Understanding the Role of Stereoelectronic Effects in Determining Collagen Stability. 1. A Quantum Mechanical Study of Proline, Hydroxyproline, and Fluoroproline Dipeptide Analogues in Aqueous Solution

Roberto Improta, Caterina Benzi, and Vincenzo Barone*

Contribution from the Dipartimento di Chimica, Università Federico II, Complesso Universitario Monte S. Angelo, Via Cintia, I-80126 Napoli, Italy

Received March 5, 2001. Revised Manuscript Received September 17, 2001

Abstract: The importance of local (intraresidue) effects in determining the stability of the collagen triple helix has been investigated with special reference to the role played by hydroxyproline. To this end the dipeptide analogues of L-proline (ProDA), 4(*R*)-hydroxy-L-proline (HypDA), and 4(*R*)-fluoro-L-proline (FlpDA) have been studied by means of quantum mechanical ab initio calculations, taking into account solvent effects by the polarizable continuum model (PCM). Our results confirm that the relative stability of up puckerings of the pyrrolidine ring increases with the electronegativity of the 4(*R*) substituent (X), whereas down puckerings are favored by 4(*S*) electronegative substituents. Calculations on model compounds show that this effect is due to the interaction between vicinal C−H bonding and C−X antibonding orbitals. Electronegative substituents on the pyrrolidine ring affect cis−trans isomerism around the peptidic bond, with trans isomers stabilized by 4(*R*) substituents and cis isomers by 4(*S*) substituents. Also the hydrogen bonding power of the carbonyl moiety following the pyrrolidine ring is affected by 4(*R*) substituents, but this effect is tuned by the polarity of the embedding medium. Finally, up puckering favors smaller values of the backbone dihedrals ϕ and ψ. All these results strongly support the proposal that the stability of triple helices containing fluorinated or hydroxylated prolines in Y positions is related to the necessity of having up puckerings in those positions.

1. Introduction

Due to its crucial biological role (it is the most abundant protein in vertebrates)[1] and very peculiar structure, collagen has been deeply investigated, together with several related polypeptides.[1−13]

Collagen is composed by approximately 300 repeats of the sequence $X_{aa}Y_{aa}$Gly, where X_{aa} and Y_{aa} are in most cases L-proline (Pro) and 4(*R*)-hydroxy-L-proline (Hyp), respectively.[1,4] The occurrence of Pro and Hyp in collagen restricts the orientational freedom of the chain in relation to the fiber axis and permits only left-handed PII ($\phi \approx -60°$, $\psi \approx 150°$) helices.[1,2] Moreover, in connective tissues, three collagen chains intertwine themselves further by twisting around the common central axis, forming a so-called "coiled-coil structure". Thanks to its triple-helix structure, collagen exhibits a great tensile strength.[1]

Hyp is the only stereoisomer present in collagen and has an important role in the thermal stability of the fiber,[5] but only if placed in the Y position. On the other hand, the presence of 4(*S*)-hydroxy-L-proline (hyp)[6] in either the X or Y position or the presence of Hyp in the X position[7] inhibits the formation of the collagen triple helix. The presence of 4(*R*)-fluoro-L-proline (Flp) in place of Hyp in the Y position increases even more the stability of the triple helix,[8a] suggesting that the electronegativity of the 4(*R*) substituent could play a significant role.[8] High-resolution X-ray studies of the tridecapeptide $(Pro\text{-}Pro\text{-}Gly)_{10}$ (hereafter $(PPG)_{10}$)[9,10] have highlighted a strong correlation between backbone and ring geometrical parameters of iminoacid residues.[9] As a matter of fact, the pyrrolidine ring can adopt

* Corresponding author. E-mail: enzo@chemistry.unina.it.

(1) (a) *Collagen*; Nimni, M. E., Ed.; CRC Press: Boca Raton, FL, 1988; Vols. 1−4. (b) *Biochemistry of Collagen*; Ramachandran, G. N., Reddi, A. H., Eds.; Plenum Press: New York, 1976.

(2) Fraser, R. D. B.; MacRae, T. P.; Suzuki, E. *J. Mol. Biol.* **1979**, *129*, 463.

(3) (a) Engel, J.; Chen, H.-T.; Prockop, D. J.; Klup, H. *Biopolymers*, **1977**, *16*, 601. (b) Tumanyan, V. G.; Esipova, N. G. *Biopolymers* **1982**, *21*, 475. (c) Bansal, M.; Ananthanarayanan, V. S. *Biopolymers* **1988**, *27*, 299.

(4) Fietzek, P. P.; Kuhn, K. *Mol. Cell Biochem.* **1975**, *8*, 141.

(5) (a) Chopra, R. K.; Ananthanarayanan, V. S. *Proc. Natl. Acad. Sci. U.S.A.* **1982**, *79*, 7180. (b) Privalov, P. L. *Adv. Protein Chem.* **1982**, *35*, 1. (c) Berg, R. A.; Prockop, D. J. *Biochem. Biophys. Res. Commun.* **1973**, *52*, 115.

(6) Inouye, K.; Sakakibara, S.; Prockop, D. J. *Biochim. Biophys. Acta* **1976**, *420*, 133.

(7) Inouye, K.; Kobayashi, Y.; Kyogoku, Y.; Kishida, Y.; Sakakibara, S.; Prockop, D. J. *Arch. Biochem. Biophys.* **1982**, *219*, 198.

(8) (a) Holmgrem, S. K.; Taylor, K. M.; Bretscher, L. E.; Raines, R. T. *Nature* **1998**, *392*, 666. (b) Eberhardt, E. S.; Panasik, N., Jr.; Raines, R. T. *J. Am. Chem. Soc.* **1996**, *118*, 2261. (c) Holmgrem, S. K.; Bretscher, L. E.; Taylor, K. M.; Raines, R. T. *Chem. Biol.* **1999**, *6*, 63.

(9) (a) Kramer, R. Z.; Vitagliano, L.; Bella, J.; Berisio, R.; Mazzarella, L.; Brodsky, B.; Zagari, A.; Berman, H. M. *J. Mol. Biol.* **1998**, *280*, 623. (b) Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A. *Biopolymers*, in press. (c) Berisio, R.; Vitagliano, L.; Sorrentino, G.; Carotenuto, L.; Piccolo, C.; Mazzarella, L. *Acta Crystallogr., Sect. D* **2000**, *56*, 55.

(10) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem* **1998**, *124*, 1117.

(11) (a) Bella, J.; Eaton, M.; Brodsky, B.; Berman, H. M. *Science* **1994**, *266*, 75. (b) Bella, J.; Brodsky, B.; Berman, H. M. *Structure* **1995**, *3*, 893.

(12) Nagarajan, V.; Kamitori, S.; Okuyama, K. *J. Biochem.* **1999**, *125*, 310.

(13) Bretscher, L. E.; Jenkins C. L.; Taylor, K. M.; DeRider, M. L.; Raines, R. T. *J. Am. Chem. Soc.* **2001**, *123*, 777.

10.1021/ja010599i CCC: $20.00
Published on Web 11/17/2001

two distinct puckerings in which the C^{β} and the C^{γ} atoms are displaced from the mean plane of the ring. The two puckered forms are generally referred to as "exo" and "endo" (more precisely C^{γ}-exo and C^{γ}-endo) or "up" and "down", respectively.[14] In the solid state, collagen-like polypeptides are always characterized by Pro with down puckering in X positions,[9] and by Pro or Hyp with up puckering in Y positions.[9,11a] Unfortunately, there is no general agreement about the explanations of the experimental results, concerning both the molecular basis for the stability of the collagen triple helix and the role played by Hyp.

A first model suggests that collagen could be stabilized by a network of hydrogen bonds involving the hydroxyl groups of Hyp, water molecules, and backbone carbonyl groups.[11] Alternatively, the electron-withdrawing effect of the hydroxyl group could stabilize the trans configuration of the proline amide bond and, thus, the collagen triple helix. As a matter of fact, Flp enhances the stability of the triple helix of collagen-like polypeptides,[8,13] fluorine being a more electronegative group but a less effective proton acceptor than the oxydril moiety.[15] Furthermore, recent studies show that hyp and 4(*S*)-fluoro-L-proline (flp) decrease the preference for the trans configuration of the amide bond.[13] This hypothesis pointed out for the first time the importance of stereoelectronic effects and explained the need of 4(*R*)-substituents in the Y position for stabilizing collagen; however, it does not account for the destabilizing effect of Hyp in the X position.

Very recently, a quite convincing proposal has been put forward, namely that the formation of a triple helix would require for the X residues ϕ dihedral angles typical of down puckerings and for the Y residues ϕ values characteristic of up puckerings.[9b] Hydroxyproline could intrinsically prefer a up puckering, thus explaining its stabilizing effect when placed in the Y position. The preference of Hyp and Flp for up puckerings is confirmed by experimental data[16−18] and has been explained[8] in terms of the so-called gauche effect: gauche (synperiplanar) conformations are preferred over anti (antiperiplanar) conformations in an X−C−C−Y moiety, whenever X and Y are two electron-withdrawing substituents.[19]

We are well aware that collagen's stability is probably enhanced by a delicate balance between local/intraresidue and nonlocal/interresidue effects. However, an assessment of the role played by intraresidue effects represents a mandatory starting point. In this connection, a quantum mechanical approach is particularly appealing and is tackled in the present paper by means of an ab initio study of the dipeptide analogues (the residue with *N*-acetyl (Ac-) and *N*′-methylamide (−NHMe) terminal blocking groups, see Figure 1) of L-proline (ProDA), 4(*R*)-hydroxy-L-proline (HypDA), and 4(*R*)-fluoro-L-proline (FlpDA) (see Figure 2), of their 4(*S*) analogues (hypDA and flpDA, respectively), and of some related compounds. We will analyze in particular the following aspects:

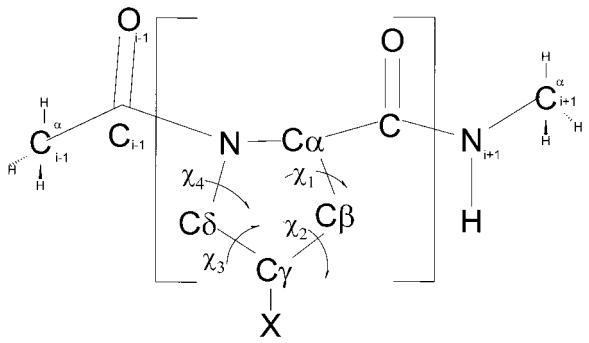

Figure 1. Atom labeling (X is the 4-substituent) and most relevant geometric parameters of the compounds under study.

(i) Is there any intrinsic preference of hydroxyproline and proline for up or down puckerings?

(ii) What is the origin of this preference?

(iii) How are the conformational equilibria of proline residues (e.g., cis−trans isomerism around the peptidic bond) in proteins and polypeptides influenced by the substituents on the pyrrolidine ring?

(iv) What is the relationship between the puckering of the pyrrolidine ring and the backbone geometrical parameters?

(v) What is the role of the chirality (*R*/*S*) of the C^{γ} atom?

(vi) How are these effects related to collagen's stability and, more generally, to the behavior of proline and hydroxyproline in polypeptides?

Since collagen and bioactive polypeptides operate in aqueous solution, it is of basic importance to consider environmental effects as well: calculations have been thus performed both in vacuo and in aqueous solution. Furthermore, since the different conformers are probably very close in energy, it is necessary to attain quite accurate results. Hartree−Fock (HF) calculations have thus been complemented by ad hoc calculations performed at the density functional and MP2 levels, to take into account electron correlation effects.

2. Methods

The geometries of ProDA, HypDA, and FlpDA have been fully optimized in aqueous solution at the Hartree−Fock (HF) level, using the standard 6-31G(d) basis set.[20] HF and MP2[20] energy calculations have been performed at these geometries using also the more extended 6-31+G(d,p) and 6-311+G(d,p) basis sets.[20]

Calculations based on density functional theory (DFT) have been performed using the PBE0[21] functional, which, despite the absence of adjustable parameters, has already shown an accuracy competitive with that of the best last-generation functionals.[21,22]

Solvent effects have been taken into account by means of the polarizable continuum model (PCM).[23,24] The CPCM variant used in the present work[25] employs conductor rather than dielectric boundary conditions, and this leads to a simpler and more robust implementation. Several studies have shown that for polar solvents the results obtained

(14) Némethy, G.; Gibson, K. D.; Palmer, K. A.; Yoon, C. N.; Paterlini, G.; Zagari, A.; Rumsey, S.; Scheraga, H. A. *J. Phys. Chem.* **1992**, *96*, 6472.
(15) Dunitz, J. D.; Taylor, R. *Chem. Eur. J.* **1997**, *3*, 89.
(16) Panasik, N., Jr.; Eberhardt, E. S.; Edison, A. S.; Powell, D. R.; Raines, R. T. *Int. J. Pept. Protein Res.* **1994**, *44*, 262.
(17) (a) Li, M.-H.; Fan, P.; Brodsky, B.; Baum, J. *Biochemistry* **1993**, *32*, 7377. (b) Dzakula, Z.; DeRider, M. L.; Markley, J. L. *J. Am. Chem. Soc.* **1996**, *118*, 12796.
(18) (a) Gerig, J. T.; McLeod, R. S. *J. Am. Chem. Soc.* **1973**, *95*, 5725. (b) Gerig, J. T.; McLeod, R. S. *J. Am. Chem. Soc.* **1976**, *98*, 3970.
(19) (a) Wolfe, S. *Acc. Chem. Res.* **1972**, *5*, 102. (b) Thibaudeau, C.; Plavec, J.; Garg, N.; Papchikhin, A.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **1994**, *116*, 4038. (c) O'Hagan, D.; Bilton, C.; Howard, J. A. K.; Knight, L.; Tozer, D. J. *J. Chem. Soc., Perkin Trans. 2* **2000**, 605.
(20) A description of basis sets and standard computational methods can be found in the following: Foresman, J. B.; Frisch, A. E. *Exploring Chemistry with Electronic Structure Methods*, 2nd ed.; Gaussian Inc.: Pittsburgh, PA, 1996.
(21) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.
(22) (a) Adamo, C.; Barone, V. *Chem. Phys. Lett.* **1998**, *298*, 113. (b) Adamo, C.; Scuseria, G. E.; Barone, V. *J. Chem. Phys.* **1999**, *111*, 2889. (c) Adamo, C.; Cossi, M.; Scalmani, G.; Barone, V. *Chem. Phys. Lett.* **1999**, *307*, 265. (d) Scalmani, G.; Bredas, J. L.; Barone, V. *J. Chem. Phys.* **2000**, *112*, 1178.
(23) (a) Miertuš, S.; Scrocco, E.; Tomasi J. *Chem. Phys.* **1981**, *55*, 117. (b) Amovilli, C.; Barone, V.; Cammi, R.; Cances, E.; Cossi, M.; Mennucci, B.; Pomelli, C. S.; Tomasi, J. *Adv. Quantum Chem.* **1998**, *32*, 227.
(24) Barone, V.; Cossi, M.; Tomasi, *J. Chem. Phys.* **1997**, *107*, 3210.
(25) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, *102*, 1995.

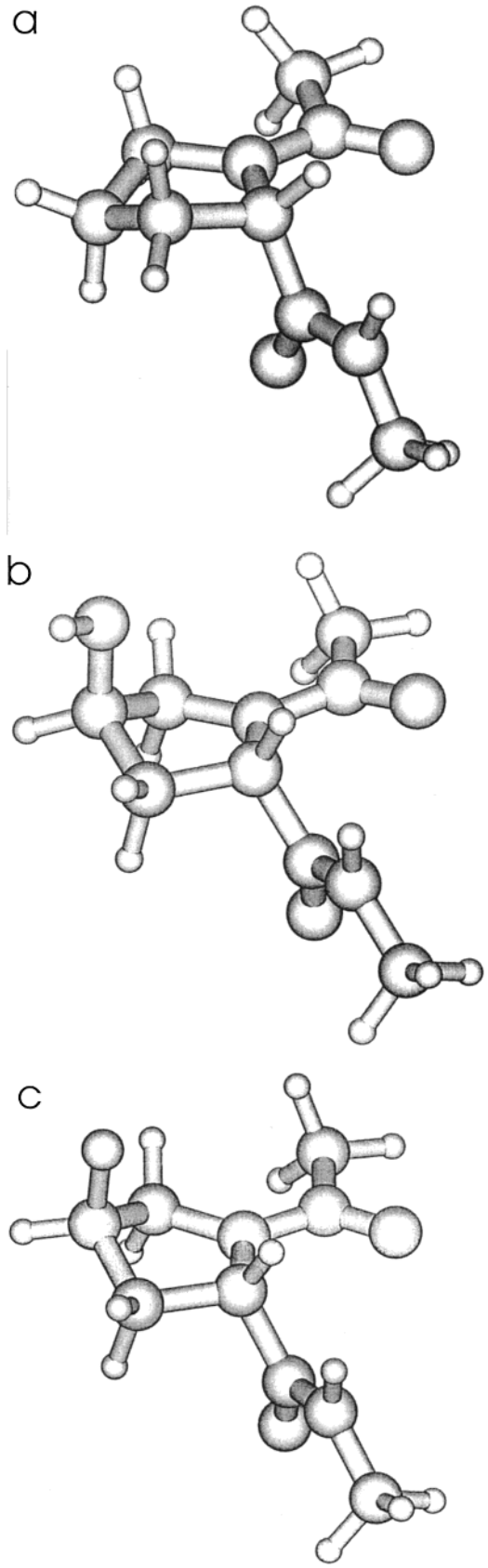

Figure 2. Minimum energy geometry of the trans isomer of (a) $ProDA_{down}$, (b) $HypDA_{up}$, and (c) $FlpDA_{up}$.

by different versions of PCM are very close and reproduce accurately the geometry of many biological molecules.[26]

Harmonic frequencies have been computed using our recent implementation of analytical CPCM second derivatives.[27]

Following the analysis introduced by Cremer and Pople,[28] ring puckerings have been described in terms of the displacements (z_i) of the atoms from the average plane of the ring. z_i's are related to two "global parameters", q (puckering amplitude) and ϕ (phase angle), describing the ring puckering.

$$z_i = q\cos(\phi + 4\pi(i-1)/5) \qquad i = 1, ..., 5 \qquad (1)$$

Ideal twist and envelope puckerings are characterized by $\phi = n \times 18°$, with n odd or even, respectively. The difference between the calculated ϕ and the value for the ideal twist (envelope), divided by 18, will then provide the percentage of twist (envelope) puckerings of a given pyrrolidine ring.

The natural bond orbital (NBO) analysis has been used to describe the electronic interactions between the pyrrolidine ring and its 4(R) substituent.[29] In this procedure, a block diagonalization of the density matrix allows the definition of natural atomic orbitals and their decomposition into bonding, Rydberg, and lone pair orbitals. After a suitable orthogonalization procedure, each bonding atomic orbital takes part in σ-bonding and σ*-antibonding localized molecular orbitals. Off-diagonal elements of the Fock matrix in the localized NBO basis thus provide a measure of the interactions among the different bonding, nonbonding, and antibonding orbitals.

All the calculations have been performed using a development version of the Gaussian package.[30]

3. Results

In a previous conformational study, we have shown that in aqueous solution ProDA and HypDA prefer a P−II conformation.[31] This result is in agreement with previous experimental determinations in the solid state[32] and in aqueous solution[33] and supports the reliability of our computational approach. Moreover, Pro and Hyp adopt a P−II conformation in collagen and in related polypeptides PPG_{10} and $(Pro\text{-}Hyp\text{-}Gly)_{10}$ (hereafter $(PHG)_{10}$). As a consequence, we will focus our attention on P−II conformers only.

Before starting our analysis, we have verified the reliability of the HF/6-31G(d) results with respect to the increase of the basis set size and to the inclusion of electron correlation. It would be indeed much less expensive to perform the bulk of our analysis at this level of the theory, using more sophisticated methods (MP2 and PBE0) only whenever it is important to

(26) (a) Jolibois, F.; Cadet, J.; Grand, A.; Subra, R.; Barone, V.; Rega, N. *J. Am. Chem. Soc.* **1998**, *120*, 1864. (b) Rega, N.; Cossi, M.; Barone, V. *J. Am. Chem. Soc.* **1998**, *120*, 5723. (c) Nielsen, P. A.; Norrby, P. O.; Liljefors, T.; Rega, N.; Barone, V. *J. Am. Chem. Soc.* **2000**, *122*, 3151.
(27) Cossi, M.; Barone, V. *J. Chem. Phys* **1998**, *109*, 6246.
(28) Cremer, D.; Pople, J. A. *J. Am. Chem. Soc.* **1975**, *97*, 1354
(29) (a) Foster, J. P.; Weinhold, F. *J. Am. Chem. Soc.* **1980**, *102*, 7211. (b) Reed, A.; Weinhold, F. *J. Chem. Phys.* **1983**, *78*, 4066. (c) Glendening, E. D.; Weinhold, F. *J. Comput. Chem.* **1998**, *19*, 593.
(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Adamo, C.; Jaramillo, J.; Cammi, R.; Pomelli, C.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 99*, Development Version (Revision C.01); Gaussian, Inc.: Pittsburgh, PA, 2000.
(31) Benzi, C.; Improta, R.; Scalmani, G.; Barone, V. *J. Comput. Chem.* in press.
(32) (a) MacArthur, M. W.; Thornton, J. M. *J. Mol. Biol.* **1991**, *218*, 397. (b) Garbay-Jaureguiberry, C.; Arnoux, B.; Prange, T., Wehri-Altenburger, S.; Pascard, C.; Roques, B. P. *J. Am. Chem. Soc.* **1980**, *102*, 1827.
(33) Madison, V.; Kopple, K. D. *J. Am. Chem. Soc.* **1980**, *102*, 4855.

Table 1. Relative Energy (in kcal/mol) in Aqueous Solution (E down trans = 0) of the Up Trans Conformers of ProDA, HypDA, and FlpDA (Single-Point Energy Evaluations at HF/6-31G(d) Geometries)

	HF		MP2		PBE0
	6-31G(d)	6-31+G(d,p)	6-31G(d)	6-31+G(d,p)	6-31G(d)
Pro	0.38	0.16	0.61	0.17	0.31(0.39)[a]
Hyp	−0.50	−0.92	−0.72	−1.49	−0.42(−0.63)[a]
Flp	−1.12	−1.52	−1.52	−2.11	−1.21

[a] Geometries optimized at the PBE0/6-31G(d) level.

obtain more accurate estimates of the relative stability of the different conformers.

Table 1 collects the relative energies of down and up forms of ProDA, HypDA, and FlpDA in their P−II minima, calculated in aqueous solution by means of the CPCM at different levels of theory. All the stability trends predicted at the HF/6-31G(d) level are confirmed by MP2/6-31+G(d,p) and PBE0/6-31G(d) calculations.

HF/6-31G(d) calculations are, in this case, quite reliable also from a quantitative point of view, the only significant modification produced by diffuse functions being a slight stabilization of up puckerings over their down counterparts (vide infra). As already shown,[31] HF and PBE0 equilibrium geometries are similar, and the relative stability of up and down conformers predicted by PBE0 calculations does not change if the geometry is optimized at the HF or at the density functional level (see ref 31 and Table 1).

3.1. Energy. Inspection of Table 1 shows that at the HF/6-31G(d) level the up form of HypDA is more stable than its down counterpart by about 0.5 kcal/mol, whereas for ProDA the situation is reversed, with the down form favored by ≈0.4 kcal/mol. This trend is in agreement with the experimental results concerning *N*-acetylproline methyl ester and *N*-acetylhydroxyproline methyl ester that, in the solid state, assume a down and an up puckering, respectively.[16] Moreover, in collagen-like peptides, hydroxyproline adopts almost exclusively an up puckering.[17] A survey of trans proline residues in polypeptides and proteins shows instead that up and down puckerings are almost equiprobable.[34] Furthermore, it is worth noting that up proline puckerings are often stabilized by interresidue $C^{\gamma}H-O$ hydrogen bonds, which are absent in dipeptide analogues.[35]

In agreement with experiments,[16,18] the up form of FlpDA is more stable than its down counterpart by 1.1 kcal/mol: this confirms that the relative stability of the up conformer increases with the electronegativity of the 4(*R*) substituent.

MP2 and DFT calculations confirm the above picture and yield results in closer agreement with experimental indications. The down conformer is still preferred for ProDA, but the two puckerings are almost isoenergetic, whereas the energy difference between up and down puckerings increases for HypDA and FlpDA.

3.2. Structures. The most relevant geometrical parameters of PII minima of ProDA, HypDA, and FlpDA calculated at the CPCM/HF/6-31G(d) level are collected in Table 2. The values of ϕ computed for down puckerings are clustered around −70° (−72.7° for ProDA), whereas those predicted for up puckerings are ≈ −60°. These values are close to the experimental averages in collagen-like polypeptides: in $(PHG)_{10}$, prolines in the X position (down puckering) have an average ϕ of −72.6°, whereas the dihedrals of Hyp's in the Y position (up puckering) are clustered around −59.6°.[11] The puckering of the cycle influences the ψ dihedral angles: for instance, in HypDA the computed ψ of the down puckering (167.7°) is ≈20° larger than that of the up puckering. Smaller differences are found for ProDA and FlpDA. As a matter of fact, in $(PHG)_{10}$ and $(PPG)_{10}$ the average ψ values for the residues in the X position are 163.8° and 161.4°, respectively, whereas the residues in the Y position have average ψ values of 149.8° and 153.3°, respectively.[9,11]

To investigate this point in more detail, we performed CPCM/HF/6-31G(d) geometry optimizations for HypDA at fixed values of ψ. For up puckerings, smaller values of ψ (≈145°) are favored over larger ones (≈160°) by ≈0.5 kcal/mol. For down puckerings, these two geometries are instead practically isoenergetic ($\Delta E \leq 0.1$ kcal/mol). Rigid geometry single-point calculations for $140° \leq \psi \leq 170°$ give a similar indication: the potential energy surface is very flat for the down conformer, whereas it exhibits a well-defined minimum around $\psi = 145°$ for the up puckering.

It is important to highlight that, despite the smaller dependence of the equilibrium ψ value on the ring puckering, ProDA and FlpDA exhibit the same behavior as HypDA. As a matter of fact, rigid geometry CPCM/HF/6-31G(d) single-point calculations show that the energy difference between geometries with ψ ≈143° and with ψ ≈ 165° is remarkably smaller for down puckerings than for up puckerings: ≈0.3 kcal/mol versus ≈1.3 kcal/mol in ProDA and ≈0.8 kcal/mol versus 1.1 kcal/mol in FlpDA. Furthermore, partial CPCM/HF/6-31G(d) geometry optimizations at fixed values of ψ predict that the potential energy surfaces for the two puckerings of ProDA are very similar to those found for HypDA.

The effects determining this result can be more easily visualized with reference to Figure 3, which sketches the most relevant interactions between the cycle and the amidic moiety. A first important repulsive "contact" occurs between N_{i+1} and C^{β} (interaction 1 in Figure 3). This effect could be more important for up puckerings, where C^{β} and N_{i+1} are on the same side with respect to the average plane of the pyrrolidine ring (see Figure 3). As a consequence, in up puckerings the ψ dihedral does not assume values too close to 180°. Furthermore, the pyrrolidine ring prefers puckerings approaching ideal "envelope" forms, where the C^{β} atom is closer to the average ring plane (vide infra).

When pyrrolidine assumes a down puckering, the contact between O and the C^{γ} group (interaction 2 in Figure 3) is switched on. This repulsion is relieved if ψ approaches 180°, thus explaining why for down conformers larger values of ψ are favored. However, as we have seen, larger values of ψ imply also an increase of the $N_{i+1}-C^{\beta}$ repulsion. So, for $\psi \rightarrow 180°$, "envelope-like" puckerings are favored also for the down conformer, with C^{γ} close to (and C^{β} well above) the average pyrrolidine plane (see HypDA). On the other hand, for smaller values of ψ the pyrrolidine ring assumes more "regular" twisted-like puckerings (see ProDA and FlpDA). Despite the smaller values of $C^{\gamma}-C^{\beta}$ and $C^{\gamma}-C^{\delta}$ bond distances, C^{γ} and O atoms are more distant (for a given value of ψ) in FlpDA than in HypDA, mostly due to the different conformation (bond and dihedral angles) of the cycle. This result can explain the larger value of ψ preferred by HypDA.

Two additional interactions are influenced by the ψ dihedral: the first interaction is a "classical" steric repulsion between O and N (interaction 3 in Figure 3). When ψ approaches 180°, those atoms are eclipsed. The second interaction (interaction 4 in Figure 3) involves the $C_{i-1}O_{i-1}$ group, which is better exposed to intermolecular interactions for larger values of ψ. So, it can

(34) Milner-White, E. J.; Bell, L. H.; Maccallum, P. H. *J. Mol. Biol.* **1992**, *228*, 725.

(35) Chakrabarti, P.; Chakrabarti, S. *J. Mol. Biol.* **1998**, *284*, 867.

Table 2. Most Relevant Geometrical Parameters of the PII Minima of Down and Up Puckerings of ProDA, HypDA, and FlpDA, Obtained at the CPCM/HF/6-31G(d) Level in Aqueous Solution[a]

	Pro		Hyp		Flp	
	down	up	down	up	down	up
ϕ	−72.7	−60.7 (−58.1)	−70.5	−63.1	−70.1	−62.0
ψ	149.2	144.6 (141.3)	167.7	146.3	147.4	146.1
ω	−178.1	−178.4 (−179.2)	176.4	−177.1	177.8	−178.1
χ_1	29.7	−26.4 (−27.3)	32.0	−26.2	28.2	−26.1
χ_2	−36.7	37.3 (−38.2)	−36.7	38.0	−36.7	36.8
χ_3	30.0	−33.6 (−33.8)	26.9	−34.7	29.8	−32.7
χ_4	−10.7	17.8 (17.5)	−6.9	19.2	−11.8	16.7
χ_5	−12.0	5.4 (6.1)	−15.8	4.4	−9.7	5.9
$C^\gamma-C^\delta-N-C_{i-1}$	168.7	−156.2 (−158.7)	164.9	−153.7	165.6	−155.9
$C_{i-1}-N-C^\alpha-C^\delta$	−179.75	174.4 (176.5)	−172.3	173.4	−177.8	173.1
C−N	1.340	1.340 (1.349)	1.341	1.340	1.342	1.342
C−O	1.215	1.215 (1.234)	1.214	1.215	1.214	1.214
$C^\alpha-N$	1.452	1.455 (1.456)	1.455	1.455	1.456	1.456
$C^\gamma-O_H$			1.398	1.406	1.377[c]	1.386[c]
$C^\beta-C^\alpha$	1.538	1.538 (1.535)	1.537	1.539	1.536	1.539
$C^\gamma-C^\beta$	1.529	1.528 (1.526)	1.529	1.525	1.517	1.513
$C^\delta-C^\gamma$	1.528	1.525 (1.524)	1.532	1.519	1.522	1.513
$C^\delta-N$	1.463	1.463 (1.460)	1.461	1.461	1.457	1.460
puckering[a]	95% T	50% T	63% T	42% T	88% T	56% T
$C_{i-1}-O_{i-1}$ stretching[b] frequency	1691.8	1691.6	1696.3	1689.4	1698.2	1696.2

[a] Geometrical parameters obtained at the CPCM/PBE0/6-31G(d) level are given in parentheses. E = envelope, T = twist, %E = 100 − %T. [b] Values (in cm^{-1}) scaled by 0.91. [c] $C^\gamma-F$ bond distance.

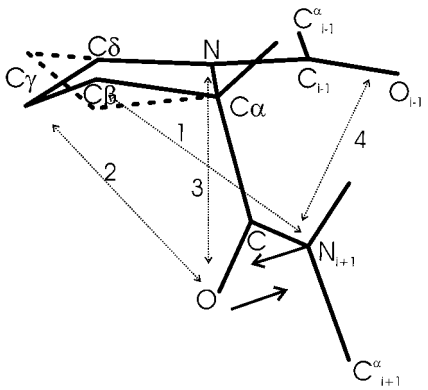

Figure 3. Most relevant interactions (dotted arrows) in proline dipeptide analogue. Dashed bonds refer to up puckering. Bold arrows indicate schematically the movement of the atoms following an increase of the ψ dihedral angle.

be expected that an increase of the solvent polarity or the formation of intermolecular H-bonds in the solid state favors structures with larger ψ values.

The nature of the 4(*R*) substituent has a strong influence on the geometry of the pyrrolidine cycle:

1. The values of the $C^\gamma-C^\delta$ and $C^\gamma-C^\beta$ bond distances are smaller in the up forms of HypDA and FlpDA than in their down counterparts. Although this effect is much smaller in ProDA, the trend is analogous. Furthermore, in agreement with experiments,[16] these bond distances decrease with the electronegativity of the substituents at C^γ. At the same time, while the down conformers exhibit a puckering very close to a twist form, the up ones are intermediate between envelope and twist forms, with the N, C^α, C^β, and C^δ atoms almost coplanar and the C^γ atom remarkably out of the plane formed by the remaining atoms of the cycle.

2. The distance between C^γ and its electronegative substituent (F or OH) is longer in the up conformer by ≈0.02 Å. In agreement with experiments,[16] a 4(*R*) electronegative substituent increases the pyramidalization of the imino nitrogen. This trend has been explained in terms of the inductive effect of the electronegative substituent, which could destabilize the "traditional" amidic resonance structure involving a nitrogen atom with a formal positive charge ($N^+=C-O^-$). This explanation can account also for the increase of the vibrational frequency associated with the stretching of the imino−carbonyl of the ester analogues of our dipeptides[8b] when going from ProDA to HypDA and FlpDA.

However, our calculations suggest that this behavior is due to the concurrence of several effects. Different geometrical constraints of up and down conformers due to their different puckering (vide supra) could explain why the nitrogen atom has a pyramidal environment in $ProDA_{up}$, whereas in $ProDA_{down}$ it is almost perfectly planar. Moreover, one of the driving forces of the N-pyramidalization seems to be the necessity of avoiding too large repulsive interactions between the nonbonding electrons of nitrogen and the X substituent (in up conformers) or O (in down conformers). This could explain why, in HypDA, N is more pyramidal for down puckering, although in this conformer the OH group is more distant than in the up one. In the PII conformation, since the equilibrium value of ψ approaches 180°, O and N are forced to be eclipsed. From this point of view it is noteworthy that the "direction" of pyramidalization of N is opposite in up and down puckerings. Shortly, the pyramidalization of N in up puckerings of HypDA and FlpDa could be due to the necessity of pulling away the electronic density of the N lone pair from that of oxygen or fluorine. As a matter of fact, the up puckerings of HypDA and FlpDA, which exhibit the largest degree of pyramidalization, have the shortest and the longest N−C bonds, and the longest and the shortest C−O bonds, respectively. This shows that any correlation between the pyramidalization of the nitrogen atom and the strength of the amidic resonance structure has to be done very cautiously and cannot be taken for granted.

The values of the imino−carbonyl stretching frequency calculated in aqueous solution at the HF/6-31G(d) level confirm these considerations (see Table 2). The lowest stretching frequency is indeed found for $HypDA_{up}$, suggesting that the $N^+=C-O^-$ resonance structure is more stable in this compound. It is also noteworthy that the frequencies of the up puckerings

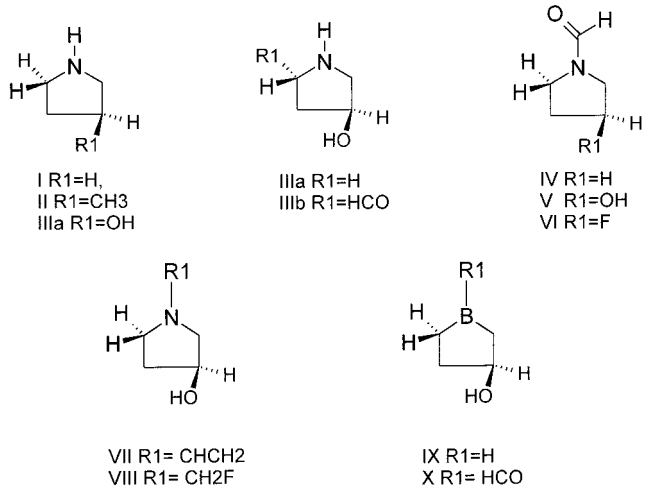

Figure 4. Pyrrolidine analogues used in the analysis of the substituent's influence on ring puckering.

Table 3. Relative Energy (E(up) − E(down) in kcal/mol) of Down and Up Puckerings of the Model Pyrrolidine Compounds of Figure 4 (All the Results Refer to HF/6-31G(d) Computations)

I	**II**	**IIIa**	**IIIb**
0.0	1.03	0.27 (0.63[a])	1.43 (1.34[a])
IV	**V**	**VI**	
0.0	−0.38 (−0.72[a])	−0.97	
VII	**VIII**	**IX**	**X**
−0.23	−0.06	−0.73	−0.71

[a] Geometrical parameters of the ring frozen to the values optimized for the corresponding dipeptide analogue.

are always lower than those of the corresponding down puckerings (vide infra).

3.3. Effects Influencing the Puckering of Proline's Derivatives. Our computations confirm that the stability of the up form of a pyrrolidine ring increases with the electronegativity of the 4(*R*) substituent. This result has been explained by invoking the so-called gauche effect.[8] However, the cyclic nature of Hyp could have some important consequences on the above picture. On the one hand, steric effects between substituents are more important: up puckering exhibits an axial-like conformation of the 4(*R*) group that could be more sterically hindered. On the other hand, the ring is a quasi planar system that can interact conjugatively with the π system of the amide bond: in the up conformation the C−X bond (X= OH, F, ...) is perpendicular to the average plane of the ring, while in the down puckering it is parallel. This effect, hitherto neglected, could play some significant role in discriminating up and down puckerings. As discussed in the preceding section, the comparison of the geometries of the pyrrolidine ring in the different conformers supports this hypothesis, suggesting that the up puckering can exhibit some kind of partial π delocalization, via an interaction with the C−X bond, which is, indeed, slightly elongated.

To get a better decomposition of the effects influencing the conformation of the pyrrolidine ring, we then studied in vacuo at the HF/6-31G(d) level some simplified models of HypDA and ProDA (see Figure 4). This allowed us to focus the attention on the pyrrolidine ring, verifying how systematic variations of the ring substituents affect the relative stability of up and down puckerings.

We first compared the energy of the up-like and down-like conformers of **I**, **II**, and **IIIa** (see Table 3), freezing the dihedral of the nitrogen substituent with respect to the pyrrolidine plane to the corresponding value of the acetyl group in the minimum of the dipeptide analogue (DA) and optimizing the remaining geometrical parameters (partial geometry opimizations). In **IIIa**, the down puckering is predicted to be more stable by 0.27 kcal/mol. When the OH group is substituted by a methyl (**II** in Figure 4), the energy difference between up and down puckerings increases to 1.03 kcal/mol. These results show that a gauche effect is operative for the HO−C^{γ}−C^{δ}−N moiety (≈0.75 kcal/mol) but that it is not large enough to overcome the larger steric hyndrances suffered by the axial-like up conformer.

The presence of an amido group at C^{α} could increase, at the second order, the electron-withdrawing power of this latter atom, eventually inducing a second C^{α}−C^{β}−C^{γ}−OH gauche effect. We have thus examined compound **IIIb**, in which the formyl group models the amido group present in the real peptide. Quasi total geometry optimizations predict that the down conformer is more stable by ≈1.4 kcal/mol, suggesting that the factors responsible for the preference for the up puckerings have to be sought in the presence of the imino group in the DA.

For **V**, the up conformer is indeed more stable than the down conformer by ≈0.4 kcal/mol, and for **VI** (its fluorinated analogue), the extra stabilization of the up puckering over the down one increases to 0.97 kcal/mol, in line with the results obtained for the dipeptide analogue.

The comparison among the geometries of up and down puckerings of **IV**, **V**, and **VI** confirms the trend already sketched at the dipeptide analogue level. The C−C bonds involving the 4(*R*)-substituted carbon atom are shorter in the up puckerings, and their length decreases when the electronegativity of the 4(*R*) substituent increases.

All these results can be explained by a larger electronic delocalization in the up puckering, due to the different orientation of the C−X bond with respect to the ring plane.

A NBO analysis of **VI** shows indeed that the stabilizing CH/CF* interaction is ≈2 kcal/mol larger for an up than for a down puckering.

The most important interactions are those between the C−F σ* antibonding orbital and the C^{β}−H_{ax} and C^{δ}−H_{ax} (ax = axial) σ bonding orbitals. This results in the formation of partial C^{γ}−C^{β} and C^{γ}−C^{δ} π bonding orbitals which stabilize the up conformer and account for the shorter C−C bonds in the pyrrolidine ring. One of the molecular orbitals issuing from that interaction is sketched in Figure 5; it is clear that in the up conformer it has a larger bonding character. An increase of the electronegativity of the X atom increases the C^{γ} contribution to the C^{γ}−X* orbital and, consequently, the stabilization coming from the CH_{ax}/CX* interaction.

These interactions are present independently from the nature of the nitrogen substituents; however, we have seen that the presence of an imino group increases the stabilization of the up conformer. This could be due to the conjugation between the π amidic bond and the partial π bond of the cycle. Furthermore, the delocalization of the nitrogen lone pair in the amidic bond could reduce its repulsion with the nonbonding electron of the electronegative substituent, and this effect could be important for the rather "diffuse" lone pairs of the oxygen atom.

The comparison between the energies of the two puckerings in compounds **VII** and **VIII** (see Figure 4) can shed some light on the relative importance of delocalization effects and of dipole−dipole interactions (in HypDA there are three strongly dipolar groups). In **VII**, the carbonyl group has been substituted by an ethylene group, thus allowing some delocalization of the nitrogen lone pair but substantially reducing dipolar interactions. In **VIII**, in contrast, the H_2C−F group should mimic the dipole moment of a carbonyl group without any delocalization effect.

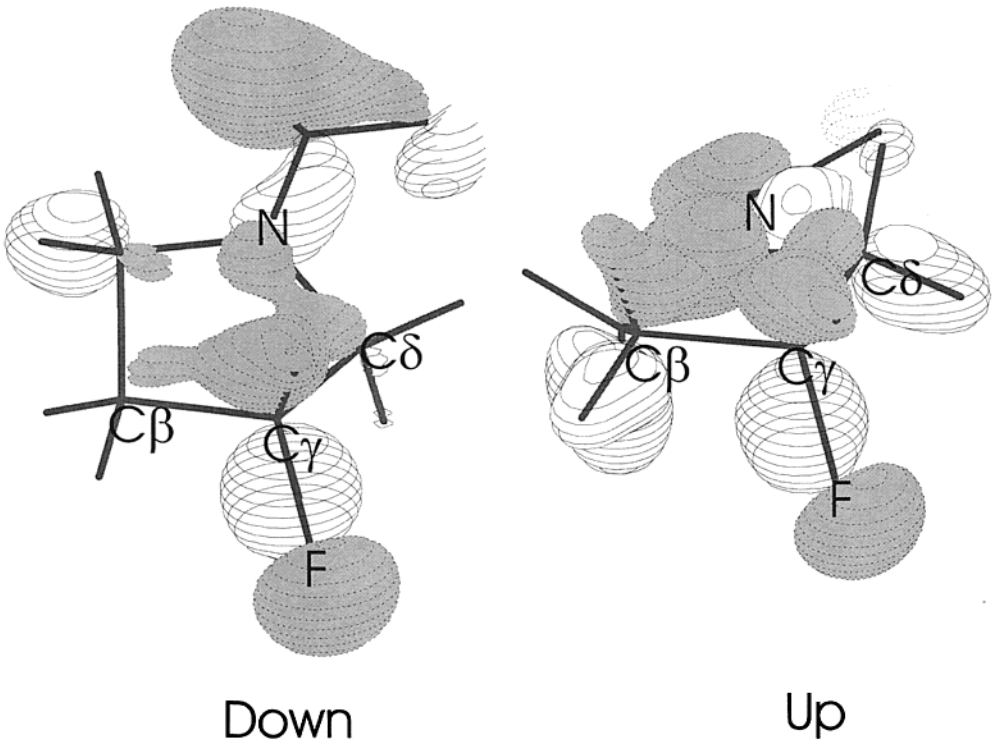

Figure 5. Molecular orbital deriving from the interaction of the C−H σ bonding orbital and the C−F σ^* antibonding orbital in 1-formyl-3(*R*)-fluoropyrrolidine (compound **VI** in Figure 4).

Partial geometry optimizations show that the difference between the up and down puckerings is reduced from 0.23 to 0.06 kcal/mol when going from **VII** to **VIII**.

It is also noteworthy that in the boron analogue of pyrrolidine (**IX** in Figure 4), the up conformer is more stable than the down one by ≈0.7 kcal/mol, despite the lower electronegativity of the boron atom. Obviously, the relative stability of the two puckerings is unaffected by the presence of a carbonyl substituent on the boron atom (**X** in Figure 4), since the absence of a nonbonding pair does not allow any conjugative effect.

Solvent effects do not modify the above picture. Quasi total geometry optimizations in aqueous solution show indeed that a polar solvent just emphasizes the delocalization effects, probably increasing the dipolar character of the $C^{\gamma}-X$ bond. Indeed, already in **IIIa** the up conformer is slightly more stable than its down counterpart (by 0.14 kcal/mol), and in **IVb** the difference increases up to 0.71 kcal/mol.

In summary, in this section we have shown that up puckerings of pyrrolidine ring are stabilized by the interactions between the antibonding $C^{\gamma}-X$ orbital and the bonding orbitals $C^{\delta}-H_{ax}$ and $C^{\beta}-H_{ax}$. The stabilization coming from this interaction increases with the electronegativity of the X substituent and with the presence of a delocalized π system including the nitrogen atom. This explanation accounts for the fact that in up puckerings (i) the $C^{\gamma}-X$ bond length increases, (ii) the $C^{\gamma}-C^{\delta}$ and $C^{\gamma}-C^{\beta}$ bond lengths decrease, and (iii) the imino−carbonyl stretching frequency is reduced.

On the other hand, the gauche effect should not play a significant role in favoring up puckerings, whose relative stabilization over down puckerings actually increases when nitrogen is replaced by the less electronegative boron atom.

3.4. Cis−Trans Isomerism around the Peptidic Bond. Table 4 collects the relative energies of the cis and trans isomers of ProDA (see Figure 6), HypDA, and FlpDA for both up and down puckerings. In agreement with previous experimental results,[8b] the presence of an electronegative substituent on the pyrrolidine ring stabilizes the trans isomer more than the cis one: the energy difference between the two isomers ($\Delta G_{trans-cis}$) increases indeed by ≈0.3 kcal/mol when going from ProDA to HypDA and FlpDA. Cis and trans isomers are closer in energy for down puckerings of the pyrrolidine ring. The largest difference is found for HypDA, where $\Delta G_{trans-cis}$ decreases from 1.54 (up puckering) to 0.81 kcal/mol (down puckering). The

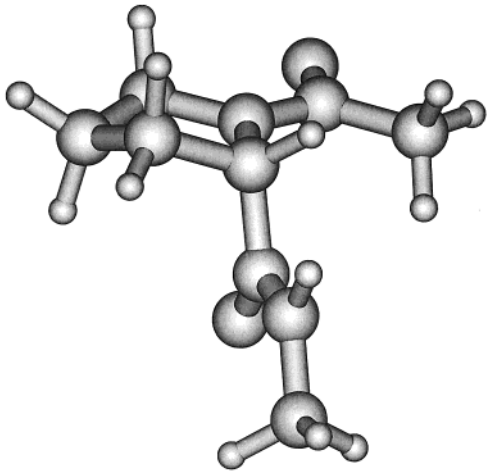

Figure 6. Minimum energy geometry of the cis isomer of $ProDA_{down}$.

Table 4. Relative Energy (in kcal/mol) in Aqueous Solution of the Cis Isomer (E trans = 0) of the Different Conformers of ProDA, HypDA, and FlpDA, Computed at the CPCM/HF/6-31G(d) and CPCM/PBE0/6-31G(d) Levels

	Pro		Hyp		Flp	
	HF	PBE0	HF	PBE0	HF	PBE0
down	0.87	0.27	0.81	−0.08	1.21	0.68
up	1.17	0.45	1.54	0.89	1.52	0.62

Table 5. Relative Energy (in kcal/mol) in Aqueous Solution of the Different Isomers of HypDA and FlpDA, Calculated at the CPCM/HF/6-31G(d) Level

	hyp		flp	
	trans	cis	trans	cis
down	0.0 (2.24[a])	1.79 (1.88[a])	0.0	−0.20
up	1.52	2.72	0.91	2.15

[a] Without intramolecular hydrogen bond.

calculated $\Delta G_{trans-cis}$ (≈1 kcal/mol) is in agreement with previous experimental determinations on the ester analogues of the dipeptides under study,[8b] especially taking into account that the experimental results account for the simultaneous down−up and cis−trans isomerism.

The larger stability of the trans isomer in FlpDA and HypDA has been explained in terms of the decrease of $C^{\gamma}-C^{\delta}$ bond distances due to the presence of a 4(*R*) electronegative substituent.[8c] The consequent reduction of the steric repulsion between C^{α}_{i-1} and C^{δ} would stabilize the trans isomer of FlpDA and HypDA with respect to the trans isomer of ProDA. Actually, several geometrical parameters concur to determine the value of the $C^{\alpha}_{i-1}-C^{\delta}$ bond distance: the $N-C_{i-1}-C^{\alpha}_{i-1}$ and $C^{\delta}-N-C_{i-1}$ bond angles, the degree of planarity of the imino moiety, and so on. As a matter of fact, $FlpDA_{trans}$ exhibits contemporarily the shortest $C^{\gamma}-C^{\delta}$ and $C^{\alpha}_{i-1}-C^{\delta}$ bond distances. Furthermore, the $C^{\alpha}_{i-1}-C^{\delta}$ bond distance is just 0.004 Å shorter in FlpDA than in HypDA (2.917 vs 2.921 Å), and it is not likely that such a difference (in a nonbonding interaction) can by itself account for a $\Delta G_{trans-cis} \approx 0.4$ kcal/mol smaller. As a matter of fact, single-point CPCM/HF/6-31G(d) calculations on both isomers of HypDA performed using the geometry of the corresponding isomers of ProDA predict that $\Delta G_{trans-cis}$ is practically unchanged (actually it is larger by ≈0.01 kcal/mol) with respect to that determined after total geometry optimizations.

It is thus likely that the influence of the 4(*R*) substituent on $\Delta G_{trans-cis}$ can be explained by a "direct" repulsion between the oxygen atom and the 4(*R*) electronegative substituents, destabilizing the cis isomer more than the trans one. This hypothesis would also explain the smaller $\Delta G_{trans-cis}$ found for down puckerings, where the electronegative substituent is less close to the imino moiety.

Inclusion of electron correlation by means of single-point PBE0/6-31G(d) calculations causes a reduction (by ≈0.6−0.8 kcal/mol) of the energy difference between trans and cis isomers. As a consequence, for down puckerings, cis and trans isomers become very close in energy. This result is not surprising, since the ester analogue of ProDA adopts a cis conformation in the crystalline state.[16]

3.5. (*S*) Diastereoisomers of Hydroxyproline and Fluoroproline. Table 5 reports the relative energies of cis and trans isomers of hypDA and flpDA. For both compounds, down puckerings are more stable than their up counterparts. In the 4(*S*) diastereoisomer, the X substituent is indeed axial-like in down puckerings and equatorial-like in up puckerings, just the opposite of what happens for 4(*R*) compounds. As a consequence, the stereoelectronic effects we have shown to favor 4(*R*) up puckerings now stabilize 4(*S*) down puckerings.

However, in hypDA and flpDA an additional effect must be taken into account. In down puckerings the $C^{\gamma}-X$ bond and the amido group are indeed eclipsed (see Figure 7), whereas in up puckerings of HypDA and FlpDA these groups are in an "anti" orientation. This repulsive interaction can explain why for flpDA up and down puckerings are closer in energy than for FlpDA.

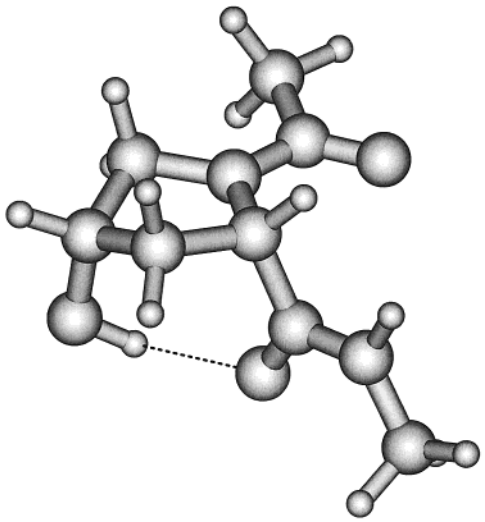

Figure 7. Minimum energy geometry of $hypDA_{down}$.

The down puckering of hypDA allows the formation of an intramolecular hydrogen bond between the oxydril hydrogen and the amidic oxygen (see Figure 7). The presence of this interaction has already been invoked to explain the low value of the carbonyl stretching frequency in the ester analogue of hypDA.[13] Due to the presence of this H-bond, $hypDA_{down}$ is more stable than $hypDA_{up}$ by ≈1.5 kcal/mol.

However, this result is not directly transferable to the analysis of collagen-like polypeptides in aqueous solution, due to the competition between inter- and intramolecular hydrogen bonds. It is thus noteworthy that for the conformations in which the intramolecular hydrogen bond is absent, $hypDA_{down}$ is less stable (by ≈0.3 kcal/mol) than its up counterpart.

In agreement with the experiments performed on analogous compounds,[13] cis isomers are relatively more stable in 4(*S*) than in 4(*R*) diastereoisomers. Cis and trans isomers are quite close in energy in flpDA and, if intramolecular hydrogen bonds are absent, also in hypDA. It is interesting to note that for the down puckering of flpDA the cis isomer is slightly more stable than the trans one. For its ester analogue, CPCM/HF/6-31G(d) calculations predict instead, in agreement with experiments,[13] the opposite trend (trans isomer more stable by ≈0.4 kcal/mol). This suggests that some caution has to be used when analyzing the behavior of polypeptides using results obtained on their ester analogues.

The trans−cis energy differences predicted for 4(*S*) up puckerings are comparable to those predicted for 4(*R*) down puckerings. The 4(*S*) down trans isomers are instead relatively destabilized with respect to their 4(*R*) up counterparts. To reduce the repulsion between the $C^{\gamma}-X$ and the amido groups, in 4(*S*) down compounds the ψ dihedrals assume quite large values (in the range ±170°−180°). This geometry disfavors trans isomers, where the carbonyl groups are too close.

As a matter of fact, when an intramolecular hydrogen bond is formed, the trans isomer is substantially stabilized over the cis one. The optimization of the OH−O hydrogen bond forces indeed the ψ dihedral to assume quite low values (≈140°).

3.6. Hydrogen Bonds. In $(PPG)_{10}$ and in its hydroxylated analogues, the hydrogen bonds connecting the three chains of collagen's triple helix always involve the carbonyl group of the proline in the X position and the amidic hydrogen of a glycine residue.[11] Raines et al. have suggested that the acidity of amidic hydrogen increases when glycine is followed by prolines bearing electronegative substituents.[8c] However, the carbonyl group of the proline in the X position is part of the imino group involving

Table 6. Interaction Energy (in kcal/mol) between a Water Molecule and the Up Puckerings of ProDA, HypDAm and FlpDA (Single-Point Energy Calculations on HF/6-31G(d) Equilibrium Geometries)

	Pro			Hyp			Flp		
	HF	MP2	PBE0	HF	MP2	PBE0	HF	MP2	PBE0
				6-31G(d)					
				Gas Phase					
ΔE	−7.96	−9.91	−9.49	−7.16	−9.05	−8.74	−6.73	−8.51	−8.16
ΔE_{corr}[a]	−6.61	−7.25	−8.14	−5.87	−6.52	−7.45	−5.53	−6.13	−6.96
				Solvent					
ΔE	−0.35	−2.81	−2.56	−0.61	−2.90	−2.76	−0.41	−2.58	−2.39
ΔE_{corr}[a]	1.0	−0.15	−1.21	0.68	−0.36	−1.61	0.79	−0.20	−1.19
				6-311+G(d,p)					
				Gas Phase					
ΔE	−7.40	−9.42	−8.62	−6.68	−8.60	−7.87	−6.38	−8.26	−7.56
ΔE_{corr}[a]	−6.77	−7.61	−7.99	−6.02	−6.80	−7.21	−5.75	−6.52	−6.93
				Solvent					
ΔE	+0.76	−1.81	−0.91	+0.40	−1.98	−1.16	+0.41	−1.86	−1.07
ΔE_{corr}[a]	+1.39	0.0	−0.28	+1.06	−0.18	−0.50	+1.04	−0.12	−0.44

[a] Corrected for the BSSE (calculated in vacuo).

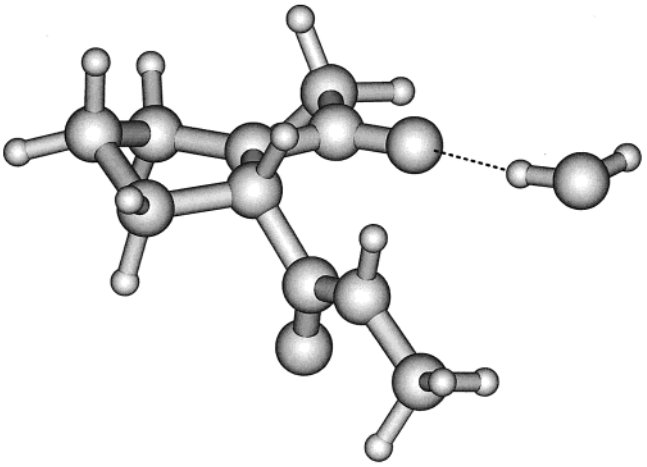

Figure 8. Minimum energy geometry of the complex formed by $ProDA_{up}$with one water molecule.

the nitrogen atom of hydroxyproline in the Y position, and we have shown that this group could participate in the electronic conjugation of the pyrrolidine ring.

It is thus interesting to verify if an increase of the electronegativity of the 4(*R*) substituents increases the "absolute" tendency of the imino−carbonyl to form hydrogen bonds. We have thus optimized at the CPCM/HF/6-31G(d) level the complex formed by a water molecule and ProDA, HypDA, and FlpDA, always with up puckered rings (see Figure 8). The hydrogen bond strength has been evaluated by single-point energy computations at HF, MP2, and PBE0 levels, using both the standard 6-31G(d) and 6-311+G(d,p) basis sets (see Table 6), and estimating the basis set superposition error (BSSE) by the counterpoise method.[36]Inspection of Table 6 shows that the relative hydrogen-bonding power strongly depends on the polarity of the embedding medium. In the gas phase, indeed, all the calculations predict that the interaction energy with the water molecule decreases in the order proline > hydroxyproline > fluoroproline. In aqueous solution, instead, HypDA and FlpDA form hydrogen bonds stronger than those of ProDA.

These results can be explained in the light of our preceding discussion on the relative stability of the amidic resonance structures. An electronegative substituent on C^{γ} could destabilize, by means of both through-bond and through-space interactions, the resonance structure with a formal positive charge on nitrogen. It can be expected that the hydrogen-bonding power of the imido moiety increases with the relative stability of this structure, in which the oxygen atom bears a formal negative charge. These considerations can account for the stability trend found in vacuo.

A polar embedding medium can modify the above picture. It shields the electron-withdrawing power of the 4(*R*) substituent. Furthermore, it could enhance the importance of the delocalization effects that are operative mostly in up conformers of HypDA and FlpDA, favoring the formation of a partial $N-C_{i-1}$ double bond. As a consequence, in aqueous solution, the hydrogen bonds of HypDA and FlpDA become stronger than those of ProDA.

4. General Discussion and Conclusions

In the preceding sections, we have shown that when proline has a 4(*R*) electronegative substituent, the up puckering is "intrinsically" stabilized over its down counterpart. Moreover, the relative stability of the up puckering increases with the electronegativity of the 4(*R*) substituent. So, the preference for up forms decreases in the same order, proline < hydroxyproline < fluoroproline, as the relative stability found in collagen-like polypetides for substitutions in the Y position.[8,13] Since Mazzarella and colleagues have shown that in $(PPG)_{10}$ prolines in the Y positions have up puckerings,[9] our calculations clearly indicate that one fundamental reason (if not the main) for the hydroxylation of the Y proline in collagen is the stabilization of the up form in that position.

Our analysis shows that the explanation for this trend put forward until now, i.e., the gauche effect,[8] is not decisive. The up form is indeed relatively more stable when nitrogen is substituted by the less electronegative boron. The interaction between vicinal CH_{ax} σ and C−X σ^* bonds seems more important, and it is maximized for an axial-like placement of the X substituent. As a consequence, up puckerings are favored by 4(*R*) substituents and down puckerings by 4(*S*) substituents.

This consideration could explain why 4(*S*) hydroxyproline destabilizes collagen's triple helix when placed in the Y position. Furthermore, 4(*S*) down conformers are characterized by ψ dihedrals quite different from those found in the collagen triple helix. Finally, in hypDA, an intramolecular hydrogen bond can be formed, perturbing not only the backbone dihedrals but also the hydration pattern of collagen.

But why should residues in Y positions assume up puckerings in collagen-like polypeptides? The present study is devoted to

dipeptide analogues, and thus it cannot give a definite answer to this question. Interresidue interactions could play a dominant role in determining the conformation of each residue most suitable to form a P−II helix and then a triple helix. As a consequence, a Gly-Pro-Pro-Gly sequence is the smallest model for any reliable analysis of this question. Notwithstanding this, our results on dipeptide analogues provide some insights into the relationship existing between backbone parameters and puckering of the pyrrolidine ring.

As a matter of fact, while it can be expected that the values of ϕ dihedrals are influenced by the puckering of the ring, it is not obvious that this is true also for the ψ dihedrals. Our calculations predict indeed that for up puckerings ψ is confined around 145° ≈ 150°, whereas for down puckerings it has a significantly larger conformational freedom, its value spanning quite easily the region 145° ≤ ψ ≤ 165°. From this point of view, it is interesting that the average value of ψ dihedrals in $(PPG)_{10}$ is 161.4° for down prolines (those in the X position) and 153.3° for up prolines (those in the Y position).[9b] The influence of the puckering on the average value of ψ seems an intrinsic feature of proline. A statistic survey of 700 protein structures[37] with a resolution of 1.8 Å or better shows indeed that for $\phi \approx -60°$, the relative probability of up puckering decreases from 0.73 to 0.46 when going from $\psi \approx 130°$ to $\psi \approx 170°$. On the other hand, for $\phi \approx -70°$, the relative probability of down puckering increases from 0.53 to 0.72 when going from $\psi \approx 130°$ to $\psi \approx 170°$. Furthermore, the large majority of up prolines exibits ψ in the range $120° < \psi < 150°$, whereas the ψ dihedrals of down prolines span quite uniformly the interval 180°−120°.[37]

In agreement with experiments,[38] our calculations predict that up conformers prefer envelope-like structures, whereas down conformers exhibit structures close to "regular twists". It is possible that even small differences in the ring puckering could influence the close packing of collagen's triple helix.

Another point worth mentioning is that, for all the DAs examined, up puckerings have a larger dipole moment than their down counterparts. It cannot be excluded, then, that the presence of up puckering in the Y position and the alternate X/down−Y/up arrangement of pyrrolidine rings in collagen is dictated by the maximization of long-range dipole−dipole interactions.[8c]

Confirming previous experimental indications on analogous compounds,[13] the relative stability of the trans isomer with respect to its cis counterpart is enhanced when going from ProDA to HypDA and FlpDA, whereas 4(*S*) substituents act in the opposite direction. However, as previously noted,[9b] it is difficult to understand why this result would be "effective" only for residues in the Y position.

Our calculations do not allow a definite answer about the role played by the puckering and by the nature of the substituents on the strength of the hydrogen bonds formed by the carbonyl following proline or hydroxyproline in collagen.

The features of a protein environment are, indeed, intermediate between the gas phase and an aqueous solution. Furthermore, 4(*R*) substituents could be involved in hydrogen bonds with solvent molecules,[11] and this would influence their electronic characteristics. Finally, collagen and collagen-like polypeptides could have a slightly different packing and, thus, a different hydrogen bond geometry, depending on the nature of the 4(*R*) substituent. It is thus very difficult to evaluate the consequences of our computational results for collagen stability. However, since the main interaction connecting the three helices of collagen is the hydrogen bond between the carbonyl following Hyp and the amido nitrogen of Gly preceding Hyp, it would surely be interesting to determine the effects of the ring puckering and of the 4(*R*) substituent on the geometry of the hydrogen bond network in collagen.

In summary, our results strongly support the proposal that the stability of the triple helix of collagen-like polypeptides fluorinated or hydroxylated in the Y position is related to the necessity of having up puckerings in that position. The reasons for this requirement could be structural, since we have shown that the ring puckering remarkably influences backbone parameters. However, the possibility that other effects (e.g., tuning of the cis−trans equilibrium or maximization of hydrogen bond's stability in triple helices) play some role cannot be ruled out.

Acknowledgment. This work was supported by the Italian Research Council (CNR), by the Italian Ministry for the University and the Scientific and Technological Research (MURST), and by Gaussian, Inc. The authors thank Dr. Luigi Vitagliano and Dr.Rita Berisio for helpful discussions and for focusing our attention on the possible role of stereoelectronic effects in collagen stability.

Supporting Information Available: Tables of total free energies in solution (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

JA010599I

(36) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553.
(37) Dunbrack, R. L., Jr.; Cohen, F. E. *Protein Sci.* **1997**, *6*, 1661. See also www.fccc.edu/research/labs/dunbrack/.
(38) Vitagliano, L.; Berisio, R.; Mazzarella, L.; Zagari, A., *Biopolymers* **2001**, *58*, 459.