Supplementary Material (ESI) for Organic & Biomolecular Chemistry

Supporting Information

Conformational ensembles of flexible β-turn mimetics in DMSO-d_6

Jari J. Koivisto, Esa T. T. Kumpulainen and Ari M. P. Koskinen*

Aalto University School of Science and Technology, Faculty of Chemistry and Materials Sciences, Department of Chemistry, P.O.Box 16100, FI-00076 Aalto, Finland.
E-mail: ari.koskinen@tkk.fi; Fax: +358 9 470 22538; Tel: +358 9 470 22526

Table of contents

Aggregation studies

Figure S1 shows the variation in the amide proton ^{1}H NMR chemical shifts for APGA in DMSO-d_6 as a function of the logarithm of peptide concentration. The chemical shift of an amide proton is often very sensitive to hydrogen bonding. Consequently, these data provide insight on the concentrations at which aggregation (*i.e.* intermolecular hydrogen bonding) occurs. None of the three amide protons display concentration dependence, suggesting that this peptide does not aggregate in the concentration range examined. Figure S2 shows the ^{1}H NMR spectra for APGA over a concentration range of 1.6 to 100 mM in DMSO-d_6. The observed chemical shifts as well as spectral line widths in this concentration range are essentially unchanged, supporting the conclusion that intermolecular association is absent over the concentration range examined. Because A(4*R*)MePGA and A(4*S*)MePGA have only one additional methyl group in the proline ring, it was concluded that these peptides exhibit virtually identical aggregation behavior as observed for APGA.

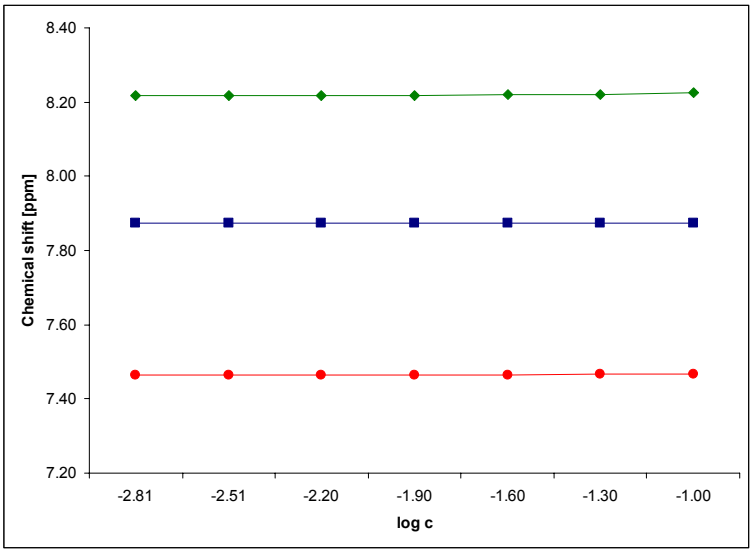


Figure S1. Amide proton chemical shifts in DMSO-d_6 at 299 K, as a function of the logarithm of APGA concentration: (♦) Gly-NH, (■) Ala^2-NH, and (●) Ala^1-NH.

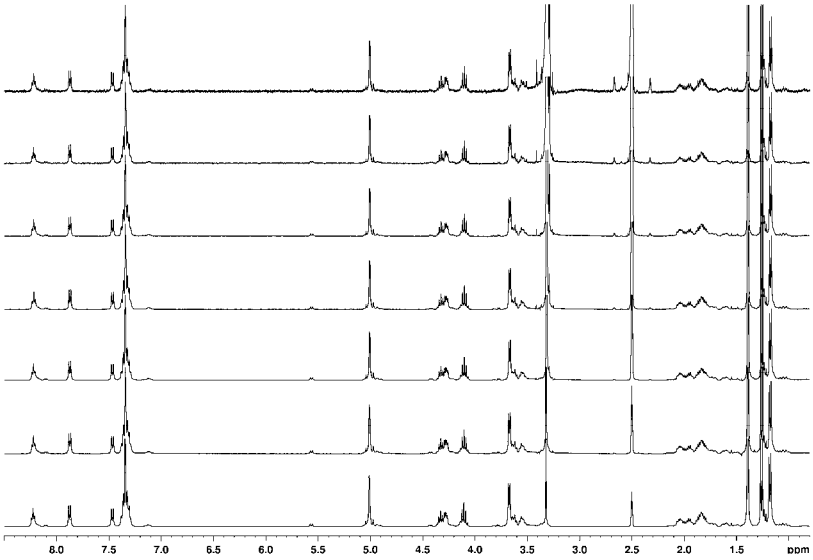


Figure S2. The ^{1}H NMR spectra for APGA over a concentration range of 1.6 (top) to 100 mM (bottom) in DMSO-d_6 at 299 K. The spectra have been normalized to approximately same peak intensity.

For a comparison, aggregation studies of APGA were performed also in $CDCl_3$ solution. Figure S3 shows the variation in the amide proton ^{1}H NMR chemical shifts for APGA in $CDCl_3$ as a function of the logarithm of peptide concentration. In Figure S4, the ^{1}H NMR spectra for APGA over a concentration range of 1.6 to 100 mM in $CDCl_3$ is displayed. These data show that APGA has a strong tendency to aggregate in $CDCl_3$, even at low concentrations.

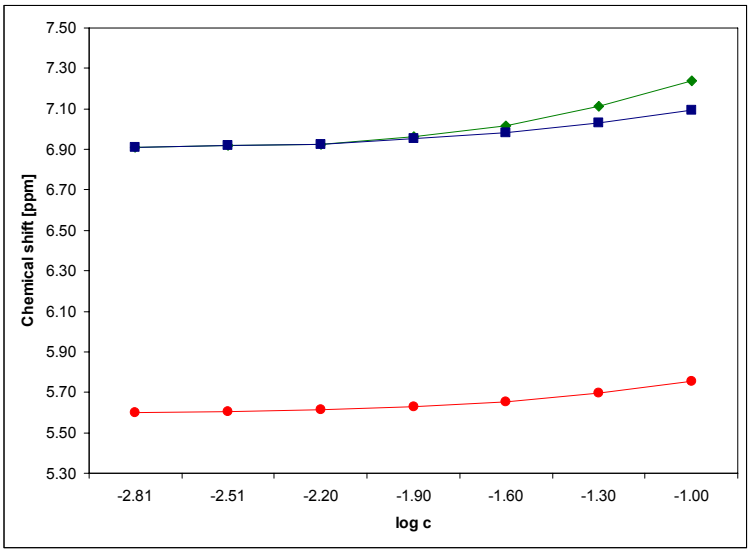


Figure S3. Amide proton chemical shifts in $CDCl_3$ at 299 K, as a function of the logarithm of APGA

concentration: (♦) Gly-NH, (■) Ala^2-NH, and (●) Ala^1-NH.

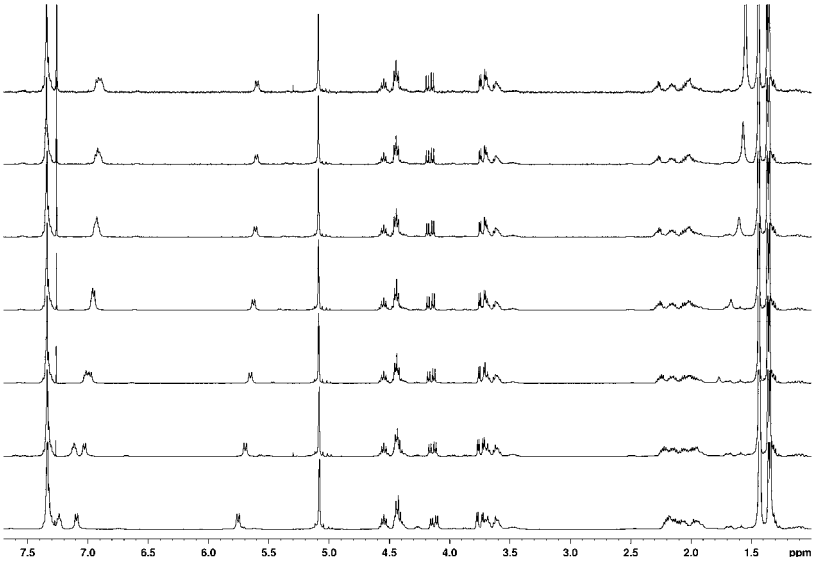


Figure S4. The ^{1}H NMR spectra for APGA over a concentration range of 1.6 (top) to 100 mM (bottom) in $CDCl_3$ at 299 K. The spectra have been normalized to approximately same peak intensity.

^{1}H and ^{13}C NMR parameters for APGA, A(4*R*)MePGA, and A(4*S*)MePGA

Table S1. ^{1}H NMR chemical shifts for APGA, A(4*R*)MePGA, and A(4*S*)MePGA in DMSO-d_6 at 299 K (ppm).

proton	APGA	A(4*R*)MePGA	A(4*S*)MePGA
CBz-Ar	7.3	7.3	7.3
CBz-CH_2	5.026; 5.005	5.016	5.006
Ala^1-NH	7.477	7.519	7.402
Ala^1-Hα	4.329	4.306	4.314
Ala^1-CH_3	1.179	1.174	1.164
Pro-Hα	4.279	4.332	4.231
Pro-$H\beta^{Pro(R)}$	2.042	1.681	2.246
Pro-$H\beta^{Pro(S)}$	1.818	1.940	1.390
Pro-$H\gamma^{Pro(R)}$	1.857		2.215
Pro-$H\gamma^{Pro(S)}$	1.955	2.407	
Pro-$H\delta^{Pro(R)}$	3.631	3.133	3.945
Pro-$H\delta^{Pro(S)}$	3.549	3.712	3.009
Pro-4-CH_3		0.970	1.026
Gly-NH	8.231	8.185	8.351
Gly-$H\alpha^{Pro(R)}$	3.674[a]	3.652	3.631
Gly-$H\alpha^{Pro(S)}$	3.674[a]	3.692	3.694
Ala^2-NH	7.882	7.863	7.875
Ala^2-Hα	4.105	4.106	4.095
Ala^2-CH_3	1.261	1.257	1.268
O*t*Bu-CH_3	1.393	1.393	1.391

[a] Degeneracy resulted in a deceptively simple triplet.

Table S2. ^{1}H NMR coupling constants for APGA, A(4*R*)MePGA, and A(4*S*)MePGA in DMSO-d_6 at 299 K (Hz).[a]

proton 1	proton 2	APGA	A(4*R*)MePGA	A(4*S*)MePGA
Ala^1-NH	Ala^1-Hα	7.36	7.55	7.13
Ala^1-Hα	Ala^1-CH_3	*6.95*	*6.93*	*7.01*
Gly-$H\alpha^{Pro(R)}$	Gly-$H\alpha^{Pro(S)}$		*-17.09*	*-16.94*
Gly-NH	Gly-$H\alpha^{Pro(R)}$	5.97	5.93	5.70
Gly-NH	Gly-$H\alpha^{Pro(S)}$	5.97	6.05	6.24
Ala^2-NH	Ala^2-Hα	6.89	6.96	6.77
Ala^2-Hα	Ala^2-CH_3	*7.31*	*7.30*	*7.31*
Pro-Hα	Pro-$H\beta^{Pro(R)}$	8.26	8.90	7.38
Pro-Hα	Pro-$H\beta^{Pro(S)}$	4.79	3.02	9.58
Pro-$H\beta^{Pro(R)}$	Pro-$H\beta^{Pro(S)}$	*-12.46*	*-12.59*	*-11.98*
Pro-$H\beta^{Pro(R)}$	Pro-$H\gamma^{Pro(R)}$	6.91		6.43
Pro-$H\beta^{Pro(R)}$	Pro-$H\gamma^{Pro(S)}$	8.09	9.53	
Pro-$H\beta^{Pro(S)}$	Pro-$H\gamma^{Pro(R)}$	6.00		11.40
Pro-$H\beta^{Pro(S)}$	Pro-$H\gamma^{Pro(S)}$	6.77	6.28	
Pro-$H\gamma^{Pro(R)}$	Pro-$H\gamma^{Pro(S)}$	*-12.21*		
Pro-$H\gamma^{Pro(R)}$	Pro-$H\delta^{Pro(R)}$	7.11		7.50
Pro-$H\gamma^{Pro(R)}$	Pro-$H\delta^{Pro(S)}$	5.69		10.24
Pro-$H\gamma^{Pro(S)}$	Pro-$H\delta^{Pro(R)}$	7.46	8.22	
Pro-$H\gamma^{Pro(S)}$	Pro-$H\delta^{Pro(S)}$	7.17	7.48	
Pro-$H\delta^{Pro(R)}$	Pro-$H\delta^{Pro(S)}$	*-9.98*	*-9.82*	*-9.80*
Pro-$H\gamma^{Pro(R)}$	Pro-4-CH_3			*6.34*
Pro-$H\gamma^{Pro(S)}$	Pro-4-CH_3		*6.63*	
CBz-CH^1	CBz-CH^2	*-12.84*		

[a] Coupling constants in italics were not used in the NAMFIS analysis.

Table S3. Off-resonance ROESY derived interproton distances for APGA, A(4*R*)MePGA, and A(4*S*)MePGA in DMSO-d_6 at 299 K (Å).

proton 1	proton 2	APGA	A(4*R*)MePGA	A(4*S*)MePGA
Ala^1-NH	Ala^1-Hα	2.60	2.60	2.58
Ala^1-NH	Ala^1-CH_3	2.75	2.76	2.76
Ala^1-NH	Pro-Hδ$^{Pro(R)}$	3.81	3.55	3.78
Ala^1-NH	Ala^2-CH_3	4.12	4.44	3.96
Ala^1-Hα	Ala^1-CH_3	2.52	2.51	2.52
Ala^1-Hα	Pro-Hδ$^{Pro(R)}$	2.18		2.02
Ala^1-Hα	Pro-Hδ$^{Pro(S)}$	2.30	2.24	2.73
Ala^1-CH_3	Pro-Hδ$^{Pro(R)}$	3.70	4.47	3.87
Ala^1-CH_3	Pro-Hδ$^{Pro(S)}$	3.14	3.12	3.10
Ala^1-CH_3	Gly-NH	4.26	4.39	4.31
Ala^1-CH_3	Ala^2-NH	4.31	4.40	4.44
Pro-Hα	Pro-Hβ$^{Pro(R)}$		2.25	
Pro-Hα	Pro-Hβ$^{Pro(S)}$		2.71	2.78
Pro-Hα	Pro-4-CH_3		3.79	
Pro-Hα	Gly-NH	2.07	2.15	2.02
Pro-Hα	Ala^2-NH	3.77	3.67	3.58
Pro-Hβ$^{Pro(R)}$	Pro-Hβ$^{Pro(S)}$		1.78	
Pro-Hβ$^{Pro(R)}$	Pro-Hγ$^{Pro(S)}$		2.90	
Pro-Hβ$^{Pro(R)}$	Pro-Hδ$^{Pro(R)}$		2.85	
Pro-Hβ$^{Pro(R)}$	Pro-4-CH_3		2.97	
Pro-Hβ$^{Pro(R)}$	Gly-NH			3.41
Pro-Hβ$^{Pro(S)}$	Pro-Hγ$^{Pro(S)}$		2.36	
Pro-Hβ$^{Pro(S)}$	Pro-Hδ$^{Pro(S)}$		3.38	2.65
Pro-Hβ$^{Pro(S)}$	Pro-4-CH_3		3.49	
Pro-Hβ$^{Pro(S)}$	Gly-NH	3.02	3.17	3.12
Pro-Hγ$^{Pro(S)}$	Pro-Hδ$^{Pro(R)}$		2.92	
Pro-Hγ$^{Pro(S)}$	Pro-Hδ$^{Pro(S)}$		2.29	
Pro-Hγ$^{Pro(S)}$	Pro-4-CH_3		2.66	
Pro-Hγ$^{Pro(S)}$	Gly-NH		3.12	
Pro-Hδ$^{Pro(R)}$	Pro-Hδ$^{Pro(S)}$		1.78[a]	1.78[a]
Pro-Hδ$^{Pro(R)}$	Pro-4-CH_3		2.94	3.58
Pro-Hδ$^{Pro(S)}$	Pro-4-CH_3		3.22	2.90
Pro-Hδ$^{Pro(S)}$	Gly-NH	3.45		3.55
Gly-NH	Ala^2-NH	2.61	2.72	2.50
Ala^2-NH	Ala^2-Hα	2.48	2.47	2.48
Ala^2-NH	Ala^2-CH_3	2.70	2.70	2.70
Ala^2-Hα	Ala^2-CH_3	2.45[a]	2.45	2.46

[a] Internal calibration distance.

Table S4. ^{13}C NMR chemical shifts for APGA, A(4*R*)MePGA, and A(4*S*)MePGA in DMSO-d_6 at 299 K (ppm).

carbon	APGA	A(4*R*)MePGA	A(4*S*)MePGA
CBz-Ar	128.31; 127.77; 127.68	128.32; 127.77; 127.65	128.30; 127.76; 127.65
CBz-Ar(q)	137.02	137.08	137.03
CBz-CH_2	65.36	65.32	65.33
CBz-CO	155.61	155.60	155.62
Ala^1-Cα	47.97	47.89	48.01
Ala^1-CH_3	16.68	16.74	16.49
Ala^1-CO	171.05	171.08	170.77
Pro-Cα	59.86	59.78	60.68
Pro-Cβ	28.93	36.49	36.90
Pro-Cγ	24.58	32.03	33.22
Pro-Cδ	46.72	53.39	53.71
Pro-4-CH_3		16.99	16.49
Pro-CO	171.86	171.80	171.73
Gly-Cα	41.71	41.67	41.83
Gly-CO	168.61	168.60	168.65
Ala^2-Cα	48.25	48.23	48.29
Ala^2-CH_3	17.02	17.04	16.97
Ala^2-CO	171.66	171.65	171.66
O*t*Bu-C(q)	80.34	80.36	80.29
O*t*Bu-CH_3	27.59	27.58	27.59

Classification of β-turns

By definition, a β-turn comprises four consecutive amino acids where the distance between $C^{\alpha}(i)$ – $C^{\alpha}(i + 3)$ is less than 7 Å, and the tetrapeptide chain is not in a helical conformation.[1] A β-turn can be stabilized by a hydrogen bond between the backbone CO(i) and the backbone NH(i + 3) or it can be "open", *i.e.* have no intraturn hydrogen bond.[2] The classification of β-turns is based on the preferred ϕ and ψ values of the central residues i + 1 and i + 2. The currently accepted nomenclature, described by Hutchinson and Thornton,[3] contains nine distinct types of β-turn: I, I', II, II', IV, VIa1, VIa2, VIb, and VIII. Each of these turn types, except type IV, have idealized ϕ,ψ values from which the maximum deviation of ±30° is allowed on three of these angles, while the fourth can deviate ±45°. If the criteria are not fulfilled, the turn is classified as type IV. However, it has been argued that a representative type IV turn does not rigorously satisfy the definition of an authentic β-turn and therefore cannot be considered one.[1] The majority of β-turns (excluding type IV) in folded proteins adopt type I and type II conformations.

References

1 K.-C. Chou, *Anal. Biochem.*, 2000, **286**, 1-16.
2 P. N Lewis, F. A. Momany and H. A. Scheraga, *Biochem. Biophys. Acta*, 1973, **303**, 211-229.
3 E. G. Hutchinson and J. M. Thornton, *Protein Sci.*, 1994, **3**, 2207-2216.

Detection of β-turns by ^{1}H NMR

The peptides under study contain only L-amino acids. Therefore, not all types of β-turn are relevant. For peptides with *trans* proline peptide bonds and a proline residue at the i + 1 position of the turn, only the type I and II β-turns need to be considered. All three peptides contain the Pro-Gly motif, which is known to strongly prefer the type II β-turn conformations over the type I β-turn conformations.[1-4] Therefore, the type II β-turn structures are expected to be significantly more populated than the type I β-turn structures. The low population of the type I β-turns can be rationalized by the lack of favorable side chain interactions in the Pro-Gly motif. The characteristic NOE connectivities for several different types of β-turns have been summarized by Wüthrich *et al*.[5] For both type I and type II β-turns, a strong $d(NH_{i+2},NH_{i+3})$ NOE connectivity should be observed together with a weaker $d(H\alpha_{i+1},NH_{i+3})$ connectivity. Type I and type II turns can be distinguished by the NOE connectivity between the backbone protons of residues i + 1 and i + 2: $d(NH_{i+1},NH_{i+2})$ connectivities

are expected for type I turns, whereas $d(H\alpha_{i+1},NH_{i+2})$ connectivities are expected for type II turns. However, detection of the $d(NH_{i+1},NH_{i+2})$ connectivity is not possible for the peptides under study, since the residue i + 1 is proline in all cases. In this case, the Pro-H$\delta^{pro(S)}$ proton may be regarded as equivalent to an NH proton. Type I and type II β-turns can also, in principle, be distinguished on the basis of ^3J(NH,Hα) coupling constants at positions i + 1 and i + 2 of the turn.[6] This coupling constant should adopt a value of approximately 5.5 – 6.5 Hz for ϕ = 80° at position i + 2 in type II turns and 7 – 8 Hz for ϕ = -90° in type I turns. In both type I and type II turns, residue i + 1 should have a ^3J(NH,Hα) of about 4 – 4.5 Hz for ϕ = -60°. Except for the peptides in which the population of β-turn structures is very large, ^3J(NH,Hα) coupling constants are not much reduced below the extended chain values. Therefore, they are less useful than other parameters in distinguishing between different types of β-turn.

Linear peptides, such as those under study, are present in solution as conformational ensembles containing a large number of widely different conformers, so that the observable NMR parameters represent population weighted averaged values.[7-9] However, the characteristic NOE/ROE connectivities and ^3J(NH,Hα) coupling constants, which are diagnostics of β-turns, are often sufficient to identify even quite small populations of these structures. The qualitative analysis of the T-ROESY spectra provided evidence for relatively stable β-turn structures in APGA, A(4*R*)MePGA, and A(4*S*)MePGA in DMSO-d_6 solution. All three peptides displayed a strong connectivity between Gly-NH and Ala^2-NH resonances together with a weak connectivity between Pro-Hα and Ala^2-NH resonances. In addition, a strong correlation between Pro-Hα and Gly-NH was observed for all peptides under study. These findings suggested that the type II β-turn conformations were highly populated in an ensemble of conformations available for these peptides. However, in addition to the above-mentioned correlations, a weak correlation between Pro-H$\delta^{pro(S)}$ and Gly-NH was observed for all three peptides. Consequently, type I β-turn conformations may also exist in the conformational pool of these peptides. The existence of the type I turn structures was supported by observation of a weak correlation between Ala^1-CH_3 and Gly-NH resonances for all peptides under study. In addition to the above-mentioned connectivities, several other inter- and intraresidue correlations were observed in the T-ROESY spectra of all three peptides. Many of these connectivities cannot be explained by a single conformation. For example, a strong sequential $d(H\alpha_i,NH_{i+1})$ connectivity (for proline and 4-methylproline, $d(H\alpha_i,H\delta_{i+1})$ connectivity) was observed together with a weak $d(H\beta_i,NH_{i+1})$ correlation between Pro and Gly residues. In addition, weak $d(H\beta_i,NH_{i+3})$ and $d(NH_i,H\beta_{i+3})$ connectivities were

observed between Ala^1 and Ala^2 residues. Among the intraresidue connectivities observed were $d(H\alpha_i,NH_i)$ connectivity for Ala^1, Gly, and Ala^2 residues, and $d(H\beta_i,NH_i)$ connectivity for Ala^1 and Ala^2 residues. Furthermore, several correlations between the protons of the proline ring were observed.

The $^3J(NH,H\alpha)$ coupling constants of flexible peptides often give equivocal results and are usually of less value than NOE/ROE connectivities in identifying different classes of β-turn structures. In the peptides under study, the $^3J(NH,H\alpha)$ coupling constant value of the pro(*R*) $C^{\alpha}H$ proton of Gly residues lies between 5.7 and 6.0 Hz, which is in the range of the expected value for a type II β-turn (5.5 – 6.5 Hz for $\phi = 80°$). However, these values are not greatly different from the value of *ca.* 7 – 8 Hz observed for unfolded peptides. The low field $C^{\alpha}H$ proton of Gly residues was tentatively assigned as pro(*S*) (except for APGA, in which degeneracy results in a deceptively simple triplet), based on a proposed connection between chemical shift and coupling constant for the assignment of diastereotopic Gly $C^{\alpha}H$ protons (according to this proposition, low field and large coupling imply pro(*S*) configuration).[10] See Table S2 for the measured $^3J(NH,H\alpha)$ coupling constants for APGA, A(4*R*)MePGA, and A(4*S*)MePGA.

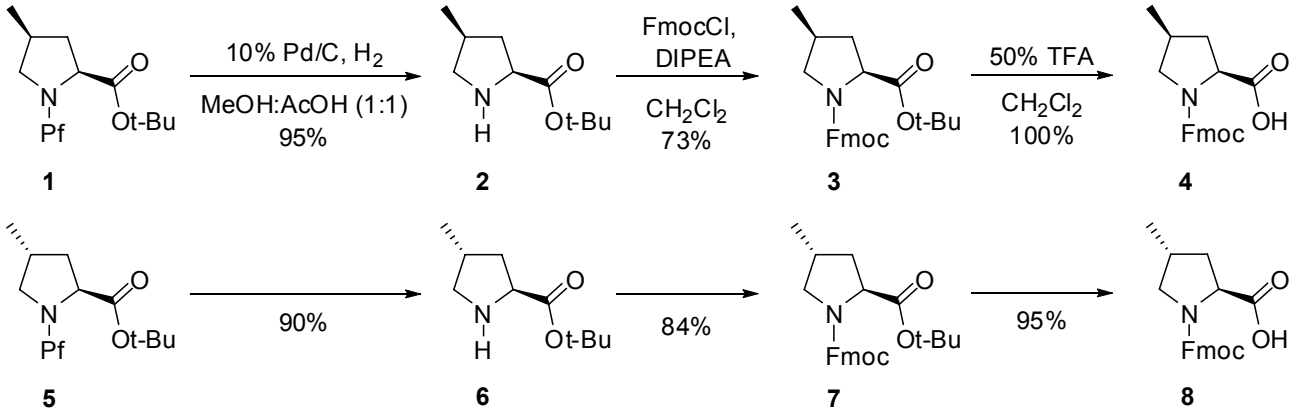
References

1 E. G. Hutchinson and J. M. Thornton, *Protein Sci.*, 1994, **3**, 2207-2216.
2 H. J. Dyson, M. Rance, R. A. Houghten, R. A. Lerner and P. E. Wright, *J. Mol. Biol.*, 1988, **201**, 161-200.
3 K. Guruprasad and S. Rajkumar, *J. Biosci.*, 2000, **25**, 143-156.
4 M. Buděšínský, J. Šebestík, L. Bednárová, V. Baumruk, M. Šafařík and P. Bouř, *J. Org. Chem.*, 2008, **73**, 1481-1489.
5 K. Wüthrich, M. Billeter and W. Braun, *J. Mol. Biol.*, 1984, **180**, 715-740.
6 G. Wagner, D. Neuhaus, E. Wörgötter, M. Vašák, J. H. R. Kägi and K. Wüthrich, *J. Mol. Biol.*, 1986, **187**, 131-135.
7 O. Jardetzky, *Biochem. Biophys. Acta*, 1980, **621**, 227-232.
8 H. Kessler, C. Griesinger, J. Lautz, A. Müller, W. F. van Gunsteren and H. J. C. Berendsen, *J. Am. Chem. Soc.*, 1988, **110**, 3393-3396.
9 X. Daura, I. Antes, W. F. van Gunsteren, W. Thiel and A. E. Mark, *Proteins Struct. Funct. Genet.*, 1999, **36**, 542-555.
10 H. Kessler, W. Hehlein and R. Schuck, *J. Am. Chem. Soc.*, 1982, **104**, 4534-4540.

Synthesis

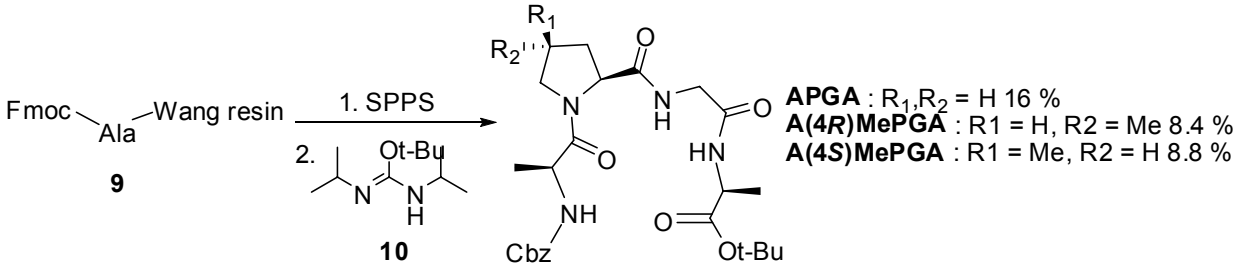
Diastereomerically pure fully protected prolinate **1** was obtained using literature procedure described by Koskinen *et al.* (Scheme S1).[1] Pf-protection from this prolinate **1** was removed by hydrogenation in excellent crude yield (95%). Reprotection with Fmoc-group in moderate yield (73%) and removal of *tert*-butyl ester with 50% TFA in dichloromethane in quantitative yield gave finally (2*S*,4*S*)-4-methylproline **4** suitable for solid phase peptide synthesis (SPPS). The (2*S*,4*R*) diastereomer **5** was reacted in the same fashion as the (2*S*,4*S*) diastereomer **1**, giving similar yields.

Scheme S1

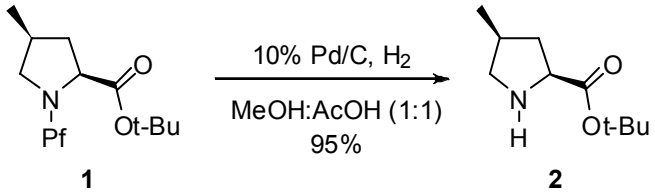


Peptide synthesis was carried out using SPPS techniques starting from commercially available Wang resin **9** (Scheme S2). First, the functionality of the synthetic route was tested using natural (2*S*)-proline. After esterification of the crude peptide with isourea **10,** it was purified using flash chromatography and fully protected APGA was obtained in 16% overall yield. The procedure was deemed reliable and it was repeated using the two 4-methylprolines. With (2*S*,4*S*)-4-methylproline, A(4*S*)MePGA was after flash chromatography in 15% yield. Further purification with the preparative HPLC lowered the yield to 8.8%. In a similar fashion with (2*S*,4*R*)-4-methylproline yields of A(4*R*)MePGA were 15% and 8.4%.

Scheme S2

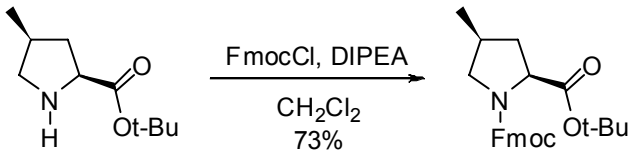


General methods. All reactions were carried out under an argon atmosphere in flame-dried glassware, unless otherwise noted. Nonaqueous reagents were transferred under argon *via* syringe or cannula and dried prior to use. MeOH was distilled from $Mg(OMe)_2$, and CH_2Cl_2 from CaH_2. Other solvents and reagents were used as obtained from supplier, unless otherwise noted. Analytical TLC was performed using Merck silica gel F_{254} (10-12 μm) plates and analyzed by UV light or by staining upon heating either with ninhydrin solution (1 g ninhydrin, 100 mL ethanol, 5 drops glacial acetic acid), PMA solution (2.5 g phosphomolybdic acid, 100 mL EtOH, 5 mL conc. H_2SO_4, 1.5 mL 85% H_3PO_4) or $KMnO_4$ solution (1 g $KMnO_4$, 2 g Na_2CO_3, 100 mL H_2O). For silica gel chromatography, the flash chromatography technique was used, with Merck silica gel 60 (40-63 μm) and p.a. grade solvents, unless otherwise noted. Melting points were measured with a Gallenkamp GMP apparatus. IR spectra were recorded on a Perkin-Elmer Spectrum One spectrometer. Optical rotations were recorded with a Perkin-Elmer 343 polarimeter. The ^1H NMR and ^{13}C NMR spectra were recorded in $CDCl_3$ at 298 K on a Bruker Avance DPX400 (^1H 400.13 MHz; ^{13}C 100.61 MHz) spectrometer. The chemical shifts are reported in ppm relative to residual $CHCl_3$ (δ 7.26) for ^1H NMR. For the ^{13}C NMR spectra, the $CDCl_3$ (δ 77.0) was used as the internal standard. Brackets "[ ]" in NMR data denote rotameric pairs of protons or carbons. HRMS spectra were recorded on Waters LCT Premier mass spectrometer. Elemental analyses were recorded on a Perkin-Elmer 2400 CHN Elemental Analyzer.



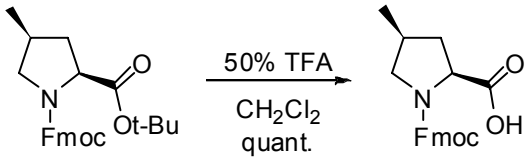
(2*S*,4*S*)-4-Methylproline *tert*-butyl ester (2). Following the procedure described by Koskinen *et al.*,[1] prolinate **1** (0.955 g, 2.24 mmol, 100 mol-%) was dissolved in MeOH (8.5 mL) and AcOH (8.5 mL). To this solution was added 10% Pd / C (0.11 g, 0.10 mmol, 5 mol-%) and flask was flushed with argon. Argon atmosphere was replaced with H_2 (1 atm) using suction. This procedure was repeated four times to remove all other gases and reaction mixture was stirred for 19 h at room temperature. The mixture

was filtered through celite followed by MeOH washings. Solvents were evaporated and the residue was partitioned between 1 M AcOH solution (20 mL) and ether (10 mL). The aqueous phase was extracted with ether. The combined organic extracts were back extracted with 1 M AcOH solution (10 mL), and the combined aqueous phases were basified with saturated Na_2CO_3 solution. This aqueous phase was extracted five times with CH_2Cl_2. The combined organic extracts were dried over anhydrous Na_2SO_4, filtered and the solvent was evaporated to give 0.390 g (95%) of amine **2** as an oil. The crude product was used as such in the next reaction. R_f = 0.30 (10% MeOH / CH_2Cl_2); ^{1}H NMR (400 MHz, $CDCl_3$) δ 1.01 (d, 3H, *J* = 6.6 Hz), 1.35 (td, 1H, *J* = 8.0 Hz, *J* = 12.3 Hz), 1.46 (s, 9H), 2.15-2.25 (m, 1H), 2.29 (td, 1H, *J* = 7.9 Hz, *J* = 12.6 Hz), 2.46 (br s, 1H), 2.60 (dd, 1H, *J* = 9.7 Hz, *J* = 8.2 Hz), 3.04 (dd, 1H, *J* = 6.7 Hz, *J* = 9.7Hz), 3.68 (t, 1H, *J* = 7.9 Hz). These data match those reported in the literature.[1]



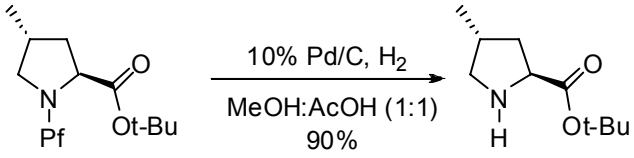
2 **3**

(2*S*,4*S*)-4-Methyl *N*-[(9-fluorenyl)-methoxycarbonyl]-proline *tert*-butyl ester (3). To a stirred solution of amine **2** (0.395 g, 2.13 mmol, 100 mol-%) in CH_2Cl_2 (20 mL) was added DIPEA (0.60 mL, 3.36 mmol, 160 mol-%) and FmocCl (0.840 g, 3.15 mmol, 150 mol-%) at room temperature. The reaction was stirred for 22 h and then diluted with CH_2Cl_2 (20 mL). The mixture was washed with 0.5 M H_3PO_4 solution (20 mL) and the organic phase was dried with anhydrous Na_2SO_4. Filtration and evaporation of the solvent in vacuo and purification with flash chromatography (5-10% EtOAc / hexanes gave) 0.760 g (87%) of prolinate **3** as off-white solids. Further purification by recrystallization from EtOAc / hexanes (1:4) gave pure **3** (0.638 g, 73%) as white granules. R_f = 0.53 (40% EtOAc / hexanes); mp. 135-136 °C; IR (thin film, cm^{-1}) 2975, 2934, 1742, 1707, 1451, 1417, 1351, 1254, 1197, 1177, 1155, 1105, 758, 740; $[\alpha]^{20}$ = -99.4 (c 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ [1.09 (d, *J* = 4.7 Hz), 1.11 (d, *J* = 4.7 Hz), 3H], [1.43 (s), 1.46 (s), 9H], 1.49-1.66 (m, 1H), 2.16-2.36 (m, 1H), [2.43 (td, *J* = 7.4 Hz, *J* = 12.7 Hz), 2.51 (td, *J* = 7.4 Hz, *J* = 13.2 Hz), 1H], [3.07, (t, *J* = 10.0 Hz), 3.11 (t, *J* = 9.8 Hz), 1H], [3.79 (dd, *J* = 7.3 Hz, *J* = 10.1 Hz), 3.86 (dd, *J* = 7.3 Hz, *J* = 10.4 Hz), 1H], 4.15-4.34 (m, 3H), 4.38-4.49 (m, 1H), 7.27-7.79 (m, 8H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 17.0, 17.1, 27.9, 27.9, 32.3, 33.2, 37.9, 39.0, 47.2, 47.2, 53.6, 54.0, 59.9, 60.2, 67.3, 67.6, 81.1, 81.3, 119.8, 125.1, 125.1, 125.4, 126.9, 126.9, 126.9, 127.0, 127.5, 127.6, 141.1, 141.2, 143.5, 143.9, 144.2, 144.3, 154.2, 154.5, 171.8, 172.0. Rotameric pairs could not be fully identified for ^{13}C spectra. Anal. calcd. for $C_{25}H_{29}NO_4$: C, 73.68; H, 7.17; N, 3.44. Found: C, 73.84; H, 7.18; N, 3.37.



3 **4**

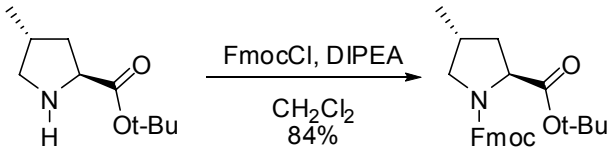
(2*S*,4*S*)-4-Methyl *N*-[(9-fluorenyl)-methoxycarbonyl]-proline (4). To a stirred solution of prolinate **3** (0.597 g, 1.46 mmol, 100 mol-%) in CH_2Cl_2 (7 mL) was added trifluoroacetic acid (7 mL). The mixture was stirred at room temperature for 1.5 h, then diluted with CH_2Cl_2 (15 mL) and distilled water (15 mL) was added. The aqueous phase was extracted with CH_2Cl_2, the combined organic phases washed with brine and the aqueous phase back extracted with CH_2Cl_2. The combined organic extracts were dried with anhydrous Na_2SO_4, filtered and evaporated to dryness. Filtration through a 2 cm pad of silica using first 0-5% MeOH / CH_2Cl_2 gave 0.530 g (quant.) of off white foamish solid. R_f = 0.22 (10% MeOH / CH_2Cl_2); IR (thin film, cm^{-1}) 3066 (broad band), 2959, 1747, 1705, 1451, 1425, 1354, 1195, 1178, 759, 737; $[\alpha]^{20}$ = -69.2 (c 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ [1.03 (d, *J* = 6.5 Hz), 1.07 (d, *J* = 6.5 Hz), 3H], 1.53-1.79 (m, 1H), 2.10-2.35 (m, 1H), 2.35-2.55 (m, 1H), [3.00 (t, *J* = 10.1 Hz), 3.04 (t, *J* = 9.7 Hz), 1H], [3.75 (dd, *J* = 7.4 Hz, *J* = 10.1 Hz), 3.82 (dd, *J* = 7.4 Hz, *J* = 10.2 Hz), 1H], 4.05-4.40 (m, 3H), [4.44 (t, *J* = 9.7 Hz), 4.46 (t, *J* = 10.4 Hz), 1H], 7.10-7.90 (m, 8H), 9.78 (br s, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 16.8, 17.0, 32.5, 33.1, 37.2, 38.9, 47.1, 53.6, 53.9, 58.8, 59.6, 67.5, 67.8, 119.8, 119.9, 124.8, 124.9, 125.0, 125.0, 127.0, 127.5, 127.7, 141.1, 141.2, 143.6, 143.6, 143.8, 144.0, 154.3, 155.5, 176.6, 178.1. Rotameric pairs could not be fully identified for ^{13}C spectra. HRMS (ESI) exact mass calc. for $C_{21}H_{21}NO_4Na$ $[M+Na]^+$: 374.1368, found: 374.1381, Δ = 3.5 ppm.



5 **6**

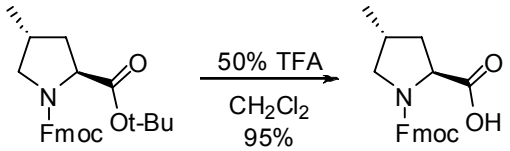
(2*S*,4*R*)-4-Methylproline *tert*-butyl ester (6). Following the procedure described for (2*S*,4*S*)-diastereomer **2** using 0.360 g (0.84 mmol, 100%) of prolinate **5** to give amine **6** (0.141 g, 90%) as an oil. The product was used as such. R_f = 0.22 (10% MeOH / CH_2Cl_2); IR (thin film, cm^{-1}) 3350, 2960, 2931, 28.71, 1728, 1456, 1367, 1339, 1244, 1157, 849; $[\alpha]^{20}$ = -27.0 (c 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ 1.01 (d, 3H, *J* = 6.5 Hz), 1.46 (s, 9H), 1.60-1.74 (m, 1H), 1.96-2.17 (m, 2H), 2.24-2.62 (m, 2H), 3.15-3.32 (m, 1H), 3.70 (br s, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 17.9, 28.0, 33.3, 38.8,

54.7, 60.5, 80.8, 174.9; HRMS (ESI) exact mass calc. for $C_{10}H_{20}NO_2$ $[M+H]^+$: 186.1494, found: 186.1487, Δ = -3.8 ppm.



6 **7**

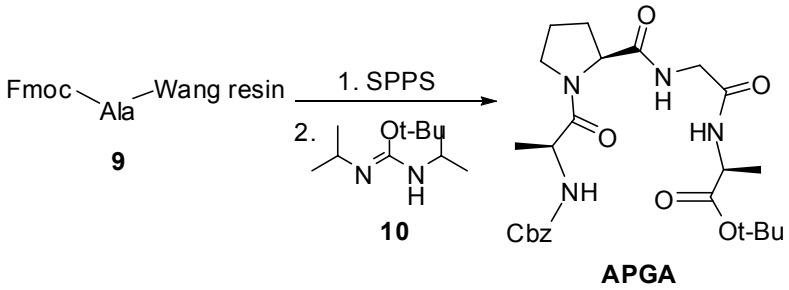
(2*S*,4*R*)-4-Methyl *N*-[(9-fluorenyl)-methoxycarbonyl]-proline *tert*-butyl ester (7). Following the procedure described for (2*S*,4*S*)-diastereomer **3**, using 0.140 g (0.75 mmol, 100%) of amine **6** to give prolinate **7** (0.258 g, 84%) as a clear glass. R_f = 0.54 (40% EtOAc / hexanes); IR (thin film, cm^{-1}) 2972, 1740, 1708, 1451, 1417, 1366, 1349, 1155, 1144, 758, 740; $[\alpha]^{20}$ = -57.0 (c 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ [1.07 (d, *J* = 4.1 Hz), 1.08 (d, *J* = 4.1 Hz), 3H], [1.43 (s), 1.46 (s), 9H], [1.83 (ddd, *J* = 9.0 Hz, *J* = 10.1 Hz, *J* = 12.7 Hz), 1.89 (ddd, *J* = 9.0 Hz, *J* = 10.4 Hz, *J* = 12.9 Hz), 1H], [2.08 (ddd, *J* = 2.9 Hz, *J* = 6.3 Hz, *J* = 12.7 Hz), 2.12 (ddd, *J* = 2.7 Hz, *J* = 6.3 Hz, *J* = 12.9 Hz), 1H], 2.35-2.54 (m, 1H), [3.04 (dd, *J* = 4.2 Hz, *J* = 8.5 Hz), 3.06 (dd, *J* = 4.5 Hz, *J* = 8.4 Hz), 1H], [3.79 (dd, *J* = 7.6 Hz, *J* = 10.6 Hz), 3.82 (dd, *J* = 7.6 Hz, *J* = 10.3 Hz), 1H], 4.15-4.37 (m, 3H), [4.41 (dd, *J* = 7.2 Hz, *J* = 10.4 Hz), 4.44 (dd, *J* = 6.3 Hz, *J* = 10.3 Hz), 1H], 7.26-7.80 (m, 8H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 17.4, 17.5, 27.9, 28.0, 30.9, 31.9, 37.6, 38.7, 47.2, 47.3, 53.3, 53.8, 59.9, 60.0, 67.3, 67.5, 81.2, 81.4, 119.9, 125.1, 125.1, 125.2, 125.4, 126.9, 127.0, 127.0, 127.6, 127.6, 141.2, 141.2, 141.2, 143.6, 143.9, 144.2, 144.3, 154.4, 154.7, 171.7, 171.8. Rotameric pairs could not be fully identified for ^{13}C spectra. HRMS (ESI) exact mass calc. for $C_{25}H_{29}NO_4Na$ $[M+Na]^+$: 430.1994, found: 430.1991, Δ = -0.7 ppm.



7 **8**

(2*S*,4*R*)-4-Methyl *N*-[(9-fluorenyl)-methoxycarbonyl]-proline (8). Following the procedure described for (2*S*,4*S*)-diastereomer **4**, using 0.243 g (0.59 mmol, 100%) of prolinate **7** to give prolinate **8** (0.199 g, 95%) as an off-white foamish solid. R_f = 0.25 (10% MeOH / CH_2Cl_2); $[\alpha]^{20}$ = -40.0 (c 1.0, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$) δ [1.02 (d, *J* = 6.5 Hz), 1.06 (d, *J* = 6.4 Hz), 3H], 1.72-1.94 (m, 1H), 2.10-2.53 (m, 2H), 2.95-3.05 (m, 1H), 3.73 (q, *J* = 9.8 Hz), 4.08-4.52 (m, 4H), 7.15-7.85 (m, 8H),

9.61 (br s, 1H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 17.3, 31.0, 32.1, 36.7, 38.5, 47.1, 47.1, 53.3, 53.6, 58.7, 59.5, 67.4, 67.9, 119.8, 119.9, 124.8, 124.9, 125.0, 125.0, 127.5, 127.7, 141.2, 143.6, 143.7, 143.8, 143.9, 154.3, 155.9, 176.0, 178.0. Rotameric pairs could not be fully identified for ^{13}C spectra. These data match those reported in the literature.[2]



APGA

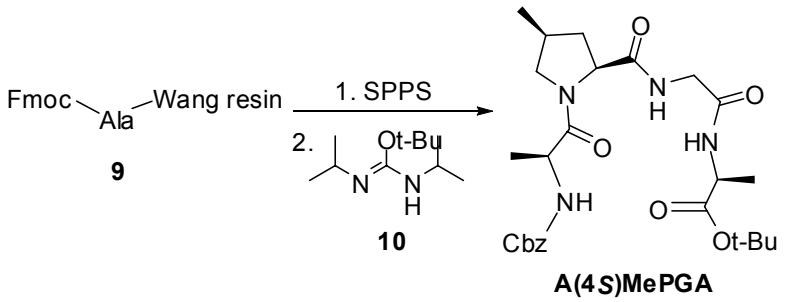
CBz-Ala-Pro-Gly-Ala-O*t*Bu (APGA). Wang resin (0.30 g, 0.20 mmol, 100 mol-%, 100-200 mesh, 0.67 mmol g^{-1}) was allowed to swell for 30 min in DMF (3 mL).

Method for attachment of amino acids: The resin was treated with 20% piperidine / DMF solution for 5 min. The solvent was removed and the resin was treated a second time with 20% piperidine / DMF solution for 30 min. The solvent was removed and the resin was washed with DMF (10 x 3 mL). A solution of Fmoc-protected amino acid (0.40 mmol, 200 mol-%), HBTU (0.152 g, 0.40 mmol, 200 mol-%), HOBt (0.062 g, 0.40 mmol, 200 mol-%) and DIPEA (0.14 mL, 0.80 mmol, 400 mol-%) in DMF was stirred for 15 min and then added onto resin. This mixture was stirred at least 4 h. The solvent was removed and the resin was washed with DMF (4 x 3 mL). The resin was treated with a solution of Ac_2O (0.057 mL, 0.20 mmol, 100 mol-%) and DIPEA (0.10 mL, 0.20 mmol, 100 mol-%) in DMF (3 mL) for 30 min. The solvent was removed and the resin was finally washed with DMF (3 x 3 mL). The procedure was repeated with 1. Fmoc-Gly-OH, 2. Fmoc-Pro-OH and 3. Cbz-Ala-OH.

Cleavage: The resin was treated with 100% TFA solution (3 mL) for 30 min and then rinsed with DCM (4 x 3 mL) and MeOH (4 mL). Collected organic extracts were evaporated to dryness in vacuo to give an oil.

Esterification: The crude product was dissolved in dry THF (3 mL) and treated with *O-tert*-butyl *N*,*N*-diisopropylisourea[3] (0.17 g, 0.85 mmol, 425 mol-%) and stirred overnight. The reaction mixture was diluted with EtOAc (5 mL) and the solids were filtered off. After evaporation of solvents, the crude product was purified by flash chromatography using 5% MeOH in EtOAc to yield 0.017 g (16%) of

APGA as a foam. R_f = 0.12 (5% MeOH / EtOAc); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.27-7.41 (m, 5H), 6.92-7.24 (m, 2H), 5.57-5.73 (m, 1H), 5.00-5.18 (m, 2H), 4.48-4.62 (m, 1H), 4.32-4.48 (m, 2H), 4.08-4.21 (m, 1H), 3.53-3.78 (m, 3H), 1.53-2.35 (m, 5-6H),1.39-1.46 (m, 9H), 1.29-1.32 (m, 6H). HRMS (ESI) exact mass calc. for $C_{25}H_{36}N_4O_7Na$ $[M+Na]^+$: 527.2482, found: 527.2483, Δ = 0.2 ppm.



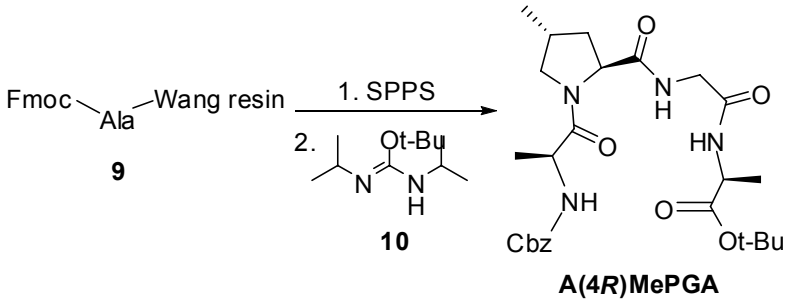
CBz-Ala-(4*S*)-MePro-Gly-Ala-O*t*Bu (A(4*S*)MePGA). Wang resin (0.746 g, 0.50 mmol, 100 mol-%, 100-200 mesh, 0.67 mmol g^{-1}) was allowed to swell for 30 min in DMF (4 mL).

Method for attachment of amino acids: The resin was treated with 20% piperidine / DMF –solution (4 mL) for 5 min. The solvent was removed and the resin treated second time with 20% piperidine / DMF –solution (4 mL) for 30 min. The solvent was removed and the resin was washed with DMF (10 x 4 mL). A solution of Fmoc-protected amino acid (1.0 mmol, 200 mol-%), HBTU (0.379 g, 1.0 mmol, 200 mol-%), HOBt (0.153 g, 1.0 mmol, 200 mol-%) and DIPEA (0.35 mL, 2.0 mmol, 400 mol-%) in DMF (3 mL) was stirred for 15 min and then added onto resin. This mixture was mixed overnight. The solvent was removed and the resin was washed with DMF (4 x 4 mL). The resin was treated with a solution of Ac_2O (0.047 mL, 0.50 mmol, 100 mol-%) and DIPEA (0.087 mL, 0.50 mmol, 100 mol-%) in DMF (4 mL) for 30 min. The solvent was removed and the resin was finally washed with DMF (4 x 4 mL). The procedure was repeated with 1. Fmoc-Gly-OH, 2. Fmoc-(4*S*)-MePro-OH using 100 mol-% of amino acid, HBTU, HOBt, and 200 mol-% DIPEA, 3. Cbz-Ala-OH.

Cleavage: The resin was treated with 100% TFA solution (3 mL) for 30 min and rinsed with DCM (4 x 4 mL) and MeOH (4 mL). Collected organic extracts were evaporated to dryness in vacuo to give oil.

Esterification: The crude product was dissolved in dry THF (5 mL) and treated with *O-tert*-butyl *N*,*N*-diisopropylisourea[3] (0.50 g, 2.5 mmol, 500 mol-%) and stirred overnight. The reaction mixture was diluted with EtOAc (5 mL) and the solids were filtered off. After evaporation of solvents, the crude product was purified by flash chromatography using 5% MeOH in EtOAc to yield 0.041 g (15%) of

A(4*S*)MePGA as oil. A second purification was done by using preparative HPLC (Shandon Hypersil 10x250 mm, 10 μm) and 5% IPA / hexanes as eluent to give 0.023 g (8.8%) of clear oil. R_f = 0.23 (5% MeOH / EtOAc); ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.27-7.41 (m, 5H), 6.87-7.09 (br s, 1H), 5.69 (d, *J* = 7.6 Hz, 1H), 5.04-5.16 (m, 2H), 4.53 (dq, *J* = 6.9 Hz, 7.1 Hz, 1H), 4.45 (dq, *J* = 7.2 Hz, 7.2 Hz, 1H), 4.28 (obs dd, *J* = 6.8 Hz, 9.1 Hz, 1H), 4.23 (obs dd, *J* = 7.3 Hz, 17.0 Hz, 1H), 3.87 (dd, *J* = 8.3 Hz, 8.3 Hz, 1H), 3.69 (dd, *J* = 4.6 Hz, 17.1 Hz, 1H), 3.12 (t, *J* = 9.9 Hz, 1H), 2.17-2.39 (m, 2H), 1.61-1.90 (m, 1H), 1.42 (s, 9H), 1.35 (obs d, *J* = 6.9 Hz, 3H), 1.33 (obs d, *J* = 6.5 Hz, 3H), 1.11 (d, *J* = 5.8 Hz, 3H). HRMS (ESI) exact mass calc. for $C_{26}H_{38}N_4O_7Na$ $[M+Na]^+$: 541.2638, found: 541.2642, Δ = 0.7 ppm.

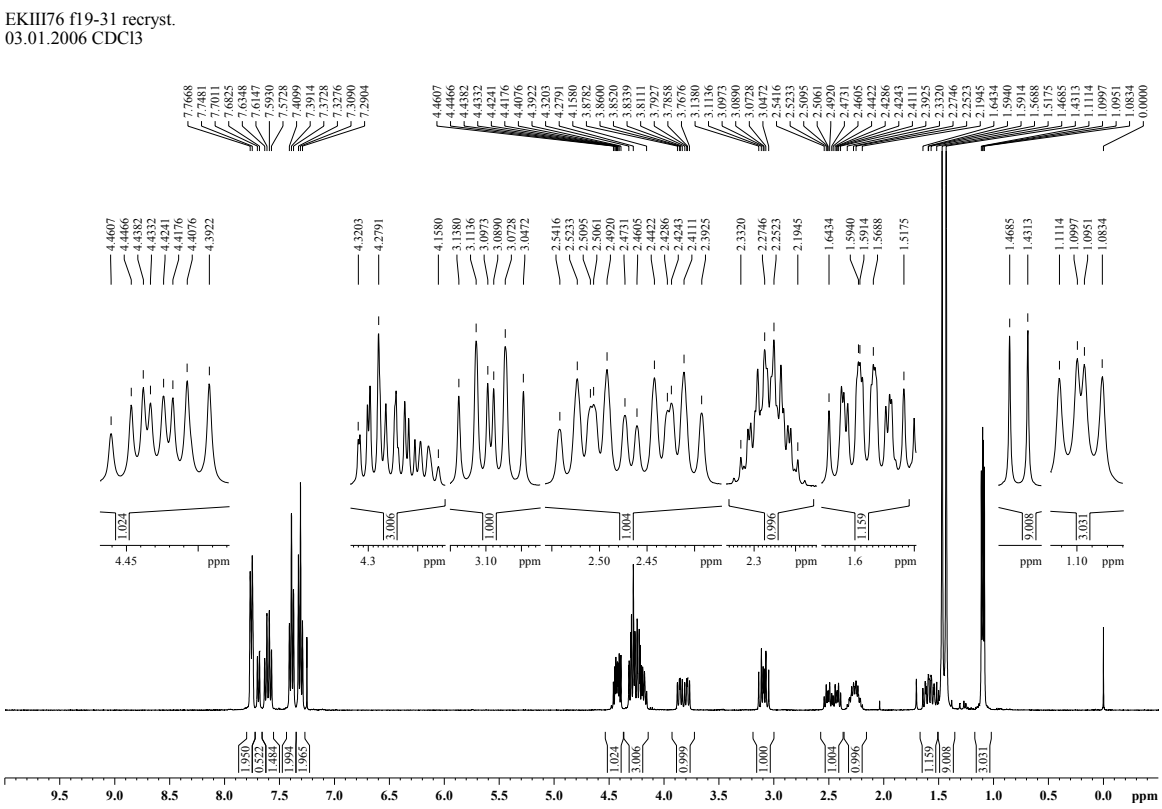


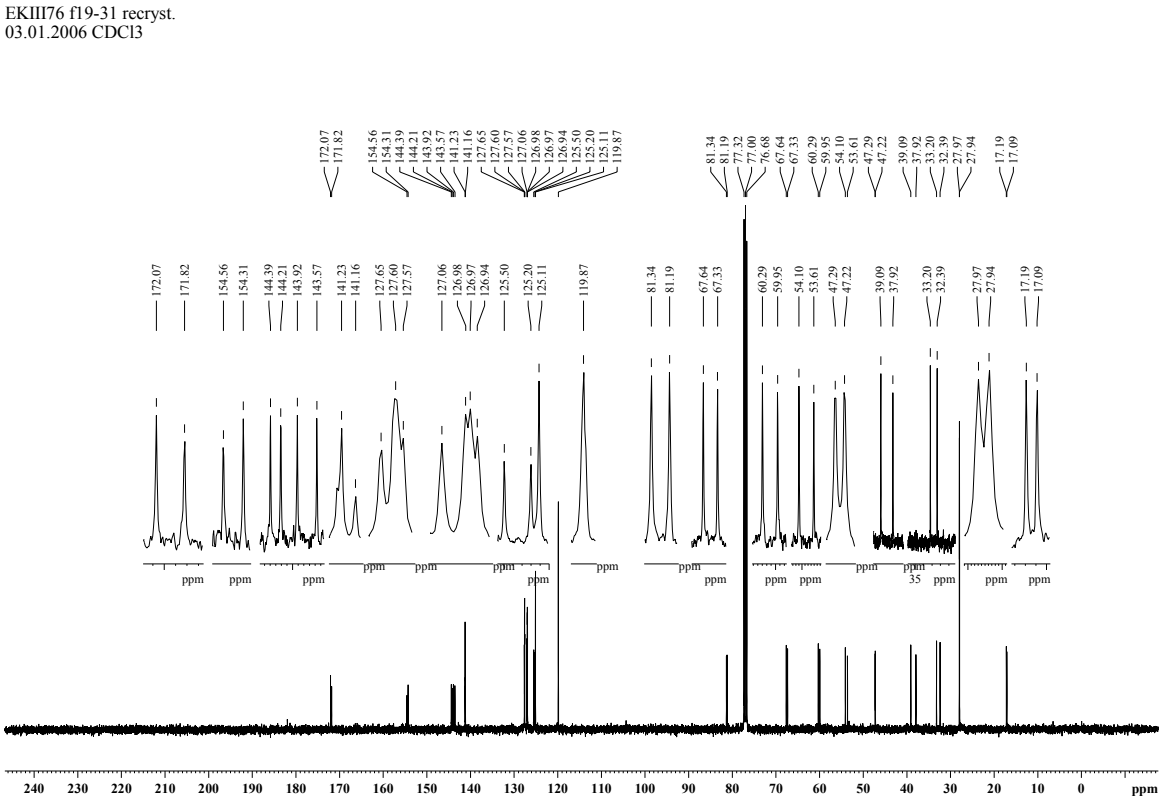
CBz-Ala-(4*R*)-MePro-Gly-Ala-O*t*Bu (A(4*R*)MePGA). In this procedure, Fmoc-(4*R*)-MePro-OH was used instead of Fmoc-(4*S*)-MePro-OH. Yields of A(4*R*)MePGA were found to be after chromatography 0.039 g (15%) and after preparative HPLC 0.022 g (8.4%). R_f = 0.21 (5 % MeOH / EtOAc); ^{1}H NMR (400 MHz, $CDCl_3$) Major conformer: δ 7.27-7.40 (m, 5H), 7.17 (t, *J* = 5.6 Hz, 1H), 6.91 (d, *J* = 7.1 Hz, 1H), 5.66 (d, *J* = 8.0 Hz, 1H), 5.05-5.13 (m, 2H), 4.34-4.58 (m, 3H), 4.11 (dd, *J* = 6.7 Hz, 16.9 Hz, 1H), 3.74 (dd, *J* = 5.2 Hz, 16.8 Hz, 2H), 3.21 (t, *J* = 9.0 Hz, 1H), 2.49-2.69 (m, 1H), 2.34 (ddd, *J* = 2.7 Hz, 6.1 Hz, 12.5 Hz, 1H), 1.57-1.73 (m, 1H), 1.43 (s, 9H), 1.35 (obs d, *J* = 6.8 Hz, 3H), 1.35 (obs d, *J* = 7.1 Hz, 3H), 1.07 (d, *J* = 6.6 Hz, 3H). HRMS (ESI) exact mass calc. for $C_{26}H_{38}N_4O_7Na$ $[M+Na]^+$: 541.2638, found: 541.2633, Δ = -0.9 ppm.

References

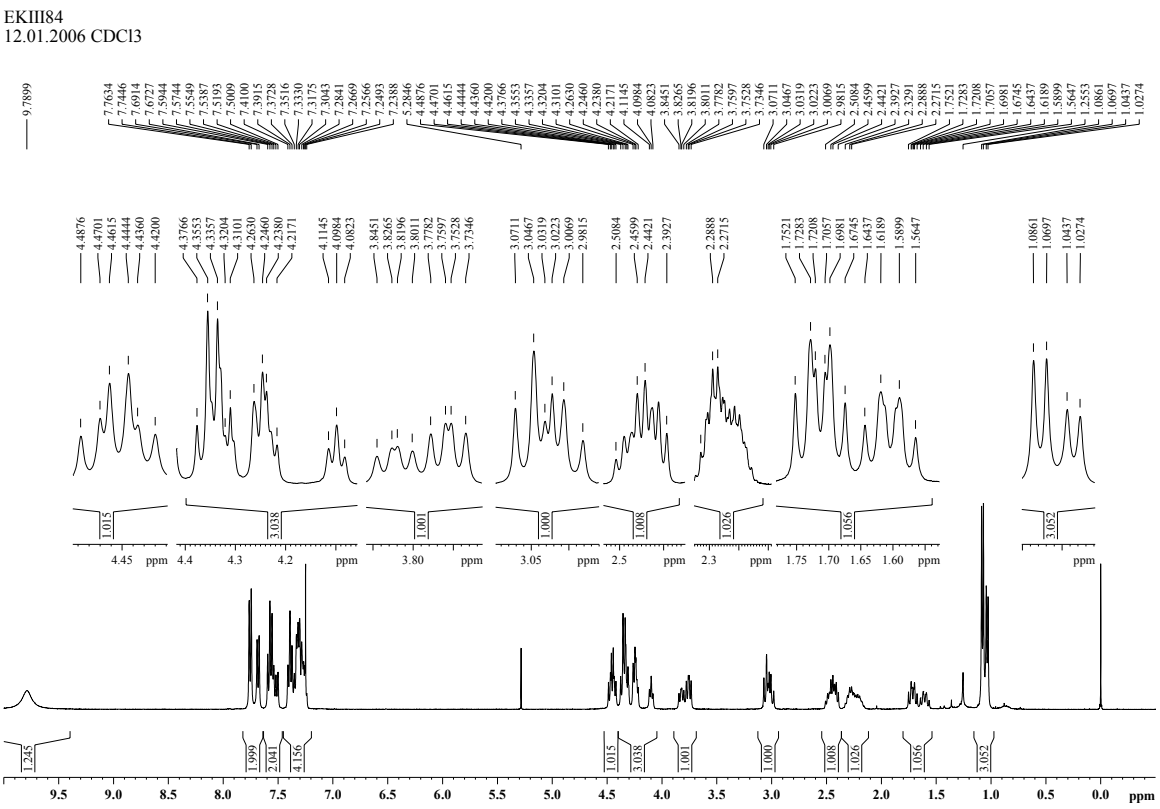
1 A. M. P. Koskinen and H. Rapoport, *J. Org. Chem.*, 1989, **54**, 1859-1866.
2 M. Nevalainen, P. M. Kauppinen and A. M. P. Koskinen, *J. Org. Chem.*, 2001, **66**, 2061-2066.
3 S. C. Bergmeier, A. A. Cobas and H. Rapoport, *J. Org. Chem.*, 1993, **58**, 2369-2376.

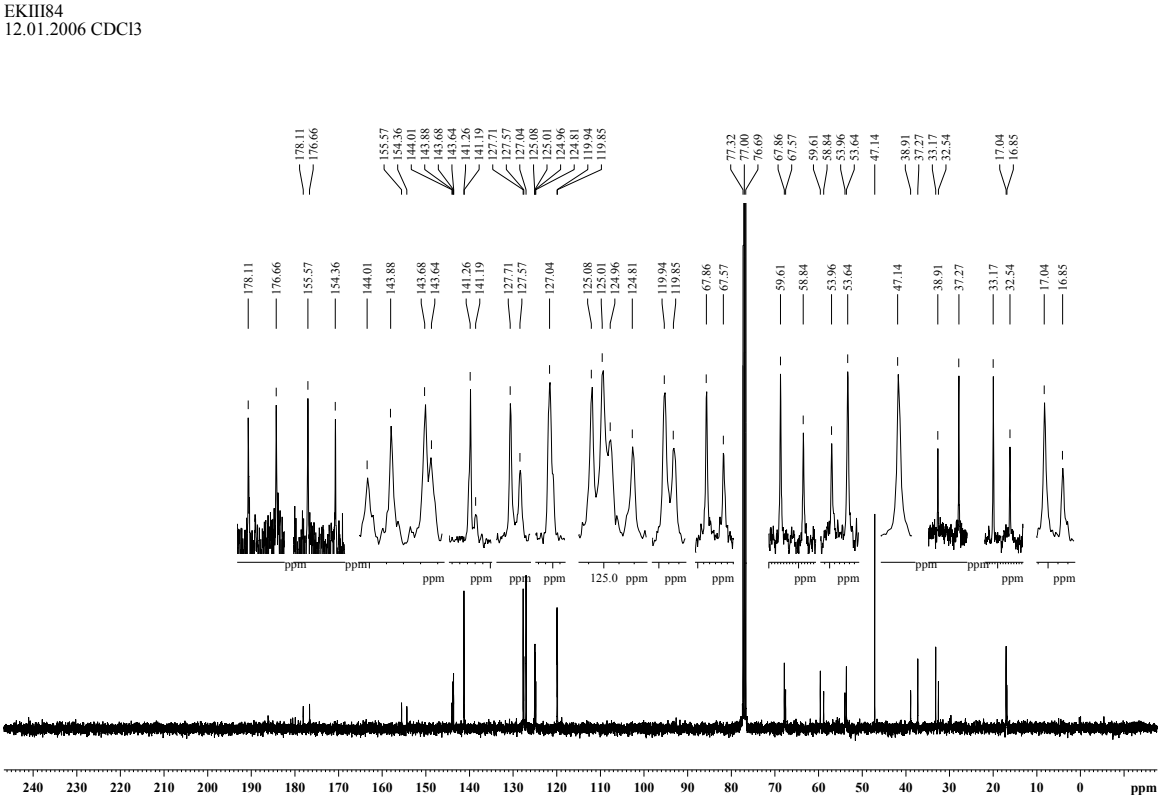
(2*S*,4*S*)-4-Methyl *N*-Fmoc-proline *tert*-butyl ester (3): ^{1}H and ^{13}C NMR ($CDCl_3$, 298 K)





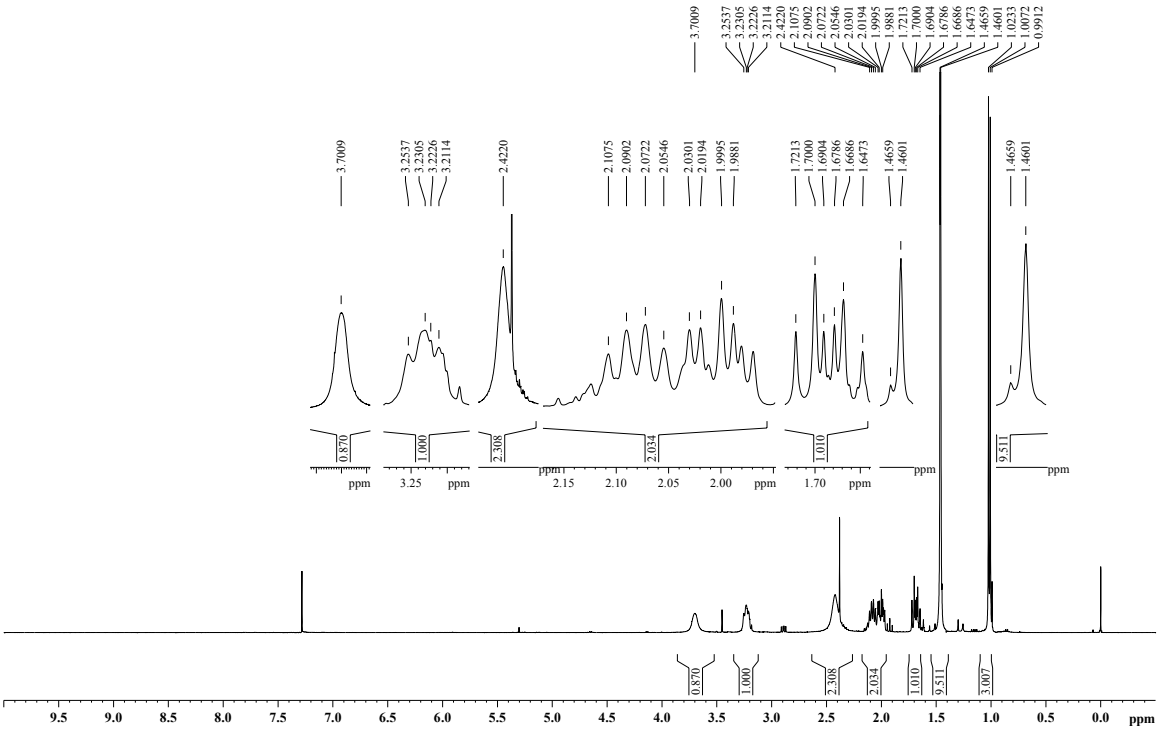
(2*S*,4*S*)-4-Methyl *N*-Fmoc-proline (4): ^{1}H and ^{13}C NMR ($CDCl_3$, 298 K)



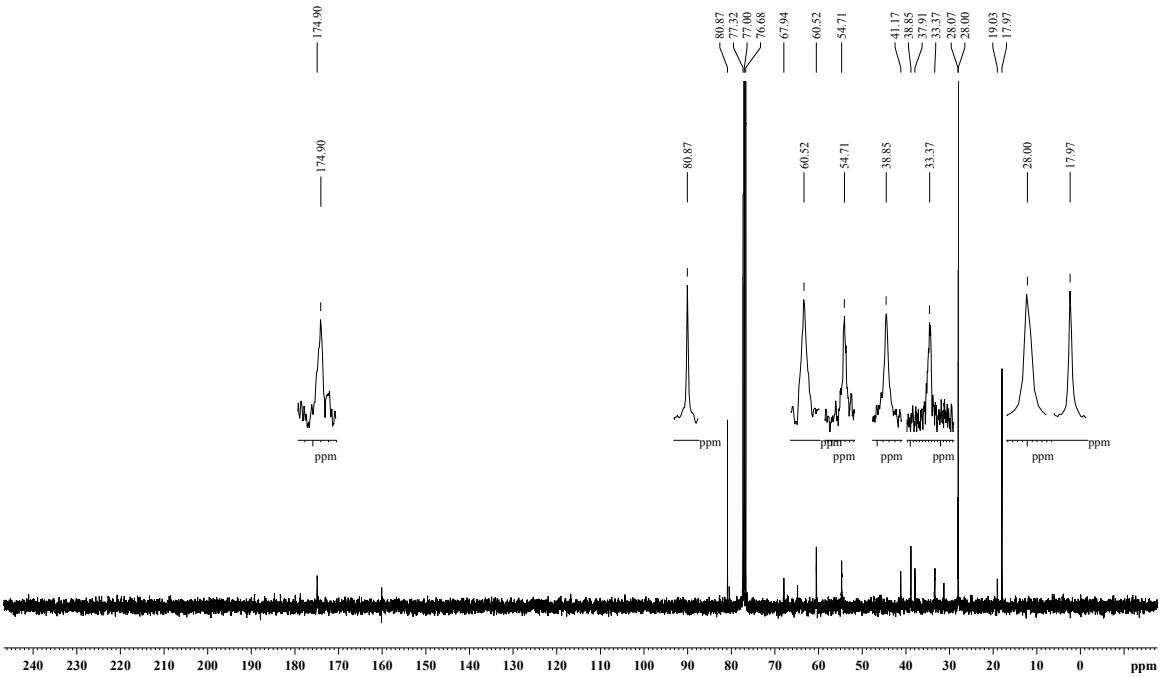


(2*S*,4*R*)-4-Methyl-L-proline *tert*-butyl ester (6): ^{1}H and ^{13}C NMR ($CDCl_3$, 298 K)

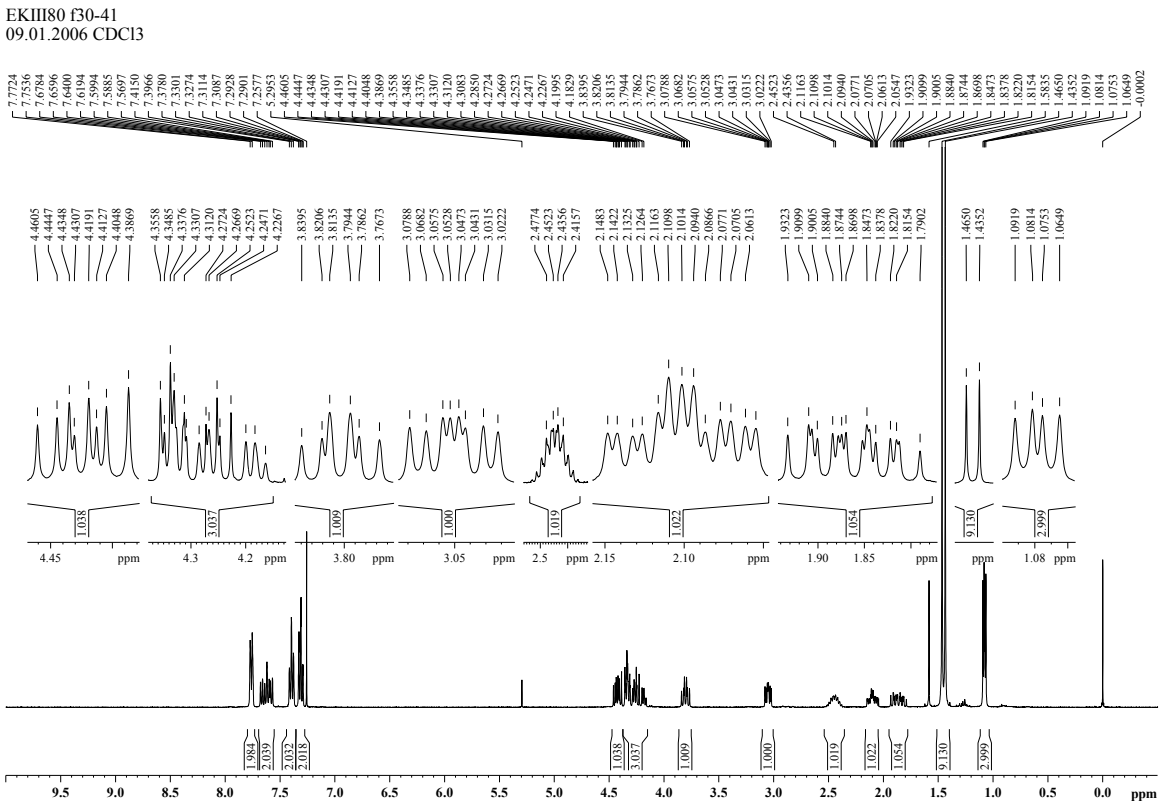
EKIII78
04.01.2006 CDCl3

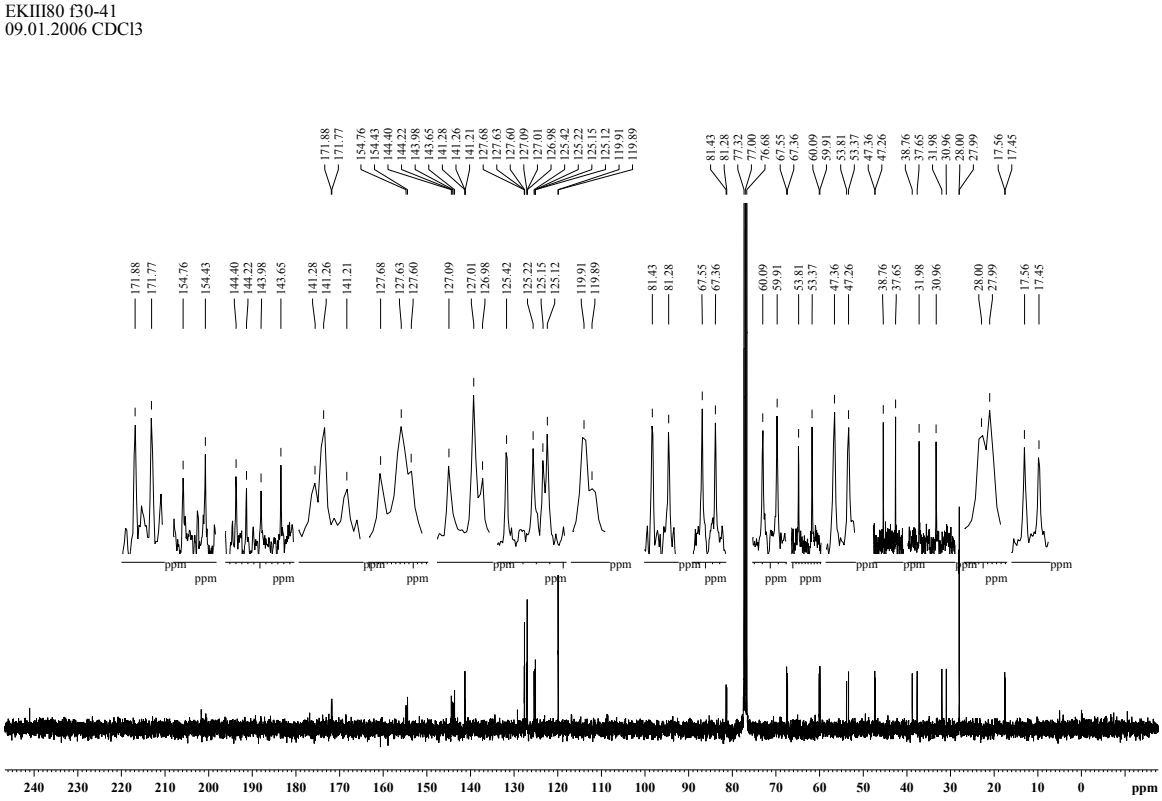


EKIII78
04.01.2006 CDCl3

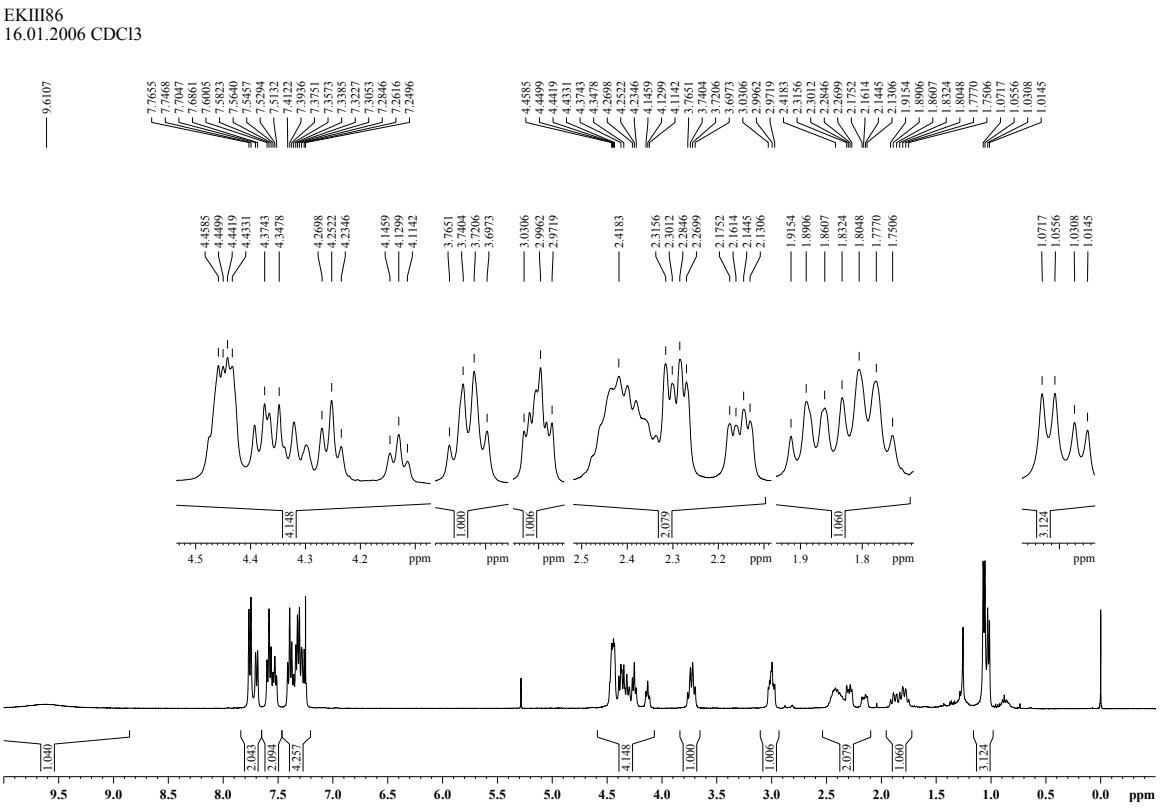


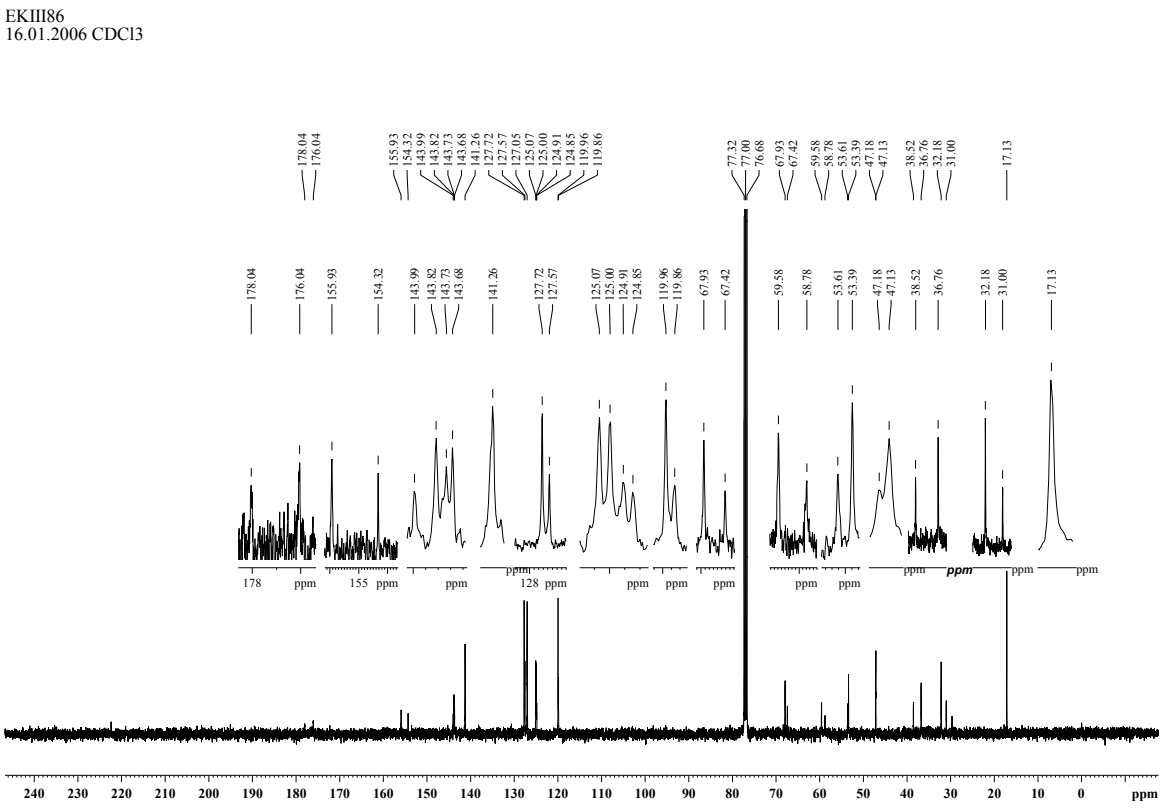
(2*S*,4*R*)-4-Methyl *N*-Fmoc-proline *tert*-butyl ester (7): ^{1}H and ^{13}C NMR ($CDCl_3$, 298 K)



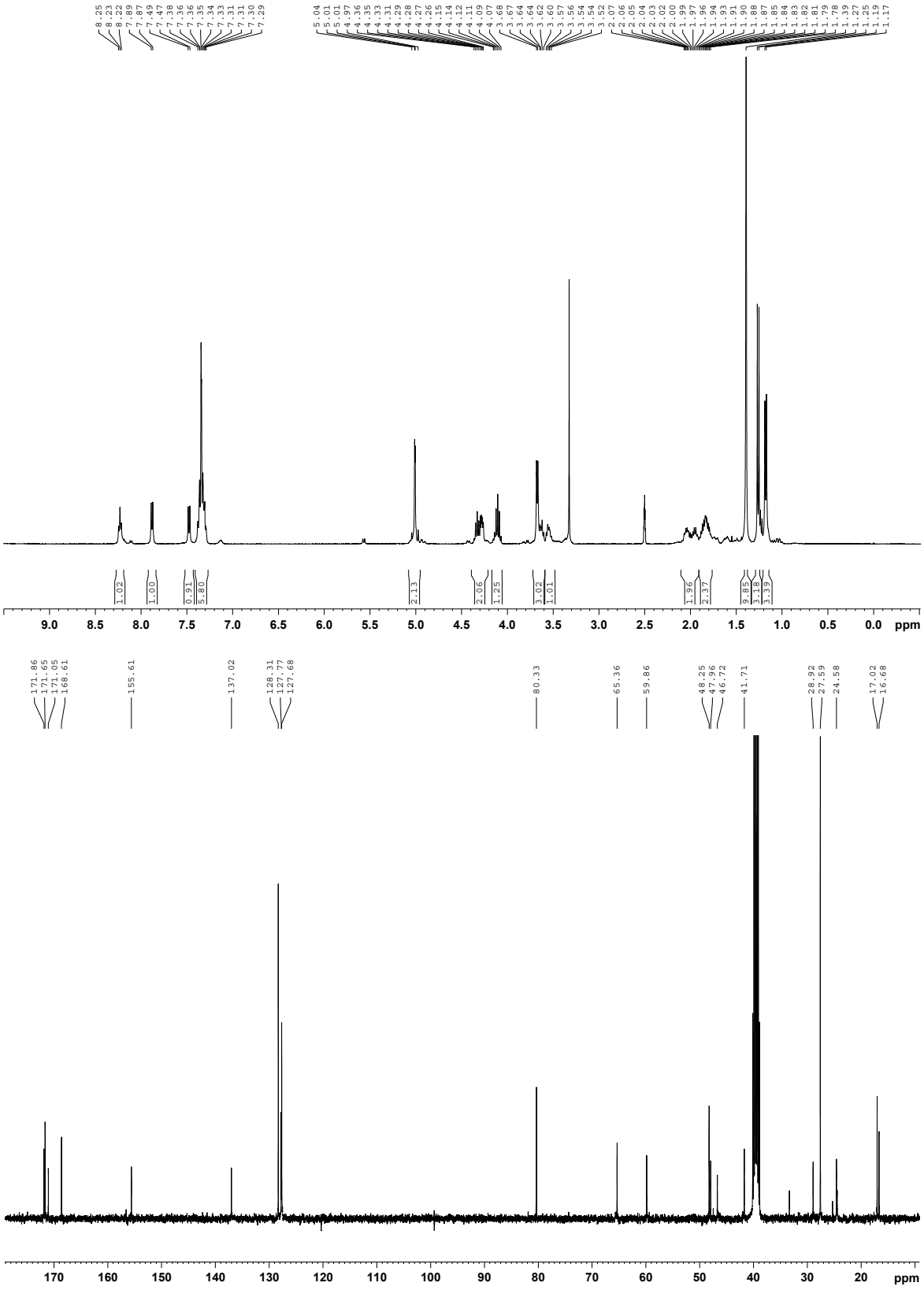


(2*S*,4*R*)-4-Methyl *N*-Fmoc-proline (8): ^{1}H and ^{13}C NMR ($CDCl_3$, 298 K)

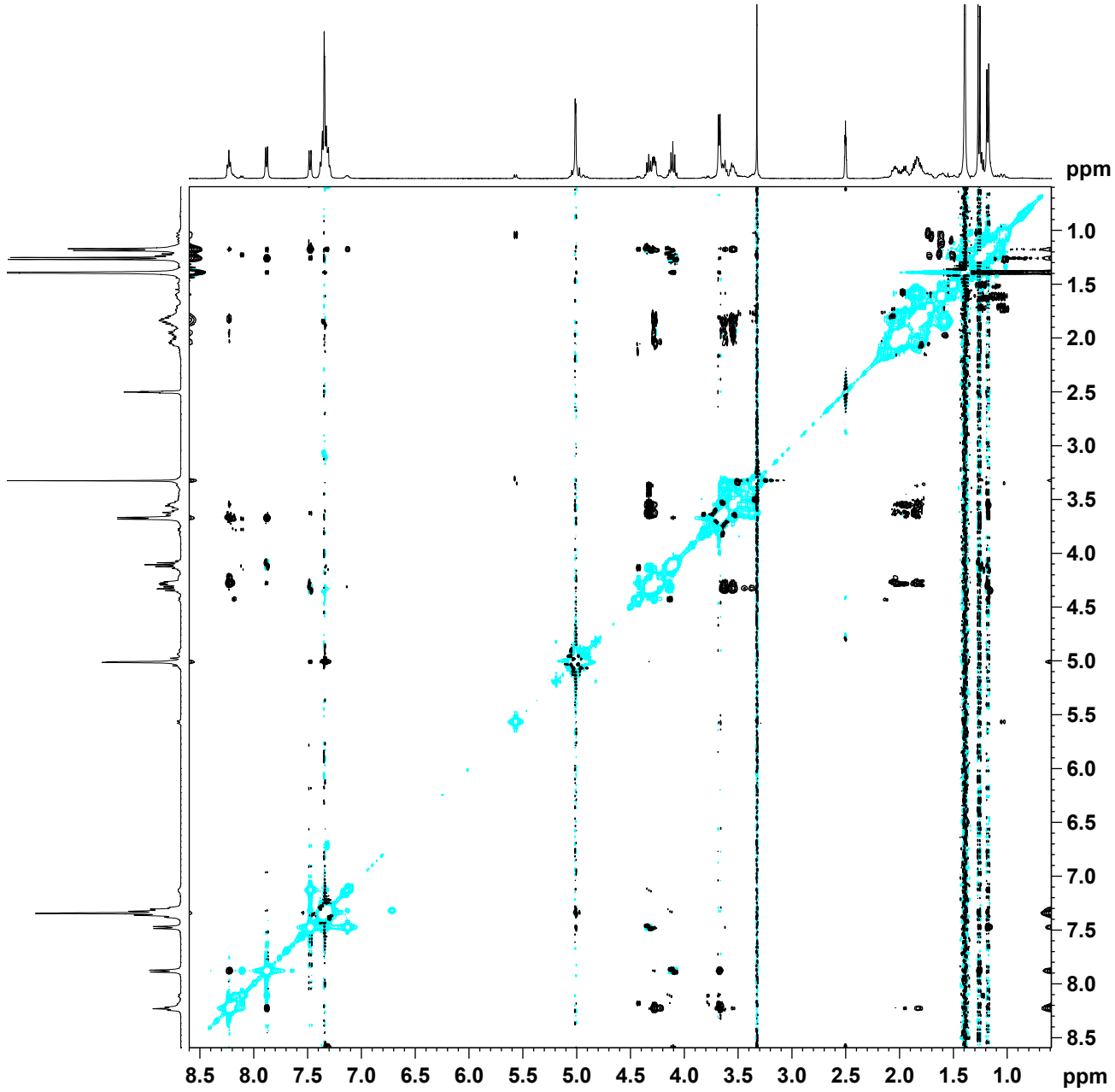




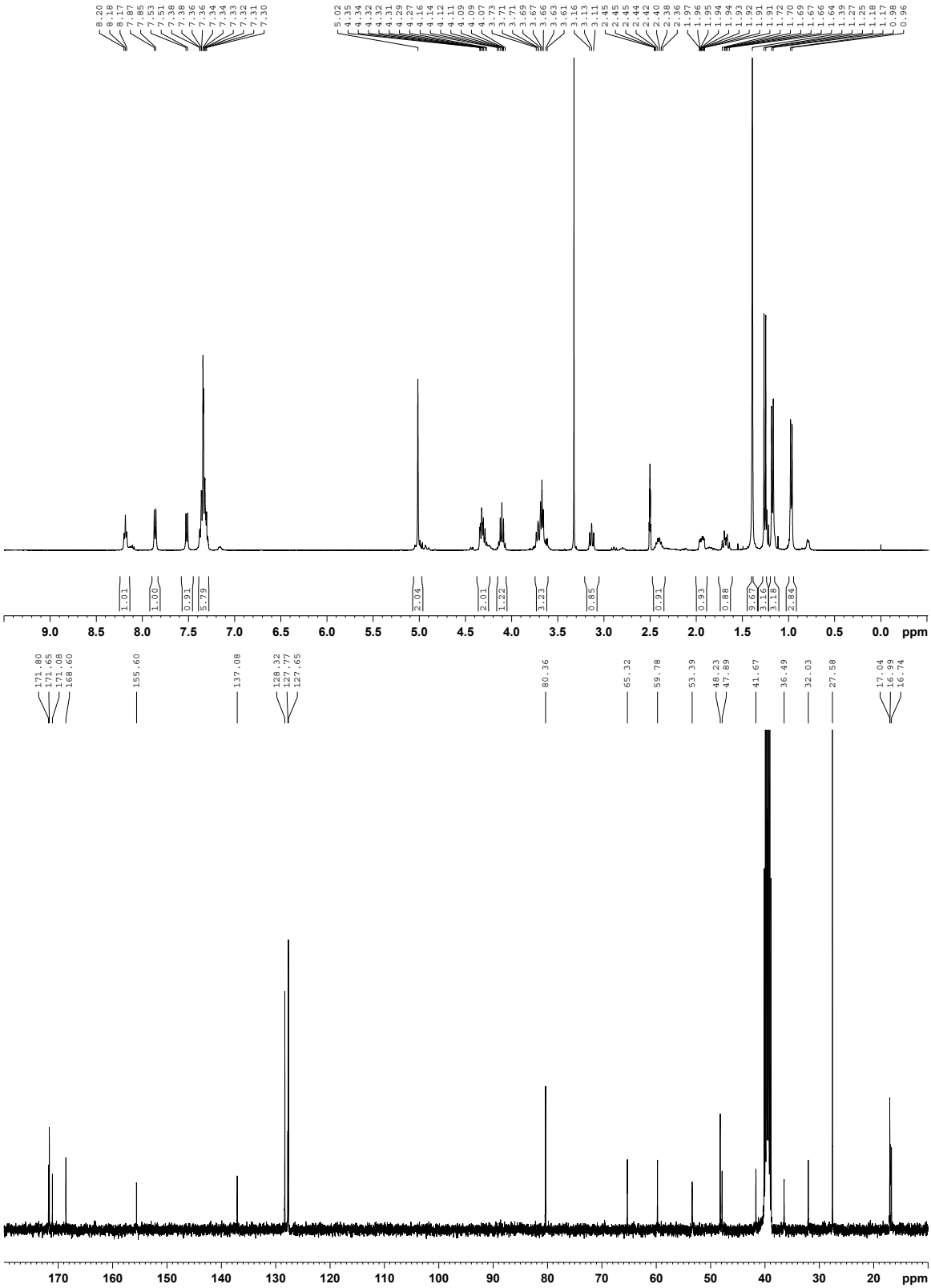
CBz-Ala-Pro-Gly-Ala-O*t*Bu (APGA): ^{1}H and ^{13}C NMR (DMSO-d_6, 299 K)



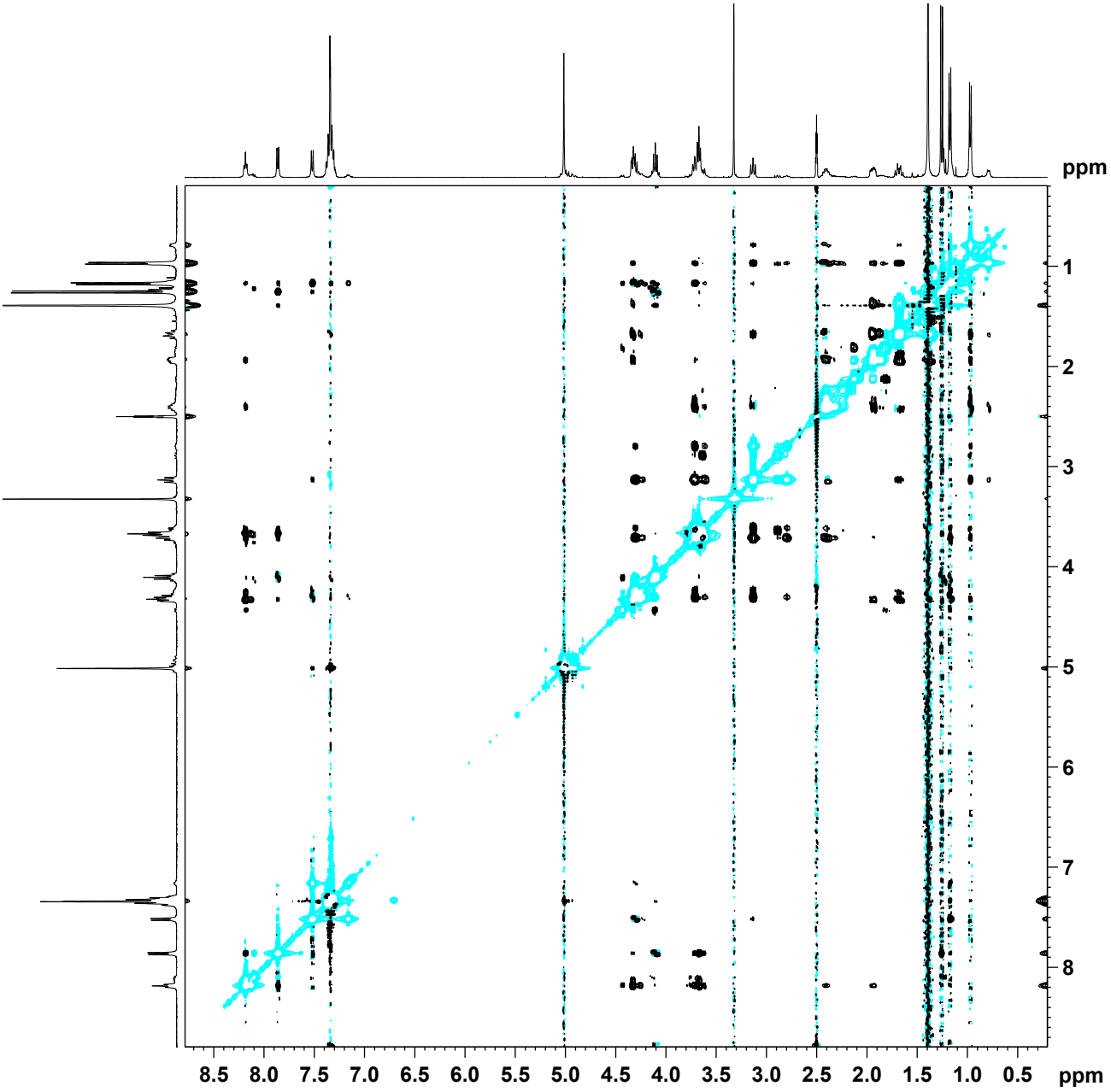
CBz-Ala-Pro-Gly-Ala-O*t*Bu (APGA): Off-Resonance ROESY (Δ = 180 ms, DMSO-d_6, 299 K)



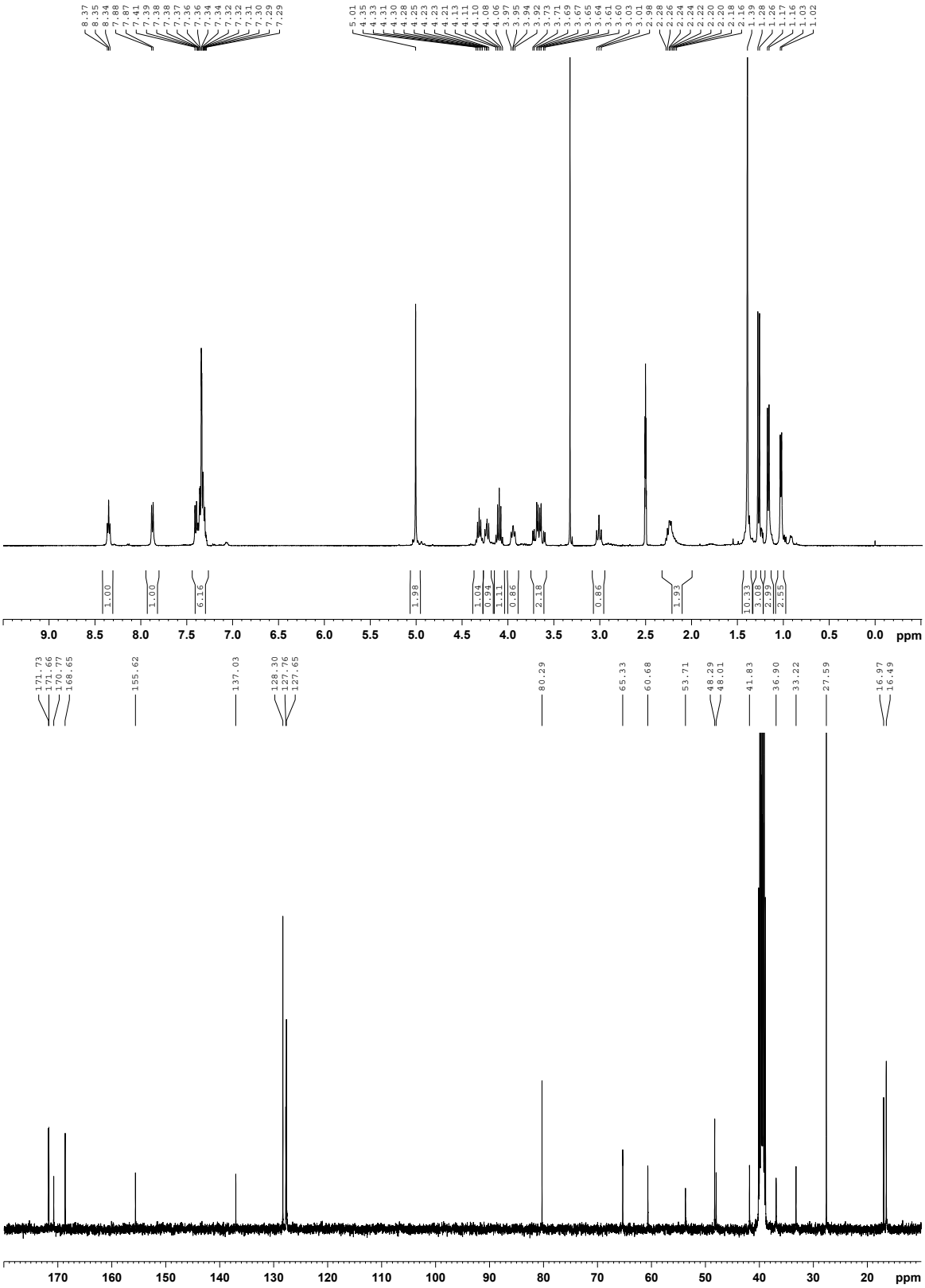
CBz-Ala-(4*R*)-MePro-Gly-Ala-O*t*Bu (A(4*R*)MePGA): ^{1}H and ^{13}C NMR (DMSO-d_6, 299 K)



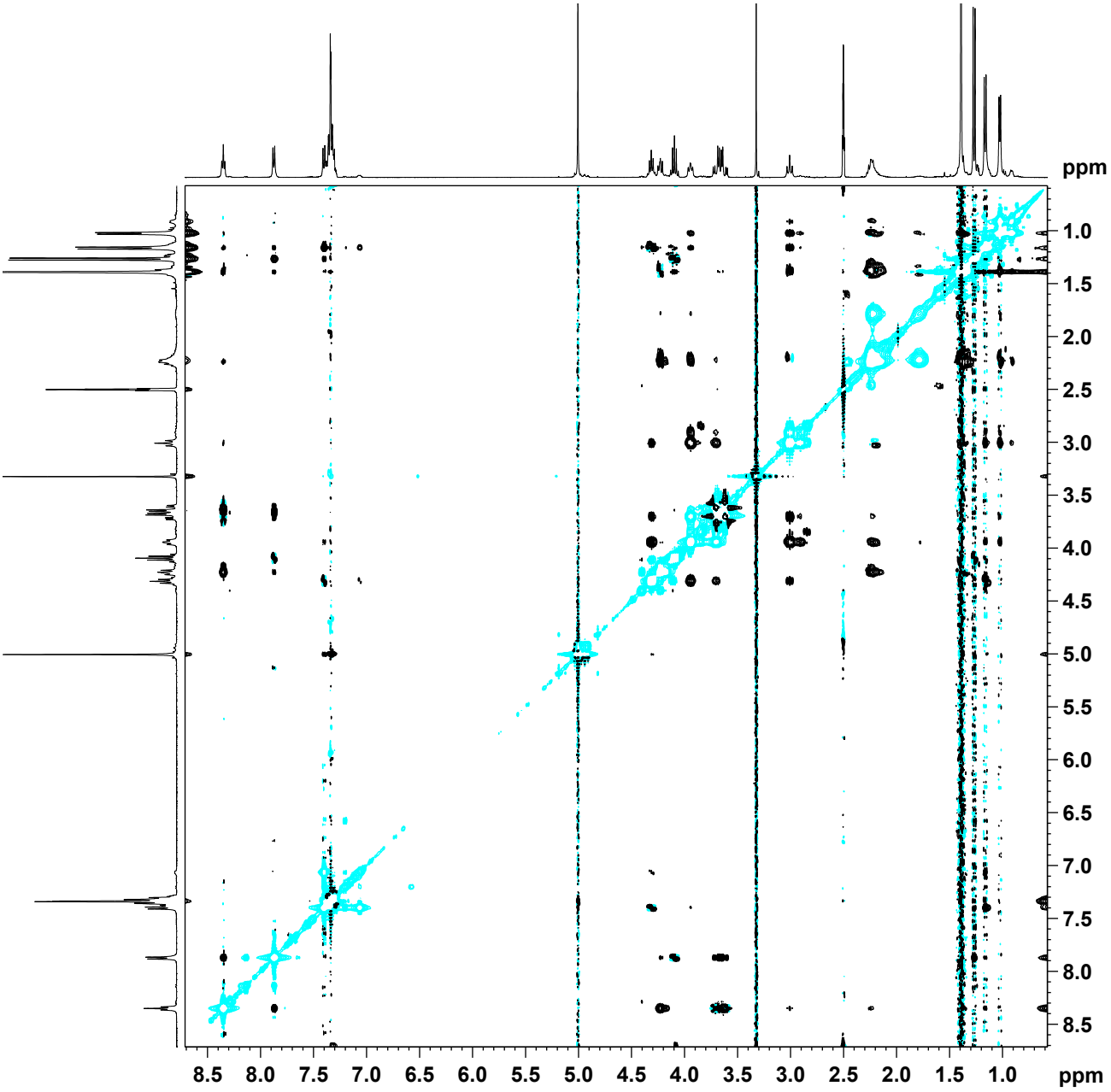
CBz-Ala-(4*R*)-MePro-Gly-Ala-O*t*Bu (A(4*R*)MePGA): Off-Resonance ROESY (Δ = 180 ms, DMSO-d_6, 299 K)



CBz-Ala-(4*S*)-MePro-Gly-Ala-O*t*Bu (A(4*S*)MePGA): ^{1}H and ^{13}C NMR (DMSO-d_6, 299 K)



CBz-Ala-(4*S*)-MePro-Gly-Ala-O*t*Bu (A(4*S*)MePGA): Off-Resonance ROESY (Δ = 180 ms, DMSO-d_6, 299 K)



DFT energies and cartesian coordinates for (4*R*)MePG-I, (4*R*)MePG-II, (4*S*)MePG-I and (4*S*)MePG-II

Table S5. DFT energies for (4*R*)MePG-I, (4*R*)MePG-II, (4*S*)MePG-I and (4*S*)MePG-II.[a]

peptide	E (Hartree)
(4*R*)MePG-I	-820.876773
(4*R*)MePG-II	-820.877489
(4*S*)MePG-I	-820.876429
(4*S*)MePG-II	-820.878225

[a] Calculated at the B3LYP/6-311++G** level of theory using the PBF solvation model with DMSO as solvent.

Cartesian coordinates for (4*R*)MePG-I

C1	16.6990830000000	-13.5889130000000	-7.2622220000000
O2	17.6602000000000	-14.2554320000000	-6.8422970000000
C3	14.3007340000000	-13.5073570000000	-8.0472640000000
N4	15.5382470000000	-14.1943330000000	-7.6150110000000
C5	15.2938690000000	-15.6066890000000	-7.2775890000000
C6	16.0433550000000	-16.6200490000000	-8.1460140000000
O7	15.9592750000000	-17.8189660000000	-7.8723500000000
C8	13.7622220000000	-15.7453840000000	-7.3661060000000
C9	13.3421150000000	-14.6627640000000	-8.3736460000000
H10	13.9057420000000	-12.8924960000000	-7.2299870000000
H11	14.4931970000000	-12.8636750000000	-8.9071940000000
H12	15.6539920000000	-15.8073860000000	-6.2650720000000
H13	13.4596580000000	-16.7506160000000	-7.6610920000000
H14	13.3240840000000	-15.5358500000000	-6.3854620000000
H15	13.5720620000000	-15.0231700000000	-9.3831310000000
H16	16.8257480000000	-15.1678680000000	-9.3450370000000
N17	16.8172820000000	-16.1583450000000	-9.1491540000000
C18	17.7595890000000	-17.0180490000000	-9.8384560000000
C19	19.0292200000000	-17.3544560000000	-9.0474820000000
O20	19.7786140000000	-18.2380030000000	-9.4844410000000
H21	17.2886370000000	-17.9691630000000	-10.0938920000000
H22	18.0638350000000	-16.5386510000000	-10.7718730000000
H23	18.6559710000000	-15.9340430000000	-7.6262300000000
N24	19.2825950000000	-16.6764740000000	-7.9222310000000
C25	11.8655240000000	-14.2812200000000	-8.3078170000000
H26	11.2347680000000	-15.1451900000000	-8.5352690000000
H27	11.6006590000000	-13.9226550000000	-7.3076380000000
H28	11.6253720000000	-13.4929750000000	-9.0277180000000
C29	16.7781180000000	-12.0890960000000	-7.4164900000000
H30	16.0322060000000	-11.5991180000000	-6.7836180000000
H31	17.7737180000000	-11.7573290000000	-7.1248600000000
H32	16.5866610000000	-11.7865370000000	-8.4501150000000
C33	20.4636870000000	-16.9553380000000	-7.1094970000000
H34	21.3818190000000	-16.7839260000000	-7.6769740000000

H35	20.4536100000000	-16.2953380000000	-6.2431490000000
H36	20.4608190000000	-17.9917970000000	-6.7628950000000

Cartesian coordinates for (4*R*)MePG-II

C1	16.6932960000000	-13.4594390000000	-7.4297390000000
O2	17.6668950000000	-14.1953460000000	-7.1841230000000
C3	14.2685600000000	-13.2320790000000	-8.0664000000000
N4	15.4748100000000	-13.9800780000000	-7.6716390000000
C5	15.2239390000000	-15.4283480000000	-7.5602160000000
C6	16.1285460000000	-16.2189900000000	-8.5117960000000
O7	16.2121810000000	-15.9429220000000	-9.7079070000000
C8	13.7266600000000	-15.5652480000000	-7.9267260000000
C9	13.3981310000000	-14.2991760000000	-8.7401860000000
H10	13.7700910000000	-12.8151390000000	-7.1815410000000
H11	14.5218750000000	-12.4132260000000	-8.7400380000000
H12	15.4158000000000	-15.7535850000000	-6.5348550000000
H13	13.5215220000000	-16.4832360000000	-8.4798690000000
H14	13.1335150000000	-15.5892260000000	-7.0077240000000
H15	13.7735570000000	-14.4374840000000	-9.7578070000000
H16	16.7136370000000	-17.4064440000000	-6.9560210000000
N17	16.8006200000000	-17.2493580000000	-7.9503880000000
C18	17.7416370000000	-18.0617720000000	-8.6936600000000
C19	19.1700630000000	-17.5195890000000	-8.7935320000000
O20	20.0270450000000	-18.2288590000000	-9.3428090000000
H21	17.8009510000000	-19.0534940000000	-8.2407890000000
H22	17.3841090000000	-18.1902220000000	-9.7181460000000
H23	18.6966750000000	-15.7426200000000	-7.8923430000000
N24	19.4387510000000	-16.3078810000000	-8.2976320000000
C25	11.9149370000000	-13.9449270000000	-8.7960500000000
H26	11.3472140000000	-14.7313430000000	-9.3015750000000
H27	11.5024790000000	-13.8205820000000	-7.7892840000000
H28	11.7549640000000	-13.0126240000000	-9.3466140000000
C29	16.8355440000000	-11.9540890000000	-7.4678420000000
H30	16.0667730000000	-11.4625610000000	-6.8661000000000
H31	17.8200930000000	-11.6866600000000	-7.0882220000000
H32	16.7446780000000	-11.5853140000000	-8.4946920000000
C33	20.7739560000000	-15.7244320000000	-8.3915340000000
H34	20.7748370000000	-14.7741920000000	-7.8597590000000
H35	21.5176220000000	-16.3861170000000	-7.9409090000000
H36	21.0580090000000	-15.5466000000000	-9.4327810000000

Cartesian coordinates for (4*S*)MePG-I

C1	-10.2097270000000	3.1664400000000	3.5461610000000
O2	-10.8374510000000	3.8384680000000	4.3820650000000
C3	-7.9890750000000	2.5256230000000	2.5237550000000
N4	-8.8833200000000	3.3670560000000	3.3510880000000
C5	-8.1124490000000	4.2307550000000	4.2671740000000
C6	-8.4827350000000	5.7087030000000	4.2661020000000
O7	-8.0857540000000	6.4259760000000	5.1865680000000
C8	-6.6443380000000	3.9950230000000	3.8567210000000

C9	-6.6505390000000	2.5733990000000	3.2756830000000
H10	-8.2772530000000	3.8978140000000	5.2961460000000
H11	-9.5292980000000	5.5600350000000	2.5190050000000
N12	-9.2342070000000	6.1864060000000	3.2538590000000
C13	-9.7323490000000	7.5482960000000	3.2655620000000
C14	-10.8456140000000	7.8363340000000	4.2788320000000
O15	-11.1555350000000	9.0173910000000	4.4891430000000
H16	-8.9240080000000	8.2464620000000	3.4911370000000
H17	-10.1129060000000	7.7943530000000	2.2721350000000
H18	-11.1689130000000	5.8568800000000	4.6682120000000
N19	-11.4458070000000	6.8091260000000	4.8886090000000
H20	-8.3793950000000	1.5131450000000	2.4316470000000
H21	-7.8908370000000	2.9558560000000	1.5201540000000
H22	-6.3459630000000	4.7143240000000	3.0862190000000
H23	-5.9690940000000	4.1086300000000	4.7049120000000
H24	-6.7074380000000	1.8630220000000	4.1086730000000
C25	-5.4373740000000	2.2374380000000	2.4131480000000
H26	-4.5182610000000	2.2953010000000	3.0034190000000
H27	-5.5096980000000	1.2242420000000	2.0063010000000
H28	-5.3461850000000	2.9362290000000	1.5746450000000
C29	-10.8989010000000	2.1140360000000	2.7096090000000
H30	-10.6061210000000	2.1664340000000	1.6589090000000
H31	-10.6435410000000	1.1162840000000	3.0821010000000
H32	-11.9758400000000	2.2523730000000	2.7959710000000
C33	-12.4891260000000	7.0202030000000	5.8889130000000
H34	-12.8340620000000	6.0492400000000	6.2415460000000
H35	-12.1034160000000	7.5894560000000	6.7397490000000
H36	-13.3358220000000	7.5669170000000	5.4660040000000

Cartesian coordinates for (4*S*)MePG-II

C1	-10.0804270000000	3.1448320000000	3.7849180000000
O2	-10.6428260000000	4.0358220000000	4.4468120000000
C3	-8.0247920000000	2.2390390000000	2.6618810000000
N4	-8.7824540000000	3.2408140000000	3.4376800000000
C5	-7.9379440000000	4.3747670000000	3.8621090000000
C6	-8.5899800000000	5.7264830000000	3.5553250000000
O7	-8.9736970000000	6.0139780000000	2.4226400000000
C8	-6.6458890000000	4.1670380000000	3.0518000000000
C9	-6.5568910000000	2.6439960000000	2.8645720000000
H10	-7.7444620000000	4.2980230000000	4.9372200000000
H11	-8.3667720000000	6.2784170000000	5.5086330000000
N12	-8.6681440000000	6.5860490000000	4.5941800000000
C13	-9.2587880000000	7.9030400000000	4.4720990000000
C14	-10.7774340000000	7.9901680000000	4.6486290000000
O15	-11.2999620000000	9.1148760000000	4.6636320000000
H16	-8.8108440000000	8.5681930000000	5.2129710000000
H17	-9.0334890000000	8.3122340000000	3.4838770000000
H18	-11.0145970000000	5.9594950000000	4.7013730000000
N19	-11.4794020000000	6.8613770000000	4.7736790000000
H20	-8.2277650000000	1.2305060000000	3.0251780000000
H21	-8.3058910000000	2.2912040000000	1.6035220000000

H22	-6.7512200000000	4.6542820000000	2.0782010000000
H23	-5.7758330000000	4.5884690000000	3.5558900000000
H24	-6.2058180000000	2.1993520000000	3.8026320000000
C25	-5.6412610000000	2.1993380000000	1.7254660000000
H26	-4.6124960000000	2.5271760000000	1.9010240000000
H27	-5.6337450000000	1.1090540000000	1.6289990000000
H28	-5.9740800000000	2.6208370000000	0.7705890000000
C29	-10.8296670000000	1.9039860000000	3.3547360000000
H30	-10.6642200000000	1.6708240000000	2.3009650000000
H31	-10.4964850000000	1.0435890000000	3.9442760000000
H32	-11.8932800000000	2.0592960000000	3.5303620000000
C33	-12.9318610000000	6.8797720000000	4.9181610000000
H34	-13.2751700000000	5.8609860000000	5.0894750000000
H35	-13.2279770000000	7.5028240000000	5.7652430000000
H36	-13.4163310000000	7.2701380000000	4.0184170000000

Additional NAMFIS analyses

In order to test the strength of the NAMFIS analyses, a series of additional NAMFIS calculations were performed. First, the 13, 16, and 13 conformation datasets of APGA, A(4*R*)MePGA, and A(4*S*)MePGA, respectively, were resubmitted to separate NAMFIS calculations to examine how well the NMR-derived parameters are reproduced. Second, the most populated members of each of the 6 structural motifs were similarly evaluated. The results, in terms of the differences between the experimental and the NAMFIS calculated NMR parameters, are given in Tables S6, S7, and S8 for APGA, A(4*R*)MePGA, and A(4*S*)MePGA, respectively. In the case of the 13, 16, and 13 member conformer pools, the maximum differences between the experimental and predicted NOE distances and coupling constant values are only 0.25 Å and 0.93 Hz for APGA, 0.24 Å and 0.53 Hz for A(4*R*)MePGA, and 0.26 Å and 0.62 Hz for A(4*S*)MePGA. The SSD values for each peptide are identical to those from the original calculations with 2259, 1996, and 2485 conformers for APGA, A(4*R*)MePGA, and A(4*S*)MePGA, respectively. In the case of the 6 member conformer pools, the maximum deviations are slightly larger, being 0.88 Å and 1.31 Hz for APGA, 0.64 Å and 1.05 Hz for A(4*R*)MePGA, and 0.96 Å and 0.70 Hz for A(4*S*)MePGA. In this case, also the SSD values are somewhat larger. These data strongly suggest that NAMFIS is capable of selecting a small collection of very discriminating conformers from a large pool. These conformers collectively retain both the important features of NMR-derived structure and characteristics of conformational families.

Table S6. Comparison between the experimental and the NAMFIS calculated NOE and coupling constant values for APGA.

NOE's		experimental (Å)	13 member conf. pool difference (Å)	6 member conf. pool difference (Å)
Ala^1-NH	Ala^1-Hα	2.60	0.07	0.10
Ala^1-NH	Ala^1-CH_3	2.75	0.18	0.10
Ala^1-NH	Pro-Hδ$^{Pro(R)}$	3.81	0.09	0.88
Ala^1-NH	Ala^2-CH_3	4.12	0.05	0.81
Ala^1-Hα	Ala^1-CH_3	2.52	0.11	0.10
Ala^1-Hα	Pro-Hδ$^{Pro(R)}$	2.18	0.12	0.10
Ala^1-Hα	Pro-Hδ$^{Pro(S)}$	2.30	0.18	0.19
Ala^1-CH_3	Pro-Hδ$^{Pro(R)}$	3.70	0.25	0.40
Ala^1-CH_3	Pro-Hδ$^{Pro(S)}$	3.14	0.17	0.17
Ala^1-CH_3	Gly-NH	4.26	0.09	0.16
Ala^1-CH_3	Ala^2-NH	4.31	0.14	0.10
Pro-Hα	Gly-NH	2.07	0.20	0.19
Pro-Hα	Ala^2-NH	3.77	0.06	0.10
Pro-Hβ$^{Pro(S)}$	Gly-NH	3.02	0.20	0.20
Pro-Hδ$^{Pro(S)}$	Gly-NH	3.45	0.08	0.11
Gly-NH	Ala^2-NH	2.61	0.07	0.16
Ala^2-NH	Ala^2-Hα	2.48	0.05	0.02
Ala^2-NH	Ala^2-CH_3	2.70	0.24	0.21
Ala^2-Hα	Ala^2-CH_3	2.45	0.18	0.17
J's				
Ala^1-NH	Ala^1-Hα	7.36	0.28	0.71
Gly-NH	Gly-Hα$^{Pro(R)}$	5.97	0.71	1.31
Gly-NH	Gly-Hα$^{Pro(S)}$	5.97	0.17	0.08
Ala^2-NH	Ala^2-Hα	6.89	0.00	0.39
Pro-Hα	Pro-Hβ$^{Pro(R)}$	8.26	0.25	0.12
Pro-Hα	Pro-Hβ$^{Pro(S)}$	4.79	0.22	0.37
Pro-Hβ$^{Pro(R)}$	Pro-Hγ$^{Pro(R)}$	6.91	0.22	0.23
Pro-Hβ$^{Pro(R)}$	Pro-Hγ$^{Pro(S)}$	8.09	0.71	1.16
Pro-Hβ$^{Pro(S)}$	Pro-Hγ$^{Pro(R)}$	6.00	0.93	0.37
Pro-Hβ$^{Pro(S)}$	Pro-Hγ$^{Pro(S)}$	6.77	0.34	0.34
Pro-Hγ$^{Pro(R)}$	Pro-Hδ$^{Pro(R)}$	7.11	0.04	0.11
Pro-Hγ$^{Pro(R)}$	Pro-Hδ$^{Pro(S)}$	5.69	0.85	0.49
Pro-Hγ$^{Pro(S)}$	Pro-Hδ$^{Pro(R)}$	7.46	0.70	1.10
Pro-Hγ$^{Pro(S)}$	Pro-Hδ$^{Pro(S)}$	7.17	0.02	0.15
SSD			8.7	13.6

Table S7. Comparison between the experimental and the NAMFIS calculated NOE and coupling constant values for A(4*R*)MePGA.

NOE's		experimental (Å)	16 member conf. pool difference (Å)	6 member conf. pool difference (Å)
Ala^1-NH	Ala^1-Hα	2.60	0.09	0.01
Ala^1-NH	Ala^1-CH_3	2.76	0.21	0.21
Ala^1-NH	Pro-Hδ$^{Pro(R)}$	3.55	0.10	0.12
Ala^1-NH	Ala^2-CH_3	4.44	0.04	0.64
Ala^1-Hα	Ala^1-CH_3	2.51	0.12	0.12
Ala^1-Hα	Pro-Hδ$^{Pro(S)}$	2.24	0.15	0.19
Ala^1-CH_3	Pro-Hδ$^{Pro(R)}$	4.47	0.02	0.02
Ala^1-CH_3	Pro-Hδ$^{Pro(S)}$	3.12	0.14	0.10
Ala^1-CH_3	Gly-NH	4.39	0.03	0.01
Ala^1-CH_3	Ala^2-NH	4.40	0.12	0.51
Pro-Hα	Pro-Hβ$^{Pro(R)}$	2.25	0.09	0.10
Pro-Hα	Pro-Hβ$^{Pro(S)}$	2.71	0.09	0.11
Pro-Hα	Pro-4-CH_3	3.79	0.01	0.17
Pro-Hα	Gly-NH	2.15	0.14	0.16
Pro-Hα	Ala^2-NH	3.67	0.14	0.07
Pro-Hβ$^{Pro(R)}$	Pro-Hβ$^{Pro(S)}$	1.78	0.01	0.01
Pro-Hβ$^{Pro(R)}$	Pro-Hγ$^{Pro(S)}$	2.90	0.06	0.02
Pro-Hβ$^{Pro(R)}$	Pro-Hδ$^{Pro(R)}$	2.85	0.12	0.19
Pro-Hβ$^{Pro(R)}$	Pro-4-CH_3	2.97	0.07	0.07
Pro-Hβ$^{Pro(S)}$	Pro-Hγ$^{Pro(S)}$	2.36	0.04	0.04
Pro-Hβ$^{Pro(S)}$	Pro-Hδ$^{Pro(S)}$	3.38	0.14	0.04
Pro-Hβ$^{Pro(S)}$	Pro-4-CH_3	3.49	0.13	0.09
Pro-Hβ$^{Pro(S)}$	Gly-NH	3.17	0.13	0.03
Pro-Hγ$^{Pro(S)}$	Pro-Hδ$^{Pro(R)}$	2.92	0.03	0.01
Pro-Hγ$^{Pro(S)}$	Pro-Hδ$^{Pro(S)}$	2.29	0.07	0.08
Pro-Hγ$^{Pro(S)}$	Pro-4-CH_3	2.66	0.03	0.03
Pro-Hγ$^{Pro(S)}$	Gly-NH	3.12	0.20	0.06
Pro-Hδ$^{Pro(R)}$	Pro-Hδ$^{Pro(S)}$	1.78	0.00	0.00
Pro-Hδ$^{Pro(R)}$	Pro-4-CH_3	2.94	0.08	0.08
Pro-Hδ$^{Pro(S)}$	Pro-4-CH_3	3.22	0.20	0.25
Gly-NH	Ala^2-NH	2.72	0.11	0.24
Ala^2-NH	Ala^2-Hα	2.47	0.05	0.02
Ala^2-NH	Ala^2-CH_3	2.70	0.24	0.29
Ala^2-Hα	Ala^2-CH_3	2.45	0.17	0.17
J's				
Ala^1-NH	Ala^1-Hα	7.55	0.46	0.91
Gly-NH	Gly-Hα$^{Pro(R)}$	5.93	0.53	0.05
Gly-NH	Gly-Hα$^{Pro(S)}$	6.05	0.10	0.24
Ala^2-NH	Ala^2-Hα	6.96	0.07	0.42
Pro-Hα	Pro-Hβ$^{Pro(R)}$	8.90	0.53	0.65
Pro-Hα	Pro-Hβ$^{Pro(S)}$	3.02	0.45	0.45
Pro-Hβ$^{Pro(R)}$	Pro-Hγ$^{Pro(S)}$	9.53	0.01	1.05
Pro-Hβ$^{Pro(S)}$	Pro-Hγ$^{Pro(S)}$	6.28	0.22	0.22
Pro-Hγ$^{Pro(S)}$	Pro-Hδ$^{Pro(R)}$	8.22	0.22	0.72
Pro-Hγ$^{Pro(S)}$	Pro-Hδ$^{Pro(S)}$	7.48	0.07	0.11
SSD			7.9	12.3

Table S7. Comparison between the experimental and the NAMFIS calculated NOE and coupling constant values for A(4*S*)MePGA.

NOE's		experimental (Å)	13 member conf. pool difference (Å)	6 member conf. pool difference (Å)
Ala^1-NH	Ala^1-Hα	2.58	0.07	0.33
Ala^1-NH	Ala^1-CH_3	2.76	0.14	0.08
Ala^1-NH	Pro-Hδ$^{Pro(R)}$	3.78	0.05	0.96
Ala^1-NH	Ala^2-CH_3	3.96	0.02	0.00
Ala^1-Hα	Ala^1-CH_3	2.52	0.11	0.11
Ala^1-Hα	Pro-Hδ$^{Pro(R)}$	2.02	0.13	0.12
Ala^1-Hα	Pro-Hδ$^{Pro(S)}$	2.73	0.06	0.06
Ala^1-CH_3	Pro-Hδ$^{Pro(R)}$	3.87	0.08	0.05
Ala^1-CH_3	Pro-Hδ$^{Pro(S)}$	3.10	0.15	0.19
Ala^1-CH_3	Gly-NH	4.31	0.15	0.01
Ala^1-CH_3	Ala^2-NH	4.44	0.07	0.25
Pro-Hα	Pro-Hβ$^{Pro(S)}$	2.78	0.24	0.26
Pro-Hα	Gly-NH	2.02	0.24	0.25
Pro-Hα	Ala^2-NH	3.58	0.14	0.13
Pro-Hβ$^{Pro(R)}$	Gly-NH	3.41	0.17	0.09
Pro-Hβ$^{Pro(S)}$	Pro-Hδ$^{Pro(S)}$	2.65	0.22	0.19
Pro-Hβ$^{Pro(S)}$	Gly-NH	3.12	0.19	0.48
Pro-Hδ$^{Pro(R)}$	Pro-Hδ$^{Pro(S)}$	1.78	0.00	0.01
Pro-Hδ$^{Pro(R)}$	Pro-4-CH_3	3.58	0.26	0.28
Pro-Hδ$^{Pro(S)}$	Pro-4-CH_3	2.90	0.01	0.00
Pro-Hδ$^{Pro(S)}$	Gly-NH	3.55	0.12	0.08
Gly-NH	Ala^2-NH	2.50	0.03	0.07
Ala^2-NH	Ala^2-Hα	2.48	0.04	0.04
Ala^2-NH	Ala^2-CH_3	2.70	0.26	0.21
Ala^2-Hα	Ala^2-CH_3	2.46	0.16	0.16
J's				
Ala^1-NH	Ala^1-Hα	7.13	0.34	0.25
Gly-NH	Gly-Hα$^{Pro(R)}$	5.70	0.62	0.56
Gly-NH	Gly-Hα$^{Pro(S)}$	6.24	0.12	0.07
Ala^2-NH	Ala^2-Hα	6.77	0.00	0.22
Pro-Hα	Pro-Hβ$^{Pro(R)}$	7.38	0.41	0.45
Pro-Hα	Pro-Hβ$^{Pro(S)}$	9.58	0.10	0.53
Pro-Hβ$^{Pro(R)}$	Pro-Hγ$^{Pro(R)}$	6.43	0.22	0.12
Pro-Hβ$^{Pro(S)}$	Pro-Hγ$^{Pro(R)}$	11.40	0.41	0.18
Pro-Hγ$^{Pro(R)}$	Pro-Hδ$^{Pro(R)}$	7.50	0.55	0.70
Pro-Hγ$^{Pro(R)}$	Pro-Hδ$^{Pro(S)}$	10.24	0.05	0.50
	SSD		9.1	14.1

Finally, a set of calculations was done to explore the sensitivity of NAMFIS populations to variations of the experimental datasets. Altogether 6 additional analyses were performed on the conformer pools of 2259 (APGA), 1996 (A(4*R*)MePGA), and 2485 (A(4*S*)MePGA) structures. The results are given in Tables S9, S10, and S11 for APGA, A(4*R*)MePGA, and A(4*S*)MePGA, respectively. The following analyses were performed:

1. **Only 17 common NOE's + all J's**: Only those 11 interresidue and 6 intraresidue NOE's that are common to all three peptides were retained (see Table S3). All coupling constants were retained.
2. **No J's**: All NOE's were retained, but all coupling constants were removed.
3. **No backbone J's**: All NOE's were retained and only backbone coupling constants were removed.
4. **No NOE's**: All NOE's were removed, but all coupling constants were retained.
5. **No intrares. NOE's**: Only intraresidue NOE's were removed. Interresidue NOE's and all coupling constants were retained.
6. **No interres**. **NOE's**: Only interresidue NOE's were removed. Intraresidue NOE's and all coupling constants were retained.

The results show that the family characteristics and populations are well preserved when at least most of the interresidue NOE's are retained. A significant redistribution of the familial populations relative to the original data set was observed when all NOE's or all interresidue NOE's were omitted. In these cases, additional turn structures appeared that are not necessarily realistic (inverse γ-turn (APGA and A(4*R*)MePGA), type VIII β-turn (APGA)).

Table S9. Family populations from additional NAMFIS analyses for APGA.

motif	original NAMFIS analysis	only 17 common NOE's + all J's	no J's	no backbone J's	no NOE's	no intrares. NOE's	no interres. NOE's
ideal βII	37.3	34.5	28.2	23.5	4.4	42.2	
distorted βII	9.3	18.8	5.8	15.4		22.7	15.0
combined βII	**46.6**	**53.3**	**34.0**	**38.9**	**4.4**	**64.9**	**15.0**
ideal βI	13.9	4.1	12.7	10.9	17.2	3.5	45.9
distorted βI	3.9	10.0	4.4	5.2	14.7	7.3	
combined βI	**17.8**	**14.1**	**17.1**	**16.1**	**31.9**	**10.8**	**45.9**
bend II	28.8	19.1	42.8	34.3		7.7	
bend I						4.2	
combined bends	**28.8**	**19.1**	**42.8**	**34.3**	**0.0**	**11.9**	**0.0**
extended	**6.7**	**13.5**	**6.2**	**10.6**	**0.0**	**12.5**	**0.0**
inverse γ-turn					60.0		39.2
distorted βVIII					3.7		
endo/*exo*	6/7	7/6	6/7	6/7	6/5	7/5	5/5

Table S10. Family populations from additional NAMFIS analyses for A(4*R*)MePGA.

motif	original NAMFIS analysis	only 17 common NOE's + all J's	no J's	no backbone J's	no NOE's	no intrares. NOE's	no interres. NOE's
ideal βII	36.6	40.7	31.2	32.6		40.7	32.6
distorted βII	10.3	1.2	10.2	9.6	21.2	19.4	23.9
combined βII	**46.9**	**41.9**	**41.4**	**42.2**	**21.2**	**60.1**	**56.5**
ideal βI		0.5			48.1	7.6	35.2
distorted βI	15.9	9.2	8.0	7.9	14.5	9.8	
combined βI	**15.9**	**9.7**	**8.0**	**7.9**	**62.6**	**17.4**	**35.2**
bend II	15.1	19.3	30.3	29.7		12.0	
bend I	4.7		15.3	15.1			4.8
combined bends	**19.8**	**19.3**	**45.6**	**44.8**	**0.0**	**12.0**	**4.8**
extended	**17.5**	**29.1**	**4.9**	**5.0**	**0.0**	**10.4**	**0.0**
inverse γ-turn					16.3		3.4
endo/*exo*	13/3	10/2	10/3	10/3	6/3	11/3	6/3

Table S11. Family populations from additional NAMFIS analyses for A(4*S*)MePGA.

motif	original NAMFIS analysis	only 17 common NOE's + all J's	no J's	no backbone J's	no NOE's	no intrares. NOE's	no interres. NOE's
ideal βII	41.3	35.6	40.3	36.4		46.9	
distorted βII	21.0	29.9	24.6	28.3	31.5	21.7	40.7
combined βII	**62.3**	**65.5**	**64.9**	**64.7**	**31.5**	**68.6**	**40.7**
ideal βI		4.0			25.5		44.2
distorted βI	8.3	8.4	7.8	8.3		5.7	15.1
combined βI	**8.3**	**12.4**	**7.8**	**8.3**	**25.5**	**5.7**	**59.3**
bend II	13.2	14.9	9.9	11.7		7.2	
bend I	5.1		5.3	5.1	10.7	6.7	
combined bends	**18.3**	**14.9**	**15.2**	**16.8**	**10.7**	**13.9**	**0.0**
extended	**11.2**	**7.2**	**12.2**	**10.2**	**32.4**	**11.9**	**0.0**
endo/*exo*	2/11	3/12	3/12	1/14	1/6	1/13	1/7