Supporting Information

Synthesis and characterization of water-soluble macrocyclic peptides stabilizing protein α-turn

Lei Wang,[a] Pascale Coric,[b] Kexin Zhu,[a] Wang-Qing Liu,[a] Michel Vidal,[a,c] Serge Bouaziz,[b,*] Sylvain Broussy[a,*]

[a] UMR COMETE 8638 CNRS, Université Paris Descartes, Faculté de Pharmacie de Paris, Sorbonne Paris Cité, 4 av. de l'Observatoire, Paris 75006, France
E-mail: sylvain.broussy@parisdescartes.fr
[b] UMR 8015 CNRS, Université Paris Descartes, Faculté de Pharmacie de Paris, Sorbonne Paris Cité 4 av. de l'Observatoire, Paris 75006, France
E-mail: serge.bouaziz@parisdescartes.fr
[c] UF Pharmacocinétique et Pharmacochimie Hôpital Cochin, AP-HP 27 rue du Faubourg Saint Jacques, Paris 75014, France
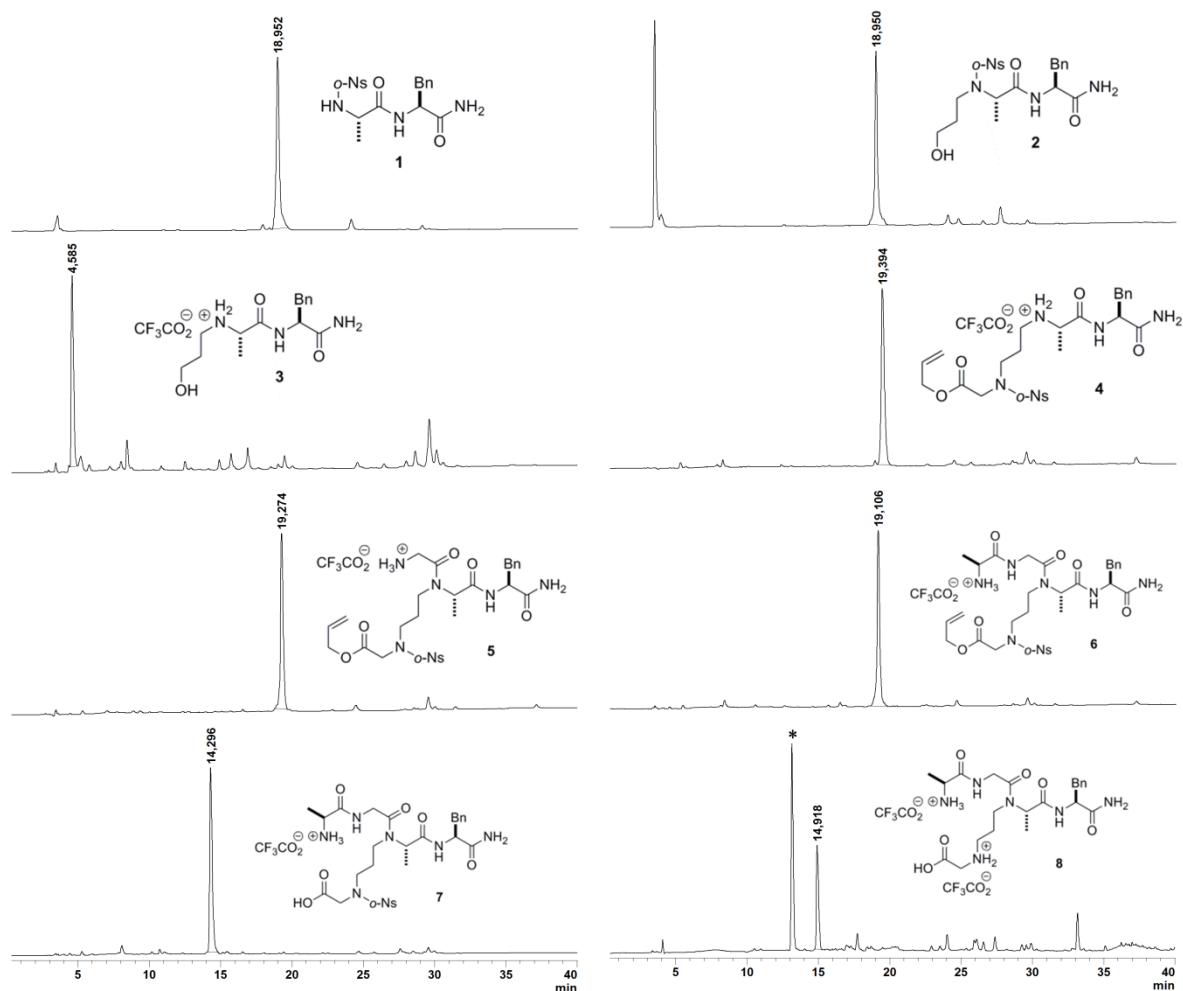
Table of Contents

HPLC chromatograms of peptides **1-8a** (AA$_1$ = AA$_3$ = Gly).

See the experimental section for details. Gradient conditions: 30-100% B in 30 min for **1-7** and 15-50% B in 30 min for **8a**.

*Denotes a non-peptide side product.

HPLC chromatograms of peptides **9a**, **10a**, **10b** and **10c**.

See the experimental section for details. Gradient conditions: 30-100% in 30 min for **9a**, **10b** and **10c**; 15-50% B in 30 min for **10a**.
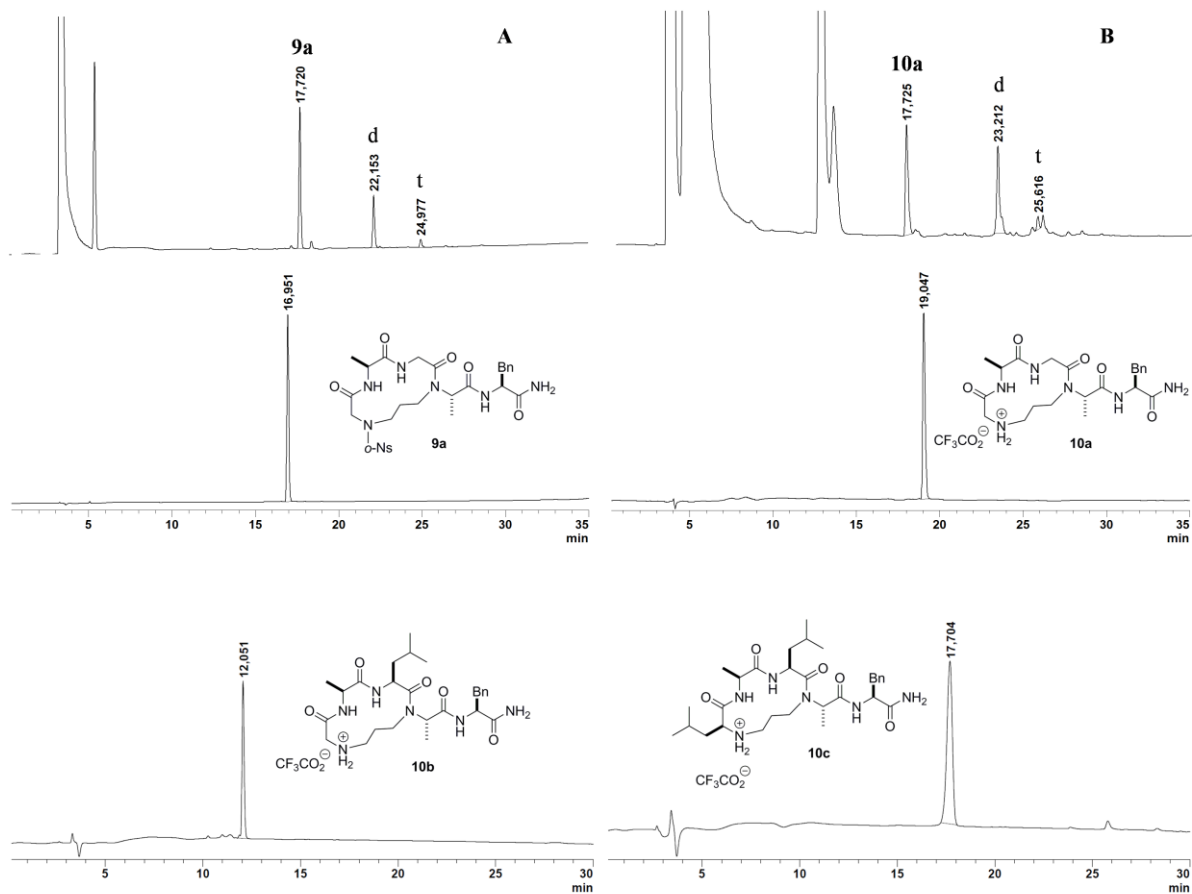
A and **B** are crude chromatograms of cyclization reactions of 10 mM solutions of **7** and **8** respectively. d: dimer; t: trimer.

CD spectra of peptides **10a** and **11** in 10 mM phosphate buffer pH 7.4 without (blue and black) and with (red and green) 30% TFE, at 293 K.
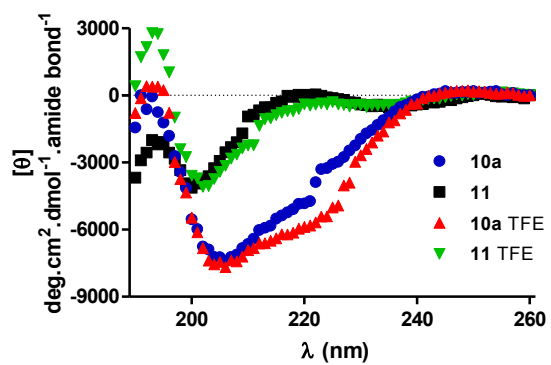
NMR spectra of peptide **9a**.
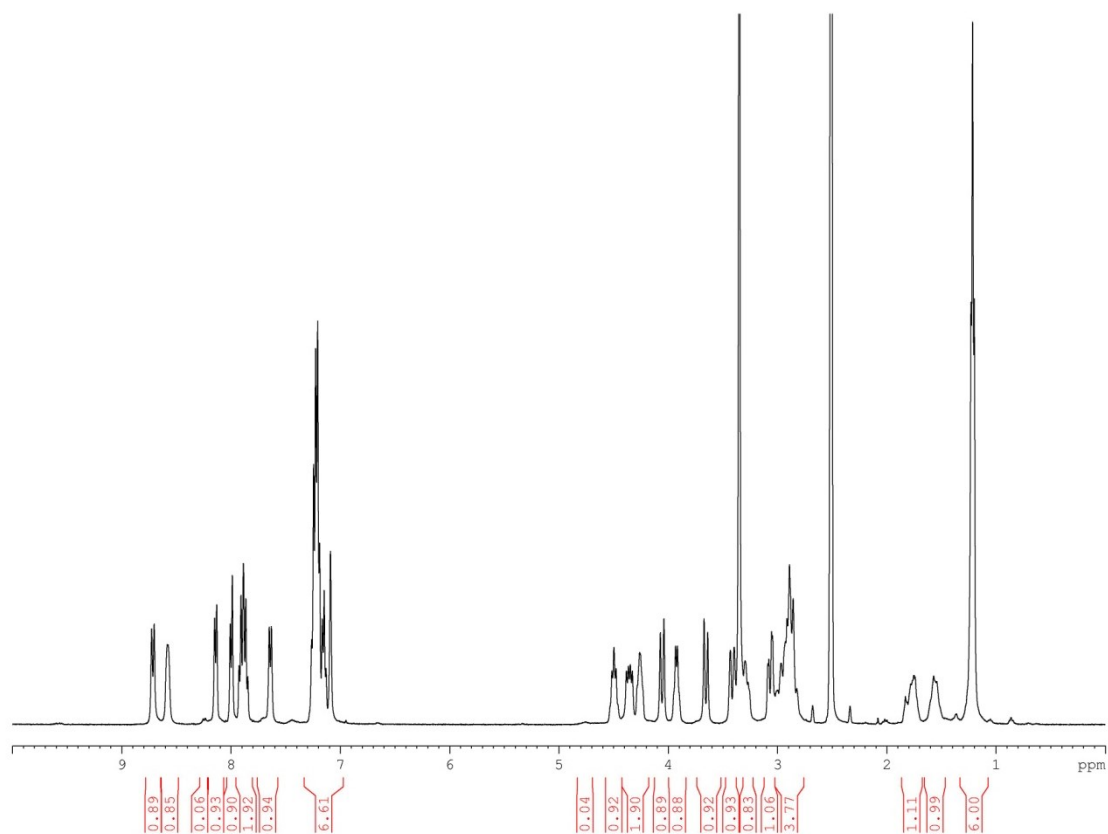
¹H NMR spectrum of peptide **9a** (400 MHz, *d6*-DMSO).

COSY spectrum of peptide **9a** (¹H, 400 MHz, *d6*-DMSO).

ROESY spectrum of peptide **9a** (^{1}H, 400 MHz, *d6*-DMSO).
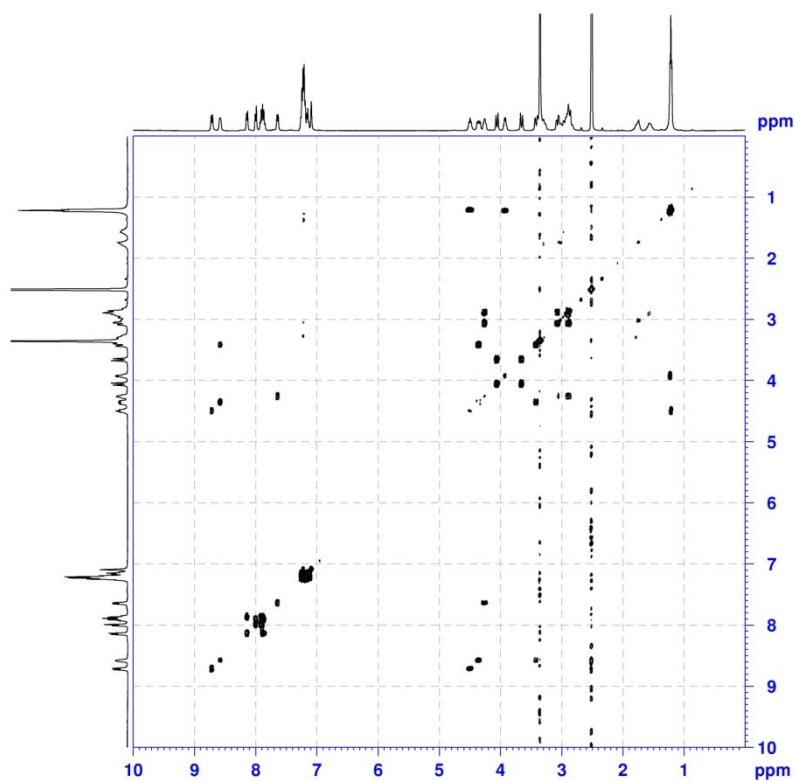
Attribution table for peptide **9a** (chemical shifts extracted from the ROESY spectrum).

	NH	H$_\alpha$	H$_\beta$	H$_{arom}$	C$_{ter}$–NH$_2$	1,1'	2,2'	3,3'
G1 **(*o*-Ns)**	-	3.66 4.05	-	7.86 7.91 8.00 8.14	-	-	-	-
A2	8.72	4.50	1.20	-	-	-	-	-
G3	8.58	3.41 4.35	-	-	-	-	-	-
A4	-	3.92	1.22	-	-	2.89 3.28	1.55 1.74	2.95 3.02
F5	7.64	4.26	2.88 3.06	7.13-7.26	7.09	-	-	-

9a
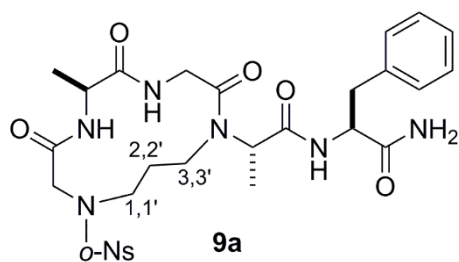
NMR spectra of peptide **10a**.
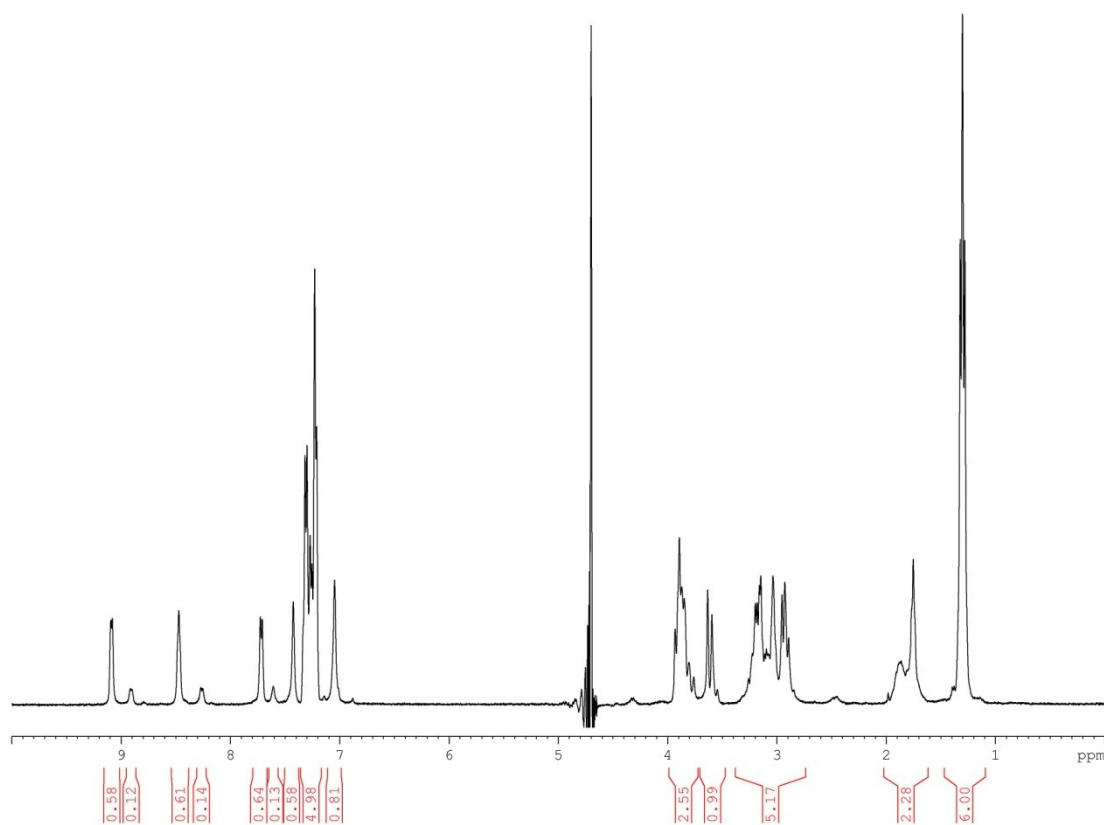
^{1}H NMR spectrum of peptide **10a** (400 MHz, H_2O/D_2O 9:1).
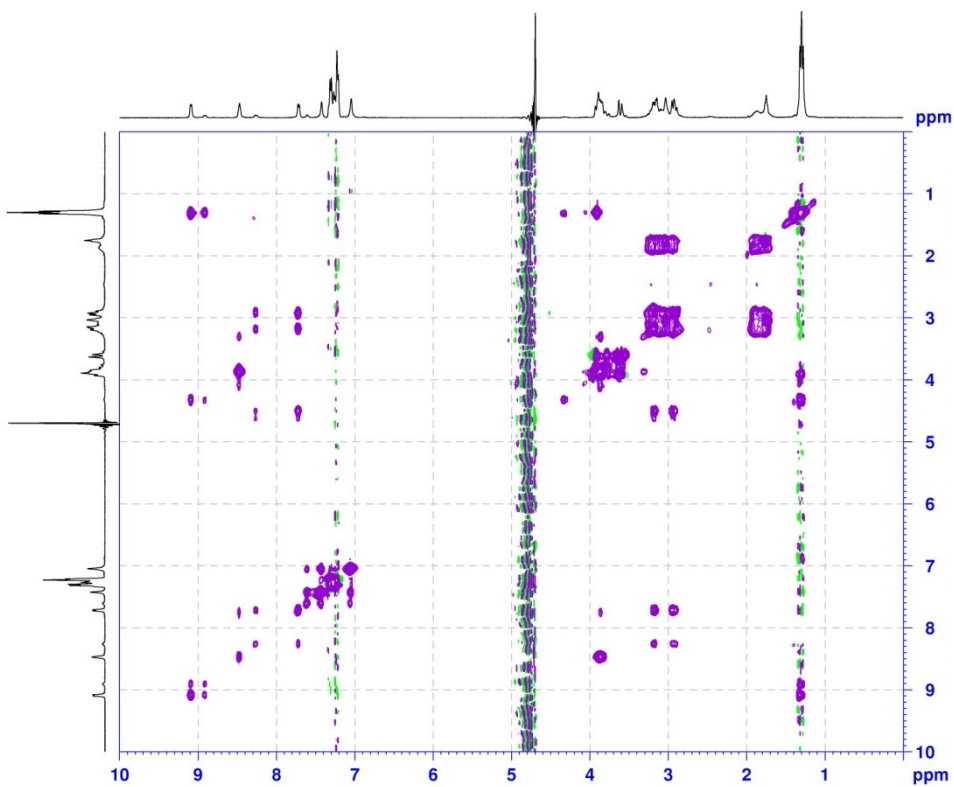
TOCSY spectrum of peptide **10a** (^{1}H, 400 MHz, H_2O/D_2O 9:1).
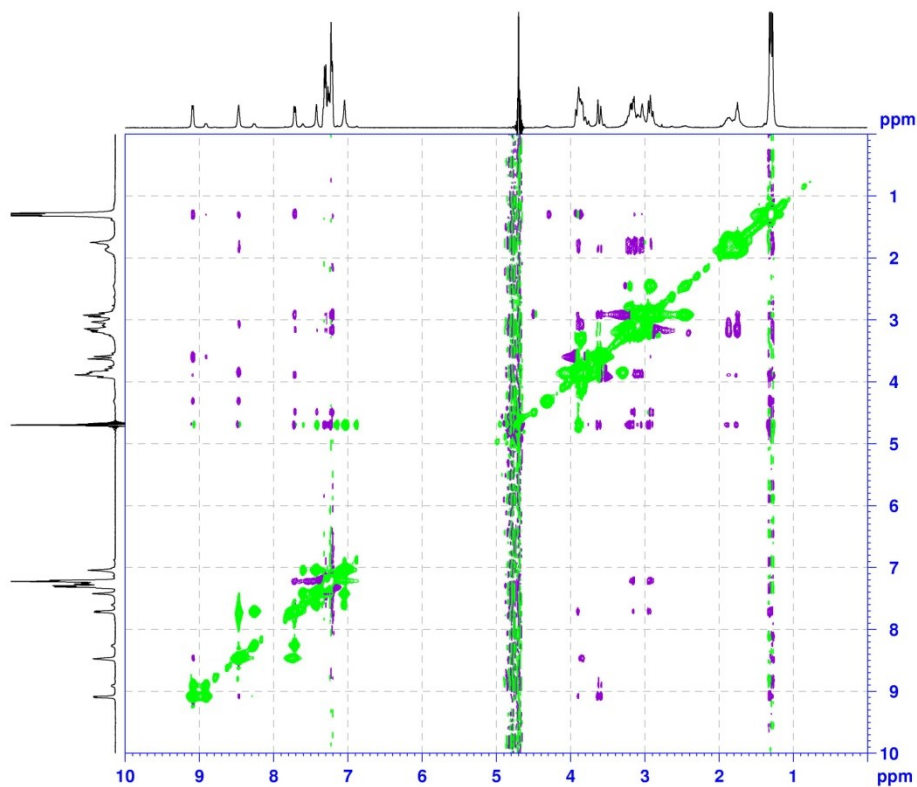
NOESY spectrum of peptide **10a** (^{1}H, 400 MHz, H_2O/D_2O 9:1).

Attribution table for peptide **10a** (major rotamer, chemical shifts extracted from the NOESY spectrum).

	NH	H$_\alpha$	H$_\beta$	H$_{arom}$	C$_{ter}$-NH$_2$	1,1'	2,2'	3,3'
G1	-	3.60 3.89	-	-	-	-	-	-
A2	9.09	4.31	1.30	-	-	-	-	-
G3	8.47	3.84	-	-	-	-	-	-
A4	-	3.89	1.29	-	-	2.85-3.25	1.75 1.86	2.95-3.15
F5	7.71	4.48	2.91 3.16	7.21-7.33	7.04 7.42	-	-	-
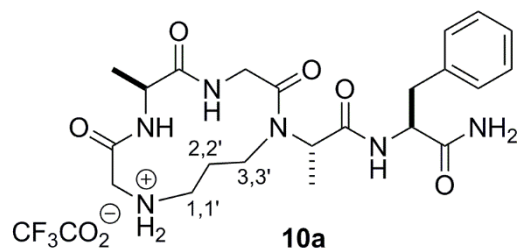
NMR spectra of peptide **10b**.
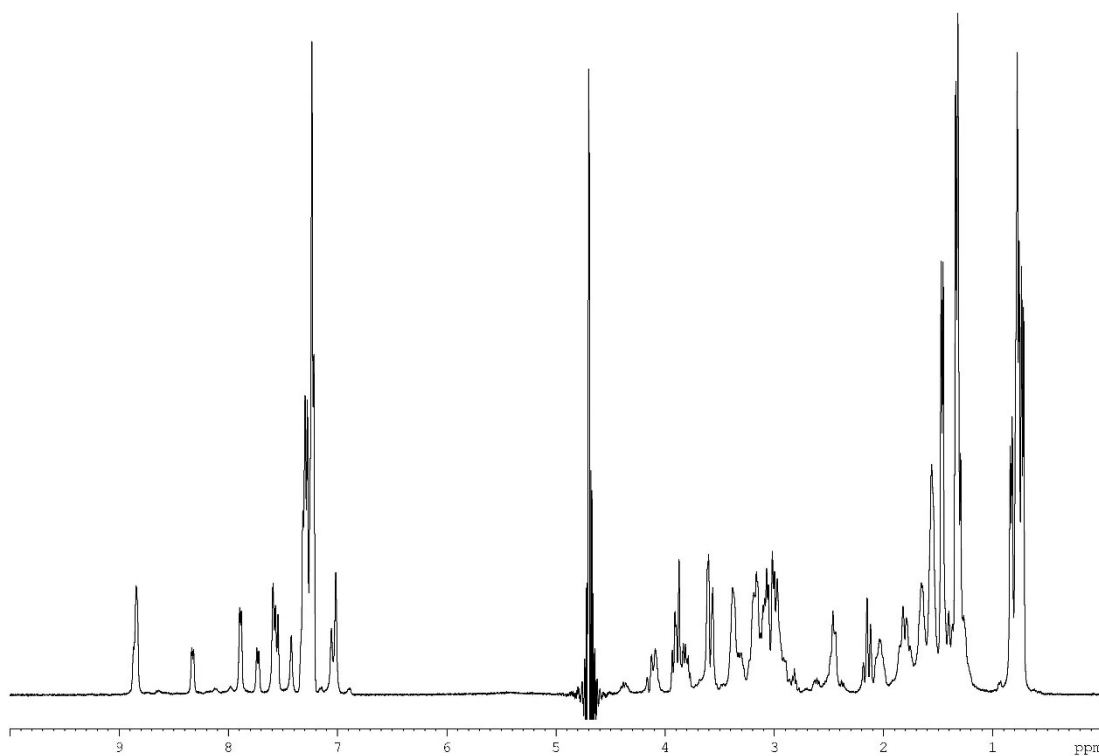
¹H NMR spectrum of peptide **10b** (400 MHz, H_2O/D_2O 9:1).
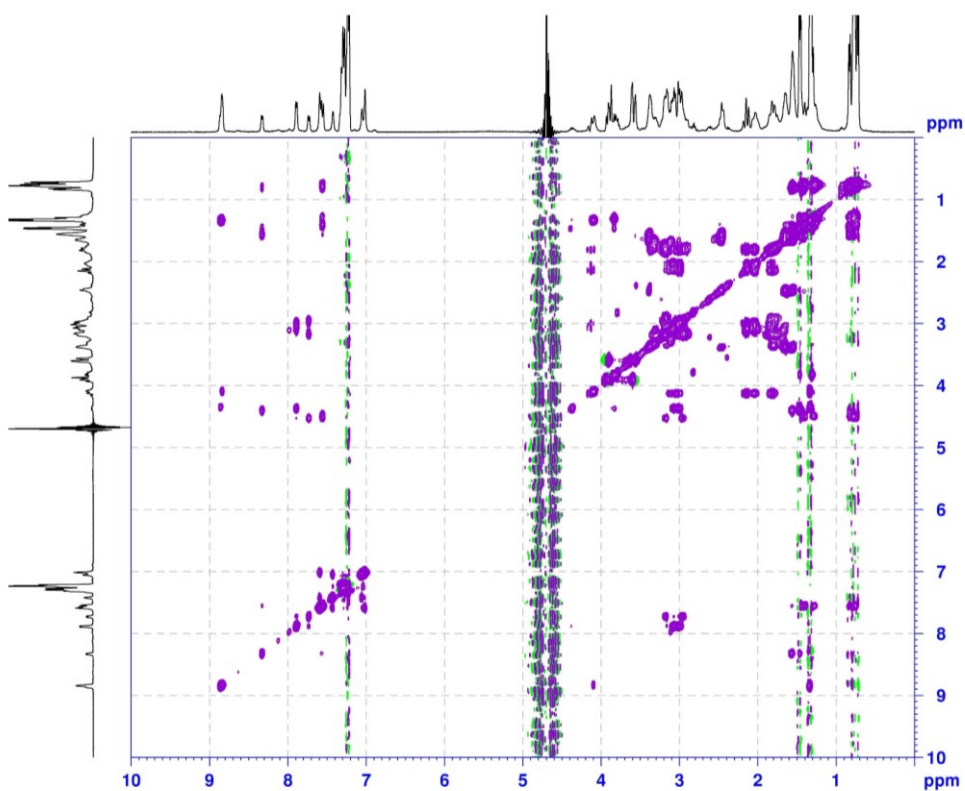
TOCSY spectrum of peptide **10b** (¹H, 400 MHz, H_2O /D_2O 9:1).
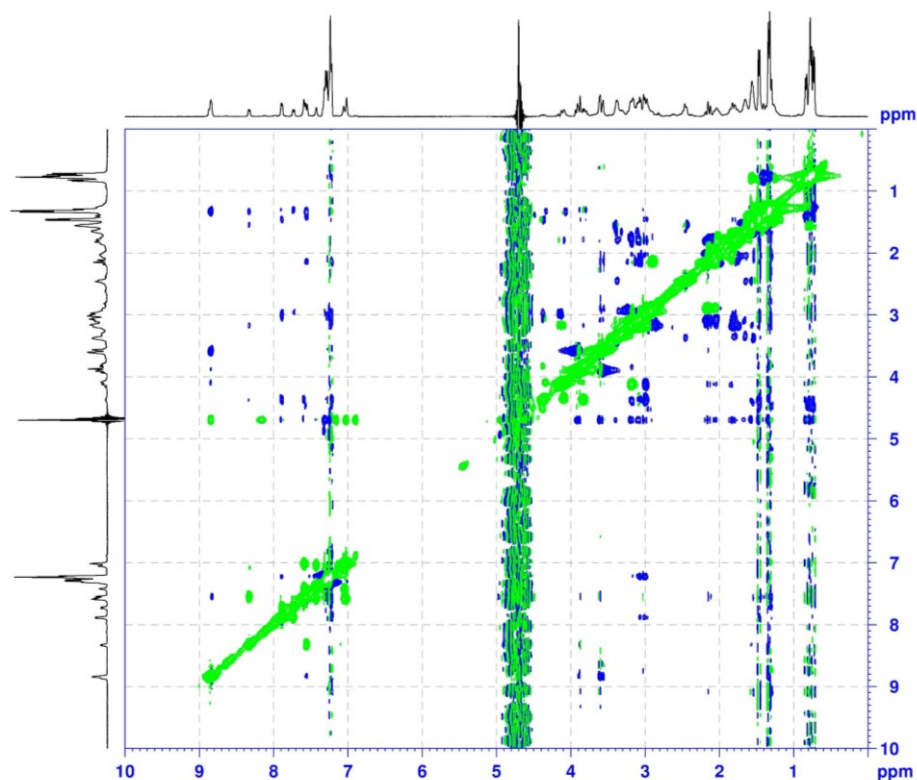
NOESY spectrum of peptide **10b** (^{1}H, 400 MHz, H_2O/D_2O 9:1).

Attribution table for peptide **10b** (major rotamer, chemical shifts extracted from the NOESY spectrum).

	NH	H$_\alpha$	H$_\beta$	H$_\gamma$	H$_\delta$	H$_{arom}$	C$_{ter}$-NH$_2$	1,1'	2,2'	3,3'
G1	-	3.58 3.88	-	-	-	-	-	-	-	-
A2	8.84	4.09	1.33	-	-	-	-	-	-	-
L3	7.56	4.47	1.39	1.27	0.78	-	-	-	-	-
A4	-	3.83	1.30	-	-	-	-	2.14 ?	1.81 2.03	2.88-3.36
F5	7.89	4.36	3.01 3.09	-	-	7.21-7.34	7.02 7.59	-	-	-

10b
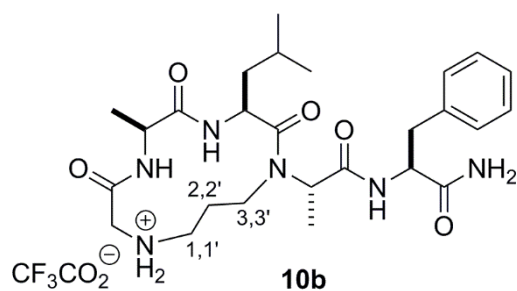
NMR spectra of peptide **10c**.
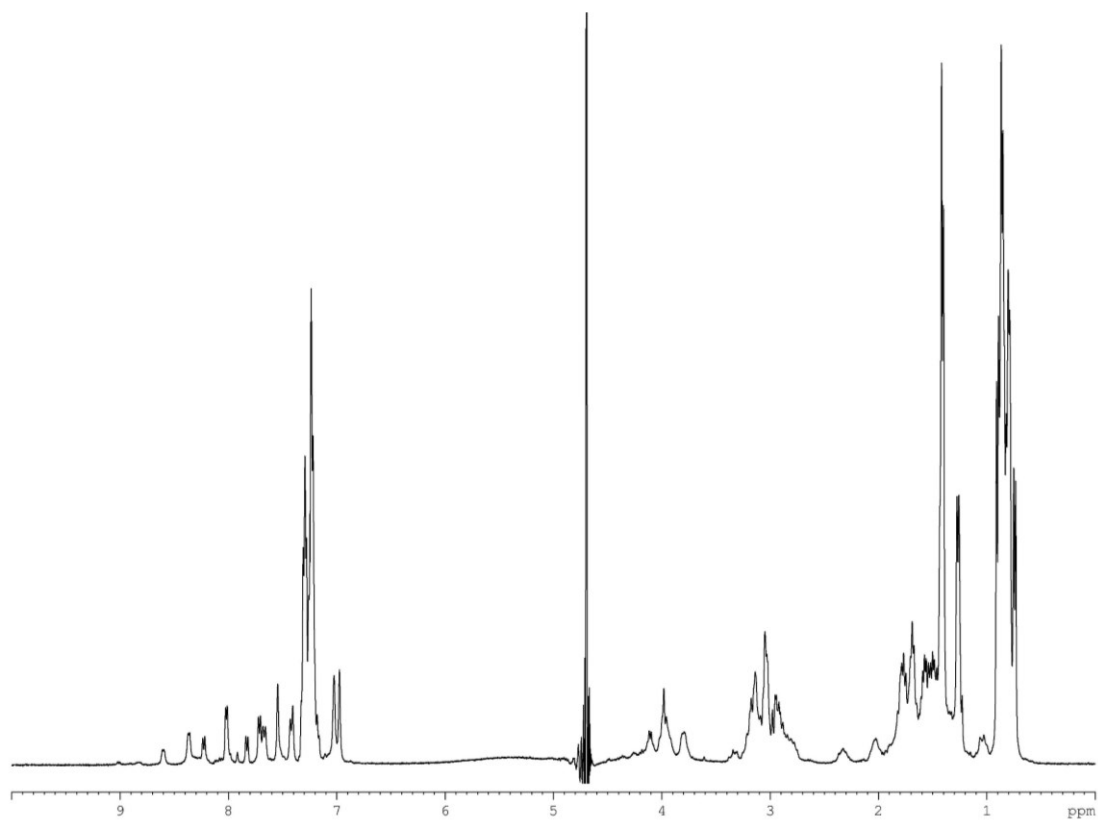
^{1}H NMR spectrum of peptide **10c** (400 MHz, H$_2$O/D$_2$O 9:1).
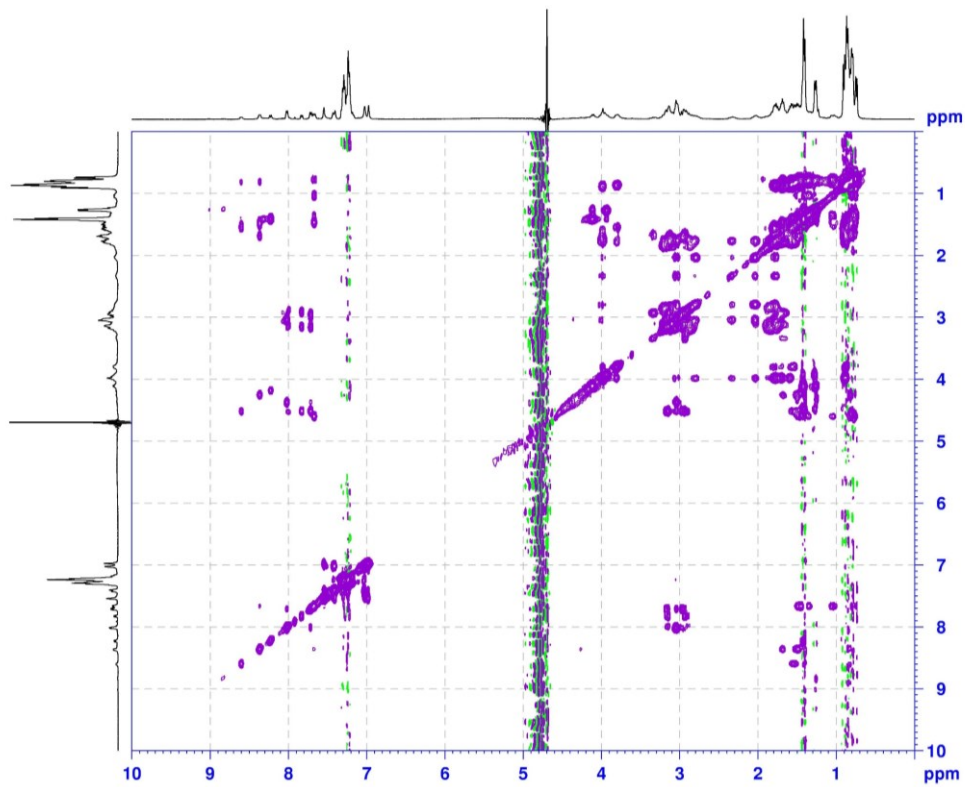
TOCSY spectrum of peptide **10c** (^{1}H, 400 MHz, H$_2$O/D$_2$O 9:1).
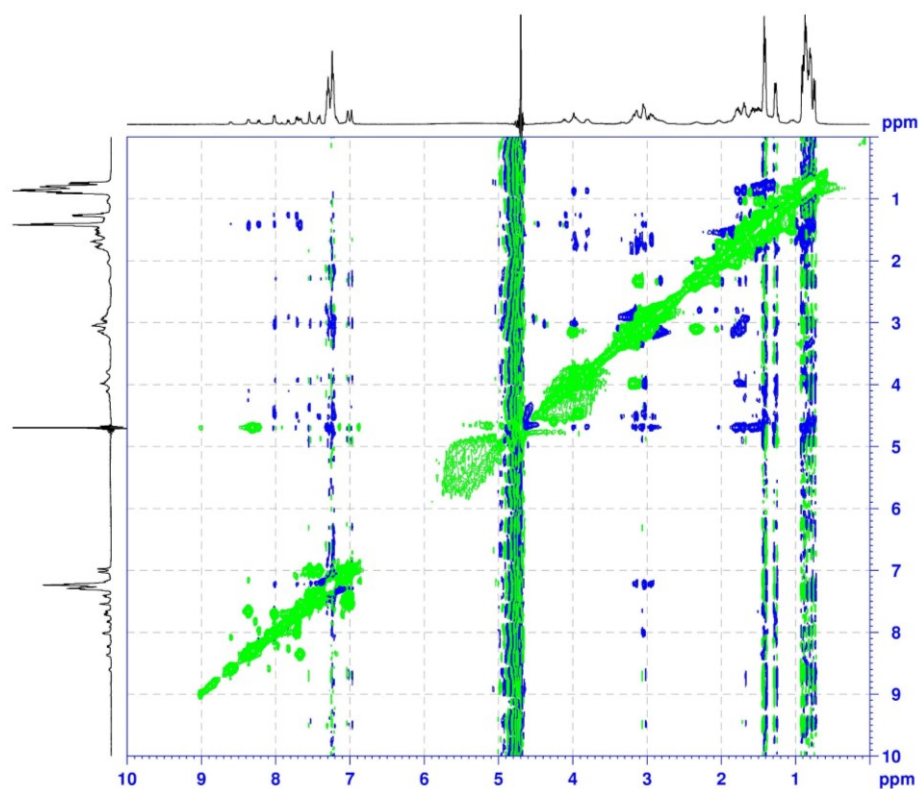
NOESY spectrum of peptide **10c** (^{1}H, 400 MHz, H_2O/D_2O 9:1).

Attribution table for peptide **10c** (major rotamer, chemical shifts extracted from the NOESY spectrum).

	NH	**H$_\alpha$**	**H$_\beta$**	**H$_\gamma$**	**H$_\delta$**	**H$_{arom}$**	**C$_{ter}$-NH$_2$**	**1,1'**	**2,2'**	**3,3'**
L1	-	4.05	1.66 1.75	?	0.94	-	-	-	-	-
A2	8.30	4.24	1.47	-	-	-	-	-	-	-
L3	8.67	4.58	1.57 1.63	1.46	0.88	-	-	-	-	-
A4	-	4.01	1.33	-	-	-	-	3.10 4.05	1.84 2.10	2.38 2.85
F5	7.78	4.58	3.02 3.24	-	-	7.34 7.52	-	-	-	-
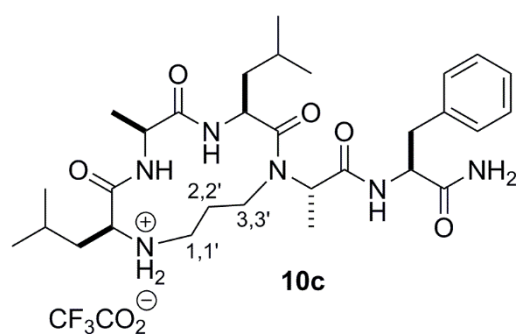
NMR spectra of peptide **11**.
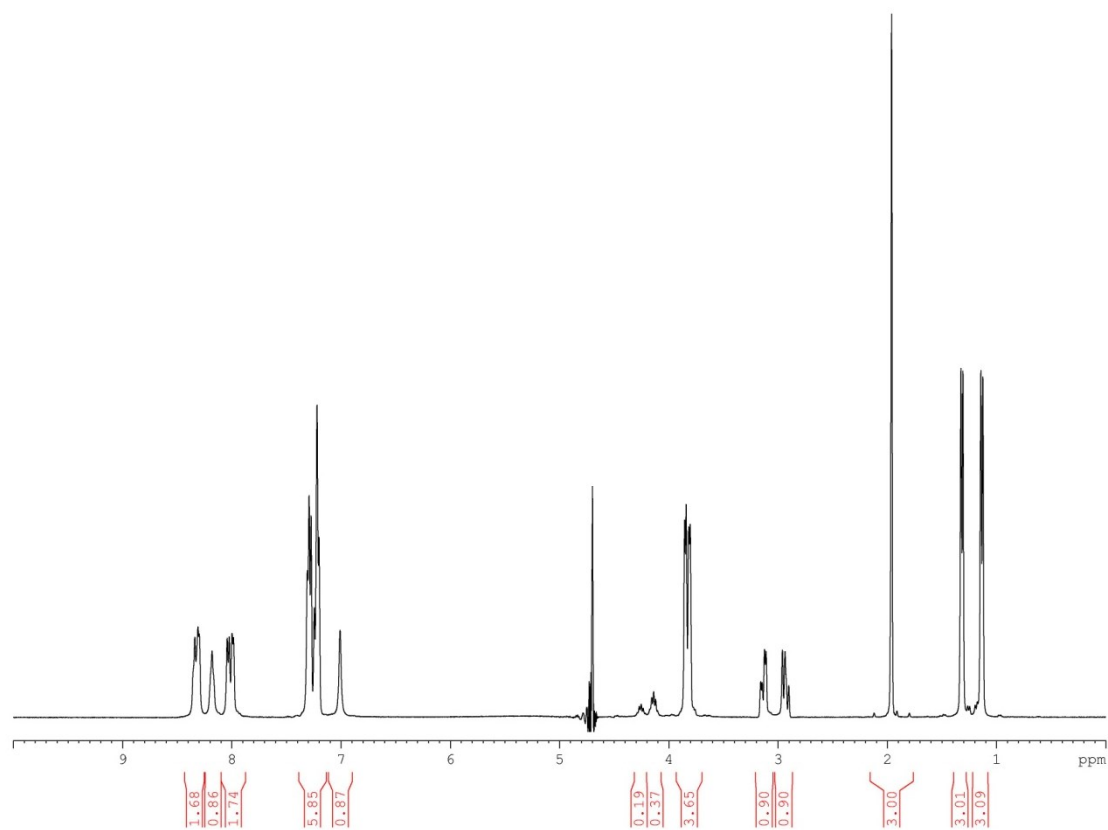
^{1}H NMR spectrum of peptide **11** (400 MHz, H$_2$O/D$_2$O 9:1).
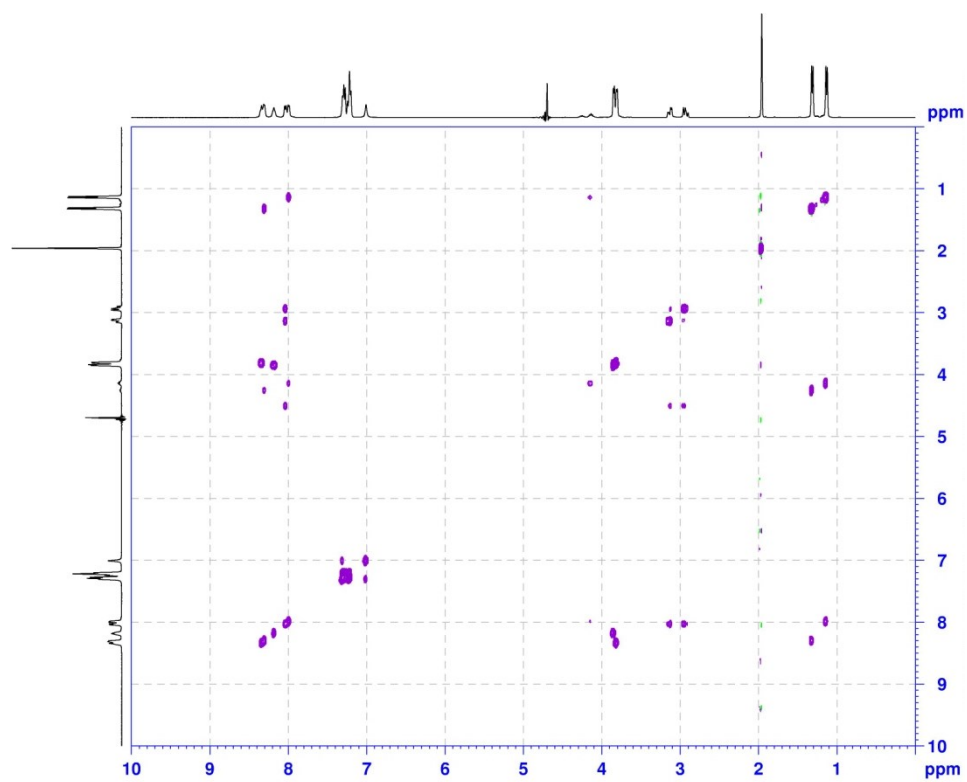
TOCSY spectrum of peptide **11** (^{1}H, 400 MHz, H$_2$O/D$_2$O 9:1).
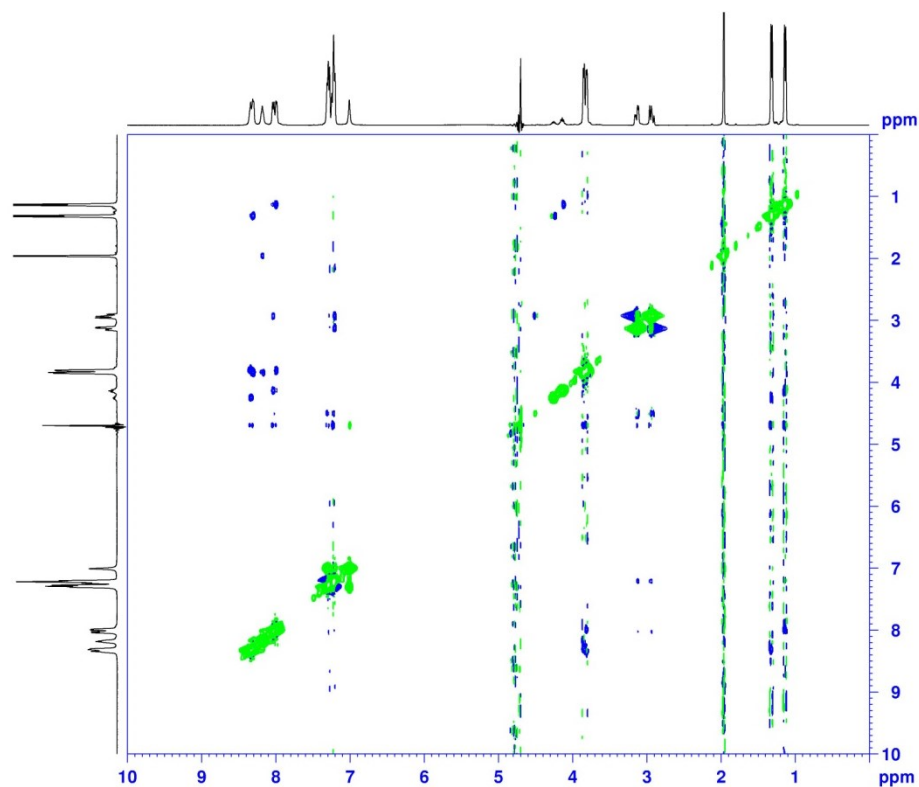
NOESY spectrum of peptide **11** (^{1}H, 400 MHz, H_2O/D_2O 9:1).

Attribution table for peptide **11**, Ac-GAGAF-NH_2 (chemical shifts extracted from the NOESY spectrum).

	NH	H_α	H_β	H_{arom}	C_{ter}-NH_2	N_{ter}-Ac
G1	8.18	3.85	-	-	-	1.96
A2	8.30	4.24	1.32	-	-	-
G3	8.34	3.80	-	-	-	-
A4	7.99	4.13	1.13	-	-	-
F5	8.03	4.50	2.93 3.12	7.20-7.31	7.01	-

Variable temperature ^{1}H NMR data for peptides **10a-c** and **11**.
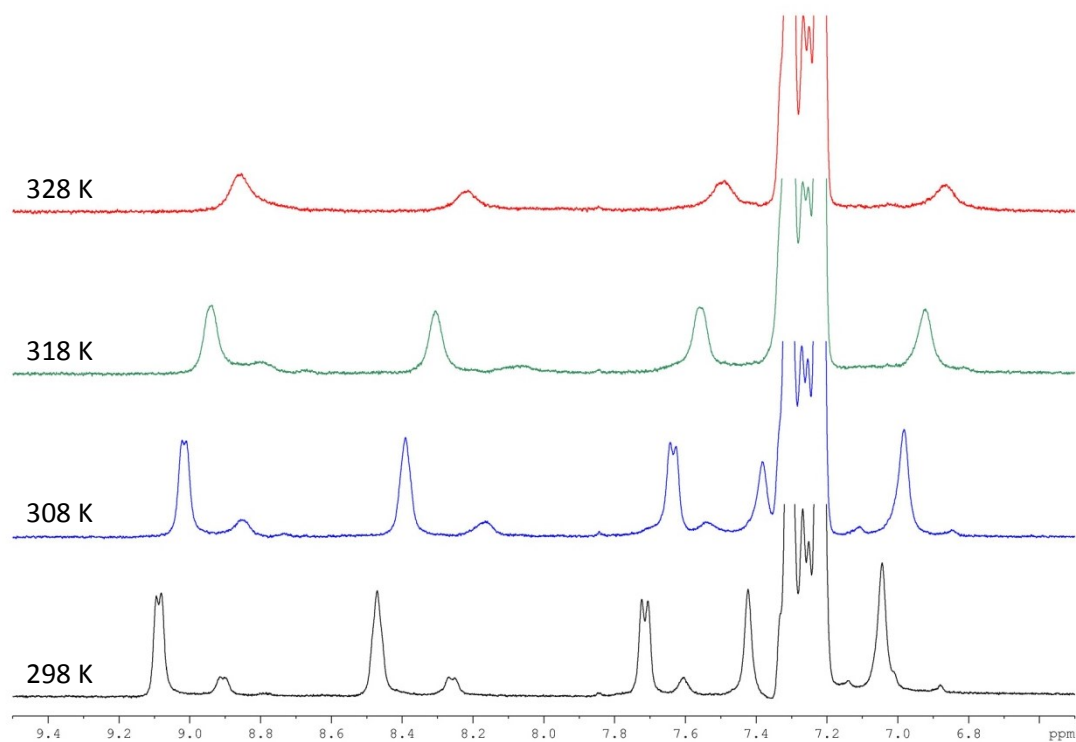
Amide region of the ^{1}H NMR spectra of peptide **10a** at the indicated temperatures (400 MHz, H_2O/D_2O 9:1).
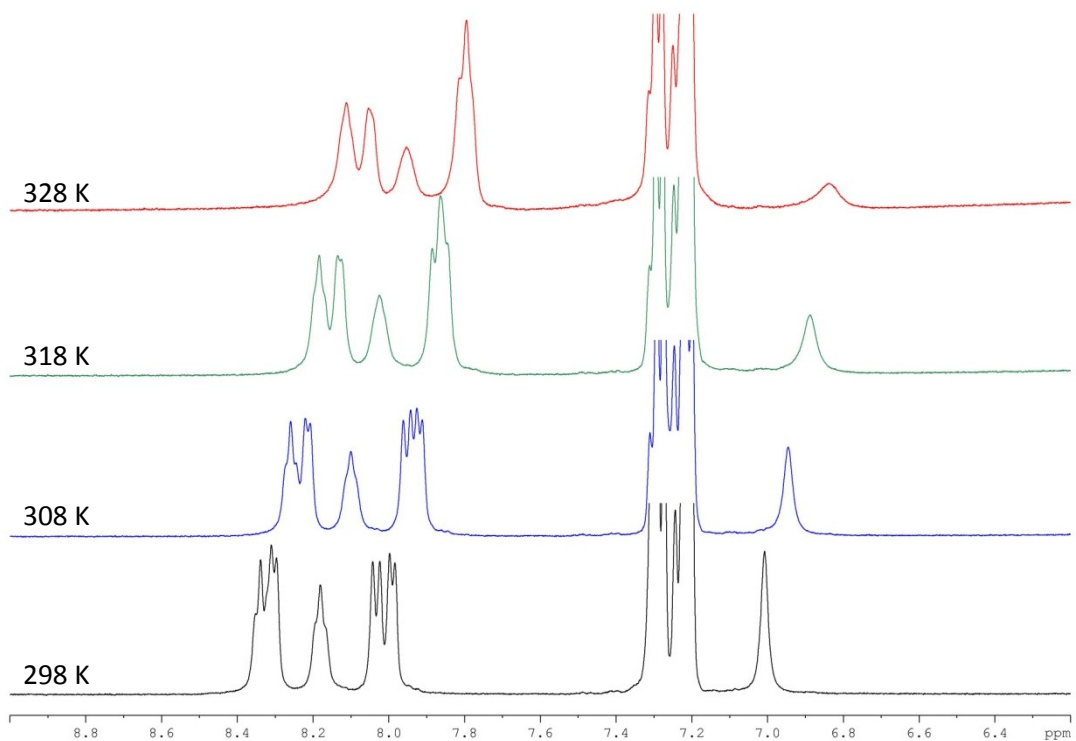
Amide region of the ^{1}H NMR spectra of peptide **11** at the indicated temperatures (400 MHz, H_2O/D_2O 9:1).
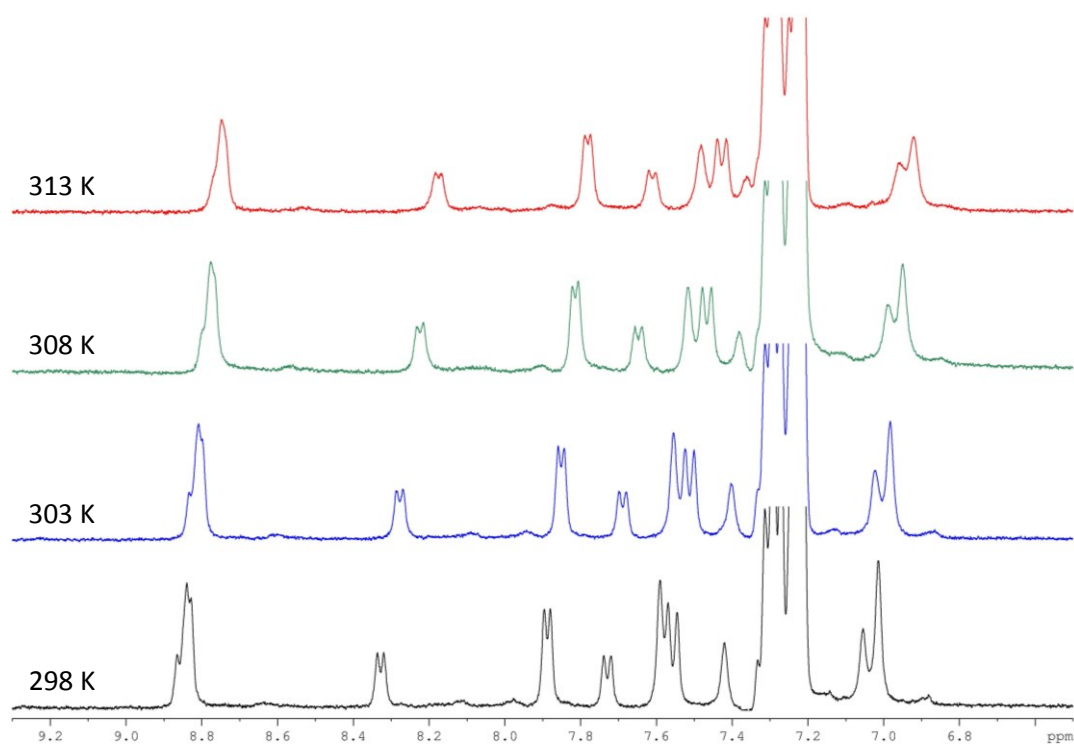
Amide region of the ^{1}H NMR spectra of peptide **10b** at the indicated temperatures (400 MHz, H_2O/D_2O 9:1).
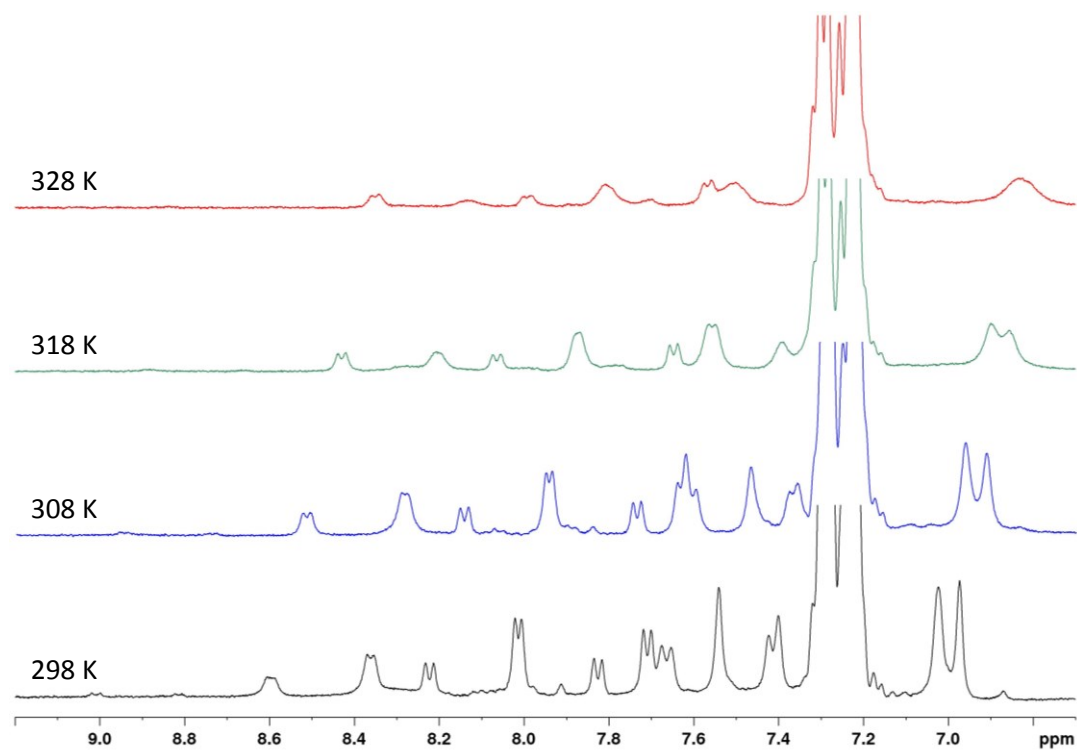
Amide region of the ^{1}H NMR spectra of peptide **10c** at the indicated temperatures (400 MHz, H_2O/D_2O 9:1).

Graphical representation of the dependence of amide protons chemical shifts on temperature for peptides **11** (left) and **10a** (right, major rotamer only).
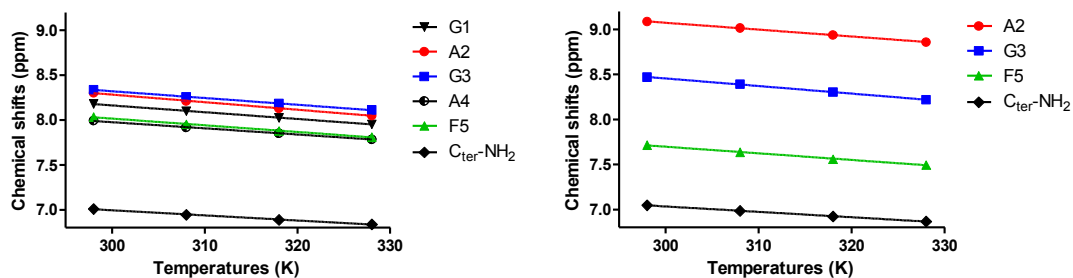
Corresponding slope values with standard errors:

Peptide 11	Slope (ppb/K)		Peptide 10a	Slope (ppb/K)
G1	-7.6 ± 0.2		A2	-7.6 ± 0.2
A2	-8.4 ± 0.2		G3	-8.4 ± 0.2
G3	-7.6 ± 0.2		F5	-7.4 ± 0.3
A4	-6.8 ± 0.2		C_{ter}-NH_2	-6.0 ± 0.1
F5	-7.4 ± 0.2			
C_{ter}-NH_2	-5.7 ± 0.2			

NOE derived distance restraints used for calculating the solution structure of the major rotamer of **10a**. Values were calculated using the CcpNmr software.[1]

GAGAF_RS5modif_run4A1		Distances (lower and upper distance errors)
4ha	4hg1	3.5 1.0 0.7
4ha	4hd1	3.3 0.8 0.7
4ha	4he1	4.2 0.8 0.8
4ha	4hd2	3.2 0.6 0.6
4ha	4hb#	3.4 0.7 0.7
1ha2	1ha1	3.3 1.4 0.7
5hn	4ha	3.1 0.6 0.6
4ha	4hg2	3.3 0.7 0.7
4ha	4he2	4.1 0.8 0.8
5hb1	5hb2	2.2 0.4 0.4
5hn	5hb2	3.0 0.6 0.6
5h1	5hb2	2.7 0.2 2.3
5hb2	4hd1	3.9 0.8 1.1
5hb2	4hd2	4.4 0.9 0.9
5hn	5hb1	3.5 0.7 0.7
5hb1	4hd1	4.3 0.9 0.9
5h1	5hb1	2.7 0.2 2.3
5hb1	4hd2	3.9 0.8 1.1
5h1	5ha	2.8 0.6 0.6
5hn	5ha	3.1 0.6 0.6
5h2	5ha	3.2 0.7 1.8
4hd2	4hb#	4.4 0.9 0.9
5hn	4hb#	3.3 0.7 0.7
4hd1	4hb#	4.4 0.9 0.9
5hn	5h2	3.8 0.8 1.2
5hn	5h1	3.5 0.7 0.7
1ha1	4hd1	4.6 0.9 0.9
4hg1	4hd1	4.5 1.8 0.9
3hn	4hd1	3.8 0.8 0.8
4hg2	4hd1	4.7 2.0 0.9
4hg1	4hd2	2.8 0.6 0.6
1ha1	4hd2	5.1 1.0 0.9
4hg2	4hd2	3.6 0.7 0.7
1ha1	4he2	4.2 0.8 0.8
3hn	4he2	4.6 0.9 0.9
5hn	2hb#	4.1 0.7 0.7
3hn	4hg2	4.0 0.8 0.8
3hn	4hg1	3.7 0.8 0.8
2hn	1ha2	3.4 0.7 0.7
2hn	2hb#	3.2 0.7 0.7
2hn	2ha	3.3 0.7 0.7
2hn	1ha1	2.9 0.6 0.6
2hn	3hn	3.4 1.0 0.7
3hn	2ha	3.4 0.7 0.7

2ha	2hb#	3.5 0.8 0.8
3hn	2hb#	4.1 0.8 0.8
3hn	3ha*	3.5 0.7 0.7
3ha2	4hg2	1.8 0.0 0.0$^\$$

Methyl groups with degenerated protons.

* Not stereospecifically assigned protons.

$^\$$ The analysis of structures indicated the presence of a steric clash between these hydrogen atoms. Therefore, a distance of 1.8 Å was imposed for the final calculation.

Numbering of the hydrogen atoms of the covalent linker for XPLOR-NIH calculations:

10a
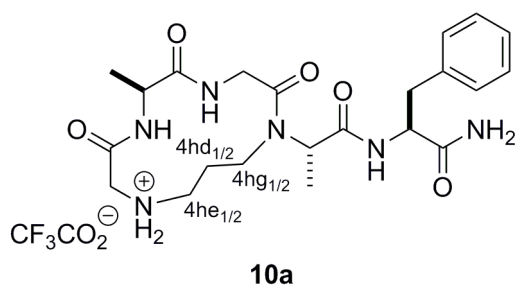
NOE derived distance restraints used for calculating the solution structure of the major rotamer of **10a**. Values were set according to the NOE intensity as strong (2.6 Å), medium (3.5 Å), weak (5.0 Å) and very weak (6.0 Å).

GAGAF_RS5modif_run4A1		Distances (lower and upper distance errors)
4ha	4hg1	5.0 3.2 0.0
4ha	4hd1	5.0 3.2 0.0
4ha	4he1	5.0 3.2 0.0
4ha	4hd2	5.0 3.2 0.0
4ha	4hb#	5.0 3.2 0.0
1ha2	1ha1	5.0 3.2 0.0
5hn	4ha	5.0 3.2 0.0
4ha	4hg2	5.0 3.2 0.0
4ha	4he2	5.0 3.2 0.0
5hb1	5hb2	2.6 0.8 0.0
5hn	5hb2	5.0 3.2 0.0
5h1	5hb2	5.0 3.2 0.0
5hb2	4hd1	5.0 3.2 0.0
5hb2	4hd2	6.0 4.2 0.0
5hn	5hb1	5.0 3.2 0.0
5hb1	4hd1	6.0 4.2 0.0
5h1	5hb1	5.0 3.2 0.0
5hb1	4hd2	5.0 3.2 0.0
5h1	5ha	3.5 1.7 0.0
5hn	5ha	5.0 3.2 0.0
5h2	5ha	5.0 3.2 0.0
4hd2	4hb#	6.0 4.2 0.0
5hn	4hb#	5.0 3.2 0.0
4hd1	4hb#	6.0 4.2 0.0
5hn	5h2	5.0 3.2 0.0
5hn	5h1	5.0 3.2 0.0
1ha1	4hd1	6.0 4.2 0.0
4hg1	4hd1	6.0 4.2 0.0
3hn	4hd1	5.0 3.2 0.0
4hg2	4hd1	6.0 4.2 0.0
4hg1	4hd2	3.5 1.7 0.0
1ha1	4hd2	6.0 4.2 0.0
4hg2	4hd2	5.0 3.2 0.0
1ha1	4he2	5.0 3.2 0.0
3hn	4he2	6.0 4.2 0.0
5hn	2hb#	5.0 3.2 0.0
3hn	4hg2	5.0 3.2 0.0
3hn	4hg1	5.0 3.2 0.0
2hn	1ha2	5.0 3.2 0.0
2hn	2hb#	5.0 3.2 0.0
2hn	2ha	5.0 3.2 0.0
2hn	1ha1	3.5 1.7 0.0

2hn	3hn	5.0 3.2 0.0
3hn	2ha	5.0 3.2 0.0
2ha	2hb#	5.0 3.2 0.0
3hn	2hb#	5.0 3.2 0.0
3hn	3ha*	5.0 3.2 0.0

Methyl groups with degenerated protons.

* Not stereospecifically assigned protons.

NOE derived distance restraints used for calculating the solution structure of the major rotamer of **10c**. Values were calculated using the CcpNmr software.[1]

LALAF_R1_run7A4		Distances (lower and upper distance errors)
4hg2	4he1	2.8 0.6 0.6
4hg2	4hd2	2.6 0.5 0.5
4hg2	4hd1	2.6 0.5 0.5
4hg2	4hg1	2.8 1.0 0.6
4he1	4hg1	3.2 0.6 0.6
1ht2	1ha or 4he1	2.9 0.6 0.6
1ha	1hd#*	2.5 0.5 0.5
4hd1	1hd#*	2.3 0.5 0.7
1hb2	1hd#*	2.3 0.5 0.7
1hb1	1hd#*	2.3 0.2 1.7
5hb2	5ha	2.4 0.5 0.5
5hn	5hb2	2.2 0.4 2.5
5hb1	5hb2	1.8 0.0 0.4
5h1	5hb2	2.1 0.3 0.4
1ha	1hb2	2.5 0.5 0.5
1ha	1hb1	2.0 0.2 0.4
2hb#	5ha	2.1 0.3 0.4
5hn	5ha	2.3 0.5 0.5
5hb1	5ha	2.6 0.5 0.5
5h1	5hb1	1.9 0.1 0.4
5hb1	2hb#	4.0 0.7 0.7
5hn	5hb1	2.5 0.5 2.0
5h2	5hb1	2.9 0.6 0.6
3hb1	3ha	2.4 0.5 0.5
3hn	3hb1	2.8 0.6 1.1
4hd2	4hd1	2.3 0.5 0.5
4he2	4hd2	2.9 0.6 0.6
4he1	4hd2	2.5 0.5 0.5
2hn	1ha	3.2 0.6 0.6
2hn	2ha	2.8 0.6 0.6
2hn	2hb#	2.3 0.5 0.5
3hn	2hn	3.0 0.6 0.6
2ha	2hb#	2.7 0.5 0.5
5hn	2hb#	2.8 0.6 0.6
3hb2	3hg	2.0 0.2 0.4
3hb2	3ha	2.4 0.5 0.5
5hn	5h2	3.1 0.6 0.6
5hn	4hb#	2.3 0.5 0.5
5hn	5h1	2.1 0.3 0.4
3hn	3hg	2.5 0.5 0.5
3hn	3ha	3.0 0.6 0.6
3hg	3ha	2.0 0.2 0.4
3hd1# ou 3hd2#	3ha	2.0 0.2 0.4
4he2	4he1	1.8 0.0 0.4

4he1	4hd1	2.5 0.4 0.4
4he2	4hd1	2.5 0.4 0.4
5h2	5h1	1.9 0.1 0.4
4ha	4hb#	2.7 0.5 0.5
4he1	2hn	2.0 0.2 0.2$^\$$
4he2	3hn	2.0 0.2 0.2$^\$$
1n	1hd22	2.0 0.2 0.2$^\$$
4hg2	3ha	2.0 0.2 0.2$^\$$
4he1	1hd21	2.0 0.2 0.2$^\$$

Methyl groups with degenerated protons.

* Not stereospecifically assigned methyl groups.

$^\$$ The analysis of structures indicated the presence of steric clashes between these hydrogen atoms. Therefore, distances of 2.0 ± 0.2 Å were imposed for the final calculation.

Numbering of the hydrogen atoms of the covalent linker for XPLOR-NIH calculations:

10c
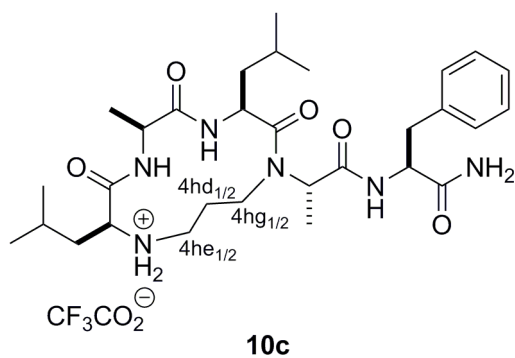
NOE derived distance restraints used for calculating the solution structure of the major rotamer of **10c**. Values were set according to the NOE intensity as strong (2.7 Å), medium (3.5 Å), weak (5.0 Å) and very weak (6.0 Å).

LALAF_R1_run7B3		Distances (lower and upper distance errors)
4hg2	4he1	3.5 1.7 0.0
4hg2	4hd2	3.5 1.7 0.0
4hg2	4hd1	3.5 1.7 0.0
4hg2	4hg1	3.5 1.7 0.0
4he1	4hg1	5.0 3.2 0.0
1ht2	1ha or 4he1	3.5 1.7 0.0
1ha	1hd#*	3.5 1.7 0.0
4hd1	1hd#*	3.5 1.7 0.0
1hb2	1hd#*	5.0 3.2 0.0
1hb1	1hd#*	5.0 3.2 0.0
5hb2	5ha	3.5 1.7 0.0
5hn	5hb2	5.0 3.2 0.0
5hb1	5hb2	2.7 0.9 0.0
5h1	5hb2	2.7 0.9 0.0
1ha	1hb2	3.5 1.7 0.0
1ha	1hb1	2.7 0.9 0.0
2hb#	5ha	2.7 0.9 0.0
5hn	5ha	3.5 1.7 0.0
5hb1	5ha	3.5 1.7 0.0
5h1	5hb1	2.7 0.9 0.0
5hb1	2hb#	5.0 3.2 0.0
5hn	5hb1	5.0 3.2 0.0
5h2	5hb1	3.5 1.7 0.0
3hb1	3ha	3.5 1.7 0.0
3hn	3hb1	5.0 3.2 0.0
4hd2	4hd1	3.5 1.7 0.0
4he2	4hd2	3.5 1.7 0.0
4he1	4hd2	3.5 1.7 0.0
2hn	1ha	5.0 3.2 0.0
2hn	2ha	3.5 1.7 0.0
2hn	2hb#	3.5 1.7 0.0
3hn	2hn	5.0 3.2 0.0
2ha	2hb#	3.5 1.7 0.0
5hn	2hb#	3.5 1.7 0.0
3hb2	3hg	2.7 0.9 0.0
3hb2	3ha	3.5 1.7 0.0
5hn	5h2	5.0 3.2 0.0
5hn	4hb#	3.5 1.7 0.0
5hn	5h1	2.7 0.9 0.0
3hn	3hg	3.5 1.7 0.0
3hn	3ha	5.0 3.2 0.0
3hg	3ha	2.7 0.9 0.0
3hd1# ou 3hd2#	3ha	2.7 0.9 0.0

4he2	4he1	2.7 0.9 0.0
4he1	4hd1	3.5 1.7 0.0
4he2	4hd1	3.5 1.7 0.0
5h2	5h1	2.7 0.9 0.0
4ha	4hb#	3.5 1.7 0.0
4he1	2hn	1.8 0.0 0.0$^\$$
4he2	3hn	1.8 0.0 0.0$^\$$
1n	1hd22	2.0 0.2 0.2$^\$$

Methyl groups with degenerated protons.

* Not stereospecifically assigned methyl groups.

$^\$$ The analysis of structures indicated the presence of steric clashes between these hydrogen atoms. Therefore, distances of 1.8 Å and 2.0 ± 0.2 Å were imposed for the final calculation.
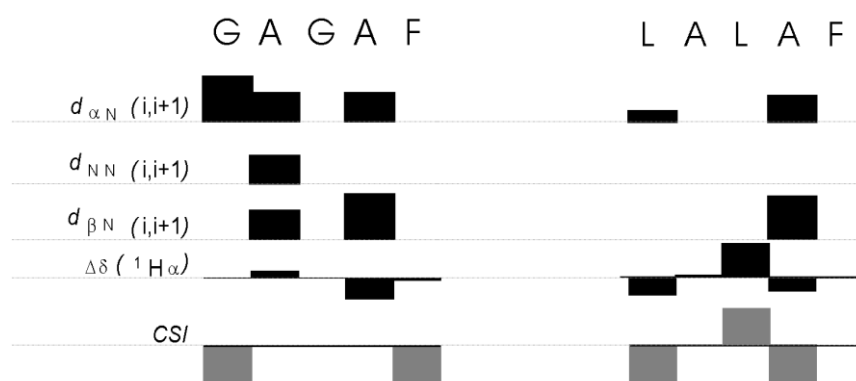
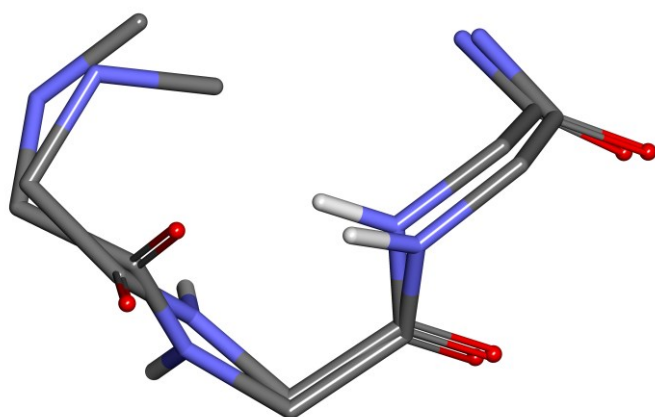Comparison of the sequential and short-range NOE connectivity and chemical shift indexes (CSI) of the major rotamers of the two cyclic peptides **10a** (left) and **10c** (right).

Overlap of backbone heavy atoms of the center member of cluster I of **10c_down** with the NMR structure of **10c**. A RMSD of 0.55 Å was calculated. Only the 14 heavy atoms constituting the turn, used for superposition and RMSD calculation, are shown.

Catalyst parameter input for conformational search:

```
// Catalyst parameter file
confAnalysis.best.max.successive.failures=1000
confAnalysis.best.torsion.factor=10
sysSearch.sp3sp3SearchOffset=10
sysSearch.sp3sp3SearchIncrement=10
sysSearch.sp3sp2SearchOffset=10
sysSearch.sp3sp2SearchIncrement=10
confAnalysis.fast.systematic.upperbound=10000
confAnalysis.fast.random.nAttempt=1000
confAnalysis.systematicFlipAxialEquatorialMaxRingSize=50
```
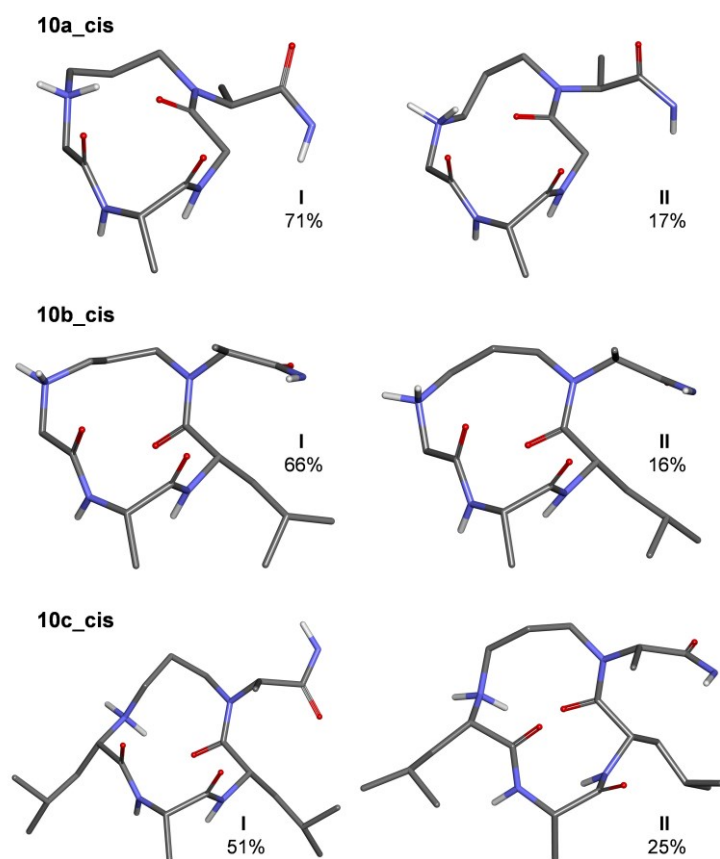
Molecular modeling of **10a-c_cis** isomers.
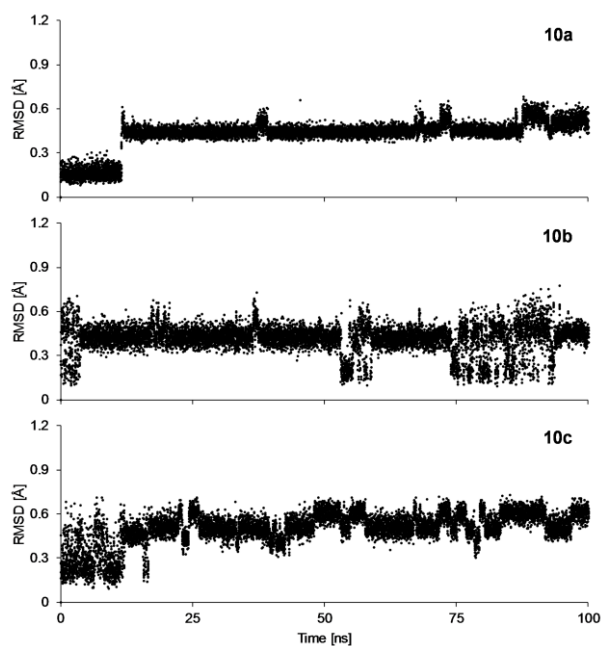
RMSD (Å) versus the first conformation of 10,000 conformations from 100 ns molecular dynamic simulations for macrocyclic peptides **10a-c_cis**.

	Cluster[a]	Φ_{i+1}	Ψ_{i+1}	Φ_{i+2}	Ψ_{i+2}	Φ_{i+3}	Ψ_{i+3}	d(Å)[b]
	I (71)	157	-153	-68	-47	-81	-136	4.5
10a	**II** (17)	117	-127	-71	-48	-81	-147	4.8
	III (12)	152	-127	-66	-63	-96	-176	4.8
	I (66)	71	-155	-68	-64	-95	134	5.1
	II (16)	93	-175	-65	-46	-94	141	4.9
	III (9)	72	-153	-55	-38	-133	111	5.2
10b	**IV** (3)	99	160	-61	-39	-107	121	4.7
	V (2)	77	-172	-69	-44	-114	151	4.7
	VI (2)	72	-152	-64	-61	-122	106	5.4
	I (51)	-55	-91	-81	-60	-87	136	5.2
	II (25)	-97	-98	-50	-52	-99	140	5.1
	III (9)	74	-132	-71	-64	-102	113	5.6
10c	**IV** (3)	67	-146	-64	-77	-91	125	5.4
	V (3)	-75	-87	-66	-64	-82	141	5.3
	VI (2)	45	-92	-74	-106	-74	126	5.9

Table title: Dihedral angles measured in center members of clusters of macrocyclic peptides **10a-c_cis**.

[a] The cluster size in % is given in parenthesis; [b] d: distance from Cα$_{i+1}$ to Cα$_{i+3}$.

10a_cis

I
71%

II
17%

10b_cis

I
66%

II
16%

10c_cis

I
51%

II
25%

Conformations of the center members of the two major clusters of **10a-c_cis** identified by molecular dynamic simulations. Cluster sizes are given in %. For clarity, only polar hydrogen atoms are shown, and the phenylalanine 5 amino acid is not represented.

[1] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides and E. D. Laue, *Proteins* **2005**, *59*, 687-696.