Dynamic combinatorial libraries of hydrazone-linked pseudo-peptides: dependence of diversity on building block structure and chirality

Jingyuan Liu, Kevin R. West, Chantelle R. Bondy and Jeremy K. M. Sanders*

Received 27th November 2006, Accepted 8th January 2007
First published as an Advance Article on the web 26th January 2007
DOI: 10.1039/b617217b

Expanding on our earlier building block architecture [$(MeO)_2CH$–Linker–Pro–X–$NHNH_2$ where X = Phe, Cha], we have produced a series of new pseudo-dipeptides [$(MeO)_2CH$–Linker–Pro–X–$NHNH_2$ where X = Val, Leu, Ile, Ala] for use in hydrazone-based dynamic combinatorial libraries (DCLs); reverse order analogues [Phe-Pro and Val-Pro] and two enantio-analogues [Pro-Phe and Pro-Val] were also prepared. The behaviours of these building blocks in DCLs, as single components and in mixtures, were studied systematically using HPLC and mass spectrometry in order to gain insight into the relationship between building block structure and good library diversity. Subtle changes in building block structure lead to significant changes in library distribution and in the ability to produce diverse libraries in mixtures.

Introduction

Dynamic Combinatorial Chemistry is a powerful approach for the discovery of new functionally successful receptors and ligands.[1] Its key feature is the dynamic combinatorial library (DCL), in which each library member is assembled from building blocks that bond reversibly to one another under specific reaction conditions. As a result of this reversibility, library members can interconvert to give a distribution that is under thermodynamic control. Addition of a guest or template molecule that can selectively recognise at least one receptor in the library will increase the concentration of that host at the expense of poorly binding members of the library.[2–4] The successful host is then isolated, identified and studied. Such systems demonstrate amplification of molecules with the desired properties in a manner that is reminiscent of the mammalian immune system. Perhaps the most remarkable example of the power of dynamic combinatorial chemistry to uncover new receptors is our unexpected discovery of a catenane that is amplified by acetylcholine chloride from the simple hydrazone DCL generated from a pseudo-dipeptide building block.[5] If this approach is to achieve its full potential, we need both a diverse library in order to maximise the probability of finding successful structures and an understanding of the relationship between building block structure and DCL behaviour.

Hydrazone exchange is one of the most successful reactions used so far in dynamic combinatorial chemistry.[1,5–8] A hydrazone linkage is formed from the acid-catalysed condensation of a hydrazide and an aldehyde. Hydrazone bonds are readily broken and formed under acidic conditions, making the process reversible, while neutralisation yields stable, isolable products. These conditions are compatible with molecular recognition events that are capable of influencing the equilibrium distribution of products, provided that all the components of the library and the template are stable under acidic conditions.

Few principles for the design of successful building blocks have emerged. Many factors need to be addressed when designing a bifunctional building block, such as solubility, rigidity/flexibility and suitable molecular recognition sites. Excessive rigidity may limit the diversity of the library members, favouring the formation of a single product or giving self-sorting in mixed libraries.[9–13] Conversely, building blocks with excessive flexibility may lead to internal cyclisation. Previously, we reported the DCL behaviour of a pseudo-dipeptide building block, **pPF**, which contains a *para*-substituted aromatic spacer (**p**), L-Pro (**P**) and L-Phe (**F**). The untemplated library produced from **pPF** has been successful at generating diversity. This may result from the Pro enforcing a degree of curvature, while the remainder of the structure gives flexibility.[6]

It is not obvious what structural changes to a building block will have a subtle or dramatic effect on the diversity of a DCL. One can imagine several possible scenarios in a DCL containing a mixture of building blocks. At one extreme they could form mixed products present in a statistical distribution with all the products having very similar geometries and properties. At the other extreme, geometrical incompatibilities could lead to self-sorting.[10–13] We have therefore investigated the effects of varying the amino acids, their chirality and the order in which they appear in the building blocks to explore the diversity of the resulting DCLs. Herein, we report how the composition of DCLs derived from these various building blocks, as single components and in mixtures, produces libraries of varying diversity. The composition of the library members in these DCLs has been determined by LC-MS, while MS-MS was used to elucidate the sequence of the constitutional isomers from mixed libraries.[14] This investigation focuses purely on analysing the diversity of the library, so not all library members were isolated for full characterisation; in a practical library one is usually interested only in successfully amplified species. Templating results from these new building blocks will be reported elsewhere.

University Chemical Laboratory, Lensfield Road, Cambridge, CB2 1EW, UK. E-mail: jkms@cam.ac.uk; Fax: +44 1223 336017; Tel: +44 1223 336411

Table 1 Summary of building blocks

Name	Amino acid 1 (AA_1)	Amino acid 2 (AA_2)	Overall yield
pPF	L-Pro	L-Phe	43% (ref. 6)
pPC	L-Pro	L-Cha	68% (ref. 6)
pPL	L-Pro	L-Leu	62%
pPV	L-Pro	L-Val	49%
pPI	L-Pro	L-Ile	45%
pPA	L-Pro	L-Ala	45%
pFP	L-Phe	L-Pro	68%
pVP	L-Val	L-Pro	39%
D-pPF	D-Pro	D-Phe	43%
D-pPV	D-Pro	D-Val	49%

Results and discussion

Syntheses of building blocks and preparation of DCLs

The structural variations made to the building blocks were restricted to the hydrophobic amino acids to ensure solubility in a single solvent system. Using the **pPF** building block as our model, the modifications explored consist of exchanging Phe for Val, Leu, Ile and Ala, as well as varying the order of the Pro amino acid between the first and second position. Enantiomeric building blocks, **D-pPF** and **D-pPV**, and a mixed chirality building block, **LD-pPF**, were also prepared; however, all DCLs prepared using the **LD-pPF** building block suffered precipitation problems, preventing further studies. Fig. 1 shows the general structure of the building blocks, while Table 1 summarises the building blocks explored and identifies the nomenclature used throughout.

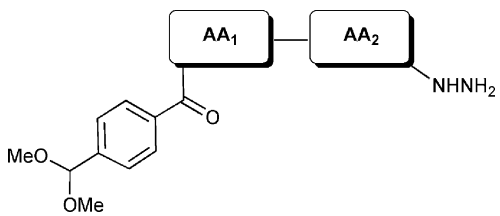
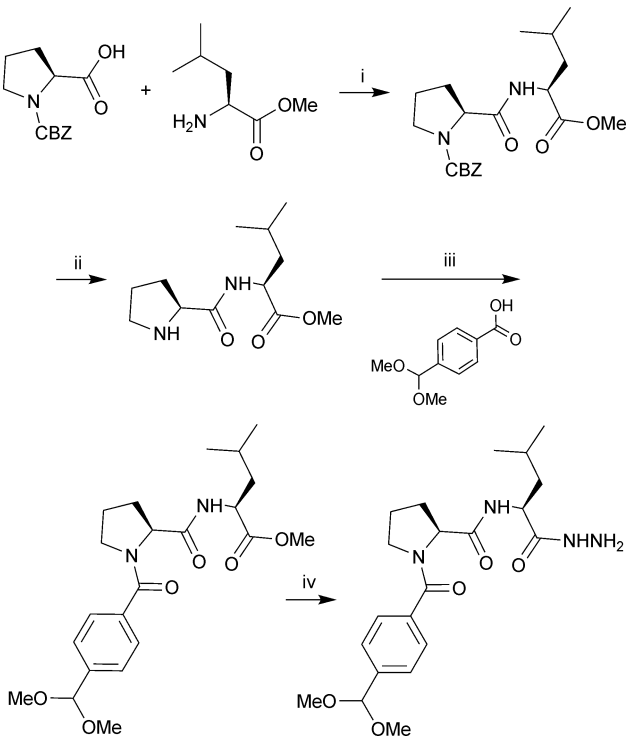


Fig. 1 General building block structure, where AA_1 and AA_2 are defined in Table 1.

The building blocks were prepared as shown in Scheme 1. The general procedure involved the EDC coupling of CBZ-protected amino acid 1 with amino acid 2 methyl ester. This was followed by the deprotection of amino acid 1 *via* Pd/C-catalysed hydrogenation. The amine of the dipeptide was then coupled to 4-carboxybenzaldehyde dimethoxyacetal (*p*DMA) using EDC, once again. The final product was obtained by hydrazinolysis of the dipeptide methyl ester, giving an overall yield that varied between 39 and 68%.

The DCLs were prepared in a $CHCl_3$–DMSO (97 : 3 v/v) solution containing a total building block concentration of 2 mM;



Scheme 1 Syntheses of building blocks using **pPL** as an example, where amino acid 1 is L-Pro and amino acid 2 is L-Leu. *Reagents and conditions*: (i) EDC, DMAP and Et_3N in dry CH_2Cl_2, inert atmosphere, 90–96%. (ii) 5% Pd/C and H_2 in MeOH, quantitative. (iii) EDC, DMAP and Et_3N in dry DCM, inert atmosphere, 54–88%. (iv) $NH_2NH_2{\cdot}H_2O$ in MeOH, 63–87%.

TFA (100 mM) was used to initiate hydrazone exchange. The solution was stirred at room temperature for 4 days, at which point thermodynamic equilibrium was reached. Control experiments were performed to confirm that DCLs were under thermodynamic control. The library members were separated using HPLC (UV = 290 nm) and identified using electrospray ionisation mass spectrometry (ESI-MS) and MS-MS to distinguish between constitutional isomers.

Product distributions in libraries of single building blocks

These DCLs gave relatively simple distributions dominated by macrocyclic dimers, trimers and tetramers. The overall distributions fell into four structure-related families (Table 2). Building blocks forming the first family contain a Pro residue as amino acid 1 and a secondary β-carbon on the side chain of amino acid 2. The building blocks **pPF**, **pPC** and **pPL** make up this family, producing similar distributions of dimers, trimers and tetramers, as exemplified by the chromatogram of **pPF** in Fig. 2a. The second family consists of two building blocks, **pPV** and **pPI**, which possess

Table 2 Building block families

Family	Building blocks	AA_2 Side-chain properties	Library composition
1	**pPF**, **pPC**, **pPL**	2° β-Carbon	Dimer ≈ trimer ≈ tetramer
2	**pPV**, **pPI**	3° β-Carbon	Tetramer
3	**pPA**	1° β-Carbon	Trimer > dimer
4	**pFP**, **pVP**	Reverse order amino acids	Trimer > dimer

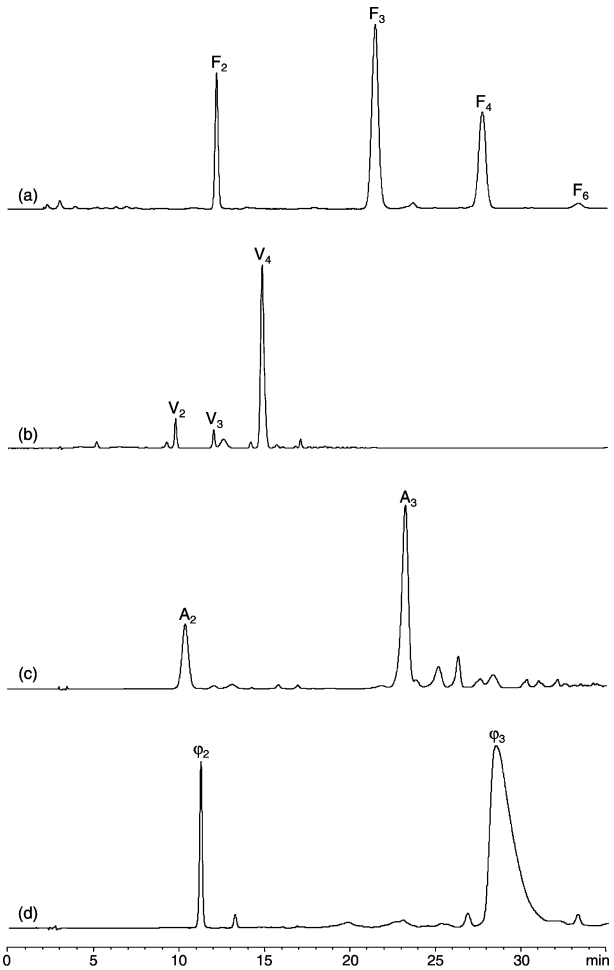


Fig. 2 HPLC traces of DCLs (a) **pPF** (F), (b) **pPV** (V), (c) **pPA** (A) and (d) **pFP** (φ). The subscript denotes the number of building blocks in each oligomer.

tertiary β-carbons on the side chains of amino acid 2 and strongly favour the formation of tetramers, although dimers and trimers are also detected (Fig. 2b). The corresponding building block with a primary β-carbon on amino acid 2, **pPA**, favours the formation of the trimer, with a modest proportion of dimer also observed (Fig. 2c); however, low solubility prevented studies of this building block in mixed libraries. The DCLs from the reverse peptide sequence building blocks, **pFP** and **pVP**, form trimers as the dominant product, with small amounts of other oligomers (Fig. 2d).

Whether the composition of the library from the first and second families resulted from stabilisation of the tetramer by intramolecular interactions or from steric destabilisation of the smaller oligomers is unclear. The much greater dominance of tetramer in the second family actually reflects a rather small energetic difference from the first family.

Mixing building blocks from the same family

As indicated above, **pPF** and **pPC** behave similarly in single building block libraries. When mixed in a DCL, these two building blocks combine in a statistical manner. As can be seen in Fig. 3a, three dimers, four trimers, and five tetramers are found. The statistical and observed ratio of possible dimers, $\mathbf{pPF_2}$: **pPF-pPC** : $\mathbf{pPC_2}$, is 1 : 2 : 1, and likewise the distributions of trimers and tetramers are found to be 1 : 3 : 3 : 1 and 1 : 4 : 6 : 4 : 1, respectively. The overall distribution of dimers, trimers and tetramers in general agrees with the distribution from the single building block libraries. Comparable diversity was also observed for the mixed DCL of **pPV** and **pPI** (Fig. 3b).

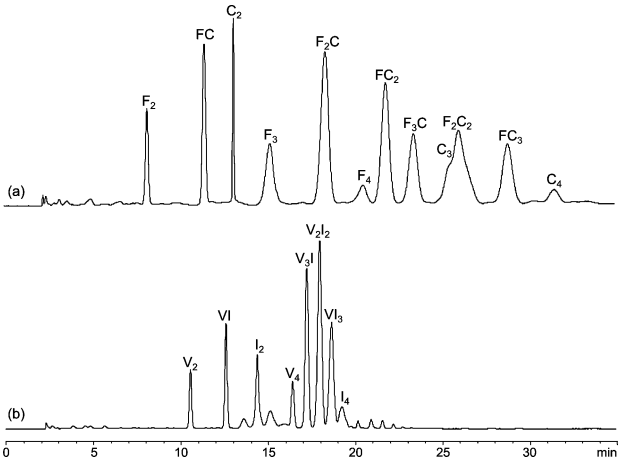


Fig. 3 HPLC traces of DCLs (a) **pPF** (F) + **pPC** (C) and (b) **pPV** (V) + **pPI** (I).

Mixing building blocks from different families

When building blocks of different families are mixed, the overall product distributions of the DCLs are unpredictable. The mixed DCL of **pPF** and **pPV** contains a broad distribution of oligomers that is dominated by mixed tetramers (Fig. 4). Two of the mixed tetramer constitutional isomers, **pPF–pPF–pPV–pPV** and **pPF–pPV–pPF–pPV**, are well separated and distinguished by their MS-MS fragmentation patterns (Fig. 5). The first-eluted tetramer (18.5 min) fragments into three distinct dimers, **pPF–pPF**, **pPF–pPV** and **pPV–pPV**, in an approximate 1 : 2 : 1 ratio, strongly suggesting the cyclic oligomer sequence of **pPF–pPF–pPV–pPV** (Fig. 5a). The second tetramer (19 min) produces only **pPF–pPV** fragments, consistent with the **pPF–pPV–pPF–pPV** sequence (Fig. 5b).

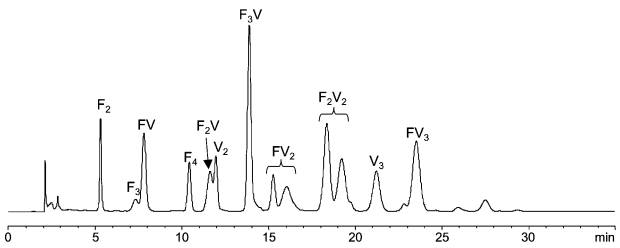


Fig. 4 HPLC trace of **pPF** (F) + **pPV** (V).

A mixture of the **pPF** and **pFP** building blocks results in a DCL with a distribution of both homo- and hetero-oligomers. Using LC-MS to compare retention times from the pure libraries, the dimers plus the homo-trimers and -tetramer were assigned (Fig. 6a). However, the composition of the hetero-oligomers cannot be identified using MS alone since **pPF** and **pFP** have identical molecular weights. To overcome this problem, **pPC** was used as a **pPF** replacement and MS marker. The mixed library of **pPC** and **pFP** shows a product distribution of mainly self-sorting

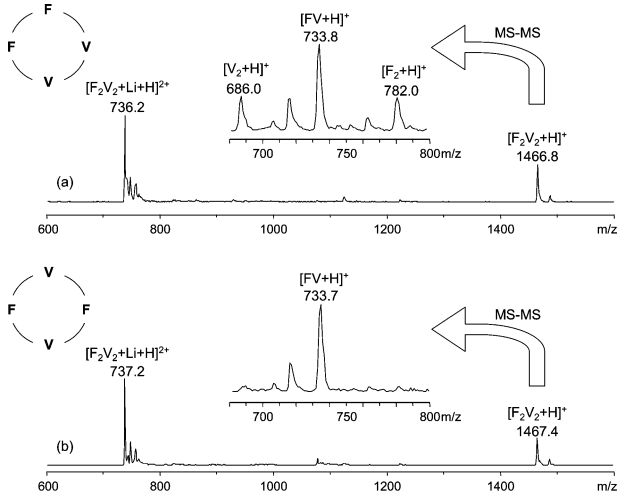


Fig. 5 MS and MS-MS of constitutional isomers of **pPF**$_2$–**pPV**$_2$ (F_2V_2) tetramers: (a) MS from LC peak at 18.5 min (inset: MS-MS); (b) MS from LC peak at 19 min (inset: MS-MS).

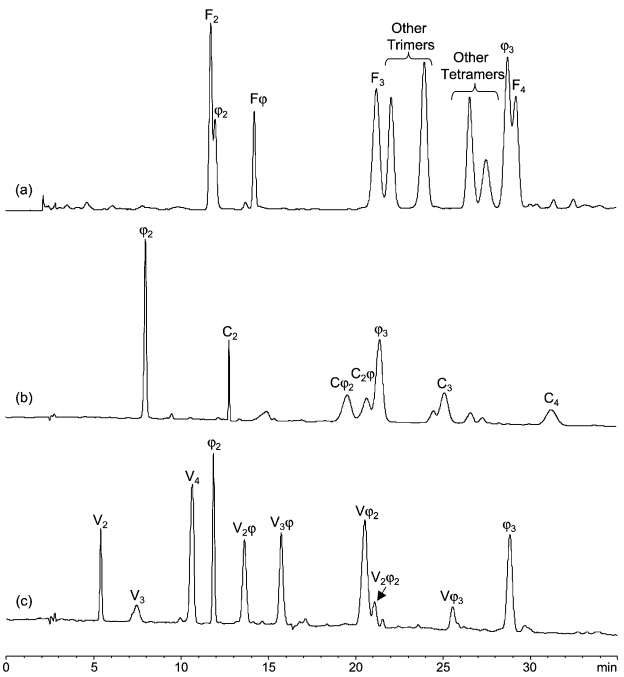


Fig. 6 HPLC traces of DCLs (a) **pPF** (F) + **pFP** (φ), (b) **pPC** (C) + **pFP** (φ) and (c) **pPV** (V) + **pFP** (φ).

homo-dimers, -trimers and -tetramers, with only trimers showing any mixing (Fig. 6b). This distribution differs from that of the mixed library of **pPF** and **pFP**. Therefore, when the order of the amino acids is changed Phe and Cha are not interchangeable, as was the case for the mixed library of **pPF** and **pPC**. This observation presumably results from inter-building block interactions that are only accessible in the mixed-order macrocycles. Similarly, the HPLC trace from the mixed library of **pPV** and **pFP** reveals the presence of homo-dimers, -trimers and -tetramers, as well as hetero-trimers and -tetramers, while the hetero-dimer is absent (Fig. 6c).

DCLs of mixed L- and D-amino acid based building blocks

The DCL of the **D-pPF** building block alone behaves identically to **pPF**, as expected. The mixed DCL of these two building blocks shows a large bias towards trimers, while the tetramers are present minimally (Fig. 7a). The homo-dimers and the mixed (*meso*) dimer are also seen in this library. Further characterisation of the diastereomeric trimers was frustrated by lack of resolution of these oligomers in the chromatogram combined with the identical masses of the building blocks. In order to overcome this problem, **pPC** was introduced as the pseudo-enantiomer to **D-pPF**.[7] Since these two building blocks are from the same family, **pPC** and **pPF** may be interchangeable.

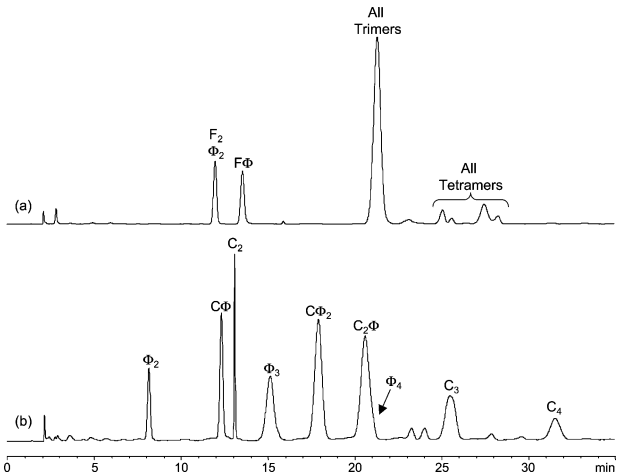


Fig. 7 HPLC traces of DCLs (a) **pPF** (F) + **D-pPF** (Φ) and (b) **pPC** (C) + **D-pPF** (Φ).

The DCL obtained from the mixture of **D-pPF** and **pPC** shows the resolution of the trimers and confirms the presence of a very small abundance of tetramers (Fig. 7b). The four possible trimers are present in a statistical proportion. This implies that the amplification of trimers in the racemic **pPF** library occurs because the mixed tetramers are in effect suppressed by their lack of stability relative to the trimers.

Similar experiments were performed by mixing **D-pPV** and **pPV** (Fig. 8a). The chromatogram reveals the presence of both homo- and hetero-dimers, -trimers and -tetramers. The mixed trimer (13.5 min) is more than three times as abundant as the homo-trimer (15 min), which indicates the preference for the mixed trimers. Once again, in order to elucidate the structure of the mixed oligomers the **pPI** building block was used as the pseudo-enantiomer to **D-pPV**, and again the mixed trimers are formed at the expense of the homo-trimers (Fig. 8b).

The final investigation into the effects of chirality on the diversity of a DCL was explored by mixing a D-building block with one L-building block from a different family. A total of six DCLs were examined, where the first three contained **D-pPF** plus one of **pPV**, **pPI** or **pFP**, and the final three contained **D-pPV** plus one of **pPF**, **pPC** or **pFP**. Generally, self-sorting is more pronounced in these DCLs compared to those of the same chirality. For example, the resulting chromatogram from mixing **D-pPF** and **pPV** shows the presence of a much lower abundance of mixed oligomers (Fig. 9) compared with the DCL from the **pPF** and **pPV** mixture (Fig. 4). The resulting DCLs from the enantiomeric and

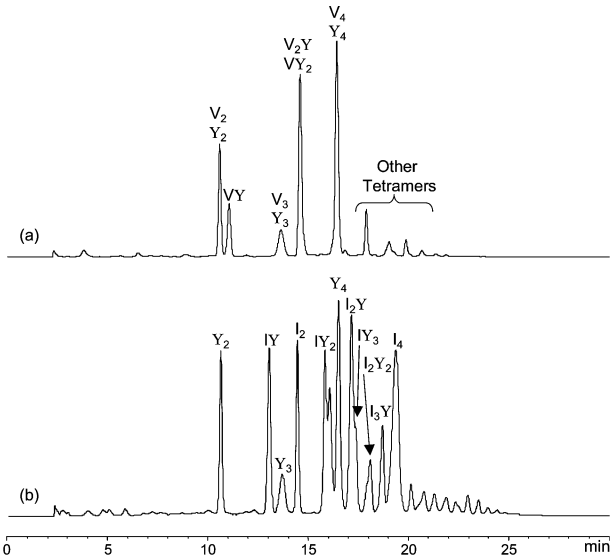


Fig. 8 HPLC traces of DCLs (a) **pPV** (V) + **D-pPV** (Y) and (b) **pPI** (I) + **D-pPV** (Y).

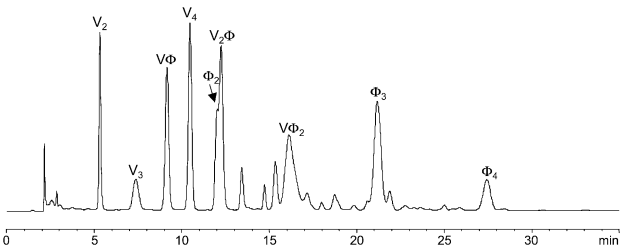


Fig. 9 HPLC trace of DCL of **pPV** (V) + **D-pPF** (Φ).

pseudo-enantiomeric building blocks are examples of how DCLs operate to optimise the free energy of the entire system.

Conclusions

Working from the starting point of an established building block, **pPF**, systematic alterations to its structure were made to produce DCLs with a view to producing greater diversity. A change to **pPF** such as substituting the Phe residue with Cha (**pPC**) does little to affect the behaviour of the building block. Indeed, in our experiments, the building blocks **pPF** and **pPC** produce similar product distributions with each other as they do with themselves. More dramatic changes in the behaviour of the DCLs are achieved by making changes that are more significant to the building block structure. For example, substitution of the Phe for Val (**pPV**) changes the distribution of library members significantly. A single building block library of **pPV** strongly favours the formation of tetramer as opposed to the relatively even distribution of dimer, trimer and tetramer seen with **pPF**. A mixed DCL of **pPF** and **pPV** favours the formation of mixed species, and the resulting constitutional isomers were identified by MS-MS.

The effects on library diversity when the chirality of the amino acid residues is mixed were also investigated. Both **L-** and **D-pPF** have a broadly even distribution of dimers, trimers and tetramers when equilibrated on their own although, when mixed, the two possible hetero-trimers are strongly favoured.

This study illustrates how relatively simple changes in the composition of an established pseudo-dipeptide building block can be used to produce subtle or major changes in the product distributions of hydrazone DCLs. The observed variation in macrocyclic product distributions at equilibrium reflects subtle and as-yet unidentified differences in thermodynamic stability. In principle this might be explored by molecular modelling, but the energy differences observed appear to be too small to allow reliable conclusions to be drawn. From a practical perspective, in searching for amplification of new receptors, it will clearly be necessary to explore a range of building block mixtures in order to explore as many potential receptors as possible.

Experimental

General

Analytical HPLC was carried out on either an Agilent 1100 or 1050 instrument coupled to a UV analyser, set to 290 nm with a reference wavelength of 550 nm, at 45 °C. The data was processed using HP Chemstation software. All separations were performed using H_2O–acetonitrile gradients, which were adjusted according to the polarity of the libraries. A 5 μL library solution was injected and pumped through a symmetry C_{18} column (25.0 cm × 4.6 mm, 5 μm) at a flow rate of 1.00 mL min^{-1}. Preparative scale isolation was performed using a Nucleodur C_{18} preparative column (25.0 cm × 2.1 cm, 100 Å, 5 μm) with a Nucleodur C_{18} guard column (5.0 cm × 2.1 cm, 100 Å, 5 μm). A 2.5 mL solution of concentrated library was injected and run through the column at a flow rate of 20.00 mL min^{-1}. The desired fractions were collected and the solvent was removed.

The LC-MS was performed using an Agilent LC-MSD-Trap-XCT system. The LC is an Agilent 1100 series HPLC equipped with an online degasser, binary pump, autosampler, heated column compartment and diode array detector. The MS was performed using an Agilent XCT ion trap MSD mass spectrometer. Mass spectra (positive mode) were acquired in ultra scan mode using a drying temperature of 350 °C, a nebuliser pressure of 60.00 psi, a drying gas flow of 11.00 L min^{-1}, a capillary voltage of 4000 V and capillary current of 39 nA. HRMS-ESI was performed using a Waters LCT Premier instrument. Mass spectra (positive mode) were acquired using a capillary voltage of 3000.0 V and a sample cone of 50.0 V. The desolvation temperature was 200.0 °C with a source temperature of 100.0 °C, a desolvation gas flow of 350.0 L h^{-1} and an ion energy of 30.0 V.

The NMR spectroscopy was performed on a Bruker DRX 500 MHz spectrometer at 300 K unless otherwise stated. Chemical shifts are quoted in parts per million with reference to residual protons from the deuterated solvent for the ^1H spectra. The coupling constants are reported in hertz (Hz).

Materials

All chemicals were purchased from Aldrich, Fluka, Sigma, Lancaster or Avocado in reagent-grade quality or better and used without further purification. All solvents where distilled prior to use and dry solvents where freshly distilled from CaH_2 under argon, with the exception of DMSO (Lancaster), which was used without further purification. Thin layer chromatography was

carried out on silica gel plates loaded with Merck silica gel 60 F_{254}. Column chromatography was performed on Merck 60 silica gel (230–400 mesh). HPLC-grade $CHCl_3$ (Aldrich) and HPLC-grade acetonitrile (Romil) were filtered with a 0.45 μm Millipore filter and used without further purification. Ultrapure H_2O was obtained from a Millipore H_2O purification system.

General DCL setup and screening. The single-component libraries were prepared by making a 2 mM building block solution in $CHCl_3$–DMSO (97 : 3 v/v) containing 100 mM of TFA. Mixed libraries were prepared by mixing two of the above solutions in a 1 : 1 ratio. The resulting solutions were allowed to equilibrate for at least 4 days. The library members were then separated and identified using HPLC and LC-MS. Four gradient methods were used depending on the polarity of the library members. In all four methods, eluent A is acetonitrile and eluent B is H_2O. Method 1 (eluent A : B 35 : 65 to 80 : 20 over 45 min and 80 : 20 to 90 : 10 over 10 min) was used for all libraries containing **pPC**. Method 2 (eluent A : B 30 : 70 to 80 : 20 over 45 min) was used for all libraries containing **pPF** and **pFP**, except those containing **pPC**. Method 3 (eluent A : B 20 : 80 to 40 : 60 over 30 min, 40 : 60 to 90 : 10 over 2 min and 90 : 10 for 6 min) was used for all the libraries containing **pPV**, **pVP**, **pPI**, **pPL**, except those containing **pPC**, **pPF** and **pFP**. Method 4 (eluent A : B 10 : 90 to 20 : 80 over 10 min, 20 : 80 to 25 : 75 over 15 min, 25 : 75 to 40 : 60 over 10 min, 40 : 60 to 90 : 10 over 2 min and 90 : 10 for 6 min) was used for the library containing **pPA**.

General procedure for amide coupling (Reactions i and iii). Amide coupling reactions were performed using EDC and DMAP. The acid and amine were dissolved in 45 mL of dry DCM, under nitrogen. Dry Et_3N (2 equivalents) was added and the resulting solution was cooled to 0 °C on an ice bath for 30 min, after which EDC (1.2 equivalents) and DMAP (3 equivalents with respect to acid) were added. The reaction was kept at 0 °C for 1 hour, allowed to warm to room temperature and stirred (under nitrogen) overnight. The work-up involved the addition of DCM (3-fold dilution) and subsequent washing of the organic solution with two portions of H_2O. The organic layer was dried over $MgSO_4$, filtered and the solvent was removed *in vacuo*. The crude product was purified by silica gel column chromatography.

General procedure for hydrogenation (Reaction ii). Hydrogenation was carried out to deprotect the CBZ group. CBZ-protected compounds were dissolved in 30 mL of 4 : 1 EtOAc–MeOH, 5% Pd/C was added and H_2 was bubbled through the reaction until completion (approximately four hours; monitored by TLC with EtOAc as eluent). The resulting suspension was filtered through a pad of Celite and the solvent was removed *in vacuo* to afford the product, which was used without further purification.

General procedure for hydrazinolysis (Reaction iv). The methyl ester was dissolved in 15 mL of MeOH, and 1 mL of hydrazine monohydrate was added. The reaction was allowed to stir at room temperature until completion (monitored by TLC with EtOAc as eluent). The solvent was removed at room temperature *in vacuo*, 100 mL of DCM was added to the residue, and the organic phase was washed twice with H_2O, dried over $MgSO_4$, filtered and evaporated. Washing the solid residue with Et_2O yielded the product as a white solid.

Synthesis of pPL

CBZ-L-Pro-L-Leu methyl ester (Reaction i). CBZ-L-proline (0.25 g, 1.0 mmol) and L-leucine methyl ester hydrochloride (0.18 g, 1.0 mmol) were coupled through the standard amide coupling process. Silica gel column chromatography (EtOAc–Hex = 7 : 3, $R_f = 0.55$) yielded the product as a colourless oil (0.34 g, 90%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.38–7.24 (5H, br, Ar-*H*), 7.06 and 6.25 (1H, br, Leu-N*H*), 5.19–5.09 (2H, m, OCH_2Ph), 4.52 (1H, br, Leu-α), 4.34 (1H, br, Pro-α), 3.69–3.63 (3H, br, OCH_3), 3.53–3.39 (2H, br, Pro-β), 2.39–1.83 (4H, br, Pro-γ(2H) and Pro-δ(2H)), 1.65–1.35 (3H, br, Leu-β(2H) and Leu-γ(1H)), 0.88–0.85 (6H, br, Leu-δ(3H) and Leu-δ′(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 173.1, 171.3, 156.1, 136.4, 128.5, 128.0, 127.8, 67.3, 60.3, 52.1, 51.0, 46.9, 41.3, 27.9, 24.9, 24.6, 22.7, 21.9; HRMS (ESI) $[M + H]^+$ $C_{20}H_{29}N_2O_5$ requires 377.2076, found 377.2089.

L-Pro-L-Leu methyl ester (Reaction ii). Hydrogenation was carried out with CBZ-L-proline-L-leucine methyl ester (0.34 g, 0.90 mmol) following the general procedure above. Product was yielded as a colourless oil (0.21 g, 96%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.00 (1H, d, $J = 8.5$ Hz, Leu-N*H*), 4.46 (1H, m, Leu-α), 3.96 (1H, br, Pro-N*H*), 3.83 (1H, dd, Pro-α), 3.63 (3H, s, OCH_3), 3.01 (1H, m, Pro-δ), 2.90 (1H, m, Pro-δ), 2.11 (1H, m, Pro-β), 1.84 (1H, m, Pro-β), 1.67 (2H, m, Pro-γ), 1.54 (3H, m, Leu-β(2H) and Leu-γ(1H)), 0.84 (6H, t, $J = 6.5$ Hz, Leu-δ(3H) and Leu-δ′(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 173.9, 173.2, 60.2, 52.0, 50.3, 47.0, 41.1, 30.7, 25.7, 24.8, 22.7, 21.7; HRMS (ESI) $[M + H]^+$ $C_{12}H_{23}N_2O_3$ requires 243.1709, found 243.1719.

***N*-(4-(Dimethoxymethyl)benzoyl)-L-Pro-L-Leu methyl ester (Reaction iii).** L-Proline-L-leucine methyl ester (0.21 g, 0.87 mmol) and *p*DMA (0.17 g, 0.87 mmol) were coupled through the general amide coupling procedure. Silica gel chromatography (EtOAc, $R_f = 0.45$) yielded the product as colourless oil (0.30 g, 82%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.46 (4H, br, Ar-*H*), 7.26 (1H, d, $J =$ 7.5 Hz, Leu–N*H*), 5.37 (1H, s, C*H*(OCH$_3$)$_2$), 4.76 (1H, m, Pro-α), 4.53 (1H, m, Leu-α), 3.69 (3H, s, OCH_3), 3.29 (6H, s, CH(OCH_3)$_2$), 3.48 (1H, m, Pro-δ), 3.40 (1H, m, Pro-δ), 2.41 (1H, m, Pro-β), 2.01 (2H, m, Pro-β), 1.79 (2H, m, Pro-γ), 1.63–1.53 (3H, m, Leu-β(2H) and Leu-γ(1H)), 0.87 (6H, t, $J = 6$ Hz, Leu-δ(3H) and Leu-δ′(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 173.1, 170.8, 170.7, 140.2, 136.3, 126.9, 126.8, 102.5, 59.6, 52.7, 52.1, 51.1, 50.3, 41.2, 27.0, 25.4, 24.9, 22.7, 21.8; HRMS (ESI) $[M + Na]^+$ $C_{22}H_{32}N_2O_6Na$ requires 443.2158, found 443.2176.

pPL: *N*-(4-(Dimethoxymethyl)benzoyl)-L-Pro-L-Leu carboxylic acid hydrazide (Reaction iv). The hydrazinolysis was carried out with *N*-(4-(dimethoxymethyl)benzoyl)-L-proline-L-leucine methyl ester (0.30 g, 0.71 mmol) following the general procedure above, yielding the product as a white solid (0.26 g, 87%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.77 (1H, br, N*H*NH$_2$), 7.49 (4H, br, Ar-*H*), 7.17 (1H, d, $J = 7.5$ Hz, Leu-N*H*), 5.39 (1H, s, C*H*(OCH$_3$)$_2$), 4.72 (1H, m, Pro-α), 4.41–4.36 (1H, m, Leu-α), 3.86 (2H, br, NHNH_2), 3.54 (1H, m, Pro-δ), 3.47 (1H, m, Pro-δ), 3.31 (6H, s, CH(OCH_3)$_2$) 2.34 (1H, m, Pro-β), 2.10 (1H, m, Pro-β), 2.02 (1H, m, Pro-γ), 1.84 (1H, m, Pro-γ), 1.73 (1H, m, Leu-β), 1.63–1.53 (2H, m, Leu-β(1H) and Leu-γ(1H)), 0.87 (6H, m, Leu-δ(3H) and Leu-δ′(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.4, 171.4, 171.1, 140.6, 135.8, 127.0, 126.9, 102.4, 60.1, 52.8, 50.8, 50.5, 40.2, 27.5, 25.5, 24.9,

22.9, 21.8; HRMS (ESI) $[M + H]^+$ $C_{21}H_{33}N_4O_5$ requires 421.2451, found 421.2445.

Synthesis of pPV or D-pPV (with L-amino acids as an example)

CBZ-L-Pro-L-Val methyl ester (Reaction i). CBZ-L-proline (0.25 g, 1.0 mmol) and L-valine methyl ester hydrochloride (0.17 g, 1.0 mmol) were coupled through the standard amide coupling process. Silica gel column chromatography (EtOAc–Hex = 7 : 3, $R_f = 0.7$) yielded the product as colourless oil (0.33 g, 91%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.32–7.25 (5H, br, Ar-*H*), 7.18 and 6.38 (1H, br, Val-N*H*), 5.19–5.07 (2H, m, OCH_2Ph), 4.44 (1H, br, Val-α), 4.35 (1H, br, Pro-α), 3.69–3.63 (3H, br, OCH_3), 3.57–3.39 (2H, br, Pro-δ), 2.39–1.83 (4H, br, Pro-β and Pro-γ), 2.12 (1H, br, Val-β), 0.85–0.79 (6H, br, Val-γ(3H) and Val-γ′(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.0, 171.4, 156.0, 136.4, 128.4, 127.9, 127.8, 67.3, 60.2, 57.3, 51.9, 46.9, 31.1, 27.9, 24.5, 18.9, 17.6; HRMS (ESI) $[M + H]^+$ $C_{19}H_{27}N_2O_5$ requires 363.1920, found 363.1910.

L-Pro-L-Val methyl ester (Reaction ii). Hydrogenation was carried out with CBZ-L-proline-L-valine methyl ester (0.33 g, 0.91 mmol) following the general procedure above. Product was yielded as colourless oil (0.21 g, 100%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.38 (1H, d, $J = 8.0$ Hz, Val-N*H*), 4.53 (1H, br, Pro-α), 4.28 (1H, m, Val-α), 3.63 (3H, s, OCH_3), 3.30–3.23 (2H, m, Pro-δ), 2.40 (1H, m, Pro-β), 2.15 (1H, m, Val-β), 1.97 (1H, m, Pro-β), 1.95–1.80 (2H, m, Pro-γ), 0.90 (6H, m, Val-γ(3H) and Val-γ′(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 171.7, 170.7, 59.7, 58.1, 51.9, 46.8, 30.7, 30.3, 24.7, 19.0, 18.0; HRMS (ESI) $[M + H]^+$ $C_{11}H_{21}N_2O_3$ requires 229.1552, found 229.1546.

***N*-(4-(Dimethoxymethyl)benzoyl)-L-Pro-L-Val methyl ester (Reaction iii).** L-Proline-L-valine methyl ester (0.21 g, 0.92 mmol) and *p*DMA (0.18 g, 0.92 mmol) were coupled through the general amide coupling procedure. Silica gel chromatography (EtOAc, $R_f = 0.40$) yielded the product as colourless oil (0.27 g, 72%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.46 (4H, br, Ar-*H*), 7.37 (1H, d, $J = 8.5$ Hz, Val-N*H*), 5.35 (1H, s, CH(OCH$_3$)$_2$), 4.77 (1H, m, Pro-α), 4.46 (1H, m, Val-α), 3.69 (3H, s, OCH_3), 3.49 (1H, m, Pro-δ), 3.39 (1H, m, Pro-δ), 3.28 (6H, s, CH(OCH_3)$_2$), 2.41 (1H, m, Pro-β), 2.14 (1H, m, Val-β), 2.08–2.00 (2H, m, Pro-β(1H) and Pro-γ(1H)), 1.79 (1H, m, Pro-γ(1H)), 0.87 (6H, m, Val-γ(3H) and Val-γ′(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.1, 170.8, 170.7, 140.2, 136.3, 126.8, 126.7, 102.4, 59.7, 57.4, 52.7, 52.0, 50.3, 31.0, 27.0, 25.3, 19.0, 17.6; HRMS (ESI) $[M + H]^+$ $C_{21}H_{31}N_2O_6$ requires 407.2182, found 407.2188.

pPV: *N*-(4-(Dimethoxymethyl)benzoyl)-L-Pro-L-Val carboxylic acid hydrazide (Reaction iv). The hydrazinolysis was carried out with *N*-(4-(dimethoxymethyl)benzoyl)-L-proline-L-valine methyl ester (0.27 g, 0.67 mmol) following the general procedure above, yielding the product as a white solid (0.20 g, 74%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.87 (1H, br, N*H*NH$_2$), 7.49 (4H, br, Ar-*H*), 7.34 (1H, d, $J = 8.5$ Hz, Val-N*H*), 5.38 (1H, s, CH(OCH$_3$)$_2$), 4.75 (1H, m, Pro-α), 4.27 (1H, m, Val-α), 3.93–3.87 (2H, br, NHNH_2), 3.56 (1H, m, Pro-δ), 3.46 (1H, m, Pro-δ), 3.31 (6H, s, CH(OCH_3)$_2$) 2.34 (1H, m, Pro-β), 2.24 (1H, m, Val-β), 2.12–1.96 (2H, m, Pro-β(1H) and Pro-γ(1H)), 1.83 (1H, m, Pro-γ), 0.90 (6H, m, Val-γ(3H) and Val-γ′(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 171.6, 171.4, 171.0, 140.5, 136.0, 127.0, 102.5, 60.1, 57.7, 52.8, 50.5, 30.1, 27.6, 25.5, 19.3, 17.7; HRMS (ESI) $[M + Na]^+$ $C_{20}H_{30}N_4O_5Na$ requires 429.2114, found 429.2117.

Synthesis of pPI

CBZ-L-Pro-L-Ile methyl ester (Reaction i). CBZ-L-proline (0.25 g, 1.0 mmol) and L-isoleucine methyl ester hydrochloride (0.18 g, 1.0 mmol) were coupled through the standard amide coupling process. Silica gel column chromatography (EtOAc–Hex = 7 : 3, $R_f = 0.55$) yielded the product as colourless oil (0.34 g, 90%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.36–7.25 (5H, br, Ar-*H*), 7.19 and 6.39 (1H, br, Ile-N*H*), 5.19–5.08 (2H, m, OCH_2Ph), 4.49 (1H, br, Ile-α), 4.36 (1H, br, Pro-α), 3.72–3.64 (3H, br, OCH_3), 3.58–3.39 (2H, br, Pro-δ), 2.36–1.80 (5H, br, Pro-β(2H), Pro-γ(2H) and Ile-β), 1.34 (1H, br, Ile-γ), 1.09 (1H, br, Ile-γ), 0.87–0.77 (6H, br, Ile-β′(3H) and Ile-δ(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 173.1, 171.3, 156.1, 136.4, 128.5, 128.0, 127.8, 67.3, 60.3, 52.1, 51.0, 46.9, 41.3, 27.9, 24.9, 24.6, 22.7, 21.9; HRMS (ESI) $[M + H]^+$ $C_{20}H_{29}N_2O_5$ requires 377.2076, found 377.2086.

L-Pro-L-Ile methyl ester (Reaction ii). Hydrogenation was carried out with CBZ-L-proline-L-isoleucine methyl ester (0.34 g, 0.90 mmol) following the general procedure above, yielding the product as a colourless oil (0.22 g, 100%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.11 (1H, d, $J = 9.0$ Hz, Ile-N*H*), 4.46 (1H, m, Ile-α), 3.85 (1H, br, Pro-α), 3.72 (1H, br, Pro-N*H*), 3.65 (3H, s, OCH_3), 3.04 (1H, m, Pro-δ), 2.93 (1H, m, Pro-δ), 2.15 (1H, m, Pro-β), 1.88 (2H, m, Pro-β(1H) and Ile-β), 1.70 (2H, m, Pro-γ), 1.36 (1H, m, Ile-γ), 1.12 (1H, m, Ile-γ), 0.90 (6H, m, Ile-β′(3H) and Ile-δ(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 174.1, 172.3, 60.3, 56.1, 51.8, 47.1, 37.6, 30.8, 25.8, 25.0, 15.5, 11.4; HRMS (ESI) $[M + H]^+$ $C_{12}H_{23}N_2O_3$ requires 243.1709, found 243.1720.

***N*-(4-(Dimethoxymethyl)benzoyl)-L-Pro-L-Ile methyl ester (Reaction iii).** L-Proline-L-isoleucine methyl ester (0.22 g, 0.91 mmol) and *p*DMA (0.18 g, 0.92 mmol) were coupled through the general amide coupling procedure. Silica gel chromatography (EtOAc, $R_f = 0.50$) yielded the product as colourless oil (0.26 g, 68%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.46 (4H, br, Ar-*H*), 7.37 (1H, d, $J = 7.5$ Hz, Val-N*H*), 5.36 (1H, s, CH(OCH$_3$)$_2$), 4.77 (1H, m, Pro-α), 4.51 (1H, m, Ile-α), 3.69 (3H, s, OCH_3), 3.49 (1H, m, Pro-δ), 3.39 (1H, m, Pro-δ), 3.29 (6H, s, CH(OCH_3)$_2$), 2.40 (1H, m, Pro-β), 2.02 (2H, m, Pro-β(1H) and Pro-γ(1H)), 1.90 (1H, m, Ile-β), 1.79 (1H, m, Pro-γ(1H)), 1.39 (1H, m, Ile-γ), 1.16 (1H, m, Ile-γ), 0.88–0.82 (6H, m, Ile-β′(3H) and Ile-δ(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.1, 170.8, 170.7, 140.2, 136.3, 126.8, 102.5, 59.7, 56.8, 52.7, 51.9, 50.3, 37.6, 27.1, 25.3, 25.0, 15.5, 11.5; HRMS (ESI) $[M + Na]^+$ $C_{22}H_{32}N_2O_6Na$ requires 443.2158, found 443.2178.

pPI: *N*-(4-(Dimethoxymethyl)benzoyl)-L-Pro-L-Ile carboxylic acid hydrazide (Reaction iv). The hydrazinolysis was carried out with *N*-(4-(dimethoxymethyl)benzoyl)-L-proline-L-isoleucine methyl ester (0.26 g, 0.62 mmol) following the general procedure above, yielding the product as a white solid (0.19 g, 73%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.55 (1H, br, N*H*NH$_2$), 7.49 (4H, m, Ar-*H*), 7.26 (1H, d, $J = 8.0$ Hz, Ile-N*H*), 5.39 (1H, s, CH(OCH$_3$)$_2$), 4.75 (1H, m, Pro-α), 4.27 (1H, m, Ile-α), 3.87 (2H, br, NHNH_2), 3.54 (1H, m, Pro-δ), 3.46 (1H, m, Pro-δ), 3.32 (6H, s, CH(OCH_3)$_2$), 2.38 (1H, m, Pro-β), 2.12–1.96 (3H, m, Pro-β(1H), Pro-γ(1H) and Ile-β), 1.84 (1H, m, Pro-β), 1.44 (1H, m, Ile-γ), 1.10 (1H, m, Ile-γ), 0.89 (3H, d, $J = 7.0$ Hz, Ile-β′), 0.84 (3H, t, $J = 7.5$ Hz, Ile-δ);

^{13}C NMR (125 MHz, $CDCl_3$) δ 171.6, 171.3, 171.2, 140.5, 135.9, 127.0, 126.9, 102.5, 60.2, 57.0, 52.8, 50.5, 36.1, 27.5, 25.5, 24.5, 15.7, 11.3; HRMS (ESI) $[M + Na]^+$ $C_{21}H_{32}N_4O_5Na$ requires 443.2270, found 443.2285.

Synthesis of pPA

CBZ-L-Pro-L-Ala methyl ester (Reaction i). CBZ-L-proline (0.25 g, 1.0 mmol) and L-alanine methyl ester hydrochloride (0.14 g, 1.0 mmol) were coupled through the standard amide coupling process. Silica gel column chromatography (EtOAc–Hex = 7 : 3, $R_f = 0.5$) yielded the product as colourless oil (0.31 g, 93%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.36–7.25 (5H, br, Ar-*H*), 7.10 and 6.42 (1H, br, Ala-N*H*), 5.16–5.08 (2H, m, OCH_2Ph), 4.47 (1H, br, Ala-α), 4.29 (1H, br, Pro-α), 3.72–3.60 (3H, br, OCH_3), 3.56–3.38 (2H, br, Pro-δ), 2.30–1.80 (4H, br, Pro-β(2H) and Pro-γ(2H)), 1.36–1.20 (3H, br, Ala-β); ^{13}C NMR (125 MHz, $CDCl_3$) δ 173.0, 171.3, 155.9, 136.4, 128.4, 127.9, 127.8, 67.2, 60.3, 52.2, 48.0, 46.9, 28.3, 24.4, 18.0; HRMS (ESI) $[M + H]^+$ $C_{17}H_{23}N_2O_5$ requires 335.1607, found 335.1597.

L-Pro-L-Ala methyl ester (Reaction ii). Hydrogenation was carried out with CBZ-L-proline-L-alanine methyl ester (0.31 g, 0.93 mmol) following the general the procedure above, yielding the product as a colourless oil (0.18 g, 97%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 8.60 (1H, d, $J = 7.0$ Hz, Ala-N*H*), 4.50 (1H, br, Pro-α), 4.42 (1H, m, Ala-α), 3.67 (3H, s, OCH_3), 3.32 (2H, m, Pro-δ), 2.43 (1H, m, Pro-β), 2.04 (1H, m, Pro-β), 1.93 (2H, m, Pro-γ), 1.42 (3H, d, $J = 7.0$ Hz, Ala-β); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.8, 170.2, 59.8, 52.3, 48.6, 46.8, 30.6, 24.7, 17.2; HRMS (ESI) $[M + H]^+$ $C_9H_{17}N_2O_3$ requires 201.1239, found 201.1240.

***N*-(4-(Dimethoxymethyl)benzoyl)-L-Pro-L-Ala methyl ester (Reaction iii).** L-proline-L-alanine methyl ester (0.18 g, 0.91 mmol) and *p*DMA (0.18 g, 0.91 mmol) were coupled through the general amide coupling procedure. Silica gel chromatography (EtOAc, $R_f = 0.4$) yielded the product as colourless oil (0.27 g, 79%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.46–7.38 (4H, m, Ar-*H*), 7.32 (1H, d, $J = 7.0$ Hz, Ala-N*H*), 5.31 (1H, s, $CH(OCH_3)_2$), 4.65 (1H, m, Pro-α), 4.42 (1H, m, Ala-α), 3.63 (3H, s, OCH_3), 3.49 (1H, m, Pro-δ), 3.39 (1H, m, Pro-δ), 3.22 (6H, s, $CH(OCH_3)_2$), 2.24 (1H, m, Pro-β), 2.02 (1H, m, Pro-β), 1.94 (1H, m, Pro-γ), 1.72 (1H, m, Pro-γ), 1.30 (3H, d, $J = 7.0$ Hz, Ala-β); ^{13}C NMR (125 MHz, $CDCl_3$) δ 173.0, 170.8, 170.2, 140.1, 136.1, 126.9, 126.5, 102.3, 59.7, 52.5, 52.1, 50.2, 48.1, 27.8, 25.1, 17.6; HRMS (ESI) $[M + Na]^+$ $C_{19}H_{26}N_2O_6Na$ requires 401.1689, found 401.1705.

pPA: *N*-(4-(Dimethoxymethyl)benzoyl)-L-Pro-L-Ala carboxylic acid hydrazide (Reaction iv). The hydrazinolysis was carried out with *N*-(4-(dimethoxymethyl)benzoyl)-L-proline-L-alanine methyl ester (0.27 g, 0.71 mmol) following the general procedure above, yielding the product as a white solid (0.17 g, 63%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.86 (1H, br, N*H*NH_2), 7.51 (4H, m, Ar-*H*), 7.16 (1H, d, $J = 7.0$ Hz, Ala-N*H*), 5.38 (1H, s, $CH(OCH_3)_2$), 4.65 (1H, m, Pro-α), 4.42 (1H, m, Ala-α), 3.86 (2H, br, NHNH_2), 3.58 (1H, m, Pro-δ), 3.49 (1H, m, Pro-δ), 3.31 (6H, s, $CH(OCH_3)_2$), 2.28 (1H, m, Pro-β), 2.14 (1H, m, Pro-β), 2.01 (1H, m, Pro-γ), 1.83 (1H, m, Pro-γ), 1.38 (3H, d, $J = 7.0$ Hz, Ala-β); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.5, 171.3, 171.0, 140.6, 135.8, 127.2, 126.9, 102.4, 60.4, 52.7, 50.6, 48.0, 28.0, 25.5, 17.5; HRMS (ESI) $[M + H]^+$ $C_{18}H_{27}N_4O_5$ requires 379.1981, found 379.1980.

Synthesis of pVP

CBZ-L-Val-L-Pro methyl ester (Reaction i). CBZ-L-valine (0.25 g, 1.0 mmol) and L-proline methyl ester hydrochloride (0.17 g, 1.0 mmol) were coupled through the standard amide coupling process. Silica gel column chromatography (EtOAc–Hex = 7 : 3, $R_f = 0.55$) yielded the product as colourless oil (0.35 g, 97%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.33–7.25 (5H, m, Ar-*H*), 5.47 (1H, d, $J = 9.0$ Hz, Val-N*H*), 5.09–5.01 (2H, m, OCH_2Ph), 4.49 (1H, m, Pro-α), 4.30 (1H, m, Val-α), 3.76 (1H, m, Pro-δ), 3.67 (3H, s, OCH_3), 3.62 (1H, m, Pro-δ), 2.17 (1H, m, Pro-β), 2.02 (2H, m, Pro-γ(1H) and Val-β), 1.94 (2H, m, Pro-β(1H) and Pro-γ(1H)), 1.01 (3H, d, $J = 7.0$ Hz, Val-γ), 0.92 (3H, d, $J = 7.0$ Hz, Val-γ′); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.3, 170.7, 156.4, 136.4, 128.4, 127.9, 127.8, 66.7, 58.7, 57.4, 52.0, 47.1, 31.3, 28.9, 24.9, 19.1, 17.4; HRMS (ESI) $[M + H]^+$ $C_{19}H_{27}N_2O_5$ requires 363.1920, found 363.1938.

L-Val-L-Pro methyl ester (Reaction ii). Hydrogenation was carried out with CBZ-L-valine-L-proline methyl ester (0.35 g, 0.97 mmol) following the general procedure above, yielding the product as a colourless oil (0.22 g, 100%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 4.59 (1H, br, Pro-α), 4.20 (1H, d, $J = 6.0$ Hz, Val-α), 3.95 (1H, m, Pro-δ), 3.66 (3H, s, OCH_3), 3.51 (1H, m, Pro-δ), 2.32 (1H, m, Val-β), 2.25 (1H, m, Pro-β), 2.01 (1H, m, Pro-γ), 1.94 (1H, m, Pro-γ), 1.88 (1H, m, Pro-β), 1.11 (6H, m, Val-γ(3H) and Val-γ′(3H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.3, 167.9, 59.3, 57.0, 52.0, 47.8, 30.1, 29.0, 25.0, 18.6, 17.8; HRMS (ESI) $[M + H]^+$ $C_{11}H_{21}N_2O_3$ requires 229.1552, found 229.1541.

***N*-(4-(Dimethoxymethyl)benzoyl)-L-Val-L-Pro methyl ester (Reaction iii).** L-Valine-L-proline methyl ester (0.22 g, 0.96 mmol) and *p*DMA (0.19 g, 0.97 mmol) were coupled through the general amide coupling procedure. Silica gel column chromatography (EtOAc, $R_f = 0.50$) yielded the product as colourless oil (0.21 g, 54%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.73 (2H, d, $J = 8.0$ Hz, Ar-*H*), 7.44 (2H, d, $J = 8.0$ Hz, Ar-*H*), 7.03 (1H, d, $J = 8.5$ Hz, Val-N*H*), 5.36 (1H, s, $CH(OCH_3)_2$), 4.79 (1H, m, Val-α), 4.43 (1H, m, Pro-α), 3.87 (1H, m, Pro-δ), 3.67 (3H, s, OCH_3), 3.66 (1H, m, Pro-δ), 3.24 (6H, s, $CH(OCH_3)_2$), 2.19 (2H, m, Pro-β(1H) and Val-β), 2.04–1.88 (3H, m, Pro-β(1H) and Pro-γ(2H)), 1.04 (3H, d, $J = 6.5$ Hz, Val-γ), 0.96 (3H, d, $J = 6.5$ Hz, Val-γ′); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.2, 170.8, 166.9, 141.5, 134.0, 126.9, 126.8, 102.2, 58.8, 55.8, 52.4, 52.0, 47.2, 31.5, 28.9, 24.9, 19.2, 17.7; HRMS (ESI) $[M + Na]^+$ $C_{21}H_{30}N_2O_6Na$ requires 429.2002, found 429.1996.

pVP: *N*-(4-(Dimethoxymethyl)benzoyl)-L-Val-L-Pro carboxylic acid hydrazide (Reaction iv). The hydrazinolysis was carried out with *N*-(4-(dimethoxymethyl)benzoyl)-L-valine-L-proline methyl ester (0.21 g, 0.52 mmol) following the general procedure above, yielding the product as a white solid (0.16 g, 76%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.88 (1H, br, N*H*NH_2), 7.78 (2H, d, $J = 8.0$ Hz, Ar-*H*), 7.50 (2H, d, $J = 8.0$ Hz, Ar-*H*), 6.81 (1H, d, $J = 9.0$ Hz, Val-N*H*), 5.42 (1H, s, $CH(OCH_3)_2$), 4.81 (1H, m, Val-α), 4.48 (1H, m, Pro-α), 3.87 (1H, m, Pro-δ), 3.66 (1H, m, Pro-δ), 3.30 (6H, s, $CH(OCH_3)_2$), 2.34 (1H, m, Pro-β), 2.15 (2H, m, Val-β and Pro-γ(1H)), 2.01(1H, m, Pro-γ), 1.94 (1H, m, Pro-β), 1.03 (3H, d, $J = 6.5$ Hz, Val-γ), 0.98 (3H, d, $J = 6.5$ Hz, Val-γ′); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.1, 171.9, 167.0, 141.8, 134.0, 127.1, 127.0, 102.2, 58.4, 55.9, 52.6, 47.8, 31.8, 27.2, 25.2, 19.4,

17.9; HRMS (ESI) $[M + Na]^+$ $C_{20}H_{30}N_4O_5Na$ requires 429.2114, found 429.2134.

Synthesis of pFP

CBZ-L-Phe-L-Pro methyl ester (Reaction i). CBZ-L-phenylalanine (0.30 g, 1.0 mmol) and L-proline methyl ester hydrochloride (0.17 g, 1.0 mmol) were coupled through the standard amide coupling process. Silica gel column chromatography (EtOAc–Hex = 7 : 3, $R_f = 0.55$) yielded the product as colourless oil (0.38 g, 93%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.33–7.16 (10H, m, Ar-*H* and Phe-Ar), 5.64 (1H, d, $J = 8.5$ Hz, Phe-N*H*), 5.08–4.96 (2H, m, OCH_2Ph), 4.69 (1H, m, Phe-α), 4.47 (1H, m, Pro-α), 3.70 (3H, s, OCH_3), 3.59 (1H, m, Pro-δ), 3.17 (1H, m, Pro-δ), 3.09 (1H, m, Phe-β), 2.91 (1H, m, Phe-β), 2.16–1.60 (4H, m, Pro-β(2H) and Pro-γ(2H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.1, 171.7, 170.2, 136.3, 136.1, 129.6, 128.3, 128.2, 127.9, 127.8, 126.7, 66.6, 58.8, 53.7, 52.0, 46.7, 38.9, 28.9, 24.7; HRMS (ESI) $[M + H]^+$ $C_{23}H_{27}N_2O_5$ requires 411.1920, found 411.1933.

L-Phe-L-Pro methyl ester (Reaction ii). Hydrogenation was carried out with CBZ-L-phenylalanine-L-proline methyl ester (0.38 g, 0.92 mmol) following the general procedure above, yielding the product as a colourless oil (0.25 g, 98%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.29–7.18 (5H, m, Phe-Ar), 4.50 (2H, br, Pro-α and Phe-α), 3.65 (3H, s, OCH_3), 3.60 (1H, m, Pro-δ), 3.46 (1H, m, Phe-β), 3.21 (1H, m, Phe-β), 2.54 (1H, m, Pro-δ), 2.10 (1H, m, Pro-β), 1.79 (2H, m, Pro-γ(1H) and Pro-β(1H)), 1.66 (1H, m, Pro-γ); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.0, 167.9, 134.4, 130.3, 128.5, 127.3, 59.4, 53.4, 52.1, 47.1, 37.3, 28.9, 24.8; HRMS (ESI) $[M + H]^+$ $C_{15}H_{21}N_2O_3$ requires 277.1552, found 277.1564.

***N*-(4-(Dimethoxymethyl)benzoyl)-L-Phe-L-Pro methyl ester (Reaction iii).** L-Phenylalanine-L-proline methyl ester (0.25 g, 0.91 mmol) and *p*DMA (0.18 g, 0.92 mmol) were coupled through the general amide coupling procedure. Silica gel column chromatography (EtOAc, $R_f = 0.55$) yielded the product as colourless oil (0.36 g, 88%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.64 (2H, d, $J = 8.5$ Hz, Ar-*H*), 7.38 (2H, d, $J = 8.5$ Hz, Ar-*H*), 7.32–7.16 (6H, m, Phe-Ar and Phe-N*H*), 5.34 (1H, s, C*H*$(OCH_3)_2$), 5.11 (1H, m, Phe-α), 4.48 (1H, m, Pro-α), 3.71 (1H, m, Pro-δ), 3.68 (3H, s, OCH_3), 3.25 (6H, s, CH$(OCH_3)_2$), 3.24–3.08 (3H, m, Phe-β(2H) and Pro-δ(1H)), 2.16 (1H, m, Pro-β), 1.92 (2H, m, Pro-β(1H) and Pro-γ(1H)), 1.65 (1H, m, Pro-γ); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.1, 170.5, 166.3, 141.4, 136.2, 133.8, 129.7, 128.3, 126.9, 126.8, 126.7, 102.3, 59.0, 52.5, 52.5, 52.1, 46.9, 38.3, 28.9, 24.9; HRMS (ESI) $[M + Na]^+$ $C_{25}H_{30}N_2O_6Na$ requires 477.2002, found 477.1996.

pFP: *N*-(4-(Dimethoxymethyl)benzoyl)-L-Phe-L-Pro carboxylic acid hydrazide (Reaction iv). The hydrazinolysis was carried out with *N*-(4-(dimethoxymethyl)benzoyl)-L-phenylalanine-L-proline methyl ester (0.36 g, 0.79 mmol) following the general procedure above, yielding the product as a white solid (0.31 g, 86%). ^{1}H NMR (500 MHz, $CDCl_3$) δ 7.75 (2H, d, $J = 8.0$ Hz, Ar-*H*), 7.69 (1H, br, N*H*NH_2), 7.50 (2H, d, $J = 8.0$ Hz, Ar-*H*), 7.32–7.16 (5H, m, Phe-Ar), 6.92 (1H, d, $J = 8.0$ Hz, Phe-N*H*), 5.41 (1H, s, C*H*$(OCH_3)_2$), 5.16 (1H, m, Phe-α), 4.48 (1H, m, Pro-α), 3.64 (1H, m, Pro-δ), 3.30 (6H, s, CH$(OCH_3)_2$), 3.15 (2H, m, Phe-β), 3.06 (1H, m, Pro-δ), 2.30 (1H, m, Pro-β), 1.88 (3H, m, Pro-β(1H) and Pro-γ(2H)); ^{13}C NMR (125 MHz, $CDCl_3$) δ 171.5, 171.1, 166.4, 141.9, 135.6, 133.7, 129.5, 128.7, 127.4, 127.1, 127.0, 102.2, 58.6, 52.6, 52.3, 47.4, 39.1, 27.0, 25.0; HRMS (ESI) $[M + Na]^+$ $C_{24}H_{30}N_4O_5Na$ requires 477.2114, found 477.2116.

Acknowledgements

We thank the Engineering and Physical Sciences Research Council for studentships to J. L. and K. R. W., the National Sciences and Engineering Research Council of Canada for a post-doctoral fellowship for C. R. B., Dr Ana Belenguer for managing the HPLC facility and Dr Sijbren Otto for fruitful discussions.

References

1 P. T. Corbett, J. Leclaire, L. Vial, K. R. West, J.-L. Wietor, J. K. M. Sanders and S. Otto, *Chem. Rev.*, 2006, **106**, 3652–3711.
2 P. T. Corbett, S. Otto and J. K. M. Sanders, *Chem. Eur. J.*, 2004, **10**, 3139–3143.
3 K. Severin, *Chem. Eur. J.*, 2004, **10**, 2565–2580.
4 I. Saur and K. Severin, *Chem. Commun.*, 2005, 1471–1473.
5 R. T. S. Lam, A. Belenguer, S. L. Roberts, C. Naumann, T. Jarrosson, S. Otto and J. K. M. Sanders, *Science*, 2005, **308**, 667–669.
6 S. L. Roberts, R. L. E. Furlan, S. Otto and J. K. M. Sanders, *Org. Biomol. Chem.*, 2003, **1**, 1625–1633.
7 S. M. Voshell, S. J. Lee and M. R. Gagné, *J. Am. Chem. Soc.*, 2006, **128**, 12422–12423.
8 T. Bunyapaiboonsri, O. Ramstrom, S. Lohmann, J. M. Lehn, L. Peng and M. Goeldner, *ChemBioChem*, 2001, **2**, 438–444.
9 M. G. Simpson, S. P. Watson, N. Feeder, J. E. Davies and J. K. M. Sanders, *Org. Lett.*, 2000, **2**, 1435–1438.
10 S. J. Rowan, D. G. Hamilton, P. A. Brady and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1997, **119**, 2578–2579.
11 S. J. Rowan, P. S. Lukeman, D. J. Reynolds and J. K. M. Sanders, *New J. Chem.*, 1998, **22**, 1015–1018.
12 I. Saur, R. Scopelliti and K. Severin, *Chem. Eur. J.*, 2006, **12**, 1058–1066.
13 S. Liu, C. Ruspic, P. Mukhopadhyay, S. Chakrabarti, P. Y. Zavalij and L. Isaacs, *J. Am. Chem. Soc.*, 2005, **127**, 15959–15967.
14 S.-A. Poulsen, P. J. Gates, G. R. L. Cousins and J. K. M. Sanders, *Rapid Commun. Mass Spectrom.*, 2000, **14**, 44–48.

