ARTICLE

OBC
www.rsc.org/obc

Metal-ion induced amplification of three receptors from dynamic combinatorial libraries of peptide-hydrazones†

Sarah L. Roberts,[a] Ricardo L. E. Furlan,[b] Sijbren Otto[a] and Jeremy K. M. Sanders*[a]

[a] *University Chemical Laboratory, Lensfield Road, Cambridge, UK CB2 1EW. E-mail: jkms@cam.ac.uk; Fax: ++44 (0)1223-336017; Tel: ++44 (0)1223-336343*

[b] *Universidad Nacional de Rosario, Facultad de Ciencias Bioquímicas y Farmacéuticas, Suipacha 531, 2000 Rosario, Argentina. E-mail: rfurlan@fbioyf.unr.edu.ar; Fax: ++54 (0)3414-375315; Tel: ++54 (0)3414-375315*

Received 24th January 2003, Accepted 18th March 2003
First published as an Advance Article on the web 9th April 2003

Three building blocks of general structure $(MeO)_2CH$–aromatic linker–Pro–amino acid–$NHNH_2$ have been prepared and tested in acid-catalysed dynamic combinatorial libraries. Exposure of these libraries to LiI and NaI led to the amplification of three macrocyclic pseudopeptide receptors. The receptors were isolated and their interactions with LiI and NaI were analysed using NMR, IR and ITC. Binding of the metal ions to the receptors is invariably entropy-driven. Nevertheless, all receptors were found to be flexible with substantial conformational rearrangements accompanying guest binding. This type of receptor is extremely difficult to access through rational design and the fact that dynamic combinatorial chemistry allows facile access to these challenging molecules underlines the power of the dynamic approach.

Introduction

Dynamic combinatorial chemistry has attracted interest over recent years as a new strategy for the preparation and identification of novel host–guest systems.[1] The approach integrates molecular evolution into combinatorial chemistry, and in principle merges the preparation and screening of libraries into one single process.

Dynamic combinatorial libraries (DCLs) are generated by the assembly of building blocks through reversible bonds, so the library product distribution is thermodynamically controlled and responsive to external influences. Stabilization of one particular member of the library by selective binding to a guest molecule will shift the equilibrium, leading to an increased concentration of the selected compound at the expense of the other library members. In this way, a guest can be used as a template to select and amplify its preferred host from a library of potential receptors.

Several different covalent and non-covalent reactions have been used to generate DCLs. These include base- and Pd-catalyzed transesterifications,[2] transimination of imines, oximes and hydrazones,[3] alkene metathesis,[4] disulfide exchange,[5] enzyme-catalyzed aldol condensation and transamination,[6] hydrogen-bond exchange,[7] metal–ligand coordination[8] and *cis–trans* isomerisation.[9] However, despite these many reports, the efficient amplification of covalently-assembled receptors or guests seems to be more elusive and few examples of significant shifts in library distribution have been reported so far.[3a,10] We have previously reported the amplification of a receptor for Li cations from a library of pseudopeptide-hydrazones;[10d] we now describe (a) the isolation and binding properties of that receptor in more detail, and (b) the amplification, isolation and properties of two new cation-binding receptors from related building blocks.

Our approach to hydrazone DCLs is summarized in Fig. 1. A range of interconverting macrocycles can be generated from bifunctional building blocks containing one hydrazide and one protected aldehyde. The addition of acid results in rapid deprotection of the aldehyde and induces hydrazone formation and exchange, leading to the assembly and interconversion of macrocycles. Neutralization of the reaction medium switches off exchange, converting the library into a 'static' mixture from which the individual members can be isolated.

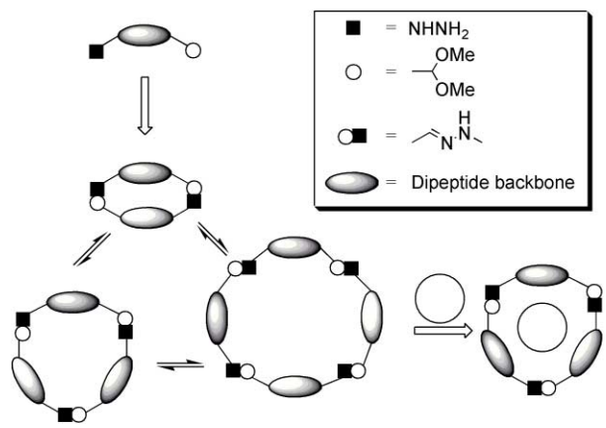
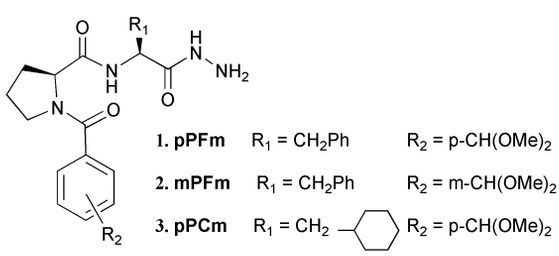


Fig. 1 Generation and templating of a dynamic combinatorial library of macrocyclic hydrazones.

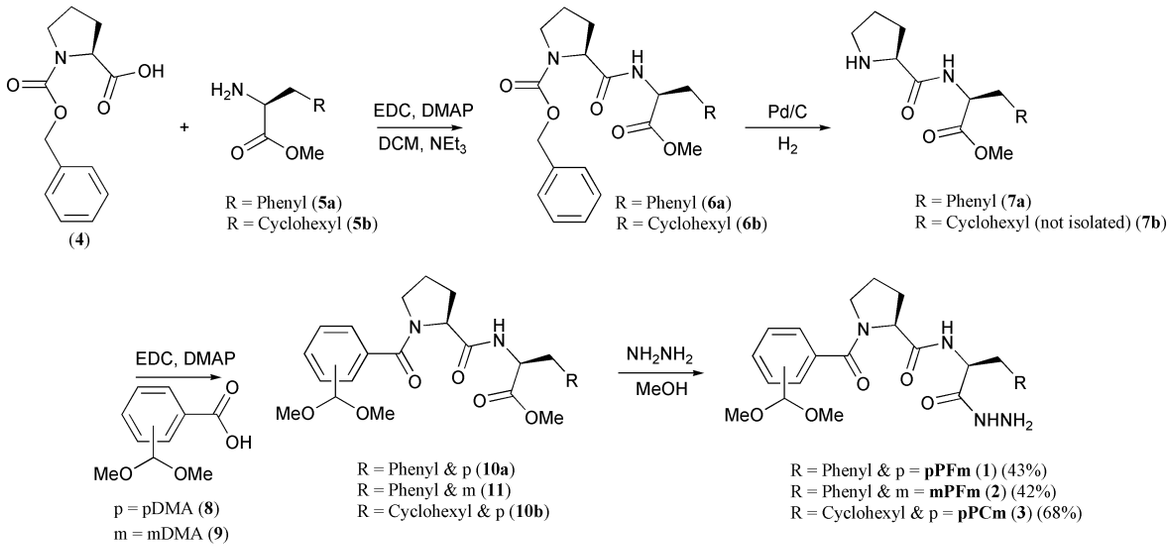
Previously we described some properties of a library of macrocycles prepared from the building block **pPFm**[11] (**1**), containing L-proline (**P**), L-phenylalanine (**F**), and a *p*-substituted (**p**) aromatic linker (Scheme 1).[10d] This building block has the potential to engage in a range of non-covalent interactions including hydrogen-bonding, Lewis acid–base interactions,



Scheme 1 Dipeptide building blocks.

† Electronic supplementary information (ESI) available: NMR and IR data of the Li^+ complexes of **(mPF)$_2$**, **(pPF)$_3$** and **(pPC)$_3$**. See http://www.rsc.org/suppdata/ob/b3/b300956d/

DOI: 10.1039/b300956d



Scheme 2 Synthesis of the dipeptide building blocks **1**–**3**.

π–π and cation–π interactions. We have shown that LiI and NaI select and amplify the cyclic trimer.[10d]

In order to explore which features of the **pPFm** system are important in the formation of the metal ion·trimer complex, we have now studied the response of libraries generated from two structurally related building blocks. Firstly, we prepared the monomer **mPFm** (**2**) which has a different linker geometry, and secondly, the phenyl side chain was replaced by a cyclohexyl side chain, to give the monomer **pPCm** (**3**).

Results and discussion

Building block synthesis

The dipeptide building blocks were prepared using standard peptide chemistry (Scheme 2) starting with EDC coupling of CBZ-protected L-proline (**4**) to the appropriate amino acid methyl ester (**5a** or **5b**), followed by deprotection by hydrogenation and a second EDC coupling to the appropriate carboxybenzaldehyde dimethoxyacetal (**8** or **9**). The final bifunctional monomers **1**–**3** were then generated by the hydrazinolysis of the methyl esters introducing the hydrazide functionality.

Preparation and screening of the DCLs

Acid-catalysed cyclisation of the building blocks in $CHCl_3$:MeOH (98:2 v/v)[12] at room temperature generated the desired DCLs. The mixtures were analysed by electrospray ionisation mass spectrometry (ESI-MS) and found to consist primarily of macrocyclic hydrazone oligomers.[13] A combination of ESI-MS and HPLC allowed the assignment of the major peaks in the chromatograms as shown in Fig. 2(a), 3(a) and 4(a). The HPLC traces were recorded once thermodynamic equilibrium had been reached after 1–5 days. After this point no further changes in the distributions of cyclic products were observed, although the reactions were monitored for 7 or 14 days. The equilibrium composition of the **mPFm** library is dominated by the cyclic dimer (Fig. 2a), whereas the **pPFm** (Fig. 3a) and **pPCm** (Fig. 4a) libraries contain more of the higher oligomers.

Addition of LiI and NaI as templates affects the product distribution in all three libraries (traces b and c in Fig. 2–4). In the **pPFm** and **pPCm** libraries the salts induce the amplification of cyclic trimer. In the case of the **pPCm** library NaI gives rise to a slightly greater response than LiI leading to approximately 95% of the peptide library material being present as trimer. However, in the **pPFm** library both NaI and LiI give rise to very similar templating responses, both leading to the trimer representing approximately 98% of the peptide material.

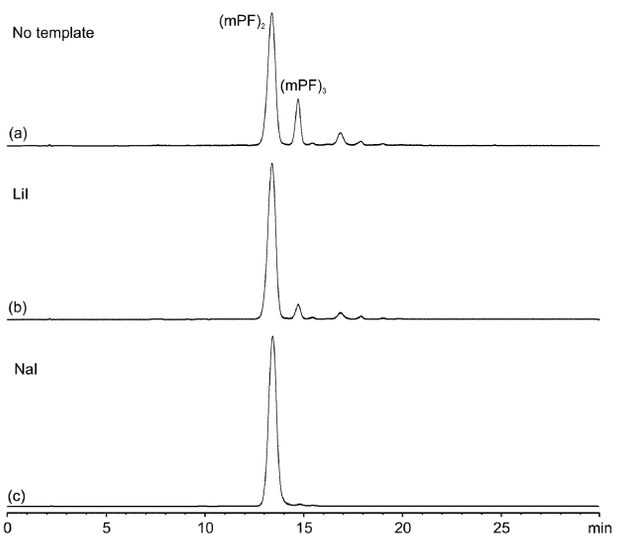


Fig. 2 HPLC traces of the DCL made from 5.0 mM **mPFm** in $CHCl_3$:MeOH (98:2 v/v) after stirring for 24 h at rt (a) in the absence of template, (b) in the presence of 15 mM LiI, (c) in the presence of 15 mM NaI.

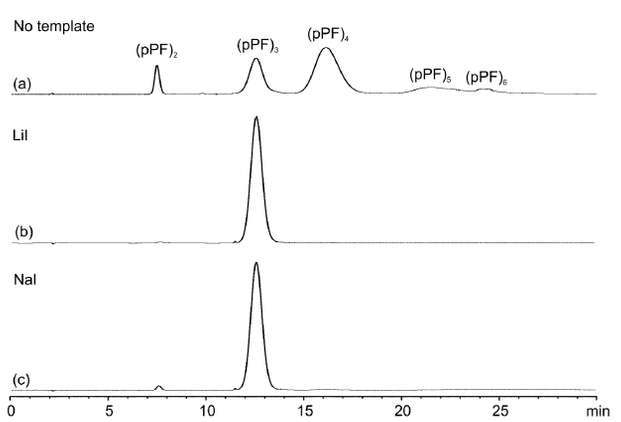


Fig. 3 HPLC traces of the DCL made from 0.3 mM $\mathbf{pPF_3}$ in $CHCl_3$:MeOH (98:2 v/v) after stirring for 5 days at rt (a) in the absence of template, (b) in the presence of 2.7 mM LiI (c) in the presence of 2.7 mM NaI.

The HPLC results shown here for the **pPFm** system with NaI differ slightly from those reported previously.[10d] Those initial HPLC traces turned out not to represent the true equilibrium position in the presence of NaI, due to precipitation over time.

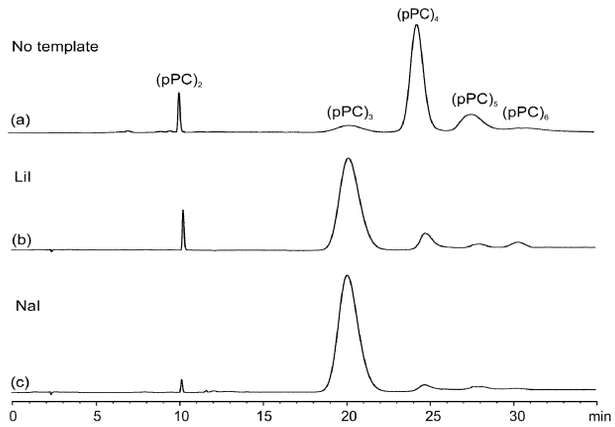


Fig. 4 HPLC traces of the DCL made from 1.0 mM **pPCm** in $CHCl_3$:MeOH (98:2 v/v) after stirring for 5 days at rt (a) in the absence of template, (b) in the presence of 3.0 mM LiI, (c) in the presence of 3.0 mM NaI.

The HPLC traces shown here are for libraries generated at lower concentration and starting from pure trimer not monomer, conditions which avoid problems with precipitation.[14] In the presence of the NaI the 'library' distribution shown in Fig. 3c remains unchanged for at least 14 days, whereas in the absence of NaI the distribution reverts to its untemplated equilibrium position in the presence of acid in the course of 5 days (Fig. 3a).

Introduction of LiI and NaI into the **mPFm** library induced a different response. Here it is the cyclic dimer, already dominant in the absence of template, which is amplified by both the LiI and NaI. Sodium iodide induces the greatest response, leading to the dimer representing >98% of the peptide material in the library.

To prove that the libraries were all operating under thermodynamic control, the same templated library distributions were generated *via* two independent routes: firstly by cyclisation of the monomers in the presence of the templates and secondly, by cyclisation of the monomers in the absence of template, followed by addition of the template to the pre-equilibrated mixtures. In all cases essentially identical product distributions were obtained, except for **pPFm** in the presence of NaI where solubility problems were encountered (see above).[14]

No significant differences in templating effects were observed using templates with different anions, *i.e.* LiCl, NaCl, Li trifluoroacetate and Na trifluoroacetate, suggesting that iodide anion is not involved in the recognition process.[15]

Binding studies

The selected receptors were isolated using semi-preparative HPLC and characterised by HPLC, NMR and MS analysis. The interactions between the amplified cyclic compounds and the metal ions were investigated using ^{1}H NMR, ^{7}Li NMR, ^{13}C NMR, FT-IR and isothermal titration calorimetry (ITC). The limited solubility of NaI in our solvent system restricted NMR binding studies to the LiI template.

The ^{1}H NMR spectrum of **(mPF)$_2$** in $CDCl_3$:MeOD (98:2 v/v) shows that the compound adopts on average a C_2 symmetric conformation. The spectrum was fully assigned with the aid of a COSY spectrum. Addition of LiI led to shifts in all ^{1}H NMR receptor signals (Fig. 5), the system being in fast exchange on the ^{1}H NMR chemical shift timescale. The Li·**(mPF)$_2$** complex also showed averaged C_2 symmetry. A binding constant of 5×10^2 M^{-1} was determined for this system by titrating LiI into a solution of **(mPF)$_2$** and assuming a 1:1 stoichiometry.[16]

Addition of LiI to a solution of **(pPC)$_3$** showed that this system is in slow exchange and forms a 1:1 complex. Both the **(pPC)$_3$** and Li·**(pPC)$_3$** complex were shown to have averaged C_3 symmetric conformations although the two species have very different ^{1}H NMR spectra (Fig. 6). Similar behaviour was observed upon titrating **(pPF)$_3$** with LiI. All the NMR spectra were assigned using COSY, and from the integrals of the bound and unbound ^{1}H NMR signals of a 1:1 host/guest mixture binding constants of 1.7×10^5 M^{-1} and 4×10^4 M^{-1} for **(pPC)$_3$** and **(pPF)$_3$**, respectively, were obtained.

For all three receptors addition of LiI leads to significant shifts in every resonance in the ^{1}H NMR spectra. This suggests that there is a substantial change in the geometry of the

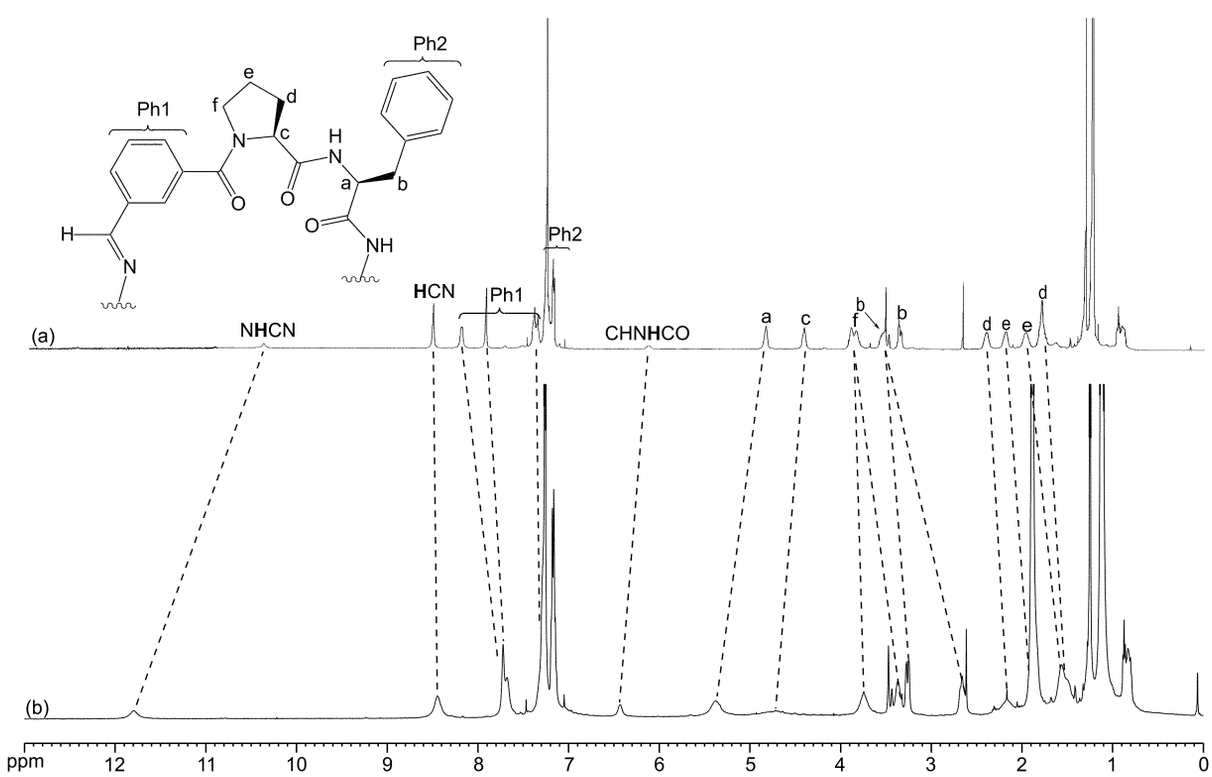


Fig. 5 ^{1}H-NMR in $CDCl_3$:CD_3OD (98:2 v/v) of **(mPF)$_2$** (a) in the absence of LiI and (b) in the presence of 1 equivalent LiI.

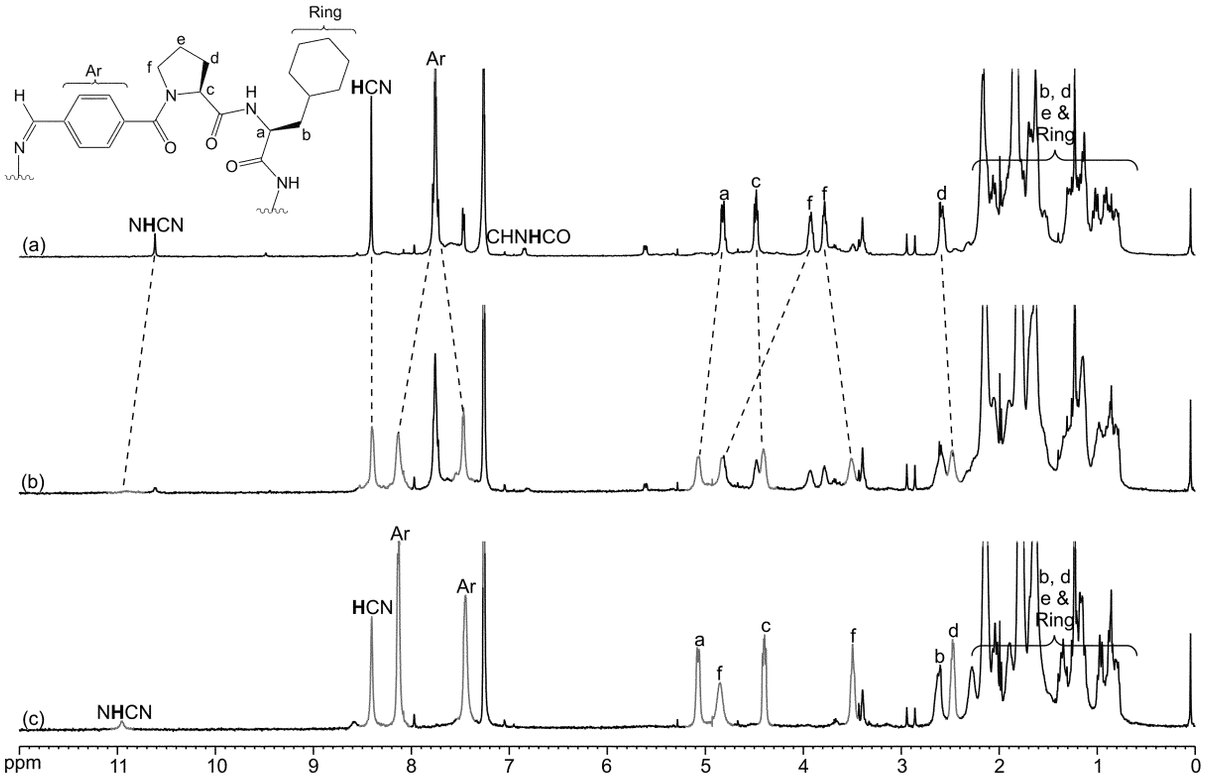


Fig. 6 ^{1}H-NMR in $CDCl_3$:CD_3OD (98:2 v/v) of **(pPC)$_3$** (a) in the absence of LiI; (b) in the presence of 0.5 equivalent LiI and (c) in the presence of 1 equivalent LiI.

receptors upon complexation with lithium, and therefore that they are *not* preorganised.

Further information about the complexation of the receptors with lithium was obtained from ^{7}Li NMR studies. Titrating **(pPC)$_3$** into a solution of LiI showed a system in fast exchange on the ^{7}Li NMR timescale, giving rise to an averaged signal for bound and free lithium (Fig. 7). However, a similar experiment with **(pPF)$_3$** showed a system in slow exchange on the ^{7}Li NMR timescale, giving rise to separate signals for the bound and free lithium (Fig. 8). This difference in behaviour seems counter-intuitive, given that the binding constant for **(pPC)$_3$** is higher than that for **(pPF)$_3$**. It suggests that there is a difference in the kinetic stabilities of the two complexes, which does not correspond to their thermodynamic stabilities.[17] This may be due to dissimilarities in the rate constants for binding (k_b) and release (k_r) of the two complexes. Since both the binding constant (K) and the rate constant of release of Li (k_r) are greater for the **(pPC)$_3$** system, then as $K = k_b/k_r$ the rate constant of binding (k_b) to **(pPC)$_3$** must also be much higher than that for **(pPF)$_3$**, suggesting that the enhanced affinity of **(pPC)$_3$** arises from the greater rate of binding of lithium, possibly due to a less demanding structural rearrangement for complex formation in this system (see below).

There are many precedents for coordination of lithium to the carbonyl oxygens of peptides.[18] Study of changes in the ^{13}C

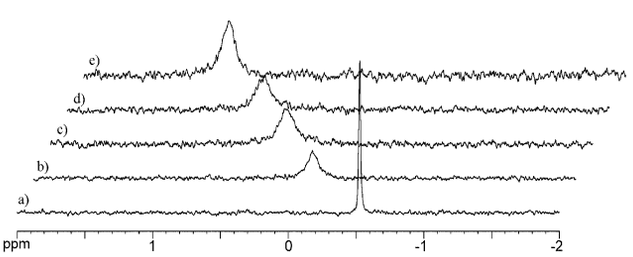


Fig. 7 ^{7}Li NMR in $CDCl_3$:CD_3OD (98:2 v/v) of 2 mM LiI (a) in the absence of **(pPC)$_3$** and in the presence of (b) 0.4; (c) 0.6; (d) 0.8; (e) 1.0 equivalents of **(pPC)$_3$**.

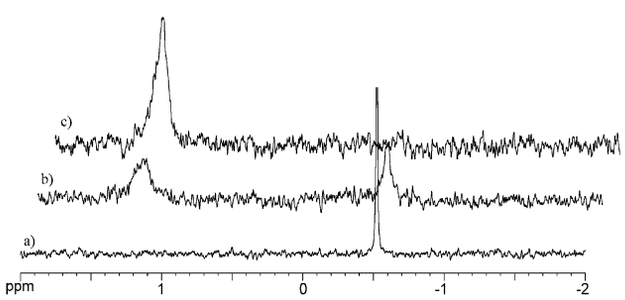


Fig. 8 ^{7}Li NMR in $CDCl_3$:CD_3OD (98:2 v/v) of 2 mM LiI (a) in the absence of **(pPF)$_3$** and in the presence of (b) 0.5 and (c) 1.0 equivalent of **(pPF)$_3$**.

NMR and FT-IR spectra of the three receptors upon complexation with LiI suggests similar interactions are occurring in these systems. The ^{13}C NMR spectra of the three cyclic species show downfield shifts of the carbonyl carbons in the presence of LiI. In the **(mPF)$_2$** spectrum, all three carbonyls shift, by 1.0, 2.0 and 2.0 ppm. In the cases of **(pPF)$_3$** and **(pPC)$_3$** only the carbonyl carbons of the proline and phenylalanine units shift: by 1.0 ppm and 4.5 ppm for **(pPF)$_3$** and by 2.0 ppm and 2.0 ppm for **(pPC)$_3$**.

In addition, in the presence of LiI and NaI the IR carbonyl stretching frequencies of the receptors shift to lower wavenumbers (Table 1) also suggesting that the metal ions are coordinating the carbonyl oxygens. For the **(pPF)$_3$** and **(pPC)$_3$**[19] systems the symmetry of the C=O peaks in the IR spectra is reduced, which supports the ^{13}C NMR evidence that in these two cases not all of the carbonyl groups are involved in metal-ion binding. No such effect is observed for the **(mPF)$_2$** system. In the **(mPF)$_2$** system NaI induces a much larger shift in the IR absorbances than does LiI, which correlates with the greater templating efficiency observed for sodium over lithium in this system.

Binding constants for the three receptors with the metal-ion templates were measured using ITC.[20] This technique involves

Table 1 IR carbonyl stretching frequencies (cm^{-1}) for the dipeptide receptors in the absence and presence of the metal ion templates

	No salt	LiI[a]	NaI[a]
(mPF)$_2$	1681	1675 (−5)	1670 (−11)
(pPF)$_3$	1673	1668 (−5)	1669 (−4)
(pPC)$_3$	1674	1667 (−7)	1668 (−6)

[a] Values in brackets indicate the change in wavenumber from the values in the absence of template.

the measurement of heat uptake or liberation upon the titration of one species into another, which, after heats of dilution are subtracted, gives the enthalpy change upon binding ($\Delta H°$) from the peak areas; the free energy of binding ($\Delta G°$) from the shape of the titration curve and the stoichiometry from the inflexion point. The entropy of binding ($\Delta S°$) can then be calculated from the Gibbs–Helmholtz equation.

All the titrations were carried out with the cyclic receptors being titrated into the metal-ion solution in $CHCl_3$:MeOH 98:2 (v/v). The reverse experiment led to large heat effects probably due to de-aggregation of the salts upon dilution. A representative ITC titration is shown in Fig. 9 and the results are summarised in Table 2 together with the binding constants obtained by ^{1}H NMR. The stoichiometries obtained from the curve-fitting procedure were found to be in agreement with 1:1 complexes being formed in all cases confirming our observations using NMR. Also the binding constants measured by NMR and microcalorimetry are in good agreement. For all receptors the binding constants for Na^+ are higher than those for Li^+, a result that correlates well with the templating experiments.

The two orders of magnitude selectivity of the relatively small and rigid **(mPF)$_2$** for Na^+ over Li^+ is much higher than that of the more flexible trimers. This difference is also reflected in the IR results for the cyclic compounds in the presence and absence of the metal ions. Significant selectivity for Na^+ over Li^+ has been reported for other mainly crown ether based receptors,[21] and is generally thought to arise from size differentiation.

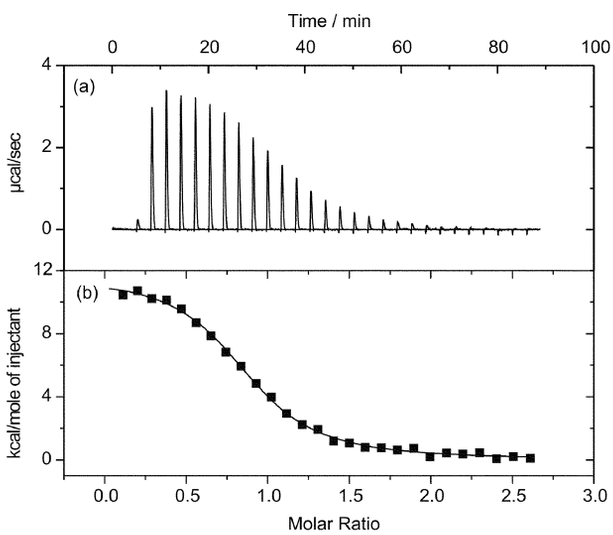


Fig. 9 Heat effects (a) and enthalpy changes (b) upon titrating 0.7 mM **(pPC)$_3$** into 0.075 mM LiI in $CHCl_3$:MeOH (98:2 v/v), measured at 298 K.

Binding of the metal ions to the receptors is invariably entropy driven, and accompanied by an unfavourable enthalpy change. Favourable entropy contributions are usually associated with the liberation of solvent molecules from around both host and guest, compensating for the cost of the immobilisation of the substrates. Enthalpic effects are balanced between solvation of the substrates, interactions within the bulk solvent, and interactions between the host and guest.[22] Comparison of the individual enthalpy and entropy contributions for the formation of the complexes of the **(pPF)$_3$** and **(pPC)$_3$** receptors highlights a striking difference in their otherwise very similar behaviour. In the case of the **(pPC)$_3$** receptor the enthalpy contributions are much more unfavourable for complexation with both Li^+ and Na^+ and the entropy gains are much greater compared to the **(pPF)$_3$** receptor. In addition, it is the Li^+ and not the Na^+ which gives rise to the greatest unfavourable enthalpic effect and the greatest entropy gain. This behaviour suggests that the **(pPC)$_3$** receptor has a greater level of preorganisation for the Li^+ binding compared to the **(pPF)$_3$** receptor. This conclusion is further supported by differences in the kinetics of complex formation as observed by ^{7}Li NMR which were much faster for Li^+ binding to **(pPC)$_3$** than to **(pPF)$_3$**.

Conclusions and summary

The results presented here have shed some light on the interactions leading to efficient templating. Replacing the phenyl group of **pPFm** by a cyclohexyl group to give **pPCm** leads to no significant change in templating behaviour: the trimer species is still templated efficiently and indeed has slightly higher affinities for Li^+ and Na^+. Despite similar affinities for the templates the thermodynamics and kinetics of binding of Li^+ to **(pPC)$_3$** and **(pPF)$_3$** are surprisingly different suggesting that **(pPC)$_3$** is the better preorganised receptor for this metal ion.

Changing the size and shape of the monomer building block does have a significant effect on the size of receptor chosen by the metal ion templates. In the case of **mPFm**, a dimer rather than a trimer is selected by Li^+ and Na^+. The ^{13}C NMR and IR data suggest that in the **(mPF)$_2$** receptor all six of the C=O groups are involved in the metal ion binding, whilst in both trimers only two carbonyls from each monomer unit are involved. Li^+ and Na^+ prefer six hard ligands in their first coordination sphere. For the **mPFm** library the smallest macrocycle that can satisfy this arrangement is the dimer. Inspection of CPK models of **(pPF)$_2$** and **(pPC)$_2$** show that an octahedral arrangement of the six carbonyls is not possible for these macrocycles and only by moving to the trimers can these two systems bind and fully satisfy the metal ion templates. In summary we have identified three receptors with high degrees of flexibility that were found to undergo substantial

Table 2 Equilibrium constants and thermodynamic data for the binding of LiI and NaI to the peptide receptors measured in $CHCl_3$:MeOH (98:2 v/v) at 298 K.

Receptor	Guest	K_{NMR}/M^{-1}	K_{ITC}/M^{-1}	$\Delta H°$/kJ mol^{-1}	$T\Delta S°$/kJ mol^{-1}	$\Delta G°$/kJ mol^{-1}
(mPF)$_2$	Li [a]	5.0×10^2	—	—	—	—
	Na [b]	—	6.2×10^4	6.1	33.4	−27.3
(pPF)$_3$	Li	4.0×10^4	1.0×10^5	4.7	34.0	−29.3
	Na [b]	—	1.9×10^5	16.3	46.4	−30.1
(pPC)$_3$	Li	1.7×10^5	2.2×10^5	50.6	81.2	−30.5
	Na [c]	—	6.0×10^5	22.4	55.2	−32.8

[a] Assuming a 1:1 stoichiometry. [b] 3.0 mM receptor titrated into approximately 0.3 mM NaI or LiI. [c] 0.7 mM trimer titrated into approximately 0.075 mM NaI/LiI.

conformational rearrangements upon binding of the metal-ion guests. Receptors with these characteristics are difficult to access through rational design, and form an underexplored but exciting area in supramolecular chemistry. Our results demonstrate that dynamic combinatorial chemistry is a very powerful tool for the discovery of these challenging 'induced-fit' receptors.

Experimental

General

HPLC analyses were carried out on a Hewlett Packard 1050 or 1100 system coupled to a UV analyser and processed using HP Chemstation software. Separations were achieved using a reversed phase Supelcosil ABZ+Plus column 15.0 cm × 4.6 mm, 3 μm particle size (analytical scale); 25 cm × 1 cm, 5 μm particle size, (semi-preparative scale). All separations were performed with gradient elution of water and acetonitrile mixtures, typically 75% to 10% water over 25 min, with a flow rate of 1.0 ml min^{-1} for analytical scale and 5.0 ml min^{-1} for semi-preparative scale.

NMR spectroscopy was performed on Bruker DRX 400, DPX 500 or DRX 800 instruments and chemical shifts are quoted in parts per million with respect to TMS.

Electrospray mass spectra were recorded on a Micromass Quattro-LC triple quadrupole mass spectrometer (QUATTRO) fitted with a z-spray electrospray source. The source was heated to 100 °C and the sampling cone voltage (V_c) was kept between 30–65 V. Samples were introduced into the mass spectrometer source with an LC pump (Shimadzu LC-9A pump) at a rate of 4 μl min^{-1} of MeCN:H_2O 1:1. Calibration was performed using protonated horse myoglobin. Scanning was carried out from *m*/*z* = 200 to 3000 in 8 s and several scans were summed to obtain the final spectrum which was processed using MassLynx V3.0 software. Electrospray mass spectra were also recorded on a Micromass Q-TOF instrument (QTOF), incorporating time-of-flight analysis with electrospray ionisation through a standard z-spray source. Calibration was performed using erythromycin as the standard. The MSMS spectra were collected on a Micromass Quattro-LC triple quadrupole apparatus (QUATTRO) by selecting the *m*/*z* peak for the desired parent ion isotopomer at the cone voltage where ion current for that peak was maximum. The collision cell voltage was manually increased to induce fragmentation and the daughter ion spectra measured by the second quadrupole analyser. A scan time of 8 s per spectrum and a low resolution setting were used and several scans were summed to obtain the final spectrum.

The isothermal calorimetry measurements (ITC) were conducted by using a MCS-ITC calorimeter from MicroCal, LLC, Northampton, MA, USA. Aliquots of 10 μl were titrated into the calorimetric cell every 3 min over a one and a half hour period. The data were analysed using the customised ITC module of the Origin 5.0 software package and a least squares fitting procedure to fit the data to the appropriate binding model. All measurements were carried out at 25 °C. For each system studied a blank run was carried out in which the titrant was titrated into the cell containing solvent only, to allow corrections for the heat effects due to dilution to be made.

Solution and solid state infrared spectra were recorded on a Perkin Elmer Paragon 1000 FT-IR spectrometer at 4 cm^{-1} resolution or better.

Materials

All chemicals were purchased from Aldrich, Fluka, Novabiochem or Senn Chemicals as reagent grade and used without further purification. Thin layer chromatography was carried out on glass sheets precoated with silica gel 60 F_{254} (Merck) which were initially inspected by UV light before developing with I_2 in $CHCl_3$ or Ninhydrin in EtOH. Column chromatography was carried out using silica gel 60 F (Fluorochem). All solvents were distilled prior to use and dry solvents freshly distilled from CaH_2 under argon, with the exception of DMSO (Lancaster) which was used without further purification. HPLC grade $CHCl_3$ (Aldrich), and HPLC grade acetonitrile (Fisher) were filtered with a 0.45 μm Millipore filter and used without further purification. Ultrapure water was obtained from a Millipore water purification system.

4-Carboxybenzaldehyde dimethoxyacetal (8).[23,24] To a suspension of 4-carboxybenzaldehyde (2.00 g, 1.33 mmol) in dry MeOH (40 ml) was added ammonium chloride (4.00 g, 74.8 mmol) and the reaction was heated under reflux for 20 h. The solvent was removed under reduced pressure and the resulting white solid was recrystallised from boiling hexane (1.55 g, 60%). IR ($CHCl_3$) ν = 1694 (C=O) cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$) δ = 8.11 (2H, d, J = 7.9 Hz, Ar–H), 7.57 (2H, d, J = 7.9 Hz, Ar–H), 5.46 (1H, s, C**H**$(OMe)_2$), 3.34 (6H, s, CH(O**Me**$)_2$); ^{13}C NMR (100 MHz, $CDCl_3$) δ = 171.7 (C=O), 143.8 (Ar*quat*), 130.2 (Ar–H), 129.3 (Ar*quat*), 126.9 (Ar–H), 102.2 (**C**H$(OMe)_2$), 52.7 (CH(O**Me**$)_2$); HRMS (EI) $[M]^+$ $C_{10}H_{12}O_4$ requires 196.0736, found 196.0740; m.pt. = 119–120 °C.

3-Carboxybenzaldehyde dimethoxyacetal (9).[23,24] Synthesised following the procedure for the synthesis of **8**, starting from 3-carboxybenzaldehyde (2.00 g, 1.33 mmol). Yield: (1.81 g, 70%). IR ($CHCl_3$) ν = 1697 (C=O) cm^{-1}; ^{1}H NMR.(400 MHz, $CDCl_3$) δ = 8.10 (1H, s, Ar–H), 8.07 (1H, d, J = 7.8 Hz, Ar–H), 7.72 (1H, d, J = 7.8 Hz, Ar–H), 7.50 (1H, t, J = 7.8 Hz, Ar–H), 5.46 (1H, s, C**H**$(OMe)_2$), 3.35 (6H, s, CH(O**Me**$)_2$); ^{13}C NMR (100 MHz, $CDCl_3$) δ = 171.9 (C=O), 138.8, 132.1 (Ar*quat*), 130.3, 129.4, 128,7, 128.5 (Ar–H), 102.4 (**C**H$(OMe)_2$), 52.7 (CH(O**Me**$)_2$); HRMS (EI) $[M]^+$ $C_{10}H_{12}O_4$ requires 196.0736, found 196.0741; m.pt. = 68–69 °C.

General procedure for amide coupling. Amide coupling reactions were performed using EDC and DMAP. The acid and amine components were dissolved in dry DCM under argon. In examples where the starting amine was purchased as the hydrochloride salt, dry Et_3N (1.5 equivalents) was added. The resulting solution was cooled to 0 °C on an ice bath for 30 min after which time the EDC (1.2 equivalents) and DMAP (20% by weight with respect to EDC) were added. The reaction was kept at 0 °C for 1 h and allowed to warm to room temperature and stirred under argon overnight. The work up involved the addition of DCM (2-fold dilution) and subsequent washing of the organic solution with three portions of H_2O. The organic layer was dried over $MgSO_4$, filtered, and the solvent removed *in vacuo* to give the crude product which was purified by silica gel column chromatography.

General procedure for hydrogenation. Deprotection of the CBZ protected intermediates was carried out using hydrogenation. The CBZ protected compounds were dissolved in a EtOAc:MeOH (4:1) mixture. To the solution was added 5% Pd/C and the flask was evacuated and flushed with an atmosphere of hydrogen three times before allowing the reaction to stir under a hydrogen atmosphere for 6 h. The suspensions were then filtered through a pad of celite and the solvent removed *in vacuo* to afford the products which were used without further purification.

General procedure for hydrazinolysis. The methyl esters were dissolved in MeOH and to the resulting solutions were added 10 equivalents of hydrazine monohydrate. The reactions were allowed to stir at room temperature overnight or until the reactions had gone to completion, as ascertained by TLC. The solvent was removed *in vacuo* to afford the crude products as oils and were purified by silica gel column chromatography.

(*S*,*S*)-*N*-(Carbobenzyloxy-proline)phenylalanine methyl ester (6a). Carbobenzyloxy-L-proline (1.0 g, 4.01 mmol) and L-phenyalanine methyl ester hydrochloride (0.87 g, 4.01 mmol) were coupled in DCM (40 ml) using the standard procedure described above. Silica gel column chromatography [EtOAc:Hex 6:4] yielded **6a** as a white solid (1.38 g, 84%). R_f = 0.43 [silica gel, EtOAc:Hex 6:4] (UV); IR ($CHCl_3$) ν = 1740, 1736 (ester C=O), 1624 (amide C=O) cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ = 7.36–7.21 (8H, br m, Ar–H), 7.06–7.03 (2H, m, Ar–H), 6.33 (1H, br s, NH), 5.17–5.08 (2H, m, C**H**$_2$Ph), 4.83 (1H, t, J = 9.7 Hz, α-H), 4.30 (1H, t, J = 8.5 Hz, Pro–α–H), 3.71 (3H, s, O**Me**), 3.35 (2H, br m, Pro–NCH_2), 3.15 (1H, br m, C$H_a$$H_b$Ph), 2.99 (1H, br m, C$H_a$$H_b$Ph), 2.04 (1H, m, Pro–C$H_a$$H_b$), 1.81 (2H, br m, Pro–C$H_2$), 1.61 (1H, br m, Pro–C$H_a$$H_b$); ^{13}C NMR (100 MHz, $CDCl_3$) δ = 171.8, 171.2 (ester, C=O), 156.1 (amide C=O), 136.4, 136.0 (Ar*quat*), 129.2, 128.5, 128.4, 128.1, 127.9, 126.9 (Ar–H), 67.4 (**C**H_2Ph), 60.4 (Pro–**C**α–H), 52.8 (**C**α–H), 50.8 (O**Me**), 46.8 (Pro–N**C**H_2), 37.9 (**C**H_2Ph), 27.9, 23.4 (Pro–**C**H_2); HRMS (QTOF) $[M + Na]^+$ $C_{23}H_{26}N_2O_5Na$ requires 433.1739, found 433.1750; m.pt. = 67–71 °C.

(*S*,*S*)-Proline-phenylalanine methyl ester (7a). *N*-(Carbobenzyloxy-proline)phenylalanine methyl ester (**6a**) (2.63 g, 6.41 mmol) was dissolved in a EtOAc:MeOH (4:1) mixture (100 ml) and hydrogenated according the procedure given above. Yield: (1.43 g, 85%). R_f = 0.32 [silica gel, DCM:MeOH 95:5] (UV, Ninhydrin); IR ($CHCl_3$) ν = 1731 (ester C=O), 1624 (amide C=O) cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ = 7.28–7.20 (3H, br m, Ar–H), 7.10 (2H, m, Ar–H), 6.34 (1H, d, J = 8.1 Hz, NH), 4.85 (1H, dd, J_1 = 10.5 Hz, J_2 = 6.8 Hz, α-H), 3.71 (3H, s, O**Me**), 3.68 (1H, dd, J_1 = 8.7 Hz, J_2 = 6.7 Hz, Pro–α–H), 3.18 (1H, m, C$H_a$$H_b$Ph), 3.03 (1H, m, C$H_a$$H_b$Ph), 2.92 (1H, m, Pro–NC$H_a$$H_b$), 2.74 (1H, m, Pro–NC$H_a$$H_b$), 2.03 (1H, m, Pro–C$H_a$$H_b$), 1.76 (1H, m, Pro–C$H_a$$H_b$), 1.63–1.47 (2H, br m, Pro–CH_2), 1.35 (1H, br, NH); ^{13}C NMR (100 MHz, $CDCl_3$) δ = 173.5 (ester C=O), 160.7 (amide C=O), 134.7 (Ar*quat*), 127.8, 127.0, 125.5 (Ar–H), 58.8 (Pro–**C**α–H), 50.9 (**C**α–H), 50.8 (O**Me**), 45.7 (Pro–N**C**H_2), 36.7 (**C**H_2Ph), 29.2, 24.6 (CH_2); HRMS (QUATTRO) $[M + H]^+$ $C_{15}H_{21}N_2O_3$ requires 277.1153, found 277.1140.

(*S*,*S*)-*N*-(4-Dimethoxymethyl-benzoyl)proline-phenylalanine methyl ester (10a). Proline-phenylalanine methyl ester (**7a**) (1.08 g, 3.91 mmol) and 4-carboxybenzaldehyde dimethoxyacetal (**8**) (0.77 g, 3.91 mmol) were reacted together in DCM (40 ml) following the amide coupling procedure given above. Silica gel column chromatography [EtOAc:Hex 8:2] yielded **10a** as a white solid (1.23 g, 69%). R_f = 0.43 [silica gel, EtOAc:Hex 8:2] (UV); IR ($CHCl_3$) ν = 1730 (ester C=O), 1674, 1625 (amide C=O) cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ = 7.48 (2H, d, J = 8.1 Hz, Ar–H), 7.41 (2H, d, J = 8.1 Hz, Ar–H), 7.30 (1H, d, J = 7.4 Hz, NH), 7.18–7.13 (5H, br m, Ar–H), 5.41 (1H, s, C**H**$(OMe)_2$), 4.84 (1H, t, J = 9.8 Hz, α-H), 4.75 (1H, t, J =8.3 Hz, Pro–α–H), 3.70 (3H, s, O**Me**), 3.41–3.37 (2H, br m, Pro–NCH_2), 3.32 (6H, s, CH(O**Me**)$_2$), 3.20 (1H, d, J = 14.0 Hz, C$H_a$$H_b$Ph), 3.03 (1H, d, J = 14.0 Hz, C$H_a$$H_b$Ph), 2.35 (1H, m, Pro–C$H_a$$H_b$), 2.02–1.90 (2H, m, Pro–CH_2), 1.78 (1H, m, Pro–C$H_a$$H_b$); ^{13}C NMR (125 MHz, $CDCl_3$) δ = 172.3 (ester C=O), 169.9, 168.7 (amide C=O), 139.4, 134.5, 133.6 (Ar*quat*), 127.5, 127.3, 125.7, 125.5, 125.3 (Ar–H), 101.2 (**C**H$(OMe)_2$), 60.9 (Pro–**C**α–H), 53.4 (CH(O**Me**)$_2$), 53.3 (**C**α–H), 50.7 (Pro–N**C**H_2), 37.5 (**C**H_2Ph), 28.4, 25.7 (CH_2); HRMS (QUATTRO) $[M + H]^+$ $C_{25}H_{31}N_2O_6$ requires 455.2183, found 455.2190; m.pt. = 127–129 °C.

pPFm: (*S*,*S*)-*N*-(4-Dimethoxymethyl-benzoyl)proline-phenylalanine carboxylic acid hydrazide (1). *N*-(4-Dimethoxymethyl-benzoyl)proline-phenylalanine methyl ester (**10a**) (1.13 g, 2.49 mmol) underwent hydrazinolysis in MeOH (25 ml) *via* the procedure given above. Silica gel column chromography [DCM:MeOH/95:5] afforded **pPFm** (**1**) as a white solid (1.00 g, 88%). R_f = 0.38 [silica gel, DCM:MeOH 96:4] (UV, Ninhydrin); IR ($CHCl_3$) ν = 1681 (hydrazide C=O), 1672, 1624 (amide C=O) cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ = 7.79 (1H, br, N**H**NH_2), 7.49 (2H, d, J = 8.2 Hz, Ar–H), 7.32 (2H, d, J = 8.2 Hz, Ar–H), 7.18–7.14 (5H, br m, Ar–H), 6.89 (1H, d, J = 8.2 Hz, NH), 5.42 (1H, s, C**H**$(OMe)_2$), 4.70 (1H, dd, J_1 = 10.1 Hz, J_2 = 6.7 Hz, α-H), 4.63 (1H, dd, J_1 = 8.5 Hz, J_2 = 6.1 Hz, Pro–α–H), 3.85 (2H, br s, NHN**H**$_2$), 3.46–3.38 (2H, br m, Pro–NCH_2), 3.34 (6H, s, CH(O**Me**)$_2$), 3.23–3.10 (2H, m, C**H**$_2$Ph), 2.22–2.11 (2H, m, Pro–CH_2), 1.92–1.76 (2H, m, Pro–CH_2); ^{13}C NMR (100 MHz, $CDCl_3$) δ = 170.1 (hydrazine C=O), 169.5, 169.2 (amide C=O), 139.1, 134.8, 133.8 (Ar*quat*), 127.6, 127.1, 125.6, 125.5, 125.2 (Ar–H), 100.8 (**C**H$(OMe)_2$), 61.1 (Pro–**C**α–H), 53.2 (CH(O**Me**)$_2$), 53.1 (**C**α–H), 50.9 (Pro–N**C**H_2), 37.5 (**C**H_2Ph), 28.5, 25.8 (CH_2); HRMS (QTOF) $[M + Na]^+$ $C_{24}H_{30}N_4O_5Na$ requires 477.2114, found 477.2105; m.pt. = 58–63 °C.

(*S*,*S*)-*N*-(3-Dimethoxymethyl-benzoyl)proline-phenylalanine methyl ester (11). Proline-phenylalanine methyl ester (**7a**) (1.25 g, 4.54 mmol) and 3-carboxybenzaldehyde dimethoxyacetal (**9**) (0.89 g, 4.54 mmol) were coupled in DCM (40 ml) following the amide coupling procedure described above. Silica gel column chromatography [EtOAc:Hex 8:2] yielded **11** as a colourless oil (1.42 g, 69%). R_f = 0.41 [silica gel, EtOAc:Hex 8:2] (UV); IR ($CHCl_3$) ν =1739 (ester C=O), 1678, 1624 (amide C=O) cm^{-1}; ^1H NMR (500 MHz, $CDCl_3$) δ = 7.57 (1H, s, Ar–H), 7.53 (1H, d, J = 6.7 Hz, Ar–H), 7.43–7.37 (2H, br m, Ar–H), 7.34 (1H, d, J = 7.9 Hz, NH), 7.22–7.19 (5H, br m, Ar–H), 5.41 (1H, s, C**H**$(OMe)_2$), 4.85 (1H, dd, J_1 = 10.5 Hz, J_2 = 7.1 Hz, α-H), 4.75 (1H, t, J = 8.4 Hz, Pro–α–H), 3.72 (3H, s, O**Me**), 3.43–3.38 (2H, m, Pro–NCH_2), 3.23 (6H, s, CH(O**Me**)$_2$), 3.20 (1H, m, C$H_a$$H_b$Ph), 3.03 (1H, m, C$H_a$$H_b$Ph), 2.37 (1H, m, Pro–C$H_a$$H_b$), 2.03–1.97 (2H, br m, Pro–CH_2), 1.78 (1H, m, Pro–C$H_a$$H_b$); ^{13}C NMR (125 MHz, $CDCl_3$) δ = 171.8 (ester C=O), 170.9, 169.7 (amide C=O), 138.4, 136.1, 136.0 (Ar*quat*), 129.3, 128.8, 128.6, 128.4, 127.3, 126.9, 125.6 (Ar–H), 102.4 (**C**H$(OMe)_2$), 59.8 (Pro–**C**α–H), 53.5 (**C**α–H), 52.7 (CH(O**Me**)$_2$), 52.3 (O**Me**), 50.2 (Pro–N**C**H_2), 38.0 (**C**H_2Ph), 27.3, 25.3 (CH_2); HRMS (QTOF) $[M + Na]^+$ $C_{25}H_{30}N_2O_6Na$ requires 477.2002, found 477.2012.

mPFm: (*S*,*S*)-*N*-(3-Dimethoxymethyl-benzoyl)proline-phenylalanine carboxylic acid hydrazide (2). *N*-(3-Dimethoxymethyl-benzoyl)proline-phenylalanine methyl ester (**11**) (1.42 g, 3.13 mmol) in MeOH (30 ml) was reacted under the standard hydrazinolysis conditions given above. Silica gel column chromography [DCM:MeOH 95:5] afforded **mPFm** (**2**) as a white solid (1.21 g, 85%). R_f = 0.31 [silica gel, DCM:MeOH 95:5] (UV, Ninhydrin); IR ($CHCl_3$) ν = 1681 (hydrazide C=O), 1672, 1624 (amide C=O) cm^{-1}; ^1H NMR (400 MHz, $CDCl_3$) δ = 7.68 (1H, br, N**H**NH_2), 7.55 (1H, s, Ar–H), 7.53 (1H, d, J = 7.8 Hz, Ar–H), 7.39 (1H, t, J = 7.8 Hz, Ar–H), 7.30 (1H, d, J = 7.8 Hz, Ar–H), 7.28–7.14 (5H, m, Ar–H), 6.88 (1H, d, J = 8.3 Hz, NH), 5.40 (1H, s, C**H**$(OMe)_2$), 4.68 (1H, m, α-H), 4.62 (1H, m, Pro–α–H), 3.83 (2H, br, NHN**H**$_2$), 3.40 (2H, m, Pro–NCH_2), 3.32 (6H, s, CH(O**Me**)$_2$), 3.15 (2H, m, C**H**$_2$Ph), 2.17 (1H, m, Pro–C$H_a$$H_b$), 2.10 (1H, m, Pro–C$H_c$$H_d$), 1.90–1.75 (2H, m, Pro–CH_2); ^{13}C NMR (125 MHz, $CDCl_3$) δ = 171.3 (hydrazine C=O), 171.0, 170.9 (amide C=O), 138.5, 136.4, 135.3 (Ar*quat*), 129.3, 129.0, 128.7, 128.4, 127.4, 127.1, 125.6 (Ar–H), 102.4 (**C**H$(OMe)_2$), 60.8 (Pro–**C**α–H), 53.1 (**C**α–H), 52.8 (CH(O**Me**)$_2$), 50.5 (Pro–N**C**H_2), 37.0 (**C**H_2Ph), 28.1, 25.3 (CH_2); HRMS (QTOF) $[M + Na]^+$ $C_{24}H_{30}N_4O_5Na$ requires 477.2114, found 477.2114; m.pt. = 61–65 °C.

(*S*,*S*)-*N*-(Carbobenzyloxy-proline)cyclohexylalanine methyl ester (6b). Carbobenzyloxy-L-proline (2.69 g, 10.8 mmol) and L-cyclohexylalanine methyl ester hydrochloride (2.00 g, 9.02 mmol) in DCM (100 ml) were reacted using the standard amide

coupling reaction described above. Silica gel column chromatography [EtOAc:Hex 6:4] yielded **6b** as a white solid (3.41 g, 91%). R_f = 0.43 [silica gel, EtOAc:Hex 6:4] (UV); IR ($CHCl_3$) ν = 1740, 1693 (ester C=O), 1618 (amide C=O) cm^{-1}; ^1H NMR (500 MHz, $CDCl_3$) δ = 7.39–7.25 (5H, m, Ar–H), 7.07 (1H, br, NH), 5.14 (2H, m, C**H**$_2$Ph), 4.55 (1H, dd, J_1 = 8.4 Hz, J_2 =6.3 Hz, Pro–α–H), 4.34 (1H, dd, J_1 = 10.5 Hz, J_2 = 8.3 Hz, α-H), 3.69 (3H, s, O**Me**), 3.57–3.38 (2H, m, Pro–NCH$_2$), 2.35 (1H, m, Pro–C**H**$_a$H$_b$), 2.14 (1H, m, Pro–CH$_a$**H**$_b$), 1.91 (2H, m, Pro–CH$_2$), 1.73–1.53 (8H, br m, ring and chain CH$_2$), 1.28 (1H, m, C**H**(CH$_2$)$_5$), 1.21–1.06 (2H, m, ring CH$_2$), 0.95–0.80 (2H, m, ring CH$_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ = 173.2, 171.2 (ester C=O), 156.2 (amide C=O), 136.4 (Ar*quat*), 128.5, 128.0, 127.8 (Ar–H), 67.3 (**C**H$_2$Ph), 60.2 (Pro–**C**α–H), 52.1 (O**Me**), 50.3 (**C**α–H), 46.8 (Pro–N**C**H$_2$), 39.7 (**C**H$_2$(C$_6$H$_{11}$)), 34.2 (**C**H(CH$_2$)$_5$), 33.3, 32.4 (ring CH$_2$), 27.8 (Pro–CH$_2$), 26.2 (ring CH$_2$), 24.6 (Pro–CH$_2$); HRMS (QTOF) [M + Na]$^+$ $C_{23}H_{32}N_2O_5Na$ requires 439.2209, found 439.2209; m.pt. = 83–84 °C.

(*S*,*S*)-*N*-(4-Dimethoxymethyl-benzoyl)proline-cyclohexylalanine methyl ester (10b). *N*-(Carbobenzyloxy-proline)cyclohexylalanine methyl ester (**6b**) (3.41 g, 8.19 mmol) was dissolved in a EtOAc:MeOH (4:1) mixture (100 ml) and hydrogenated as described above, yielding the free amine as a colourless oil (2.20 g, 95%) which was used without further purification. R_f = 0.28 [silica gel, EtOAc:Hex 6:4] (Ninhydrin). The prolinecyclohexylalanine methyl ester (2.20 g, 7.78 mmol) and 4-carboxybenzaldehyde dimethoxyacetal (**8**) (1.78 g, 9.34 mmol) were then coupled in DCM (100 ml) according to the procedure given above. Silica gel column chromatography [EtOAc:Hex 8:2] yielded **10b** as a colourless oil (2.97 g, 83%). R_f = 0.36 [silica gel, EtOAc:Hex 8:2] (UV); IR ($CHCl_3$) ν = 1740 (ester C=O), 1678, 1618 (amide C=O) cm^{-1}; ^1H NMR (500 MHz, $CDCl_3$) δ = 7.47 (4H, m, Ar–H), 7.31 (1H, d, J = 8.2 Hz, NH), 5.39 (1H, s, C**H**(OMe)$_2$), 4.81 (1H, dd, J_1 = 8.4 Hz, J_2 = 6.4 Hz, Pro–α–H), 4.57 (1H, dd, J_1 = 10.1 Hz, J_2 = 8.5 Hz, α-H), 3.71 (3H, s, O**Me**), 3.51 (1H, m, Pro–NC**H**$_a$H$_b$), 3.43 (1H, m, Pro–NCH$_a$**H**$_b$), 3.29 (6H, s, CH(O**Me**)$_2$), 2.48 (1H, m, Pro–C**H**$_a$H$_b$), 2.03 (2H, m, Pro–CH$_2$), 1.81 (1H, m, Pro–CH$_a$**H**$_b$), 1.72–1.52 (8H, br m, ring and chain CH$_2$), 1.29 (1H, m, C**H**(CH$_2$)$_5$), 1.11 (2H, m, ring CH$_2$), 0.87 (2H, m, ring CH$_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ = 173.2 (ester C=O), 170.8, 170.7 (amide C=O), 140.3, 136.2 (Ar*quat*), 129.9, 127.0 (Ar–H), 102.4 (**C**H(OMe)$_2$), 59.5 (Pro–**C**α–H), 52.6 (CH(O**Me**)$_2$), 52.1 (O**Me**), 50.4 (Pro–N**C**H$_2$), 50.3 (**C**α–H), 39.7 (**C**H$_2$(C$_6$H$_{11}$)), 34.2 (**C**H(CH$_2$)$_5$), 33.4, 32.3 (ring CH$_2$), 26.8 (Pro–CH$_2$), 26.1 (ring CH$_2$), 25.4 (Pro–CH$_2$); HRMS (QTOF) [M + Na]$^+$ $C_{25}H_{36}N_2O_6Na$ requires 483.2471, found 483.2471.

pPCm: (*S*,*S*)-*N*-(4-Dimethoxymethyl-benzoyl)proline-cyclohexylalanine carboxylic acid hydrazide (3). *N*-(4-Dimethoxymethyl-benzoyl)proline-cyclohexylalanine methyl ester (**10b**) (2.97 g, 6.46 mmol) was treated with hydrazine monohydrate in MeOH (30 ml) according to the standard hydrazinolysis procedure. Silica gel column chromography [DCM:MeOH 95:5] afforded **pPCm** (**3**) as a white solid (2.68 g, 90%). R_f = 0.38 [silica gel, DCM:MeOH 95:5] (UV, Ninhydrin); IR ($CHCl_3$) ν = 1681 (hydrazide C=O), 1672, 1618 (amide C=O) cm^{-1}; ^1H NMR (500 MHz, $CDCl_3$) δ = 7.61 (1H, br, N**H**NH$_2$), 7.52 (4H, m, Ar–H), 7.20 (1H, d, J = 8.1 Hz, NH), 5.42 (1H, s, C**H**(OMe)$_2$), 4.75 (1H, dd, J_1 = 8.5 Hz, J_2 = 6.1 Hz, Pro–α–H), 4.43 (1H, dd, J_1 = 10.3 Hz, J_2 = 8.5 Hz, α-H), 3.89 (2H, br, NHN**H**$_2$), 3.45 (2H, m, Pro–NCH$_2$), 3.32 (6H, s, CH(O**Me**)$_2$), 2.43 (1H, m, Pro–C**H**$_a$H$_b$), 2.15–2.01 (2H, m, Pro–CH$_2$), 1.89 (1H, m, Pro–C**H**$_c$H$_d$), 1.73 (1H, m, C**H**$_a$H$_b$(C$_6$H$_{11}$)), 168 (1H, m, CH$_a$**H**$_b$(C$_6$H$_{11}$)), 1.67–1.48 (6H, br m, ring CH$_2$), 1.36 (1H, m, C**H**(CH$_2$)$_5$), 1.11 (2H, m, ring CH$_2$), 0.96–0.81 (2H, m, ring CH$_2$); ^{13}C NMR (125 MHz, $CDCl_3$) δ = 171.3 (hydrazide C=O), 171.2, 170.9 (amide C=O), 140.7, 135,7 (Ar*quat*), 127.2, 126.9 (Ar–H), 102.4 (**C**H(OMe)$_2$), 60.0 (Pro–**C**α–H), 52.7 (CH(O**Me**)$_2$), 50.6 (Pro–N**C**H$_2$), 50.2 (**C**α–H), 38.6 (**C**H$_2$-(C$_6$H$_{11}$)), 34.2 (**C**H(CH$_2$)$_5$), 33.5, 32.3 (ring CH$_2$), 27.1 (Pro–CH$_2$), 26.2 (ring CH$_2$), 25.5 (Pro–CH$_2$); HRMS (QTOF) [M + Na]$^+$ $C_{24}H_{36}N_4O_5Na$ requires 483.2584, found 483.2583; m.pt. = 70–73 °C.

General procedures for the preparation of DCLs

All cyclisation reactions were performed at room temperature on a 1 ml scale. The monomer(s) were dissolved in $CHCl_3$:MeOH 98:2 (v/v) to give a solution of 20 mM. The stock solution was then diluted to give 1 ml of a 5 mM solution. TFA was added (43 equivalents in 20–100 µl of a stock solution in $CHCl_3$:MeOH 98:2 (v/v)) to the reaction to initiate deprotection and cyclisation. HPLC or mass spectrometry analysis was performed by injection of 10–50 µl of the reaction mixture. Unless otherwise stated cyclisation reactions were stirred at rt and followed by HPLC analysis every 24 h for up to 7 days.

General procedure for the large scale isolation of cyclic oligomers

Large scale cyclisations were carried out on a 25 ml scale with 5 mM monomer, 43 equivalents of TFA and 3 or 5 equivalents template with respect to monomer, in $CHCl_3$:MeOH 98:2 (v/v). The reactions were stirred at rt and monitored by HPLC until thermodynamic equilibrium was reached (24 h). The exchange reaction was quenched by the addition of 15 g of solid phase base Amberlyst-21, which was washed thorougly with water, MeOH and chloroform before use. After stirring for 15 min the base was filtered off and washed with $CHCl_3$ (100 ml) and MeOH (100 ml). The filtrates were combined and evaporated to dryness before redissolution in approximately 2 ml of DMSO. The crude product was purified by semi-preparative HPLC which yielded the products as white solids.

(mPF)$_2$:(*S*,*S*)-*N*-(3-Dimethoxymethyl-benzoyl)proline-phenylalanine carboxylic acid hydrazide dimer. Yield: (22 mg, 71%). IR ($CDCl_3$:MeOD 98:2) ν = 1681 (br, C=Os) cm^{-1}; ^1H NMR (500 MHz (cryoprobe), $CDCl_3$:MeOD 98:2) δ = 10.52 (2H, br, N**H**NC), 8.44 (2H, s, **H**CN), 8.11 (2H, br, Ar–H), 7.93 (2H, s, Ar–H), 7.47–7.35 (4H, m, Ar–H), 7.33–7.15 (10H, br m, Ar–H), 6.20 (2H, br, N**H**), 4.76 (2H, br m, α-H), 4.37 (2H, br m, Pro–α–H), 3.90–3.83 (4H, br m, Pro–NCH$_2$), 3.50 (2H, br m, C**H**$_a$H$_b$Ph), 3.30 (2H, br m, CH$_a$**H**$_b$Ph), 2.32 (2H, br m, Pro–C**H**$_a$H$_b$), 2.12 (2H, br m, Pro–C**H**$_c$H$_d$), 1.89 (2H, br m, Pro–CH$_c$**H**$_d$), 1.70 (2H, br m, Pro–CH$_a$**H**$_b$); ^{13}C NMR (125 MHz (cryoprobe), $CDCl_3$:MeOD 98:2) δ = 171.0, 169.4, 167.6 (C=O), 149.0 (CN), 137.0, 134.8, 133.5, 131.2, 130.2, 129.5, 129.3, 128.5, 126.9, 123.5 (Ar), 63.6 (Pro–**C**α–H), 51.3 (**C**α–H), 50.7 (Pro–N**C**H$_2$), 35.7 (**C**H$_2$Ph), 29.6, 26.0 (CH$_2$); HRMS (QTOF) [M + Na]$^+$ $C_{44}H_{44}N_8O_6Na$ requires 803.3281, found 803.3282; m.pt. 258 °C (decomposition).

(pPC)$_3$: (*S*,*S*)-*N*-(4-Dimethoxymethyl-benzoyl)proline-cyclohexylalanine carboxylic acid hydrazide trimer. Yield: (34 mg, 59%). IR ($CDCl_3$:MeOD 98:2) ν = 1674 (br, C=Os) cm^{-1}; ^1H NMR (500 MHz (cryoprobe), $CDCl_3$:MeOD 98:2) δ = 10.57 (3H, br, N**H**NC), 8.40 (3H, s, **H**CN), 7.75 (12H, m, Ar–H), 4.81 (3H, dd, J_1 = 11.0 Hz, J_2 = 4.0 Hz, α-H), 4.49 (3H, dd, J_1 = 10.4 Hz, J_2 = 7.3 Hz, Pro–α–H), 3.91 (3H, m, Pro–NC**H**$_a$H$_b$), 3.78 (3H, m, Pro–NCH$_a$**H**$_b$), 2.58 (3H, m, Pro–C**H**$_a$H$_b$), 2.10 (3H, m, Pro–C**H**$_c$H$_d$), 2.07 (3H, m, C**H**$_a$H$_b$(C$_6$H$_{11}$)), 1.91 (3H, m, Pro–CH$_c$**H**$_d$), 1.84 (3H, m, Pro–CH$_a$**H**$_b$), 1.60 (3H, m, CH$_a$**H**$_b$-(C$_6$H$_{11}$)), 1.18–0.78 (11H, br m, ring CH$_2$); ^{13}C NMR (125 MHz (cryoprobe), $CDCl_3$:MeOD 98:2) δ = 171.5, 171.1, 170.0 (C=O), 147.8 (CN), 137.2, 135.3, 128.4, 127.4 (Ar), 64.0 (Pro–**C**α–H), 51.2 (Pro–N**C**H$_2$), 49.9 (**C**α–H), 37.5, 34.7, 33.8, 31.9, 31.6, 29.6, 27.5 (ring and chain CH$_2$); MS (QUATTRO) [M + Na]$^+$ $C_{66}H_{84}N_{12}O_9Na$ requires 1212, found 1212; MSMS (QUATTRO) [trimer + H]$^+$ 1190 ($C_{66}H_{85}N_{12}O_9$), [dimer + H]$^+$ 793 ($C_{44}H_{57}N_8O_6$), [monomer− cyclohexylalanine unit]$^+$ 243 ($C_{13}H_{13}N_3O_2$); m.pt. 267 °C (decomposition).

(pPF)$_3$: (*S*,*S*)-*N*-(4-Dimethoxymethyl-benzoyl)proline-phenylalanine carboxylic acid hydrazide trimer. Isolated as described previously.[10d] R_f = 0.45 [silica gel, DCM:MeOH 95:5] (UV); IR ($CDCl_3$:MeOD 98:2) ν = 1673 (br, C=Os) cm^{-1}; ^{1}H NMR (500 MHz (cryoprobe), $CDCl_3$:MeOD 98:2) δ = 10.29 (3H, br, N**H**NC), 8.24 (3H, s, **H**CN), 7.67 (6H, d, J = 8.0 Hz, Ar–H), 7.43–7.14 (21H, m, Ar–H), 4.99 (3H, t, J = 5.6 Hz, α-H), 4.40 (3H, dd, J_1 = 10.3 Hz, J_2 = 7.0 Hz, Pro–α–H), 3.67 (3H, dd, J_1 = 14.4 Hz, J_2 = 7.0 Hz, C**H**$_a$H$_b$Ph), 3.58 (6H, m, Pro–NCH$_2$), 3.11 (3H, dd, J_1 = 14.4 Hz, J_2 = 10.3 Hz, C**H**$_a$H$_b$Ph), 2.47 (3H, m, Pro–C**H**$_a$H$_b$), 2.05 (3H, m, Pro–C**H**$_c$H$_d$), 1.87 (3H, m, Pro–CH$_a$**H**$_b$), 1.78 (3H, m, Pro–CH$_c$**H**$_d$); ^{13}C NMR (125 MHz (cryoprobe), $CDCl_3$:MeOD 98:2) δ = 171.4, 170.1, 167.2 (C=O), 148.3 (CN), 137.1, 135.7, 135.4, 129.4, 129.2, 127.9, 127.5, 127.2 (Ar), 63.2 (Pro–**C**α–H), 51.2 (Pro–N**C**H$_2$), 50.1 (**C**α–H), 36.3 (**C**H$_2$Ph), 29.6, 26.3 (CH$_2$); MS (QUATTRO) [M + Na]$^+$ $C_{66}H_{84}N_{12}O_9Na$ requires 1212, found 1212; MSMS (QUATTRO) [trimer + H]$^+$ 1172 ($C_{66}H_{67}N_{12}O_9$), [trimer − phenylalanine unit]$^+$ 1025 ($C_{57}H_{58}N_{11}O_8$), [dimer + H]$^+$ 781 ($C_{44}H_{45}N_8O_6$), [monomer − phenylalanine unit]$^+$ 243 ($C_{13}H_{14}N_3O_2$); m.pt. 262 °C (decomposition).

Acknowledgements

We are grateful for support from GlaxoSmithKline and BBSRC to SLR, from the Fundación Antorchas and the Consejo Nacional de Investigaciones Científicas y Tecnicas (Argentina) to RLEF, from the Royal Society (University Research Fellowship) to SO, and from EPSRC to JKMS.

References

1 For reviews, see: (*a*) S. Otto, R. L. E. Furlan and J. K. M. Sanders, *Curr. Opin. Chem. Biol.*, 2002, **6**, 321; (*b*) R. L. E. Furlan, S. Otto and J. K. M. Sanders, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 4801; (*c*) S. J. Rowan, S. J. Cantrill, G. R. L. Cousins, J. K. M. Sanders and J. F. Stoddart, *Angew. Chem., Int. Ed.*, 2002, **41**, 898; (*d*) O. Ramström, T. Bunyapaiboonsri, S. Lohmann and J.-M. Lehn, *Biochim. Biophys. Acta*, 2002, **1572**, 178; (*e*) J.-M. Lehn and A. V. Eliseev, *Science*, 2001, **291**, 2331; (*f*) C. Karan and B. L. Miller, *Drug Discov. Today*, 2000, **5**, 67; (*g*) A. Ganesan, *Angew. Chem., Int. Ed.*, 1998, **37**, 2828.

2 (*a*) P. A. Brady, R. P. Bonar-Law, S. J. Rowan, C. J. Suckling and J. K. M. Sanders, *Chem. Commun.*, 1996, 319; (*b*) S. J. Rowan, P. A. Brady and J. K. M. Sanders, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 2143; (*c*) S. J. Rowan, P. A. Brady and J. K. M. Sanders, *Tetrahedron Lett.*, 1996, **37**, 6013; (*d*) S. J. Rowan and J. K. M. Sanders, *J. Org. Chem.*, 1998, **63**, 1536; (*e*) P. A. Brady and J. K. M. Sanders, *J. Chem. Soc., Perkin Trans. 1*, 1997, 3237; (*f*) G. Kaiser and J. K. M. Sanders, *Chem. Commun.*, 2000, 1763.

3 Imine exchange: (*a*) I. Huc and J.-M. Lehn, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 2106; (*b*) A. Star, I. Goldberg and B. Fuchs, *Angew. Chem., Int. Ed.*, 2000, **39**, 2685; (*c*) O. Storm and U. Lüning, *Chem. Eur. J.*, 2002, **8**, 793; Oxime exchange (*d*) N. Nazarpack-Kandlousy, J. Zweigenbaum, J. Henion and A. V. Eliseev, *J. Comb. Chem.*, 1999, **1**, 199 . Hydrazone exchange; (*e*) G. R. L. Cousins, S.-A. Poulsen and J. K. M. Sanders, *Chem. Commun.*, 1999, 1575; (*f*) V. Berl, I. Huc, J.-M. Lehn, A. DeCian and J. Fischer, *Eur. J. Inorg. Chem.*, 1999, 3089; (*g*) T. Bunyapaiboonsri, O. Ramström, S. Lohmann, J.-M. Lehn, L. Peng and M. Goeldner, *ChemBioChem*, 2001, **2**, 438.

4 T. Giger, M. Wigger, S. Audetat and S. A. Benner, *Synlett*, 1998, 688.

5 (*a*) S. Otto, R. L. E. Furlan and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2000, **122**, 12063; (*b*) O. Ramström and J.-M. Lehn, *ChemBioChem*, 2000, **1**, 41.

6 (*a*) P. G. Swann, R. A. Casanova, A. Desai, M. M. Frauenhoff, M. Urbancic, U. Slomczynska, A. J. Hopfinger, G. C. Le Breton and D. L. Venton, *Biopolymers*, 1996, **40**, 617; (*b*) R. J. Lins, S. L. Flitsch, N. J. Turner, E. Irving and S. A. Brown, *Angew. Chem., Int. Ed.*, 2002, **41**, 3405.

7 (*a*) M. Crego Calama, R. Hulst, R. Fokkens, N. M. M. Nibbering, P. Timmerman and D. N. Reinhoudt, *Chem. Commun.*, 1998, 1021; (*b*) M. Crego Calama, P. Timmerman and D. N. Reinhoudt, *Angew. Chem., Int. Ed.*, 2000, **39**, 755; (*c*) F. Hof, C. Nuckolls and J. Rebek, *J. Am. Chem. Soc.*, 2000, **122**, 4251.

8 (*a*) D. M. Epstein, S. Choudhary, R. M. Churchill, K. M. Keil, A. V. Eliseev and J. R. Morrow, *Inorg. Chem.*, 2001, **40**, 1591; (*b*) S. Choudhary and J. R. Morrow, *Angew. Chem., Int. Ed.*, 2002, **41**, 4096; (*c*) I. Huc, M. J. Krische, D. P. Funeriu and J.-M. Lehn, *Eur. J. Inorg. Chem.*, 1999, 1415; (*d*) E. Stulz, Y.-F. Ng, S. M. Scott and J. K. M. Sanders, *Chem. Commun.*, 2002, 524; (*e*) M. Ziegler, J. J. Miranda, U. N. Andersen, D. W. Johnson, J. A. Leary and K. N. Raymond, *Angew. Chem., Int. Ed.*, 2001, **40**, 733; (*f*) M. Albrecht, O. Blau and R. Fröhlich, *Chem. Eur. J.*, 1999, **5**, 48; (*g*) S. Hiraoke and M. Fujita, *J. Am. Chem. Soc.*, 1999, **121**, 10239; (*h*) Y. Kubota, S. Sakamoto, K. Yamaguchi and M. Fujita, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 4854; (*i*) S. Sakai, Y. Shigemasa and T. Sasaki, *Tetrahedron Lett.*, 1997, **38**, 8145; (*j*) B. Klekota, M. H. Hammond and B. L. Miller, *Tetrahedron Lett.*, 1997, **38**, 8639; (*k*) B. Klekota and B. L. Miller, *Tetrahedron*, 1999, **55**, 11687; (*l*) C. Karan and B. L. Miller, *J. Am. Chem. Soc.*, 2001, **123**, 7455; (*m*) M. A. Case and G. L. McLendon, *J. Am. Chem. Soc.*, 2000, **122**, 8089; (*n*) E. C. Constable, C. E. Housecroft, T. Kulke, C. Lazzarini, E. R. Schofield and Y. Zimmermann, *J. Chem. Soc., Dalton Trans.*, 2001, 2864; (*o*) V. Goral, M. I. Nelen, A. V. Eliseev and J.-M. Lehn, *Proc. Natl. Acad. Sci. USA*, 2001, **98**, 1347.

9 (*a*) A. V. Eliseev and M. I. Nelen, *J. Am. Chem. Soc.*, 1997, **119**, 1147; (*b*) A. V. Eliseev and M. I. Nelen, *Chem. Eur. J.*, 1998, **4**, 825.

10 (*a*) S. Otto, R. L. E. Furlan and J. K. M. Sanders, *Science*, 2002, **297**, 590; (*b*) R. L. E. Furlan, Y.-F. Ng, G. R. L. Cousins, J. E. Redman and J. K. M. Sanders, *Tetrahedron*, 2002, **58**, 771; (*c*) M. Hochgürtel, H. Kroth, D. Piecha, M. W. Hofmann, C. Nicolau, S. Krause, O. Schaaf, G. Sonnenmoser and A. V. Eliseev, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 3382; (*d*) R. L. E. Furlan, Y.-F. Ng, S. Otto and J. K. M. Sanders, *J. Am. Chem. Soc.*, 2001, **123**, 8876; (*e*) G. R. L. Cousins, R. L. E. Furlan, Y.-F. Ng, J. E. Redman and J. K. M. Sanders, *Angew. Chem., Int. Ed.*, 2001, **40**, 423; (*f*) H. Hioki and W. C. Still, *J. Org. Chem.*, 1998, **63**, 904.

11 Nomenclature of the dipeptide building blocks: the first **m/p** denotes the *meta* or *para* position of the protected aldehyde group, the last character indicates the size of the macrocycle (**m** for monomer) and the middle two characters are the one letter codes for the amino acids, *i.e.* **P** for proline, **F** for phenylalanine, and **C** for cyclohexylalanine.

12 HPLC grade chloroform (Aldrich), stabilised by 0.1% (by weight) amylenes was used throughout.

13 Linear oligomers are also present at concentrations which are generally too low to be detected easily, but they can be amplified by binding the terminal protonated hydrazinium ion to 18-crown-6. See: R. L. E. Furlan, G. R. L. Cousins and J. K. M. Sanders, *Chem. Commun.*, 2000, 1761.

14 In the presence of NaI libraries generated from **pPFm** pass through an intermediate stage with relatively large amounts of cyclic dimer, which is only sparingly soluble. The precipitates formed upon starting equilibration from the monomer are slow to redissolve, acting as a kinetic trap, preventing thermodynamic equilibrium from being reached.

15 The role of the trifluoroacetate anion cannot be assessed, as TFA is always present in large excess.

16 L. Fielding, *Tetrahedron*, 2000, **56**, 6151.

17 A. J. Goshe, I. M. Steele, C. Ceccarelli, A. L. Rheingold and B. Bosnich, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 4823.

18 (*a*) D. L. Boger, M. A. Patane and J. Zhou, *J. Am. Chem. Soc.*, 1995, **117**, 7357; (*b*) G. Müller, G.-M. Maier and M. Lutz, *Inorg. Chim. Acta*, 1994, **218**, 121; (*c*) D. Seebach and H. G. Bossler, *Helv. Chim. Acta*, 1994, **77**, 291; (*d*) D. Seebach, A. K. Beck, H. G. Bossler, C. Gerber, S. Y. Ko, C. W. Murtiashaw, R. Naef, S. Shoda and A. Thaler, *Helv. Chim. Acta*, 1993, **76**, 1565; (*e*) M. Köck, H. Kessler, D. Seebach and A. Thaler, *J. Am. Chem. Soc.*, 1992, **114**, 2676; (*f*) H. Kessler, W. Hehlein and R. Schuck, *J. Am. Chem. Soc.*, 1982, **104**, 4534.

19 The shifts in the carbonyl stretching frequencies of **(pPC)$_3$** measured here in $CHCl_3$:MeOH (98:2 v/v) are much smaller than those reported previously (1684 to 1666 cm^{-1}) measured in pure chloroform (see reference 10*d*).

20 M. J. Blandamer, P. M. Cullis and J. B. F. N. Engberts, *J. Chem. Soc., Faraday Trans.*, 1998, **94**, 2261.

21 J. S. Bradshaw, R. M. Izatt, A. V. Bordunov, C. Y. Zhu and J. K. Hathaway, in *Comprehensive Supramolecular Chemistry, Vol. 1*, eds. J. L. Atwood, J. E. D. Davies, D. D. Macnicol and F. Vögtle, Elsevier, Oxford, 1996, pp. 65–67 and references therein.

22 B. R. Linton, M. S. Goodman, E. Fan, S. A. van Arman and A. D. Hamilton, *J. Org. Chem.*, 2001, **66**, 7313.

23 A. Clerici, N. Pastori and O. Porta, *Tetrahedron*, 1998, **54**, 15679.

24 J. I. De Graw, L. Goodman and B. R. Baker, *J. Org. Chem.*, 1961, **26**, 1156.