Stereoselective synthesis and structure determination of a bicyclo[3.3.2]decapeptide

Marco Bartoloni,[a] Sandro Waltersperger,[b] Mario Bumann,[a] Achim Stocker,[a] Tamis Darbre,[a] and Jean-Louis Reymond[a]*

[a]*Department of Chemistry and Biochemistry, University of Bern, Freiestrasse 3, 3012 Bern, Switzerland*
[b]*Paul Scherrer Institut, 5232 Villigen PSI, Switzerland*
E-mail: jean-louis.reymond@ioc.unibe.ch

This paper is dedicated to Prof. Pierre Vogel, a master of bicyclic molecules, on the occasion of his 70th Birthday

Abstract

By analogy to the structural diversity of covalent bond networks between atoms within organic molecules, one can design topologically diverse peptides from mathematical graphs by assigning amino acids to graph nodes and peptide bonds to graph edges. The key is to use diamino acids or amino diacids as equivalents of trivalent graph nodes, which enables a variety of graph topologies beyond the standard linear and monocyclic graphs in natural peptides. Here the bicyclic decapeptide $A^1FGk^2VFPE^1AG^2$ (**1b**) was prepared and crystallized to assign its bridge stereochemistry. The bridge configuration appears as planned by the chirality of the branching amino acids. Bicyclization furthermore depends on the presence of matched chiralities in the branching amino acids. The stereoselective formation of the second bridge opens the way for the synthesis of a large family of bicyclic peptides as promising new scaffolds for drug design.

Keywords: Bicyclic peptides, orthogonality, X-ray crystallography, stereochemistry

Introduction

In analogy with the structural diversity of covalent bond networks between atoms within organic molecules,[1,2] topologically diverse peptides can be designed from mathematical graphs by assigning amino acids to graph nodes and peptide bonds to graph edges. The key is to use diamino acids (e.g. lysine) or amino diacids (e.g. glutamate) as equivalents of trivalent graph nodes, which enables branched peptide chains and therefore a variety of graph topologies beyond

the standard linear and monocyclic graphs found in naturally occurring peptides and proteins. This graph based strategy opens vast possibilities for molecular design in the size range of MW 1000-5000 Da, which is usually quite difficult to access by synthesis yet is required for various applications where small molecules fail, such as disrupting protein-protein interactions.[3,4] Branched peptide chains are expected to be conformationally constrained or even completely rigid, resulting in stronger and more selective target binding and increased resistance to proteolytic degradation compared to linear peptides. The use of standard amino acid building blocks furthermore ensures relatively low toxicity and unproblematic metabolism, which is one key advantage of peptides when considering therapeutic use.[5]

Surprisingly, relatively little attention has been paid to a systematic exploration of graph based peptide design, except for the preparation of bi- or polycyclic peptides using side chain cross-links[6] as found in cysteine knot proteins[7] and lantibiotics.[8] We have shown that peptide equivalents of multi-branched graphs can be prepared in good yields and purities by solid phase peptide synthesis (SPPS), a robust method to prepare peptides in the laboratory as well as on an industrial scale.[9] For instance, peptide dendrimers can be obtained in excellent yields by the successive use of branching diamino acids producing sequence-symmetrical molecular trees,[10] where multivalency effects enable a variety of properties inaccessible to small molecules such as enhancement of organocatalysis,[11] remote control of metal coordination properties,[12] lectin and biofilm inhibition,[13] membrane disruptive antimicrobial effects,[14] and DNA transfection activities.[15]

In a similar design, we recently showed that bicyclic peptides comprising two branching residues can be produced efficiently by SPPS using a simple orthogonal protection scheme.[16] In this synthesis, a first peptide cyclization is carried out on the solid support by amide bond formation between the *N*-terminal amino group and the γ-carboxyl group of a side chain allyl ester protected glutamate residue after palladium deprotection. The second cyclization is then carried out between the free C-terminal carboxyl group and the ε-amino group of a Boc protected lysine residue after acidic deprotection and simultaneous cleavage from the support (Scheme 1). The SPPS of bicyclic peptides gave good results for various combinations such as bicyclo[3.3.2]decapeptide **1b**, norbornapeptide **2b** and bicyclo[2.2.2]octapeptide **3b**. In all three sequences, one proline residue was included to facilitate on-resin cyclization, while the remaining positions were diversified with amino acids carrying non-protected side chains.

When designing the bicyclic peptides the chirality of the branching diamino acid was planned to enable formation of the bridge above or below the plane of the first cycle in a stereoselective manner.[17] The obtainment of single products after bicyclization indeed strongly pointed to a stereoselective process. In the case of the smaller bicyclic peptides such as **2b** and **3b**, structure determination by ^{1}H-NMR showed unambiguously that the process was indeed stereoselective and that the expected bridge geometry was formed. In the case of the larger bicyclo[3.3.2]decapeptide **1b**, by constrast, a single product was also observed but the ^{1}H-NMR data was inconclusive as to the bridge stereochemistry due to the much smaller number of observed NOE-distance constraints, probably caused by the larger and somewhat more flexible

tripeptide loops. Thus, the issue of which of the two possible bridge stereoisomers, **1b** or **1b***, was formed could not be resolved (Figure 1). We therefore set out to test whether the structure could be determined by crystallography. Herein we report the structure of **1b**, which represents one of the few X-ray crystal structures of a homodetic bicyclic peptide.[18] The data shows that for this larger bicyclic system compared to the norbornapeptides the expected bridge configuration is indeed formed, as planned by the chirality of the branching amino acids. The stereoselective bridge formation as demonstrated here opens the way for the planned synthesis of a large family of conformationally constrained bicyclic peptides, which constitute a promising new class of scaffolds for drug design.
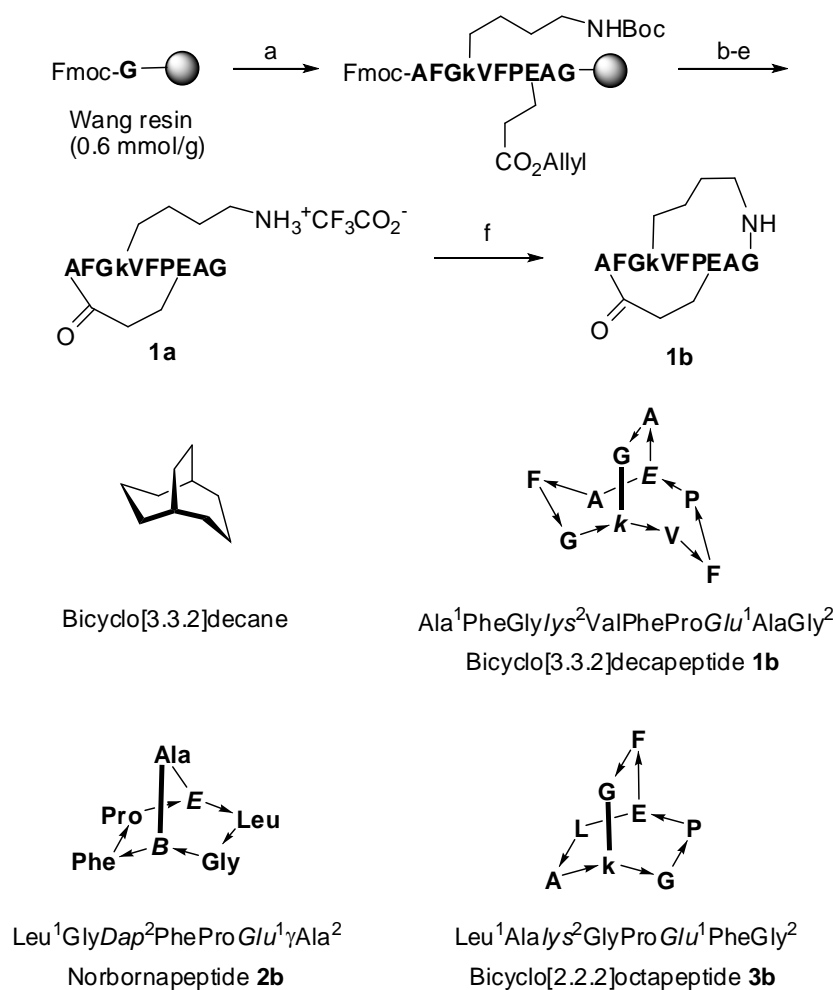
Scheme 1. Solid-phase synthesis of **1b** and graphs of peptides **1-3b**. Conditions: a- Fmoc SPPS; b- Pd(PPh₃)₄, PhSiH₃, dry CH₂Cl₂; c- piperidine/DMF 1:4; d- HATU, DIEA, NMP/DMSO 4:1; e-TFA/TIS/H₂O 94:5:1; f- PyBOP, DIEA, DCM/DMF 10:1.
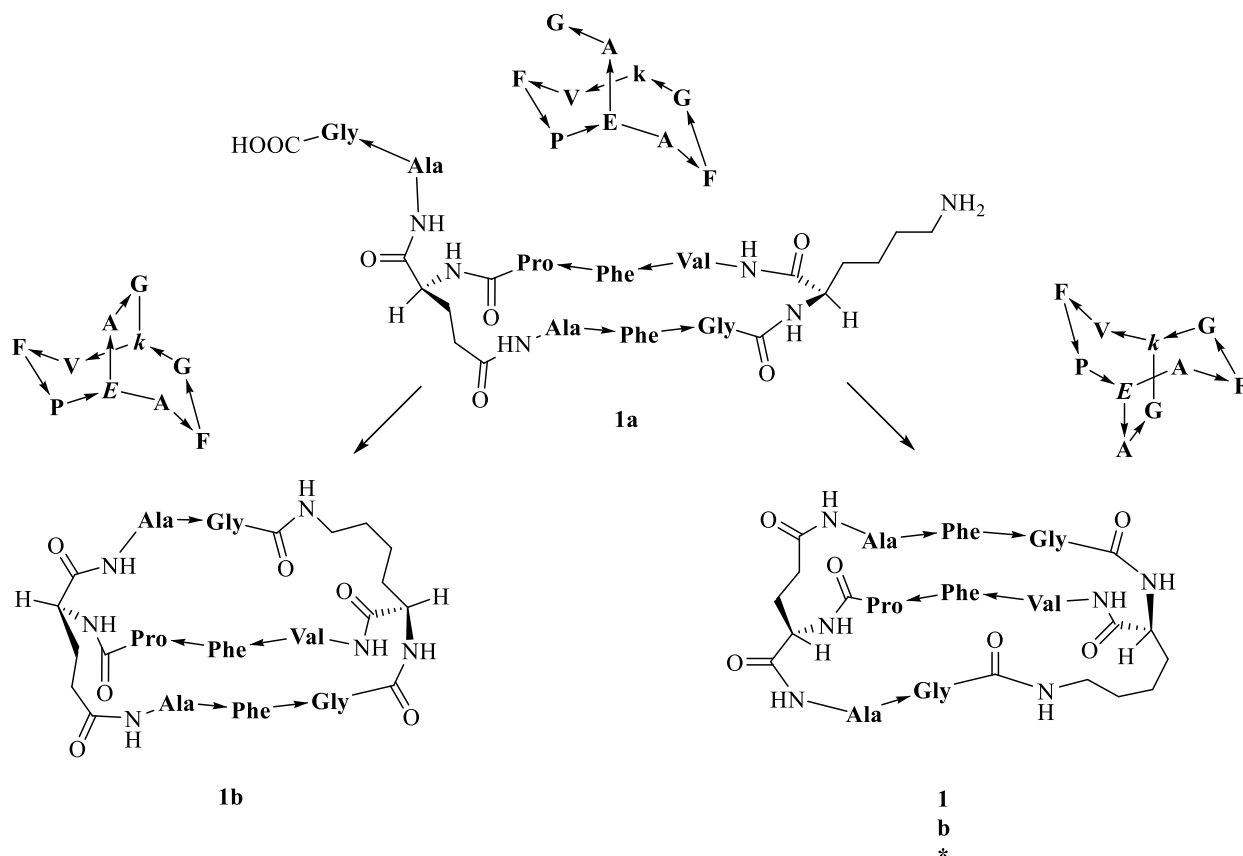
Figure 1. Possible cyclization diastereoisomers and their graph representations. Arrows in the structures indicate the direction of the amide bonds.

Results and Discussion

Synthesis

The linear sequence of peptide **1b** was first assembled by Fmoc SPPS from commercially available building blocks. The sequence was designed to allow the first cyclization between the N-terminal alanine residue and the side-chain carboxyl group of the glutamate residue to form the larger octapeptide ring. The cyclization was performed on resin using HATU as coupling agent. After completion of the reaction as detected by the TNBS test, the resin was treated with trifluoroacetic acid, which effected cleavage of the Wang-linker and removal of the lysine side chain Boc protecting group. The branched cyclic octapeptide **1a** was obtained in 27% yield after purification by preparative RP-HPLC.

The second cyclization to form bicyclic peptide **1b** was performed by subjecting the pure monocyclic peptide **1a** to amide bond formation condition in high dilution (0.7 mM in DCM/DMF 8:1), using PyBOP as coupling agent. The reaction proceeded smoothly over 2 h as followed by HPLC and gave a single product corresponding to the mass of **1b**. The product was

isolated in 38% yield after preparative RP-HPLC (Table 1) and its purity was confirmed by ^{1}H-NMR.

To test the sensitivity of the second cyclization to the stereochemistry of the branching amino acids, the stereoisomeric monocyclic peptide **1a'** was prepared bearing an L-lysine branching residue at position 7, which would be expected to hamper the formation of the second cycle since the stereochemistry of the bridging pair is mismatched. Indeed, while the monocyclic octapeptide **1a'** was obtained in comparable yield as **1a**, attempts to form the second cycle from **1a'** under the same conditions only gave traces of bicyclic product even after prolonged incubation, showing that formation of the second peptide ring is subjected to stereochemical constraints. We could indeed isolate a peptide with the right mass, but with a complex ^{1}H-NMR spectrum, possibly due to the presence of a conformational mixture (Figure S1).

Table 1. Data for the formation of **1a**, **1a'** and **1b**. Superscript 1 and 2 indicate residues connected in the 1st and 2nd cyclization, respectively

| No.	nr. AA	Sequence	mg	%	MS calc.	MS obs.
1a	10	A^{1}FGkVFPE^{1}AG	67.5	27	1004.52	1004.52
1a'	10	A^{1}FGKVFPE^{1}AG	43.8	17	1004.52	1004.52
1b	10	A^{1}FG*k^2*VFP*E^1*AG^2	22.8	38	986.51	986.51

Crystal structure

Peptide **1b** was subjected to crystallization screening using typical protein crystallization conditions (Crystal Screen and Crystal Screen 2, Hampton Research). After screening 98 different conditions, single crystals suitable for structure determination were obtained with Crystal Screen solution 47 (0.1 M sodium acetate trihydrate pH 4.6, 2.0 M ammonium sulfate). The presence of two phenylalanines, one valine and one alanine at one face of the peptide surface may have facilitated crystallization by limiting solubility in aqueous buffers and has yielded extremely well-diffracting rod-shaped crystals. Several X-ray data sets were collected at the Swiss Light Source (Paul Scherrer Institute, Villigen) at beamlines X06DA and X06SA. The obtained model represents one of the few X-ray structures of a bicyclic homodetic peptide determined at atomic resolution, the corresponding data statistics of which are given in Table 2. Overall, the structure of peptide **1b** represents an assembly of two protomers being placed in the asymmetric unit as a homodimer. Both protomers adopt overall globular shape as predicted for bicyclic peptide scaffolds. The structure also unambiguously confirms the stereochemistry for the bridgehead residues as predicted by synthetic design. Crystal packing within the homodimer is governed predominantly by burying the hydrophobic residues of the two adjacent **1b** protomers into a hydrophobic surface patch (Figure 2).

Table 2. Crystallographic data for **1b**

Space group	P 21 21 21
a, b, c (Å)	12.82 23.42 41.92
α, β, γ (°)	90 90 90
Protomers/a.s.u[a]	2
Data collection (XDS)	
Beamlines	X06DA / X06SA
Wavelength (Å)	0.73
Resolution range (Å)	20.5 – 0.81
	(0.83-0.81)
No. observations	135390
No. unique reflections	12759
Completeness (%)[b]	95.4 (64.7)
R_{meas} (%)[b,c]	7.3 (31.2)
$I/\sigma I^2$	21.1 (4.3)
Refinement (Phenix)	
Resolution range (Å)	20.5 – 0.81
No. reflections	12720
No. reflections test set	636
R/R_{free} (%)	6.65 / 8.35
No. atoms (without H)	
Peptide	336
Additives	4
Water	25
B-factors (Å^2)	
Overall	4.94
Main Chain	2.25
Solvent and Side chains	5.64

[a] a.s.u., asymmetric unit. [b] The values in parentheses of resolution range, completeness, R_{meas} and $I/\sigma(I)$ correspond to the outermost resolution shell. [c] $R_{meas} / R_{rim} = \Sigma_{hkl}(N/N-1)^{1/2}\Sigma_j |I(hkl; j) - \langle I(hkl)\rangle|/(\Sigma_{hkl}\Sigma_j I(hkl;j))$, where I(hkl;j) is the jth measurement of the intensity of the unique reflection (hkl) and ‹I(hkl)› is the mean over all symmetry-related measurements.
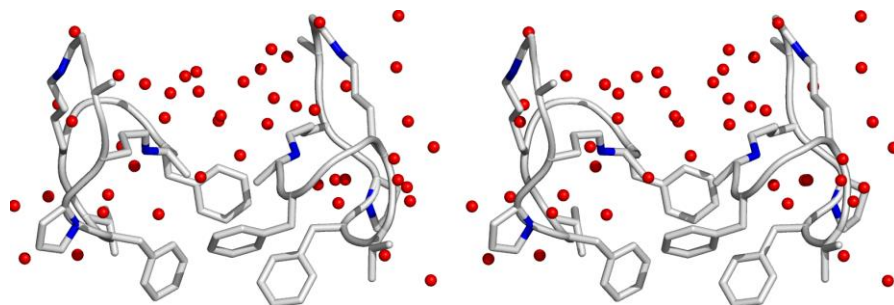
Figure 2. Stereoplot highlighting the hydrophobic interactions between the two protomers of bicyclic peptide **1b** within the unit cell. Co-crystallized waters are shown as red spheres. Note the absence of water molecules around the phenyl rings.

A stereochemical analysis of all residues of the two protomers reveals full agreement with most favorable regions in the Ramachandran plot. Accordingly, bond lengths, bond angles and dihedral angles are mostly consistent with parameter ranges of known peptide structures (Tables S1-S2). Moreover, secondary structure analysis reveals the presence of five intramolecular H-bonds strongly stabilizing the overall backbone trace of peptide **1b**. Among those, two involve hydrogen bond donor/acceptor pairs from the same peptide loop and define different β-turns. The one between εNH of D-Lys4 and CO of Glu8 defines a type II β-turn ($\varphi_9 = -58°$, $\psi_9 = 140°$, $\varphi_{10} = 83°$, $\psi_{10} = 5°$), whereas the one between NH of D-Lys4 and CO of Ala1 defines a type I' β-turn ($\varphi_2 = 55°$, $\psi_2 = 41°$, $\varphi_3 = 74°$, $\psi_3 = 1°$). The remaining three hydrogen bonds, on the other hand, cross-link different loops of the bicyclic structure: they connect, respectively, NH of Ala1 to CO of Phe6, NH of Val5 to CO of Phe2 and NH of Phe6 to CO of Ala1, which is therefore a double hydrogen bond acceptor (Figure 3).
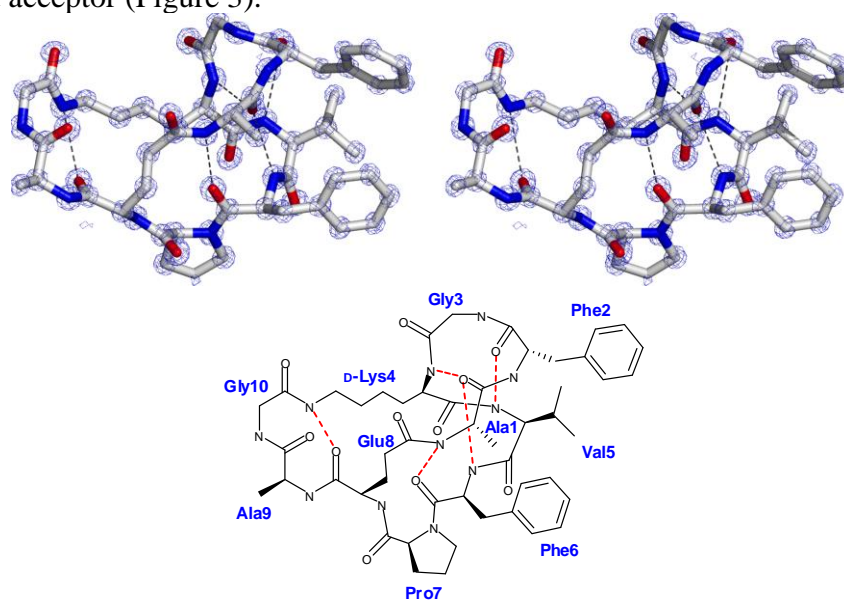
Figure 3. Stereoplot and 2D map of the crystal structure of **1b**. Hydrogen bonds donor/acceptor pairs are connected by dashed lines. $2F_o–F_c$ electron density map of **1b** contoured at 3.0 σ.

Conclusions

We have reported the synthesis and crystal structure of a bicyclic homodetic peptide. The data show that for this macrobicyclic system the expected bridge configuration is indeed formed, as planned by the chirality of the branching amino acids. Furthermore, bicyclization only took place efficiently when bridgehead amino acid chiralities were properly matched. The stereoselective bridge formation opens the way for the planned synthesis of a large family of conformationally constrained bicyclic peptides, which constitute a promising new class of scaffolds for drug design.

Experimental Section

General. Fmoc-amino acids and reagents were purchased from commercial sources. Fmoc-Gly-Wang resin (loading: 0.56 mmol/g) was purchased from Fluorochem UK. NMR data were acquired in DMSO-d_6 at 298 K using a Bruker AvanceII 400 MHz NMR spectrometer. Proton resonances were assigned using TOCSY and HSQC data. Analytical RP-HPLC was performed with an Ultimate 3000 Rapid Separation LC using an Acclaim RSLC 120 C18 column (2.2 μm, 120 Å, 3×50 mm, flow 1.2 mL/min) from Dionex. Preparative RP-HPLC was performed with a Waters Prep LC Controller System using a Dr. Maisch GmbH Reprospher column (C18-DE, 100×30 mm, particle size 5 µm, pore size 100 Å, flow 60 mL/min). The following eluents were used: **A** H_2O/TFA (100:0.1); **D** H_2O/MeCN/TFA (10:90:0.1). MS spectra were recorded on a Thermo Scientific LTQ OrbitrapXL.

Solid phase synthesis. 400 mg of Fmoc-Gly-Wang resin were placed in a polypropylene syringe fitted with a polyethylene frit, a stopcock and a stopper. The resin was swollen in DCM for 15 min. After removal of DCM, the Fmoc protecting group was removed. Stirring of the reaction mixture at any given step described below was performed by attaching the closed syringe to a rotating axis. Elongation of the peptide chain was performed using standard Fmoc-SPPS conditions. After the linear peptide chain was assembled, the polypropylene syringe was equipped with a septum and dried under vacuum for 1 hour. It was then swollen in dry DCM for 15 min under argon. After removal of the solvent, Pd(PPh$_3$)$_4$ (0.1 eq) was diluted in 5 mL of dry DCM and added to the resin under argon. Phenylsilane (10 eq) was then added to the resin. The reaction was stirred under argon bubbling for 40 min. The reagents were removed by filtration and the resin washed with dry DCM. The resin was treated with sodium diethylthiocarbamate (20 mM in DMF, 5 mL) for 15 min to remove metal traces and finally washed with NMP, MeOH and DCM. The last Fmoc protecting group was removed with DMF/piperidine (4:1).

On-resin cyclization. HATU (3 eq) and DIPEA (6 eq) were added to the deprotected peptidyl-resin in 6 mL of NMP/DMSO (4:1) and the mixture was stirred at room temperature for 8 h. The reagents were removed by filtration and the resin washed with NMP, MeOH and DCM.

TFA cleavage and purification. The cleavage was carried out using a TFA/TIS/H_2O (95:4:1) solution for 2 h. The peptide solution was separated from the resin by filtration, evaporated and dried under high vacuum and finally lyophilized from a freshly-prepared water/acetonitrile (4:1) solution. The lyophilized crude was then dissolved in a water/acetonitrile mixture, purified by preparative RP-HPLC and lyophilized again. Yields were calculated for the TFA salts of **1a** and **1a'**.

Second cyclization. 1b was synthesized by subjecting **1a** to amide bond formation conditions in high dilution (0.7 mm). To a solution of PyBOP (3 eq) and DIEA (6 eq) in DCM/DMF (10:1), a 10 mm solution of **1a** in DMF was slowly added dropwise at 0 °C under stirring. The reaction was left stirring for 16 h at r.t. and monitored via RP-HPLC, until complete disappearance of the starting material. The reaction was then quenched by evaporation of the solvent. The crude was then dissolved in a water/acetonitrile mixture, purified by preparative RP-HPLC and lyophilized. The yield was calculated considering the trifluoroacetate salt of **1a** as the starting material.

X-ray diffraction experiments. The peptide was dissolved at 20 mg/mL in water and screened for suitable crystallization conditions using commercially available crystallization kits (Hampton Research; Aliso Viejo, CA, USA). Colorless rod-shaped crystals were obtained overnight by sitting drop vapor diffusion after mixing equal volumes of compound solution with 0.1 M sodium acetate trihydrate pH 4.6, 2.0 M ammonium sulphate at 18 °C. Diffraction data were collected at 100 K at the Swiss Light Source (Paul Scherrer Institute, Villigen) at beamlines X06DA and X06SA. Data processing, scaling and merging was performed by XDS and XSCALE[19]. The structure was determined by direct methods using SHELXD[20]. Structure refinement and model building was performed by iterative cycles of phenix.refine without imposing any stereochemical restrains in the last steps[21] and Coot[22].

cyclo(Glu(Ala-Phe-Gly-lys-Val-Phe-Pro))-Ala-Gly-OH (**1a, A[1]FGkVFPE[1]AG).** Foamy white powder (67.5 mg, 60.4 µmol, 27%). Analytical RP-HPLC: r.t. = 1.765 min (A/D 100:0 to 0:100 in 2.2 min, λ = 214 nm). MS (ESI+): $C_{49}H_{69}N_{11}O_{12}$ calc./obs. 1004.52/1004.52 $[M+H]^+$, 502.76/502.76 $[M+2H]^{2+}$.

cyclo(Glu(Ala-Phe-Gly-Lys-Val-Phe-Pro))-Ala-Gly-OH (**1a', A[1]FGKVFPE[1]AG).** Foamy white powder (43.8 mg, 39.2 µmol, 17%). Analytical RP-HPLC: r.t. = 1.704 min (A/D 100:0 to 0:100 in 2.2 min, λ = 214 nm). MS (ESI+): $C_{49}H_{69}N_{11}O_{12}$ calc./obs. 1004.52/1004.52 $[M+H]^+$, 502.76/502.76 $[M+2H]^{2+}$.

cyclo(Ala[1]Phe[2]Gly[3]lys[4]Val[5]Phe[6]Pro[7]Glu[8])cyclo(8α→4ε)Ala[9]Gly[10] (1b, A[1]FG*k*[2]VFP*E*[1]AG[2]). Foamy white powder (22.8 mg, 23.1 µmol, 38%). Analytical RP-HPLC: r.t. = 3.91 min (A/D 100:0 to 0:100 in 7.5 min, λ = 214 nm). MS (ESI+): $C_{49}H_{67}N_{11}O_{11}$ calc./obs. 986.51/986.51 $[M+H]^+$, 1008.49/1008.49 $[M+Na]^+$. See supporting information for full ^{1}H-NMR characterization.

Acknowledgements

References

1. Reymond, J.-L.; Blum, L. C.; van Deursen, R. *Chimia* **2011**, *65*, 863.
    http://dx.doi.org/10.2533/chimia.2011.863
    PMid:22289373
2. Ruddigkeit, L.; Van Deursen, R.; Blum, L. C.; Reymond, J.-L. *J. Chem. Inf. Model.* **2012**, *52*, 2864.
    http://dx.doi.org/10.1021/ci300415d
    PMid:23088335
3. Robinson, J. A. *Chembiochem* **2009**, *10*, 971.
    http://dx.doi.org/10.1002/cbic.200900055
    PMid:19266524
4. Rubinstein, M.; Niv, M. Y. *Biopolymers* **2009**, *91*, 505.
    http://dx.doi.org/10.1002/bip.21164
    PMid:19226619
5. Stevenson, C. L. *Curr. Pharm. Biotechnol.* **2009**, *10*, 122.
    http://dx.doi.org/10.2174/138920109787048634
    PMid:19149594
6. Karskela, T.; Virta, P.; Lonnberg, H. *Curr. Org. Synth.* **2006**, *3*, 283.
    http://dx.doi.org/10.2174/157017906777934917
7. Cascales, L.; Craik, D. J. *Org. Biomol. Chem.* **2010**, *8*, 5035.
    http://dx.doi.org/10.1039/c0ob00139b
    PMid:20835453
8. Chatterjee, C.; Paul, M.; Xie, L.; van der Donk, W. A. *Chem. Rev.* **2005**, *105*, 633.
    http://dx.doi.org/10.1021/cr030105v
    PMid:15700960
9. Merrifield, B. *Methods Enzymol.* **1997**, *289*, 3.
    PMid:9353714
10. Reymond, J.-L.; Darbre, T. *Org. Biomol. Chem.* **2012**, *10*, 1483.
    http://dx.doi.org/10.1039/c2ob06938e
    PMid:22274649

11. Uhlich, N. A.; Darbre, T.; Reymond, J.-L. *Org. Biomol. Chem.* **2011**, *9*, 7071.
    http://dx.doi.org/10.1039/c1ob05877k
    PMid:21879129
12. Geotti-Bianchini, P.; Darbre, T.; Reymond, J.-L. *Org. Biomol. Chem.* **2013**, *11*, 344.
    http://dx.doi.org/10.1039/c2ob26551f
    PMid:23172354
13. Kadam, R. U.; Garg, D.; Schwartz, J.; Visini, R.; Sattler, M.; Stocker, A.; Darbre, T.;
    Reymond, J.-L. *ACS Chem. Biol.* **2013**, *8*, 1925.
    http://dx.doi.org/10.1021/cb400303w
    PMid:23869965
14. Ravi, H. K.; Stach, M.; Soares, T. A.; Darbre, T.; Reymond, J.-L.; Cascella, M. *Chem. Commun.* **2013**, *49*, 8821.
    http://dx.doi.org/10.1039/c3cc44912b
    PMid:23959139
15. Kwok, A.; Eggimann, G. A.; Reymond, J.-L.; Darbre, T.; Hollfelder, F. *ACS Nano* **2013**, *7*, 4668.
    http://dx.doi.org/10.1021/nn400343z
    PMid:23682947 PMCid:PMC3715887
16. Bartoloni, M.; Kadam, R. U.; Schwartz, J.; Furrer, J.; Darbre, T.; Reymond, J.-L. *Chem. Commun.* **2011**, *47*, 12634.
    http://dx.doi.org/10.1039/c1cc15704c
    PMid:22031227
17. Wareham, R. S.; Kilburn, J. D.; Turner, D. L.; Rees, N. H.; Holmes, D. S. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 2660.
    http://dx.doi.org/10.1002/anie.199526601
18. Di Blasio, B.; Benedetti, E.; Pavone, V.; Pedone, C.; Saviano, M.; Zanotti, G.; Blout, E. R. *Biopolymers* **1990**, *30*, 509.
    http://dx.doi.org/10.1002/bip.360300504
    PMid:2265225
19. Kabsch, W. *Acta Crystallogr. Sect. D - Biol. Crystallogr.* **2010**, *66*, 125.
20. Sheldrick, G. M. *Acta Crystallogr. Sect. A* **2008**, *64*, 112.
    http://dx.doi.org/10.1107/S0108767307043930
    PMid:18156677
21. Adams, P. D *et al. Acta Crystallogr. Sect. D - Biol. Crystallogr.* **2010**, *66*, 213.
22. Emsley, P.; Lohkamp, B.; Scott, W. G. and Cowtan, K. *Acta Crystallogr. Sect. D - Biol. Crystallogr.* **2010**, *66*, 486.