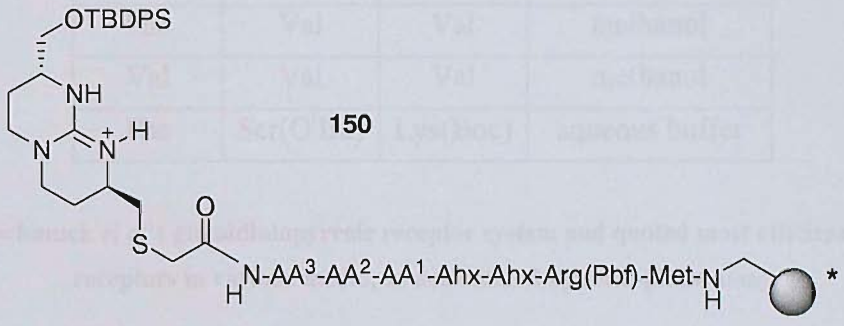
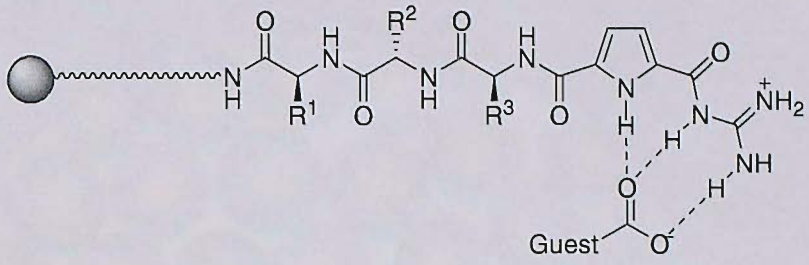


Fig 4-1 Single-Armed receptor library as abbreviated structure 150. Synthesis is described in Chapter 2. The * notation denotes the presence of the 'ladder' sequences

4.3.1 Screening with guest 190 (Val-Val-Ile-Ala)

Single-armed receptor library **149** is analogous to a solid-phase library of receptors synthesised by Schmuck *et al* and evaluated for binding of the *C*-terminal tetrapeptide sequence of the amyloid-β protein (Aβ), derived from Alzheimer precursor protein[136], which is known to form protein plaques in the brains of Alzheimer's patients. It is the *C*-terminal Val-Val-Ile-Ala sequence that is thought to promote precipitation of Aβ. Certain guanidiniopyrrole-linked tripeptides were found to bind strongly to the sequence Val-Val-Ile-Ala, in particular the structures indicated in Fig 4-2. Hence this guest was selected to test the viability of screening receptor libraries indexed by 'ladder' sequences, since results should be comparable to those of Schmuck *et al*, whose binding unit is of similar nature and size to the de Mendoza CBS **145**.



68

R^1	R^2	R^3	Medium
Phe	Val	Val	methanol
Val	Val	Val	methanol
Phe	Ser(O^{t}Bu)	Lys(Boc)	aqueous buffer

Fig 4-2 Schmuck *et al*'s guanidiniopyrrole receptor system and quoted most efficient binding receptors in various media, as determined by solid-phase assay

In a typical screening experiment, ~8mg of the resin library **149** (of sufficient mass such that the number of beads sampled represents a statistical likelihood of *at least* five copies of the library, often more) was allowed to swell in a mixture of 50% DMSO/buffer and titrated with aliquots of dye-labelled tetrapeptide **190** in 50% DMSO/buffer. Such a high proportion of DMSO was used to ensure complete solubilisation of the guest **190**. The presence of the buffering agent which ensures that the guest is present in the form of the free carboxylate, also exchanges the nominal counter-anion of the guanidine unit for the buffering anion. The system was allowed to equilibrate for at least 24 hours before inspection under a low-magnification hand lens. Concentration of peptide guest could be adjusted by dilution or titration of further guest to give an 'optimum' selectivity amongst the beads as judged by eye. An example system is shown in Fig 4-3

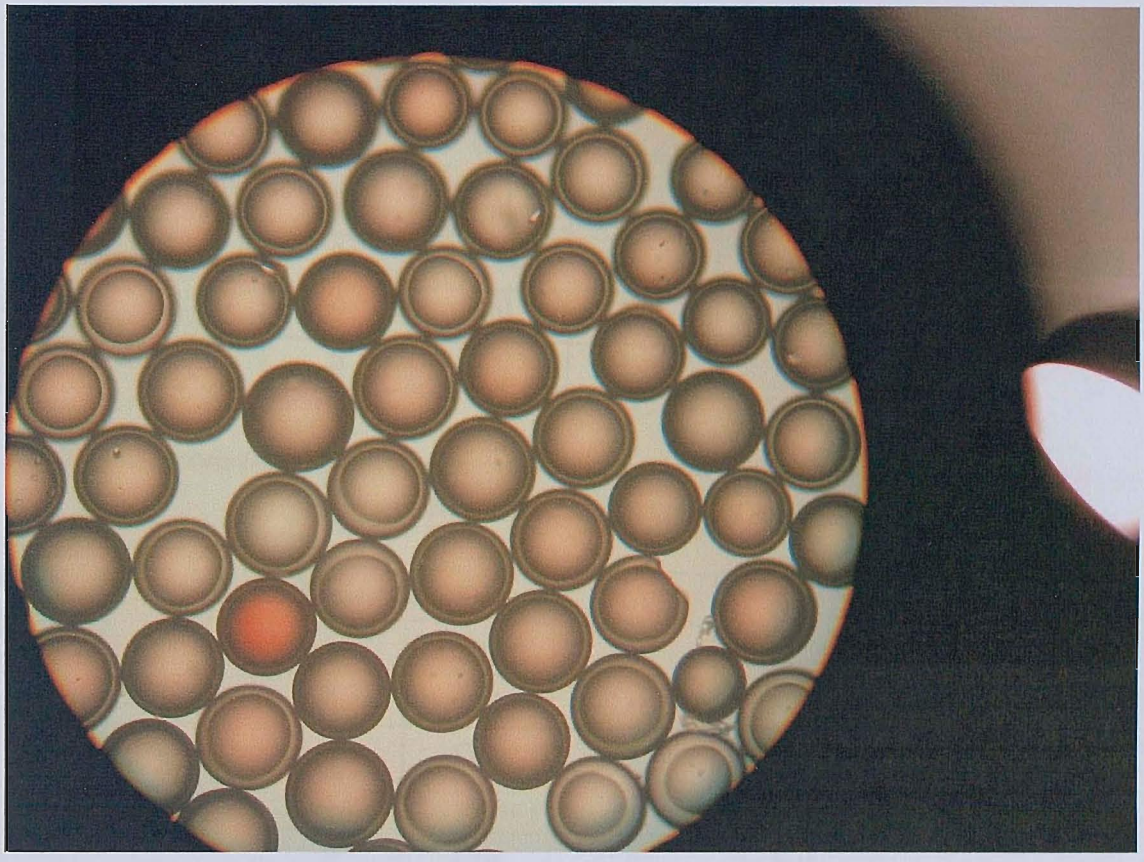


Fig 4-3 Sample screening experiment shown under magnification. The presence of a single heavily stained bead is obvious. Slight colouration is noted on other beads

The screening was carried out using two different buffers; borax (a borate-based buffer) and HEPES (an organic sulphonate-based buffer). Individual, heavily stained beads were removed from the screen medium by micropipette, Pbf-deprotected using 50% TFA/DCM, cleaved using standard CNBr/TFA/H_2O treatment and analysed by MALDI-TOF-MS to determine the structure of the receptors on the beads. Results from different screens are indicated in the following Tables (4-1 to 4-3).

Bead	AA^1	AA^2	AA^3
1	Val	Val	Phe
2	Val	Val	Glu(O^{t}Bu)
3	Ser(O^{t}Bu)	Ala	Gly
4	Val	Val	Leu
5	Val	Leu	Leu

Table 4-1 Sequences determined from a screening experiment of library 149 against guest 190 at a concentration of 29.5μM in 50% DMSO/borax

Bead	AA^1	AA^2	AA^3
1	Phe	Val	Glu(O^{t}Bu)
2	Val	Leu	Leu
3	Gly	Asn	Ala
4	Lys(Boc)	Val	Val
5	Leu	Leu	Val
6	Asn	Val	Leu
7	Val	Val	Leu
8	Val	Glu(O^{t}Bu)	Val
9	Phe	Val	Val
10	Val	Leu	Leu
11	Val	Val	Leu
12	Val	Leu	Ala

Table 4-2 Sequences determined from a screening experiment of library 149 against guest 190 at a concentration of 42.4μM in 50% DMSO/borax

Bead	AA^1	AA^2	AA^3
1	Val	Leu	Val
2	Val	Val	Val
3	Leu	Leu	Glu(O^tBu)
4	Val	Leu	Glu(O^tBu)
5	Val	Leu	Val
6	Val	Leu	Glu(O^tBu)
7	Leu	Val	Val
8	Leu	Lys(Boc)	Lys(Boc)
9	Val	Phe	Val
10	Val	Val	Leu
11	Val	Val	Glu(O^tBu)
12	Glu(O^tBu)	Val	Leu
13	Val	Leu	Leu
14	Val	Val	Phe
15	Val	Phe	Val
16	Glu(O^tBu)	Val	Val

Table 4-3 Sequences determined from a screening experiment of library 149 against Val-Val-Ile-Ala guest 190 at a concentration of 36.3μM in 50% DMSO/HEPES

Certain observations can immediately be made. Firstly, there is a high proportion of Val-residues at all positions (48% of all residues), including a very great preference for AA^1 = Val. The second most-abundant residue in the screens is Leu (25% of all residues). Leu and Val have highly similar side-chains of lipophilic nature – Leu could be considered as '*homo*-Val' – hence it is perhaps unsurprising to find both occurring frequently together. It is also to be noted that the receptor structures Val/Val/Leu and Val/Leu/Leu ($AA^1/AA^2/AA^3$) both occur more than once across the screens *and* within an individual screen, suggesting that either of these two structures is a highly promising receptor for guest **190**. Phe (6%) and Glu(O^tBu) (9%) occur occasionally throughout the screens, including a slight preference for AA^3 = Glu(O^tBu) in the DMSO/HEPES screen. There is an almost total lack of the remaining possible residues [Gly, Ala, Asn, Ser(O^tBu),

Lys(Boc)], with none occurring more than three times across all the screens (all five making 11% of all noted residues).

That Val and Leu are highly favoured residues is perhaps unsurprising. The guest **190** is highly hydrophobic in nature, with bulky non-polar side-chains (Val, Ile). The frequent occurrence of the lipophilic Val and Leu in the library members selected by the screening likely represents the hydrophobic nature of the guest, which prefers to be in the more non-polar environment provided by a receptor with lipophilic side-chains, rather than polar bulk solvent. This result is also comparable to Schmuck *et al*'s results[136], whereby one of the strongest-binding receptors in MeOH is the Val-Val-Val receptor (indicated in Fig 4-2).

Schmuck *et al*'s results also indicate a strong preference, in a totally aqueous medium, for the presence of Lys(Boc) close to the guanidiniopyrrole binding site. This is rationalised on the basis of molecular dynamics calculations, which indicate that the guest and receptor associate into a β-sheet structure, with the *tert*-butyl group of Lys(Boc) positioned between the first Val-group of the guest and the Ile-group. This will minimise hydrophobic contact with the surrounding medium, a favoured situation which would explain the preference for this residue (see Fig 1-49, Chapter 1).

In a test of this hypothesis, screening of library **149** and dye-labelled guest **190** was repeated in 20% DMSO/buffer, a medium with a greater proportion of aqueous buffer. The guest **190** remained solubilised throughout the course of the screening experiment. Heavily stained beads were again removed, Arg-deprotected, cleaved and analysed by MALDI-TOF-MS under the standard conditions. Results are shown in Table 4-4 (Sequencing was not possible for one of the selected beads).

Bead	AA^1	AA^2	AA^3
1	Leu	Leu	Val
2	Val	Leu	Leu
3	Val	Lys(Boc)	Leu
4	Leu	Val	Leu
5	Leu	Lys(Boc)	Val
6	Val	Val	Ser(O^{t}Bu)
7	Glu(O^{t}Bu)	Val	Phe
8	x	x	x
9	Phe	Val	Leu
10	Val	Lys(Boc)	Leu

Table 4-4 Sequences determined from a screening experiment of library 149 against guest 190 at a concentration of 36.3μM in 20% DMSO/borax

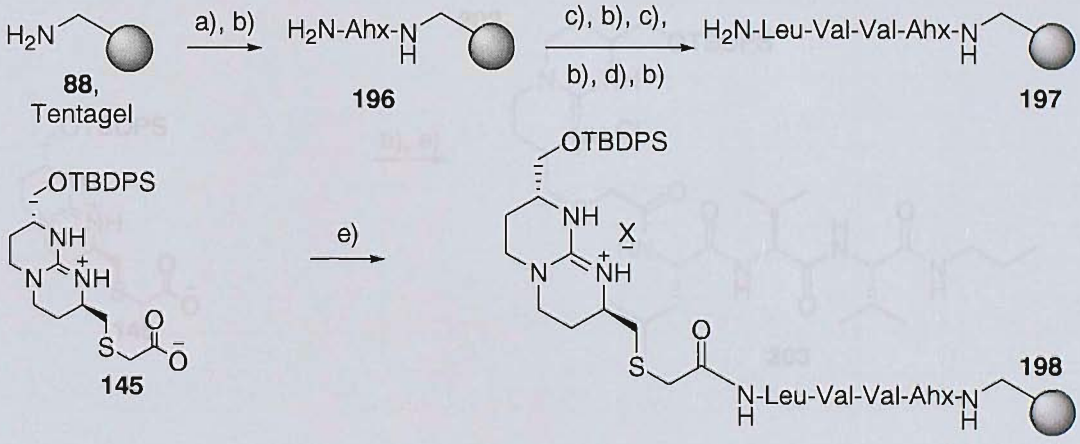
Again a high proportion of Val and Leu residues are noted (33% Val, 33% Leu of all residues), and again a distinct lack of Gly, Ala and Asn (0%, all residues). Phe (3%), Glu(O^{t}Bu) (3%) and Ser(O^{t}Bu) (3%) all occur a limited number of times. Interestingly, there appears to be a partial preference for AA^2 = Lys(Boc). Lys(Boc) appears *only* at this position within the screen, and in 33% of all the residues at AA^2, whereas previously it had appeared with lesser frequency and in all three positions. This tendency towards Lys(Boc) in position AA^2 perhaps demonstrates a similar situation to that of the Schmuck system, accounting for the different sizes and conformational demands of the two receptors. It is encouraging to find that the current results reflect similar conclusions about known receptors for the Val-Val-Ile-Ala system.

Since the combination [AA^1 = Val, AA^2 = Val, AA^3 = Leu] occurred most frequently in all the screens, and reflects the general trend of the results, this particular structure was selected for further investigation.

4.3.2 Resynthesis of selected receptor structure

In order to evaluate the binding of the most promising receptor selected by the screening experiments, resynthesis of this structure in quantity was required on both solid phase and as an isolated entity (*without* the presence of the Arg-containing linker). Methods are described below.

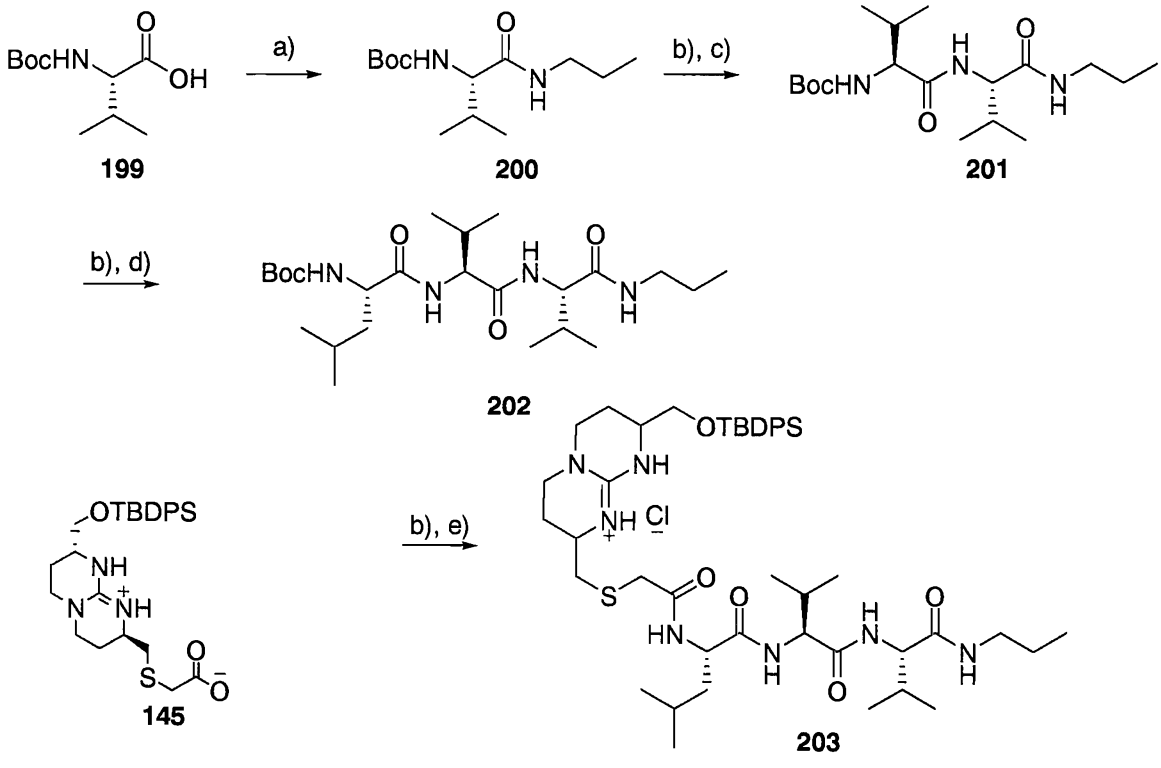
Solid-phase bound receptor **198** was synthesised according to Scheme 4-3, using standard Fmoc-SPPS procedures. Tentagel resin was first derivatised with an Ahx residue and then the sequence Leu-Val-Val by sequential coupling-deprotection steps, giving **197**. All steps were monitored using the Kaiser ninhydrin test. Finally, **145** was coupled using DIC/HOBt to give receptor-derivatised resin **198**.



Reagents and Conditions: a) Fmoc-Ahx-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% piperidine/DMF, 20 mins; c) Fmoc-Val-OH, DIC, HOBt, DIPEA, DMF, 3 hours; d) Fmoc-Leu-OH, DIC, HOBt, DIPEA, DMF, 3 hours; e) **145**, DIC, HOBt, DIPEA, DMF, 18 hours

Scheme 4-3 Solid-phase synthesis of receptor 198

Receptor **203** was synthesised in solution according to Scheme 4-4. Boc-Val **199** was converted into the propyl amide **200** using CDI and propylamine. After Boc-deprotection, a further Boc-Val unit was coupled using CDI and Et_3N, giving dipeptide **201**. Further deprotection and coupling gave tripeptide **202**. **202** was then deprotected and coupled to the CBS **145** using EDC and HOBt., giving the receptor **203**.



Reagents and conditions: a) CDI, propylamine, THF, 3 hours, 92%; b) 20% TFA/DCM, 1 hour; c) Boc-Val-OH, CDI, Et_3N, THF/CMF, 18 hours, 80%; d) Boc-Leu-OH.H_2O, CDI, Et_3N, THF, 18 hours, 70%; e) **145**, EDC, HOBt, DIPEA, DMAP, DMF, 1 hour, 9%

Scheme 4-4 Synthesis of receptor 203

4.3.3 Evaluation of Binding

Binding was investigated for both the solid-phase linked receptor **198** and the isolated receptor **203** with guest **190**.

In an identical manner to that previously detailed (Chapter 2), a UV uptake assay was performed. Calibration curves were made by measuring the UV absorption of the dye-chromophore of guest **190** at varying concentration in 50% DMSO/borax. Binding experiments were then carried out as detailed above (Chapter 2) using receptor-derivatised resin **198**, Tentagel and **190**. Binding constants were estimated using the calculation method described, assuming a 20% error in resin loading. Results are displayed in Table 4-5

	1	2	3	4
m_{rec}/mg	6.8	8.5	9.3	9.4
m_{TG}/mg	5.9	7.3	8.0	8.0
c_{init}/M	1.85×10^{-4}	1.85×10^{-4}	1.32×10^{-4}	1.32×10^{-4}
V/ml	1.2	1.2	1.2	1.2
c_{rec}/M	4.23×10^{-5}	4.73×10^{-5}	3.95×10^{-5}	3.81×10^{-5}
c_{TG}/M	1.34×10^{-4}	1.32×10^{-4}	1.05×10^{-4}	1.08×10^{-4}
$[HG]$/M	9.20×10^{-5}	8.44×10^{-5}	6.51×10^{-5}	7.04×10^{-5}
$[G]$/M	4.23×10^{-5}	4.73×10^{-5}	3.95×10^{-5}	3.81×10^{-5}
$[H]$/M	1.04×10^{-3}	1.33×10^{-3}	1.48×10^{-3}	1.50×10^{-3}
K_a/M^{-1}	$2.09 \times 10^{3} \pm 580$	$1.34 \times 10^{3} \pm 360$	$1.11 \times 10^{3} \pm 290$	$1.23 \times 10^{3} \pm 330$

Table 4-5 Results obtained in solid-phase binding assay of resin-linked receptor 198 and dye-labelled guest 190, assuming a nominal loading of 0.2mmol/g. Mean $K_a = 1.44 \times 10^3 \pm 440$

Though merely an estimate of binding affinity, these calculated values are reassuring as they are of the same order of magnitude as those obtained by Schmuck *et al* for their guanidiniopyrrole receptors (Fig 4-2). The absolute value is high for a binding interaction in aqueous medium, and reflects the strength of the guanidinium-carboxylate ion-pair. The errors in the calculated values show the sensitivity of the measurements to the loading of the resin.

Although a positive binding interaction has been observed and quantified on the solid-phase, demonstration of binding in the solution phase is required, since the usefulness of such receptors (e.g. in biochemical investigations) depends on ability to function in solution. It is possible that observed binding is an artefact of the receptor in tandem with the resin-matrix, perhaps owing to some conformational restriction imposed by the resin or a pre-association of guest with resin before true binding with the receptor can occur.

UV-titration experiments were carried out using receptor **203** and dye-labelled guest **190**. UV absorption of the dye chromophore of **190** was monitored during additions of receptor **203** at constant concentration of **190**, in 50% DMSO/borax solution. A typical binding isotherm is plotted in Fig 4-4.

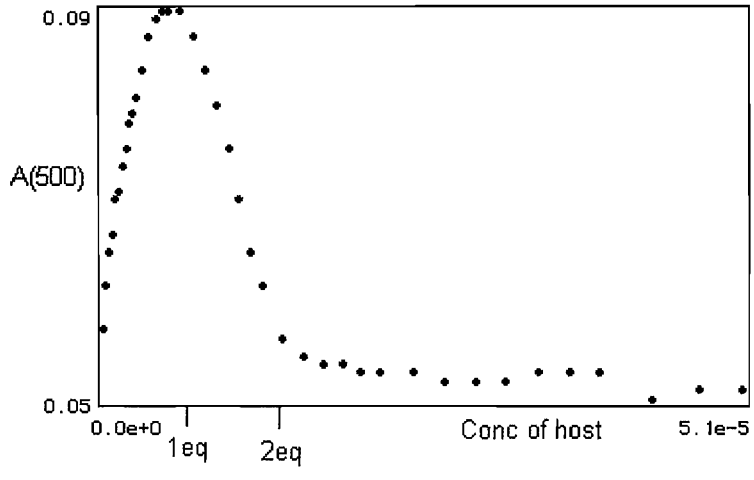


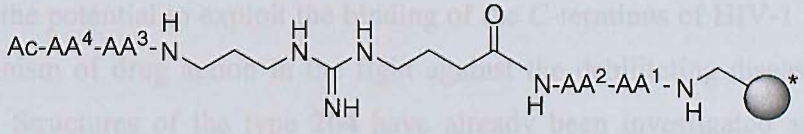
Fig 4-4 Typical UV-titration curve for addition of receptor 203 to dye-labelled guest 190 in 50% DMSO/borax at constant concentration of 190, monitoring the absorption at 500nm.

It is obvious from these data that binding events are occurring; there are significant changes in the absorption of the dye-chromophore. First the intensity of the dye-chromophore increases to a maximum after the addition of one equivalent of **203**, followed by a reduction and approach of a limit after the addition of two equivalents of **203**. However, it has proven impossible to fit such data convincingly to the natural 1:1 model, nor to 2:1/1:2 models. It is perhaps indicative of a complex situation in solution with many competing equilibria. For example, it is quite reasonable to assume that the

guest **190** is at least partially aggregated in solution, since it is the spontaneous formation of aggregates of the full Aβ peptide (of which **190** represents the *C*-terminus) which makes the guest of interest. Hence guests must potentially deaggregate before 1:1 binding can occur. Simple 1:1 binding may also be in competition with the binding of aggregates. This is further supported by the logarithmic relationship derived as a calibration curve; **190** does not strictly obey the Beer-Lambert law, which indicates complex behaviour in solution. Further work is required to evaluate a numerical, solution-based binding constant. However it has been proved that binding in solution *does* occur, and hence the screening of the library **149** (containing ladder sequences) to find receptors that bind in solution is proven as an effective methodology.

4.4 Tweezer Library

Encouraged by the success observed in screening library **149**, the screening of the tweezer library **184** was attempted (shown in abbreviated form **185**). This library again relies on a primary guanidinium-carboxylate ion-pair interaction. However, with the presence of two peptide side arms, the potential exists of a more sophisticated array of secondary interactions, potentially leading to the formation of pleated β-sheet structures between the peptide chain of the guest and the two side-arms (Fig 4-5).



185

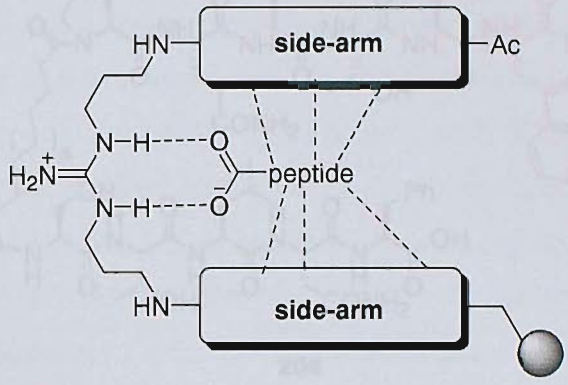
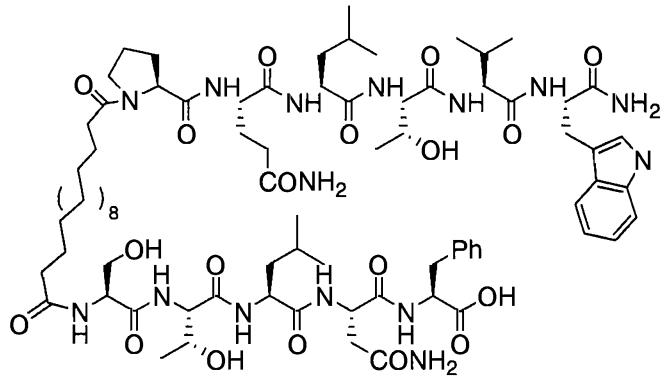


Fig 4-5 Potential binding mode of tweezer library 184 with primary guanidinium-carboxylate interaction and secondary interactions with side-arms

The greater demands of this indicated binding conformation (plus the potential for a greater number of secondary interactions) suggest that a greater selectivity will be possible amongst library members. However, owing to the flexibility of the binding-site (in comparison with the bicyclic guanidinium of library **149**), library members may be less efficient at binding than expected.

4.4.1 Screening with guest 195 (Leu-Asn-Phe)

There exists the potential to exploit the binding of the *C*-terminus of HIV-1 Protease as a novel mechanism of drug action in the fight against the debilitating disease AIDS (see Chapter 1). Structures of the type **204** have already been investigated as *C*-terminus binding agents; library members of **184** represent a further development of such structures by the incorporation of the guanidinium binding site.



204

The screening of an appropriate guest to represent the C-terminus of HIV-1 Protease (Leu-Asn-Phe) thus has potential to generate receptors of use in both biochemical investigation of the HIV-virus. Eventual development into drugs is also envisaged.

In a typical screening experiment, ~20mg of the resin library **184** (of sufficient mass such that the number of beads sampled represents a statistical likelihood of *at least* five copies of the library, often more) was allowed to swell in a mixture of 20% DMSO/buffer and titrated with aliquots of dye-labelled tripeptide **195** in 20% DMSO/buffer. The presence of the buffering agent ensures that the guest is present in the form of the free carboxylate, also acts as the counter-anion of the guanidinium obtained in aqueous solution. The system was allowed to equilibrate for at least 24 hours before inspection under a low-magnification hand lens. Concentration of peptide guest could be adjusted by dilution or titration of further guest to give an 'optimum' selectivity amongst the beads as judged by eye, similar to results obtained previously for single-armed receptor library **149**.

It was noted during the screens that large numbers (>50) became stained with the dye-labelled guest, even at low concentrations. The adjudication of which beads to select was made by eye; it should be appreciated that the situation reached is more closely considered as a continuum of staining, rather than a stark difference between stained/unstained beads. Thus it is already noted that though selectivity for guest **195** exists amongst library members, that selectivity is less pronounced than noted between single-armed library **149** and guest **190**.

Beads which were selected as being highly stained were individually removed from the screen medium using a micropipette and the Mtr-group (Arg-linker-protection) was removed by treatment with 1M TMSBr in TFA for 1 hour. Cleavage of the bead was then performed using the standard CNBr/TFA/H_2O mixture and the cleavage products were analysed by MALDI-TOF-MS as standard. Sequencing results of the screens are listed in Tables 4-6 to 4-9 below. Note that results from one screen (Table 4-8) arise from incubation and removal of a certain number of beads, followed by further titration of **195** and selection of further beads; hence the sequences are not determined under identical conditions for this case. A summary of the frequency of occurrence of residues accompanies each table. Certain selected beads could not be sequenced; these are included for completeness.

Bead	AA^1	AA^2	AA^3	AA^4
1	Phe	Val	Gly	Thr(OBn)
2	x	x	x	x
3	Phe	Phe	Lys(Cbz)	Ala
4	Thr(OBn)	Lys(Cbz)	Phe	Asn
5	Lys(Cbz)	Phe	Thr(OBn)	Val
6	Phe	Ala	Ala	Phe
7	Phe	Gly	Lys(Cbz)	Asn
8	Gly	Phe	Thr(OBn)	Val
9	Val	Phe	Lys(Cbz)	Thr(OBn)
10	Val	Asn	Ala	Asn
11	Lys(Cbz)	Thr(OBn)	Val	Phe
12	Val	Gly	Lys(Cbz)	Val
13	Val	Phe	Lys(Cbz)	Thr(OBn)

Residue	AA^1	AA^2	AA^3	AA^4
Thr(OBn)	8%	8%	15%	23%
Lys(Cbz)	15%	8%	38%	0%
Glu(OBn)	0%	0%	0%	0%
Gly	8%	15%	8%	0%
Ala	0%	8%	15%	8%
Val	31%	8%	8%	23%
Phe	31%	38%	8%	15%
Asn	0%	8%	0%	15%

Table 4-6 Sequences determined from a screening experiment of library 184 against guest 195 at a concentration of 11.34μM in 20% DMSO/borax, and % occurrence of each residue at each position

Bead	AA^1	AA^2	AA^3	AA^4
1	Val	Asn	Ala	Val
2	Phe	Thr(OBn)	Ala	Gly
3	Thr(OBn)	Thr(OBn)	Val	Val
4	Asn	Lys(Cbz)	Val	Ala
5	Thr(OBn)	Thr(OBn)	Ala	Thr(OBn)
6	x	x	x	x
7	Phe	Thr(OBn)	Ala	Val
8	Thr(OBn)	Ala	Lys(Cbz)	Lys(Cbz)
9	Phe	Thr(OBn)	Phe	Phe

Residue	AA^1	AA^2	AA^3	AA^4
Thr(OBn)	33%	44%	0%	11%
Lys(Cbz)	0%	11%	11%	11%
Glu(OBn)	0%	0%	0%	0%
Gly	0%	0%	0%	11%
Ala	0%	11%	44%	11%
Val	11%	0%	22%	33%
Phe	33%	0%	11%	11%
Asn	11%	11%	0%	0%

Table 4-7 Sequences determined from a screening experiment of library 184 against guest 195 at a concentration of 1.89μM in 20% DMSO/borax

Bead	AA^1	AA^2	AA^3	AA^4
1	Val	Thr(OBn)	Lys(Cbz)	Asn
2	Lys(Cbz)	Ala	Thr(OBn)	Gly
3	Gly	Val	Asn	Glu(OBn)
4	Phe	Gly	Phe	Ala
5	Phe	Val	Val	Ala
6	Asn	Val	Thr(OBn)	Asn
7	Ala	Ala	Phe	Gly
8	Gly	Ala	Gly	Gly
9	x	x	x	x
10	x	x	x	x
11	Ala	Val	Thr(OBn)	Val
12	Val	Asn	Val	Ala
13	Val	Gly	Gly	Val
14	Gly	Gly	Thr(OBn)	Gly
15	Val	Lys(Cbz)	Lys(Cbz)	Lys(Cbz)
16	x	x	x	x
17	Lys(Cbz)	Val	Val	Glu(OBn)

Residue	AA^1	AA^2	AA^3	AA^4
Thr(OBn)	0%	6%	18%	0%
Lys(Cbz)	12%	6%	12%	6%
Glu(OBn)	0%	0%	0%	12%
Gly	18%	18%	12%	24%
Ala	12%	18%	0%	18%
Val	18%	24%	18%	12%
Phe	12%	0%	12%	0%
Asn	6%	6%	6%	12%

Table 4-8 Sequences determined from a screening experiment of library 184 against guest 195 at varying concentration in 20% DMSO/borax

Bead	AA^1	AA^2	AA^3	AA^4
1	x	x	x	x
2	Val	Phe	Ala	Gly
3	Gly	Ala	Val	Thr(OBn)
4	Ala	Asn	Phe	Ala
5	Phe	Asn	Phe	Ala
6	Phe	Val	Val	Lys(Cbz)
7	x	x	x	x
8	Asn	Ala	Phe	Glu(OBn)
9	Phe	Val	Val	Asn
10	Gly	Val	Phe	Gly
11	x	x	x	x
12	Asn	Ala	Phe	Gly
13	Val	Phe	Thr(OBn)	Gly
14	Gly	Val	Ala	Thr(OBn)
15	Phe	Gly	Val	Ala
16	x	x	x	x
17	x	x	x	x
18	Asn	Glu(OBn)	Asn	Phe
19	Ala	Ala	Ala	Thr(OBn)
20	x	x	x	x
21	Phe	Asn	Asn	Ala
22	Phe	Lys(Cbz)	Asn	Asn
23	Ala	Phe	Thr(OBn)	Asn
24	x	x	x	x
25	Val	Asn	Phe	Lys(Cbz)
26	Asn	Thr(OBn)	Phe	Gly
27	Lys(Cbz)	Ala	Lys(Cbz)	Lys(Cbz)

Residue	AA^1	AA^2	AA^3	AA^4
Thr(OBn)	0%	4%	7%	11%
Lys(Cbz)	4%	4%	4%	11%
Glu(OBn)	0%	4%	0%	4%
Gly	11%	4%	0%	19%
Ala	11%	19%	11%	15%
Val	11%	15%	15%	0%
Phe	22%	11%	26%	4%
Asn	15%	15%	11%	11%

Table 4-9 Sequences determined from a screening experiment of library 184 against guest 195 at a concentration of 8.76μM in 20% DMSO/borax

Residues occur (across all screens) with the following listed frequency; Thr(OBn) 13%, Lys(Cbz) 12%. Glu(OBn) 2%, Gly 12%, Ala 15%, Val 19%, Phe 17%, Asn 11%. The only immediately obvious point is the paucity of Glu(OBn); all other residues appear roughly as one-eighth of the total proportion. Val, Phe and Ala appear slightly more than this, but not significantly so. The occurrence of each residue in each position is given in Table 4-10

Residue	**AA^1**	**AA^2**	**AA^3**	**AA^4**
Thr(OBn)	7%	16%	15%	11%
Lys(Cbz)	13%	7%	17%	9%
Glu(OBn)	0%	2%	0%	6%
Gly	13%	11%	6%	19%
Ala	9%	19%	17%	15%
Val	22%	17%	20%	17%
Phe	28%	15%	19%	7%
Asn	7%	13%	7%	17%

Table 4-10 Percent occurrence of each residue at each position from all beads sequenced

Looking at Table 4-10, we can see a distinct preference for certain residues in certain positions; AA^1 favours Val and Phe; AA^2 favours Ala, Val and Thr(OBn); AA^3 favours Ala, Val, Phe and Lys(Cbz); AA^4 favours Gly, Val, Asn and Ala. This generates ninety-six possible sequences that might be considered effective receptors for the guest **195** (This is greater than the total number of beads sequenced). Of these ninety-six, five sequences appear individually once across all the screens; Val-Thr(OBn)-Lys(Cbz)-Asn, Phe-Val-Val-Asn, Phe-Thr(OBn)-Ala-Gly, Phe-Thr(OBn)-Ala-Val and Phe-Val-Val-Ala (for AA^1-AA^2-AA^3-AA^4). However, the selection of any of these sequences is fundamentally unsatisfying since the number of potential 'best receptors' is so large; the results from the screening do not converge on the residues indicated to any great degree.

It is noted that each screen appears to give rise to a different distribution of residues and sequences. For example, three screens showed a preference for AA^1 = Phe which was not evident in the remaining screen. A further example is the frequent occurrence of AA^3 = Lys(Cbz) in one screen but not in the remaining screens. Such results are counter-

intuitive and puzzling. No adequate explanation is offered for these results, since the selectivity evident in the screens (as a differentiation in degree of staining) is undeniable.

A single sequence, Val-Phe-Lys(Cbz)-Thr(OBn), is duplicated in a single screen (Table 4-6). This recurrence lends credit to the view that this is an effective receptor for **195**. Of this sequence, residues AA^1, AA^2 and AA^3 also belong to the most frequent residues listed above, the sole exception being AA^4= Thr(OBn) (which occurs roughly the expected statistical 12%).

Considering residues AA^2 and AA^3 only (the two residues flanking the binding site, which are likely to be heavily implicated in binding conformations owing to proximity), it is to be noted that the combinations [AA^2-AA^3] Thr(OBn)-Ala, Phe-Thr(OBn) and Val-Val occur four times each throughout all screens and the combination Phe-Lys(Cbz) occurs three times throughout all screens. Obviously, two occurrences of the Phe-Lys(Cbz) motif are due to the duplicate sequence [Val- Phe-Lys(Cbz)-Thr(OBn)] previously mentioned.

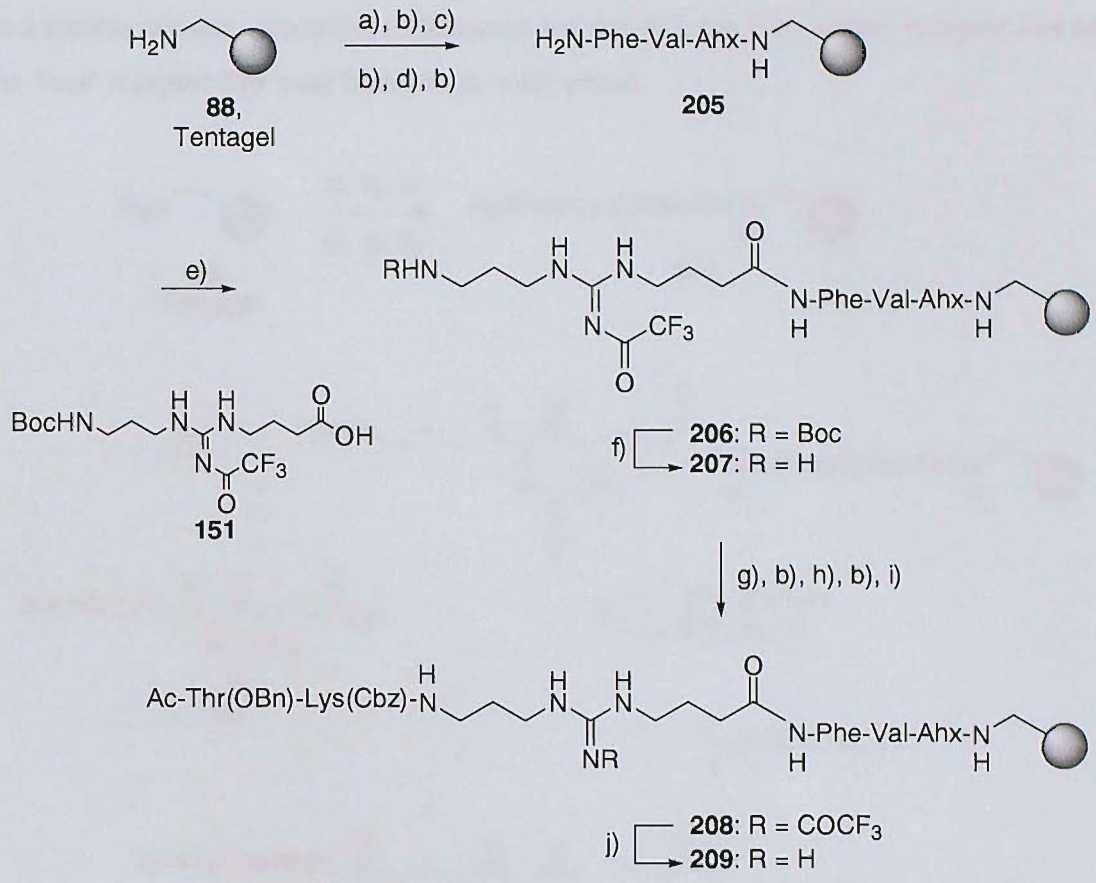
In view of the above considerations, it was decided to further investigate the receptor sequence [AA^1 = Val, AA^2 = Phe, AA^3 = Lys(Cbz), AA^4 = Thr(OBn)]. This is the duplicate result (within a single screen), thus perhaps representing the greatest selectivity amongst library members. The selected residues for $AA^{1/2/3}$ also represent a greater than random occurrence amongst all the screens at the given position. It is highlighted that this is a compromise choice.

Some investigation was also made of a 'bad' receptor sequence [AA^1 = AA^2 = AA^3 = AA^4 =Glu(OBn)] and a 'scrambled' receptor sequence [AA1 = Phe, AA^2 = Thr(OBn), AA^3 = Val, AA^4 = Lys(Cbz)] in order to make some assessment of the success of the screening experiment.

4.4.2 Resynthesis of selected receptor structures

In order to evaluate the binding of the most promising receptor sequence Val-Phe-Lys(Cbz)-Thr(OBn) selected as described above, resynthesis of this structure was required on both solid phase and as an isolated entity (*without* the presence of the Arg-containing linker). Methods are described below.

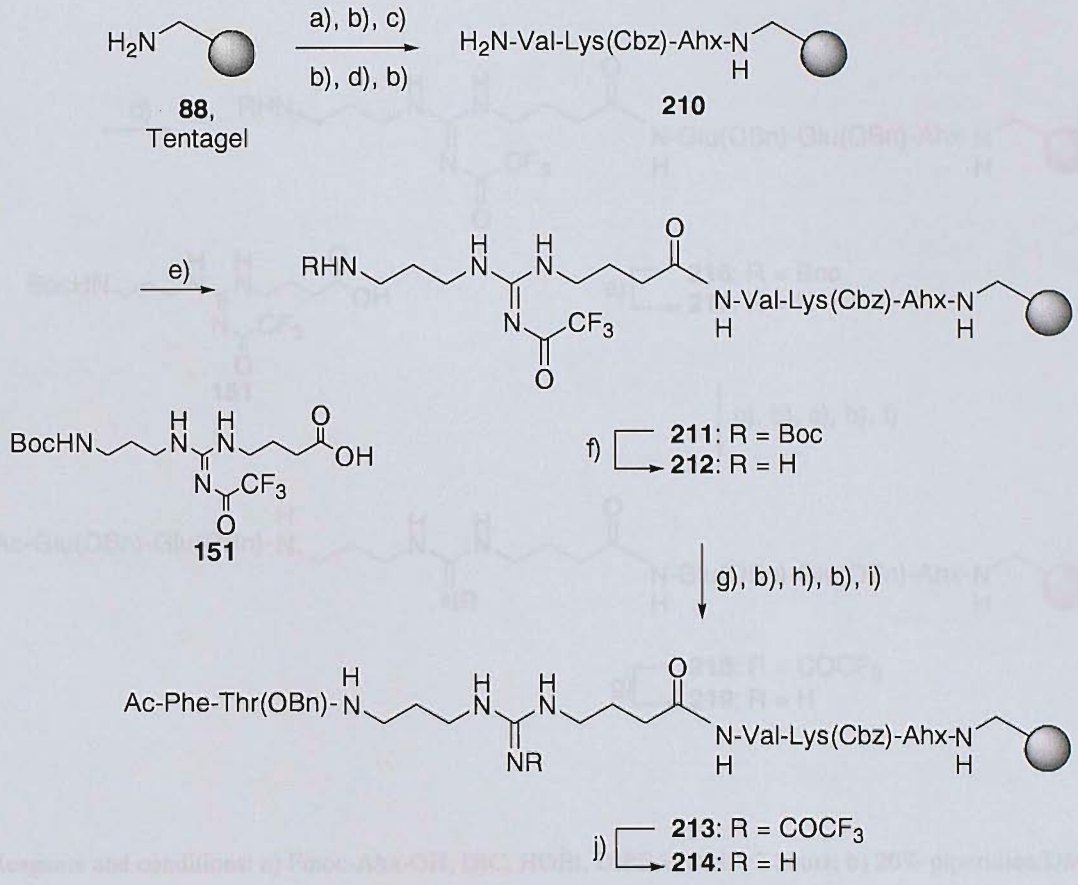
Solid-phase bound receptor **209** was synthesised according to Scheme 4-5, using standard Fmoc-SPPS procedures. Tentagel resin **88** was first derivatised with an Ahx residue and then the sequence Phe-Val by sequential coupling-deprotection steps, giving **205**. The CBS **151** was then coupled giving **206**. This was then deprotected with dilute TFA and Fmoc-SPPS then continued to install the sequence Thr(OBn)-Lys(Cbz) by sequential coupling-deprotection. A final Fmoc-deprotection and reaction with Ac_2O and DIPEA gave the acetyl-capped sequence **208**. All steps were monitored using the Kaiser ninhydrin test. The guanidine was then deprotected by treatment with 0.15M K_2CO_3 in MeOH/DMF/H_2O giving receptor **209**.



Reagents and conditions: a) Fmoc-Ahx-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% piperidine/DMF, 20mins; c) Fmoc-Val-OH, DIC, HOBt, DIPEA, DMF, 3 hours; d) Fmoc-Phe-OH, DIC, HOBt, DIPEA, DMF, 3 hours; e) **151**, DIC, HOBt, DIPEA, DMF, 3 hours; f) 20% TFA/DCM, 1 hour; g) Fmoc-Lys(Cbz)-OH, DIC, HOBt, DIPEA, DMF, 3 hours; h) Fmoc-Thr(OBn)-OH, DIC, HOBt, DIPEA, DMF, 3 hours; i) Ac_2O, DIPEA, DMF; j) 0.15M K_2CO_3, MeOH/DMF/H_2O, 3 hours

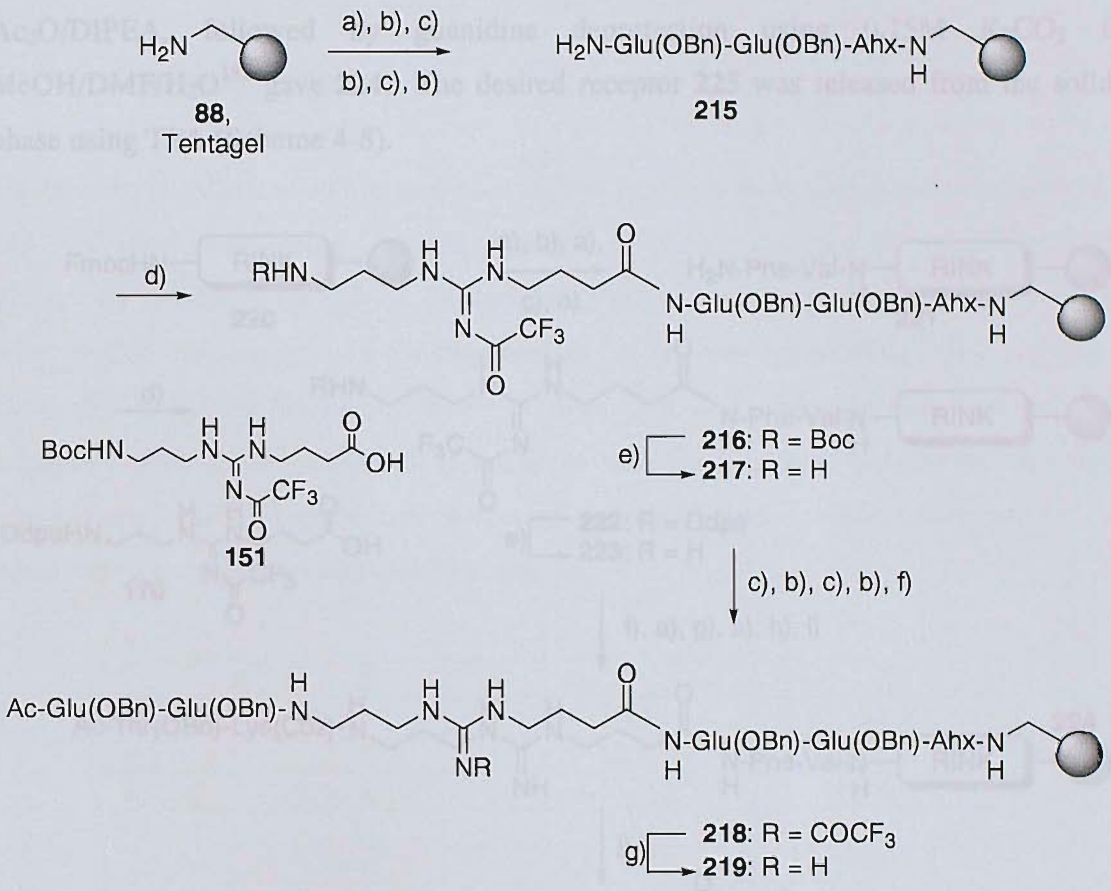
Scheme 4-5 Synthesis of solid-phase-linked receptor 209

In a similar manner, according to Schemes 4-6 and 4-7, the 'scrambled' receptor **214** and the 'bad' receptor **219** were built on the solid-phase.



Reagents and conditions: a) Fmoc-Ahx-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% piperidine/DMF, 20mins; c) Fmoc-Lys(Cbz)-OH, DIC, HOBt, DIPEA, DMF, 3 hours; d) Fmoc-Val-OH, DIC, HOBt, DIPEA, DMF, 3 hours; e) **151**, DIC, HOBt, DIPEA, DMF, 3 hours; f) 20% TFA/DCM, 1 hour; g) Fmoc-Thr(OBn)-OH, DIC, HOBt, DIPEA, DMF, 3 hours; h) Fmoc-Phe-OH, DIC, HOBt, DIPEA, DMF, 3 hours; i) Ac_2O, DIPEA, DMF; j) 0.15M K_2CO_3, MeOH/DMF/H_2O, 3 hours

Scheme 4-6 Synthesis of solid-phase-linked, 'scrambled' receptor 214

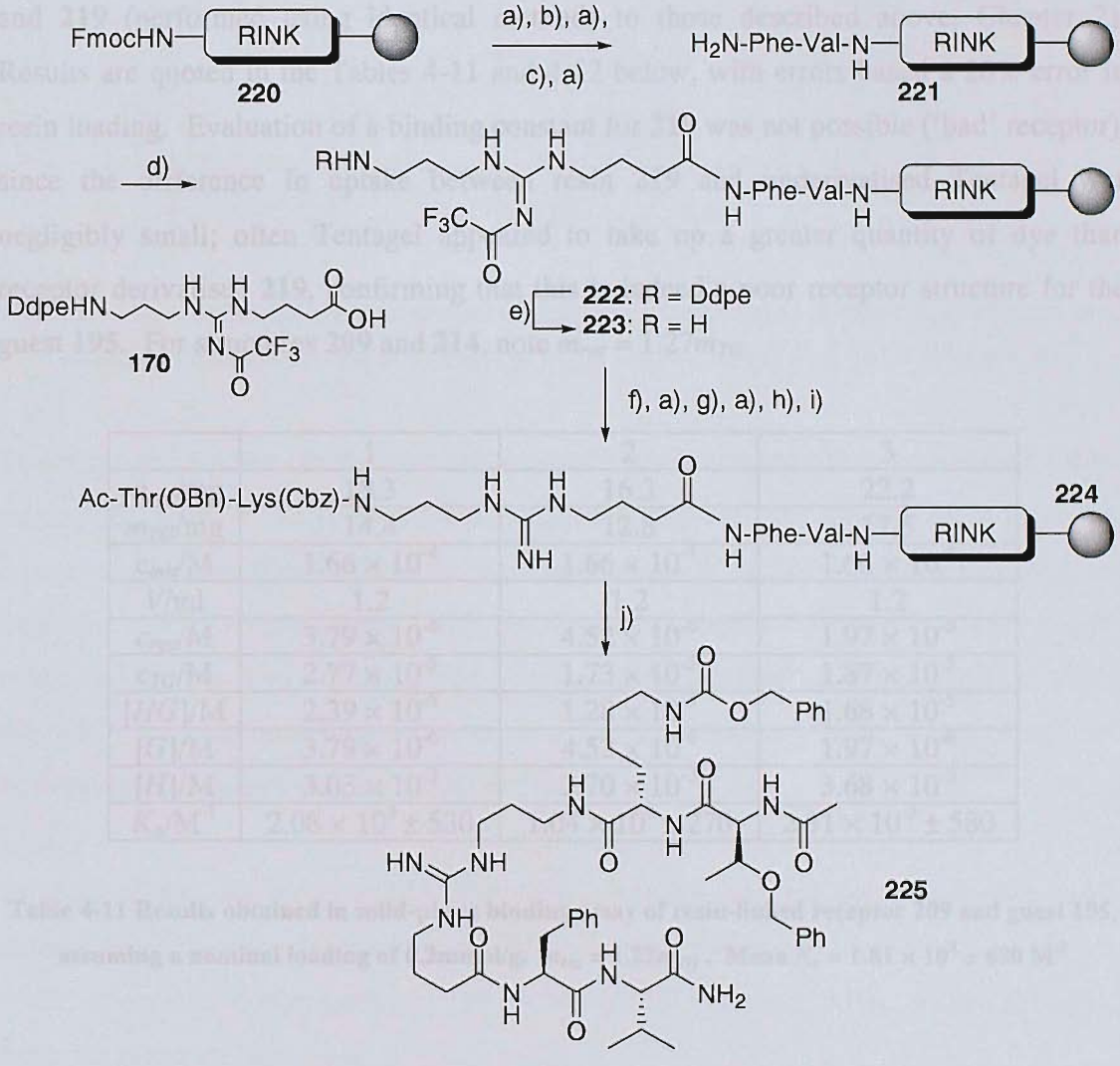


Reagents and conditions: a) Fmoc-Ahx-OH, DIC, HOBt, DIPEA, DMF, 3 hours; b) 20% piperidine/DMF 20 mins; c) Fmoc-Glu(OBn)-OH, DIC, HOBt, DIPEA, DMF, 3 hours; d) **151,** DIC, HOBt, DIPEA, DMF, 3 hours; e) 20% TFA/DCM, 1 hour; f) Ac_2O, DIPEA, DMF, 3 hours; g) 0.15M K_2CO_3, MeOH/DMF/H_2O, 3 hours

Scheme 4-7 Synthesis of solid-phase-linked 'bad' receptor 219

In order to prepare the desired receptor for solution-phase binding evaluation, a solid-phase synthesis was undertaken using Rink-amide resin **220**, which is cleaved using TFA to give peptide amides. Thus Rink-resin was first Fmoc-deprotected. Fmoc-SPPS was then used to couple the sequence Phe-Val. After deprotection, the Ddpe-protected CBS **170** was incorporated using DIC/HOBt and DIPEA. Deprotection of the Ddpe group proceeded simply using aqueous hydrazine, giving **223**. Further Fmoc-SPPS was carried out to install the sequence Thr(OBn)-Lys(Cbz). A final deprotection and reaction with

Ac_2O/DIPEA, followed by guanidine deprotection using 0.15M K_2CO_3 in MeOH/DMF/H_2O[182] gave **224**. The desired receptor **225** was released from the solid-phase using TFA (Scheme 4-8).



Reagents and conditions: a) 20% piperidine/DMF; b) Fmoc-Val-OH, DIC, HOBt, DIPEA, DMF, 3 hours; c) Fmoc-Phe-OH, DIC, HOBt, DIPEA, DMF, 3 hours; d) **170**, DIC, HOBt, DIPEA, DMF, 3 hours; e) aqueous hydrazine, 3 hours; f) Fmoc-Lys(Cbz)-OH, DIC, HOBt, DIPEA, DMF, 3 hours; g) Fmoc-Thr(OBn)-OH, DIC, HOBt, DIPEA, DMF, 3 hours; h) Ac_2O, DIPEA, DMF, 3 hours; i) 0.15M K_2CO_3, MeOH/DMF/H_2O, 3 hours; j) TFA, 2 hours, 22% based on initial resin loading

Scheme 4-8 Solid-phase synthesis of receptor 225

4.4.3 Evaluation of Binding

1:1 Binding constants were estimated for each of the three resin-linked receptors **209**, **214** and **219** (performed using identical methods to those described above, Chapter 2). Results are quoted in the Tables 4-11 and 4-12 below, with errors based a 20% error in resin loading. Evaluation of a binding constant for **219** was not possible ('bad' receptor), since the difference in uptake between resin **219** and underivatised Tentagel was negligibly small; often Tentagel appeared to take up a greater quantity of dye than receptor derivatised **219**, confirming that this is indeed a poor receptor structure for the guest **195**. For structures **209** and **214**, note $m_{rec} = 1.27 m_{TG}$.

	1	2	3
m_{rec}/mg	18.3	16.3	22.2
m_{TG}/mg	14.4	12.8	17.5
c_{init}/M	1.66×10^{-4}	1.66×10^{-4}	1.66×10^{-4}
V/ml	1.2	1.2	1.2
c_{rec}/M	3.79×10^{-6}	4.52×10^{-6}	1.97×10^{-6}
c_{TG}/M	2.77×10^{-5}	1.73×10^{-5}	1.87×10^{-5}
$[HG]$/M	2.39×10^{-5}	1.28×10^{-5}	1.68×10^{-5}
$[G]$/M	3.79×10^{-6}	4.52×10^{-6}	1.97×10^{-6}
$[H]$/M	3.03×10^{-3}	2.70×10^{-3}	3.68×10^{-3}
K_a/M^{-1}	$2.08 \times 10^{3} \pm 530$	$1.04 \times 10^{3} \pm 270$	$2.31 \times 10^{-3} \pm 580$

Table 4-11 Results obtained in solid-phase binding assay of resin-linked receptor 209 and guest 195, assuming a nominal loading of 0.2mmol/g. $m_{rec} = 1.27 m_{TG}$. Mean $K_a = 1.81 \times 10^3 \pm 680$ M^{-1}

	1	2	3
m_{rec}/mg	26.3	13.7	27.3
m_{TG}/mg	20.7	10.8	21.5
c_{init}/M	1.66×10^{-4}	1.66×10^{-4}	1.66×10^{-4}
V/ml	1.2	1.2	1.2
c_{rec}/M	1.84×10^{-6}	2.84×10^{-6}	1.17×10^{-6}
c_{TG}/M	5.01×10^{-6}	8.52×10^{-6}	5.68×10^{-6}
$[HG]$/M	3.17×10^{-6}	5.68×10^{-6}	4.51×10^{-6}
$[G]$/M	1.84×10^{-6}	2.84×10^{-6}	1.17×10^{-6}
$[H]$/M	4.83×10^{-3}	2.28×10^{-3}	4.55×10^{-3}
K_a/M^{-1}	$3.94 \times 10^{2} \pm 99$	$8.78 \times 10^{2} \pm 222$	$8.49 \times 10^{2} \pm 211$

Table 4-12 Results obtained in solid-phase binding assay of resin-linked receptor 214 and guest 195, assuming a nominal loading of 0.2mmol/g. $m_{rec} = 1.27 m_{TG}$. Mean $K_a = 7.07 \times 10^2 \pm 270$ M^{-1}

The estimated binding constant for the nominal 'good' receptor **209** is greater than for the 'scrambled' structure **214**. Both are of satisfying magnitude for the binding of a guest in a polar aqueous medium. Since the estimated constant for **209** is greater than **214**, the importance of residue sequence within the receptor (as opposed to merely residue nature) is shown. Since analysis of the screen data lead to the proposal of **209** as a good receptor, the utility of combinatorial screening is again demonstrated, even when using non-homogeneous beads displaying ladder sequences.

Binding of **195** by receptor **225** in solution was investigated by UV-titration. The absorption of the dye-chromophore of **195** was monitored during the addition of aliquots of **225** at constant concentration of **195**, in 20% DMSO/borax solution. A typical binding isotherm is shown in Fig 4-6.

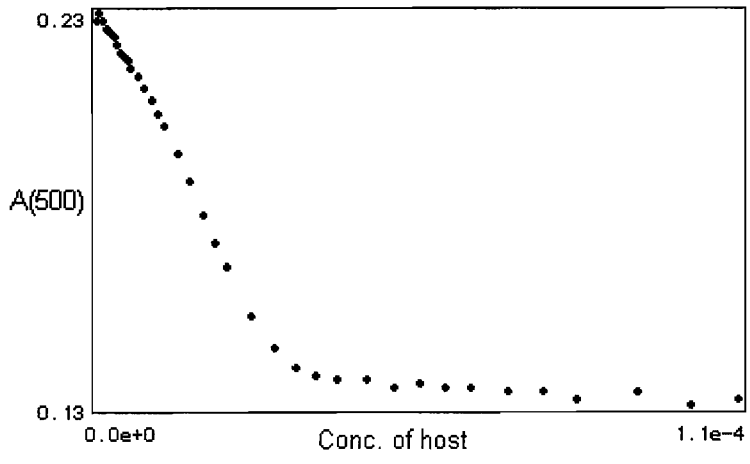


Fig 4-6 Typical UV-titration curve for addition of tweezer receptor 225 to guest 195 in 20% DMSO/borax at constant concentration of 195, monitoring the absorption at 500nm.

The data shows appropriate behaviour, with a reduction in the intensity of the chromophore and the approach of a saturation limit, typical of 1:1 behaviour. However, it did not prove possible to fit the data using a standard 1:1 binding model; numerically derived curves did not fit satisfactorily across the entire data range. That binding is occurring is indisputable, but the precise nature of the interaction remains unspecified. It is likely that the primary interaction is indeed 1:1, with alternative binding equilibria arising at greater concentrations of host (perhaps self-aggregation of host). Further studies of solution binding need to be performed to further investigate the binding situation. However, the identification of a useful, strongly binding receptor in aqueous medium of the *C*-terminal sequence of HIV-protease has been accomplished, illustrating the effectiveness of the screening methodology.

4.5 Conclusions and Future Work

This chapter has described the screening of solid-phase libraries of receptors for carboxylates **149** and **184** against oligopeptide carboxylate guests of biological significance. Structure determination of the identified 'hit' beads was accomplished by MALDI-TOF-MS analysis of the 'ladder sequence' index incorporated into the library

beads during synthesis (as described in Chapter 3). An evaluation of the 'hit' sequences has been made and a designated effective receptor identified in both cases.

Screening of single-armed receptor library **149** gave a high consensus amongst screening results, and led to a measured binding constant on the solid phase of $K_a \sim 1.4 \times 10^3\ M^{-1}$. Solution-phase binding was also observed but could not be quantified, owing to unusual behaviour perhaps arising from the known propensity of the guest to form aggregates in aqueous medium.

The screening results reported for library **184** are disappointing with regards to selectivity. There is no clearly identifiable sequence which emerges from the data as the most selective receptor. The lack of duplication amongst 'hit' sequences is also of concern, since the quantity of resin used should represent duplicate copies of all library members; duplication of results would be expected but is not in fact observed. This lack of selectivity perhaps reflects the flexibility inherent in the receptor design; no rigidity has been incorporated. This lack of rigidity (and hence pre-organisation for binding) perhaps means that the library members are not enveloping guest molecules in the predicted fashion and no β-sheets are forming. If there is little contribution from secondary interactions (i.e. the guanidinium-carboxylate ion pair is the only significant interaction), then selectivity would not be expected. However, this contradicts the visually evident selectivity within a screen, where uptake of dye-guest **195** differs significantly between the most heavily stained beads and the least heavily stained beads.

Some comfort is to be drawn from the severe lack of Glu(OBn) throughout the results, suggesting that this residue is highly disfavoured in the binding of **195**. This could potentially be of significance, perhaps leading to the design of true effective receptors by minimising the effects of disfavoured residues (by thorough investigation of their properties).

Comparing with previous literature results, it is noted that Chmielewski *et al*'s tweezer receptors[155-159] based on an alkyl spacer head-group (e.g. **204**) are effective binders of

HIV-1-Protease (as measured by bio-assay). In this case the alkyl spacer is much larger than the CBS used in library **184**. This suggests that inherent flexibility is of no consequence for effective binding, but that perhaps the receptors of library **184** are too congested around the binding-site to allow efficient access by the carboxylate guest (with bulky Phe as the *C*-terminal residue of the guest). Receptors (e.g. **79-82** by Sicsic *et al*[161,162]) utilise rigid head-groups of much smaller size; such receptors are also effective at binding HIV-1 Protease as determined by bioassay. In this case, it is perhaps the pre-organisation of the receptor which is conducive to binding. Such arguments point towards the need for a larger, more rigidified CBS than **151**.

The designated effective receptor was resynthesised on the solid-phase (**209**) and estimation of binding constant made. **209** binds guest **195** with $K_a \sim 1.8 \times 10^3\ M^{-1}$ on the solid phase. A binding event in solution was also demonstrated by UV titration for receptor **225** with guest **195**. Unfortunately no absolute binding constant could be calculated based on the data obtained, suggesting alternative titration methods are necessary; ITC and/or NMR methods are indicated.

Further work in the area is anticipated. Screening of the current libraries with a greater variety of carboxylate guests is an obvious application. Libraries of greater scope utilising more diversity elements (different amino acids) or longer side-arms are possible using the current methodology. Of particular interest would be the use of alternative, rigidified binding sites, which could potentially give rise to greater selectivity than evident in the current work. Rigidity could also be achieved by performing on-bead cyclisation reactions of appropriate substrates to give libraries of macrocycles (which can still be analysed by MS in a similar fashion). Of particular relevance to the binding of the HIV-1 Protease *C*-terminal strand, a binding site head-group of greater dimensions (of similar size to the Chmielewski alkyl-spaced tweezers[155]) would be worth investigating.

Chapter 5

Experimental

5.1 General Experimental

Solvents for synthesis were purchased from Fisher. Tentagel S NH_2 resin, 2-chlorotrityl resin or Rink Amide resin were used as solid support in peptide synthesis and purchased from Rapp Polymere (Germany) or NovaBiochem. Amino acids were purchased from NovaBiochem or BAChem. All other chemicals were purchased from Aldrich, Avocado, Fluka, Lancaster or Novabiochem. Peptide synthesis on solid phase was performed in polypropylene filtration tubes with polyethylene frits on a Visiprep SPE Vacuum Manifold from Supelco. The reaction containers were agitated on a blood tube rotator (Stuart Scientific Blood Tune Rotator SB1). Coupling was monitored using the Kaiser Ninhydrin Test (described below). Thin layer chromatography (TLC) was performed on aluminium-backed plates Merck silica gel 60 F254 or Macherey-Nagel Alugram Sil G/UV_{254}. Silica 60A (particle size 35-70 micron) was used for column chromatography.

5.2 General Instrumentation

Melting points were determined in open capillary tubes using a Gallenkamp Electrothermal Melting Point Apparatus, all are uncorrected. Infrared Spectra were recorded using a Bio-Rad FT-IR spectrometer or a Nicolet 380 FT-IR spectrometer as neat solids or oils. Proton NMR spectra were obtained at 300 MHz on a Bruker AC 300 or Bruker AV 300 and at 400 MHz on a Bruker DPX 400. Carbon NMR spectra were recorded at 75 MHz on a Bruker AC300 and at 100 MHz on a Bruker DPX 400. Analytical RP-HPLC was performed on a HP1100 system equipped with a Phenomenex Prodigy C_{18} reverse phase column (150 × 4.6 mm i.d.) with a flow rate of 1ml/min. UV Spectra were recorded using a Shimadzu UV-1601 spectrometer in quartz cells. Low-resolution ES mass spectra were recorded using a Micromass Platform II quadrupole mass analyser (Fisons VG platform through a Hewlett Packard 1050 HPLC system).

High-resolution ES mass spectra were obtained on a Bruker Apex III FT-ICR mass spectrometer.

5.3 General procedures for solid-phase reactions and solid-phase tests

5.3.1 Resin preparation and washing

Prior to reaction, resins were swollen for 15 mins in DCM. After completion of reaction, resins were drained of the reaction medium and extensively washed; first with DCM, then DMF, then DCM. Each washing consisted of suspending resin in solvent (~ 15ml per g of resin) and allowing to stand for one minute before draining and then repeating the process twice. Any additional/alternate washing is noted in the appropriate section.

5.3.2 Standard coupling of amino acids

Coupling of *N*-protected amino acids to an NH_2-functionalised resin was carried out according to the following procedure. A protected amino acid (in excess; precise amount varies depending on substrate, see individual cases) and HOBt (excess) are dissolved in DMF (~1ml per 50mg of amino acid). DIC (excess) is then added and the mixture added to the pre-swollen resin. Finally excess DIPEA is added, the reaction vessel sealed and vented, and the mixture agitated for at least one hour. The reaction was repeated as needed until resin gave a negative ninhydrin test.

5.3.3 General protocol for solid-phase Fmoc-deprotection

Resin (in preswollen state) to be deprotected is treated with 20% piperidine/DMF (~20ml per g of resin) for 10 mins. The process is repeated before washing as described above.

5.3.4 General protocol for solid-phase Boc-deprotection

Resin (in preswollen state) to be deprotected is treated with 20% TFA/DCM (~20ml per g of resin) for 20 mins. The resin is drained and washed with 50% DIPEA/DMF (~ 20ml per g of resin), before washing as described above.

5.3.5 General protocol for solid-phase Ddpe-deprotection

Resin (in preswollen state) to be deprotected is treated with 30% aqueous hydrazine for 3 hours. The resin is then drained, washed with DMF (~20ml per g of resin) before washing as described above.

5.3.6 General protocol for the capping reaction

Resin (in preswollen state) to be capped is shaken with a solution of *p*-Br PhCOOH in DMF (quantities specified in individual experiments) for 1 hour, before addition of excess DIC, HOBt and DIPEA and further shaking for 18 hours. The resin is then drained and washed according to general procedure described above.

5.3.7 Kaiser Ninhydrin test[178]

Qualitative ninhydrin tests were carried out by placing a small sample of resin (mgs) into a test-tube and adding 3-5 drops of Reagent A and 1-2 drops of Reagent B. The tube was then incubated for 5 mins at 100°C. Blue beads and/or blue solution indicates the presence of free $-NH_2$ residues on the resin ('positive' result). For comparison, a blank consisting of the reagent mixture without resin was also prepared.

Quantitative ninhydrin tests were carried out by accurately weighing a mass of dry resin into a test-tube (3-6mg). 3 drops of Reagent A and 1 drop of Reagent B were added and the mixture incubated at 100°C for 5 mins. A control was simultaneously prepared using only the reagent mixture without resin. The tubes were cooled, 60% $EtOH/H_2O$ was

added and the solutions mixed thoroughly. The mixtures were filtered through a glass wool plug and the resin rinsed with 0.5M Et_4NCl. The solutions were then made up to known volume. Absorbance of the test solution at 570nm was measured using the control as a blank. NH_2-loading was calculated using the following equation;

$$S = \frac{A_{570}.V.10^3}{E.m}$$

where S represents NH_2-loading in mmol/g, V represents volume of test solution in ml, A_{570} is the absorption at 570nm, m is the mass of the tested resin in mg and E is the extinction coefficient of the chromophore observed (= 1.5×10^4 Mcm^{-1}).

Reagent A: *Solution 1* - Reagent grade phenol (40g) was dissolved in absolute EtOH under mild heating. After cooling, IWT TMD-8 ion exchange resin (4g) was added and the mixture stirred for one hour before filtering. *Solution 2* – KCN (65mg) was dissolved in distilled water (100ml). A 2ml aliquot of this solution was diluted to 100ml with pyridine. IWT TMD-8 ion exchange resin (4g) was added and the mixture stirred for one hour before filtering. Solution 1 and solution 2 were then mixed.

Reagent B: Ninhydrin (2.5g) was dissolved in absolute ethanol. The solution was stored in the dark.

5.3.8 Quantitative Fmoc-test[179]

A known mass of dry resin was placed in a test-tube. Freshly prepared 20% piperidine/DMF (1ml) was then added and the mixture agitated for 1 min and allowed to stand for 10 min. The mixture was then filtered through a glass-wool plug and the resin rinsed with freshly prepared 20% piperidine/DMF. The solutions were then made up to known volume with freshly prepared 20% piperidine/DMF. Absorbance of the test solution at 302nm was measured using freshly prepared 20% piperidine/DMF as a blank. Fmoc-loading was calculated according to the following equation;

$$S = \frac{A_{302}.V.10^3}{E.m}$$

where S represents Fmoc-loading in mmol/g, V represents volume of test solution in ml, A_{302} is the absorption at 302nm, m is the mass of the tested resin in mg and E is the extinction coefficient of the chromophore observed (= 7800 Mcm^{-1}).

5.3.9 Chloranil test[180]

To a small sample of resin were added two drops of 2% acetaldehyde/DMF and two drops 2% *p*-chloranil/DMF. The sample was allowed to stand for 5 mins. Blue stained beads indicate the presence of secondary amines on the resin ('positive' result).

5.4 Mass spectrometric methods for single bead analysis

5.4.1 Deprotection and cleavage of material from single resin beads[93]

A single resin bead to be analysed was placed in a glass conical insert (100µl volume) using tweezers or a micropipette. The insert was placed inside an Eppendorf tube. If Arg-deprotection was required, it was carried out as follows;

Pbf-deprotection: the bead was treated with 50µl of 50% TFA/DCM for 1 hour. The solvent was then removed under vacuum.

Mtr-deprotection: the bead was treated with 40µl of 1M TMSBr in TFA for 1 hour. The reagent mixture was then removed under vacuum.

Cleavage was performed by treatment of the bead with 20µl of a solution of CNBr in 1:1 TFA/H_2O (concentration 50-80mg/ml) and incubation in the dark for 18 hours, sealed.

5.4.2 Sample preparation

The cleavage solution was evaporated from the sample tubes *in vacuo*. 20μl of either MeCN or 1:1 MeCN/H_2O was then added to the tube. The samples were sonicated for 5 mins. 1μl of each sample was spotted onto the MALDI-platen. 1μl of a solution of matrix solution was then applied on top and the sample mixed on the platen surface. The spot was allowed to dry under a stream of warm air before analysis.

Matrix: α-cyanohydroxycinnamic acid was prepared as a saturated solution in either acetone or 1:1 MeCN/H_2O for application as above. 1:1 MeCN/H_2O gave consistently better results for more complex mixtures.

5.4.3 MALDI acquisition and data processing

Spectra were acquired on a Micromass-Tofspec 2E reflectron MALDI-TOF machine, recording in reflectron positive ion mode at a pulse voltage 3000V. Spectra were obtained by summing multiple laser shots. The following compounds were used for external calibration; Oxybutinin (*m/z* 358.24 $[M+H]^+$), Terfenadine (*m/z* 472.32 $[M+H]^+$), Bradykinin (*m/z* 1060.57 $[M+H]^+$), Substance P (*m/z* 1347.74 $[M+H]^+$), Renin Substrate Tetradecapeptide (Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu-Leu-Val-Tyr-Ser) (*m/z* 1758.93 $[M+H]^+$). In the presence of the Arg-linker, species were generally observed as $[M+H]^+$. In the absence of the Arg-linker, species were generally observed as $[M+Na]^+$.

Analysis of raw spectra was performed using Masslynx v3.2 (Micromass Ltd., Manchester, UK). The cluster analysis function was used with the following parameters; First mass difference 2.00, First ratio 1.0, Mass tolerance 0.05-0.1, Ratio tolerance 15-50%, Threshold 7-10%.

5.5 Preparation of buffer solutions

5.5.1 Borax buffer

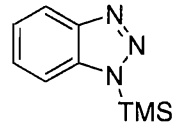
Solution A: 1.24g of boric acid were dissolved in 100ml distilled water. *Solution B*: 3.81g of sodium tetraborate were dissolved in 200ml distilled water. 50ml of solution A and 115ml of solution B were mixed to create borax buffer, pH 9.2

5.5.2 HEPES buffer

0.1M solutions of HEPES and sodium HEPES were prepared in distilled water. Equal volumes of each solution were mixed to create HEPES buffer, pH 8.0

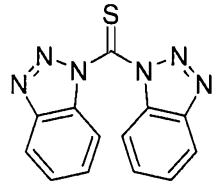
5.6 Experimental for Chapter 2

1-(1,1,1-Trimethylsilyl)-1*H*-1,2,3-benzotriazole (98)



According to a procedure by Lesiv *et al.*[181]
Benzotriazole (10.25g, 86mmol) was suspended in HMDS (50ml) and the mixture heated at reflux for 4 hours. After cooling, the product was obtained by vacuum distillation (fraction distilling at 78°C at 0.08Torr) as a colourless oil, **98** (12.3g , 75%): b.p.: 78°C at 0.08Torr; ^{1}H NMR (300MHz, $CDCl_3$): δ 8.02 (d, 1H, *J* = 8.0Hz, arom), 7.55 (d, 1H, *J* = 8.0Hz, arom), 7.34 (t, 1H, *J* = 8.0Hz, arom), 7.23 (t, 1H, *J* = 8.0Hz, arom), 0.66 (s, 9H, TMS); ^{13}C NMR (75MHz, $CDCl_3$): 146.5 (C), 137.6 (C), 126.9 (CH), 123.3 (CH), 119.5 (CH), 111.0 (CH), -0.70 (CH_3); LRMS: product decomposes under MS conditions
Data agrees with literature results.[181]

Di-1*H*-1,2,3-benzotriazol-1-ylmethanethione (99)

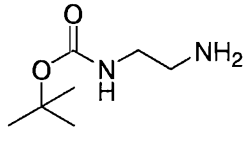


According to a procedure modified from Larsen *et al.*[169]

1-(1,1,1-trimethylsilyl)-1*H*-1,2,3-benzotriazole **98** (8.36g, 43.6mmol) was dissolved in DCE (200ml). Thiophosgene (1.7ml, 21.8mmol) was dissolved in DCE (100ml) and added dropwise to the solution over 3 hours. The mixture was stirred for 2 hours. A white precipitate forms, which is removed by filtration. The filtrates are concentrated *in vacuo* and then diluted with petrol to precipitate the product, which was collected by filtration; **99** Yellow solid (3.41g, 56%); m.p. 170-171°C (lit. value: 170-171°C)[169], IR ν_{max} (cm^{-1}) 1596 (w), 1482 (m), 1446 (m), 1401 (m), 1371 (m), 1282 (m), 1192 (s), 1062 (m), 925 (s), 894 (s), 833 (s), 735 (s); ^{1}H NMR (300MHz, $CDCl_3$) 8.27 (2H, m), 8.22 (2H, m), 7.74 (2H, m), 7.60 (2H, m); ^{13}C NMR (75MHz, $CDCl_3$) 168.0 (C), 145.1 (C), 131.4 (C), 129.1 (CH), 125.5 (CH), 119.2 (CH), 112.3 (CH); LRMS: decomposes under MS conditions.

Data agrees with literature results.[169]

***tert*-Butyl *N*-(2-aminoethyl)carbamate (102)**

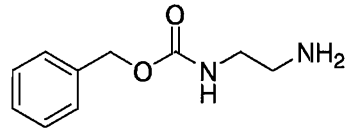


Di-*tert*-butyl dicarbonate (24.2g, 0.111mol) was dissolved in $CHCl_3$ (700ml) and added dropwise to 1,2-ethylenediamine (22.2ml, 0.322mol) in $CHCl_3$ (60ml) over the course of 12 hours. Stirring was continued for a further 12 hours. The reaction mixture was then concentrated *in vacuo* to half its volume and washed twice with sat. $NaHCO_3$(aq). The organic phase was dried over $MgSO_4$ and evaporated to give **102** as a colourless oil (13.66g, 77%). IR ν_{max} (cm^{-1}): 3352 (m), 2974 (m), 2931 (m), 2863 (w), 1688 (s), 1520

(s), 1250 (m), 1168 (s); ^{1}H NMR (300MHz, $CDCl_3$): δ 4.96 (br s, 1H, N*H*Boc), 3.17 (q, 2H, *J* = 6.0Hz, C*H*$_2$NH$_2$), 2.80 (t, 2H, *J* = 6.0Hz, BocNHC*H*$_2$), 1.09 (s, 9H, Boc); ^{13}C NMR (75MHz, $CDCl_3$): δ 156.4 (C), 79.5 (C), 43.3 (CH_2), 41.0 (CH_2), 28.5 (CH_3); LRMS: (ES^+) *m/z* 160.7 $[M+H]^+$.

Data agrees with literature results.[182]

Benzyl *N*-(2-aminoethyl)carbamate (100)

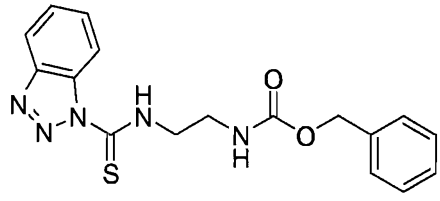


According to a procedure by Pittelkow *et al*[183]

1,2-ethylenediamine (5ml, 74.8mmol) was dissolved in absolute EtOH (200ml). Benzylphenylcarbonate (17.1g, 74.8mmol) was dissolved in absolute EtOH (50ml) and added to the diamine solution in three portions over ¾ hours. The mixture was then stirred for 18 hours. A white precipitate was noted. The reaction mixture was evaporated *in vacuo*, the residue taken up in distilled water and acidified with 2M HCl(aq). The aqueous phase was extracted twice with DCM. After phase separation, the aqueous phase was basified with 2M KOH(aq) and extracted twice with DCM. The basic organic fractions were combined, dried over $MgSO_4$ and evaporated to give the product **100** as colourless oil (8.68g, 60%). ^{1}H NMR (300MHz, $CDCl_3$): δ 7.34 (m, 5H, arom), 5.23 (br s, 1H, N*H*Cbz), 5.10 (s, 2H, C*H*$_2$Ph), 3.23 (m, 2H, C*H*$_2$CH$_2$), 2.81 (m, 2H, C*H*$_2$CH$_2$), 1.77 (br s, 2H, N*H*$_2$); ^{13}C NMR (75MHz, $CDCl_3$): 156.8 (C), 136.7 (C), 129.7 (CH), 128.6 (CH), 128.2 (CH), 66.8 (CH_2), 43.9 (CH_2), 41.8 (CH_2); LRMS: (ES^+) *m/z* 194.9 $[M+H]^+$.

Data agrees with literature results.[183]

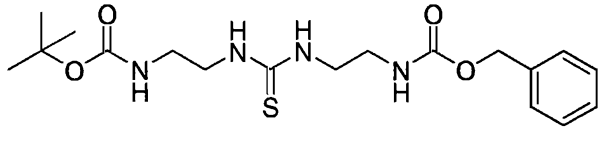
Benzyl *N*-2-[(1*H*-1,2,3-benzotriazol-1-ylcarbothioyl)amino]ethylcarbamate (103)



According to a procedure by Katritzky *et al.*[184]

Protected amine **100** (72mg, 0.37mmol) was dissolved in freshly distilled DCM (1ml) and added dropwise to a solution of di-1*H*-1,2,3-benzotriazol-1-ylmethanethione **99** (103.6mg, 0.37mmol) in freshly distilled DCM (2ml). The mixture was stirred for 18 hours and then evaporated *in vacuo*. The residue was taken up in EtOAc and the organic phase was washed sequentially with saturated $NaHCO_3$(aq), water and brine. The organic phase was dried over $MgSO_4$ and evaporated. This residue was redissolved in DCM. Product precipitated upon the addition of petrol and was collected by filtration to give **103** as a white solid (97.2mg, 74%); m.p. 142°C; IR ν_{max} (cm^{-1}) 3330 (w), 3246 (br w), 1682 (m), 1519 (m), 1451 (m), 1266 (m), 1208 (m), 1158 (m), 1002 (m), 739 (m); ^{1}H NMR (300MHz, $CDCl_3$): δ 9.76 (br s, 1H, CSN*H*), 8.86 (d, 1H, *J* = 8.5Hz, arom), 8.10 (d, 1H, *J* = 8.5Hz, arom), 7.62 (ddd, 1H, *J* = 8.5, 7.0, 1.0Hz, arom), 7.46 (ddd, 1H, *J* = 8.5, 7.0, 1.0Hz, arom), 7.30 (m, 5H, arom), 5.39 (br s, 1H, CbzN*H*), 5.11 (s, 2H, CH_2Ph), 4.01 (dt, 2H, *J* = 6.0, 5.5Hz), 3.64 (dt, 2H, *J* = 6.0, 5.5Hz); ^{13}C NMR (75MHz, $CDCl_3$): δ 175.3 (C), 157.4 (C), 147.1 (C), 136.3 (C), 132.6 (C), 130.4 (CH), 128.6 (CH), 128.3 (CH), 125.8 (CH), 120.4 (CH), 116.1 (CH), 67.3 (CH_2), 45.7 (CH_2), 39.6 (CH_2); LRMS: decomposes under MS conditions.

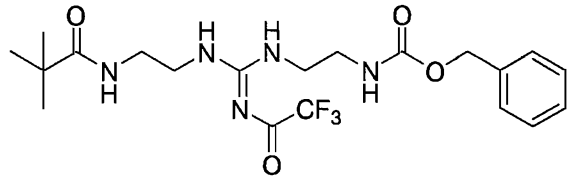
Benzyl *N*-(2-[(2-[(*tert*-butoxycarbonyl)amino]ethylamino)carbothioyl]aminoethyl) carbamate (104)



According to a procedure by Katritzky *et al.*[184]

Amine **102** (367mg, 2.29mmol) was dissolved in freshly distilled DCM (20ml) and added dropwise to a solution of carbamate **103** (814mg, 2.29mmol) in freshly distilled DCM (60ml). Et_3N (0.64ml, 4.58mmol) was added and the mixture stirred for 18 hours. The reaction mixture was evaporated and the residue taken up in EtOAc. The organic phase was washed sequentially with saturated $NaHCO_3$(aq), 1M HCl(aq), water and brine, then dried over $MgSO_4$ and evaporated. The residue was purified by column chromatography (eluent 40-50% EtOAc/petrol) to give the product thiourea **104** as a yellow oil (533.4mg, 59%); IR ν_{max} (cm^{-1}): 3302 (m), 2961 (m), 2932 (m), 2883 (w), 1679 (s), 1521 (s), 1245 (s), 1161 (s); ^1H NMR (400MHz, $CDCl_3$): δ 7.34 (m, 5H, arom), 6.93 (br s, 2H, N*H* thiourea), 5.56 (br s, 1H, carbamate N*H*), 5.02 (m, 3H, CH_2Ph + carbamate N*H*), 3.53 (m, 2H, CH_2CH_2), 3.42 (m, 2H, CH_2CH_2), 3.33 (m, 2H, CH_2CH_2), 3.20 (m, 2H, CH_2CH_2), 1.35 (s, 9H, Boc); ^{13}C NMR (100MHz, $CDCl_3$): δ 182.3 (C), 157.8 (C), 157.7 (C), 136.7 (C), 128.9 (CH), 128.6 (CH), 128.5 (CH), 80.7 (C), 67.4 (CH_2), 40.7 (CH_2), 40.1 (CH_2), 38.2 (CH_2), 37.3 (CH_2), 28.8 (CH_3); LRMS (ES^+) *m/z* 397.2 $[M+H]^+$, 419.2 $[M+Na]^+$
Data agrees with literature results.[168]

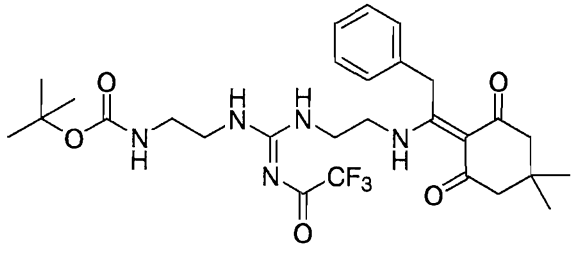
Benzyl {2-[*N*-(2-*tert*-butoxycarbonylamino-ethyl)-*N*'-(2,2,2-trifluoroacetyl)guanidinio]-propyl}carbamate (106)



According to a procedure by Kilburn *et al.*[168]

Methyl iodide (1.3ml, 20.5mmol) was added to a solution of thiourea **104** (813.9mg, 2.05mmol) in acetone (30ml) and the mixture stirred for 2 hours. The solvent and all volatile components were then removed *in vacuo* to give a white foam which was redissolved in 1:1 MeOH/DCM (40ml). Ammonium hexafluorophosphate (350mg, 2.15mmol) was then added and the resulting solution stirred for 18 hours. The solvents were removed *in vacuo* to give a white foam which was redissolved in DCM and washed once with water. The organic phase was dried over $MgSO_4$ and evaporated to give a white foam. This foam was redissolved in 4:1 toluene/$CHCl_3$ (100ml). Trifluoroacetamide (1.16g, 10.25mmol) and DBU (0.77ml, 5.13mmol) were added and the mixture heated at reflux for 18 hours. The solution was then evaporated and the residue purified by column chromatography (eluent 50% EtOAc/petrol gradient to 100% EtOAc) to give the product guanidine **106** as a white, hygroscopic foam (496.9mg, 51%); IR ν_{max} (cm^{-1}): 3319 (w), 1689 (m), 1626 (s), 1520 (m), 1238 (s), 1193 (s), 1139 (s), 843 (s), 733 (s); ^1H NMR: (300MHz, $CDCl_3$) δ 9.25 (br s, 1H, N*H*), 7.19 (m, 5H, arom), 7.00 (br s, 1H, N*H*), 6.03 (br s, 1H, N*H*), 5.42 (br s, 1H, N*H*), 4.96 (s, 2H, CH_2Ph), 3.58-2.96 (m, 8H, CH_2 × 4), 1.28 (s, 9H, Boc); ^{13}C NMR: (75MHz, $CDCl_3$) δ 165.9 (q, J = 35Hz, *C*OCF_3), 160.9 (C), 157.4 (C), 156.4 (C), 136.1 (C), 128.5 (CH), 128.1 (CH), 127.8 (CH), 117.0 (q, J = 285Hz, *C*F_3), 80.4 (C), 67.0 (CH_2), 45.0 (CH_2), 42.1 (CH_2), 36.9 (CH_2), 36.0 (CH_2), 28.1 (CH_3); LRMS: (ES^+) *m/z* 476.4 $[M+H]^+$, 498.4 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{20}H_{28}N_5O_5F_3$ $[M+Na]^+$ 498.1934, found 498.1934

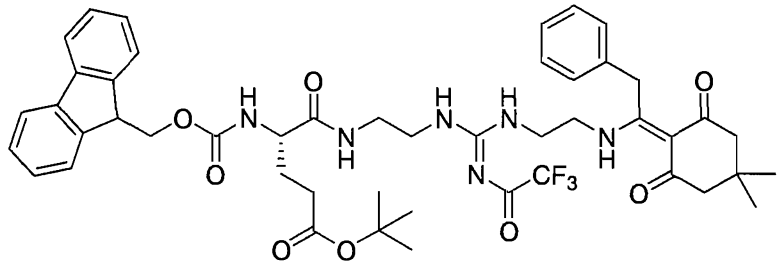
tert-Butyl {2-[*N*'-{3-[1-(2,6-dioxocyclohexylidene)-2-phenylethylamino]ethyl}-*N*''-(2,2,2-trifluoroacetylguanidinio]-ethyl}carbamate (107)



Carbamate **106** (525.6mg, 1.11mmol) was dissolved in DMF (15ml). Palladium on carbon (10% by weight, 156mg) was added and the mixture stirred under a hydrogen atmosphere for 3 hours. The reaction mixture was filtered through a plug of Celite and the palladium rinsed with DMF. Diketone **108** (571mg, 2.22mmol) was added to the filtrates and the mixture stirred for 2 days. The reaction mixture was evaporated and the residue purified by column chromatography (eluent gradient 40-70% EtOAc/petrol) to give the product **107** as a white hygroscopic foam (170.3mg, 27%); IR ν_{max} (cm^{-1}): 3271 (br w), 2957 (br w), 1790 (w), 1631 (m), 1566 (s), 1451 (m), 1137 (s), 732 (m); ^{1}H NMR (400MHz, d_3-MeCN): δ 13.70 (br s, 1H, DdpeN*H*), 9.62 (br s, 1H, N*H*), 7.52 (br s, 1H, N*H*), 7.28-7.11 (m, 5H, arom), 5.12 (br s, 1H, BocN*H*), 4.57 (s, 2H, CH_2Ph), 3.65-3.15 (m, 8H, CH_2 × 4), 2.37 (s, 4H, CH_2 × 2, Ddpe), 1.42 (s, 9H, Boc), 1.03 (s, 6H, CH_3 × 2, Ddpe); ^{13}C NMR is highly complex and has not been assigned; LRMS: (ES^+) *m/z* 582.5 $[M+H]^+$, 604.5 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{28}H_{38}N_5O_5F_3$ $[M+Na]^+$ 604.2717, found 604.2721.

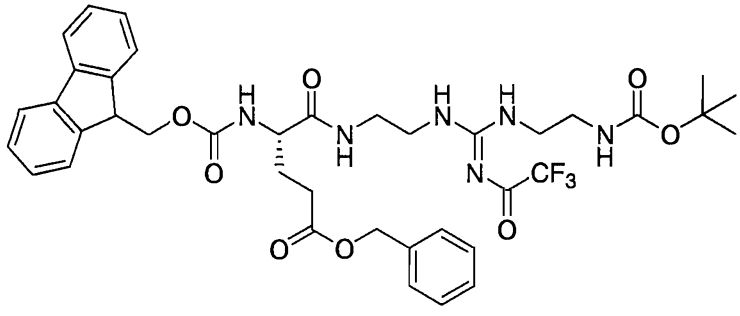
Data agrees with literature values.[168]

tert-Butyl 4-{2-[N'-(2-_tert_-butoxycarbonylamino-ethyl)-N''-(2,2,2-trifluoro-acetyl)-guanidino]-ethylcarbamoyl}-4-(9H-fluoren-9-ylmethoxycarbonylamino)-butyrate (109)



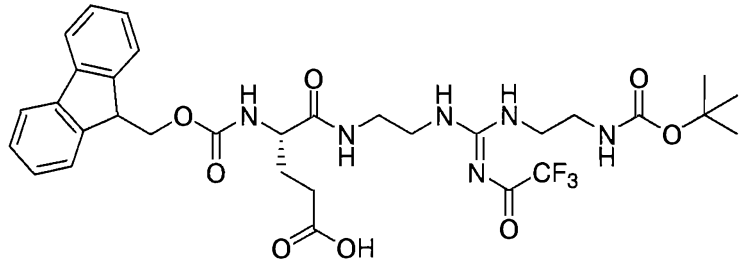
Carbamate **107** (139.1mg, 0.24mmol) was stirred in 20% TFA/DCM for 2 hours. The reaction mixture was rigourously evaporated *in vacuo* using toluene as an azeotrope. The residue was redissolved in DMF(6ml). *N*-α-Fmoc-glutamic acid-ω-*tert*-butyl ester (112mg, 0.26mmol), HOBt (97mg, 0.72mmol), DMAP (3mg, 24μmol) and DIPEA were dissolved in DCM (4ml) and EDC (51mg, 0.26mmol) was added. This DCM solution was added to the DMF solution above and the entire mixture stirred for 18 hours. The reaction mixture was then evaporated to dryness and the residue purified by column chromatography (eluent gradient 50-80% EtOAc/petrol) to give the product ester **109** as an off-white hygroscopic foam (168.0mg, 79%); IR ν_{max} (cm^{-1}) 3285 (br w), 2957 (w), 2358 (w), 1722 (m), 1632 (m), 1567 (s), 1450 (w), 1229 (m), 1140 (s), 738 (m); ^{1}H NMR (400MHz, d_3-MeCN) δ 13.51 (br s, 1H, DdpeN*H*), 9.53 (br s, 1H, NH), 7.68 (d, 2H, *J* = 7.5Hz, arom) 7.48 (d, 2H, *J* = 7.0Hz), 7.33-7.08 (m, 11H, arom + N*H* × 2), 6.10 (br s, 1H, N*H*), 4.54-4.40 (m, 2H, Fmoc C*H* + α-C*H*), 4.29 (s, 2H, Fmoc CH_2), 4.12 (t, 2H, *J* = 7.0Hz, Glu CH_2COO^tBu), 3.54-3.17 (m, 8H, CH_2 × 4), 2.29 (s, 4H, Ddpe CH_2 × 2) 2.29 (m, 1H, Glu β-C*H*H'), 1.97 (s, 2H, CH_2Ph), 1.87 (m, 1H, Glu β-CH*H*'), 1.37 (s, 9H, ^{t}Bu), 0.92 (s, 6H, Ddpe CH_3 × 2) ; ^{13}C NMR (100MHz, d_3-MeCN) is highly complex, apparently due to many rotameric forms in solution; LRMS: (ES^+) *m/z* 889.5 $[M+H]^+$, 911.4 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{47}H_{55}N_6O_8F_3$ $[M+Na]^+$ 911.3925, found 911.3915.

Benzyl 4-{2-[N'-(2-*tert*-butoxycarbonylamino-ethyl)-N''-(2,2,2-trifluoro-acetyl)-guanidino]-ethylcarbamoyl}-4-(9H-fluoren-9-ylmethoxycarbonylamino)-butyrate (111)



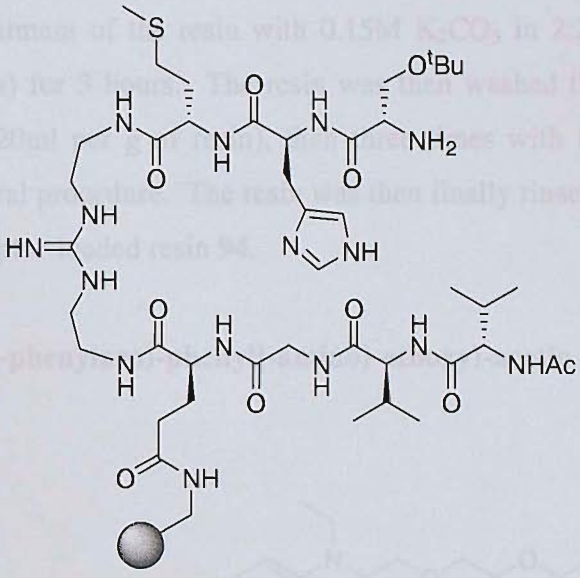
Carbamate **106** (224 mg, 0.47 mmol) was dissolved in DMF (10ml). Pd/C (10% by wt., 10 mol%, 50mg) was added and the mixture stirred under hydrogen gas for 16 hours. The mixture was then filtered through Celite and the filtrates evaporated to give a pale yellow oil. The oil was redissolved in DMF (10ml) and *N*-α-Fmoc-L-glutamic acid-ω-benzyl ester (238mg, 0.52mmol), HOBt (127mg, 0.94mmol), DIPEA (410µl, 2.36mmol), DMAP (6mg, 47µmol) and EDC (99mg, 0.52mmol) were added. The mixture was stirred for 18 hours. Evaporation of all solvent and purification of the residue by column chromatography (eluent 50% EtOAc/petrol) afforded ester **111** as a white hygroscopic foam (73 mg, 20 %): I.R. (neat) ν_{max} = 1629 (m), 1522 (m), 1449 (m), 1242 (m), 1166 (m), 1140 (m), 848 (w), 739 (m), 514 (s) cm^{-1}; ^1H NMR (300 MHz, $CDCl_3$) δ 9.65 (br s, 1H, N*H*), 7.76 (d, 2H, *J* = 7.5 Hz, arom), 7.57 (d, 2H, *J* = 7.0Hz, arom), 7.42-7.16 (m, 9H, arom), 7.16 (br s, 1H, N*H*), 7.01 (br s, 1H, N*H*), 5.76 (br s, 1H, N*H*), 5.12 (s, 2H, CH_2Ph), 4.45-4.30 (m, 2H, H_9 Fmoc and C*H* Glu), 4.19 (m, 2H, CH_2 Fmoc), 3.60-3.15 (m, 8H, CH_2CH_2), 2.49 (m, 2H, CH_2COO Glu), 2.13 (m, 1H, CHC*H*H' Glu), 2.01 (m, 1H, CHCH*H'* Glu), 1.40 (s, 9H, Boc) ; ^{13}C NMR (75 MHz, $CDCl_3$) δ 173.3 (C), 173.2 (C), 162.7 (C), 157.6 (C), 157.0 (C), 156.6 (C), 143.8 (C), 141.4 (C), 135.7 (C), 128.7 (CH), 128.4 (CH), 128.3 (CH), 127.9 (CH), 127.2 (CH), 125.3 (CH), 120.1 (CH), 116.7 (C), 80.6 (C), 67.3 (CH_2), 66.8 (CH_2), 54.8 (CH), 47.1 (CH_2), 41.9 (CH_2), 41.0 (CH_2), 39.6 (CH_2), 38.9 (CH_2), 30.4 (CH_2), 28.4 (CH_3); LRMS: (ES^+) *m/z* 783.4 $[M+H]^+$, 805.4 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{39}H_{45}N_6O_8F_3$ $[M+H]^+$ 783.3324, found 783.3332.

4-{2-[N'-(2-tert-Butoxycarbonylamino-ethyl)-N''-(2,2,2-trifluoro-acetyl)-guanidino]-ethylcarbamoyl}-4-(9H-fluoren-9-ylmethoxycarbonylamino)-butyric acid (112)



Ester **111** (55mg, 70μmol) was dissolved in EtOH (5ml) and Pd/C was added (10% by wt., 10mol%, 8mg). The mixture was stirred under hydrogen for 2 hours. The mixture was then filtered through Celite and the solvent removed by evaporation under reduced pressure to afford **112** as an off-white hygroscopic foam (43mg, 88%): I.R. (neat) ν_{max} = 2977 (br w), 1628 (s), 1524 (m), 1449 (m), 1243 (m), 1139 (s), 909 (m), 844 (m), 734 (s) cm^{-1}; ^1H NMR (300 MHz, $CDCl_3$) δ 9.47 (br s, 1H, N*H*), 7.74 (d, 2H, *J* = 7.5 Hz, arom), 7.55 (d, 2H, *J* = 6.5Hz, arom), 7.44 (br s, 1H, N*H*), 7.42 – 7.22 (m, 4H, arom), 6.18 (br s, 1H, N*H*), 5.88 (br s, 1H, N*H*), 4.45-4.38 (m, 3H, CH_2 Fmoc and C*H* Glu), 4.17 (m, 1H, C*H* Fmoc), 3.70-2.95 (m, 8H, CH_2CH$_2$), 2.44 (m, 2H, CH_2COO Glu), 2.18-1.82 (m, 2H, CHCH_2 Glu), 1.40 (s, 9H, Boc); ^{13}C NMR (75 MHz, $CDCl_3$) δ 180.1 (C), 176.2 (C), 173.7 (C), 171.7 (C), 159.9 (C), 157.8 (C), 156.7 (C), 143.7 (C), 141.4 (C), 127.9 (CH), 127.3 (CH), 125.3 (CH), 120.1 (CH), 80.7 (C), 67.5 (CH_2), 55.7 (CH), 54.6 (CH), 47.1 (CH_2), 41.9 (CH_2), 41.1 (CH_2), 39.6 (CH_2), 38.7 (CH_2), 30.0 (CH_2), 28.4 (CH_3); LRMS: (ES^+) *m/z* 693.3 $[M+H]^+$, 715.3 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{32}H_{39}N_6O_8F_3$ $[M+H]^+$ 693.2854, found 693.2852.

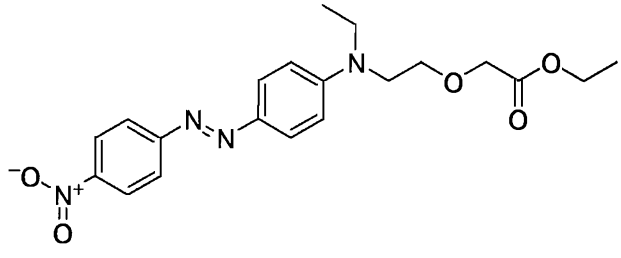
Solid-phase-linked tweezer-receptor (94)



Tentagel resin **88** (0.2mmol/g, 162.2mg, 32.4μmol) was swollen according to general procedure. A solution of acid **112** (34mg, 48.7μmol), DIC (15μl, 97.3μmol), HOBt (13mg, 97.3μmol) and DIPEA (17μl, 97.3μmol) in DMF (2ml) was added to the resin and the mixture shaken for 18 hours. Any remaining amine residues were capped by treating the resin with an excess of acetic anhydride. The resin was filtered and washed according to general procedure. A qualitative ninhydrin test was negative. There followed three sequential cycles of standard amino acid coupling/Fmoc-deprotection (according to general procedure) to add Fmoc-Gly, Fmoc-Val and Fmoc-Val to the resin (97.3μmol of Fmoc-AA, DIC, HOBt and DIPEA in 2ml DMF). Ninhydrin tests were used to monitor progress. After deprotection of the second Val-residue, the chain was capped with acetyl by treatment of the resin with acetic anhydride (9μl, 97.3μmol) and DIPEA (17μl, 97.3μmol) in DMF (2ml). Boc-deprotection was then carried out according to the general procedure. A ninhydrin test was positive. There followed three sequential cycles of standard amino acid coupling/Fmoc-deprotection (according to general procedure) to add, in order, Fmoc-Met, Fmoc-His(Trt) and Fmoc-Ser(O^{t}Bu) to the resin (97.3μmol of Fmoc-AA, DIC, HOBt and DIPEA in 2ml DMF). Ninhydrin tests were used to monitor progress. Side-chain deprotection was then performed by treatment of the resin with 95%

TFA/DCM (2 × 1h), followed by washing with 50% DIPEA/DMF (~20ml per g of resin) and then washing according to general procedure. Deprotection of the trifluoroacetyl was then achieved by treatment of the resin with 0.15M K_2CO_3 in 2:2:1 MeOH/DMF/H_2O (~20ml per g of resin) for 3 hours. The resin was then washed three times with 2:2:1 MeOH/DMF/H_2O (~20ml per g of resin), then three times with DMF before washing according to the general procedure. The resin was then finally rinsed with Et_2O and dried *in vacuo* to yield receptor-loaded resin **94**.

(2-{Ethyl-[4-(4-nitro-phenylazo)-phenyl]-amino}-ethoxy)-acetic acid ethyl ester (116)



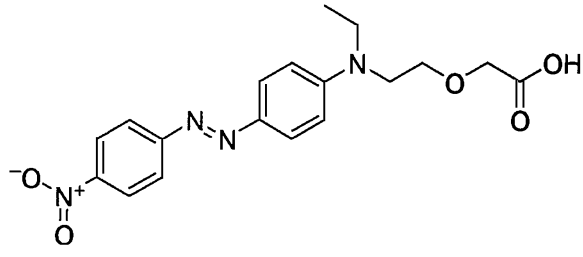
According to a procedure by Davis *et al.*[170]

To a solution of Disperse Red 1 (2g, 6.36mmol) and $Rh_2(OAc)_4$ (48mg, 17 mol%) in 1:1 DCM/ toluene (65ml) at 40°C was added a solution of ethyl diazoacetate (2.67ml, 25.4mmol) in toluene (13ml) over the course of one hour. The mixture was then stirred at room temp. for 24 hours. After evaporation of the solvent, the residue was purified by column (eluent 20% ethyl acetate/ petrol). Fractions containing the desired product were concentrated to a total volume of ~10ml. 100ml of hexane was then added and the solution refrigerated overnight. The resultant precipitate was collected by filtration and washed with hexane to yield **116** as a red solid (1.13g, 44%); m.p.: 79-80°C (lit. value = 80°C)[170]; ^{1}H-NMR: (300 MHz, $CDCl_3$) δ 8.32 (2H, d, J = 9.0 Hz, arom), 8.01 (4H, m, arom), 6.66 (2H, d, J = 9.0 Hz, arom), 4.21 (2H, q, J = 7.0 Hz, OCH_2CH_3), 4.1 (2H, s, OCH_2CO), 3.80 (2H, t, J = 6.0 Hz, NCH_2CH_2O), 3.75 (2H, t, J = 6.0 Hz, NCH_2CH_2O), 3.62 (2H, q, J = 7.0 Hz, NCH_2CH_3), 1.29 (3H, t, J = 7.0 Hz, OCH_2CH_3), 1.25 (3H, t, J = 7.0 Hz, NCH_2CH_3); ^{13}C-NMR: (75 MHz, $CDCl_3$) δ 170.3 (C), 157.0 (C), 151.5 (C),

147.4 (C), 143.8 (C), 126.4 (CH), 124.9 (H), 122.69 (CH), 111.5 (CH), 69.2 (CH_2), 68.8 (CH_2), 61.1 (CH_2), 50.3 (CH_2), 46.1 (CH_2), 14.3 (CH_3), 12.3 (CH_3); LRMS: (ES^+) *m/z* 401.3 $[M+H]^+$.

Data agrees with literature values.[170]

2-(2-ethyl-4-[(*E*)-2-(4-nitrophenyl)-1-diazenyl]anilinoethoxy)acetic acid (117)

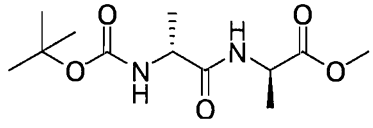


According to a procedure by Davis *et al.*[170]

LiOH(aq) (1M, 110ml) was added to a solution of ester **116** (1.109g, 2.77mmol) in dioxane (60ml). The mixture was stirred for 18 hours. The reaction mixture was then concentrated *in vacuo*, acidified to ~pH 4 with 1M $KHSO_4$(aq) and extracted with DCM. The organic phase was dried over $MgSO_4$ and evaporated to give **117** as a deep red solid (0.912g, 88%); m.p.: 160-161°C (lit. value 161°C)[183]; ^{1}H NMR: (300 MHz, DMSO-d_6) δ 8.37 (2H, d, J = 9 Hz, arom), 7.92 (2H, d, J = 9 Hz, arom), 7.82 (2H, d, J = 9 Hz), 6.89 (2H, d, J = 9 Hz), 4.06 (2H, s, O*CH*$_2$COOH), 3.67 (4H, m, N*CH*$_2$*CH*$_2$O), 3.54 (2H, q, J = 7 Hz, N*CH*$_2$CH3), 1.16 (3H, t, J = 7 Hz, NCH$_2$*CH*$_3$); ^{13}C NMR: (75 MHz, DMSO-d_6) δ 171.7 (C), 156.3 (C), 151.8 (C), 146.9 (C), 142.7 (C), 126.2 (CH), 125.1 (CH), 122.5 (CH), 111.7 (CH), 68.3 (CH_2), 67.8 (CH_2), 49.6 (CH_2), 45.3 (CH_2), 12.1 (CH_3); LRMS: (ES^+) *m/z* 373.4 $[M+H]^+$.

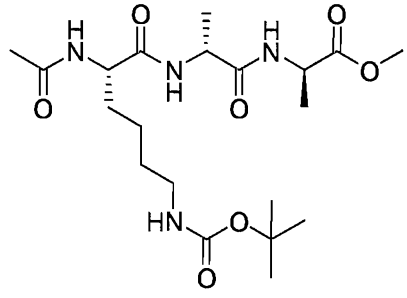
Data agrees with literature values.[170]

Methyl (2*R*)-2({(2*R*)-2[(*tert*-butoxycarbonyl)amino]-propanoyl}amino)propanoate (119)



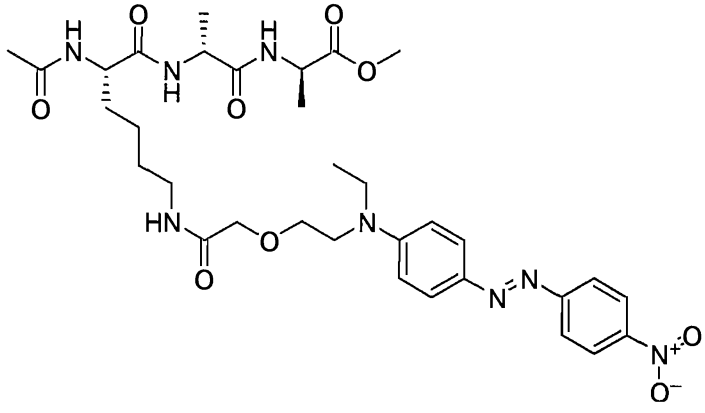
D-Alanine (473mg, 5.31mmol) was dissolved in MeOH (30ml) and thionyl chloride (2ml) was added. The mixture was heated at reflux for 3 hours and then concentrated under reduced pressure to give a yellow oil. This oil was dissolved in chloroform (30ml) and added dropwise to a mixture of Boc-D-Alanine (1.00g, 5.31mmol), HOBt (0.86g, 6.37mmol), EDC (1.06g, 5.58mmol) and triethylamine (2.76ml, 15.93mmol) in chloroform (40ml) and the entire mixture stirred for 18 hours. The reaction mixture was then concentrated under reduced pressure. The residues were taken up in DCM and washed with sat. $NaHCO_3$ (aq), 1M $KHSO_4$ (aq), water and sat. NaCl (aq). The organic phase was dried over $MgSO_4$ and evaporated to give a colourless oil which was purified by column chromatography (eluent 40% EtOAc/petrol) to give **119** as a white solid (0.778g, 53%); m.p. 111-112°C (lit. value: 109-110°C)[185]; IR ν_{max} (cm^{-1}): 3311 (m), 3263 (m), 2981 (w), 1739 (m), 1680 (s), 1654 (s), 1521 (s), 1252 (s), 1165 (s), 1071 (s), 1024 (m), 858 (w); ^1H NMR: (300MHz, $CDCl_3$) δ 6.97 (1H, m, N*H*), 5.33 (1H, m, N*H*), 4.50 (1H, dq, J = 7.4, 7.4Hz, α-C*H*), 4.19 (1H, m, α-C*H*), 3.68 (3H, s, OCH_3), 1.38 (9H, s, Boc), 1.33 (3H, d, J = 7.4Hz, CHCH_3), 1.30 (3H, d, J = 6.6Hz, CHCH_3); ^{13}C NMR: (75MHz, $CDCl_3$) δ 173.3 (C), 172.6 (C), 155.5 (C), 80.0 (C), 52.4 (CH_3), 49.9 (CH), 48.0 (CH), 28.3 (CH_3), 18.5 (CH_3), 18.1 (CH_3); LRMS: (ES^+) *m/z* 297.3 $[M+Na]^+$ 571.4 $[2M+Na]^+$; HRMS: (ES^+) calculated for $C_{12}H_{22}N_2O_5$ $[M+Na]^+$ 297.1421, found 297.1421. Data according to literature values.[185]

Methyl (2*R*)-2-{[(2*R*)-2-({(2*S*)-2-(acetylamino)-6-[(*tert*-butoxycarbonyl)amino]hexanoyl}amino)-propanoyl]amino}propanoate (120)



Dipeptide **119** (93mg, 0.34mmol) was stirred in 20% TFA/DCM (5ml) for 30 mins before being evaporated to dryness using toluene as an azeotrope. The oily residue was taken up in chloroform (20ml) containing a drop of DMF. Triethylamine (94µl, 0.68mmol) was added and the resulting mixture added to a mixture of Ac-Lys(Boc)-OH (98mg, 0.34mmol), HOBt (92mg, 0.68mmol), EDC (68mg, 0.36mmol) and triethylamine (47µl, 0.34mmol) in chloroform (20ml) at 0°C. The whole was stirred for 18 hours and then concentrated under reduced pressure. The residues were taken up in DCM and washed with sat. $NaHCO_3$ (aq), 1M $KHSO_4$ (aq), water and sat. NaCl (aq). The organic phase was dried over $MgSO_4$ and evaporated to give a white solid. This solid was dissolved in DCM and a white precipitate was crashed out using Et_2O and collected by filtration; **120** (41.5mg, 27%); m.p. 177-180°C; IR: ν_{max} (cm^{-1}) 3307 (m), 2942 (w), 2360 (w), 1681 (s), 1629 (s), 1536 (s), 1168 (s); ^{1}H NMR: (400MHz, $CDCl_3$) δ 7.20 (1H, d, J = 7.0Hz, N*H* [Ala]), 7.10 (1H, d, J = 7.0Hz, N*H* [Ala]), 6.79 (1H, d, J = 7.5Hz, N*H* [Lys]), 4.79 (1H, br s, N*H*Boc), 4.51 (2H, m, α-C*H* [Ala]), 4.39 (1H, m, α-C*H* [Lys]), 3.72 (3H, s, OCH_3), 3.08 (2H, m, ε-CH_2 [Lys]), 2.01 (3H, s, CH_3CO), 1.80 (1H, m, β-C*H*H' [Lys]), 1.67 (1H, m, β-CH*H*' [Lys]), 1.42 (9H, s, Boc), 1.38 (4H, d, J = 7.0Hz, CH_3CH [Ala]), 1.52-1.35 (4H, m, γ,δ-CH_2 [Lys]); ^{13}C NMR: (100MHz, $CDCl_3$) δ 173.6 (C), 172.2 (C), 171.4 (C), 171.0 (C), 156.1 (C), 79.7 (C), 53.8 (CH), 52.9 (CH_3), 49.3 (CH), 48.5 (CH), 40.4 (CH_2), 32.0 (CH_2), 30.0 (CH_2), 28.8 (CH_3), 23.4 (CH_3), 22.9 (CH_2), 18.6 (CH_3), 18.4 (CH_3); LRMS: (ES+) *m/z* 467.3 $[M+Na]^+$ 911.5 $[2M+Na]^+$; HRMS: (ES^+) calculated for $C_{20}H_{36}N_4O_7$ $[M+Na]^+$ 467.2476, found 467.2470.

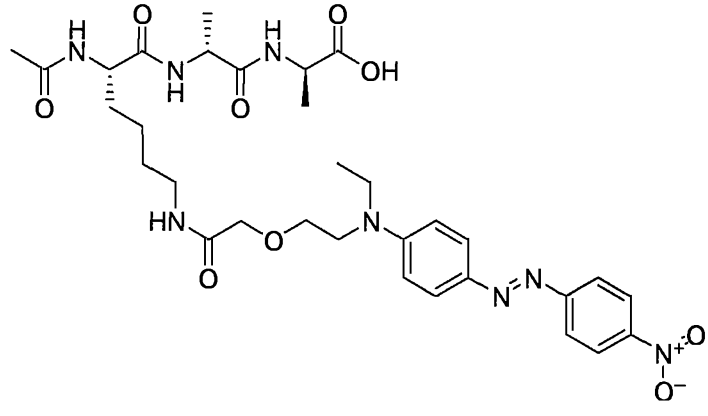
Methyl (2*R*)-2-((2*R*)-2-[((2*S*)-2-(acetylamino)-6-[2-(2-ethyl-4-[(*E*)-2-(4-nitrophenyl)-1-diazenyl]anilinoethoxy)acetyl]aminohexanoyl)amino]propanoylamino)propanoate (121)



Tripeptide **120** (39.3mg, 88.5μmol) was dissolved in 20% TFA/DCM and stirred for 3 hours. The mixture was evaporated using toluene as an azeotrope and the residues were taken up in chloroform (10ml) containing a drop of DMF and triethylamine was added (25μl, 176.9μmol). This mixture was added dropwise to a solution of dye acid **117** (36.2mg, 97.3μmol), HOBt (30mg, 176.9μmol), EDC (18mg, 92.9μmol) and triethylamine (13μl, 88.5μmol) in chloroform (10ml) at 0°C. The whole was stirred for 18 hours at room temperature. The reaction mixture was then concentrated under reduced pressure and the residues taken up in DCM. The organic layer was washed with sat. $NaHCO_3$ (aq), 1M $KHSO_4$ (aq), water and sat. NaCl (aq). The organic phase was separated and evaporated to give a red residue. The residue was taken up in MeOH and a red solid crashed out upon addition of Et_2O, which was collected by filtration; **121** (20.4mg, 33%); m.p.: 194-197°C; IR: ν_{max} (cm^{-1}) 3277 (m), 1743 (m), 1631 (s), 1513 (s), 1339 (s), 1141 (s), 857 (m), 824 (m); ^1H NMR: (400MHz, $CDCl_3$) δ 8.33 (2H, d, J = 9.0Hz, arom), 7.98 (4H, m, arom), 7.00 (2H, m, arom), 6.88 (1H, m, N*H*), 6.79 (1H, m, N*H*), 6.65 (2H, br m, N*H*), 4.51 (1H, dq, J = 7.3, 7.3Hz, α-C*H* Ala), 4.44 (1H, dq, J = 7.3, 7.3Hz, α-C*H* Ala), 4.27 (1H, m, α-C*H* Lys), 3.96 (1H, d, J = 15.3Hz, OC*H*H'CO), 3.92 (1H, d, J = 15.3Hz, OCH*H'*CO), 3.71 (3H, s, OCH_3), 3.80-3.65 (4H, br m, NCH_2CH_2O), 3.57 (2H, q, J = 7.0Hz, NCH_2CH_3), 3.26 (1H, m, ε-C*H*H'NH Lys), 3.14 (1H, m, ε-

CH*H'*NH Lys), 2.01 (3H, s, CH_3CO), 1.85-1.61 (2H, br m, β-CH_2 Lys), 1.39 (6H, d, J = 7.3Hz, CH_3CH Ala), 1.30 (3H, t, J = 7.0Hz, NCH_2CH_3), 1.46-1.25 (4H, br m, γ,δ-CH_2 Lys); ^{13}C NMR: (100MHz, $CDCl_3$) δ 173.5 (C), 172.0 (C), 171.5 (C), 169.6 (C), 169.3 (C), 159.4 (C), 131.3 (C), 130.2 (C), 129.3 (C), 128.0 (CH), 127.5 (CH), 125.2 (CH), 122.9 (CH), 71.3 (CH_2), 71.1 (CH_2), 53.9 (CH), 52.8 (CH_3), 51.6 (CH), 51.2 (CH), 49.5 (CH_2), 48.5 (CH_2), 38.2 (CH_3), 31.1 (CH_2), 29.3 (CH_2), 23.4 (CH_3), 22.8 (CH_2), 18.5 (CH_3), 18.3 (CH_3), 12.5 (CH_3); LRMS: (ES^+) *m/z* 699.3 $[M+H]^+$, 721.3 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{33}H_{46}N_8O_9$ $[M+Na]^+$ 721.3280, found 721.3283.

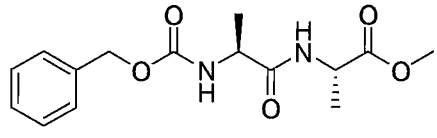
(2*R*)-2-((2*R*)-2-[((2*S*)-2-(Acetylamino)-6-[2-(2-ethyl-4-[(*E*)-2-(4-nitrophenyl)-1-diazenyl]anilinoethoxy)acetyl]aminohexanoyl)amino]propanoylamino)propanoic acid (122)



LiOH(aq) (1M, 3.2ml, 3.20mmol) was added to a solution of dye-labelled ester **121** (223mg, 0.32mmol) in dioxane (20ml) and the mixture stirred for 18 hours. The reaction mixture was evaporated *in vacuo* and the highly coloured residue taken up in distilled water. The aqueous phase was acidified with conc. HCl and repeatedly extracted with DCM and EtOAc. All organic phases were combined, dried over $MgSO_4$ and evaporated to give a highly coloured solid, which was purified by column chromatography (eluent gradient 1-20% MeOH/DCM) to give **122** as a deep red solid (23.2mg, 11%); m.p. 201-204°C, IR ν_{max} (cm^{-1}): 3273 (br w), 2932 (br w), 1646 (m), 1601 (m), 1513 (m), 1340 (m), 1138 (m), 858 (m), 825 (m); ^1H NMR: (300MHz, d_6-DMSO) δ 8.36 (d, 2H, *J* =

9.0Hz, arom), 8.09 (m, 1H, N*H*), 8.05-7.98 (m, 2H, N*H*), 7.93 (d, 2H, *J* = 9.0Hz, arom), 7.84 (d, 2H, *J* = 9.0Hz, arom), 7.66 (m, 1H, N*H*), 6.92 (d, 2H, *J* = 9.0Hz, arom), 4.30-4.15 (m, 2H, α-C*H* × 2), 4.04 (m, 1H, α-C*H*), 3.89 (s, 2H, OCH_2CO), 3.68 (m, 4H, NCH_2CH_2O), 3.56 (q, 2H, *J* = 7.0Hz, NCH_2CH_3), 3.05 (dt, 2H, *J* = 7.0, 6.0 Hz, Lys $NHCH_2CH_2$), 1.83 (s, 3H, Ac), 1.60-1.45 (m, 2H, Lys β-CH_2), 1.37 (m, 2H, Lys CH_2), 1.25-1.10 (m, 11H, CH_3 × 3 + Lys CH_2); ^{13}C NMR: (75MHz, d_6-DMSO) δ 171.6 (C), 171.3 (C), 169.4 (C), 168.7 (C), 168.7 (C), 156.2 (C), 151.7 (C), 146.8 (C), 142.7 (C), 126.1 (CH), 124.9 (CH), 122.4 (CH), 111.6 (CH), 70.1 (CH_2), 68.5 (CH_2), 52.7 (CH), 49.2 (CH_2), 48.1 (CH), 47.8 (CH), 45.1 (CH_2), 37.9 (CH_2), 31.5 (CH_2), 28.9 (CH_2), 22.7 (CH_2), 22.4 (CH_3), 18.1 (CH_3), 17.6 (CH_3), 12.0 (CH_3); LRMS: (ES^+) *m/z* 707.4 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{32}H_{44}N_8O_9$ $[M+Na]^+$ 707.3123, found 707.3132.

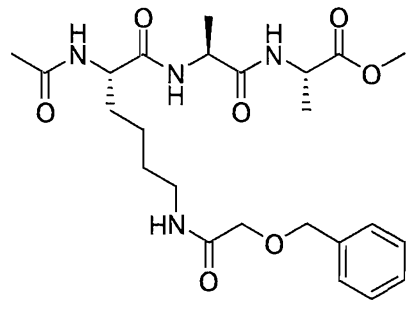
Methyl (2*S*)-2-[((2*S*)-2-[(benzyloxy)carbonyl]aminopropanoyl)amino]propanoate (124)



Cbz-L-alanine (456.1mg, 2.04mmol), L-alanine methyl ester hydrochloride (0.2g, 1.94mmol), HOBt (828mg, 6.13mmol) and DIPEA (1.1ml, 6.13mmol) were dissolved in $CHCl_3$ (10ml). EDC (431mg, 2.25mmol) was added and the mixture stirred for 2 hours. The reaction mixture was then sequentially washed with saturated $NaHCO_3$(aq), 1M $KHSO_4$(aq), water and brine. The organic phase was dried over $MgSO_4$ and evaporated to give dipeptide **124** as a white solid (367.9mg, 58%); m.p. 100-105°C (lit. value: 105°C)[186]; ^1H NMR (300MHz, $CDCl_3$) δ 7.39-7.28 (m, 5H, arom), 6.54 (m, 1H, N*H*), 5.34 (m, 1H, N*H*), 5.11 (s, 2H, CH_2Ph), 4.56 (dq, 1H, *J* = 7.5, 7.0Hz, α-CH), 4.23 (m, 1H, α-CH), 3.75 (s, 3H, OMe), 1.39 (d, 3H, *J* = 7.0Hz), 1.37 (d, 3H, *J* = 7.0Hz); ^{13}C NMR: (75MHz, $CDCl_3$) δ 173.3 (C), 171.8 (C), 155.8 (C), 135.6 (C), 128.7 (CH), 128.4 (CH), 128.2 (CH), 67.2 (CH_2), 52.7 (CH_3), 50.5 (CH), 48.2 (CH), 18.9 (CH_3), 18.7 (CH_3); LRMS: (ES^+) 331.2 $[M+Na]^+$.

Data agrees with literature values.[187]

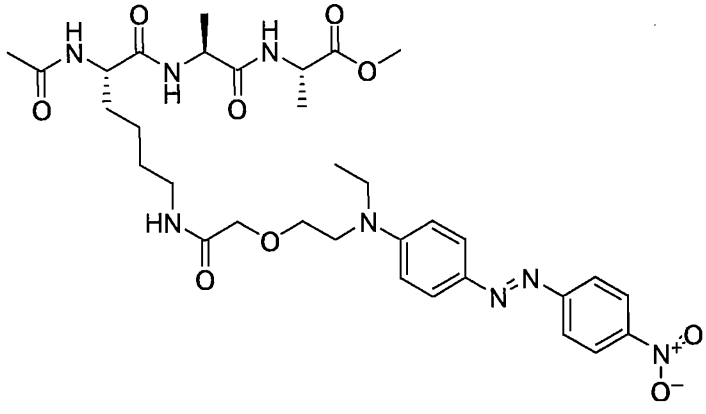
Methyl (2*S*)-2-((2*S*)-2-[((2*S*)-2-(acetylamino)-6-[(benzyloxy)carbonyl]-aminohexanoyl)amino]propanoylamino)propanoate (125)



Palladium on carbon (10% by weight, 71mg) was added to a solution of dipeptide **124** (205.2mg, 0.67mmol) in MeOH (4ml). The mixture was stirred under a hydrogen atmosphere for 2 hours, and then filtered through a Celite plug. The palladium was rinsed with MeOH and the combined filtrates evaporated. The residue was redissolved in $CHCl_3$ (10ml) containing a few drops of DMF and added to a solution of *N*-α-acetyl-lysine-*N*-ω-benzyl carbamate (225mg, 0.70mmol), HOBt (270mg, 2.00mmol), DIPEA (348μl, 2.00mmol) and EDC (140mg, 0.73mmol) in $CHCl_3$ (10ml). The mixture was stirred for 16 hours and then evaporated to dryness *in vacuo*. The residue was redissolved in EtOAc and sequentially washed with saturated $NaHCO_3$(aq), 1M $KHSO_4$(aq), water and brine. The organic phase was dried over $MgSO_4$ and evaporated to dryness. The residue was redissolved in MeOH and tripeptide **125** precipitated upon addition of Et_2O, and was collected by filtration. White solid (76.6mg, 24%); m.p.: 183-185°C; IR ν_{max} (cm^{-1}): 3285 (m), 1687 (m), 1629 (s), 1537 (s), 1453 (m), 1372 (m), 1268 (m), 1212 (m), 694 (m); ^{1}H NMR: (400MHz, $CDCl_3$) δ 7.33 (m, 5H, arom), 7.05 (d, 1H, *J* = 6.0Hz, N*H*), 6.97 (d, 1H, *J* = 7.5Hz, N*H*), 6.67 (m, 1H, N*H*), 5.21 (m, 1H, N*H*), 5.08 (s, 2H, CH_2Ph), 4.57-4.46 (m, 3H, α-C*H* × 3), 3.72 (s, 3H, OMe), 3.17 (m, 2H, Lys NHCH_2CH$_2$), 2.00 (s, 3H, Ac), 1.80 (m, 1H, Lys β-CHH'), 1.67 (m, 1H Lys β-CH*H'*), 1.50 (m, 2H, Lys CH_2), 1.42-1.34 (m, 2H, Lys CH_2), 1.37 (m, 6H, Ala CH_3 × 2); ^{13}C NMR: (100MHz, $CDCl_3$) δ 173.6 (C), 172.1 (C), 171.3 (C), 171.0 (C), 157.1 (C), 137.0 (C), 128.9 (CH), 128.5 (CH), 128.4 (CH), 67.0 (CH_2), 53.7 (CH_3), 53.5 (CH), 52.8 (CH), 49.3 (CH), 48.5 (CH_2), 40.7 (CH_2), 29.8 (CH_2), 23.4 (CH_2), 22.8 (CH_3), 18.5 (CH_3), 18.4

(CH_3); LRMS: (ES^+) *m/z* 501.2 $[M+Na]^+$, HRMS: (ES^+) calculated for $C_{23}H_{34}N_4O_7$ $[M+Na]^+$ 501.2320, found 501.2316

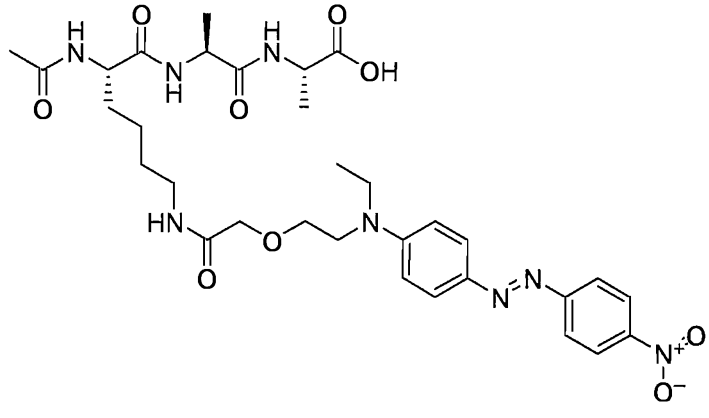
Methyl (2*S*)-2-((2*S*)-2-[((2*S*)-2-(acetylamino)-6-[2-(2-ethyl-4-[(*E*)-2-(4-nitrophenyl)-1-diazenyl]anilinoethoxy)acetyl]aminohexanoyl)amino]propanoylamino)propanoate (126)



Palladium on carbon (10% by weight, 12mg) was added to a solution of tripeptide **125** (54.5mg, 0.11mmol) in MeOH (5ml) and the mixture stirred under an atmosphere of hydrogen for 2 hours. The reaction mixture was then filtered through a plug of Celite and the palladium residue rinsed with MeOH. Combined filtrates were evaporated *in vacuo*. The residues were redissolved in $CHCl_3$ (3ml) and added to a solution of dye acid **117** (47mg, 0.13mmol), HOBt (46mg, 0.33mmol), DIPEA (60µl, 0.33mmol), DMAP (1mg) and EDC (24mg, 0.13mmol) in $CHCl_3$ (20ml). The mixture was stirred for 16 hours. The reaction mixture was sequentially washed with saturated $NaHCO_3$(aq), 1M $KHSO_4$(aq), water and brine. The organic phase was evaporated to dryness. The residue was redissolved in DCM and a red solid precipitated on addition of Et_2O, which was collected by filtration, **126** (66.5mg, 86%); m.p.: 171-172°C; IR ν_{max} (cm^{-1}): 3279 (w), 1633 (m), 1602 (m), 1513 (s), 1382 (m), 1336 (s), 1134 (s), 1106 (s), 857 (m); ^1H NMR: (400MHz, d_6-DMSO) δ 8.35 (d, 2H, *J* = 9.0Hz, arom), 8.17 (d, 1H, *J* = 7.0Hz, N*H*), 8.12 (m, 1H, N*H*), 7.99 (m, 1H, N*H*), 7.92 (d, 2H, *J* = 9.0Hz, arom), 7.83 (d, 2H, *J* = 9.0Hz, arom), 7.62 (m, 1H, N*H*), 6.91 (d, 2H, *J* = 9.0Hz, arom), 4.33-4.07 (m, 3H, α-C*H* × 3), 3.88 (s, 2H, OCH_2CO), 3.67 (m, 4H, OCH_2CH_2N), 3.60 (s, 3H, OMe), 3.55 (q, 2H, *J* =

7.0Hz, NCH_2CH_3), 3.05 (m, 2H, Lys $NHCH_2CH_2$), 1.83 (s, 3H, Ac), 1.58 (m, 1H, Lys β-C*H*H'), 1.47 (m, 1H, Lys β-CH*H*'), 1.39 (m, 2H, Lys CH_2), 1.28 (t, 3H, *J* = 7.0Hz, NCH_2CH_3), 1.22-1.13 (m, 8H, Ala $CH_3 \times 2$ + Lys CH_2); ^{13}C NMR: (100MHz, d_6-DMSO) δ 182.4 (C), 181.5 (C), 180.9 (C), 178.8 (C), 178.2 (C), 165.7 (C), 161.2 (C), 156.3 (C), 152.2 (C), 135.6 (CH), 134.4 (CH), 131.9 (CH), 121.1 (CH), 79.6 (CH_2), 78.0 (CH_2), 62.2 (CH_2), 62.0 (CH_2), 61.3 (CH_2), 58.7 (CH), 57.0 (CH), 54.6 (CH), 47.4 (CH_2), 41.0 (CH_2), 38.4 (CH_2), 31.9 (CH_3), 27.5 (CH_3), 26.2 (CH_3), 21.5 (CH_3); LRMS: (ES^+) *m/z* 699.3 $[M+H]^+$ 721.3 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{33}H_{46}N_8O_9$ $[M+Na]^+$ 721.3279, found 721.3268.

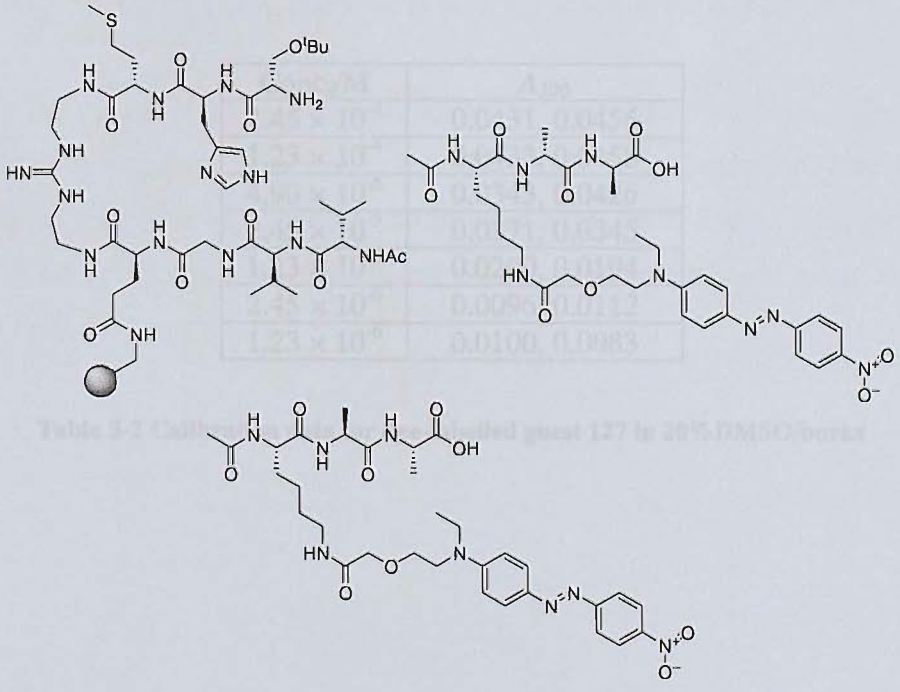
(2*S*)-2-((2*S*)-2-[((2*S*)-2-(Acetylamino)-6-[2-(2-ethyl-4-[(*E*)-2-(4-nitrophenyl)-1-diazenyl]anilinoethoxy)acetyl]aminohexanoyl)amino]propanoylamino)propanoic acid (127)



LiOH(aq) (1M, 741μl, 741μmol) was added to a solution of dye-labelled tripeptide **126** (51.8mg, 74.1μmol) in dioxane (3ml). The mixture was stirred for 4 hours and then evaporated *in vacuo*. The residue was taken up in DCM and washed with 1M $KHSO_4$(aq). The organic phase was evaporated and the residue redissolved in MeOH. Red solid precipitated upon addition of Et_2O, which was collected by filtration, **127** (3.7mg, 7%); m.p.: 200-203°C; IR ν_{max} (cm^{-1}): 3286 (br w), 2933 (br w), 1650 (s), 1600 (m), 1513 (m), 1138 (m), 858 (m); ^1H NMR: (400MHz, d_4-MeOH) δ 8.37 (d, 2H, *J* = 9.0Hz, arom), 7.98 (d, 2H, *J* = 9.0Hz, arom), 7.91 (d, 2H, *J* = 9.0Hz, arom), 6.96 (d, 2H,

J = 9.0Hz, arom) , 4.50-4.20 (m, 3H, α-C*H* × 3), 4.01 (s, 2H, OCH_2CO), 3.80 (m, 4H, NCH_2CH_2O), 3.65 (q, 3H, J = 7.0Hz, NCH_2CH_3), 3.21 (m, 2H, Lys NHCH_2CH_2), 2.00 (s, 3H, Ac), 1.78 (m, 1H, Lys β-C*H*H') 1.66 (m, 1H, Lys β-CH*H*'), 1.48 (m, 2H, Lys CH_2), 1.46-1.33 (m, 8H, Ala CH_3 × 2 + Lys CH_2), 1.29 (t, 3H, J = 7.0Hz, NCH_2CH_3); ^{13}C NMR: (100MHz, d_4-MeOH) δ 174.2 (C), 174.1 (C), 173.4 (C), 172.2 (C), 168.5 (C), 158.2 (C), 153.3 (C), 148.8 (C), 144.9 (C), 137.3 (CH), 125.7 (CH), 123.6 (CH), 112.8 (CH), 71.4 (CH_2), 70.5 (CH_2), 55.0 (CH_2), 54.8 (CH_2), 50.8 (CH), 50.2 (CH), 46.6 (CH), 39.6 (CH_2), 32.6 (CH_2), 32.3 (CH_2), 30.1 (CH_2), 24.1 (CH_3), 22.4 (CH_3), 18.3 (CH_3), 18.1 (CH_3); LRMS: (ES^+) *m/z* 707.4 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{32}H_{44}N_8O_9$ $[M+H]^+$ 685.3304, found 685.3285

Solid-phase binding assay: Solid-phase-linked tweezer receptor 94 vs. dye-labelled guests 122 and 127



Adapted from a procedure by Wennemers *et al.*[167]

Calibration curves were plotted for the dye labelled tripeptides **122** and the diastereomeric **127**, in 20% DMSO/borax buffer. Results are shown in Tables 5-1 and 5-

2. Logarithmic least squares regression gave the relationship between the absorption at 500nm A_{500} and the concentration c as;

$$c = e^{\frac{A_{500}-0.112}{0.0078}}$$

for both solutions. A plot of this curve is shown in Fig 5-1.

Conc./M	A_{500}
1.34×10^{-4}	0.0425, 0.0434
6.70×10^{-5}	0.0407
1.34×10^{-5}	0.0290, 0.0206
6.70×10^{-6}	0.0197, 0.0178
1.34×10^{-6}	0.0047, 0.0075
6.70×10^{-7}	0.0015, 0.0077

Table 5-1 Calibration data for dye-labelled guest 122 in 20%DMSO/borax

Conc./M	A_{500}
2.45×10^{-4}	0.0431, 0.0456
1.23×10^{-4}	0.0433, 0.0458
4.90×10^{-5}	0.0343, 0.0426
2.45×10^{-5}	0.0371, 0.0345
1.23×10^{-5}	0.0200, 0.0194
2.45×10^{-6}	0.0096, 0.0112
1.23×10^{-6}	0.0100, 0.0083

Table 5-2 Calibration data for dye-labelled guest 127 in 20%DMSO/borax

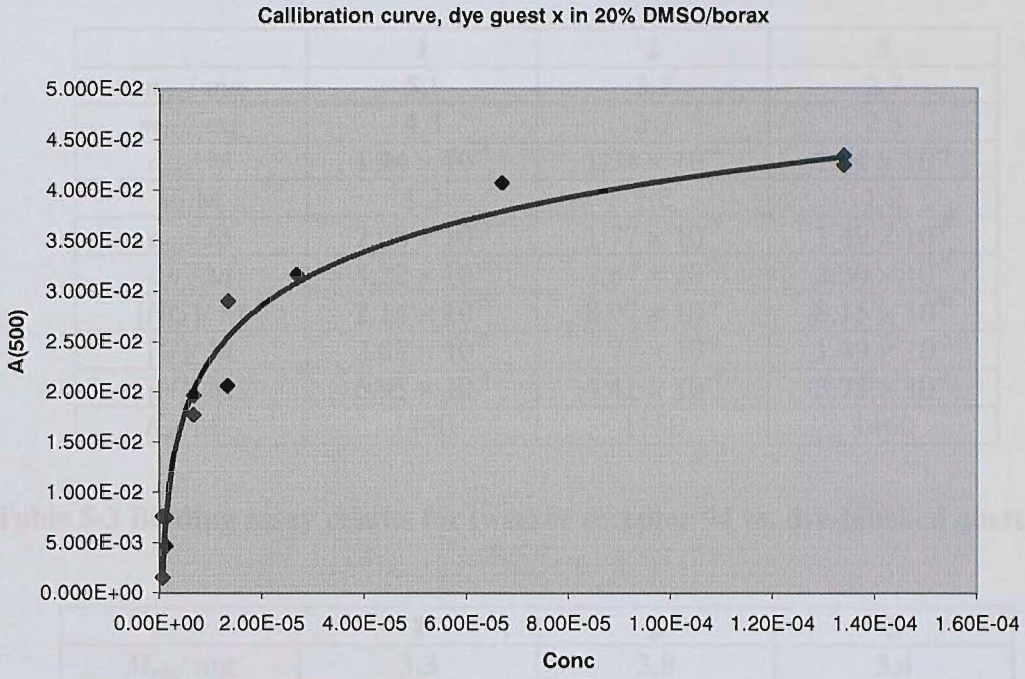


Fig 5-1 Calibration curve of guest 122 in 20% DMSO/borax

Incubation of a mass of **94**, m_{rec} and an equivalent mass of underivatised Tentagel m_{TG} with equivalent solutions of either guest **122** or **127** was carried out for 18 hours. Samples were then filtered through a cotton wool plug and UV absorption measured at 500nm. Binding constants were evaluated as described in the main text. Results are shown in Tables 5-3 and 5-4.

	1	2	3
m_{rec}/ mg	5.1	3.2	2.7
m_{TG}/ mg	4.3	2.7	2.3
c_{init}/ M	1.34×10^{-4}	1.34×10^{-4}	1.34×10^{-4}
V/ M	1.2	1.2	1.2
c_{rec}/ M	2.08×10^{-5}	1.77×10^{-5}	1.49×10^{-5}
c_{TG}/ M	4.22×10^{-5}	2.67×10^{-5}	2.30×10^{-5}
$[HG]$/ M	2.14×10^{-5}	8.97×10^{-6}	8.15×10^{-6}
$[G]$/ M	2.08×10^{-5}	1.77×10^{-5}	1.49×10^{-5}
$[H]$/ M	6.95×10^{-4}	4.41×10^{-4}	3.75×10^{-4}
K_a/ M^{-1}	1480	1150	1460

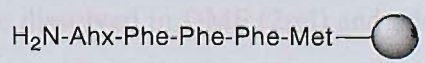
Table 5-3 Binding assay results for tweezer receptor 94 vs. dye-labelled guest 122

	1	2	3
M_{rec}/ mg	3.3	3.8	3.4
m_{TG}/ mg	2.8	3.2	2.9
c/ mol.dm^{-3}	1.23×10^{-4}	1.23×10^{-4}	4.90×10^{-5}
V/ ml	1.2	1.2	1.2
c_{rec}/ mol.dm^{-3}	1.93×10^{-5}	2.88×10^{-5}	5.21×10^{-6}
c_{TG}/ mol.dm^{-3}	2.58×10^{-5}	3.14×10^{-5}	6.33×10^{-6}
$[HG]$/ mol.dm^{-3}	6.51×10^{-6}	2.66×10^{-6}	1.12×10^{-6}
$[G]$/ mol.dm^{-3}	1.93×10^{-5}	2.88×10^{-5}	5.21×10^{-6}
$[H]$/ mol.dm^{-3}	4.60×10^{-4}	5.31×10^{-4}	4.82×10^{-4}
K_a/ mol^{-1}dm^3	734	174	446

Table 5-4 Binding assay results for tweezer receptor 94 vs. dye-labelled guest 127

5.7 Experimental for Chapter 3

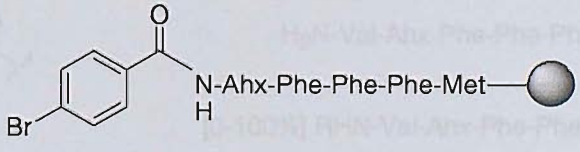
Resin construct 129



Tentagel resin **88** (0.2mmol/g, 641.9mg, 0.11mmol) was preswollen according to the general procedure. Coupling of Fmoc-Met-OH (122mg, 0.33mmol) using DIC (51µl,

0.33mmol), HOBt (44mg, 0.33mmol) and DIPEA (57μl, 0.33mmol) in DMF (5ml) was carried out according to the general procedure. The resin was washed according to general procedure and gave a negative ninhydrin test. An Fmoc-deprotection was carried out according to the general procedure to yield a resin which gave a positive ninhydrin test. Three cycles of standard amino acid coupling/Fmoc-deprotection were carried out according to the general procedure with the following reagent mixture; Fmoc-Phe-OH (127mg, 0.33mmol), DIC (51μl, 0.33mmol), HOBt (44mg, 0.33mmol) and DIPEA (57μl, 0.33mmol) in DMF (5ml). Washing according to the general procedure occurred after every step, and reaction progression was monitored using ninhydrin tests. A final standard amino acid coupling according to the general procedure using Fmoc-Ahx-OH (116mg, 0.33mmol), DIC (51μl, 0.33mmol), HOBt (44mg, 0.33mmol) and DIPEA DIPEA (57μl, 0.33mmol) in DMF (5ml) was carried out. Washing was carried out according to the general procedure and the resin was found to give a negative ninhydrin test. An Fmoc-deprotection was then carried out according to the general procedure, followed by washing according to the general procedure. The resin gave a positive ninhydrin test. The resin **129** was finally rinsed with Et_2O and dried *in vacuo*. Quantitative ninhydrin test gave the resin loading as 0.13mmol/g.

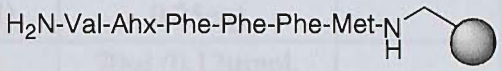
Loading of the capping group on to resin 129, giving 130



Resin **129** (25.2mg, 0.13mmol/g, 3.28μmol) was preswollen according to the general procedure. *para*-Bromobenzoic acid (2.0mg, 9.83μmol), HOBt (1.3mg, 9.83μmol), and DIC (1.5μl, 9.83μmol) were dissolved in DMF (2ml) and added to the resin with DIPEA (1.7μl, 9,83μmol) and the mixture shaken for 1 hour. The resin was then drained and washed according to the general procedure. A ninhydrin test gave a negative result. The resin was then rinsed with Et_2O and dried *in vacuo*. Single bead cleavage and analysis was carried out according to the general procedure.

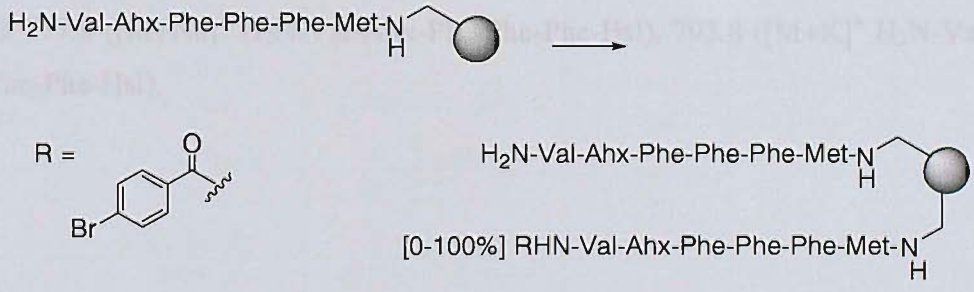
MALDI comments: Spectrum noisy. Cluster analysis extracted the following significant peaks; *m/z* 860.6/862.6 ($[M+Na]^+$ *p*-Br PhCONH-Ahx-Phe-Phe-Phe-Hsl), 876.7/878.7 ($[M+K]^+$ *p*-Br PhCONH-Ahx-Phe-Phe-Phe-Hsl).

Resin construct 131



Resin **129** (100mg, 0.13mmol/g, 13μmol) was preswollen according to general procedure, coupled to Val according to the general procedure [Fmoc-Val-OH (13.2mg, 39μmol), HOBt (5.3mg, 39μmol), DIC (6μl, 39μmol), DIPEA (6.8μl, 39μmol) in DMF (4ml)], washed according to the resin according to general procedure, and then Fmoc-deprotected according to the general procedure. All steps were monitored using the ninhydrin test. The resin was then rinsed with Et_2O and dried *in vacuo*.

Capping experiments at various capping loadings

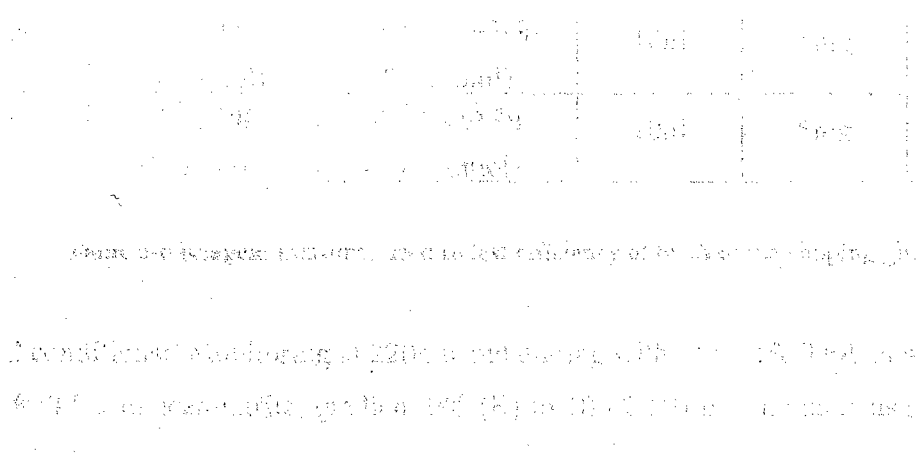


Five samples of resin **131** were each treated with a mixture of *p*-Br PhCOOH and excess DIC/HOBt/DIPEA as detailed in Table 5-5. Each sample was shaken for 18 hours before draining and washing as above. Single beads were chosen from each sample for cleavage and MALDI-TOF-MS analysis according to the general procedure.

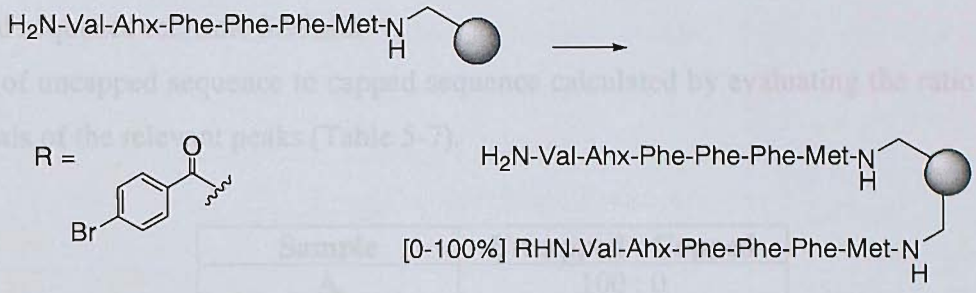
Sample	Mass resin x	*p*-Br PhCOOH (1.79mM/DMF)	DIC (excess)	HOBt (excess)	DIPEA (excess)
A	10.6mg (1.38μmol)	847μl (1.52μmol, 1.1eq)	10μl	2mg	10μl
B	10.6mg (1.38μmol)	385μl (0.69μmol, 0.5eq)	10μl	2mg	10μl
C	6.6mg (0.86μmol)	120μl (0.21μmol, 0.25eq)	10μl	2mg	10μl
D	9.6mg (1.25μmol)	70μl (0.12μmol, 0.1eq)	10μl	2mg	10μl

Table 5-5 Reagent mixtures used to test loading of the capping group

MALDI comments (single bead cleavage): All spectra were noisy. Cluster analysis extracted Br-doublets. Sample **A**: *m/z* 959.9/961.9 ($[M+Na]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 975.9/977.8 ($[M+K]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl); Samples **B**, **C**: *m/z* 777.8 ($[M+Na]^+$ H_2N-Val-Ahx-Phe-Phe-Phe-Hsl), 793.8 ($[M+K]^+$ H_2N-Val-Ahx-Phe-Phe-Phe-Hsl), 959.9/961.9 ($[M+Na]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 975.9/977.8 ($[M+K]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl); Sample **D**: *m/z* 777.8 ($[M+Na]^+$ H_2N-Val-Ahx-Phe-Phe-Phe-Hsl), 793.8 ($[M+K]^+$ H_2N-Val-Ahx-Phe-Phe-Phe-Hsl).



Evaluation of capping efficiency



Five resin samples of **131** were weighed and each was treated with a mixture of *p*-Br PhCOOH and excess DIC/HOBt/DIPEA as detailed in Table 5-6. Each sample was shaken for 18 hours before draining and washing according to the general procedure. The entire resin sample was then cleaved by shaking the resin with a solution of 100mg/ml CNBr in 50% TFA/water for 16 hours. Each sample was filtered, the filtrates evaporated to dryness and the residues dissolved in 1ml of acetonitrile and analysed by HPLC.

Sample	Mass resin	*p*-Br PhCOOH (0.45mM/DMF)	DIC (excess)	HOBt (excess)	DIPEA (excess)
A	10.6mg (2.1μmol)	-	10μl	5mg	10μl
B	11.5mg (2.3μmol)	256μl (0.05eq, 0.12μmol)	10μl	5mg	10μl
C	15.8mg (3.1μmol)	702μl (0.1eq, 0.32μmol)	10μl	5g	10μl
D	7.5mg (1.5μmol)	833μl (0.25eq, 0.38μmol)	10μl	5mg	10μl
E	17.9mg (3.6μmol)	2.2mg (3 eq 10.74μmol)	10μl	5mg	10μl

Table 5-6 Reagent mixtures used to test efficiency of loading the capping group

HPLC conditions: Monitoring at 220nm and eluting with (A) 0.1% TFA in water and (B) 0.042% TFA in acetonitrile, gradient 0% (B) to 100% (B) in 5 minutes using a HP1100

system equipped with a Phenomenex Prodigy C_{18} reverse phase column (150 × 4.6 mm i.d.) with a flow rate of 1ml/min. Uncapped sequence ret. time 3.20min, *p*-Br PhCOO-capped sequence ret. time 3.99min

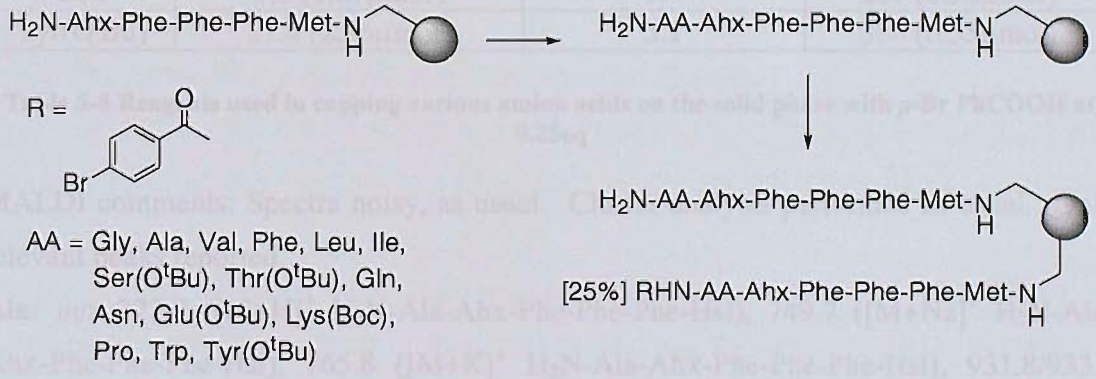
Ratio of uncapped sequence to capped sequence calculated by evaluating the ratio of the integrals of the relevant peaks (Table 5-7).

Sample	Uncapped : Capped
A	100 : 0
B	93 : 7
C	85 : 15
D	72 : 28
E	0 : 100

Table 5-7 Ratio of capped to uncapped sequence as determined by HPLC

Within experimental error (due to variability of bead loading and in measurement of bead loading), these values correspond to complete incorporation of the added *p*-Br PhCOOH.

Capping of various amino acids



Samples of resin **129** were first loaded with a variety of amino acids using the general procedures; resin was preswollen, then standard coupling was carried out using DIC (4μl), HOBt (2mg), DIPEA (4μl) and Fmoc-AA-OH (as in Table 5-8) in DMF. The samples were and washed according to the general procedure, then Fmoc-deprotected according to the general procedure. Reactions were monitored using the ninhydrin test and the chloranil test (to test for deprotection of Pro).

Each sample was then subjected to a capping reaction according to the general procedure using *p*-Br PhCOOH (according to Table 5-8) and DIC (5μl), HOBt (2mg) and DIPEA (5μl). Single beads were then selected from each sample, cleaved and analysed by MALDI-TOF-MS according to the general procedure.

AA	Mass resin/mg	Mass Fmoc-AA (3eq)/mg	Vol *p*-Br PhCOOH (1.47mM/ DMF, 0.25eq)/μl
Ala	12.5 (1.63μmol)	1.5	276 (0.41μmol)
Phe	20.3 (2.64μmol)	3.1	449 (0.66μmol)
Val	18.9 (2.46μmol)	2.5	418 (0.61μmol)
Ser(O^{t}Bu)	11.0 (1.43μmol)	1.6	243 (0.36μmol)
Thr(O^{t}Bu)	21.2 (2.76μmol)	3.3	469 (0.69μmol)
Glu(O^{t}Bu)	10.4 (1.35μmol)	1.7	230 (0.34μmol)
Asn	12.5 (1.63μmol)	1.7	276 (0.41μmol)
Lys(Boc)	15.9 (2.07μmol)	2.9	352 (0.52μmol)
Pro	11.5 (1.50μmol)	1.5	254 (0.37μmol)
Gln	10.7 (1.39μmol)	1.5	237 (0.35μmol)
Trp	7.9 (1.03μmol)	1.3	175 (0.26μmol)
Gly	11.7 (1.52μmol)	1.4	259 (0.38μmol)
Ile	19.8 (2.58μmol)	2.7	438 (0.64μmol)
Leu	9.8 (1.27μmol)	1.4	217 (0.32μmol)
Tyr(O^{t}Bu)	17.4 (2.26μmol)	3.1	384 (0.55μmol)

Table 5-8 Reagents used in capping various amino acids on the solid phase with *p*-Br PhCOOH at 0.25eq

MALDI comments: Spectra noisy, as usual. Cluster analysis performed as usual. Only relevant peaks reported.

Ala: *m/z* 727.7 ($[M+H]^+$ H_2N-Ala-Ahx-Phe-Phe-Phe-Hsl), 749.7 ($[M+Na]^+$ H_2N-Ala-Ahx-Phe-Phe-Phe-Hsl), 765.8 ($[M+K]^+$ H_2N-Ala-Ahx-Phe-Phe-Phe-Hsl), 931.8/933.8 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Ahx-Phe-Phe-Phe-Hsl), 947.8/949.8 ($[M+K]^+$ *p*-Br PhCONH-Ala-Ahx-Phe-Phe-Phe-Hsl)

Phe: *m/z* 803.9 ($[M+H]^+$ H_2N-Phe-Ahx-Phe-Phe-Phe-Hsl), 825.7 ($[M+Na]^+$ H_2N-Phe-Ahx-Phe-Phe-Phe-Hsl), 841.9 ($[M+K]^+$ H_2N-Phe-Ahx-Phe-Phe-Phe-Hsl), 1007.7/1009.7 ($[M+Na]^+$ *p*-Br PhCONH-Phe-Ahx-Phe-Phe-Phe-Hsl)

Val: *m/z* 755.7 ($[M+H]^+$ H_2N-Val-Ahx-Phe-Phe-Phe-Hsl), 777.7 ($[M+Na]^+$ H_2N-Val-Ahx-Phe-Phe-Phe-Hsl), 793.9 ($[M+K]^+$ H_2N-Ala-Ahx-Phe-Phe-Phe-Hsl), 959.8/961.8

([M+Na]$^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 975.8/977.7 ([M+K]$^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl)

Ser(O^{t}Bu): *m/z* 743.6 ([M+H]$^+$ H_2N-Ser-Ahx-Phe-Phe-Phe-Hsl), 765.8 ([M+Na]$^+$ H_2N-Ser-Ahx-Phe-Phe-Phe-Hsl), 799.6 ([M+H]$^+$ H_2N-Ser(O^{t}Bu)-Ahx-Phe-Phe-Phe-Hsl), 821.6 ([M+Na]$^+$ H_2N-Ser(O^{t}Bu)-Ahx-Phe-Phe-Phe-Hsl), 1003.6/1005.6 ([M+Na]$^+$ *p*-Br PhCONH-Ser(O^{t}Bu)-Ahx-Phe-Phe-Phe-Hsl), 1019.6/1021.6 ([M+K]$^+$ *p*-Br PhCONH-Ser(O^{t}Bu)-Ahx-Phe-Phe-Phe-Hsl)

Thr(O^{t}Bu): *m/z* 757.7 ([M+H]$^+$ H_2N-Thr-Ahx-Phe-Phe-Phe-Hsl), 779.7 ([M+Na]$^+$ H_2N-Thr-Ahx-Phe-Phe-Phe-Hsl), 961.8/963.8 ([M+Na]$^+$ *p*-Br PhCO-Thr-Ahx-Phe-Phe-Phe-Hsl)

Glu(O^{t}Bu): *m/z* 785.5 ([M+H]$^+$ H_2N-Glu-Ahx-Phe-Phe-Phe-Hsl), 807.5 ([M+Na]$^+$ H_2N-Glu-Ahx-Phe-Phe-Phe-Hsl), 823.6 ([M+K]$^+$ H_2N-Glu-Ahx-Phe-Phe-Phe-Hsl), 989.7/991.7 ([M+Na]$^+$ *p*-Br PhCONH-Glu-Ahx-Phe-Phe-Phe-Hsl)

Asn: *m/z* 770.5 ([M+H]$^+$ H_2N-Asn-Ahx-Phe-Phe-Phe-Hsl), 792.6 ([M+Na]$^+$ H_2N-Asn-Ahx-Phe-Phe-Phe-Hsl), 808.6 ([M+K]$^+$ H_2N-Asn-Ahx-Phe-Phe-Phe-Hsl), 974.7/976.7 ([M+Na]$^+$ *p*-Br PhCONH-Asn-Ahx-Phe-Phe-Phe-Hsl)

Lys(Boc): *m/z* 784.6 ([M+H]$^+$ H_2N-Lys-Ahx-Phe-Phe-Phe-Hsl), 806.6 ([M+Na]$^+$ H_2N-Lys-Ahx-Phe-Phe-Phe-Hsl), 822.7 ([M+K]$^+$ H_2N-Lys-Ahx-Phe-Phe-Phe-Hsl), 988.7/990.7 ([M+Na]$^+$ *p*-Br PhCONH-Lys-Ahx-Phe-Phe-Phe-Hsl)

Pro: *m/z* 753.6 ([M+H]$^+$ H_2N-Pro-Ahx-Phe-Phe-Phe-Hsl), 775.6 ([M+Na]$^+$ H_2N-Pro-Ahx-Phe-Phe-Phe-Hsl, 957.7/957.8 ([M+Na]$^+$ *p*-Br PhCONH-Pro-Ahx-Phe-Phe-Phe-Hsl)

Gln: *m/z* 789.7 ([M+H]$^+$ H_2N-Gln-Ahx-Phe-Phe-Phe-Hsl), 806.7 ([M+Na]$^+$ H_2N-Gln-Ahx-Phe-Phe-Phe-Hsl), 822.7 ([M+K]$^+$ H_2N-Gln-Ahx-Phe-Phe-Phe-Hsl), 988.8/990.8 ([M+Na]$^+$ *p*-Br PhCONH-Gln-Ahx-Phe-Phe-Phe-Hsl)

Trp: *m/z* 842.7 ([M+H]$^+$ H_2N-Trp-Ahx-Phe-Phe-Phe-Hsl), 864.8 ([M+Na]$^+$ H_2N-Trp-Ahx-Phe-Phe-Phe-Hsl), 880.8 ([M+K]$^+$ H_2N-Trp-Ahx-Phe-Phe-Phe-Hsl), 1046.9/1048.9 ([M+Na]$^+$ *p*-Br PhCONH-Trp-Ahx-Phe-Phe-Phe-Hsl), 1062.9/1064.9 ([M+K]$^+$ *p*-Br PhCONH-Trp-Ahx-Phe-Phe-Phe-Hsl)

Gly: *m/z* 713.5 ([M+H]$^+$ H_2N-Gly-Ahx-Phe-Phe-Phe-Hsl), 735.6 ([M+Na]$^+$ H_2N-Gly-Ahx-Phe-Phe-Phe-Hsl), 751.6 ([M+K]$^+$ H_2N-Gly-Ahx-Phe-Phe-Phe-Hsl), 917.7/919.6

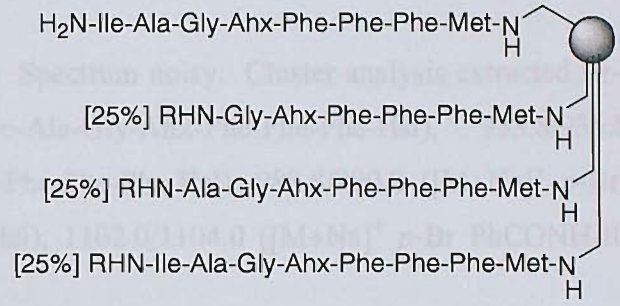
($[M+Na]^+$ *p*-Br PhCONH-Gly-Ahx-Phe-Phe-Phe-Hsl), 933.6/935.6 ($[M+K]^+$ *p*-Br PhCONH-Gly-Ahx-Phe-Phe-Phe-Hsl)

Ile: *m/z* 769.8 ($[M+H]^+$ H_2N-Ile-Ahx-Phe-Phe-Phe-Hsl), 791.7 ($[M+Na]^+$ H_2N-Ile-Ahx-Phe-Phe-Phe-Hsl), 807.7 ($[M+K]^+$ H_2N-Ile-Ahx-Phe-Phe-Phe-Hsl), 973.8/975.8 ($[M+Na]^+$ *p*-Br PhCONH-Ile-Ahx-Phe-Phe-Phe-Hsl), 989.8/991.8 ($[M+K]^+$ *p*-Br PhCONH-Ile-Ahx-Phe-Phe-Phe-Hsl)

Leu: *m/z* 769.7($[M+H]^+$ H_2N-Leu-Ahx-Phe-Phe-Phe-Hsl), 791.6 ($[M+Na]^+$ H_2N-Leu-Ahx-Phe-Phe-Phe-Hsl), 807.6 ($[M+K]^+$ H_2N-Leu-Ahx-Phe-Phe-Phe-Hsl), 973.7/975.7 ($[M+Na]^+$ *p*-Br PhCONH-Leu-Ahx-Phe-Phe-Phe-Hsl), 989.7/991.7 ($[M+K]^+$ *p*-Br PhCONH-Leu-Ahx-Phe-Phe-Phe-Hsl)

Tyr(O^{t}Bu): *m/z* 819.6 ($[M+H]^+$ H_2N-Tyr-Ahx-Phe-Phe-Phe-Hsl), 841.5 ($[M+Na]^+$ H_2N-Tyr-Ahx-Phe-Phe-Phe-Hsl), 857.6 ($[M+K]^+$ H_2N-Tyr-Ahx-Phe-Phe-Phe-Hsl), 1023.7/1025.7 ($[M+Na]^+$ *p*-Br PhCONH-Tyr-Ahx-Phe-Phe-Phe-Hsl), 1039.7/1041.7 ($[M+K]^+$ *p*-Br PhCONH-Tyr-Ahx-Phe-Phe-Phe-Hsl)

Single tripeptide sequence on solid phase with 'ladder sequences' (137)

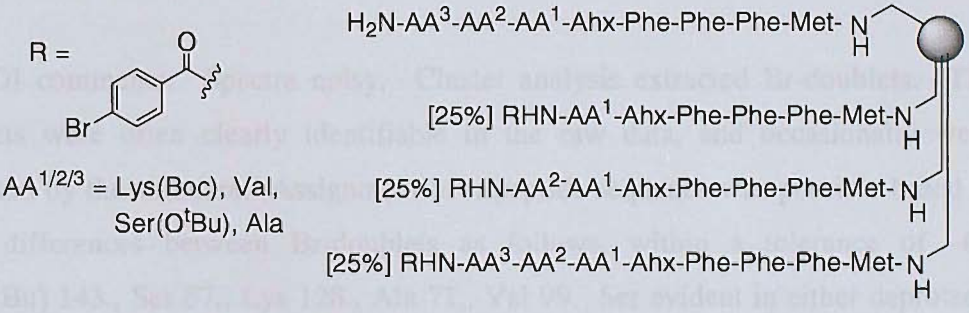


Resin **129** (11.7mg, 0.13mmol/g, 1.52μmol) was preswollen according to the general procedure. Standard amino acid coupling was then used to incorporate Gly using the following reagents; Fmoc-Gly-OH (1.4mg, 4.56μmol), DIC (5μl, excess), HOBt (5mg, excess) and DIPEA (5μl, excess) in DMF (2ml). The resin was then washed according to the general procedure and Fmoc-deprotected according to the general procedure. Reactions were monitored using the ninhydrin test. A capping reaction according to the general procedure was then performed using *p*-Br PhCOOH (259μl, 1.47mM/DMF, 0.38μmol) and DIC/HOBt/DIPEA (5μl/5mg/5μl). Standard amino acid coupling according to the general procedure was then carried out using the following reagents; Fmoc-Ala-OH (1.4mg, 4.56μmol), DIC (5μl, excess), HOBt (5mg, excess) and DIPEA (5μl, excess) in DMF (2ml). The resin was then washed according to general procedure. Fmoc-deprotection was then carried out according to the general procedure. Reactions were monitored using the ninhydrin test. A capping reaction according to the general procedure was then carried out using *p*-Br PhCOOH (330μl, 1.15mM/DMF, 0.38μmol) and DIC/HOBt/DIPEA (5μl/5mg/5μl). Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Ile-OH (1.6mg, 4.56μmol), DIC (5μl, excess), HOBt (5mg, excess) and DIPEA (5μl, excess) in DMF (2ml). Fmoc-deprotection according to the general procedure was carried out. Reactions were monitored using the ninhydrin test. A capping reaction according to the general procedure was then carried out using *p*-Br PhCOOH (330μl, 1.15mM/DMF, 0.38μmol) and DIC/HOBt/DIPEA (5μl/5mg/5μl). The resin was finally rinsed in Et_2O

and dried *in vacuo*. A single bead was picked for cleavage and analysis according to the general method.

MALDI comments: Spectrum noisy. Cluster analysis extracted Br-doublets. *m/z* 919.9 ([M+Na]$^+$ H_2N-Ile-Ala-Gly-Ahx-Phe-Phe-Phe-Hsl), 933.8/935.8 ([M+K]$^+$ *p*-Br PhCONH-Gly-Ahx-Phe-Phe-Phe-Hsl), 988.8/990.8 ([M+Na]$^+$ *p*-Br PhCONH-Ala-Gly-Ahx-Phe-Phe-Phe-Hsl), 1102.0/1104.0 ([M+Na]$^+$ *p*-Br PhCONH-Ile-Ala-Gly-Ahx-Phe-Phe-Phe-Hsl).

Restricted tripeptide library with 'ladder sequences' (140)



Resin **129** (80mg, 0.13mmol) was preswollen according to the general procedure. The bulk resin was then divided into four roughly equal portions (~20mg each, 0.13mmol, 2.6μmol). Each portion was coupled with an Fmoc-AA-OH according to the standard general procedure using DIC (10μl, excess), HOBt (5mg, excess) and DIPEA (10μl, excess) in DMF (2ml) and one of the following amino acids; Fmoc-Ser(O^{t}Bu)-OH, (3.0mg, 7.8μmol), Fmoc-Ala-OH (2.4mg,7.8μmol), Fmoc-Val-OH (2.6mg, 7.8μmol), Fmoc-Lys(Boc) (3.7mg, 7.8μmol). The resin was then washed according to general procedure. Fmoc-deprotection according to the general procedure was then carried out on each sample. All reactions were monitored using the ninhydrin test. Each sample was then subjected to a capping reaction according to the general procedure using *p*-Br PhCOOH (565μl, 1.15mM/DMF, 0.65μmol), DIC (10μl, excess), HOBt (5mg, excess) and DIPEA (10μl, excess). The resin portions were then combined, suspended in DCM and shaken for 10 mins to ensure thorough mixing. The bulk resin was then again

divided into four roughly equal portions. Standard amino acid coupling, washing, Fmoc-deprotection and capping reaction was then carried out in order on each sample according to the general procedures, with the same quantities of reagents. All reactions were monitored using the ninhydrin test. The resin was again combined, suspended in DCM and shaken for 10 mins to ensure thorough mixing. A final iteration of dividing into four roughly equal portions, standard amino acid coupling, Fmoc-deprotection and capping reaction (all according to general procedures using same quantities as above) was performed. Reactions were again monitored using the ninhydrin test. The resin samples were again combined, suspended in DCM and shaken for 10 mins to ensure thorough mixing. Finally the resin was rinsed with Et_2O and dried *in vacuo*. 20 single beads were selected, cleaved and analysed according to the general procedure.

MALDI comments: Spectra noisy. Cluster analysis extracted Br-doublets. The Br-doublets were often clearly identifiable in the raw data, and occasionally were not extracted by the software. Assignment of tripeptide sequence was possible based on the mass differences between Br-doublets as follows, within a tolerance of ~0.3mu; Ser(O^{t}Bu) 143., Ser 87., Lys 128., Ala 71., Val 99. Ser evident in either deprotected or protected form, occasionally in both. 18 out of 20 cleaved beads could be assigned.

Bead 1: *m/z* 931.9/933.9 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Ahx-Phe-Phe-Phe-Hsl), 1003.0/1005.0 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Ala-Ahx-Phe-Phe-Phe-Hsl), 1074.1/1076.1 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Ala-Ala-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Ala-Ala-Ala-linker

Bead 2: *m/z* 931.8/933.8 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Ahx-Phe-Phe-Phe-Hsl), 1002.9/1004.9 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Ala-Ahx-Phe-Phe-Phe-Hsl), 1146.1/1148.1 ($[M+Na]^+$ *p*-Br PhCONH-Ser(O^{t}Bu)-Ala-Ala-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Ser(O^{t}Bu)-Ala-Ala-linker

Bead 3: 959.9/961.9 ($[M+Na]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 1031.0/1033.0 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Val-Ahx-Phe-Phe-Phe-Hsl) 1159.2/1161.2 ($[M+Na]^+$ *p*-Br PhCONH-Lys-Ala-Val-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Lys(Boc)-Ala-Val-linker

Bead 4: *m/z* 960.1/962.1 ($[M+Na]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 976.1/978.1 ($[M+K]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 1059.2/1061.3 ($[M+Na]^+$ *p*-Br PhCONH-Val-Val-Ahx-Phe-Phe-Phe-Hsl), 1075.2/1077.2 ($[M+K]^+$ *p*-Br PhCONH-Val-Val-Ahx-Phe-Phe-Phe-Hsl), 1146.3/1148.3 ($[M+Na]^+$ *p*-Br PhCONH-Ser-Val-Val-Ahx-Phe-Phe-Phe-Hsl) 1202.5/1204.4 ($[M+Na]^+$ *p*-Br PhCONH-Ser(O^tBu)-Val-Val-Ahx-Phe-Phe-Phe-Hsl), 1218.4/1220.5 ($[M+K]^+$ *p*-Br PhCONH-Ser(O^tBu)-Val-Val-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Ser(O^tBu)-Val-Val-linker

Bead 5: *m/z* 1019.1/1021.1 ($[M+K]^+$ *p*-Br PhCONH-Ser(O^tBu)-Ahx-Phe-Phe-Phe-Hsl), 1147.2/1149.1 ($[M+K]^+$ *p*-Br PhCONH-Lys-Ser(O^tBu)-Ahx-Phe-Phe-Phe-Hsl), 1246.1/1248.1 ($[M+K]^+$ *p*-Br PhCONH-Val-Lys-Ser(O^tBu)-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Val-Lys(Boc)-Ser(O^tBu)-linker

Bead 6: *m/z* 989.1/991.1 ($[M+Na]^+$ *p*-Br PhCONH-Lys-Ahx-Phe-Phe-Phe-Hsl), 1060.2/1062.2 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Lys-Ahx-Phe-Phe-Phe-Hsl), 1188.3/1190.3 ($[M+Na]^+$ *p*-Br PhCONH-Lys-Ala-Lys-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Lys(Boc)-Ala-Lys(Boc)-linker

Bead 7: *m/z* 932.0/934.0 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Ahx-Phe-Phe-Phe-Hsl), 1031.2/1033.1 ($[M+Na]^+$ *p*-Br PhCONH-Val-Ala-Ahx-Phe-Phe-Phe-Hsl), 1159.2/1161.3 ($[M+Na]^+$ *p*-Br PhCONH-Lys-Val-Ala-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Lys(Boc)-Val-Ala-linker

Bead 8: *m/z* 989.1/991.1 ($[M+Na]^+$ *p*-Br PhCONH-Lys-Ahx-Phe-Phe-Phe-Hsl), 1004.1/1006.1 ($[M+K]^+$ *p*-Br PhCONH-Lys-Ahx-Phe-Phe-Phe-Hsl), 1060.3/1062.3 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Lys-Ahx-Phe-Phe-Phe-Hsl), 1146.4/1148.3 ($[M+Na]^+$ *p*-Br PhCONH-Ser-Ala-Lys-Ahx-Phe-Phe-Phe-Hsl) 1202.4/1204.4 ($[M+Na]^+$ *p*-Br PhCONH-Ser(O^tBu)-Ala-Lys-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Ser(O^tBu)-Ala-Lys(Boc)-linker

Bead 9: *m/z* 960.0/962.0 ($[M+Na]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 1031.0/1033.0 ($[M+Na]^+$ *p*-Br PhCONH-Val-Val-Ahx-Phe-Phe-Phe-Hsl), 1047.1/1049.1 ($[M+K]^+$ *p*-Br PhCONH-Val-Val-Ahx-Phe-Phe-Phe-Hsl), 1159.2/1161.3 ($[M+Na]^+$ *p*-Br PhCONH-Lys-Val-Val-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Lys(Boc)-Val-Val-linker

Bead 10: *m/z* 1019.9/1021.9 ($[M+K]^+$ *p*-Br PhCONH-Ser(O^tBu)-Ahx-Phe-Phe-Phe-Hsl), 1119.1/1121.1 ($[M+K]^+$ *p*-Br PhCONH-Val-Ser(O^tBu)-Ahx-Phe-Phe-Phe-Hsl), 1231.2/1233.2 ($[M+Na]^+$ *p*-Br PhCONH-Lys-Val-Ser(O^tBu)-Ahx-Phe-Phe-Phe-Hsl), 1247.2/1249.2 ($[M+K]^+$ *p*-Br PhCONH-Lys-Val-Ser(O^tBu)-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Lys(Boc)-Val-Ser(O^tBu)-linker

Bead 11: *m/z* 988.9/990.9 ($[M+Na]^+$ *p*-Br PhCONH-Lys-Ahx-Phe-Phe-Phe-Hsl), 1004.8/1006.8 ($[M+K]^+$ *p*-Br PhCONH-Lys-Ahx-Phe-Phe-Phe-Hsl), 1133.0/1135.0 ($[M+K]^+$ *p*-Br PhCONH-Lys-Lys-Ahx-Phe-Phe-Phe-Hsl), 1204.1/1206.1 ($[M+K]^+$ *p*-Br PhCONH-Ala-Lys-Lys-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Ala-Lys(Boc)-Lys(Boc)-linker

Bead 12: *m/z* 988.9/990.9 ($[M+Na]^+$ *p*-Br PhCONH-Lys-Ahx-Phe-Phe-Phe-Hsl), 1059.9/1061.9 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Lys-Ahx-Phe-Phe-Phe-Hsl), 1203.0/1205.1 ($[M+Na]^+$ *p*-Br PhCONH-Ser(O^tBu)-Ala-Lys-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Ser(O^tBu)-Ala-Lys(Boc)-linker

Bead 13: *m/z* 959.9/961.8 ($[M+Na]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 975.8/977.8 ($[M+K]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 1058.9/1060.9 ($[M+Na]^+$ *p*-Br PhCONH-Val-Val-Ahx-Phe-Phe-Phe-Hsl), 1129.9/1131.9 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Val-Val-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N- Ala-Val-Val-linker

Bead 14: *m/z* 959.9/961.9 ($[M+Na]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 1031.0/1033.0 ($[M+Na]^+$ *p*-Br PhCONH-Ala-Val-Ahx-Phe-Phe-Phe-Hsl), 1174.2/1176.2 ($[M+Na]^+$ *p*-Br PhCONH-Ser(O^tBu)-Ala-Val-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Ser(O^tBu)-Ala-Val-linker

Bead 15: *m/z* 989.0/991.0 ($[M+Na]^+$ *p*-Br PhCONH-Lys-Ahx-Phe-Phe-Phe-Hsl), 1117.1/1119.1 ($[M+Na]^+$ *p*-Br PhCONH-Lys-Lys-Ahx-Phe-Phe-Phe-Hsl), 1260.3/1262.3 ($[M+Na]^+$ *p*-Br PhCONH-Ser(O^tBu)-Lys-Lys-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Ser(O^tBu)-Lys(Boc)-Lys(Boc)-linker

Bead 16: *m/z* 959.8/961.8 ($[M+Na]^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 1103.0/1105.1 ($[M+Na]^+$ *p*-Br PhCONH-Ser(O^tBu)-Val-Ahx-Phe-Phe-Phe-Hsl), 1245.9/1247.9 ($[M+Na]^+$ *p*-Br PhCONH-Ser(O^tBu)-Ser(O^tBu)-Val-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Ser(O^{t}Bu)-Ser(OtBu)-Val-linker

Bead 17: *m/z* 975.9/977.9 ([M+K]$^+$ *p*-Br PhCONH-Val-Ahx-Phe-Phe-Phe-Hsl), 1047.0/1049.0 ([M+K]$^+$ *p*-Br PhCONH-Ala-Val-Ahx-Phe-Phe-Phe-Hsl), 1175.1/1177.0 M+K]$^+$ *p*-Br PhCONH-Lys-Ala-Val-Ahx-Phe-Phe-Phe-Hsl),

Hence tripeptide sequence is H_2N-Lys(Boc)-Ala-Val-linker

Bead 18: m/z 1003.8/1005.8 ([M+Na]$^+$ *p*-Br PhCONH-Ser(O^{t}Bu)-Ahx-Phe-Phe-Phe-Hsl), 1019.8/1021.8 ([M+K]$^+$ *p*-Br PhCONH-Ser(O^{t}Bu)-Ahx-Phe-Phe-Phe-Hsl), 1147.0/1149.0 ([M+Na]$^+$ *p*-Br PhCONH-Ser-(O^{t}Bu)-Ser(O^{t}Bu)-Ahx-Phe-Phe-Phe-Hsl), 1163.0/1165.0 ([M+K]$^+$ *p*-Br PhCONH-Ser-(O^{t}Bu)-Ser(O^{t}Bu)-Ahx-Phe-Phe-Phe-Hsl), 1234.0/1236.0 ([M+K]$^+$ *p*-Br PhCONH-Ala-Ser-(O^{t}Bu)-Ser(O^{t}Bu)-Ahx-Phe-Phe-Phe-Hsl)

Hence tripeptide sequence is H_2N-Ala-Ser(O^{t}Bu)-Ser(O^{t}Bu)-linker

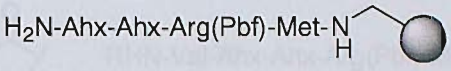
Statistical analysis:

Residue	**Times occurring**
Ser	12
Ala	15
Val	13
Lys	14

	Probability	Times occurring	Times occurring %
3 identical residues	0.0625	1	0.05
2 identical residues	0.5625	8	0.40
All different	0.375	9	0.45

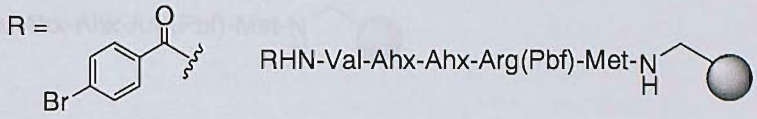
Thus both residues and tripeptides represent a roughly random distribution.

Resin with Arg-containing linker (143)



Tentagel resin **88** (0.2mmol/g, 551mg, 121μmol) was preswollen according to general procedure. Standard amino acid coupling was then performed using the following reagents; Fmoc-Met-OH (90mg, 242μmol), DIC (57μl, 363μmol), HOBt (49mg, 363μmol) and DIPEA (63μl, 363μmol) in DMF (5ml). The resin was washed according to general procedure. Fmoc-deprotection according to the general procedure was then carried out. Reactions were monitored using the ninhydrin test. Standard amino acid coupling according to the general procedure was then carried out using the following reagents; Fmoc-Arg(Pbf)-OH (157mg, 242μmol), DIC (57μl, 363μmol), HOBt (49mg, 363μmol) and DIPEA (63μl, 363μmol) in DMF (5ml). The resin was then washed according to the general procedure. Fmoc-deprotection was carried out according to the general procedure. Reactions were monitored using the ninhydrin test. The resin was then washed with 50% DIPEA in DMF, followed by washing according to the general procedure. Standard amino acid coupling according to the general procedure was then carried out using the following reagents; Fmoc-Ahx-OH (86mg, 242μmol), DIC (57μl, 363μmol), HOBt (49mg, 363μmol) and DIPEA (63μl, 363μmol) in DMF (5ml). The resin was washed according to the general procedure, then subjected to Fmoc-deprotection according to the general procedure. Reactions were monitored using the ninhydrin test. The resin was then washed with 50% DIPEA/DMF, then washed according to the general procedure. Standard amino acid coupling according to the general procedure was then performed using the following reagents; Fmoc-Ahx-OH (86mg, 242μmol), DIC (57μl, 363μmol), HOBt (49mg, 363μmol) and DIPEA (63μl, 363μmol) in DMF (5ml). The resin was washed according to general procedure and then subjected to Fmoc-deprotection according to the general procedure. Resin **143** was then finally rinsed with Et_2O and rigourously dried *in vacuo*. A quantitative ninhydrin test gave the resin loading as 0.15mmol/g.

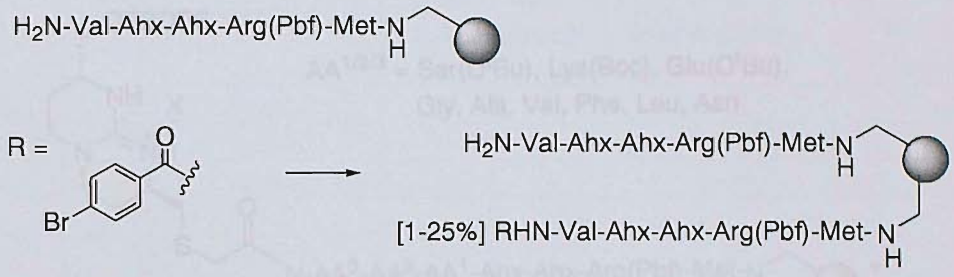
Loading of the Arg-resin 143 with capping group, giving 145



Resin construct **143** (0.15mmol/g, 146.1mg, 21.9μmol) was preswollen according to the general procedure. Standard amino acid coupling was performed using the following reagents; Fmoc-Val-OH (22.3mg, 65.7μmol), DIC (11μl, 65.7μmol), HOBt (9mg, 65.7μmol) and DIPEA (12μl, 65.7μmol) in DMF (5ml). The resin was then washed according to the general procedure and subjected to Fmoc-deprotection according to the general procedure. Reactions were monitored using the ninhydrin test. A sample of this resin (13.2mg, 1.98μmol) was treated with *p*-Br-PhCOOH (3mg, 14.9μmol) and DIC/HOBt/DIPEA (10μl/10mg/10μl, all in excess) in DMF (1ml) and shaken for 18 hours. The resin was then drained and washed according to the general procedure. A ninhydrin test gave a negative result. The resin was then shaken with 50% TFA/DCM for 1 hour (to remove the Pbf group), then drained and washed with 50% DIPEA/DMF (~20ml per g of resin, three times) and then according to the general procedure. The resin was then finally rinsed with Et_2O and dried *in vacuo*. 2 single beads were picked out for cleavage and analysis by MALDI according to the general method.

MALDI comments: Spectra noisy. Cluster analysis was performed to extract Br-doublets. Both spectra clearly showed the presence of 764.7/766.7 $[M+H]^+$ of the cleavage product *p*-Br-PhCOO-Val-Ahx-Arg-Ahx-Hsl.

Capping experiments at various loadings onto Arg-construct



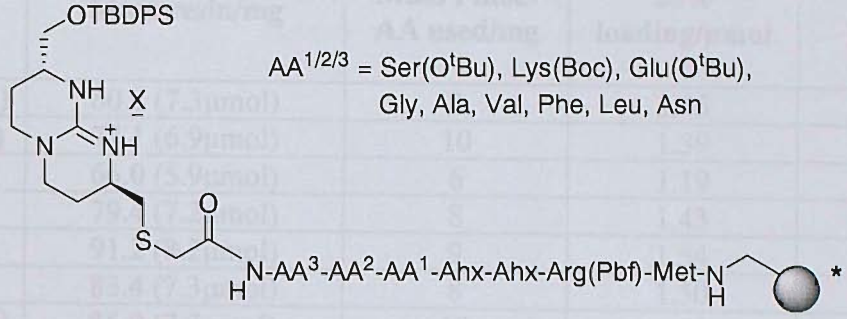
Samples of resin **144** (prepared in previous experiment) were preswollen according to the general procedure. Each sample was then subjected to a capping reaction using *p*-Br-PhCOOH and excess DIC/HOBt/DIPEA in DMF (as detailed in the Table 5-9). Each sample was treated with 50% TFA/DCM (3ml) for 1 hour (to remove the Pbf group), then washed with 50% DIPEA/DMF (20ml per g of resin, three times), and then washed according to the general procedure. The resin was finally rinsed with Et_2O and dried *in vacuo*. 2 resin beads were selected from each sample and cleaved according to the standard procedure.

MALDI comments: Spectra noisy. Cluster analysis was performed to extract Br-doublets. Br-doublets were clearly evident at 764.7/766.7 $[M+H]^+$ of the cleavage product *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl in samples D, E and F, but were absent from samples A, B, C, indicating that the limit of detectability of the capping group on the Arg-construct lies between 5-10%

Sample	Mass resin 15	% capping	*p*-Br PhCOOH in DMF added/µl
A	17.4mg, 2.61µmol	1	105 (0.25mM, 0.026µmol)
B	15.7mg, 2.36µmol	2	189 (0.25mM, 0.047µmol)
C	19.5mg, 2.93µmol	5	222 (0.66mM, 0.146µmol)
D	15.9mg, 2.39µmol	10	159 (1.51mM, 0.24µmol)
E	15.7mg, 2.36µmol	15	231 (1.51mM, 0.35µmol)
F	15.6mg, 2.34µmol	20	311 (1.51mM, 0.47µmol)

Table 5-9 Reagents used in capping resin x with *p*-Br PhCOOH various loadings

Single-armed receptor library (149)



Resin **143** (measured as 0.09mmol/g, 735mg, 66μmol) was split into nine portions according to Table 5-10. Each portion was then preswollen according to the general procedure and subjected to standard amino acid coupling according to the general procedure with a different Fmoc-amino acid and DIC (10μl, excess), HOBt (10mg, excess), and DIPEA (10μl, excess) in DMF (1ml) according to the table. Each sample was then washed according to general procedure and Fmoc-deprotected according to general procedure. Reactions were monitored using the ninhydrin test. The resin was then subjected to a capping reaction according to the general procedure using *p*-Br PhCOOH (as detailed in Table 5-10) and DIC (10μl, excess), HOBt (10mg, excess) and DIPEA (10μl, excess). The samples were then combined, suspended in DMF and shaken for 10 mins to ensure complete mixing.

AA	Mass resin/mg	Mass Fmoc-AA used/mg	20% loading/μmol	Vol. 2.33mM *p*-Br PhCOOH used/μl
Ser(O^{t}Bu)	80.8 (7.3μmol)	8	1.45	624
Lys(Boc)	77.1 (6.9μmol)	10	1.39	596
Leu	66.0 (5.9μmol)	6	1.19	510
Phe	79.4 (7.2μmol)	8	1.43	614
Asn	91.2 (8.2μmol)	9	1.54	704
Ala	83.4 (7.3μmol)	8	1.50	645
Glu(O^{t}Bu)	85.0 (7.7μmol)	10	1.53	657
Gly	86.9 (7.8μmol)	7	1.56	671
Val	85.5 (7.7μmol)	8	1.54	661

Table 5-10 Quantities of reagents used in single-armed receptor library synthesis, first iteration of split-and-mix.

The resin was then divided into nine portions according to Table 5-11. Each portion was then preswollen according to the general procedure and subjected to standard amino acid coupling according to the general procedure with a different Fmoc-amino acid and DIC (10μl, excess), HOBt (10mg, excess), and DIPEA (10μl, excess) in DMF (1ml) according to Table 5-11 (note correction for average added mass). Each sample was then washed according to general procedure and Fmoc-deprotected according to general procedure. Reactions were monitored using the ninhydrin test. The resin was then subjected to a capping reaction according to the general procedure using *p*-Br PhCOOH (as detailed in Table 5-11) and DIC (10μl, excess), HOBt (10mg, excess) and DIPEA (10μl, excess). The samples were then combined, suspended in DMF and shaken for 10 mins to ensure complete mixing.

Residue	Mass resin/mg	Equivalent mass at known loading/mg	Mass Fmoc-AA used/mg	20% loading/ μmol	Vol 3.32mM *p*-Br PhCOOH used/μl
Ser(O^{t}Bu)	82.5	80.8 (7.27μmol)	8	1.45	438
Lys(Boc)	76.7	75.2 (6.77μmol)	10	1.35	407
Leu	88.9	87.1 (7.84μmol)	8	1.57	472
Glu(O^{t}Bu)	81.8	80.2 (7.22μmol)	9	1.44	435
Gly	86.5	84.8 (7.63μmol)	7	1.52	459
Ala	91.9	90.0 (8.10μmol)	8	1.62	488
Val	99.6	97.6 (8.78μmol)	9	1.76	528
Phe	97.9	95.9 (8.63μmol)	10	1.73	519
Asn	95.2	93.9 (8.40μmol)	9	1.68	506

Table 5-11 Quantities of reagents used in single-armed receptor library synthesis, second iteration of split-and-mix.

The resin was then divided into nine portions according to Table 5-12. Each portion was then preswollen according to the general procedure and subjected to standard amino acid coupling according to the general procedure with a different Fmoc-amino acid and DIC (10μl, excess), HOBt (10mg, excess), and DIPEA (10μl, excess) in DMF (1ml) according to the table. Each sample was then washed according to general procedure and Fmoc-deprotected according to general procedure. Reactions were monitored using the ninhydrin test. The resin was then subjected to a capping reaction according to the general procedure using *p*-Br PhCOOH (as detailed in Table 5-12) and DIC (10μl, excess), HOBt (10mg, excess) and DIPEA (10μl, excess). The samples were then combined, suspended in DMF and shaken for 10 mins to ensure complete mixing.

Residue	Mass resin/mg	Equivalent mass at known loading/mg	Mass Fmoc-AA used/mg	30% loading/ μmol	Vol 3.66mM *p*-Br PhCOOH used/μl
Ser(O^{t}Bu)	83.5	80.2 (7.22μmol)	8	1.44	395
Lys(Boc)	87.8	84.3 (7.59μmol)	11	1.52	415
Leu	104.4	100.3 (9.03μmol)	10	1.81	493
Glu(O^{t}Bu)	71.3	68.5 (6.16μmol)	8	1.23	337
Gly	81.8	78.6 (7.07μmol)	6	1.41	386
Ala	96.4	92.6 (8.33μmol)	8	1.67	455
Val	79.9	76.8 (6.91μmol)	7	1.38	378
Phe	111.4	107.0 (9.63μmol)	11	1.93	526
Asn	82.4	79.2 (7.12μmol)	8	1.42	389

Table 5-12 Quantities of reagents used in single-armed receptor library synthesis, third iteration of split-and-mix.

Resin was then preswollen according to the general procedure and subjected to coupling according to the general procedure using CBS-acid **145** (55mg, 108μmol) and DIC (34μl, 216μmol), HOBt (29mg, 216μmol) and DIPEA (38μl, 216μmol) in DMF (10 ml). The resin was then washed according to the general procedure. A ninhydrin test gave a negative result, indicating that all remaining free peptide had been capped by the CBS. Resin was then rinsed with Et_2O and dried *in vacuo*. Three single beads were selected, and cleaved and analysed according to the general procedure.

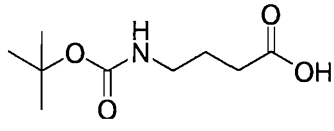
MALDI comments: Spectra messy. Bromine doublets identified and sequences assigned according to mass difference;

Bead 1: 753.1/755.1 ($[M+H]^+$ *p*-Br PhCONH-Ser-Ahx-Ahx-Arg-Hsl), 900.4/902.3 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ser-Ahx-Ahx-Arg-Hsl), 999.5/1001.5 ($[M+H]^+$ *p*-Br PhCONH-Val-Phe-Ser-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence is Val-Phe-Ser(O^{t}Bu)

Bead 2: 812.3/814.5 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 869.3/871.2 ($[M+H]^+$ *p*-Br PhCONH-Gly-Phe-Ahx-Ahx-Arg-Hsl), 926.2/928.1 ($[M+H]^+$ *p*-Br PhCONH-Gly-Gly-Phe-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence is Gly-Phe-Phe

Bead 3: 753.4/755.2 ($[M+H]^+$ *p*-Br PhCONH-Ser-Ahx-Ahx-Arg-Hsl), 852.1/853.1 ($[M+H]^+$ *p*-Br PhCONH-Val-Ser-Ahx-Ahx-Arg-Hsl), 922.9/924.9 ($[M+H]^+$ *p*-Br PhCONH-Ala-Val-Ser-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence is Ala-Val-Ser

4-[(*tert*-butoxycarbonyl)amino]butanoic acid (153)

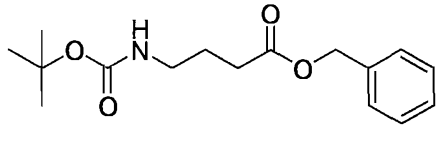


According to a procedure by Castonguay *et al.*[175]

Di-*tert*-butyl dicarbonate (9.60g, 44.0mmol) was dissolved in MeOH (10ml) and added to a solution of 4-aminobutyric acid (2.27g, 22.0mmol) and Et_3N (18.4ml, 132mmol) in MeOH (20ml). The mixture was heated at reflux for 3 hours and then evaporated to dryness. The residue was redissolved in saturated $NaHCO_3$(aq) and extracted with hexane. The aqueous phase was then acidified with conc. HCl and extracted with EtOAc. The organic phase was dried over $MgSO_4$ and evaporated to give **153** as a colourless oil (4.04g, 90%); ^{1}H NMR: (300MHz, $CDCl_3$) δ 4.70 (br s, 1H, N*H*), 3.17 (m, 2H, NHCH_2), 2.39 (t, 2H, *J* = 7.0Hz, CH_2COOH), 1.81 (quin, 2H, *J* = 7.0Hz, $CH_2CH_2CH_2$), 1.44 (s, 9H, Boc); ^{13}C NMR: (75MHz, $CDCl_3$) δ 178.4 (C), 156.4 (C), 79.7 (C), 39.9 (CH_2), 31.4 (CH_2), 28.5 (CH_3), 25.3 (CH_2); LRMS: (ES^+) *m/z* 203.9 $[M+H]^+$ 225.9 $[M+Na]^+$

Data agrees with literature values.[175]

Benzyl 4-[(*tert*-butoxycarbonyl)amino]butanoate (154)

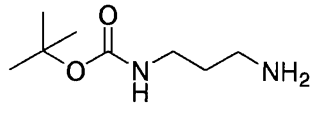


Acid **153** (120mg, 0.59mmol), K_2CO_3 (193mg, 1.16mmol), and benzyl bromide (70μl, 0.59mmol) were suspended in a mixture of acetone (5ml) and water (100μl). The mixture was heated at reflux for 2 hours. After cooling to room temperature, the reaction

mixture was filtered and the filtrates evaporated. The residue was purified by column chromatography (eluent 10% EtOAc/petrol) to give the product **154** as a white waxy solid (157.7mg, 91%); m.p.: 64-65°C (lit. value: 64°C)[178]; ^{1}H NMR: (300MHz, $CDCl_3$) δ 7.30-7.22 (m, 5H, arom), 5.03 (s, 2H, C*H*$_2$Ph), 4.66 (br s, 1H, N*H*), 3.06 (m, 2H, NHC*H*$_2$), 2.31 (t, 2H, *J* = 7.0Hz, C*H*$_2$COOBn), 1.74 (quin, 2H, *J* = 7.0Hz, $CH_2CH_2CH_2$), 1.35 (s, 9H, Boc); ^{13}C NMR: (75MHz, $CDCl_3$) δ 173.1 (C), 156.0 (C), 136.0 (C), 128.6 (CH), 128.3 (CH), 128.2 (CH), 79.2 (C), 66.3 (CH_2), 40.0 (CH_2), 31.5 (CH_2), 28.4 (CH_3), 25.4 (CH_2); LRMS: (ES$^+$) *m/z* 294.0 $[M+H]^+$ 316.1 $[M+Na]^+$

Data agrees with literature values.[188]

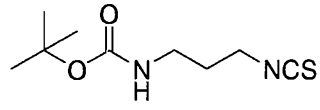
tert-Butyl _N_-(3-aminopropyl)carbamate (156)



Di-*tert*-butyl dicarbonate (9.8g, 0.045mmol) was dissolved in $CHCl_3$ (300ml) and added dropwise to a solution of 1,3-diaminopropane (11.3ml, 0.135mmol) in $CHCl_3$ (30ml) over the course of 4 hours. A white precipitate was noted. The reaction mixture was concentrated to half-volume and then washed with saturated $NaHCO_3$(aq). The aqueous phase was extracted with $CHCl_3$. The combined organic phases were dried over $MgSO_4$ and evaporated to give a colourless oil which solidified on standing to give **156** as a white waxy solid (6.82g, 87%); m.p.: 53-55°C (lit. value: 54-56°C)[189]; ^{1}H NMR: (300MHz, $CDCl_3$) δ 4.91 (br s, 1H, N*H*), 3.19 (dt, 2H, *J* = 6.5, 6.5Hz, NHC*H*$_2$CH$_2$), 2.74 (t, 2H, *J* = 6.5Hz, C*H*$_2$NH$_2$), 1.59 (quin, 2H, *J* = 6.5Hz, $CH_2CH_2CH_2$), 1.42 (s, 9H, Boc); ^{13}C NMR: (75MHz, $CDCl_3$) δ 156.3 (C), 79.2 (C), 39.9 (CH_2), 38.6 (CH_2), 33.6 (CH_2), 28.5 (CH_3); LRMS: (ES$^+$) *m/z* 174.8 $[M+H]^+$

Data agrees with literature values.[190]

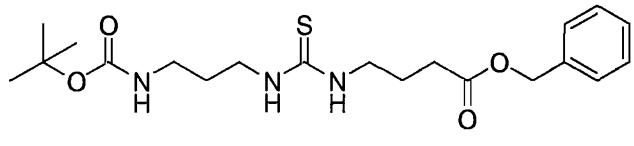
tert-Butyl _N_-(3-isothiocyanatopropyl)carbamate 157



Thiophosgene (5.1ml, 66.3mmol) was added to a biphasic solution of amine **156** (2.89g, 16.6mmol) in $CHCl_3$ (150ml) and 2M K_2CO_3(aq) (40ml). The mixture was stirred for 18 hours. The layers were then separated and the organic phase was washed with 2M K_2CO_3, dried over $MgSO_4$ and evaporated. The residue was purified by column chromatography (eluent 20% EtOAc/petrol) to give a yellow solid (3.02g, 84%); m.p. 76-78°C; IR ν_{max} (cm^{-1}): 3376 (m), 2977 (m), 2933 (m), 2186 (m), 2139 (s), 2089 (s), 1672 (s), 1513 (s), 1365 (s), 1245 (s), 1158 (s); ^1H NMR: (300MHz, $CDCl_3$) δ 4.70 (br s, 1H, N*H*), 3.59 (t, 2H, *J* = 6.5Hz, CH_2NCS), 3.23 (m, 2H, NHCH_2CH_2), 1.89 (quin, 2H, *J* = 6.5Hz, $CH_2CH_2CH_2$), 1.44 (s, 9H, Boc); ^{13}C NMR: (75MHz, $CDCl_3$) δ 173.8 (C), 155.9 (C), 79.9 (C), 42.6 (CH_2), 37.7 (CH_2), 30.3 (CH_2), 28.4 (CH_3); LRMS: decomposes under MS conditions.

Data agrees with literature results.[177]

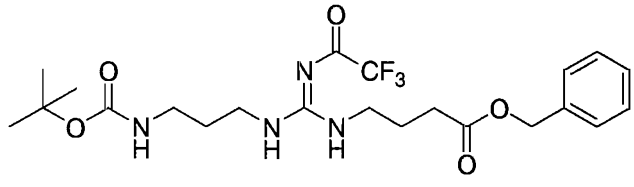
Benzyl 4-[(3-[(_tert_-butoxycarbonyl)amino]propylamino)carbothioyl]aminobutanoate (155)



Ester **154** (2.12g, 7.22mmol) was stirred in 20% TFA/DCM (30ml) for 1 hour. The solvents were then removed *in vacuo* by azeotropic distillation with toluene and the residue redissolved in 1:1 MeOH/DCM (40ml). Isothiocyanate **157** (1.56g, 7.22mmol) was then added followed by Et_3N (5ml, 36.1mmol). The mixture was stirred for 18 hours and then evaporated *in vacuo*. The residue was purified by column chromatography (eluent 30% EtOAc/petrol) to give the product thiourea **155** as a yellow oil (2.22g, 75%); IR ν_{max} (cm^{-1}): 3325 (br w), 2974 (w), 2932 (w), 1687 (m), 1548 (m), 1516 (m), 1251

(m), 1163 (s); ^{1}H NMR: (300 MHz, $CDCl_3$) δ 7.34 (m, 5H, arom), 6.40 (br s, 1H, N*H*), 5.12 (s, 2H, C*H*$_2$Ph), 4.90 (br s, 1H, N*H*), 3.61 (m, 2H, SCNHC*H*$_2$), 3.41 (m, 2H, SCNHC*H*$_2$), 3.17 (m, 2H, BocHNC*H*$_2$), 2.46 (t, 2H, *J* = 7.0Hz, C*H*$_2$COOBn), 1.93 (quin, 2H, *J* = 7.0Hz, $CH_2CH_2CH_2$); 1.71 (quin, 2H, *J* = 6.0Hz, $CH_2CH_2CH_2$); 1.42 (s, 9H, Boc); ^{13}C NMR: (75MHz, $CDCl_3$) δ 181.5 (C), 173.5 (C), 157.1 (C), 135.8 (C), 128.7 (CH), 128.5 (CH), 128.4 (CH), 79.8 (C), 66.8 (CH_2), 43.2 (CH_2), 41.3 (CH_2), 37.3 (CH_2), 31.5 (CH_2), 29.8 (CH_2), 28.5 (CH_3), 24.1 (CH_2); LRMS: (ES^+) *m/z* 410.2 $[M+H]^+$ 432.2 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{19}H_{31}N_3O_4S_1$ $[M+Na]^+$ 432.1927, found 432.1926.

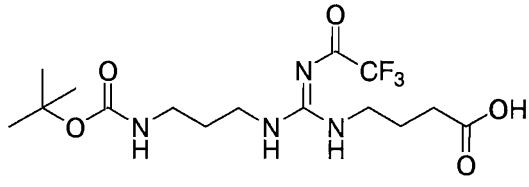
Benzyl 4-((3-[(*tert*-butoxycarbonyl)amino]propylamino)[(2,2,2-trifluoroacetyl)-imino]methylamino)butanoate (160)



Iodomethane (3.4ml, 54.1mmol) was added to a solution of thiourea **155** (2.22g, 5.41mmol) in acetone (40ml). The mixture was stirred for 1 hour and then evaporated *in vacuo*. The residue was redissolved in 1:1 MeOH/DCM (40ml) and ammonium hexafluorophosphate (1.77g, 10.8mmol) was added. The mixture was stirred for 18 hours and then evaporated *in vacuo*. The residue was taken up in DCM and washed once with distilled water. The organic phase was dried over $MgSO_4$ and evaporated to give a white foam. This foam was then dissolved in 4:1 toluene/$CHCl_3$ (50ml). Trifluoroacetamide (3.06g, 27.1mmol) and DBU (1.6ml, 10.8mmol) were added and the mixture heated at reflux for 1 hour. After cooling, the solvents were removed *in vacuo* and the residue purified by column chromatography (20% EtOAc/petrol) to give the product **160** as a yellow oil (1.31g, 50%); IR ν_{max} (cm^{-1}): 3333 (br w), 2979 (w), 1690 (m), 1624 (s), 1517 (m), 1434 (m), 1167 (s), 1137 (s); ^{1}H NMR: (300MHz, $CDCl_3$) δ 9.56 (br s, 1H, N*H*), 7.34 (m, 5H, arom), 6.39 (br s, 1H, NH), 5.14 (br s, 1H, N*H*) 5.13 (s, 2H, C*H*$_2$Ph), 3.52 (m, 2H, BocNHC*H*$_2$), 3.26 (m, 2H, NHC*H*$_2$), 3.17 (dt, 2H, *J* = 6.5, 6.0Hz, NHC*H*$_2$), 2.47

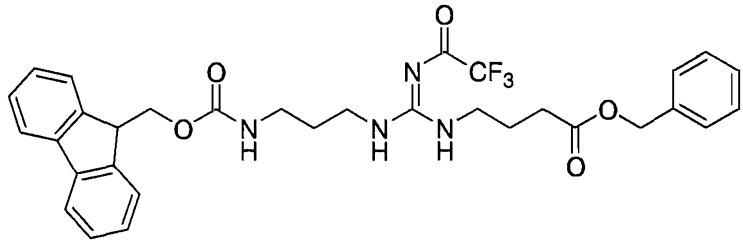
(t, 2H, J = 7.0Hz, CH_2COOBn), 1.94 (quin, 2H, $CH_2CH_2CH_2$), 1.65 (m, 2H, $CH_2CH_2CH_2$), 1.42 (s, 9H, Boc); ^{13}C NMR: (75MHz, $CDCl_3$) δ 173.3 (C), 166.7 (q, J = 35Hz, C), 160.8 (C), 157.2 (C), 135.6 (C), 128.7 (CH), 128.5 (CH), 128.3 (CH), 117.1 (q, J = 286Hz, CF_3), 79.7 (C), 66.9 (CH_2), 61.1 (CH_2), 40.2 (CH_2), 37.5 (CH_2), 36.6 (CH_2), 30.8 (CH_2), 28.4 (CH_3), 23.8 (CH_2); LRMS: (ES^+) m/z 489.3 $[M+H]^+$, 511.3 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{22}H_{31}N_4O_5F_3$ $[M+H]^+$ 489.2320, found 489.2317.

4-((3-[(*tert*-Butoxycarbonyl)amino]propylamino)[(2,2,2-trifluoroacetyl)-imino]methylamino)butanoic acid 161



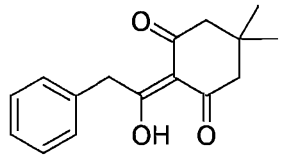
Palladium on carbon (10% by weight, 175mg) was added to a solution of carbamate **160** (804.3mg, 1.65mmol) in EtOAc (20ml) and the mixture stirred under a hydrogen atmosphere for 3 hours. The reaction mixture was filtered through Celite and the residue washed with EtOAc. The combined filtrates were evaporated to give product **161** as a white solid (619.3mg, 94%); m.p. 104-105°C; IR ν_{max} (cm^{-1}) 3281 (w), 2960 (w), 1742 (w), 1682 (w), 1637 (s), 1521 (s), 1448 (m), 1389 (m), 1366 (m), 1158 (s), 660 (m); ^{1}H NMR: (300MHz, d_6-DMSO) δ 12.14 (br s, 1H, COO*H*), 9.04 (br s, 1H, N*H*), 7.78 and 7.63 (br s, 1H, N*H*), 6.85 and 6.75 (br s, 1H, N*H*), 3.29 (m, 2H, NHCH_2), 3.19 (m, 2H, NHCH_2), 2.95 (m, 2H, BocNHCH_2), 2.22 (t, 2H, J = 7.5Hz, CH_2COOH), 1.72 (m, 2H, $CH_2CH_2CH_2$), 1.60 (m, 2H, $CH_2CH_2CH_2$), 1.37 (s, 9H, Boc); ^{13}C NMR: δ 174.2 (C), 160.0 (C), 155.9 (C), 120.8 (q, J = 260Hz, CF_3), 77.5 (C), 40.4 (CH_2), 37.2 (CH_2), 31.1 (CH_2), 29.7 (CH_2), 28.2 (CH_3), 24.6 (CH_2), 23.8 (CH_2) *C*OCF_3 not visible; LRMS: (ES^+) m/z 303.0 $[M+H\text{-}COCF_3]^+$ 421.0 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{15}H_{25}N_4O_5F_3$ $[M+Na]^+$ 421.1669, found 421.1665.

Benzyl 4-([(3-[(9*H*-9-fluorenylmethoxy)carbonyl]aminopropyl)amino][(2,2,2-trifluoroacetyl)imino]methylamino)butanoate (162)



Carbamate **160** (1.25g, 2.56mmol) was stirred in 20% TFA/DCM (30ml) for 1 hour. The solvents were then removed *in vacuo* by azeotropic distillation with toluene. The residue was redissolved in DCM (40ml) and 9-Fluorenylmethoxycarbonyl chloride (0.73g, 2.81mmol) and Et_3N (713µl, 5.12mmol) were added. The mixture was stirred for 1 hour and then evaporated *in vacuo*. The residue was purified by column chromatography (eluent 25% EtOAc/petrol) to give **162** as a colourless oil (1.103g, 71%); IR ν_{max} (cm^{-1}): 2941 (w), 1705 (m), 1621 (s), 1518 (m), 1433 (m), 1185 (s), 1132 (s), 901 (m), 791 (m), 735 (m), 696 (m); ^1H NMR: (300MHz, $CDCl_3$) δ 9.56 (br s, 1H, N*H*), 7.68 (d, 2H, *J* = 7.5Hz, arom), 7.52 (d, 2H, *J* = 7.5Hz, arom), 7.34-7.17 (m, 9H, arom), 6.54 (br s, 1H, N*H*), 5.35 (br s, 1H, N*H*), 5.04 (s, 2H, C*H*$_2$Ph), 4.34 (d, 2H, *J* = 6.5Hz, Fmoc C*H*$_2$), 4.12 (t, 1H, *J* = 6.5Hz, Fmoc C*H*), 3.20-3.02 (m, 4H, C*H*$_2$ × 2), 2.39 (t, 2H, *J* = 6.5Hz, C*H*$_2$COOBn), 1.92-1.53 (m, 6H, C*H*$_2$ × 3); ^{13}C NMR: (75MHz, $CDCl_3$) δ 173.8 (C), 166.9 (q, *J* = 35Hz, *C*OCF_3), 161.0 (C), 157.4 (C), 144.0 (C), 141.5 (C), 135.5 (C), 128.8 (CH), 128.6 (CH), 128.4 (CH), 127.8 (CH), 127.2 (CH), 125.1 (CH), 120.1 (CH), 117.1 (q, *J* = 285Hz), 67.0 (CH_2), 66.9 (CH_2), 47.4 (CH), 40.1 (CH_2), 37.7 (CH_2), 37.1 (CH_2), 30.6 (CH_2), 30.2 (CH_2), 23.7 (CH_2); LRMS: (ES^+) *m/z* 611.2 $[M+H]^+$ 633.2 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{32}H_{33}N_4O_5F_3$ $[M+H]^+$ 611.2476, found 611.2466.

2-(1-Hydroxy-2-phenylethylidene)-5,5-dimethyl-1,3-cyclohexanedione (108)

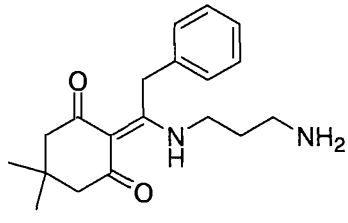


According to a procedure by Bycroft *et al.*[191]

Dimedone (1.634g, 11.7mmol), DCC (2.19g, 10.6mmol) and DMAP (1.31g, 10.7mmol) were added to a solution of phenylacetic acid (1.443g, 10.6mmol) in DMF (100ml). A yellow colour was noted and a white precipitate formed. The mixture was stirred for 3 days and then filtered. The residue was rinsed with DMF, and the combined filtrates evaporated. Recrystallisation of the residue from MeOH gave **108** as white crystals (1.28g, 42%); m.p.: 102-104°C; IR ν_{max} (cm^{-1}): 2961 (w), 2879 (w), 1650 (s), 1556 (s), 1455 (m), 1430 (m), 1407 (m), 1033 (m); ^{1}H NMR: (300MHz, $CDCl_3$) δ 14.15 (br s, 1H, O*H*), 7.24 (m, 5H, arom), 4.33 (s, 2H, C*H*$_2$Ph), 2.51 (s, 2H, COC*H*$_2$), 2.30 (s, 2H, COC*H*$_2$), 1.02 (s, 6H, $CH_3 \times 2$); ^{13}C NMR: (75MHz, $CDCl_3$) δ 203.0 (C), 197.6, (C), 195.3 (C), 134.7 (C), 130.0 (CH), 128.6 (CH), 127.1 (C), 52.7 (CH_2), 46.7 (CH_2), 46.4 (CH_2), 30.9 (C), 28.3 (CH_2); LRMS: decomposes under MS conditions.

Data agrees with literature values.[177]

2-1-[(3-Aminopropyl)amino]-2-phenylethylidene-5,5-dimethyl-1,3-cyclohexanedione (163)

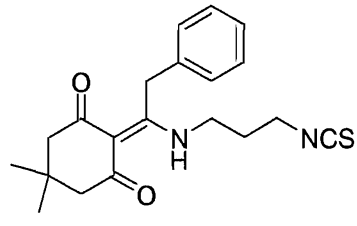


According to a procedure modified from Kilburn *et al.*[168]

Diketone **108** (2.60g, 10.0mmol) was dissolved in DCM (200ml) and added dropwise to a solution of 1,3-diaminopropane (3.4ml, 40.2mmol) and TFA (77μl, 1.00mmol) in DCM

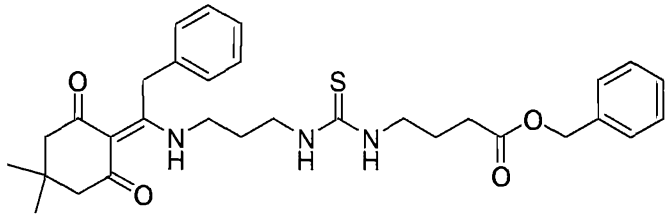
(10ml) over the course of 4 hours. The reaction mixture was stirred for a further 1 hour and then washed with 2M Na_2CO_3(aq) and water. The organic phase was dried over MgSO4 and evaporated to give a yellow oil **163** (3.82g, quant.); ^{1}H NMR: (300MHz, $CDCl_3$) δ 13.65 (br s, 1H, DdpeN*H*), 7.33-7.03 (m, 5H, arom), 4.58 (s, 2H, C*H*$_2$Ph), 3.44 (m, 2H, C*H*$_2$), 2.74 (t, 2H, *J* = 7.0Hz, C*H*$_2$NH$_2$), 2.39 (s, 4H, COC*H*$_2$ × 2), 1.68 (tt, 2H, *J* = 7.0, 7.0Hz, CH$_2$C*H*$_2$CH$_2$), 1.04 (s, 6H, C*H*$_3$ × 2); ^{13}C NMR: (75MHz, $CDCl_3$) 173.7 (C), 136.0 (C), 128.9 (CH), 128.1 (CH), 126.7 (CH), 108.1 (C), 41.0 (CH_2), 40.3 (CH_2), 39.2 (CH_2), 35.2 (CH_2), 32.7 (CH_2), 30.2 (C), 28.4 (CH_3); LRMS: (ES$^+$) 315.1 $[M+H]^+$; HRMS: (ES$^+$) calculated for $C_{19}H_{26}N_2O_2$ $[M+H]^+$ 315.2067, found 315.2063.
Data agreed with literature.[177]

2-1-[(3-Isothiocyanatopropyl)amino]-2-phenylethylidene-5,5-dimethyl-1,3-cyclohexanedione (164)



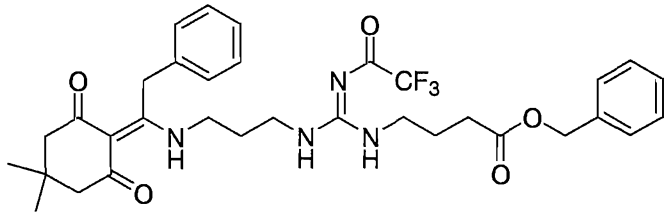
Thiophosgene (1.6ml, 21.0mmol) was added to a biphasic solution of amine **163** (1.65g, 5.25mmol) in DCM (50ml) and 2M K_2CO_3 (35ml). The mixture was stirred for 18 hours and then the phases separated. The organic phase was washed with 2M K_2CO_3(aq), dried over $MgSO_4$ and evaporated. The residue was purified by column chromatography to give **164** as a yellow oil (850mg, 45%); IR ν_{max} (cm^{-1}): 2954 (w), 2867 (w), 2189 (w), 2105 (m), 1636 (m), 1565 (s), 1450 (m), 1366 (m), 1282 (m), 728 (m); ^{1}H NMR (300MHz, $CDCl_3$) δ 13.67 (br s, 1H, DdpeN*H*), 7.29-7.04 (m, 5H, arom), 4.53 (s, 2H, CH_2Ph), 3.48 (t, 2H, *J* = 6.5Hz, C*H*$_2$NCS), 3.41 (dt, 2H, *J* = 6.5, 6.5Hz, NHC*H*$_2$), 2.35 (s, 4H, Ddpe C*H*$_2$ × 2), 1.81 (tt, 2H, *J* = 6.5, 6.5Hz, $CH_2CH_2CH_2$), 0.99 (s, 6H, Ddpe C*H*$_3$ × 2); ^{13}C NMR: (75MHz, $CDCl_3$) δ 198.4 (C), 174.1 (C), 135.6 (C), 129.1 (CH), 128.1 (CH), 126.9 (CH), 108.5 (C), 53.1 (CH_2), 42.4 (CH_2), 40.2 (CH_2), 35.2 (CH_2), 30.2 (C), 29.6 (CH_2), 28.4 (CH_3); LRMS: decomposes under MS conditions.

Benzyl *N*-[3-([(3-[1-(4,4-dimethyl-2,6-dioxocyclohexyliden)-2-phenylethyl]aminopropyl)amino]carbothioylamino)propyl]carbamate (167)



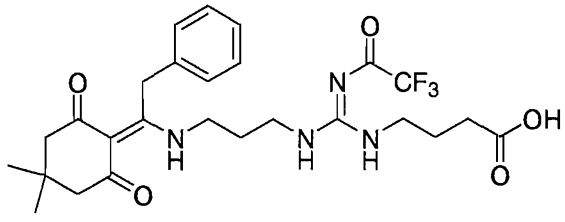
Ester **167** (392.8mg, 1.34mmol) was stirred in 20% TFA/DCM (10ml) for 1 hour. The solvents were then removed *in vacuo* by azeotropic distillation with toluene. The residue was redissolved in 1:1 MeOH/DCM (30ml) and isothiocyanate **164** (0.477g, 1.34mmol) was added followed by Et_3N (373µl, 2.68mmol). The reaction mixture was stirred for 18 hours and then evaporated *in vacuo.* The residue was purified by column chromatography (eluent 40% EtOAc/petrol) to give thiourea **167** as a pale yellow oil (617.9mg, 84%); IR ν_{max} (cm^{-1}): 3272 (br w), 2953 (w), 1731 (m), 1626 (m), 1564 (s), 1450 (m), 1342 (m), 1284 (m), 1164 (m), 730 (s), 696 (s); ^1H NMR: (300MHz, $CDCl_3$) δ 13.54 (br s, 1H, DdpeN*H*), 7.39-6.96 (m, 10H, arom), 6.16 (m, 2H, N*H* × 2), 5.05 (s, 2h, OCH_2Ph), 4.49 (s, 2H, CH_2Ph), 3.48 (m, 2H, CH_2), 3.41-3.22 (m, 4H, CH_2 × 2), 2.35 (q, 2H, *J* = 6.5Hz, CH_2), 2.32 (s, 4H, Ddpe CH_2 × 2), 1.89-1.73 (m, 4H, CH_2 × 2), 0.97 (s, 6H, Ddpe CH_3 × 2); ^{13}C NMR: (75MHz, $CDCl_3$) 182.4 (C), 174.1 (C), 173.9 (C), 135.9 (C), 135.6 (C), 129.0 (CH), 128.8 (CH), 128.6 (CH), 128.4 (CH), 128.2 (CH), 126.8 (CH), 108.3 (C), 67.0 (CH_2), 53.1 (CH_2), 42.0 (CH_2), 41.1 (CH_2), 40.1 (C), 35.3 (CH_2), 31.2 (CH_2), 30.2 (CH_2), 28.9 (CH_2), 28.4 (CH_3), 23.9 (CH_2), 22.9 (CH_2); LRMS: (ES^+) 550.4 $[M+H]^+$ 572.3 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{31}H_{39}N_3O_4S_1$ $[M+Na]^+$ 572.2553, found 572.2557.

Benzyl 4-([(3-[1-(4,4-dimethyl-2,6-dioxocyclohexyliden)-2-phenylethyl]-aminopropyl)amino][(2,2,2-trifluoroacetyl)imino]methylamino)butanoate (169)



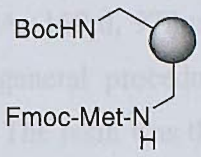
Iodomethane (685µl, 11mmol) was added to a solution of thiourea **167** (604.8mg, 1.10mmol) in acetone (20ml). The mixture was stirred for 3 hours and then all volatilities were removed *in vacuo*. The residue was redissolved in 1:1 MeOH/DCM (30ml) and ammonium hexafluorophosphate was added (360mg, 2.20mmol). The mixture was stirred for 18 hours and then evaporated *in vacuo*. The residue was taken up in DCM and washed once with water. The organic phase was dried over $MgSO_4$ and evaporated *in vacuo*. The residue was redissolved in 4:1 toluene/$CHCl_3$ and trifluoroacetamide (622mg, 5.50mmol) and DBU (329µl, 2.20mmol) were added. The mixture was heated at reflux for two hours. After cooling, the mixture was evaporated *in vacuo*. The residue was purified by column chromatography (eluent 60% EtOAc/petrol) to give carbamate **169** as a colourless, viscous oil (224.2mg, 32%); IR ν_{max} (cm^{-1}): 2955 (w), 1688 (w), 1626 (s), 1563 (s), 1541 (m), 1344 (m), 1171 (s), 1132 (s), 732 (m), 696 (m); ^1H NMR: (300 MHz, $CDCl_3$) δ 13.61 (br s, 1H, DdpeN*H*), 9.56 (br s, 1H, N*H*), 7.29-7.03 (m, 10H, arom), 6.35 (br s, 1H, N*H*), 5.06 (s, 2H, OCH_2Ph), 4.48 (s, 2H, CH_2Ph), 3.55-3.29 (m, 4H, CH_2 × 2), 3.05 (m, 2H, CH_2), 2.35 (m, 2H, CH_2), 2.31 (s, 4H, Ddpe CH_2 × 2) 1.82-1.75 (m, 4H, CH_2 × 2), 0.96 (s, 6H, CH_3 × 2); ^{13}C NMR: (75MHz, $CDCl_3$) δ 198.1 (C), 174.5 (C), 173.9 (C), 166.9 (q, *J* = 35Hz, *C*OCF_3), 160.9 (C), 135.7 (C), 135.4 (C), 128.9 (CH), 128.8 (CH), 128.7 (CH), 128.4 (CH), 128.1 (CH), 126.7 (CH), 117.0 (q, *J* = 285Hz, CF_3), 108.2 (C), 67.2 (CH_2), 53.1 (CH_2), 41.0 (CH_2), 39.9 (CH_2), 39.0 (CH_2), 35.2 (CH_2), 30.2 (CH_2), 29.6 (CH_2), 28.4 (CH_3), 23.6 (CH_2), 4° of Ddpe not visible; LRMS: (ES^+) 629.4 $[M+H]^+$ 651.4 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{33}H_{39}N_4O_5F_3$ $[M+H]^+$ 629.2945, found 629.2943.

4-([(3-[1-(4,4-Dimethyl-2,6-dioxocyclohexyliden)-2-phenylethyl]-aminopropyl)amino][(2,2,2-trifluoroacetyl)imino]methylamino)butanoic acid 170



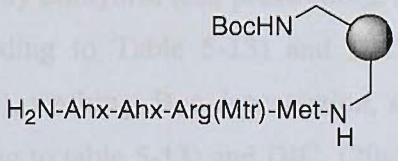
Palladium on carbon (10% by weight, 5mg) was added to a solution of carbamate **169** (30mg, 47.7μmol) in EtOAc (2ml). The mixture was stirred under a hydrogen atmosphere for 2 hours. The reaction mixture was filtered through a plug of Celite, and the residue washed with EtOAc. The combined filtrates were evaporated to dryness giving acid **170** as a hygroscopic white solid (25.2mg, 98%); m.p.: 88-91°C; IR ν_{max} (cm^{-1}): 2930 (w), 1625 (s), 1560 (s), 1541 (m), 1345 (m), 1178 (s), 1129 (s), 799 (m), 720 (m); ^1H NMR: (300MHz, d_6-DMSO) δ 13.46 (br s, 1H, DdpeN*H*), 12.13 (br s, 1H, COO*H*), 9.05 (br s, 1H, N*H*), 7.61 (br s, 1H, N*H*), 7.29-7.10 (m, 5H, arom), 4.54 (s, 2H, C*H*$_2$Ph), 3.47-3.06 (m, 6H, C*H*$_2$ × 3), 2.32 (s, 4H, Ddpe C*H*$_2$ × 2), 2.29-2.17 (m, 2H, C*H*$_2$), 1.80-1.63 (m, 4H, C*H*$_2$ × 2), 0.97 (s, 6H, C*H*$_3$ × 2); ^{13}C NMR: (75MHz, d_6-DMSO) δ 196.8 (C), 174.0 (C), 172.7 (C), 163.7 (q, *J* = 34Hz, *C*OCF$_3$), 159.9 (C), 136.1 (C), 128.5 (CH), 127.9 (CH), 126.2 (CH). 116.9 (q, *J* = 286Hz), 107.0 (C), 52.4 (CH_2), 33.9 (CH_2), 30.8 (CH_2), 29.6 (CH_2), 28.9 (CH_2), 27.8 (CH_3), 24.6 (CH_2), 23.9 (CH_2), 23.6 (CH_2), 4° of Ddpe not visible; LRMS: (ES^+) 539.4 $[M+H]^+$ 561.4 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{26}H_{33}N_4O_5F_3$ $[M+Na]^+$ 561.2295, found 561.2284.

Orthogonally-protected resin (172)



Tentagel resin **88** (loading 0.3mmol/g, 1.08g, 324μmol) was preswollen according to the general procedure. A solution of Fmoc-Met-OH (100.3mg, 270μmol) in DMF (20ml) was added to the resin and the mixture shaken for 1 hour. DIC (152μl, 972μmol), HOBt (131mg, 972μmol) and DIPEA (169μl, 972μmol) were then added to the mixture and shaking continued a further 18 hours. The resin was then washed according to the general procedure, rinsed with Et_2O and dried in vacuo. A quantitative ninhydrin test gave the remaining free NH_2-loading as 0.16mmol/g. The process was then repeated using Fmoc-Met-OH (40mg, 108μmol) to obtain resin which gave a free NH_2-loading of 0.10mmol/g by quantitative ninhydrin test. The resin was then preswollen again according to general procedure and treated with a solution of di-*tert*-butyl dicarbonate (94mg, 434μmol) and DIPEA (169μl, 973μmol) in DMF (20ml) for 2 hours. The resin **172** was then washed according to general procedure and gave a negative ninhydrin test.

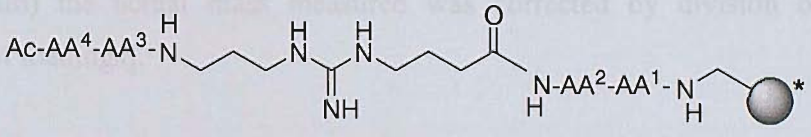
Arg-linker-derivatised bifunctional resin 174



Resin **172** was preswollen according to general procedure. Fmoc-deprotection was then carried out according to the general procedure, followed by standard amino acid coupling according to the general procedure using the following reagents; Fmoc-Arg(Mtr)-OH (394mg, 648μmol), DIC (152μl, 972μmol), HOBt (131mg, 972μmol) and DIPEA (169μl, 972μmol) in DMF (20ml). Washing was then carried out according to the general procedure, followed by Fmoc-deprotection according to the general procedure. The resin

was then subjected to standard amino acid coupling according to the general procedure using the following reagents; Fmoc-Ahx-OH (229mg, 648μmol), DIC (152μl, 972μmol), HOBt (131mg, 972μmol) and DIPEA (169μl, 972μmol) in DMF (20ml). Washing was then carried out according to the general procedure, followed by Fmoc-deprotection according to the general procedure. The resin was then subjected to standard amino acid coupling according to the general procedure using the following reagents; Fmoc-Ahx-OH (229mg, 648μmol), DIC (152μl, 972μmol), HOBt (131mg, 972μmol) and DIPEA (169μl, 972μmol) in DMF (20ml). Washing was then carried out according to the general procedure, followed by Fmoc-deprotection according to the general procedure. All reactions were monitored using the ninhydrin test. The resin **174** was then rinsed with Et_2O and dried *in vacuo*. A quantitative ninhydrin test gave the free NH_2-loading of the resin to be 0.13mmol/g.

Tweezer library 184



Linker-derivatised resin **174** was divided into eight portions. Each portion was then subjected to the following sequence of processes, all according to general procedures, with all reactions monitored by ninhydrin test; preswelling, standard amino acid coupling using Fmoc-AA-OH (according to Table 5-13) and DIC (20μl), HOBt (10mg) and DIPEA (40μl) in DMF (2ml), washing, Boc-deprotection, standard amino acid coupling using Boc-AA-OH (according to table 5-13) and DIC (20μl), HOBt (10mg) and DIPEA (40μl) in DMF (2ml), washing, Fmoc-deprotection and capping reaction (according to Table 5-13). The resin was the combined, suspended in DMF and shaken for 5 mins to ensure complete mixing of the library.

Residue	Mass resin/ mg	Fmoc-strand loading/μmol	Mass Boc-AA-OH/mg	Mass Fmoc-AA-OH/mg	15% Fmoc-loading/μ mol	Vol. 4.72mM *p*-Br PhCOOH soln/μl
Thr(OBn)	140.9	18.3	13	24	2.74	581
Lys(Cbz)	150.2	19.5	17	29	2.93	621
Glu(OBn)	157.0	20.4	16	28	3.06	648
Gly	147.3	19.1	8	17	2.87	608
Ala	142.7	18.6	8	17	2.79	591
Val	151.7	19.7	10	20	2.96	627
Phe	159.9	20.8	13	24	3.12	661
Asn	178.9	23.3	12	25	3.50	742

Table 5-13 Preparation of library 184; Reagents used in first cycle of split-and-mix

The above sequence of procedures was then repeated (using reagents according to Table 5-14) to incorporate the second randomised position. [NB Since mass equivalent to one average amino acid has been added across the entire bead, (i.e. 106 × 0.3 = 32mg per gram of resin) the actual mass measured was corrected by division by 1.032 for calculation of loadings].

Residue	Mass resin/ mg	Fmoc-strand loading/μmol	Mass Boc-AA-OH/mg	Mass Fmoc-AA-OH/mg	15% Fmoc-loading/μ mol	Vol. 4.42mM *p*-Br PhCOOH soln/μl
Thr(OBn)	191.9	24.2	17	31	3.63	821
Lys(Cbz)	142.0	17.9	16	27	2.69	609
Glu(OBn)	131.1	16.5	13	23	2.48	561
Gly	197.0	24.8	10	22	3.72	842
Ala	190.1	23.9	10	22	3.59	812
Val	139.9	17.6	9	18	2.64	597
Phe	188.4	23.7	14	28	3.56	805
Asn	151.0	19.0	10	20	2.85	645

Table 5-14 Preparation of library 184; Reagents used in second cycle of split-and-mix

The resin was then divided into eight portions and subjected to the following sequence of processes, all according to general procedure, with all reactions monitored by ninhydrin

test; preswelling, standard amino acid coupling using Fmoc-AA-OH (according to Table 5-15) and DIC (20μl), HOBt (10mg) and DIPEA (40μl) in DMF (2ml), washing, Boc-deprotection, standard amino acid coupling using acid **151** (according to Table 5-15) and DIC (20μl), HOBt (10mg) and DIPEA (40μl) in DMF (2ml), Boc-deprotection, standard amino acid coupling using Boc-AA-OH (according to Table 5-15) and DIC (20μl), HOBt (10mg) and DIPEA (40μl) in DMF (2ml). Fmoc deprotection and capping reaction (according to Table 5-15). The resin was the combined, suspended in DMF and shaken for 5 mins to ensure complete mixing of the library. [NB. Since mass equivalent to two average amino acid has been added across the entire bead, (i.e. 2 × 106 × 0.3 = 64mg per gram of resin) the actual mass measured was corrected by division by 1.064 for calculation of loadings].

Residue	Mass resin/ mg	Fmoc-strand loading/μmol	Mass Boc-AA-OH/mg	Mass Fmoc-AA-OH/mg	20% Fmoc-loading/μ mol	Vol. 5.75mM *p*-Br PhCOOH soln/μl
Thr(OBn)	156.4	19.1	14	25	3.82	664
Lys(Cbz)	154.8	18.9	17	28	3.78	657
Glu(OBn)	170.2	20.7	16	29	4.14	720
Gly	183.8	22.4	9	20	4.48	779
Ala	161.7	19.7	9	18	3.94	685
Val	178.2	21.7	11	22	4.34	755
Phe	173.8	21.2	13	25	4.24	737
Asn	144.3	17.6	9	19	3.52	612

Residue	Mass 151/mg
Thr(OBn)	15
Lys(Cbz)	15
Glu(OBn)	16
Gly	18
Ala	16
Val	17
Phe	17
Asn	14

Table 5-15 Preparation of library 184; Reagents used in third cycle of split-and-mix

The resin was the divided into eight portions and subjected to the following sequence of processes, all according to general procedure, with all reactions monitored by ninhydrin test; preswelling, standard amino acid coupling using Fmoc-AA-OH (according to Table 5-16) and DIC (20μl), HOBt (10mg) and DIPEA (40μl) in DMF (2ml), washing, Boc-deprotection, standard amino acid coupling using Boc-AA-OH (according to table 5-16) and DIC (20μl), HOBt (10mg) and DIPEA (40μl) in DMF (2ml), washing, Fmoc-deprotection and capping reaction (according to Table 5-16). The resin was the combined, suspended in DMF and shaken for 5 mins to ensure complete mixing of the library. [NB. Since mass equivalent to three average amino acid has been added across the entire bead, plus the CBS in the Boc-strand only (i.e. [3 × 106 × 0.3] + [294 × 0.1] = 124mg per gram of resin) the actual mass measured was corrected by division by 1.124 for calculation of loadings].

Residue	**Mass resin/ mg**	**Fmoc-strand loading/μmol**	**Mass Boc-AA-OH/mg**	**Mass Fmoc-AA-OH/mg**	**Vol. 99mM *p*-Br PhCOOH soln/μl**
Thr(OBn)	173.8	20.0	14	26	1000
Lys(Cbz)	142.6	16.4	14	25	1000
Glu(OBn)	213.6	24.6	19	34	1000
Gly	187.6	21.6	9	19	1000
Ala	147.9	17.0	7	16	1000
Val	190.5	22.0	11	22	1000
Phe	161.8	18.6	11	22	1000
Asn	155.8	18.0	10	19	1000

Table 5-16 Preparation of library 184; Reagents used in fourth (final) cycle of split-and-mix

The resin was then subjected to Boc-deprotection according to general procedure. The resin gave a positive ninhydrin test. The resin was then treated with a solution of Ac_2O (100μl) and DIPEA (100μl) in DMF (2ml) for 2 hours. The resin was washed according to general procedure and gave a negative ninhydrin test. Deprotection of the trifluoroacetyl group was then performed by shaking the resin with a solution of 0.15M K_2CO_3 in 2:2:1 MeOH/DMF/H_2O for 3 hours. The resin was then washed with 2:2:1 MeOH/DMF/H_2O (~20ml per g of resin, three times), then DMF (~20ml per g of resin,

three times) and then according to general procedure. The resin **184** was finally rinsed with Et_2O and dried in vacuo. 5 single beads were selected, cleaved and analysed according to the general procedures.

MALDI comments: Spectra noisy. Sequences were identified from each bead as follows;

Bead 1: 812.4/814.4 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 883.4/885.4 ($[M+H]^+$ *p*-Br PhCONH-Ala-Phe-Ahx-Ahx-Arg-Hsl), 954.3/956.3 ($[M+H]^+$ *p*-Br PhCONH-Ala-Ala-Phe-Ahx-Ahx-Arg-Hsl), 1011.2/1013.2 ($[M+H]^+$ *p*-Br PhCONH-Gly-Ala-Ala-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer structure runs Gly-Ala-CBS-Ala-Phe

Bead 2: 764.3/766.3 ($[M+H]^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 865.1/867.1 ($[M+H]^+$ *p*-Br PhCONH-Thr-Val-Ahx-Ahx-Arg-Hsl), 1011.9/1013.9 ($[M+H]^+$ *p*-Br PhCONH-Phe-Thr-Val-Ahx-Ahx-Arg-Hsl), 1158.6/1160.6 ($[M+H]^+$ *p*-Br PhCONH-Phe-Phe-Thr-Val-Ahx-Ahx-Arg-Hsl), hence tweezer structure runs Phe-Phe-CBS-Thr(OBn)-Val

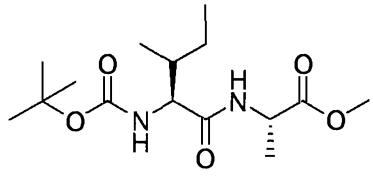
Bead 3: 812.0/814.0 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 958.8/960.9 ($[M+H]^+$ *p*-Br PhCONH-Phe-Phe-Ahx-Ahx-Arg-Hsl), 1015.8/1017.7 ($[M+H]^+$ *p*-Br PhCONH-Gly-Phe-Phe-Ahx-Ahx-Arg-Hsl), 1114.6/1116.5 ($[M+H]^+$ *p*-Br PhCONH-Val-Gly-Phe-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer structure runs Val-Gly-CBS-Phe-Phe

Bead 4: 779.2/781.3 ($[M+H]^+$ *p*-Br PhCONH-Asn-Ahx-Ahx-Arg-Hsl), 836.2/838.1 ($[M+H]^+$ *p*-Br PhCONH-Gly-Asn-Ahx-Ahx-Arg-Hsl), 893.1/895.1 ($[M+H]^+$ *p*-Br PhCONH-Gly-Gly-Asn-Ahx-Ahx-Arg-Hsl), 1021.8/1023.8 ($[M+H]^+$ *p*-Br PhCONH-Glu-Gly-Gly-Asn-Ahx-Ahx-Arg-Hsl), hence tweezer structure Glu(OBn)-Gly-CBS-Gly-Asn

Bead 5: 794.1/796.0 ($[M+H]^+$ *p*-Br PhCONH-Glu-Ahx-Ahx-Arg-Hsl), 907.7/909.7 ($[M+H]^+$ *p*-Br PhCONH-Asn-Glu-Ahx-Ahx-Arg-Hsl), 1006.5/1008.6 ($[M+H]^+$ *p*-Br PhCONH-Val-Asn-Glu-Ahx-Ahx-Arg-Hsl), 1077.5/1079.4 ($[M+H]^+$ *p*-Br PhCONH-Ala-Val-Asn-Glu-Ahx-Ahx-Arg-Hsl), hence tweezer structure runs Ala-Val-CBS-Asn-Glu(OBn)

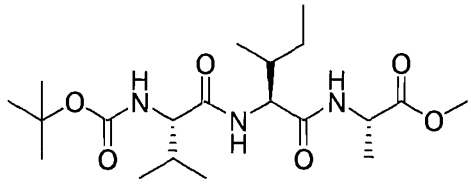
5.8 Experimental for Chapter 4

Methyl (2*S*)-2-((2*S*,3*R*)-2-[(*tert*-butoxycarbonyl)amino]-3-methylpentanoyl-amino)propanoate (186)



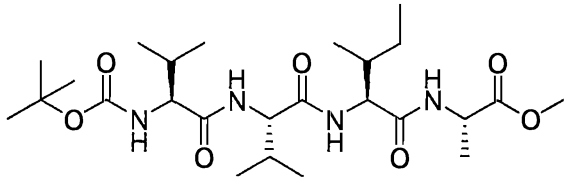
DIC (918μl, 5.86mmol) was added to a solution of *N*-α-Boc-Ile-OH (1.281g, 5.33mmol), alanine methyl ester hydrochloride (706mg, 5.06mmol), HOBt (1.44g, 10.66mmol) and DIPEA (1.9ml, 10.66mmol) in freshly distilled THF (10ml) and freshly distilled DCM (10ml). The mixture was stirred for 3.5 hours and then evaporated *in vacuo*. The residue was redissolved in $CHCl_3$ and washed sequentially with saturated $NaHCO_3$(aq), 1M $KHSO_4$(aq), water and brine. The organic phase was dried over $MgSO_4$ and evaporated *in vacuo*. The residue was then purified by column chromatography (50% EtOAc/petrol) to give **186** as a white solid (1.39g, 82%); m.p.: 145-146°C; IR ν_{max} (cm^{-1}): 3303 (w), 2967 (w), 1747 (m), 1678 (m), 1650 (m), 1525 (m), 1295 (m), 1251 (m), 1168 (m), 631 (m); ^{1}H NMR: (300MHz, $CDCl_3$) δ 6.44 (d, 1H, *J* = 6.5Hz, Ala N*H*), 5.05 (br s, 1H, Ile N*H*), 4.58 (dq, 1H, *J* = 7.5, 6.5Hz, Ala α-C*H*), 3.95 (m, 1H, Ile α-C*H*), 3.74 (s, 3H, OMe), 1.86 (m, 1H, Ile β-C*H*), 1.44 (s, 9H, Boc), 1.40 (d, 3H, *J* = 7.5Hz, Ala CH_3), 1.23-1.04 (m, 2H, Ile CH_2), 0.94 (d, 3H, *J* = 7.0 Hz, Ile CH_3CH), 0.91 (t, 3H, *J* = 7.5Hz, Ile CH_3CH_2); ^{13}C NMR: (75MHz, $CDCl_3$) δ 173.1 (C), 171.1 (C), 155.7 (C), 79.9 (C), 59.2 (CH), 52.4 (CH_3), 48.0 (CH), 37.4 (CH), 28.3 (CH_3), 24.7 (CH_2), 18.3 (CH_3), 15.5 (CH_3), 11.4 (CH_3); LRMS: (ES^+) *m/z* 317.1 $[M+H]^+$ 339.1 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{15}H_{28}N_2O_5$ $[M+Na]^+$ 339.1890, found 339.1887.

Methyl (2*S*)-2-[(2*S*,3*R*)-2-((2*S*)-2-[(*tert*-butoxycarbonyl)amino]-3-methylbutanoyl-amino)-3-methylpentanoyl]aminopropanoate (187)



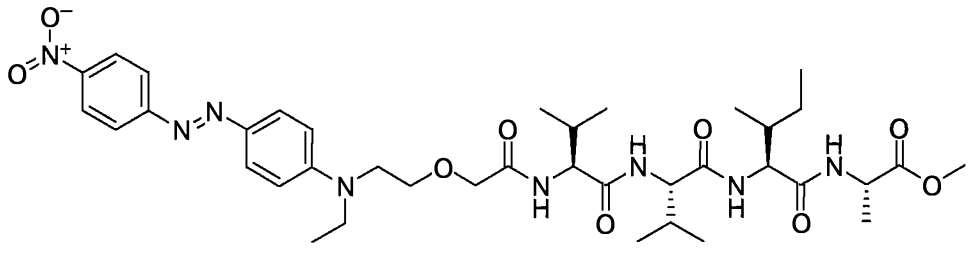
Dipeptide **186** (69mg, 0.22mmol) was stirred in 20% TFA/DCM (2ml) for 1 hour. The solvents were then removed *in vacuo* by azeotropic distillation with toluene. The residue was redissolved in THF (2ml) and Et_3N (61µl, 0.44mmol) was added. The resulting solution was then added to a solution of *N*-α-Boc-Val-OH (52mg, 0.24mmol) and CDI (42mg, 0.26mmol) in THF (which had already been stirred for 1 hour). The resulting solution was stirred a further 18 hours and then evaporated *in vacuo*. The residue was redissolved in EtOAc and washed sequentially with 1M $KHSO_4$(aq), saturated $NaHCO_3$(aq) and brine. The organic phase was dried over $MgSO_4$ and evaporated to give **187** as a white solid (54.1mg, 59%); m.p.: 172-174°C; IR ν_{max} (cm^{-1}), 3282 (m), 2961 (m), 1742 (m), 1682 (m), 1638 (s), 1523 (s), 1448 (m), 1387 (m), 1365 (m), 1245 (m), 1211 (m), 1158 (s), 1049 (m), 1024 (m), 661 (s); ^{1}H NMR: (400MHz, $CDCl_3$) δ 6.58-6.50 (m, 2H, N*H* × 2), 5.03 (br s, 1H, N*H*), 4.56 (dq, 1H, *J* = 7.5, 7.5Hz), 4.30 (dd, 1H, *J* = 8.5, 7.0Hz, Ile α-C*H*), 3.91 (dd, 1H, *J* = 7.0, 6.5Hz, Val α-C*H*), 3.74 (s, 3H, OMe), 2.14 (m, 1H, Val β-C*H*), 1.92 (m, 1H, Ile β-C*H*), 1.51 (m, 1H, Ile γ-C*H*H'), 1.44 (s, 9H, Boc), 1.39 (d, 3H, *J* = 7.5Hz, Ala C*H*$_3$), 1.14 (m, 1H, Ile γ-CH*H'*), 0.97-0.87 (m, 12H, C*H*$_3$ × 4); ^{13}C NMR: (100MHz, $CDCl_3$) δ 173.1 (C), 171.8 (C), 170.6 (C), 156.0 (C), 80.2 (C), 61.2 (CH), 57.8 (CH), 52.6 (CH), 48.2 (CH_3), 37.1 (CH), 30.6 (CH), 28.4 (CH_3), 24.9 (CH_2), 19.5 (CH_3), 18.3 (CH_3), 18.0 (CH_3), 15.6 (CH_3), 11.4 (CH_3); LRMS: (ES^+) *m/z* 416.1 $[M+H]^+$ 438.1 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{20}H_{37}N_3O_6$ $[M+Na]^+$ 438.2574, found 438.2568.

Methyl (2*S*)-2-[((2*S*,3*R*)-2-[(2*S*)-2-((2*S*)-2-[(*tert*-butoxycarbonyl)amino]-3-methylbutanoylamino)-3-methylbutanoyl]amino-3-methylpentanoyl)-amino]propanoate (188)



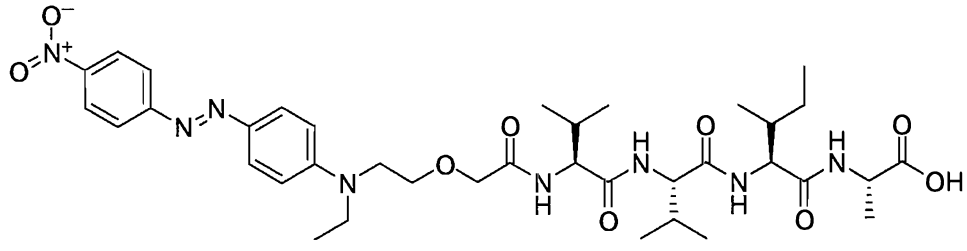
Tripeptide **187** (745mg, 1.79mmol) was stirred in 20% TFA/DCM (20ml) for 1 hour. The solvents were then removed *in vacuo* by azeotropic distillation with toluene. The residue was redissolved in DMF (20ml) and *N*-α-Boc-Val-OH (428mg,1.97mmol), HOBt (484mg, 3.58mmol), DIPEA (624μl, 3.58mmol), DMAP (10mg) were added followed by EDC (378mg, 1.97mmol). The reaction mixture was stirred for 18 hours and then evaporated *in vacuo*. The residue was redissolved in $CHCl_3$ and sequentially washed with saturated $NaHCO_3$(aq), 1M $KHSO_4$(aq) and brine. The organic phase was dried over $MgSO_4$ and evaporated to give **188** as a white solid (809.9mg, 88%); m.p.: >250°C; IR ν_{max} (cm^{-1}) 3279 (br w), 2965 (w), 1636 (s), 1520 (m), 1455 (m), 1366 (m), 1214 (m), 1161 (s), 754 (w); ^1H NMR: (300MHz, $CDCl_3$) δ 8.29 (m, 2H, N*H* × 2), 8.19 (d, 1H, *J* = 7.0Hz, N*H*), 6.02 (d, 1H, *J* = 9.0Hz, N*H*), 4.70 (m, 1H, α-C*H*), 4.58-4.53 (m, 2H, α-C*H* × 2), 4.32 (m, 1H, α-C*H*), 3.71 (s, 3H, OMe), 2.00 (m, 2H, β-C*H* × 2), 1.86 (m, 1H, β-C*H*), 1.50 (m, 1H, Ile γ-C*H*H'), 1.40 (s, 9H, Boc), 1.33 (d, 3H, *J* = 7.0Hz, Ala CH_3), 1.06 (m, 1H, Ile γ-CH*H'*), 0.95-0.81 (m, 15H, CH_3 × 5), 0.77 (t, 3H, *J* = 7.5Hz, Ile CH_3CH_2); ^{13}C NMR: (75MHz, $CDCl_3$) δ 173.2 (C), 172.4 (C), 172.0 (C), 171.7 (C), 156.1 (C), 78.9 (C), 60.0 (CH), 58.1 (CH), 57.6 (CH), 52.3 (CH_3), 48.0 (CH), 36.5 (CH), 31.8 (CH), 31.6 (CH), 28.6 (CH_3), 25.1 (CH_2), 19.1 (CH_3), 19.0 (CH_3), 18.9 (CH_3), 18.7 (CH_3), 17.6 (CH_3), 15.2 (CH_3), 11.0 (CH_3); LRMS: (ES^+) *m/z* 537.5 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{25}H_{46}N_4O_7$ $[M+H]^+$ 515.3439, found 515.3442.

Methyl (2*S*)-2-[(2*S*,3*R*)-2-((2*S*)-2-[((2*S*)-2-[2-(2-ethyl-4-[(*E*)-2-(4-nitrophenyl)-1-diazenyl]anilinoethoxy)acetyl]amino-3-methylbutanoyl)amino]-3-methylbutanoyl-amino)-3-methylpentanoyl]aminopropanoate (189)



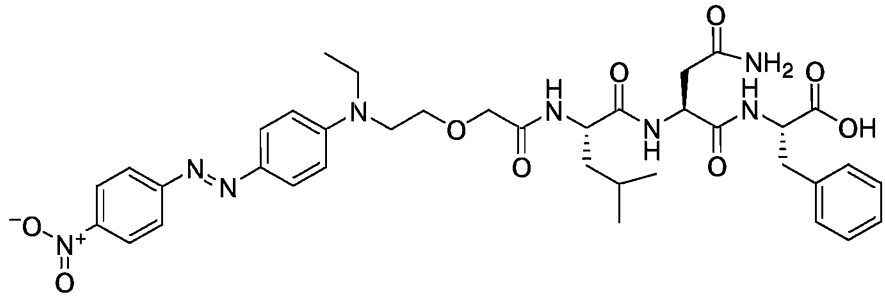
Tetrapeptide **188** (263.7mg, 0.51mmol) was stirred in 20% TFA/DCM (10ml) for 1 hour. The solvents were then removed *in vacuo* by azeotropic distillation with toluene. The residue was redissolved in 1:1 THF/DCM (20ml) and **117** (210mg, 0.56mmol), HOBt (138mg, 1.02mmol), DIPEA (446µl, 2.56mmol) and DMAP (6mg) were added followed by PyBOP (293mg, 0.56mmol). The mixture was stirred overnight and then evaporated *in vacuo*. The residue was purified by column chromatography (eluent 5% MeOH/DCM) to give **189** as a red solid (217.2mg, 55%); m.p.: 171-174°C; IR ν_{max} (cm^{-1}): 3276 (w), 2965 (w), 1632 (s), 1515 (s), 1385 (m), 1337 (s), 1137 (s), 858 (m); ^{1}H NMR: (300MHz, d_6-DMSO) δ 8.36 (d, 2H, *J* = 9.0Hz, arom), 8.32 (m, 1H, N*H*), 8.03 (d, 1H, *J* = 7.5Hz, N*H*), 7.92 (d, 2H, *J* = 8.5Hz, arom), 7.83 (d, 2H, *J* = 9.0Hz, arom), 7.74 (d, 1H, *J* = 9.0Hz, N*H*), 7.40 (d, 1H, *J* = 9.0Hz, N*H*), 6.91 (d, 2H, *J* = 8.5Hz, arom), 4.39-4.11 (m, 4H, α-C*H* × 4), 4.05 (s, 3H, OMe), 3.96 (s, 2H, OCH_2CO), 3.70 (m, 4H, CH_2 × 2), 3.59 (m, 2H, CH_2), 1.91 (m, 2H), 1.73 (m, 2H), 1.41 (m, 1H), 1.26 (d, 3H, *J* = 7.5Hz, CH_3), 0.89-0.66 (m, 21H, CH_3 × 7); ^{13}C NMR: (75MHz, d_6-DMSO) δ 172.8 (C), 170.9 (C), 170.7 (C), 170.6 (C), 168.5 (C), 156.3 (C), 151.8 (C), 146.9 (C), 142.8 (C), 126.2 (CH), 125.0 (CH), 122.5 (CH), 111.7 (CH), 69.9 (CH_2), 68.7 (CH_3), 58.0 (CH_2), 56.6 (CH_2), 56.3 (CH_2), 51.7 (CH), 49.3 (CH), 47.5 (CH), 45.2 (CH), 36.7 (CH), 30.9 (CH), 30.2 (CH), 24.1 (CH_3), 19.3 (CH_3), 19.1 (CH), 18.3 (CH_3), 17.8 (CH_3), 16.8 (CH_3), 15.0 (CH_3), 12.0 (CH_3), 10.8 (CH_3); LRMS: (ES^+) m/z 791.3 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{38}H_{56}N_8O_9$ $[M+Na]^+$ 791.4063, found 791.4073

(2*S*)-2-[(2*S*,3*R*)-2-((2*S*)-2-[((2*S*)-2-[2-(2-Ethyl-4-[(*E*)-2-(4-nitrophenyl)-1-diazenyl]anilinoethoxy)acetyl]amino-3-methylbutanoyl)amino]-3-methylbutanoyl-amino)-3-methylpentanoyl]aminopropanoic acid (190)



LiOH(aq) (1M, 4.7ml, 4.7mmol) was added to a solution of dye-labelled ester **189** (148mg, 0.19mmol) in dioxane (3ml). The mixture was stirred for 18 hours and then concentrated *in vacuo*. The residue was acidified with conc. HCl and extracted with copious amounts of EtOAc. The organic phase was dried over $MgSO_4$ and evaporated to give **190** as a red solid (78.3mg, 55%); m.p.: 238-242°C, IR ν_{max} (cm^{-1}): 3272 (w), 2963 (w), 1632 (s), 1550 (m), 1512 (m), 1337 (s), 1258 (s), 1144 (m), 1105 (s), 1017 (s), 853 (m), 801 (s); ^{1}H NMR: (300MHz, d_6-DMSO) δ 8.36 (d, 2H, *J* = 9.0Hz, arom), 8.13 (d, 2H, *J* = 7.0Hz, N*H*), 8.03 (d, 1H, *J* = 9.0Hz, N*H*), 7.93 (d, 2H, *J* = 9.0Hz, arom), 7.83 (d, 2H, *J* = 9.0Hz, arom), 7.72 (d, 1H, *J* = 9.0Hz, N*H*), 7.41 (d, 1H, *J* = 9.0Hz, N*H*), 6.91 (d, 2H, *J* = 9.0Hz, arom), 4.31 (m, 1H, α-C*H*), 4.25-4.10 (m, 3H, α-C*H* × 3), 3.97 (s, 2H, OCH_2CO), 3.70 (m, 4H, NCH_2CH_2O), 3.44 (q, 2H, *J* = 7.0Hz, NCH_2CH_3), 1.90 (m, 2H, β-C*H* × 2), 1.71 (m, 1H, β-C*H*), 1.48 (m, 1H, Ile γ-C*H*H'), 1.24 (d, 3H, *J* = 7.5Hz, Ala CH_3), 1.17 (t, 3H, *J* = 7.0Hz, NCH_2CH_3), 1.06 (m, 1H, Ile γ-CH*H'*), 0.98-0.67 (m, 15H, CH_3 × 5); ^{13}C NMR: (75MHz, d_6-DMSO) δ 173.8 (C), 170.7 (C), 170.5 (C), 168.9 (C), 168.4 (C), 156.2 (C), 151.7 (C), 146.8 (C), 142.7 (C), 126.1 (CH), 124.9 (CH), 122.4 (CH), 111.6 (CH), 69.9 (CH_2), 68.6 (CH_2), 57.9 (CH), 56.6 (CH), 56.3 (CH), 49.3 (CH_2), 47.4 (CH), 45.1 (CH_2), 36.7 (CH), 30.9 (CH), 30.2 (CH), 24.0 (CH_2), 19.2 (CH_3), 19.1 (CH_3), 18.3 (CH_3), 17.8 (CH_3), 17.0 (CH_3), 15.1 (CH_3), 12.0 (CH_3), 10.8 (CH_3); LRMS: (ES^+) *m/z* 755.4 $[M+H]^+$ 777.4 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{38}H_{56}N_8O_8$ $[M+H]^+$ 755.4086, found 755.4079.

(2*S*)-2-((2*S*)-4-Amino-2-[((2*S*)-2-[2-(2-ethyl-4-[(*E*)-2-(4-nitrophenyl)-1-diazenyl]-anilinoethoxy)acetyl]amino-4-methylpentanoyl)amino]-4-oxobutanoylamino)-3-phenylpropanoic acid (195)



2-chlorotrityl resin (200mg, 0.2mmol) was preswollen in freshly distilled DCM for 10 mins and then drained. A solution of *N*-α-Fmoc-Phe-OH (155mg, 0.4mmol) in freshly distilled DCM (8ml) was added to the resin with DIPEA (105μl, 0.6mmol) and the mixture shaken for 18 hours. The resin was then drained, washed with MeOH (~20ml per g of resin, three times) to react any remaining sites. The resin was then washed according to the general procedure. Fmoc-deprotection according to the general procedure was then carried out. Amino acid coupling using Fmoc-Asn-OH (213mg, 0.6mmol), DIC (94μl, 0.6mmol), HOBt (81mg, 0.6mmol) and DIPEA (105μl, 0.6mmol) in DMF (8ml) was then performed according to the general procedure. The resin was then washed according to the general procedure and subjected to Fmoc-deprotection according to the general procedure. Amino acid coupling using Fmoc-Leu-OH (212mg, 0.6mmol), DIC (94μl, 0.6mmol), HOBt (81mg, 0.6mmol) and DIPEA (105μl, 0.6mmol) in DMF (8ml) was then performed according to the general procedure. The resin was washed according to general procedure and subjected to Fmoc-deprotection according to the general procedure. All reactions were monitored using the ninhydrin test. Standard amino acid coupling according to the general procedure was then performed using the following reagents; **117** (0.179g, 0.48mmol), HOBt (92mg, 0.68mmol) PyBOP (354mg, 0.68mmol) and DIPEA (118μl, 0.68mmol) in DMF (8ml), for 18 hours. The resin was then washed twice according to the general procedure; the resin was noted to be stained a deep red colour. The resin was then treated with 2% TFA/DCM (8ml) for 90 mins. The resin was drained and rinsed with chloroform, and the combined filtrates evaporated to

give **195** as a red solid (178mg, quant.); m.p.: 180-184°C; ^{1}H NMR: (400MHz, DMSO-d_6) δ 8.35 (d, 2H, *J* = 7.5Hz, arom), 8.22 (d, 1H, 7.3Hz, N*H*), 7.91 (d, 2H, *J* = 7.5Hz, arom), 7.85 (m, 1H, N*H*), 7.83 (d, 2H, *J* = 9.0Hz, arom), 7.54 (d, 1H, *J* = 8.5Hz, N*H*), 7.28 – 7.19 (m, 5H, arom), 6.91 (d, 2H, *J* = 9.0Hz, arom), 4.56 (m, 1H, α-C*H*), 4.39 (m, 2H, α-C*H*), 3.95 (s 2H, COCH_2O), 3.69 (m, 4H, NCH_2CH_2O), 3.56 (q, 2H, *J* = 7.0Hz, NCH_2CH$_3$), 3.01 (dd, 1H, *J* = 13.5, 5.0Hz, β-C*H*H'), 2.91 (dd, 1H, *J* = 13.5, 7.0Hz, β-CH*H*'), 2.53 (dd, 1H, *J* = 15.5, 5.0Hz, β-C*H*H'), 2.39 (dd, 1H, *J* = 15.5, 8.0Hz, β-CH*H*'), 1.48 (m, 1H, Leu γ-C*H*), 1.38 (m, 2H, Leu β-CH_2), 1.16 (t, 3H, *J* = 7.0Hz, NCH$_2$CH_3), 0.79 (m, 6H, Leu CH_3 × 2); ^{13}C NMR: (100MHz, DMSO-d_6) δ 178.1 (C), 177.2 (C), 177.0 (C), 176.4 (C), 174.4 (C), 161.9 (C), 157.4 (C), 152.3 (C) 148.4 (C), 142.8 (C), 134.8 (CH), 133.8 (CH), 132.1 (CH), 131.8 (CH), 130.6 (CH), 128.1 (CH), 117.3 (CH), 75.5 (CH_2), 74.3 (CH_2), 59.1 (CH_2), 55.8 (CH_2), 55.1 (CH), 54.9 (CH), 50.7 (CH), 46.6 (CH_2), 42.5 (CH_2), 42.3 (CH_2), 29.8 (CH), 28.7(CH_3), 27.1 (CH_3), 17.6 (CH_3); LRMS: (ES^+) *m/z* 747.3 $[M+H]^+$ 769.3 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{37}H_{46}N_8O_9$ $[M+H]^+$ 747.3461, found 747.3464.

Screening experiments with Single-armed receptor library 149

General procedure for all screens: A mass of single-armed receptor library resin **149** (sufficient to represent five copies of the library, ~5mg) was placed in a flat-bottomed glass dish and pre-swollen in the buffer indicated for 1 hour. A solution of dye-labelled guest **190** in the same buffer was then added to the mixture and the system incubated for the indicated time, with further aliquots of guest **190** added as indicated. Heavily stained beads were then removed from the mixture using a micropipette and each individual bead subjected to cleavage and analysis according to the general procedure.

First screen:

Mass resin **149** = 6.4mg

Buffer system 50% DMSO/borax

Guest **190** concentration = 68.9μM

Library resin swollen in 400μl buffer, then 10μl dye titrated followed by 24 hours incubation. This cycle repeated for a further

5 days (total addition 60μl) followed by the addition of 20μl guest and 24 hours incubation (repeated for 4 days, total addition 140μl), followed by the addition of 100μl dye and 24 hours incubation (repeated 2 days, total addition 340μl), followed by a final addition of 300μl guest and 24 hours incubation. Total addition of dye, 640μl, final concentration 42.4μM. 12 beads were selected, results detailed below.

Bead 1: 812.3/814.3 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 911.3/913.2 ($[M+H]^+$ *p*-Br PhCONH-Val-Phe-Ahx-Ahx-Arg-Hsl), 1040.0/1042.1 ($[M+H]^+$ *p*-Br PhCONH-Glu-Val-Phe-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Glu(O^{t}Bu)-Val-Phe

Bead 2: 764.5/766.5 ($[M+H]^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 877.4/879.4 ($[M+H]^+$ *p*-Br PhCONH-Leu-Val-Ahx-Ahx-Arg-Hsl), 990.2/992.2 ($[M+H]^+$ *p*-Br PhCONH-Leu-Leu-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Leu-Val

Bead 3: 722.5/724.5 ($[M+H]^+$ *p*-Br PhCONH-Gly-Ahx-Ahx-Arg-Hsl), 836.3/838.3 ($[M+H]^+$ *p*-Br PhCONH-Asn-Gly-Ahx-Ahx-Arg-Hsl), 907.2/909.1 ($[M+H]^+$ *p*-Br PhCONH-Ala-Asn-Gly-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Ala-Asn-Gly

Bead 4: 793.4/795.4 ($[M+H]^+$ *p*-Br PhCONH-Lys-Ahx-Ahx-Arg-Hsl), 892.2/894.3 ($[M+H]^+$ *p*-Br PhCONH-Val-Lys-Ahx-Ahx-Arg-Hsl), 991.1/993.1 ($[M+H]^+$ *p*-Br PhCONH-Val-Val-Lys-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Val-Val-Lys(Boc)

Bead 5: 778.3/780.3 ($[M+H]^+$ *p*-Br PhCONH-Leu-Ahx-Ahx-Arg-Hsl), 891.1/893.2 ($[M+H]^+$ *p*-Br PhCONH-Leu-Leu-Ahx-Ahx-Arg-Hsl), 989.8/991.7 ($[M+H]^+$ *p*-Br PhCONH-Val-Leu-Leu-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Val-Leu-Leu

Bead 6: 779.4/781.4 ($[M+H]^+$ *p*-Br PhCONH-Asn-Ahx-Ahx-Arg-Hsl), 877.1/879.2 ($[M+H]^+$ *p*-Br PhCONH-Val-Asn-Ahx-Ahx-Arg-Hsl), 990.2/992.1 ($[M+H]^+$ *p*-Br PhCONH-Leu-Val-Asn-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Val-Asn

Bead 7: 764.3/766.3 ($[M+H]^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 863.2/865.1 ($[M+H]^+$ *p*-Br PhCONH-Val-Val-Ahx-Ahx-Arg-Hsl), 975.9/977.9 ($[M+H]^+$ *p*-Br PhCONH-Leu-Val-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Val-Val

Bead 8: 764.3/766.3 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 893.1/895.1 ([M+H]$^+$ *p*-Br PhCONH-Glu-Val-Ahx-Ahx-Arg-Hsl), 992.0/994.0 ([M+H]$^+$ *p*-Br PhCONH-Val-Glu-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Val-Glu(O^{t}Bu)-Val

Bead 9: 812.2/814.3 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 911.1/913.1 ([M+H]$^+$ *p*-Br PhCONH-Val-Phe-Ahx-Ahx-Arg-Hsl), 1010.0/1012.0 ([M+H]$^+$ *p*-Br PhCONH-Val-Val-Phe-Ahx-Ahx-Arg-Hsl) hence tripeptide sequence Val-Val-Phe

Bead 10: 764.3/766.3 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 877.1/879.2 ([M+H]$^+$ *p*-Br PhCONH-Leu-Val-Ahx-Ahx-Arg-Hsl), 990.0/992.0 ([M+H]$^+$ *p*-Br PhCONH-Leu-Leu-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Leu-Val

Bead 11: 764.5/766.4 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 863.4/865.3 ([M+H]$^+$ *p*-Br PhCONH-Val-Val-Ahx-Ahx-Arg-Hsl), 976.2/978.1 ([M+H]$^+$ *p*-Br PhCONH-Leu-Val-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Val-Val

Bead 12: 764.5/766.5 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 877.2/879.2 ([M+H]$^+$ *p*-Br PhCONH-Leu-Val-Ahx-Ahx-Arg-Hsl), 948.1/950.0 ([M+H]$^+$ *p*-Br PhCONH-Ala-Leu-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Ala-Leu-Val

Second screen: Mass resin **149** = 7.9mg
Buffer system 50% DMSO/borax
Guest **190** concentration = 68.9μM
Library resin swollen in 400μl buffer. 300μl of guest added and incubated for 3 days. Final concentration 29.5μM. 6 beads selected, results detailed below.

Bead 1: 764.5/766.5 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 863.5/865.5 ([M+H]$^+$ *p*-Br PhCONH-Val-Val-Ahx-Ahx-Arg-Hsl), 1010.2/1012.2 ([M+H]$^+$ *p*-Br PhCONH-Phe-Val-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Phe-Val-Val

Bead 2: 764.5/766.5 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 863.5/865.5 ([M+H]$^+$ *p*-Br PhCONH-Val-Val-Ahx-Ahx-Arg-Hsl), 992.2/994.2 ([M+H]$^+$ *p*-Br PhCONH-Glu-Val-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Glu(O^{t}Bu)-Val-Val

Bead 3: 752.5/754.5 ([M+H]$^+$ *p*-Br PhCONH-Ser-Ahx-Ahx-Arg-Hsl), 823.4/825.4 ([M+H]$^+$ *p*-Br PhCONH-Ala-Ser-Ahx-Ahx-Arg-Hsl), 880.3/882.3 ([M+H]$^+$ *p*-Br PhCONH-Gly-Ala-Ser-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Gly-Ala-Ser(O^{t}Bu)
Bead 4: 764.4/766.4 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 863.3/865.3 ([M+H]$^+$ *p*-Br PhCONH-Val-Val-Ahx-Ahx-Arg-Hsl), 976.0/978.1 ([M+H]$^+$ *p*-Br PhCONH-Leu-Val-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Val-Val
Bead 5: no sequencing results obtained
Bead 6:764.3/766.3 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 877.3/879.2 ([M+H]$^+$ *p*-Br PhCONH-Leu-Val-Ahx-Ahx-Arg-Hsl), 990.2/991.8 ([M+H]$^+$ *p*-Br PhCONH-Leu-Leu-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Leu-Val

Third screen:

Mass resin **149** = 7.2mg
Buffer system 50% DMSO/HEPES
Guest **190** concentration = 278.19μM
Library resin swollen in 400μl buffer. 40μl of guest titrated, followed by 24 hours incubation. A further 20μl dye added, followed by 2 days incubation. Total addition 60μl, final concentration 36.3μM. 16 beads were selected, results detailed below.

Bead 1: 764.4/766.6 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 877.1/879.2 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Val-Ahx-Ahx-Arg-Hsl), 976.2/978.1, ([M+H]$^+$ *p*-Br-PhCONH-Val-Leu-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Val-Leu-Val
Bead 2: 795.2/797.2 ([M+H]$^+$ *p*-Br-PhCONH-Glu-Ahx-Ahx-Arg-Hsl), 894.5/896.4 ([M+H]$^+$ *p*-Br-PhCONH-Val-Glu-Ahx-Ahx-Arg-Hsl), 993.5/995.6 ([M+H]$^+$ *p*-Br-PhCONH-Val-Val-Glu-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Val-Val-Glu(O^{t}Bu)
Bead 3: 778.4/780.4 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Ahx-Ahx-Arg-Hsl), 891.2/893.2 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Leu-Ahx-Ahx-Arg-Hsl), 1019.9/1021.9 ([M+H]$^+$ *p*-Br-PhCONH-Glu-Leu-Leu-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Glu(O^{t}Bu)-Leu-Leu

Bead 4: 764.3/766.3 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 877.1/879.1 M+H]$^+$ *p*-Br-PhCONH-Leu-Val-Ahx-Ahx-Arg-Hsl), 1005.8/1007.9 M+H]$^+$ *p*-Br-PhCONH-Glu-Leu-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Glu(O^{t}Bu)-Leu-Val

Bead 5: 764.3/766.3 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 877.1/879.1 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Val-Ahx-Ahx-Arg-Hsl), 975.9/977.9 ([M+H]$^+$ *p*-Br-PhCONH-Val-Leu-Val-Ahx-Ahx-Arg-Hsl)

Bead 6: 764.3/766.3 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 877.1/879.1 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Val-Ahx-Ahx-Arg-Hsl), 1006.0/1007.9 ([M+H]$^+$ *p*-Br-PhCONH-Glu-Leu-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Glu(O^{t}Bu)-Leu-Val

Bead 7: 778.2/780.2 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Ahx-Ahx-Arg-Hsl), 877.0/879.0 ([M+H]$^+$ *p*-Br-PhCONH-Val-Leu-Ahx-Ahx-Arg-Hsl), 975.9/977.9 ([M+H]$^+$ *p*-Br-PhCONH-Val-Val-Leu-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Val-Val-Leu

Bead 8: 778.2/780.3 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Ahx-Ahx-Arg-Hsl), 906.0/907.9 ([M+H]$^+$ *p*-Br-PhCONH-Lys-Leu-Ahx-Ahx-Arg-Hsl), 1033.7/1035.8 ([M+H]$^+$ *p*-Br-PhCONH-Lys-Lys-Leu-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Lys(Boc)-Lys(Boc)-Leu

Bead 9: 764.3/766.3 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 910.9/912.9 ([M+H]$^+$ *p*-Br-PhCONH-Phe-Val-Ahx-Ahx-Arg-Hsl), 1009.6/1011.6 ([M+H]$^+$ *p*-Br-PhCONH-Val-Phe-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Val-Phe-Val

Bead 10: 764.1/766.1 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 862.9/864.9 ([M+H]$^+$ *p*-Br-PhCONH-Val-Val-Ahx-Ahx-Arg-Hsl), 975.6/977.7 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Val-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Val-Val

Bead 11: 764.5/766.5 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 863.5/865.5 ([M+H]$^+$ *p*-Br-PhCONH-Val-Val-Ahx-Ahx-Arg-Hsl), 991.2/993.2 ([M+H]$^+$ *p*-Br-PhCONH-Glu-Val-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Glu(O^{t}Bu)-Val-Val

Bead 12: 794.5/796.4 ([M+H]$^+$ *p*-Br-PhCONH-Glu-Ahx-Ahx-Arg-Hsl), 893.3/895.3 ([M+H]$^+$ *p*-Br-PhCONH-Val-Glu-Ahx-Ahx-Arg-Hsl), 1006.2/1008.2 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Val-Glu-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Val-Glu

Bead 13: 764.5/766.5 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 877.1/879.2 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Val-Ahx-Ahx-Arg-Hsl), 990.0/992.0 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Leu-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Leu-Val

Bead 14: 764.4/766.4 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 863.2/865.2 ([M+H]$^+$ *p*-Br-PhCONH-Val-Val-Ahx-Ahx-Arg-Hsl), 1009.9/1011.9 ([M+H]$^+$ *p*-Br-PhCONH-Phe-Val-Val-Ahx-Ahx-Arg-Hsl)

Bead 15: 764.5/766.3 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 911.0/913.0 ([M+H]$^+$ *p*-Br-PhCONH-Phe-Val-Ahx-Ahx-Arg-Hsl), 1010.0/1011.9 ([M+H]$^+$ *p*-Br-PhCONH-Val-Phe-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Val-Phe-Val

Bead 16: 794.4/796.3 ([M+H]$^+$ *p*-Br-PhCONH-Glu-Ahx-Ahx-Arg-Hsl), 892.3/894.3 ([M+H]$^+$ *p*-Br-PhCONH-Val-Glu-Ahx-Ahx-Arg-Hsl), 991.1/993.1 ([M+H]$^+$ *p*-Br-PhCONH-Val-Val-Glu-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Val-Val-Glu(O^{t}Bu)

Fourth screen:

Mass resin **149** = 6.7mg

Buffer system 20% DMSO/borax

Guest **190** concentration = 185.5μM

Library resin swollen in 400μl buffer. 60μl of guest titrated and incubated for 3 days. Final guest concentration 24.2μM. 10 beads selected, results detailed below.

Bead 1: 778.5/780.5 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Ahx-Ahx-Arg-Hsl), 891.4/893.4 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Leu-Ahx-Ahx-Arg-Hsl), 990.2/999.2 ([M+H]$^+$ *p*-Br-PhCONH-Val-Leu-Leu-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Val-Leu-Leu

Bead 2: 764.7/766.6 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 877.4/879.4 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Val-Ahx-Ahx-Arg-Hsl), 990.2/992.2 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Leu-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Leu-Val

Bead 3: 764.4/766.4 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 892.4/894.3 ([M+H]$^+$ *p*-Br-PhCONH-Lys-Val-Ahx-Ahx-Arg-Hsl), 1005.1/1007.1 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Lys-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Lys(Boc)-Val

Bead 4: 778.3/780.3 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Ahx-Ahx-Arg-Hsl), 877.1/879.1 ([M+H]$^+$ *p*-Br-PhCONH-Val-Leu-Ahx-Ahx-Arg-Hsl), 989.9/992.0 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Val-Leu-Ahx-Ahx-Arg-Hsl)

Bead 5: 778.3/780.3 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Ahx-Ahx-Arg-Hsl), 906.0/908.0 ([M+H]$^+$ *p*-Br-PhCONH-Lys-Leu-Ahx-Ahx-Arg-Hsl), 1004.8/1006.8 ([M+H]$^+$ *p*-Br-PhCONH-Val-Lys-Leu-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Val-Lys(Boc)-Leu

Bead 6: 764.3/766.3 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 863.1/865.1 ([M+H]$^+$ *p*-Br-PhCONH-Val-Val-Ahx-Ahx-Arg-Hsl), 949.9/952.0 ([M+H]$^+$ *p*-Br-PhCONH-Ser-Val-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Ser(O^{t}Bu)-Val-Val

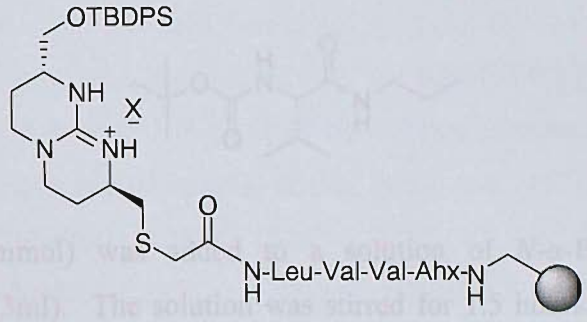
Bead 7: 794.2/796.1 ([M+H]$^+$ *p*-Br-PhCONH-Glu-Ahx-Ahx-Arg-Hsl), 892.9/894.9 ([M+H]$^+$ *p*-Br-PhCONH-Val-Glu-Ahx-Ahx-Arg-Hsl), 1039.6/1041.6 ([M+H]$^+$ *p*-Br-PhCONH-Phe-Val-Glu-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Phe-Val-Glu(O^{t}Bu)

Bead 8: no sequencing results obtained

Bead 9: 812.1/814.1 ([M+H]$^+$ *p*-Br-PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 910.9/912.8 ([M+H]$^+$ *p*-Br-PhCONH-Val-Phe-Ahx-Ahx-Arg-Hsl), 1023.6/1025.6 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Val-Phe-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Val-Phe

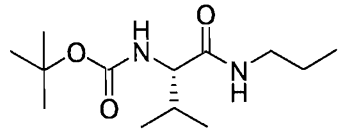
Bead 10: 764.1/766.1 ([M+H]$^+$ *p*-Br-PhCONH-Val-Ahx-Ahx-Arg-Hsl), 891.7/893.8 ([M+H]$^+$ *p*-Br-PhCONH-Lys-Val-Ahx-Ahx-Arg-Hsl), 1004.6/1006.6 ([M+H]$^+$ *p*-Br-PhCONH-Leu-Lys-Val-Ahx-Ahx-Arg-Hsl), hence tripeptide sequence Leu-Lys(Boc)-Val

Solid-phase-linked receptor (198)



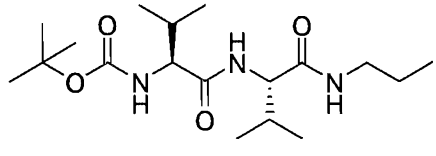
Tentagel resin **88** (119mg, 0.2mmol/g, 23.8μmol) was preswollen according to the general procedure. Standard amino acid coupling was then performed according to the general procedure using the following reagents; Fmoc-Ahx-OH (17mg, 47.6μmol), DIC (11μl, 71.4μmol), HOBt (10mg, 71.4μmol) and DIPEA (12μl, 71.4μmol) in DMF (3ml). The resin was then washed according to the general procedure and subjected to Fmoc-deprotection according to the general procedure. Reactions were monitored using the ninhydrin test. The resin was then submitted to two iterations of the following sequence; standard amino acid coupling according to the general procedure using Fmoc-Val-OH (24mg, 71.4μmol), DIC (11μl, 71.4μmol), HOBt (10mg, 71.4μmol) and DIPEA (12μl, 71.4μmol) in DMF (3ml), washing according to the general procedure, then Fmoc-deprotection according to the general procedure; this installs two Val-residues. Reactions were monitored using the ninhydrin test. Next the resin was subjected to standard amino coupling using the following reagents; Fmoc-Leu-OH (25mg, 71.4μmol), DIC (11μl, 71.4μmol), HOBt (10mg, 71.4μmol) and DIPEA (12μl, 71.4μmol) in DMF (3ml). After washing according to the general procedure, Fmoc-deprotection was carried out according to the general procedure. Reactions were monitored using the ninhydrin test. Finally, the resin was subjected to standard amino acid coupling using the following reagents; acid **145** (18mg, 35.7μmol), DIC (11μl, 71.4μmol), HOBt (10mg, 71.4μmol) and DIPEA (12μl, 71.4μmol) in DMF (3ml). After washing according to general procedure, the resin gave a negative ninhydrin test. The resin was then rinsed with Et_2O and dried *in vacuo*.

tert-Butyl _N_-(1_S_)-2-methyl-1-[(propylamino)carbonyl]propylcarbamate (200)



CDI (169mg, 1.04mmol) was added to a solution of *N*-α-Boc-Val-OH (206mg, 0.95mmol) in THF (3ml). The solution was stirred for 1.5 hours. Propylamine (78μl, 0.95mmol) was then added and the mixture stirred a further 1.5 hours. The reaction mixture was then evaporated *in vacuo* and the residue redissolved in DCM. The solution was washed with 1M $KHSO_4$(aq). The organic phase was dried over $MgSO_4$ and evaporated to give **200** as a white solid (225mg, 92%); m.p.: 128-131°C, IR ν_{max} (cm^{-1}): 3285 (m), 2963 (m), 1678 (s), 1644 (s), 1522 (s), 1247 (m), 1164 (s), 1018 (m); ^{1}H NMR: (300MHz, $CDCl_3$) δ 6.00 (br s, 1H, N*H*), 5.06 (br s, 1H, N*H*), 3.83 (dd, 1H, *J* = 9.0, 6.5Hz, α-C*H*), 3.22 (m, 2H, NHCH_2), 2.13 (m, 1H, β-CH_2), 1.52 (tq, 2H, *J* = 7.5, 7.5Hz, $CH_2CH_2CH_3$), 1.44 (s, 9H, Boc), 0.95 (d, 3H, *J* = 7.0Hz, Val CH_3), 0.92 (d, 3H, *J* = 7.0Hz, Val CH_3), 0.92 (t, 3H, *J* = 7.5Hz, $CH_2CH_2CH_3$); ^{13}C NMR: (75MHz, $CDCl_3$) δ 171.7 (C), 157.4 (C), 80.0 (C), 60.5 (CH), 41.3 (CH_2), 30.8 (CH), 28.5 (CH_3), 22.9 (CH_2), 19.5 (CH_3), 18.0 (CH_3), 11.5 (CH_3); LRMS: (ES^+) *m/z* 258.9 $[M+H]^+$ 280.9 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{13}H_{26}N_2O_3$ $[M+Na]^+$ 281.1835, found 281.1833.

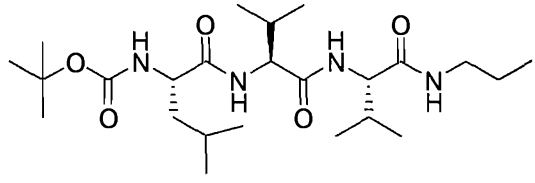
tert-Butyl _N_-(1_S_)-2-methyl-1-[((1_S_)-2-methyl-1-[(propylamino)carbonyl]propyl-amino)carbonyl]propylcarbamate (201)



Carbamate **200** (1.02g, 3.94mmol) was stirred in 20% TFA/DCM (10ml) for 2 hours. The reaction mixture was rigourously evaporated *in vacuo* using toluene as an azeotrope. The residue was redissolved in DMF (5ml) and Et_3N (1.1ml, 7.90mmol) and added to a

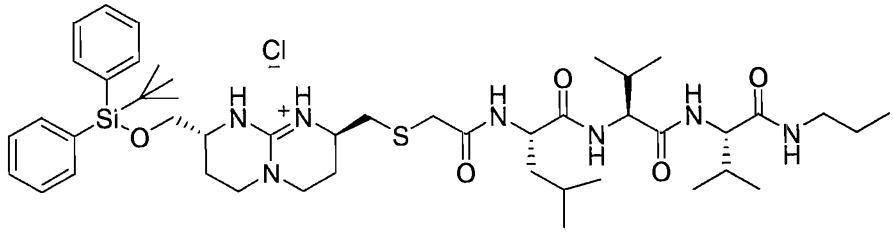
solution of CDI (768mg, 4.74mmol) and *N*-α-Boc-Val-OH (944mg, 4.34mmol) in THF (5ml) which had been stirred for 2 hours. The entire mixture was stirred for a further 18 hours. The reaction mixture was then evaporated *in vacuo* and the residue redissolved in $CHCl_3$. The organic phase was then washed with 1M $KHSO_4$(aq), dried over $MgSO_4$ and evaporated. The residue was then purified by column chromatography (eluent 20% EtOAc/petrol) to give **201** as a white solid (1.13g, 80%); m.p.: 153-155°C; IR ν_{max} (cm^{-1}): 3283 (w), 2962 (w), 1684 (m), 1637 (s), 1518 (s), 1467 (m), 1365 (s), 1245 (s), 1158 (s), 674 (m); ^{1}H NMR: (300MHz, $CDCl_3$) δ 6.69 (br s, 1H, N*H*), 6.33 (br s, 1H, N*H*), 5.06 (d, 1H, *J* = 7.5Hz, N*H*), 4.20 (dd, 1H, *J* = 8.0, 7.0Hz, α-C*H*), 3.93 (dd, 1H, *J* = 7.5, 6.0Hz, α-C*H*), 3.20 (m, 2H, NHCH_2), 2.21-2.11 (m, 2H, β-C*H* × 2), 1.52 (tq, 2H, *J* = 7.5,7.5Hz, $NHCH_2CH_2$), 1.44 (s, 9H, Boc), 1.10-0.08 (m, 15H, CH_3 × 5); ^{13}C NMR: (100MHz, $CDCl_3$) δ 174.7 (c), 171.1 (C), 156.3 (C), 80.5 (C), 60.7 (CH), 59.0 (CH), 41.5 (CH_2), 31.3 (CH), 30.5 (CH), 28.4 (CH_3), 22.8 (CH_2), 19.5 (CH_3), 19.1 (CH_3), 17.9 (CH_3), 17.7 (CH_3), 11.5 (CH_3); LRMS: (ES^+) *m/z* 358.1 $[M+H]^+$ 380.1 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{18}H_{35}N_3O_4$ $[M+Na]^+$ 380.2520, found 380.2521.

tert-Butyl _N_-(1_S_)-3-methyl-1-[((1_S_)-2-methyl-1-[((1_S_)-2-methyl-1-[(propylamino)-carbonyl]propylamino)carbonyl]propylamino)carbonyl]butylcarbamate (202)



Carbamate **201** (624.6mg, 1.75mmol) was stirred in 20% TFA/DCM (10ml) for 2 hours. The reaction mixture was rigourously evaporated *in vacuo* using toluene as an azeotrope. The residue was redissolved in THF (5ml) and Et_3N (5ml) and the whole added to a solution of CDI (340mg, 2.10mmol) and *N*-α-Boc-Leu-OH.H_2O (479mg, 1.92mmol) in THF (10ml) which had been stirred for 2 hours. The entire mixture was stirred a further 18 hours and then evaporated *in vacuo*. The residue was redissolved in EtOAc and washed sequentially with 1M $KHSO_4$(aq), saturated $NaHCO_3$ and brine. The organic phase was dried over $MgSO_4$ and evaporated to give a white solid **202** (580mg, 70%); m.p.:235-236°C; IR ν_{max} (cm^{-1}): 3277 (w), 2959 (w), 1690 (m), 1631 (s), 1523 (s), 1365 (m), 1231 (m), 1158 (s), 698 (s); ^{1}H NMR: (300MHz, $CDCl_3$) δ 7.78 (br s, 1H, N*H*), 7.48 (br s, 1H, N*H*), 7.29 (br s, 1H, N*H*), 5.61 (br s, 1H, N*H*), 4.50 (m, 1H, α-C*H*), 4.41 (m, 1H, α-C*H*), 4.34 (m, 1H, α-C*H*), 3.19 (m, 2H, NHCH_2), 2.15-2.11 (m, 2H, Val β-C*H* × 2), 2.00 (m, 1H, Leu β-C*H*H'), 1.69-1.47 (m, 2H, Leu β-CH*H'* + Leu γ-C*H*), 1.51 (tq, 2H, *J* = 7.5, 7.5Hz, $NHCH_2CH_2$), 1.42 (s, 9H, Boc), 0.95-0.84 (m, 21H, CH_3 × 7); ^{13}C NMR: (100MHz, $CDCl_3$) δ 173.2 (C), 171.6 (C), 171.4 (C), 156.0 (C), 79.7 (C), 59.2 (CH), 58.7 (CH), 53.5 (CH), 41.8 (CH_2), 41.3 (CH_2), 31.5 (CH), 30.4 (CH), 28.5 (CH_3), 25.0 (CH_3), 23.1 (CH_3), 22.8 (CH_2), 22.6 (CH), 19.4 (CH_3), 19.3 (CH_3), 18.5 (CH_3), 18.4 (CH_3), 11.5 (CH_3); LRMS: (ES^+) *m/z* 471.3 $[M+H]^+$ 493.3 $[M+Na]^+$; HRMS: (ES^+) calculated for $C_{24}H_{46}N_4O_5$ $[M+Na]^+$ 493.3360, found 493.3358.

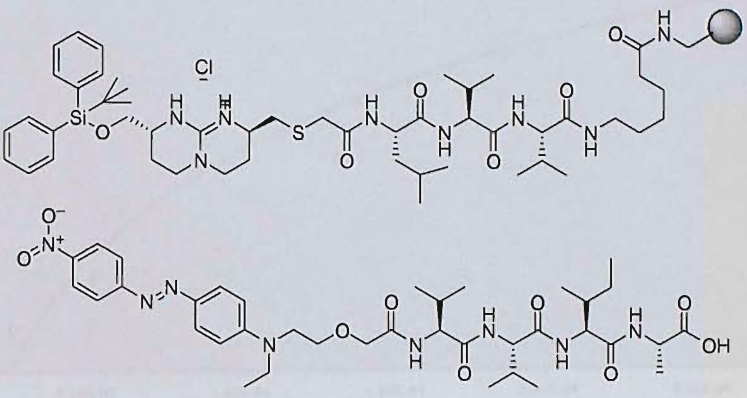
(8*R*,2*R*)-8-(*tert*-Butyl-diphenyl-silanyloxymethyl)-2-({(2*S*)-3-methyl-1-[(2*S*)-2-methyl-1-((2*S*)- 2-methyl-1-propylcarbamoyl-propylcarbamoyl)-propylcarbamoyl]-butylcarbamoyl}-methylsulfanylmethyl)-3,4,6,7,8,9-hexahydro-2H-pyrimido[1,2-a]pyrimidin-1-ium chloride (203)



Carbamte **202** (62mg, 0.13mmol) was stirred in 20% TFA/DCM (2ml) for 2 hours. The reaction mixture was rigourously evaporated *in vacuo* using toluene as an azeotrope. The residue was redissolved in DMF (3ml) and acid **145** (67mg, 0.13mmol), HOBt (20mg, 0.14mmol), DIPEA (45µl, 0.26mmol) and DMAP (2mg) were added followed by EDC (28mg, 0.14mmol). The mixture was stirred for 1 hour and then evaporated *in vacuo*. The residue was redissolved in MeOH. White solid precipitated on addition of Et_2O. This precipitate was recrystallised from MeCN to give product **203** as a white solid (10mg, 9%); m.p.: 233-235°C; IR ν_{max} (cm^{-1}): 2957 (w), 1626 (s), 1543 (m), 1459 (m), 1310 (m), 1221 (m), 1101 (s), 741 (m), 701 (s); ^1H NMR: (400MHz, d_6-DMSO) δ 8.20 (d, 1H, *J* = 8.0Hz, N*H*), 7.94 (d, 1H, *J* = 9.0Hz, N*H*), 7.92 (d, 1H, *J* = 8.0Hz, N*H*), 7.85 (t, 1H, *J* = 5.5Hz, N*H*), 7.65-7.58 (m, 6H, 5 C*H* arom + N*H*), 7.50-7.41 (m, 6H, 5 C*H* arom + N*H*), 4.37 (dt, 1H, *J* = 8.0, 8.0Hz, Leu α-C*H*), 4.16 (dd, 1H, *J* =8.5, 7.5Hz, α-C*H*), 4.07 (dd, 1H, *J* = 8.5, 7.5Hz, α-C*H*), 3.64 (m, 2H), 3.57 (m, 2H), 3.42–3.21 (m, 3H), 3.23 (s, 2H, SCH_2CO), 3.08-2.92 (m, 2H), 2.80-2.69 (m, 2H), 2.36-1.85 (m, 4H), 1.84-1.70 (m, 2H), 1.59 (m, 1H), 1.46-1.34 (m, 4H), 1.03 (s, 9H, ^{t}BuSi), 0.90-0.79 (m, 21H, CH_3 × 7); ^{13}C NMR: (100MHz, d_6-DMSO) 171.8 (C), 170.5 (C), 170.4, 168.7, 150.3, 135.1 (CH), 132.6 (C), 132.4 (C), 130.0 (CH), 128.0 (CH), 126.4 (CH), 124.1 (CH), 118.9 (CH), 109.9 (C), 65.8 (CH_2), 57.9 (CH), 57.8 (CH), 51.2 (CH), 49.4 (CH), 47.2 (CH), 44.4 (CH_2), 40.8 (CH_2), 36.3 (CH_2), 34.2 (CH_2), 30.5 (CH), 30.1 (CH), 26.6 (CH_3), 25.2 (CH_2) 24.7 (CH_2), 24.1 (CH_2), 23.0 (CH_2), 22.6 (CH), 22.2 (CH_2), 22.1 (CH_3), 21.6 (CH_3), 19.1 (CH_3), 18.8 (CH_3), 18.2 (CH_3), 18.1 (CH_3), 11.3 (CH_3); LRMS:

(ES$^+$) *m/z* 864.7 [M-Cl]$^+$; HRMS: (ES$^+$) calculated for $C_{46}H_{74}N_7O_5S_1Si_1$ [M$^+$] 864.5236, found 864.5228.

Solid-phase assays: Solid-phase-linked 'pseudo-tweezer' receptor 198 vs. dye-labelled guest 190



Adapted from a procedure by Wennemers *et al.*[167]

Calibration curves were plotted for the dye labelled tetrapeptide **190**. Results are shown in Table 5-17. Logarithmic least squares regression gave the relationship between the absorption at 500nm A_{500} and the concentration c as;

$$c = e^{\frac{A_{500} - 3.8133}{0.3603}}$$

A plot of this curve is shown in Fig 5-2.

Conc./M	A_{500}
2.65×10^{-4}	1.026
1.85×10^{-4}	0.618
1.32×10^{-4}	0.474
5.30×10^{-5}	0.221
2.65×10^{-5}	0.099

Table 5-17 Calibration data for dye-labelled guest 190 in 50%DMSO/borax

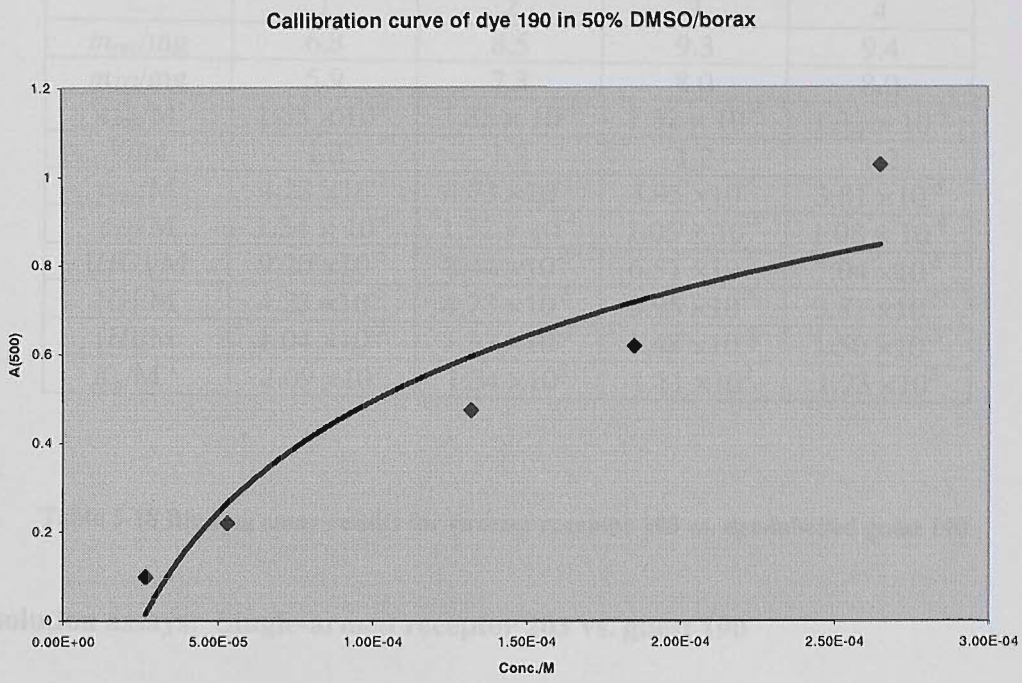


Fig 5-2 Calibration curve of guest 190 in 50% DMSO/borax

Incubation of a mass of **198**, m_{rec} and an equivalent mass of underivatised Tentagel m_{TG} with equivalent solutions of guest **190** was carried out for 18 hours. Samples were then filtered through a cotton wool plug and UV absorption measured at 500nm. Binding constants were evaluated as described in the main text. Results are shown in Table 5-18.

	1	2	3	4
m_{rec}/mg	6.8	8.5	9.3	9.4
m_{TG}/mg	5.9	7.3	8.0	8.0
c_{init}/M	1.85×10^{-4}	1.85×10^{-4}	1.32×10^{-4}	1.32×10^{-4}
V/ml	1.2	1.2	1.2	1.2
c_{rec}/M	4.23×10^{-5}	4.73×10^{-5}	3.95×10^{-5}	3.81×10^{-5}
c_{TG}/M	1.34×10^{-4}	1.32×10^{-4}	1.05×10^{-4}	1.08×10^{-4}
$[HG]$/M	9.20×10^{-5}	8.44×10^{-5}	6.51×10^{-5}	7.04×10^{-5}
$[G]$/M	4.23×10^{-5}	4.73×10^{-5}	3.95×10^{-5}	3.81×10^{-5}
$[H]$/M	1.04×10^{-3}	1.33×10^{-3}	1.48×10^{-3}	1.50×10^{-3}
K_a/M^{-1}	2.09×10^{3}	1.34×10^{3}	1.11×10^{3}	1.23×10^{3}

Table 5-18 Binding assay results for tweezer receptor 198 vs. dye-labelled guest 190

UV solution assays: Single-armed receptor 203 vs. guest 190

General procedure for UV titration:[192] A solution of dye-labelled guest **190** was prepared in 50% DMSO/borax (Solution A). A solution of single-armed receptor **203** was prepared in Solution A (Solution B). Thus solutions A and B are equal in concentration of **190**. 2ml of Solution A were placed in a quartz cell and the UV absorbance at 500nm was monitored following addition of aliquots of solution B, with thorough mixing.

Titration 1 **203** vs. **190**.

Conc. of guest **190** = 5μM, conc. of host **203** = 111μM

Titre/μl	A_{500}	Titre/μl	A_{500}
0	0.074	200	0.074
5	0.067	220	0.069
10	0.080	240	0.065
15	0.083	260	0.062
20	0.086	280	0.061
25	0.084	300	0.060
30	0.091	350	0.060
35	0.091	400	0.058
40	0.093	450	0.058
45	0.094	500	0.059
50	0.096	550	0.059
60	0.097	600	0.057
70	0.099	650	0.057
80	0.098	700	0.058
90	0.097	800	0.057
100	0.098	900	0.056
120	0.094	1000	0.057
140	0.089	1100	0.055
160	0.085	1200	0.056
180	0.079	-	-

Titration 2 **203** vs. **190**.

Conc. of guest **190** = 5μM, conc. of host **203** = 111μM

Titre/μl	A_{500}	**Titre/μl**	A_{500}
0	0.056	200	0.077
5	0.061	220	0.071
10	0.065	240	0.065
15	0.067	260	0.061
20	0.071	300	0.055
25	0.072	340	0.053
30	0.075	380	0.052
35	0.077	420	0.052
40	0.080	460	0.051
45	0.081	500	0.051
50	0.083	580	0.051
60	0.086	660	0.050
70	0.090	740	0.050
80	0.092	820	0.050
90	0.093	920	0.051
100	0.093	1020	0.051
120	0.093	1120	0.051
140	0.090	1320	0.048
160	0.086	1520	0.049
180	0.082	1720	0.049

Titration 3 **203** vs. **190**.

Conc. of guest **190** = 5μM, conc. of host **203** = 111μM

Titre/μl	A_{500}	Titre/μl	A_{500}
0	0.048	200	0.082
5	0.052	220	0.075
10	0.055	240	0.073
15	0.060	260	0.068
20	0.063	280	0.064
25	0.067	300	0.062
30	0.067	360	0.059
35	0.071	420	0.059
40	0.069	480	0.058
45	0.078	540	0.058
50	0.082	600	0.058
60	0.086	700	0.055
70	0.090	800	0.055
80	0.093	900	0.056
90	0.093	1000	0.055
100	0.097	1100	0.057
120	0.097	1200	0.055
140	0.096	1400	0.055
160	0.092	1600	0.055
180	0.088	-	-

Screening experiments with Tweezer library 184

General procedure for all screens: A mass of tweezer library resin **184** (sufficient to represent five copies of the library, ~20mg) was placed in a flat-bottomed glass dish and pre-swollen in the buffer indicated for 1 hour. A solution of dye-labelled guest **195** in the same buffer was then added to the mixture and the system incubated for the indicated time, with further aliquots of guest **195** added as indicated. Heavily stained beads were then removed from the mixture using a micropipette and each individual bead subjected to cleavage and analysis according to the general procedure.

First screen: Mass resin **184** = 32.8mg
Buffer system 20% DMSO/borax
Guest **195** concentration = 96.4μM
Library resin swollen in 2000μl buffer. 400μl of guest solution added and incubated for 24 hours. 1000μl buffer added and incubated a further 48 hours. Final guest concentration 11.3μM. 13 beads picked, cleaved and analysed according to general procedure. Results detailed below.

Bead 1: 812.1/814.1 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 911.1/913.1 ($[M+H]^+$ *p*-Br PhCONH-Val-Phe-Ahx-Ahx-Arg-Hsl), 968.0/969.9 ($[M+H]^+$ *p*-Br PhCONH-Gly-Val-Phe-Ahx-Ahx-Arg-Hsl), 1068.9/1070.8 ($[M+H]^+$ *p*-Br PhCONH-Thr-Gly-Val-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Thr(OBn)-Gly-CBS-Val-Phe
Bead 2: no sequencing results obtained
Bead 3: 812.1/814.1 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 958.9/961.0 ($[M+H]^+$ *p*-Br PhCONH-Phe-Phe-Ahx-Ahx-Arg-Hsl), 1086.8/1088.8 ($[M+H]^+$ *p*-Br PhCONH-Lys-Phe-Phe-Ahx-Ahx-Arg-Hsl), 1157.6/1159.7 ($[M+H]^+$ *p*-Br PhCONH-Ala-Lys-Phe-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Ala-Lys(Cbz)-CBS-Phe-Phe
Bead 4: 766.2/768.1 ($[M+H]^+$ *p*-Br PhCONH-Thr-Ahx-Ahx-Arg-Hsl), 893.8/895.9 ($[M+H]^+$ *p*-Br PhCONH-Lys-Thr-Ahx-Ahx-Arg-Hsl), 1040.6/1042.7 ($[M+H]^+$ *p*-Br

PhCONH-Phe-Lys-Thr-Ahx-Ahx-Arg-Hsl), 1154.4/1156.3 ($[M+H]^+$ *p*-Br PhCONH-Asn-Phe-Lys-Thr-Ahx-Ahx-Arg-Hsl), hence tweezer Asn-Phe-CBS-Lys(Cbz)-Thr(OBn)

Bead 5: 793.2/795.2 ($[M+H]^+$ *p*-Br PhCONH-Lys-Ahx-Ahx-Arg-Hsl), 939.9/941.8 ($[M+H]^+$ *p*-Br PhCONH-Phe-Lys-Ahx-Ahx-Arg-Hsl), 1040.7/1042.7 ($[M+H]^+$ *p*-Br PhCONH-Thr-Phe-Lys-Ahx-Ahx-Arg-Hsl), 1139.4/1141.5 ($[M+H]^+$ *p*-Br PhCONH-Val-Thr-Phe-Lys-Ahx-Ahx-Arg-Hsl), hence tweezer Val-Thr(OBn)-CBS-Phe-Lys(Cbz)

Bead 6: 812.6/814.6 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 883.6/885.6 ($[M+H]^+$ *p*-Br PhCONH-Ala-Phe-Ahx-Ahx-Arg-Hsl), 954.5/956.5 ($[M+H]^+$ *p*-Br PhCONH-Ala-Ala-Phe-Ahx-Ahx-Arg-Hsl), 1101.3/1103.3 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ala-Ala-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Phe-Ala-CBS-Ala-Phe

Bead 7: 811.9/813.9 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 868.9/870.8 ($[M+H]^+$ *p*-Br PhCONH-Gly-Phe-Ahx-Ahx-Arg-Hsl), 996.6/998.5 ($[M+H]^+$ *p*-Br PhCONH-Lys-Gly-Phe-Ahx-Ahx-Arg-Hsl), 1110.3/1112.3 ($[M+H]^+$ *p*-Br PhCONH-Asn-Lys-Gly-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Asn-Lys(Cbz)-CBS-Gly-Phe

Bead 8: 722.3/724.2 ($[M+H]^+$ *p*-Br PhCONH-Gly-Ahx-Ahx-Arg-Hsl), 869.0/870.9 ($[M+H]^+$ *p*-Br PhCONH-Phe-Gly-Ahx-Ahx-Arg-Hsl), 969.7/971.8 ($[M+H]^+$ *p*-Br PhCONH-Thr-Phe-Gly-Ahx-Ahx-Arg-Hsl), 1068.5/1070.6 ($[M+H]^+$ *p*-Br PhCONH-Val-Thr-Phe-Gly-Ahx-Ahx-Arg-Hsl), hence tweezer Val-Thr(OBn)-CBS-Phe-Gly

Bead 9: 764.2/766.2 ($[M+H]^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 910.9/912.8 ($[M+H]^+$ *p*-Br PhCONH-Phe-Val-Ahx-Ahx-Arg-Hsl), 1038.7/1040.6 ($[M+H]^+$ *p*-Br PhCONH-Lys-Phe-Val-Ahx-Ahx-Arg-Hsl), 1139.4/1141.5 ($[M+H]^+$ *p*-Br PhCONH-Thr-Lys-Phe-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Thr(OBn)-Lys(Cbz)-CBS-Phe-Val

Bead 10: 764.1/766.1 ($[M+H]^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 877.9/879.8 ($[M+H]^+$ *p*-Br PhCONH-Asn-Val-Ahx-Ahx-Arg-Hsl), 948.6/950.7 ($[M+H]^+$ *p*-Br PhCONH-Ala-Asn-Val-Ahx-Ahx-Arg-Hsl), 1062.4/1064.4 ($[M+H]^+$ *p*-Br PhCONH-Asn-Ala-Asn-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Asn-Ala-CBS-Asn-Val

Bead 11: 793.3/795.4 ($[M+H]^+$ *p*-Br PhCONH-Lys-Ahx-Ahx-Arg-Hsl), 894.3/896.4 ($[M+H]^+$ *p*-Br PhCONH-Thr-Lys-Ahx-Ahx-Arg-Hsl), 993.2/995.3 ($[M+H]^+$ *p*-Br PhCONH-Val-Thr-Lys-Ahx-Ahx-Arg-Hsl), 1140.0/1141.9 ($[M+H]^+$ *p*-Br PhCONH-Phe-Val-Thr-Lys-Ahx-Ahx-Arg-Hsl), hence tweezer Phe-Val-CBS-Thr(OBn)-Lys(Cbz)

Bead 12: 764.4/766.4 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 821.3/823.3 ([M+H]$^+$ *p*-Br PhCONH-Gly-Val-Ahx-Ahx-Arg-Hsl), 949.1/951.1 ([M+H]$^+$ *p*-Br PhCONH-Lys-Gly-Val-Ahx-Ahx-Arg-Hsl), 1048.0/1049.9 ([M+H]$^+$ *p*-Br PhCONH-Val-Lys-Gly-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Val-Lys(Cbz)-CBS-Gly-Val

Bead 13: 764.4/766.4 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 911.1/913.1 ([M+H]$^+$ *p*-Br PhCONH-Phe-Val-Ahx-Ahx-Arg-Hsl), 1038.9/1041.0 ([M+H]$^+$ *p*-Br PhCONH-Lys-Phe-Val-Ahx-Ahx-Arg-Hsl), 1139.7/1141.7 ([M+H]$^+$ *p*-Br PhCONH-Thr-Lys-Phe-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Thr(OBn)-Lys(Cbz)-CBS-Phe-Val

Second screen:

Mass resin **184** = 24.1mg

Buffer system 20% DMSO/borax

Guest **195** concentration = 96.4μM

Library swollen in 3000μl buffer. 40μl guest solution added and incubated for 48 hours. A further 80μl dye added and incubated a further 48 hours. 17 beads picked, cleaved and analysed according to general procedure. Results detailed below.

Bead 1: 764.3/766.4 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 865.1/867.1 ([M+H]$^+$ *p*-Br PhCONH-Thr-Val-Ahx-Ahx-Arg-Hsl), 993.0/995.0 ([M+H]$^+$ *p*-Br PhCONH-Lys-Thr-Val-Ahx-Ahx-Arg-Hsl), 1106.7/1108.8 ([M+H]$^+$ *p*-Br PhCONH-Asn-Lys-Thr-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Asn-Lys(Cbz)-CBS-Thr(OBn)-Val

Bead 2: 793.3/795.4 ([M+H]$^+$ *p*-Br PhCONH-Lys-Ahx-Ahx-Arg-Hsl), 864.3/866.2 ([M+H]$^+$ *p*-Br PhCONH-Ala-Lys-Ahx-Ahx-Arg-Hsl), 965.0/967.0 ([M+H]$^+$ *p*-Br PhCONH-Thr-Ala-Lys-Ahx-Ahx-Arg-Hsl), 1021.9/1023.9 ([M+H]$^+$ *p*-Br PhCONH-Gly-Thr-Ala-Lys-Ahx-Ahx-Arg-Hsl), hence tweezer Gly-Thr(OBn)-CBS-Ala-Lys(Cbz)

Bead 3: 722.4/724.4 ([M+H]$^+$ *p*-Br PhCONH-Gly-Ahx-Ahx-Arg-Hsl), 821.2/823.3 ([M+H]$^+$ *p*-Br PhCONH-Val-Gly-Ahx-Ahx-Arg-Hsl), 935.0/936.9 ([M+H]$^+$ *p*-Br PhCONH-Asn-Val-Gly-Ahx-Ahx-Arg-Hsl), 1063.7/1065.6 ([M+H]$^+$ *p*-Br PhCONH-Glu-Asn-Val-Gly-Ahx-Ahx-Arg-Hsl), hence tweezer Glu(OBn)-Asn-CBS-Val-Gly

Bead 4: 812.1/814.2 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 869.1/871.1 ([M+H]$^+$ *p*-Br PhCONH-Gly-Phe-Ahx-Ahx-Arg-Hsl), 1015.8/1017.8 ([M+H]$^+$ *p*-Br PhCONH-Phe-Gly-Phe-Ahx-Ahx-Arg-Hsl), 1086.7/1088.5 ([M+H]$^+$ *p*-Br PhCONH-Ala-Phe-Gly-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Ala-Phe-CBS-Gly-Phe

Bead 5: 812.1/814.3 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 910.9/912.9 ([M+H]$^+$ *p*-Br PhCONH-Val-Phe-Ahx-Ahx-Arg-Hsl), 1010.0/1012.0 ([M+H]$^+$ *p*-Br PhCONH-Val-Val-Phe-Ahx-Ahx-Arg-Hsl), 1080.9/1082.8 ([M+H]$^+$ *p*-Br PhCONH-Ala-Val-Val-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Ala-Val-CBS-Val-Phe

Bead 6: 779.1/781.2 ([M+H]$^+$ *p*-Br PhCONH-Asn-Ahx-Ahx-Arg-Hsl), 878.0/880.1 ([M+H]$^+$ *p*-Br PhCONH-Val-Asn-Ahx-Ahx-Arg-Hsl), 978.8/980.8 ([M+H]$^+$ *p*-Br PhCONH-Thr-Val-Asn-Ahx-Ahx-Arg-Hsl), 1092.5/1094.6 ([M+H]$^+$ *p*-Br PhCONH-Asn-Thr-Val-Asn-Ahx-Ahx-Arg-Hsl), hence tweezer Asn-Thr(OBn)-CBS-Val-Asn

Bead 7: 736.3/738.2 ([M+H]$^+$ *p*-Br PhCONH-Ala-Ahx-Ahx-Arg-Hsl), 807.1/809.1 ([M+H]$^+$ *p*-Br PhCONH-Ala-Ala-Ahx-Ahx-Arg-Hsl), 953.7/955.8 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ala-Ala-Ahx-Ahx-Arg-Hsl), 1010.6/1012.6 ([M+H]$^+$ *p*-Br PhCONH-Gly-Phe-Ala-Ala-Ahx-Ahx-Arg-Hsl), hence tweezer Gly-Phe-CBS-Ala-Ala

Bead 8: 722.4/724.4 ([M+H]$^+$ *p*-Br PhCONH-Gly-Ahx-Ahx-Arg-Hsl), 793.3/795.2 ([M+H]$^+$ *p*-Br PhCONH-Ala-Gly-Ahx-Ahx-Arg-Hsl), 850.2/852.1 ([M+H]$^+$ *p*-Br PhCONH-Gly-Ala-Gly-Ahx-Ahx-Arg-Hsl), 907.0/909.0 ([M+H]$^+$ *p*-Br PhCONH-Gly-Gly-Ala-Gly-Ahx-Ahx-Arg-Hsl), hence tweezer Gly-Gly-CBS-Ala-Gly

Bead 9: no sequencing results obtained

Bead 10: no sequencing results obtained

Bead 11: 736.5/738.5 ([M+H]$^+$ *p*-Br PhCONH-Ala-Ahx-Ahx-Arg-Hsl), 835.3/837.3 ([M+H]$^+$ *p*-Br PhCONH-Val-Ala-Ahx-Ahx-Arg-Hsl), 936.2/938.2 ([M+H]$^+$ *p*-Br PhCONH-Thr-Val-Ala-Ahx-Ahx-Arg-Hsl), 1035.0/1037.0 ([M+H]$^+$ *p*-Br PhCONH-Val-Thr-Val-Ala-Ahx-Ahx-Arg-Hsl), hence tweezer Val-Thr(OBn0-CBS-Val-Ala

Bead 12: 764.1/766.2 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 878.8/880.9 ([M+H]$^+$ *p*-Br PhCONH-Asn-Val-Ahx-Ahx-Arg-Hsl), 977.7/979.7 ([M+H]$^+$ *p*-Br PhCONH-Val-Asn-Val-Ahx-Ahx-Arg-Hsl), 1048.5/1050.6 ([M+H]$^+$ *p*-Br PhCONH-Ala-Val-Asn-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Ala-Val-CBS-Asn-Val

Bead 13: 764.1/766.2 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 821.0/823.0 ([M+H]$^+$ *p*-Br PhCONH-Gly-Val-Ahx-Ahx-Arg-Hsl), 877.9/880.0 ([M+H]$^+$ *p*-Br PhCONH-Gly-Gly-Val-Ahx-Ahx-Arg-Hsl), 976.7/978.7 ([M+H]$^+$ *p*-Br PhCONH-Val-Gly-Gly-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Val-Gly-CBS-Gly-Val

Bead 14: 722.3/724.2 ([M+H]$^+$ *p*-Br PhCONH-Gly-Ahx-Ahx-Arg-Hsl), 779.1/781.2 ([M+H]$^+$ *p*-Br PhCONH-Gly-Gly-Ahx-Ahx-Arg-Hsl), 880.0/881.9 ([M+H]$^+$ *p*-Br PhCONH-Thr-Gly-Gly-Ahx-Ahx-Arg-Hsl), 936.8/938.8 ([M+H]$^+$ *p*-Br PhCONH-Gly-Thr-Gly-Gly-Ahx-Ahx-Arg-Hsl), hence tweezer Gly-Thr(OBn)-CBS-Gly-Gly

Bead 15: 764.3/766.3 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 892.0/893.8 ([M+H]$^+$ *p*-Br PhCONH-Lys-Val-Ahx-Ahx-Arg-Hsl), 1020.0/1021.8 ([M+H]$^+$ *p*-Br PhCONH-Lys-Lys-Val-Ahx-Ahx-Arg-Hsl), 1147.6/1149.5 ([M+H]$^+$ *p*-Br PhCONH-Lys-Lys-Lys-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Lys(Cbz)-Lys(Cbz)-CBS-Lys(Cbz)-Val

Bead 16: no sequencing results obtained

Bead 17: 793.5/795.2 ([M+H]$^+$ *p*-Br PhCONH-Lys-Ahx-Ahx-Arg-Hsl), 892.2/894.2 ([M+H]$^+$ *p*-Br PhCONH-Val-Lys-Ahx-Ahx-Arg-Hsl), 991.1/993.1 ([M+H]$^+$ *p*-Br PhCONH-Val-Val-Lys-Ahx-Ahx-Arg-Hsl), 1119.8/1121.8 ([M+H]$^+$ *p*-Br PhCONH-Glu-Val-Val-Lys-Ahx-Ahx-Arg-Hsl), hence tweezer Glu(OBn)-Val-CBS-Val-Lys(Cbz)

Third screen:

Mass resin **184** = 31.0mg

Buffer system 20% DMSO/borax

Guest **195** concentration = 96.4μM

Library resin swollen in 3000μl buffer. 300μl guest solution added and incubated for 48 hours. Final guest concentration 8.76μM. 27 beads picked, cleaved and analysed according to general procedure. Results detailed below.

Bead 1: no sequencing results obtained

Bead 2: 764.3/766.3 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 911.0/913.0 ([M+H]$^+$ *p*-Br PhCONH-Phe-Val-Ahx-Ahx-Arg-Hsl), 981.9/983.8 ([M+H]$^+$ *p*-Br

PhCONH-Ala-Phe-Val-Ahx-Ahx-Arg-Hsl), 1038.7/1040.7 ([M+H]$^+$ *p*-Br PhCONH-Gly-Ala-Phe-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Gly-Ala-CBS-Phe-Val

Bead 3: 722.4/724.3 ([M+H]$^+$ *p*-Br PhCONH-Gly-Ahx-Ahx-Arg-Hsl), 793.2/795.2 ([M+H]$^+$ *p*-Br PhCONH-Ala-Gly-Ahx-Ahx-Arg-Hsl), 892.1/894.0 ([M+H]$^+$ *p*-Br PhCONH-Val-Ala-Gly-Ahx-Ahx-Arg-Hsl), 992.9/994.9 ([M+H]$^+$ *p*-Br PhCONH-Thr-Val-Ala-Gly-Ahx-Ahx-Arg-Hsl), hence tweezer Thr(OBn)-Val-CBS-Ala-Gly

Bead 4: 736.3/738.4 ([M+H]$^+$ *p*-Br PhCONH-Ala-Ahx-Ahx-Arg-Hsl), 850.2/852.0 ([M+H]$^+$ *p*-Br PhCONH-Asn-Ala-Ahx-Ahx-Arg-Hsl), 948.9/950.9 ([M+H]$^+$ *p*-Br PhCONH-Val-Asn-Ala-Ahx-Ahx-Arg-Hsl), 1062.8/1064.7 ([M+H]$^+$ *p*-Br PhCONH-Asn-Val-Asn-Ala-Ahx-Ahx-Arg-Hsl), hence tweezer Asn-Val-CBS-Asn-Ala

Bead 5: 811.9/813.9 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 925.6/927.6 ([M+H]$^+$ *p*-Br PhCONH-Asn-Phe-Ahx-Ahx-Arg-Hsl), 1072.2/1074.2 ([M+H]$^+$ *p*-Br PhCONH-Phe-Asn-Phe-Ahx-Ahx-Arg-Hsl), 1143.1/1145.0 ([M+H]$^+$ *p*-Br PhCONH-Ala-Phe-Asn-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Ala-Phe-CBS-Asn-Phe

Bead 6: 811.9/813.9 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 910.9/912.8 ([M+H]$^+$ *p*-Br PhCONH-Val-Phe-Ahx-Ahx-Arg-Hsl), 1009.6/1011.5 ([M+H]$^+$ *p*-Br PhCONH-Val-Val-Phe-Ahx-Ahx-Arg-Hsl), 1137.3/1139.3 ([M+H]$^+$ *p*-Br PhCONH-Lys-Val-Val-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Lys(Cbz)-Val-CBS-Val-Phe

Bead 7: no sequencing results obtained

Bead 8: 779.5/781.4 ([M+H]$^+$ *p*-Br PhCONH-Asn-Ahx-Ahx-Arg-Hsl), 850.4/852.2 ([M+H]$^+$ *p*-Br PhCONH-Ala-Asn-Ahx-Ahx-Arg-Hsl), 996.9/999.1 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ala-Asn-Ahx-Ahx-Arg-Hsl), 1125.6/1127.8 ([M+H]$^+$ *p*-Br PhCONH-Glu-Phe-Ala-Asn-Ahx-Ahx-Arg-Hsl), hence tweezer Glu(OBn)-Phe-CBS-Ala-Asn

Bead 9: 812.3/814.2 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 911.1/913.1 ([M+H]$^+$ *p*-Br PhCONH-Val-Phe-Ahx-Ahx-Arg-Hsl), 1010.0/1012.0 ([M+H]$^+$ *p*-Br PhCONH-Val-Val-Phe-Ahx-Ahx-Arg-Hsl), 1123.8/1125.8 ([M+H]$^+$ *p*-Br PhCONH-Asn-Val-Val-Phe-Ahx-Ahx-Arg-Hsl)

Bead 10: 722.5/724.5 ([M+H]$^+$ *p*-Br PhCONH-Gly-Ahx-Ahx-Arg-Hsl), 821.4/823.4 ([M+H]$^+$ *p*-Br PhCONH-Val-Gly-Ahx-Ahx-Arg-Hsl), 968.1/970.0 ([M+H]$^+$ *p*-Br PhCONH-Phe-Val-Gly-Ahx-Ahx-Arg-Hsl), 1025.0/1027.0 ([M+H]$^+$ *p*-Br PhCONH-Gly-Phe-Val-Gly-Ahx-Ahx-Arg-Hsl), hence tweezer Gly-Phe-CBS-Val-Gly

Bead 11: no sequencing results obtained

Bead 12: 779.4/781.4 ($[M+H]^+$ *p*-Br PhCONH-Asn-Ahx-Ahx-Arg-Hsl), 850.2/852.2 ($[M+H]^+$ *p*-Br PhCONH-Ala-Asn-Ahx-Ahx-Arg-Hsl), 997.0/999.0 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ala-Asn-Ahx-Ahx-Arg-Hsl), 1053.9/1055.9 ($[M+H]^+$ *p*-Br PhCONH-Gly-Phe-Ala-Asn-Ahx-Ahx-Arg-Hsl), hence tweezer Gly-Phe-CBS-Ala-Asn

Bead 13: 764.4/766.4 ($[M+H]^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 911.1/913.1 ($[M+H]^+$ *p*-Br PhCONH-Phe-Val-Ahx-Ahx-Arg-Hsl), 1012.0/1014.0 ($[M+H]^+$ *p*-Br PhCONH-Thr-Phe-Val-Ahx-Ahx-Arg-Hsl), 1068.7/1070.8 ($[M+H]^+$ *p*-Br PhCONH-Gly-Thr-Phe-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Gly-Thr(OBn)-CBS-Phe-Val

Bead 14: 722.4/724.5 ($[M+H]^+$ *p*-Br PhCONH-Gly-Ahx-Ahx-Arg-Hsl), 821.4/823.4 ($[M+H]^+$ *p*-Br PhCONH-Val-Gly-Ahx-Ahx-Arg-Hsl), 892.2/894.3 ($[M+H]^+$ *p*-Br PhCONH-Ala-Val-Gly-Ahx-Ahx-Arg-Hsl), 993.1/995.1 ($[M+H]^+$ *p*-Br PhCONH-Thr-Ala-Val-Gly-Ahx-Ahx-Arg-Hsl), hence tweezer Thr(OBn)-Ala-CBS-Val-Gly

Bead 15: 812.1/814.2 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 869.2/871.1 ($[M+H]^+$ *p*-Br PhCONH-Gly-Phe-Ahx-Ahx-Arg-Hsl), 968.1/970.0 ($[M+H]^+$ *p*-Br PhCONH-Val-Gly-Phe-Ahx-Ahx-Arg-Hsl), 1038.8/1040.7 ($[M+H]^+$ *p*-Br PhCONH-Ala-Val-Gly-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Ala-Val-CBS-Gly-Phe

Bead 16: no sequencing results obtained

Bead 17: no sequencing results obtained

Bead 18: 779.5/781.5 ($[M+H]^+$ *p*-Br PhCONH-Asn-Ahx-Ahx-Arg-Hsl), 908.4/910.3 ($[M+H]^+$ *p*-Br PhCONH-Thr-Asn-Ahx-Ahx-Arg-Hsl), 1022.0/1024.1 ($[M+H]^+$ *p*-Br PhCONH-Lys-Thr-Asn-Ahx-Ahx-Arg-Hsl), 1168.7/1170.8 ($[M+H]^+$ *p*-Br PhCONH-Phe-Lys-Thr-Asn-Ahx-Ahx-Arg-Hsl), hence tweezer Phe-Lys(Cbz)-CBS-Thr(OBn)-Asn

Bead 19: 736.2/738.2 ($[M+H]^+$ *p*-Br PhCONH-Ala-Ahx-Ahx-Arg-Hsl), 807.0/809.1 ($[M+H]^+$ *p*-Br PhCONH-Ala-Ala-Ahx-Ahx-Arg-Hsl), 877.9/879.8 ($[M+H]^+$ *p*-Br PhCONH-Ala-Ala-Ala-Ahx-Ahx-Arg-Hsl), 978.7/980.7 ($[M+H]^+$ *p*-Br PhCONH-Thr-Ala-Ala-Ala-Ahx-Ahx-Arg-Hsl), hence tweezer Thr(OBn)-Ala-CBS-Ala-Ala

Bead 20: no sequencing results obtained

Bead 21: 812.5/814.5 ($[M+H]^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 926.3/928.3 ($[M+H]^+$ *p*-Br PhCONH-Asn-Phe-Ahx-Ahx-Arg-Hsl), 1040.0/1042.0 ($[M+H]^+$ *p*-Br

PhCONH-Asn-Asn-Phe-Ahx-Ahx-Arg-Hsl), 1111.0/1113.0 ([M+H]$^+$ *p*-Br PhCONH-Ala-Asn-Asn-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Ala-Asn-CBS-Asn-Phe

Bead 22: 812.1/814.1 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 939.9/941.8 ([M+H]$^+$ *p*-Br PhCONH-Lys-Phe-Ahx-Ahx-Arg-Hsl), 1053.7/1055.7 ([M+H]$^+$ *p*-Br PhCONH-Asn-Lys-Phe-Ahx-Ahx-Arg-Hsl), 1167.4/1169.5 ([M+H]$^+$ *p*-Br PhCONH-Asn-Asn-Lys-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Asn-Asn-CBS-Lys(Cbz)-Phe

Bead 23: 736.4/738.4 ([M+H]$^+$ *p*-Br PhCONH-Ala-Ahx-Ahx-Arg-Hsl), 883.2/885.2 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ala-Ahx-Ahx-Arg-Hsl), 984.0/986.0 ([M+H]$^+$ *p*-Br PhCONH-Thr-Phe-Ala-Ahx-Ahx-Arg-Hsl), 1097.7/1099.8 ([M+H]$^+$ *p*-Br PhCONH-Asn-Thr-Phe-Ala-Ahx-Ahx-Arg-Hsl), hence tweezer Asn-Thr(OBn)-CBS-Phe-Ala

Bead 24: no sequencing results obtained

Bead 25: 764.4/766.4 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 878.2/880.2 ([M+H]$^+$ *p*-Br PhCONH-Asn-Val-Ahx-Ahx-Arg-Hsl), 1025.0/1027.0 ([M+H]$^+$ *p*-Br PhCONH-Phe-Asn-Val-Ahx-Ahx-Arg-Hsl), 1152.9/1154.8 ([M+H]$^+$ *p*-Br PhCONH-Lys-Phe-Asn-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Lys(Cbz)-Phe-CBS-Asn-Val

Bead 26: 779.2/781.2 ([M+H]$^+$ *p*-Br PhCONH-Asn-Ahx-Ahx-Arg-Hsl), 880.1/882.0 ([M+H]$^+$ *p*-Br PhCONH-Thr-Asn-Ahx-Ahx-Arg-Hsl), 1026.9/1028.8 ([M+H]$^+$ *p*-Br PhCONH-Phe-Thr-Asn-Ahx-Ahx-Arg-Hsl), 1083.7/1085.7 ([M+H]$^+$ *p*-Br PhCONH-Gly-Phe-Thr-Asn-Ahx-Ahx-Arg-Hsl), hence tweezer Gly-Phe-CBS-Thr(OBn)-Asn

Bead 27: 793.3/795.3 ([M+H]$^+$ *p*-Br PhCONH-Lys-Ahx-Ahx-Arg-Hsl), 864.1/866.2 ([M+H]$^+$ *p*-Br PhCONH-Ala-Lys-Ahx-Ahx-Arg-Hsl), 991.9/994.0 ([M+H]$^+$ *p*-Br PhCONH-Lys-Ala-Lys-Ahx-Ahx-Arg-Hsl), 1119.8/1121.6 ([M+H]$^+$ *p*-Br PhCONH-Lys-Lys-Ala-Lys-Ahx-Ahx-Arg-Hsl), hence tweezer Lys(Cbz)-Lys(Cbz)-CBS-Ala-Lys(Cbz)

Fourth screen:

Mass resin **184** = 12.9mg

Buffer system 20% DMSO/borax

Guest **195** concentration = 96.4μM

Library swollen in 1000μl buffer. 20μl guest solution added and incubated 48 hours. Final guest concentration 1.89 μM. 9 beads picked, cleaved and analysed according to general procedure. Results detailed below.

Bead 1: 764.2/766.2 ([M+H]$^+$ *p*-Br PhCONH-Val-Ahx-Ahx-Arg-Hsl), 835.1/837.1 ([M+H]$^+$ *p*-Br PhCONH-Ala-Val-Ahx-Ahx-Arg-Hsl), 948.9/950.8 ([M+H]$^+$ *p*-Br PhCONH-Asn-Ala-Val-Ahx-Ahx-Arg-Hsl), 1047.8/1049.7 ([M+H]$^+$ *p*-Br PhCONH-Val-Asn-Ala-Val-Ahx-Ahx-Arg-Hsl), hence tweezer Val-Asn-CBS-Ala-Val

Bead 2: 812.2/814.2 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 912.9/914.9 ([M+H]$^+$ *p*-Br PhCONH-Thr-Phe-Ahx-Ahx-Arg-Hsl), 984.0/986.0 ([M+H]$^+$ *p*-Br PhCONH-Ala-Thr-Phe-Ahx-Ahx-Arg-Hsl), 1040.8/1042.8 ([M+H]$^+$ *p*-Br PhCONH-Gly-Ala-Thr-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Gly-Ala-CBS-Thr(OBn)-Phe

Bead 3: 766.3/768.3 ([M+H]$^+$ *p*-Br PhCONH-Thr-Ahx-Ahx-Arg-Hsl), 867.1/869.1 ([M+H]$^+$ *p*-Br PhCONH-Thr-Thr-Ahx-Ahx-Arg-Hsl), 966.0/968.0 ([M+H]$^+$ *p*-Br PhCONH-Val-Thr-Thr-Ahx-Ahx-Arg-Hsl), 1064.7/1066.8 ([M+H]$^+$ *p*-Br PhCONH-Val-Val-Thr-Thr-Ahx-Ahx-Arg-Hsl), hence tweezer Val-Val-CBS-Thr(OBn)-Thr(OBn)

Bead 4: 779.2/781.3 ([M+H]$^+$ *p*-Br PhCONH-Asn-Ahx-Ahx-Arg-Hsl), 907.1/909.1 ([M+H]$^+$ *p*-Br PhCONH-Lys-Asn-Ahx-Ahx-Arg-Hsl), 1007.9/1009.9 ([M+H]$^+$ *p*-Br PhCONH-Val-Lys-Asn-Ahx-Ahx-Arg-Hsl), 1078.7/1080.7 ([M+H]$^+$ *p*-Br PhCONH-Ala-Val-Lys-Asn-Ahx-Ahx-Arg-Hsl), hence tweezer Ala-Val-CBS-Lys(Cbz)-Asn

Bead 5: 766.3/768.3 ([M+H]$^+$ *p*-Br PhCONH-Thr-Ahx-Ahx-Arg-Hsl), 867.1/869.2 ([M+H]$^+$ *p*-Br PhCONH-Thr-Thr-Ahx-Ahx-Arg-Hsl), 938.0/940.0 ([M+H]$^+$ *p*-Br PhCONH-Ala-Thr-Thr-Ahx-Ahx-Arg-Hsl), 1038.8/1040.8 ([M+H]$^+$ *p*-Br PhCONH-Thr-Ala-Thr-Thr-Ahx-Ahx-Arg-Hsl), hence tweezer Thr(OBn)-Ala-CBS-Thr(OBn)-Thr(OBn)

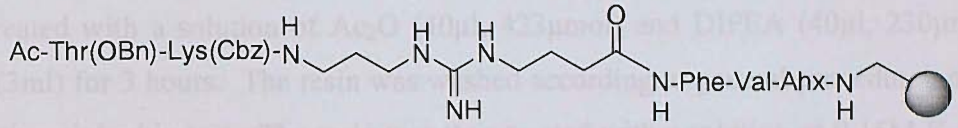
Bead 6: no sequencing results obtained

Bead 7: 812.1/814.0 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 913.0/914.9 ([M+H]$^+$ *p*-Br PhCONH-Thr-Phe-Ahx-Ahx-Arg-Hsl), 983.8/985.8 ([M+H]$^+$ *p*-Br PhCONH-Ala-Thr-Phe-Ahx-Ahx-Arg-Hsl), 1082.7/1084.6 ([M+H]$^+$ *p*-Br PhCONH-Val-Ala-Thr-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Val-Ala-CBS-Thr(OBn)-Phe

Bead 8: 766.2/768.2 ([M+H]$^+$ *p*-Br PhCONH-Thr-Ahx-Ahx-Arg-Hsl), 837.0/839.0 ([M+H]$^+$ *p*-Br PhCONH-Ala-Thr-Ahx-Ahx-Arg-Hsl), 964.9/966.9 ([M+H]$^+$ *p*-Br PhCONH-Lys-Ala-Thr-Ahx-Ahx-Arg-Hsl), 1092.7/1094.7 ([M+H]$^+$ *p*-Br PhCONH-Lys-Lys-Ala-Thr-Ahx-Ahx-Arg-Hsl), hence tweezer Lys(Cbz)-Lys(Cbz)-CBS-Ala-Thr(OBn)

Bead 9: 812.1/814.0 ([M+H]$^+$ *p*-Br PhCONH-Phe-Ahx-Ahx-Arg-Hsl), 913.0/914.9 ([M+H]$^+$ *p*-Br PhCONH-Thr-Phe-Ahx-Ahx-Arg-Hsl), 1059.6/1061.5 ([M+H]$^+$ *p*-Br PhCONH-Phe-Thr-Phe-Ahx-Ahx-Arg-Hsl), 1206.6/1208.5 ([M+H]$^+$ *p*-Br PhCONH-Phe-Phe-Thr-Phe-Ahx-Ahx-Arg-Hsl), hence tweezer Phe-Phe-CBS-Thr(OBn)-Phe

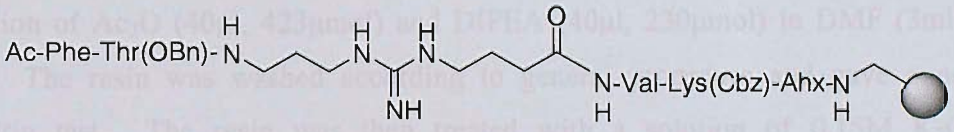
Solid-phase-linked receptor (209)



Tentagel resin **88** (165mg, 0.2mmol/g, 33μmol) was preswollen according to the general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Ahx-OH (30mg, 85.8μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Val-OH (44mg, 129μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Phe-OH (50mg, 129μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; acid **151** (34mg, 85.8μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Boc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Lys(Cbz)-OH (65mg, 129μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl,

230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Thr(OBn)-OH (55mg, 129μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. All of these reactions were monitored using the ninhydrin test. The resin was then treated with a solution of Ac_2O (40μl, 423μmol) and DIPEA (40μl, 230μmol) in DMF (3ml) for 3 hours. The resin was washed according to general procedure and gave a negative ninhydrin test. The resin was then treated with a solution of 0.15M K_2CO_3 in MeOH/DMF/H_2O (2:2:1) (~20ml per g of resin, 3 hours), and then washed with 2:2:1 MeOH/DMF/H_2O (~20ml per g of resin, three times), then DMF (~20ml per g of resin, three times) and then according to the general procedure. The resin **209** was finally rinsed with Et_2O and dried *in vacuo*.

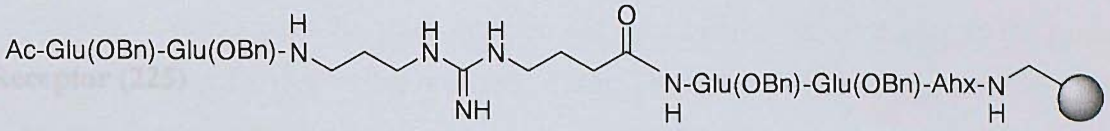
Solid-phase-linked receptor (214)



Tentagel resin **88** (201mg, 0.2mmol/g, 40μmol) was preswollen according to the general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Ahx-OH (37mg, 105μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Phe-OH (61mg, 157μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following

reagents; Fmoc-Thr(OBn)-OH (68mg, 157μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; acid **151** (42mg, 105μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Boc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Val-OH (53mg, 157μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Lys(Cbz)-OH (79mg, 157μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. All of these reactions were monitored using the ninhydrin test. The resin was then treated with a solution of Ac_2O (40μl, 423μmol) and DIPEA (40μl, 230μmol) in DMF (3ml) for 3 hours. The resin was washed according to general procedure and gave a negative ninhydrin test. The resin was then treated with a solution of 0.15M K_2CO_3 in MeOH/DMF/H_2O (2:2:1) (~20ml per g of resin, 3 hours), and then washed with 2:2:1 MeOH/DMF/H_2O (~20ml per g of resin, three times), then DMF (~20ml per g of resin, three times) and then according to the general procedure. The resin **214** was finally rinsed with Et_2O and dried *in vacuo*.

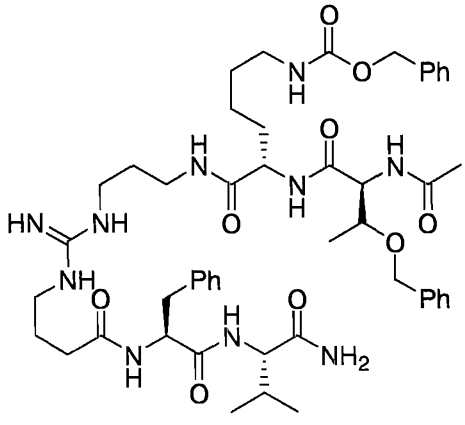
Solid-phase-linked receptor 219



Tentagel resin **88** (205mg, 0.2mmol/g, 41μmol) was preswollen according to the general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Ahx-OH (38mg, 106μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to the general procedure and subjected to Fmoc-deprotection according to the general procedure. Two glutamate residues were then incorporated by two iterations of the following reaction sequence; standard amino acid coupling according to the general procedure [using the following reagents; Fmoc-Glu(OBn)-OH (73mg, 160μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml)] washing according to the general procedure and Fmoc-deprotection according to the general procedure. Standard amino acid coupling according to the general procedure was then performed using the following reagents; acid **151** (42mg, 106μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Boc-deprotection according to general procedure. Two further glutamate residues were incorporated by twice repeating the following reaction sequence; standard amino acid coupling according to the general procedure [using Fmoc-Glu(OBn)-OH (73mg, 160μmol), DIC (40μl, 238μmol), HOBt, (20mg, 148μmol) and DIPEA (40μl, 230μmol) in DMF (3ml)], washing according to the general procedure and Fmoc-deprotection according to the general procedure. All of the above steps were monitored using the ninhydrin test. The resin was then treated with a solution of Ac_2O (40μl, 423μmol) and DIPEA (40μl, 230μmol) in DMF (3ml) for 3 hours. The resin was washed according to general procedure and gave a negative ninhydrin test. The resin was then treated with a solution of 0.15M K_2CO_3 in MeOH/DMF/H_2O (2:2:1) (~20ml per g of resin, 3 hours), and then washed with 2:2:1 MeOH/DMF/H_2O (~20ml per g of resin,

three times), then DMF (~20ml per g of resin, three times) and then according to the general procedure. The resin **219** was finally rinsed with Et_2O and dried *in vacuo*.

Receptor (225)

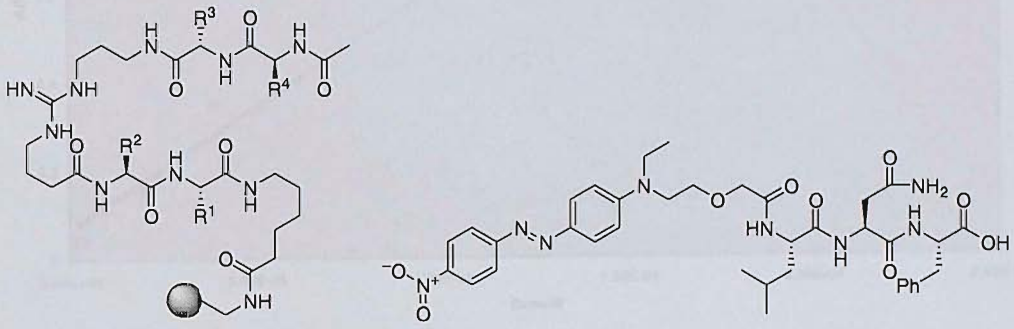


Rink amide resin (227mg, 0.69mmol/g, 156μol) was preswollen according to the general procedure, and then subjected to Fmoc-deprotection according to the general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Val-OH (159mg, 469μmol), DIC (73μl, 469μmol), HOBt, (63mg, 469μmol) and DIPEA (82μl, 469μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Phe-OH (182mg, 469μmol), DIC (73μl, 469μmol), HOBt, (63mg, 469μmol) and DIPEA (82μl, 469μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; acid **170** (168mg, 313μmol), DIC (73μl, 469μmol), HOBt, (63mg, 469μmol) and DIPEA (82μl, 469μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Ddpe-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Lys(Cbz)-OH (236mg, 469μmol), DIC (73μl, 469μmol), HOBt, (63mg,

469μmol) and DIPEA (82μl, 469μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. Standard amino acid coupling was then carried out according to the general procedure using the following reagents; Fmoc-Thr(OBn)-OH (202mg, 469μmol), DIC (73μl, 469μmol), HOBt, (63mg, 469μmol) and DIPEA (82μl, 469μmol) in DMF (3ml). The resin was then washed according to general procedure and subjected to Fmoc-deprotection according to general procedure. All of these reactions were monitored using the ninhydrin test. The resin was then treated with a solution of Ac_2O (40μl, 423μmol) and DIPEA (40μl, 230μmol) in DMF (3ml) for 3 hours. The resin was washed according to general procedure and gave a negative ninhydrin test. The resin was then treated with a solution of 0.15M K_2CO_3 in MeOH/DMF/H_2O (2:2:1) (~20ml per g of resin, 3 hours), and then washed with 2:2:1 MeOH/DMF/H_2O (~20ml per g of resin, three times), then DMF (~20ml per g of resin, three times) and then according to the general procedure. The resin was then suspended in TFA for 2 hours and then filtered. The residual resin was washed with TFA and DCM. The combined filtrates were rigourously evaporated *in vacuo* using toluene as an azeotrope. The residue was redissolved in MeCN and a white solid **225** precipitated on addition of Et_2O, which was collected by filtration (33mg, 22% based on initial resin loading); m.p.: 119-125°C; IR ν_{max} (cm^{-1}): 3272 (w), 1634 (s), 1538 (s), 1258 (m), 1201 (s), 1177 (s), 1129 (s), 833 (m), 737 (m), 697 (s); ^1H NMR: (300MHz, d_6-DMSO) δ 8.17 (d, 1H, *J* = 8.0Hz, N*H*), 7.96 (d, 1H, *J* = 8.5Hz, N*H*), 7.89-7.74 (m, 3H, N*H* × 3), 7.51-7.13 (m, 19H, arom C*H* × 15 + N*H* × 4), 7.05 (br s, 1H, N*H*), 4.99 (s, 2H, OCH_2Ph), 4.62 (m, 1H, α-C*H*), 4.53-4.38 (m, 4H, Lys CH_2 + OCH_2Ph), 4.24-4.08 (m, 2H, α-C*H* × 2), 3.93 (dd, 1H, *J* = 6.5, 4.0Hz, α-C*H*), 3.38 (dt, 2H, *J* = 7.0, 7.0Hz, CH_2), 3.14-2.87 (m, 12H, CH_2 × 6), 2.15-2.05 (m, 2H, CH_2), 1.92 (s, 3H, Ac), 1.72-1.46 (m, 6H, Phe β-CH_2 + Lys CH_2 × 2), 1.44-1.17 (m, 2H, β-C*H* × 2), 1.10 (d, 3H, *J* = 6.0Hz, CH_3), 0.85 (d, 3H, *J* = 7.0Hz, CH_3), 0.83 (d, 3H, *J* = 7.0Hz, CH_3); ^{13}C NMR: (75MHz, d_6-DMSO) δ 172.9 (C), 172.6 (C), 171.7 (C), 171.5 (C), 171.1 (C), 170.0 (C), 169.7 (C), 156.0 (C), 155.5 (C), 138.5 (C), 137.9 (C), 137.2 (C), 129.1 (CH), 128.3 (CH), 128.0 (CH), 127.9 (CH), 127.7 (CH), 127.7 (CH), 127.5 (CH), 127.3 (CH), 126.1 (CH), 74.8 (CH), 70.4 (CH_2), 65.1 (CH_2), 57.3 (CH), 56.8 (CH), 53.8 (CH), 52.8 (CH), 40.3 (CH_2), 39.2 (CH_2), 37.2 (CH_2), 35.8 (CH_2), 31.8 (CH_2), 31.6 (CH_2), 30.6 (CH_3), 29.1

(CH_2), 28.6 (CH_2), 24.6 (CH_2), 22.7 (CH_2), 22.5 (CH_2), 19.2 (CH_3), 17.8 (CH), 16.3 (CH_3), 15.1 (CH_3); LRMS: (ES^+) *m/z* 943.7 $[M+H]^+$; HRMS: (ES^+) calculated for $C_{49}H_{70}N_{10}O_9$ $[M+H]^+$ 943.5400, found 943.5412.

Solid-phase assays: Solid-phase-linked tweezer receptors 209, 214 and 219 vs. dye-labelled guest x



Adapted from a procedure by Wennemers *et al.*[167]

Calibration curves were plotted for the dye labelled tripeptide **195**. Results are shown in Table 5-19. Linear least squares regression gave the relationship between the absorption at 500nm A_{500} and the concentration c as;

$$c = \frac{A_{500} - 0.0682}{5488.2}$$

Conc/M	A_{500}
1.93×10^{-4}	1.074
9.64×10^{-5}	0.683
3.86×10^{-5}	0.349
1.93×10^{-5}	0.156
9.64×10^{-6}	0.117
3.86×10^{-6}	0.048
1.93×10^{-6}	0.041

Table 5-19 Calibration data for dye-labelled guest 195 in 20%DMSO/borax

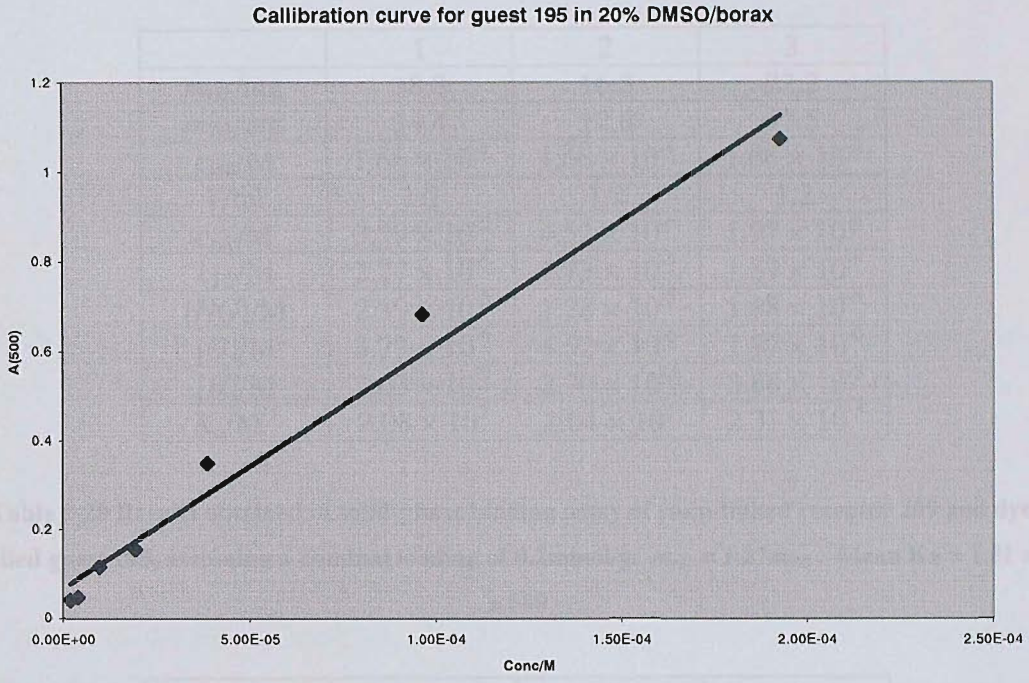


Fig 5-3 Calibration curve of guest 195 in 20% DMSO/borax

Incubation of a mass of **209**, **214**, or **219** m_{rec} and an equivalent mass of underivatised Tentagel m_{TG} with equivalent solutions of guest **195** was carried out for 18 hours. Samples were then filtered through a cotton wool plug and UV absorption measured at 500nm. Binding constants were evaluated as described in the main text. Results are shown in Tables 5-20 to 5-22.

	1	2	3
m_{rec}/mg	18.3	16.3	22.2
m_{TG}/mg	14.4	12.8	17.5
c_{init}/M	1.66×10^{-4}	1.66×10^{-4}	1.66×10^{-4}
V/ml	1.2	1.2	1.2
c_{rec}/M	3.79×10^{-6}	4.52×10^{-6}	1.97×10^{-6}
c_{TG}/M	2.77×10^{-5}	1.73×10^{-5}	1.87×10^{-5}
[*HG*]/M	2.39×10^{-5}	1.28×10^{-5}	1.68×10^{-5}
[*G*]/M	3.79×10^{-6}	4.52×10^{-6}	1.97×10^{-6}
[*H*]/M	3.03×10^{-3}	2.70×10^{-3}	3.68×10^{-3}
K_a/M^{-1}	2.08×10^{3}	1.04×10^{3}	2.31×10^{-3}

Table 5-20 Results obtained in solid-phase binding assay of resin-linked receptor 209 and dye-labelled guest 195, assuming a nominal loading of 0.2mmol/g. $m_{rec} = 1.27 m_{TG}$. Mean Ka = 1.81×10^{3} ± 680

	1	2	3
m_{rec}/mg	26.3	13.7	27.3
m_{TG}/mg	20.7	10.8	21.5
c_{init}/M	1.66×10^{-4}	1.66×10^{-4}	1.66×10^{-4}
V/ml	1.2	1.2	1.2
c_{rec}/M	1.84×10^{-6}	2.84×10^{-6}	1.17×10^{-6}
c_{TG}/M	5.01×10^{-6}	8.52×10^{-6}	5.68×10^{-6}
[*HG*]/M	3.17×10^{-6}	5.68×10^{-6}	4.51×10^{-6}
[*G*]/M	1.84×10^{-6}	2.84×10^{-6}	1.17×10^{-6}
[*H*]/M	4.83×10^{-3}	2.28×10^{-3}	4.55×10^{-3}
K_a/M^{-1}	3.94×10^{2}	8.78×10^{2}	8.49×10^{2}

Table 5-21 Results obtained in solid-phase binding assay of resin-linked receptor 214 and dye-labelled guest 195, assuming a nominal loading of 0.2mmol/g. $m_{rec} = 1.27 m_{TG}$. Mean Ka = 7.07×10^{2} ± 270

	1	2	3
m_{rec}/mg	16.8	20.1	26.6
m_{TG}/mg	12.8	15.3	20.2
c_{init}/M	1.66×10^{-4}	1.66×10^{-4}	1.66×10^{-4}
V/ml	1.2	1.2	1.2
c_{rec}/M	1.07×10^{-5}	8.35×10^{-6}	3.28×10^{-7}
c_{TG}/M	1.07×10^{-5}	6.71×10^{-6}	n/a

Table 5-22 Results obtained in solid-phase binding assay of resin-linked receptor 219 and dye-labelled guest 195. [*H*], [*HG*] and K_a cannot be meaningfully calculated

UV solution assays: Tweezer receptor 225 vs. guest 195

General procedure for UV titration:[192] A solution of dye-labelled guest **195** was prepared in 20% DMSO/borax (Solution A). A solution of tweezer receptor **225** was prepared in Solution A (Solution B). Thus solutions A and B are equal in concentration of **195**. 2ml of Solution A were placed in a quartz cell and the UV absorbance at 500nm was monitored following addition of aliquots of solution B, with thorough mixing.

Titration 1 **225** vs. **195**.

Conc. of guest **195** = 5.01μM, conc. of host **225** = 254μM

Titre/μl	A_{500}	Titre/μl	A_{500}
0	0.231	200	0.162
5	0.229	240	0.151
10	0.227	280	0.143
15	0.225	320	0.138
20	0.223	360	0.136
25	0.222	400	0.135
30	0.221	460	0.134
35	0.218	520	0.134
40	0.218	580	0.134
45	0.216	640	0.133
50	0.214	700	0.133
60	0.211	800	0.132
70	0.209	900	0.132
80	0.205	1000	0.132
90	0.201	1200	0.131
100	0.198	1400	0.130
120	0.189	1600	0.129
140	0.182	-	-
160	0.173	-	-
180	0.168	-	-

Titration 2 **225** vs. **195**.

Conc. of guest **225** = 5.01μM, conc. of host **195** = 254μM

Titre/μl	A_{500}	**Titre/μl**	A_{500}
0	0.228	200	0.171
5	0.226	240	0.158
10	0.225	280	0.149
15	0.224	320	0.141
20	0.223	360	0.136
25	0.221	400	0.135
30	0.220	460	0.134
35	0.219	520	0.135
40	0.218	580	0.134
45	0.217	640	0.134
50	0.215	700	0.133
60	0.213	800	0.133
70	0.211	900	0.132
80	0.209	1000	0.132
90	0.206	1200	0.132
100	0.205	1400	0.131
120	0.199	1600	0.131
140	0.192	-	-
160	0.187	-	-
180	0.179	-	-

Titration 3 **225** vs. **195**.

Conc. of guest **195** = 5.01μM, conc. of host **225** = 254μM

Titre/μl	A_{500}	Titre/μl	A_{500}
0	0.226	200	0.165
5	0.228	240	0.153
10	0.226	280	0.145
15	0.224	320	0.140
20	0.223	360	0.138
25	0.222	400	0.137
30	0.220	460	0.137
35	0.218	520	0.135
40	0.217	580	0.136
45	0.216	640	0.135
50	0.214	700	0.135
60	0.212	800	0.134
70	0.209	900	0.134
80	0.206	1000	0.132
90	0.203	1200	0.134
100	0.200	1400	0.131
120	0.193	1600	0.132
140	0.186	-	-
160	0.178	-	-
180	0.171	-	-

References

1. Lehn, J.M., *Angew. Chem. Int. Ed. Engl.* **1988**, *27*, 89
2. Pedersen, C.J., *J. Am. Chem. Soc.* **1967**, *89*, 2495
3. Pedersen, C.J., *J. Am. Chem. Soc.* **1967**, *89*, 7017
4. Fischer, E., *Ber. Deutsch. Chem. Ges.* **1894**, *27*, 2985
5. a) Bennett N.T.; Schultz, G.S.; *Am. J. Surg.* **1993**, *165*, 728-737; b) Waterfield, M.D., *Lancet* **1989**, *3*, 1243-1246
6. Cheng, B.; Fu, X.B.; Sun, T.Z.; Sun, X.Q.; Shuz, Z.Y., *Chinese Med. J.* **2002**, *115*, 1525-1528
7. Damjanov, I.; Milner, B.; Knowles, B.B., *Laboratory Investigation* **1986**, *55*, 588-592
8. Kim, J.S.; McKinnis, V.S.; Nawrocki, A.; White, S.R.; *Am. J. Resp. Cell. Mol.* **1998**, *18*, 66-74
9. Badjii, J.D.; Balzani, V.; Credi, A.; Silvii, S.; Stoddart, J.F., *Science* **2004**, *303*, 1845-1849
10. Tseng, H.R.; Vignon, S.A.; Stoddart, J.F., *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 1491-1495
11. Cram, D.J.; Katz, H.E.; Dicker, I.B.; *J. Am. Chem. Soc.* **1984**, *106*, 4987-5000
12. Cram, D.J.; Dicker, I.B.; Lauer, M.; Knobler, C.B.; Trueblood K.N., *J. Am. Chem. Soc.* **1984**, *106*, 7150-7167
13. Cram, D.J.; Lam, P.Y.S., *Tetrahedron Symposium-in-Print* **1986**, *42*, 1607-1615
14. Newcomb, M.; Torer, J.L.; Helgeson, R.C.; Cram, D.J., *J. Am. Chem. Soc.* **1979**, *101*, 4941-4947
15. Lein, G.M.; Cram, D.J., *J. Chem. Soc. Chem. Commun.* **1982**, *5*, 301-304
16. Piode, L.A.R.; Simon, W.; *Chemica* **1979**, *24*, 72
17. Lima, J.L.F.C.; Machado, A.A.S.C., *Analyst* **1986**, *111*, 799
18. Allinger, N.L.; Cava, M.P.; De Jongh, D.C.; Johnson, C.R.; Lebbel, N.A.S., "Organic Chemistry", Worth Publishers, New York, 1971
19. Jeffry, G.A., "An Introduction to Hydrogen Bonding", Oxford University Press, Oxford, 1997

20. Jorgenson, W.L.; Pranata, J., *J. Am. Chem. Soc.* **1990**, *112*, 2008
21. Chang, S.K.; Van Engen, D.; Fan, E.; Hamilton, A.D., *J. Am. Chem. Soc.* **1991**, *113*, 7640
22. Schneider, H.J.; Yatsimirsky, A., "Principles and Methods in Supramolecular Chemistry", John Willis and Sons Ltd, Chichester, 2000
23. Szetjli, J.J., *Chem. Rev.* **1998**, *98*, 1743
24. Smithrud, D.B.; Wyman, T.B.; Diederich, F., *J. Am. Chem. Soc.* **1991**, *113*, 5420
25. "Excitatory Amino Acid Receptors" ed. Krogsgaard-Larsen, P.; Hansen, J.J., Ellis Harwood, Chichester, 1992
26. Fitzmaurice, R.J.; Kyne, G.M.; Douheret, D.; Kilburn, J.D., *J. Chem. Soc. Perkin Trans. 1* **2002**, 841
27. Hartley, J.H.; James, T.D.; Ward, C.J.; *J. Chem. Soc. Perkin Trans. 1* **2000**, 3155
28. Peczuh, M.W.; Hamilton A.D., *Chem. Rev.* **2000**, *100*, 2479
29. Ilioudis, C.A.; Steed, J.W., *J. Supramol. Chem.* **2001**, *1*, 165
30. Miller, G.; Riede, J.; Schmitchen, F.P.; *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 1516
31. Berlinck, R.G.S., *Nat. Prod. Rep.* **1999**, *16*, 339
32. Linton, B.; Hamilton, A.D., *Tetrahedron* **1999**, *55*, 6027
33. Schmuck, C., *Chem Commun.* **1999**, 843
34. Schmuck, C., *Chem. Eur. J.* **2000**, *6*, 709
35. Schmuck, C.; Machon, U., *Chem. Eur. J.* **2005**, *11*, 1109-1118
36. Metzger, A.; Anslyn, E.V., *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 649
37. Metzger, A.; Lynch, V.M.; Anslyn, E.V., *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 862
38. Snowden, T.S.; Anslyn, E.V., *Curr. Opp. Chem. Biol.* **1999**, *3*, 740
39. Galan, A.; Andrea, D.; Echavarren, A.M.; Prados, P; de Mendoza, J., *J. Am. Chem. Soc.* **1992**, *114*, 1511-1512
40. Rossi, S.; Kyne, G.M.; Turner, D.L.; Wells, N.J.; Kilburn, J.D., *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 4233
41. Schneider, H-J., *Adv. Supramol. Chem.* **2000**, *6*, 185-216
42. Wennemers, H.; Yoon, S.S.; Still, W.C., *J. Org. Chem.* **1995**, *60*, 1108

43. Torneiro, M.; Still, W.C., *J. Am. Chem. Soc.* **1995**, *117*, 5887
44. Peczuh, M; Hamilton, A.D.; Sanchez-Quesada, J.; de Mendoza, J.; Haack, T.; Giralt, E., *J. Am. Chem. Soc.* **1997**, *119*, 9327
45. Sanchez-Quesada, J.; Seel, C.; Prados, P.; de Mendoza, J.; *J. Am. Chem. Soc.* **1996**, *118*, 277
46. Rensing, S.; Schrader, T.; *Org. Lett.* **2002**, *4(13)*, 2161-2164
47. Albelda, S.M.; Buck, C.A., *FASEB J* **1990**, *4*, 2868
48. Chen, C.W.; Whitlock, H.W. Jr, *J. Am. Chem. Soc.* **1978**, *100*, 4921
49. Smith, C.K.; Regan, L., *Acc. Chem. Res.* **1997**, *30*, 153
50. Fokkens, M.; Schrader, T.; Klärner, F-G., *J. Am. Chem. Soc.* **2005**, *127(41)*, 14415-14421
51. Torneiro, M.; Still, W.C., *Tetrahedron* **1997**, *53*, 8739
52. Schmuck, C.; Geiger, L., *J. Am. Chem. Soc.* **2004**, *126(29)*, 8898-8899
53. Nowick, J.S.; Insaf, S., *J. Am. Chem. Soc.* **1997**, *119*, 10903-10908
54. Chitnumsub, P.; Fiori, W.R.; Lashmel, H.A.; Diaz, H.; Kelly, J.W., *Bioorg. Med. Chem.* **1999**, *7*, 39-59
55. Merrifield, R.B., *J. Am. Chem. Soc.* **1963**, *85*, 2149
56. Fenniri, H., "Combinatorial Chemistry", Oxford University Press, Oxford, 2000
57. Terrett, N.K., "Combinatorial Chemistry", Oxford University Press, Oxford, 1998
58. Estrada, N.M.; Garrido, F.M.D.; Medeiros, M.E.; Machado, S.D., *Quim. Nova* **2005**, *28(4)*, 670-677
59. Houghten, R.A.; Pinilla, C.; Blondella, S.E.; Appel, J.R.; Dooley, C.T.; Cuervo, J.H., *Nature* **1991**, *354*, 84
60. Dooley, C.T.; Chung, N.N.; Schiller, P.V.; Houghten, R.A., *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10811
61. Erb, S.; Janda, K.D.; Bremer, S., *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 11422-11426
62. Geysen, H.M.; Meloen, R.H.; Bartelling, S.J., *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 3998
63. Lam, K.S.; Lebl, M.; Krchnak, V., *Chem. Rev.* **1997**, *97*, 411-448

64. Furka, A.; Sebestyen, F.; Aszedom, M.; Dibo, G. *in* "Highlights of Modern Biochemistry," Proceedings of the 14th International Congress of Biochemistry, USP, Utrecht, Netherlands 1988, Vol. 5, p47
65. Furka, A.; Sebestyen, F.; Aszedom, M.; Dibo, G. *in* "More peptides by less labour", poster presented at Xth International Synposium on Medicinal Chemistry, Budapest 1988
66. Wang, X.; Peng, X.; Liu, R.; Gill, S.S.; Lam, K.S.; *J. Comb. Chem.* **2005**, *7*, 197-209
67. Nilsson, J.W.; Thorstensson, F.; Kvarnström, I.; Oprea, T.; Samuelsson, B.; Nilsson, I., *J. Comb. Chem.* **2001**, *3*, 546-553
68. Czarnik, A.W., *Curr. Opp. Chem. Biol.* **1997**, *1*, 60-66
69. Houghten, R.A., *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 5131-5135
70. Czarnik, A.W., Abstract at The Assosciation for Laboratory Automation *Labautomation '97* Conference, 1997 January 18-22, San Diego (Available online at http://labautomation.org)
71. Edman, P.; Begg, G., *Eur. J. Biochem.* **1967**, *1*, 80
72. Needels, M.C.; Jones, D.G.; Tate, E.H.; Heinkel, G.L.; Kochersperger, I.M.; Dower, W.J.; Barrett, R.W.; Gallop, M.A., *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 10700
73. Nielsen, J.; Janda, K.D., *Methods(San Diego)* **1994**, *6*, 361
74. Nikolaiev, V.; Stierandova, A.; Krchnak, V.; Seligman, B.; Lam, K.S.; Salmon, S.E.; Lebl, M., *Pept. Res.* **1993**, *6*, 161
75. Kerr, J.M.; Banvilla, S.C.; Zuckermann, R.N., *J. Am. Chem. Soc.* **1993**, *115*, 2529
76. Nestler, H.P.; Bartlett, P.A.; Still, W.C., *J. Org. Chem.* **1994**, *59*, 4723-4724
77. Ni, Z-J.; Maclean, D.; Holmes, C.P.; Murphy, M.M.; Ruhland, B.; Jacobs, J.W.; Gordon, E.M.; Gallop, M.A., *J. Med. Chem.* **1996**, *39*, 1601-1608
78. Swali, V.; Langley, G.J.; Bradley, M., *Curr. Opp. Chem. Biol.* **1999**, *3*, 337-341
79. Geysen, H.M.; Wagner, C.D.; Bodnar, W.M.; Markworth, C.J.; Parke, G.J., Schoenen, F.J.; Wagner, D.S.; Kinder, D.S., *Chem. Biol.* **1996**, *3(8)*, 679
80. Zambias, R.A.; Boulton, D.A.; Griffin, P.R., *Tetrahedron Lett.* **1994**, *35*, 4283-4286

81. Brummel, C.L.; Lee, I.N.W.; Zhou, Y.; Benkovic, S.J.; Winograd, N., *Science* **1994**, *264*, 399-402
82. Cheng, X.; Hochlawski, J.; *Anal. Chem.* **2002**, *74*, 2679-2690
83. Franz, A.H.; Liu, R.; Song, A.; Lam, K.S.; Lebrilla, C.B., *J. Comb. Chem.* **2003**, *5*, 125-137
84. Youngquist, R.S.; Fuentes, G.R.; Lacey, M.P.; Keough, T., *J. Am. Chem. Soc.* **1995**, *117*, 3900-3906
85. St. Hilaire, P.M.; Lowary, T.M.; Meldal, M.; Bock, K., *J. Am. Chem. Soc.* **1998**, *120*, 13312-13320
86. Davies, M.; Bradley, M., *Tetrahedron Lett.* **1997**, *38(49)*, 8565-8568
87. Sweeney, M.C.; Pei, D., *J. Comb. Chem.* **2003**, *5*, 218-222
88. Wang, P.; Arabaci, G.; Pei, D., *J. Comb. Chem.* **2001**, *3*, 251-254
89. Hoffman, C.; Blechschmidt, D.; Krüger, R.; Karas, M.; Griesinger, C., *J. Comb. Chem.* **2002**, *4*, 79-86
90. Nitz, M.; Franz, K.J.; Maglachlin, R.L., Imperiali, B., *ChemBioChem* **2003**, *4*, 272-276
91. Wang, X.; Zhang, J.; Song, A.; Lebrilla, C.B.; Lam, K.S., *J. Am. Chem. Soc.* **2004**, *126*, 5740-5749
92. Tornøe, C.W.; Sanderson, S.J.; Mottram, J.C.; Coombs, G.H.; Meldal, M., *J. Comb. Chem.* **2004**, *6*, 312-324
93. Pastor, J.J.; Lingard, I.; Bhalay, G.; Bradley, M., *J. Comb. Chem.* **2003**, *5*, 85-90
94. Miguel, Y.R.; Sanders, J.K.M.; *Curr. Opp. Chem. Biol.* **1998**, *2*, 417-421
95. Lovigne, J.J., Anslyn, E.V., *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 3118-3130
96. Linton, B.; Hamilton, A.D., *Curr. Opp. Chem. Biol.* **1999**, *3*, 307-312
97. Srinivasan, N.; Kilburn, J.D., *Curr. Opp. Chem. Biol.* **2004**, *8*, 305-310
98. Borchardt, A.; Still, W.C., *J. Am. Chem. Soc.* **1994**, *116*, 373
99. Burger, M.T.; Still, W.C., *J. Org. Chem.* **1995**, *60*, 7382
100. Boyce, R.; Li, G.; Nestler, H.P.; Suenaga, T.; Still, W.C., *J. Am. Chem. Soc.* **1994**, *116*, 7955
101. Cheng, Y.; Suenaga, T.; Still, W.C., *J. Am. Chem. Soc.* **1996**, *118*, 1813

102. Braxmeier, T.; Demarcus, M.; Fessman, T.; McAteer, S.; Kilburn, J.D., *Chem. Eur. J.* **2001**, *7*, 1889-1898

103. Arienzo, R.; Kilburn, J.D., *Tetrahedron* **2002**, *58*, 711-719

104. Jensen, K.B.; Braxmeier, T.B.; Demarcus, M.; Kilburn, J.D., *Chem. Eur. J.* **2002**, *8*, 1300-1309

105. Wennemers, H.; Conza, M., *J. Org. Chem.* **2002**, *67*, 2696-2698

106. Wennemers, H.; Nold, M.C.; Conza, M.M.; Kulicke, K.J.; Neuburger, M. *Chem. Eur. J.* **2003**, *9*, 442-448

107. Chamorro, C.; Liskamp, R.M.J., *J. Comb. Chem.* **2003**, *5*, 794-801

108. Chamorro, C.; Hofmann, J.W.; Liskamp, R.M.J., *Tetrahedron* **2004**, *60*, 8691-8697

109. Chamorro, C.; Liskamp, R.M.J., *Tetrahedron* **2004**, *60*, 11145-11157

110. Schmuck, C., *Chemie in Unserer Zeit* **2001**, *35(6)*, 356-366

111. Neu, H.C., *Science* **1992**, *257*, 1065-1072

112. Süssmuth, R.D., *ChemBioChem* **2002**, *3*, 295, 298

113. Nicolaou, K.C.; Boddy, C.N.C.; Bräse, S.; Winssinger, N., *Angew. Chem. Int. Ed. Engl.* **1999**, *38*, 2096-2152

114. Axelsen, P.H.; Li, D.; *Bioorg. Med. Chem.* **1998**, *6*, 877-881

115. Xu, R.; Greiveldinger, G.; Marenus, L.E.; Cooper, A.; Ellman, J.A., *J. Am. Chem. Soc.* **1999**, *121*, 4898-4899

116. Pant, N.; Hamilton, A.D., *J. Am. Chem. Soc.* **1988**, *110*, 2002-2003

117. Mann, M.J.; Pant, N.; Hamilton, A.D., *J. Chem. Soc. Chem. Commun.* **1986**, 158

118. Liu, R.; Sanderson, P.E.J.; Still, W.C., *J. Org. Chem.* **1990**, *55*, 5184-5186

119. Dowden, J.; Edwards, P.D.; Flack, S.S.; Kilburn, J.D., *Chem. Eur. J.* **1999**, *5*, 79-89

120. Waymark, C.P.; Kilburn, J.D.; Gillies, I., *Tetrahedron Lett.* **1992**, *33*, 3939

121. Henley, P.D; Kilburn, J.D., *Chem. Commun.* **1999**, 1335-1336

122. Casnati, A.; Sansone, F.; Ungaro, R., *Acc. Chem. Res.* **2003**, *36*, 246-254

123. Botana, E.; Ongari, S.; Arienzo, R.; Demarcus, M.; Frey, J.G.; Piarulli, U.; Potenza, D.; Gennari, C.; Kilburn, J.D., *Chem. Commun.* **2001**, 1358-1359

124. Schmuck, C.; Heil, M.; Scheiber, J.; Baumann, K., *Angew. Chem. Int. Ed. Engl.* preprint, DOI: 10.1002/anie.200501812
125. Chippingdale, A.B.; Wade, J.D.; Barrow, C.J., *J. Pept. Sci.* **2001**, *7*, 227-249
126. Zerovnik, E., *Eur. J. Biochem.* **2002**, *269*, 3362-3371
127. Hardy, J.; Selkoe, D.J., *Science* **2002**, *297*, 353
128. Yankner, B.A., *Neuron* **1996**, *16*, 921-932
129. Lansbury Jr, P.T., *Acc. Chem. Res.* **1996**, *29*, 317-321
130. Jarrett, J.T.; Berger, E.P.; Lansbury Jr, P.T., *Biochemistry* **1993**, *32(18)*, 4693
131. Lansbury Jr, P.T., *Curr. Opp. Chem. Biol.* **1997**, *1*, 260-267
132. Ono, K.; Hasezawa, K.; Naiki, H.; Yamada, M., *J. Neurosci. Res.* **2004**, *75*, 742-750
133. Lashmel, H.A.; Hartley, D.M.; Balakhareh, D.; Aggarwal, A.; Teichberg, S.; Callaway, D.J.E., *J. Biol. Chem.* **2002**, *277*, 42881-42890
134. Soto, C.; Kindy, M.S.; Baumann, M.; Frangione, B.; *Biochem. Biophys. Res. Commun.* **1996**, *226*, 672-680
135. Schmuck, C.; Heil M.; *ChemBioChem* **2003**, *4*, 1232-1238
136. Schmuck, C.; Heil, M.; *Org. Biomol. Chem.* **2003**, *1*, 633-636
137. Schmuck, C.; Frey, P.; Heil, M., *ChemBioChem* **2005**, *6*, 628-631
138. "AIDS epidemic update", UNAIDS/WHO, 2001
139. Barre-Sinoussi, F.; Cherman, J.C.; Rey, P.; Nugyeyne, M.T.; Chamaret, S.; Gruest, J.; Dauget, C.; Axler-Blin, C.; Vezinet-Brun, F.; Rouzioux, C.; Rozenbaum, W.; Montagnier, L.; *Science* **1983**, *220*, 868-871
140. de Clercq, E., *J. Med. Chem.* **1995**, *38*, 2491-2517
141. Wainberg, M.A.; Friedland, G., *JAMA* **1998**, *279*, 1977
142. Maschera, B.; Darby, G.; Palù, G.; Wright, L.L.; Tisdale, M; Myers, R.; Blair, E.D., *J. Biol. Chem.* **1996**, *271*, 33231-33235
143. Wlodawer, A.; Miller, M; Jaskolski, M.; Sathyanarayana, B.; Baldwin, E.; Weber, I.; Selk, L.; Clawson, L.; Schneider, J.; Kent, S., *Science* **1989**; *245*, 616-621
144. Todd, M.J.; Semo, N.; Freire, E., *J. Mol Biol.* **1998**, *283*, 475-488
145. King, R.W.; Winslow, D.L.; Garber, S.; Scarnati, H.T.; Bachelor, L.; Stack, S.; Otoo, M., *J. Antiviral Res.* **1995**, *28*, 13-24

146. Darke, P.L.; Jordan, S.P.; Hall, D.L.; Zugag, J.A.; Shafer, J.A.; Ku, L.C., *Biochemistry* **1994**, *33(1)*, 98-105

147. Kuzmic, P., *Biochem. Biophys. Res. Commun.* **1993**, *265*, 10492

148. Gustchina, A.; Weber, I.T., *Proteins* **1991**, *10*, 325-329

149. Babe, L.M.; Pichuantes, S.; Craik, C.S., *Biochemistry* **1990**, *30*, 106-111

150. Cheng, Y-S.E.; Yin, F.H.; Foundly, S.; Blomstrom, D.; Kettner, C.A., *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9660-9664

151. Schramm, H.J.; de Rosny, E.; Reboud-Ravaux, M.; Büttner, J.; Dick, A.; Schramm, W., *Biol. Chem.* **1999**, *380*, 593-596

152. Schramm H.J.; Boetzel, J.; Büttner, J.; Fritsche, E.; Göhring, W.; Jaegar, E.; König, S.; Thumfart, O.; Wenger, T.; Nagel, N.E.; Schramm, W., *Antiviral Res.* **1996**, *30*, 155-170

153. Bogetto, N.; Reboud-Ravaux, M., *Biol. Chem.* **2002**, *383*, 1321-1324

154. Bowman, M.J.; Chmielewski, J., *Biopolymers* **2002**, *66*, 126-133

155. Zutshi, R.; Franciskovich, J.; Shultz, M.; Schweitzer, B.; Bishop, P.; Wilson, M.; Chmielewski, J., *J. Am. Chem. Soc.* **1997**, *119*, 4841-4845

156. Shultz, M.D.; Bowman, M.J.; Ham, Y-W.; Zhao, X.; Tora, G.; Chmielewski, J., *Angew. Chem. Int. Ed. Engl.* **2000**, *39(15)*, 2710

157. Shultz, M.D.; Chmielewski, J., *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2431-2436

158. Bowman, M.J.; Chmielewski, J., *Bioorg. Med. Chem. Lett.* **2004**, *14*, 1395-1398

159. Shultz, M.D.; Ham, Y-W.; Lee, S-G.; Davis, D.A.; Brown, C.; Chmielewski, J., *J. Am. Chem. Soc.* **2004**, *126(32)*, 9886-9887

160. Hwang, Y-S.; Chmielewski, J., *J. Med. Chem.* **2005**, *48*, 2239-2242

161. Bouras, A.; Bogetto, N.; Beratalah, Z.; de Rosny, E.; Sicsic, S.; Reboud-Ravaux, M., *J. Med. Chem.* **1999**, *42*, 957-962

162. Merabet, N.; Dumond, J.; Collinet, B.; Van Baelingham, L.; Bogetto, N.; Ongeri, S.; Ressad, F.; Reboud-Ravaux, M.; Sicsic, S., *J. Med. Chem.* **2004**, *47*, 6392-6400

163. Song, M-C.; Rajesh, S.; Hayashi, Y.; Kiso, Y., *Bioorg. Med. Chem. Lett.* **2001**, *11*, 2465-2468

164. LaBrenz, S.R.; Kelly, J.W., *J. Am. Chem. Soc.* **1995**, *117*, 1655

165. Breccia, P.; Bogetto, N.; Perz-Fernandez, R.; Van Gool, M.; Takahashi, M.; René, L.; Prados, P.; Badet, B.; Reboud-Ravaux, M.; de Mendoza, J., *J. Med. Chem.* **2003**, *46*, 5196-5207

166. Shepherd, J.; Gale, T.; Jensen, K.B.; Kilburn, J.D., *Chem. Eur. J.* preprint, DOI: 10.1002/chem..2005.00876

167. Wennemers, H.; Conza, M., *Chem. Commun.* **2003**, *7*, 866-867

168. Bartoli, S.; Jensen, K.B.; Kilburn, J.D., *J. Org. Chem.* **2003**, *68*, 9416-9422

169. Larsen, C.; Steliou, K.; Harpp, D.N., *J. Org. Chem.* **1978**, *43(2)*, 337-339

170. Madder, A.; Farcy, N.; Hosteen, N.G.C.; de Muynck, H.; de Clercq, P.J.; Barry, J.; Davis, A.P., *Eur. J. Org. Chem.* **1999**, 2787-2791

171. Carpino, L.A.; Shroff, H.; Triolo, S.A.; Mansour, E-S.M.E.; Wenschuh, H.; Albericio, F., *Tetrahedron Lett.* **1993**, *34(49)*, 7829-7832

172. Atherton, E.; Sheppard, R.C.; Wade, J.D., *J. Chem. Soc. Chem. Commun.* **1983**, 1060-1062

173. Bedford, J. *et al*, *Int. J. Peptide Protein Res.* **1992**, *40*, 300

174. a) Münster, I.; Rolle, U.; Madder, A.; de Clercq, P.J.; *Tetrahedron: Assymetry* **1995**, *6(11)*, 2673-2674, b) Peschke, W.; Schiessl, P.; Schmidtchen, F.P.; Bissinger, P.; Schier, A., *J. Org. Chem.* **1995**, *60*, 1039-1043, c) Kurzmeier, H.; Schmidtchen, F.P., *J. Org. Chem.* **1990**, *55*, 3749-3755

175. Castonguay, R.; L'Herbert, C.; Keillor, J.W.; *Bioorg. Med. Chem.* **2002**, *10*, 4185-4191

176. Bretonniere, Y.; Cann, M.J.; Parker, D.; Slater, R.; *Org. Biomol. Chem.* **2004**, *2(11)*, 1624-1632

177. Bartoli, S., PhD thesis, University of Southampton, 2005

178. Sarin, V.K.; Kent, S.B.H.; Tam, J.P.; Merrifield, R.B., *Anal. Biochem.* **1981**, *117*, 147

179. Chan, W.C.; White, P.D., "Fmoc solid-phase peptide synthesis: a practical approach", Oxford University Press, Oxford, 2000

180. Vojkovsky, T., *Pept. Res.* **1995**, *8*, 236

181. Lesiv, A.V.; Ioffe, S.L.; Strelenko, Y.A.; Tartakovsky, V.A., *Helv. Chim. Acta* **2002**, *85(10)*, 3489-3507

182. Karki, L.; Lakshmi, K.V.; Szalai, V.A.; Brudvig, G.W., *J. Am. Chem. Soc.* **2000**, *122*, 5180

183. Pittelkov, M.; Lewinsky, R.; Christensen, J.B., *Synthesis* **2002**, *15*, 2195-2202

184. Katritzsky, A.R.; Ledoux, S.; Witek, R.M., Nair, S.K., *J. Org. Chem.* **2004**, *69*, 2976-2982

185. Ariyoshi, Y., *Bull. Chem. Soc. Jap.* **1984**, *57(11)*, 3197-3202

186. Jahreis, G.; Fittkau, S., *J. Prakt. Chem.* **1984**, *326(1)*, 35-40

187. Larsen, C. ; Steliou, K. ; Harpp, D.N., *J. Org. Chem.* **1978**, *43(2)*, 337-339

188. Barton, D.H.R.; Herve, Y.; Potier, P.; Thierry, J., *Tetrahedron* **1988**, *44(17)*, 5479-5486

189. Montero, A.; Goya, P.; Jagerovic, N.; Callado, L.F. ; Meana, J.J. ; Giron, R. ; Goicoechea, C.; Martin, I., *Bioorg. Med. Chem.* **2002**, *10(4)*, 1009-1018

190. Wang, Q.X.; Phanstiel, O. IV, *J. Org. Chem.* **1998**, *63*, 1491

191. Chhabra, S.R. ; Hothi, B. ; Evans, D.J. ; White, P.D. ; Bycroft, B.W., Chan, W.C., *Tetrahedron Lett.* **1998**, *39*, 1603

192. Connors, K.A., "Binding Constants", Wiley, New York, 1987