

## Cyclopentapeptides as Flexible Conformational Templates

Gregory V. Nikiforovich,<sup>\*,†</sup> Katalin E. Kövér,<sup>‡</sup> Wei-Jun Zhang,<sup>§</sup> and Garland R. Marshall<sup>†,§</sup>

Contribution from the Center for Molecular Design and Department of Molecular Biology and Pharmacology, Washington University, St. Louis, Missouri 63130, and L. Kossuth University, H-4010 Debrecen, Hungary

Received May 24, 1999. Revised Manuscript Received February 1, 2000

**Abstract:** Studies of 3D models for cyclopentapeptides (CPP's) employing *only* NMR spectroscopy encounter a serious problem. Because of conformer averaging, 3D structure(s) derived directly from NMR data may not correspond to the energy minimum (minima) with low relative conformational energy. At the same time, independent energy calculations can determine all low-energy conformers for the CPP backbone. The two approaches are compared in this study by results obtained for *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>). Contrary to the conclusion (predominance of the  $\beta$ II' $\gamma$  type conformer) of earlier NMR studies, independent energy calculations found a different family of low-energy 3D structures that are consistent both with the NMR data in DMSO and with the known X-ray data on CPP's. The preferable Ala<sup>4</sup> conformations were found in the  $\alpha$ <sub>R</sub>/ $\alpha$ <sub>L</sub> regions suggesting studies of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>) which was synthesized. Further NMR studies confirmed the results of the independent energy calculations. The independent energy calculations have been applied also to *cyclo*(Arg<sup>1</sup>-Gly<sup>2</sup>-Asp<sup>3</sup>-D-Phe<sup>4</sup>-Val<sup>5</sup>) and *cyclo*(Arg<sup>1</sup>-Gly<sup>2</sup>-Asp<sup>3</sup>-Phe<sup>4</sup>-D-Val<sup>5</sup>). Both peptides are almost equally potent inhibitors of binding of  $\alpha$ <sub>Ib</sub> $\beta$ <sub>3</sub> integrins to fibrinogen and of  $\alpha$ <sub>v</sub> $\beta$ <sub>3</sub> integrins to vitronectin. If both of them possess a NMR-predicted conformer of the  $\beta$ II' $\gamma$  type, however, the conformations of the active sequence, Arg<sup>1</sup>-Gly<sup>2</sup>-Asp<sup>3</sup>, should be dissimilar in these two peptides. This discrepancy is eliminated in the 3D pharmacophore model proposed by independent energy calculations. The model is also in good agreement with the model by other authors that was confirmed by X-ray studies.

## Introduction

Rational design of pharmaceuticals derived from naturally occurring peptides has been enhanced recently by two major breakthroughs. First, peptide and peptidomimetic libraries have been instrumental in producing hundreds of thousands of different compounds for biological screening. Second, cloning and expressing transmembrane peptide receptors has created mutant and chimeric receptors, thus allowing an opportunity to study peptide–receptor interaction “from the receptor side”. However, so far both techniques have generated more questions than potential pharmaceuticals. The enormous amount of screening data coming from biological testing of libraries needs to be rationalized (see, e.g., refs 1 and 2). The same is true for the data obtained on peptide binding to mutant and chimeric receptors; it is enough to mention the observed differences in binding sites/modes for agonists and antagonists (for review see, e.g., ref 3 edited by Schwartz et al.).

In both cases, one of the main obstacles for drug design is the absence of reliable information on the 3D structures of peptide within the ligand–receptor complex. Therefore, it would be extremely useful to develop a variety of “conformational templates”, i.e., model ligands, which should satisfy at least three

requirements: (i) they should possess only one 3D structure (or just a few well-determined 3D structures) and (ii) they should be readily accessible synthetically; and (iii) they should be able to position the regular peptide side chains which are believed to transfer most information during peptide–receptor interaction.

Excellent candidates for such conformational templates are cyclopentapeptides (CPP's). First, they are expected to be relatively conformationally rigid. Second, different types of CPP's can reproduce different types of conformational elements of peptide backbone, as various  $\beta$ -turns,  $\gamma$ -turns, and even  $\alpha$ -helical-like structures (see, e.g., ref 4). Third, CPP's can be easily modified to include a large variety of side chains. And, fourth, they are synthetically accessible. A recent review<sup>5</sup> points out that CPP's containing D- or nonchiral amino acids in addition to L-amino acids are readily prepared. All-L-amino acid CPP's also can be prepared by solid-phase synthesis using reagents derived from 7-hydroxyazabenzotriazole with quite reasonable yields (see refs 6 and 7).

Extensive experimental studies of the 3D structures of CPP's have been performed in the last two decades both by X-ray spectroscopy and by NMR spectroscopy. The X-ray studies have

\* To whom correspondence should be addressed.

<sup>†</sup> Center for Molecular Design.

<sup>‡</sup> L. Kossuth University.

<sup>§</sup> Department of Molecular Biology and Pharmacology.

(1) Lam, K. S. *Anti-Cancer Drug Des.* **1997**, *12*, 145–167.

(2) Pinilla, C.; Appel, J.; Blondelle, S.; Dooley, C.; Dorner, B.; Eichler, J.; Ostresh, J.; Houghten, R. *Biopolymers* **1995**, *37*, 221–240.

(3) *Structure and Function of 7TM Receptors*; Alfred Benzon Symposium 39; Schwartz, T. W., Hjorth, S. A., Kastrup, J. S., Eds.; Munksgaard: Copenhagen, 1996; p 428.

(4) Weisshoff, H.; Wieprecht, T.; Henklein, P.; Antz, C.; Muge, C. *Biochem. Biophys. Res. Commun.* **1995**, *213*, 506–512.

(5) Schmidt, U.; Langner, J. *J. Pept. Res.* **1997**, *49*, 67–73.

(6) Ehrlich, A.; Rothmund, S.; Brudel, M.; Beyermann, M.; Carpino, L. A.; Bienert, M. In *Peptides. Chemistry, Structure and Biology*; Proceedings of the Thirteenth American Peptide Symposium; Hodges, R. S., Smith, J. A., Eds.; ESCOM: Leiden, 1995; pp 95–96.

(7) Ehrlich, A.; Brudel, M.; Beyermann, M.; Winter, R.; Carpino, L. A.; Bienert, M. In *Peptides 1994*; Proceedings of the Twenty-Third European Peptide Symposium; Maia, H. L. S., Ed.; ESCOM: Leiden, 1995; pp 167 and 168.

been performed mostly by the Karle group (e.g., refs 8–12), by the Italian groups (e.g., refs 13–15), and by the Gierasch group.<sup>16,17</sup> The X-ray structures are now available for several CPP's, including those containing unusual amino acids.<sup>14,18</sup> As to NMR studies, two groups of researchers should be mentioned as, perhaps, the more productive ones. They are the Gierasch group, which accumulated a large amount of information concerning CPP's with one or two proline residues (see ref 19 for a review), and the Kessler group, which studied mostly CPP's containing D-amino acid residues (see, e.g., a paper on the RGD-related CPP's and references therein<sup>20</sup>).

In fact, employment of CPP's as receptor probes with known 3D structures was initiated by the Kessler group in the early nineties (e.g., refs 21–23). On the basis of extensive NMR measurements, they proposed a "conformational template" of the (*abcde*) type (the lower case denotes D-amino acids) that possessed a single conformation characterized by a  $\beta$ II'-turn centered at the *ab* fragment, and a  $\gamma$ -turn at the *D* residue (see one of the first papers<sup>24</sup>). Moving the position of the D-amino acid residue along the sequence, it would be possible to obtain new conformational templates of the same type, and to use the data of their biological testing for elucidation of a peptide pharmacophore. The Kessler group applied the above approach to RGD peptides (see the following papers and references therein<sup>25,26</sup>), and have designed several types of corresponding peptidomimetics.<sup>27,28</sup>

However, this approach suffers from a serious drawback. The problem is that most short peptides, even cyclic ones, exist in solution as a mixture of different interconverting conformers. As a consequence, there are unavoidable difficulties in employing *only* experimental techniques for determining 3D structures of CPP's. X-ray studies produce knowledge of a very few 3D structures stabilized during the process of crystallization by intermolecular interactions in the crystal lattice; these 3D structures do not necessarily correspond to the "receptor-bound"

conformer(s). On the other hand, each value of the conformational parameter measured by NMR spectroscopy (like the vicinal coupling constants, NOE's, etc.) represents an average over an unknown number of conformers with significant statistical weights. An attempt to fit all measured parameters into a single 3D structure imposing the corresponding restrictions can be justified only in the very unlikely case that one conformer exists in solution with a highly predominant statistical weight. Many researchers tackle this problem of conformational averaging either by relaxing the NMR-derived limitations imposed on the single conformer (see, e.g., ref 23) or by generating a random family of conformers that satisfy the NMR limitations as a whole (see, e.g., refs 29 and 30). In both cases, the suggested 3D structures are refined by some procedures involving energy calculations, such as molecular dynamics simulations. As a result, the molecule is forced into the nearest energetic minimum (minima) which is (are) not necessarily of low relative conformational energy. A vivid example is provided by a recent study by Zanotti et al. showing that the same cyclopentapeptide [cyclo-(Phe-Phe-Aib-Leu-Pro)] possesses different conformations in the crystal state and in various apolar solutions, none of which conformations are of the  $\beta$ II' $\gamma$  type.<sup>15</sup>

On the other hand, all low-energy conformers for a peptide backbone of a short peptide can be elucidated by independent energy calculations, and then may be evaluated as possible solution conformations. At the same time, the calculated sets of low-energy conformers can always be validated by NMR and/or X-ray spectroscopy. Moreover, the combined use of the independent NMR measurements and energy calculations allows an estimation of the statistical weights for the actual conformers observed in solution. This approach was developed by us earlier,<sup>31</sup> and has been successfully applied in the cases of spin-labeled angiotensin,<sup>32,33</sup> enkephalin,<sup>34</sup> dermenkephalin,<sup>35</sup> and DPDPE.<sup>36</sup>

Accordingly, the main goal of this study is to outline the advantages of applying independent energy calculations to CPP's as possible receptor probes in comparison to other approaches based on NMR measurements only.<sup>21–23</sup> Comparison of results obtained on 3D structures for a simple *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) peptide by the two approaches clearly shows inconsistency of the  $\beta$ II' $\gamma$  model. Our results provide the more realistic view on flexibility of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>); this view is substantiated also by synthesis, energy calculations, and NMR studies of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>). Finally, we analyze inevitable discrepancies in elucidation of peptide pharmacophores using NMR measurements only as proposed for the RGD peptides;<sup>25</sup> these discrepancies do not exist when independent energy calculations are applied.

(8) Gierasch, L. M.; Karle, I.; Rockwell, A. L.; Yenai, K. *J. Am. Chem. Soc.* **1985**, *107*, 3321–3327.

(9) Karle, I. L. *J. Am. Chem. Soc.* **1978**, *100*, 1286–1289.

(10) Karle, I. L. *J. Am. Chem. Soc.* **1979**, *101*, 181–184.

(11) Karle, I. L. In *The Peptides, Analysis, Synthesis, Biology*; Gross, E., Meienhofer, J., Eds.; Academic Press: New York, 1981; Vol. 4, pp 1–54.

(12) Karle, I. L. *Int. J. Pept. Protein Res.* **1986**, *28*, 420–427.

(13) Toniolo, C. *CRC Crit. Rev. Biochem.* **1980**, *9*, 1–44.

(14) Lombardi, A.; Saviano, M.; Nastri, F.; Maglio, O.; Mazzeo, M.; Pedone, C.; Isernia, C. V. P. *Biopolymers* **1996**, *38*, 683–691.

(15) Zanotti, G.; Saviano, M.; Saviano, G.; Tancredi, T.; Rossi, F.; Pedone, C.; Benedetti, E. *J. Pept. Res.* **1998**, *51*, 460–466.

(16) Stroup, A. N.; Rheingold, A. L.; Rockwell, A. L.; Gierasch, L. M. *J. Am. Chem. Soc.* **1987**, *109*, 7146–7150.

(17) Stroup, A. N.; Rockwell, A. L.; Rheingold, A. L.; Gierasch, L. M. *J. Am. Chem. Soc.* **1988**, *110*, 5157–5161.

(18) Anwer, M. K.; Sherman, D. B.; Spatola, A. F. *Int. J. Pept. Protein Res.* **1990**, *36*, 392–399.

(19) Stradley, S. J.; Rizo, J.; Bruch, M. D.; Stroup, A. N.; Gierasch, L. M. *Biopolymers* **1990**, *29*, 263–287.

(20) Koppitz, M.; Huenges, M.; Gratiias, R.; Kessler, H.; Goodman, S. L.; Jonczyk, A. *Helv. Chim. Acta* **1997**, *80*, 1280–1300.

(21) Kessler, H.; Gurrath, M.; Muller, G.; Aumailley, M.; Timpl, R. In *Peptides 1992*; Proceedings of the Twenty-Second European Peptide Symposium; Schneider, C. H., Eberle, A. N., Eds.; ESCOM: Leiden, 1993; pp 75 and 76.

(22) Aumailley, M.; Gurrath, M.; Muller, G.; Calvete, J.; Timpl, R.; Kessler, H. *FEBS Lett.* **1991**, *291*, 50–54.

(23) Muller, G.; Gurrath, M.; Kessler, H.; Timpl, R. *Angew. Chem., Int. Ed. Engl.* **1992**, *31*, 326–328.

(24) Gurrath, M.; Muller, G.; Kessler, H.; Aumailley, M.; Timpl, R. *Eur. J. Biochem.* **1992**, *210*, 911–921.

(25) Haubner, R.; Gratiias, R.; Diefenbach, B.; Goodman, S. L.; Jonczyk, A.; Kessler, H. *J. Am. Chem. Soc.* **1996**, *118*, 7461–7472.

(26) Wermuth, J.; Goodman, S. L.; Jonczyk, A.; Kessler, H. *J. Am. Chem. Soc.* **1997**, *119*, 1328–1335.

(27) Haubner, R.; Schmitt, W.; Holzemann, G.; Goodman, S. L.; Jonczyk, A.; Kessler, H. *J. Am. Chem. Soc.* **1996**, *118*, 7881–7891.

(28) Haubner, R.; Finsinger, D.; Kessler, H. *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 1374–1389.

(29) Mierke, D. F.; Kurz, M.; Kessler, H. *J. Am. Chem. Soc.* **1994**, *116*, 1042–1049.

(30) Cuniassie, P.; Raynal, I.; Yioutakis, A.; Dive, V. *J. Am. Chem. Soc.* **1997**, *119*, 5239–5248.

(31) Nikiforovich, G. V.; Vesterman, B. G.; Betins, J. *Biophys. Chem.* **1988**, *31*, 101–106.

(32) Nikiforovich, G. V.; Vesterman, B.; Betins, J.; Podins, L. J. *Biomolecular Structure & Dynamics* **1987**, *4*, 1119–1135.

(33) Vesterman, B. G.; Betins, J. R.; Nikiforovich, G. V. *Biophysics* **1988**, *33*, 999–1006.

(34) Vesterman, B.; Saulitis, J.; Betins, J.; Liepins, E.; Nikiforovich, G. V. *Biochim. Biophys. Acta* **1989**, *998*, 204–209.

(35) Nikiforovich, G. V.; Prakash, O.; Gehrig, C. A.; Hruby, V. J. *J. Am. Chem. Soc.* **1993**, *115*, 3399–3406.

(36) Nikiforovich, G. V.; Prakash, O.; Gehrig, C.; Hruby, V. J. *Int. J. Peptide Protein Res.* **1993**, *41*, 347–361.

**Table 1.** Dihedral Angle Values (rounded up to degrees) for the X-ray Structures (refs 8–10, 12, 16, 17, first row) and for the Closest Calculated Low-Energy Conformers (second row)<sup>a</sup>

peptide	dihedral angles												$\Delta E$ , kcal/mol	rms, Å
	$\phi_1$	$\psi_1$	$\omega_{12}$	$\phi_2$	$\psi_2$	$\phi_3$	$\psi_3$	$\phi_4$	$\psi_4$	$\omega_{45}$	$\phi_5$	$\psi_5$		
c(fPGaP)	70	-131	-172	-83	-9	-152	97	73	-133	173	-50	136	1.5	0.75
	83	-127	174	-75	-46	-115	107	82	-103	153	-75	132		
c(GPSaP)	58	-128	-175	-75	-20	-167	114	86	-123	160	-66	165	0.0	0.65
	84	-126	175	-75	-46	-116	108	81	-103	155	-75	130		
c(APGfP)	69	86	14	-89	154	74	34	124	-68	175	-80	2	0.5	1.11
	65	84	29	-75	141	72	-135	-54	-60	175	-75	-8		
c(GPfAP)	109	94	9	-74	170	64	-143	-68	-45	177	-74	-30	6.3	0.47
	68	81	29	-75	143	77	-138	-61	-51	175	-75	-17		
c(GPGaP)	83	-134	174	-52	126	74	12	134	-69	178	-86	70	1.6	0.94
	87	-129	155	-75	123	69	94	80	-113	166	-75	118		
c(GPfGA)	104	-176	-179	-65	112	105	0	126	115	-173	58	65	1.5	1.12
	139	-117	161	-75	73	132	-115	-62	-70	180	-119	71		
c(GPfGV)	-91	-149	-179	-57	125	67	17	-165	-132	167	-70	-39	2.1	1.05
	-98	-81	169	-75	90	87	-90	-48	-92	180	-89	-61		

<sup>a</sup> Table contains also the relative energy values for calculated conformers, and the rms differences between these conformers and the corresponding X-ray structures. The rms differences involve spatial positions of all heavy atoms of the backbone, as well as of all C <sup>$\beta$</sup>  atoms. The one-letter symbols for D-amino acid residues are shown in the lower case.

## Results

**Energy Calculations for CPP's Employing the ECEPP Force Field: Validation.** The energy calculations in this study were performed using the ECEPP force field. Since the choice of the force field is one of the most disputable problems in any energy calculation (see, e.g., ref 37 and subsequent discussion), our first priority was to validate the use of the ECEPP force field for conformational studies of isolated CPP's.

The most detailed experimental information on the 3D structures of CPP's has been obtained from X-ray studies. Obviously, we should not expect that our energy calculations would reproduce the observed X-ray structures for CPP's as the one with lowest energy. It is clearly improbable for at least two reasons: (i) any force field possesses its own inherent inaccuracies in energy estimations and (ii) the isolated 3D molecular structures in the crystal cells may be distorted by strong intermolecular interactions, i.e., packing interactions. However, it is legitimate to ask another question, namely, will the 3D structures found by X-ray crystallography for different cyclopentapeptides be geometrically similar to at least one of the low-energy conformers found by energy calculations employing the ECEPP force field?

We have performed energy calculations for seven cyclopentapeptides (see Table 1) with known X-ray structures starting from two model sequences, *cyclo*(Gly-Pro-Gly-Gly-Pro) and *cyclo*(Gly-Pro-Gly-Gly-Ala), and exploring all combinations of local energetic minima of all amino acid residues in both sequences, including the trans/cis conformers for Pro residues. We have used the ECEPP/2 force field<sup>38,39</sup> with rigid valence geometry and with the value of the macroscopic dielectric constant  $\epsilon = 45$ . The ring closure has been ensured by the overlapping of the C <sup>$\alpha$</sup> -C' valence bond for the "zeroth" residue in the CPP with that for the fifth one, the "zeroth" residue including the dummy atoms for this particular bond. This overlapping was achieved with the use of parabolic closing potentials between the C <sup>$\alpha$</sup>  and C' atoms of the fifth residue and the corresponding dummy atoms of the "zeroth" residue, employing the  $U_0$  value for the potentials of 1000 kcal/mol.

(37) Roterman, I. K.; Lambert, M. H.; Gibson, K. D.; Scheraga, H. A. *J. Biomol. Struct. Dyn.* **1989**, *7*, 421–453.

(38) Dunfield, L. G.; Burgess, A. W.; Scheraga, H. A. *J. Phys. Chem.* **1978**, *82*, 2609–2616.

(39) Nemethy, G.; Pottle, M. S.; Scheraga, H. A. *J. Phys. Chem.* **1983**, *87*, 7, 1883–1887.

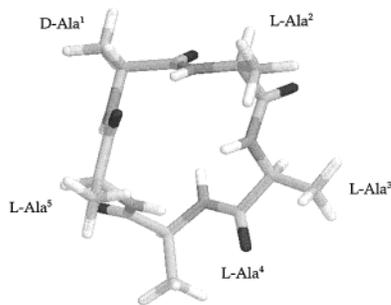
(The same technique has been used in energy calculations for all other CPP's in this study.) Low-energy backbone conformers (those with the relative energies  $\Delta E < 10$  kcal/mol) of the model sequences have been subjected to further energy minimization with side chains incorporated at the proper positions. Geometrical similarity between the X-ray structures and the low-energy conformers was assessed by calculating the rms values between them involving all heavy backbone atoms, as well as all C <sup>$\beta$</sup>  atoms.

Results in Table 1 show that low-energy conformers obtained by the ECEPP force field can be considered geometrically similar to the corresponding X-ray structures practically in all cases. Notably, two cases with the  $\omega_{12}$  angle in the cis-conformation [*c*(APGfP) and *c*(GPfAP)] have been reproduced. In four cases, the rms values have been less than 1 Å, which is a very good similarity. Some peptide bond planes in the found low-energy conformers are rotated almost by 180° compared to the corresponding X-ray structures, namely the bonds connecting residues 3 and 4 for *c*(APGfP), *c*(GPfGV), and *c*(GPfGV), as well as residues 4 and 5 for *c*(GPfGA). However, for those peptides, the rms values also are less than 1.2 Å. In all these cases, strong hydrogen bonds between adjacent molecules have been observed within the crystal lattice (see refs 8, 12, and 16) that have not been taken into account in the energy calculations.

Table 1 contains the calculated conformations which are most similar to the corresponding X-ray structures; in all cases but one, *c*(GpfAP), they possess relative energy values less than 5 kcal/mol. (For *c*(GpfAP), too, there is a conformer possessing a  $\Delta E$  value of 3.7 kcal/mol, and the rms value of 0.99 Å.) Therefore, it seems logical to retain the energy threshold of ca. 1 kcal/mol/residue defining the "low-energy" conformations obtained by the ECEPP calculations performed for isolated molecules (see also ref 40), i.e., to 5 kcal/mol in the case of CPP's.

***cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) and  $\beta$ II' $\gamma$  Model.** The  $\beta$ II' $\gamma$  model has been proposed as a single 3D structure for *cyclo*(DLLLL) peptides in solution in earlier papers by the Kessler group.<sup>21–23</sup> According to this model, residues 1 and 2 form a  $\beta$ II'-like turn with  $\phi_1, \psi_1; \phi_2, \psi_2$  values of ca. 60°, -120°; -80°, 0°, and residue 4 adopts a  $\gamma$ -turn conformation that corresponds to the  $\phi_4, \psi_4$  values of ca. 70°, -70° (see Figure 1).

(40) Nikiforovich, G. V. *Int. J. Pept. Protein Res.* **1994**, *44*, 513–531.



**Figure 1.** The  $\beta\text{II}'\gamma$  model for *cyclo*(D-Ala<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) (residue numbering is clockwise starting from the upper left corner) based on dihedral angle values for *c*(RGDFV).<sup>55</sup>

This model seems to completely ignore the fact that it is highly unusual to find a residue of the L-configuration in a conformation with positive  $\phi$  and negative  $\psi$  values. None of the known X-ray structures of cyclopentapeptides possess this feature. Moreover, in all known X-ray structures, only the Pro residue in the corresponding position possesses an actual “inverted  $\gamma$ -turn” values for  $\phi, \psi$  (ca.  $-70^\circ, 70^\circ$ ) (see Table 7 in ref 16 and Tables 1 and 2 in ref 19). All known “ $\gamma$ -turn” values belong to D-Ala residues (see Table 1 and references therein).

The  $\beta\text{II}'\gamma$  model has been constructed by applying all experimental NMR restraints to a single possible 3D structure of CPP.<sup>23</sup> Obviously, this assumption is restrictive, and, in most cases, it is rather far from reality. Later, the same group has modified the initial approach trying to introduce possibilities for multiple 3D structures of CPP in solution. As a simple example, they have studied 3D structures in DMSO of the *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) peptide.<sup>29</sup> The authors employed the following approach: (a) 20 interatomic distances derived from proton–proton NOE's and 4 values of the  $J(\text{HC}^\alpha\text{-NH})$  coupling constants have been established by NMR spectroscopy (Tables 1 and 2 in ref 29); (b) 50 structures have been randomly generated and subjected to optimizing using distance and angle driven dynamics (DADD) with the “force field” ensuring the proper valence geometry; (c) 38 structures selected for nonviolating experimental data were copied 10 times to produce an ensemble, and the average values of the calculated interproton distances and the coupling constants were fit as close as possible to the experimental ones (see Table 1 in 29); and (d) 5 selected families of structures were subjected to energy minimization (the GROMOS force field), and then to molecular dynamics simulations in DMSO. Energies of all five families were found approximately the same.

As the final result, five possible 3D structures have been proposed for *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>).<sup>29</sup> All of them contain the  $\beta\text{II}'$  turn encompassing the D-Pro<sup>1</sup>-Ala<sup>2</sup> fragment, and differ in the conformations of the Ala<sup>4</sup> residue. The  $\phi, \psi$  values of the Ala<sup>4</sup> for these five structures are as follows: ( $90^\circ, -60^\circ$ ); ( $0^\circ, -60^\circ$ ); ( $-120^\circ, -60^\circ$ ); ( $30^\circ, 120^\circ$ ); ( $-170^\circ, 120^\circ$ ). In other words, the sterically impossible  $\gamma$ -turn conformation still survived the selection procedure based on some refinement of NMR data by energy calculations.

We have decided to study this simple, but characteristic example with our approach. Independent energy calculations for *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) included 3895 peptide conformers geometrically allowed to close the pentapeptide rings that were then subjected to energy minimization. Five of them possessed relative energies  $\leq 5$  kcal/mol, the criterion chosen for selection of “low-energy” 3D structures. The Ala<sup>4</sup> residue in the five conformers possesses  $\phi, \psi$  values as follows: ( $-90^\circ, -16^\circ$ ); ( $-139^\circ, 19^\circ$ ); ( $-110^\circ, 6^\circ$ ); ( $65^\circ, 15^\circ$ ); ( $70^\circ, 141^\circ$ ).

The five 3D structures are depicted in Figure 2 and are described also in Table 2. None of our 3D structures possess the pronounced  $\beta\text{II}'$  turn in the D-Pro<sup>1</sup>-Ala<sup>2</sup> region, but all of them are geometrically similar to the discussed  $\beta\text{II}'\gamma$  type.

***cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>): Comparison with NMR Data.** Low-energy 3D structures of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) obtained by independent energy calculations do not contradict the NMR data obtained by the Kessler group. Table 3 contains the interatomic distances derived from NMR data<sup>29</sup> and calculated in this study using an approach developed earlier.<sup>31</sup>

A cornerstone of this approach is an assumption that experimentally measured and calculated parameters are in good agreement when their mean values are statistically indistinguishable. In other words, for each structural parameter,  $A$ , the following condition between the experimental value,  $\langle A^{\text{exp}} \rangle$ , and the weighted sum of calculated values,  $\langle A^{\text{calc}} \rangle$ , should be satisfied:

$$\frac{|\sum_{i=1}^N w_i \langle A^{\text{calc}} \rangle_{ik} - \langle A^{\text{exp}} \rangle_k|}{(\sum_{i=1}^N (w_i D^{\text{calc}}_{ik})^2 + (D^{\text{exp}}_k)^2)^{1/2}} < \mathbf{t}_k \quad (1)$$

Here  $i$  and  $k$  are indexes related to the number of low-energy conformers and to the number of measured parameters, respectively, whereas  $\mathbf{t}$  is the Student's coefficient at the chosen confidence level,  $w_i$  are statistical weights of low-energy conformers, and  $D$  are standard deviations of the mean values for calculated and experimentally measured parameters. Thus, for  $N$  conformers and  $M$  measured parameters, one can randomly generate the sets  $\{w_i\}$  of  $N$  statistical weight values and determine whether each set  $\{w_i\}$  will satisfy to  $M$  inequalities of the above type with the obvious conditions of  $w_i > 0$  and  $\sum_{i=1}^N w_i = 1$ . If such sets are found, it means that agreement between the calculated and experimentally measured parameters is achieved. Notably, in this approach, one obtains not a constant value for the statistical weights of conformers, but their distributions, each with its mean value  $\langle w_i \rangle$  and its upper  $w_i^{\text{up}}$  and lower  $w_i^{\text{low}}$  levels.

In the particular case of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>), we have used 20 values for interproton distances and 4 values for  $J(\text{HC}^\alpha\text{NH})$  coupling constants as measured by the Kessler group (Table 1 in ref 29) as  $\langle A^{\text{exp}} \rangle$ 's. The values of  $D^{\text{exp}}$ 's for the interproton distances have been assigned according to the upper and lower values provided in ref 29 (ca.  $\pm 10\%$ ), and  $\pm 1.0$  Hz for all the coupling constants. For each of the five low-energy conformers, the same 20 interatomic distances and 4 values of coupling constants were calculated and used as  $\langle A^{\text{calc}} \rangle$ 's. It was assumed also that the  $D^{\text{calc}}_{ik}$  values were  $\pm 0.5$  Å for all interatomic distances. For the coupling constants, they have been chosen as limits of the  $J(\text{HC}^\alpha\text{NH})$  variations due to variations of the corresponding  $\phi$  angles by  $20^\circ$  estimated according to the most reliable  $J(\phi)$  dependencies.<sup>41–44</sup> Out of ca. 1000000 random trials, 10000 different  $\{w_i\}$  points of five statistical weights were chosen randomly, all satisfying the above conditions with  $\mathbf{t} = 1.645$  (the confidence interval of 90%).

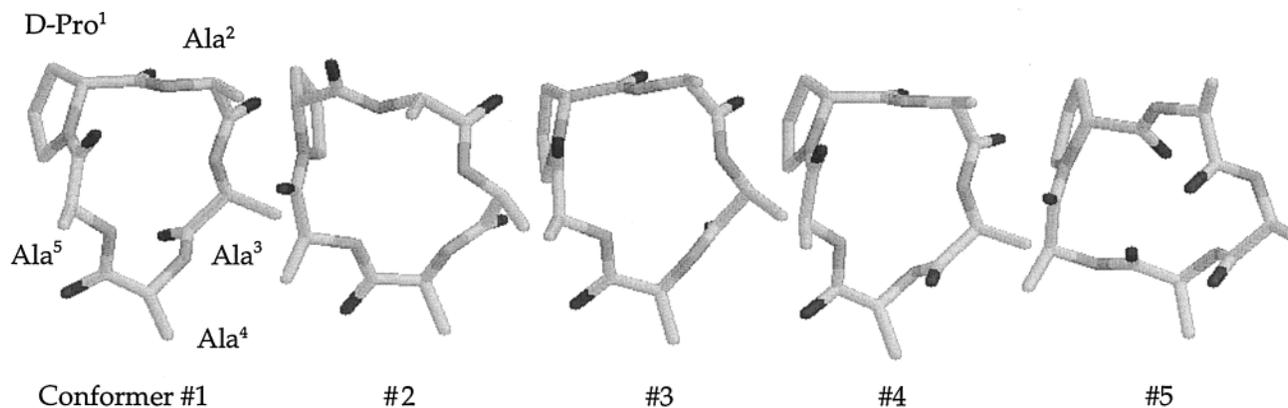
It appeared that conformer #4 is the predominant one with a mean statistical weight value of 80% (see also the lower, mean,

(41) Bystrov, V. F. *Prog. NMR Spectrosc.* **1976**, *10*, 41–81.

(42) Demarco, A.; Llinas, M.; Wuthrich, K. *Biopolymers* **1978**, *17*, 617–636.

(43) Ludvigsen, S.; Andersen, K. V.; Poulsen, F. M. *J. Mol. Biol.* **1991**, *217*, 731–736.

(44) Pardi, A.; Billeter, M.; Wuthrich, K. *J. Mol. Biol.* **1984**, *180*, 741–751.



**Figure 2.** Low-energy conformers for *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>). Residue numbering is clockwise starting from the upper left corner. Conformers are depicted from 1 to 5 from left to right. All hydrogens are omitted for clarity.

**Table 2.** Dihedral Angle Values and Relative Energies for the Calculated Low-Energy Conformers of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>)

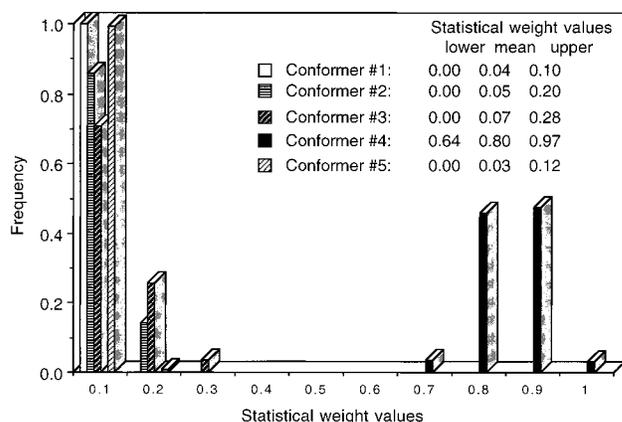
conf no.	dihedral angles											$\Delta E$ , kcal/mol
	$\phi_1$	$\psi_1$	$\phi_2$	$\psi_2$	$\phi_3$	$\psi_3$	$\phi_4$	$\psi_4$	$\phi_5$	$\psi_5$	$\omega_{51}$	
1	75	-73	-50	-53	-153	-66	-90	-16	-151	81	-169	0.0
2	75	34	-179	-68	-115	-55	-139	19	-172	68	-179	3.4
3	75	-67	-61	-78	-105	-75	-110	6	-147	71	-167	3.4
4	75	-78	-62	-65	-144	104	65	15	-171	61	-164	3.9
5	75	-142	-57	131	50	65	70	141	50	70	-158	4.8

**Table 3.** Interatomic Distances in *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) Calculated by Averaging over the Statistical Weight Distributions (Figure 3) and Experimentally Measured in Ref 29<sup>a</sup>

interproton contact <sup>b</sup>	averaged limits, Å		measd limits, Å <sup>29</sup>		conf #4
	lower	upper	lower	upper	
ProH <sup>α</sup> -Ala <sup>2</sup> NH	1.9	3.1	2.0	2.4	2.4
ProH <sup>α</sup> -Ala <sup>3</sup> NH	4.3	5.4	3.3	4.0	<b>4.8</b>
Proq $\delta_2$ -Ala <sup>5</sup> NH	2.5	3.7	3.6	4.9	<b>3.0</b>
Proq $\delta_2$ -Ala <sup>5</sup> H <sup>α</sup>	2.1	3.1	2.1	3.0	2.7
Proq $\delta_2$ -Ala <sup>5</sup> q $\beta_3$	3.9	4.9	2.6	4.6	4.4
Ala <sup>2</sup> NH-Ala <sup>2</sup> H <sup>α</sup>	2.3	3.3	2.7	3.0	2.8
Ala <sup>2</sup> NH-Ala <sup>2</sup> q $\beta_3$	1.9	3.1	2.6	3.1	2.4
Ala <sup>2</sup> NH-Ala <sup>3</sup> NH	2.3	3.5	2.3	2.8	2.9
Ala <sup>2</sup> H <sup>α</sup> -Ala <sup>3</sup> NH	2.9	4.1	2.6	3.1	3.6
Ala <sup>2</sup> q $\beta_3$ -Ala <sup>3</sup> NH	2.2	3.4	2.8	3.8	2.8
Ala <sup>3</sup> H <sup>α</sup> -Ala <sup>3</sup> NH	2.3	3.4	2.4	3.0	2.9
Ala <sup>3</sup> H <sup>α</sup> -Ala <sup>4</sup> NH	1.6	3.1	2.6	3.1	2.1
Ala <sup>3</sup> NH-Ala <sup>3</sup> q $\beta_3$	2.3	3.4	2.7	3.5	2.9
Ala <sup>3</sup> NH-Ala <sup>3</sup> NH	2.1	3.7	3.5	4.2	<b>2.8</b>
Ala <sup>4</sup> NH-Ala <sup>4</sup> H <sup>α</sup>	1.7	2.9	2.4	3.0	2.2
Ala <sup>4</sup> NH-Ala <sup>5</sup> NH	1.9	3.3	2.2	2.8	2.6
Ala <sup>4</sup> H <sup>α</sup> -Ala <sup>5</sup> NH	2.6	3.7	3.0	3.6	3.1
Ala <sup>4</sup> q $\beta_3$ -Ala <sup>5</sup> NH	3.2	4.3	2.6	3.7	3.9
Ala <sup>5</sup> NH-Ala <sup>5</sup> H <sup>α</sup>	2.2	3.3	2.6	3.0	2.8
Ala <sup>5</sup> NH-Ala <sup>5</sup> q $\beta_3$	2.5	3.6	2.6	3.6	3.1

<sup>a</sup> The last column contains the corresponding distances for the indispensable conformer #4. <sup>b</sup> q represents pseudoatoms (i.e., the corresponding C <sup>$\beta$</sup>  or C <sup>$\delta$</sup>  atoms).

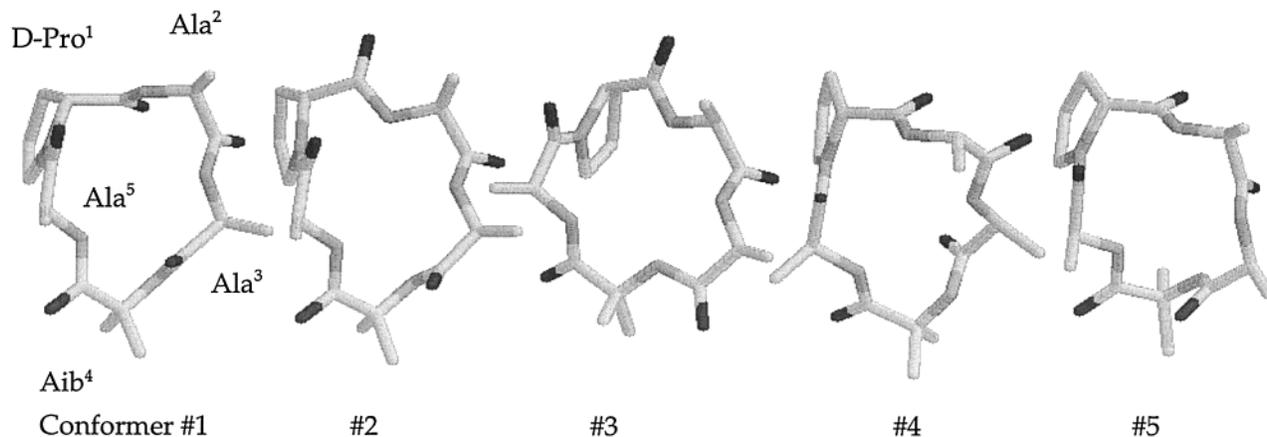
and upper statistical weight values for each conformer in Figure 3). However, the maximal limit of this weight is not equal to 100%, which means that this conformer alone cannot account for good agreement with all experimental data. The obtained distributions of statistical weights are depicted in Figure 3, and the corresponding 20 interproton distances averaged over the distributions of Figure 3 are listed in Table 3 together with the experimental data. There is no disagreement between the experimental and calculated data. At the same time, conformer #4, though the predominant one, disagrees with the experimental data in three cases (within the distance limits of  $\pm 0.5$  Å, see the last column of Table 3, in bold).



**Figure 3.** Distributions of statistical weight values for low-energy conformers of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>). Frequencies of occurrence of a given statistical weight value for each particular conformer are normalized.

These results do not mean that all five low-energy conformers are necessary for achieving agreement with the experimental data. Obviously, each of the four conformers with the lower statistical weight limit of zero can be discarded. The additional separate runs of our procedure showed that the “minimal” possible sets of the low-energy conformers of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) consist of two structures, the indispensable conformer #4 and any one of the other conformers (i.e., conformer #1, #2, #3, or #5). All of those “minimal” sets demonstrate the same level of agreement with the experimental data; in each pairwise combination, we are able to find statistical weight values satisfying the inequalities (1).

***cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>).** The independent energy calculations were able to find low-energy 3D structures of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) that are consistent both with the NMR data in DMSO and with the available X-ray data on CPP's. Contrary to the conclusions of the Kessler group, we have found that the preferable Ala<sup>4</sup> conformations are in the regions corresponding either to right or to left  $\alpha$ -helices, but not in the  $\gamma$ -turn conformation. On the other hand, the Aib



**Figure 4.** Low-energy conformers of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>). Residue numbering is clockwise starting from the upper left corner. Conformers are depicted from 1 to 5 from left to right. All hydrogens are omitted for clarity.

**Table 4.** Dihedral Angle Values and Relative Energies for the Calculated Low-Energy Conformers of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>)

conf no.	dihedral angles											$\Delta E$ , kcal/mol
	$\phi_1$	$\psi_1$	$\phi_2$	$\psi_2$	$\phi_3$	$\psi_3$	$\phi_4$	$\psi_4$	$\phi_5$	$\psi_5$	$\omega_{51}$	
1	75	-80	-61	-67	-139	106	60	19	-174	61	-164	0.0
2	75	11	-175	-55	-134	76	70	15	-170	67	178	0.4
3	75	26	-161	-81	-126	12	175	-39	-130	73	-178	2.2
4	75	-82	-50	-53	-154	-78	-66	-26	-157	96	-169	3.0
5	75	-71	-88	-88	-82	82	53	25	-155	62	-153	3.1

residue ( $\alpha$ -methylalanine, MeA) is known to limit conformational flexibility of the backbone either to the right or to the left  $\alpha$ -helix.<sup>45</sup> Therefore, studies of the *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>) peptide should verify the efficiency and reliability of independent energy calculations for conformational studies of CPP's even more convincingly.

**Energy Calculations.** Energy calculations for *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>) included 2840 peptide conformers geometrically allowed to close the pentapeptide ring. Five of them possessed relative energies  $\leq 5$  kcal/mol. The Aib<sup>4</sup> residue in the five conformers possesses the  $\phi, \psi$  values as follows: ( $60^\circ, -19^\circ$ ); ( $70^\circ, 15^\circ$ ); ( $175^\circ, -39^\circ$ ); ( $-66^\circ, -26^\circ$ ); ( $53^\circ, 25^\circ$ ). These 3D structures are depicted in Figure 4 and described in Table 4. Again, none of the structures possesses the pronounced  $\beta II'$  turn in the D-Pro<sup>1</sup>-Ala<sup>2</sup> region, but all of them are geometrically similar to the  $\beta II'\gamma$  type. Energy calculations confirmed the initial assumption that the low-energy conformers of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) will retain either right or left  $\alpha$ -helices as preferential conformations for the Aib<sup>4</sup> residue as found in four conformers out of five (Table 4).

**Synthesis.** The *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>) peptide was synthesized by solid phase synthesis using routine manual methods. However, attempts to couple Boc-Aib to Boc-Ala-resin resulted in forming the diketopiperazine structure leading to cleavage from the support. Accordingly, we have synthesized Boc-Ala-Aib-OH dipeptide separately in solution and then incorporated it using the solid-phase approach. All details of the synthesis are described in the Experimental Section. The overall yield was ca. 20% for the linear pentapeptide, and ca. 20% for the cyclization step.

**NMR Measurements.** We have obtained NMR data for *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>) in DMSO solution employing techniques of 1D- and 2D-NMR <sup>1</sup>H and <sup>13</sup>C spectroscopy. Sequential assignment of proton resonances was obtained in a straightforward manner by the combined use of 2D TOCSY

**Table 5.** <sup>1</sup>H and <sup>13</sup>C<sup>(a)</sup> Chemical Shifts (ppm) and Homonuclear <sup>1</sup>H Coupling Constants (*J*, Hz) for *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>) (300 K, DMSO-*d*<sub>6</sub>)

residue	NH	H <sup><math>\alpha</math></sup> /C <sup><math>\alpha</math></sup>	H <sup><math>\beta</math></sup> / $\beta'$ /C <sup><math>\beta</math></sup>	H <sup><math>\gamma</math></sup> / $\gamma'$ /C <sup><math>\gamma</math></sup>	H <sup><math>\delta</math></sup> /C <sup><math>\delta</math></sup>	CO
D-Pro <sup>1</sup>		4.27 $J_{\alpha\beta} = 4.3$ $J_{\alpha\beta'} = 8.3$ 60.4 <sup>a</sup>	1.97 ( $\beta$ ) 1.85 ( $\beta'$ ) 28.4 <sup>a</sup>	2.13 ( $\gamma$ ) 1.86 ( $\gamma'$ ) 25.7 <sup>a</sup>	3.53 47.2 <sup>a</sup>	172.6 <sup>a</sup>
Ala <sup>2</sup>	8.72 $J_{NH\alpha} = 7.5$	4.02 $J_{\alpha\beta} = 7.6$	1.24 17.9 <sup>a</sup>			172.3 <sup>a</sup>
Ala <sup>3</sup>	7.63 $J_{NH\alpha} = 7.3$	4.23 $J_{\alpha\beta} = 6.6$ 49.6 <sup>a</sup>	1.22 17.8 <sup>a</sup>			171.9 <sup>a</sup>
Aib <sup>4</sup>	7.92	-	1.5 ( $\beta$ ) 1.23 ( $\beta'$ ) 57.9 <sup>a</sup> 25.9/24.9 <sup>a</sup>			174.6 <sup>a</sup>
Ala <sup>5</sup>	7.80 $J_{NH\alpha} = 8.7$	4.54 $J_{\alpha\beta} = 7.0$ 46.9 <sup>a</sup>	1.19 18.9 <sup>a</sup>			172.2 <sup>a</sup>

<sup>a</sup> Chemical shifts were referenced to the residual DMSO solvent signal at 2.49 ppm for <sup>1</sup>H and 39.5 ppm for <sup>13</sup>C. <sup>b</sup> Coupling constants were measured from the resolution enhanced 1D spectrum or highly digitized 1D traces of the 2D TOCSY experiment with an accuracy of ca.  $\pm 0.2$  Hz.

and ROESY spectra.<sup>46</sup> <sup>1</sup>H chemical shifts and homonuclear coupling constants (reported in Table 5) were extracted from resolution enhanced 1D proton and/or 1D TOCSY spectra or, in case of signal overlap, from the highly digitized 1D traces of the gradient-enhanced TOCSY spectrum.<sup>47-50</sup> ROE peak intensities measured with the mixing time of 120 ms and the corresponding estimated interproton distances (NH-H <sup>$\alpha$</sup>  dis-

(46) Wüthrich, K. *NMR of proteins and nucleic acids*; Wiley-Interscience: New York, 1986.

(47) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *65*, 355-360.

(48) Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1983**, *53*, 521-528.

(49) Davis, A. L.; Estcourt, G.; Keeler, J.; Laue, E. D.; Titman, J. J. *J. Magn. Res. A.* **1993**, *105*, 167-183.

(50) Kövér, K. E.; Uhrin, D.; Hruby, V. J. *J. Magn. Reson.* **1998**, *130*, 162-168.

(45) Marshall, G. R. *Tetrahedron* **1993**, *49*, 3547-3558.

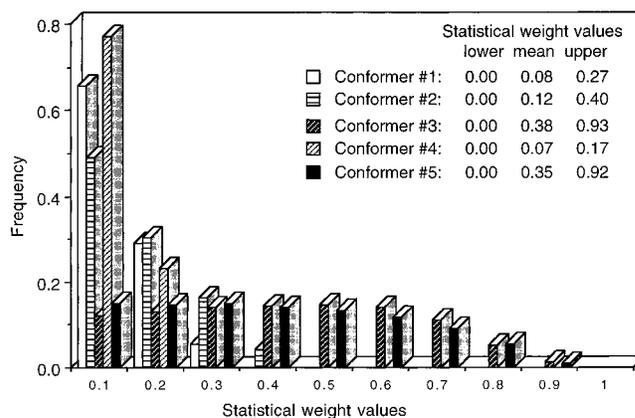
**Table 6.** Interatomic Distances and Values of Vicinal Coupling Constants in *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>) Calculated by Averaging over the Statistical Weight Distributions (Figure 5) and Experimentally Measured

interproton contact/ $J(\text{HC}^\alpha\text{NH})$	averaged limits, Å/Hz		measd limits, Å/Hz	
	lower	upper	lower	upper
D-ProH <sup>α</sup> -Ala <sup>2</sup> NH	2.3	3.4	1.9	2.3
Ala <sup>2</sup> H <sup>α</sup> -Ala <sup>2</sup> NH	2.4	3.4	2.5	3.1
Ala <sup>3</sup> H <sup>α</sup> -Aib <sup>4</sup> NH	2.4	3.4	2.1	2.5
Ala <sup>3</sup> H <sup>α</sup> -Ala <sup>3</sup> NH	2.1	3.5	2.5	3.1
Ala <sup>3</sup> H <sup>α</sup> -Ala <sup>3</sup> NH	1.6	4.4	2.4	3.0
Ala <sup>3</sup> NH-Ala <sup>2</sup> H <sup>α</sup>	2.2	3.8	2.8	3.4
Ala <sup>2</sup> NH-Ala <sup>3</sup> NH	1.5	3.2	2.3	2.7
Aib <sup>4</sup> NH-Ala <sup>3</sup> NH	2.0	4.0	3.1	3.7
Ala <sup>3</sup> NH-Ala <sup>5</sup> NH	3.1	4.1	2.9	3.5
Aib <sup>4</sup> NH-Ala <sup>5</sup> NH	1.9	3.6	2.6	3.2
Ala <sup>3</sup> H <sup>α</sup> -Ala <sup>5</sup> NH	3.3	5.4	3.4	4.2
Ala <sup>5</sup> H <sup>α</sup> -Ala <sup>3</sup> NH	4.7	6.3	4.3	5.3
Ala <sup>2</sup> NH-Ala <sup>5</sup> NH	2.2	3.9	3.2	4.0
D-ProH <sup>β</sup> -Ala <sup>5</sup> H <sup>α</sup>	1.5	2.9	1.8	2.2
D-ProH <sup>β</sup> -Ala <sup>3</sup> NH	3.8	6.5	4.2	5.2
D-ProH <sup>β</sup> -Ala <sup>5</sup> NH	2.3	3.7	3.7	4.5
Ala <sup>5</sup> $J(\text{HC}^\alpha\text{NH})$	3.0	9.6	8.5	8.9
Ala <sup>2</sup> $J(\text{HC}^\alpha\text{NH})$	1.9	9.5	7.3	7.7
Ala <sup>3</sup> $J(\text{HC}^\alpha\text{NH})$	3.5	9.9	7.1	7.5

tances of ca. 2.8 in Ala<sup>2</sup> and Ala<sup>5</sup> were used as internal reference) are reported in Table 6. Carbon assignment of protonated carbons has been deduced from the sensitivity-enhanced gradient HSQC experiment<sup>51-53</sup> relying on previously assigned proton resonances. Characteristic carbon chemical shifts of D-Pro<sup>1</sup> could be used as further support of the *trans*-amide bond Ala<sup>5</sup>-D-Pro<sup>1</sup>, already assigned based on ROE connectivities. All carbon chemical shift data are also given in Table 5.

**Estimation of Statistical Weights for Low-Energy Conformers in DMSO.** We have estimated possible statistical weights for the five low-energy conformers of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>) by the same approach used above in the case of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>). We have used 16 values for interproton distances and 3 values for  $J(\text{HC}^\alpha\text{NH})$  coupling constants listed in Tables 5 and 6 as  $\langle A^{\text{exp}} \rangle$ 's. The values of  $D^{\text{exp}}$ 's for the interproton distances have been calculated as ca.  $\pm 10\%$  of the corresponding distance. The coupling constants were chosen with an experimentally defined error of ca.  $\pm 0.2$  Hz (see Table 5). For each of the five low-energy conformers, the same 16 interatomic distances and 3 values of coupling constants were calculated and used as  $\langle A^{\text{calc}} \rangle$ 's. It was assumed also that the  $D^{\text{calc}}_{ik}$  values were  $\pm 0.5$  Å for all interatomic distances. The coupling constants were chosen as limits of the  $J(\text{HC}^\alpha\text{NH})$  variations due to variations of the corresponding  $\phi$  angles by 20° as in the previous case. Out of ca. 1000000 random trials, 10000 different  $\{w_i\}$  points of five statistical weights were chosen randomly, all satisfying the above conditions with  $t = 1.645$  (the confidence interval of 90%).

The obtained distributions of statistical weights together with the lower, mean and upper statistical weight values for each conformer are depicted in Figure 5. In this case two conformers can be regarded as more significant than others, namely conformers #3 and #5, with mean statistical weight values of

**Figure 5.** Distributions of statistical weight values for low-energy conformers of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>). Frequencies of occurrence of a given statistical weight value for each particular conformer are normalized.

38% and 35%, respectively (the maximal limits of these weights are equal to 93% and 92%, respectively). Distributions of statistical weights for these conformers are highly correlated with the correlation coefficient of  $-0.91$ , which means that these two conformers can, in a sense, represent each other. Interestingly, conformer #5 is very similar to the predominant conformer #4 of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>); compare Tables 2 and 4. The corresponding 16 interproton distances and the 3 values of coupling constants averaged over the distributions of Figure 5 are listed in Table 6 together with the experimental data. There is no disagreement between the two data sets in Table 6.

As in the previous case, the “minimal” possible sets of the low-energy conformers of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>) consist of two structures. The first structure is either conformer #3 or conformer #5, and the second structure may be any other conformer (with the exception of the pair consisting of conformer #1 and conformer #5). Again, all of those “minimal” pairs of structures demonstrate the same level of agreement with the experimental data.

## Discussion

In our view, these results lead to several important conclusions. First, CPP's still possess certain conformational flexibility in solution, since it was impossible to accommodate all NMR restraints for the simple *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) peptide with a single 3D structure without steric hindrance. Second, as a consequence, the  $\beta\text{II}'\gamma$  model suggested for CPP's of the (aBCDE) type earlier by NMR measurements refined by energy calculations is invalid, being an artifact of conformational averaging. Third, the low-energy conformers of CPP's obtained by independent energy calculations find experimental confirmation when confronted with NMR data. In total, it suggests that the low-energy conformations of CPP's obtained by independent energy calculations can be used as a good starting point for assessing actual 3D structures of CPP's in solution. This approach, of course, has its own inherent limitations; some of them are discussed below followed by discussion on applicability of 3D structures of CPP's as conformational templates for pharmacophore models.

**Limitations of Combining NMR Data and Independent Energy Calculations.** Our approach interprets the NMR data for CPP's with the aid of two computational procedures: the search for all low-energy backbone conformers employing the ECEPP force field and the procedure for statistical weight estimation described above. As has been pointed out earlier,<sup>40</sup>

(51) Kay, L. E.; Keifer, P.; Saareinen, T. *J. Am. Chem. Soc.* **1992**, *114*, 10663.

(52) Kontaxis, G.; Stonehouse, J.; Laue, E. D.; Keeler, J. *J. Magn. Reson. A* **1994**, *111*, 70-76.

(53) Palmer, A. G., III; Cavanagh, J.; Wright, P. E.; Rance, M. *J. Magn. Reson.* **1991**, *93*, 151-170.

general results of the search for all low-energy backbone conformers depend mainly on the applied computational protocols and on the employed force field. One can add the third factor, which is the energy threshold established for selection of "low-energy" structures.

The computational protocol used in this study was, in fact, a variant of a systematic search procedure (see Methods). It considered 3895 peptide conformers geometrically allowed that closed the pentapeptide ring for *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) and 2840 conformers for *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>). Since only  $\psi_1$ ,  $\omega_{12}$ ,  $\phi_{2-4}$ , and  $\psi_{2-4}$  have been considered as variables for these CPP's in assuming rigid valence geometry (see Methods), the number of independent variables has been  $10 - 6 = 4$  (ref 65). Obviously, this four-dimensional conformational space has been searched very exhaustively.

Selection of the energy threshold of 5 kcal/mol for calculations employing the ECEPP force field has already been partly validated by energy calculations performed for seven CPP's with the known X-ray structures (see Results). However, we performed an additional validation of the energy threshold value by including in the procedure for estimating statistical weights not 5, but 10 low-energy structures of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>), thus increasing the energy threshold to 7.5 kcal/mol. The resulting distributions of statistical weight values showed that conformer #4 retains its domination and indispensability, the lower limit, the mean value, and the upper limits being 0.35, 0.64, and 0.83, respectively. At the same time, the total mean statistical weight of the added five conformers was ca. 0.10. In other words, the energy threshold of 5 kcal/mol is quite sufficient for selection of possible low-energy conformers in solution as judged by these limited studies. Evidently, that value is large enough to account for uncertainties in our energy calculations performed with the rather nonsophisticated ECEPP force field in the absence of explicitly described solvent molecules.

On the other hand, the choice of the ECEPP force field is, perhaps, much more important. To investigate this point, we performed a Monte Carlo driven conformational search employing the AMBER\* force field implemented in the commercially available MacroModel program for *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>). The run included 5000 conformers selected as the starting structures for energy minimization, and found 7 conformers with relative energies less than 2 kcal/mol. (Note that application of the energy threshold of 2 kcal/mol in the AMBER\* force field yielded almost the same number of the "low-energy" conformers as the energy threshold of 5 kcal/mol in the ECEPP force field calculations.) The AMBER\* force field calculations yielded 13 conformers within 3 kcal/mol, and 22 conformers within 5 kcal/

mol; a cyclopentapeptide can hardly possess so many "low-energy" conformers. Our procedure for estimating statistical weights has been applied to the above 7 conformers, and has yielded good agreement with the experimental data showing that none of the conformers is indispensable, and their mean statistical weight values range from 0.07 to 0.25. There was, however, a serious problem. Four out of those seven conformers (and 14 out of 22 conformers within 5 kcal/mol) feature the distinct  $\gamma$ -turn conformation for one of the Ala residues, the conformation that, as has been argued above, is sterically forbidden. The remaining three conformers alone did not yield agreement with the NMR data; they also possessed the relative energies well above 10 kcal/mol when re-minimized with the ECEPP force field. One need to add the "non- $\gamma$ -turn" conformer #21 to the set of the "low-energy" conformers to achieve agreement with the experimental data. This conformer possesses a relative energy of 4.5 kcal/mol in the AMBER\* force field and 12.4 kcal/mol in the ECEPP force field that excluded it from the list of "low-energy" structures.

The second main computational procedure in question is estimating the statistical weights of the low-energy conformers by selecting sets of statistical weight values,  $\{w_i\}$ , to ensure that parameters that are experimentally measured and averaged over all low energy conformers, are statistically indistinguishable. It is important that in our procedure the above condition is satisfied for *each separate* parameter, not for the weighted sum of them, as in the approaches of others.<sup>66,67</sup> In this way, the level of agreement between experimental and calculated data is the same for each parameter; otherwise, agreement may be better for some parameters than for others. However, this advantage has its price, the inability to find the single all-satisfying set of statistical weight values. Instead, our procedure produces the possible distributions of statistical weight values for each low-energy conformer revealing which one(s) of them is (are) the most important for achieving agreement with the experimental data.

Obviously, distributions of statistical weight values for low-energy conformers provide only qualitative estimations. They depend on the assumed values of some parameters such as  $t_k$ ,  $D^{\text{calc}}$ ,  $D^{\text{exp}}$ , etc. For instance, to check the stability of our estimations, we have run the procedure of statistical weight selection for low-energy conformers of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>) several times with somewhat different values for the mentioned parameters ( $\pm 50\%$  of the values described above). The obtained mean values of statistical weights slightly varied, but in all cases the general results were qualitatively the same showing the zero lower limits for statistical weights of all conformers, as well as moderate predominance of the same conformer #5, as in the previous results. Therefore, we regard the distributions of statistical weight values obtained as fairly reliable. Generally, however, such distributions should be inspected for their stability on a case-to-case basis.

**3D Structures of Cyclopentapeptides as Templates for Pharmacophore Models.** The best known case of employing NMR spectroscopy for elucidating CPP pharmacophores is the design of RGD-containing CPP's by the Kessler group. They found that both *cyclo*(Arg-Gly-Asp-D-Phe-Val) and *cyclo*(Arg-Gly-Asp-Phe-D-Val) (*c*(RGDFV) and *c*(RGDFv), respectively) are almost equally potent inhibitors of binding  $\alpha_{\text{IIb}}\beta_3$  integrins to fibrinogen and of  $\alpha_v\beta_3$  integrins to vitronectin at a level of

(54) Pfaff, M.; Tangemann, K.; Muller, B.; Gurrath, M.; Muller, G.; Kessler, H.; Timpl, R.; Engel, J. *J. Biol. Chem.* **1994**, *269*, 20233–20238.

(55) Muller, G.; Gurrath, M.; Kessler, H. *J. Comput.-Aided Mol. Des.* **1994**, *8*, 709–730.

(56) Kopple, K. D.; Bures, P. W.; Bean, J. W.; D'Ambrosio, C. A.; Hughes, J. L.; Peishoff, C. E.; Eggleston, D. S. *J. Am. Chem. Soc.* **1992**, *114*, 9615–9623.

(57) Bothner-By, A. A.; Stephens, R. L.; Lee, J.; Warren, C. D.; Jeanloz, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 811–813.

(58) Bax, A.; Davis, D. G. *J. Magn. Reson.* **1985**, *63*, 207–213.

(59) Dezheng, Z.; Fujiwara, T.; Nagayama, K. *J. Magn. Reson.* **1989**, *81*, 628–630.

(60) Desvaux, H.; Berthault, P.; Birlirakis, N.; Goldman, M. *J. Magn. Reson. A* **1994**, *108*, 219–229.

(61) Kuwata, K.; Schleich, T. *J. Magn. Reson. A* **1994**, *111*, 43–49.

(62) Hurd, R. E.; John, B. K. *J. Magn. Reson.* **1991**, *91*, 648–653.

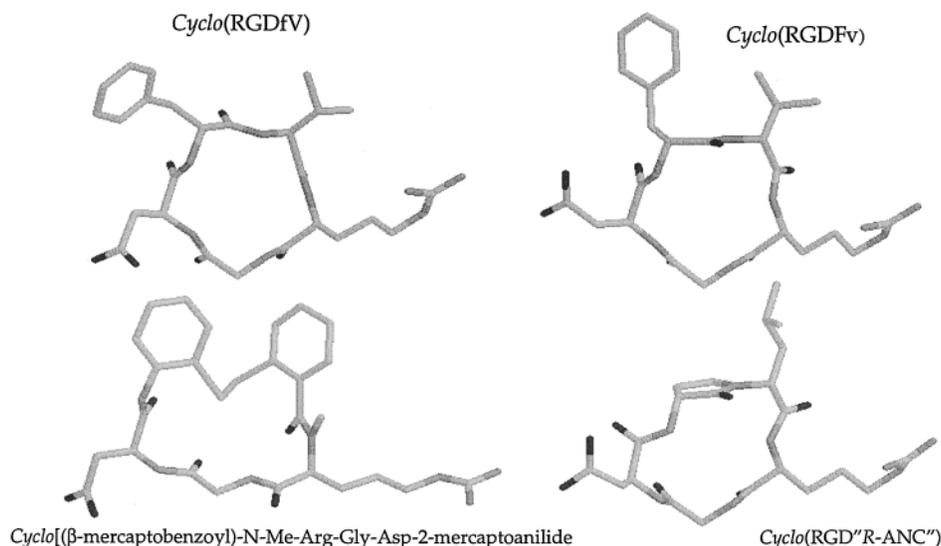
(63) Zimmerman, S. S.; Scheraga, H. A. *Biopolymers* **1977**, *16*, 811–843.

(64) Nikiforovich, G. V.; Hrubby, V. J.; Prakash, O.; Gehrig, C. A. *Biopolymers* **1991**, *31*, 941–955.

(65) Go, N.; Scheraga, H. A. *Macromolecules* **1975**, *8*, 750–761.

(66) Cicero, D. O.; Barbato, G.; Bazzo, R. *J. Am. Chem. Soc.* **1995**, *117*, 1027–1033.

(67) Nevins, N.; Cicero, D.; Snyder, J. P. *J. Org. Chem.* **1999**, *64*, 3979–3986.



**Figure 6.** Geometrically similar low-energy conformers of *c*(RGDfV) and *c*(RGDFv) (the upper row), as well as the X-ray structure of *c*[(β-mercaptobenzoyl)-N-Me-Arg-Gly-Asp-2-mercaptoanilide]<sup>56</sup> and the 3D structure of *c*(RGD'R-ANC'').

**Table 7.** Dihedral Angles of the Backbone for Geometrically Similar Conformers of *c*(RGDfV) and *c*(RGDFv)

peptide	$\phi_1$	$\psi_1$	$\phi_2$	$\psi_2$	$\phi_3$	$\psi_3$	$\phi_4$	$\psi_4$	$\phi_5$	$\psi_5$
<i>c</i> (RGDfV)	-84	-75	-76	-76	-104	-57	-60	-81	-111	-52
<i>c</i> (RGDFv)	-102	-72	-74	-78	-103	-52	-115	90	113	-73

a few hundreds nanomolar.<sup>22,24,54</sup> However, since both peptides, according to their interpretation of the NMR data,<sup>22</sup> should possess a single conformation of the  $\beta\text{II}'\gamma$  type, the conformations of the bioactive motif, RGD, are dissimilar to each other in these two peptides (see, e.g., Figure 1 comparing conformations of residues 3–4–5 (clockwise from the right lower corner) and 2–3–4 (clockwise from the right upper corner) that corresponds to the  $\beta\text{II}'\gamma$  conformation for *c*(RGDfV) and *c*(RGDFv), respectively). To explain this discrepancy, it was suggested that the  $\beta\text{II}'\gamma$  conformation of *c*(RGDfV) may undergo a conformational transition via an intermediate  $\beta\text{I}\gamma'$  turn shifted along the sequence by one residue, and then via conversion of the  $\gamma'$ -turn into a  $\gamma$ -turn that will finally lead to the  $\beta\text{II}'\gamma$  type conformation of *c*(RGDFv).<sup>22</sup> It was noted, however, that the reverse conformational transition, from the  $\beta\text{II}'\gamma$  type conformation of *c*(RGDFv) to the  $\beta\text{II}'\gamma$  type conformation of *c*(RGDfV), would be not possible.<sup>22</sup> Nevertheless, the  $\beta\text{II}'\gamma$  type conformation of *c*(RGDfV) has been proposed as the 3D model of the pharmacophore for the RGD-containing CPP's. Several years later, the above discrepancy was explained in a different manner, namely by similarity of spatial arrangements of the  $\text{C}^\alpha\text{--C}^\beta$  vectors for Arg and Asp in both peptides,<sup>55</sup> so the  $\beta\text{II}'\gamma$  type conformation of *c*(RGDfV) was still suggested as the most probable 3D model of the pharmacophore for RGD-containing CPP's.<sup>25</sup> However, the authors noted that the results of their "vector analysis" are not in agreement with the independent 3D model for the RGD pharmacophore confirmed by X-ray studies.<sup>56</sup> Moreover, introduction of a rigid peptidomimetic element stabilizing the suggested  $\beta\text{II}'\gamma$  type structure resulted in the complete loss of inhibition of binding of  $\alpha_{\text{IIIb}}\beta_3$  integrins to fibrinogen and of  $\alpha_{\text{V}}\beta_3$  integrins to vitronectin (compound *c*(RGD"spiro"), PA4; see Table 2 and Figure 3 in ref 27), whereas stabilizing a different 3D structure yields the best of the tested compounds (see *c*(RGD'R-ANC'), PA2; see Table 2 and Figure 3 in ref 27). Therefore, finding the *c*(RGD'R-ANC'') compound showing an excellent level of inhibition of vitronectin binding to the

$\alpha_{\text{V}}\beta_3$  receptors ( $\text{IC}_{50} = 0.85 \text{ nM}^{28}$ ) can hardly be regarded as a result of rational drug design based on NMR studies of CPP's.

At the same time, our energy calculations have revealed seven low-energy backbone conformers ( $\Delta E \leq 5 \text{ kcal/mol}$ ) for *c*(RGDfV), and six low-energy backbone conformers for *c*(RGDFv). Since the detailed NMR data for both peptides are unavailable in the literature, we have performed additional calculations for the very similar peptide, *c*(RGDWv), instead of *c*(RGDFv) to confront our results with independent NMR data. We have considered 12 measured interproton distances within the peptide backbone (Table 2 in<sup>25</sup>) as  $\langle A^{\text{exp}} \rangle$ 's, with the values of  $D^{\text{exp}}$ 's of  $\pm 0.3 \text{ \AA}$ . For each of the six low-energy conformers of *c*(RGDWv), the same interproton distances were calculated and used as  $\langle A^{\text{calc}} \rangle$ 's, with the  $D^{\text{calc}}_{\text{ik}}$  values of  $\pm 0.3 \text{ \AA}$ . Out of ca. 1000000 trials, 10000 different  $\{w_i\}$  points of six statistical weights were chosen randomly, all satisfying the above-described conditions with  $t = 1.645$ . It appeared that no low-energy conformer alone satisfies all NMR data, since all mean statistical weight values were of 15–18%, and all minimal limits were equal to zero. However, the six conformers together ensure excellent agreement with the experimental data.

Geometrical similarity of low-energy conformers for *c*(RGDfV) and *c*(RGDFv) (i.e., for 42 pairs of conformers) has been inspected by achieving the best fit of spatial arrangements of the  $\text{C}^\alpha$  and  $\text{C}^\beta$  atoms for the RGD sequence and of the  $\text{C}^\alpha$  atoms for the L/D-Phe and L/D-Val residues. The distances between all seven corresponding atoms have been less than  $0.50 \text{ \AA}$  only for the one pair of conformers for each peptide. These conformers are depicted in Figure 6 and described in Table 7. It is noteworthy that the conformer of *c*(RGDfV) described in Table 7 possesses all negative values of the  $\phi$  and  $\psi$  dihedral angles, presenting a somewhat distorted  $\alpha$ -helical conformation. (A similar  $\alpha$ -helical-like conformation has been reported earlier for the *c*(DWMDf) peptide.<sup>4</sup>) The difference between the conformers of *c*(RGDfV) and *c*(RGDFv) in Table 7 is mainly in the flip-flop rotation of the Phe-Val peptide bond plane, but not in the RGD region. This particular conformer could be

regarded as the 3D pharmacophore model for the RGD-containing CPP's, which is not only devoid of the discrepancy discussed above, but is also in good agreement with the model for the RGD pharmacophore proposed by other authors<sup>56</sup> (see Figure 6).

Figure 6 contains also the 3D model of the best RGD peptidomimetic obtained in ref 27, namely *c*(RGD"R-ANC"), which is also in very good agreement with both conformers in Table 7. The model has been built introducing the mimetic block into the above conformer of *c*(RGDFV) with subsequent energy minimization employing the SYBYL 6.3 package. Interestingly, it was impossible to introduce the same conformer to an enantiomer of the same mimetic (i.e., the "S-ANC" block) as the corresponding chiral center underwent inversion during the very first steps of energy minimization. Obviously, the enantiomeric block does not fit 3D structures depicted in Figure 6: compare orientations of the CO groups of the D-Phe residue in *c*(RGDFV) and of the ANC block in *c*(RGD"R-ANC"). On the other hand, the enantiomeric compound, *c*(RGD"S-ANC"), showed about 500-fold less potency than *c*(RGD"R-ANC");<sup>27</sup> These results further corroborate the suggestion that the 3D structures described above are reliable models for the RGD pharmacophore inhibiting binding of  $\alpha_{IIb}\beta_3$  integrins to fibrinogen.

The above example shows that independent energy calculations have been able to find an internally consistent 3D model of the pharmacophore for the RGD-containing CPP's that is in agreement with the model proposed by other authors for highly potent RGD-related compounds as confirmed by X-ray studies.<sup>56</sup> This conclusion was not achieved by approaches employing only NMR spectroscopy data refined by energy calculations.<sup>22,23,25,55</sup> Therefore, it may be concluded that cyclopentapeptides are indeed very convenient compounds for use as receptor probes (see also the recent paper on CPP's as scaffolds for interactions with G-protein coupled receptors<sup>68</sup>).

## Conclusions

Apart from the obvious general conclusion that cyclopentapeptides should be regarded as rather flexible systems, and, as such, their studies will benefit from employing independent energy calculations along with NMR measurements, it is important to emphasize one particular conclusion of this paper. Namely, our results point out that the  $\beta II'\gamma$  model for CPP's is not valid if the  $\gamma$ -turn is centered at the amino acid residue of L-configuration as repeatedly suggested by the Kessler lab. In addition to our data, we can support this conclusion by the results of other authors, both experimental and theoretical. First, the distribution of the  $\phi, \psi$  values for individual residues from 462 proteins studied by the X-ray crystallography (121870 residues) shows that 82% of them are located inside of three "core" regions in the Ramachandran map; these regions do not include the region of the  $\gamma$ -turn that is populated with less than 1% of all  $\phi, \psi$  points.<sup>69</sup> This is not a special feature of crystallized proteins; similar studies for 3D structures of proteins solved by NMR showed that there are ca. 67% of all points in the core regions of the Ramachandran map (data for 21 proteins;<sup>70</sup> more recent data for 97 proteins estimate the same population of somewhat less than 90%<sup>71</sup>). The cyclic constraint

imposed on CPP's also does not force the residues of L-/D-configuration into the  $\gamma$ -/ $\gamma'$ -turn regions in the Ramachandran map. The  $\phi, \psi$  values for 110 chiral residues included into 29 cyclopentapeptides, whose 3D structures have been studied either by the X-ray crystallography or by NMR spectroscopy, were thoroughly examined in this respect; only for two residues were the values outside the core regions of the Ramachandran map.<sup>72</sup> Therefore, in our view, there is no definitive experimental evidence that the  $\gamma$ -turn conformation actually exists for L-amino acid residues, either in the crystalline state or in solution.

On the other hand, one of the possible reasons to deduce this type of conformation for the L-amino acid residues in cyclopentapeptides was pointed out in the recent paper from the group in Berlin.<sup>73</sup> In this paper, NMR studies performed for cyclopentapeptides consisting of four L-amino acid residues and of one D-amino residue revealed contradictions in different sets of experimental data. Namely, whereas the temperature dependencies of chemical shifts of the amide protons favor the existence of the  $\gamma$ -turn conformation for one of the L-amino acid residues, the close contact between the amide protons of the residues flanking the suspected " $\gamma$ -turn" residue rules out this possibility, suggesting for this residue an "open" conformation ( $\phi \approx -75^\circ$ ,  $\psi \approx -70^\circ$ ). A delicate balance of restraints assigned to either the first or the second set of experimental data in subsequent MD simulations led either to the  $\gamma$ -turn or to the open conformation. Interestingly, the NOE data obtained by the Kessler group on *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) clearly show the contact between the NH protons of Ala<sup>3</sup> and Ala<sup>5</sup> (see Table 3) that is, according to the above,<sup>73</sup> inconsistent with the  $\gamma$ -turn conformation of Ala<sup>4</sup>. However, the final experimental resolution of the " $\gamma$ -turn" problem in CPP's is yet to come; in our view, solid-state NMR spectroscopy of peptides is a very promising approach in this regard (e.g., ref 74).

At the same time, if the  $\gamma$ -turn conformation indeed has no experimental confirmation, one more important conclusion may be drawn. Our results obtained for *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Ala<sup>4</sup>-Ala<sup>5</sup>) using our energy calculation procedure and the MacroModel program suggest that one needs to be rather cautious employing the AMBER-like force fields in energy calculations for CPP's. Such force fields regard the  $\gamma$ -turn conformations for the L-amino acid residues as much more probable compared with results produced by the ECEPP force field (see a recent comparison of several force fields<sup>75</sup>). The reason for this difference has been thoroughly discussed earlier,<sup>37</sup> and has been attributed to the excessive flexibility of the peptide molecule described in terms of flexible valence geometry. As was shown, the set of "low-energy" structures of CPP's obtained by the AMBER force field calculations may yield good agreement with the experimental data, but this result can be misleading in view of overestimation of the poorly experimentally supported  $\gamma$ -turn conformation. (Note that the GROMOS force field employed in ref 29 is an AMBER-like one.) Therefore, the ECEPP force field that assumes rigid valence geometry still seems preferred for energy calculations of CPP's, especially in the process of a conformational search. Besides, there are indications that the ECEPP force field satisfies the distribution of experimentally

(72) Viles, J. H.; Mitchell, J. B. O.; Gough, S. L.; Doyle, P. M.; Harris, C. J.; Sadler, P. J.; Thornton, J. M. *Eur. J. Biochem.* **1996**, *242*, 352–362.

(73) Weisshoff, H.; Präsang, C.; Henklein, P.; Frömmel, C.; Zschunke, A.; Mügge, C. *Eur. J. Biochem.* **1999**, *259*, 776–788.

(74) Beusen, D. D.; McDowell, L. M.; Slomczynska, U.; Schaefer, J. J. *Med. Chem.* **1995**, *38*, 27–42–2747.

(75) Rodriguez, A. M.; Baldoni, H. A.; Suvire, F.; Vázquez, R. N.; Zamarbide, G.; Enriz, R. D.; Farkas, Ö.; Perczel, A.; McAllister, M. A.; Torday, L. L.; Papp, J. G.; Csizmadia, I. G. *J. Mol. Struct. (THEOCHEM)* **1998**, *455*, 275–301.

(68) Porcelli, M.; Casu, M.; Lai, A.; Saba, G.; Pinori, M.; Cappelletti, S.; Mascagni, P. *Biopolymers* **1999**, *50*, 211–219.

(69) Morris, A. L.; MacArthur, M. W.; Hutchinson, E. G.; Thornton, J. M. *Proteins* **1992**, *12*, 345–364.

(70) MacArthur, M. W.; Thornton, J. M. *Proteins* **1993**, *17*, 232–251.

(71) Doreleijers, J. F.; Rullmann, J. A. C.; Kaptein, R. *J. Mol. Biol.* **1998**, *281*, 149–164.

observed  $\phi, \psi$  values better than many other force fields including AMBER.<sup>75</sup>

## Experimental Section

**Synthesis: General Procedures.** The *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>) peptide was synthesized by solid-phase techniques using routine manual methods except for the solution synthesis of Boc-Ala-Aib-OH. The solid support used was Boc-Ala Merrifield resin (0.69 mM/g). Boc-Aib-OH, Boc-Ala-OH, and Boc-D-Pro-OH as well as the Merrifield resin were purchased from Advanced Chemtech (Louisville, KY). Solvents, DMF and CH<sub>2</sub>Cl<sub>2</sub>, were of HPLC grade and were dried in 4 Å molecular sieve prior to use. Reagents, Bop, Hobt, and HATu, were purchased from Richelieu Biotechnologies (St-Hyacinthe, Canada). The coupling indicator used was the Kaiser test. TLC was performed on silica gel plates (Analtech, 250 mm) using the indicated developing solvent. Plates were visualized by UV irradiation; by spraying with 0.5% ninhydrin solution in acetone and heating to 100 °C; and by placing them in a chamber containing Cl<sub>2</sub> vapor and then sprayed with 1% KI and 1% starch solution. The melting points were taken on a Thomas-Hooving melting point apparatus. The preparative HPLC chromatography was performed using a Dynamax instrument (Varian) equipped with a Dynamax C18 column (300 Å, 5 μM, 10 × 250 mm). The mobile phase consisted of two solvents, A (0.1% TFA in H<sub>2</sub>O) and B (acetonitrile). The purity of the peptide was determined using an analytical HPLC instrument (SP8800 Spectra-Physics, Houston, TX) with a C18 column microsorb-MV (300 Å, 5 μM, 4.3 × 250 mm). The mobile phase was as follows: A (0.05% TFA in H<sub>2</sub>O) and B (0.038% TFA in 10% H<sub>2</sub>O/90% acetonitrile). The purity and identity of the peptide was confirmed by electrospray mass spectrometry.

**Aib(Obz)·CF<sub>3</sub>COOH.** Boc-Aib-OH (2.03 g, 10 mM) was dissolved in DMF and stirred in an ice bath. Then 264 mg of NaH (11 mM) was added. After half an hour, 1.69 g of BzIbR (10 mM) was added. After removing the mixture from the ice bath, the solution was stirred at the room temperature for 6 h. Then the DMF was evaporated. The oily solid was dissolved in ethyl acetate and the solution was washed sequentially twice with a 5% solution of NaHCO<sub>3</sub> and several times with NaCl-saturated water and then dried with Mg<sub>2</sub>SO<sub>4</sub>. A solid residue was obtained after evaporation of ethyl acetate. The Boc group was deprotected with 50% TFA/CH<sub>2</sub>Cl<sub>2</sub> for 45 min. CH<sub>2</sub>Cl<sub>2</sub> was evaporated, and then ethyl ether was added for crystallization. Yield was 80%.

**Boc-Ala-Aib-Obz.** Aib-Obz·CF<sub>3</sub>COOH (1.017 g, 3.32 mM) was dissolved in DMF and neutralized with 0.577 mL of DIEA (3.32 mM). Boc-Ala-OH (753.4 mg, 3.984 mM) and 1.226 g of HATu (3.32 mM) were added into the solution, and then 1.30 mL of DIEA (6.64 mM) was added. The pH was adjusted to 7 and the solution was stirred over two nights and then distilled to remove DMF. The residual was dissolved in ethyl acetate. It was then washed sequentially with a 5% solution of NaHCO<sub>3</sub>, NaCl-saturated H<sub>2</sub>O, 0.1 N citric acid, and H<sub>2</sub>O and dried with MgSO<sub>4</sub>. Finally, the solution was evaporated to dryness to obtain an oil. The TLC plate showed only one spot (positive in UV and Cl<sub>2</sub> tests). Yield was 90%.

**Boc-Ala-Aib-OH.** Boc-Ala-Aib(Obz) (1.15 g, 3.17 mM) oil was hydrogenated with Pd/C in methanol for 3 h. After that Pd/C was filtered and the methanol evaporated. Boc-Ala-Aib-OH (0.873 g, 3.18 mM) oil was obtained. Crystalline compound (3.17 mM) was obtained by recrystallization with ethyl acetate and petroleum ether. Yield was 97%.

**D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>.** The linear peptide was assembled via standard solid-phase peptide synthesis with two couplings at each step. At each step, the procedure included the following: (1) deprotection with 50% TFA/CH<sub>2</sub>Cl<sub>2</sub> (1 × 2 min; 1 × 25 min); (2) washing the resin (CH<sub>2</sub>Cl<sub>2</sub>; 5 × 1 min); (3) neutralizing the resin with 5% DIEA/CH<sub>2</sub>Cl<sub>2</sub> (2 × 5 min); (4) washing the resin with CH<sub>2</sub>Cl<sub>2</sub> (3 × 1 min) and with DMF (3 × 1 min); (5) coupling of 3 equiv of Boc amino acid (3 equiv of HATu/6 equiv of DIEA with 1 equiv of resin) twice for 2 h, while for the Boc-Ala-Aib-OH segment, coupling was performed twice (8 h, 2 h); and (6) washing the resin with DMF (2 × 1 min), 2-propanol (2 × 1 min), CH<sub>2</sub>Cl<sub>2</sub> (2 × 1 min), and DMF (3 × 1 min). The linear peptide was cleaved from the solid support using HF containing 5% anisole at 0 °C for 1 h. The peptide was purified according to the general procedure with the B gradient of 5–30% during

25 min. The retention time was 10.88 min. The identity of the linear peptide was confirmed by mass spectroscopy ( $M + 1 = 415$ ). The overall yield was 20% based on the initial loading of the polymer.

**cyclo(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>).** The peptide was cyclized in DMF, using 2 equiv of Bop, Hobt and 4 equiv of DIEA in dilute solution 1 mg/1 mL. The peptide was purified according to a general procedure with the B gradient of 5–30% during 25 min. The retention time was 14.90 min. The identity of the cyclic peptide was confirmed by mass spectroscopy ( $M + 1 = 396.8$ ). The overall yield was 20% relative to the linear peptide.

**NMR Measurements.** NMR experiments have been carried out at 300 K using a Bruker Avance DRX 500 (500.13 /125.76 MHz for <sup>1</sup>H/<sup>13</sup>C) spectrometer equipped with a 5 mm triple-resonance probe (<sup>1</sup>H/<sup>13</sup>C/<sup>15</sup>N) and an actively shielded z-gradient coil. The sample contained 10 mg of *cyclo*(D-Pro<sup>1</sup>-Ala<sup>2</sup>-Ala<sup>3</sup>-Aib<sup>4</sup>-Ala<sup>5</sup>), dissolved in 0.5 mL of DMSO-*d*<sub>6</sub>. The <sup>1</sup>H NMR spectrum remained the same after 10-fold dilution, which excludes the occurrence of peptide self-aggregation. Chemical shifts are referenced to the residual solvent signal (for <sup>1</sup>H,  $\delta_{\text{DMSO-}d_6} = 2.49$  ppm and for <sup>13</sup>C,  $\delta_{\text{DMSO-}d_6} = 39.5$  ppm).

All <sup>1</sup>H and <sup>13</sup>C NMR data used in the present study were extracted from 1D and 2D experiments, respectively. The 2D TOCSY experiment<sup>50</sup> was run using a MLEV 17 sequence<sup>47</sup> for isotropic mixing with a duration of 60 ms. Spin-lock pulses with simultaneously switched gradients<sup>49</sup> were applied to generate pure absorption signals for coupling constant measurement. The 2D data matrix consisted of 4K × 512 complex data points. Zero-filling in  $F_1$  and a squared cosine function in both  $F_1$  and  $F_2$  were applied prior to Fourier transformation. Eight transients were accumulated for each of the  $t_1$  increments with a relaxation delay of 2 s. A spin-lock field of 8300 Hz was used for the TOCSY transfer.

ROESY spectra were recorded for different mixing times (60 and 120 ms) using the conventional ROESY experiment<sup>57,58</sup> with a CW spin-lock field of 3300 Hz. In addition, a small-flip angle (30°), offset compensated experiment of Nagayama,<sup>59</sup> and the recently proposed off-resonance ROESY<sup>60,61</sup> (axis tilt angle  $\Theta = 55^\circ$ ) were also carried out at each mixing time. A relaxation delay of 2 s was allowed between successive transients. Thirty-two transients were recorded with 2K complex data points each for a total number of 512 experiments. For processing the matrices were zero filled and apodized by a squared cosine function in both dimensions. A polynomial baseline correction was also applied. The HSQC proton–carbon correlation map<sup>51–53</sup> was recorded using the standard Bruker pulse sequence. Thirty-two scans were collected for each of the 256 experiments. A relaxation delay of 2 s was allowed and 2 K complex data points were acquired in  $F_2$ . Zero-filling and apodization was performed as indicated above. HMBC correlations allowed an unambiguous assignment of all quaternary carbons including the C- $\alpha$  of Aib<sup>4</sup> and also provided additional support of amino acid sequence. The gradient HMBC experiment<sup>62</sup> was performed allowing 70 ms for long-range coupling evolution. Sixteen scans were accumulated for each of the 512 experiments and 2 K data points were acquired in the acquisition domain.

**Molecular Modeling.** Energy calculations for all cyclic peptides were performed employing the ECEPP/2 potential field<sup>38,39</sup> assuming rigid valence geometry with planar trans-peptide bonds. Both trans and cis conformations were examined for peptide bonds in the D-proline residue. In this case, the  $\omega$  angle was allowed to vary. Aliphatic and aromatic hydrogens were generally included in united atomic centers of CH<sub>n</sub> type; H $^\alpha$ -atoms and amide hydrogens, as well as H $^\delta$ -atoms in D-Pro, were described explicitly. All calculations were performed with the value of the dielectric constant  $\epsilon = 45$  (the macroscopic  $\epsilon$  value for DMSO) to mimic to some extent solution effects (see also ref 35).

The calculation scheme involved estimation of C $^\alpha$ –C $^\alpha$  distances that were  $\leq 5$  Å for all possible combinations of local minima for the peptide backbone in a L<sup>4</sup>-L<sup>5</sup>-D<sup>1</sup>-L<sup>2</sup>-L<sup>3</sup> sequence. These local minima were the energetic minima in the Ramachandran map of E, F, C, A, and A\* types (according to the notation in ref 63) for all L-residues; of E\*, F\*, C\*, A, and A\* types for all D-residues; of E\*, F\*, C\*, A, E, F, C, and A\* types for the Gly residue; and of F\*, C\*, and A\* types for D-Pro. Conformers selected at the first step were subjected to energy minimization; the cycle closing was ensured by employing the two

parabolic potential functions keeping together the C<sub>3</sub>' and C<sup>α</sup><sub>3</sub> atoms at the C-terminus of the molecule to their dummy equivalents at the N-terminus. The side chain dihedral angle values were optimized before energy minimization to achieve their most favorable spatial arrangements, employing an algorithm described previously.<sup>64</sup>

**Acknowledgment.** The authors are grateful to Dr. Milo Budešinsky (Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic) for sharing the data on CPP structures and to Dr. Anders Berglund for valuable

help in running the MacroModel program. This work was supported in part by NIH grant RO1 GM48184 and the Washington University NIH Mass Spectrometry Resource. K.E.K. thanks the Hungarian National Research Foundation for support (OTKA T 029089 and OTKA D 23749). The purchase of the spectrometer used in the study was supported by OMF B Mec-93-0098, Phare-Accord H-9112-0198, and OTKA A084.

JA991728M