The conformational landscape of α-aminoglycine

Carlota Rodríguez, Alberto Macario, Juan Carlos López, Susana Blanco

Manuscript Number:	
Article Type:	VSI:Norman C. Craig Issue
Keywords:	microwave spectroscopy; computational chemistry; conformational analysis; amino acids; hydrogen bonding
Corresponding Author:	Alberto Macario, Dr. Universidad de Valladolid SPAIN
First Author:	Carlota Rodríguez
Order of Authors:	Carlota Rodríguez
	Alberto Macario
	Juan Carlos López
	Susana Blanco
Abstract:	The conformational landscape of α-aminoglycine has been studied. A conformational search has been first done using a set of different low-cost computational methods. The conformers obtained have been finally optimized at MP2 and B3LYP-D3/6-311++G(2d,p) levels of theory. A conformational landscape of 15 conformers with relative energies lower than 2000 cm^{-1} has been obtained. For the global minimum (IIa) a cooperative network formed by an O-H···N-H, NH···N and N-H···O=C hydrogen bonds, closing a sequential cycle stabilizes the structure. The next conformer in energy (Ia) also has two N-H···O=C, and N-H··N and N-H···O-H hydrogen bonds. The potential energy surfaces associated with different torsions have been investigated, to rationalize the stability of some conformers. The relative population ratio in a supersonic jet has been estimated to calculate the relative intensity of the rotational spectra of the predicted conformers.
Suggested Reviewers:	Garry Smitty Grubbs II, Dr. Assocaite Professor, Missouri S&T: Missouri University of Science and Technology grubbsg@mst.edu He is an expert in rotational spectroscopy, mainly dedicated to undergraduate students.
	Assimo Maris, Dr. Associate Professor, Universita di Bologna assimo.maris@unibo.it She is an expert in rotatioal spectroscopy and molecular structure determination, expecialy focused on biomolecules such as amino acids.
Opposed Reviewers:	

Universidad de Valladolid

Dr. Alberto Macario
Postdoctoral researcher
Department of Physical Chemistry
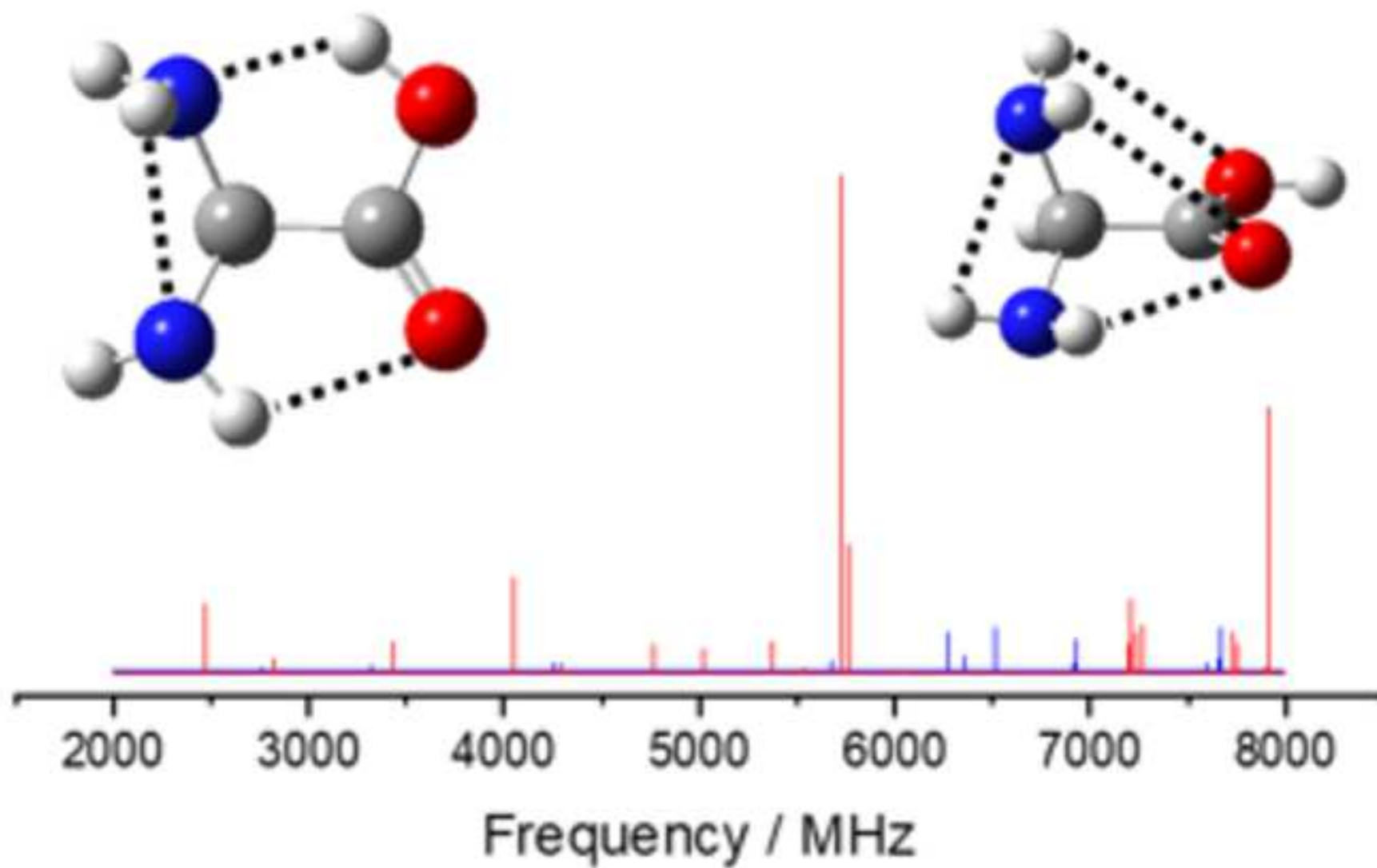Faculty of Science, University of Valladolid
Spain

April, 27th, 2022

Dear Prof. Miller,

Please find enclosed the manuscript of the paper entitled "The conformational landscape of α-aminoglycine" by Carlota Rodríguez, Juan Carlos López, Susana Blanco and myself which we like to submit for publication in The Journal of Molecular Spectroscopy in the special issue dedicated in memorial to Prof. Norman C. Craig.

This work shown the first steps towards the rotational spectroscopy investigation of the molecular structure of a molecule given by one of our undergraduate students: the computational investigation of the conformational landscape of α-aminoglycine non-natural amino acid. This search allows to understand how the molecule behaves as a subunit in other systems and to unravel how interacts with itself and with other species. We also predict the rotational spectra and the equilibrium structure of the most stable conformers of this molecule in isolation.

Sincerely yours,

Alberto Macario

THE CONFORMATIONAL LANDSCAPE OF α-AMINOGLYCINE

Carlota Rodríguez,[a] Alberto Macario,[a,b]* Juan Carlos López,[a] Susana Blanco[a]

[a] Departamento de Química Física y Química Inorgánica, IU-CINQUIMA, Universidad de Valladolid, Valladolid, 47011, Spain

[b] Current adress: Département de Physique Moléculaire, IPR (Institut de Physique de Rennes), CNRS-UMR 6251, Université de Rennes 1, F-35000 Rennes, France

* Corresponding author: Alberto Macario (alberto.macario@univ-rennes1.fr)

KEYWORDS: microwave spectroscopy, computational chemistry, conformational analysis, amino acids, hydrogen bonding.

ABSTRACT

The conformational landscape of α-aminoglycine has been studied. A conformational search has been first done using a set of different low-cost computational methods. The conformers obtained have been finally optimized at MP2 and B3LYP-D3/6-311++G(2d,p) levels of theory. A conformational landscape of 15 conformers with relative energies lower than 2000 cm^{-1} has been obtained. For the global minimum (IIa) a cooperative network formed by an O-H···N-H, NH···N and N-H···O=C hydrogen bonds, closing a sequential cycle stabilizes the structure. The next conformer in energy (Ia) also has two N-H···O=C, and N-H··N and N-H···O-H hydrogen bonds. The potential energy surfaces associated with different torsions have been investigated, to rationalize the stability of some conformers. The relative population ratio in a supersonic jet has been estimated to calculate the relative intensity of the rotational spectra of the predicted conformers.

1. Introduction

The molecular structure of the proteinogenic amino acids plays an important role in the secondary structure of proteins, that determines molecular recognition properties and their three-dimensional arrangement, and consequently governs the activity of proteins [1–3]. Therefore, a great effort has been made in the last decades in order to investigate the molecular structure of proteinogenic α-amino acids by means of high-resolution gas-phase microwave spectroscopy [4–17]. However, most of amino acids present high melting points, low vapor pressures and low thermal stabilities which makes them to decompose upon heating being thus the gas phase spectrum not accessible by traditional vaporization methods. Thus, the study of the molecular structure of most amino acids [8–17] has been possible thanks to the use of laser ablation techniques coupled with high-resolution rotational spectroscopy [4,18]. On the other hand, prior to any experimental study it is necessary to

develop a conformational search in order to analyze the intramolecular interactions involved, together with the conformational relative stability. Moreover, this allows to shed light on the possible interactions with other systems or solvents that could be present in their natural medium [19–22].

Glycine (Gly) is the simplest proteinogenic amino acid and the only one which is not chiral. Their molecular structure has been investigated by means of high-resolution molecular spectroscopy [5–7]. Due to its small size, glycine plays an important role during chain compaction early in the secondary proteins structure folding, affecting the backbone flexibility and the formation of loops, turns, and helices of polypeptide chains [23,24]. Also because of its size, is the major participant in the formation of collagen triple-helix [25]. Being the simplest amino acid, their molecular structure has been taken as a model for the rest of amino acids. Ab initio computations predict several conformers but only two of them were observed in a supersonic jet [7]. The most stable has an N-H···O=C intramolecular hydrogen bond with the carboxylic group is in the most stable *cis*-COOH arrangement while the second conformer presents a O-H···N intramolecular interaction with the acid group in a *trans*-COOH disposition. A third plausible conformer with an N-H···O-H bond was reported to relax to the most stable form in the supersonic expansion.

Protein design and engineering is trying to expand the activity and functions of proteins [26] by incorporating new amino acids beyond the natural ones and studying the molecular structure of these new species in order to reveal their chemical behavior [27]. α-Aminoglycine (Agl) is one of such non-natural α-amino acids derived from glycine replacing one of its α hydrogen atoms by an amino group. The presence of two amino groups linked to the same carbon atom increases the number of possible intramolecular interactions, and consequently affects its conformational panorama, but also makes the system unstable. Therefore, its synthesis can only be achieved in its protected form [28]. Protected aminoglycine derivatives have been identified as useful scaffolds for the introduction of chemical diversity into bioactive polypeptides [29], for example in the formation of analogs of thyrotropin-releasing hormone (TRH) [30]. It also presents anticonvulsant properties [31]. Nevertheless, the chemical and biological properties associated to the molecular structure of Agl remain currently poorly understood.

The starting point of an experimental rotational spectroscopy study is the computational investigation of the conformational landscape. This search allows to understand how the molecule behaves as a subunit in other systems, such as in those in which it is in its protected form, and to unravel how interacts with itself and with other species. First, it is worth to characterize the intramolecular interactions which stabilize the molecule and unravel the large amplitude motions associated to its conformational groups which may provide the paths to interconvert and/or connect the different conformations. The conformational panorama also provides clues about the possible interactions with molecules in its environment, defining the functional groups affected by those interactions and the moyeties that could be

unaltered by the media. This research allows also to determine or predict the possible intermolecular interactions of Agl with another species in isolation. In this work we aim to investigate the equilibrium structure and the conformational panorama of Agl.

2. Computational methods

In a first step of the conformational search of Agl, we employed the computational methods implemented in different programs [32,33]. These computational chemistry methods include molecular mechanics, semiempirical, *ab initio* and DFT. In a second step for each predicted conformer a geometry optimization was done employing Gaussian G16 [33] at *ab initio* MP2/6-311++G(2d,p) and DFT B3LYP-D3/6-311++G(2d,p) levels of theory. This has allowed us to obtain the equilibrium structure for each conformer and to calculate the different spectroscopic parameters needed to predict the rotational spectra. Complementary vibrational frequency calculations were also done in order to ensure that the predicted conformers are true minima and to evaluate Gibss energies.

Molecular mechanics are based on classic mechanics concepts and considers the molecule as a matrix of atoms governed by a classical force field. In this work we employed the Merck molecular force field MMFF94 [34], which uses parameters derived from *ab initio* computational data, mainly calculated at HF/6-31G* level of theory, except torsional parameters which are taken from MP4SDQ/TZP calculations on selected models.

Semiempirical methods are simplified versions of Hartree-Fock theory using empirical corrections in order to improve performance. They are mainly based on the Neglect of Differential Diatomic Overlap (NDDO) integral approximation. We have used the modified neglect of differential overlap MNDO [35], the Austin model 1 AM1 [36] and the parametric model number 3 PM3 [37–39] methods.

Ab initio methods are based on the computational resolution of the electronic Schrodinger equation. The first approximation uses the Hartree-Fock HF method [40], which rejects the electron correlation energy. Post-Hartree-Fock methods, as those based on the Møller-Plesset perturbation theory, start from the HF results and introduce corrections for the electron correlation through the different perturbation orders. In this work, we have employed the second-order correction in the frozen core MP2 approximation [41].

In contrast with *ab* initio methods, density functional theory (DFT) methods are based on the Hohenberg–Kohn paradigm [42,43]. From the wide range of functionals, B3LYP is one of the most popular arising from the combination of Becke's exchange functional [44] together with Lee, Yang and Parr correlation functional [45]. However, this functional has a poor description of dispersion forces and it is necessary to include them through different empirical corrections as that used in this work, the GD3 correction proposed by Grimme (B3LYP-D3) [46].

In this work we are mainly employed Pople´s split-valence basis sets [47] 6-31G, 6-31+G(d,p) and 6-311++G(2d,p). Some of these bases incorporate diffuse functions to describe the heavy atoms (+) and the hydrogen atoms (++) and d or 2d polarization functions for the heavy atoms and p-type functions for the hydrogen atoms. These further improve the description of non-covalent interactions in order to obtain better accuracy and reliability of the predictions.

3. Results and discussion

3.1. Conformational analysis

As for other amino acids [6,8,9,11–14], Agl presents a complex conformational panorama due to the combinations of intramolecular hydrogen bonds between the acid and the amino groups. In a general way, for amino acids, these interactions, schematized in Figure 1, have been labeled I, II and III. When the acid group presents a *cis*-COOH disposition it acts as a proton acceptor, as in I and III configurations, while for the *trans*-COOH arrangement it acts as a proton donor, as in type II configuration. The interaction that stabilizes type I conformers is a single or bifurcated N-H···O=C hydrogen bond. In type II the main interaction is an O-H···N hydrogen bond between the hydroxy moiety of the carboxylic group and the nitrogen of the amino group. In the III family, the main interaction is an N-H···O-H hydrogen bond. These interactions give rise to the three basic structures found for Gly, although conformer III has not been observed experimentally due to conformational relaxation to form I in the supersonic jet [7]. However, the presence of a second amino group in the C_α of Agl increases considerably the interaction possibilities because both amino groups can interact simultaneously with the acid group and between them. As a consequence, we have named the conformers of Agl following the same labeling scheme as used in natural amino acids by assigning the main interaction in each case, to that with the smallest hydrogen bond length (see Figure 2).
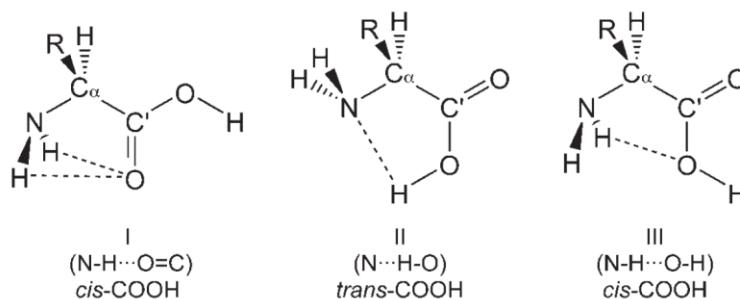

Fig. 1. Labeling of the main hydrogen bond intramolecular interaction in amino acids (Gly R=H, Agl R= NH_2).

The conformational search on Agl was done using different computational methods as it is shown in Table 1. After the search, a geometry optimization was done for each one of them at MP2 and B3LYP-D3 methods with 6-311++G(2d,p) basis set in order to improve the

results and have a realistic prediction of the final number of conformers, because some configurations obtained with low-level calculations converge to other forms already predicted. Harmonic frequency calculations were also done at the same levels of theory to corroborate that all the predicted conformers are true minima. This allows us to discover the reliability of each method initially employed in the conformer search. The number of conformers found in the first step and those found after optimization are given in Table 1.

Table 1. The number of stable species found for each different low-level computational method employed in the conformational search, and the number of stable conformers determined from them after their geometry optimization.

Computational Method	Conformational Search	Geometry Optimization
MMFF94	23	11
AM1	20	11
PM3	22	13
MNDO	22	12
HF/6-31G	24	14
HF/6-31+G(d,p)	27	17

As we can see in Table 1, HF/6-31+G(d,p), the highest level computational method is predicting the highest number of conformations, and therefore the optimum method to analyze the conformational landscape of Agl. We compared the conformers obtained between the different methods to corroborate that all of them predict the same structures. Surprisingly, apart from the 17 conformers obtained with the HF/6-31+G(d,p) method, there is an extra conformer (Ic, see Figure 2) that is only predicted by the low-level methods, molecular mechanics MMFF94 force field, and the PM3 or MNDO semiempirical methods. The 15 conformers with energies relative to the global minimum (IIa) lower than 2000 cm^{-1}/25 kJ·mol^{-1} are shown in Figure 2.
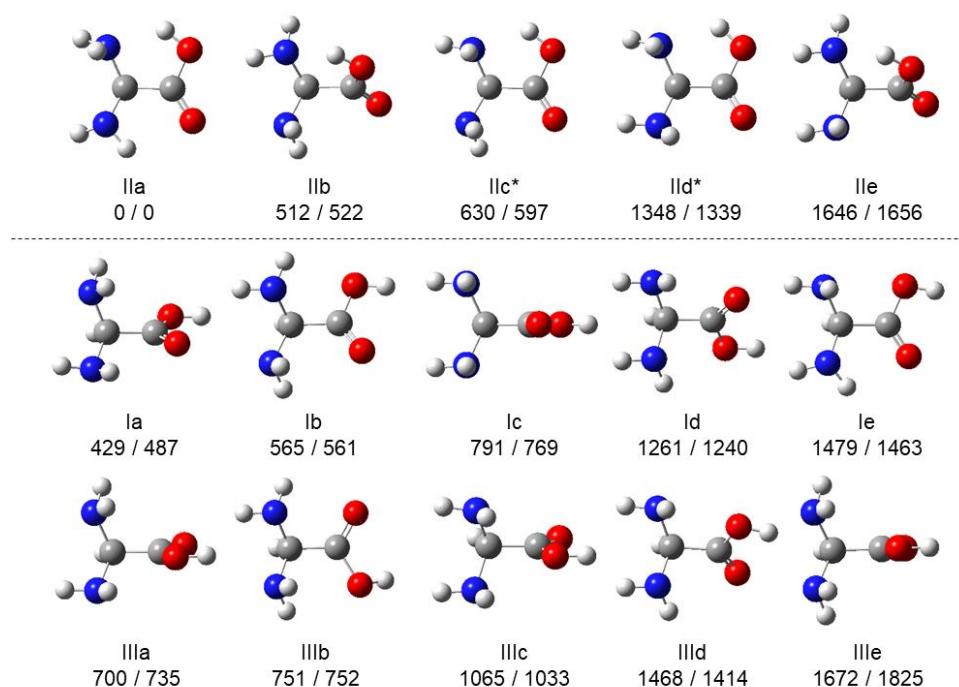
Fig. 2. The most stable conformers predicted for AgI, labeled following the type I, II and III criteria (see text) and their relative energies (cm⁻¹) predicted at the MP2/6-311++G(2d,p) / B3LYP-D3/6-311++G(2d,p) levels of theory. The conformers marked with an * might relax species IIa in a supersonic expansion (see text).

At this point, observing the different conformers we noticed that some of them only differ in the position of the hydrogen atoms of one of the amino groups, for example, IIa, IIc, and IId. We have analyzed the interconversion between these species through the $N-C_\alpha$ torsion. We have calculated from a relaxed scan the potential energy as a function of the dihedral angle φ ($H-N-C_\alpha-C'$) which allows the rotation of the amino group over the $N-C_\alpha$ bond. This energy profile predicts the interconversion pathway between IIa, IIc and IId species (see Figure 3). The small interconversion barriers of tens of wavenumbers from IId to IIc and IIc to IIa would allow their subsequent conformational relaxation to the most stable IIa form in a supersonic expansion [48].
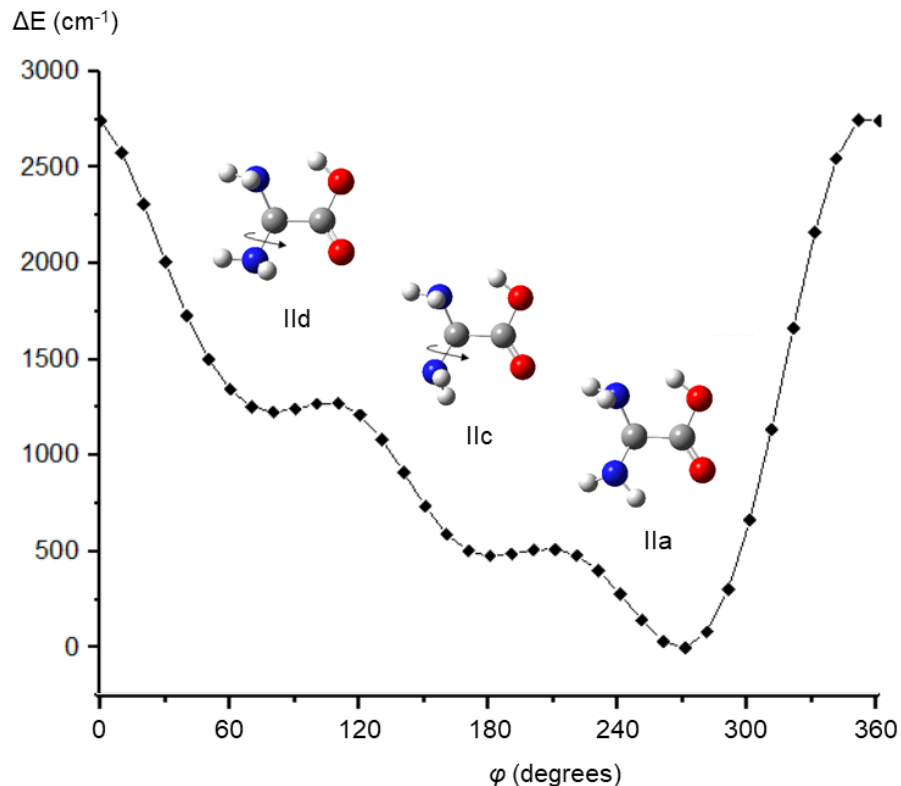
Fig. 3. Potential energy function for the rotation around the N-C$_\alpha$ bond ($\varphi = \angle$H-N-C$_\alpha$-C') connecting the IId, IIc and IIa conformers calculated at B3LYP-D3/6-311++G(2d,p) level of theory.

On the other hand, type I and type III conformers can be interconverted by the exchange of the OH and the C=O groups through the internal rotation of the COOH group, while the rest of the structure remains almost unaltered, as happens for Ia and IIIa conformers. Thus, as well as the other I and III conformers, they can be interconverted through the torsion of the C$_\alpha$-C' bond, which allows the rotation of the carboxylic group. The Ia-IIIa interconversion potential energy profile was calculated from a relaxed scan varying the dihedral angle τ (N-C$_\alpha$-C'=O) (see Figure 4). This predicts an interconversion energy barrier of more than 1200 cm^{-1}, which hinders the rotation of the acid group, and consequently the torsion over the C$_\alpha$-C' bond, avoiding the interconversion between the III and the I families. Generally, the stabilization of the III type conformers is caused by the existence of some interaction blocking the movement of the acid group. In the amino acids with aliphatic chains, this interaction does not exist or is extremely weak, such as it happens in glycine [7], alanine [8], valine [11], leucine [12] and isoleucine [13]. However, in amino acids with polar functional groups, such as serine [9] or threonine[14], there are intramolecular interactions that further stabilize and fix the arrangement of the acid group. In the case of IIIa conformer of Agl, it is stabilized by the interaction between the two amino groups and the acid group.
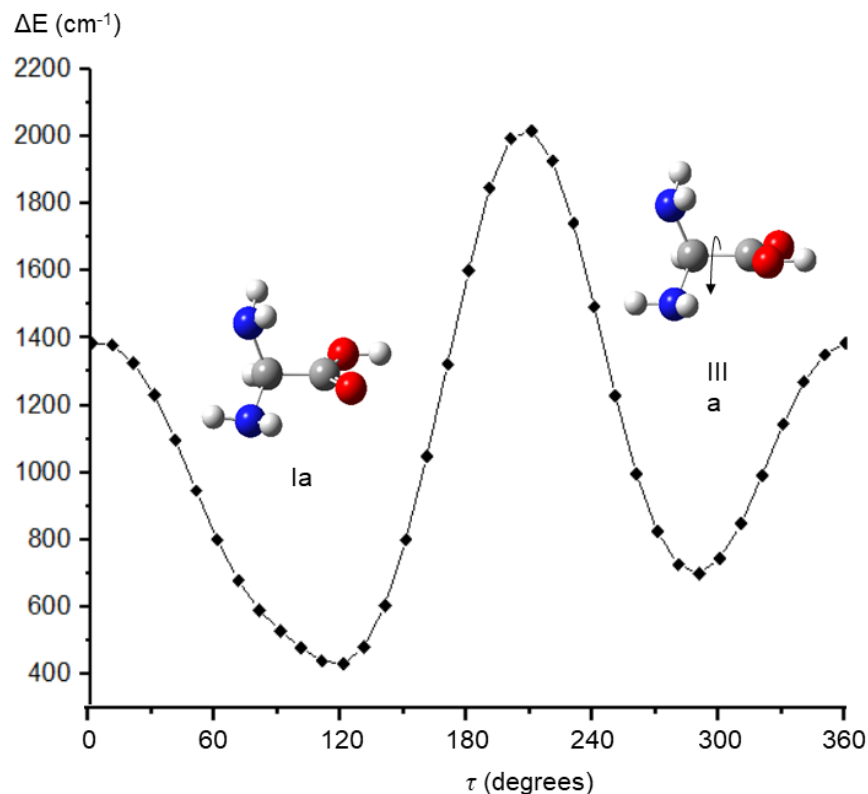
Fig. 4. Potential energy function for the rotation around the C_α- C' bond ($\tau = \angle N\text{-}C_\alpha\text{-}C'\text{=}O$) connecting the IIIa and Ia conformers calculated at B3LYP-D3/6-311++G(2d,p) level of theory.

3.2. Prediction of the rotational spectra

Previous studies of the molecular structure of amino acids [8–10,27], reported that MP2 and B3LYP-D3/6-311++G(2d,p) computational calculations predict equilibrium structures that are in good agreement with the experimentally determined molecular structures. Due to the absence of experimental information regarding the molecular structure of Agl, we have considered these calculations as a reliable prediction of the equilibrium structure of Agl. In addition, the quadrupole coupling constants due to the presence of ^{14}N atoms also give information on the orientation of the amino groups, which will be helpful for distinguishing conformers with similar rotational constants and are necessary for the accurate prediction of the rotational spectra. Consequently, the complete set of parameters, including rotational constants, quadrupole coupling constants, and electric dipole moment components, for the lower energy conformers of Agl, necessaries for the prediction of their rotational spectra, are listed in Table 2.

Table 2. Rotational parameters and relative energies for the conformers of AgI predicted at B3LYP-D3/6-311++G(2d,p) level of theory.

Param.[a]	IIa	Ia	IIb	Ib	IIc	IIIa	IIIb
A/MHz	5013.2	5215.2	5149.1	5242.2	4999.2	5158.0	5188.2
B/MHz	3537.9	2939.9	3162.0	3266.8	3467.0	2845.1	3371.5
C/MHz	2189.4	2448.5	2384.2	2206.1	2166.9	2576.5	2150.8
$^{14}N_a$ $3/2(\chi_{aa})$/MHz	2.807	-1.712	1.917	3.089	-6.421	-6.752	3.346
$^{14}N_a$ $1/4(\chi_{bb}-\chi_{cc})$/MHz	1.179	-0.693	0.568	1.592	0.109	-	1.879
$^{14}N_b$ $3/2(\chi_{aa})$/MHz	-6.575	2.544	-6.920	-5.576	-1.627	1.699	-6.167
$^{14}N_b$ $1/4(\chi_{bb}-\chi_{cc})$/MHz	0.015	1.507	-0.013	0.126	-0.645	1.064	0.353
μ_a/D	4.6	-0.5	-3.8	-0.7	-3.3	-0.6	0.7
μ_b/D	1.4	1.7	1.0	1.4	1.8	-0.3	1.2
μ_c/D	1.4	0.5	-0.4	1.3	-0.1	2.3	0.9
ΔE/cm^{-1}	0.0	487.2	522.4	561.6	597.3	734.8	752.0
ΔGibbs/cm^{-1}	0.0	188.7	549.1	350.1	378.8	488.8	514.4

Param.	Ic	Id	IId	IIId	Ie	IIe	IIIe
A/MHz	5209.7	5203.5	4980.5	5145.2	5154.3	5151.7	5132.4
B/MHz	2735.2	3211.6	3519.5	3285.4	3401.5	3093.5	2723.5
C/MHz	2644.4	2295.5	2189.1	2269.4	2153.1	2423.8	2659.5
$^{14}N_1$ $3/2(\chi_{aa})$/MHz	0.609	2.872	-1.911	3.279	-0.289	2.297	-6.754
$^{14}N_1$ $1/4(\chi_{bb}-\chi_{cc})$/MHz	-0.706	0.349	-0.702	1.346	-0.855	0.583	2.443
$^{14}N_2$ $3/2(\chi_{aa})$/MHz	0.618	3.232	3.792	1.658	1.757	-0.725	-6.755
$^{14}N_2$ $1/4(\chi_{bb}-\chi_{cc})$/MHz	-0.691	1.330	0.552	-3.990	3.098	-1.220	0.244
μ_a/D	1.7	-0.9	4.7	1.6	-1.2	-5.7	1.1
μ_b/D	0.0	2.6	2.7	0.2	-0.4	1.8	0.0
μ_c/D	-0.6	0.9	-0.6	0.0	1.1	-0.6	1.8
ΔE/cm^{-1}	768.8	1239.8	1339.2	1414.2	1462.8	1656.1	1825.1
ΔGibbs/cm^{-1}	460.5	1082.4	1040.7	1065.8	965.0	1533.5	1184.1

[a] A, B and C are rotational constants. χ_{aa}, χ_{bb} and χ_{cc} are ^{14}N nuclear quadrupole coupling constants ($^{14}N_1$ is the nitrogen of the amino group facing upwards according to Figure 2, while $^{14}N_2$ is the one of the amino group facing downwards). μ_a, μ_b and μ_c are the electric dipole moment components along the principal inertial axes. ΔE are the calculated electronic energies relative to the global minimum IIa conformer. ΔGibbs are the calculated Gibbs energies relative to the global minimum IIa conformer.

As it was concluded in previous studies of laser-ablated amino acids such as alanine [8], it can be expected that the population of the different conformers of AgI in the supersonic jet is close to that of thermodynamic equilibrium at the temperature of the carrier gas, given that a high collisional rate exists at the beginning of the expansion. In addition, we assume that the collisional relaxation processes occurring in the supersonic jet would alter the equilibrium population of the molecular systems in vibrational states of IId and IIc forms bringing them to the lowest vibrational state of conformer IIa. As a result, it is possible to relate the equilibrium distribution of the conformers with the relative population ratio of them in the

expansion. Thus, we have predicted the equilibrium relative populations of the conformers of AgI at 298K from the calculated Gibbs energies. The predicted relative populations for the most stable forms are N IIa/N Ia/ N Ib/ N Ic/ N IIIa/ N IIIb/ N IIb = 100:34:16:9:8:7:6. This population can be related to the relative intensities of the free jet rotational spectra. The intensities of the $\mu_{a/b/c}$-type transitions of the i conformer, with population Ni, are proportional to $(\mu_{a/b/c})_i^2 \cdot$Ni in a chirped-pulse Fourier transform microwave spectrometer [49], where $(\mu_{a/b/c})_i$ are the corresponding components of the electric dipole moment along the a, b or c principal inertial axis [50] (see Table 2). Therefore, the obtained relative intensity ratios are IIa/Ia/Ib/Ic/IIIa/IIIb/IIb = 100:0:0:1:0:0:4 for the μ_a-type transitions, IIa/Ia/Ib/Ic/IIIa/IIIb/IIb = 100:51:16:0:0:5:3 for the μ_b-type and IIa/Ia/Ib/Ic/IIIa/IIIb/IIb = 100:4:14:2:22:3:0 for the μ_c-type transitions. According to this, the conformers IIa and Ia will have the most intense spectra but given the sensitivity of the CP-FTMW spectrometers other forms could be expected to be observed. The rigid rotor rotational spectra [50] of the IIa and Ia conformers scaled for the relative intensities have been predicted and are shown in Figure 5 for the 2-8 GHz frequency range.
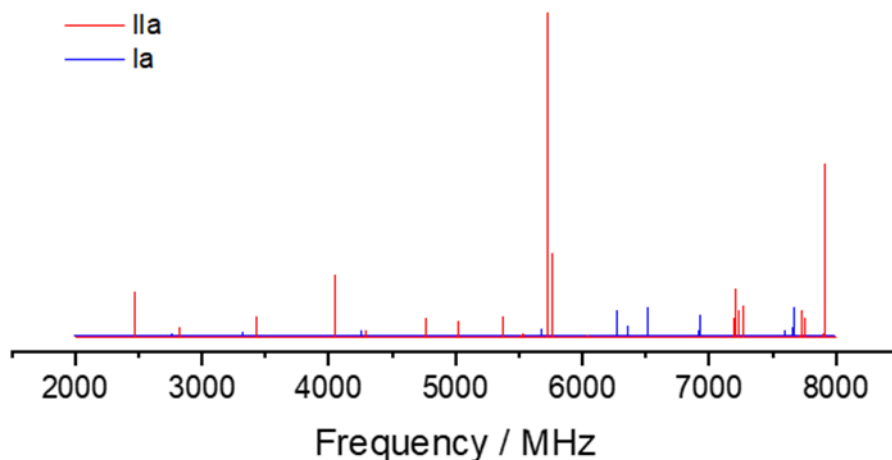

Fig. 5. Predicted rotational spectra of the IIa and Ia conformers of AgI in the 2-8 GHz frequency range.

3.3. Molecular structure

The molecular structure and the stability of the conformers of AgI depend of the intramolecular interactions, mainly hydrogen bonding, established between the different functional groups. The hydrogen bonding network predicted for the two most populated forms IIa and Ia are depicted in Figure 6. For the global minimum,conformer IIa, the strongest interaction is reasonably the O-H$\cdots$N$_a$-H hydrogen bond defining type II configurations, but it also has a type I N$_b$-H$\cdots$O=C bond. A weaker interaction N$_a$-H$\cdots$N$_b$-H between the two amino groups further stabilizes the structure closing a sequential cycle between the three functional groups. The enhanced cooperativity [51] due to the formation of this sequential

cycle explains the higher stability of IIa form. For Ia form, the main interaction is a type I N$_b$-H···O=C hydrogen bond. However, in this case the two amino groups act as double proton donors. As a result, the same amino group which establishes the main type I interaction also presents a N$_b$-H···N$_a$-H weaker hydrogen bond with the second amino group which in turn forms a second type I N$_a$-H···O=C interaction and a weaker type III N$_a$-H···O-H hydrogen bond with the acid group. This network of interactions contribute to the stability of Ia form against the rest of the predicted I family conformers.
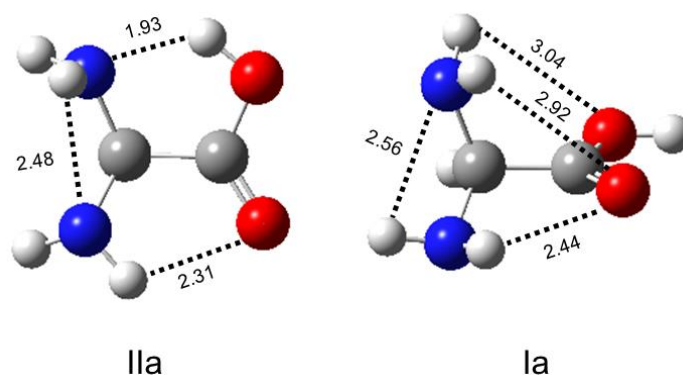

Fig. 6. Predicted molecular structures of IIa and Ia conformers of Agl. The doted lines shown the possible hydrogen bonds formed and their calculated distances (in Å) at B3LYP-D3/6-311++G(2d,p) level of theory.

Comparing the structures and the conformational stability for Agl with other amino acids it is interesting to remark that considering the two most stable structures in this case the type II structure prevails over the type I one in stability. This is in contrast with the aliphatic amino acids glycine [6], alanine [8], valine [11], leucine [12] and isoleucine [13], for which the most stable conformer, type I, is stabilized by a bifurcated N-H···O=C intramolecular hydrogen bond with a *cis*-COOH arrangement. Moreover, in all of these cases type II conformers were observed to be of higher energy, with a N···H-O intramolecular interaction and consequently a *trans*-COOH arrangement of the acid group. This is also the case of one structural derivative of glycine, the noncoded aliphatic α-aminobutyric acid [15]. Interestingly, although the N···H-O interaction is predicted to have a larger stabilization effect (favouring the type II conformers) [52] the prevalence of the lowest energy type-I forms has been attributed to a cooperative stabilizing effect caused by the *cis*-arrangement of the acid group [53]. This *cis*-COOH stabilization effect could also be the cause of the dominant I forms in the cases of serine [9] and threonine [14], where they have only one N-H···O=C intramolecular hydrogen bond but are still more stable than the II type conformers, even in the case that they are further stabilized by an intramolecular sequential cycle. This was also observed for phenylglycine [16], one of the structural derivatives of glycine amino acid. However, in the case of our glycine´s derivative, Agl, we consider that the hydrogen bond networks s causes a higher stabilization in form IIa that in conformer Ia. This enhanced stabilization occurs even if the former has a *trans*-COOH arrangement of the carboxylic group also because it presents a sequential cycle of O-H···N, N-H···N and N-H···O=C intermolecular interactions. The tipe

II N⋯H-O interaction also prevails in the observed conformers of phenylalanine molecule [17] but in this case this structure is further stabilized for a N-H⋯π intramolecular interaction between the amino and the ring groups.

Conclusions

We have investigated theoretically the conformational landscape of α-aminoglycine and the equilibrium structures of their most stable conformers. From the comparison of the results reached with the different low-cost methods employed in the conformational search, we have concluded that HF/6-31+G(d,p) is the best method in the first steps of the conformational search and analysis, but it should also be complemented with other methods such as molecular mechanics or semiempirical methods to reach a complete conformational landscape. We obtain 15 conformers in an energy interval of 2000 cm^{-1} from the global minimum. Further investigation of the potential energy surface to investigate possible interconversion coordinates concludes that some of the predicted conformers are associated with potential energy profiles suggesting that some of them presumably do not hold vibrational states and have barriers that allow their relaxation in a supersonic jet to a more stable conformer. Taking into account this, it is possible to conclude that Agl molecule presents a maximum number of 12 stable energy conformers. Their equilibrium structures have been predicted at MP2/6-311++G(2d,p) and B3LYP-D3/6-311++G(2d,p) levels of theory. These can be taken as good approximations of the molecular structure of Agl given the lack of experimental results in this case. The relative populations in the supersonic jet for a potential experiment in rotational spectroscopy, have been calculated, and, considering the possible conformational relaxations, we conclude that only the most populated conformers, would have observable spectra. The rotational spectra of the IIa and Ia species have been predicted taking into account their relative intensity estimations. The main intramolecular interaction of the global minimum IIa form is an O-H⋯N-H type II interaction, but it also presents an N-H⋯O=C type I hydrogen bond and an N-H⋯N-H weaker interaction closes a sequential cycle which further stabilizes this species due to cooperativity. We think that this is the main reason that in this molecule type II conformers prevail in stability against type I ones, in opposition to the observed in most of the other amino acids studied experimentally until now.

References

[1] A. Meister, *Biochemistry of the amino acids*, 2nd ed., vol. 2. New York: Academic, 1965. Accessed: Apr. 27, 2022. [Online]. Available: https://www.elsevier.com/books/biochemistry-of-the-amino-acids/meister/978-0-12-395523-4

[2] G. C. Barrett, *Chemistry and Biochemistry of the Amino Acids*. London: Chapman and Hall, 1985. Accessed: Apr. 27, 2022. [Online]. Available: https://link.springer.com/book/10.1007/978-94-009-4832-7

[3] G. C. Barrett and D. T. Elmore, *Amino Acids and Peptides*. Cambridge: Cambridge University Press, 1998.

[4] J. L. Alonso and J. C. López, 'Microwave spectroscopy of biomolecular building blocks', *Top. Curr. Chem.*, vol. 364, pp. 335–401, 2015, doi: 10.1007/128_2014_601.

[5] R. D. Suenram and F. J. Lovas, 'Millimeter wave spectrum of glycine', *J. Mol. Spectrosc.*, vol. 72, no. 3, pp. 372–382, Sep. 1978, doi: 10.1016/0022-2852(78)90137-6.

[6] R. D. Suenram and F. J. Lovas, 'Millimeter wave spectrum of glycine. A new conformer', *J. Am. Chem. Soc.*, vol. 102, no. 24, pp. 7180–7184, Nov. 1980, doi: 10.1021/ja00544a002.

[7] P. D. Godfrey and R. D. Brown, 'Shape of Glycine', *J. Am. Chem. Soc.*, vol. 117, no. 7, pp. 2019–2023, Feb. 1995, doi: 10.1021/ja00112a015.

[8] S. Blanco, A. Lesarri, J. C. López, and J. L. Alonso, 'The Gas-Phase Structure of Alanine', *J. Am. Chem. Soc.*, vol. 126, no. 37, pp. 11675–11683, Sep. 2004, doi: 10.1021/ja048317c.

[9] S. Blanco, M. E. Sanz, J. C. López, and J. L. Alonso, 'Revealing the multiple structures of serine', *Proc. Natl. Acad. Sci.*, vol. 104, no. 51, pp. 20183–20188, Dec. 2007, doi: 10.1073/pnas.0705676104.

[10] S. Blanco, J. C. López, S. Mata, and J. L. Alonso, 'Conformations of γ-Aminobutyric Acid (GABA): The Role of the n→π* Interaction', *Angew. Chem.*, vol. 122, no. 48, pp. 9373–9378, 2010, doi: 10.1002/ange.201002535.

[11] A. Lesarri, E. J. Cocinero, J. C. López, and J. L. Alonso, 'The Shape of Neutral Valine', *Angew. Chem. Int. Ed.*, vol. 43, no. 5, pp. 605–610, 2004, doi: 10.1002/anie.200352543.

[12] E. J. Cocinero, A. Lesarri, J.-U. Grabow, J. C. López, and J. L. Alonso, 'The Shape of Leucine in the Gas Phase', *ChemPhysChem*, vol. 8, no. 4, pp. 599–604, 2007, doi: 10.1002/cphc.200600730.

[13] A. Lesarri, R. Sánchez, E. J. Cocinero, J. C. López, and J. L. Alonso, 'Coded Amino Acids in Gas Phase: The Shape of Isoleucine', *J. Am. Chem. Soc.*, vol. 127, no. 37, pp. 12952–12956, Sep. 2005, doi: 10.1021/ja0528073.

[14] J. L. Alonso, C. Pérez, M. E. Sanz, J. C. López, and S. Blanco, 'Seven conformers of L-threonine in the gas phase: a LA-MB-FTMW study', *Phys. Chem. Chem. Phys.*, vol. 11, no. 4, pp. 617–627, Jan. 2009, doi: 10.1039/B810940K.

[15] E. J. Cocinero, A. Lesarri, M. E. Sanz, J. C. López, and J. L. Alonso, 'Conformations of α-Aminobutyric Acid in the Gas Phase', *ChemPhysChem*, vol. 7, no. 7, pp. 1481–1487, 2006, doi: 10.1002/cphc.200600091.

[16] M. E. Sanz, V. Cortijo, W. Caminati, J. C. López, and J. L. Alonso, 'The Conformers of Phenylglycine', *Chem. – Eur. J.*, vol. 12, no. 9, pp. 2564–2570, 2006, doi: 10.1002/chem.200501013.

[17] C. Pérez, S. Mata, S. Blanco, J. C. López, and J. L. Alonso, 'Jet-Cooled Rotational Spectrum of Laser-Ablated Phenylalanine', *J. Phys. Chem. A*, vol. 115, no. 34, pp. 9653–9657, Sep. 2011, doi: 10.1021/jp200800a.

[18] S. Blanco, A. Macario, and J. C. López, 'The structure of isolated thalidomide as reference for its chirality-dependent biological activity: a laser-ablation rotational study', *Phys. Chem. Chem. Phys.*, vol. 23, no. 24, pp. 13705–13713, 2021, doi: 10.1039/D1CP01691A.

[19] J. C. López, A. Macario, and S. Blanco, 'Conformational equilibria in *o* -anisic acid and its monohydrated complex: the prevalence of the *trans* -COOH form', *Phys. Chem. Chem. Phys.*, vol. 21, no. 13, pp. 6844–6850, 2019, doi: 10.1039/C9CP00552H.

[20] A. Macario, S. Blanco, J. Thomas, Y. Xu, and J. C. López, 'Competition Between Intra- and Intermolecular Hydrogen Bonding: *o* -Anisic Acid···Formic Acid Heterodimer', *Chem. – Eur. J.*, vol. 25, no. 53, pp. 12325–12331, Sep. 2019, doi: 10.1002/chem.201902086.

[21] S. Blanco, P. Pinacho, and J. C. López, 'Structure and Dynamics in Formamide–(H2O)3: A Water Pentamer Analogue', *J. Phys. Chem. Lett.*, vol. 8, no. 24, pp. 6060–6066, Dec. 2017, doi: 10.1021/acs.jpclett.7b02948.

[22] P. Pinacho, J. C. López, and S. Blanco, 'Prediction of the rotational spectra of microsolvated complexes with low cost DFT methods', *J. Mol. Spectrosc.*, vol. 337, pp. 145–152, Jul. 2017, doi: 10.1016/j.jms.2017.02.018.

[23] F. Krieger, A. Möglich, and T. Kiefhaber, 'Effect of Proline and Glycine Residues on Dynamics and Barriers of Loop Formation in Polypeptide Chains', *J. Am. Chem. Soc.*, vol. 127, no. 10, pp. 3346–3352, Mar. 2005, doi: 10.1021/ja042798i.

[24] J. Jacob, H. Duclohier, and D. S. Cafiso, 'The Role of Proline and Glycine in Determining the Backbone Flexibility of a Channel-Forming Peptide', *Biophys. J.*, vol. 76, no. 3, pp. 1367–1376, Mar. 1999, doi: 10.1016/S0006-3495(99)77298-X.

[25] R. Berisio, L. Vitagliano, L. Mazzarella, and A. Zagari, 'Crystal structure of the collagen triple helix model [(Pro-Pro-Gly)(10)](3)', *Protein Sci. Publ. Protein Soc.*, vol. 11, no. 2, pp. 262–270, Feb. 2002, doi: 10.1110/ps.32602.

[26] T. S. Young and P. G. Schultz, 'Beyond the canonical 20 amino acids: expanding the genetic lexicon', *J. Biol. Chem.*, vol. 285, no. 15, pp. 11039–11044, Apr. 2010, doi: 10.1074/jbc.R109.091306.

[27] J. C. López, A. Macario, A. Verde, A. Pérez-Encabo, and S. Blanco, 'Decoding the Structure of Non-Proteinogenic Amino Acids: The Rotational Spectrum of Jet-Cooled Laser-Ablated Thioproline', *Molecules*, vol. 26, no. 24, p. 7585, Dec. 2021, doi: 10.3390/molecules26247585.

[28] M. G. Bock, R. M. DiPardo, and R. M. Freidinger, 'Differentially protected .alpha.- aminoglycine', *J. Org. Chem.*, vol. 51, no. 19, pp. 3718–3720, Sep. 1986, doi: 10.1021/jo00369a035.

[29] S. Far and O. Melnyk, 'Synthesis of glyoxylyl peptides using an Fmoc-protected α,α'- diaminoacetic acid derivative', *J. Pept. Sci.*, vol. 11, no. 7, pp. 424–430, 2005, doi: 10.1002/psc.632.

[30] D. A. Kirby, W. Wang, M. C. Gershengorn, and J. E. Rivier, 'Analogs of thyrotropin-releasing hormone using an aminoglycine-based template', *Peptides*, vol. 19, no. 10, pp. 1679–1683, Jan. 1998, doi: 10.1016/S0196-9781(98)00124-7.

[31] H. Kohn, K. N. Sawhney, D. W. Robertson, and J. David Leander, 'Anticonvulsant Properties ofN-Substituted α,α-Diamino Acid Derivatives', *J. Pharm. Sci.*, vol. 83, no. 5, pp. 689–691, May 1994, doi: 10.1002/jps.2600830519.

[32] Y. Shao *et al.*, 'Advances in methods and algorithms in a modern quantum chemistry program package', *Phys Chem Chem Phys*, vol. 8, no. 27, pp. 3172–3191, 2006, doi: 10.1039/B517914A.

[33] M. J. Frisch *et al.*, *Gaussian 16 Rev. C.01*. Wallingford, CT, 2016.

[34] T. A. Halgren, 'Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94', *J. Comput. Chem.*, vol. 17, no. 5–6, pp. 490–519, 1996, doi: 10.1002/(SICI)1096-987X(199604)17:5/6<490::AID-JCC1>3.0.CO;2-P.

[35] M. J. S. Dewar and W. Thiel, 'Ground states of molecules. 38. The MNDO method. Approximations and parameters', *J. Am. Chem. Soc.*, vol. 99, no. 15, pp. 4899–4907, Jun. 1977, doi: 10.1021/ja00457a004.

[36] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, 'Development and use of quantum mechanical molecular models. 76. AM1: a new general purpose quantum mechanical molecular model', *J. Am. Chem. Soc.*, vol. 107, no. 13, pp. 3902–3909, Jun. 1985, doi: 10.1021/ja00299a024.

[37] J. J. P. Stewart, 'Optimization of parameters for semiempirical methods I. Method', *J. Comput. Chem.*, vol. 10, no. 2, pp. 209–220, 1989, doi: 10.1002/jcc.540100208.

[38] J. J. P. Stewart, 'Optimization of parameters for semiempirical methods II. Applications', *J. Comput. Chem.*, vol. 10, no. 2, pp. 221–264, 1989, doi: 10.1002/jcc.540100209.

[39] J. J. P. Stewart, 'Optimization of parameters for semiempirical methods. III Extension of PM3 to Be, Mg, Zn, Ga, Ge, As, Se, Cd, In, Sn, Sb, Te, Hg, Tl, Pb, and Bi', *J. Comput. Chem.*, vol. 12, no. 3, pp. 320–341, 1991, doi: 10.1002/jcc.540120306.

[40] J. C. Slater, 'A Simplification of the Hartree-Fock Method', *Phys. Rev.*, vol. 81, no. 3, pp. 385–390, Feb. 1951, doi: 10.1103/PhysRev.81.385.

[41] Chr. Møller and M. S. Plesset, 'Note on an Approximation Treatment for Many-Electron Systems', *Phys. Rev.*, vol. 46, no. 7, pp. 618–622, Oct. 1934, doi: 10.1103/PhysRev.46.618.

[42] P. Hohenberg and W. Kohn, 'Inhomogeneous Electron Gas', *Phys. Rev.*, vol. 136, no. 3B, pp. B864–B871, Nov. 1964, doi: 10.1103/PhysRev.136.B864.

[43] W. Kohn and L. J. Sham, 'Self-Consistent Equations Including Exchange and Correlation Effects', *Phys. Rev.*, vol. 140, no. 4A, pp. A1133–A1138, Nov. 1965, doi: 10.1103/PhysRev.140.A1133.

[44] A. D. Becke, 'Density-functional exchange-energy approximation with correct asymptotic behavior', *Phys. Rev. A*, vol. 38, no. 6, pp. 3098–3100, Sep. 1988, doi: 10.1103/PhysRevA.38.3098.

[45] C. Lee, W. Yang, and R. G. Parr, 'Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density', *Phys. Rev. B*, vol. 37, no. 2, pp. 785–789, Jan. 1988, doi: 10.1103/PhysRevB.37.785.

[46] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, 'A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu', *J. Chem. Phys.*, vol. 132, no. 15, p. 154104, Apr. 2010, doi: 10.1063/1.3382344.

[47] R. Ditchfield, W. J. Hehre, and J. A. Pople, 'Self-Consistent Molecular-Orbital Methods. IX. An Extended Gaussian-Type Basis for Molecular-Orbital Studies of Organic Molecules', *J. Chem. Phys.*, vol. 54, no. 2, pp. 724–728, Jan. 1971, doi: 10.1063/1.1674902.

[48] R. S. Ruoff, T. D. Klots, T. Emilsson, and H. S. Gutowsky, 'Relaxation of conformers and isomers in seeded supersonic jets of inert gases', *J. Chem. Phys.*, vol. 93, no. 5, pp. 3142–3150, Sep. 1990, doi: 10.1063/1.458848.

[49] G. G. Brown, B. C. Dian, K. O. Douglass, S. M. Geyer, S. T. Shipman, and B. H. Pate, 'A broadband Fourier transform microwave spectrometer based on chirped pulse excitation', *Rev. Sci. Instrum.*, vol. 79, no. 5, Art. no. 5, May 2008, doi: 10.1063/1.2919120.

[50]   W. Gordy and R. L. Cook, *Microwave molecular spectra*, 3rd ed. New York: Wiley, 1984.

[51]   G. Gilli, F. Bellucci, V. Ferretti, and V. Bertolasi, 'Evidence for resonance-assisted hydrogen bonding from crystal-structure correlations on the enol form of the .beta.-diketone fragment', *J. Am. Chem. Soc.*, vol. 111, no. 3, pp. 1023–1028, Feb. 1989, doi: 10.1021/ja00185a035.

[52]   A. G. Császár, 'Conformers of Gaseous α-Alanine', *ACS Publications*, Feb. 29, 1996. https://pubs.acs.org/doi/pdf/10.1021/jp9533640 (accessed Apr. 24, 2022).

[53]   A. G. Császár, W. D. Allen, and H. F. Schaefer, 'In pursuit of the *ab initio* limit for conformational energy prototypes', *J. Chem. Phys.*, vol. 108, no. 23, pp. 9751–9764, Jun. 1998, doi: 10.1063/1.476449.

Declaration of interests

☐The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Alberto Macario, Carlota Rodriguez, Susana Blanco, Juan Carlos Lopez reports financial support was provided by Junta de Castilla y Leon Consejeria de Educacion. Alberto Macario, Carlota Rodriguez, Susana Blanco, Juan Carlos Lopez reports financial support was provided by Ministry of Economic Affairs and Digital Transformation.