An Exhaustive Conformational Analysis of *N*-Acetyl-L-cysteine-*N*-methylamide. Identification of the Complete Set of Interconversion Pathways on the ab Initio and DFT Potential Energy Hypersurface

J. A. Bombasaro, M. A. Zamora, H. A. Baldoni, and R. D. Enriz*

Department of Chemistry, National University of San Luis, Chacabuco 971, 5700 SAN LUIS, Argentina

Received: September 1, 2004; In Final Form: November 5, 2004

The full conformational space of *N*-acetyl-L-cysteine-*N*-methylamide was explored by ab initio (RHF/6-31G(d)) and DFT (B3LYP/6-31G(d)) computations. Multidimensional conformational analysis predicts 81 structures in *N*-acetyl-L-cysteine-*N*-methylamide, but only 47 relaxed structures were previously determined at the RHF/3-21G level of theory. These structures were now optimized using RHF/6-31G(d) and B3LYP/6-31G(d) approaches. Seven conformational migrations were observed when recalculated at higher level of theory. Besides these major changes, only smaller conformational shifts were operative for the remaining stationary points. The exploration of the whole conformational space of *N*-acetyl-L-cysteine-*N*-methylamide, including the transition-state structures allowing the conformational interconversion among the low-energy forms, was analyzed in this study. Our results offer new insights into the influence of polar side chains on the conformational preferences of peptide structures.

1. Introduction

There are many arguments to explain why amino acids and simple model peptides form an attractive target for studies by methods of molecular electronic structure theory. Probably the main ones are (a) there is a finite number of amino acids and they have tremendous biological and biochemical significance; (b) amino acids and peptides contain a variety of intramolecular and intermolecular interactions most easily accessible by calculations; (c) these systems are very flexible conformationally and thus not easily amenable to experiments; and (d) some of them are of tractable size even for high-level ab initio and DFT calculations. Nevertheless, the application of quantum mechanics for the analysis of peptides and protein fragments is less popular in comparison with other approaches. Ab initio calculations, although now accepted, still suffer from some skepticism. One of the most commonly asked questions is "Why do we calculate, if we can measure?". Because of the intrinsic flexibility of fragments of peptides and proteins, measuring a single conformer by any spectroscopic method is often straightforward.

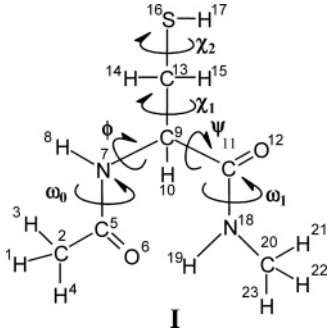
A second inquiry could be stated: "Why do we perform expensive ab initio calculation instead of applying only a commonly used force-field?". In this sense, it is clear that, if we stick to the state of the art of the approach, the size, and not the type of the molecule, will predetermine the method to be used. In response to the challenge provided by experimental observations, computations have been carried out on single amino acids ($NH_2-CHR-COOH$), on their diamide derivatives (PCO−NHCHR−CONHQ), on more complex peptides, and even on proteins at different levels of theory. Force-field methodologies, that is, molecular mechanics (MM), molecular dynamics (MD), and Monte Carlo (MC) method are of great importance in the study of the structural dynamical and equilibrium thermodynamic properties of proteins. These techniques are based on the implicit assumption that the energy of a system can be represented by a sum of classical terms. The reliability of the calculations based on the force-field techniques depends not only on the formalism used to describe the different contributions to the energy but also on the quality of the parameters incorporated within the force-field. Comparative studies have been published[1−3] from time to time. Some of the force-fields HYPERCHEM,[4] MM+,[5] AMBER,[6] CHARMM,[7] and OPLS,[8,9] as well as the common semiempirical approaches (MINDO/3,[10] MNDO,[11] AM1,[12] and PM3[13]) have been widely used to investigate molecules too large for ab initio or DFT studies. However, the results obtained in these calculations have too often differed from each other and from available ab initio data. Disturbingly, the number of allowed conformers, their locations, and their relative energies vary from one method to the other considerably.[1−3] Therefore, ab initio and DFT calculations would be useful to better understand the potential energy hypersurfaces of single amino acids in their diamide forms and to test the quality of the parameters incorporated within the different force-fields.

Another frequent objection is "Why do we try to calculate all possible conformers if they are not readily energetically accessible?". The reason for looking for the less easily found conformers is that some of them may have interesting structures or other properties. Ease of rearrangement is one important property of conformers. Several conformers could be easily distorted and then interconverted in other lower energy conformers. It might be tempting to guess that the conformers one misses are likely to be either high in energy or kinetically very unstable and thus less important. However, it is more prudent to locate all the conformers one can possibly find on an energy hypersurface first and then to examine their stability and their rates of conversion to other conformers to determine their relative importance.

Cysteine and serine are the smallest amino acids that have polar side chains. Their diamides (e.g., $HCO-NH-L-CH(CH_2-SH)CO-NH_2$) are a suitable model for better understanding the

* Corresponding author. E-mail: denriz@unsl.edu.ar.

10.1021/jp0460386 CCC: $30.25
Published on Web 01/19/2005



I

Figure 1. A skeletal diagram showing the numbering of atoms and tosional angles definitions of *N*-acetyl-L-cysteine-*N*-methylamide.

conformational properties of peptide building units. Cysteine is directly involved in a variety of biochemical interactions. As a residue with a polar side chain, it is often located on the surface of proteins. Cystein can mediate structural changes, improve hydratation of accessible surface area, or be involved in intermolecular interactions. Thus, cysteine is frequently involved in a wide range of enzymatic reactions including those of cysteine proteases.[14−16]

To our best knowledge, no experimental studies are available on the structures of the conformers of *N*-acetyl-L-cysteine-*N*-methylamide ((**I)** Figure 1). Ab initio calculations for cysteine include studies on conformational behavior[17−21] and on various physical properties (e.g., proton affinities and ionization potentials)[22] and comparison of some PCILO and RHF results.[23] Schafer et al.[17,18] investigated 10 conformers of Cys at several ab initio levels of theory including RHF/6-31G* and MP2/6-31+G*. We reported 47 different conformers for *N*-acetyl-L-cysteine-*N*-methylamide from RHF/3-21G calculations.[20] More recently, we reported a conformational study of both *N*-formyl-L-cysteine-*N*-methylamide and *N*-acetyl-L-cysteine-*N*-methylamide using DFT calculations, but only the C_7^{eq} conformations of the backbone were considered in this paper.[21]

Molecular conformational changes can be described theoretically in terms of transitions between local minima on the potential energy hypersurface (PEHS), which represent the stable conformers. These transitions are mediated by transition states. Even small molecules can have a relatively complex network of connection between stable conformers and, to understand the overall kinetics of such systems, it is essential to be able to map the topology and barrier distributions of these pathways.[24] There are interesting properties of potential energy surfaces for biological molecules, in particular for *N*-acetyl-triptophan-*N*-methylamide and *N*-acetyl-triptophan-amide.[25−27] However, in cysteine, in general, previously reported investigations[17−19,21] have concentrated on finding the lowest energy conformers, since these are the ones which one might expect to be populated at room temperature. However, as was previously pointed out, it is more prudent to locate all the conformers one can possibly find on an energy surface. Moreover, high-energy conformers could serve as intermediates in the interconversion of low energy species. In addition to the variation in the difficulty of locating conformers, there are considerable differences in their kinetic stability. The rate at any temperature at which a conformer will rearrange to other conformer is a function of the activation free energy (or energies) for the isomerization processes. Locating all of the transition states on the energy surface would allow evaluation of all these rates. Exploration of a conformational energy surface by methods that ignore all features other than local energy minima does not always give a satisfactory picture. In general, it is very desirable to determine the lowest energy transition states linking all pairs of conformations.

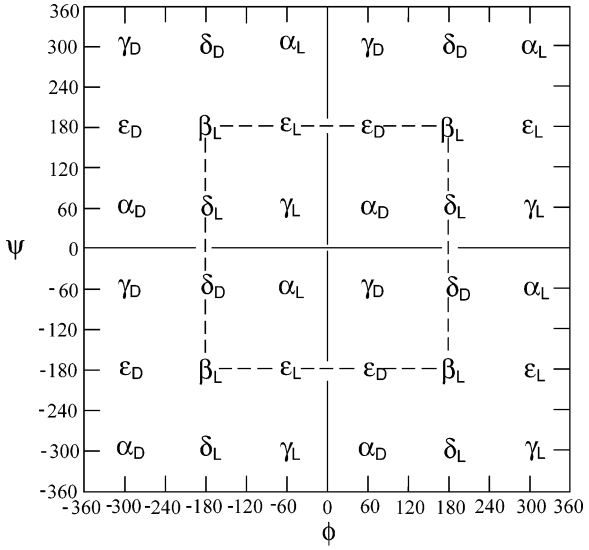


Figure 2. Topological representation of the Ramachandran map for an N- and C-protected amino acid PCO−NH−CHR−CO−NHQ (P and Q may be H or CH_3) showing two full cycles of rotation: $-360° \le \phi \le +360°$; $-360° \le \psi \le +360°$. The central box, denoted by a broken line, represents the cut suggested by the IUPAC convention. The four quadrants denoted by solid lines are the conventional cuts. Most peptide residues exhibit nine unique conformations labeled as α_{LEFT} (α_D), ϵ_D, C_7^{ax} (γ_D), β_2 (δ_L), C_5 (β_L), α' (δ_D), C_7^{eq} (γ_L), Polyproline II (PPII (ϵ_L)), and α_{RIGHT} (α_L).

The interaction between side chain and backbone in peptides is a fundamental question that has not been satisfactorily answered yet. Side chain folding is not only interesting but also important because side chain orientation can influence backbone-folding via side chain/backbone interaction. Because of the rather large dipole moment of an amide plane, it is obvious that the polar side chain may have a capacity for influencing the backbone conformation. Clearly, a better understanding of these topics could be enhanced by explicit knowledge of the quantum mechanical conformational properties of compound **I**. Thus, in this paper, we wish to report the exploration of the conformational space of compound **I**, including the transition-state structures allowing the conformational interconversion among the low energy forms.

2. Methods

2.1. Nomenclature and Abbreviations. IUPAC−IUB[28] rules recommend the use of $0° \rightarrow +180°$ for clockwise rotation and $0° \rightarrow -180°$ for counterclockwise rotation. For side chain rotation, this implies the following range: $-180° \le \chi_1 \le 180°$ and $-180° \le \chi_2 \le 180°$. On the Ramachandran map (Figure 2), the central box denoted by a broken line ($-180° \le \phi \le 180°$ and $-180° \le \psi \le 180°$) represents the cut suggested by the IUPAC convention. The four quadrants denoted by solid lines are the traditional cuts. Most peptide residues exhibit nine unique conformations, labeled as α_{LEFT} (α_D), ϵ_D, C_7^{ax} (γ_D), β_2 (δ_L), C_5 (β_L), α′ (δ_D), C_7^{eq} (γ_L), Polyproline II (PPII (ϵ_L)), and α_{RIGHT} (α_L).

However, for graphical presentation of the side chain conformational potential energy surface (PES), we use the traditional cut ($0° \le \chi_1 \le 360°$ and $0° \le \chi_2 \le 360°$), similar to that previously suggested by Ramachandran and Sasisekharan.[29]

Side chain conformations were characterized using the following nominations:

folded forms: gauche plus (g^+) ($0° \le \chi_1$ or $\chi_2 \le 120°$) and gauche minus (g^-) ($240° \le \chi_1$ or $\chi_2 \le 360°$);

extended forms: anti (a) ($120° \le \chi_1$ or $\chi_2 \le 240°$).

TABLE 1: Total Energy Values of the Component Molecules for Isodesmic Reactions

		energy (hartrees)			
molecular system		RHF/6-31G(d)		B3LYP/6-31G(d)	
MeCONH−CH**R**−CONHMe	C_7^{eq}	−453.8237500		−456.5375160	
	C_5	−453.8241110		−456.5361650	
CH_3−**R**		**R** = CH_2−SH	**R** = H	**R** = CH_2−SH	**R** = H
		−476.7362245	−40.195167	−478.0137708	−40.5183890

2.2. Computations of Molecular Conformers. Molecular geometry optimizations were performed at three levels of theory: RHF/3-21G, RHF/6-31G(d), and B3LYP/6-31G(d), using the Gaussian 03[30] program employing standard basis sets with no modifications. The importance of including electronic correlations in the conformational study has been previously reported.[21] Recently, Improta et al.[31] reported that conventional density functional theory (DFT) methods employing periodic boundary conditions give an accurate description of both the geometry and the relative energy on this kind of molecular systems. Correlation effects were included in the present work using DFT with the Becke3−Lee−Yang−Parr (B3LYP)[32] functional and the 6-31G(d) basis set. Conformations were optimized at each level of theory. All stationary points were converged to a root-mean-squared gradient of less than 10^{-6} kcal mol^{-1} $Å^{-1}$, in agreement with the limits imposed internally by Gaussian 03. With any conformational search, it is very important to examine the structures obtained to make sure that they are true minima and not transition structures or other structures with very low or zero forces on the atoms (stationary points). Thus, minima and transition-state structures were characterized through harmonic frequency analysis employing RHF/6-31G(d) and B3LYP/6-31G(d) calculations.

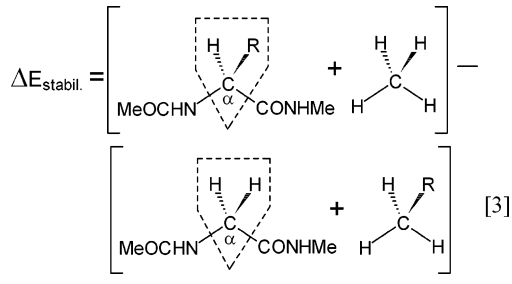
2.3. Stabilization Energies. The following isodesmic reaction [1], where **R** = CH_2−SH, was used to calculate the stabilization energies [2] with respect to the C_7^{eq} as well as to the C_5 backbone conformation of N- and C- protected glycine.[11,33]

$$\begin{aligned}&\text{MeCONH}-\text{CH}_2-\text{CONHMe} + \text{CH}_3-\mathbf{R} \rightarrow \\ &\quad \text{reference conformation } C_7^{eq} \text{ or } C_5 \\ &\qquad\qquad \text{MeCONH}-\text{CH}\mathbf{R}-\text{CONHMe} + \text{CH}_3-\text{H} \quad (1) \\ &\qquad\qquad\quad \text{conformation X}\end{aligned}$$

The stabilization energy is defined as follows:

$$\begin{aligned}&\Delta E_{\text{stabilization}} = \\ &\quad \{E[\text{MeCONH}-\text{CH}\mathbf{R}-\text{CONHMe}]_X + E[\text{CH}_3-\mathbf{H}]\} - \\ &\quad \{E[\text{MeCONH}-\text{CH}_2-\text{CONHMe}]_{\text{C7eq or C5}} + \text{CH}_3-\mathbf{R}]\} \quad (2)\end{aligned}$$

This equation is also illustrated graphically below [3] where the molecular structures are symbolizing their total energy values.



[3]

The components' energy values are summarized in Table 1.

2.4. Surface Cross Sections of Hypersurfaces. The comprehensive structural analysis of the alternative side chain geometries associated with a common backbone conformation type α_{RIGHT} or C_5 or C_7^{eq}, and so forth, resulted in characteristic (ϕ, ψ) values at the RHF/3-21G level of theory. The evaluation of these ab initio conformers provided characteristic backbone torsional angles used for the systematic grid search of the side chain conformational subspace $E = E\{[\phi, \psi][\chi_1, \chi_2]\}$.

The different side chain surfaces reported here, with 144 grid points, were generated using a 30° incrementation along both dimensions (i.e., χ_1 and χ_2). Since these maps are associated with a particularly fixed backbone (BB) conformation (i.e., fixed ϕ and ψ), they are called backbone-fixed side chain maps, even though [(3-n-6)-4] internal coordinates are fully relaxed.

3. Results and Discussion

3.1. Locating the Minima at the PEHS. Limiting our considerations only to trans-peptide bonds (i.e., $\omega_0 \cong \omega_1 \cong 180°$), the full conformational space includes four torsional angles: ϕ, ψ, χ_1, and χ_2 as defined in Figure 1. Thus, the potential energy hypersurface (PEHS) is a function of four independent variables:

$$E = E(\phi, \psi, \chi_1, \chi_2) \quad (4)$$

On the $E = E(\phi, \psi, \chi_1, \chi_2)$ potential energy hypersurface associated with **I**, $3^4 = 81$ stationary points are expected: a maximum of nine different side chain orientations for all nine backbone conformers. Forty-seven relaxed structures have been previously located at 3-21G RHF level of theory.[20] The residual 34 input structures have migrated, upon attempted optimization, to one of the 47 minima. To avoid misleading terminology, the term "migration" is used when an actually optimized conformer turned out to be in a different catchment region from the one its input was located in.

The 47 RHF/3-21G structures were reoptimized at RHF/6-31G(d) and DFT levels of theory. The following seven additional migrations were observed from RHF/3-21G to RHF/6-31G(d) calculations: ϵ_D (g^- g^+) ⇒ C_7^{ax} (g^- g^+); ϵ_D (g^- a) ⇒ C_7^{ax} (g^- a); ϵ_D (g^- g^-) ⇒ C_7^{ax} (g^- g^-); C_7^{ax} (g^+ g^+) ⇒ α_{LEFT} (g^+ g^+); C_5 (g^- g^+) ⇒ C_7^{eq} (g^- g^+); C_5 (g^- a) ⇒ C_7^{eq} (g^- a); and α' (g^- g^-) ⇒ β_2 (g^- g^-). On the other hand, three additional conformations were also obtained, they are C_5 (g^- g^-), PPII (g^+ a), and α_{RIGHT} (g^- g^+). Thus, RHF/6-31G(d) calculations predicted 43 conformations for compound **I**; these data are shown in Supporting Information in Table 1S.

Considering migrations from RHF/6-31G(d) to DFT calculations, only four cases were noted: C_7^{ax} (g^- g^-) ⇒ C_7^{eq} (g^+ g^+), β_2 (g^+ g^+) ⇒ C_7^{eq} (g^+ g^+), PPII (g^+ a) ⇒ C_7^{eq} (g^+ a), and C_5 (g^- g^-) ⇒ C_5 (g^+ g^-). DFT calculations predicted the existence of conformations C_7^{ax} (g^+ g^+) and α' (g^- g^-) which were previously reported from RHF/3-21G calculations but which are not minima at the RHF/6-31G(d) PEHS. In addition, a PPII (g^- g^+) conformation was obtained.

Interestingly enough, one PPII and one α_{RIGHT} conformation were found using DFT calculations. However, these forms display 5.65 and 12.78 kcal/mol above the global minimum, respectively. Thus, both RHF/6-31G(d) and DFT calculations

TABLE 2: Torsional Angles and Total Energy Values for Backbone and Side Chain Conformers of $CH_3CONH-Cys-CONHCH_3$ Optimized at B3LYP/6-31G(d) Level of Theory. The Calculated Relative Energies ($\Delta E_{rel.}$)[a] and Stabilization Energies ($\Delta E_{stabil.}$) Are Also Shown

conformation	ϕ (°)	ψ (°)	χ_1 (°)	χ_2 (°)	ω_0 (°)	ω_1 (°)	total energy (Hartree)	$\Delta E_{rel.}$ (kcal/mol)	ΔE_{stabil} (C_7^{eq}) (kcal/mol)	ΔE_{stabil} (C_5) (kcal/mol)
α_{LEFT} (g^+g^+)	42.96	52.17	54.45	61.58	161.53	−171.2 2	−894.0256463	10.27	4.55	3.70
(g^+a)	53.90	35.00	53.57	−169.5 2	162.93	−175.1 9	−894.0201343	13.73	8.01	7.16
(ag^+)	61.45	39.54	−162.84	116.21	169.00	−176.2 5	−894.0246707	10.88	5.16	4.31
(ag^-)	63.33	37.03	−156.69	−66.03	168.10	−176.6 8	−894.0296026	7.79	2.07	1.22
(g^-g^+)	68.00	26.33	−62.59	78.95	168.72	−177.7 2	−894.0298656	7.62	1.90	1.05
(g^-a)	69.20	26.97	−59.92	−171.2 2	164.20	−177.1 5	−894.0289934	8.17	2.45	1.60
(g^-g^-)	67.00	28.82	−61.45	−75.74	164.05	−176.9 1	−894.0304170	7.28	1.56	0.71
ϵ_D (g^+a)	47.24	−128.07	65.03	172.50	−168.0 4	176.64	−894.0219259	12.60	6.88	6.04
(g^+g^-)	53.19	−141.99	82.21	−34.08	−160.0 8	179.60	−894.0261393	9.96	4.24	3.39
(ag^+)	64.80	−168.11	−157.71	59.89	−158.5 1	−177.1 0	−894.0277268	8.96	3.24	2.40
(aa)	65.51	−163.85	−150.76	−165.5 6	−159.6 6	−176.8 9	−894.0265938	9.67	3.96	3.11
C_7^{ax} (g^+g^+)	49.61	−23.22	60.18	59.60	165.75	−178.9 9	−894.0260734	10.00	4.28	3.43
(g^+a)	60.65	−29.14	70.65	161.95	171.85	−175.1 9	−894.0248764	10.75	5.03	4.19
(g^+g^-)	66.00	−36.28	87.11	−55.15	179.67	−174.0 8	−894.0275040	9.10	3.38	2.54
(ag^+)	72.73	−66.59	−176.97	60.29	177.99	178.71	−894.0308689	6.99	1.27	0.43
(aa)	72.96	−61.01	−171.18	179.15	178.02	−179.2 9	−894.0289664	8.19	2.47	1.62
(ag^-)	73.24	−51.22	−160.67	−44.87	174.23	−177.1 3	−894.0318398	6.38	0.66	−0.18
(g^-g^+)	73.67	−57.44	−61.34	79.05	175.81	−179.3 5	−894.0337985	5.15	−0.57	−1.41
(g^-a)	73.87	−56.17	−57.35	−161.8 6	170.75	−179.8 2	−894.0321735	6.17	0.45	−0.39
β_2 (g^+g^-)	−133.3 3	25.94	56.30	−98.81	−169.9 0	176.27	−894.0334318	5.38	−0.34	−1.18
(g^-g^-)	−124.1 9	15.35	−56.94	−59.26	−164.5 0	175.42	−894.0324115	6.02	0.31	−0.54
C_5 (g^+g^+)	−162.9 5	154.09	55.82	57.00	−175.7 4	177.47	−894.0339875	5.04	−0.68	−1.53
(g^+a)	−161.8 5	163.96	64.54	−166.9 7	179.48	176.34	−894.0319263	6.33	0.61	−0.24
(g^+g^-)	−155.3 4	174.45	65.23	−59.66	171.85	177.83	−894.0345520	4.68	−1.04	−1.89
(ag^+)	−160.0 7	172.87	−163.03	72.92	174.15	177.24	−894.0382025	2.39	−3.33	−4.18
(ag^-)	−160.9 3	156.86	−172.79	−84.39	177.74	177.26	−894.0363516	3.55	−2.17	−3.02
α'(g^+g^+)	−161.3 7	−40.31	61.45	92.58	172.09	−174.5 6	−894.0276439	9.02	3.30	2.45
(g^+g^-)	−163.1 7	−47.06	54.26	−76.57	170.49	−174.4 9	−894.0291602	8.06	2.35	1.50
(ag^+)	−154.0 3	−65.36	177.20	61.42	172.88	−175.2 7	−894.0249119	10.73	5.01	4.16
(aa)	−162.6 9	−48.46	179.85	165.10	168.46	−173.0 9	−894.0227071	12.11	6.39	5.55
(ag^-)	−170.3 7	−37.89	−168.59	−49.63	167.24	−173.1 8	−894.0252046	10.55	4.83	3.98
(g^-g^+)	−143.2 3	−60.95	−167.47	29.20	166.88	−175.5 2	−894.0226771	12.13	6.41	5.57
(g^-g^-)	−131.9 3	−66.01	−61.18	−14.40	171.66	−177.3 8	−894.0225877	12.19	6.47	5.62
C_7^{eq} (g^+g^+)	−82.28	66.10	52.71	64.82	−176.4 3	−178.0 2	−894.0420115	0.00	−5.72	−6.57
(g^+a)	−82.30	66.86	56.14	−122.6 5	−178.9 0	−177.6 2	−894.0370298	3.13	−2.59	−3.44
(ag^+)	−83.18	80.08	−166.86	77.74	179.38	−174.0 7	−894.0350375	4.38	−1.34	−2.19
(ag^-)	−83.32	73.27	−170.39	−70.81	−177.6 5	−176.5 2	−894.0381960	2.39	−3.32	−4.17
(g^-g^+)	−82.60	72.70	−68.96	54.59	179.00	−176.6 7	−894.0357285	3.94	−1.78	−2.62
(g^-a)	−84.26	71.00	−55.53	−176.6 3	−172.2 7	−176.3 1	−894.0363518	3.55	−2.17	−3.02
(g^-g^-)	−84.61	73.20	−49.80	−57.82	−170.9 7	−175.8 1	−894.0372915	2.96	−2.76	−3.60
PPII (g^-g^+)	−111.1 0	143.42	−65.82	53.85	167.86	178.80	−894.0330154	5.65	−0.07	−0.92
α_{RIGHT} (g^-g^+)	−107.2 3	−83.63	−68.42	50.38	168.56	−179.4 9	−894.0216403	12.78	7.06	6.22

[a] The global minimum corresponds to C_7^{eq} (g^+g^+) conformation having −894.0420115 hartrees total energy. This value is taken a reference value, corresponding to relative energy 0.00 kcal. mol^{-1}.

confirm the presence of all nine types of backbone conformations. This is a striking difference with respect to the previously reported RHF/3-21G results.[20]

The DFT results of geometry optimizations of the title compounds at B3LYP/6-31G(d) level of theory including geometrical parameters, total energies, relative energies, and stabilization energies are given in Table 2. The total energies are given in hartrees, and relative and stabilization energies are given in kcal/mol.

We located a total of 42 conformers on the PEHS (eq 4) at the DFT level of theory, instead of the expected 81 structures. The distribution of conformers found for each backbone conformer is given in Figure 3. The global minimum is the C_7^{eq} (g^+ g^+) conformation; this backbone conformation is a folded structure (a C_7 form), and the side chain conformation corresponds to the gauche rotamers. Conformations C_7^{eq} (a g^-) and C_5 (a g^+) are the lowest-energy local minimum displaying the same energy gap (2.39 kcal/mol) above the global minimum. These three conformations are the preferred forms for the two levels of theory reported here. However, the energy gaps are substantially different. It is clear that inclusion of electron correlation tends to increase the energy differences between the conformers in **I**. Considering an energy window of 3 kcal/mol above the global minimum, only four conformations are present from DFT calculations; they are (1) C_7^{eq} (g^+ g^+), (2) C_7^{eq} (a g^-), (3) C_5 (a g^+), and (4) C_7^{eq} (g^- g^-). In contrast, 10 conformations might be included in this energy range from RHF/6-31G(d) calculations. They are (1) C_7^{eq} (g^+ g^+), (2) C_7^{eq} (a g^-), (3) C_5 (a g^+), (4) β_2 (g^+ g^+), (5) C_5 (a g^-), (6) C_7^{eq} (g^- g^-), (7) C_7^{eq} (g^- a), (8) C_7^{eq} (g^- g^+), (9) C_7^{eq} (a g^+), and (10) C_5 (g^+ g^-). Two additional conformations, C_7^{eq} (g^+ a) and C_5 (g^+ g^+), display only 3.01 and 3.06 kcal/mol above the global minimum, respectively. The RHF and DFT energy results are different possibly because of the much weaker interactions present in Cys. While in Gly and Ala, inclusion of electron correlation tends to decrease the energy differences between

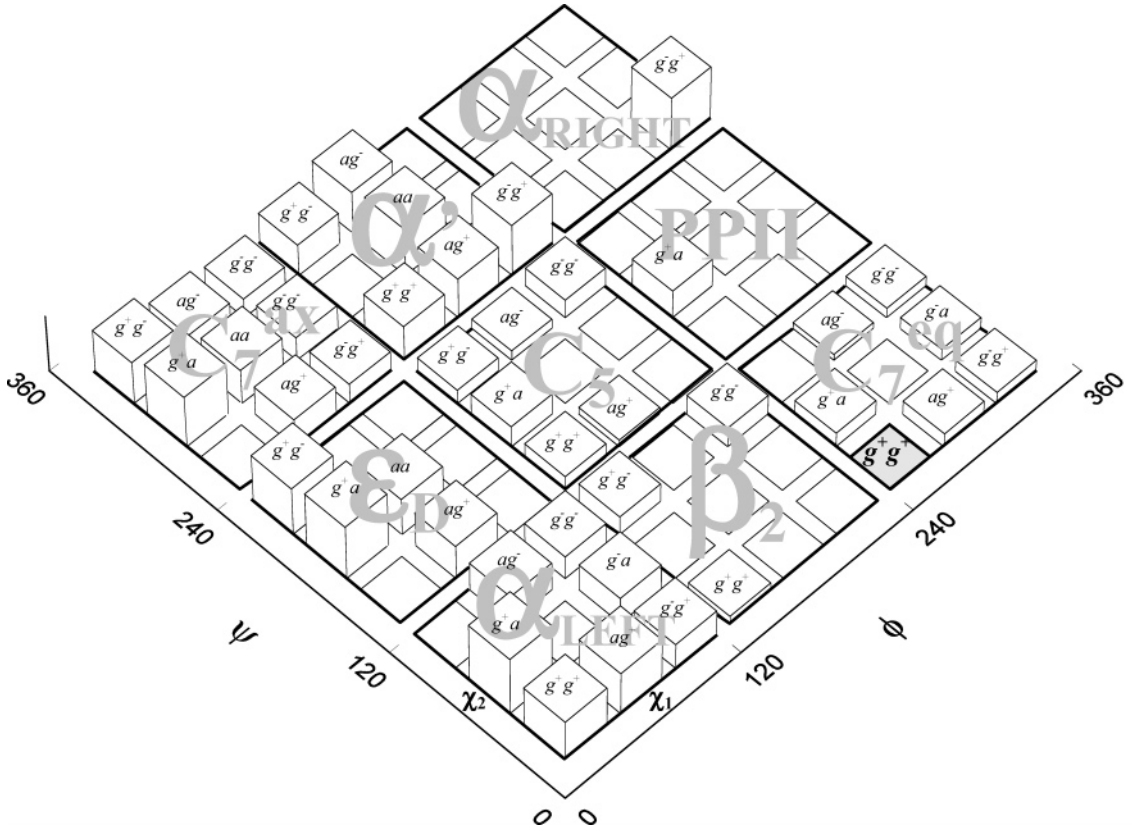


Figure 3. Schematic representations of the different minima on the PEHS of four independent variables: $E = E(\phi, \psi, \chi_1, \chi_2)$. Conformations were obtained from RHF/6-31G(d) calculations. The global minimum is denoted in bold (square in gray) and the energy gap above the global minimum is in relationship with the highest of the block.

the conformers; in Cys, it increases the energy differences in almost all cases. These results are in agreement with previously reported data, indicating that the RHF and MP2 single-point energy results[19] are drastically different for several conformers.

Conformation β_2 ($g^+ g^+$) which is the fourth preferred conformation from RHF/6-31G(d) calculations is not a minimum in the DFT-PEHS. This is another striking difference between the two methods utilized here. In contrast to Ser[34], the global minimum of Cys seems to have a H-bonding interaction between the carboxyl proton and the NH group; however, conformers with syn carbonyls are nearly as stable.

3.2. Locating the TS Structures at the PEHS. The most important point on the conformational potential energy surface is undoubtly the global minimum; the various local minima that are relatively close in energy to the global minimum are second in importance. Next are the lowest energy transition states between the above energy minima. Considerably less significant are, in most applications, the higher local energy minima, their associated transition states, and other points. Once it is established that conformations C_7^{eq}, C_5, and in a smaller proportion β_2 conformers are the more probable structures for compound **I**, it is certainly of interest to investigate the conformational interconversion process among these preferred forms.

Figure 4a shows the PES obtained for **I** rotating ϕ versus ψ keeping χ_1 and χ_2 in the $g^+ g^+$ form. This surface was obtained using the modest basis set 3-21G and therefore these results were used only in a preliminary and exploratory form. Next, we recalculate all the critical points of the PES using RHF/6-31G(d) more accurate calculations.

Keeping the side chain in the $g^+ g^+$ form, there is not a direct connection between the global minimum C_7^{eq} with the C_5 spatial ordering. TS_a is connecting C_7^{eq} ($g^+ g^+$) with β_2 ($g^+ g^+$) whereas TS_b connects β_2 ($g^+ g^+$) with C_5 ($g^+ g^+$). Thus, the conformational interconvertion between C_7^{eq} ($g^+ g^+$) and C_5 ($g^+ g^+$) involves, at least, two transition states and an energy requirement of 3.90 kcal/mol (see Figure 4b). These results indicate a significant molecular flexibility for compound **I**.

We are particularly interested in the conformational intricacies of the polar side chain of **I**. However, the question which arises is the following: is it valid to study the side chain conformational behavior for a determinated backbone conformation? To answer this question, we evaluate the conformational change of χ_1 as a consequence of the systematic conformational variations of ϕ and ψ, respectively (Figure 5). We chose the C_7^{eq} backbone conformation for this analysis considering that the global minimum displays this spatial arrangement. Even though there are noticeable changes, the backbone does not migrate to another catchment region. Thus, the results shown in Figure 5 justify, at least in part, the reason we can study side chain conformation separate from backbone.

Altogether, 92 critical points (43 energy minima (Table 2) and 49 transition states (Table 2S)) were found to be important for a description of the conformational features of the polar side chain of **I**.

3.3. Locating the TS Structures at the PESs. Figure 6a shows the PES obtained from RHF/3-21G calculations rotating the torsional angles χ_1 versus χ_2 keeping ϕ and ψ in the C_7^{eq} form. RHF/6-31G(d) calculations indicate 10 distinct TS structures which are required to describe the conformational dynamics of the C_7^{eq} form (see Figure 6a and 6b). The global minimum C_7^{eq} ($g^+ g^+$) possesses only direct connections with C_7^{eq} ($g^+ a$) and C_7^{eq} ($g^- g^+$). Transition-structure TS_1 (clock-

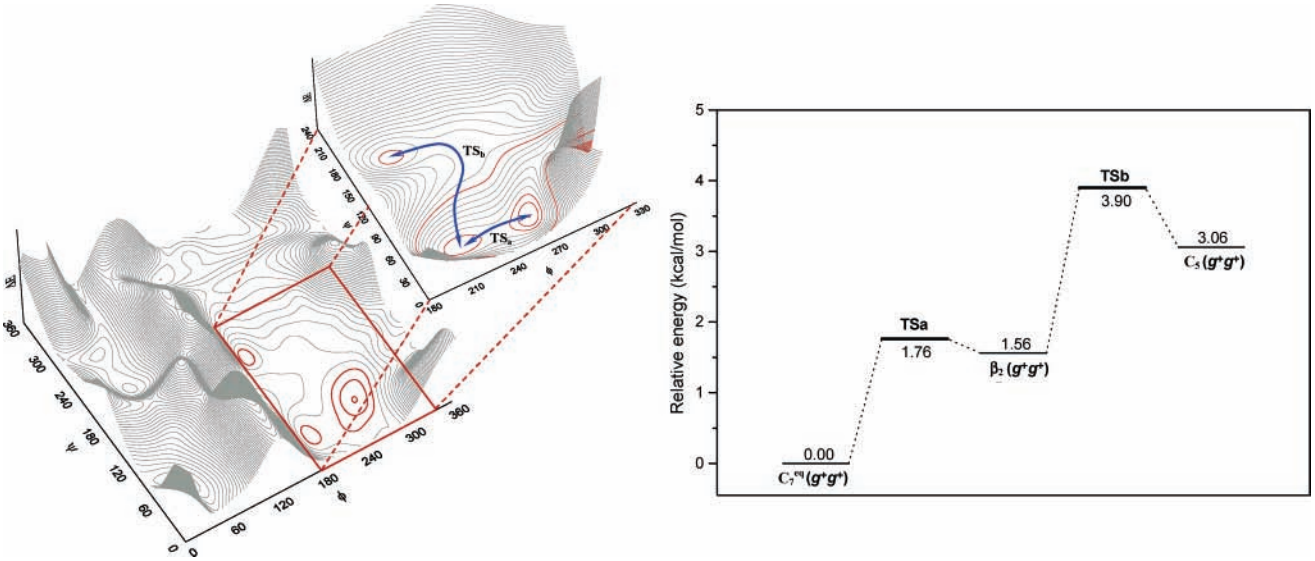


Figure 4. (a) RHF/3-21G potential energy surface (PES) obtained rotating ϕ vs ψ keeping χ_1 and χ_2 in the g^+ g^+ form. The TS structures (TSa and TSb) interconnecting the three low-energy conformations are remarked in the top. (b) Potential energy diagram for the conformational interconversions of C_7^{eq} (g^+g^+), β_2 (g^+g^+) and C_5 (g^+g^+) conformers. Calculations performed at RHF/6-31G(d) level.

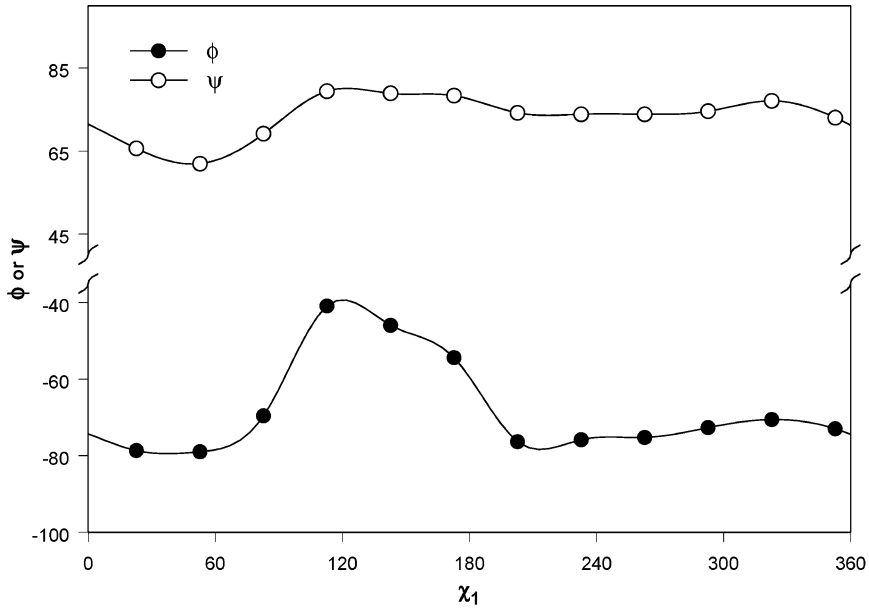


Figure 5. Conformational change of χ_1 as a consequence of the systematic conformational variations of ϕ and ψ. The torsional angle χ_2 was kept fixed in 64.82°. The global minimum displays χ_1 at 52.7°.

wise) connects C_7^{eq} (g^+ g^+) with C_7^{eq} (g^+ a), whereas TS_3 (counterclockwise) is connecting C_7^{eq} (g^+ g^+) with C_7^{eq} (g^- g^+).

C_7^{eq} (a g^-) possesses only 1.29 kcal/mol above the global minimum; however, there is not a direct connection between these forms. In fact, there are two different pathways interconnecting C_7^{eq} (g^+ g^+) with C_7^{eq} (a g^-):

$$C_7^{eq}(g^+g^+) - TS_1 \rightarrow C_7^{eq}(g^+\,a) - TS_2 \rightarrow C_7^{eq}(g^-\,g^+) - TS_7 \rightarrow C_7^{eq}(a\,g^-) \quad \text{(a)}$$

$$C_7^{eq}(g^+g^+) - TS_3 \rightarrow C_7^{eq}(g^-\,g^+) - TS_6 \rightarrow C_7^{eq}(a\,g^+) - TS_5 \rightarrow C_7^{eq}(a\,g^-) \quad \text{(b)}$$

The transition-state structures interconnecting the different C_7^{eq} forms with their respective energy profiles are shown in Figure 6b. RHF/6-31G(d) calculations predict an energy requirement of 7.17 kcal/mol for the conformational interconversion of the different C_7^{eq} forms. These results indicate a significant molecular flexibility for the polar side chain of **I**.

There are eight transition-state structures interconnecting the different six low-energy forms with the backbone in the C_5 conformation (Figure 7a). The C_5 (g^- g^-) conformation is not a minimum at the RHF/3-21G hypersurface.[20] In fact, this form collapses into C_5 (a g^-). Thus, it is not possible to observe in Figure 7b the TS_{18} obtained from RHF/6-31G(d) calculations which is connecting C_5 (a g^-) with C_5 (g^- g^-). This critical point is denoted in this figure with an arrow. Figure 7b gives the energy relationship for the six minima and eight TS structures obtained for **I** with the backbone in C_5. In this case, the whole conformational interconversion process requires 9.86 kcal/mol.

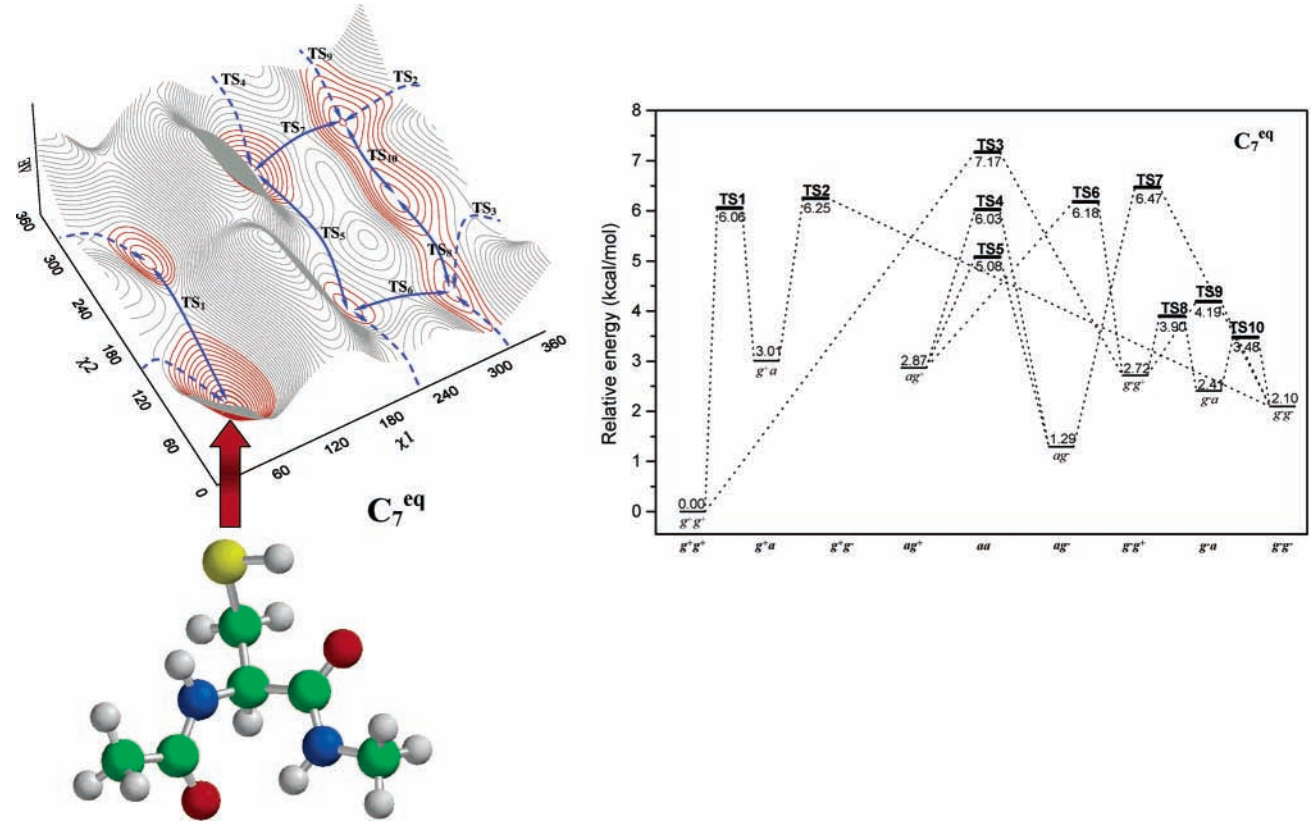


Figure 6. (a) RHF/3-21G PES obtained rotating χ_1 vs χ_2 keeping ϕ and ψ in the C_7^{eq} form. The different TS structures interconnecting the low-energy conformations and their respective pathways are shown in this figure. Counterclockwise TS are denoted in dot lines. (b) Potential energy diagram for the conformational interconversions of the backbone C_7^{eq} conformers. Calculations performed at RHF/6-31G(d) level.

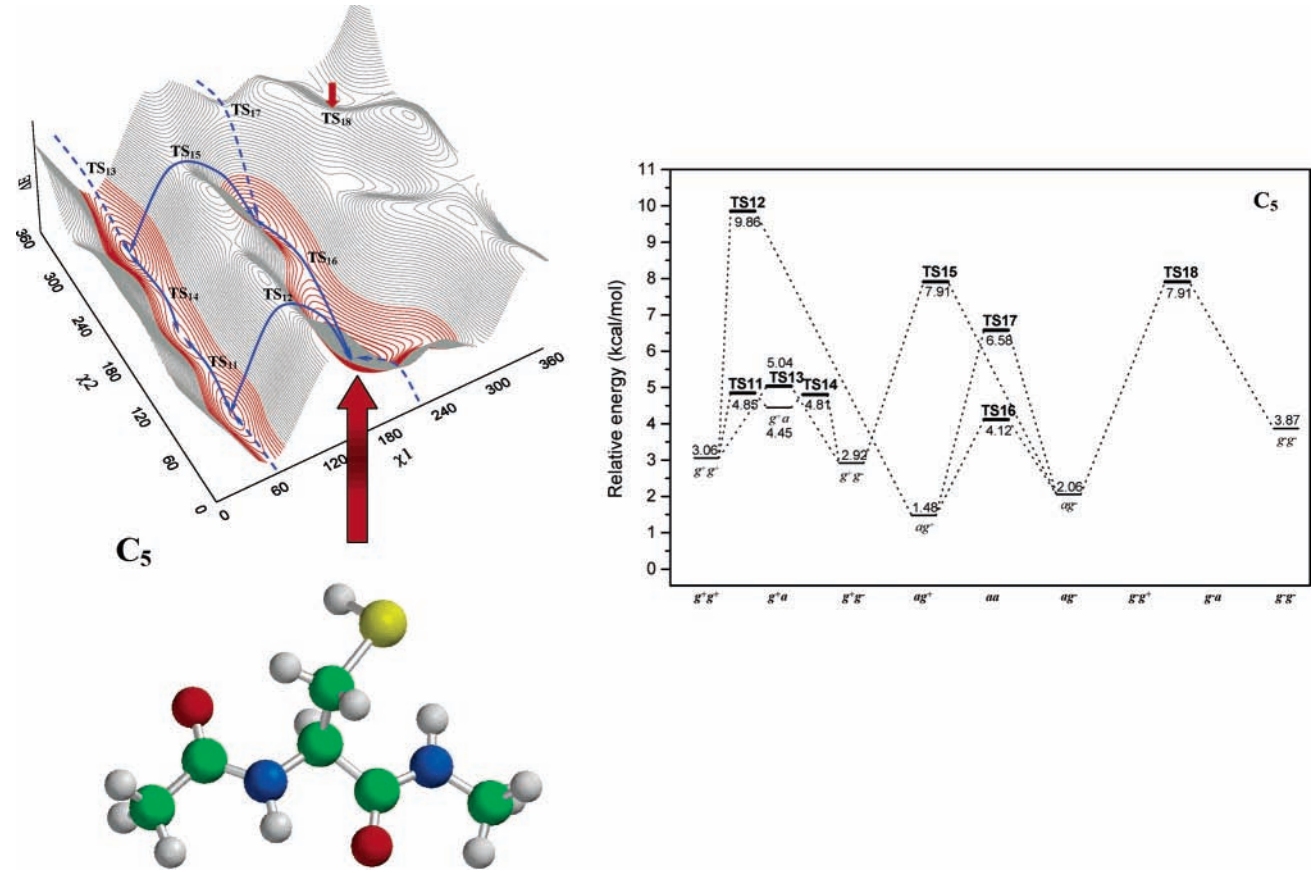


Figure 7. (a) RHF/3-21G PES obtained rotating χ_1 vs χ_2 keeping ϕ and ψ in the C_5 form. The different TS structures interconnecting the low-energy conformations and their respective pathways are shown in this figure. (b) Potential energy diagram for the conformational interconversions of the backbone C_5 conformers. Calculations performed at RHF/6-31G(d) level.

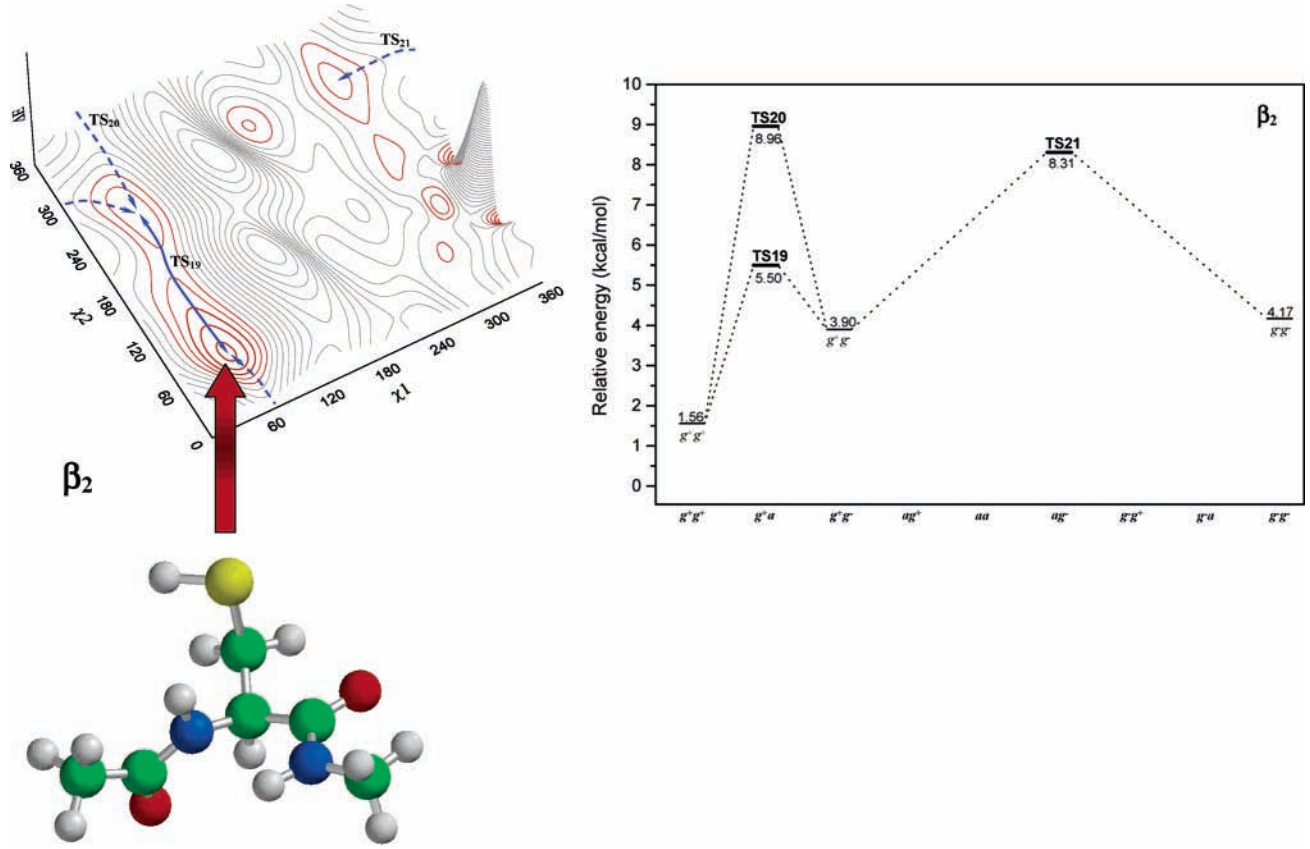


Figure 8. (a) RHF/3-21G PES obtained rotating χ_1 vs χ_2 keeping ϕ and ψ in the β_2 form. The different TS structures interconnecting the low-energy conformations and their respective pathways are shown in this figure. (b) Potential energy diagram for the conformational interconversions of the backbone β_2 conformers. Calculations performed at RHF/6-31G(d) level.

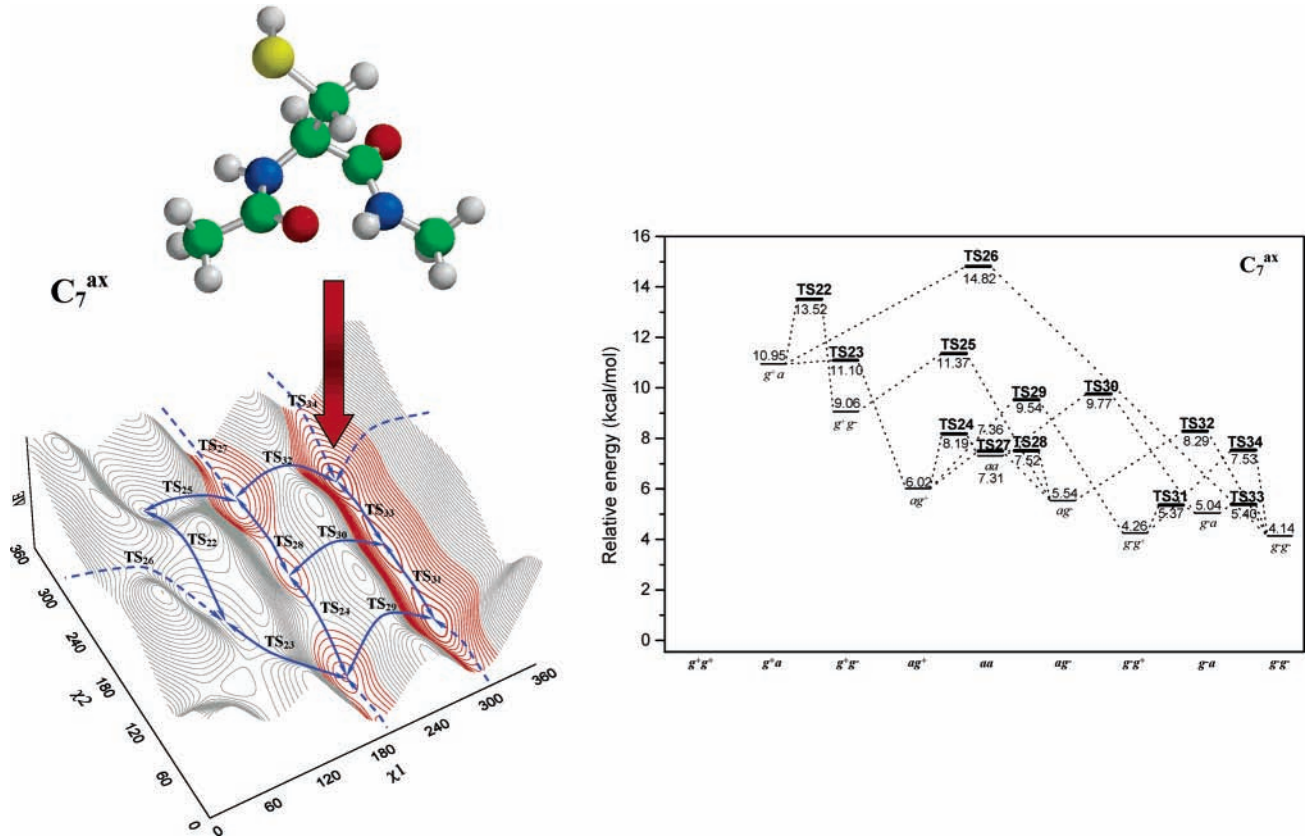


Figure 9. (a) RHF/3-21G PES obtained rotating χ_1 vs χ_2 keeping ϕ and ψ in the C_7^{ax} form. The different TS structures interconnecting the low-energy conformations and their respective pathways are shown in this figure. (b) Potential energy diagram for the conformational interconversions of the backbone C_7^{ax} conformers. Calculations performed at RHF/6-31G(d) level.

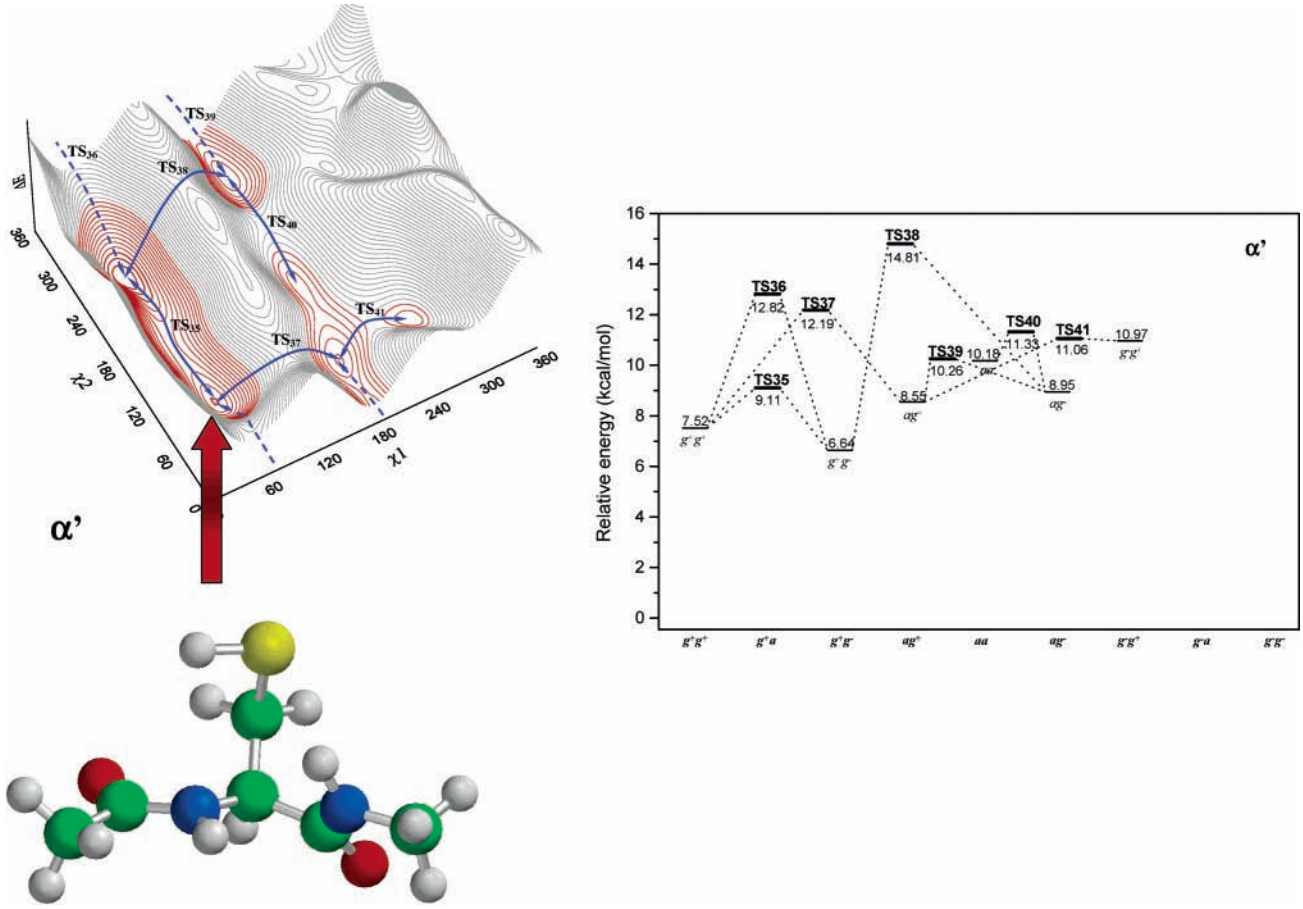


Figure 10. (a) RHF/3-21G PES obtained rotating χ_1 vs χ_2 keeping ϕ and ψ in the α' form. The different TS structures interconnecting the low-energy conformations and their respective pathways are shown in this figure. (b) Potential energy diagram for the conformational interconversions of the backbone α' conformers. Calculations performed at RHF/6-31G(d) level.

Only three minima and three transition states (TS_{19}, TS_{20}, and TS_{21}) were located at the β_2 PES (Figure 8a). These six critical points with their respective energy profiles are shown in Figure 8b. In this case, the energy gap for the conformational interconversions is 8.96 kcal/mol at RHF/6-31G(d) level of calculation.

The PESs of α_{RIGHT} and PPII forms possess only one conformer and therefore they were not analyzed in more details.

RHF/6-31G(d) calculations indicate that the conformational interconversions of the polar side chain for the preferred forms of **I** require less than 10 kcal/mol suggesting a moderate but significant molecular flexibility.

In *N*-acetyl-L-cysteine-*N*-methylamide, as in all previous cases of L-amino acids studied,[35−38] α_{LEFT}, ϵ_D, C_7^{ax}, and α' conformers are not preferred because of their relative high-energy values.

Figure 9a shows the PES obtained for the C_7^{ax} backbone conformations. A total of 20 critical points (8 minima and 12 TSs) were located in this surface. The preferred forms are C_7^{ax} ($g^- g^+$), C_7^{ax} ($g^- a$), and C_7^{ax} ($g^- g^-$) located at the right-hand side of Figure 8a. TS_{31} connects C_7^{ax} ($g^- g^+$) with C_7^{ax} ($g^- a$), whereas TS_{33} connects C_7^{ax} ($g^- a$) with C_7^{ax} ($g^- g^-$); the energetic requirement for these interconversions are 5.37 and 5.40 kcal/mol, respectively. The rest of the conformations possesses more than 5.54 kcal/mol above the global minimum and the conformational interconversions vary in a range between 7.52 and 14.82 kcal/mol (see Figure 9b).

The PES obtained for the α' forms possesses 13 critical points; six minima and seven TS structures (see Figure 10a). The local minimum with the lowest energy in this surface possesses 6.64 kcal/mol above the global minimum; the conformational interconversions vary in a range between 9.11 and 14.81 kcal/mol (Figure 10b).

The PES obtained for the α_{LEFT} forms displays 15 critical points: seven minima and eight TS structures (Figure 11a). RHF/6-31G(d) calculations predict that the whole conformational interconversions for the α_{LEFT} forms require 13.82 kcal/mol (Figure 11b). The results obtained for ϵ_D conformations (not shown) are closely related to those obtained for α_{LEFT}, C_7^{ax}, and α' conformers.

The theoretical total number of saddle points for **I** is unknown. All that can be said is that our results do not violate the so-called Morse inequalities.[39] Since the PEHS of **I** is highly multidimensional, it is not an easy task to explore all possibilities. Thus, we cannot state that the transition states reported here are the only possible structures for interconversions among all the side chain conformers of this molecule. However, we have carried out sufficient calculations to claim that the lowest-energy paths, or something close to them, have been obtained in this study.

Figure 1S presents a disconnectivity graph[40,41] of *N*-acetyl-L-cysteine-*N*-methylamide's potential energy hypersurface, produced with RHF/6-31G(d) calculations, showing a topological summary of the entire PEHS.

4. Conclusions

The potential energy hypersurface (PEHS) for *N*-acetyl-L-cysteine-*N*-methylamide (**I**) was comprehensively investigated at the Hartree−Fock (HF) employing the 3-21G and 6-31G(d)

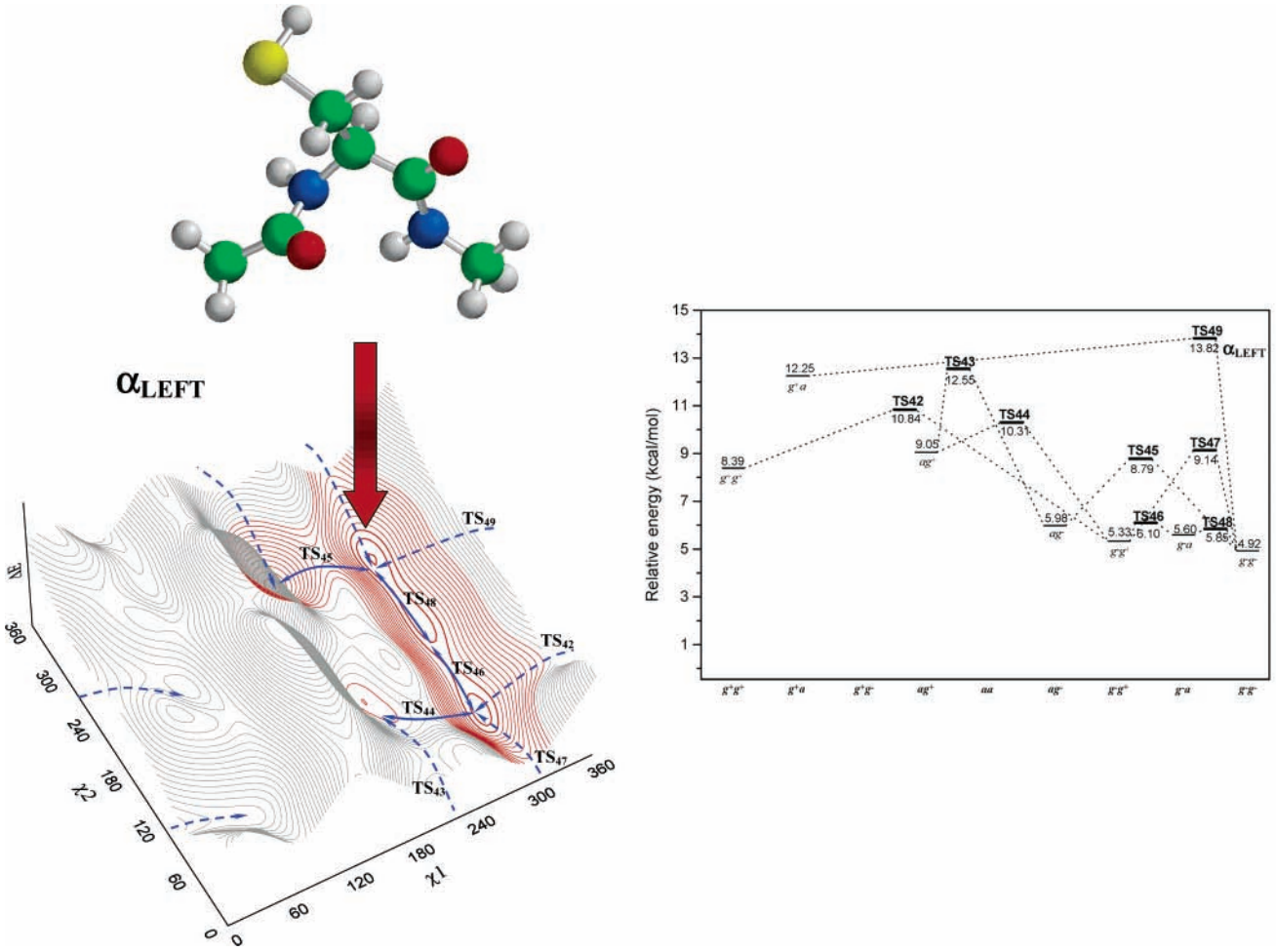


Figure 11. (a) RHF/3-21G PES obtained rotating χ_1 vs χ_2 keeping ϕ and ψ in the α_{LEFT} form. The different TS structures interconnecting the low-energy conformations and their respective pathways are shown in this figure. (b) Potential energy diagram for the conformational interconversions of the backbone α_{LEFT} conformers. Calculations performed at RHF/6-31G(d) level.

basis set and DFT (using the B3LYP/6-31G(d) approach). Theoretical calculations provide a clear picture for the conformational energy hypersurface of **I** from both structural and energetic points of view.

The three levels of theory reported here (RHF/3-21G, RHF/6-31G(d), and B3LYP/6-31G(d)) displayed qualitatively similar results. On the basis of our results, the PESs obtained at RHF/3-21G level are closely related to those attained at RHF/6-31G(d). Thus, RHF/3-21G calculations are sufficient to use in a preliminary exploratory conformational analysis of this kind of molecules. However, higher levels of theory are necessary to confirm critical points and to assign the conformational preferences. This is particularly apparent considering that the global minimum, as well as the number of conformations, varies as a function of the basis set or level of theory.

Multidimensional conformational analysis predicts 81 structures in this compound. Forty-three and 42 relaxed structures were determined at the RHF/6-31G(d) and DFT levels of theory, respectively.

The exploration of the whole conformational space of the side chain of **I**, including the transition-state structures allowing the conformational interconversion among the low-energy forms, was analyzed in this study. Thus, 92 critical points (43 energy minima and 49 transition states) were found to be important for a description of the conformational features of this portion of **I**. Ab initio calculations proved that polar side chains are able to interact with a peptide backbone, eliminating some other legitimate minima through unfavorable backbone and side chain torsional angle combinations. In addition, our results indicate that the molecular flexibility of **I** is moderate but significant and, therefore, many different conformations are available via conformational interconversion. These results offer new insights into the influence of polar side chains on the conformational preferences of peptide structures. Thus, this study can contribute to a better understanding of some less noticeable effects, which might influence the structure of a polypeptide or a protein possessing this residue in their structures.

Acknowledgment. This work was supported by grants from Universidad Nacional de San Luis. R. D. Enriz is a member of the Research Career of CONICET, Argentina.

Supporting Information Available: Tables of torsional angles and total energies and Figure 1S showing the disconnectivity graph for $CH_3CONH-Cys-CONHCH_3$ using the RHF/6-31G(d) calculations. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Roterman, I.; Lambert, M.; Gibson, K.; Scheraga, H. *J. Biomol. Struct. Dyn.* **1989**, *7*, 421.

(2) McAllister, M.; Perczel, A.; Császár, P.; Viviani, W.; Rivail, J.; Csizmadia, I. *J. Mol. Struct. (THEOCHEM)* **1993**, *288*, 161.

(3) Rodríguez, A.; Baldoni, H.; Suvire, F.; Nieto-Vazquez, R.; Zamarbide, G.; Enriz, R.; Farkas, Ö.; Perczel, A.; Csizmadia, I. *J. Mol. Struct. (THEOCHEM)* **1998**, *455*, 275−301.

(4) HYPERCHEM 4.5, Hypercube, Inc. 419 Phillip St., Waterloo, Ontario, Canada N21, 3X2, 1994.

(5) Allinger, N. *J. Am. Chem. Soc.* **1977**, *99*, 8127.

(6) Wiener, S.; Singh, U.; O'Donnel, T.; Kollman, P. *J. Am. Chem. Soc.* **1984**, *106*, 6243.

(7) Brooks, B.; Bruccoleri, R.; Olafson, B.; States, D.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.

(8) Jorgensen, W.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.

(9) Pranate, J.; Wierschke, S.; Jorgenson, W. *J. Am. Chem. Soc.* **1991**, *113*, 2810.

(10) Bringham, R.; Dewar, M.; Lo, D. *J. Am. Chem. Soc.* **1975**, *97*, 1302.

(11) Dewar, M.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899.

(12) Dewar, M.; Zoebisch, E.; Healy, E.; Stewardt, J. *J. Am. Chem. Soc.* **1986**, *107*, 3902.

(13) Stewardt, J. *J. Comput. Chem.* **1989**, *10*, 221.

(14) Polgar, L. *FEBS Lett.* **1974**, *47*, 15.

(15) Fersht, A. *In Enzyme Structure and Mechanism*, 2nd ed.; Fine, H., Ed.; W. Freeman and Company: New York, 1984; pp 413−416.

(16) Storer, A.; Menard, R. *Methods Enzymol.* **1994**, *244*, 486.

(17) Schäfer, L.; Siam, K.; Klimkowski, V. J.; Ewbank, J. D.; Van Alsenoy, C. *J. Mol. Struct. (THEOCHEM)* **1990**, *204*, 361.

(18) Schäfer, L.; Kulp-Newton, S. Q.; Siam, K.; Klimkowski, V. J.; Van Alsenoy, C. *J. Mol. Struct. (THEOCHEM)* **1990**, *209*, 373.

(19) Gronert, S.; O'Hair, R. A. J. *J. Am. Chem. Soc.* **1995**, *117*, 2071.

(20) Zamora, M. A.; Baldoni, H. A.; Bombasaro, J. A.; Mak, M. L.; Prczel, A.; Farkas, O.; Enriz, R. D. *J. Mol. Struct. (THEOCHEM)* **2001**, *540*, 271.

(21) Zamora, M. A.; Baldoni, H. A.; Rodríguez, A. M.; Enriz, R. D.; Sosa, C. P.; Prczel, A.; Farkas, O.; Deretey, E.; Vank, J. C.; Csizmadia, I. G. *Can. J. Chem.* **2002**, *80*, 832.

(22) Wright, L.; Borkman, R. *J. Am. Chem. Soc.* **1980**, *102*, 6207.

(23) Laurence, P. R.; Thomson, C. *Theor. Chim. Acta* **1981**, *58*, 121.

(24) Wales, D. J.; Doye, J. P. K.; Miller, M. A.; Mortenson, P. N.; Walsh, T. R. *Adv. Chem. Phys.* **2000**, *115*, 1.

(25) Dian, B. C.; Longarte, A.; Mercier, S.; Evans, D. A.; Wales; D. J.; Zwier, T. S. *J. Chem. Phys.* **2002**, *117*, 10688.

(26) Dian, B. C.; Longarte, A.; Winter, P. R.; Zwier, T. S. *J. Chem. Phys.* **2004**, *120*, 133.

(27) Evans, D. A.; Wales, D. J.; Dian, B. C.; Zwier, T. S. *J. Chem. Phys.* **2000**, *120*, 148.

(28) IUPAC−IUB. Comission on Biochemical Nomenclature. *Biochemistry* **1970**, *9*, 3471.

(29) Ramachandran, I.; Sasisekharan, V. *Adv. Protein Chem.* **1968**, *23*, 283.

(30) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; 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.; 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.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision B.05; Gaussian, Inc.: Pittsburgh, PA, 2003.

(31) Improta, R.; Boarone, V.; Kudin, K. N.; Scuseria, G. E. *J. Chem. Phys.* **2001**, *114*, 2541.

(32) (a) Becke, A. D. *Phys. Rev. A* **1998**, *38*, 3098. (b) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5618. (c) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1998**, *37*, 785.

(33) McAllister, M.; Endredi, G.; Viviani, W.; Perczel, A.; Csazar, P.; Ladik, J.; Rivail, J.-L.; Csizmadia, I. *Can. J. Chem.* **1995**, *73*, 563.

(34) Jákli, I.; Perczel, A.; Farkas, Ö; Sosa, C.; Csizmadia, I. G. *J. Comput. Chem.* **2000**, *21*, 626.

(35) Masman, M. F.; Zamora, M. A.; Rodríguez, A. M.; Fidanza, N. G.; Peruchena, N. M.; Enriz, R. D.; Csizmadia, I. G. *Eur. Phys. J. D* **2002**, *20*, 531.

(36) Calasa, F. C.; Rigo, M. V.; Rinaldoni, A. N.; Masman, M. F.; Rodríguez, A. M.; Enriz, R. D. *J. Mol. Struct. (THEOCHEM)*. **2003**, *634*, 201−213.

(37) Ceci, M. L.; López-Verrilli, M. A.; Vallcaneras, S. S.; Bombasaro, J. A.; Rodríguez, A. M.; Enriz, R. D. *J. Mol. Struct. (THEOCHEM)*. **2003**, *631*, 277−290.

(38) Klipfel, M. W.; Zamora, M. A.; Rodriguez, A. M.; Fidanza, N. G.; Enriz, R. D.; Csizmadia, I. G. *J. Phys. Chem. A* **2003**, *107*, 5079−5091.

(39) Mezey, P. G. *Potential Energy Hypersurfaces*; Elsevier: Amsterdam, 1987; pp 78−81.

(40) Wales, D. J.; Doye, J. P. K.; Miller, M. A.; Mortenson, P. N.; Walsh, T. R. In *Advances in Chemical Physics*; Priggogine, I., Rice, S. A., Eds.; Wiley: New York, 2000; Vol. 115, p 1.

(41) Becker, O. M.; Karplus, M. *J. Chem. Phys.* **1977**, *106*, 1495.