Journal of Computer-Aided Molecular Design, 7 (1993) 241–250
ESCOM

J-CAMD 196

SCF-MO study of the polyglycine II structure

Carlos Alemán and Juan J. Perez*

Departament d'Enginyeria Química, UPC, E.T.S. d'Enginyers Industrials, Av. Diagonal 647, E-08028 Barcelona, Spain

Received 15 May 1992
Accepted 17 November 1992

Key words: Polyglycine II; Packing effects; Conformational study; Hydrogen bond; AM1

SUMMARY

In order to get insight into the conditions that make polyglycine (PG)II a stable structure, the conformational features of three model molecules closely related to the PGII conformation were investigated. The model molecules selected were glycine dipeptide (AGN), glycine tripeptide (AGGN), and glycine tetrapeptide (AGGGN). Environmental effects were mimicked by means of formaldehyde molecules. The calculations were carried out at the SCF semiempirical level, using the AM1 method. The calculations show that of the three systems considered, only the AGGGN molecule presents a minimum energy conformation which corresponds to a PGII structure. The environmental conditions in which this conformation is found were also analyzed.

INTRODUCTION

It is well established that packing effects play an important role in determining the conformations adopted by molecules in crystals. For small molecules, this has been well documented in the past years [1]. In the case of medium and large molecules, although crystal environmental effects may affect the conformations exhibited by globular proteins, they are particularly important in peptides [2] and fibrous proteins [3].

Conformations found in the solid state are usually understood to be accessible energy minima of the molecule that do not necessarily correspond to any of its preferred conformations in vacuo or in solution [2]. Thus, the conformations adopted by a molecule in the crystal state are low-energy structures which result from the balance between inter- and intra-atomic contributions. However, in addition to producing changes in the energy ordering of the different minima, environmental effects may change the topology of the potential energy surface, thereby generating new minimum energy structures [4].

Polyglycine (PG) is an example of a molecule where packing effects play an important role in determining its accessible conformations. In the solid state the molecule presents a characteristic

*To whom correspondence should be addressed.

helical conformation referred to as polyglycine II (PGII). The structure was first proposed by Crick and Rich in 1955 from X-ray diffraction studies [5]. The crystal can be described as composed of 3_1-type helices packed in an hexagonal arrangement, yielding neighboring chains 4.79 Å apart. The packing is stabilized by a network of intermolecular hydrogen bonds in six directions perpendicular to the chain axis, 60° apart. The helix repeat is 9.3 Å (3.1 Å by each glycine residue) and the values for φ and ψ torsional angles proposed by Crick and Rich were −77° and 145°, respectively. However, as the PGII structure was determined with poor resolution, the values for φ and ψ were later standardized to −80° and 150° by different authors [6,7].

Several conformational studies on both the glycine residue and its dipeptide, i.e. the acetyl, *N'*-methylamide derivative, have been carried out in vacuo, using either quantum mechanics [8–15] or empirical force-field [10,16] calculations. Two important observations can be drawn from these studies. First, the topology of the conformational energy surface is clearly dependent on the computational method used. Thus, different numbers of minimum energy structures as well as their location on the map are found depending on whether ab initio, semiempirical or empirical methodologies are used. Moreover, differences are also found within a given method, since different results are obtained using different basis sets, semiempirical approximations, or force-field potentials. This points out the delicate balance between the different contributions to the intramolecular energy. Second, the PGII torsional angles have never been reported as an energy minimum on the φ–ψ conformational map.

Some calculations have also been made on PG chains. Molecular orbital calculations of one chain found the α-helix conformation to be the most stable [17]. These results are consistent with those previously published by Momany et al. [18], who, using molecular mechanics, also predicted the α-helix conformation as the lowest energy structure for a single strand of PG. On the other hand, calculations performed at the same level of theory considering two [19] or three [20] PG chains reproduced the experimental geometries of the PGII structure.

Although the structural parameters of PGII are satisfactorily reproduced by force-field calculations, there are still some questions that remain open. Packing interactions play an important role in the stabilization of this conformation, but the interplay between inter- and intramolecular interactions on the stability of the structure is not clear. The present paper addresses the questions of how environmental effects influence the accessible conformations of PGII and what are the conditions which determine whether the experimental structure has a minimum potential energy surface.

METHODS AND COMPUTATIONS

To address the problem, conformational studies of glycine and of two glycine oligomers in vacuo and including crystal environmental effects were carried out.

The calculations were performed using a semiempirical quantum-mechanical method. Owing to the size of the molecular systems studied in the present work, ab initio calculations using extended basis sets are not computationally affordable nowadays. So, because of the importance of hydrogen-bonding interactions, the Austin Model 1 (AM1) semiempirical method [21] was selected for the present study, and all the calculations were carried within this framework. This is a well-known method which provides molecular geometries very similar to those provided by ab initio calculations [22], as has been reported in conformational studies on glycine residues [23]. In

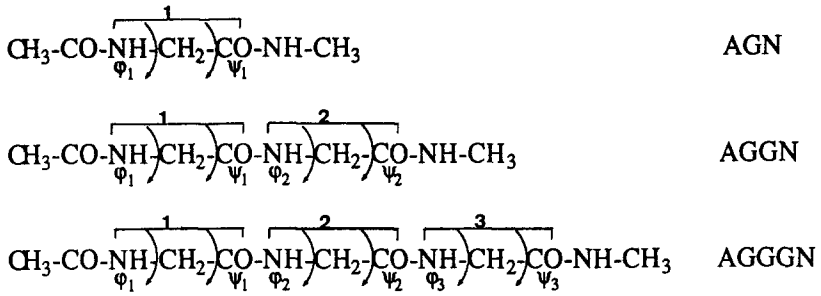

Fig. 1. Model molecules selected for the present study: the glycine dipeptide (AGN); the glycine tripeptide (AGGN); and the glycine tetrapeptide (AGGGN). The hydrogen bonds formed with environment molecules for each molecule are also shown.

addition, the AM1 method provides a good description of hydrogen-bonding interactions [24–27], which are of enormous importance for the present investigations. Different studies have shown the suitability of the method for this kind of study [28,29]. The advantage of this methodology is that all the effects due to mutual polarization of the molecules involved in the interaction or even due to conformational changes are included at this level of the theory.

In order to assess the reliability of the AM1 method for studying the accessible conformations of glycine, the fully relaxed conformational energy map of the glycine dipeptide (*N*-acetyl-*N'*-methylamide-glycine) (AGN) was computed as a function of the dihedral angles, φ–ψ, using a grid of 20° and in vacuo. Standard first-order optimization methods were used to locate all stationary points on the map.

The helical preferences of PG in vacuo were investigated further by studying the accessible conformations of glycine oligomers with two and three residues in the vicinity of the 3_1 helical region of the Ramachandran map. Model molecules selected for this purpose were the glycine tripeptide (AGGN) and the glycine tetrapeptide (AGGGN), in which the glycine chains were also blocked with an acetyl group at the N-terminus and with a methylamide group at the C-terminal end to mimic the progression of the polypeptide chain of the molecule. Figure 1 depicts the three model molecules whose conformational space was investigated.

The crystalline environment was simulated by surrounding the amide groups of the different glycine oligomers with formaldehyde molecules, which form hydrogen bonds. This way of simulating the crystal environment has been successfully used by our group for studying the conformational preferences of a related class of polyamides [30]. The accessible conformations in the vicinity of the 3_1 helical region of the three model molecules of Fig. 1 were re-investigated under these conditions.

All the calculations were carried out with a documented version [31] of the MOPAC computer program [32], using the standard parameters [23]. The calculations were carried out on an IBM 3090/600 of the 'Centre de Supercomputacio de Catalunya' (CESCA).

RESULTS AND DISCUSSION

The φ–ψ conformational energy map of AGN is depicted in Fig. 2. Considering only the asymmetric part of the plot, four minima were found, which are shown with filled circles on the

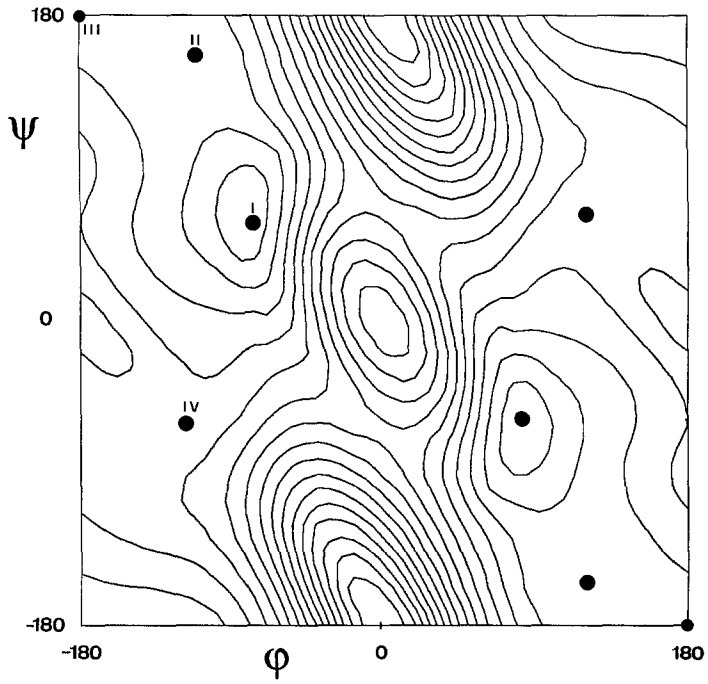

Fig. 2. Conformational map of the *N*-acetyl-*N'*-methylamide-glycyl molecule. Stationary points are shown with a full circle (see also Table 1). Contour lines are spaced 1 kcal/mol.

map and listed in Table 1. These minima were characterized as such after diagonalization of the Hessian matrix by showing all their positive eigenvalues. The lowest energy conformation is a C_7 structure, the second one is located in the β-sheet region of the Ramachandran map, the third low-energy conformation is a C_5 structure, and the last one is found in a flat region close to the α-helix. From ab initio calculations only minima I and III are found, suggesting that minima II and IV should be labeled as spurious. Furthermore, when a STO-3G or 4-21G basis set is used, the same relative stability between conformations I and III is reported [10,11]. Moreover, this behaviour is also found in force-field calculations [10,16]. Only when ab initio calculations are performed with more extended basis sets, of the quality of a double zeta plus polarization or

TABLE 1
DIHEDRAL ANGLES OF THE MINIMA FOUND FOR THE GLYCINE DIPEPTIDE (AGN) WITH THE AM1 METHOD

Minima	ω_1	φ	ψ	ω_2	ΔH_f	ΔE
I	178.1	−80.9	62.6	176.2	−86.1	0.0
II	170.1	−108.1	156.0	−179.6	−84.2	1.9
III	−179.8	−179.6	179.7	180.0	−82.9	3.2
IV	−173.5	−121.9	−59.0	−175.4	−81.8	4.3

ΔH_f and ΔE correspond to the heats of formation (kcal/mol) and relative energies (kcal/mol), respectively.

higher [14,15], the order of the lowest energy minima is reversed, being the C_7 conformation about 0.5 kcal/mol higher than that of C_5.

Next, we proceeded to evaluate the energy of the model molecules AGN, AGGN and AGGGN in PGII conformation. Then, the geometries of model molecules were fully optimized starting from these structures. After minimizations, all the structures were found to have a helical nature. This can be characterized either by the number of residues per turn (n) and the rise per residue (h) or by the dihedral angles, φ and ψ, the former being easily calculated from the latter [33]. Table 2 displays the torsional angles, φ and ψ, and the helical parameters, h and n. Furthermore, the heats of formation for the different model molecules in both conformations are also shown in Table 2.

These results show that the φ and ψ angles of the PGII conformation do not correspond to a stationary point of the conformational potential surface and that, after full geometry optimization, the molecules (i.e. AGN, AGGN and AGGGN) achieve minimum II in the AGN Ramachandran plot.

In order to study the effect of the crystal environment on the preferred conformations of PG, the model molecules were surrounded by formaldehyde molecules. Previous studies conducted in this laboratory [30] suggest that formaldehyde molecules produce an effective hydrogen-bonding environment for different classes of polyamides. Each surrounding formaldehyde molecule forms two hydrogen bonds with the model molecule: one geometrically more favorable, between the oxygen atom of formaldehyde and the N–H group of the peptide group; and a second one, between one of the hydrogens of the formaldehyde molecule and the C=O group of the following peptide group. In this way, each formaldehyde molecule saturates two directions of hydrogen bonding of the model molecule. As the PGII structure presents six directions of hydrogen bonding, three formaldehydes will be necessary per PG turn. Figure 1 shows the scheme of hydrogen bonds for each of the model molecules considered in the present work.

The positions of the formaldehydes around the model molecules were established in a stepwise manner. First, a formaldehyde molecule was placed in each of the regions suitable for hydrogen bonding formation with PG. Under these conditions the geometry of the system was optimized, constraining the dihedral angles of the peptide at fixed values of the PGII conformation. Next, a grid of calculations was performed in which the O...H distance between the two interacting atoms

TABLE 2
DIHEDRAL ANGLES, HELICAL PARAMETERS AND HEATS OF FORMATION IN KCAL/MOL OF AGN, AGGN AND AGGGN MODEL MOLECULES WITHOUT CONSIDERING ENVIRONMENTAL EFFECTS

Molecules		φ_1	ψ_1	φ_2	ψ_2	φ_3	ψ_3	h	n	ΔH_f
AGN	(i)	−80	150	–	–	–	–	3.1	3.0	−83.6
	(ii)	−108.6	154.0	–	–	–	–	3.4	2.6	−84.2
AGGN	(i)	−80	150	−80	150	–	–	3.1	3.0	−119.6
	(ii)	−111.5	158.5	−108.4	157.2	–	–	3.4	2.6	−120.6
AGGGN	(i)	−80	150	−80	150	−80	150	3.1	3.0	−155.3
	(ii)	−113.4	155.5	−112.4	157.6	−111.2	157.4	3.4	2.6	−157.0

For each molecule, row (i) corresponds to the energy associated with a PG II conformation and row (ii) is the conformation adopted after full relaxation of the geometry.

was varied by increments of 0.15 Å and the angle O...H–N was changed by 10°. Once the lowest energy hydrogen-bonding complex was identified, a final gradient optimization was performed under the same conditions, in one step. After it, we proceeded to a full geometry optimization.

The results of these minimizations are shown in Table 3. The final positions of the formaldehyde molecules have also been included in the Table for completeness. Interaction energies between the model molecule and the environment (ΔE_{int}) were calculated according to supermolecule approach, based on the fact that the interaction between environmental formaldehydes is negligible. These values are shown in Table 3.

Optimization of the conformation of the AGN model molecule within the environmental shows a deviation with respect to the starting geometry. The φ_1 torsional angle deviation is about 20° although, ψ_1 is approximately kept in its 3_1 value, exhibiting a small deviation of 4°. The interaction energy is 5.4 kcal/mol. This value can be assigned to the formation of a hydrogen bond between the model molecule and the surrounding formaldehyde molecule. Similar values have been obtained using a 6–31G' basis set, and can be considered typical for this type of hydrogen bonds [31]. As a result of the minimization process the molecule is taken to the transition structure II, which corresponds precisely to the same point as in vacuo.

When the AGGN model molecule is minimized in the presence of two formaldehyde molecules, the results are very similar to those found for AGN. The torsional angles, φ_i, deviate about 25° from their initial values and the ψ_i values are maintained close to their initial values. The minimizer takes the molecule again to the structure II, whose helical parameters indicate that the structure corresponds to an extended helix with only 2.5 residues per turn close to the one obtained without environmental effects. The interaction energy is −10.2 kcal/mol which is just twice the value found for AGN.

In contrast to the results presented above, optimization of the AGGGN model molecule with three formaldehyde molecules, yields the PGII conformation as a stable minimum energy structure. The angles φ_i and ψ_i are −86° and 150°, and the corresponding helical parameters are n = 2.9 and h = 3.2 Å which compare well with those of the PGII structure, n = 3 and h = 3.1 Å. The interaction energy is −15.1 kcal/mol, which is precisely three times the interaction of

TABLE 3
CONFORMATIONS OF THE MODEL MOLECULES AGN, AGGN, AGGGN AND RELATIVE POSITIONS OF THE SURROUNDING FORMALDEHYDE MOLECULES (DISTANCE AND ANGLE OF THE HYDROGEN BOND BETWEEN FORMALDEHYDE AND AMIDE GROUP) AFTER FULLY RELAXED OPTIMIZATION

Molecules	φ_1	ψ_1	φ_2	ψ_2	φ_3	ψ_3	h	n	d(O...H)	< OHN	ΔE_{int}
AGN	−101.7	146.5	–	–	–	–	3.3	2.6	2.19	144.2	−5.4
AGGN	−106.4	152.6	−108.9	144.8	–	–	3.4	2.5	2.19	153.9	−10.2
									2.17	141.6	
AGGGN	−87.1	150.0	−85.4	150.1	−85.8	150.0	3.2	2.9	2.22	147.9	−15.1
									2.18	144.3	
									2.16	143.8	

Interaction energies (ΔE_{int}) in kcal/mol are also included.

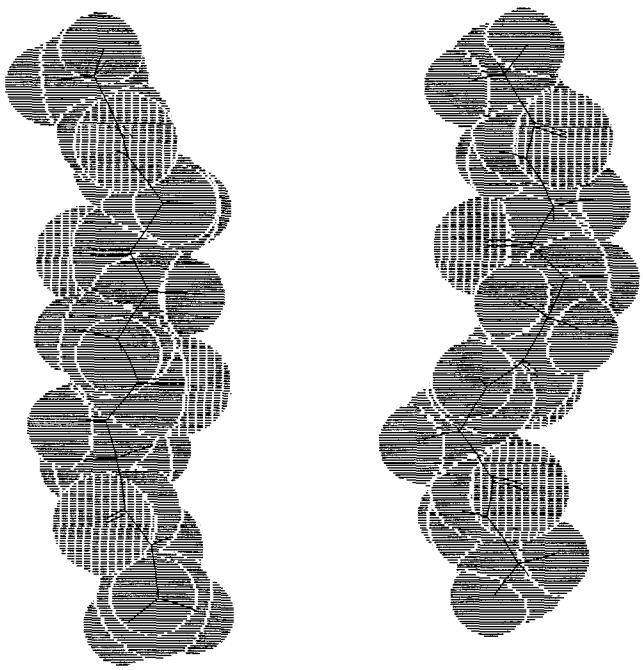

Fig. 3. Comparison between the AGGGN model molecule optimized with environment (left) and without environment (right).

AGN...HCHO, as can be seen in Table 3. Comparison between the structures of AGGGN minimized with and without environment is shown in Fig. 3.

In order to gain further insight into the role of the environment, we carried out the same minimizations for AGGN and AGGGN with a smaller number of saturating formaldehyde molecules. In this case the minimizations were carried out with only one formaldehyde molecule for AGGN and three calculations were performed for AGGGN: a first one with one formaldehyde in the N-terminal glycine residue, the second one with one formaldehyde in the middle glycine residue and finally, the third one with two formaldehyde molecules. The results are shown in Table 4. The first and the second calculations provide very similar results and therefore, only averaged values are reported in Table 4. Final geometries and the positions of the formaldehyde

TABLE 4
SAME AS IN TABLE 3, BUT USING ONLY ONE FORMALDEHYDE MOLECULE AS ENVIRONMENT FOR AGGN, AGGGN AND AGGGGN, RESPECTIVELY

Molecules	φ_1	ψ_1	φ_2	ψ_2	φ_3	ψ_3	φ_4	ψ_4	h	n	d(O...H)	< OHN	ΔE_{int}
AGGN	−107.1	145.6	−111.1	160.0	–	–	–	–	3.4	2.5	2.21	146.4	−5.1
AGGGN	−85.0	149.7	−85.1	150.3	−85.1	150.0	–	–	3.2	2.9	2.21	146.4	−5.1
AGGGGN	−85.8	150.0	−85.0	150.1	−85.1	150.1	−84.6	−149.8	3.2	2.9	2.21	146.0	−5.1

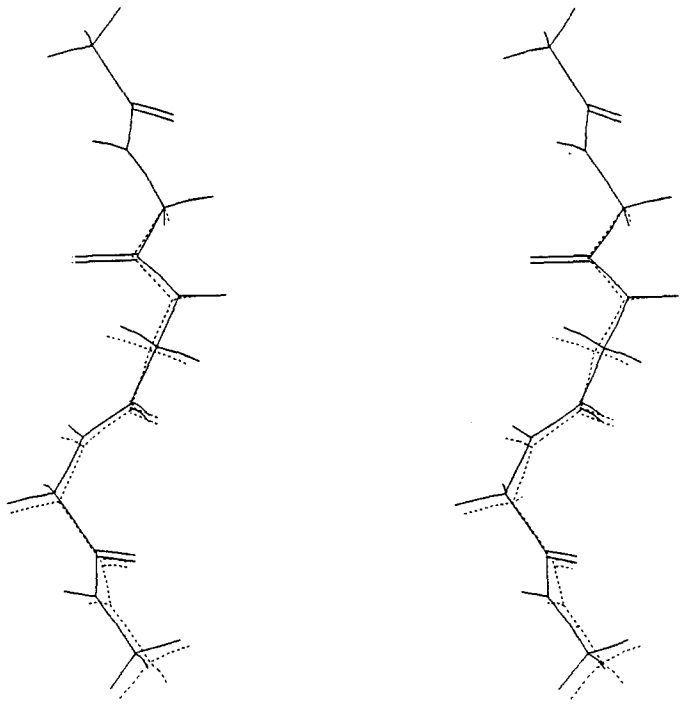

Fig. 4. Stereo diagram of the overlap between AGGGN model molecule optimized with one formaldehyde and a PGII conformation.

molecules are also included. The calculations showed surprisingly that the results do not change with respect to those of Table 3. The structure obtained for AGGGN using one formaldehyde molecule and crystallographic PGII are compared and depicted in Fig. 4.

From the results described above it appears that formaldehyde environmental molecules do not provide a change on the conformational behaviour of AGN and AGGN model molecules. Thus, starting from a PGII structure in both situations with and without environment, the molecules end up in a distorted structure after geometry relaxation. In contrast, the PGII structure is stable for the AGGGN molecule, even when the environment is mimicked with only one formaldehyde molecule.

In order to demonstrate that the present results were not due to an artefact of the method used, the model molecule AGGGGN was also calculated. The isolated molecule went to the same conformation of AGN, AGGN and AGGGN, starting from the PGII structure. According to the result discussed, a single molecule of formaldehyde was considered to mimic the environment. This calculation was carried out independently for a formaldehyde molecule placed at each of the four residues of glycine of the model molecule. The results show a consistency in getting a PGII structure regardless of the position of the formaldehyde. Averaged values are included in Table 4.

Unfortunately, quantitative analysis of the results does not provide us with further information. So, if we compare the heats of formation shown in Tables 2 and 3, an extra stabilization due to intramolecular interactions does not seem to arise when three residues of glycine in the AGGGN model molecule are considered. On the other hand, the fact that hydrogen bonds

formed with formaldehydes show an additive behaviour indicates that no extra stabilization is obtained because the hydrogen bonding with AGGGN is more effective. In summary, no special inter- or intra-molecular interactions are shown from the present calculations as responsible for the stability of the PGII structure of AGGGN compared with other molecules.

CONCLUSIONS

The conformational space of AGN, AGGN and AGGGN model molecules has been studied with and without environmental effects included. These effects have been mimicked using formaldehyde molecules, which form hydrogen bonds. When the model molecules were minimized, neither AGN nor AGGN stayed in a PGII structure, even when environmental effects were included. However, AGGGN exhibits a PGII conformation after minimization only when environmental effects are included.

ACKNOWLEDGEMENTS

C.A. acknowledges financial support from the Departament d'Ensenyament de la Generalitat de Catalunya. This work has been supported in part by CICYT Grant Number MAT 89-1136-CE.

REFERENCES

1 Bernstein, J., In Domenicano, A. and Hargittai, I. (Eds.) Accurate Molecular Structures, Oxford University Press, Oxford, 1992.
2 a. Fernandez, Y., Ricart, J.M. and Perez, J.J., Int. J. Biol. Macromol., 14 (1992) 154.
b. Perez, J.J., Villar, H.O. and Loew, G.H., J. Comput.-Aided Mol. Design, 6 (1992) 175.
3 Alemán, C., Subirana, J.A. and Perez, J.J., Biopolymers, 32 (1992) 631.
4 Ooi, T., Oobatake, M., Nemethy, G. and Scheraga, H.A., Proc. Natl. Acad. Sci. USA, 84 (1987) 3086.
5 Crick, F.H.C. and Rich, A., Nature, 176 (1955) 780.
6 Ramachandran, G.N. and Sasisekharan, V., Adv. Protein Chem., 23 (1968) 283.
7 Maigret, B., Pullman, B. and Perahia, D., J. Theor. Biol., 31 (1971) 269.
8 Vishreshwara, S. and Pople, J.A., J. Am. Chem. Soc., 99 (1977) 2422.
9 Peters, D. and Peters, J., J. Mol. Struct., 53 (1979) 103.
10 Palla, P., Petrongolo, C. and Tomasi, J., J. Phys. Chem., 84 (1980) 435.
11 Schäfer, L., Van Alsenoy, C. and Scarsdale, J.N., J. Chem. Phys., 76 (1982) 1439.
12 Luke, B.T., Gupta, A.G., Loew, G.H., Lawless, J.G. and White, D.H., Int. J. Quantum Chem., Quantum Biol. Symp., 11 (1984) 117.
13 Klimkowski, V.J., Schäfer, L., Momany, F.A. and Van Alsenoy, C., J. Mol. Struct. (Theochem), 124 (1985) 143.
14 Head-Gordon, T., Head-Gordon, M., Frisch, M.J., Brooks III, C.L. and Pople, J.A., J. Am. Chem. Soc., 113 (1991) 5989.
15 Jensen, J.H. and Gordon, M.S., J. Am. Chem. Soc., 113 (1991) 7917.
16 Lau, W.F. and Pettit, B.M., Biopolymers, 26 (1987) 1817.
17 a. Kleier, D.A. and Lipscomb, W.N., Int. J. Quantum Chem., 4 (1977) 73.
b. Scheiner, S. and Kern, C.W., J. Am. Chem. Soc., 100 (1978) 7539.
c. Scheiner, S. and Kern, C.W., Proc. Natl. Acad. Sci. USA, 75 (1978) 75.
18 Momany, F.A., McGuire, R.F., Burgess, A.W. and Scheraga, H.A., J. Phys. Chem., 79 (1975) 2361.
19 a. Hopfinger, A.J., Biopolymers, 10 (1971) 1299.
b. Balaji, V.N., Int. J. Quantum Chem., 20(1981) 347.

20 Nakamura, H. and Go, N., Int. J. Pept. Protein Res., 25 (1985) 232.
21 Dewar, M.J.S., Zoebisch, E.G., Healy, E.F. and Stewart, J.J.P., J. Am. Chem. Soc., 107 (1986) 3902.
22 a. Topiol, S., J. Comput. Chem., 8 (1987) 142.
b. Bofill, J.M., Castells, J., Olivella, S. and Solé, A., J. Org. Chem., 53 (1988) 5148.
23 Masamura, M., J. Mol. Struct. (Theochem), 164 (1988) 299.
24 Dannenberg, J.J. and Vinson, L.K., J. Phys. Chem., 92 (1988) 5635.
25 Dannenberg, J.J., J. Phys. Chem., 92 (1988) 6869.
26 Galera, S., Lluch, J.M., Oliva, A. and Bertran, J., J. Mol. Struct. (Theochem), 40 (1988) 101.
27 Vinson, L.K. and Dannenberg, J.J., J. Am. Chem. Soc., 111 (1989) 2777.
28 Novoa, J.J. and Whangbo, M., J. Am. Chem. Soc., 113 (1991) 9017.
29 Palmer, D.E., Pattaromi, C., Nunami, K., Chadha, R.K. and Goodman, M.J., J. Am. Chem. Soc., 114 (1992) 5634.
30 Alemán, C. and Perez, J.J., J. Mol. Struct., submitted.
31 Olivella, S., QCPE Bull., 4 (1984) 10.
32 Stewart, J.J.P., QCPE Bull., 3 (1983) 431.
33 Fasman, G.D., Poly(α-amino acids): Protein Models for Conformational Studies (Biological Macromolecules Series), Marcel Dekker, New York, NY, 1967.