JOC Article

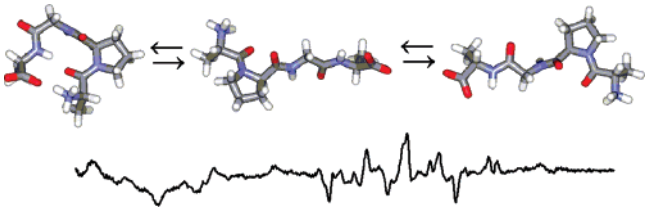
Conformational Properties of the Pro-Gly Motif in the D-Ala-L-Pro-Gly-D-Ala Model Peptide Explored by a Statistical Analysis of the NMR, Raman, and Raman Optical Activity Spectra

Miloš Buděšínský,[†] Jaroslav Šebestík,[†] Lucie Bednárová,[†] Vladimír Baumruk,[‡] Martin Šafařík,[†] and Petr Bouř*,[†]

Institute of Organic Chemistry and Biochemistry, Academy of Sciences, Flemingovo nám. 2, 16610, Prague 6, Czech Republic, and Charles University, Faculty of Mathematics and Physics, Institute of Physics, Ke Karlovu 5, 12116, Prague 2, Czech Republic

bour@uochb.cas.cz.

Received October 23, 2007



The Pro-Gly sequence in designed peptides and proteins is often used to mimic natural β-hairpin turns. Shorter peptides containing this moiety, however, adopt multiple conformations, and their propensity to form the turn is not obvious. In this study, conformational flexibility of Pro-Gly was investigated with the aid of NMR, Raman scattering, and Raman optical activity (ROA). The spectra of a model tetrapeptide NH_3^+-D-Ala-L-Pro-Gly-D-Ala-CO_2^- were analyzed on the basis of statistical methods and density functional computations. Other peptide derivatives were adopted as supporting theoretical and experimental models. The results suggest that the loop conformation of the Pro-Gly core is not inherently stable in vacuum. On the other hand, in aqueous environment the propensity to form the β-hairpin loops is an intrinsic property of the Pro-Gly sequence. It was observed also in a shorter Ac-D-Pro-Gly-NH-Me dipeptide. The attached alanine residues in the tetrapeptide stabilize the structure only partially. Thus an inclusion of the solvent in the calculations is important for correct description of peptide folding in the aqueous environment. The agreement of the optical spectra with the experiment was determined from overlaps between simulated and measured spectral curves. The comparison of the computed NMR, Raman, and ROA was hampered by experimental noise and limited accuracy of the computations. However, the statistical analysis of the spectroscopic data provided conformer distribution consistent with the computation of the relative energies. The combination of the NMR and Raman techniques with the quantum computations appeared very beneficial for the investigation of Pro-Gly conformational behavior, and can be recommended for future peptide folding studies.

Introduction

The −D-Pro-Gly− fragment often induces a β-hairpin-like turn in a large variety of peptide and protein sequences.[1−4] This knowledge enhances rational synthesis of models mimicking diverse peptide and protein folding patterns, such as well-defined antiparallel β-sheets.[5] The Pro-Gly sequence can also modify biological activity of potential therapeutic drugs as a nucleation agent.[6] Many important features of the structure and dynamics

[†] Academy of Sciences.
[‡] Charles University.

(1) Roy, R. S.; Gopi, H. N.; Raghothama, S.; Karle, I. L.; Balaram, P. *Chem. Eur. J.* **2006**, *12*, 3295−3302.
(2) Mahalakshmi, R.; Raghothama, S.; Balaram, P. *J. Am. Chem. Soc.* **2006**, *128*, 1125−1138.
(3) Syud, F. A.; Stanger, H. E.; Mortell, H. S.; Espinosa, J. F.; Fisk, J. D.; Fry, C. G.; Gellman, S. H. *J. Mol. Biol.* **2003**, *326*, 553−568.
(4) Karle, I.; Gopi, H. N.; Balaram, P. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5160−5164.
(5) Chakraborty, K.; Shivakumar, P.; Raghothama, S.; Varadarajan, R. *Biochem. J.* **2005**, *390*, 573−581.
(6) Krishnakumari, V.; Sharadadevi, A.; Singh, S.; Nagaraj, R. *Biochemistry* **2003**, *42*, 9307−9315.

10.1021/jo702297y CCC: $40.75
Published on Web 01/19/2008

of this sequence in aqueous solutions were studied with the aid of relevant model peptides by means of the NMR and optical spectroscopies previously.[7,8]

Relatively fewer works were dedicated solely to conformational properties of this moiety.[9,10] Multidimensionality of the problem inhibits a simple abstract modeling, as the peptide secondary structure generally depends on non-additive mutual interactions of multiple amino acid residues. However, many experiments indicate that the propensity to form the turn is an intrinsic property of the −D-Pro-Gly− residue or involves a limited ensemble of interactions with other molecular parts.[11,12] In this work, we concentrate on this aspect and adopt minimal synthetic and theoretical peptide models, vibrational and NMR spectra of which can be compared to computations performed at a relatively high level of approximation.

Simpler peptide sequences proved to be convenient models for β-hairpins that can be thoroughly investigated in aqueous solutions.[7,13] We have chosen the D-Ala-L-Pro-Gly-D-Ala tetrapeptide as a basic model that contains both the Pro-Gly sequence and a minimal stabilization of the loop via one hydrogen bond connecting the alanine residues. It is small enough to allow for precise quantum chemical computations of NMR and vibrational properties. Because of a high solubility of the peptide the noise level in the experimental spectra can be decreased significantly. For practical reasons (large stock of D-alanine available for the synthesis) the L-Pro-Gly sequence was used instead of the D-Pro-Gly moiety. Obviously, properties of L-Ala-D-Pro-Gly-L-Ala peptide that would normally correspond to a β-hairpin containing natural L-amino acids can be derived directly from our enantiomeric D-Ala-L-Pro-Gly-D-Ala sequence. The Raman and NMR spectra are identical,[14] whereas the Raman optical activity (ROA) spectra of enantiomers are equal in magnitude and opposite in sign. Therefore, we calculate and report all properties for the L-Ala-D-Pro-Gly-L-Ala enantiomer, except for the ROA spectra. The same is true for the shorter Ac-L-Pro-Gly-NHMe sequence (Figure 1), which enabled us to assess the role of the peptide chain length in formation of the turn.

Peptides exhibit several features that complicate a direct analysis of their liquid-phase NMR and vibrational spectra. Strong interaction of the peptide chain with polar solvents, particularly with water, broadens the experimental signal. Multiple conformers of similar energies are often present in the sample. This makes approximate quantum mechanical predictions even less reliable.[15,16] In the past, vibrational spectra of seemingly simple protopeptide systems, such as the alanine or proline zwitterions, could be assigned only when the molecular flexibility was considered and many conformers of similar

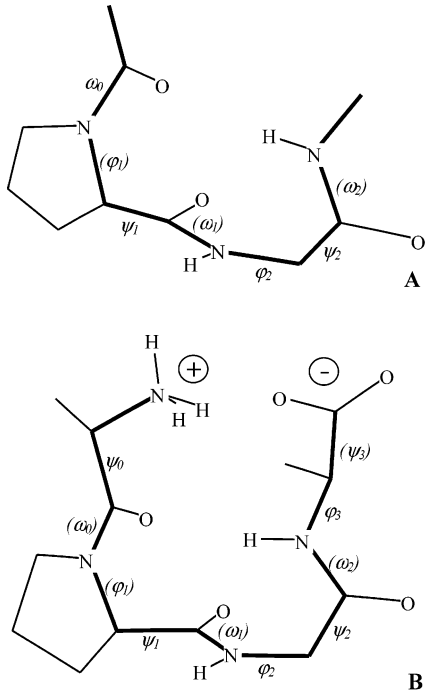


FIGURE 1. Equilibrium structures of the two principal experimental and theoretical peptide models (**A** and **B**) analyzed in this work, and definition of the main torsion angles.

energy averaged according to the Boltzmann statistics.[17,18] Dipeptide NMR and ROA spectra, such as those of *cyclo*-diproline[19] or dialanine,[20,21] were analyzed utilizing a statistical agreement between computed and experimental parameters. In this context, the tetrapeptide zwitterion, nowadays amendable to a relatively accurate theoretical modeling, represents an interesting intermediate between flexible, but still relatively short peptides, and bigger systems where many spectral interpretation techniques developed for single small molecules are not applicable. Because of the band overlap, we analyze the Raman and ROA spectra on the basis of a direct comparison of the calculated and experimental spectral shapes, instead of in terms of individual normal mode contributions. In this approach the curve similarity can be quantified, and thus more objective information can be obtained than by the usual visual comparison of the spectra.

The Raman optical activity technique has been established as an incisive probe of biomolecular structure and behavior.[22−24] It provides an enhanced conformational sensitivity, if compared to unpolarized Raman measurements. Unlike for the infrared

(7) Hilario, J.; Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2003**, *125*, 7562−7574.
(8) Hilario, J.; Kubelka, J.; Syud, F. A.; Gellman, S. H.; Keiderling, T. A. *Biospectroscopy* **2002**, *67*, 233−236.
(9) Colombo, G.; De Mori, G. M. S.; Roccatano, D. *Protein Sci.* **2003**, *12*, 538−550.
(10) Hassan, S. A.; Mehler, E. L. *Int. J. Quantum Chem.* **2001**, *83*, 193−202.
(11) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Protein Sci.* **2002**, *11*, 1492−1505.
(12) Espinosa, J. F.; Gellman, S. H. *Angew. Chem., Int. Ed.* **2000**, *39*, 2330−2333.
(13) Kuznetsov, S. V.; Hilario, J.; Keiderling, T. A.; Ansari, A. *Biochemistry* **2003**, *42*, 4321−4332.
(14) Harris, R. A.; Jameson, C. J. *J. Chem. Phys.* **2006**, *124*, 096101.
(15) Bouř, P.; Keiderling, T. A. *J. Chem. Phys.* **2003**, *119*, 11253−11262.
(16) Bouř, P. *J. Chem. Phys.* **2004**, *121*, 7545−7548.
(17) Kapitán, J.; Baumruk, V.; Kopecký, V., Jr.; Bouř, P. *J. Phys. Chem. A* **2006**, *110*, 4689−4696.
(18) Kapitán, J.; Baumruk, V.; Kopecký, V., Jr.; Pohl, R.; Bouř, P. *J. Am. Chem. Soc.* **2006**, *128*, 13451−13462.
(19) Bouř, P.; Sychrovský, V.; Maloň, P.; Hanzlíková, J.; Baumruk, V.; Pospíšek, J.; Buděšínský, M. *J. Phys. Chem. A* **2002**, *106*, 7321−7327.
(20) Bouř, P.; Buděšínský, M.; Špirko, V.; Kapitán, J.; Šebestík, J.; Sychrovský, V. *J. Am. Chem. Soc.* **2005**, *127*, 17079−17089.
(21) Bouř, P.; Kapitán, J.; Baumruk, V. *J. Phys. Chem. A* **2001**, *105*, 6362−6368.
(22) Barron, L. D.; Zhu, F.; Hecht, L. *Vib. Spectrosc.* **2006**, *42*, 15−24.
(23) Zhu, F.; Isaacs, N. W.; Hecht, L.; Barron, L. D. *J. Am. Chem. Soc.* **2005**, *127*, 6142−6143.
(24) Haesler, J.; Schindelholz, I.; Riguet, E.; Bochet, C. G.; Hug, W. *Nature* **2007**, *446*, 526−529.

TABLE 1. Characteristic Torsion Angles (deg) and Relative Conformer Energies *E* (kcal/mol) Calculated for the Ten Lowest Energy Conformers of the Peptide A

(a) Vacuum Computations

conformer (level)	ω_0	ψ_1	φ_2	ψ_2	*E*
1 (BPW91/6-31G**)	173	−73	−79	65	0
1 (B3LYP/6-31G**)	173	−73	−80	66	0
1 (BPW91/6-311++G**)	173	−72	−80	64	0
2 (BPW91/6-31G**)	172	−69	83	−63	0.1
2 (B3LYP/6-31G**)	172	−70	84	−64	0.1
2 (BPW91/6-311++G**)	173	−70	83	−63	0.0
BPW91/6-31G** only:					
3	177	−94	−112	4	1.2
4	173	−75	122	−14	1.4
5	176	−71	−96	−163	3.3
6	171	19	105	−18	3.3
7	173	−75	141	−156	3.7
8	−13	13	−81	66	4.6
9	−14	16	81	−63	4.6
10	171	14	−83	66	4.9

(b) The CPCM Water Model

conformer	$\omega_0{}^b$	$\psi_1{}^b$	$\varphi_2{}^b$	$\psi_2{}^b$	E^a	E^b
1′	179	−130	−86	4	0.0	0.0
2′	3	−154	−115	142	1.9	2.2
3′	180	−136	98	14	1.9	1.8
4′	179	−148	108	151	2.0	1.4
5′	176	22	79	10	2.2	2.4
6′	2	−154	−112	143	2.4	2.3
7′	3	−146	108	146	2.5	2.7
8′	2	−152	92	−144	2.5	1.9
9′	3	−141	90	16	2.8	2.7
10′	172	19	76	36	2.8	4.7

[a] BPW91/6-31G**. [b] B3LYP/6-311++G**.

and vibrational dichroism spectroscopy,[25] a broad range of vibrational frequencies in the aqueous environment is accessible. Application of the ROA spectroscopy to longer peptides is somewhat hindered by the difficult interpretation, requiring lengthy quantum chemical computations and a complex modeling of the sample.[26,27] On the other hand, such a modeling brings enhanced information about the molecules of interest. The scattering techniques also nicely complement the NMR spectroscopy, being inherently much faster and sampling directly the conformational composition.[19] For example, two labile conformers providing an average NMR signal can often be distinguished as individual species by ROA. For the tetrapeptide, due to the presence of the proline in position 2, some NH(2)-NOE NMR contacts are not observable at all. Thus the Raman and ROA spectra add unique information that cannot be obtained from NMR.

Results and Discussion

Conformer Geometries. Systematic conformational analysis provided ten lowest energy conformers of the model dipeptide **A**, as summarized in Table 1. Three approximation levels were applied, BPW91/6-31G**, B3LYP/6-31G**, and BPW91/6-311++G**; only for the first default method are the geometric parameters for all the conformers listed. It is apparent that the

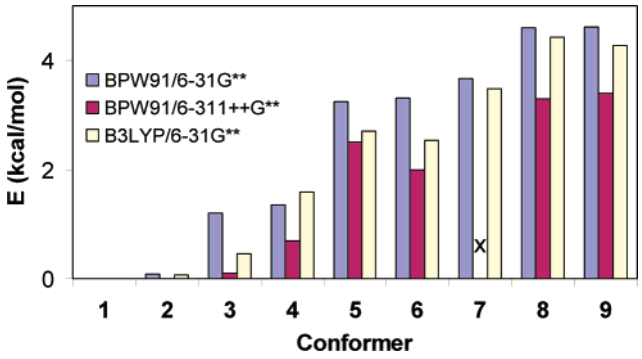


FIGURE 2. Computed relative conformer energies of the model peptide **A**. Conformer number 7 was not stable at the BPW91/6-31G** approximation level.

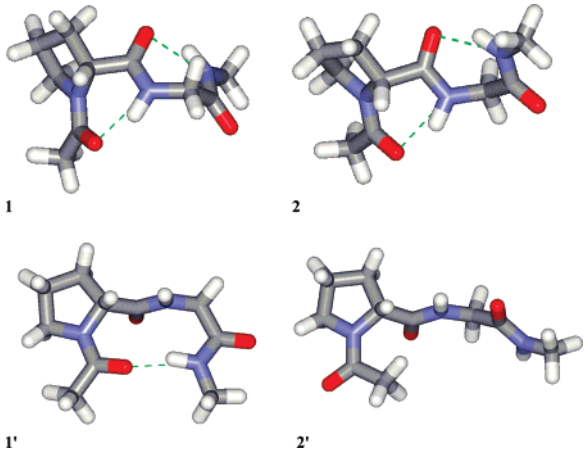


FIGURE 3. Two most stable conformers of peptide **A** found at the BPW91/6-31G** (vacuum, **1**, **2**) and BPW91/CPCM/6-31G** (water, **1′**, **2′**) approximations. The strongest hydrogen bonds are indicated by the dashed green lines.

vacuum computations (part a of Table 1, Figure 2) provide similar energy ordering. However, conformers 1−4 differ by less than 1.4 kcal/mol. Because an expected error of the calculations[28] is comparable, an exact prediction of conformer populations would be difficult. Equilibrium geometries predicted at different levels practically do not differ, neither does the usage of the bigger 6-311++G** basis set lead to significant changes. Neither the two lowest energies (structures 1 and 2, displayed in Figure 3) nor the following conformers resemble standard peptide turns (typical angles for type I′ turn would be $\psi_1 \approx 30°$, $\varphi_2 \approx 90°$, and $\psi_2 \approx 0°$; for type II′ $\psi_1 \approx -120°$, $\varphi_2 \approx -80°$ and $\psi_2 \approx 0°$).[29] Thus the hairpin turn is not preferred by the Ac-L-Pro-Gly-NH−Me peptide in vacuum. Similarly, the proline cis-conformers 8 and 9 ($\omega_0 \approx 0°$) are predicted with energies too high (~4.6 kcal/mol) to be significantly populated under room temperature.

The turn, however, becomes the most preferred conformation when the CPCM aqueous environment is added (conformer 1′ in Table 1b and Figure 3). It would also become the prevalent conformation in water, as the next two conformers (2′ and 3′) are ~2 kcal/mol higher in energy. Most of the ten lowest energy conformations (1′−10′) cannot be associated with those found

(25) Keiderling, T. A.; Kubelka, J.; Hilario, J. In *Vibratonal spectrroscopy of polymers and biological systems*; Braiman, M., Gregoriou, V., Eds.; Marcel Dekker Publ.: New York, 2005.
(26) Ruud, K.; Helgaker, T.; Bouř, P. *J. Phys. Chem. A* **2002**, *106*, 7448−7455.
(27) Reiher, M.; Liegeois, V.; Ruud, K. *J. Phys. Chem. A* **2005**, *109*, 7567−7574.
(28) van Mourik, T.; Karamertzanis, P. G.; Price, S. L. *J. Phys. Chem. A* **2006**, *110*, 8−12.
(29) Stanger, H. E.; Syud, F. A.; Espinosa, J. F.; Giriat, I.; Muir, T.; Gellman, S. H. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12015−12020.

TABLE 2. Characteristic Torsion Angles (deg) and Relative Energies *E* (kcal/mol) Calculated for the Ten Lowest Energy Conformers of the Tetrapeptide B

(a) *trans*-Pro structures ($\omega_0 \approx 180°$)

conformer	ψ_0[a]	ψ_1[a]	φ_2[a]	ψ_2[a]	φ_3[a]	E[a]	E[b]	E[c]	E[d]	η[e]
1	117	−123	−98	15	−111	0.0	0.0	0.0	0.0	19
2	153	−144	168	−162	−160	3.6	5.0	4.3	0.4	10
3	152	−144	111	150	−155	3.7	7.9	5.3	0.1	16
4	152	−154	−128	20	−155	4.0	6.9	5.7	1.6	1
5	152	−143	136	−145	−160	4.0	5.3	4.5	0.1	16
6	153	−146	−134	152	−154	4.2	*f*	4.7	1.3	2
7	153	−142	110	−11	−159	4.4	5.7	4.9	0.8	5
8	152	−147	−148	−168	−157	4.6	5.1	4.2	0.3	11
9	153	−144	102	152	−157	4.7	6.3	5.2	0.6	7
10	153	−141	81	−155	−159	4.8	6.2	5.4	0.5	8

(b) *cis*-Pro structures ($\omega_0 \approx 0°$)

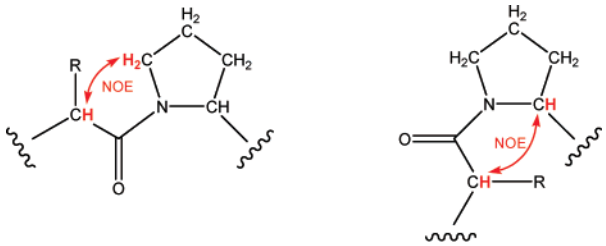
conformer	ψ_0[a]	ψ_1[a]	φ_2[a]	ψ_2[a]	φ_3[a]	E[a]	E[d]	η[e]
1′	143	177	176	−145	143	4.3	2.3	0
2′	143	−162	72	−156	143	4.5	2.0	1
3′	142	154	−149	−151	142	4.8	1.9	1
4′	143	−149	−116	−141	143	5.3	2.3	0
5′	145	0	102	−144	145	5.3	2.5	0
6′	144	149	101	−148	144	5.4	2.9	0
7′	143	−127	116	−143	143	5.5	7.9	0
8′	143	163	125	−144	143	5.6	2.2	0
9′	141	167	−90	−152	141	5.6	2.1	1
10′	145	−150	65	−148	145	5.9	1.9	1

[a] BPW91/6-31G*/CPCM. [b] B3LYP/6-31G*/CPCM. [c] BPW91/6-31G*/PCM. [d] BPW91/6-31++G*/CPCM. [e] Relative populations (%) at 298 K. [f] Not converged.

in vacuum, although conformer 1′ remotely resembles the vacuum structure number 3, and 5′ is somewhat similar to 6. The behavior of this dipeptide nicely demonstrates a primitive peptide denaturation within the solvent → vacuum environmental change. Indeed, most hairpin peptides change structure when moved from water to a less polar environment.[30] However, the second most convenient conformer of **A** in water (2′) does not form an intramolecular hydrogen bond and its predominantly extended structure is stabilized by the solvent polarization only. This is in accord with the known experimental behavior of β-hairpins, as the turn-form is often accompanied by an unordered structure and the melting transition temperature is found mostly within 0−100 °C, indicating relative conformer free energies of the order of 1 kcal/mol.[30] The aqueous environment apparently also enhances the transition of the tertiary proline amide group into the cis-conformation ($\omega_0 \approx 0°$, conformers 2′ and 6′−9′). In water, the cis-form becomes more convenient than in vacuum as the relative energy calculated in vacuum as 4.6 kcal/mol decreases to ∼2 kcal/mol in CPCM. It is shown below that this is in agreement with the experimental results indicating a significant presence of the cis-form.

The zwitterionic peptide **B** does not exist in vacuum and a solvent must be added for meaningful simulations. As can be seen from the summary of the conformational scan in Table 2, the replacement of the methyl termini in **A** by the charged $-CH(CH_3)\text{-}NH_3^+$ and $-CH(CH_3)\text{-}CO_2^-$ groups did not change the turn geometry in the lowest energy conformer. Particularly, the angles ψ_1, φ_2, and ψ_2 in Table 1b for **A** (−130°, −86°, 4°) mimic very well those in **B** (−123°, −98°, 15°). Similarly, conformers 2′, 3′, and 4′ of **A** correspond to forms 6, 7, and 3 of **B**, respectively. The conformer 1 of **B** (see Figure 4 with the 10 lowest energy trans-conformers) corresponds to the peptide turn type II′.[29] Although the D-Pro-Gly residue was reported to have a propensity to form the turn type I′ as well, we did not observe it for our models. Only the conformer 5′ of **A** remotely resembles the usual turn type I′ geometry.[8,31−33]

The *trans*-/*cis*-Pro Ratio. It is well-known that linear peptides containing the tertiary amide group X-Pro can exist in solution as mixtures of the trans- and cis-forms.[34,35] Populations of both isomers depend on the peptide structure as well as on the solvent environment. Because of the relatively slow trans ↔ cis transition, the two isomers can be distinguished on the basis of the characteristic NOE contacts observed in 2D-NOESY or 2D-ROESY NMR spectra:



***trans*- X-Pro** ***cis*- X-Pro**

The cis- and trans-conformer ratios determined from the NOE contacts (see the Supporting Information (SI) for details) for the peptides **A**−**D** are consistent with the calculated relative conformer energies for **A** and **B**. Experimentally, the less polar compounds (**A**, **C**, and **D**) exhibit an increased population of the cis-form, in agreement with the relatively small relative energies of many *cis*-**A** conformers predicted in Table 1. However, the higher trans-/cis-conformer ratios observed in $CDCl_3$ than DMSO do not correlate with the relative electric permittivity of these two solvents. In water, a minor population of the cis-conformer was observed, about 22% for **A** and 5% for **B**. The computations with the CPCM water model (Tables 1 and 2) predict a 2 kcal/mol energy gap between the cis- and trans-forms for both peptides. At room temperature (298 K) this corresponds to $\exp(-\Delta E/kT) = 4\%$ ratio of the cis-form. Thus the estimation based solely on the Boltzmann ratio (without taking into account relative numbers of all trans and cis conformers) is in good agreement with the observation for the **B** peptide, while it somewhat underestimates the cis-population of **A**.

In principle, the two basic I′ and II′ turn types can also be distinguished from the NOE patterns (see the contact list in the SI). Due to the presence of the proline in position 2, however, some NH(2)NOE contacts are not observable in our peptides. In **A** and **B**, three NOE contacts—$d_{NH(3),NH(4)}$, $d_{\alpha H(2),NH(3)}$, and $d_{\alpha H(3),NH(4)}$—were observed, while $d_{\alpha H(2),NH(4)}$ (with the longest interproton distance) was not detected. Under these circumstances the turn type cannot be determined from NOE unambiguously. Compounds **C** and **D** were originally proposed in

(30) Setnička, V.; Huang, R.; Thomas, C. L.; Etienne, M. A.; Kubelka, J.; Hammer, R. P.; Keiderling, T. A. *J. Am. Chem. Soc.* **2005**, *127*, 4992−4993.

(31) Chen, P. Y.; Lin, C. K.; Lee, C. T.; Jan, H.; Chan, S. I. *Protein Sci.* **2001**, *10*, 1794−1800.

(32) Geierstanger, B.; Jamin, M.; Volkman, B. F.; Baldwin, R. L. *Biochemistry* **1998**, *37*, 4254−4265.

(33) Syud, F. A.; Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **2001**, *123*, 8667−8677.

(34) Dukor, R. K.; Keiderling, T. A. *Biospectroscopy* **1996**, *2*, 83−100.

(35) Vitagliano, L.; Berisio, R.; Mastrangelo, A.; Mazzarella, L.; Zagari, A. *Protein Sci.* **2006**, *10*, 2627−2632.

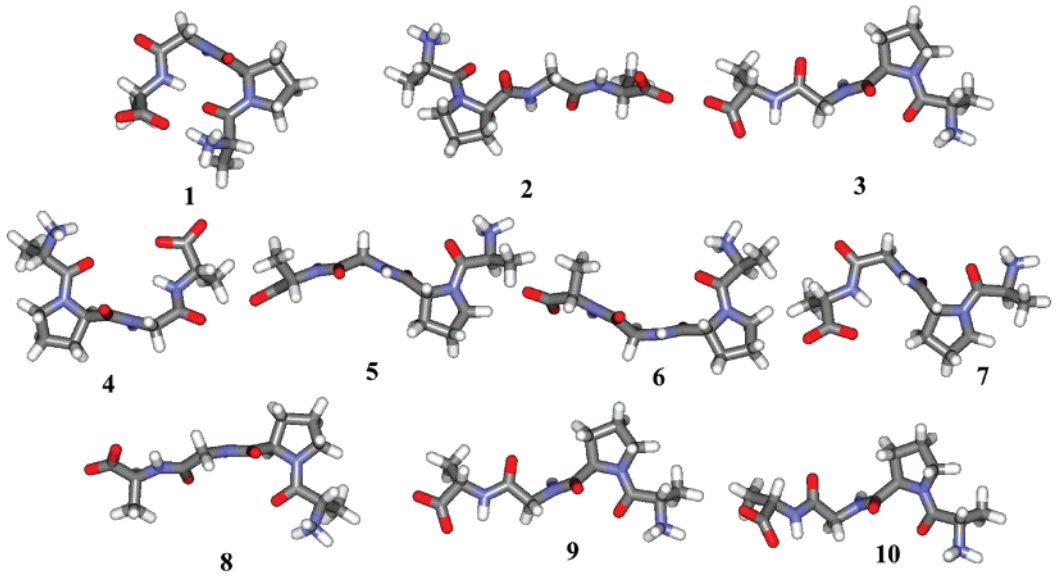


FIGURE 4. Ten most stable (BPW91/CPCM/6-31G**) conformers of the tetrapeptide **B**.

order to stabilize the turn, but the NOE did not reveal any particular turn preference, similarly as for **A** and **B**. Moreover, the conformer analysis above indicates significant flexibility of the peptide chain. Thus the rigid NOE model may not be appropriate for the conformational characterization of the compounds and the analysis of the NMR shifts and spin−spin coupling constants reported below might be more adequate.

NMR Shifts and Couplings. Generally, the calculated spin−spin coupling constant and chemical shifts for peptide **A** (Tables 4s−7s in the SI) agree best with the experiment for the lowest energy conformers. Also the conformation 9′ provides coupling constants very close to the experiment (Table 7s), but this agreement is most probably accidental given the limited number of NMR parameters that could be measured and compared to the calculations. Then we can conclude that the β-hairpin type loop is formed as the most-populated conformer of peptide **A**, although it cannot be further stabilized by main- and side-chain interactions. Additionally, the experimental data suggest that in DMSO the conformation behavior is very similar to that observed in H_2O.

Further on, we will concentrate on peptide **B**, where a complete set of ROA and NMR data is available. Only data of the trans-isomer are analyzed because of its relation to the hairpin peptide and protein structures. First, for each conformer, we compared the chemical shifts calculated at the B3LYP/6-311++G** level with the CPCM dielectric water model to the experimental. The computed values (Table 8s, SI) follow excellently the experimental data for various hydrogen and heavy atom types. However, as was observed already for similar systems, the accuracy of the computations is not sufficient for a distinction of individual conformers.[18−20,36] The accuracy cannot be increased by an improvement of the solvent model (e.g., by an inclusion of explicit water molecules) or by a basis set extension.[20] Most probably, the error stems from the density functional approximation[37] and will not significantly change until better theoretical methods are developed.

On the other hand, the indirect spin−spin coupling constants (Table 9s, SI) do allow for a conformer distinction. For example, the (NH,α1H) interaction at the glycine residue for conformer 2 (J = 0.9 Hz) is clearly too low if compared to experiment (J = 5.1 Hz), which indicates a low probability of finding this conformer in the sample. Some constants are rather insensitive to the conformations. Overall, however, we can see that the Boltzmann-averaged values fit the experiment much better (see the root-mean-square deviation Δ = 1.6 Hz in Table 9s) than any of the individual conformers 1−10 (with Δ = 1.9−2.9 Hz). Thus we conclude that also tetrapeptide **B** is present in several conformations, relative ratios of which approximately correspond to the relative conformer energies calculated at the B3LYP/6-311++G**/CPCM level.

Vibrational Spectra. The previous conclusions can further be supported by the analysis of the Raman and Raman optical activity spectra. Similarly as for the NMR parameters, the ab initio computations are indispensable for vibrational spectra interpretation. The computed harmonic frequencies (given in Table 3 for the first conformer of **B**) allow assigning most of the observed experimental Raman bands. The Raman intensities (Figure 5) are too complex and the vibrational transitions too dense to enable a band-to-band assignment. In spite of these difficulties, the basic shape of the Raman spectrum is predicted correctly by the BPW91/6-31G*, BPW91/6-31+G**, and Becke3LYP/6-31+G** methods. The scattering of the polar groups (C=O stretching including the amide I mode and the N−H bending at the NH_3^+ group) is identifiable as a broad experimental signal within 1550−1700 cm^{-1}. The bandwidth broadening can be explained due to the strong interaction with the solvent, particularly via hydrogen bonds,[15] but also due to the flexibility and rotational motion of the NH_3^+ and COO^- groups.[17] The region of 1200−1500 cm^{-1} comprises a plethora of frequencies, some of them significantly dependent on the peptide conformation (see the frequency dispersions for the ten conformers given in the parentheses in Table 3). On the other hand, most of the experimental peaks within 400−1200 cm^{-1} can be assigned, at least to certain groups of vibrational normal modes. At the lowest frequency region below 400 cm^{-1} the experimental Raman signal becomes very broad and unstruc-

(36) Bouř, P.; Navrátilová, H.; Setnička, V.; Urbanová, M.; Volka, K. *J. Org. Chem.* **2002**, *67*, 161−168.
(37) Kümmel, S.; Kronik, L.; Perdew, J. P. *Phys. Rev. Lett.* **2004**, *93*, 213002.

TABLE 3. Calculated (B3LYP/6-31+G/CPCM) Vibrational Frequencies (cm^{-1}) for the First Conformer of Peptide B and Assignment to the Experiment[a]**

mode	ω_{cal}[b]	ω_{exp}		mode	ω_{cal}[b]	ω_{exp}	
23	1701 (5)	1661	ν(C═O), A I	68	1058 (4)		ν(C−C_α)
24	1680 (2)	1655	ν(C═O) , A I	69	1054 (2)	1048	ν(C−C_α)
25	1662 (7)		ν(C═O) , A I	70	1040 (9)		ν(C−C_α)
26	1660 (15)	1621	$\delta(NH_3^+)$, scissor	71	1013 (2)		ν(C−C_α)
27	1625 (12)		$\delta(NH_3^+)$, scissor	72	998 (7)		ν(C−C_α)
28	1597 (5)	1552	ν(C═O), asymm CO_2^-	73	985 (5)		ν(C−C_α)
29	1582 (7)		ν(C−N), A II	74	972 (4)	971	ν(C−C_α)
30	1552 (6)		ν(C−N), A II	75	954 (8)	923	ν(N−C_α)
31	1545 (10)		$\delta(NH_3^+)$, umbrella	76	938 (2)		Pro, breathing I
32	1510 (2)	1483	δ(CH), scissor	77	917 (2)		Pro, def.
33	1505 (4)		δ(CH), scissor	78	912 (1)		deloc.
34	1500 (2)		δ(CH), scissor	79	893 (7)		deloc.
35	1493 (1)		δ(CH), scissor	80	882 (2)		Pro, breathing II
36	1489 (1)		δ(CH), scissor	81	865 (4)	877	δ(O═C═O)
37	1488 (2)		δ(CH), scissor	82	856 (2)	864	ν(N−C_α)
38	1485 (1)	1457	δ(CH), scissor	83	792 (6)	784	deloc, oop C═O Gly
39	1480 (1)	1449	A II, Pro	84	775 (3)		deloc, oop CO_2
40	1442 (15)	1413	δ(CH), scissor, Gly	85	761 (4)		NH_3 wagg, oop CO Ala
41	1419 (2)	1372	ν(C═O), symm CO_2^-	86	752 (5)		oop amide, deloc
42	1416 (2)	1389	$\delta(NH_3^+)$, umbrella	87	734 (5)	742	oop amide, deloc
43	1391 (3)	1372	$\delta(NH_3^+)$, umbrella	88	729 (14)		oop amide, deloc
44	1383 (4)		δ(CH), δ(CC)	89	687 (4)		oop amide, deloc
45	1375 (3)		$\delta(^{\alpha}CH)$, CH_2 wag	90	683 (16)	702	deloc
46	1369 (4)		$\delta(^{\alpha}CH)$, CH_2 wag	91	660 (15)		deloc
47	1361 (4)		$\delta(^{\alpha}CH)$, CH_2 wag	92	599 (3)		Pro def.
48	1359 (2)	1331	$\delta(^{\alpha}CH)$, CH_2 wag	93	569 (9)		oop NH
49	1358 (3)		$\delta(^{\alpha}CH)$, CH_2 wag	94	549 (11)		oop NH
50	1348 (1)		$\delta(^{\alpha}CH)$, CH_2 twist	95	542 (6)		oop NH
51	1344 (3)		$\delta(^{\alpha}CH)$, CH_2 twist	96	517 (9)	544	deloc
52	1310 (3)		$\delta(^{\alpha}CH)$, CH_2 twist	97	496 (22)	520	δ(C−C−N)
53	1304 (4)		$\delta(^{\alpha}CH)$, CH_2 twist	98	452 (12)	454	δ(C−C−N)
54	1298 (9)	1289	ν(C−N), A III	99	388 (5)	370	δ(C−C−N)
55	1283 (10)		ν(C−N), A III	100	382 (7)		NH_3 rot.
56	1272 (7)	1271	ν(C−N), A III, Pro	101	368 (8)		oop amide N
57	1249 (6)	1256	ν(C_0-C_α)	102	335 (4)		oop amide N
58	1224 (2)		ν(C_0-C_α)	103	317 (3)		deloc.
59	1216 (2)		deloc, Pro	104	310 (8)		deloc.
60	1206 (2)	1193	deloc	105	295 (7)		deloc.
61	1170 (2)	1167	deloc, Pro	106	285 (6)		deloc.
62	1163 (4)		ν(N−C_α)	107	268 (5)		deloc.
63	1123 (4)		deloc	108	262 (9)		Pro def.
64	1118 (3)	1116	ν(N−C_α)	109	245 (6)		deloc.
65	1106 (3)	1104	ν(N−C_α), ν(C−C_α)	110	235 (11)		CH_3 rot.
66	1104 (2)		ν(N−C_α)	111	234 (13)		CH_3 rot.
67	1093 (2)	1091	ν(N−C_α)	112	220 (10)		CH_3 group wag.
68	1058 (4)		ν(C−C_α)	113	204 (13)		deloc.

[a] ν, stretching; δ, bending; oop, out of plane. [b] Average rms dispersions in the ten conformer set are given in parentheses.

tured due to the large-amplitude motions of peptide parts[17] as well as to the contribution of the water present in the first hydration sphere.[18]

Both the BPW91 and B3LYP functionals provide reasonable vibrational frequencies. The former gives more realistic values of the higher frequency vibrations, particularly the C═O stretching modes, and as a pure GGA method independent of the Hartree−Fock exchange,[38] it is more economic with respect to computer time. However, the B3LYP frequencies are more accurate on average (cf. the deviations δ in Figure 5) and the intensity profiles are more faithful. As can also be seen in Figure 5, less extended changes in the spectra are caused by the improvement of the basis (6-31G*→6-31+G**) than by the functional change, although the frequency and intensity agreement for the bigger basis is better.

The ROA intensities are significantly more sensitive to the computational parameters than the unpolarized Raman scatter-

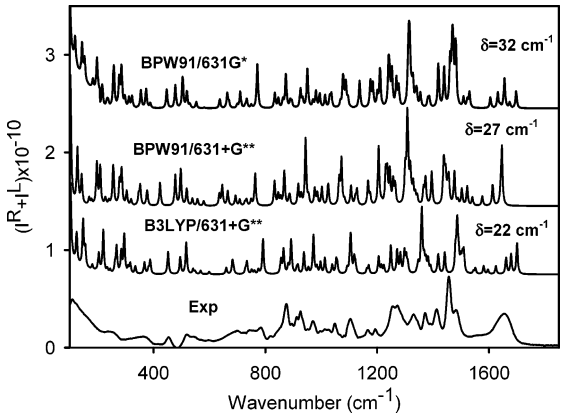


FIGURE 5. Comparison of the Raman spectra of the tetrapeptide **B** simulated at three approximation levels with the experiment. Average rms frequency deviations according to the assignments in Table 3 are indicated.

(38) Becke, A. *Phys. Rev. A* **1988**, *38*, 3098−3100.

TABLE 4. Statistical Comparison of the Raman and ROA Simulated Spectra of Ten B Conformers to the Experiment: Average Deviations (δ) and the Decomposition Coefficients (b_i)[a]

	Raman		ROA	
conformation	δ	b_i	δ	b_i
1	0.41	0.23	1.40	0.06
2	0.47	0.04	1.36	0.02
3	0.44	0.10	1.34	0.04
4	0.46	0.05	1.23	0.12
5	0.45	0.07	1.34	0.11
6	0.45	0.12	1.22	0.16
7	0.43	0.14	1.24	0.08
8	0.45	0.08	1.27	0.02
9	0.44	0.12	1.30	0.08
10	0.53	0.06	1.05	0.31

[a] The B3LYP/6-31+G**/CPCM force field and Raman intensities were used for the simulated spectra with the HF/6-31+G**/CPCM ROA tensors.

ing. Therefore, for the first conformer of **B**, we performed several tests summarized in the figure on page S12 of the SI. With the BPW91 functional, for example, ROA sign patterns obtained with smaller (6-31G*) and bigger (6-31+G**) bases are mostly conserved, although relative intensities differ notably. Because the ROA intensity computation is the most expensive part of the simulation, we also explored the possibility of combining a higher quality force field with a lower level intensity estimation.[26] The best simulation could be obtained from the Becke3LYP/6-31+G** frequency and the BPW91/6-31+G** intensity calculation via the Cartesian components-based transfer of the molecular property tensors.[39] Also other models provided reasonable spectral patterns. These include the polarization approximation, which allows the estimation of the ROA shape very quickly from the electric dipole polarizabilities only,[40−42] the Hartree−Fock, and lower level BPW91 computations.

As the one conformer spectrum cannot fully explain the experiment, normalized experimental Raman and ROA curves were decomposed into the theoretical spectra of the ten lowest energy conformers. The average deviations δ and decomposition coefficients b_i are listed in Table 4. The lowest energy conformer suits best the Raman spectrum ($b_1 = 0.23$), but other conformations must be included to explain all spectral features. In fact, spectral decomposition even more consistent with the Boltzmann ratios in Table 2 is obtained on a limited frequency interval (Table 11s in the SI). Due to the limited precision we do not consider the decomposition of the ROA spectra reliable at this point, although it provides results in qualitative agreement with the Raman pattern.

In spite of these problems, the multiconformational model clearly explains the observed spectrum better than a one-conformer model, similarly as for the NMR spin−spin coupling constants. Most of the features of the experimental ROA spectrum (trace d in Figure 6) are reproduced with correct signs by the decomposition fit (c), although relative intensities of several bands still do not agree with the experiment. The Boltzmann-averaged (b) and fitted (c) spectra are quite similar,

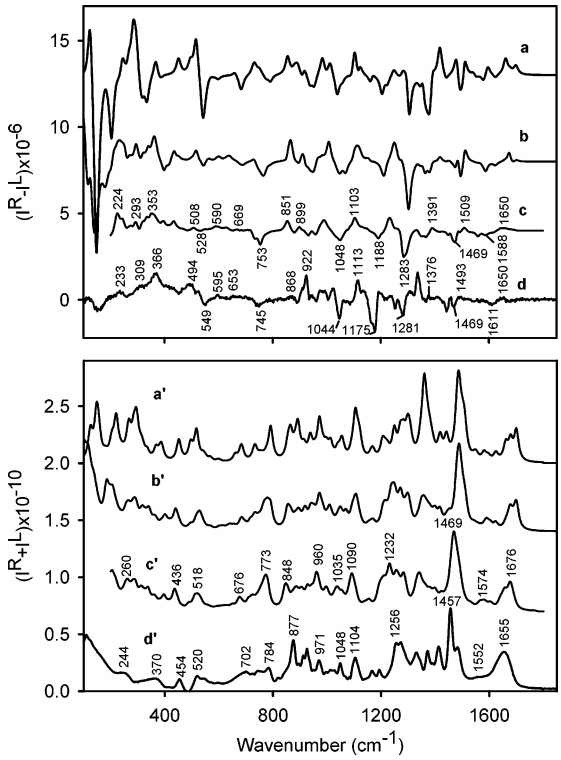


FIGURE 6. ROA (a−d) Raman (a′−d′) and spectra of the tetrapeptide (**B,** L-Ala-D-Pro-Gly-L-Ala): (a, a′) first conformer only, (b, b′) Boltzmann average (Table 2), (c, c′) intensity fit (Table 4), and (d, d′) the experiment. A bandwidth of 10 cm^{-1} was used for the simulated spectra.

which indicates that the spectral decomposition is not unique due to the peptide flexibility and multiconformational equilibrium. The averaging reproduces the relative intensity decrease above 1500 cm^{-1}, as well as the signal broadening of the lower frequency (ca. <800 cm^{-1}) bands. For the corresponding Raman spectra (a′−d′) where the averaging cannot significantly modify integrated intensities (the ROA bands can theoretically average to zero), the agreement between the simulated and experimental spectra, at least for the human eye, is even better. We can thus see that even for a system of multiple conformations the Raman and ROA spectra provide relatively accurate information about the conformer composition and geometry. It can be significantly increased if better computational methods and lower noise levels in the experimental spectra are available in the future.

Conclusions

The results indicate that the Pro-Gly residue does have an internal propensity to form the hairpin turn structures, but only in a polar environment. The flexibility of this residue is also significantly modified by other peptide parts. Since the sequence acts relatively independently, the turn most probably does not induce a particular conformation of the rest of the peptide chain. Specific loop structures in design sequences then must be stabilized by other means. The *cis*-/*trans*-Pro conformer ratios determined by the NMR data are in agreement with the computed relative conformer energies. For the Ac-L-Pro-Gly-NH-Me and NH_3^+-D-Ala-L-Pro-Gly-D-Ala-CO_2^- peptides in solutions the hairpin forms were found as parts of greater conformer ensembles only. The statistical analysis of the spectra was found consistent with the ab initio computations and

(39) Bouř, P.; Sopková, J.; Bednárová, L.; Maloň, P.; Keiderling, T. A. *J. Comput. Chem.* **1997**, *18*, 646−659.
(40) Bouř, P.; Baumruk, V.; Hanzliková, J. *Collect. Czech. Chem. Commun.* **1997**, *62*, 1384−1395.
(41) Kapitán, J.; Baumruk, V.; Kopecký, V., Jr.; Bouř, P. *J. Am. Chem. Soc.* **2006**, *128*, 2438−2443.
(42) Barron, L. D. *Molecular Light Scattering and Optical Activity*; Cambridge University Press: Cambridge, UK, 2004.

confirmed the ability of the NMR, Raman, and ROA techniques to decipher multiconformational equilibria of peptides in solutions. Boltzmann averaging provided better spectral data than any individual conformer. The interpreting of the spectra is limited by the accuracy of present computational methods and experiments.

Experimental Section

Synthesis. The blocked Ac-L-Pro-Gly-NH-Me (**A**), Ac-D-Ala-L-Pro-Gly-D-Ala-NH-Me (**C**), Bz-D-Ala-L-Pro-Gly-D-Ala-NH-Bn (**D**), and Z-D-Ala-L-Pro-Gly-D-Ala-OBn (**E**, with the Z- and Bn-protecting groups) peptides were synthesized by the Fmoc/tBu strategy on Merrifields resin with the indole-3-carboxaldehyde linker.[43] The resin was converted to aldimine with either 2.5 M Bzl-NH_2 in THF or 0.4 M CH_3NH_2 in a THF/EtOH mixture (1:1). The volume (mL) of the agent solution was ten times bigger than the weight (g) of the indole-3-carboxaldehyde resin. The *N*-aminoacid termini were blocked either by the Ac_2O or Bz_2O groups. Synthesized peptides were cleaved from the resin by the TFA/anisole/DCM (2:1:20) mixture and purified by a column chromatography on silica gel with gradient 10:1 to 5:1 $CHCl_3$/MeOH mobile phase. The target unblocked H-D-Ala-L-Pro-Gly-D-Ala-OH (**B**) peptide was prepared on the Wang resin using the same Fmoc/tBu strategy. It was purified by the reverse-phase (RP) HPLC on a 25 × 1 cm^2 column, 10 μm in gradient 0−50% ACN with 0.05% TFA. The zwitterion was obtained after ionex chromatography on Dowex-50 and an elution with 2% ammonia solution, evaporation, and lyophilization.

NMR. The NMR spectra of peptides **A**−**E** were measured with two frequency setups (^{1}H at 500.13 MHz and ^{13}C at 125.77 MHz frequency, and/or ^{1}H at 600.13 MHz and ^{13}C at 150.90 MHz frequency) in d_6-DMSO (compounds **A**, **C**, and **D**), H_2O/D_2O 9:1 (compounds **A**, **B**, and **C**) and $CDCl_3$ (compound **E**). Proton and carbon-13 signals were structurally assigned by means of the series of 1D- (^{1}H, ^{13}C-APT) and 2D-NMR spectra (^{1}H,^{1}H-COSY, ^{1}H,^{1}H-TOCSY, ^{1}H,^{1}H-ROESY, ^{1}H,^{13}C-HMQC, ^{1}H,^{13}C-HMBC). Temperature coefficients of the amide NH protons were determined from the set of 1D-proton NMR spectra in the temperature interval of 20−40 °C. An attempt to obtain other temperature-dependent NMR parameters was made, but did not provide useful results.

Vibrational Spectra. The backscattered Raman and incident circular polarization (ICP) ROA spectra of tetrapeptide **B** (the **A**, **C**, **D**, and **E** derivatives were not soluble enough for the measurement) were recorded on the spectrometer located at Charles University described elsewhere.[21,44] The peptide was dissolved in deionized water at room temperature with final concentrations of 0.24 mol/L. The solution was filtered into a 5 × 5 mm^2 quartz cell. Residual fluorescence from impurities, which could otherwise cause large and unstable background signal, were quenched by leaving the sample in the laser beam for 5 h before the spectra acquisition. The laser wavelength was 532.15 nm, laser power at the sample 500 mW, spectral resolution 6.5 cm^{-1}, and total acquisition times 36 h. The solvent signal was subtracted from the Raman spectra and fine baseline ROA artifacts were corrected according to ref 45.

Computations. The conformational spaces of the Ac-D-Pro-Gly-NH−Me (**A**) and zwitterionic forms of the D-Ala-L-Pro-Gly-D-Ala (**B**) peptides were explored by a systematic analysis: For **A**, 54 = 2 × 3 × 3 × 3 initial structures were generated with all combinations of the torsion angles defined in Figure 1, ω_0(0°, 180°), ψ_1(60, 180, −60), ψ_2(60, 180, −60°), and φ_2(60, 180, −60°). For each structure, an energy minimization was performed combining the BPW91[38] and Becke3LYP[46] functionals, 6-31G*, 6-31G**, and 6-311++G** basis sets, and PCM and CPCM solvent (H_2O) models as specified below. GAUSSIAN software[47] was used for the calculation. For **B**, 486 = 2 × 3 × 3 × 3 × 3 initial structures were generated combining the angles of ω_0(0°, 180°), ψ_0(60, 180, −60), ψ_1(60, 180, −60), φ_2(60, 180, −60°), ψ_2(60, 180, −60°), and φ_3(60, 180, −60°) (Figure 1), and energy minimizations were performed at the BPW91/6-31G**/CPCM level. Geometries of selected lowest energy conformers were reoptimized with the bigger 6-311++G** basis set. It is important to note that the solvent model had to be used for zwitterion **B** as it is not stable in vacuum. This brought some practical problems. Many conformer optimizations did not converge because of unrealistic starting geometries or problems with numerical stability. Such cases were treated separately via change of the GAUSSIAN default parameters or a visual inspection of the geometries, so that realistic conformers did not escape from the automatic conformer selection. For selected conformers chemical shifts and spin−spin coupling constants were also calculated by GAUSSIAN, using the default analytical method[48] with the gauge invariant atomic orbitals (GIAO).[49] The harmonic vibrational frequencies and Raman and ROA intensities[42] for fully optimized geometries were computed also with GIAOs in the static approximation. Control computation with the actual laser excitation frequency did not provide significant change of spectral intensities. Therefore, static approximation requiring less computer time was retained as a default.

Vibrational Spectra Analysis. As a supplement to the usual visual comparison, the agreement of a simulated Raman or ROA spectrum $S_i(\omega)$ with an experimental curve $S_e(\omega)$ was determined on the basis of the sum of the absolute deviations:

$$\delta = \sum_j |\tilde{S}_i(\omega_j) - k\tilde{S}_e(a\omega_j + \Delta)| + a_k(k-1)^2 + a_a(a-1)^2 + a_\Delta\Delta^2 \quad (1)$$

where $\tilde{S}(\omega)$ is the normalized spectral curve ($\int|\tilde{S}(\omega)|d\omega = 1$), and the adjustment parameters (k, a, Δ) were determined by a nonlinear iterative minimization of δ.[50] Although for perfect agreement $a = k = 1$ and $\Delta = 0$, the adjusting of frequencies (a, Δ) was introduced in order to allow for a systematic frequency shift typical for ab initio computations,[51] while the k-parameter allowed for intensity scale inaccuracies and noise. The barriers of the quadratic penalty functions were chosen as $a_k = 100$, $a_a = 10$, and $a_\Delta = 0.002$. A bandwidth of 10 cm^{-1} was used for the simulated spectra for the automatic comparisons with the experiment.

Raw calculated spectra were used as well as those obtained with a scaled force field $\mathbf{F}' = \mathbf{S}^t\omega^2\mathbf{S}$, where the Cartesian-normal mode transformation matrix[52] **S** was obtained with the raw force field, and the diagonal matrix ω^2 contained squares of experimentally assigned or interpolated frequencies (listed in Table 3). The results

(43) Šebestík, J.; Matějka, P.; Hlaváček, J.; Stibor, I. *Tetrahedron Lett.* **2004**, *45*, 1203−1205.
(44) Kapitán, J.; Baumruk, V.; Gut, V.; Hlaváček, J.; Dlouhá, H.; Urbanová, M.; Wunsch, E.; Maloň, P. *Collect. Czech. Chem. Commun.* **2005**, *70*, 403−409.
(45) Kapitán, J.; Baumruk, V.; Hulačová, H.; Maloň, P. *Vib. Spectrosc.* **2006**, *42*, 88−92.
(46) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(47) Frisch, M. J.; et al. Gaussian, Inc.: Wallingford CT, 2004.
(48) Sychrovský, V.; Gräfenstein, J.; Cremer, D. *J. Chem. Phys.* **2000**, *113*, 3530−3547.
(49) Ruud, K.; Helgaker, T.; Bak, K. L.; Jorgensen, P.; Jensen, H. J. A. *J. Chem. Phys.* **1993**, *99*, 3847−3859.
(50) Press, W. H.; Teukolsky, S. A.; Vetteling, W. T.; Flannery, B. P. *Numerical Recipes in Fortran*, 2nd ed.; Cambridge University Press: New York, 1992.
(51) Pulay, P. In *Modern electronic structure theory*; Yarkony, D. R., Ed.; World Scientific: Singapore, 1995; Vol. 2, pp 1191−1240.
(52) Papoušek, D.; Aliev, M. R. *Molecular Vibrational/Rotational Spectra*; Academia: Prague, Czech Republic, 1982.

with the scaled and unscaled force fields, however, were very similar (see the SI, Table 11s).

The minimization of the δ-parameter enables one to find the most similar spectrum, but does not account for a conformer mixture. Therefore, the experimental Raman and ROA spectra were decomposed into the theoretical curves of *m*-conformations:

$$\tilde{S}_e = \sum_{i=1...m} b_i s_i \quad (2)$$

where $\sum_{i=1...m} b_i = 1$ and $s_i = k\tilde{S}_i(a\omega_j + \Delta)$. The ratios a_i were obtained by minimization of the square sum, $\Delta = \sum_j(\tilde{S}_e(\omega_j) - \sum_{i=1..m} b_i s_i(\omega_j))^2 + \alpha \sum_{i=1..m}(b_i - 1/m)^2$. The penalty function (last term, with $\alpha = 0.001$ and 0.0002, for Raman and ROA, respectively) was added to prevent large negative decomposition coefficients b_i. The coefficients were then obtained by a constrained minimization of Δ utilizing the standard Lagrange multiplier method solvable by matrix inversion.

Acknowledgment. The work was supported by the Grant Agency of the Czech Republic (grants 203/06/0420, 202/07/0732) and the Grant Agency of the Academy of Sciences (A400550702).

Supporting Information Available: Experimental and computational details. This material is available free of charge via the Internet at http://pubs.acs.org.

JO702297Y