JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Characterization of the Elusive Conformers of Glycine from State-of-the-Art Structural, Thermodynamic, and Spectroscopic Computations: Theory Complements Experiment

Vincenzo Barone,[†] Malgorzata Biczysko,[‡] Julien Bloino,*[,§,†] and Cristina Puzzarini[∥]

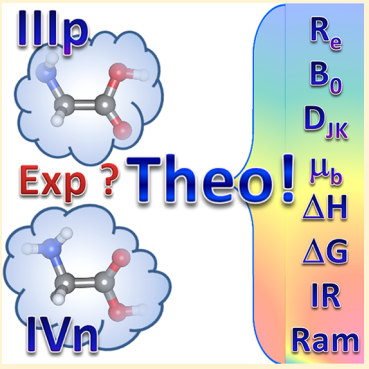
[†]Scuola Normale Superiore, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

[‡]Center for Nanotechnology Innovation @NEST, Istituto Italiano di Tecnologia, Piazza San Silvestro 12, I-56127 Pisa, Italy

[§]Consiglio Nazionale delle Ricerche, Istituto di Chimica dei Composti OrganoMetallici (ICCOM-CNR), UOS di Pisa, Area della Ricerca CNR, Via G. Moruzzi 1, I-56124 Pisa, Italy

[∥]Dipartimento di Chimica "G. Ciamician," Università di Bologna, Via F. Selmi 2, 40126 Bologna, Italy

ABSTRACT: A state-of-the-art computational strategy for the evaluation of accurate molecular structures as well as thermodynamic and spectroscopic properties along with the direct simulation of infrared (IR) and Raman spectra is established, validated (on the basis of the experimental data available for the Ip glycine conformer) and then used to provide a reliable and accurate characterization of the elusive IVn/gtt and IIIp/tct glycine conformers. The integrated theoretical model proposed is based on accurate post-Hartree−Fock computations (involving composite schemes) of energies, structures, properties, and harmonic force fields coupled to DFT corrections for the proper inclusion of vibrational effects at an anharmonic level (as provided by general second-order perturbative approach). It is shown that the approach presented here allows the evaluation of structural, thermodynamic, and spectroscopic properties with an overall accuracy of about, or better than, 0.001 Å, 20 MHz, 1 kJ·mol^{-1}, and 10 cm^{-1} for bond distances, rotational constants, conformational enthalpies, and vibrational frequencies, respectively. The high accuracy of the computational results allows one to support and complement experimental studies, thus providing (*i*) an unequivocal identification of several conformers concomitantly present in the experimental mixture and (*ii*) data not available or difficult to experimentally derive.



1. INTRODUCTION

As early as 1965, Eliel and co-workers[1] pointed out that "it is almost platitudinous to say that a chemist who does not understand conformational analysis does not understand organic chemistry. Even the area of physical chemistry related to molecular structure and physical properties has fallen heavily under the sway of conformational concepts." Since then, this statement has been systematically confirmed and enlarged to cover, on one side, biochemistry and molecular biology and, on the other side, molecular mechanics and, more recently, *ab initio* quantum chemistry. From the experimental point of view, there is an increasing consensus that a detailed knowledge of the conformational behavior of the main building blocks of biomolecules (amino acids, nucleic bases, and carbohydrates, not to mention lipids etc.) without the perturbing effect of environment (unavoidable in condensed phases) is a mandatory prerequisite toward the understanding of the role played by different interactions in determining the biological activity in terms of structure−activity relationships. This has stimulated an increasing number of experimental investigations, mainly based on mass spectrometry (MS) and thus restricted to charged species.[2−4] Investigation of neutral species is more problematic and relies on spectroscopic methods (especially IR and Raman but also rotational or double resonance (IR−UV, UV−UV) spectroscopies),[5−9] whose interpretation is more indirect and often needs quantum-chemical computations for supporting the analysis, verifying the results, and/or providing the final answer among different hypotheses. Fortunately, thanks to impressive developments of both hardware and software, contemporary quantum-chemical methods can rival experiments at least for small- to medium-sized molecules.[5,10−14] Thus, the integration of experimental and computational techniques is becoming more and more the method of choice for the conformational analysis of flexible molecules.

Before proceeding with the specific topic of the present study, it is necessary to point out that a reliable, complete characterization supplemented by the comparison with experimental data requires the determination of accurate molecular structures and of the corresponding energetics as well as the prediction of the spectra considered, also including their intensities. Energies, geometries, and spectroscopic properties can be determined at a very high level of accuracy by means of composite schemes that rely on extrapolation techniques (based on a hierarchic basis set families) and the

Received: December 5, 2012

   dx.doi.org/10.1021/ct3010672 | *J. Chem. Theory Comput.* XXXX, XXX, XXX−XXX

proper inclusion of electron-correlation effects.[15−19] These approaches suffer from a significantly high computational cost, but reliable and accurate results for biomolecule building blocks are of paramount significance in view of the limitations of experimental determinations, especially when several short-life conformers can be present. Our previous investigations on uracil[20,21] suggest that the strategy mentioned above can be pursued in the present study of two elusive conformers of glycine. The situation is more involved for vibrational frequencies (also needed to derive zero point and temperature effects on thermodynamic functions) and, especially, for intensities when the sought accuracy implies going beyond the harmonic-oscillator level, thus including anharmonic effects and vibro-rotation couplings. Several studies have demonstrated that electron correlation should be included at a very refined level for harmonic frequencies, while lower computational levels (especially density functional theory (DFT) within the hybrid functional approximation) perform very well for anharmonic terms, provided that functionals and basis sets are carefully chosen.[21−25] Much less experience is available for IR and, in particular, Raman intensities beyond the harmonic level,[26−28] but our general implementation provides encouraging results.[14,24,29,30]

In summary, we have at our disposal state-of-the-art integrated approaches to derive very accurate structural, thermodynamic, and spectroscopic results for biomolecule building blocks characterized by the contemporaneous presence of different nearly iso-energetic conformers. Among different possible case studies, the most obvious one is represented by glycine, which is the prototypical amino acid and has been the subject of systematic experimental[31−37] and theoretical investigations (e.g., refs 30, 38, and 39 and references therein). As a result of these studies, two conformers (Ip/ttt and IIn/ccc) have been unequivocally characterized in terms of relative stability and molecular structure.[32,38,39] A third rotamer (IIIp/tct) was claimed to be present on the basis of IR spectroscopy,[32] and very recently, the presence of a fourth conformer (IVn/gtt) has been suggested in a careful Raman study.[34,35] On these grounds, we decided to perform a comprehensive study of the most stable Ip and two "elusive" IIIp and IVn glycine conformers in view of providing accurate energies, structures, and spectroscopic outcomes to be directly compared with their experimental counterparts, when available. The agreement obtained for the most stable, unequivocally determined glycine conformer allowed us to estimate the accuracy and reliability of the predictions for the other two conformers, for which the experimental determination of accurate molecular structures and parameters faces intrinsic difficulties. As an overall conclusion, the approach employed is expected to provide structural, thermodynamic, and spectroscopic properties with an overall accuracy of about, or better than, 0.001 Å, 20 MHz, 1 kJ·mol^{-1}, and 10 cm^{-1} for bond distances, rotational constants, conformational enthalpies, and vibrational frequencies, respectively.

2. METHODOLOGY AND COMPUTATIONAL DETAILS

2.1. Molecular Structure. Equilibrium structures of the Ip/ttt, IVn/gtt, and IIIp/tct conformers of glycine have been investigated using a purely theoretical approach and, for Ip, by means of a mixed experimental-theoretical procedure. The former is based on a composite scheme to account simultaneously for basis-set and electron-correlation effects. The latter is the mixed experimental-theoretical approach based on rotational constants, formerly introduced by Pulay and coworkers[40] long ago, which provides the so-called semi-experimental equilibrium structure.

In the composite scheme, the various contributions are evaluated separately at the highest possible level and then combined in order to obtain the best theoretical estimates. Since the additivity approximation is directly applied to geometrical parameters, this scheme mainly involves geometry optimizations at the second-order Møller−Plesset perturbation theory (MP2)[41] level. The MP2 method has been used in conjunction with the standard cc-pVnZ[42] basis sets (n = T, Q), as well as a triple-ζ basis set augmented by diffuse functions, aug-cc-pVTZ.[42,43] In both cases, the frozen-core (fc) approximation has been adopted while, to account for core-correlation effects, the core−valence correlation-consistent cc-pCVTZ basis set[42,44] has been employed. The coupled-cluster (CC) singles and doubles approximation augmented by a perturbative treatment of triple excitations [CCSD(T)][45] method has been used together with the cc-pVTZ basis set in order to improve the electron-correlation treatment. To take into account basis-set truncation effects, the complete basis set (CBS) limit has been evaluated by making the assumption that the convergence behavior of the structural parameters mimics that of the energy. The consolidated n^{-3} extrapolation form[46] has been applied to the case n = T and Q:

$$r(\mathrm{CBS}) = \frac{n^3 r(n) - (n-1)^3 r(n-1)}{n^3 - (n-1)^3} \quad (1)$$

where n = 4, and thus $r(n)$ and $r(n-1)$ denote the MP2/cc-pVQZ and MP2/cc-pVTZ optimized parameters, respectively. Even though this procedure is only empirically based, involves small- to medium-sized basis sets, and is applied to the whole parameter and not only to the correlation contribution (as it should be), the investigation carried out in ref 47 has demonstrated its reliability and the good agreement of the corresponding results with those derived from more rigorous extrapolation schemes.[47] The effects due to core−valence (CV) electron correlation have been included by means of the corresponding correction, Δr(CV), derived as

$$\Delta r(\mathrm{CV}) = r(\mathrm{CVTZ, all}) - r(\mathrm{CVTZ, fc}) \quad (2)$$

where r(CVTZ,all) and r(CVTZ,fc) are the geometries optimized at the MP2/cc-pCVTZ level correlating all and only valence electrons, respectively. Analogously, to account for the effect of diffuse functions (aug), the corresponding correction, Δr(aug), has been evaluated by the following difference

$$\Delta r(\mathrm{aug}) = r(\mathrm{augVTZ, fc}) - r(\mathrm{VTZ, fc}) \quad (3)$$

where r(augVTZ,fc) and r(VTZ,fc) are the optimized structures at the MP2 level employing the aug-cc-pVTZ and cc-pVTZ basis sets, respectively, within the frozen-core approximation. The higher-order correlation energy contributions on molecular structure have also been considered using CCSD(T).[45] The corresponding correction, Δr(T), has been thus derived from the comparison of the geometries optimized at the MP2 and CCSD(T) levels, both with the cc-pVTZ basis:

$$\Delta r(\mathrm{T}) = r(\mathrm{CCSD(T)}) - r(\mathrm{MP2}) \quad (4)$$

On the whole, our best-estimated equilibrium structure has been determined as

dx.doi.org/10.1021/ct3010672 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

$$r(\text{best}) = r(\text{CBS}) + \Delta r(\text{CV}) + \Delta r(\text{aug}) + \Delta r(\text{T}) \quad (5)$$

This structure will be denoted as "best estimate" or, in equations and tables, more simply "best."

For the conformers of C_s symmetry, thanks to the reduced computational cost that the higher symmetry implies, the best estimate of the equilibrium structure could be determined at a higher level of theory, i.e., by exclusively employing CCSD(T) calculations. Furthermore, a more rigorous approach has been used, which is based on additivity at an energy-gradient level.[15,16] The contributions considered are the Hartree−Fock self-consistent-field (HF-SCF) energy extrapolated to the basis-set limit, the valence correlation energy at the CCSD(T) level extrapolated to the basis-set limit as well, and the core-correlation correction. This scheme has been employed because, despite the fact that it is more computationally demanding, it is well tested and its accuracy clearly assessed (see, for instance, refs 12, 15, and 48−50 and references therein). This structure, which is usually referred to as "CCSD(T)/CBS+CV," in analogy to the first approach, will be simply denoted as "bestCC".

The energy gradient used in the geometry optimization is given by

$$\frac{dE_{\text{CBS+CV}}}{dx} = \frac{dE^{\infty}(\text{HF} - \text{SCF})}{dx} + \frac{d\Delta E^{\infty}(\text{CCSD(T)})}{dx} + \frac{d\Delta E(\text{CV})}{dx} \quad (6)$$

where $dE^{\infty}(\text{HF-SCF})/dx$ and $d\Delta E^{\infty}(\text{CCSD(T)})/dx$ are the energy gradients corresponding to the $\exp(-Cn)$ extrapolation scheme for HF-SCF[51] and to the n^{-3} extrapolation formula for the CCSD(T) correlation contribution,[46] respectively. In the expression given above, n = T, Q, and 5 have been chosen for the HF-SCF extrapolation, while n = T and Q have been used for CCSD(T). Core-correlation effects have been included by adding the corresponding correction, $d\Delta E(\text{CV})/dx$, where the core-correlation energy correction, $\Delta E(\text{CV})$, is obtained as the difference between the all-electron and frozen-core CCSD(T) energies using the core−valence cc-pCVTZ basis set.

For the Ip conformer, the availability of experimental rotational constants of various isotopic species allows the determination of the so-called semiexperimental structure (shortly denoted "SE" in tables). This has been obtained by a least-squares fit of the molecular structural parameters to the equilibrium moments of inertia, I_e^i. The latter are straightforwardly obtained from the corresponding equilibrium rotational constants, B_e^i, which in turn are derived from the experimental ground-state constants, B_0^i, by correcting them for vibrational effects:

$$B_e^i = B_0^i + \frac{1}{2}\sum_r \alpha_r^i \quad (7)$$

In the fitting procedure, the weighting scheme has been chosen in order to have the moments of inertia equally weighted. Experimental ground-state rotational constants for five isotopic species are known,[31] which means that, in addition to the main isotopologue, data for the isotopic substitution at both carbons, nitrogen, and hydrogens bonded to carbon are available. In eq 7, α_r^i are the computed vibration−rotation interaction constants, with r and i denoting the normal mode and the inertial axis, respectively. These constants have been obtained by means of vibrational second-order perturbation theory (VPT2).[27,52−56] The required cubic force field has been computed at the DFT level. Within the DFT approach, the standard B3LYP functional has been used in conjunction with the double-ζ SNSD[57] basis set. Due to the lack of experimental data (i.e., not all the possible isotopic substitution are available), a partial structure has actually been derived, with the nondeterminable geometrical parameters kept fixed at the corresponding best-estimated values.

The fitting procedure deserves a few comments. The fit has been carried out in three steps. First, only the determinable angles have been fitted with all the other parameters kept fixed at the best-estimated values. Then, the obtained angle values have been introduced in the input and kept fixed during the fitting of all determinable distances. Once the obtained values have been included in the input, all the determinable distances and angles have been relaxed. For all three steps, no convergence problems have been encountered. This positive outcome also depends on a proper Z-matrix definition whose free parameters do not include any dihedral angle. As a matter of fact, a more conventional Z-matrix involving dihedral angles led to serious convergence problems.

2.2. Rotational Spectra. Moving to the field of rotational spectroscopy, best-estimated ground-state rotational constants have been obtained by adding to the equilibrium rotational constants corresponding to the best-estimated equilibrium structure the vibrational corrections, ΔB_0^i, at the B3LYP/SNSD level. From eq 7, the corresponding expression is

$$\Delta B_0^i = -\frac{1}{2}\sum_r \alpha_r^i \quad (8)$$

By making use of the harmonic force fields obtained at the levels of theory considered above (i.e., in the frame of the best-estimated molecular structure determination), best estimates for quartic centrifugal-distortion constants, D(best), have been derived by means of the following composite scheme:

$$\begin{aligned} D(\text{best}) &= D(\text{CCSD(T)/VTZ}) \\ &+ [D(\text{MP2/CVTZ, all}) - D(\text{MP2/CVTZ, fc})] \\ &+ [D(\text{MP2/augVTZ, fc}) - D(\text{MP2/VTZ, fc})] \\ &+ [D(\text{MP2/VQZ, fc}) - D(\text{MP2/VTZ, fc})] \end{aligned} \quad (9)$$

where D denotes a generic quartic centrifugal-distortion constant. The first difference (in square brackets) provides the CV correction ($\Delta D(\text{CV})$), the second one the contribution of diffuse functions ($\Delta D(\text{aug})$), and the latter the effect of enlarging the basis set from a triple-ζ to a quadruple-ζ set. Watson's S-reduced Hamiltonian in the I^r representation[58] has been employed.

An analogous composite scheme

$$\begin{aligned} \chi_{ij}(\text{best}) &= \chi_{ij}(\text{CCSD(T)/VTZ}) \\ &+ \chi_{ij}(\text{MP2/CVTZ, all}) - \chi_{ij}(\text{MP2/CVTZ, fc}) \\ &+ \chi_{ij}(\text{MP2/augVTZ, fc}) - \chi_{ij}(\text{MP2/VTZ, fc}) \\ &+ \chi_{ij}(\text{MP2/VQZ, fc}) - \chi_{ij}(\text{MP2/VTZ, fc}) \end{aligned} \quad (10)$$

has been considered for the nitrogen quadrupole-coupling constants χ_{ij}, where ij refers to the principal inertial axes. The additivity scheme has actually been applied to the electric field-gradient components q_{ij}, which are the quantities computed by

dx.doi.org/10.1021/ct3010672 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

electronic-structure calculations. The latter are transformed to nuclear quadrupole-coupling constants by means of

$$\chi_{ij} = eQq_{ij} \tag{11}$$

where eQ in the present case is the nitrogen quadrupole moment, $Q(^{14}N) = 0.02044(3)$ barn, taken from ref 59. The electric field-gradient tensors have been initially obtained for the main isotopic species and then transformed in order to obtain those for the trideuterated [OD,ND_2] species considered. The reasons why this isotopic species has also been investigated will be made clear later in the text.

In ref 60, Halkier et al. showed that the molecular electric dipole moment can be extrapolated to the CBS limit (provided that a hierarchical sequence of basis sets is employed) using the following n^{-3} extrapolation form for the correlation contribution:

$$\Delta\mu^{corr}(n) = \Delta\mu_{\infty}^{corr} + An^{-3} \tag{12}$$

In our investigation, this formula has been applied with n = 3 (triple-ζ, i.e., MP2/cc-pVTZ) and 4 (quadruple-ζ, i.e., MP2/cc-pVQZ). To obtain the extrapolated dipole moment, the CBS limit of the correlation contribution should then be added to the HF-SCF CBS limit, which is assumed in the present case to be reached at the HF-SCF/cc-pV5Z level:

$$\mu(CBS) = \mu_{\infty}^{SCF} + \Delta\mu_{\infty}^{corr} \tag{13}$$

The best estimate of the dipole moment has then been derived by applying a composite scheme aiming at accounting for core-correlation, diffuse function, and high-order electron-correlation effects. In analogy to what has been done for the determination of the best-estimated molecular structure, the best estimates for the electric dipole-moment components have been obtained by the following expression:

$$\mu(best) = \mu(CBS) + \Delta\mu(CV) + \Delta\mu(aug) + \Delta\mu(T) \tag{14}$$

To directly predict experimental results, zero-point vibrational (ZPV) corrections have been computed and added to the best-estimated values. ZPV corrections have been obtained at the B3LYP/SNSD level using the perturbational approach described in ref 56. The latter have been found to be non-negligible, as they are on the order of 5−10%.

2.3. Energetics. In view of establishing accurate energy differences of the IVn and IIIp conformers with respect to Ip, single-point energy calculations at the best-estimated equilibrium structure have been carried out at the CBS+CV level of theory. This means that for both the extrapolation to the CBS limit and inclusion of core-correlation corrections, the CCSD(T) method has been used. CBS total energies have been determined by extrapolating the CCSD(T) correlation contribution to the CBS limit by means of the n^{-3} formula:[46]

$$\Delta E^{corr}(n) = \Delta E_{\infty}^{corr} + A'n^{-3} \tag{15}$$

and by adding the HF-SCF CBS limit, evaluated by the expression[51]

$$E^{SCF}(n) = E_{\infty}^{SCF} + B' \exp(-C'n) \tag{16}$$

The cc-pVTZ and cc-pVQZ basis sets have been employed in the former equation, whereas the cc-pVnZ sets, with n = T, Q, and 5, have been used in the latter.

Making use of the additivity approximation, corrections to take into account CV effects have then been added to CBS energies. This involves carrying out energy computations using the core−valence correlation-consistent cc-pCVTZ basis set in conjunction with the CCSD(T) method. The CV corrections to the total energies are thus given as

$$\Delta E_{CV} = E_{core+val} - E_{val} \tag{17}$$

where $E_{core+val}$ is the CCSD(T) total energy obtained by correlating all electrons and E_{val} is the CCSD(T) total energy obtained in the frozen core approximation, both in the same basis set.

2.4. Vibrational Spectra. Best-estimated harmonic force fields for the three conformers of glycine considered (both main and trideuterated isotopologues) have been evaluated by means of the same composite scheme used to evaluate the best-estimated equilibrium structures. For each conformer, harmonic force fields have been obtained at the various levels of theory using analytic second derivatives.[61] Following the procedure introduced in ref 62, the harmonic frequencies, ω, have been extrapolated to the CBS limit starting from the results obtained at the MP2/cc-pVTZ and MP2/cc-pVQZ levels. The extrapolated correlation contribution has been added to the HF-SCF CBS limit, which is assumed to be reached at the HF/cc-pV5Z level. As for geometries, corrections due to core correlation and effects of diffuse functions in the basis set have then been evaluated at the MP2/cc-pCVTZ ($\Delta\omega$(CV) = ω(MP2/cc-pCVTZ,all) − ω(MP2/cc-pCVTZ,fc)) and MP2/aug-cc-pVTZ levels ($\Delta\omega$(aug) = ω(MP2/aug-cc-pVTZ,fc) − ω(MP2/cc-pVTZ,fc)), respectively. The latter correction has been introduced since diffuse functions are required to properly describe electronegative atoms and also to recover the corresponding limitations affecting the extrapolation procedure when small- to medium-sized basis sets are employed. Higher-order electron-correlation energy contributions, $\Delta\omega$((T)), have been derived by comparing the harmonic frequencies at the MP2 and CCSD(T) levels, both in the cc-pVTZ basis set. The best-estimated harmonic frequencies, ω(best), are then provided by

$$\omega(best) = \omega(CBS(T, Q)) + \Delta\omega(CV) + \Delta\omega(aug) + \Delta\omega((T)) \tag{18}$$

A special comment is deserved for the IIIp conformer, as it has an imaginary frequency at the SCF level (with all basis sets considered). As a consequence, for this frequency the extrapolation to the CBS limit has been carried out by applying the n^{-3} formula to the entire terms, and not only to the correlation contribution (as required). Since the imaginary frequency has been obtained only at the SCF level, we expect that the problem is solved once electron correlation is included in the treatment.

An analogous composite scheme has also been used to determine best estimates for the IR intensities, I(best), within the harmonic approximation. As extrapolation schemes have not been yet formulated for such a quantity, eq 18 has been rearranged as follows:

$$I(best) = I(CCSD(T)/VTZ) + \Delta I(CV) + \Delta I(QZ - TZ) + \Delta I(aug) \tag{19}$$

where ΔI(QZ−TZ) is the correction due to the enlargement of the basis set from triple- to quadruple-ζ.

The computations of vibrational spectra beyond the double-harmonic approximation, the vibrational corrections to rota-

dx.doi.org/10.1021/ct3010672 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 1. Equilibrium Structures of the Ip, IVn, and IIIp Conformers of Glycine[a]

	Ip					IVn	IIIp	
parameters	bestCC[c]	best	r_e^{SEc}	r_e^{SEd}	r_e^{SEe}	best	bestCC	best
C1−O2	1.3476	1.3486	1.34827(31)	1.34839(31)	1.353(1)	1.3461	1.3482	1.3483
C1−O3	1.2021	1.2025	1.20331(62)	1.20319(62)	1.207(2)	1.2028	1.2024	1.2034
O3−C1−O2	123.05	123.07	122.990(31)	123.016(31)	123.2(1)	123.21	122.69	122.73
O2−H4	0.9645	0.9651	0.9645^{fix}	0.9651^{fix}	0.9660^{fix}	0.9651	0.9650	0.9656
H4−O2−C1	106.64	107.10	106.64^{fix}	107.10^{fix}	106.04^{fix}	106.96	105.89	106.38
C1−C5	1.5128	1.5141	1.51318(47)	1.51328(48)	1.511(1)	1.5035	1.5165	1.5174
C5−C1−O2	111.35	111.16	111.482(41)	111.483(42)	111.9(1)	111.62		113.47
C5−N6	1.4430	1.4420	1.44245(14)	1.44220(12)	1.441(1)	1.4488	1.4453	1.4446
N6−C5−C1	115.25	115.38	115.285(17)	115.292(16)	115.4(1)	110.05	118.64	118.63
C5−H7	1.0903	1.0905	1.09078(7)	1.09077(7)	1.0907(2)	1.0968	1.0898	1.0899
C5−H8	1.0903	1.0905	1.09078(7)	1.09077(7)	1.0907(2)	1.0896	1.0898	1.0899
H7−C5−C1	107.49	107.37	[107.36]	[107.36]		105.64	106.03	105.96
H8−C5−C1	107.49	107.37	[107.36]	[107.36]		108.31	106.03	105.96
H7−C5−C1−O3	−123.18	−123.13	[−123.21]	[−123.21]		−106.19	56.10	56.20
H8−C5−C1−O3	123.18	123.13	[123.21]	[123.21]		139.45	−56.10	−56.20
N6−H9	1.0109	1.0109	1.01044(10)	1.01063(11)	1.0065(2)	1.0086	1.0098	1.0097
N6−H10	1.0109	1.0109	1.01044(10)	1.01063(11)	1.0065(2)	1.0119	1.0098	1.0097
H9−N6−C5	109.86	110.55	[110.10]	[109.97]		111.65	110.61	111.35
H10−N6−C5	109.86	110.55	[110.10]	[109.97]		110.28	110.61	111.35
H9−N6−C5−C1	57.93	58.674	[57.43]	[57.34]		−155.36	58.87	59.72
H10−N6−C5−C1	−57.93	−58.674	[−57.43]	[−57.34]		−34.32	−58.87	−59.72
H4−O2−C1−O3	0.0	0.0	0.0	0.0	0.0	1.1206	0.0	0.0
C5−C1−O3−O2	180.0	180.0	180.0	180.0	180.0	178.03	180.0	180.0
N6−C5−C1−O3	0.0	0.0	0.0	0.0	0.0	18.05	180.0	180.0
N6−C5−C1−O2	180.0	180.0	180.0	180.0	180.0			
H4−O2−C1−C5	180.0	180.0	180.0	180.0	180.0			
$\theta(CH_2$ scissor$)^f$			105.988(10)	105.990(10)	105.95(3)			
$\theta(NH_2$ scissor$)^f$					104.98^{fix}			
$\gamma(NH_2$ wag$)^g$					57.67^{fix}			
$\gamma(CH_2$ wag$)^h$					5.4(1)			

[a]Distances in Å, angles in degrees. [b]CBS+CV structure: ref 85. [c]Semiexperimental equilibrium structure: ref 85. Values in square brackets have not been directly fitted. [d]Semiexperimental equilibrium structure: This work. Values in square brackets have not been directly fitted. [e]Semiexperimental equilibrium structure: ref 38. [f]$\theta(XH_2$ scissor) = H−X−H bond angle. [g]$\gamma(NH_2$ wag) = out-of-plane angle: 180°, ∠C5−N6−(H−N−H plane). [h]$\gamma(CH_2$ wag) = out-of-plane angle: ∠N6−C5−(H−C−H plane)−∠C1−C5−(H−C−H plane).

tional constants, and the vibrational contributions to thermodynamic properties have been performed by means of a Hindered-Rotor Anharmonic Oscillator (HRAO) model, within the vibrational second-order perturbation theory (VPT2).[27,52−56] In view of accurately computing the vibrational spectra and vibrational corrections for the conformers considered, a hybrid CCSD(T)/DFT approach has been employed.[21,22,24,63−67] It is based on the assumption that the differences between CCSD(T) and B3LYP results are mainly due to the harmonic terms, and it has already been validated for vibrational frequencies of several closed- and open-shell systems (see for instance refs 21,22 and 65−67) as well as for thermodynamic properties.[68] Anharmonic DFT semidiagonal quartic force fields (i.e., the cubic (K_{ijk}) and semidiagonal quartic (K_{iijj} and K_{iijk}) force constants) have been obtained by numerical differentiation of the analytical second derivatives of the energy (with the standard 0.01 Å step), starting from equilibrium structures optimized using tight convergence criteria. Within the DFT approach, the standard B3LYP functional[69] has been used in conjunction with the double-ζ SNSD[57] basis set, developed for spectroscopic studies of medium-to-large molecular systems. This basis set has been constructed from the polarized double-ζ N07D basis set[57,70−72] by consistently including diffuse *s* functions on all atoms, and one set of diffuse polarized functions (*d* on heavy atoms and *p* on hydrogens). This basis set allows cost-effective prediction of a broad range of spectroscopic properties, including electron-spin resonance (ESR),[70−73] vibrational (IR, Raman, VCD),[14,24,30,73,74] and electronic (absorption, emission, ECD)[14,24,73,75] spectra.

With respect to VPT2 computations, the generalized second-order vibrational perturbation model (GVPT2), as well as the corresponding deperturbed approach (DVPT2),[26,53,55,56,76] as implemented in the Gaussian package,[28−30,56,77,78] have been applied to compute anharmonic frequencies and IR intensities, respectively. In the present work, the criteria chosen to define Fermi resonances are those proposed by Martin et al.,[79] while for the 1−1 resonances present in the perturbative treatment of intensities,[28−30] the criteria proposed by Bloino and Barone[29] have been adopted. In some cases, showing strong anharmonic resonances, it is necessary to include Fermi and/or Darling-Dennison resonances in a variational treatment of polyad couplings,[80] and our GVPT2 model allows such computations for vibrational energies.[56] However, for molecular systems devoid of the problem, the DVPT2 approach represents a versatile way of calculating anharmonic IR intensities. Vibrational contributions to thermodynamic properties have been evaluated by means of a resonance-free perturbative approach

dx.doi.org/10.1021/ct3010672 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

within the hybrid degeneracy corrected second-order perturbation theory (HDCPT2).[68] The latter provides an automatic treatment of internal rotations through the hindered-rotor[81] model in conjunction with simple perturbation theory (SPT)[82] reformulated to treat consistently both energy minima and transition states.[68]

In all cases, the best-estimated harmonic frequencies have been introduced directly into the VPT2 computations along with the cubic and quartic force constants obtained at the DFT level. In the view that the DFT, MP2, and CCSD(T) normal modes are very similar (as expected in most cases), DFT cubic and quartic force constants have been used without any transformation.

MP2 and CCSD(T) calculations have been carried out with the quantum-chemical CFOUR program package.[83] All DFT and VPT2 computations have been performed employing a locally modified version of the GAUSSIAN suite of programs for quantum chemistry.[84]

3. RESULTS AND DISCUSSION

Molecular structures of the Ip, IVn and IIIp glycine conformers, as obtained by the composite scheme mentioned above, are collected in Table 1, where the atom labeling is given in Figure 1. The results for Ip are reported because the availability of the CCSD(T)/CBS+CV structure[85] as well as of semiexperimental equilibrium structures[38,85] allow us to verify the accuracy of our computed best estimates for the other two conformers. From the comparison reported in Table 1, we first note that all our structures agree very well with each other, the largest deviation being 0.001 Å for distances and 0.3° for angles. As already discussed in ref 85, our anharmonic force field together with improved geometrical parameters enabled the revision of the semiexperimental equilibrium structure. As mentioned in section 2.1, since only a partial molecular structure can be determined, the accuracy of the nondeterminable parameters might play a role. For this reason, we decided to compare the semiexperimental structures obtained by fixing them at either the corresponding best or bestCC values but also at the fixed parameters of ref 38. While there is a very good agreement between our two semiexperimental equilibrium geometries (ref 85 and this work), we note non-negligible discrepancies for the semiexperimental structure of ref 38. In particular, the latter overestimates the C−O distances by about 0.005 Å, whereas it underestimates the N−H bond lengths by about the same extent. Furthermore, the H4−O2−C1 angle was kept fixed at a value, which results in being underestimated by about 0.6−1°. For the semiexperimental structure determination carried out by fixing the undeterminable parameters to the values of ref 38, a slight improvement with respect to ref 38 itself is noted, but still discrepancies of about 0.002−0.003 Å are observed. We therefore mainly ascribe the improved agreement to the best description of the large amplitude motions provided by the force field at the B3LYP/SNSD level [this topic will be addressed in detail in a manuscript in preparation, which concerns a thorough investigation on the performance of the semiexperimental approach in the derivation of equilibrium structures], even if more accurate values for the fixed parameters provide non-negligible effects. For the conformers of C_s symmetry (Ip and IIIp), Table 1 also reports the structural parameters obtained from the second type of composite scheme applied (eq 6), i.e., the CCSD(T)/CBS+CV results (under the notation "bestCC"). In view of the high accuracy that characterizes such structural determinations (see, for instance, refs 12, 15, and 48−50 and references therein), their comparison with the results from the first composite scheme allows us to point out the accuracy of the latter. We note that the differences in the bond lengths are well within 0.001 Å. For angles, the deviations are usually smaller than 0.1° and well within 0.5° in almost all cases. The present comparison thus confirms the conclusions drawn in ref 20; that is, the less expensive composite scheme can be easily applied to rather large molecules allowing one to obtain very accurate structure determinations. Given the comparisons discussed above and on the basis of the literature on this topic (see, for example, refs 12, 49, 86, and 87 and references therein), we are then confident that the molecular structures obtained for the IVn and IIIp conformers have an accuracy of about 0.001 Å for bond lengths and 0.5° for angles.

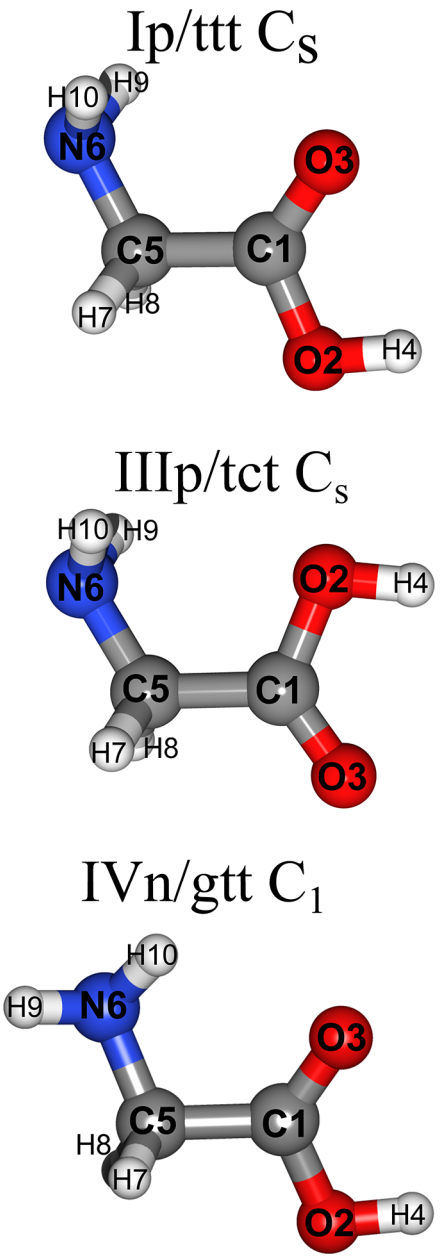


Figure 1. Geometry structures and atom labeling of the Ip/ttt, IIIp/tct, and IVn/gtt conformers of free glycine.

We can take advantage of this great accuracy for providing the rotational parameters required for predicting the rotational

dx.doi.org/10.1021/ct3010672 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 2. Best-Estimated Rotational Parameters for the Main and Trideuterated [OD,ND_2] Species of the Ip, IVn and IIIp Conformers

	Ip			IVn		IIIp	
	main						
	theo	exptl[a]	trideut	main	trideut	main	trideut
A_0/MHz	10319.9	10341.521(89)	9918.5	10425.0	10163.8	9903.4	9082.4
B_0/MHz	3870.4	3876.1785(12)	3458.7	3962.1	3510.1	3983.3	3655.9
C_0/MHz	2908.4	2912.3509(10)	2679.1	2945.4	2675.4	2953.0	2740.4
D_J /kHz	0.762	0.7434(33)	0.737	0.698	0.571	10.057	14.177
D_{JK} /kHz	4.184	3.986(17)	2.992	4.019	3.098	−14.056	−23.683
D_K /kHz	3.243		2.964	5.703	5.289	12.040	15.589
d_1 /kHz	−0.197	−0.190l(28)	−0.157	−0.187	−0.143	−0.222	−0.211
d_2 /kHz	−0.014	−0.0158(17)	0.030	−0.056	−0.035	4.601	6.681
χ_{aa} /MHz	−1.278		−1.057	2.613	2.639	−1.596	−1.574
χ_{bb} /MHz	−0.464		−0.685	2.089	2.073	−0.218	−0.239
χ_{cc} /MHz	1.742		1.742	−4.701	−4.713	1.814	1.814
$\lvert\chi_{ab}\rvert$ /MHz	3.830		3.847	0.655	0.649	3.838	3.842
$\lvert\chi_{ac}\rvert$ /MHz				0.661	0.591		
$\lvert\chi_{bc}\rvert$ /MHz				−0.072	−0.097		
μ_a/D	0.809			0.156		−0.600	
μ_b/D	−0.838			1.678		−1.683	

[a]Ref 31.

Table 3. Theoretical[a] Thermodynamic Properties of the Two Higher-Energy Conformers of Glycine (IVn and IIIp) Compared to Available Experimental Results

	IVn				IIIp			
	HO	HO+HR[b]	SPT(HRAO)[b,c]	exptl.	HO	HO+HR[b]	SPT(HRAO)[b,c]	exptl.
ΔE_{ele} [kJ·mol^{-1}]	4.87				7.44			
ΔE_{ZPVE} [kJ·mol^{-1}]	4.81		4.74		7.48		7.94	
temperature = 15 K								
ΔH [kJ·mol^{-1}]	4.81	4.82	4.75		7.59	7.55	7.90	
ΔG [kJ·mol^{-1}]	4.81	4.82	4.75		7.34	7.28	7.87	
temperature = 410 K								
ΔH [kJ·mol^{-1}]	4.59	4.68	4.62	4.81[d]	7.61	6.59	6.62	5.8[e]
ΔG [kJ·mol^{-1}]	5.97	5.99	5.78		−1.17	0.04	9.72	

[a]Conformational energies. All values (except ΔE_{ele}) at 1 atm. [b]The two lowest vibrations have been described by hindered-rotor contributions computed by automatic procedure.[81] [c]Contributions computed by the HDCPT2[68] model using the hybrid CC/DFT force field, in conjunction with simple perturbation theory (SPT)[68,82] (see text for the details). [d]Gas-phase thermodynamic data from ref 35, obtained from the Raman band ratios using the van't Hoff scheme. [e]Experimental low-temperature matrix data from ref 91, based on the integrated IR intensities of ν(C═O) from the samples evaporated at 358 and 438 K.

spectra, namely, the vibrational ground-state rotational constants and the quartic centrifugal-distortion constants. As described in detail in section 2.2, the former are derived by adding the vibrational corrections to the equilibrium rotational constants corresponding to the best-estimated geometries, while the centrifugal-distortion constants are obtained from the best-estimated harmonic force field. As for the molecular structure, we used the available experimental data for the Ip conformer to verify the accuracy obtainable from our computations. The results for the main isotopic species as well as those for the [OD,ND_2] trideuterated isotopologue of the Ip, IVn, and IIIp conformers are collected in Table 2. From the comparison to experimental data for Ip, we note that the rotational constants are well reproduced, with discrepancies of about 0.1−0.2%, and the quartic centrifugal-distortion constants agree well within 5−10%. These findings are in line with our previous work on uracil[20] and denote a great accuracy (see ref 12 for a review on this topic). The trideuterated species has also been considered because, as will be explained later on in the discussion of the IR spectra, it might provide a route to observing the elusive IVn conformer. Since for flexible molecules nuclear quadrupole-coupling constants might help in elucidating which conformer has been observed,[88−90] in Table 2 we also report the nitrogen quadrupole-coupling constants obtained by means of the composite scheme introduced in section 2.2. From Table 2, it is evident that, while Ip and IIIp present similar nitrogen quadrupole-coupling constants, those of the IVn conformer are quite different in magnitude and opposite in sign. In the fortunate case that the rotational spectra of both conformers are measured, the nuclear quadrupole-coupling constants would be more effective than rotational constants in the discrimination of the conformers.

A large set of first-order properties has been obtained as a byproduct of the geometry optimizations carried out in the frame of the composite scheme aiming at the determination of the best-estimated equilibrium structure. As a consequence, the additivity scheme could also be applied to the determination of the best-estimated dipole moment components. In addition to its intrinsic importance as molecular properties, the latter provide information of the type of rotational spectrum that can

dx.doi.org/10.1021/ct3010672 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

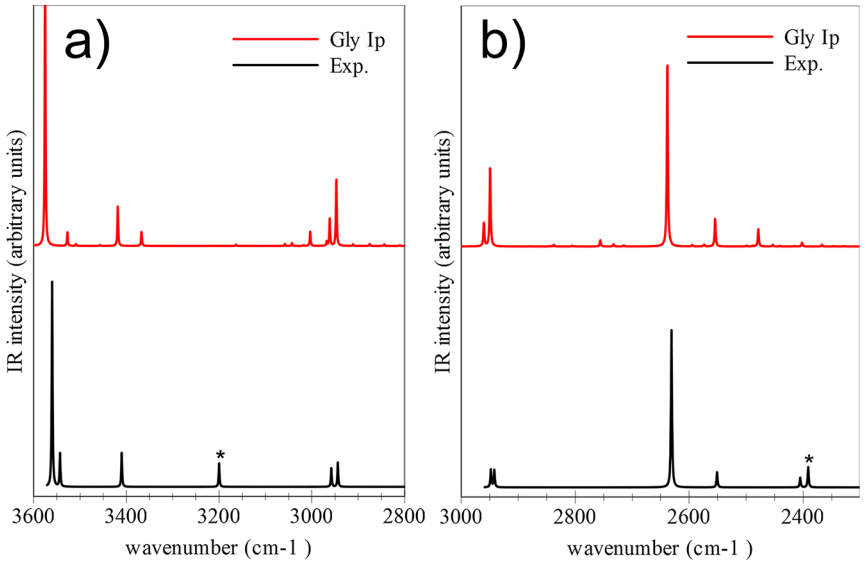


Figure 2. Best-estimated IR spectra in the high-frequency region for the (a) main and (b) trideuterated [OD,ND_2] species of the Ip glycine conformer. Experimental IR spectra recorded in low-temperature Ar matrix generated using the data of Tables 5 and 7 of ref 32. Bands marked by an asterisk have been assigned to the IIn conformer.

be observed. From the results collected in Table 2, it is clear that the *b*-type rotational spectrum is expected to be intense for all conformers, about 5 times more intense for IVn and IIIp than for Ip. The opposite trend is observed for the *a*-type rotational spectrum, which is expected to be rather weak for IVn and about 11 times more intense for IIIp and about 27 times more intense for Ip. For the latter, the *a*-type spectrum is the most intense.

The gas-phase (15 and 410 K, 1 atm) thermodynamic properties of the IVn and IIIp conformers of glycine, computed as differences with respect to the Ip conformer (i.e., conformational energies), are listed in Table 3 along with the experimental enthalpy value for IVn and IIIp, obtained from the analysis of the Raman band ratios using the van't Hoff scheme and integrated IR intensities of ν(C═O), respectively; for all details the reader is referred to refs 35 and 91. The energy differences at 0 K have been computed by means of the composite scheme[15,16] explained in section 2.3. According to the literature on this topic (see for instance refs 92−94 and references therein), for energy differences between rotational conformers, the post-CCSD(T) contributions are unimportant, and an uncertainty of 1 kJ·mol^{-1} for our results turns out to be a conservative estimate. We furthermore note that our accurate computations essentially confirm the stability order for the IVn and IIIp conformers already pointed out in refs 38 and 39, with IVn more stable by less that 3 kJ·mol^{-1}. Vibrational contributions to thermodynamic properties have been included as mentioned above and thoroughly detailed in ref 68. While computational studies of thermodynamic properties for biomolecule building blocks can nowadays provide a typical accuracy of about 10 kJ·mol^{-1},[95] the methodology presented in this work has been demonstrated to allow their evaluation with an overall accuracy of about, or better than, 1 kJ·mol^{-1} for conformational enthalpies.[68,85] This is also confirmed by the present results; in fact, all approaches show a very good agreement with the experimental data available, with the values nearly within experimental error bars (4.6 vs 4.8 ± 0.3 kJ mol^{-1} and 6.6 vs 5.8 ± 0.6 kJ mol^{-1}, for IVn and IIIp, respectively). However, it should be pointed out that accurate computations of entropy contributions (1 J·mol^{-1}·K^{-1}) to free energies require the proper treatment of low-frequency torsional motions and anharmonic effects.[68] In fact, Table 3 shows that only the full anharmonic-oscillator/hindered-rotor approaches lead to a reasonable value for the IIIp conformer.

For both the isotopologues (main and trideuterated species) of the Ip, IVn, and IIIp conformers considered in this work, IR spectra have been computed according to the procedure described in section 2.4. The quality of the simulated IR spectra for the Ip conformer is demonstrated in Figure 2 by their comparison with the experimental ones recorded in low-temperature Ar matrix.[32]

A brief comment on the expected accuracy of our best-estimated vibrational frequencies and intensities is warranted. In the literature, various papers demonstrate the reliability and accuracy of the hybrid CC/DFT approaches in accurately reproducing or predicting IR spectra.[21,22,63−67] In particular, our recent work on uracil shows its successful application to biomolecules,[21] while ref 68 demonstrates that the issues deriving from Fermi and Coriolis resonances in the frame of VPT2 can be readily solved, thus leading to reliable and accurate values for both vibrational frequencies and related thermodynamic properties. As far as best-estimated harmonic frequencies are concerned, the role of higher-order effects in the correlation treatment beyond CCSD(T) and their impact on accuracy should be addressed. From the available literature (see for example refs 96−99), it turns out that the full CC singles, doubles, and triples method (CCSDT)[100−102] is expected to provide negligible improvements with respect to CCSD(T) (i.e., in general smaller than 1 cm^{-1}), while corrections due to quadruple excitations are expected to be larger (i.e., on the order of 2−6 cm^{-1}). On the other side, the importance of including core-correlation effects is well recognized.[97,98,103] In conclusion, on the basis of the approximations made and of the estimates for neglected contributions, we expect that the accuracy of our best-estimated harmonic frequencies is a few wavenumbers: from 4 to 11 cm^{-1}, where the latter value essentially applies to the larger frequency values. Regarding IR intensities, in view of the lack of literature

dx.doi.org/10.1021/ct3010672 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 4. Vibrational Frequencies Computed by Different Computational Models for the Most Stable (Ip) Glycine Conformer Compared to Selected Experimental Results

mode	symm.	B3LYP/SNSD		M05-2X/6-31+G(d)		B2PLYP/B3LYP[a]	CC/B3LYP	exptl.[b]		assignment[a,c]
		scaled[d]	anh	scaled[e]	anh[f]	anh	anh	ν	ref.	
υ_1	A′	3590	3551	3582	3634	3566	3575	3585[g]	33	OH stretch
υ_2	A″	3434	3403	3443	3473	3422	3418	3410	32, 37	NH_2 (A) stretch
υ_3	A′	3363	3356	3432	3428	3401	3367	3359	37	NH_2 (S) stretch
υ_4	A″	2958	2931	3045	3030	2957	2961	2969	91	CH_2 (A) stretch
υ_5	A′	2923	2924	3008	3023	2963	2947	2943	37	CH_2 (S) stretch
υ_6	A′	1745	1784	1798	1841	1768	1774	1779	32	C═O stretch
υ_7	A′	1606	1604	1651	1677	1637	1612	1608	37	NH_2 bend
υ_8	A′	1395	1418	1440	1461	1450	1435	1429	32, 37	CH_2 bend x
υ_9	A′	1340	1373	1386	1402	1366	1387	1405	37	CH_2 bend z
υ_{10}	A″	1328	1343	1350	1366	1357	1353	1340	91	CH_2 NH_2 twist
υ_{11}	A′	1251	1280	1275	1283	1296	1286	1297	37	(OH+CH_2x) bend
υ_{12}	A″	1136	1155	1157	1183	1169	1164	1166	37	CH_2 NH_2 twist
υ_{13}	A′	1116	1131	1153	1150	1123	1144	1136	32, 37	CN str. + OH bend
υ_{14}	A′	1080	1089	1124	1125	1094	1103	1101	32, 37	CO str. + OH bend
υ_{15}	A′	883	862	919	721	870	863	883	32, 37	CC str. + NH_2 bend
υ_{16}	A″	879	907	899	917	915	907	907	32, 37	CH_2 NH_2 twist
υ_{17}	A′	788	796	819	755	809	802	801	32, 37	CC stretch
υ_{18}	A″	628	616	649	354	626	603	619	32, 37	OH oop bend
υ_{19}	A′	609	629	622	631	631	633	615	37	(NH_2 + OCO) bend
υ_{20}	A″	488	491	500	370	500	494	500	32	OH oop bend
υ_{21}	A′	445	458	454	461	462	461	458[h]	34	
υ_{22}	A′	246	256	254	221	262	255	250[h]	34	
υ_{23}	A″	204	202	248		229	203	204[h]	34	
MIN[i]		−65	−38	−22	−265	−39	−20			
MAX[i]		24	14	76	80	42	18			
MAE[i]		18	11	25	57	15	8			

[a]Hybrid B2PLYP/aug-cc-pVTZ//B3LYP/aug-N07D anharmonic GVPT2 frequencies from ref 30. [b]Experimental data from IR spectrum recorded in low-temperature Ar matrix. [c](A) and (S) refer to "asymmetric" and "symmetric," respectively; oop stands for out-of-plane. [d]B3LYP/SNSD harmonic frequencies scaled by a unique 0.96 scaling factor (see for example ref 112). [e]M05-2X/6-31+G(d) harmonic frequencies scaled by 0.952 for OH stretch, 0.943 for NH stretch (ref 106), by the average best-fit scale factors 0.9578 for the amide I (C═O) vibrations and 0.9538 for the amide II (NH_2+CN) vibrations (ref 105), and by 0.96 for all other modes. [f]Anharmonic frequencies for modes showing unreliable anharmonic corrections at the M05-2X level have been crossed out [g]Experimental data from IR spectra recorded in helium nanodroplets. [h]Experimental data from Raman spectrum of jet-cooled glycine. [i]MIN and MAX stands for signed errors, largest positive (MAX) and largest negative (MIN). MAE stands for Mean Absolute Error.

on this topic, it is difficult to estimate the accuracy of our best estimated values, we can only mention the successful comparison with experimental data pointed out by previous works (see for example refs 14, 24, 29, and 30). To fill this lack, a benchmark investigation on IR intensities for small- to medium-sized molecules is in progress.

A more detailed account on the accuracy of vibrational frequencies can be derived from Tables 4 and 5, where vibrational frequencies computed at different levels are compared to the experimental data obtained in low-temperature rare-gas matrices[32] and nanodroplets,[33] as well as in the gas phase.[34] The comparison of frequency values obtained in different environments[32−34,36,37] shows that the matrix effects affect the O−H stretching vibration (e.g., they red-shift the corresponding frequency up to 25 cm^{-1} for the Ip conformer) but are much less pronounced for the other observed transitions. It has already been noted that B3LYP/SNSD GVPT2 computations provide accurate results, with maximum deviations (in absolute terms) not exceeding 38 cm^{-1} and the mean absolute errors (MAE) of 11 cm^{-1} and 12 cm^{-1} for the Ip and IIIp conformers, respectively. As demonstrated in ref 30, improved results are obtained by replacing harmonic B3LYP frequencies with their B2PLYP/aug-cc-pVTZ counterparts.[23] The approach proposed in the present work leads to further improved results, thus in better agreement with experimental data. In detail, maximum discrepancies are reduced to 20 cm^{-1}, and MAE decreases below 10 cm^{-1}. As an example, we mention the O−H stretching frequency of the Ip conformer, for which the B3LYP level yields 3551 cm^{-1}, hybrid B2PLYP/B3LYP 3568 cm^{-1}, and CC/B3LYP 3575 cm^{-1}, to be compared with 3585 cm^{-1} (measured for glycine trapped in helium clusters[33]).

Tables 4 and 5 also report scaled harmonic frequencies. We decided to additionally provide this information since such a scaling methodology is still largely applied to the conformational analysis of spectroscopic data for biomolecule building blocks. For this reason, in addition to B3LYP, the M05-2X[104] functional, recently employed in the conformation-specific spectroscopy studies of model polypeptides,[105,106] has also been considered. On average, the scaled B3LYP frequencies differ from experimental data by about 15−20 cm^{-1}, while the M05-2X results show MAE's of 20−35 cm^{-1} and maximum discrepancies over 100 cm^{-1}. The latter results (and the corresponding anharmonic frequencies) further confirm that the M05-2X functional, despite its success in predicting accurate structural data for weakly bound flexible systems,[107,108] is not suitable for anharmonic vibrational analysis.[23,109]

dx.doi.org/10.1021/ct3010672 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 5. Vibrational Frequencies Computed by Different Computational Models for the IIIp Glycine Conformer Compared to Available Experimental Results

mode	symm.	B3LYP/SNSD		M05-2X/6-31+G(d)		B2PLYP/B3LYP[a]	CC/B3LYP	exptl[b]		assignment[a,c]
		scaled[d]	anh	scaled[e]	anh[f]	anh	anh	ν	ref.	
υ_1	A′	3589	3557	3586	3456	3562	3576	3579[g]	33	OH stretch
υ_2	A″	3446	3419	3295.9	3495	3440	3434	3430	37	NH_2 (A) stretch
υ_3	A′	3372	3364	3255.0	3452	3385	3391			NH_2 (S) stretch
υ_4	A″	2965	2963	3052.5	2991	2959	2967	2968	37	CH_2 (A) stretch
υ_5	A′	2930	2912	3015.3	2976	2955	2953			CH_2 (S) stretch
υ_6	A′	1739	1778	1796	1829	1761	1770	1767	32, 37	C═O stretch
υ_7	A′	1603	1638	1648.6	1693	1652	1674	1624	37	NH_2 bend
υ_8	A′	1392	1417	1440.0	1445	1438	1432			CH_2 bend
υ_9	A″	1326	1341	1350.2	1363	1353	1351			CH_2 NH_2 twist
υ_{10}	A′	1303	1318	1339.7	1314	1330	1343	1352	37	CH2 bend z
υ_{11}	A′	1291	1332	1326.8	1330	1346	1336	1325	37	OH bend, C−O str
υ_{12}	A″	1141	1171	1163.3	1172	1167	1172	1177	37	CH_2 NH_2 twist
υ_{13}	A′	1124	1122	1162.4	1140	1123	1147	1147	32	(CO+CC) str. + OH bend
υ_{14}	A′	1086	1102	1125.2	1129	1111	1135	1101	37	CN str.
υ_{15}	A″	870	895	893.5	920	903	902	903	37	CH_2 NH_2 twist
υ_{16}	A′	854	828	890.9	783	828	832	852	32, 37	CC str. + NH_2zbend
υ_{17}	A′	767	771	796.3	758	771	779	777	32, 37	CC stretch
υ_{18}	A″	656	652	674.0	625	660	646	648	32, 37	COOH tor
υ_{19}	A′	571	600	581.7	585	604	596	594	37	(NH_2+OCO) bend
υ_{20}	A″	493	518	506.4	467	503	509			
υ_{21}	A′	478	481	490.0	502	489	499			
υ_{22}	A′	251	259	279.4		263	269			
υ_{23}	A″	237	227	262.3		224	246	237	91	
MIN[h]		−49	−34	−134	−123	−24	−20			
MAX[h]		16	14	85	69	28	18			
MAE[h]		17	12	33	42	13	7			

[a]Hybrid B2PLYP/aug-cc-pVTZ//B3LYP/aug-N07D anharmonic GVPT2 frequencies from ref 30. [b]Experimental data from IR spectrum recorded in low-temperature Ar matrix. [c](A) and (S) refer to "asymmetric" and "symmetric," respectively; oop stands for out-of-plane. [d]B3LYP/SNSD harmonic frequencies scaled by a unique 0.96 scaling factor (see for example ref 112). [e]M05-2X/6-31+G(d) harmonic frequencies scaled by 0.952 for OH stretch, 0.943 for NH stretch (ref 106), by the average best-fit scale factors 0.9578 for the amide I (C═O) vibrations and 0.9538 for the amide II (NH_2+CN) vibrations (ref 105), and by 0.96 for all other modes. [f]Anharmonic frequencies for modes showing unreliable anharmonic corrections at the M05-2X level have been crossed out [g]Experimental data from IR spectra recorded in helium nanodroplets. [h]MIN and MAX stands for signed errors, largest positive (MAX) and largest negative (MIN). MAE stands for Mean Absolute Error.

Moreover, harmonic M05-2X (and M06-2X) frequencies should not be used to estimate anharmonic vibrational frequencies based on a single scaling factor, and only more sophisticated scaling methods[110] lead to reliable results for these functionals. Furthermore, although scaled B3LYP harmonic frequencies on average agree well with experimental data, they show quite large errors, of about 30 cm^{-1}, for the most intense mode (the C═O stretch), also known as the "amide I mode" and widely used in the analysis of complex biomolecules. On the contrary, all B3LYP and hybrid anharmonic models present much smaller discrepancies, at a maximum of 11 cm^{-1}, and within 5 cm^{-1} for the best estimates computed in this work. By comparing all amide I bands from ref 32 (1790, 1779, 1767 cm^{-1}) to our best-estimated values, we note that it is possible to assign each band to each glycine conformer considered with an error of ±5 cm^{-1}: IIn/1793 cm^{-1}, Ip/1774 cm^{-1}, and IIIp/1770 cm^{-1}. In our opinion, for larger biomolecules, where several conformational minima can be present concomitantly, improved accuracy of estimated vibrational frequencies becomes even more important for a reliable analysis of experimental data. Therefore, in view of the feasibility of anharmonic force field computations at the DFT level (reduced-dimensionality schemes are also available[111]), we recommend to apply the VPT2 approach and improve, whenever possible, the harmonic part by means of high-level *ab initio* calculations. It should be pointed out that, at variance with simple methodologies based on double-harmonic approximation, anharmonic spectra take also into account intensities of overtones and combination bands. This allows one to dissect between low-intensity features related to nonfundamental transitions of the most abundant conformer and the fundamental transitions of the less abundant ones. The same computational procedure can be easily applied also to a set of isotopologues in view of obtaining an unequivocal identification of several concomitantly present conformers when the situation is not so clear for the main isotopic species.

Once the accuracy of our best-estimated vibrational frequencies is established, we can focus on the search for spectral fingerprints of the "IR-missing" IVn conformer. The best candidate is the 2800−2900 cm^{-1} range and the C5−H7 stretching frequency, which differs significantly from that of the Ip conformer. Indeed, several weak features (a small peak at 2878 cm^{-1} and a small broad band between 2835 and 2850 cm^{-1}) have been observed,[36] but the presence of small amounts of the IVn conformer in the deposited matrix could not be established without ambiguities, mostly because of overtones of Ip, which may show intensities comparable to the small- to medium-intensity fundamental band of the low-concentration

dx.doi.org/10.1021/ct3010672 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

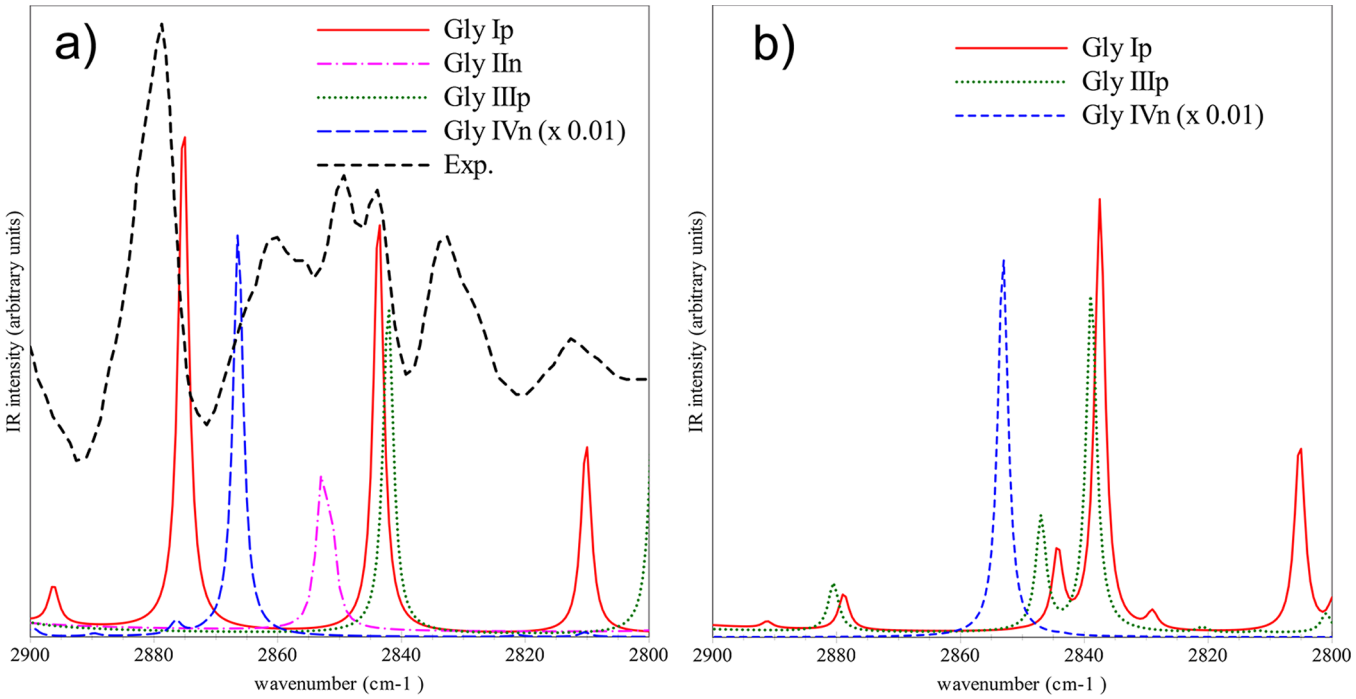


Figure 3. Best-estimated IR spectra in the 2800−2900 cm^{-1} frequency region for the (a) main and (b) trideuterated [OD,ND_2] species of the Ip, IVn, and IIIp glycine conformers. For the main isotopologue, the spectrum of the IIn conformer is also shown. Experimental IR spectra recorded in low-temperature Ar matrix from Figure 5 of ref 36.

IVn conformer. Figure 3 and Table 6 show that some spectral features can be indeed related to overtones and combination bands of Ip, namely, the first overtone of ν_8 together with $\nu_8+\nu_9$ and $\nu_6+\nu_{14}$ combination bands at 2844, 2810, 2875 cm^{-1}, respectively. In particular, the band at 2875 cm^{-1} corresponding to the combination of the C═O stretching and C−O/O−H bending is rather intense and shows a frequency value very close to the C5−H7 stretching (ν_5) of IVn at 2866 cm^{-1}. However, the quite complex band pattern is best described by taking also into account the C5−H7 stretching of the IVn conformer, whose intensity has been scaled to 1% due to the possibility of conformational cooling by either tunneling effects or transformation over a classical barrier that might lower the IVn abundance with respect to the 10% obtained from a direct calculation of the Boltzmann population at 410 K. The former processes are also present in low-temperature matrices, explaining the more feasible detection in the nonequilibrium conditions of jet-cooled molecular beams, in particular close to the entrance of the nozzle.[34] Our computations suggest a more promising situation for the trideuterated [OD,ND_2] species, for which (*i*) the spectrum related to Ip shows less intense features and (*ii*) the energy gap between the C5−H7 stretching (ν_2) of the IVn conformer and the most intense (and close-lying) band of the Ip conformer increases to 15 cm^{-1}. The former band, which is slightly shifted with respect to the nondeuterated species, is predicted at 2853 cm^{-1}, while the latter (first overtone of ν_7, CH_2 bending) is predicted at 2837 cm^{-1}. Furthermore, a higher percentage of IVn in the experimental mixture can be foreseen due to a lower probability of tunneling effects for ND_2 with respect to NH_2, which are considered responsible for the low abundance of IVn.[36,37] Thus, supported by the detailed computational analysis reported in Table 6, we suggest a reinvestigation of the 2800−2900 cm^{-1} energy range for trideuterated [OD,ND_2] glycine in search for the "missing" IVn conformer, which has not been considered in the study by Adamowicz et al.[32]

The presence of overtones and the related problems is removed at low frequency, and recent studies of the 150−500 cm^{-1} jet-cooled Raman spectrum indeed provided evidence for the IVn conformer. Our best-estimated harmonic and anharmonic frequencies of low-frequency bands (see Table 7) confirm the conclusions based on the standard methodologies applied by Balabin,[34] also showing a good agreement with available experimental results from both Raman[34] and IR[32] spectra. We note that in general the lowest frequencies do not show strong anharmonic effects. Therefore the best-estimated harmonic frequencies agree well with the experimental values, and in most cases, the agreement is further improved by accounting for anharmonicity. Only in the case of frequencies related to the hindered rotations the VPT2 approach might overestimate the corresponding anharmonic correction, so that the harmonic and anharmonic values can be considered as upper and lower limits, respectively, as in the case of ν_{23} for the IVn conformer. Taking into account the overall accuracy of the best-estimated frequencies, the data reported for the Ip, IIIp, and IVn conformers of the trideuterated [OD,ND_2] glycine can support further Raman spectroscopy investigations of these species.

4. CONCLUDING REMARKS

In the present work, state-of-the-art quantum-chemical computations allowed us to draw a fully consistent interpretation of the most recent experimental data available for glycine and to suggest new routes for a more complete characterization of the elusive IVn and IIIp conformers. In detail, our computations showed that the trideuterated [OD,ND_2] species could be a good candidate for observing the IR spectrum of the IVn conformer, since in such a case the Ip conformer is predicted to interfere less and a lower

dx.doi.org/10.1021/ct3010672 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

Table 6. Best-Estimated Vibrational Frequencies (cm^{-1}) and IR Intensities (km/mol) of Glycine in the High-Frequency Region of IR Spectra[a]

NH_2-CH_2-COOH				ND_2-CH_2-COOD			
freq	int[b]	conf	assignment	freq	int[c]	conf	assignment
3579	65.5	IVn	υ_1	2970	7.4	IIIp	υ_1
3576	63.7	IIIp	υ_1	2962	0.8	IVn	$\upsilon_2+\upsilon_{24}$
3575	58.6	Ip	υ_1	2961	11.0	IIIp	υ_2
3538	2.8	IVn	$2\upsilon_6$	2960	5.0	Ip	υ_1
3527	2.8	Ip	$2\upsilon_6$	2958	14.1	IVn	υ_1
3520	3.2	IIIp	$2\upsilon_6$	2949	16.6	Ip	υ_2
3509	0.4	Ip	$\upsilon_2+\upsilon_{24}$	2853	29.8	IVn	υ_2
3507	3.7	IIIp	$\upsilon_2+\upsilon_{24}$	2837	0.3	Ip	$2\upsilon_7$
3472	7.5	IIIp	$\upsilon_3+\upsilon_{24}$	2805	0.2	Ip	$\upsilon_7+\upsilon_8$
3457	0.2	Ip	$\upsilon_4+\upsilon_{20}$	2779	1.1	IVn	$\upsilon_6+\upsilon_{13}$
3441	15.0	IVn	υ_2	2772	2.7	IVn	$2\upsilon_8$
3434	6.0	IIIp	υ_2	2756	1.3	Ip	$\upsilon_6+\upsilon_{14}$
3418	8.2	Ip	υ_2	2739	1.7	IIIp	$\upsilon_6+\upsilon_{14}$
3391	3.6	IIIp	υ_3	2733	0.5	Ip	$2\upsilon_8$
3367	3.0	Ip	υ_3	2715	0.2	Ip	$\upsilon_4+\upsilon_{23}$
3361	6.9	IVn	υ_3	2706	3.3	IVn	$\upsilon_6+\upsilon_{15}$
3163	0.3	Ip	$\upsilon_4+\upsilon_{23}$	2639	40.5	IVn	υ_3
3103	0.9	IIIp	$\upsilon_6+\upsilon_{10}$	2638	39.6	Ip	υ_3
3058	0.4	Ip	$\upsilon_4+\upsilon_{24}$	2636	40.7	IIIp	υ_3
3047	4.8	IIIp	$\upsilon_5+\upsilon_{24}$	2632	1.3	IIIp	$\upsilon_4+\upsilon_{24}$
3043	0.6	Ip	$\upsilon_6+\upsilon_{11}$	2632	0.2	Ip	$\upsilon_4+\upsilon_{24}$
3040	8.2	IIIp	$\upsilon_4+\upsilon_{24}$	2595	0.3	Ip	$\upsilon_8+\upsilon_{10}$
3017	0.2	Ip	$\upsilon_5+\upsilon_{24}$	2585	0.9	IVn	$\upsilon_5+\upsilon_{23}$
3013	1.9	IVn	$\upsilon_7+\upsilon_9$	2573	0.3	Ip	$\upsilon_8+\upsilon_{11}$
3008	0.7	IVn	$\upsilon_6+\upsilon_{11}$	2573	2.7	IIIp	$\upsilon_5+\upsilon_{24}$
3004	3.0	Ip	$\upsilon_7+\upsilon_9$	2567	4.3	IIIp	υ_4
2973	0.7	IIIp	$\upsilon_7+\upsilon_9$	2566	10.2	IVn	υ_4
2968	0.9	Ip	$\upsilon_7+\upsilon_{10}$	2554	5.8	Ip	υ_4
2967	7.5	IIIp	υ_4	2536	1.4	IIIp	$2\upsilon_{10}$
2961	5.6	Ip	υ_4	2499	0.2	Ip	$\upsilon_9+\upsilon_{10}$
2956	12.1	IVn	υ_4	2495	2.9	IIIp	υ_5
2953	10.0	IIIp	υ_5	2489	0.1	Ip	$\upsilon_7+\upsilon_{13}$
2947	14.2	Ip	υ_5	2479	3.6	Ip	υ_5
2915	0.8	IIIp	$\upsilon_6+\upsilon_{13}$	2478	3.5	IVn	υ_5
2911	0.3	Ip	$\upsilon_6+\upsilon_{13}$	2453	0.4	Ip	$\upsilon_8+\upsilon_{12}$
2900	0.7	IVn	$\upsilon_6+\upsilon_{13}$	2437	0.9	IVn	$2\upsilon_{10}$
2876	1.0	IVn	$\upsilon_8+\upsilon_9$	2402	0.8	Ip	$2\upsilon_{11}$
2875	0.4	Ip	$\upsilon_6+\upsilon_{14}$				
2866	32.1	IVn	υ_5				
2844	0.3	Ip	$2\upsilon_8$				
2810	0.2	Ip	$\upsilon_8+\upsilon_9$				

[a]Computed frequencies for the main and tri-deuterated [OD,ND_2] species of the Ip, IVn, and IIIp conformers. [b]All IR transitions in the 2800−3600 cm^{-1} frequency range with intensities larger than 0.1 km/mol for the Ip conformer and 0.5 km/mol for IIIp and IVn conformers are reported. [c]All IR transitions in the 2400−3000 cm^{-1} frequency range with intensities larger than 0.1 km/mol for the Ip conformer and 0.5 km/mol for IIIp and IVn conformers are reported.

probability of tunneling effects for ND_2 with respect to NH_2 is also expected. Thus, on the basis of our computational results, we suggest a reinvestigation of the IR spectrum in the 2800−2900 cm^{-1} region for trideuterated [OD,ND_2] glycine.

Thanks to the accurate and reliable experimental data available for the well characterized Ip conformer, our computational strategy has been validated, thus showing that it enables us to evaluate structural, thermodynamic, and

Table 7. Best-Estimated Vibrational Frequencies of Glycine in the Low-Frequency Region of the Raman/IR Spectra[a]

		NH_2-CH_2-COOH			ND_2-CH_2-COOD		
	mode	harm	anh	exptl[b]	harm	anh	exptl[c]
Ip/ttt	υ_{23}	200	203	204	157	161	
	υ_{22}	254	255	250	237	241	
	υ_{21}	467	461	458(463[c])	409	394	
	υ_{20}	508	494	498[c]	440	435	437
IIIp/tct	υ_{23}	238	246	237[d]	179	164	
	υ_{22}	263	269		243	245	
	υ_{21}	502	499		428	418	419
	υ_{20}	515	509		483	480	481
IVn/gtt	υ_{23}	183	156	171	133	122	
	υ_{22}	282	278	275	265	269	
	υ_{21}	469	467		414	397	
	υ_{20}	518	502		451	446	

[a]Computed frequencies (in cm^{-1}) for the main and trideuterated [OD,ND_2] species of the Ip, IVn, and IIIp conformers. [b]Experimental data from Raman spectrum of jet-cooled glycine from ref 34. [c]Experimental data from IR spectrum recorded in low-temperature Ar matrix from ref 32. [d]Experimental data from IR spectrum recorded in low-temperature Ar matrix from ref 91.

spectroscopic properties with an overall accuracy of about, or better than, 0.001 Å, 20 MHz, 1 kJ·mol^{-1}, and 10 cm^{-1} for bond distances, rotational constants, conformational enthalpies, and vibrational frequencies, respectively. In conclusion, the present work demonstrates that integrated experimental and computational studies are nowadays able to characterize medium-sized molecular systems of biological and/or technological interest with an accuracy so far reached only for very small rigid/semirigid molecules.

AUTHOR INFORMATION

Corresponding Author

*E-mail: julien.bloino@pi.iccom.cnr.it.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by Italian MIUR (PRIN and FIRB funds). The high performance computer facilities of the DREAMS center (http://dreamshpc.sns.it) are acknowledged for providing computer resources. The COST-CMTS Action CM1002 "COnvergent Distributed Environment for Computational Spectroscopy (CODECS)" is also acknowledged.

REFERENCES

(1) Eliel, E. L.; Allinger, N. L.; Angyal, S. J.; Morrison, G. A. *Conformational Analysis*; American Chemical Society: Washington, DC, 1965.

(2) Bythell, B. J.; Csonka, I. P.; Suhai, S.; Barofsky, D. F.; Paizs, B. Gas-Phase Structure and Fragmentation Pathways of Singly Protonated Peptides with N-Terminal Arginine. *J. Phys. Chem. B* **2010**, *114*, 15092−15105.

(3) Albrieux, F.; Calvo, F.; Chirot, F.; Vorobyev, A.; Tsybin, Y. O.; Lepere, V.; Antoine, R.; Lemoine, J.; Dugourd, P. Conformation of Polyalanine and Polyglycine Dications in the Gas Phase: Insight from Ion Mobility Spectrometry and Replica-Exchange Molecular Dynamics. *J. Phys. Chem. A* **2010**, *114*, 6888−6896.

(4) Chingin, K.; Balabin, R. M.; Barylyuk, K.; Chen, H.; Frankevich, V.; Zenobi, R. Rhodamines in the gas phase: cations, neutrals, anions,

and adducts with metal cations. *Phys. Chem. Chem. Phys.* **2010**, *12*, 11710−11714.
(5) Quack, M.; Merkt, F. *Handbook of High-Resolution Spectroscopy*; John Wiley & Sons: New York, 2011.
(6) Laane, J. *Frontiers of Molecular Spectroscopy*; Elsevier: Amsterdam, 2009.
(7) Zwier, T. S. Laser Probes of Conformational Isomerization in Flexible Molecules and Complexes. *J. Phys. Chem. A* **2006**, *110*, 4133−4150.
(8) de Vries, M. S.; Hobza, P. Gas-phase spectroscopy of biomolecular building blocks. *Annu. Rev. Phys. Chem.* **2007**, *58*, 585−612.
(9) Astrid, G.; Rudolf, R.; Jerker, W. *Single Molecule Spectroscopy in Chemistry, Physics and Biology: Nobel Symposium*; Springer-Verlag: Berlin, 2010.
(10) Barone, V. *Computational Strategies for Spectroscopy, from Small Molecules to Nano Systems.*; John Wiley & Sons, Inc.: Hoboken, NJ, 2011.
(11) Jensen, P.; Bunker, P. R. *Computational Molecular Spectroscopy*; John Wiley and Sons Ltd: Chichester, U. K., 2000.
(12) Puzzarini, C.; Stanton, J. S.; Gauss, J. Quantum-chemical calculation of spectroscopic parameters for rotational spectroscopy. *Int. Rev. Phys. Chem.* **2010**, *29*, 273−367.
(13) Grunenberg, J. *Computational Spectroscopy*; Wiley-VCH Verlag GmbH & Co. KGaA: New York, 2010.
(14) Barone, V.; Baiardi, A.; Biczysko, M.; Bloino, J.; Cappelli, C.; Lipparini, F. Implementation and validation of a multi-purpose virtual spectrometer for large systems in complex environments. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12404−12422.
(15) Heckert, M.; Kállay, M.; Gauss, J. Molecular equilibrium geometries based on coupled-cluster calculations including quadruple excitations. *Mol. Phys.* **2005**, *103*, 2109−2115.
(16) Heckert, M.; Kállay, M.; Tew, D. P.; Klopper, W.; Gauss, J. Basis-set extrapolation techniques for the accurate calculation of molecular equilibrium geometries using coupled-cluster theory. *J. Chem. Phys.* **2006**, *125*, 044108.
(17) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. HEAT: High accuracy extrapolated ab initio thermochemistry. *J. Chem. Phys.* **2004**, *121*, 11599−11613.
(18) Puzzarini, C. Accurate thermochemistry and spectroscopy of the oxygen-protonated sulfur dioxide isomers. *Phys. Chem. Chem. Phys.* **2011**, *13*, 21319−21327.
(19) Puzzarini, C.; Barone, V. Toward spectroscopic accuracy for open-shell systems: Molecular structure and hyperfine coupling constants of H_2CN, H_2CP, NH_2, and PH_2 as test cases. *J. Chem. Phys.* **2010**, *133*, 184301/1−11.
(20) Puzzarini, C.; Barone, V. Extending the molecular size in accurate quantum-chemical calculations: the equilibrium structure and spectroscopic properties of uracil. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7189−7197.
(21) Puzzarini, C.; Biczysko, M.; Barone, V. Accurate anharmonic vibrational frequencies for uracil: the performance of composite schemes and hybrid CC/DFT model. *J. Chem. Theory Comput.* **2011**, *7*, 3702−3710.
(22) Puzzarini, C.; Biczysko, M.; Barone, V. Accurate Harmonic/Anharmonic Vibrational Frequencies for Open-Shell Systems: Performances of the B3LYP/N07D Model for Semirigid Free Radicals Benchmarked by CCSD(T) Computations. *J. Chem. Theory Comput.* **2010**, *6*, 828−838.
(23) Biczysko, M.; Panek, P.; Scalmani, G.; Bloino, J.; Barone, V. Harmonic and Anharmonic Vibrational Frequency Calculations with the Double-Hybrid B2PLYP Method: Analytic Second Derivatives and Benchmark Studies. *J. Chem. Theory Comput.* **2010**, *6*, 2115−2125.
(24) Biczysko, M.; Bloino, J.; Brancato, G.; Cacelli, I.; Cappelli, C.; Ferretti, A.; Lami, A.; Monti, S.; Pedone, A.; Prampolini, G.; Puzzarini, C.; Santoro, F.; Trani, F.; Villani, G. Integrated computational approaches for spectroscopic studies of molecular systems in the gas phase and in solution: pyrimidine as a test case. *Theor. Chem. Acc.* **2012**, *131*, 1201/1−19.
(25) Cappelli, C.; Biczysko, M. In *Computational Strategies for Spectroscopy, from Small Molecules to Nano Systems*; Barone, V., Ed.; John Wiley & Sons, Inc.: New York, 2011; Chapter Time-Independent Approach to Vibrational Spectroscopies, pp 309−360.
(26) Gaw, F.; Willetts, A.; Handy, N.; Green, W. In *SPECTRO - A Program for Derivation of Spectroscopic Constants from Provided Quartic Force Fields and Cubic Dipole Fields*; Bowman, J. M., Ed.; JAI Press: Greenwich, CT, 1991; Vol. *1B*, pp 169−185.
(27) Vázquez, J.; Stanton, J. F. Simple(r) algebraic equation for transition moments of fundamental transitions in vibrational second-order perturbation theory. *Mol. Phys.* **2006**, *104*, 377−388.
(28) Bloino, J.; Guido, C.; Lipparini, F.; Barone, V. A fully automated implementation of VPT2 Infrared intensities. *Chem. Phys. Lett.* **2010**, *496*, 157−161.
(29) Bloino, J.; Barone, V. A second-order perturbation theory route to vibrational averages and transition properties of molecules: General formulation and application to infrared and vibrational circular dichroism spectroscopies. *J. Chem. Phys.* **2012**, *136*, 124108.
(30) Biczysko, M.; Bloino, J.; Carnimeo, I.; Panek, P.; Barone, V. Fully ab initio IR spectra for complex molecular systems from perturbative vibrational approaches: Glycine as a test case. *J. Mol. Spectrosc.* **2012**, *1009*, 74−82.
(31) Godfrey, P. D.; Brown, R. D. Shape of Glycine. *J. Am. Chem. Soc.* **1995**, *117*, 2019−2023.
(32) Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Rosado, M. T. S.; Duarte, M. L. T. S.; Fausto, R.; Adamowicz, L. Matrix-Isolation Infrared and Theoretical Studies of the Glycine Conformers. *J. Phys. Chem. A* **1998**, *102*, 1041−1054.
(33) Huisken, F.; Werhahn, O.; Ivanov, A. Y.; Krasnokutski, S. A. The O−H stretching vibrations of glycine trapped in rare gas matrices and helium clusters. *J. Chem. Phys.* **1999**, *111*, 2978−2984.
(34) Balabin, R. M. Conformational Equilibrium in Glycine: Experimental Jet-Cooled Raman Spectrum. *J. Phys. Chem. Lett.* **2010**, *1*, 20−23.
(35) Balabin, R. M. Experimental thermodynamics of free glycine conformations: the first Raman experiment after twenty years of calculations. *Phys. Chem. Chem. Phys.* **2012**, *14*, 99−103.
(36) Bazso, G.; Magyarfalvi, G.; Tarczay, G. Near-infrared laser induced conformational change and UV laser photolysis of glycine in low-temperature matrices: Observation of a short-lived conformer. *J. Mol. Struct.* **2012**, *1025*, 33−42.
(37) Bazso, G.; Magyarfalvi, G.; Tarczay, G. Tunneling Lifetime of the ttc/VIp Conformer of Glycine in Low-Temperature Matrices. *J. Phys. Chem. A* **2012**, *116*, 10539−10547.
(38) Kasalová, V.; Allen, W. D.; Schaefer, H. F., III; Czinki, E.; Császár, A. G. Molecular structures of the two most stable conformers of free glycine. *J. Comput. Chem.* **2007**, *28*, 1373−1383.
(39) Balabin, R. M. Conformational equilibrium in glycine: Focal-point analysis and ab initio limit. *Chem. Phys. Lett.* **2009**, *479*, 195−200.
(40) Pulay, P.; Meyer, W.; Boggs, J. E. Cubic force constants and equilibrium geometry of methane from Hartree−Fock and correlated wavefunctions. *J. Chem. Phys.* **1978**, *68*, 5077−5085.
(41) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618−622.
(42) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(43) Kendall, A.; Dunning, T. H., Jr.; Harrison, R. J. Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796−6806.
(44) Woon, D. E.; Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *J. Chem. Phys.* **1995**, *103*, 4572−4585.
(45) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479−483.

(46) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639−9646.
(47) Puzzarini, C. Extrapolation to the Complete Basis Set Limit of Structural Parameters: Comparison of Different Approaches. *J. Phys. Chem. A* **2009**, *113*, 14530−14535.
(48) Puzzarini, C.; Heckert, M.; Gauss, J. The accuracy of rotational constants predicted by high-level quantum-chemical calculations. I. molecules containing first-row atoms. *J. Chem. Phys.* **2008**, *128*, 194108.
(49) Demaison, J. Experimental, semi-experimental and ab initio equilibrium structures. *Mol. Phys.* **2007**, *105*, 3109−3138.
(50) Feller, D.; Peterson, K. A. Probing the limits of accuracy in electronic structure calculations: Is theory capable of results uniformly better than "chemical accuracy"? *J. Chem. Phys.* **2007**, *126*, 114105.
(51) Feller, D. The use of systematic sequences of wave functions for estimating the complete basis set, full configuration interaction limit in water. *J. Chem. Phys.* **1993**, *98*, 7059−7071.
(52) Nielsen, H. H. The Vibration-Rotation Energies of Molecules. *Rev. Mod. Phys.* **1951**, *23*, 90−136.
(53) Mills, I. M. In *Molecular Spectroscopy: Modern Research*; Rao, K. N., Mathews, C. W., Eds.; Academic: New York, 1972.
(54) Isaacson, A. D.; Truhlar, D. G.; Scanlon, K.; Overend, J. Tests of approximation schemes for vibrational energy levels and partition functions for triatomics: H_2O and SO_2. *J. Chem. Phys.* **1981**, *75*, 3017−3024.
(55) Amos, R. D.; Handy, N. C.; Green, W. H.; Jayatilaka, D.; Willets, A.; Palmieri, P. Anharmonic vibrational properties of CH_2F_2: A comparison of theory and experiment. *J. Chem. Phys.* **1991**, *95*, 8323−8336.
(56) Barone, V. Anharmonic vibrational properties by a fully automated second-order perturbative approach. *J. Chem. Phys.* **2005**, *122*, 014108/1−10.
(57) Double and triple-ζ basis sets of SNS and N07 families are available for download. Visit http://dreamslab.sns.it (accessed December 1, 2012).
(58) Watson, J. K. G. *Vibrational Spectra and Structure*; Elsevier: New York, 1977.
(59) Puzzarini, C. A theoretical investigation on the HCCS radical and its ions. *Chem. Phys.* **2008**, *346*, 45−52.
(60) Halkier, A.; Klopper, W.; Helgaker, T.; Jørgensen, P. Basis-set convergence of the molecular electric dipole moment. *J. Chem. Phys.* **1999**, *111*, 4424−4430.
(61) Gauss, J.; Stanton, J. F. Analytic CCSD(T) second derivatives. *Chem. Phys. Lett.* **1997**, *276*, 70−77.
(62) Tew, D. P.; Klopper, W.; Heckert, M.; Gauss, J. Basis Set Limit CCSD(T) Harmonic Vibrational Frequencies. *J. Phys. Chem. A* **2007**, *111*, 11242−11248.
(63) Puzzarini, C.; Barone, V. Toward spectroscopic accuracy for organic free radicals: Molecular structure, vibrational spectrum, and magnetic properties of F_2NO. *J. Chem. Phys.* **2008**, *129*, 084306/1−7.
(64) Puzzarini, C.; Barone, V. Assessment of a computational strategy approaching spectroscopic accuracy for structure, magnetic properties and vibrational frequencies of organic free radicals: the F_2CN and F_2BO case. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6991−6997.
(65) Carbonniere, P.; Lucca, T.; Pouchan, C.; Rega, N.; Barone, V. Vibrational computations beyond the harmonic approximation: performances of the B3LYP functional for semirigid molecules. *J. Comput. Chem.* **2005**, *26*, 384−388.
(66) Begue, D.; Carbonniere, P.; Pouchan, C. Calculations of Vibrational Energy Levels by Using a Hybrid ab Initio and DFT Quartic Force Field: Application to Acetonitrile. *J. Phys. Chem. A* **2005**, *109*, 4611−4616.
(67) Begue, D.; Benidar, A.; Pouchan, C. The vibrational spectra of vinylphosphine revisited: Infrared and theoretical studies from CCSD(T) and DFT anharmonic potential. *Chem. Phys. Lett.* **2006**, *430*, 215−220.
(68) Bloino, J.; Biczysko, M.; Barone, V. General Perturbative Approach for Spectroscopy, Thermodynamics, and Kinetics: Methodological Background and Benchmark Studies. *J. Chem. Theory Comput.* **2012**, *8*, 1015−1036.
(69) Becke, D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(70) Barone, V.; Cimino, P.; Stendardo, E. Development and Validation of the B3LYP/N07D Computational Model for Structural Parameter and Magnetic Tensors of Large Free Radicals. *J. Chem. Theory Comput.* **2008**, *4*, 751.
(71) Barone, V.; Cimino, P. Accurate and feasible computations of structural and magnetic properties of large free radicals: The PBE0/N07D model. *Chem. Phys. Lett.* **2008**, *454*, 139−143.
(72) Barone, V.; Cimino, P. Validation of the B3LYP/N07D and PBE0/N07D Computational Models for the Calculation of Electronic g-Tensors. *J. Chem. Theory Comput.* **2009**, *5*, 192−199.
(73) Barone, V.; Bloino, J.; Biczysko, M. Validation of the DFT/N07D computational model on the magnetic, vibrational and electronic properties of vinyl radical. *Phys. Chem. Chem. Phys.* **2010**, *12*, 1092−1101.
(74) Carnimeo, I.; Biczysko, M.; Bloino, J.; Barone, V. Reliable structural, thermodynamic, and spectroscopic properties of organic molecules adsorbed on silicon surfaces from computational modeling: the case of glycine@Si(100). *Phys. Chem. Chem. Phys.* **2011**, *13*, 16713−16727.
(75) Bloino, J.; Biczysko, M.; Santoro, F.; Barone, V. General Approach to Compute Vibrationally Resolved One-Photon Electronic Spectra. *J. Chem. Theory Comput.* **2010**, *6*, 1256−1274.
(76) Vázquez, J.; Stanton, J. F. Treatment of Fermi resonance effects on transition moments in vibrational perturbation theory. *Mol. Phys.* **2007**, *105*, 101−109.
(77) Barone, V. Vibrational zero-point energies and thermodynamic functions beyond the harmonic approximation. *J. Chem. Phys.* **2004**, *120*, 3059−3065.
(78) Barone, V. Characterization of the potential energy surface of the HO_2 molecular system by a density functional approach. *J. Chem. Phys.* **1994**, *101*, 10666−10676.
(79) Martin, J. M. L.; Lee, T. J.; Taylor, P. R.; Francois, J.-P. The anharmonic force field of ethylene, C_2H_4, by means of accurate ab initio calculations. *J. Chem. Phys.* **1995**, *103*, 2589−2602.
(80) Begue, D.; Pouchan, C.; Guillemin, J.-C.; Benidar, A. Anharmonic treatment of vibrational resonance polyads, the diborane: a critical case for numerical methods. *Theor. Chim. Acta* **2012**, *131*, 1122/1−11.
(81) Ayala, P. Y.; Schlegel, H. B. Identification and treatment of internal rotation in normal mode vibrational analysis. *J. Phys. Chem.* **1998**, *108*, 2314−2325.
(82) Truhlar, D. G.; Isaacson, A. D. Simple perturbation theory estimates of equilibrium constants from force fields. *J. Chem. Phys.* **1991**, *94*, 357−359.
(83) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G.; Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C.; Bernholdt, D. E.; Bomble, Y. J.; Christiansen, O.; Heckert, M.; Heun, O.; Huber, C.; Jagau, T.-C.; Jonsson, D.; J. Jusélius, Klein, K.; Lauderdale, W. J.; Matthews, D.; Metzroth, T.; Mueck, L. A.; D. P. O'Neill, Price, D. R.; Prochnow, E.; Puzzarini, C.; Ruud, K.; Schiffmann, F.; Schwalbach, W.; Stopkowicz, S.; Tajti, A.; J. Vázquez, Wang, F.; Watts, J. D. *CFour*; 2011. Integral packages: Almloef, J.; Taylor, P. R. *MOLECULE*; Taylor, P. R. *PROPS*; Helgaker, T.; Jensen, H. J. Aa.; Jørgensen, P.; Olsen, J. *ABACUS*; Mitin, A. V.; van Wuellen, C. *ECP routines*. For the current version, see http://www.cfour.de (accessed September 13, 2012).
(84) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. R.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. R., Jr., Peralta, J. A.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.;

Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, R.; Cammi, A. J.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision C.01.; Gaussian Inc.: Wallingford, CT, 2009.

(85) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. Glycine conformers: a never-ending story? *Phys. Chem. Chem. Phys.* **2013**, *15*, 1358−1363.

(86) Helgaker, T.; Jørgensen, P.; Olsen, J. *Electronic-Structure Theory*; Wiley: Chichester, U. K., 2000.

(87) Pawlowski, F.; Jørgensen, P.; Olsen, J.; Hegelund, F.; Helgaker, T.; Gauss, J.; Bak, K. L.; Stanton, J. F. Molecular equilibrium structures from experimental rotational constants and calculated vibration−rotation interaction constants. *J. Chem. Phys.* **2002**, *116*, 6482−6496.

(88) López, J. C.; Peña, M. I.; Sanz, M. E.; Alonso, J. L. Probing thymine with laser ablation molecular beam Fourier transform microwave spectroscopy. *J. Chem. Phys.* **2007**, *126*, 191103/1−4.

(89) Alonso, J. L.; Pérez, C.; Sanz, M. E.; López, J. C.; Blanco, S. Seven conformers of l-threonine in the gas phase: a LA-MB-FTMW study. *Phys. Chem. Chem. Phys.* **2009**, *11*, 617−627.

(90) Peña, M. I.; Sanz, M. E.; López, J. C.; Alonso, J. L. Preferred Conformers of Proteinogenic Glutamic Acid. *J. Am. Chem. Soc.* **2012**, *134*, 2305−2312.

(91) Ivanov, A.; Sheina, G.; Blagoi, Y. FTIR spectroscopic study of the UV-induced rotamerization of glycine in the low temperature matrices (Kr, Ar, Ne). *Spectrochim. Acta A* **1998**, *55*, 219−228.

(92) Barna, D.; Nagy, B.; Csontos, J.; Csaszar, A. G.; Tasi, G. Benchmarking Experimental and Computational Thermochemical Data: A Case Study of the Butane Conformers. *J. Chem. Theory Comput.* **2012**, *8*, 479−486.

(93) Gruzman, D.; Karton, A.; Martin, J. M. L. Performance of Ab Initio and Density Functional Methods for Conformational Equilibria of C_nH_{2n+2} Alkane Isomers (n = 4−8). *J. Phys. Chem. A* **2009**, *113*, 11974−11983.

(94) Puzzarini, C. A Theoretical Study of the CH_2N Isomers: Molecular Structure and Energetics. *Int. J. Quantum Chem.* **2010**, *110*, 2483−2494.

(95) Uddin, K. M.; Warburton, P. L.; Poirier, R. A. Comparisons of Computational and Experimental Thermochemical Properties of α-Amino Acids. *J. Phys. Chem. B* **2012**, *116*, 3220−3234.

(96) Feller, D. A.; Sordo, J. A. A CCSDT study of the effects of higher order correlation on spectroscopic constants. I. First row diatomic hydrides. *J. Chem. Phys.* **2000**, *112*, 5604−5610.

(97) Ruden, T. A.; Helgaker, T.; Jørgensen, P.; Olsen, J. Coupled-cluster connected quadruples and quintuples corrections to the harmonic vibrational frequencies and equilibrium bond distances of HF, N_2, F_2, and CO. *J. Chem. Phys.* **2004**, *121*, 5874−5884.

(98) Martin, J. M. L. Spectroscopic quality ab initio potential curves for CH, NH, OH and HF. A convergence study. *Chem. Phys. Lett.* **1998**, *292*, 411−420.

(99) Cortez, M. H.; Brinkmann, N. R.; Polik, W. F.; Taylor, P. R.; Bomble, Y. J.; Stanton, J. F. Factors Contributing to the Accuracy of Harmonic Force Field Calculations for Water. *J. Chem. Theory Comput.* **2007**, *3*, 1267−1274.

(100) Noga, J.; Bartlett, R. J. The full CCSDT model for molecular electronic structure. *J. Chem. Phys.* **1987**, *86*, 7041.

(101) Scuseria, G. E.; Schaefer, H. F., III. A new implementation of the full CCSDT model for molecular electronic-structure. *Chem. Phys. Lett.* **1988**, *152*, 382−386.

(102) Watts, J. D.; Bartlett, R. J. The coupled-cluster single, double, and triple excitation model for open-shell single reference functions. *J. Chem. Phys.* **1993**, *93*, 6104.

(103) Pawlowski, F.; Halkier, A.; Jørgensen, P.; Bak, K. L.; Helgaker, T.; Klopper, W. Accuracy of spectroscopic constants of diatomic molecules from ab initio calculations. *J. Chem. Phys.* **2003**, *118*, 2539−2549.

(104) Zhao, Y.; Schults, N. E.; Truhlar, D. G. Design of Density Functionals by Combining the Method of Constraint Satisfaction with Parametrization for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions. *J. Chem. Theory Comput.* **2006**, *2*, 364−382.

(105) Buchanan, E. G.; James, W. H., III; Choi, S. H.; Guo, L.; Gellman, S. H.; Muller, C. W.; Zwier, T. S. Single-conformation infrared spectra of model peptides in the amide I and amide II regions: Experiment-based determination of local mode frequencies and inter-mode coupling. *J. Chem. Phys.* **2012**, *137*, 094301.

(106) Dean, J. C.; Buchanan, E. G.; Zwier, T. S. Mixed 14/16 Helices in the Gas Phase: Conformation-Specific Spectroscopy of Z-(Gly)n, n = 1, 3, 5. *J. Am. Chem. Soc.* **2012**, *134*, 17186−17201.

(107) Pietraperzia, G.; Pasquini, M.; Schiccheri, N.; Piani, G.; Becucci, M.; Castellucci, E.; Biczysko, M.; Bloino, J.; Barone, V. The Gas Phase Anisole Dimer: A Combined High-Resolution Spectroscopy and Computational Study of a Stacked Molecular System. *J. Phys. Chem. A* **2009**, *113*, 14343−14351.

(108) Pietraperzia, G.; Pasquini, M.; Mazzoni, F.; Piani, G.; Becucci, M.; Biczysko, M.; Michalski, D.; Bloino, J.; Barone, V. Noncovalent Interactions in the Gas Phase: The Anisole-Phenol Complex. *J. Phys. Chem. A* **2011**, *115*, 9603−9611.

(109) Biczysko, M.; Panek, P.; Barone, V. Toward spectroscopic studies of biologically relevant systems: Vibrational spectrum of adenine as a test case for performances of long-range/dispersion corrected density functionals. *Chem. Phys. Lett.* **2009**, *475*, 105−110.

(110) Fabri, C.; Szidarovszky, T.; Magyarfalvi, G.; Tarczay, G. Gas-Phase and Ar-Matrix SQM Scaling Factors for Various DFT Functionals with Basis Sets Including Polarization and Diffuse Functions. *J. Phys. Chem. A* **2011**, *115*, 4640−4649.

(111) Barone, V.; Biczysko, M.; Bloino, J.; Borkowska-Panek, M.; Carnimeo, I.; Panek, P. Toward anharmonic computations of vibrational spectra for large molecular systems. *Int. J. Quantum Chem.* **2012**, *112*, 2185−2200.

(112) Baquero, E. E.; James, W. H.; Choi, S. H.; Gellman, S. H.; Zwier, T. S. Single-Conformation Ultraviolet and Infrared Spectroscopy of Model Synthetic Foldamers:β-Peptides Ac-β^3-hPhe-β^3-hAla-NHMe and Ac-β^3-hAla-β^3-hPhe-NHMe. *J. Am. Chem. Soc.* **2008**, *130*, 4795−4807.