DOI: 10.1002/ejoc.200601004

Rapid Sequencing of Split-and-Mix Peptide Receptor Libraries – Identification of Binding Partners for Val-Val-Ile-Ala in Aqueous Solution

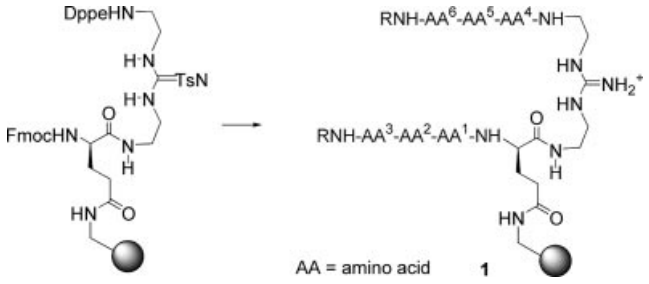
Jon Shepherd,[a] G. John Langley,[a] Julie M. Herniman,[a] and Jeremy D. Kilburn*[a]

Keywords: Peptides / Coding / Sequencing / Combinatorial chemistry / Receptors / Screening

Split-and-mix peptide libraries have been encoded by capping cleavable, laddered sequences with *p*-bromobenzoic acid, allowing rapid analysis of the cleaved ladder fragments by mass spectrometry. The encoding methodology has been applied to the synthesis of libraries of peptide receptors and a library of tripeptides functionalised with a bicyclic guanidinium unit has been successfully screened to identify binding partners for Val-Val-Ile-Ala, the C-terminal tetrapeptide sequence of the amyloid-β protein, Aβ(1–42).

Introduction

The combinatorial, split-and-mix approach[1] has been used by several research groups to prepare combinatorial libraries which have been successfully screened to identify novel receptors for small molecules[2] and in particular sequence-selective receptors for peptides, in both organic[3] and aqueous[4,5] solvent systems. In this context we have used bis(aminoalkyl)guanidiniums as a scaffold to prepare resin-bound libraries of two-armed "tweezer" receptor structures **1** (Scheme 1).[4a,4i] Screening experiments with these libraries involve incubation of a sample of the library with a dye-labelled guest peptide and selection of "hit" beads, which are strongly coloured as a result of selective binding of the dye-labelled guest peptide to the receptor on the particular bead. The "hit" beads are manually selected and the structure of the receptor on each "hit" bead is identified. This approach has been successfully used to identify tweezer receptors of general structure **1** for peptides such as the bacterial cell wall precursor tripeptide Lys-D-Ala-D-Ala in aqueous media.[4a]



Scheme 1.

These results, and related studies by others,[4] emphasise the potential of such tweezer structures as strong and selective peptide receptors even in competitive aqueous solvents, despite the inherent flexibility of such structures.

An added advantage of this split-and-mix approach has been that once a library is synthesised, a series of screening experiments can be carried out very rapidly which in prinicple allows one to probe the binding selectivities of the receptors as a function of both the structure of the peptide guest and the solvent used in the screening experiment. However, screening experiments with libraries of several thousand receptor structures often lead to many "hit" (strongly dyed) beads. A limitation of this approach, in our hands at least, has been that the identification of receptor structures on these "hit" beads has relied on Edman sequencing,[6] either of the peptide arms of the receptor itself, or of a separately synthesized peptide coding strand. The Edman sequencing methodology has proved to be both expensive and time-consuming which has greatly restricted the number of beads from a screening experiment that can be conveniently sequenced, and this in turn has restricted the amount of statistical information about the preferred structure of potent receptors that can be obtained from a given screening experiment. Although reasonably strong and selective receptors have none-the-less been identified, sequencing of greater numbers of beads should lead to more reliable identification of the most potent receptor structure from a given library. These considerations are even more relevant if larger libraries incorporating greater diversity than used to date are to be screened.

With this in mind an alternative sequencing strategy has been developed which allows rapid, reliable and inexpensive sequencing of large numbers of beads.[7] Herein we describe the results of these studies, and the application of the new sequencing methodology to the screening of a single-armed

[a] School of Chemistry, University of Southampton, Southampton, SO17 1BJ, UK
Fax: +44-2380-596805
E-mail: jdk1@soton.ac.uk

receptor library to identify binding partners for Val-Val-Ile-Ala, the C-terminal tetrapeptide sequence of the amyloid-β protein, Aβ(1–42).

Results and Discussion

Development of Capping Methodology

A very useful method on the basis of mass spectrometry for the determination of the sequence of bead-based peptides has been described by Youngquist and by Lebl.[8] During synthesis of a library, a small percentage of the growing peptide chains are capped at each step by an acylating group to generate a family of termination products on the bead – a "ladder sequence". Cleavage of an individual bead and analysis by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), gives the mass of each of the termination products and allows the full peptide sequence to be read from the N-terminus as the mass difference between adjacent peaks equates to the mass of each consecutive amino acid component. In their original patent[8c] Lebl et al. also suggested using bromobenzoic acid as a capping group to generate mass peaks with a characteristic bromine isotope pattern. The concept of introducing a substituent with a readily identifiable isoptope pattern was successfully developed by Bradley[9] who used an equimolar mixture of AcOH and $[D_3]$AcOH as the capping reagent, so that the relevant peaks of the ladder sequence are readily identified by the presence of two peaks of equal intensity at $[M + H]^+$ (for Ac-capped fragments) and $[M + 3 + H]^+$ (for the corresponding $[D_3]$Ac-capped fragment). Identification of relevant peaks was greatly facilitated by application of data reduction software which can select peaks with defined isotopic patterns, and as a consequence greatly increases the effective signal-to-noise ratio of the relevant signals in the mass spectra.

As a coding strategy for our receptor libraries, we decided to generate similar ladder sequences using a single molecular identity, *p*-bromobenzoic acid – as suggested by Lebl – as the capping reagent, as a modification to the well described methodology developed by Bradley.[9,10]

The utility of *p*-bromobenzoic acid was investigated using tentagel resin (0.2 mmol/g) and standard Fmoc SPPS methods for peptide synthesis. Thus, following the procedure described by Bradley,[9] methionine, which is selectively cleaved under conditions of excess CNBr/TFA/H_2O[11] generating the C-terminal homoserine lactone (Hsl), was used as the linker and several aminohexanoic acid, or phenylalanine residues were introduced to provide sufficient mass such that cleavage products of interest were out of range of low-mass noise and matrix ions inherent in MALDI-TOF MS analysis. Incorporation of an arginine unit was also investigated as the readily ionisable guanidine group increases the sensitivity of the mass spectrometric analysis. In Youngquist's original development of ladder sequences[8] the side-chain guanidine of the arginine used for this purpose was unprotected, but here a protected guanidine, Arg(Pbf) or Arg(Mtr), which can be deprotected before cleavage and MS analysis, was used to avoid competitive binding of a guest peptide carboxylate in screening experiments.

In preliminary experiments to ascertain the limit of detection of the capped fragments, tentagel resin was derivatised, by sequential coupling of methionine, arginine(Pbf) and three aminohexanoic acid (Ahx) or phenylalanine (Phe) moieties. Samples of the resulting resin were coupled with single amino acids [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)] and then capped with varying amounts (from 0–100 mol-%) of *p*-bromobenzoic acid relative to the resin loading and excess DIC/HOBt/DIPEA for 18 h. Single beads were selected from each sample, cleaved using an excess of CNBr, and the cleavage products were analysed by MALDI-TOF MS. The resultant data were processed using data reduction software.[12] It was thus determined that, with the inclusion of the arginine in the linker construct, the capped sequences (as the $[M + H]^+$ adducts) could be detected using as little as 10 mol-% of the *p*-bromobenzoic acid in each capping reaction. The amount of capping in each step is clearly important as each capping reaction clearly consumes some of the peptide of interest, but using only 10 mol-% would in principle allow one to construct a suitably encoded nonapeptide, and certainly should satisfactorily allow encoding of shorter peptides. Of course preparing encoded peptides in this way, even using 10 mol-% capping, means that only a portion of the material on a single bead is represented by the final structure and accordingly any screening experiments should take account of contributions from the truncated peptide sequences. However, in the studies described herein only the full peptide structure incorporates the crucial guanidinium functionality that serves as a carboxylate binding site. As this study and our earlier studies have shown, the presence of the guanidinium is essential in achieving binding of peptides with a free carboxylate terminus. Hence, using the laddered peptide encoding strategy is particularly suitable for the sorts of libraries we wish to prepare.

During these preliminary investigations it was discovered that when partial capping was attempted by simultaneous addition of *p*-bromobenzoic acid and coupling reagents (DIC/HOBt/DIPEA) to the resin bound peptide some beads within a sample of resin were not being capped at all whereas other beads were being completely capped, presumably due to inefficient mixing of the capping group throughout the resin sample in conjunction with a rapid and efficient coupling reaction. To avoid this problem a solution of *p*-bromobenzoic acid in DMF was equilibrated with the resin sample (for one hour) before addition of coupling reagents (DIC/HOBt/DIPEA) thus allowing an even distribution of the capping group throughout the resin sample.

To demonstrate the effectiveness of the protocol a tripeptide (Ile-Ala-Gly) was prepared on the solid phase from the resin **2**, using Fmoc SPPS, with the inclusion of capping reactions after each cycle of coupling–deprotection, to generate the resin **3** (Scheme 2).

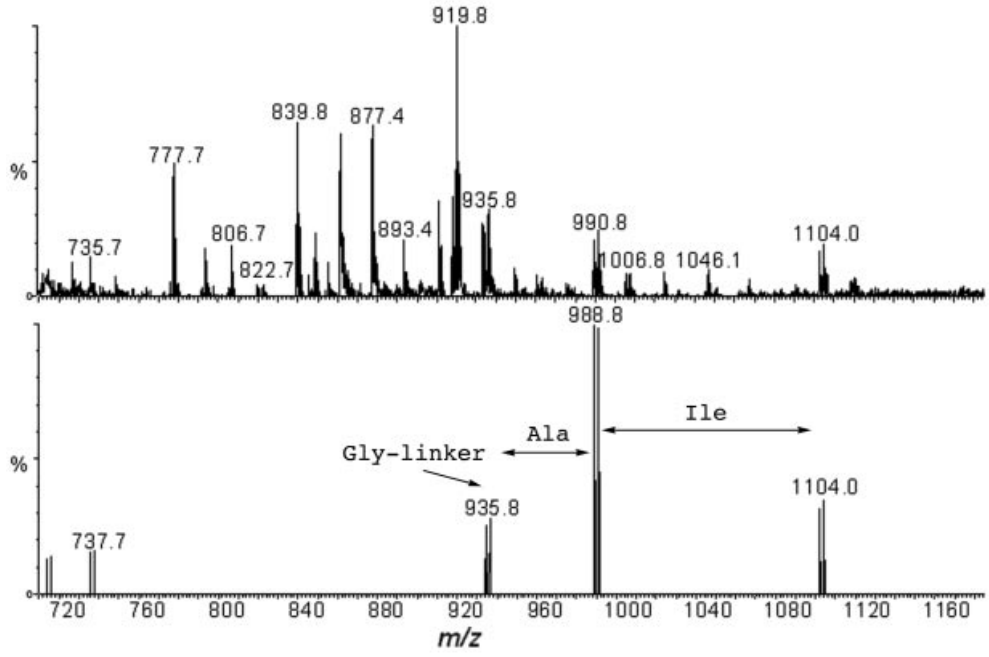
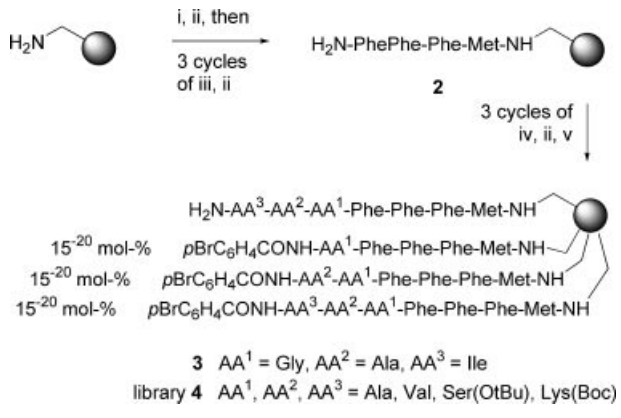


Figure 1. Mass spectra of a sample from a single cleaved bead of **3** a) raw data, b) processed to identify Br-isotope patterns.

MALDI-TOF MS of the single bead cleavage followed by processing with the appropriate software[12] gave easily interpretable data (Figure 1) and the peptide sequence could be read off from the spectrum by calculating the mass differences of the amino acid residues.

In a further demonstration of the methodology, a small library **4** of 64 tripeptides was synthesised on resin **2** by split-and-mix synthesis, using the four amino acids Ala, Val, Ser(O*t*Bu) and Lys(Boc) (Scheme 2). Twenty single beads were selected from the library, cleaved and analysed by MALDI-TOF MS. The sequence of eighteen of these beads was readily assigned. The remaining two beads showed no relevant characteristic signals in the mass spectrum, suggesting either unsuccessful coupling or cleavage reactions.



Scheme 2. i) Fmoc-Met, DIC, HOBt, DIPEA, DMF, 3 h; ii) 20% piperidine/DMF, 20 min; iii) Fmoc-Phe, DIC, HOBt, DIPEA, DMF, 3 h; iv) Fmoc-AA, DIC, HOBt, DIPEA, DMF, 3 h; v) *p*-BrC_6H_4COOH (15–20 mol-%), DIC, HOBt, DIPEA, DMF, 18 h.

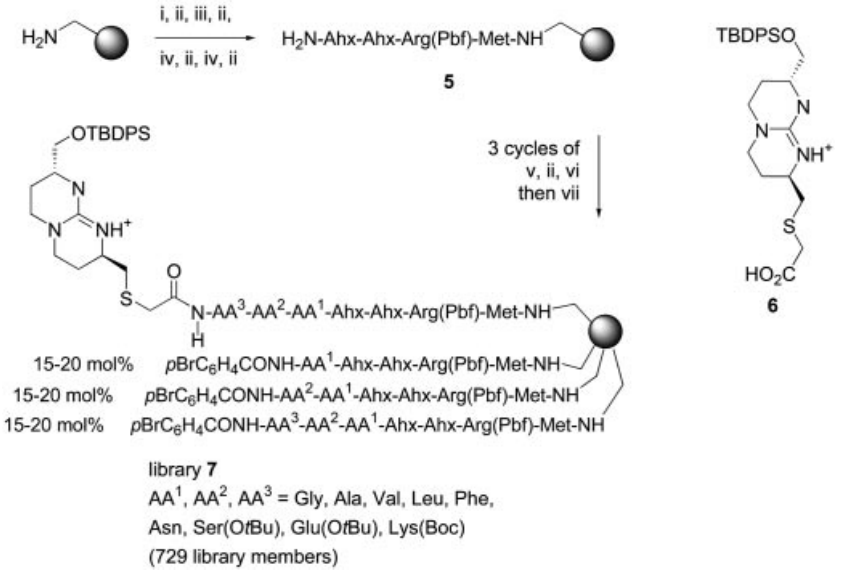
Synthesis of Receptor Libraries

With a functioning methodology for generating "ladder sequences" (and hence "readable" receptor libraries) achieved, synthesis of two libraries of receptors for carboxylates was undertaken: a single-armed receptor library **7** and a tweezer receptor library **17**.

The basic design of the single-armed receptor consists of a bicyclic guanidinium **6** (which has been developed extensively by de Mendoza[13] and by Schmitdchen[14] as a potent carboxylate binding site), linked to a randomised peptide sequence, and attached to Tentagel resin through the methionine linker, and with an arginine residue for greater sensitivity in the mass spectrometry analysis.

The resin **5** was thus prepared by sequential coupling of methionine, arginine(Pbf) and three aminohexanoic acid moieties on to tentagel resin (Scheme 3). This was followed by three rounds of split-and-mix synthesis using nine amino acids [Gly, Ala, Val, Leu, Phe, Asn, Ser(O*t*Bu), Glu(O*t*Bu), Lys(Boc)], and incorporating a capping reaction with *p*-bromobenzoic acid (15–20 mol-% of resin loading) to generate a 729-membered tripeptide library with a capped "ladder" sequence. The synthesis was completed by attaching the bicyclic guanidinium **6**[15] (DIC/HOBt/DIPEA in DMF) to generate single-armed receptor library **7**. To assess 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 using MALDI-TOF MS.

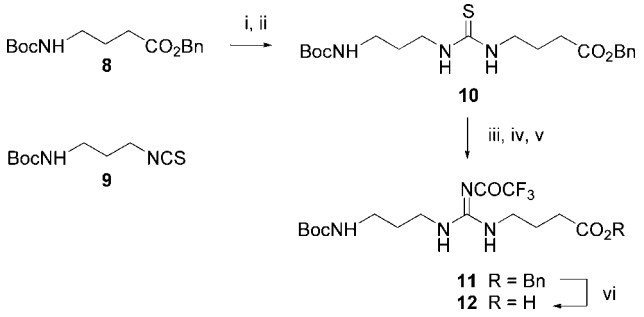
In order to generate tweezer receptors on the solid-phase and make use of the new coding methodology we prepared the functionalised guanidine **12** which, with its amino and carboxylic acid functionality, could be incorporated into a



Scheme 3. i) Fmoc-Met, DIC, HOBt, DIPEA, DMF, 3 h; ii) 20% piperidine/DMF, 20 min; iii) Fmoc-Arg(Pbf), DIC, HOBt, DIPEA, DMF, 3 h; iv) Fmoc-Ahx, DIC, HOBt, DIPEA, DMF, 3 h; v) Fmoc-AA, DIC, HOBt, DIPEA, DMF, 3 h; vi) *p*-BrC_6H_4COOH (15–20 mol-%), DIC, HOBt, DIPEA, DMF, 18 h; vii) **6**, DIC, HOBt, DIPEA, DMF, 3 h.

growing peptide chain using standard solid phase peptide synthesis (SPPS). This approach differs from our previous syntheses of tweezer receptors on the solid-phase, which utilised an orthogonally protected bis(amino)guanidine derivative as the starting scaffold for attachment of two peptidic arms[4a,4i] (see Scheme 1).

The guanidine **12** was prepared by coupling of the isothiocyanate **9** and an amino ester, prepared by Boc deprotection of the ester **8**, to give the thiourea **10** (Scheme 4). Alkylation of **10** gave the thiouronium salt which was converted into the trifluoroacetyl-protected guanidine **11** by reaction with trifluoroacetamide and DBU.[16] Hydrogenation of the benzyl ester gave the acid **12**. The trifluoroacetyl protecting group can be easily removed on the solid phase using mild aqueous base.[16]



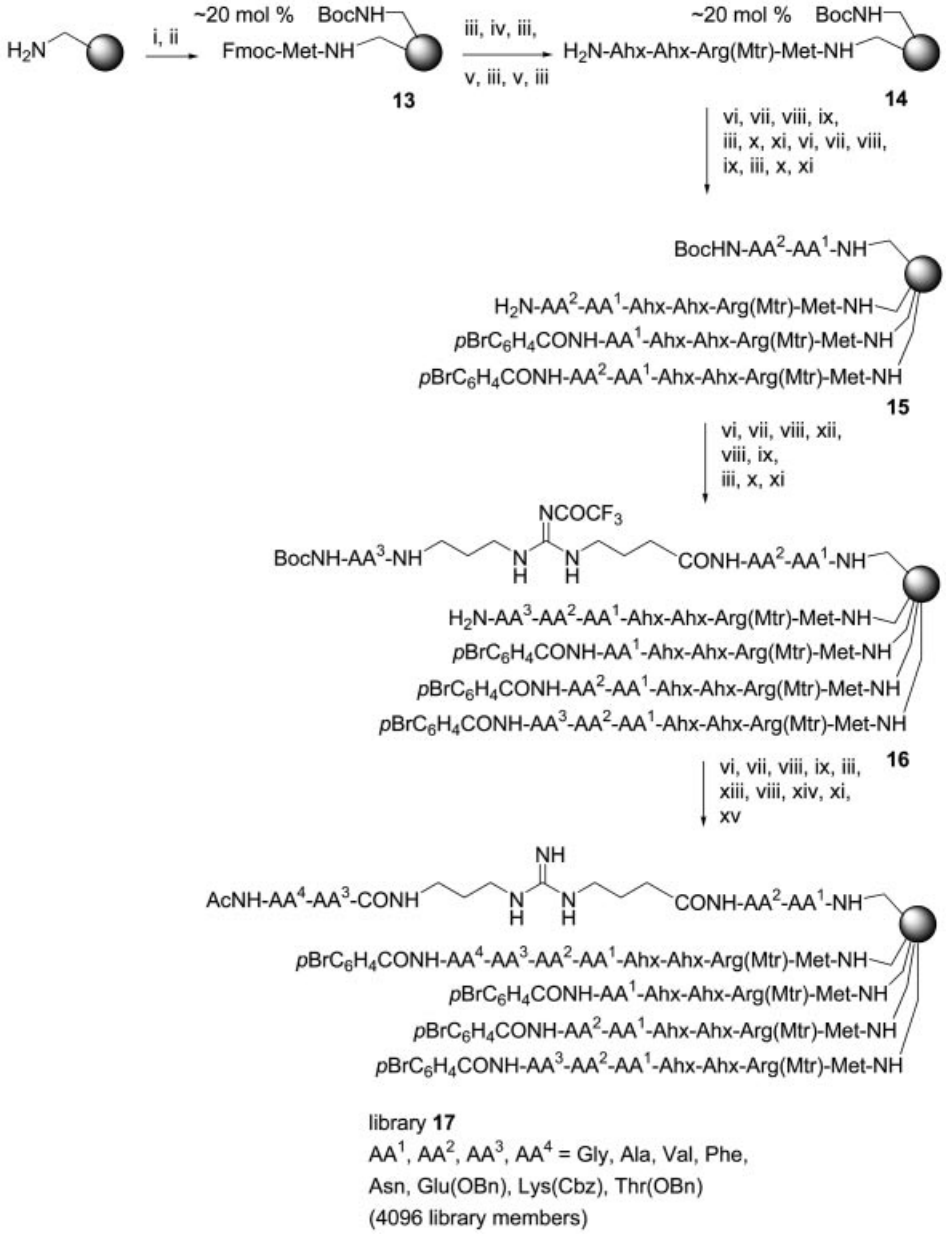
Scheme 4. i) CF_3CO_2H, DCM; ii) **9**, Et_3N, DCM/MeOH, 18 h; iii) MeI, acetone, 1 h; iv) NH_4PF_6, DCM/MeOH, 18 h; v) F_3CCONH_2, DBU, toluene/$CHCl_3$, 2 h; vi) H_2, Pd/C, EtOAc, 3 h.

Furthermore, in order to avoid generating truncated sequences on each resin bead that contained the guanidine moiety **12**, when implementing the capping strategy, it was necessary to generate distinct "strands" (receptor-strand and coding-strands) on each bead, using orthogonal protecting groups.

Solid-phase library synthesis began with partial derivatisation of Tentagel resin (0.47 mmol/g) with Fmoc-Met (ca. 80 mol-%). A quantitative ninhydrin test[17] indicated a residual-free NH_2 loading of 0.10 mmol/g (ca. 20%). The remaining free sites were treated with excess Boc_2O to generate the bifunctional resin **13** (Scheme 5). Deprotection of the methionine residue was followed by coupling of Arg(Mtr), then two aminohexanoic acid moieties giving **14**. The more acid-stable Arg(Mtr) was selected because of its compatability[18] with the several Boc-deprotections that followed.

Resin **14** was then subjected to two cycles of split-and-mix synthesis. Thus the resin was split into eight portions; each portion was independently coupled using an Fmoc-amino acid [Gly, Ala, Val, Phe, Asn, Glu(OBn), Lys(Cbz), Thr(OBn)]; Boc-deprotection using dilute TFA; coupling with the same amino acid as before, but with a Boc rather than Fmoc protecting group; Fmoc-deprotection; capping reaction (15% of nominal loading of the Fmoc-deprotected strand); mixing of the portions to give the resin **15**. 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.

Resin **15** was again split into eight portions and coupled with an Fmoc-amino acid (eight amino acids as before). Boc-deprotection was carried out using dilute TFA and the acid **12** was coupled to each portion. 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. Finally Fmoc-deprotection was carried out, followed by capping reaction (20% of original loading of Fmoc-strand), and mixing of the resin portions to give **16**. In this manner the guanidine is incorporated into only one strand, with a ladder sequence generated separately.



Scheme 5. i) Fmoc-Met (0.9 equiv.), DMF, DIC, HOBt, DIPEA, 3 h; ii) xs. Boc_2O, DIPEA, DMF, 3 h; iii) 20% piperidine/DMF, 20 min; iv) Fmoc-Arg(Mtr), DIC, HOBt, DIPEA, 3 h; v) Fmoc-Ahx, DIC, HOBt, DIPEA, 3 h; vi) split; vii) Fmoc-AA, DIC, HOBt, DIPEA, DMF, 3 h; viii) 20% TFA/DCM, 20 min; ix) Boc-AA, DIC, HOBt, DIPEA, DMF, 3 h; x) *p*-BrC_6H_4COOH (15 mol-%), DIC, HOBt, DIPEA, DMF, 18 h; xi) mix; xii) **12**, DIC, HOBt, DIPEA, DMF, 3 h; xiii) *p*-BrC_6H_4COOH (xs), DIC, HOBt, DIPEA, DMF, 18 h; xiv) Ac_2O, DIPEA, DMF, 3 h; xv) K_2CO_3, MeOH, DMF, H_2O.

A final iteration of split-and-mix synthesis was performed similar to the first two iterations: Fmoc-coupling; Boc-deprotection; Boc-coupling; Fmoc-deprotection as before, followed by a complete termination of the Fmoc-strand using excess *p*-bromobenzoic acid. A final Boc-deprotection was carried out, the Boc-strand capped using an excess of acetic anhydride and all resin portions were then mixed together. Final deprotection of the trifluoroacetyl group was carried out using 0.15 M K_2CO_3 in MeOH/DMF/H_2O[16] to give the 4096-membered library **17**.

To test the success of the library synthesis, 5 beads were selected at random, individually treated to deprotect the arginine side chain (1 M TMSBr in TFA), cleaved (CNBr, TFA, H_2O) and analysed by MALDI-TOF MS. Successful sequencing of all five beads indicated successful implementation of the synthesis and a suitable diversity in the members of the library.

Screening Experiments

In order to demonstrate the utility of the new coding strategy for the rapid sequencing of hit beads from screening experiments, library **7** has been screened with a peptide

guest using the same methodology we have described previously.[4a,4i] For the peptide guest we used the dye-labelled tetrapeptide Val-Val-Ile-Ala **18**, synthesised using standard peptide couplings (Figure 2). Val-Val-Ile-Ala is the C-terminal tetrapeptide sequence of the amyloid-β protein, Aβ(1–42). Amyloid-β proteins, consisting of 39–43 amino acids constitute the major proteinaceous component of protein plaques found in the brains of Alzheimer's patients.[19] Inhibition of the self-assembly of amyloid-β proteins, and consequent formation of β-amyloid fibrils, provides a potential therapeutic approach for treating Alzheimer's disease[20] and a number of small molecules and designed β-sheet breakers have indeed been reported to inhibit Aβ aggregation.[21]

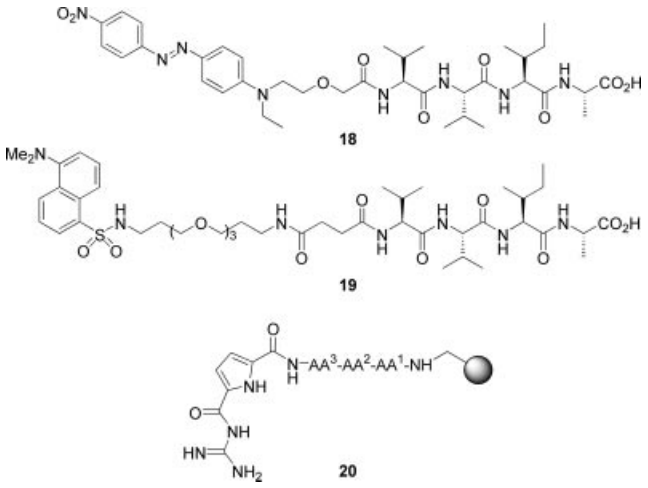


Figure 2. Dye-labelled peptide **18** used in screening experiments with library **7**, and dye-lablled peptide **19** used by Schmuck in screening experiment with library **20**.

In particular, Schmuck has described[4g,4h] screening experiments using the dansyl-labelled tetrapeptide Val-Val-Ile-Ala **19** and a tripeptide receptor library **20** capped with a guanidinopyrrole.[22] Schmuck successfully identified guanidinopyrrole-linked tripeptides, which bound strongly to the sequence Val-Val-Ile-Ala and subsequently showed that such guanidinopyrrole-linked tripeptides could inhibit the formation of β-amyloid fibrils.[21a] Given the similarities between Schmuck's library **20** and our own bicyclic guanidinium-linked tripeptide library **7**, it was of interest to screen our library with dye-labelled tetrapeptide Val-Val-Ile-Ala in order to verify the viability of screening receptor libraries indexed by "ladder" sequences as well as comparing screening results with those of Schmuck.

In a typical screening experiment, about 8 mg of the resin-bound library **7** (sufficient such that the number of beads sampled ensures a statistical likelihood of all library members being represented in the screening experiment[23]) was allowed to swell in a mixture of DMSO/buffer and titrated with aliquots of the dye-labelled tetrapeptide **18** (60–300 μM) also in DMSO/buffer. DMSO was used to ensure solubilisation of the guest tetrapeptide **18**, and the buffer ensured that the guest was present as the free carboxylate, and the guanidinium of the receptor remained protonated. The system was allowed to equilibrate for at least 24 h before visualisation under a microscope or using a hand-lens. Additional aliquots of peptide guest could be added to increase the peptide concentration to provide optimal selectivity, as adjudged by the number of highly stained beads (typically 0.1–1% of beads) against a background of non- or lightly stained beads.

Control experiments confirmed that selective binding observed was not due to difference in resin composition of individual beads (by incubating dye-labelled guest with underivatised resin and noting no selective uptake), and that selective binding is not due to selective recognition of the dye portion (by incubating the library with the dye chromophore alone and noting no selective uptake). Screening the dye-labelled guest peptide against simple peptide libraries lacking any guanidinium carboxylate binding site also showed no selectivity, emphasising the importance of the interaction between the guest carboxylate and the receptor guanidinium, and consistent with all of our earlier work on such receptor libraries.[4a,4i]

Screening experiments were carried out using solvent compositions of 50% DMSO/buffer and 20% DMSO/buffer, and with two different buffers: borax (pH = 9.2) and HEPES (pH = 8.0). The most strongly dyed beads were manually picked from each screening experiment and sequenced using MALDI-TOF MS.

Results from three different screening experiments are summarised in Table 1, Table 2, and Table 3.

Table 1. Sequencing results for highly-stained beads found in library screening experiment using library **7** and guest peptide **18** in 50% DMSO/borax buffer (pH 9.2).

Bead	AA^3	AA^2	AA^1
1	Glu(O*t*Bu)	Val	Phe
2	Leu	Leu	Val
3	Ala	Asn	Gly
4	Val	Val	Lys(Boc)
5	Val	Leu	Leu
6	Leu	Val	Asn
7	Leu	Val	Val
8	Val	Glu(O*t*Bu)	Val
9	Val	Val	Phe
10	Leu	Leu	Val
11	Leu	Val	Val
12	Ala	Leu	Val
13	Phe	Val	Val
14	Glu(O*t*Bu)	Val	Val
15	Gly	Ala	Ser(O*t*Bu)
16	Leu	Val	Val
17	Leu	Leu	Val

From the screening experiments the high preponderence of valine at all three positions (AA^1–AA^3) and structurally very similar leucine at AA^2 and AA^3, is immediately apparent in all screening experiments. Certain receptor structures featured several times across all the screening experiments and in particular the sequence BiG-Leu-Val-Val (BiG = bicyclic guanidinium) was found 4 times, and Val-Leu-Leu was found 5 times out of a total of 43 beads sequenced. Overall the sequences BiG-xxx-Val-Val and BiG-xxx-Leu-Val account for 20 of the 43 beads sequenced. The Glu-

Table 2. Sequencing results for highly-stained beads found in library screening experiment using library **7** and guest peptide **18** in 50% DMSO/HEPES buffer (pH 8.0).

Bead	AA3	AA2	AA1
1	Val	Leu	Val
2	Val	Val	Val
3	Glu(O*t*Bu)	Leu	Leu
4	Glu(O*t*Bu)	Leu	Val
5	Val	Leu	Val
6	Glu(O*t*Bu)	Leu	Val
7	Val	Val	Leu
8	Lys(Boc)	Lys(Boc)	Leu
9	Val	Phe	Val
10	Leu	Val	Val
11	Glu(O*t*Bu)	Val	Val
12	Leu	Val	Glu(O*t*Bu)
13	Leu	Leu	Val
14	Phe	Val	Val
15	Val	Phe	Val
16	Val	Val	Glu(O*t*Bu)

Table 3. Sequencing results for highly-stained beads found in library screening experiment using library **7** and guest peptide **18** in 20% DMSO/borax buffer (pH 9.2).

Bead	AA^3	AA^2	AA^1
1	Val	Leu	Leu
2	Leu	Leu	Val
3	Leu	Lys(Boc)	Val
4	Leu	Val	Leu
5	Val	Lys(Boc)	Leu
6	Ser(O*t*Bu)	Val	Val
7	Phe	Val	Glu(O*t*Bu)
8	–	–	–[a]
9	Leu	Val	Phe
10	Leu	Lys(Boc)	Val

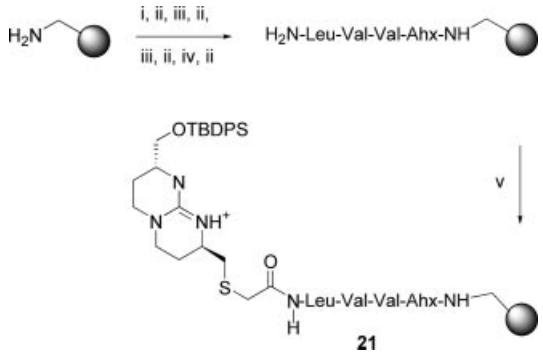
[a] Sequencing was unsuccessful suggesting either unsuccessful coupling or cleavage reactions.

(O*t*Bu) residue was found at AA^3 for 4 out of 16 beads identified using HEPES buffer, and Lys(Boc) was found at AA^2 for 3 of the 10 beads selected from the screening experiment using the lower proportion (20%) of DMSO.

Overall these results compare very favourably with those of Schmuck[4g,4h] who found that in his system receptors featuring the sequences Gua-Val-Val-Val (Gua = guanidinopyrrole) and Gua-Val-Val-Phe were the strongest receptors in MeOH as solvent ($K_a \approx 9000\ M^{-1}$), whereas in water (using bis-tris buffer) receptors of the general structure Lys(Boc)-Lys(Boc)-xxx-Gua were the strongest. With Schmuck's system, for example, the receptor Gua-Leu-Lys-(Boc)-Lys(Boc) bound dansyl-labelled tetrapeptide **19** with $K_{ass} = 2000\ M^{-1}$ (measured using an on-bead binding assay). The analogous sequence BiG-Leu-Lys(Boc)-Lys(Boc) was identified in the screening experiment with library **7** using HEPES buffer, and again from the screening experiment with **7** in the more aqueous solvent (20% DMSO), the sequence Val/Leu-Lys(Boc)-Val/Leu was found for 3 of the 10 beads sequenced.

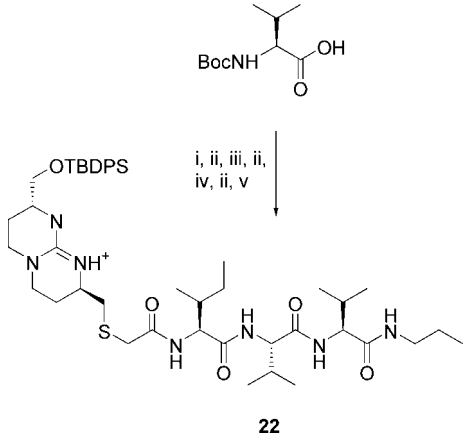
Binding Studies

Given the high occurrence of the sequence BiG-Leu-Val-Val in the screening experiments, this structure was chosen for further investigation. The resin-bound receptor **21** was synthesised using standard Fmoc-SPPS procedures (Scheme 6).



Scheme 6. i) Fmoc-Ahx, DIC, HOBt, DIPEA, DMF, 3 h; ii) 20% piperidine/DMF, 20 min; iii) Fmoc-Val, DIC, HOBt, DIPEA, DMF, 3 h; iv) Fmoc-Leu, DIC, HOBt, DIPEA, DMF, 3 h; v) **6**, EDC, HOBt, DIPEA, DMAP, DMF, 1 h.

The free receptor **22** was synthesised in solution using Boc-protected amino acid couplings (Scheme 7).



Scheme 7. i) CDI, propylamine, THF, 3 h; ii) 20% TFA/DCM, 1 h; iii) Boc-Val-OH, CDI, Et_3N, THF/CMF, 18 h; iv) Boc-Leu-OH, CDI, Et_3N, THF, 18 h; v) **6**, EDC, HOBt, DIPEA, DMAP, DMF, 1 h.

The solid-phase bound receptor **21** was used to determine an on-bead binding constant with the dye-labelled tetrapeptide **18** by determining the ability of the resin-bound receptor to absorb the peptide guest from solution, using previously described methodology.[3c,4a,4h,24] Thus a known mass of resin **21** was incubated with a solution of guest **18** in an aqueous buffer and the change in the concentration of the guest in free solution was monitored by UV spectroscopy. To account for non-specific absorption of **18** to the polystyrene resin matrix, an equivalent mass of underivatised tentagel resin was also incubated with **18**. This allowed the amount of **18** specifically bound by the tweezer receptor to be determined, and hence an association con-

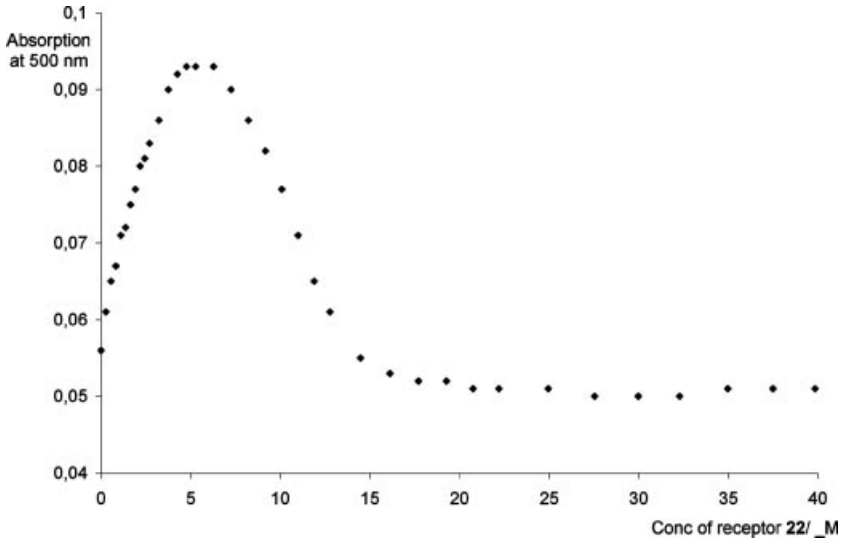


Figure 3. Plot of UV absorption (at 500 nm) of peptide **18** vs. concentration of added receptor **22**.

stant based on a 1:1 receptor-substrate stoichiometry was estimated as $K_{ass} = 1440 \pm 440$ M^{-1}.

Binding of the receptor **22** with the dye-labelled tetrapeptide **18** in free solution was investigated using a reverse titration, monitoring the UV absorption of the dye chromophore of guest **18** during additions of the receptor **22** at constant guest concentration, in 50% DMSO/borax solution. Monitoring the absorption at 500 nm showed a significant increase in the absorption as the first equivalent of the receptor was added but a subsequent decrease as a second equivalent was added (Figure 3).

The data could not be fitted convincingly to simple binding models assuming 1:1, 1:2 or 2:1 binding stoichiometries. The failure to obtain clear binding data for receptors identified from such library screening experiments when transferred to free solution is a common, if frustrating, occurrence, and in part indicates that the environment provided by the solid-phase support can influence the binding selectivities and potency.[4a] Furthermore, in the case of the receptor-substrate combination described in this work (i.e. using models for inhibition of amyloid-β peptide self-assembly) it is not surprising that binding studies in free solution (off-bead) may be complicated by aggregation phenomena, and indeed Schmuck reported[4h] that the simple tetrapeptide **19** formed insoluble, fibril-like aggregates upon standing in solution despite its hydrophilic triethylene glycol chain. None-the-less both the on-bead binding assay and pronounced changes in UV spectrum observed during the titration experiment clearly indicate significant interaction between the receptor and the guest tetrapeptide. Attempted investigation of this interaction by ^{1}H NMR was unsuccessful due to solubility problems with the receptor **22** and peptide **18** at suitable concentrations, again providing evidence of the likely aggregation of these compounds.

Conclusions

In summary a modified capping methodology has been developed to facilitate the rapid sequencing of "hit" structures identified from combinatorial libraries of peptide receptors. The methodology has been adapted for the synthesis of a library of single-armed peptide receptors **7** and a library of small "tweezer" receptors **17**. The synthesis of the latter library required the tandem synthesis of "ladder sequences" separated from the full receptor. Library **7** was successfully screened to identify binding partners for the *C*-terminal tetrapeptide sequence of the amyloid-β protein (Aβ), with notable structural similarities to similar amyloid-β protein binding partners identified by Schmuck. This capping methodology is now being used for the synthesis of structurally more complex encoded receptor libraries and results from this work will be reported in due course.

Experimental Section

General Methods: Commercially available compounds were used without further purification. Tentagel-S-NH_2 resin was used as solid support and purchased from Rapp Polymere (Germany). Amino acid derivatives were purchased from NovaBiochem or BAChem. 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). All coupling reactions were monitored with the Kaiser Ninhydrin Test.[17] 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. Melting points were determined in open capillary tubes using a Gallenkamp Electrothermal Melting Point Apparatus, all are uncorrected. Infrared spectra were recorded with 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 with a Bruker AC 300 or Bruker AV 300 and at 400 MHz with a Bruker DPX 400. Carbon NMR spectra were recorded at 75 MHz with a Bruker AC300 and at 100 MHz with a Bruker DPX 400. UV Spectra were recorded with a Shimadzu UV-1601 spectrometer in quartz cells. Low-resolution ESI mass spectra were recorded with a Micromass Platform II quadrupole mass analyser (Fisons VG platform through a Hewlett Packard 1050 HPLC

system). High-resolution ESI mass spectra were obtained with a Bruker Apex III FT-ICR mass spectrometer.

Abbrevations: Ahx = 6-aminohexanoyl, Boc = *tert*-butyloxycarbonyl, CDI = carbonyldiimidazole, DBU = 1,8-diazabicyclo[5.4.0]-undec-7-ene, DIC = *N*,*N*′-diisopropylcarbodiimide, DIPEA = diisopropylethylamine, DMAP = 4-(dimethylamino)pyridine, DMF = dimethylformamide, Fmoc = 9-fluorenyloxycarbonyl, HOBt = 1-hydroxybenzotriazole, Mtr = 4-methoxy-2,3,6-trimethylbenzenesulphonyl, Pbf = 2,2,4,6,7-pentamethyldihydrobenzofuran-5-sulfonyl, PyBOP = benzotriazol-1-yloxy-tris(pyrrolino)phosphonium hexafluorophosphate, SPPS = solid-phase peptide synthesis, TFA = trifluoroacetic acid, TMS = trimethylsilyl.

Benzyl 4-[{3-[(*tert*-Butoxycarbonyl)amino]propylamino}thiocarbonyl]amino]butanoate (10): Ester **8** (2.12 g, 7.22 mmol) was stirred in 20% TFA/DCM (30 mL) for 1 h. The solvents were removed in vacuo by azeotropic distillation with toluene and the residue redissolved in 1:1 MeOH/DCM (40 mL). Isothiocyanate **9**[16] (1.56 g, 7.22 mmol) was added, followed by Et_3N (5 mL, 36.1 mmol). The mixture was stirred for 18 h and then evaporated in vacuo. The residue was purified by column chromatography (eluent 30% EtOAc/petroleum ether) to give the thiourea **10** as a yellow oil (2.22 g, 75%). IR (cm^{-1}): $\tilde{\nu}_{max}$ = 3325 (br. w), 2974 (w), 2932 (w), 1687 (m), 1548 (m), 1516 (m), 1251 (m), 1163 (s). ^1H NMR (300 MHz, $CDCl_3$): δ = 7.34 (m, 5 H, arom), 6.40 (br. s, 1 H, N*H*), 5.12 (s, 2 H, CH_2Ph), 4.90 (br. s, 1 H, N*H*), 3.61 (m, 2 H, SCNHCH_2), 3.41 (m, 2 H, SCNHCH_2), 3.17 (m, 2 H, BocHNCH_2), 2.46 (t, J = 7 Hz, 2 H, CH_2COOBn), 1.93 (quin, J = 7 Hz, 2 H, $CH_2CH_2CH_2$); 1.71 (quin, J = 7 Hz, 2 H, $CH_2CH_2CH_2$); 1.42 (s, 9 H, Boc) ppm. ^{13}C NMR (75 MHz, $CDCl_3$): 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) ppm. LRMS (ESI^+): m/z = 410.2 $[M + H]^+$ 432.2 $[M + Na]^+$. HRMS: (ESI^+) 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 (11): Iodomethane (3.4 mL, 54.1 mmol) was added to a solution of the thiourea **10** (2.22 g, 5.41 mmol) in acetone (40 mL). The mixture was stirred for 1 h and then evaporated in vacuo. The residue was redissolved in 1:1 MeOH/DCM (40 mL) and ammonium hexafluorophosphate (1.77 g, 10.8 mmol) was added. The mixture was stirred for 18 h and then evaporated in vacuo. The residue was taken up in DCM and washed once with distilled water. The organic phase was dried with $MgSO_4$ and evaporated to give a white foam. The foam was dissolved in 4:1 toluene/$CHCl_3$ (50 mL). Trifluoroacetamide (3.06 g, 27.1 mmol) and DBU (1.6 mL, 10.8 mmol) were added and the mixture heated at reflux for 1 h. After cooling, the solvents were removed in vacuo and the residue purified by column chromatography (20% EtOAc/petroleum ether) to give the guanidine **11** as a yellow oil (1.31 g, 50%). IR (cm^{-1}): $\tilde{\nu}_{max}$ = 3333 (br. w), 2979 (w), 1690 (m), 1624 (s), 1517 (m), 1434 (m), 1167 (s), 1137 (s). ^1H NMR (300 MHz, $CDCl_3$): δ = 9.56 (br. s, 1 H, N*H*), 7.34 (m, 5 H, arom), 6.39 (br. s, 1 H, NH), 5.14 (br. s, 1 H, N*H*) 5.13 (s, 2 H, CH_2Ph), 3.52 (m, 2 H, BocNHCH_2), 3.26 (m, 2 H, NHCH_2), 3.17 (dt, J = 6, 7 Hz, 2 H, NHCH_2), 2.47 (t, J = 7 Hz, 2 H, CH_2COOBn), 1.94 (quin, 2 H, $CH_2CH_2CH_2$), 1.65 (m, 2 H, $CH_2CH_2CH_2$), 1.42 (s, 9 H, Boc) ppm. ^{13}C NMR (75 MHz, $CDCl_3$): δ = 173.3 (C), 166.7 (q, J = 35 Hz, C), 160.8 (C), 157.2 (C), 135.6 (C), 128.7 (CH), 128.5 (CH), 128.3 (CH), 117.1 (q, J = 286 Hz, 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) ppm. LRMS (ESI^+): m/z = 489.3 $[M + H]^+$, 511.3 $[M + Na]^+$. HRMS: (ESI^+) calculated for $C_{22}H_{31}F_3N_4O_5$ $[M + H]^+$ 489.2320, found 489.2317.

4-({3-[(*tert*-Butoxycarbonyl)amino]propylamino}[(2,2,2-trifluoroacetyl)imino]methylamino)butanoic Acid (12): Palladium on carbon (10% by weight, 175 mg) was added to a solution of the carbamate **11** (804 mg, 1.65 mmol) in EtOAc (20 mL) and the mixture stirred under hydrogen for 3 h. The reaction mixture was filtered through celite and the residue washed with EtOAc. The combined filtrates were evaporated to give the acid **12** as a white solid (619.3 mg, 94%); m.p. 104–105 °C. IR (cm^{-1}): $\tilde{\nu}_{max}$ = 3281 (w), 2960 (w), 1742 (w), 1682 (w), 1637 (s), 1521 (s), 1448 (m), 1389 (m), 1366 (m), 1158 (s), 660 (m). ^1H NMR (300 MHz, [D$_6$]DMSO): δ = 12.14 (br. s, 1 H, COO*H*), 9.04 (br. s, 1 H, N*H*), 7.78 and 7.63 (br. s, 1 H, N*H*), 6.85 and 6.75 (br. s, 1 H, N*H*), 3.29 (m, 2 H, NHCH_2), 3.19 (m, 2 H, NHCH_2), 2.95 (m, 2 H, BocNHCH_2), 2.22 (t, J = 7.5 Hz, 2 H, CH_2COOH), 1.72 (m, 2 H, $CH_2CH_2CH_2$), 1.60 (m, 2 H, $CH_2CH_2CH_2$), 1.37 (s, 9 H, Boc) ppm. ^{13}C NMR: δ = 174.2 (C), 160.0 (C), 155.9 (C), 120.8 (q, J = 260 Hz, 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) ppm, *C*OCF_3 not visible. LRMS (ESI^+): m/z = 303.0 $[M + H - COCF_3]^+$, 421.0 $[M + Na]^+$. HRMS: (ESI^+) calculated for $C_{15}H_{25}F_3N_4O_5$ $[M + Na]^+$ 421.1669, found 421.1665.

General Procedures for Solid-Phase Reactions and Solid-Phase Tests

Resin Preparation and Washing: Prior to reaction, resins were swollen for 15 min in DCM. After completion of reaction, resins were drained of the reaction medium and extensively washed with DCM, DMF, and again with DCM. Each washing consisted of suspending resin in solvent (≈ 15 mL per g of resin) which was allowed to stand for one minute before draining and then repeating the process twice. Any additional/alternate washing is noted in the appropriate section.

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. Protected amino acid (3 equiv. based on resin loading) and HOBt (3 equiv.) were dissolved in DMF (≈ 1 mL per 50 mg of amino acid). DIC (3 equiv.) was added and the mixture added to the pre-swollen resin. Finally excess DIPEA (3–5 equiv.) was added, the reaction vessel sealed and vented, and the mixture agitated for at least one hour. The reaction was repeated as necessary until the resin gave a negative ninhydrin test, and finally washed according to the general procedure.

Solid-Phase Fmoc-Deprotection: Pre-swollen resin was treated with 20% piperidine/DMF (≈ 20 mL per g of resin) for 10 min. The process was repeated before washing as described above.

Solid-Phase Boc-Deprotection: Pre-swollen resin was treated with 20% TFA/DCM (≈ 20 mL per g of resin) for 20 min. The resin was drained and washed with 50% DIPEA/DMF (≈ 20 mL per g of resin), before washing as described above.

General Protocol for the Capping Reaction: Pre-swollen resin to be capped was shaken with a solution of *p*-BrC_6H_4COOH in DMF (for precise quantities see below) for 1 h, before addition of excess DIC, HOBt and DIPEA (>10 equiv.) and further shaking for 18 h. The resin was then drained and washed according to the general procedure described above.

Synthesis of Resin 2: Tentagel resin (0.2 mmol/g, 642 mg, 0.11 mmol) was pre-swollen using DCM and coupled with Fmoc-Met-OH (122 mg, 0.33 mmol) using DIC (51 μL, 0.33 mmol), HOBt (44 mg, 0.33 mmol) and DIPEA (57 μL, 0.33 mmol) in DMF (5 mL). The resin was washed and gave a negative ninhydrin

test. Fmoc-deprotection was followed by three cycles of amino acid coupling/Fmoc-deprotection using Fmoc-Phe-OH (127 mg, 0.33 mmol), DIC (51 μL, 0.33 mmol), HOBt (44 mg, 0.33 mmol) and DIPEA (57 μL, 0.33 mmol) in DMF (5 mL). A final amino acid coupling using Fmoc-Ahx-OH (116 mg, 0.33 mmol), DIC (51 μL, 0.33 mmol), HOBt (44 mg, 0.33 mmol) and DIPEA DIPEA (57 μL, 0.33 mmol) in DMF (5 mL) was carried out followed by Fmoc-deprotection. The resin **2** was finally rinsed with Et_2O and dried in vacuo. Quantitative ninhydrin test gave the resin loading as 0.13 mmol/g.

Synthesis of Resin 5: Tentagel resin (0.2 mmol/g, 551 mg, 121 μmol) was pre-swollen using DCM. Amino acid coupling was carried out using Fmoc-Met-OH (90 mg, 242 μmol), DIC (57 μL, 363 μmol), HOBt (49 mg, 363 μmol) and DIPEA (63 μL, 363 μmol) in DMF (5 mL) followed by Fmoc-deprotection. A second amino acid coupling was carried out using Fmoc-Arg(Pbf)-OH (157 mg, 242 μmol), DIC (57 μL, 363 μmol), HOBt (49 mg, 363 μmol) and DIPEA (63 μL, 363 μmol) in DMF (5 mL) followed by Fmoc-deprotection. Two cycles of amino acid coupling, followed by Fmoc-deprotection was carried out using Fmoc-Ahx-OH (86 mg, 242 μmol), DIC (57 μL, 363 μmol), HOBt (49 mg, 363 μmol) and DIPEA (63 μL, 363 μmol) in DMF (5 mL). Resin **5** was finally rinsed with Et_2O and rigorously dried in vacuo. A quantitative ninhydrin test gave the resin loading as 0.15 mmol/g.

Synthesis of Resin 3: Resin **2** (12 mg, 0.13 mmol/g, 1.5 μmol) was pre-swollen using DCM. Amino acid coupling using Fmoc-Gly-OH (1.4 mg, 4.56 μmol), DIC (5 μL, excess), HOBt (5 mg, excess) and DIPEA (5 μL, excess) in DMF (2 mL) was followed by Fmoc-deprotection. A capping reaction was then performed using *p*-BrC_6H_4COOH (259 μL, 1.47 mM/DMF, 0.38 μmol) and DIC (5 μL), HOBt (5 mg), DIPEA (5 μL). A second amino acid coupling using Fmoc-Ala-OH (1.4 mg, 4.56 μmol), DIC (5 μL, excess), HOBt (5 mg, excess) and DIPEA (5 μL, excess) in DMF (2 mL) was followed by Fmoc-deprotection and again a capping reaction using *p*-BrC_6H_4COOH (259 μL, 1.47 mM/DMF, 0.38 μmol) and DIC (5 μL), HOBt (5 mg), DIPEA (5 μL). A third amino acid coupling using Fmoc-Ile-OH (1.6 mg, 4.56 μmol), DIC (5 μL, excess), HOBt (5 mg, excess) and DIPEA (5 μL, excess) in DMF (2 mL) was followed by Fmoc-deprotection and a capping reaction using *p*-BrC_6H_4COOH (330 μL, 1.15 mM/DMF, 0.38 μmol) and DIC (5 μL), HOBt (5 mg), DIPEA (5 μL). The resin **3** was finally rinsed in Et_2O and dried in vacuo.

Synthesis of Library 4: Resin **2** (80 mg, 0.13 mmol/g) was divided into four equal portions (≈ 20 mg each) and pre-swollen using DCM. Each portion was coupled with an Fmoc-protected amino acid (7.8 μmol) using DIC (10 μL, excess), HOBt (5 mg, excess) and DIPEA (10 μL, excess) in DMF (2 mL) and one of the following amino acids: Fmoc-Ser(O*t*Bu)-OH, Fmoc-Ala-OH, Fmoc-Val-OH, Fmoc-Lys(Boc) followed by Fmoc-deprotection. Each resin sample was then subjected to a capping reaction using *p*-BrC_6H_4COOH (565 μL of a 1.15 mM solution in DMF, 0.65 μmol), DIC (10 μL, excess), HOBt (5 mg, excess) and DIPEA (10 μL, excess). The resin portions were then combined, suspended in DCM and shaken for 10 min to ensure thorough mixing. Two further iterations of the same procedure were carried out (division into four roughly equal portions; amino acid coupling; washing; Fmoc-deprotection; capping reaction) with the same quantities of reagents as before. Finally the resin **4** was rinsed with Et_2O and dried in vacuo. 20 single beads were selected, cleaved and analysed by MALDI-TOF MS.

Synthesis of Single-Armed Receptor Library 7: Resin **5** (0.09 mmol/g, 735 mg) was split into nine equal portions (≈ 82 mg each) and was pre-swollen according to the general procedure. Each portion was coupled with an Fmoc-protected amino acid (22 μmol) and DIC (10 μL, excess), HOBt (10 mg, excess), and DIPEA (10 μL, excess) in DMF (2 mL) and one of the following amino acids: Fmoc-Gly-OH, Fmoc-Ala-OH, Fmoc-Val-OH, Fmoc-Leu-OH, Fmoc-Phe-OH, Fmoc-Glu(OBn)-OH, Fmoc-Asn-OH, Fmoc-Ser(O*t*Bu)-OH, Fmoc-Lys(Boc), followed by Fmoc-deprotection. The resin was then subjected to a capping reaction using *p*-BrC_6H_4COOH (1.27 mL of a 1.15 mM solution in DMF, 1.47 μmol), DIC (10 μL, excess), HOBt (5 mg, excess) and DIPEA (10 μL, excess). The resin portions were then combined, suspended in DCM and shaken for 10 min to ensure thorough mixing. Two further iterations of the same procedure were carried out (division into four roughly equal portions; amino acid coupling; Fmoc-deprotection; capping reaction; mixing) with the same quantities of reagents as before. The resin was finally coupled with bicyclic guanidinium **6** (55 mg, 108 μmol) and DIC (34 μL, 216 μmol), HOBt (29 mg, 216 μmol) and DIPEA (38 μL, 216 μmol) 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 bicyclic guanidinium **6**. Resin **7** was rinsed with Et_2O and dried in vacuo.

Solid-Phase-Linked Receptor 21: Tentagel resin (119 mg, 0.2 mmol/g, 23.8 μmol) was pre-swollen according to the general procedure. Amino acid coupling was carried out using Fmoc-Ahx-OH (17 mg, 47.6 μmol), DIC (11 μL, 71.4 μmol), HOBt (10 mg, 71.4 μmol) and DIPEA (12 μL, 71.4 μmol) in DMF (3 mL) followed by Fmoc-deprotection. The resin was subjected to two iterations of amino acid coupling using Fmoc-Val-OH (24 mg, 71.4 μmol), DIC (11 μL, 71.4 μmol), HOBt (10 mg, 71.4 μmol) and DIPEA (12 μL, 71.4 μmol) in DMF (3 mL), followed by Fmoc-deprotection. The resin was subjected to a further amino acid coupling using Fmoc-Leu-OH (25 mg, 71.4 μmol), DIC (11 μL, 71.4 μmol), HOBt (10 mg, 71.4 μmol) and DIPEA (12 μL, 71.4 μmol) in DMF (3 mL) followed by Fmoc-deprotection. Finally, the resin was coupled with acid **6** (18 mg, 35.7 μmol), DIC (11 μL, 71.4 μmol), HOBt (10 mg, 71.4 μmol) and DIPEA (12 μL, 71.4 μmol) in DMF (3 mL). After washing the resin gave a negative ninhydrin test. The resin **21** was rinsed with Et_2O and dried in vacuo.

Receptor 22: *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 (62 mg, 0.13 mmol) was stirred in 20% TFA/DCM (2 mL) for 2 h. The reaction mixture was rigorously evaporated in vacuo using toluene as an azeotrope. The residue was redissolved in DMF (3 mL) and acid **12** (67 mg, 0.13 mmol), HOBt (20 mg, 0.14 mmol), DIPEA (45 μL, 0.26 mmol) and DMAP (2 mg) were added followed by EDC (28 mg, 0.14 mmol). The mixture was stirred for 1 h and then evaporated in vacuo. The residue was redissolved in MeOH and a white solid precipitated on addition of Et_2O. The precipitate was recrystallised from MeCN to give the receptor **22** as a white solid (10 mg, 9%); m.p. 233–235 °C. IR (cm^{-1}): $\tilde{\nu}_{max}$ = 2957 (w), 1626 (s), 1543 (m), 1459 (m), 1310 (m), 1221 (m), 1101 (s), 741 (m), 701 (s). ^1H NMR (400 MHz, [D$_6$]DMSO): δ = 8.20 (d, J = 8.0 Hz, 1 H, N*H*), 7.94 (d, J = 9.0 Hz, 1 H, N*H*), 7.92 (d, J = 8.0 Hz, 1 H, N*H*), 7.85 (t, J = 5.5 Hz, 1 H, N*H*), 7.65–7.58 (m, 6 H, 5 C$H_{arom.}$ + N*H*), 7.50–7.41 (m, 6 H, 5 C$H_{arom.}$ + N*H*), 4.37 (q, J = 8 Hz, 1 H, Leu α-C*H*), 4.16 (dd, J = 8.5, 7.5 Hz, 1 H, α-C*H*), 4.07 (dd, J = 8.5, 7.5 Hz, 1 H, α-C*H*), 3.64 (m, 2 H), 3.57 (m, 2 H), 3.42–3.21 (m, 3 H), 3.23 (s, 2 H, SCH_2CO), 3.08–2.92 (m, 2 H), 2.80–2.69 (m, 2 H), 2.36–1.85 (m, 4 H), 1.84–1.70 (m, 2 H), 1.59 (m, 1 H), 1.46–1.34 (m, 4 H), 1.03 (s, 9 H, *t*BuSi), 0.90–0.79 (m, 21 H, 7CH_3) ppm. ^{13}C NMR (100 MHz, [D$_6$]DMSO): δ = 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) ppm. LRMS (ESI^+): *m/z* = 864.7 $[M - Cl]^+$. HRMS: (ESI^+) calculated for $C_{46}H_{74}N_7O_5S_1Si_1$ $[M^+]$ 864.5236, found 864.5228.

Mass Spectrometric Methods for Single Bead Analysis: Deprotection and Cleavage of Material from Single Resin Beads: A single resin bead to be analysed was placed in a glass conical insert (100 μL 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 50 μL of 50% TFA/DCM for 1 h and the solvent was removed in vacuo.

Mtr Deprotection: The bead was treated with 40 μL of 1 M TMSBr in TFA for 1 h and the reagent mixture removed in vacuo. Cleavage was performed by treatment of the bead with 20 μL of a solution of CNBr in 1:1 TFA/H_2O (concentration 50–80 mg/mL) and incubation in the dark for 18 h.

Sample Preparation: The cleavage solution was evaporated from the sample tube in vacuo. 20 μL of either MeCN or 1:1 MeCN/H_2O was then added to the tube. The samples were sonicated for 5 min. 1 μL of solution was spotted onto the MALDI-platen. 1 μL of a solution of matrix was then applied on top and mixed with the sample on the platen surface. The spot was allowed to dry under a stream of warm air before analysis. Matrix was prepared as a saturated solution of α-cyanohydroxycinnamic acid 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.

MALDI Acquisition and Data Processing: Spectra were acquired with a Micromass-Tofspec 2E reflectron MALDI-TOF mass spectrometer, recording positive ion data at an accelerating voltage of 20 kV and a pulse voltage of 3000 V. 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 $[M + Na]^+$. Data was acquired and processed using Masslynx v3.2.[12] 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%.

Screening Experiments: Receptor library resin **7** (≈ 8 mg) was placed in a flat-bottomed glass dish and pre-swollen in DMSO/buffer (0.4 mL) for 1 h. A solution of the dye-labelled guest **18** (typically 60–300 μM) in the same solvent was then added to the mixture and the system incubated for 24 h. Further aliquots of guest **18** could be added as to the library, followed by further incubation periods of 24 h, to provide optimal selectivity as adjudged by the number of highly stained beads against a background of non- or lightly stained beads. 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.

Acknowledgments

We thank the Commission of the European Union (TMR Network grant "Enantioselective Separations" ERB FMRX-CT-98-0233) for financial support and BBSRC for a postgraduate studentship (J. S.). We also thank Prof. J. de Mendoza, Institute of Chemical Research of Catalonia (ICIQ), Tarragona, Spain and his research group for providing us with a sample of the bicyclic guanidinium derivative **6**.

[1] K. S. Lam, M. Lebl, V. Krchnak, *Chem. Rev.* **1997**, *97*, 411–448.
[2] Reviews: a) N. Srinivasan, J. D. Kilburn, *Curr. Opin. Chem. Biol.* **2004**, *8*, 305–310; b) B. Linton, A. D. Hamilton, *Curr. Opin. Chem. Biol.* **1999**, *3*, 307–312; c) Y. R. Miguel, J. M. K. Sanders, *Curr. Opin. Chem. Biol.* **1998**, *2*, 417–421; d) W. C. Still, *Acc. Chem. Res.* **1996**, *29*, 155–163.
[3] a) H. Wennemers, M. C. Nold, M. Conza, K. J. Kulicke, M. Neuburger, *Chem. Eur. J.* **2003**, *9*, 442–448; b) H. Wennemers, M. Conza, *J. Org. Chem.* **2002**, *67*, 2696–2698; c) R. Arienzo, J. D. Kilburn, *Tetrahedron* **2002**, *58*, 711–719; d) T. Braxmeier, M. Demarcus, T. Fessmann, S. McAteer, J. D. Kilburn, *Chem. Eur. J.* **2001**, *7*, 1889–1898; e) J. D. Kilburn, T. Fessmann, *Angew. Chem.* **1999**, *111*, 2170–2174; *Angew. Chem. Int. Ed. Angew. Chem. Int. Ed. Engl.* **1999**, *38*, 1993–1996.
[4] a) J. Shepherd, T. Gale, K. B. Jensen, J. D. Kilburn, *Chem. Eur. J.* **2006**, *12*, 713–720; b) C. Schmuck, M. Heil, *Chem. Eur. J.* **2006**, *12*, 1339–1348; c) C. Schmuck, P. Wich, *Angew. Chem.* **2006**, *118*, 4383–4387; *Angew. Chem. Int. Ed. Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 4277–4281; d) C. Schmuck, M. Heil, J. Scheiber, K. Baumann, *Angew. Chem.* **2005**, *117*, 7374–7379; *Angew. Chem. Int. Ed. Angew. Chem. Int. Ed. Engl.* **2005**, *44*, 7208–7212; e) C. Chamorro, R. M. J. Liskamp, *Tetrahedron* **2004**, *60*, 11145–11157; f) C. Chamorro, J.-W. Hofman, R. M. J. Liskamp, *Tetrahedron* **2004**, *60*, 8691–8699; g) C. Schmuck, M. Heil, *Org. Biomol. Chem.* **2003**, *5*, 633–636; h) C. Schmuck, M. Heil, *ChemBioChem* **2003**, *4*, 1232–1238; i) K. B. Jensen, T. M. Braxmeier, M. Demarcus, J. D. Kilburn, *Chem. Eur. J.* **2002**, *8*, 1300–1309.
[5] Other recent examples of sequence selective peptide receptors: a) C. Schmuck, L. Geiger, *J. Am. Chem. Soc.* **2004**, *126*, 8898–8899; b) S. Rensing, T. Schrader, *Org. Lett.* **2002**, *4*, 2161; c) R. Xu, G. Greiveldinger, L. E. Marenus, A. Cooper, J. A. Ellman, *J. Am. Chem. Soc.* **1999**, *121*, 4898–4899; d) M. Sirish, H.-J. Schneider, *Chem. Commun.* **1999**, 907–908; e) M. A. Hossain, H.-J. Schneider, *J. Am. Chem. Soc.* **1998**, *120*, 11208–11209; f) R. Breslow, Z. Yang, R. Ching, G. Trojandt, F. Odobel, *J. Am. Chem. Soc.* **1998**, *120*, 3536–3537.
[6] P. Edman, G. Begg, *Eur. J. Biochem.* **1967**, *1*, 80–91.
[7] For other approaches to encoding and sequencing of combinatorial libraries see, for example: a) M. G. Paulick, K. M. Hart, K. M. Brinner, M. Tiandra, D. H. Charych, R. N. Zuckermann, *J. Comb. Chem.* **2006**, *8*, 417–426; b) X. Wang, L. Peng, R. Liu, S. S. Gill, K. S. Lam, *J. Comb. Chem.* **2005**, *7*, 197–209; c) X. Wang, J. Zhang, A. Song, C. B. Lebrilla, K. S. Lam, *J. Am. Chem. Soc.* **2004**, *126*, 5740–5749; d) R. M. Sanchez-Martin, S. Mittoo, M. Bradley, *Curr. Top. Med. Chem.* **2004**, *4*, 653–669; e) R. L. Affleck, *Curr. Opin. Chem. Biol.* **2001**, *5*, 257–263; f) C. Barnes, S. Balasubramanian, *Curr. Opin. Chem. Biol.* **2000**, *4*, 346–350; g) A. W. Czarnik, *Curr. Opin. Chem. Biol.* **1997**, *1*, 60–66.
[8] a) R. S. Youngquist, G. R. Fuentes, M. P. Lacey, T. Keough, *Rapid Commun. Mass Spectrom.* **1994**, *8*, 77–81; b) R. S. Youngquist, G. R. Fuentes, M. P. Lacey, T. Keough, *J. Am. Chem. Soc.* **1995**, *117*, 3900–3906; c) N. Sepetov, O. Issakova, V. Krchnak, M. Lebl, US Patent 5470753, **1995**.
[9] J. J. Pastor, I. Lingard, G. Bhalay, M. Bradley, *J. Comb. Chem.* **2003**, *5*, 85–90.

[10] During the course of our studies Lam et al. have described the introduction of a 3-(4-bromophenyl)-β-alanine into the laddered peptide sequence which also provides the characteristic bromine doublet: see ref.[7b,7c]
[11] M. Davies, M. Bradley, *Tetrahedron Lett.* **1997**, *38*, 8565–8568.
[12] Masslynx v3.2, Micromass Ltd., Manchester, UK.
[13] A. M. Echavarren, A. Galán, J. de Mendoza, A. Salmerón, J. M. Lehn, *Helv. Chim. Acta* **1988**, *71*, 685–693.
[14] a) W. Peschke, P. Scheissl, F. P. Schmidtchen, P. Bissinger, A. Schier, *J. Org. Chem.* **1995**, *60*, 1039–1043; b) H. Kurzmeier, F. P. Schmidtchen, *J. Org. Chem.* **1990**, *55*, 3749.
[15] Compound **6** was kindly provided by Prof. J. de Mendoza, Institute of Chemical Research of Catalonia (ICIQ), Tarragona, Spain.
[16] S. Bartoli, K. B. Jensen, J. D. Kilburn, *J. Org. Chem.* **2003**, *68*, 9416–9422.
[17] V. K. Sarin, S. B. H. Kent, J. P. Tam, R. B. Merrifield, *Anal. Biochem.* **1981**, *117*, 147–157.
[18] a) E. Atherton, R. C. Sheppard, J. D. Wade, *J. Chem. Soc., Chem. Commun.* **1983**, 1060–1062; b) L. Guo, S. Funakoshi, N. Fujii, H. Yajima, *Chem. Pharm. Bull.* **1988**, *36*, 4989–4992.
[19] J. Hardy, D. J. Selkoe, *Science* **2002**, *297*, 353–356.
[20] a) D. I. Dominguez, B. De Strooper, *Trends Pharmacol. Sci.* **2002**, *23*, 324–330; b) P. T. Lansbury Jr, *Curr. Opin. Chem. Biol.* **1997**, *1*, 260–267.
[21] See for example: a) C. Schmuck, P. Frey, M. Heil, *ChemBioChem* **2005**, *6*, 628–631; b) Y. Kuroda, K. Maeda, H. Hanaoka, K. Miyamoto, T. Nakagawa, *J. Pept. Sci.* **2004**, *10*, 8–17; c) P. Rzepicki, H. Gallmeier, N. Gieb, K. Cernovska, B. König, T. Schrader, *J. Org. Chem.* **2004**, *69*, 5168–5178; d) C. Hetényi, T. Körtvélyesi, B. Penke, *Bioorg. Med. Chem.* **2002**, *10*, 1587–1593; e) T. L. Lowe, A. Strzelec, L. L. Kiessling, R. M. Murphy, *Biochemistry* **2002**, *40*, 7882–7889; and references cited therein.
[22] Guanidinopyrroles as carboxylate receptors: C. Schmuck, L. Geiger, *J. Am. Chem. Soc.* **2004**, *126*, 8898–8899; C. Schmuck, *Chem. Commun.* **1999**, 843–844.
[23] K. Burgess, A. I. Liaw, N. Wang, *J. Med. Chem.* **1994**, *37*, 2985–2987.
[24] a) M. Conza, H. Wennemers, *Chem. Commun.* **2003**, 866–867; b) M. C. F. Monnee, A. J. Brouwer, L. M. Verbeek, A. M. A. van Wageningen, R. M. J. Liskamp, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 1521–1525; c) Y. Cheng, T. Suenaga, W. C. Still, *J. Am. Chem. Soc.* **1996**, *118*, 1813–1814; d) P. W. Smith, G. Chang, W. C. Still, *J. Am. Chem. Soc.* **1988**, *53*, 1587–1590.

Received: November 20, 2006
Published Online: January 10, 2007

