Conformers of Gaseous Proline

Eszter Czinki and Attila G. Császár*[a]

Abstract: Accurate geometries, relative energies, rotational and quartic centrifugal distortion constants, dipole moments, harmonic vibrational frequencies, and infrared intensities were determined from ab initio electronic structure calculations for eighteen conformers of the neutral form of the amino acid L-proline. Only four conformers have notable population at low and moderate temperature. The second most stable conformer is only 2 ± 2 kJ mol^{-1} above the global minimum, while the third and fourth conformers are nearly degenerate and have an excess energy of 7 ± 2 kJ mol^{-1} relative to the global minimum. All four conformers have one hydrogen bond: N···HO in the lower energy pair of conformers, and NH···O in the higher energy pair of conformers. The conformer pairs differ only in their ring puckering. The relative energies of the conformers include corrections for valence electron correlation, extrapolated to the complete basis set limit, as well as core correlation and relativistic effects. Structural features of the pyrrolidine ring of proline are discussed by using the concept of pseudorotation. The accurate rotational and quartic centrifugal distortion constants as well as the vibrational frequencies and infrared intensities should aid identification and characterization of the conformers of L-proline by rotational and vibrational spectroscopy, respectively. Bonding features of L-proline, especially intramolecular hydrogen bonds, were investigated by the atoms-in-molecules (AIM) technique.

Keywords: ab initio calculations · amino acids · conformation analysis · L-proline · pseudorotation

Introduction

The naturally occurring amino acid L-proline (Pro) plays an important role in determining the structures of proteins[1, 2] and peptides.[3] Proline is unique among the natural amino acids in that it has an imino group fixed within a pyrrolidine ring, and this makes Pro conformationally somewhat less flexible than most other amino acids. In proteins Pro residues occur[1] at a much higher than average frequency, often as part of repetitive sequences. Structural studies on Pro-containing peptides frequently find Pro in β turns,[4] with a hydrogen bond between residues 1 and 4, and in γ turns,[5] with a hydrogen bond between C=O of residue 1 and NH of residue 3. Due to the absence of an N–H bond, Pro in peptides and proteins rarely occurs in the middle of hydrogen-bond-stabilized helices. Pro residues relatively often exhibit a *cis* peptide bond, and this prompted investigations on the relationship of *cis–trans* isomerism and ring puckering.[3]

The structural features of Pro and its derivatives have been studied by experimental techniques (X-ray crystallography,[6–8] NMR[9] and IR[10, 11] spectroscopy) and theoretical (molecular mechanics,[8, 12] perturbed configuration interaction with localized orbitals (PCILO),[13] and ab initio[12, 14–17]) calculations. Nevertheless, no high-quality structural investigation of the parent amino acid proline, similar to those available for some smaller neutral amino acids such as glycine[18–20] and α-alanine,[21] has been performed prior to the present study.

There are several reasons why free amino acids are an attractive target for structural studies by modern methods of electronic structure theory: a) there is a finite number of natural amino acids and, at the same time, they have tremendous chemical, biochemical, and biological significance; b) amino acids undergo a variety of intramolecular interactions, which are best understood by means of appropriate calculations; c) these compounds exhibit extreme conformational flexibility, and hence study of their potential energy surface (PES) is not readily amenable to experiments, which are hindered further by the low volatility of amino acids and their ready thermal decomposition; and d) most amino acids are of tractable size even for high-level ab initio electronic structure calculations .

Matrix-isolation techniques and infrared spectroscopy confirmed that the molecular structure of free proline is non-zwitterionic.[10] In the same study Reva et al.[10] reported AM1 optimized geometries (resulting, unfortunately, in almost planar rings, similar to STO-3G and 6-31G RHF optimizations[14]) and consequent relative energies, which all proved to be inadequate. The earliest ab initio study for

---

[a] E. Czinki, Prof. A. G. Császár
Department of Theoretical Chemistry, Eötvös University
1518 Budapest 112, P.O. Box 32 (Hungary)
Fax: (+36) 1-209-0602
E-mail: csaszar@chem.elte.hu

0947-6539/03/0904-1008 $ 20.00+.50/0

proline, at the STO-3G and 6-31G restricted Hartree – Fock (RHF) level, was performed by Sapse et al.[14] This study used severe model restrictions, including the assumption of bond-length and -angle equalities: $N{-}C^{\alpha} = N{-}C^{\delta}$, $C^{\alpha}{-}C^{\beta} = C^{\gamma}{-}C^{\delta}$, and $N\text{-}C^{\alpha}\text{-}C^{\beta} = N\text{-}C^{\gamma}\text{-}C^{\delta}$ (see Figure 1 for labeling of the atoms of Pro), which result in an almost planar pyrrolidine ring. The corresponding results[14] quickly became outdated. With the rapid development of computer resources several more elaborate theoretical investigations followed: Ramek et al.[12] located ten distinct conformers of Pro at the 6-311++G** RHF level and determined reaction paths and potential barriers connecting stationary points on the PES of Pro; Császár et al.[15] reported structural results for 12 conformers of neutral proline, including relative energies obtained at correlated levels of electronic structure theory; and Stepanian et al.[17] determined 15 stable minima on the PES of Pro using density functional theory (DFT) and reassigned the vibrational spectra of unionized proline obtained from matrix-isolation IR spectroscopy.[10]

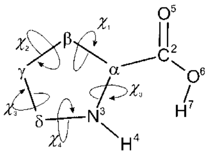

Figure 1. Labeling of the atoms of neutral proline.

In this study detailed structural and energetic results are presented for the conformers of neutral proline obtained at high levels of ab initio electronic-structure theory. Due to a more systematic search than before, altogether 18 minima were located and characterized on the PES of Pro. The relative energies of the conformers, computed using valence-only correlated treatments, are corrected, for the first time, for core correlation effects and effects due to special relativity. The pyrrolidine ring was characterized by using the pseudorotation concept.[22] Rotational and quartic centrifugal distortion constants were calculated to aid detection of as many conformers as possible by microwave (MW) and millimeter-wave (MMW) spectroscopy. Interpretation of the IR spectra of matrix-isolated Pro was extended by quadratic force fields obtained at the 6-311++G** DFT (B3LYP) level. Hydrogen-bonding features characteristic of the conformers of Pro were investigated with the atoms-in-molecules (AIM) approach of Bader.[23, 24]

Abstract in Hungarian: *Magas szintű ab initio elektronszerkezet számítások segítségével meghatároztuk a semleges* L-*prolin aminosav 18 konformerének pontos geometriai, relatív energia, forgási és centrifugális torzulási állandó, dipólusmomentum, harmonikus rezgési frekvencia és infravörös intenzitás adatait. Alacsony illetve közepes hőmérsékleten négy konformer jelenléte várható. A globális minimumhoz képest a második konformer csak* $2 \pm 2\,kJ\,mol^{-1}$*-lal magasabb energiájú, míg a harmadik és negyedik konformer majdnem degenerált és a globális minimumhoz képest* $7 \pm 2\,kJ\,mol^{-1}$*-lal magasabb energiájú. A négy konformer mindegyikében található H-kötés, az alacsony energiájú konformer-párban N···HO, míg a harmadik és negyedik konformerben NH···O típusú. Ezen konformer-párok a gyűrű beállásában különböznek. A konformerek relatív energiájának számításakor figyelembe vettük az elektronkorrelációt, melyet a teljes bázisra extrapoláltunk, valamint kiszámítottuk a törzs-vegyérték korreláció, valamint a relativisztikus effektusok hatását is. A pszeudorotáció elmélete segítségével tanulmányoztuk a prolin pirrolidin-gyűrűjét. A számított nagypontosságú forgási és negyedrendű centrifugális torzulási állandók, valamint a rezgési frekvenciák és infravörös intenzitások segíthetik az* L-*prolin konformereinek azonosítását forgási illetve rezgési spektroszkópiák segítségével. Az* L-*prolin kötésviszonyait - különös tekintettel a H-kötésekre- az "atoms-in-molecules" (AIM) módszer segítségével tanulmányoztuk.*

Methods of Calculation

Several Gaussian basis sets were selected for this study. Most of them contain both polarization and diffuse functions, as the need for including such functions in the basis set for calculations on hydrogen-bonded systems has long been recognized.[25] The 6-311++G** basis,[26] used in this study for geometry optimizations and subsequent quadratic force-field determinations, contains 239 contracted Gaussian functions (CGFs) for proline. The correlation-consistent, polarized-valence (aug)-cc-p(C)V*n*Z ($n = 2$ (D), 3 (T), 4 (Q), and 5) basis sets of Dunning et al.[27, 28] have been employed extensively for single-point energy calculations. (Note that only the augmented (aug) basis set contains diffuse functions, while tight functions, necessary for calculation of core correlation[29] and relativistic[30] effects are only part of core-polarized (C) basis sets.) For some valence-only focal-point[31] calculations, a mixed cc-pV(T/D)Z basis was used, which employs the cc-pVTZ basis set for all but the six CH_2 hydrogen atoms, for which the cc-pVDZ basis is used instead. For proline, the cc-pV(T/D)Z, cc-pVTZ, aug-cc-pVTZ, cc-pVQZ, and cc-pV5Z basis sets contain 303, 366, 575, 710, and 1223 CGFs, respectively. Only the pure spherical harmonics were employed in all basis sets.

Electronic wave functions were determined by the single-configuration, self-consistent field, restricted Hartree – Fock (RHF) method,[32–34] by second-order Møller – Plesset theory (MP2),[35] by coupled-cluster (CC) methods[36] including all single and double excitations (CCSD)[37] and, in some cases, an additional perturbative correction for contributions from connected triple excitations [CCSD(T)],[38] and by a density functional theory (DFT) approach usually abbreviated as B3LYP.[39] The T_1 diagnostic values of coupled cluster theory[40] are around 0.013 for the different conformers of proline, and this suggests that they can adequately be described by single-reference-based electron-correlation methods. The eight lowest 1s-like core orbitals were kept frozen in all post-Hartree – Fock treatments unless otherwise noted.

In this study the geometrical structures of the conformers of proline were optimized at the 3-21G RHF and 6-311++G** DFT(B3LYP) levels of theory, but only results obtained at the latter level are reported. The residual Cartesian gradients were in all cases less than 1×10^{-5} hartree bohr^{-1}. To determine whether the optimized structures correspond to minima on the PES of free proline, to avoid problems arising from the nonzero force dilemma,[41] and to obtain a quadratic force-field representation of the PES at the minima, the 6-311++G** DFT(B3LYP) quadratic force constants in Cartesian coordinates were determined at the respective optimized geometries by employing analytic second-derivative techniques.[42, 43]

Theoretical calculation of accurate relative energies for the conformers of proline followed the focal-point approach,[31, 44] which was employed successfully for determining the conformational preferences of neutral glycine[18, 19] and α-alanine.[15, 21] This procedure takes further advantage of the observation[18, 21] that in neutral amino acids in most cases higher-order [CCSD and CCSD(T)] contributions to the energy differences of the conformers investigated are either small or seem to cancel out, so that MP2 calculations with large basis sets should result in highly accurate relative energies. Extrapolation of the finite-basis RHF energies and MP2 energy increments to the complete basis set (CBS) limits were performed with the formulas

$$E_X = E_{CBS} + a\exp(-bX)$$

and

$$E_X = E_{CBS} + cX^{-3}$$

respectively,[31] where X is the cardinal number of the correlation-consistent basis set. For estimation of the core correlation energy correction two sets of cc-pCVTZ MP2 calculations were performed, one with the usual frozen-core approximation and one with all electrons correlated. Relativistic energy corrections were obtained at the all-electron cc-pCVDZ MP2 level within the mass-velocity and one-electron Darwin (MVD1) approximation,[45] which, for molecules composed of first-row atoms, should provide an excellent approximation of the full Dirac treatment.[30]

Electronic-structure calculations were performed with the program packages ACES II,[46] Gaussian 98,[47] and Turbomole.[48] For calculation of the quartic centrifugal distortion constants from quadratic force fields the package ASYM40[49] was employed.

The bonding patterns of the conformers of Pro, especially those associated with hydrogen bonds, were examined with tools provided by the atoms-in-molecules (AIM) theory,[23, 24] which leads to a wealth of information obtained principally from the molecular electron density $\rho(\boldsymbol{r})$. Intramolecular nonbonded effects were explored by analysis of the topological features of $\rho(\boldsymbol{r})$, and the atomic properties integrated in the basins, Ω, defined by the gradient vector field of the density. Eight criteria have been proposed for inter-[50] and intramolecular hydrogen bonds.[51] Existence of a hydrogen bond implies: a) presence of a bond critical point (BCP) with the proper $(3,-1)$ topology; b) the electron density at the BCP, ρ_b, lies in the range [0.002,0.04] au, that is, much lower than typical for covalent bonds; c) the Laplacian of $\rho(\boldsymbol{r})$ at the BCP is positive and within the range [0.015,0.15] au; d) mutual penetration of hydrogen and acceptor atoms; e) an increase in the net positive atomic charge, $q(\Omega)$, on the H atom; f) energetic destabilization of the H atom, $E(\Omega)$; g) decrease in the dipolar polarization of the H atom, $\mu(\Omega)$; and h) a decrease in the atomic volume of the H atom, $V(\Omega)$. For our optimized conformers, 6-311++G** MP2 electron densities were obtained with Gaussian 98.[47] The critical points were located and classified with the program Extreme,[52] and the atomic properties integrated in the basins Ω defined by the gradient vector field of $\rho(\boldsymbol{r})$ were calculated with the program Proaim.[53] The integral of $\nabla^2\rho$ over the basin of an H atom $L(\Omega)$ should be zero. In this work the values of $L(\Omega)$ are in the range of 2.1×10^{-6} to 1.2×10^{-4}, that is, sufficiently close to zero.

Results and Discussion

The stable conformers of proline can be built up by considering orientation (puckering) of the pyrrolidine ring, which is known to undergo pseudorotation,[22] orientation of the COOH group (Z vs E), orientation of the imino NH group (up or down with respect to the mean pyrrolidine ring), and torsions about the C$^{\alpha}$–C and C–O bonds. All conformers thus built belong to the C_1 point group. After careful consideration and using a large number of starting structures generated on the basis of the above structural considerations, we were able to locate 18 conformers on the PES of neutral proline at the 3-21G RHF and 6-311++G** DFT(B3LYP) levels. Figure 2 shows the numbering of the conformers, which reflects the extrapolated final relative energies of the distinct conformers obtained at 6-311++G** DFT(B3LYP) optimized geometries. Figures 1 and 2 also give pseudorotation parameters (see below) for each conformer.

Each conformer of free, neutral proline has a large number of independent geometry parameters, but structural results, in form of conformational averages and their standard deviations, are given in Table 1 only for bond lengths and bond angles. Table 2 lists valence-only relative energies for all 18 conformers of free proline investigated in this study at different levels of electronic-structure theory. Table 3 lists extrapolated valence-only relative energies plus small correction energies due to the effects of core correlation and special relativity. Table 4 lists theoretical rotational and quartic centrifugal distortion constants as well as total dipole moments for all the conformers with a view to future microwave (MW) and/or millimeter-wave (MMW) spectroscopic studies on free proline. Harmonic vibrational frequencies and infrared intensities are presented for the lowest energy conformers in Table 5. Characteristic AIM parameters are given in Table 6 for the most important hydrogen-bonding interactions.

Geometries: As expected, differences in bond lengths among the conformers of free proline are rather small. Standard deviations of characteristic C–C, C–N, C–O, and O–H single bond lengths are 0.01 Å or less (Table 1). Standard deviations of the N–H and C=O bond lengths are especially small (0.002 and 0.003 Å, respectively).

The standard deviations of the bond angles can also be regarded as small. There are, however, notable differences

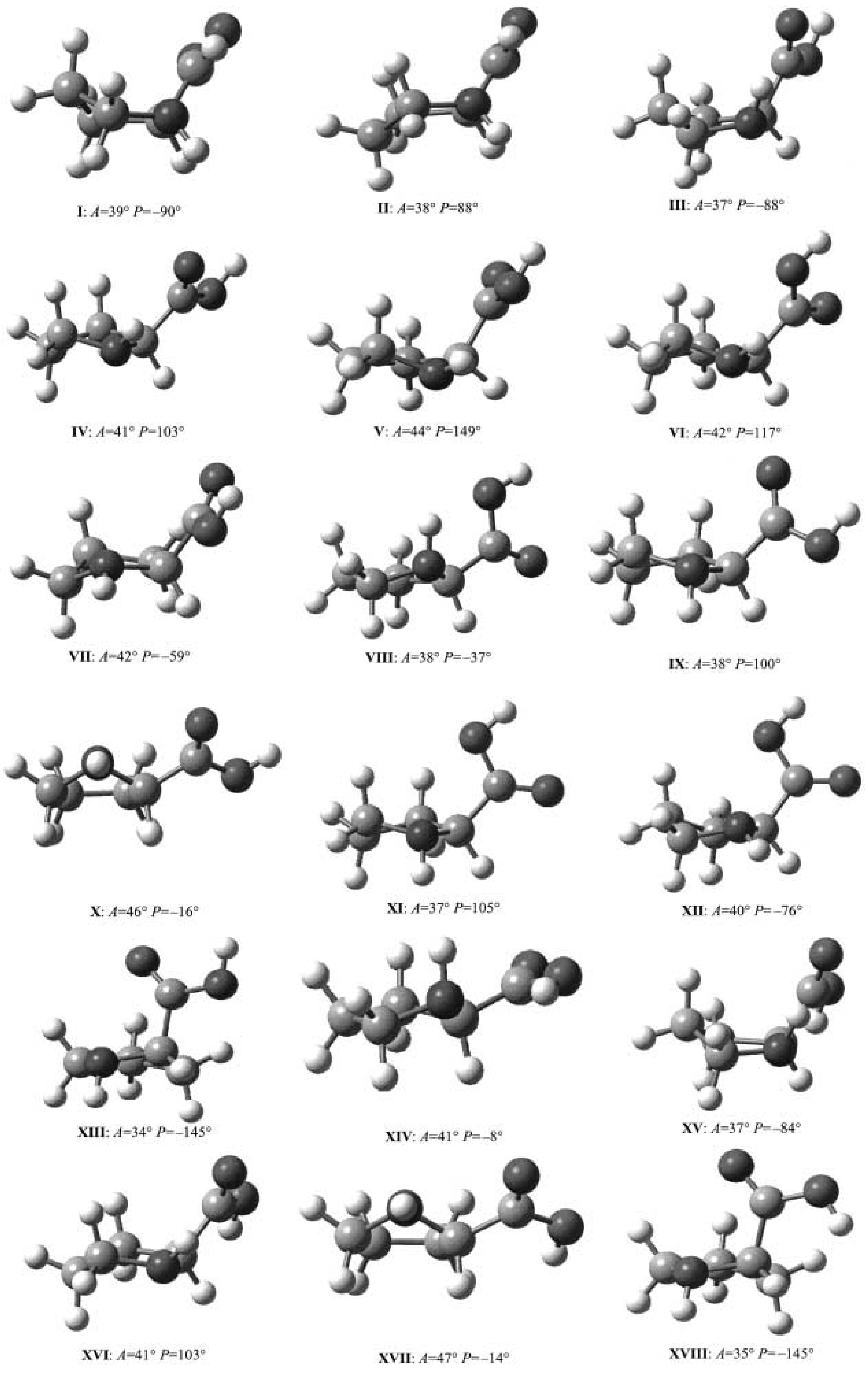
I: A=39° P=−90°
II: A=38° P=88°
III: A=37° P=−88°
IV: A=41° P=103°
V: A=44° P=149°
VI: A=42° P=117°
VII: A=42° P=−59°
VIII: A=38° P=−37°
IX: A=38° P=100°
X: A=46° P=−16°
XI: A=37° P=105°
XII: A=40° P=−76°
XIII: A=34° P=−145°
XIV: A=41° P=−8°
XV: A=37° P=−84°
XVI: A=41° P=103°
XVII: A=47° P=−14°
XVIII: A=35° P=−145°

Figure 2. Conformers of neutral L-proline numbered according to their relative energy order, with simple pseudorotation coordinates (see text).

among the bond angles, probably the most important being the dependence of the N-C$^{\alpha}$-C^{2}, N-C$^{\alpha}$-H, and C$^{\alpha}$-C^{2}-O^{6} angles on the orientation of the NH and OH groups. The N-C$^{\alpha}$-C^{2} angle varies between 107.4° (**IX**) and 116.8° (**VI**), the N-C$^{\alpha}$-H angle between 106.4° (**XIV**) and 112.8° (**X**), and the C$^{\alpha}$-C^{2}-O^{6} angle between 110.6° (**VII**) and 115.4° (**XV**). While part of this spread in bond angles should be attributed to sizable through-space repulsions, the *trans*-angle rule proposed by Räsänen et al.,[54] stating that "if in a conformer of a primary alcohol or amine a C−C or C−H bond is *trans* to an X−H bond

Table 1. Average bond lengths [Å] and bond angles [°] with standard deviations for neutral proline, obtained at the 6-311++G** B3LYP level.

Bond[a]	Average	Standard dev.	Angle[a]	Average	Standard dev.
$N-C^{\alpha}$	1.469	0.011	N-C^{α}-C^{2}	111.61	2.39
$C^{\alpha}-C^{\beta}$	1.556	0.009	N-C^{α}-H	110.62	1.55
$C^{\beta}-C^{\gamma}$	1.542	0.007	C^{2}-C^{α}-C^{β}	112.53	2.31
$C^{\gamma}-C^{\delta}$	1.538	0.009	C^{α}-$C^{2}{=}O^{5}$	124.97	1.10
$C^{\delta}-N$	1.474	0.007	C-$^{\alpha}C^{2}$-O^{6}	113.02	1.59
$C^{\alpha}-C^{2}$	1.526	0.008	O-C^{2}-O	121.98	1.12
$C^{2}{=}O^{5}$	1.203	0.003	C^{2}-O-H	107.67	1.84
$C^{2}-O^{6}$	1.356	0.007	C^{α}-C^{β}-C^{γ}	103.74	0.60
$O-H$	0.970	0.006	C^{β}-C^{γ}-C^{δ}	103.30	1.13
$N-H$	1.013	0.002	C^{γ}-C^{δ}-N	104.41	1.84
			C^{δ}-N-C^{α}	107.66	1.86
			N-C^{α}-C^{β}	105.58	1.35

[a] See Figure 1 for labeling of the atoms.

(X = O, N), the corresponding X-C-C or X-C-H angle will be considerably smaller than that for other configurations", is clearly valid for the conformers of proline. For example, the N-C-H angle of conformers **XIV** and **VIII**, in which the C−H bond is *trans* to the N−H bond, are 106.4 and 109.0°, respectively, while that of **I**, in which the C−H bond is in *cis* arrangement with the N−H bond, is 112.0°. The C^{α}-C^{2}-O^{6} angle of **VII**, in which the C−C bond is *trans* to the O−H bond, is 110.6°, while those of **XV** and **XVI**, in which the C−C bond is in *cis* arrangement with the O−H bond, are 115.4 and 115.3°, respectively, again in agreement with the *trans*-angle rule.

Changes in the C^{2}-O-H bond angle are of particular interest. While the *trans*-angle rule is clearly operative in this case, so that the C^{2}-O-H angle in conformers having an *E* COOH arrangement are larger than average, the smallest C^{2}-O-H angles are observed in conformers **I** and **II**, which also have an *E* COOH arrangement, due to the strong hydrogen bond in which the OH group participates. The difference between the smallest (**I**) and largest (**XVII**) C^{2}-O-H angle is more than 6°. Changes in the O-C^{2}-O bond angle can be rationalized by using the same arguments, but in this case strong hydrogen bonding increases the bond angle which is smallest in an *E* arrangement. The standard deviation of bond angles corresponding to linking the COOH group to the ring (the C^{2}-C^{α}-N and C^{2}-C^{α}-C^{β} angles) are larger than the standard deviation of the bond angles of the COOH group (C^{α}-C^{2}-O^{5}, C^{α}-C^{2}-O^{6}, C^{2}-O^{6}-H) and the bond angles of the ring. Among the bond angles of the ring, those involving N have larger standard deviations than the others.

The COOH group remains basically planar in all conformers investigated, and the largest deviation from planarity is a mere 2.9°. Interestingly, seven of the 18 conformers, and most importantly the conformer pair **I**/**II**, which have the same hydrogen-bonding characteristics but differ in ring puckering, have an *E* (*trans*) COOH arrangement, despite the fact that the energy of this form is some 20 $kJ\,mol^{-1}$ higher[21, 31] than that of *Z* (*cis*) form. Nevertheless, favorable hydrogen bonding can make up for this energy penalty.

The pyrrolidine ring in proline was calculated to be almost planar at the RHF level with the STO-3G and 6-31G basis sets.[14] Nevertheless, more elaborate electronic-structure results, especially those employing polarization functions in the basis set,[25] show that the ring is considerably puckered in all conformers. It is clear that in all conformers the character of the N atom is pyramidal.

The strength of the N⋯H−O and N−H⋯O hydrogen bonds, which are principally responsible for the increased stability of certain Pro conformers, is of considerable interest and is investigated in below in detail, using the AIM approach of Bader.[23] It is only noted here that a) the N⋯H−O hydrogen bond is part of a five-membered ring that is close to being planar in the case of a strong interaction; for example, for conformers **I** and **II** the largest deviations of individual dihedral angles from planarity are 0.7 and 2.9°, respectively; and b) the N−H⋯O=C hydrogen bond forms part of a five-membered ring which is never close to being planar; the largest deviations from planarity are 12 and 18° for conformers **III** and **IV**, respectively.

Pseudorotation: For a description of the puckering of the pyrrolidine ring of Pro, the pseudorotation concept,[22, 55–62] first introduced by Kilpatrick et al.[22] for cyclopentane, was employed in this study. According to the concept of pseudorotation the ring is involved, without producing any angular momentum, in a rotationlike movement that passes through envelope (E), twist (T), and half-twist (H) forms separated by low barriers. For the pyrrolidine molecule, these barriers were estimated on the basis of electronic structure calculations of different sophistication to be on the order of 2–4 $kJ\,mol^{-1}$.[55–57]

Pseudorotation coordinates can be defined in several ways. The original definition, due to Kilpatrick et al.,[22] as well as its generalization by Cremer and Pople,[58] requires specification of an appropriate mean plane and uses the displacements of each atom from that plane to define puckering coordinates. Geise, Altona, and Romers[59] and their followers use endocyclic torsion angles and thus avoid the need to define a mean plane. Dunitz[60] proved that in case of infinitesimal displacements of a pentagon from planarity there is a direct linear relationship between torsion angles and displacements, so that in this limit the pseudorotation coordinates derived from torsional angles can be rigorously related to those in the original definition of Kilpatrick et al.[22] Nevertheless, for finite displacements there are significant deviations from these linear relationships. The pseudorotation concept has been further modified by introduction of correction terms for describing nonequilateral rings.[57]

Our intent in using the concept of pseudorotation is the qualitative description of the puckering of the ring by grouping rings characteristic of the conformers into subtypes specified by pseudorotation coordinates. Therefore, we apply the simplest model with only two derived coordinates, a puckering amplitude *A* and a phase angle *P*. Amplitude *A* provides a maximum for any of the endocyclic torsional angles, while the phase angle *P* characterizes their ratios.

In its simplest form the pseudorotation coordinates *A* and *P* can be derived from endocyclic torsion angles χ_0, χ_1, χ_2, χ_3, and χ_4 (see Figure 1) and can be obtained by solving Equations (1) and (2).

$$A\cos P = \chi_0 \tag{1}$$

$$A\sin P = \frac{\chi_1 - \chi_2 + \chi_3 - \chi_4}{-2\,(\sin 4\pi/5 + \sin 2\pi/5)} \tag{2}$$

As can be seen in Table 2 and Figure 2, the puckering amplitude has a well-defined value of $A = 40 \pm 4^\circ$. This means that the maximum deviation of ring atoms from the mean plane is approximately the same for all conformers. This result is very similar to that observed for the pyrrolidine molecule.[56] Consequently, the phase angle P, which ideally has a periodicity of 36°, is sufficient to describe the different ring structures of neutral proline.

Although several nomenclatures have been suggested in the literature to distinguish conformations along a pseudorotation path,[61] in Figure 2 we simply use the phase angle P to distinguish the conformers. The rings of the most stable pair of conformers **I** and **II**, with phase angles of -90 and $+88^\circ$, are clearly of the envelope type and mark the origin of the pseudorotation path. All other conformers can be grouped into types E, T, and H based on the periodicity of the five-membered ring and their phase angle P.

In the literature statements such as "the conformation analysis of pyrrolidine ring would eventually lead to a better molecular understanding of the relationship between internal motions and functions of proteins"[63] and "the structure of [the] pyrrolidine [molecule] is of great interest since this nitrogen heterocycle is a basic building block of amino acids and peptides"[55] can be found. The pseudorotation profile of pyrrolidine was first studied, at the 4-21G(N*) RHF level, by Pfafferott et al.[55] who compared their calculated results with structures obtained by gas-phase electron diffraction. The pseudorotation profile of the proline peptide analogues For-L-Pro-NH_2 and Ac-L-Pro-NH_2 were investigated by Hudáky et al.[64] and Ramek et al.,[65] respectively. Therefore, we are in the position to address the above comments in some detail. Naturally, given the basic similarity of pyrrolidine to proline and For-L-Pro-NH_2, pseudorotation profiles in all three compounds are expected to be similar. However, even though all studies concerning the barrier to pseudorotation agree on almost free pseudorotation, the lowest energy minima on the PESs of the three compounds show rather different conformational characteristics. The pyrrolidine molecule adopts an envelope conformation with axial NH group.[55, 66] This spatial arrangement is not preferred in the more stable conformers of Pro and For-L-Pro-NH_2. The E-type ring conformations of the most stable Pro conformers (**I** and **II**) are not preferred in Ac-L-Pro-NH_2 either. Therefore, it is clear that the ring conformations in the minima on the PES of Pro are determined mostly by the interactions of the NH and COOH moieties, and the conformational preference of the ring is the result of a second-order effect.

Following a notation applicable for peptides,[15, 67] in which ϕ and ψ are defined as the N-C^α-C′-O(H) and H^N-N-C^α-C′ torsion angles, respectively, the conformers of Pro can be labeled as α_L, ε_L, and γ_L, while the other six idealized conformers are missing. It is notable that in simple peptide models For-Xxx-NH_2 the α_L and ε_L conformers are usually missing.[15, 68] The appearance of these conformers in Pro prove that the concept of multidimensional conformational analysis (MDCA)[69] is of considerable general utility. Nevertheless, there are certain problems with this simple technique: a) it is not capable of predicting the stable conformers of Pro, as it predicts 24 conformers but considerably fewer are found (this is a rather general observation when MDCA is used to generate structures for geometry optimizations, but it causes no real difficulty); and b) disturbingly, conformer pairs **XIII/X** and **XVIII/XVII** belong to the same MDCA-predicted forms, and thus if the geometry optimizations were started according to the conventions of MDCA one of the minima would have been missed.

In summary, no preferred conformation exists for the pyrrolidine ring which could be transferred in the series pyrrolidine – proline – For-L-Pro-NH_2 – Ac-L-Pro-NH_2 and further to peptides and proteins.

Energies: The high-quality relative energy results presented in Tables 2 and 3 clearly confirm that at normal temperatures only four conformers of neutral proline should coexist, namely, conformer pairs **I/II** and **III/IV**. The conformers within the pairs have similar bonding patterns, in particular the conformer pair **I/II** differ only in the puckering of the E-type pyrrolidine ring.

Table 2 reveals disturbing failures of RHF theory in predicting the energy order of the 18 conformers of proline: a) at the RHF level of theory, independent of the basis set used, conformer **IV** is predicted to be more stable than the true global minimum, conformer **I**, by some 6 kJ mol^{-1}, and even **III** is predicted to be more stable than **I** by almost 4 kJ mol^{-1}; b) where adequate pairs can be found, again independent of the basis set employed, the pucker-up conformer is always more stable than its pucker-down analogue at correlated levels, but in the case of conformer pairs **III/IV** and **XV/XVI** RHF theory reverses the stability order; and c) the RHF energy spread is considerably smaller than the MP2 spread, which is somewhat suprising in the light of opposite results for glycine[18] and α-alanine.[21] Inclusion of electron correlation in the energy calculation, even at the simplest B3LYP level, remedies the problems of RHF theory, and there is basic agreement between B3LYP, MP2, CCSD, and CCSD(T) energy predictions, although the relative energies of the conformer pairs **III/IV**, **VI/VII**, and **XV/XVI** may be slightly different at the B3LYP level of theory.

For Pro much larger basis sets can be afforded at the RHF and MP2 levels than at coupled cluster levels. Therefore, the ultimate (complete basis set) valence-only energies can be calculated, following recommendations of the focal-point approach,[31, 44] by extrapolating the RHF and MP2 energies to the CBS limit and appending coupled-cluster energy increments to the extrapolated results. RHF and MP2 results extrapolated[31, 70] to the CBS limit can be found in Table 2. Although convergence of the relative valence-only energies obtained is somewhat uneven (most likely due to the lack of diffuse functions in the basis set), the accuracy of the extrapolated results should be better than 0.2 and 1.0 kJ mol^{-1} at the RHF and MP2 levels, respectively. The coupled-cluster

Table 2. Relative energies [kJ mol^{-1}] of conformers of proline.[a]

Conformer	6-311++G** B3LYP	cc-pV(T/D)Z				cc-pVTZ		aug-cc-pVTZ		cc-pVQZ		cc-pV5Z	CBS	
		RHF	MP2	CCSD	CCSD(T)	RHF	MP2	RHF	MP2	RHF	MP2	RHF	RHF	MP2
I	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
II	1.98	1.34	2.95	2.15	2.24	1.47	3.25	1.09	2.77	1.14	2.96	1.04	1.00	2.85
III	8.56	– 4.00	8.26	5.03	6.96	– 3.74	8.59	– 3.52	8.35	– 3.53	8.33	– 3.59	– 3.64	7.89
IV	7.66	– 6.53	8.92	5.02	7.61	– 6.37	8.96	– 6.28	8.78	– 6.30	8.59	– 6.40	– 6.48	8.08
V	14.38	2.87	15.44	11.88	14.00	2.84	15.16	3.32	15.15	3.26	15.01	3.24	3.20	14.54
VI	14.95	2.85	16.37	13.10	15.36	2.91	16.15	2.99	15.97	3.03	15.98	2.94	2.85	15.59
VII	14.83	0.93	17.77	12.98	15.88	0.82	12.10	0.91	17.24	0.84	17.01	0.71	0.62	20.37
VIII	16.34	6.34	17.94	14.52	16.23	6.52	18.25	6.02	17.18	6.21	17.63	5.98	5.84	17.02
IX	16.39	4.83	18.88	14.73	16.88	4.85	18.67	4.87	18.39	4.87	18.36	4.72	4.61	17.86
X	17.06	3.01	19.42	14.75	17.69	3.25	19.79	2.83	18.76	2.88	18.76			
XI	18.08	5.64	19.89	15.74	18.01	5.74	19.91	5.35	19.15	5.48	19.50			
XII	18.25	3.92	20.52	16.08	19.01	3.94	20.29	3.60	19.63	3.72	19.80			
XIII	20.52	9.59	20.82	17.48	19.31	9.47	20.22	9.43	19.79	9.51	20.14			
XIV	23.20	17.12	21.47	18.60	19.17	17.29	22.01	17.26	22.08	17.17	22.10			
XV	29.80	19.43	27.77	24.85	26.02	19.57	27.93	18.93	27.00	18.90	27.08			
XVI	28.91	17.16	28.30	24.80	26.59	17.21	28.23	16.59	27.59	16.51	27.44			
XVII	39.70	28.67	40.38	36.11	38.16	28.74	40.61	27.58	38.98	27.56	39.10			
XVIII	45.65	36.70	44.27	41.13	42.27	36.55	43.55	35.61	42.19	35.74	42.77			

[a] See Figure 2 for the shapes of the conformers. The reference geometries employed for all energy calculations have been optimized at the 6-311++G** DFT(B3LYP) level. Zero-point vibrational energy corrections were not added. Note that conformer **VII** may not be a minimum at higher levels of theory.

Table 3. Predicted relative energies, including auxiliary energy corrections, for the lowest energy conformers of proline.[a]

Conformer	CBS CC	Core correlation[b]	Relativistic[c]	Final[d]
I	0.000	0.000	0.000	0.000
II	2.152	0.065	– 0.020	2.199
III	6.591	0.088	– 0.037	6.642
IV	6.779	0.031	– 0.007	6.803
V	13.093	0.130	– 0.033	13.190
VI	14.582	0.110	– 0.012	14.680
VII	18.483	0.074	– 0.013	18.544
VIII	15.309	0.246	– 0.076	15.479
IX	15.863	0.166	– 0.058	15.971
X		0.091	– 0.018	
XI		0.181	– 0.052	
XII		0.060	0.008	
XIII		0.193	– 0.066	
XIV		0.268	– 0.097	
XV		0.150	– 0.045	
XVI		0.084	– 0.014	
XVII		0.135	– 0.018	
XVIII		0.275	– 0.070	

[a] All values in kJ mol^{-1}. CBS CC = complete basis set CCSD(T). [b] All values obtained at the cc-pCVTZ MP2 level. [c] All values obtained at the cc-pCVDZ MP2 level. [d] The expected uncertainty of the final computed energies is 2kJ mol^{-1}.

energy increments should have an even better accuracy; therefore, the valence-only CBS CCSD(T) results presented in Table 3, obtained by adding the CBS MP2, cc-pV(T/D)Z CCSD, and CCSD(T) energy increments to the CBS RHF values, should have an accuracy of better than 2 kJ mol^{-1}.

To move beyond valence-only energies one has to consider several small correction terms. It has been shown in many recent high-level computational studies (see, e.g., refs. [29–31, 44, 71]) that the largest corrections to valence-only energies come from effects due to core–core and core–valence interactions and special relativity. Therefore, we computed these relative energy increments for the conformers of Pro (Table 3).

Based on high-quality results for formic acid,[21, 31] core correlation effects for the relative energies of the conformers of Pro were expected to be small. This expectation is confirmed by the present calculations, in which core correlation effects change the relative energies by less than 0.28 kJ mol^{-1}. Relativistic energy corrections turned out to be less than 0.1 kJ mol^{-1} in a relative sense for all conformers. Therefore, these energy corrections may be ignored for most conformational studies of biomolecules, even at the level of precision provided by state-of-the-art electronic-structure calculations.

It is clear from the results of Table 3 that MP2 theory, similar to the cases of neutral glycine[18] and alanine,[21] provides an accurate description of the energetic preferences of the conformers of proline. It is perhaps even more important for future studies that the 6-311++G** DFT(B3LYP) level of theory provides relative energies very close to those obtained at the extrapolated CBS CCSD(T) level. Our definitive energy results are in some contrast to a statement of Stepanian et al.,[17] who concluded that "unlike in other amino acids, the DFT/B3LYP and MP2 methods are not capable of predicting correct relative stabilities of the proline conformers".

Rotational spectra: Future identification of proline conformers by microwave (MW) and millimeter-wave (MMW) spectroscopy is greatly aided by the availability of accurate theoretical predictions for rotational and quartic centrifugal distortion (QCD) constants.[72] Table 4 lists theoretical rotational and QCD constants and dipole moments calculated at the 6-311++G** DFT(B3LYP) level.

Unfortunately, no experimental data are available for comparison with the calculated values. Nevertheless, observation of conformers **I** and **II** of proline is aided by their substantial dipole moments. Since the rotational constants of **I** and **II** differ substantially, it should be straightforward to distinguish the rotational lines of the two conformers.

Table 4. Rotational and quartic centrifugal distortion constants and total dipole moments for conformers of proline.[a]

Conformer	Rotational constants [MHz]			Quartic centrifugal distortion constants [kHz]					μ_T [D]
	A_e	B_e	C_e	Δ_J	Δ_{JK}	Δ_K	δ_J	δ_K	
I	3730.15	1641.52	1369.52	0.51	−2.07	4.80	0.10	0.54	5.86
II	4030.41	1551.28	1246.90	0.37	−1.36	5.56	0.11	0.95	5.98
III	3929.01	1541.38	1330.23	1.10	−5.88	13.47	0.23	2.10	1.96
IV	4089.31	1520.74	1242.87	0.45	−2.40	6.90	0.06	0.87	1.79
V	3608.04	1681.65	1351.83	0.71	1.05	2.22	0.02	0.92	1.36
VI	3972.84	1549.09	1306.52	1.09	−4.22	11.71	0.05	5.43	2.05
VII	4050.00	1516.99	1241.39	0.89	−3.58	8.92	−0.15	2.65	2.29
VIII	4217.00	1407.83	1282.82	0.91	−3.90	10.50	0.08	5.76	2.63
IX	4129.26	1435.18	1310.79	0.56	−1.30	3.59	−0.06	7.32	2.43
X	4309.74	1455.52	1156.98	0.13	0.44	0.26	0.02	1.29	2.25
XI	4101.27	1434.65	1372.62	0.59	−1.69	3.72	0.06	10.80	2.73
XII	4016.93	1495.96	1372.58	0.69	−1.43	2.89	−0.04	14.21	2.45
XIII	3702.83	1585.86	1492.12	0.82	−2.31	5.08	0.22	9.80	2.60
XIV	4210.24	1491.25	1157.23	0.08	0.05	0.61	0.02	0.17	4.78
XV	3935.20	1528.37	1328.51	1.87	−10.88	23.64	0.42	2.56	4.02
XVI	4061.52	1521.36	1243.90	0.46	−2.36	6.80	0.08	1.01	4.02
XVII	4273.02	1458.54	1151.65	0.11	0.44	0.12	0.02	0.93	5.01
XVIII	3666.92	1600.38	1493.19	0.72	−1.63	3.85	0.16	6.50	5.28

[a] The quartic centrifugal distortion (QCD) data refer to A-reduction. The rotational constants given in the table refer to equilibrium values.

The theoretical estimates given for the yet unobserved conformers should be most valuable in searching for these conformers by MW and MMW spectroscopy.

Vibrational spectra: The second-derivative results obtained at the 6-311++G** B3LYP level confirm that all optimized structures **I**–**XVIII** correspond to minima on the PES of Pro. The plethora of information contained in the subsequently obtained vibrational frequencies and infrared intensities, part of which is presented in Table 5, for the Pro conformers allows interpretation of carefully executed experimental investigations of the vibrational spectra of neutral proline.

To the best of our knowledge only one relevant experimental study has been performed.[10, 17] Reva et al.[10] obtained the vibrational spectra of matrix-isolated Pro (proline vapor from a Knudsen cell at 152 °C was deposited in an Ar matrix). They proved that Pro exists in its neutral form in the matrix (and hence in the gas phase). More importantly, while in the first report[10] interpretation of the spectra was of dubious quality, in a recent study[17] some of the same authors

Table 5. Normal modes and the corresponding scaled IR frequencies $\tilde{\nu}$ [cm^{-1}] and intensities A [km mol^{-1}] dominated (>70%) by a single internal coordinate for conformers **I**–**IV** of proline.[a]

Mode	**I**		**II**		Exptl.		**III**		**IV**		Exptl	
	A	$\tilde{\nu}$	A	$\tilde{\nu}$	I_{obsd}	$\tilde{\nu}$	A	$\tilde{\nu}$	A	$\tilde{\nu}$	I_{obsd}	$\tilde{\nu}$
NH str	3399.2	7.2	3394.7	6.6	3393	0.54	3367.2	25.2	3385.1	17.5	3369	0.50
OH str	3237.9	351.3	3267.3	326.8	3025	5.20	3585.0	74.5	3583.8	70.0	3559	2.69
$C^{\beta}H_2$ as str	2976.3	8.8	2974.5	16.6			2963.7	23.2	2965.6	29.0		
$C^{\beta}H_2$ s str							2913.2	30.0				
$C^{\beta}H-C^{\gamma}H$ as str[b,c]	2908.0	18.6							2904.8	19.0		
$C^{\beta}H-C^{\gamma}H$ s str[b,c]	2900.8	11.0							2900.8	15.5		
$C^{\alpha}H-C^{\beta}H$ as str[b]			2916.5	15.0								
$C^{\alpha}H-C^{\beta}H$ s str[b]			2908.2	9.5								
$C^{\alpha}H$ str	2923.0	26.7					2890.0	21.4	2868.7	27.3		
$C^{\gamma}H-C^{\delta}H$ as str[b]			2892.5	34.4			2945.8	41.6				
$C^{\gamma}H-C^{\delta}H$ s str[b]			2885.0	28.3			2940.7	28.5				
$C^{\gamma}H_2$ as str	2958.2	28.5							2952.2	28.1		
$C^{\gamma}H_2$ s str							2895.2	20.9				
$C^{\delta}H_2$ as str	2932.2	33.4							2922.0	54.3		
$C^{\delta}H$ str[d]	2858.3	61.1					2867.4	57.7	2820.0	74.9		
C=O str	1791.4	380.5	1792.6	390.0	1789	7.22	1759.9	295.3	1761.0	308.7	1766	6.03
$C^{\delta}H_2$ sci	1480.3	2.7	1475.0	0.7	1488	0.05	1472.7	0.3	1482.8	2.2	1488	0.05
$C^{\gamma}H_2$ sci									1444.7	4.8	1451	0.52
NH wag			1398.9	29.0	1405	1.40			1423.8	12.3	1412	2.28
$C^{\delta}H_2$ wag							1326.1	3.4			1330	0.10

[a] Contributions of internal coordinates are given according to the total energy distribution (TED).[68] Abbreviations employed for the description of normal modes: str = stretching, as = antisymmetric, s = symmetric, sci = scissor, wag = wagging. The CH str region cannot be assigned unambiguously. The observed frequencies are 2984, 2959, 2934, 2916, 2885, 2865, 2846 cm^{-1}. I_{obsd} are taken from Table 6 of ref. [17] and refer to experimental relative integral intensities measured for single bands or for the group of merged bands. [b] In the antisymmetric/symmetric CH str modes presented in the table the H atom on the same/opposite side as the atom out of the mean ring plane provides the larger contribution. [c] In conformers **I** and **IV** different H atoms attached to C^{β} and C^{γ} are involved in this mode. [d] Individual CH str, the H atom opposite to the atom out of the mean ring plane moves.

presented a rather careful analysis based on the original experimental data and a theoretical harmonic vibrational analysis performed at the aug-cc-pVDZ B3LYP level. Nevertheless, possible formation of dimers of the highly polar proline molecules in the matrix was not discussed.

To complement the revised vibrational analysis of Reva et al.,[17] in Table 5 we present selected theoretical vibrational results (scaled harmonic frequencies and intensities) for the four most stable conformers of Pro, which are likely to be present in the matrix according to our relative energy data. The selection criteria for inclusion of the normal modes in Table 5 was simply that a given normal mode should be dominated (>70%)[73] by a single internal coordinate. Scaling of the quadratic force field $\boldsymbol{F}$ employed the formula[74, 75]

$$F_{ij}^{\text{scaled}} = \sqrt{\gamma_i \gamma_j} F_{ij}^{\text{unscaled}}$$

and the following scale factors: 0.96 for the C=O stretch, 0.91 for all other stretches, and 0.94 for all bends. The scale factors 0.96, 0.91, and 0.94 were optimized for the observed C=O frequencies,[17] OH and NH stretching modes,[17] and CH_2 scissor and NH wagging modes,[17] respectively. This simplified scaling of the quadratic force field is expected to result in calculated vibrational band origins in error of no more than about 20 cm^{-1}.

The following important conclusions can be drawn from the frequency and intensity results obtained in this study: a) Based on the positions of the NH stretching vibrations all four conformers could in principle be distinguished in the vibrational spectra of Pro. Nevertheless, due to the small calculated intensities, this may be problematic in the case of conformers **I** and **II**. In fact, Reva et al.[17] reported only two weak bands (3393 and 3369 cm^{-1}) in this region. b) The very similar position and intensity of the OH stretching vibrations in conformers **III** and **IV** preclude their distinction in a low-resolution vibrational spectrum. The fact that Reva et al.[17] observed bands at 3559 and 3545 cm^{-1} is a clear proof of the presence of the conformer pair **III/IV** in the matrix. The origin of the relatively large observed splitting is unclear to us. Due to strong hydrogen bonds the OH stretching vibrations of the conformer pair **I/II** lie some 350 cm^{-1} lower, in fact lower than the NH stretching vibrations of conformers **I**–**IV**. These vibrations have very large calculated intensities and therefore it is somewhat strange that they were not detected by Reva et al.[17] at around 3200 cm^{-1}. A possible explanation is the expected broadness of the OH stretching bands. c) The region made up of CH stretching motions involves interesting features. In several cases the individual CH stretching motions of neighboring atoms couple strongly with each other, both in the antisymmetric and symmetric stretches. This coupling of CH stretches attached to different C atoms means that the difference between the appropriate antisymmetric and symmetric stretches can be as small as 4 cm^{-1}. Furthermore, some individual (>90%) CH stretching motions are observed. d) There is almost no difference in the very strong C=O stretching vibrations within the pairs **I/II** and **III/IV**, although the difference between the pairs is a substantial 32 cm^{-1}. Therefore, the C=O stretching region should be clearly indicative of the presence of the less stable conformer pair **III/IV**. Indeed, Reva et al.[17] observed two very intense bands in this region (1789 and 1766 cm^{-1}), separated by 23 cm^{-1}. This is further clear proof that both conformer pairs were present in the matrix they studied. e) According to our calculations, the CH_2 scissor motion appears at about the same frequency in all four conformers and is thus characteristic of the presence of Pro. However, the low intensity of these bands makes this observation of little practical use. f) Finally, the NH wag vibration appears as an almost pure normal mode in the pucker-down conformers **II** and **IV**, while it is strongly mixed with other internal coordinates in the pucker-up conformers **I** and **III**.

Future experiments involving variation of the temperature of the Knudsen cell and photochemical excitation of the matrix-isolated proline molecules should yield even more reliable band assignments and important additional data on the dynamical structure of the lowest energy conformers of proline.

Atoms in molecules (AIM): The results presented in Table 6, based on the AIM theory, allow us to draw the following conclusions about the different possible hydrogen bonds of Pro: a) For the $H^7 \cdots N^3$ hydrogen bonds, found in the conformer pair **I/II**, the BCPs were easily found, and all eight criteria mentioned above are satisfied. In conformers **I** and **II** the hydrogen bonds seem to have the same strength. This result is expected since the structures differ only in their ring puckering. b) For $H^4 \cdots O^5$ hydrogen bonds, present in conformer pair **III/IV**, the same statements hold. These hydrogen bonds seem to have approximately the same strength. This is again expected, as these conformers differ only in their ring puckering. For the conformer pair **XV/XVI** the AIM results suggest the existence of a relatively strong hydrogen bond. The energy difference between conformers **III** and **XV**, as well as between **IV** and **XVI**, simply reflect the energy difference between the *Z* and *E* arrangements of the COOH group. The hydrogen-bond strengths show excellent correlation with the hydrogen-bond length for $H^7 \cdots N^3$ and $H^4 \cdots O^5$ hydrogen bonds. c) No hydrogen bonds were found corresponding to the *Z* arrangement of the COOH group; therefore, its increased stability cannot be explained in this way. d) For $H^4 \cdots O^6$ hydrogen bonds, characteristic of conformers **V** and **VIII**, BCPs were found and all eight criteria are satisfied. The shorter hydrogen-bonding distance in **VIII** is clearly characteristic of a larger hydrogen-bond effect. No BCPs were found for conformers not mentioned so far, and this suggests that they do not have hydrogen bonds.

Conclusion

The detailed high-level ab initio structural investigation of L-proline allows us to draw the following conclusions about this neutral amino acid.

In contrast to simpler amino acids like glycine and α-alanine, only a few low-energy conformers exist on the PES of neutral proline. Only the conformer pairs **I/II** and **III/IV** (see Figure 2) have relative energies of less than 10 $kJ\,mol^{-1}$ relative to the global minimum **I**.

Table 6. Properties of [3, – 1] bond critical points of selected conformers and atomic integrated properties of hydrogen atoms in hydrogen bonds.[a]

Bond type	Conformer	Distance/Å	ρ_b	$\nabla^2\rho_b$	$q(\Omega)$	$E(\Omega)$	$\mu(\Omega)$	$V(\Omega)$
					integrated atomic properties of H^7			
$H^7\cdots N^3$	**I**	1.877	0.040	0.108	0.6145	– 0.3414	1.1411	15.7943
	II	1.898	0.038	0.107	0.6159	– 0.3412	0.1411	15.9538
	XIV[b]	2.138	0.022	0.084	0.6048	– 0.3475	0.1636	19.1474
	VIII	3.684	no BCP	no BCP	0.6021	– 0.3532	0.1609	21.6907
$H^7\cdots O^5$	**VI**	2.286	no BCP	no BCP	0.5990	– 0.3549	0.1625	21.9015
	XVII	3.012	no BCP	no BCP	0.5827	– 0.3675	0.1650	21.5794
					Integrated atomic properties of H^4			
$H^4\cdots O^5$	**III**	2.248	0.017	0.070	0.3778	– 0.4792	0.1894	28.4904
	IV	2.363	0.017	0.064	0.3762	– 0.4777	0.1839	29.6748
	XV	2.225	0.018	0.073	0.3833	– 0.4759	0.1797	28.0751
	XVI	2.338	0.014	0.055	0.3821	– 0.4747	0.1817	29.2649
	X	2.761	no BCP	no BCP	0.3443	– 0.4974	0.1865	33.1957
$H^4\cdots O^6$	**V**	2.363	0.015	0.067	0.3707	– 0.4810	0.1837	30.2565
	VIII	2.273	0.015	0.073	0.3649	– 0.4852	0.1859	29.8856
	XVIII	4.310	no BCP	no BCP	0.3498	– 0.4920	0.2003	32.9870

[a] See text for definition of headings. For each possible hydrogen-bond type investigated, $H^7\cdots N^3$, $H^7\cdots O^5$, $H^4\cdots O^5$, and $H^4\cdots O^6$, reference values are given for a conformer which does not have that particular hydrogen-bond interaction. [b] $L(\Omega)$ is unusually large, -3.1×10^{-3}.

The conformational flexibility of proline is illustrated by the fact that we found, at the 6-311++G** B3LYP level, 18 minima on the PES of proline. Considering all 18 conformers, standard deviations of characteristic single-bond lengths are 0.01 Å or less. Standard deviations of bond angles are also small, especially if the effect of the *trans*-angle rule[54] is taken into account.

It is clear from our detailed comparisons that the ring conformations in the minima on the PES of proline are determined mostly by the interactions of the NH and COOH moieties, and the conformational preference of the pyrrolidine ring is the result of a second-order effect. No preferred conformation exists for the pyrrolidine ring which could be transferred in the series pyrrolidine – proline – For-L-Pro-NH_2 – Ac-L-Pro-NH_2 and further to peptides and proteins.

The lowest energy conformer pair **I/II** of Pro has the largest total dipole moment, close to 6 D, which should make detection of these conformers by MW and MMW spectroscopy straightforward. The rotational and quartic centrifugal distortion constants presented in this study should greatly aid future experimental studies.

Similarly to previous studies on neutral amino acids,[18, 21] RHF energy predictions can be highly inaccurate, especially for the lowest energy conformers of Pro. Nevertheless, it is reassuring that the simple hybrid DFT functional B3LYP is able to recover the most important structural and energetic features on the PES of Pro, and this suggests that structural investigations on much larger peptides and even proteins can reliably be performed with DFT theory. Furthermore, inclusion of electron correlation in the relative energy calculations, even at the simplest B3LYP level, remedies the problems of RHF theory, and there is basic agreement between B3LYP, MP2, CCSD, and CCSD(T) energy predictions, although the relative energies of the conformer pairs **III/IV**, **VI/VII**, and **XV/XVI** may be slightly different at some of these levels of theory.

Core correlation and relativistic effects for the relative energies of the conformers of Pro are small, consistently less than 0.3 and 0.1 kJ mol^{-1}, respectively. Therefore, for most conformational studies of biomolecules these effects can be ignored at the level of precision provided by even state-of-the-art electronic-structure calculations.

Our calculations of the vibrational spectra of Pro, at the 6-311++G** B3LYP level, complement a previous study of Reva and et al.[17] and confirm that they observed both conformer pairs **I/II** and **III/IV** in their matrix-isolation study.

Our AIM results confirm the existence of hydrogen bonds of types $H^7\cdots N^3$, $H^4\cdots O^5$, and $H^4\cdots O^6$ in the conformers of Pro. The strengths of these hydrogen bonds, obtained by mathematical analysis of the electron density, correlate well with the length of the hydrogen bonds.

Appendix

Some time ago Tomimoto and Go (TG)[62] developed an analytical theory of pseudorotation in constrained five-membered rings. To describe geometrical variations within the ring TG used (constrained) bond lengths, valence bond angles, and torsional (dihedral) angles. During derivation of algebraic equations necessary to describe pseudorotation TG presented intermediate results (see Eqs. (1) – (3) in ref. [62]) in detail which, as shown here, could have been presented in a general and still simple closed form. Furthermore, derivation of the resulting equations with the aid of sophisticated computer algebra systems (CAS) such as Mathematica,[76] applied in the present study, is straightforward and allows generalization of the results of TG to six- and (perhaps) n-membered rings with arbitrary constraints.

In an unrelated paper,[77] concerned with the general analytical form of vibrational kinetic energy operators of sequentially bonded molecules in valence stretch (r), bend (θ), and torsion (τ) internal coordinates, the vector position of a fourth atom A_{i+1} in terms of the position of earlier atoms A_i, A_{i-1}, A_{i-2} and the new internal coordinates r_i, θ_{i-1}, and τ_{i-2} was given as Equation (A1)

$$\begin{aligned}\underline{A_{i+1}} = \underline{A_i} - r_i\cos\Theta_{i-1}\overrightarrow{A_{i-1}A_i} + r_i\sin\Theta_{i-1}\sin\tau_{i-2}\frac{\overrightarrow{A_{i-1}A_i}\times\overrightarrow{A_{i-1}A_{i-2}}}{\sin\Theta_{i-2}} \\ + r_i\sin\Theta_{i-1}\cos\tau_{i-2}\frac{\overrightarrow{A_{i-1}A_{i-2}} - \cos\Theta_{i-2}\overrightarrow{A_{i-1}A_i}}{\sin\Theta_{i-2}}\end{aligned} \tag{A1}$$

where $\overrightarrow{A_iA_{i-1}}$ and $\overrightarrow{A_iA_{i+1}}$ are appropriate unit vectors, and definition of the atoms and the internal coordinates employed is given in Figure 3. This simple analytical expression providing all atom–atom distances as functions of $\boldsymbol{r}$, $\boldsymbol{\Theta}$, and $\boldsymbol{\tau}$ results in Equations (1)–(3) of TG and allows extension of the results to n-membered rings. Using this formula one could simply define the geometry of a general five-membered ring in terms of four bond stretches, three bends, and two torsions (i.e., $9 = 3\times 5 - 6$ coordinates). The only difference between sequentially bonded open-chain molecules and the ring molecules treated by TG is that in the latter case ring-closure constraints exist. An example is presented in Equations (4) and (5) of TG: an idealized cyclopentane-type ring with equal bond lengths as the constraint. Use of Equation (A1) in a CAS program immediately yields the distance r_{15} in terms of the nine independent internal coordinates [Eq. (A2)]

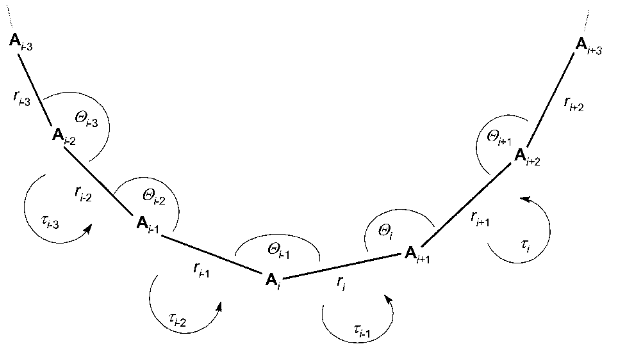

Figure 3. Definition of the atoms and the internal coordinates r, Θ, and τ.

$$\begin{aligned} r_{15}^2 = {} & r_1^2 + r_2^2 + r_3^2 + r_4^2 - 2r_1r_2\cos\Theta_1 - 2r_2r_3\cos\Theta_2 - 2r_3r_4\cos\Theta_3 \\ & + 2r_1r_3\cos\Theta_1\cos\Theta_2 + 2r_2r_4\cos\Theta_2\cos\Theta_3 - 2r_1r_4\cos\Theta_1\cos\Theta_2\cos\Theta_3 \\ & - 2r_1r_3\cos\tau_1\sin\Theta_1\sin\Theta_2 + 2r_2r_4\cos\tau_2\sin\Theta_2\sin\Theta_3 \\ & + 2r_1r_4\cos\tau_1\cos\Theta_3\sin\Theta_1\sin\Theta_2 + 2r_1r_4\cos\tau_2\cos\Theta_1\sin\Theta_2\sin\Theta_3 \\ & + 2r_1r_4\cos\tau_1\cos\tau_2\cos\Theta_2\sin\Theta_1\sin\Theta_3 + 2r_1r_4\sin\tau_1\sin\tau_2\sin\Theta_1\sin\Theta_3 \end{aligned} \tag{A2}$$

Assuming all bond lengths to be equal results in Equation (4) of TG, which contains 13 terms. Note that a similar expression for an idealized six-membered ring contains 34 terms, and the number of terms increases rapidly with increasing size of the ring.

After calculating through-ring distances (e.g., r_{24}) the two possible ways a deceptively simple generalized form of Equation (5) of TG, that is, a relation between torsion angles and valence bond lengths and bond angles in arbitrary five-membered rings (or chains) can be given by Equation (A3)

$$\cos\varphi_1 = \cot\Theta_1\cot\Theta_2 + \frac{r_1^2 + r_2^2 + r_3^2 - r_4^2 - r_5^2 - 2r_1r_2\cos\Theta_1 + 2r_4r_5\cos\Theta_4 - 2r_2r_3\cos\Theta_2}{2r_1r_3\sin\Theta_1\sin\Theta_3} \tag{A3}$$

The simplicity of this formula is misleading as it contains r_5 and Θ_4, dependent parameters (they are defined by cyclically extending the internal coordinate definitions of Figure 1) whose dependence on the nine independent internal coordinates is rather complicated. Nevertheless, if all r_i are equal Equation (A3) will simply lead to Equation (5) of TG.

In summary, use of Equation (A1) together with state-of-the-art CAS systems, such as Mathematica, is suggested to obtain analytical expressions for pseudorotations in constrained n-membered rings.

Acknowledgement

The authors wish to thank Mrs. I. Hudáky (Budapest, Hungary) and Dr. H. P. Reisenauer (Giessen, Germany) for helpful discussions. This work was supported by the Hungarian Scientific Research Fund (OTKA T033074).

[1] M. MacArthur, J. M. Thornton, *J. Mol. Biol.* **1991**, *218*, 397.
[2] J. S. Richardson, D. C. Richardson, *Science* **1988**, *240*, 1648–1652.
[3] E. J. Milner-White, L. H. Bell, P. H. Maccallum, *J. Mol. Biol.* **1992**, *228*, 725.
[4] J. S. Richardson, in *Advances in Protein Chemistry*, *Vol. 34* (Eds.: C. B. Anfinsen, J. T. Edsall, F. M. Richards), Academic Press, New York, **1981**.
[5] G. Nemethy, M. P. Printz, *Macromolecules* **1972**, 6, 755.
[6] R. L. Kayushina, B. K. Vainstein, *Sov. Phys. Crystallogr.* **1966**, *10*, 698.
[7] R. Balasubramanian, A. V. Lakshminarayanan, M. N. Sabesan, G. Tegoni, K. Venkatesan, G. N. Ramachandran, *Int. J. Pept. Protein Res.* **1971**, *3*, 25.
[8] D. F. DeTar, N. P. Luthra, *J. Am. Chem. Soc.* **1977**, *99*, 1232.
[9] C. A. G. Haasnoot, F. A. A. M. De Leeuw, H. P. M. De Leeuw, C. Altona, *Biopolymers* **1981**, *20*, 1211.
[10] I. D. Reva, S. G. Stepanian, A. M. Plokhotnichenko, E. D. Radchenko, G. G. Sheina, Yu. P. Blagoi, *J. Mol. Struct.* **1994**, *318*, 1.
[11] A. W. Herlinger, T. V. Long II, *J. Am. Chem. Soc.* **1970**, *92*, 6481.
[12] M. Ramek, A.-M. Kelterer, S. Nikolic, *Int. J. Quantum Chem.* **1997**, *65*, 1033.
[13] D. Cabrol, H. Brock, D. Vasilescu, *Int. J. Quantum Chem.* **1979**, *6*, 365.
[14] A.-M. Sapse, L. Mallah-Levy, S. B. Daniels, B. W. Erickson, *J. Am. Chem. Soc.* **1987**, *109*, 3526.
[15] A. G. Császár, A. Perczel, *Progr. Biophys. Mol. Biol.* **1999**, *71*, 243.
[16] P. Tarakeshwar, S. Manogaran, *J. Mol. Struct. (THEOCHEM)* **1996**, *365*, 167.
[17] S. G. Stepanian, I. D. Reva, E. D. Radchenko, L. Adamowicz, *J. Phys. Chem. A* **2001**, *105*, 10664.
[18] A. G. Császár, *J. Am. Chem. Soc.* **1992**, *114*, 9568.
[19] C.-H. Hu, M. Shen, H. F. Schaefer III, *J. Am. Chem. Soc.* **1993**, *115*, 2923.
[20] A. G. Császár, *J. Mol. Struct.* **1995**, *346*, 141.
[21] A. G. Császár, *J. Phys. Chem.* **1996**, *100*, 3541.
[22] J. E. Kilpatrick, K. S. Pitzer, R. Spitzer, *J. Am. Chem. Soc.* **1947**, *69*, 2483.
[23] R. F. W. Bader, *Atoms in Molecules. A Quantum Theory*, Clarendon Press, Oxford, **1990**.
[24] P. Popelier, *Atoms in Molecules. An Introduction*, Prentice Hall, Harlow, England, **2000**.
[25] M. J. Frisch, J. A. Pople, J. E. Del Bene, *J. Phys. Chem.* **1985**, *89*, 3664.
[26] a) R. Krishnan, J. S. Binkley, R. Seeger, J. A. Pople, *J. Chem. Phys.* **1980**, *72*, 650; b) T. Clark, J. Chandrasekhar, G. W. Spitznagel, P. von R. Schleyer, *J. Comput. Chem.* **1983**, *4*, 294; c) M. J. Frisch, J. A. Pople, J. S. Binkley, *J. Chem. Phys.* **1984**, *80*, 3265.
[27] T. H. Dunning, Jr., *J. Chem. Phys.* **1989**, *90*, 1007.
[28] R. A. Kendall, T. H. Dunning, Jr., R. J. Harrison, *J. Chem. Phys.* **1992**, *96*, 6796.
[29] A. G. Császár, W. D. Allen, *J. Chem. Phys.* **1996**, *104*, 2746.
[30] G. Tarczay, A. G. Császár, W. Klopper, H. M. Quiney, *Mol. Phys.* **2001**, *99*, 1769.
[31] A. G. Császár, W. D. Allen, H. F. Schaefer III, *J. Chem. Phys.* **1998**, *108*, 9751.
[32] C. C. J. Roothaan, *Rev. Mod. Phys.* **1951**, *23*, 69; C. C. J. Roothaan, *Rev. Mod. Phys.* **1960**, *32*, 179.
[33] P. Pulay, *Mol. Phys.* **1969**, *17*, 197.
[34] W. J. Hehre, L. Radom, P. von R. Schleyer, J. A. Pople, *Ab Initio Molecular Orbital Theory*, Wiley-Interscience, New York, **1986**.
[35] a) C. Møller, M. S. Plesset, *Phys. Rev.* **1934**, *46*, 618; b) R. Krishnan, M. J. Frisch, J. A. Pople, *J. Chem. Phys.* **1980**, *72*, 4244.
[36] a) J. Cízek, *J. Chem. Phys.* **1966**, *45*, 4256; b) J. Cízek, J. Paldus, *Physica Scripta* **1980**, *21*, 251.
[37] R. J. Bartlett, C. E. Dykstra, J. Paldus in *Advanced Theories and Computational Approaches to the Electronic Structure of Molecules* (Ed.: C. E. Dykstra), Reidel, Dordrecht, **1984**, p. 127.
[38] K. Raghavachari, G. W. Trucks, J. A. Pople, M. Head-Gordon, *Chem. Phys. Lett.* **1989**, *157*, 479.
[39] a) P. A. M. Dirac, *Proc. Cambridge Philos. Soc.* **1930**, *26*, 376; b) A. D. Becke, *Phys. Rev. B* **1988**, *38*, 3098; c) A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648; d) C. Lee, W. Yang, R. G. Parr, *Phys. Rev. B* **1988**, *37*, 785; e) S. H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **1980**, *58*, 1200.
[40] T. J. Lee, P. R. Taylor, *Int. J. Quantum Chem. Symp.* **1989**, *23*, 199.
[41] W. D. Allen, A. G. Császár, *J. Chem. Phys.* **1993**, *98*, 2983.

0947-6539/03/0904-1018 $ 20.00+.50/0

[42] J. A. Pople, R. Krishnan, H. B. Schlegel, J. S. Binkley, *Int. J. Quantum Chem. Symp.* **1979**, *13*, 225.

[43] B. G. Johnson, M. J. Frisch *J. Chem. Phys.* **1994**, *100*, 7429.

[44] A. L. L. East, W. D. Allen, A. G. Császár in *Structures and Conformations of Non-Rigid Molecules* (Eds.: J. Laane, M. Dakkouri, B. van der Veken, H. Oberhammer), Kluwer, Dordrecht, **1993**, p. 343.

[45] R. D. Cowan, D. C. Griffin, *J. Opt. Soc. Am.* **1976**, *66*, 1010.

[46] J. F. Stanton, J. Gauss, J. D. Watts, W. J. Lauderdale, R. J. Bartlett, *Int. J. Quant. Chem. Symp.* **1992**, *26*, 879.

[47] Gaussian98, Revision B.2, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, Jr., R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, Ö. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komáromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, J. A. Pople, Gaussian, Inc., Pittsburgh, PA, **1998**.

[48] a) R. Ahlrichs, M. Bar, M. Haser, H. Horn, C. Kolmel *Chem. Phys. Lett.* **1989**, *162*, 165; b) Turbomole User's Manual, version 5, **2000**.

[49] L. Hedberg, I. M. Mills, *J. Mol. Spectrosc.* **2000**, *203*, 82.

[50] U. Kock, P. L. A. Popelier, *J. Phys. Chem*, **1995**, *99*, 9747.

[51] L. F. Pacios, P. C. Gómez, *J. Comput. Chem.* **2001**, *22*, 702.

[52] F. W. Biegler-Konig, R. F. W. Bader, T. H. Tang, *J. Comput. Chem.* **1982**, *3*, 317.

[53] R. F. W. Bader, AIMPAC95 program package, McMaster University, Hamilton, Ontario, Canada, **1998**.

[54] M. Räsänen, A. Aspiala, L. Homanen, J. Murto, *J. Mol. Struct.* **1982**, *96*, 81.

[55] G. Pfafferott, H. Oberhammer, J. E. Boggs, W. Caminati, *J. Am. Chem. Soc.* **1985**, *107*, 2305.

[56] L. Carballeira, I. Pérez-Juste, *J. Chem. Soc. Perkin Trans. 2* **1998**, 1339.

[57] S. J. Han, Y. K. Kang, *J. Mol. Struct. (THEOCHEM)* **1996**, *369*, 157.

[58] D. Cremer, J. A. Pople, *J. Am. Chem. Soc.* **1975**, *97*, 1354.

[59] H. J. Geise, C. Altona, C. Romers, *Tetrahedron Lett.* **1967**, *15*, 1383.

[60] J. D. Dunitz, *Tetrahedron* **1972**, *28*, 5459.

[61] C. Altona, M. Sundaralingam, *J. Am. Chem. Soc.* **1972**, *94*, 8205.

[62] M. Tomimoto, N. Go, *J. Phys. Chem.* **1995**, *99*, 563.

[63] W. Westhof, M. Sundaralingam, *J. Am. Chem. Soc.* **1983**, *105*, 970.

[64] I. Hudáky, H. A. Baldoni, A. Perczel, *J. Mol. Struct. (THEOCHEM)* **2002**, *582*, 233.

[65] M. Ramek, A.-M. Kelterer, B. J. Teppen, L. Schäfer, *J. Mol. Struct. (THEOCHEM)* **1995**, *352*, 59.

[66] W. Caminati, H. Oberhammer, G. Pfafferott, R. R. Filgueira, C. H. Gomez, *J. Mol. Spectrosc.* **1984**, *106*, 217.

[67] A. Perczel, J. G. Ángyán, M. Kajtár, W. Viviani, J. L. Rivail, J. F. Marcoccia, I. G. Csizmadia, *J. Am. Chem. Soc.* **1991**, *113*, 6256.

[68] A. Perczel, A. G. Császár, *J. Comput. Chem.* **2000**, *21*, 882.

[69] a) M. R. Peterson, I. G. Csizmadia, *Progr. Theor. Org. Chem.* **1982**, *3*, 10; b) P. G. Mezey, *Potential Energy Hypersurfaces*, Elsevier, **1987**; c) A. Perczel, J. G. Ángyán, M. Kajtár, W. Viviani, J. L. Rivail, J. F. Marcoccia, I. G. Csizmadia, *J. Am. Chem. Soc.* **1991**, *113*, 6256.

[70] D. Feller, *J. Chem. Phys.* **1992**, *96*, 6104.

[71] A. G. Császár, G. Tarczay, M. L. Leininger, O. L. Polyansky, J. Tennyson, W. D. Allen in *Spectroscopy from Space* (Eds.: J. Demaison, K. Sarka, E. A. Cohen), Kluwer, Dordrecht, **2001**, pp. 317.

[72] A. G. Császár, G. Fogarasi, *J. Chem. Phys.* **1988**, *89*, 7646.

[73] a) P. Pulay, F. Török, *Acta Chim. Hung.* **1965**, *44*, 287; b) G. Keresztury, G. Jalsovszky, *J. Mol. Struct.* **1971**, *10*, 304.

[74] P. Pulay, G. Fogarasi, G. Pongor, J. E. Boggs, A. Vargha, *J. Am. Chem. Soc.* **1983**, *105*, 7037.

[75] W. D. Allen, A. G. Császár, D. A. Horner, *J. Am. Chem. Soc.* **1992**, *114*, 6834.

[76] a) S. Wolfram, *Mathematica: A System for Doing Mathematics by Computer*, 2nd ed., Addison-Wesley, Redwood City, **1991**; b) Mathematica v2.2.3, Wolfram Research, Inc., **1994**.

[77] A. G. Császár, N. C. Handy, *J. Chem. Phys.* **1995**, *102*, 3962.

Received: June 12, 2002 [F4173]