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

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

Received 19th October 2009, Accepted 19th February 2010
First published as an Advance Article on the web 11th March 2010
DOI: 10.1039/b921794k

β-Turns play an important role in peptide and protein chemistry, biophysics, and bioinformatics. The aim of this research was to study short linear peptides that have a high propensity to form β-turn structures in solution. In particular, we examined conformational ensembles of β-turn forming peptides with a general sequence CBz-L-Ala-L-Xaa-Gly-L-Ala-O*t*Bu. These tetrapeptides, APGA, A(4*R*)MePGA, and A(4*S*)MePGA, incorporate proline, (4*R*)-methylproline, and (4*S*)-methylproline, respectively, at the Xaa position. To determine the influence of the 4-methyl substituted prolines on the β-turn populations, the NAMFIS (NMR analysis of molecular flexibility in solution) deconvolution analysis for these three peptides were performed in DMSO-d_6 solution. The NBO (natural bond orbital) method was employed to gain further insight into the results obtained from the NAMFIS analysis. The emphasis in the NBO analysis was to characterize remote intramolecular interactions that could influence the backbone-backbone interactions contributing to β-turn stability. NAMFIS results indicate that the enantiospecific incorporation of the methyl substituent at the C^{γ} (C4) position of the proline residue can be used to selectively control the pyrrolidine ring puckering propensities and, consequently, the preferred ϕ,ψ angles associated with the proline residue in β-turn forming peptides. The NAMFIS analyses show that the presence of (4*S*)-methylproline in A(4*S*)MePGA considerably increased the type II β-turn population with respect to APGA and A(4*R*)MePGA. The NBO calculations suggest that this observation can be rationalized based on an $n \rightarrow \pi^*$ interaction between the N-terminus alanine carbonyl oxygen and the proline carbonyl group. Several other interactions between remote orbitals in these peptides provide a more detailed explanation for the observed population distributions.

Introduction

Reverse turns are elements of secondary structure that have an important role both in peptides and proteins.[1] They are broadly defined as those regions of the polypeptide where the amino acid chain turns back onto itself. The most common type of reverse turn, the β-turn, represents about 25% of all residues in proteins.[2] β-Turns play a critical role in the topology of polypeptides by orienting structural elements such as α-helices and β-strands. It has also been proposed that they might be important in the initiation of the protein folding process.[3] Furthermore, β-turns play many biological roles in peptides and proteins.[1] They are predominantly constituted of hydrophilic residues and thus frequently occur on the surfaces of globular proteins.[4] This and the higher than random proportion of structural or functional side chains such as proline, aspartic acid, serine, threonine, and lysine makes them viable candidates for structural motifs for molecular recognition.[1] The β-turn structure is able to orient the side chains of the central residues i + 1 and i + 2 in such a manner that they can be recognized electronically and sterically by the complementary receptor molecule. Therefore, β-turns are widely implicated in hormone-receptor interactions as well as recognition sites for immunological, metabolic, genomic and endocrinological mechanisms.[1] β-Turns are also involved in the biological activity of peptides, often postulated as the bioactive conformation that interacts with another molecule. The above-mentioned properties have led to great interest in mimicking β-turns for the synthesis of medicines in the field of medicinal and pharmacological chemistry.[5]

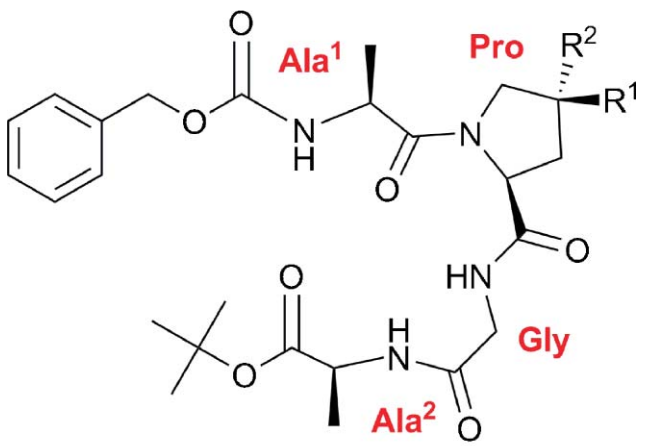
An understanding of the conformational behavior of the β-turn forming peptides in solution has the potential to provide valuable insights into the conformational requirements for biological activity. In the present work, we have analyzed the conformational profiles of short linear β-turn forming peptides in solution by using nuclear magnetic resonance (NMR) spectroscopy. Generally, in this method, interproton distances are derived from NOESY (nuclear Overhauser spectroscopy) or ROESY (rotating frame Overhauser spectroscopy) experiments and torsional angles are calculated from *J*-coupling constants using empirical Karplus-type equations.[6] These parameters are used as geometrical restraints to define the conformation(s) of the molecule by the means of various protocols including distance geometry and simulated annealing. These methods assume the existence of a single, unique conformation. However, most short peptides, even cyclic ones, are present in solution as conformational ensembles containing a large number of widely different conformers that are in rapid interconversion on the NMR time scale and are characterized by

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

† Electronic supplementary information (ESI) available: Aggregation studies. ^{1}H and ^{13}C NMR chemical shift assignments, ^{1}H,^{1}H coupling constants, and off-resonance ROESY derived interproton distances for APGA, A(4*R*)MePGA, and A(4*S*)MePGA. Classification of β-turns. Detection of β-turns by ^{1}H NMR. Synthetic procedures. ^{1}H, ^{13}C, and off-resonance ROESY spectra. DFT energies and Cartesian coordinates for (4*R*)MePG-I, (4*R*)MePG-II, (4*S*)MePG-I, and (4*S*)MePG-II. Additional NAMFIS analyses. See DOI: 10.1039/b921794k

an averaged NMR spectrum.[7–9] Consequently, NOE/ROE and *J*-coupling constants correspond to population weighted averaged values which, when used as restraints in the above-mentioned calculations, will most probably lead to a virtual conformation that combines multiconformer features in a single geometry. Therefore, for flexible molecules, different types of procedures are needed to identify and quantitate the ensemble of different conformations in solution.

In this research, we have employed the NAMFIS (NMR analysis of molecular flexibility in solution)[10–12] method to analyze the conformational profiles of linear tetrapeptides in solution and to determine the estimated populations of the β-turn containing conformers among the structures in the conformational pool. NAMFIS is an approach that deconvolutes the averaged NMR data into a small family of conformations that optimally represents the torsions derived from *J*-coupling constants and the distances derived from NOEs/ROEs. In the NAMFIS analysis, explicit force field energies are ignored during the deconvolution. Therefore, artifacts emerging from energy misevaluation introduced by long-range interactions among polar functional groups are bypassed. NAMFIS has been used, for example, to analyze the solution structures of short peptides,[11,13–15] a tricyclic ketone,[10] Taxol,[16] discodermolide,[17] several epothilones,[18,19] laulimalide,[12] and geldanamycin and radiacol.[20] The research described in this paper involves three tetrapeptides, with the general amino acid sequence CBz-L-Ala-L-Xaa-Gly-L-Ala-O*t*Bu, in which Xaa is proline, or the nonproteinogenic amino acid (4*R*)- or (4*S*)-methylproline. These peptides are illustrated in Scheme 1, along with the shorthand designations used in the text.



APGA : R^1, R^2 = H
A(4*R*)MePGA : R^1 = H, R^2 = CH_3
A(4*S*)MePGA : R^1 = CH_3, R^2 = H

Scheme 1 The structures of the peptides under study, along with the shorthand designations used in the text.

A substituent at the C^γ (C4) position of the pyrrolidine ring of proline can provide a large degree of conformational control over the side chain. The pyrrolidine ring has two predominant puckering modes: C^γ *exo* (up) and *endo* (down) envelope conformers.[21] In the case of the unsubstituted *trans* proline residue in peptides and proteins, these two modes are almost equally probable.[22] The puckering propensity can be modified by introduction of a substituent at the C^γ position of the pyrrolidine ring. It has been shown that the (4*R*) epimers of hydroxy-, fluoro- and chloroproline favor the *exo* puckering, while the (4*S*) epimers favor the *endo* puckering.[23–25] The relative stability of the *exo* puckering increases with the electronegativity of the (4*R*) substituent, whereas the *endo* puckering is favored by an electronegative (4*S*) substituent. In the preferred puckering, these substituents adopt a pseudoaxial orientation. However, in the case of (4*R*)- and (4*S*)-mercapto-L-proline and related derivatives, the puckering propensities are reversed: the (4*R*) epimers induce the *endo* puckering, while the (4*R*) epimers generate the *exo* puckering, orienting the substituents pseudoequatorially.[26] In addition to conformational control over the proline side chain, substitutions on C^γ are known to influence the equilibrium conformational populations of the *cis*/*trans* prolyl bonds.[27-29] Extensive computational and NMR studies have suggested that the conformational effects of the electronegative substituent is dictated by inductive and stereoelectronic factors.[23–30]

Results and discussion

In this study, we have used the non-electronegative methyl substituent at the C^γ position to control the puckering propensity of the pyrrolidine ring and, consequently, the preferred ϕ,ψ angles associated with the proline residue in β-turn forming peptides. DFT calculations have previously indicated that the pyrrolidine ring of (4*R*)-methyl-L-proline has a strong preference for the C^γ *endo* puckering, while that of (4*S*)-methyl-L-proline has a strong preference for the C^γ *exo* puckering.[31] These ring puckering preferences have also been observed in crystal structures of Ac-(4*R*)-MePro-NHMe and Ac-(4*S*)-MePro-NHMe.[32] In addition, the predominant ring puckering observed in (4*R*)- and (4*S*)-*tert*-butylprolines follow the same trend.[33] In the preferred conformations, the methyl or *tert*-butyl group adopts a pseudoequatorial orientation. The ring puckering preferences of these substituents are strongly determined by steric effects.[31]

The puckering preference of the pyrrolidine ring could elicit changes in the polypeptide chain ϕ, ψ and ω angles, leading to a correlation of proline ring puckering with polypeptide chain conformation: the C^γ *exo* puckering is more common in compact structures, while the C^γ *endo* puckering is favored in extended structures.[22–24,28,34–37] These observations are related to the fact that the C^γ *endo* conformations are accompanied by more negative ϕ angles and larger ψ angles, whereas the opposite is true for the C^γ *exo* conformations.[23,24] Therefore, selective control of the pyrrolidine ring puckering provides the possibility to influence the polypeptide chain conformation. 4-Substituted proline derivatives have been incorporated into collagen[25,26,31,38–42] and elastin-mimetic[30,43] peptides, and into globular proteins,[37,44,45] to selectively control the backbone conformations and stability of these structures by defining the ring puckering of critical proline residues using stereoelectronic effects.

The simple model peptides used in this study are equally accessible to both experimental and theoretical investigation. They can be used to analyze and quantify the different interactions contributing to β-turn stability and, therefore, β-turn formation in peptides and proteins. The peptides examined consist of alanine, glycine, and proline; thus side-chain interactions are minimal,

and the primary interactions contributing to β-turn stability should result from the peptide backbone. In addition to the above, these peptides provide a means to study the relationship between the proline ring puckering and the peptide backbone conformation. Our working hypothesis was that incorporation of the 4-methylproline epimers with altered puckering propensities into these linear tetrapeptides influences the bias in the populations of the β-turn conformations in solution, thus allowing a means to control the solution conformation of even a small peptide in solution. In this investigation, we employed the NAMFIS[10–12] method to analyze the conformational profiles of proline and 4-methylproline containing tetrapeptides in DMSO-d_6. By using the NAMFIS, we determined the ring puckering preferences of the proline and 4-methylproline residues incorporated into these peptides, and analyzed the influence of the 4-methylproline epimers on the estimated populations of the β-turn structures in solution. In addition, we used the natural bond orbital (NBO)[46] method to gain further insight into the results obtained from the NAMFIS analysis. The emphasis in the NBO analysis was to characterize remote intramolecular interactions that could influence the peptide backbone conformation, thus providing information on the underlying principles responsible for the conformational preferences of β-turns.

^{1}H NMR spectroscopy

Complete ^{1}H chemical shift assignments for APGA, A(4*R*)MePGA, and A(4*S*)MePGA were made using a combination of standard TOCSY and T-ROESY experiments. Interproton distances were obtained from analysis of off-resonance ROESY experiments.[47] The ^{1}H NMR parameters were analyzed using the PERCH NMR software.[48,49] Because the peptides were protected as a benzyl carbamate and a *tert*-butyl ester, they were insoluble in water. Therefore, the NMR solvent in this study was chosen to be DMSO-d_6, because it has similar polarity and dielectric constant as water. Aggregation studies over a concentration range of 1.6 to 100 mM revealed that the chemical shifts were essentially insensitive to concentration (see Supporting Information for details). This observation indicates that the intermolecular associations were negligible over this concentration range. Therefore, the samples were prepared at relatively high concentrations (*ca.* 80 mM). Accordingly, the signal-to-noise level of the spectra was high, enabling accurate determination of NMR parameters needed for conformational analysis. One might question the relevance of conformations identified in DMSO in the biological context. However, based on the observation that the liquid phases in the cytoplasm, intersynaptic space, and receptor-binding sites are characterized by a higher viscosity than water, it has been suggested that DMSO, or another high viscosity solvent, might be more realistic than water in modeling the above-mentioned environments.[50] In addition, it has been reported that DMSO does not necessarily alter either biological activity or conformational preferences of peptides and, therefore, is a pragmatic choice.[51–54] Furthermore, a preliminary study by O'Connell *et al.*[55] on unprotected APGA peptide in aqueous solution showed that this peptide behaves in the same way both in aqueous solution and in DMSO.

The 1D ^{1}H NMR spectrum of APGA, A(4*R*)MePGA, and A(4*S*)MePGA displayed two sets of signals in the amide proton region, which gave rise to negative exchange cross-peaks (same sign as the diagonal) in the T-ROESY spectra and corresponded to the slow *cis*/*trans* isomerization of the Ala^1-Pro/4MePro peptide bond. However, the integration of these signals in the 1D ^{1}H spectrum indicated that the population of the minor isomer was less than 20% of the major one. Therefore, only the major isomers were analyzed further. Based on the T-ROESY spectra, the proline peptide bond conformation of the major species corresponded to the *trans* isomer, as detected through the strong correlation between Ala^1-Hα and Pro-H$\delta^{pro(R)/(S)}$ resonances for all three peptides. ^{1}H NMR chemical shifts and coupling constants for APGA, A(4*R*)MePGA, and A(4*S*)MePGA are listed in Tables S1 and S2 (Supporting Information). Please see Supporting Information for the classification of β-turns and for the detailed description of the detection of β-turns by ^{1}H NMR.

Solution structures by NAMFIS analysis

In order to take conformational averaging into account, the NAMFIS (NMR analysis of molecular flexibility in solution)[10–12] procedure was utilized to evaluate the population distribution of the conformational ensemble based on the distances derived from NOEs/ROEs and torsion angles derived from the *J*-coupling constants. The NAMFIS method deconvolutes an averaged NMR spectrum into a small family of conformers, identifies individual conformational minima, and assigns weights to the conformers that optimally represent the experimentally obtained NOEs/ROEs and *J*-couplings. Consequently, the burden of accommodating the complete set of averaged NMR data is shared among about 5 to 20 conformers. Interproton distances for the NAMFIS analysis were obtained from the adiabatic version of off-resonance ROESY experiments.[47] The backbone ^3J(NH,Hα) and the pyrrolidine ring ^3J(H,H) coupling constants were extracted from the 1D ^{1}H NMR spectra using the PERCH NMR software.[48,49] The latter coupling constants were included into the NAMFIS deconvolution in order to analyze the puckering propensity of the pyrrolidine ring in solution. The ^3J(H,H) coupling constants and the ROE-derived distances used in the NAMFIS analysis are provided in Supporting Information in Tables S2 and S3, respectively. Extensive Monte Carlo conformational searches for all three peptides were carried out using OPLS_2005, AMBER*, and MMFFs force fields in MacroModel 9.5[56] combined with the GB/SA H_2O continuum solvent model.[57] The fully optimized structures were relieved from duplicate conformations and loaded into a common structural database, yielding 2259, 1996, and 2485 conformers for APGA, A(4*R*)MePGA, and A(4*S*)MePGA, respectively. These conformers were then subjected to the NAMFIS deconvolution procedure using interproton distances and torsion angles obtained from the averaged NMR data. This analysis, which couples calculated structures with structures determined by NMR, resulted in an ensemble of conformations with varying populations for the peptides under study. NAMFIS populations and geometric parameters for APGA, A(4*R*)MePGA, and A(4*S*)MePGA are provided in Tables 1–3, respectively.

The obtained NAMFIS conformations were grouped into 7 structural motifs (Table 4): ideal and distorted type I β-turn, ideal and distorted type II β-turn, type I bend, type II bend, and extended structures. In general, a structure was classified as a β-turn if the distance between C^α(i)–C^α(i + 3) was <7 Å. More

Table 1 NAMFIS populations and geometric parameters for APGA

conformer	pop (%)	ϕ(i + 1) (deg)	ψ(i + 1) (deg)	ϕ(i + 2) (deg)	ψ(i + 2) (deg)	*d*(Ala^1-C$^\alpha$–Ala^2-C$^\alpha$) (Å)	Pro conf.	H-bond Ala^1-CO–Ala^2-NH
1	28.7	−73.5	129.6	171.8	−141.6	9.0	*endo*	−
2	12.3	−68.0	−26.5	−94.1	5.1	5.2	*exo*	+
3	11.5	−74.2	108.5	87.9	15.6	4.9	*endo*	+
4	10.7	−73.7	130.9	88.8	3.4	5.8	*endo*	+
5	9.3	−47.2	109.0	143.9	−43.8	5.8	*exo*	+
6	7.4	−65.5	121.4	93.8	−0.6	5.4	*exo*	+
7	5.5	−77.7	−176.0	−103.0	5.2	8.6	*endo*	−
8	4.8	−67.3	123.4	88.0	10.9	5.3	*exo*	+
9	3.9	−69.1	−25.6	−88.9	−86.4	5.4	*exo*	+
10	2.2	−72.9	110.3	86.8	22.1	4.9	*endo*	+
11	1.5	−63.7	−16.2	−122.9	24.0	4.8	*exo*	+
12	1.2	−76.6	161.0	84.6	21.9	7.0	*endo*	−
13	0.6	−66.4	120.7	88.7	8.3	5.2	*exo*	+

Table 2 NAMFIS populations and geometric parameters for A(4*R*)MePGA

conformer	pop (%)	ϕ(i + 1) (deg)	ψ(i + 1) (deg)	ϕ(i + 2) (deg)	ψ(i + 2) (deg)	*d*(Ala^1-C$^\alpha$–Ala^2-C$^\alpha$) (Å)	Pro conf.	H-bond Ala^1-CO–Ala^2-NH
1	16.4	−72.0	122.7	95.7	4.9	5.2	*endo*	+
2	15.3	−73.9	131.8	88.2	3.8	5.8	*endo*	+
3	11.3	−76.6	−28.1	−123.5	−144.4	6.6	*endo*	−
4	9.7	−76.7	167.2	−88.1	157.4	9.8	*endo*	−
5	8.6	−59.7	125.0	165.8	−140.1	8.4	*exo*	−
6	7.0	−50.8	109.8	144.9	−43.5	5.8	*exo*	+
7	6.3	−68.2	153.7	169.2	−127.0	9.4	*endo*	−
8	5.5	−81.0	99.1	173.2	−147.7	8.1	*endo*	−
9	4.7	−90.4	−21.4	179.0	163.5	7.7	*endo*	−
10	3.9	−72.1	−2.4	−116.0	−60.7	5.1	*endo*	+
11	3.4	−70.2	127.4	96.2	1.1	5.4	*endo*	+
12	3.3	−49.3	106.4	144.3	−48.7	5.6	*exo*	+
13	1.5	−70.6	122.6	100.3	−1.5	5.5	*endo*	+
14	1.5	−79.2	−173.2	−102.9	5.6	8.5	*endo*	−
15	1.0	−57.8	133.5	162.1	−120.2	8.2	*endo*	−
16	0.7	−73.2	−15.9	−87.0	−89.6	5.4	*endo*	+

Table 3 NAMFIS populations and geometric parameters for A(4*S*)MePGA

conformer	pop (%)	ϕ(i + 1) (deg)	ψ(i + 1) (deg)	ϕ(i + 2) (deg)	ψ(i + 2) (deg)	*d*(Ala^1-C$^\alpha$–Ala^2-C$^\alpha$) (Å)	Pro conf.	H-bond Ala^1-CO–Ala^2-NH
1	21.0	−48.9	109.2	143.8	−42.7	5.8	*exo*	+
2	15.7	−65.8	121.9	89.3	7.4	5.2	*exo*	+
3	14.2	−67.1	123.6	87.9	10.9	5.3	*exo*	+
4	13.2	−57.9	130.9	171.9	−143.6	9.0	*exo*	−
5	5.9	−69.8	167.8	−91.1	−124.8	10.2	*exo*	−
6	5.1	−71.4	−32.5	−93.4	−162.2	7.2	*exo*	−
7	4.8	−66.4	−14.9	−129.9	41.6	4.7	*exo*	+
8	4.7	−66.3	123.0	89.4	5.1	5.4	*exo*	+
9	4.2	−74.2	−176.9	−97.2	52.7	8.2	*endo*	−
10	3.5	−70.6	−34.6	−171.2	29.5	5.8	*exo*	−
11	3.4	−44.3	116.8	117.4	−18.7	5.6	*exo*	+
12	3.3	−63.3	132.6	103.9	−23.6	5.7	*exo*	+
13	1.1	−75.3	173.5	−112.5	12.2	9.1	*endo*	−

Table 4 Total populations of the NAMFIS motifs for APGA, A(4*R*)MePGA, and A(4*S*)MePGA[a]

	APGA		A(4*R*)MePGA		A(4*S*)MePGA	
motif	pop, %	members	pop, %	members	pop, %	members
ideal βII	37.3	3, 4, 6, 8, 10, 13	36.6	1, 2, 11, 13	41.3	2, 3, 8, 11, 12
distorted βII	9.3	5	10.3	6, 12	21.0	1
ideal βI	13.9	2, 11				
distorted βI	3.9	9	15.9	3, 10, 16	8.3	7, 10
bend II	28.8	1	15.1	5, 8, 15	13.2	4
bend I			4.7	9	5.1	6
extended	6.7	7, 12	17.5	4, 7, 14	11.2	5, 9, 13

[a] Populations are the normalized sum of the individual conformer populations.

precisely, a structure was considered as an ideal type I or type II β-turn if three of the ϕ,ψ angles of the central residues i + 1 and i + 2 deviated less than ±30° from the idealized values for type I or II β-turn, and the fourth angle deviated less than ±45°. Idealized ϕ,ψ values, as determined by Lewis *et al.*,[58] for type I and type II β-turns are as follows: ϕ(i + 1) = −60°, ψ(i + 1) = −30°, ϕ(i + 2) = −90°, ψ(i + 2) = 0° for type I β-turn, and ϕ(i + 1) = -60°, ψ(i + 1) = 120°, ϕ(i + 2) = 80°, ψ(i + 2) = 0° for type II β-turn. In a distorted β-turn structure, ϕ(i + 2) and/or ψ(i + 2) angles were allowed to deviate more than ±45° from the idealized values. In addition, a structure was considered as distorted if both of these angles deviated more than ±30°, but less than ±45° from the idealized values. A structure was classified as type I bend, type II bend or extended, if the distance between C^{α}(i)–C^{α}(i + 3) was ≥7 Å. In a type I or type II bend, the ϕ(i + 1) and ψ(i + 1) angles were allowed to deviate less than ±30° from the idealized values for type I or II β-turn. In these structures, the ϕ(i + 2) and ψ(i + 2) angles were allowed to take values greater than ±45° from the idealized β-turn values. In extended structures, all ϕ and ψ angles were allowed to adopt values greater than ±45° from the idealized values. Hydrogen bonds in the NAMFIS structures were defined by relations between four atoms: the donor hydrogen atom (H), the acceptor atom (A), the donor atom (D) bonded to H, and another neighbor atom (B) bonded to A. The following values were used determine a valid hydrogen bond: H···A distance < 2.50 Å, D–H···A angle > 120°, and H···A–B angle > 90°. The structures of the most populated members of each motif for APGA, A(4*R*)MePGA, and A(4*S*)MePGA are given in Fig. 1–3, respectively.

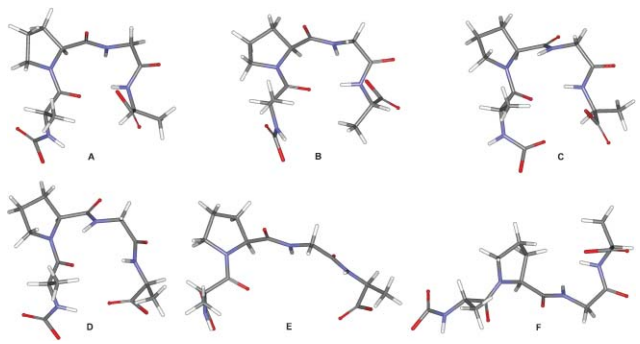


Fig. 1 The structures of the most populated members of each motif for APGA. (A) Ideal type II β-turn, (B) distorted type II β-turn, (C) ideal type I β-turn, (D) distorted type I β-turn, (E) type II bend, and (F) extended motif. Protecting groups have been omitted for clarity.

In addition to the backbone conformation, the proline ring puckering (C^{γ} *endo*/*exo*) of the NAMFIS structures was determined, and the influence of the proline ring conformation on the estimated populations of the β-turn structures among the other structures in solution was analyzed. As discussed above, the pyrrolidine ring of the (4*R*)- and (4*S*)-methylproline residues have alternative ring puckering preferences: the (4*R*) epimer of methylproline favors *endo* puckering, while the (4*S*) epimer favors *exo* puckering. The puckering preference of the pyrrolidine ring, in turn, has an effect on the polypeptide chain ϕ,ψ angles, leading to a correlation of proline ring puckering with polypeptide chain conformation. As discussed below, incorporation of 4-

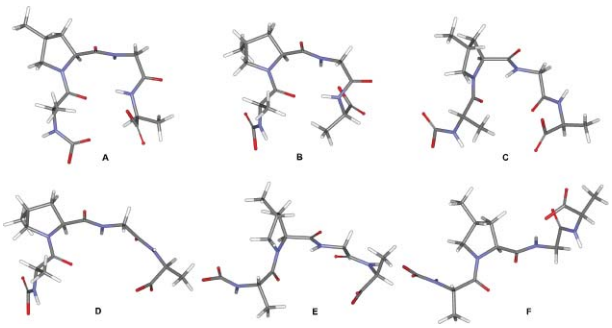


Fig. 2 The structures of the most populated members of each motif for A(4*R*)MePGA. (A) Ideal type II β-turn, (B) distorted type II β-turn, (C) distorted type I β-turn, (D) type II bend, (E) type I bend, and (F) extended motif. Protecting groups have been omitted for clarity.

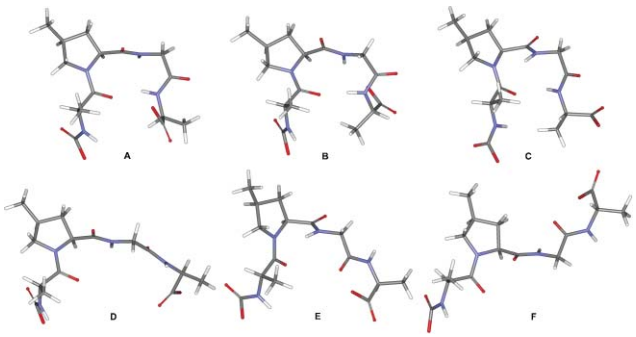


Fig. 3 The structures of the most populated members of each motif for A(4*S*)MePGA. (A) Ideal type II β-turn, (B) distorted type II β-turn, (C) distorted type I β-turn, (D) type II bend, (E) type I bend, and (F) extended motif. Protecting groups have been omitted for clarity.

methylproline epimers with altered puckering propensities into linear tetrapeptides introduces a bias in the populations of the β-turn conformations in solution.

NAMFIS analysis of APGA

The NAMFIS analysis of APGA in DMSO-d_6, using 19 ROE distances, 4 ^3J(NH,Hα) backbone couplings, and 10 ^3J(H,H) pyrrolidine ring couplings, resulted in 13 conformers above a molar fraction of 0.01 with estimated populations ranging from 1 to 29% (Table 1). The most populated conformer (29%) represents the type II bend motif, while the second (12%) and third most populated (12%) conformers represent the ideal type I β-turn and the ideal type II β-turn motifs, respectively. The most plentiful motif among the NAMFIS conformers is the ideal type II β-turn (37%), although its top contributor is only the third most populated structure. This motif contains 6 out of 13 NAMFIS conformers. The second most populated motif is the type II bend (29%) containing only the most populated conformer. The remaining 6 conformers are divided between 4 motifs: ideal type I β-turn (14%), distorted type II β-turn (9%), extended (7%), and distorted type I β-turn (4%). None of the NAMFIS conformers belong to the type I bend motif. The estimated combined population of the ideal type II β-turn and the distorted type II β-turn motifs is 46%, and that of the ideal type I β-turn and the

distorted type I β-turn motifs is 18%. Thus, the combined β-turn population of APGA in DMSO-d_6 solution is 64%. As expected, the type II β-turn conformations are much more populated than the type I β-turn conformations. The bend structures represent 29% and the extended structures 7% of the total population of the NAMFIS conformers. Hydrogen bonding between Ala^1-CO and Ala^2-NH is present in all NAMFIS structures of APGA belonging to the ideal or distorted type I/II β-turn motif. The conformation of the pyrrolidine ring of the unsubstituted proline residue in APGA is expected to be a time average of a series of different puckered conformers. Among these conformers, the energetically equivalent C^γ *endo* and C^γ *exo* puckers are preferred. These two puckerings are essentially equivalently populated among the NAMFIS conformers of APGA: in 7 out of 13 structures, the pyrrolidine ring is in the C^γ *exo* puckering, while in the remaining 6 structures it is in the C^γ *endo* puckering. The C^γ *exo* puckering is more common in β-turn structures, evidenced by the observation that the pyrrolidine ring is in the *exo* conformation in 7 out of 10 β-turn conformers. On the contrary, in the bend and extended structures, the C^γ *endo* puckering prevails. This is consistent with the observation that the C^γ *exo* puckering is more common in compact structures, while the C^γ *endo* puckering is favored in extended structures.

NAMFIS analysis of A(4*R*)MePGA

A deconvolution of the average structure of A(4*R*)MePGA in DMSO-d_6 with the NAMFIS procedure, using 34 ROE distances, 4 ^3J(NH,Hα) backbone couplings, and 6 ^3J(H,H) pyrrolidine ring couplings, delivered 16 conformations above a molar fraction of 0.01 that accommodate the NMR data with estimated populations ranging from 1 to 16% (Table 2). The most (16%) and the second most populated (15%) conformers represent the ideal type II β-turn motif, while the third most populated conformer (11%) represents the distorted type I motif. The most populated motif is the ideal type II β-turn (37%), although it contains only 4 out of 16 NAMFIS conformers. However, as mentioned above, the two most populated conformers belong to this motif. The second most populated motif is the extended motif (18%), containing 3 conformers. The rest of the 9 conformers are divided between 4 motifs: the distorted type I β-turn (16%), type II bend (15%), distorted type II β-turn (10%), and type I bend (5%). The ideal type I motif contains no NAMFIS conformers. The combined population of the ideal type II β-turn and the distorted type II β-turn motifs is 47%, while the combined population of the ideal type I β-turn and the distorted type I β-turn motifs is 16%. Consequently, the combined β-turn population of A(4*R*)MePGA in DMSO-d_6 solution is approximately 63%. As in the case of APGA, the type II β-turn conformations are strongly favored over the type I β-turn conformations. The bend structures represent 20% and the extended structures 18% of the total population of the NAMFIS conformers. The conformation of the pyrrolidine ring in A(4*R*)MePGA strongly favors C^γ *endo* puckering: in 13 out of 16 NAMFIS structures, the pyrrolidine ring is in the C^γ *endo* puckering mode. Consequently, the (4*R*) methyl substituent adopts a pseudoequatorial orientation in these structures. All except one NAMFIS structure of A(4*R*)MePGA belonging to the ideal or distorted type I/II β-turn motif exhibit a hydrogen bond between Ala^1-CO and Ala^2-NH.

NAMFIS analysis of A(4*S*)MePGA

The NAMFIS deconvolution of A(4*S*)MePGA in DMSO-d_6, using 25 ROE distances, 4 ^3J(NH,Hα) backbone couplings, and 6 ^3J(H,H) pyrrolidine ring couplings, yielded 13 conformations above a molar fraction of 0.01 with estimated populations ranging from 1 to 21% (Table 3). The most populated conformer (21%) represents the distorted type II β-turn motif, while the second (16%) and third most populated (14%) conformers represent the ideal type II β-turn motif. Similarly with APGA and A(4*R*)MePGA, the most populated motif is the ideal type II β-turn (41%), containing 5 out of 13 NAMFIS conformers. The second most populated motif is the distorted type II β-turn (21%), containing only one conformer. The remaining 7 conformers are divided between 4 motifs: type II bend (13%), extended (11%), distorted type I β-turn (8%), and type I bend (5%). None of the NAMFIS conformers represent the ideal type I β-turn motif. The estimated combined population of the ideal type II β-turn and the distorted type II β-turn motifs is 62%, and that of the ideal type I β-turn and the distorted type I β-turn motifs is 8%. Therefore, the combined β-turn population of A(4*S*)MePGA in DMSO-d_6 solution is 70%. As expected, the type II β-turn conformations are very strongly favored over the type I β-turn conformations. The bend structures represent 18% and the extended structures 11% of the total population of the NAMFIS conformers. The pyrrolidine ring of the proline residue in A(4*S*)MePGA strongly favors C^γ *exo* puckering: in 11 out of 13 NAMFIS-derived structures the pyrrolidine ring exhibits the C^γ *exo* conformation. In the preferred puckering mode, the (4*S*) methyl substituent adopts a pseudoequatorial orientation. Hydrogen bonding between Ala^1-CO and Ala^2-NH is present in all except one NAMFIS conformer of A(4*S*)MePGA belonging to the ideal or distorted type I/II β-turn motif.

Effect of $n \rightarrow \pi^*$ interaction on proline ψ angles

The most populated structural motifs for all peptides under study are those in which the ψ(i + 1) dihedral angle is close to values compatible with the type II β-turn conformation, *i.e.* in which the ψ(i + 1) deviates less than ±30° from the idealized value for type II β-turn (120°). We propose that this observation can be rationalized based on an energetic stabilization associated with a non-bonded $n \rightarrow \pi^*$ interaction between the lone pair of the Ala^1 carbonyl oxygen and the antibonding orbital of the proline carbonyl group. The C^γ *exo* ring puckering preorganizes the main chain ϕ,ψ angles of the proline residue for this favorable O(i)$\cdots$C′=O(i + 1) interaction, which is strongest when O(i) is positioned proximally and along the Bürgi-Dunitz trajectory to C′=O(i + 1).[59-62] This type of $n \rightarrow \pi^*$ interaction has been proposed to have an important role in stabilizing *trans* peptide bond.[23] In addition, the $n \rightarrow \pi^*$ interaction has been suggested to stabilize polypeptide structures such as α-helix and polyproline type II (PPII) helix.[63]

A significant $n \rightarrow \pi^*$ interaction can be defined as one in which the O(i)$\cdots$C′(i + 1) distance is $\delta_{BD} < 3.00$ Å and the O(i)$\cdots$C′=O(i + 1) angle is $\tau_{BD} = 107 \pm 10°$. These parameters are listed for APGA, A(4*R*)MePGA, and A(4*S*)MePGA in Table 5. The NAMFIS structures of A(4*S*)MePGA belonging to the ideal and distorted type II β-turn motif and those belonging to the type II bend motif have values of $\delta_{BD} \leq 3.00$ Å and τ_{BD} within the range

Table 5 The O(i)···C′(i + 1) distance (δ_{BD}) and the O(i)···C′=O(i + 1) angle (τ_{BD}) for APGA, A(4*R*)MePGA, and A(4*S*)MePGA

APGA			A(4*R*)MePGA			A(4*S*)MePGA		
conformer	δ_{BD} (Å)	τ_{BD} (deg)	conformer	δ_{BD} (Å)	τ_{BD} (deg)	conformer	δ_{BD} (Å)	τ_{BD} (deg)
1	2.99	108.8	1	3.08	116.5	1	2.85	114.2
2	3.05	120.4	2	3.13	111.6	2	2.99	113.3
3	3.11	126.1	3	3.30	113.0	3	3.00	113.4
4	3.12	112.0	4	3.23	90.5	4	2.83	100.5
5	2.84	113.6	5	2.89	107.9	5	3.12	86.5
6	2.97	113.8	6	2.86	114.8	6	3.14	116.2
7	3.32	79.2	7	3.06	92.9	7	2.99	125.0
8	3.01	113.6	8	3.10	137.9	8	2.99	113.4
9	3.11	120.4	9	3.29	117.7	9	3.27	78.0
10	3.10	124.4	10	3.23	133.7	10	3.09	114.9
11	2.94	125.6	11	3.03	111.0	11	2.92	104.0
12	3.23	94.2	12	2.82	114.9	12	2.93	104.2
13	3.00	114.4	13	3.03	115.4	13	3.27	83.7
			14	3.34	78.8			
			15	2.93	103.0			
			16	3.18	125.4			

of 100–113°, thus producing a favorable $n \rightarrow \pi^*$ interaction. In all these structures, the proline ring is in the C^γ *exo* puckering. The structures belonging to the distorted type I β-turn motif, type I bend motif, or extended motif exhibit O(i)···C′(i + 1) distances longer than 3.00 Å and/or significant deviation of the O(i)···C′=O(i + 1) angle from the ideal value of 107°. Therefore, no significant energetic stabilization from the $n \rightarrow \pi^*$ interaction could be reasonably expected for the latter structures.

In A(4*R*)MePGA, in which the pyrrolidine ring strongly prefers the C^γ *endo* puckering, only 4 NAMFIS structures have $\delta_{BD} \leq$ 3.00 Å. All three structures in which the pyrrolidine ring has C^γ *exo* puckering belong to this group. The structures in which the ψ(i + 1) dihedral angle is close to values compatible with the type II β-turn conformation exhibits the shortest O(i)···C′(i + 1) distances ($\delta_{BD} \leq$ 3.13 Å). Excluding one structure, these conformations also exhibit the O(i)···C′=O(i + 1) angle within the range of 103–116°. The structures belonging to the distorted type I β-turn motif, type I bend motif, or extended motif exhibit longer O(i)···C′(i + 1) distances ($\delta_{BD} \geq$ 3.18 Å), thus lacking the ability to form a meaningful $n \rightarrow \pi^*$ interaction. In only one extended structure is the O(i)···C′(i + 1) distance is shorter (δ_{BD} = 3.06 Å). In all these latter structures, the pyrrolidine ring is in the C^γ *endo* puckering. The trend observed for A(4*S*)MePGA and A(4*R*)MePGA is also evident for APGA. The NAMFIS structures in which the pyrrolidine ring is in the C^γ *exo* puckering and in which the ψ(i + 1) dihedral is close to values compatible with the type II β-turn conformation typically exhibit δ_{BD} and τ_{BD} values that can provide a significant energetic stabilization from the $n \rightarrow \pi^*$ interaction. On the contrary, with one exception, the structures in which the pyrrolidine ring is in the C^γ *endo* puckering exhibit δ_{BD} and τ_{BD} values that cannot provide a meaningful $n \rightarrow \pi^*$ interaction.

NBO analysis of Pro-Gly turns

The selected stabilizing hyperconjugative interactions in A(4*R*)MePGA and A(4*S*)MePGA were analyzed with density functional theory (DFT) calculations using the natural bond orbital (NBO) paradigm.[46] In the NBO analysis, the electronic wave functions are interpreted in terms of a set of occupied Lewis-type natural bond orbitals and a set of non Lewis-type localized orbitals (either two-center antibonding or one-center Rydberg). Delocalization effects can be identified from the presence of off diagonal elements of the Fock matrix in the NBO basis. Approximate stabilization energies, $E_{i,j^*}^{(2)}$, from these delocalization interactions can be determined for specific bond and antibond or Rydberg orbital pairs using second-order perturbation theory. Another NBO method, natural steric analysis (NSA),[64,65] was also applied to study certain repulsive interactions in these peptides. In this method, based on NBOs, steric interaction energies, $E_{i,j}^{(st)}$, can be calculated by identifying interatomic exchange repulsion energy components associated with wave function antisymmetry.

The tetrapeptide unit was truncated to a capped moiety corresponding to the fragment MeCO-Xaa-Gly-NHMe (Xaa = (4*R*)- or (4*S*)-MePro), and the resulting dipeptides were conformationally manipulated to produce the type I and type II β-turn structures. The puckering of the pyrrolidine ring was set to match C^γ *endo* in the (4*R*)-MePro peptides and C^γ *exo* in the (4*S*)-MePro peptides. Consequently, the methyl substituent is oriented pseudoequatorially in the corresponding four peptides. These structures were minimized without any constraints using the OPLS_2005 force field combined with the GB/SA H_2O continuum solvent model.[57] Each β-turn structure was further optimized, first at B3LYP/6-31+G* level of theory in gas phase, and second at B3LYP/6-311++G** level of theory using the PBF solvation model with DMSO as solvent.[66,67] Finally, NBO calculations were carried out for all four structures at the B3LYP/6-311++G** level in DMSO using the PBF solvation model and geometries obtained at the same level of theory. For ease of discussion, the (4*R*)-MePro and (4*S*)-MePro peptides have been designated (4*R*)MePG-I, (4*R*)MePG-II, (4*S*)MePG-I and (4*S*)MePG-II, in which the I and II implies the type I or type II β-turn conformation, respectively. All four MeCO-Xaa-Gly-NHMe units exhibit ϕ,ψ angles for proline and glycine residues that are within 30° of the idealized ϕ,ψ values corresponding to the (i +1) and (i + 2) residues of the respective turn type. In addition, on the grounds of geometric parameters listed in the section "Solution structures by NAMFIS analysis", all four peptides exhibit hydrogen bonding between O(i)

Table 6 Relative energies and geometric parameters for (4*R*)MePG-I, (4*R*)MePG-II, (4*S*)MePG-I and (4*S*)MePG-II[a]

peptide	ΔE (kcal mol^{-1})	ϕ(i + 1) (deg)	ψ(i + 1) (deg)	ϕ(i + 2) (deg)	ψ(i + 2) (deg)	δ_{BD} (Å)	τ_{BD} (deg)
(4*R*)MePG-I	0.9	−75.7	−1.8	−74.6	−11.2	3.147	132.4
(4*R*)MePG-II	0.5	−59.0	130.0	84.7	−5.3	2.868	104.8
(4*S*)MePG-I	1.1	−65.7	−15.1	−71.2	−12.2	3.009	125.8
(4*S*)MePG-II	0.0	−50.8	128.0	85.6	−6.3	2.805	101.8

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

and the NHMe group. The relative energies and the geometric parameters for (4*R*)MePG-I, (4*R*)MePG-II, (4*S*)MePG-I, and (4*S*)MePG-II are listed in Table 6. The B3LYP/6-311++G** optimized geometries of these dipeptides are shown in Fig. 4.

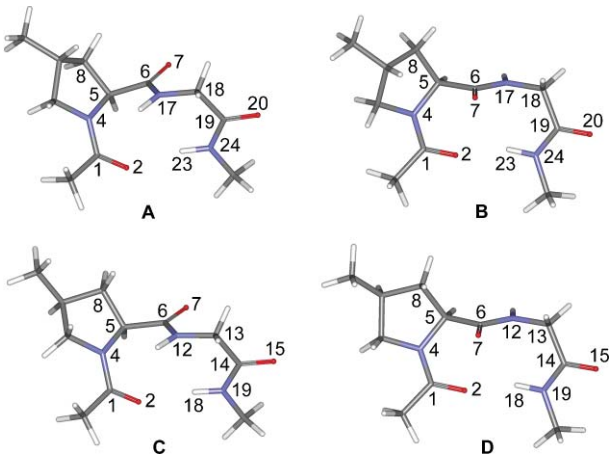


Fig. 4 The B3LYP/6-311++G** optimized geometries of (A) (4*R*)MePG-I, (B) (4*R*)MePG-II, (C) (4*S*)MePG-I, and (D) (4*S*)MePG-II.

The strongest donor character is shown by the lone pair of nitrogen atoms and the lone pair of oxygen atoms. The $n \rightarrow \sigma^*$ and $n \rightarrow \pi^*$ interactions between the lone pair of nitrogen atoms and the antibonding orbitals of vicinal C=O bonds correspond to the $E_{i,j^*}^{(2)}$ value as high as about 84 kcal mol^{-1}. The $n \rightarrow \sigma^*$ interactions between the lone pair of oxygen atoms and the vicinal C–C and C–N bonds reach the $E_{i,j^*}^{(2)}$ value of approximately 24 kcal mol^{-1}. The interactions between the above-mentioned lone pair orbitals and their vicinal C–C, C–N, or C=O filled orbitals exhibit the largest steric energies, $E_{i,j}^{(st)}$. Most stabilizing and destabilizing interactions in these peptides occur between vicinal NBOs. In addition, several long-range interactions between remote orbitals are also present. This analysis is focused on the latter interactions owing to the influence they can play on the peptide backbone conformation. Selected hyperconjugative and steric repulsive interactions between remote orbitals are given in Table 7.

First, the stabilizing $n \rightarrow \pi^*$ interaction between O(i) and C′=O(i + 1) is considered. In both (4*R*)MePG-II and (4*S*)MePG-II, the τ_{BD} angle from the amide oxygen O(i) to the proline carbonyl group C′=O(i + 1) is similar and very close to the optimal Bürgi-Dunitz trajectory ($\tau_{BD} = 107°$) for the attack of a nucleophile on a carbonyl group. On the contrary, in (4*R*)MePG-I and (4*S*)MePG-I, the τ_{BD} angle is much larger (>125°) and, therefore, not optimal for the nucleophilic attack. The distance δ_{BD} between the nucleophile and the electrophilic carbonyl carbon plays an important role in the stabilization that arises from the $n \rightarrow \pi^*$ interaction. This distance increases in the order of (4*S*)MePG-II < (4*R*)MePG-II < (4*S*)MePG-I < (4*R*)MePG-I. (Table 6). Thus, taking the δ_{BD} and τ_{BD} values into consideration, the $n \rightarrow \pi^*$ stabilization is expected to be strongest for (4*S*)MePG-II and weakest for (4*R*)MePG-I. Second-order perturbation estimates from NBO do indeed show increasing $n \rightarrow \pi^*$ stabilization energies, $E_{i,j^*}^{(2)}$, with decreasing distance between O(i) and C′(i + 1) (Tables 6 and 7). The type II β-turn peptides, (4*S*)MePG-II and (4*R*)MePG-II, show the highest stabilization energies (1.94 and 1.39 kcal mol^{-1}, respectively), while the type I β-turn peptides, (4*S*)MePG-I and (4*R*)MePG-I, show clearly reduced $n \rightarrow \pi^*$ stabilization (0.59 and 0.21 kcal mol^{-1}, respectively). Consequently, the stabilizing $n \rightarrow \pi^*$ interaction between O(i) and C′=O(i + 1) is more effective in the type II β-turn peptides than in the type I β-turn peptides. Furthermore, the (4*S*)MePro peptides exhibit stronger stabilization than the corresponding (4*R*)MePro peptides. A representation of the $n \rightarrow \pi^*$ interaction between O(i) and C′=O(i + 1) is given in Fig. 5.

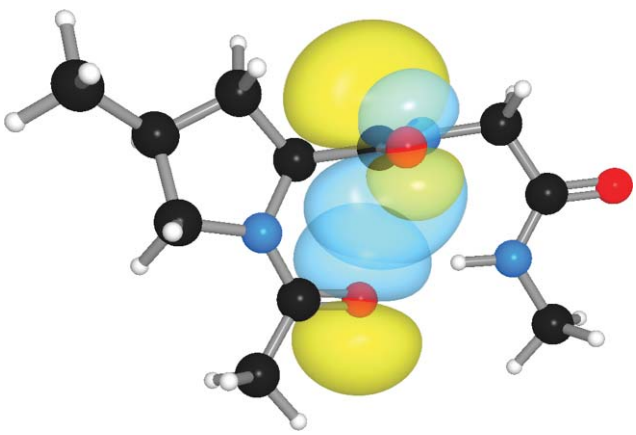


Fig. 5 3D representation of the $n \rightarrow \pi^*$ interaction between O(i) and C′=O(i + 1) in (4*S*)MePG-II.

In addition to the $n \rightarrow \pi^*$ interaction between O(i) and C′=O(i + 1), other stabilizing interactions between remote filled and unfilled orbitals are also present in MeCO-Xaa-Gly-NHMe (Xaa = (4*R*)- or (4*S*)-MePro) peptides (Table 7). The strongest remote interaction in all four peptides involves the O(i) lone pairs as donor and the N–H antibond orbital, σ*(N–H), of the NHMe group as acceptor, with n(Lp_1O) being a stronger donor than n(Lp_2O). The stabilization energies, $E_{i,j^*}^{(2)}$, associated

Table 7 Selected hyperconjugative and steric repulsive interactions between remote orbitals for (4*R*)MePG-I, (4*R*)MePG-II, (4*S*)MePG-I, and (4*S*)MePG-II[a]

peptide	stabilizing interaction	$E_{i,j^*}^{(2)}$ (kcal mol^{-1})	repulsive interaction	$E_{i,j}^{(st)}$ (kcal mol^{-1})
(4*R*)MePG-I	$n(Lp_2O_2) \rightarrow \pi^*(C_6O_7)$	0.21		
	$n(Lp_1O_2) \rightarrow \sigma^*(N_{24}H_{23})$	2.76	$n(Lp_1O_2) \leftrightarrow \sigma(N_{24}H_{23})$	2.26
	$n(Lp_2O_2) \rightarrow \sigma^*(N_{24}H_{23})$	1.71	$n(Lp_2O_2) \leftrightarrow \sigma(N_{24}H_{23})$	1.04
	$n(Lp_2O_7) \rightarrow \sigma^*(N_4C_5)$	1.08	$n(Lp_2O_7) \leftrightarrow \sigma(N_4C_5)$	0.39
	$n(Lp_2O_{20}) \rightarrow \sigma^*(N_{17}C_{18})$	1.08	$n(Lp_2O_{20}) \leftrightarrow \sigma(N_{17}C_{18})$	0.28
	$n(Lp_2O_2) \rightarrow \sigma^*(C_5C_8)$	0.68		
	$n(LpN_{17}) \rightarrow \pi^*(C_{19}O_{20})$	0.62		
	$\pi(C_1O_2) \rightarrow \sigma^*(N_{24}H_{23})$	1.10	$\pi(C_1O_2) \leftrightarrow \sigma(N_{24}H_{23})$	0.84
(4*R*)MePG-II	$n(Lp_2O_2) \rightarrow \pi^*(C_6O_7)$	1.39	$n(Lp_2O_2) \leftrightarrow \pi(C_6O_7)$	0.87
	$n(Lp_1O_2) \rightarrow \sigma^*(N_{24}H_{23})$	4.94	$n(Lp_1O_2) \leftrightarrow \sigma(N_{24}H_{23})$	4.19
	$n(Lp_2O_2) \rightarrow \sigma^*(N_{24}H_{23})$	2.14	$n(Lp_2O_2) \leftrightarrow \sigma(N_{24}H_{23})$	1.17
	$n(Lp_2O_{20}) \rightarrow \sigma^*(N_{17}C_{18})$	1.13	$n(Lp_2O_{20}) \leftrightarrow \sigma(N_{17}C_{18})$	0.31
	$n(Lp_2O_2) \rightarrow \sigma^*(C_5C_8)$	0.63		
	$n(LpN_4) \rightarrow \pi^*(C_6O_7)$	0.69	$n(LpN_4) \leftrightarrow \pi(C_6O_7)$	0.51
	$n(LpN_{17}) \rightarrow \pi^*(C_{19}O_{20})$	0.54		
	$\pi(C_1O_2) \rightarrow \sigma^*(N_{24}H_{23})$	0.86	$\pi(C_1O_2) \leftrightarrow \sigma(N_{24}H_{23})$	0.89
(4*S*)MePG-I	$n(Lp_2O_2) \rightarrow \pi^*(C_6O_7)$	0.59	$n(Lp_2O_2) \leftrightarrow \pi(C_6O_7)$	0.27
	$n(Lp_1O_2) \rightarrow \sigma^*(N_{19}H_{18})$	3.44	$n(Lp_1O_2) \leftrightarrow \sigma(N_{19}H_{18})$	2.82
	$n(Lp_2O_2) \rightarrow \sigma^*(N_{19}H_{18})$	1.34	$n(Lp_2O_2) \leftrightarrow \sigma(N_{19}H_{18})$	0.83
	$n(Lp_2O_7) \rightarrow \sigma^*(N_4C_5)$	1.06	$n(Lp_2O_7) \leftrightarrow \sigma(N_4C_5)$	0.34
	$n(Lp_2O_{15}) \rightarrow \sigma^*(N_{12}C_{13})$	1.09	$n(Lp_2O_{15}) \leftrightarrow \sigma(N_{12}C_{13})$	0.29
	$n(Lp_2O_2) \rightarrow \sigma^*(C_5C_8)$	0.69		
	$n(LpN_{12}) \rightarrow \pi^*(C_{14}O_{15})$	0.60		
	$\pi(C_1O_2) \rightarrow \sigma^*(N_{19}H_{18})$	1.29	$\pi(C_1O_2) \leftrightarrow \sigma(N_{19}H_{18})$	0.99
(4*S*)MePG-II	$n(Lp_2O_2) \rightarrow \pi^*(C_6O_7)$	1.94	$n(Lp_2O_2) \leftrightarrow \pi(C_6O_7)$	1.14
	$n(Lp_1O_2) \rightarrow \sigma^*(N_{19}H_{18})$	5.55	$n(Lp_1O_2) \leftrightarrow \sigma(N_{19}H_{18})$	4.75
	$n(Lp_2O_2) \rightarrow \sigma^*(N_{19}H_{18})$	2.15	$n(Lp_2O_2) \leftrightarrow \sigma(N_{19}H_{18})$	1.14
	$n(Lp_2O_{15}) \rightarrow \sigma^*(N_{12}C_{13})$	1.13	$n(Lp_2O_{15}) \leftrightarrow \sigma(N_{12}C_{13})$	0.32
	$n(Lp_2O_2) \rightarrow \sigma^*(C_5C_8)$	0.50		
	$n(LpN_4) \rightarrow \pi^*(C_6O_7)$	0.58	$n(LpN_4) \leftrightarrow \pi(C_6O_7)$	0.40
	$n(LpN_{12}) \rightarrow \pi^*(C_{14}O_{15})$	0.52		
	$\pi(C_1O_2) \rightarrow \sigma^*(N_{19}H_{18})$	0.55	$\pi(C_1O_2) \leftrightarrow \sigma(N_{19}H_{18})$	0.63

[a] The numbering of the atoms is shown in Fig. 4. Blank cells correspond to a value lower than the threshold printing value 0.1 kcal mol^{-1}.

with these hyperconjugative interactions are obtained as 7.70, 7.08, 4.78, and 4.74 kcal mol^{-1} for (4*S*)MePG-II, (4*R*)MePG-II, (4*S*)MePG-I and (4*R*)MePG-I, respectively, which quantify the extent of hydrogen bonding between O(i) and the NHMe group in these peptides. Another remote interaction present in all four peptides is the $n \rightarrow \sigma^*$ interaction between the lone pair of O(i + 2) and the glycine N–C$^\alpha$ antibond orbital. This interaction is strongest when the N–C$^\alpha$-C′-O dihedral angle of the glycine residue is close to 180° or, in other terms, when the ψ(i + 2) angle approaches 0° (Fig. 6). Consequently, this type of hyperconjugative interaction contributes to the stability of the peptide when the ψ(i + 2) angle approaches a value compatible with the idealized type I and type II β-turn conformation. The $E_{i,j^*}^{(2)}$ value associated with the $n(Lp_2O) \rightarrow \sigma^*$(N-C$^\alpha$) interaction is about 1.1 kcal mol^{-1} for all four peptides (Table 7). This type of $n \rightarrow \sigma^*$ interaction is also present in (4*R*)MePG-I and (4*S*)MePG-I between the lone pair of O(i + 1) and the proline N–C$^\alpha$ antibond orbital. In analogy to the above, this interaction is strongest when the N–C$^\alpha$–C′–O dihedral angle of the proline residue is close to 180°. Consequently, the $n \rightarrow \sigma^*$ interaction gives the strongest stabilization when the ψ(i + 1) angle is near 0°, which is about 30° away from the idealized value for the type I β-turn conformation (Fig. 7). As above, the stabilization energy, $E_{i,j^*}^{(2)}$, associated with the $n(Lp_2O) \rightarrow \sigma^*$(N-C$^\alpha$) interaction is about 1.1 kcal mol^{-1} for both (4*R*)MePG-I and (4*S*)MePG-I, while the ψ(i + 1) angle is –1.8° in (4*R*)MePG-I and –15.1° in (4*S*)MePG-I. As can be seen from Table 7, steric repulsive interactions between the above-mentioned remote orbitals are outweighed by the stabilizing hyperconjugative interactions. Thus, the net result from these interactions is a stabilizing contribution to the β-turn formation. Besides the interactions discussed above, Table 7 also includes the energies of other selected remote interactions. However, these are not discussed further here.

A particular repulsive $n \leftrightarrow \sigma$ interaction, not included in Table 7, deserves some comment. This interaction is apparent in (4*R*)MePG-II, in which the pyrrolidine ring of the proline residue

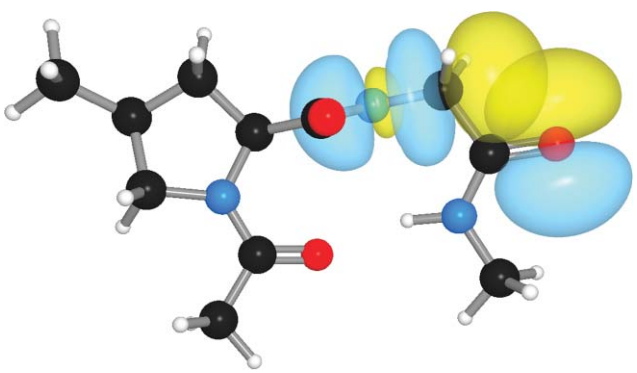


Fig. 6 3D representation of the $n \rightarrow \sigma^*$ interaction between O(i + 2) and the glycine N–C$^\alpha$ antibond orbital in (4*S*)MePG-II.

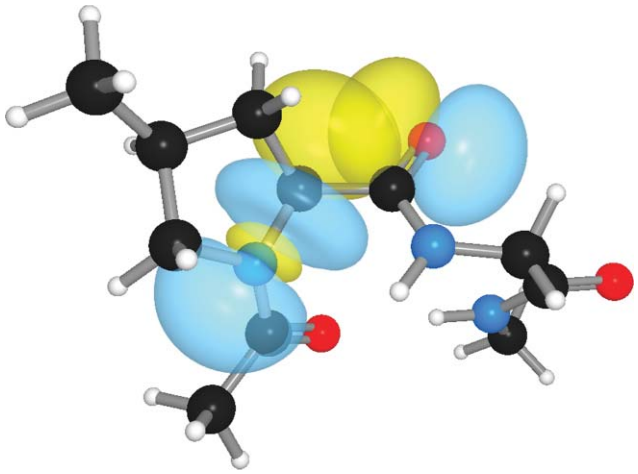


Fig. 7 3D representation of the $n \rightarrow \sigma^*$ interaction between O(i + 1) and the proline N–C$^{\alpha}$ antibond orbital in (4*S*)MePG-I.

is in the C$^{\gamma}$ *endo* conformation, thus bringing the lone pair of O(i + 1) and the proline C$^{\gamma}$–H filled orbital in close contact to each other. The distance between the oxygen lone pair and the C$^{\gamma}$–H bond, and consequently, the steric energy, $E_{i,j}^{(st)}$, associated with this interaction is dependent on the proline ϕ angle. The more negative values of proline ϕ angle bring the oxygen lone pair and the C$^{\gamma}$–H bond closer to each other, thus increasing the repulsion. The proline ϕ angle in the DFT optimized structure of (4*R*)MePG-II is −59.0°, which results in an $E_{i,j}^{(st)}$ value of 0.61 kcal mol^{-1}. If the ϕ angle alone is decreased to −65°, steric repulsion between the lone pair and the C$^{\gamma}$–H bond increases to 1.21 kcal mol^{-1}. This repulsion can be relieved in three ways (if the pyrrolidine ring stays in the *endo* conformation): 1) the proline ψ angle approaches larger values (positive or negative) compatible with the extended structures, 2) the proline ψ angle approaches small negative values compatible with type I β-turn conformations, and 3) the pyrrolidine ring adopts a more flattened conformation. The repulsive $n \leftrightarrow \sigma$ interaction between the lone pair of O(i + 1) and the proline C$^{\gamma}$–H filled orbital may provide one explanation why the combined population of type I β-turns, type I bends, and extended structures is considerably higher for A(4*R*)MePGA (38%) than that for APGA and A(4*S*)MePGA (25% for both peptides).

Effect of remote interactions on peptide conformation

The results from the NBO analysis suggest that the non-bonding interactions between remote filled and unfilled orbitals have important implications in determining solution conformations of the peptides under study. Particularly interesting is the $n \rightarrow \pi^*$ interaction between the lone pair of the Ala^1 carbonyl oxygen and the antibonding orbital of the proline carbonyl group. The C$^{\gamma}$ *exo* ring puckering in A(4*S*)MePGA preorganizes the main chain ϕ,ψ angles of the proline residue for this favorable O(i)$\cdots$C′=O(i + 1) interaction, which stabilizes conformations in which the proline ψ dihedral angle is close to values compatible with the type II β-turn conformation. This, in turn, may increase the probability that the ϕ,ψ angles of the glycine residue approach values also compatible with the type II β-turn conformation, thus making the β-turn stabilizing hydrogen bond forming between Ala^1-CO and Ala^2-NH more probable. Remote $n \rightarrow \sigma^*$ interaction between the lone pair of O(i + 2) and the glycine N–C$^{\alpha}$ antibond orbital may help to increase the probability of hydrogen bonding by stabilizing conformations in which the glycine ψ angle approaches values around 0°. As a consequence of the above, the peptide spends more time in conformations related to type II β-turn, increasing the type II turn population. The C$^{\gamma}$ *endo* ring puckering in A(4*R*)MePGA, on the contrary, favors proline ϕ,ψ angles that result in less effective $n \rightarrow \pi^*$ interaction between O(i) and C′=O(i + 1). Therefore, in A(4*R*)MePGA, conformations in which the proline ψ dihedral is close to values compatible with the type II β-turn conformation are less stabilized in comparison with A(4*S*)MePGA, in which the C$^{\gamma}$ *exo* ring puckering of the proline residue prevails. Consequently, although the above-mentioned $n \rightarrow \sigma^*$ interaction is effective also in A(4*R*)MePGA, there may be a reduced probability that the glycine residue samples ϕ,ψ angles compatible with the type II β-turn conformation. This, in turn, makes the β-turn stabilizing hydrogen bond forming between Ala^1-CO and Ala^2-NH less probable, decreasing the type II β-turn population of the peptide. The magnitude of the stabilizing $n \rightarrow \pi^*$ interaction between the lone pair of the Ala^1 carbonyl oxygen and the antibonding orbital of the proline carbonyl group is significantly reduced in structures in which the ψ(i + 1) dihedral is close to values compatible with type I β-turn conformation, *i.e.* in type I β-turns and bends. The same applies to the extended conformations. Therefore, in these structures, stabilization from the hyperconjugative interaction between O(i) and C′=O(i + 1) is very small. However, the type I β-turn conformations may be somewhat stabilized by the remote $n \rightarrow \sigma^*$ interaction between the lone pair of O(i + 1) and the proline N–C$^{\alpha}$ antibond orbital. This interaction gives the strongest stabilization when the ψ(i + 1) angle is near 0°, which is about 30° away from the idealized value for the type I β-turn conformation.

Conclusion

In this study, we have determined the ensembles of conformations for three linear tetrapeptides (APGA, A(4*R*)MePGA, and A(4*S*)MePGA) in DMSO-d_6 solution using the NAMFIS methodology. It was shown that enantiospecific incorporation of the methyl substituent at the C$^{\gamma}$ (C4) position of the proline residue can be used to selectively control the pyrrolidine ring puckering propensities and, consequently, the preferred ϕ,ψ angles associated with the proline residue in β-turn forming peptides. Thus, the populations of the β-turn conformations of the peptide in solution can be controlled by a minor chemical change on the natural amino acid. In addition, we have employed the NBO method to gain further insight into the results obtained from the NAMFIS analysis. The outcome from the NBO calculations shed light on the underlying principles responsible for the conformational preferences of β-turns.

The NAMFIS selected conformations could be reduced to 7 structural families (motifs) to reproduce the interproton distances and the torsion angles obtained from the averaged NMR data (Table 4). As expected, the NAMFIS analysis demonstrated that the pyrrolidine ring of (4*R*)-methylproline in A(4*R*)MePGA has a strong preference for the C$^{\gamma}$ *endo* puckering, whereas (4*S*)-methylproline in A(4*S*)MePGA has a strong preference for the C$^{\gamma}$

exo puckering. In the preferred conformations, the 4-methyl substituent adopts a pseudoequatorial orientation. In APGA, the two puckerings are essentially equally populated among the NAMFIS conformers. According to the NAMFIS analysis, the combined type II β-turn population of A(4*S*)MePGA is considerably higher than that of APGA and A(4*R*)MePGA. The results from the NBO analysis suggest that this observation can be rationalized by an energetic stabilization associated with a non-bonded $n \rightarrow \pi^*$ interaction between the lone pair of the Ala[1] carbonyl oxygen and the antibonding orbital of the proline carbonyl group. The dominant C^γ *exo* ring puckering in A(4*S*)MePGA preorganizes the main chain ϕ,ψ angles of the proline residue for this remote O(i)···C′=O(i + 1) interaction, whereas in A(4*R*)MePGA the prevailing C^γ *endo* ring puckering favors proline ϕ,ψ angles that results in less effective $n \rightarrow \pi^*$ interaction between O(i) and C′=O(i + 1). In APGA, the more flexible pyrrolidine ring of the proline residue results in a combined type II β-turn population that is comparable to that of A(4*R*)MePGA.

Several other remote interactions provide stabilization for both the type I and II β-turn conformations. The strongest of these occur between the O(i) lone pairs as donor and the N–H antibonding orbital, σ*(N–H), of the NHMe group as acceptor, quantifying the extent of hydrogen bonding between O(i) and the NHMe group. The $n \rightarrow \sigma^*$ interaction between the lone pair of O(i + 2) and the glycine N–C^α antibonding orbital, can potentially help the formation of the above-mentioned hydrogen bond by stabilizing conformations in which the glycine ψ angle approaches values around 0°. Type I β-turn conformations may gain some stabilization from the $n \rightarrow \sigma^*$ interaction between the lone pair of O(i + 1) and the proline N–C^α antibonding orbital. This interaction gives the strongest stabilization when the ψ(i + 1) angle approaches 0°, which is about 30° away from the idealized value for the type I β-turn conformation. Finally, the repulsive $n \leftrightarrow \sigma$ interaction between the lone pair of O(i + 1) and the proline C^γ–H filled orbital may provide an explanation for the observation that the combined population of the type I β-turns, type I bends, and extended structures is considerably higher for A(4*R*)MePGA than that for APGA and A(4*S*)MePGA.

There is a great interest in mimicking β-turns for the synthesis of medicines in the field of medical and pharmacological chemistry. Although the peptides studied are not biologically important *per se*, they provided a simple model system to investigate different backbone-backbone interactions that contribute to β-turn stability and, therefore, β-turn formation in peptides and proteins. In addition, they lend themselves to examination of the factors behind the relationship between the proline ring puckering and the peptide backbone conformation. Further, peptides such as A(4*R*)MePGA and A(4*S*)MePGA potentially offer a probe for molecular recognition studies since the 4-methyl substituent of the pyrrolidine ring strongly favors pseudoequatorial orientation, thus providing the possibility to deduce directionality requirements in the binding of β-turns and of polypeptides in general. We anticipate that this investigation has the potential to improve our understanding of structural parameters that influence conformational behavior of flexible β-turn forming peptides in liquid phase. The findings from this study could be useful in the *de novo* design of peptides, peptidomimetic compounds, and proteins, and in the development of peptide-based drugs.

Experimental section

Tetrapeptide synthesis

Detailed procedures for the peptide synthesis are provided in the Supporting Information. 4-Methylprolines suitable for solid phase synthesis were prepared according to Scheme S1 (Supporting Information). The solid phase synthesis of APGA, A(4*R*)MePGA, and A(4*S*)MePGA is described in Scheme S2 (Supporting Information).

NMR spectroscopy of tetrapeptides

NMR studies were carried out on an Bruker Avance DPX400 spectrometer operating at 400.13 and 100.61 MHz for ^{1}H and ^{13}C, respectively. All experiments were performed in DMSO-d_6 solution at 299 K. The sample concentration was 80 mM for APGA, 77 mM for A(4*R*)MePGA, and 74 mM for A(4*S*)MePGA. Chemical shifts are reported in ppm relative to TMS (δ 0.00) for ^{1}H NMR. For ^{13}C spectra, the residual DMSO (δ 39.51) was used as internal standard. All 1D and 2D spectra were recorded using standard pulse sequences from the Bruker pulse program library in TopSpin 1.3. Data processing was made using TopSpin 2.0 software. The ^{1}H chemical shift assignments were made using a combination of TOCSY and T-ROESY experiments. Interproton distances were obtained from the adiabatic version of off-resonance ROESY experiments.[47] This experiment was chosen because it has been shown to provide both efficient suppression of HOHAHA transfers and reduction of offset effects. The ^{1}H NMR spectral analyses were performed with the PERCHit iterator under PERCH software.[48,49] The Total-Line-Shape-Fitting option of the PERCHit iterator was used for the final refinement of the result. Prior to the spin system analysis, the resolution of the ^{1}H NMR spectra was enhanced using the combination of exponential and trapezoidal windowing before Fourier transformation.[68] The ^{13}C chemical shifts were assigned by a combination of HSQC and CIGAR-HMBC experiments. These assignments are given in Table S4 (Supporting Information).

One-dimensional (1D) ^{1}H spectra were acquired using a 1 s relaxation delay, 64 transients, a spectral width of 10 ppm, and 64 K data points zero-filled to 128 K data points before Fourier transformation and baseline correction. 1D ^{13}C spectra were recorded using 2 s relaxation delay, a spectral width of 180 or 190 ppm, 1 K transients, and 64 K data points. The phase-sensitive TOCSY spectra were obtained with a mixing time of 80 ms using the MLEV17 spin lock sequence. A typical experiment was recorded using a 2 s relaxation delay, a spectral width of 8.6 ppm in both dimensions, 2 K data points, 512 t_1 increments, and 8 transients. The data was zero-filled to a 2 K × 2 K matrix before Fourier transformation. A sine-bell weighting function was used in both dimensions. The phase-sensitive T-ROESY spectra were obtained with mixing times of 100, 200, and 300 ms. Typical spectra were recorded using 2 s relaxation delay, a spectral width of 8.6 ppm in both dimensions, 2 K data points, 512 t_1 increments, and 8 transients. The data was zero-filled to a 2 K × 2 K matrix before Fourier transformation. A sine-bell weighting function was used in both dimensions. The phase-sensitive off-resonance ROESY spectra were obtained with mixing times of 70, 140, 210, 280, 350, and 420 ms to check the linearity of the cross-relaxation buildup. The mixing time of 180 ms was chosen for

the quantitative off-resonance ROESY experiments. The mixing sequence was achieved by adiabatic rotation. The spin-lock field strength γB_1 was 5107 Hz and the Θ angle was equal to 45°. A typical experiment was recorded using 2 s relaxation delay, a spectral width of 8.6 ppm in both dimensions, 2 K data points, 512 t_1 increments, and 16 transients. The data was zero-filled to a 2 K × 2 K matrix before Fourier transformation. A sine-bell weighting function was used in both dimensions. Typical phase-sensitive HSQC spectra were acquired with spectral widths of 8.6 ppm for ^{1}H and 190 ppm for ^{13}C using 2 s relaxation delay, 2 K data points, 512 t_1 increments, and 4 transients. The data was zero-filled to a 2 K × 2 K matrix before Fourier transformation. A sine squared window multiplication was used in both dimensions. The CIGAR-HMBC spectra were obtained in the magnitude mode typically with spectral widths of 8.6 ppm for ^{1}H and 190 ppm for ^{13}C, 6–10 Hz optimization range for long-range heteronuclear couplings, 140–180 Hz range for suppression of one-bond coupling responses, and the *J*-scaling factor of 0. These experiments were recorded using 2 s relaxation delay, 2 K data points, 512 t_1 increments, and 16 transients. The data was zero-filled to a 2 K × 2 K matrix before Fourier transformation. A sine squared window multiplication was used in the *F*2 dimension and a sine-bell weighting function in the *F*1 dimension. A polynomial baseline correction was applied for all 2D experiments mentioned above.

NAMFIS analysis

The extended structures of APGA, A(4*R*)MePGA, and A(4*S*)MePGA were used as starting points for all conformational searches. For each peptide, three different force fields (OPLS_2005, AMBER*, and MMFFs) were used to generate conformations in MacroModel 9.5[56] combined with the GB/SA H_2O continuum solvent model.[57] This solvent model is well-parameterized for a range of organic structures and was chosen to mimic polar solvent to dampen unrealistic gas-phase dipole–dipole interactions. NAMFIS analyses, in which conformational searches have been performed with the GB/SA H_2O model and the NMR parameters have been derived in DMSO-d_6, have been published earlier.[17,20,69,70] Each force field was applied to two separate random-seeded searches of 90000 MCMM (Monte Carlo Multiple Minimum)[71] steps using the PRGC algorithm. The minimized conformers from the duplicate runs were combined and optimized to convergence with their respective force fields using the PRGC algorithm and a 6.0 kcal mol^{-1} (25 kJ mol^{-1}) energy cutoff. Duplicate conformers were eliminated at this point. The global energy minimum was found at least 179 times for each force field and peptide. Finally, all structures from the multiple force field optimizations were combined using a redundant conformer elimination to form a single conformation pool for each peptide. This resulted in 2259, 1996, and 2485 conformers for APGA, A(4*R*)MePGA, and A(4*S*)MePGA, respectively. These conformers were subjected to the NAMFIS deconvolution procedure using interproton distances and torsion angles obtained from the averaged NMR data. NAMFIS performs a least-squares fit of the above-mentioned experimental data against the same data extracted from the conformer database. Goodness of fit is expressed as the sum of square differences (SSD) between experimental and calculated variables. The result is a population of conformers (set of mole fractions) that optimally represents the experimental NMR parameters. SSD values for APGA, A(4*R*)MePGA, and A(4*S*)MePGA were 8.7, 7.9, and 9.1, respectively. Interproton distances for the NAMFIS analysis were obtained from the off-resonance ROESY experiment with a mixing time of 180 ms using the initial rate approximation.[72] Average cross-peak volumes were determined *via* integration of rectangular peak areas from both sides of the diagonal using TopSpin 2.0. In the case of APGA, the obtained distances were calibrated with respect to the distance between Ala^2-Hα and Ala^2-CH_3 (2.45 Å). For A(4*R*)MePGA and A(4*S*)MePGA, the distance between proline H$\delta^{pro(R)}$ and H$\delta^{pro(S)}$ (1.78 Å) was used. Error bars used by NAMFIS for off-resonance ROESY derived distances (d_{NOE}, in Å) were as follows: $d_{NOE} <$ 2.5 (±0.1), $2.5 \leq d_{NOE} < 3.0$ (±0.2), $3.0 \leq d_{NOE} < 3.5$ (±0.3), and $3.5 \leq d_{NOE} < 6.0$ (±0.4). An extended Karplus equation developed by Haasnoot *et al.*[73] (parameter set E) was used to calculate the pyrrolidine ring ^3J(H,H) coupling constants. For the backbone ^3J(NH,Hα) couplings, the Karplus-type equation parameterized by Pardi *et al.*[74] was employed. Both equations were associated with an error of ±1.5 Hz. NAMFIS deconvolution was performed on all possible permutations of methylene protons. Permutations were handled by an external program permute.py. Specific permutation with lowest SSD corresponded to the correct assignment. Initial conformer populations were randomized using the random option implemented in NAMFIS program in order to avoid "local minima" as well as to ensure that the solution was unique. Several additional NAMFIS analyses were also performed in order to check the robustness of the NAMFIS-selected peptide conformations and to explore the sensitivity of NAMFIS populations to variations of the experimental datasets. These data are provided as Supporting Information.

Density functional calculations

All DFT calculations were performed using the Jaguar 7.0 program package.[75] Natural bond orbital (NBO) analysis was performed using the NBO 5.0 program[76] interfaced into the Jaguar 7.0 software. Orbital graphics were produced by the NBOView 1.1 program.[77] For visualization purposes, 3D orbital diagrams display the pre-orthogonal NBO (PNBO) associated with each NBO. PNBOs are a set of localized Lewis NBO precursors that lack the final interatomic orthogonalization step. They provide a convenient way to visualize interactions between orbitals in different bond regions, because their overlap is proportional to their interaction energy.[78] Starting geometries for the calculations were build using the Maestro 8.0 program.[79] The tetrapeptide unit A(4*R*)MePGA and A(4*S*)MePGA was truncated to a capped moiety corresponding to the fragment MeCO-Xaa-Gly-NHMe (Xaa = (4*R*)- or (4*S*)-MePro), and the resulting dipeptides were conformationally manipulated to produce the type I and type II β-turn structures. The puckering of the pyrrolidine ring was set to match C$^\gamma$ *endo* in the (4*R*)-MePro peptides and C$^\gamma$ *exo* in the (4*S*)-MePro peptides. Consequently, the methyl substituent is oriented pseudoequatorially in corresponding four peptides. These structures were minimized without any constraints using the OPLS_2005 force field combined with the GB/SA H_2O continuum solvent model.[57] Each β-turn structure was further optimized, first at B3LYP/6-31+G* level of theory in gas phase, and second at B3LYP/6-311++G** level of theory using the PBF solvation model[66,67] with DMSO as solvent. Frequency

calculations at the B3LYP/6-311++G** level of theory (in gas phase) were used to confirm that the optimized structures were minima, as characterized by positive vibrational frequencies. Finally, NBO calculations were carried out for all four structures at the B3LYP/6-311++G** level in DMSO using the PBF solvent model and geometries obtained at the same level of theory. The stabilizing and destabilizing NBO interaction energies were considered down to 0.1 kcal mol^{-1}.

Acknowledgements

ETTK wishes to thank The Graduate School of Organic Chemistry and Chemical Biology, Orion Farmos Research Foundation and Tekniikan Edistämissäätiö (TES) for support. We are greatly indebted to Dr James P. Snyder (Emory University, Atlanta, GA, USA) for providing the NAMFIS program. We also thank Matthew Geballe (Emory University, Atlanta, GA, USA) for teaching the use of the NAMFIS software.

Notes and references

1 G. D. Rose, L. M. Gierasch and J. A. Smith, *Adv. Protein Chem.*, 1985, **37**, 1–109.
2 W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577–2637.
3 A. M. C. Marcelino and L. M. Gierasch, *Biopolymers*, 2008, **89**, 380–391.
4 E. G. Hutchinson and J. M. Thornton, *Protein Sci.*, 1994, **3**, 2207–2216.
5 K. S. Kee and S. D. S. Jois, *Curr. Pharm. Des.*, 2003, **9**, 1209–1224.
6 K. Wüthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, 1986.
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 N. Nevins, D. Cicero and J. P. Snyder, *J. Org. Chem.*, 1999, **64**, 3979–3986.
11 D. Cicero, G. Barbato and R. Bazzo, *J. Am. Chem. Soc.*, 1995, **117**, 1027–1033.
12 P. Thepchatri, D. O. Cicero, E. Monteagudo, A. K. Ghosh, B. Cornett, E. R. Weeks and J. P. Snyder, *J. Am. Chem. Soc.*, 2005, **127**, 12838–12846.
13 J. P. Snyder, A. S. Lakdawala and M. J. Kelso, *J. Am. Chem. Soc.*, 2003, **125**, 632–633.
14 G. Barbato, E. Bianchi, P. Ingallinella, W. H. Hurni, M. D. Miller, G. Ciliberto, R. Cortese, R. Bazzo, J. W. Shiver and A. Pessi, *J. Mol. Biol.*, 2003, **330**, 1101–1115.
15 M. J. Kelso, R. L. Beyer, H. N. Hoang, A. S. Lakdawala, J. P. Snyder, W. V. Oliver, T. A. Robertson, T. G. Appleton and D. P. Fairlie, *J. Am. Chem. Soc.*, 2004, **126**, 4828–4842.
16 J. P. Snyder, N. Nevins, D. O. Cicero and J. Jansen, *J. Am. Chem. Soc.*, 2000, **122**, 724–725.
17 E. Monteagudo, D. O. Cicero, B. Cornett, D. C. Myles and J. P. Snyder, *J. Am. Chem. Soc.*, 2001, **123**, 6929–6930.
18 J. H. Nettles, H. Li, B. Cornett, J. M. Krahn, J. P. Snyder and K. H. Downing, *Science*, 2004, **305**, 866–869.
19 M. Erdélyi, B. Pfeiffer, K. Hauenstein, J. Fohrer, J. Gertsch, K.-H. Altmann and T. Carlomagno, *J. Med. Chem.*, 2008, **51**, 1469–1473.
20 P. Thepchatri, T. Eliseo, D. O. Cicero, D. Myles and J. P. Snyder, *J. Am. Chem. Soc.*, 2007, **129**, 3127–3134.
21 D. F. DeTar and N. P. Luthra, *J. Am. Chem. Soc.*, 1977, **99**, 1232–1244.
22 L. Vitagliano, R. Berisio, A. Mastrangelo, L. Mazzarella and A. Zagari, *Protein Sci.*, 2001, **10**, 2627–2632.
23 M. L. DeRider, S. J. Wilkens, M. J. Waddell, L. E. Bretscher, F. Weinhold, R. T. Raines and J. L. Markley, *J. Am. Chem. Soc.*, 2002, **124**, 2497–2505.
24 R. Improta, C. Benzi and V. Barone, *J. Am. Chem. Soc.*, 2001, **123**, 12568–12577.
25 M. D. Shoulders, I. A. Guzei and R. T. Raines, *Biopolymers*, 2008, **89**, 443–454.
26 S. A. Cadamuro, R. Reichold, U. Kusebauch, H.-J. Musiol, C. Renner, P. Tavan and L. Moroder, *Angew. Chem., Int. Ed.*, 2008, **47**, 2143–2146.
27 E. S. Eberhardt, N. Panasik and R. T. Raines, *J. Am. Chem. Soc.*, 1996, **118**, 12261–12266.
28 L. E. Bretscher, C. L. Jenkins, K. M. Taylor, M. L. DeRider and R. T. Raines, *J. Am. Chem. Soc.*, 2001, **123**, 777–778.
29 C. Renner, S. Alefelder, J. H. Bae, N. Budisa, R. Huber and L. Moroder, *Angew. Chem., Int. Ed.*, 2001, **40**, 923–925.
30 W. Kim, A. McMillan, J. P. Snyder and V. P. Conticello, *J. Am. Chem. Soc.*, 2005, **127**, 18121–18132.
31 M. D. Shoulders, J. A. Hodges and R. T. Raines, *J. Am. Chem. Soc.*, 2006, **128**, 8112–8113.
32 J. L. Flippen-Anderson, R. Gilardi, I. L. Karle, M. H. Frey, S. J. Opella, L. M. Gierasch, M. Goodman, V. Madison and N. G. Delaney, *J. Am. Chem. Soc.*, 1983, **105**, 6609–6614.
33 A. M. P. Koskinen, J. Helaja, E. T. T. Kumpulainen, J. Koivisto, H. Mansikkamäki and K. Rissanen, *J. Org. Chem.*, 2005, **70**, 6447–6453.
34 P. Chakrabarti and D. Pal, *Prog. Biophys. Mol. Biol.*, 2001, **76**, 1–102.
35 B. K. Ho, E. A. Coutsias, C. Seok and K. A. Dill, *Protein Sci.*, 2005, **14**, 1011–1018.
36 J.-C. Horng and R. T. Raines, *Protein Sci.*, 2006, **15**, 74–83.
37 D. Naduthambi and N. J. Zondlo, *J. Am. Chem. Soc.*, 2006, **128**, 12430–12431.
38 N. B. Malkar, J. L. Lauer-Fields, J. A. Borgia and G. B. Fields, *Biochemistry*, 2002, **41**, 6054–6064.
39 M. Doi, Y. Nishi, S. Uchiyama, Y. Nishiuchi, T. Nakazawa, T. Ohkubo and Y. Kobayashi, *J. Am. Chem. Soc.*, 2003, **125**, 9922–9923.
40 A. V. Persikov, J. A. M. Ramshaw, A. Kirkpatrick and B. Brodsky, *J. Am. Chem. Soc.*, 2003, **125**, 11500–11501.
41 D. Barth, A. G. Milbradt, C. Renner and L. Moroder, *ChemBioChem*, 2004, **5**, 79–86.
42 F. W. Kotch, I. A. Guzei and R. T. Raines, *J. Am. Chem. Soc.*, 2008, **130**, 2952–2953.
43 W. Kim, A. George, M. Evans and V. P. Conticello, *ChemBioChem*, 2004, **5**, 928–936.
44 S. C. R. Lummis, D. L. Beene, L. W. Lee, H. A. Lester, R. W. Broadhurst and D. A. Dougherty, *Nature*, 2005, **438**, 248–252.
45 C. Boulègue, A. G. Milbradt, C. Renner and L. Moroder, *J. Mol. Biol.*, 2006, **358**, 846–856.
46 A. E. Reed, L. A. Curtiss and F. Weinhold, *Chem. Rev.*, 1988, **88**, 899–926.
47 H. Desvaux, P. Berthault, N. Birlirakis, M. Goldman and M. Piotto, *J. Magn. Reson., Ser. A*, 1995, **113**, 47–52.
48 R. Laatikainen, M. Niemitz, U. Weber, J.-P. Sundelin, T. Hassinen and J. Vepsäläinen, *J. Magn. Reson., Ser. A*, 1996, **120**, 1–10.
49 *PERCH NMR Software, version 1/2005*, PERCH Solutions Ltd., Kuopio, Finland, 2005.
50 P. Amodeo, A. Motta, D. Picone, G. Saviano, T. Tancredi and P. A. Temussi, *J. Magn. Reson.*, 1991, **95**, 201–207.
51 K. Sorimachi and D. J. Craik, *Eur. J. Biochem.*, 1994, **219**, 237–251.
52 O. M. A. El-Agnaf, G. B. Irvine, G. Fitzpatrick, W. K. Glass and D. J. S. Guthrie, *Biochem. J.*, 1998, **336**, 419–427.
53 P. Desai, M. Prachand, E. Coutinho, A. Saran, J. Bodi and H. Süli-Vargha, *Int. J. Biol. Macromol.*, 2002, **30**, 187–195.
54 G. Shanmugam and P. L. Polavarapu, *Biophys. J.*, 2004, **87**, 622–630.
55 J. F. O'Connell, A. M. P. Koskinen, M. Krische and H. Rapoport, unpublished work.
56 MacroModel, version 9.5, Schrödinger, LLC, New York, NY, 2007.
57 W. C. Still, A. Tempczyk, R. C. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127–6129.
58 P. N Lewis, F. A. Momany and H. A. Scheraga, *Biochem. Biophys. Acta*, 1973, **303**, 211–229.
59 H. B. Bürgi, J. D. Dunitz and E. Shefter, *J. Am. Chem. Soc.*, 1973, **95**, 5065–5067.
60 H. B. Bürgi, J. M. Lehn and G. Wipff, *J. Am. Chem. Soc.*, 1974, **96**, 1956–1957.
61 H. B. Bürgi, J. D. Dunitz, J. M. Lehn and G. Wipff, *Tetrahedron*, 1974, **30**, 1563–1572.
62 H. B. Bürgi, J. D. Dunitz and E. Shefter, *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.*, 1974, **30**, 1517–1527.
63 J. A. Hodges and R. T. Raines, *Org. Lett.*, 2006, **8**, 4695–4697.
64 J. K. Badenhoop and F. Weinhold, *J. Chem. Phys.*, 1997, **107**, 5406–5421.

65 J. K. Badenhoop and F. Weinhold, *Int. J. Quantum Chem.*, 1999, **72**, 269–280.
66 D. J. Tannor, B. Marten, R. Murphy, R. A. Friesner, D. Sitkoff, A. Nicholls, M. Ringnalda, W. A. Goddard, III and B. Honig, *J. Am. Chem. Soc.*, 1994, **116**, 11875–11882.
67 B. Marten, K. Kim, C. Cortis, R. A. Friesner, R. B. Murphy, M. N. Ringnalda, D. Sitkoff and B. Honig, *J. Phys. Chem.*, 1996, **100**, 11775–11788.
68 R. Laatikainen, K. Tuppurainen, Y. Hiltunen and S. Lötjönen, *Magn. Reson. Chem.*, 1990, **28**, 939–946.
69 L. Barboni, C. Lambertucci, G. Appendino, D. G. Vander Velde, R. H. Himes, E. Bombardelli, M. Wang and J. P. Snyder, *J. Med. Chem.*, 2001, **44**, 1576–1587.
70 L. Barboni, A. Datta, D. Dutta, G. I. Georg, D. G. Vander Velde, R. H. Himes, M. Wang and J. P. Snyder, *J. Org. Chem.*, 2001, **66**, 3321–3329.
71 G. Chang, W. C. Guida and W. C. Still, *J. Am. Chem. Soc.*, 1989, **111**, 4379–4386.
72 D. Neuhaus and M. P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, 2nd edn, Wiley-VHC, New York, 2000.
73 C. A. G. Haasnoot, F. A. A. M. de Leeuw and C. Altona, *Tetrahedron*, 1980, **36**, 2783–2792.
74 A. Pardi, M. Billeter and K. Wüthrich, *J. Mol. Biol.*, 1984, **180**, 741–751.
75 Jaguar, version 7.0, Schrödinger, LLC, New York, NY, 2007.
76 E. D. Glendening, J. K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales and F. Weinhold, *NBO 5.0*, Theoretical Chemistry Institute, University of Wisconsin, Madison, WI, 2001.
77 M. Wendt and F. Weinhold, *NBOView 1.1*, NBO Orbital Graphics, Department of Chemistry, University of Wisconsin, Madison, WI, 2001.
78 S. J. Wilkens, W. M. Westler, J. L. Markley and F. Weihold, *J. Am. Chem. Soc.*, 2001, **123**, 12026–12036.
79 Maestro, version 8.0, Schrödinger, LLC, New York, NY, 2007.

