Structural and Vibrational Properties of Amino Acids from Composite Schemes and Double-Hybrid DFT: Hydrogen Bonding in Serine as a Test Case

Structural and Vibrational Properties of Amino Acids from Composite Schemes and Double-hybrid DFT: Hydrogen Bonding in Serine as a Test Case

Mingzhu Sheng[1], Filippo Silvestrini[2], Malgorzata Biczysko[1*] and Cristina Puzzarini[2*]

[1]International Centre for Quantum and Molecular Structures, Physics Department, Shanghai University, 99 Shangda Road, Shanghai, 200444 China

[2]Department of Chemistry "Giacomo Ciamician", University of Bologna, Via F. Selmi 2, 40126 Bologna, Italy

*Corresponding authors: biczysko@shu.edu.cn, cristina.puzzarini@unibo.it

Abstract

The structures, relative stabilities, and vibrational wavenumbers of the two most stable conformers of serine, stabilized by the O-H···N, O-H···O=C and N-H···O-H intramolecular hydrogen bonds, have been evaluated by means of state-of-the-art composite schemes based on coupled-cluster (CC) theory. The so-called "cheap" composite approach (CCSD(T)/(CBS+CV)^{MP2}) allowed accurate equilibrium structures and harmonic vibrational wavenumbers, also pointing out significant corrections beyond the CCSD(T)/cc-pVTZ level. These accurate results stand as a reference for benchmarking selected hybrid and double-hybrid, dispersion corrected, DFT functionals. B2PLYP-D3 and DSDPBEP86 in conjunction with a triple-zeta basis set have been confirmed as effective methodologies for structural and spectroscopic studies of medium-sized flexible biomolecules, also showing intramolecular hydrogen bonding. These best performing double-hybrid functionals have been employed to simulate IR spectra by means of vibrational perturbation

theory, also considering hybrid CC/DFT schemes. The best overall agreement with experiment, with mean absolute error of 8 cm^{-1}, has been obtained by combining CCSD(T)/$(CBS+CV)^{MP2}$ harmonic wavenumbers with B2PLYP-D3/maug-cc-pVTZ anharmonic corrections. Finally, a composite scheme entirely based on CCSD(T) calculations (CCSD(T)/CBS+CV) has been employed for energetics, further confirming that serine II is the most stable conformer, also when zero-point vibrational energy corrections are included.

1. Introduction

Experimental [1][2][3][4][5][6][7][8][9][10] and theoretical [11][12][13][14][15][16][17][18][19] investigations of isolated amino acids (AAs), polypeptides and their analogues allow for gaining insight into their intrinsic properties without the perturbing effects of the environment. In turn, such information helps obtain a better understanding of the complicated processes at the basis of protein folding [20][21][22], develop and validate models for protein simulations [23][24][25][26] or refinement [27][28][29], or –as far as astrochemistry and astrobiology are concerned– shed light on the origin of life on Earth [30][31][32]. Focusing on the latest context, meteorites and comets have been found to contain a large variety of amino acids [33][34][35][36][37][38][39], and plausible routes for their formation in the circumstellar and interstellar environments have also been proposed [40][41][42][43][44][45][46][47][48].

In the investigation of AAs and their oligomers, one of the major challenges is their conformational flexibility, which –in combination with possible weak interactions– leads to a large number of possible three-dimensional geometries, the so-called conformers. For isolated AA and their small oligomers, spectroscopic techniques combined with theoretical analysis are proven to be powerful tools for their structural characterization. Examples are provided by rotational spectroscopy measurements in the centimeter-wave region [5][17][49][50][51][52][53][54], or indirect analysis of the ‘fingerprint’ vibrational features

[3][6][9][55][56][57][58][59][60][61]. Moreover, the combination of cryogenic matrix isolation techniques with UV-vis or NIR (near infrared) irradiation allows the detection of short-lived, highly energetic conformers, which are hardly observable in the gas phase [62][63][64]. However, all these experiments require computational supports that provide: (i) a preliminary investigation of the conformational space, (ii) the identification of the most stable conformers, (iii) the accurate characterization of their equilibrium structures and relative energetics, and finally (iv) the prediction of spectroscopic parameters and properties, also including vibrational corrections [16][65][66][67].

From the computational point of view, to achieve high accuracy, sophisticated composite schemes have been introduced [68][69][70][71][72][73][74]. These rely on the additivity approximation and, the most accurate variants, are based on coupled-cluster (CC) theory. Among them, the approach denoted as "CCSD(T)/CBS+CV" [74][75] is particular effective. This composite scheme is entirely based on the CC singles and doubles with a perturbative treatment of triples method, CCSD(T) [76][77], and incorporates the extrapolation to the complete basis set (CBS) limit and the effect of core-valence (CV) correlation. However, for medium- to large-sized molecular systems, such scheme is unaffordable because of the unfavorable scaling of high-level quantum-chemical (QC) methods. For AAs and their oligomers, in order to obtain good accuracy at an affordable computational cost, different computational strategies have been proposed [67][78][79][80][81][82]. Among them, the so-called "cheap" scheme [78][82], hereafter denoted as CCSD(T)/(CBS+CV)^{MP2}, evaluates the CBS and CV corrections using second-order Møller-Plesset perturbation theory (MP2) [83], and it has been demonstrated reliable, accurate and robust [82][84], even for rather flexible systems like glycine and glycine dipeptide analogues [16][17] as well as for weakly bonded molecular complexes [85]. While composite schemes are mostly exploited in geometry optimizations and energy evaluations [78][80][81][82], they can also be employed in the determination of spectroscopic properties, such as harmonic wavenumbers [86][87], CCSD(T)/(CBS+CV)^{MP2} being still affordable for medium-sized molecules [88].

Indeed, it has been employed, for instance, for dimethyloxirane [89], glycine [16][90][91][92], pyruvic acid [66], uracil [82][88] and thiouracil [93].

Accurate structural and spectroscopic parameters from composite schemes provide valuable reference data for benchmarking the accuracy of much cheaper QC methods. To deal with larger systems, methods based on density functional theory (DFT) are usually applied. For vibrational wavenumbers, it has been shown that the difference between high-level QC estimates and DFT computations is mainly related to the harmonic part [94][95][96][97][98], while some hybrid density functionals can provide accurate anharmonic terms [88][94][99][100]. The combination of harmonic force field at the CCSD(T) level or evaluated using a composite scheme with anharmonic terms at the DFT level leads to the so-called hybrid CC/DFT schemes [16][82]. However, for systems larger than 15 heavy atoms, the DFT needs to be also employed in the evaluation of the harmonic part, thus leading to the less expensive hybrid DHF/DFT approaches, where DHF stands for double-hybrid DFT functionals. An example is provided by the B2PLYP/B3LYP [101][102][103][104] scheme, which gives rather accurate results [105][106].

Moving to the subject of this work, serine (Ser, $C_3H_7NO_3$, 2-Amino-3-hydroxypropanoic acid) is an α-amino acid containing a -CH_2OH side chain, which extends the conformational flexibility and can lead to intramolecular hydrogen-bond interactions, thus requiring the accurate treatment of weak interactions [107][108]. The two most stable serine conformers (Ser I and Ser II), which have been experimentally characterized by both rotational [109] and infrared [64][110][111][112] spectroscopy, show structures governed by different types of hydrogen-bond interactions, namely, O-H···N, O-H···O=C and N-H···O-H. As far as the conformational analysis of neutral serine is concerned, the most recent systematic and complete theoretical investigation of the conformational space has been carried out in ref. [18] at the MP2/cc-pVTZ level, with a total of 85 unique conformers being identified. A more accurate characterization of structures, relative energies, and spectroscopic parameters has also been performed for the twelve lowest in energy conformers, which lie within 8 kJ/mol. However, only electronic energies have been

evaluated by means of a composite scheme based on CCSD(T) computations, while structural and spectroscopic parameters were determined at the MP2/aug-cc-pVTZ level.

Structural, spectroscopic and energetic computations at the DFT and MP2 levels have also been performed to support experimental analyses. Four conformers of neutral serine were first characterized by low temperature matrix isolation (MI) Fourier Transform Infrared (FTIR) measurements [111]. Subsequently, UV irradiation and annealing experiments demonstrated that nine different conformers are present in inert-gas matrices [110]. Seven conformers out of these nine have been observed by means of laser ablation Fourier transform microwave spectroscopy under jet isolation conditions [109]. More recently, serine conformations and their interconversions have been re-investigated by MI-IR spectroscopy combined with selective NIR-laser irradiation [64]. This study led to the identification of six conformers owing to the comparison of the experimental IR spectra with anharmonic computations [64]. Ser I and Ser II being the most stable conformers have been observed in all experiments mentioned above.

The present computational study aims at a more accurate determination of the hydrogen-bond characteristics that are the structural features ruling the various conformations. For this purpose, equilibrium structures and harmonic vibrational properties have been accurately determined, thereby exploiting the CCSD(T)/(CBS+CV)^{MP2} composite scheme. These best-estimated theoretical results have also been employed as a reference for benchmarking DFT approaches, thereby focusing on selected dispersion-corrected and double-hybrid DFT models. Moreover, the best-estimated harmonic wavenumbers have been combined with DFT anharmonic corrections, thus allowing the direct comparison with the experimental MI-IR results.

The manuscript is organized as follows. In the next section, a thorough account of the computational details is reported. In section 3, the results are summarized and presented in three subsections, which address (i) the best results for structural, energetic and spectroscopic properties, (ii) a benchmark study of DFT functionals,

and (iii) the best hybrid anharmonic force fields for serine. Finally, concluding remarks are provided.

2. Computational details

2.1 Composite schemes for structures, harmonic wavenumbers and energies

The equilibrium structures and harmonic force fields (in a normal mode representation), evaluated using analytical techniques [113], have been obtained within the CCSD(T)/(CBS+CV)^{MP2} composite scheme. As mentioned in the Introduction, the starting point is CCSD(T) in conjunction with the cc-pVTZ (VTZ) basis set [114][115]. This term is then corrected at the MP2 level in order to recover the basis-set truncation error, and to incorporate the effects of core-valence correlation and diffuse functions in the basis set. Overall, we have:

$$p(\text{best}) = p(\text{CCSD(T)/cc-pVTZ}) + \Delta p(CBS) + \Delta p(CV) + \Delta p(aug) \qquad (1)$$

where p denotes a generic structural parameter (bond length, angle or dihedral angle) or harmonic vibrational wavenumber. The second term of equation (1) is the correction to the CBS limit and exploits the n^{-3} formula [68] as follows:

$$\Delta p(\text{CBS}) = \frac{n^3 p(n) - (n-1)^3 p(n-1)}{p^3 - (p-1)^3} - p(n-1) \qquad (2)$$

where n=4, $p(n)$ thus corresponding to MP2/cc-pVQZ and $p(n-1)$ to MP2/cc-pVTZ. The correction due to the CV correlation (the third term of equation (1)) is computed in conjunction with the cc-pCVTZ [116] basis set, and defined as:

$$\Delta p(\text{CV}) = p(cc-pCVTZ, \text{all}) - p(cc-pCVTZ, \text{fc}) \qquad (3)$$

with *p*(cc-pCVTZ,all) and *p*(cc-pCVTZ,fc) being the MP2 values obtained correlating all electrons and within the frozen core (fc) approximation, respectively. The last term of equation (1), i.e. the correction due to the inclusion of diffuse functions in the basis set (Δ*p*(aug)), is introduced on an empirical basis to recover the limitations affecting extrapolation procedures carried out with small- to medium-sized basis sets [93][117] and is evaluated as

$$\Delta p(\text{aug}) = p(AVTZ) - p(VTZ) \tag{4}$$

where *p*(AVTZ) and *p*(VTZ) are the generic structural parameter or harmonic vibrational wavenumber computed at the MP2 level within the fc approximation using the aug-cc-pVTZ (AVTZ) [118] and cc-pVTZ [114] basis sets, respectively.

To apply the CCSD(T)/(CBS+CV)^{MP2} composite scheme, for both Ser I and Ser II, geometry optimizations and harmonic force-field calculations have been performed at different levels of theory: fc-CCSD(T)/cc-pVTZ, fc-MP2/cc-pVTZ, fc-MP2/cc-pVQZ, fc-MP2/cc-pCVTZ, all-MP2/cc-pCVTZ, and fc-MP2/cc-aug-pVTZ. For each conformer, from the CCSD(T)/(CBS+CV)^{MP2} harmonic force-field calculations the zero-point vibrational energy (ZPE) correction is derived as half the sum of all harmonic wavenumbers.

For the electronic energy, the more accurate CCSD(T)/CBS+CV composite scheme has been employed. Within this approach, the extrapolation to the CBS limit is carried out in two steps, as follows:

$$E_{elec}^{CCSD(T)/CBS} = E_{\infty}^{HF-SCF} + \Delta E_{\infty}^{corr} \tag{5}$$

where E_{∞}^{HF-SCF} is the Hartree-Fock (HF-SCF) electronic energy extrapolated to the CBS limit according to the e^{-Cn} exponential formula [119] :

$$E^{SCF}(n) = E_{\infty}^{SCF} + B' e^{-C'n} \tag{6}$$

using the HF/cc-pV*n*Z energies, with *n* = T, Q, and 5. The CCSD(T) correlation energy is extrapolated to the CBS limit by means of the n^{-3} formula [68]:

$$\Delta n^{corr}(n) = \Delta E_{\infty}^{corr} + A' n^{-3} \tag{7}$$

employing the fc-CCSD(T)/cc-pVTZ and fc-CCSD(T)/cc-pVQZ correlation energies. Best estimates for electronic energies have then been obtained by incorporating the

CV correlation correction:

$$\Delta E(\mathrm{CV}) = E[CCSD(T)/pCVTZ, all] - E[CCSD(T)/pCVTZ, fc] \quad (8)$$

where the cc-pCVTZ basis set has been employed. Overall, for both serine conformers considered, the CCSD(T)/CBS+CV energies are given by:

$$E_{elec}^{CBS+CV} = E_{\infty}^{HF-SCF} + \Delta E_{\infty}^{corr} + \ \Delta E(\mathrm{CV}) \quad (9)$$

with the single-point energy calculations being performed on top of the CCSD(T)/(CBS+CV)^{MP2} equilibrium structures.

All CCSD(T) computations have been performed with the CFOUR code [120], while the Gaussian package [121] has been employed for the MP2 computations.

2.2 DFT, Anharmonic and Non-Covalent Interaction Computations

Different dispersion-corrected DFT models [122] have been considered for structure and harmonic force-field computations. These include the hybrid B3LYP-D3 [101][102] (in tables and figures denoted as B3-D3), MN15 [123] and double-hybrid B2PLYP-D3 [103][104] (shortly denoted as B2-D3), mPW2PLYPD [124] (shortly denoted as mPW2D), DSD-PBEP86 [125] (shortly denoted as DSD), PBE0-DH [126] and PBE-QIDH [127] functionals. These have been selected because of the availability of analytical second derivatives of energy and dipole moments. For all functionals, full geometry optimizations followed by evaluation of harmonic wavenumbers have been performed in conjunction with the double-ζ SNSD [105][128] and triple-ζ maug-cc-pVTZ [129] (mAVTZ; a modified version of aug-cc-pVTZ) basis sets. All DFT calculations incorporate the dispersion correction [130]: Grimme's D3 formulation [131] combined with the Becke-Johnson (BJ) damping function [132] has been used in conjunction with B3LYP, B2PLYP and DSD-PBEP86, while the D2 version has been employed for mPW2PLYPD. Specific functional-dependent formulations have instead been used for MN15 [123], PBE0-DH [126], and PBEQIDH [127].

Preliminary benchmark studies allowed for selecting the best performing models for anharmonic computations; these are: B2PLYP-D3, DSDPBEP86, and

mPW2PLYPD, all of them in conjunction with the maug-cc-pVTZ basis set. Subsequently, the DFT cubic and semi-diagonal quartic force constants have been combined with the CCSD(T)/(CBS+CV)^{MP2} quadratic force constants (harmonic wavenumbers), thus leading to hybrid CC/DFT approaches. The correspondence between the normal modes description of different QC models was checked by the visual inspection of molecular vibrations and the computations of "Duschinsky-like" matrices [133][134][135]. Finally, second-order vibrational perturbation theory (VPT2) [136][137], within the generalized GVPT2 model [138][139][140][141], has been applied to the hybrid and DFT anharmonic force fields to evaluate anharmonic zero-point vibrational energy, wavenumbers and IR intensities. Standard criteria for anharmonic resonances have been employed, as suggested in reference [134].

The DFT and VPT2 computations have been performed with the Gaussian package [121].

The analysis of non-covalent interactions has been performed with the NCI method [142] based on the electron density and its derivatives, using the Multiwfn code [143]. The electron densities have been computed at the B2PLYP-D3/maug-cc-pVTZ level on top of the best-estimated equilibrium structures. The isosurfaces of reduced density gradient (RDG) have been visualized using the VMD package [144].

3. Results and discussion.

The two most stable serine conformers (Ser I and Ser II) are shown in Figure 1 along with the atoms labeling and the schematic representation of the intramolecular hydrogen bonds (HBs). The non-covalent interactions stabilizing these two structures are depicted in Figure 2.

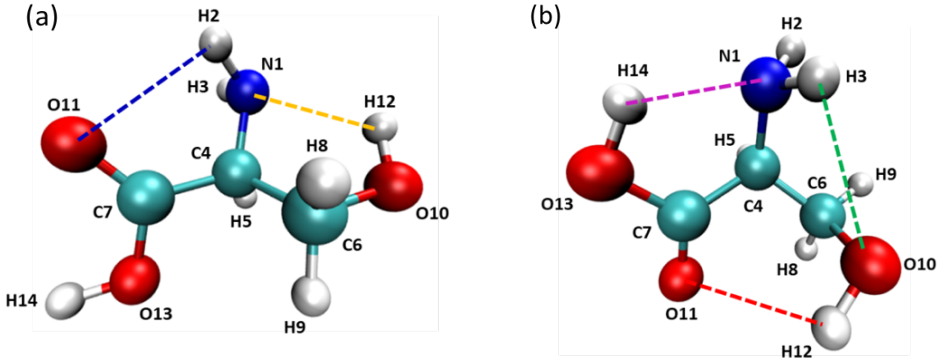


Figure 1. Structures of Ser I (a) and Ser II (b) with the atoms labelling and the hydrogen bonds marked as colored dashed lines.

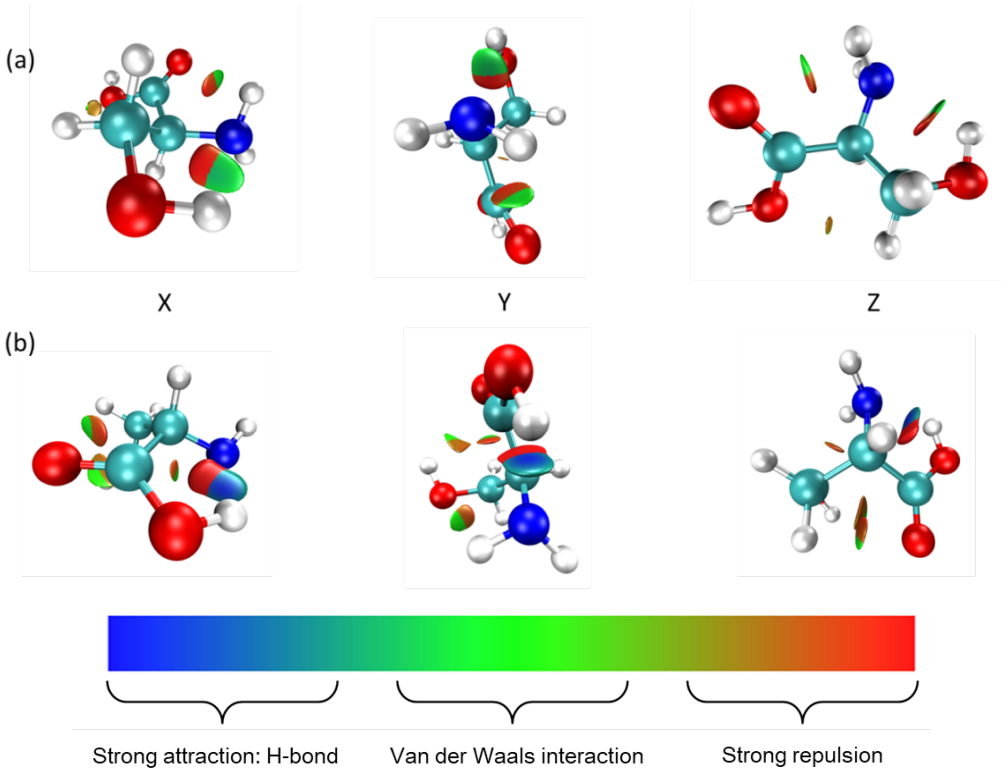


Figure 2. RDG isosurface at 0.5 a.u. for Ser I (a) and Ser II (b) in the different axis orientations, together with a schematic description of interactions.

From Figures 1 and 2, it is evident that Ser I is stabilized by two HBs, namely N1-H2···O11 and O10-H12···N1, but the NCI analysis reveals an additional weak non-covalent interaction, i.e. C6-H9···O13. Ser II is the only serine conformer that has all three hydrogen donor groups involved in formal HBs with three different acceptors [18]: N1-H3···O10, O10-H12···O11 and O13-H14···N1. Among them, the

NCI calculations points out that the latter is strongest interaction (OH group acting as donor and NH_2 as acceptor).

In the following analysis, we will focus on the influence of these interactions on the magnitude and importance of specific contributions to the best-estimated structural and spectroscopic parameters.

3.1 Best-Estimated Determinations

3.1.1 Best-Estimated Equilibrium Structural Parameters

The best-estimated CCSD(T)/(CBS+CV)^{MP2} equilibrium structural parameters p in a Z-Matrix representation are listed in Table S1 of the Supporting Information (SI), with the corresponding Cartesian coordinates being reported in Table S2. The various contributions to the composite scheme on bond lengths, angles and dihedrals are analyzed in Table 1. These are: (i) the basis set effects Δp(BSE), with two terms Δp(CBS), the extrapolation to the CBS limit, and Δp(aug), the effect of diffuse functions; (ii) the core-valence corrections Δp(CV); (iii) the effect of triple excitations Δp(T) (where Δp(T) = p(CCSD(T)/cc-pVTZ) – p(MP2/cc-pVTZ)).

Table 1. Contributions[a] to equilibrium structural parameters for Ser I and Ser II.

		Δp(CBS)	Δp(aug)	Δp(BSE)	Δp(CV)	Δp(T)
				Ser I		
	MAX	-	0.0030	0.0001	-	0.0070
Bonds / Å	MIN	-0.0072	-0.0010	-0.0082	-0.0029	-0.0014
	MAE	0.0032	0.0011	0.0023	0.0017	0.0027
	MAX	1.11	0.71	1.82	0.14	0.35
Angles / deg.	MIN	-0.06	-0.16	-0.22	-0.03	-0.30
	MAE	0.35	0.25	0.60	0.05	0.16
	MAX	1.88	1.20	3.08	0.28	1.08
Dihedrals / deg.	MIN	-1.58	-1.16	-2.24	-0.25	-0.39
	MAE	0.85	0.52	1.31	0.10	0.41
				Ser II		

	MAX	-	0.0027	0.0000	-	0.0058
Bonds / Å	MIN	-0.0066	-0.0005	-0.0071	-0.0028	-0.0042
	MAE	0.0032	0.0010	0.0023	0.0017	0.0028
	MAX	0.97	0.60	1.57	0.15	0.70
Angles / deg.	MIN	-0.08	-0.03	-0.10	-0.02	-0.48
	MAE	0.35	0.20	0.55	0.05	0.19
	MAX	2.87	2.30	5.17	0.39	1.63
Dihedrals / deg.	MIN	-2.11	-1.87	-3.99	-0.30	-1.68
	MAE	1.31	1.11	2.41	0.13	0.97

a. The largest positive (MAX), the largest negative (MIN), and the mean absolute (MAE) contributions to the CCSD(T)/(CBS+CV)^{MP2} best estimates.

From Tables 1 and S1 we note that the extrapolation to the CBS limit leads in all cases to shorter bond lengths by, on average, 0.003 Å (1 Å = 10^{-10} m). For both conformers, the largest effect is about -0.007 Å for the N1-C4 bond and the smallest -0.001 Å for the C4-H5 bond. Among the bond lengths involving hydrogens, the largest CBS contribution (about -0.0035 Å) are observed for the O-H distances involved in the O-H···N HBs, which are O10-H12···N1 for Ser I and O13-H14···N1 for Ser II. At variance, both N-H bonds show the same correction (about -0.0026 Å) in spite of the fact that in both conformers only one is involved in HB. The Δp(aug) corrections are always positive, with the only exception of the N1-C4 bond, and of the order of 0.001 Å (except for the C6-O10 bond, for which the correction is +0.003 Å). On the whole, the two basis set contributions being of opposite signs partially counterbalance one with the other, thus leading to bonds shorter, on average, by 0.0023 Å. Only for the N1-C4, with N1 being the proton acceptor in the N-H···O HB, they sum-up with a final shortening by about 0.008 Å for Ser I (0.007 Å for Ser II). The CV correlation corrections are always negative, yet smaller than the CBS term, with their values ranging between -0.0007 Å and -0.003 Å (-0.0017 Å on average). As expected, they are larger for bonds between heavy atoms, and -0.0015 Å at most for distances involving hydrogens. The triple excitation effects are either positive or negative, with largest positive contributions noted for the C-C and C-N bonds. The most significant negative correction is for O13-H14 in Ser II, which is involved in the strongest HB. Interestingly, CV corrections at the MP2/cc-pCVTZ level are very close

to those evaluated using the same basis set and the CCSD(T) method [16][89].

Moving to angles, the largest CBS terms are experienced by the angles involving the N-H or O-H groups, which increase by about 0.5 to 1 deg. The latter correction is noted for the C-O-H angles that are involved in HBs. The diffuse function corrections for these C-O-H angles are also positive, thus leading to a total basis set effect of about +1.5 deg. In general Δp(CBS) and Δp(aug) sum up and lead to angles larger, on average, by 0.6 deg, with the only exception being the C4-C7-O13 angle of Ser I, and the C4-C7-O11 and C4-C6-H angles of Ser II. For them, both contributions are negative, but small: about -0.1 deg or even less. The Δp(CV) corrections are much smaller, indeed ranging from -0.03 deg to 0.15 deg. The Δp(T) term can be either positive or negative, and on average on the order of ~0.2 deg. The largest positive corrections are noted for the C-O-H angles where the OH group is involved in the N-H···O HB, the correction being up to 0.7 deg. For angles involving NH_2 groups, the Δp(T) effect is negative, the largest correction being observed for the C-N-H^{free} angles.

For dihedrals, in all cases both negative and positive contributions are observed, which are also significantly larger than those for bond angles. Basis set effects are the largest terms, with Δp(CBS) ranging from +2.9 to -2.1 deg and Δp(aug) between +2.3 and -1.9 deg. For each dihedral angle, these two contributions always have the same sign, and thus they sum up. As a result, for Ser I the basis set corrections are between +3.1 and -2.2 deg, the average correction being 1.3 deg. For Ser II, Δp(BSE) corrections are about twice larger than those of Ser I, indeed ranging between about +5.2 and -4.0 deg, the average being 2.5 deg. Such a difference can be explained by Ser II being strongly affected by weak intramolecular interactions, which thus lead to a more “compact/folded” form than the “open” one of Ser I. Differently, Δp(CV) contributions are almost negligible and range between 0.4 deg and -0.3 deg, the average correction being 0.1 deg. The triple excitation effects are again twice larger for Ser II than for Ser I, they range between +1.6 and -1.7 deg for Ser II and between 1.1 and -0.4 deg for Ser I, the average value being about 1 and 0.4 deg, respectively.

In summary, the extrapolation to the CBS limit gives the largest contribution to

all structural parameters, CV correction is the smallest contribution for angles and dihedrals, while -for bond lengths- it provides contributions similar in magnitude to the Δp(aug) term. Moreover, the detailed analysis of all contributions points out that the parameters involved in hydrogen bonds are those most significantly affected by the corrections beyond the CCSD(T)/cc-pVTZ level. We can therefore conclude that the differences in the magnitude of the corrections between Ser I and Ser II arise from the different HB patterns.

Due to the size of the molecule, equilibrium structure determinations using composite schemes entirely based on CC theory are not feasible for serine. However, the accuracy of CCSD(T)/(CBS+CV)^{MP2} structures can be assessed from the comparison with experiment based on rotational constants [109]. Indeed, these spectroscopic parameters strongly depend on the molecular structures [67]. The computed equilibrium rotational constants can be straightforwardly obtained from the equilibrium geometry and are listed in Table 2, along with the semi-experimental equilibrium constants, which have been obtained from the experimental rotational constants (which correspond to the vibrational ground state) by subtracting the vibrational corrections reported in the reference [18] (obtained using VPT2 computations at the MP2/6-31G* level). From the inspection of Table 2 we note that the CCSD(T)/cc-pVTZ level of theory represents a suitable starting point, with an average deviation from experiment of 0.4-0.6%. However, adding the extrapolation to the CBS limit and the CV correction improves this agreement, thus reducing the discrepancy to 0.2-0.3%. This result further validates the accuracy of the equilibrium structures at the CCSD(T)/(CBS+CV)^{MP2} level. This also shows a better agreement with experiment than the previously reported most advanced theoretical results at the MP2/aug-cc-pVTZ level [18].

Table 2. Equilibrium theoretical and semi-experimental rotational constants (in MHz).

	CCSD(T)[a]	Best[b]	MP2[c][18]	Experiment[d]
			Ser I	
A_e	4507.28	4499.37	4498	4525.03
B_e	1825.82	1841.28	1833	1841.16
C_e	1452.52	1459.98	1464	1458.80

			Ser II	
A_e	3569.51	3560.24	3552	3584.20
B_e	2414.62	2410.19	2417	2415.37
C_e	1745.35	1757.62	1761	1757.92

a. fc-CCSD(T)/cc-pVTZ, this work
b. CCSD(T)/(CBS+CV)^{MP2}, this work
c. MP2/aug-cc-pVTZ from reference [18]
d. Experimental (ground-state) rotational constants [109] corrected for computed vibrational corrections at the MP2/6-31G* level [18]

3.1.2 Best-Estimated Electronic Energies

The best theoretical electronic energies obtained using the CCSD(T)/CBS+CV composite scheme, as described in the Methodology section, are collected in Table 3 together with the specific contributions. It is noted that the two serine conformers are very close in energy, with the conformer II being more stable by 2 kJ/mol. Incorporation of the ZPE corrections, which is about 1.5 kJ/mol larger for the Ser II, at either harmonic or anharmonic level, reduces the energy difference to ~0.5 kJ/mol, without altering the stability order. Interestingly, at the HF-SCF level, the Ser I conformer is more stable than Ser II by about 5.6 kJ/mol, but the incorporation of the correlation effects (~7.5 kJ/mol) inverts the stability order of the two conformers. As expected, the CV corrections that are so important for accurate structural determinations provide an almost negligible contribution. Focusing on single computations, the CCSD(T)/cc-pVTZ level overestimates the energy difference by more than 0.5 kJ/mol, which means a deviation of about 30% with respect to CCSD(T)/CBS+CV. The disagreement reduces to 0.26 kJ/mol when moving to CCSD(T) computations in conjunction with the quadruple-ζ basis set. The CCSD(T)/CBS+CV conformational energies from this work are in excellent agreement with those computed using a similar approach (2.03 kJ/mol), but obtained using smaller basis sets for the CCSD(T) terms and MP2/aug-cc-pVTZ optimized structures [18]. However, the effect of molecular structure on conformational energy is very limited, as demonstrated -for example- in references [16][66].

Table 3. Electronic energies of the Ser I and Ser II conformers.

	Ser I [a.u][a]	Ser II [a.u]	Ser I[b] [kJ/mol]
E_{∞}^{SCF}	-396.920154	-396.918037	-5.56
CCSD(T)/cc-pVTZ	-398.386468	-398.387478	2.65
CCSD(T)/cc-pVQZ	-398.511911	-398.512770	2.26
ΔE_{∞}^{corr}	-1.671367	-1.674221	7.49
CCSD(T)/CBS	-398.591521	398.592258	1.93
$\Delta E(CV)$	-0.363226	-0.363250	0.06
CCSD(T)/CBS+CV	-398.954747	-398.955508	2.00
Harm ZPE correction[c]	0.113692	0.114271	-1.52
Anharm ZPE correction[d]	0.112156	0.112737	-1.53
CCSD(T)/CBS+CV+HarmZPE	-398.841055	-398.841237	0.48
CCSD(T)/CBS+CV+AnharmZPE	-398.842591	-398.842771	0.47

a. Atomic units, 1 a.u = $4.3597447222071(85)\times10^{-18}$ J

b. Relative energy of Ser I with respect to Ser II

c. Harmonic ZPE correction at the CCSD(T)/$(CBS+CV)^{MP2}$ level

d. Anharmonic ZPE correction computed at the GVPT2 level using the hybrid CC/DFT force field obtained by combining CCSD(T)/$(CBS+CV)^{MP2}$ harmonic wavenumbers with B2PLYP-D3/maug-cc-pVTZ anharmonic corrections.

3.1.3 Best-Estimated Harmonic Wavenumbers

In addition to the structure, the main novelty of this work is the derivation of the harmonic vibrational wavenumbers at the CCSD(T)/$(CBS+CV)^{MP2}$ level, which -to the best of our knowledge- has not yet been employed for a similar task in the case of molecular systems with size and complexity comparable to that of serine. For smaller molecules such as glycine and pyruvic acid, the availability of CCSD(T)/$(CBS+CV)^{MP2}$ harmonic wavenumbers was crucial for the prediction of infrared spectra to very high accuracy [16][66].

For selected normal modes, Table 4 reports the CCSD(T)/cc-pVTZ and best-estimated harmonic wavenumbers together with the various contributions to the latter, namely the p(CBS) and Δp(aug) terms together with their sum Δp(BSE), the Δp(CV) contribution, and finally the Δp(T) term. For these selected modes, the comparison with the results at the MP2/aug-cc-pVTZ level [18], which were the most accurate data prior to this study, is also given. Additionally, for both conformers, the statistical analysis based on all normal modes is also provided. The whole set of vibrational wavenumbers is reported in the SI (Table S3). For the normal modes description, we have followed previous works where a detailed potential energy

distribution analysis was performed [18].

From the inspection of Tables 4 and S3, it is noted that the BSE term has the largest effect on the final wavenumbers, with the Δp(CBS) and Δp(aug) contributions either summing-up or cancelling out. The largest negative corrections are observed for the $\tau(\mathrm{O-H})_{sc}$ mode, and are about -50 cm^{-1} for the free O-H in Ser I and about -30 cm^{-1} for hydrogen-bonded O-H in Ser II. For the $\tau(\mathrm{O\text{-}H})_{bb}$ mode, these corrections are instead only about -10 cm^{-1} for both conformers. The next largest correction, nearly -30 cm^{-1}, is noted for the C=O stretching of both conformers, with Δp(CBS) and Δp(aug) contributions being about -10 and -20 cm^{-1}, respectively. The BSE has also a significant effect on all δ(COH)'s, the largest correction (about -25 cm^{-1}) affecting the vibrations involved in the O-H···N HB, namely $\delta(\mathrm{COH})_{sc}$ for Ser I and $\delta(\mathrm{COH})_{bb}$ for Ser II. For the O-H and N-H stretching vibrations, the two BSE contributions are opposite in sign. For all ν(OH) modes, the Δp(aug) term leads to final negative BSE corrections, which range between -1 and -15 cm^{-1}, the largest contribution being that for the free $\nu(\mathrm{O\text{-}H})_{bb}$ stretching of Ser I. For $\nu(NH_2)$, the two BSE corrections are of the same magnitude and cancel out. Considering all vibrations, the average basis set correction is about 10 cm^{-1}, and in most cases it is negative (a few positive corrections are negligible). Much smaller and always positive are the Δp(CV) corrections, with the average value being 3 cm^{-1} and with all corrections being smaller than 8 cm^{-1}. As far as the triple excitation effects are concerned, the largest corrections are again negative, and as large as nearly -40, -50 cm^{-1} for the C-H stretching vibrations of both conformers. Interestingly, large negative corrections are also observed for both the N-H stretchings of Ser I and the ν_{asym}(N-H) stretch of Ser II. At variance, large (+30cm^{-1}) and smaller (+10 cm^{-1}) blue-shifts are observed for the ν_{sym}(N-H) and $\nu(\mathrm{O\text{-}H})_{bb}$ stretching modes of Ser II. These latter vibrations are related to the strongest HB interaction, where the nitrogen atom of the NH_2 group acts as a donor, also introducing a coupling between the ν_{sym}(N-H) and $\nu(\mathrm{O\text{-}H})_{bb}$ vibrations. Further relevant corrections are for $\tau(\mathrm{O\text{-}H})_{sc}$ of Ser I and $\nu(\mathrm{N\text{-}C_\alpha})$ of Ser II, and they are again negative. For all the other vibrations, the Δp(T) term can be either positive or negative, but -in any case- smaller than 10 cm^{-1} in absolute terms.

Overall, the corrections on top of CCSD(T)/cc-pVTZ are not negligible, indeed being ~10 cm^{-1} on average for both Ser I and Ser II. For most vibrations, the Δp(BSE) contributions are the largest ones and negative, these being only partially balanced by the smaller and positive Δp(CV) corrections. Considering all vibrational modes, the corrections vary both in sign and magnitude, and therefore they cannot be recovered by using, for instance, scaling factors. The differences with the MP2/aug-cc-pVTZ harmonic values from reference [18] can be attributed to CBS, CV and triple excitation corrections; overall, they range from +38 cm^{-1} to -64 cm^{-1} and are ~15 cm^{-1} on average.

Table 4. Harmonic vibrational wavenumbers (in cm^{-1}) for selected modes of Ser I and Ser II.

Mode	Assignment[a]	Best	MP2[b]	CCSD(T)[c]	Δp(CBS)	Δp(aug)	Δp(BSE)	Δp(CV)	Δp(T)
					Ser I				
1	$\nu(O\text{-}H)_{bb}$	3757	3760	3767	4.4	-19.7	-15.4	5.4	6.3
2	$\nu(O\text{-}H)_{sc}$	3753	3737	3749	12.5	-13.8	-1.3	4.9	12.3
3	$\nu_{asym}(NH_2)$	3591	3619	3581	11.5	-8.4	3.1	7.6	-38.6
4	$\nu_{sym}(NH_2)$	3503	3521	3495	9.3	-9.3	0.1	7.1	-25.7
5	$\nu_{asym}(C_\beta H_2)$	3125	3161	3116	5.5	-2.8	2.6	5.9	-44.9
6	$\nu(C_\alpha\text{-}H)$	3088	3133	3092	-3.1	-7.1	-10.2	5.9	-40.8
7	$\nu_{sym}(C_\beta H_2)$	2997	3023	2981	8.8	1.6	10.4	5.4	-41.8
8	$\nu(C{=}O)$	1798	1815	1822	-11.0	-17.8	-28.8	5.2	6.5
11	$\delta(COH)_{sc}$	1432	1448	1455	-9.8	-14.8	-24.5	1.7	6.5
14	$\delta(COH)_{bb}$	1347	1357	1363	-8.0	-10.5	-18.5	2.7	5.2
18	$\nu(C\text{-}O)_{bb}$	1176	1182	1178	-0.8	-3.6	-4.3	3.2	-4.3
19	$\nu(N\text{-}C_\alpha)$	1141	1143	1145	-1.8	-5.7	-7.5	3.3	2.3
20	$\nu(C\text{-}O)_{sc}$	1095	1108	1104	-2.9	-10.6	-13.5	3.9	-3.6
24	$\delta_{wag}(NH_2)$	849	867	864	-7.5	-9.5	-17.1	1.7	-3.5
27	$\tau(O\text{-}H)_{bb}$	590	605	595	-0.1	-6.4	-6.6	1.4	-9.5
28	$\tau(O\text{-}H)_{sc}$	524	588	572	-24.3	-25.7	-50.0	2.0	-15.2
					Ser II				
1	$\nu(O\text{-}H)_{sc}$	3809	3825	3817	6.1	-20	-13.8	6.0	-7.8
2	$\nu_{asym}(NH_2)$	3598	3627	3588	10.3	-7.7	2.6	7.4	-38.3
3	$\nu_{syn}(NH_2)$	3555	3517	3547	9.4	-8.5	0.9	6.8	30.4
4	$\nu(O\text{-}H)_{bb}$	3480	3476	3488	8.2	-18.9	-10.7	3.0	11.6
5	$\nu_{asym}(C_\beta H_2)$	3122	3159	3111	7.4	-2.6	4.8	6.0	-48.1
6	$\nu_{sym}(C_\beta H_2)$	3065	3103	3065	-0.9	-5.8	-6.7	6.0	-38.1
7	$\nu(C_\alpha\text{-}H)$	3055	3093	3054	0.5	-5.3	-4.7	6.0	-39.6
8	$\nu(C{=}O)$	1813	1833	1839	-11.6	-20.2	-31.7	5.3	6.6
11	$\delta(COH)_{bb}$	1424	1445	1449	-12.7	-14.2	-26.9	2.3	3.5
16	$\delta(COH)_{sc}$	1246	1251	1250	-2.2	-4.6	-6.8	3.0	-1.1
17	$\nu(C\text{-}O)_{bb}$	1224	1237	1234	-5.4	-8.6	-14.0	3.5	-2.9
21	$\nu(C\text{-}O)_{sc}$	997	1006	1003	-2.2	-6.3	-8.4	2.8	-3.1
22	$\delta_{wag}(NH_2)$	941	951	953	-7.5	-5.0	-12.5	0.5	1.7
23	$\tau(O\text{-}H)_{bb}$	882	902	893	-6.3	-7.4	-13.7	2.8	-8.9
24	$\nu(N\text{-}C_\alpha)$	843	885	863	-11.4	-10.5	-21.9	1.9	-21.7
30	$\tau(O\text{-}H)_{sc}$	425	450	456	-13.4	-16.9	-30.2	0.1	5.7

	ΔMP2[d]	ΔCCSD(T)[d]	Δp(CBS)	Δp(aug)	Δp(BSE)	Δp(CV)	Δp(T)
MAX[e]	38	16	13	2	10	8	30
MIN[e]	-64	-48	-24	-26	-50	0	-48
MAE[e]	15	9	6	7	11	3	9

a. The symbols ν, δ, and τ denote the stretching, valence angle bending, and torsional motion, respectively; "bb" means backbone; "sc" side-chain; "sym" and "asym" stands for symmetric and antisymmetric vibrations, respectively

b. MP2/aug-cc-pVTZ from reference [18]

c. CCSD(T)/cc-pVTZ, this work

d. Differences with respect to the best-estimated harmonic wavenumbers

e. All normal modes of both conformers have been considered: largest positive (MAX), largest negative (MIN), and mean absolute (MAE) deviations.

3.2 DFT Benchmark

3.2.1 DFT Equilibrium Structures

The best-estimated equilibrium structures stand as a reference for less expensive methodologies based on DFT. CCSD(T)/cc-pVTZ and MP2/(aug)-cc-pVTZ results are also included for comparison purposes. The overall accuracy of the equilibrium geometries is analyzed in terms of internal parameters, such as bond lengths, angles and dihedral angles for which the maximum and mean absolute errors are reported in Figure 3 (all DFT parameters are listed in Table S4) Moreover, since the final three-dimensional structure is also defined by the weak interactions, the donor-acceptor distances (D···A) for all five HBs are compared with the corresponding best estimates in Figure 4.

For the CCSD(T)/cc-pVTZ level, absolute maximum and average errors are large (~0.01 Å and 0.004 Å, 2 deg and 0.6 deg, 5.5 and 2 deg, for bond lengths, bond angles and dihedral angles, respectively) and mainly attributed to the lack of diffuse functions in the basis set. The improved accuracy of MP2 bond lengths with both basis sets (maximum error of about 0.007-0.008 Å and the average of about 0.0025-0.003 Å with cc-pVTZ and aug-cc-pVTZ) may be due to error cancellation. For bond angles, the overall large errors are presumably associated with the lack of triple excitations. MP2/cc-pVTZ results show rather large deviations, with absolute maximum errors of 2.2 deg and mean errors of 0.6 deg, with inclusion of diffuse functions only slightly improving the results. The difference between MP2/cc-pVTZ and MP2/aug-cc-pVTZ values is more pronounced for dihedral angles. For the former, the absolute maximum and mean errors are 4.0 deg and 1.8 deg, respectively, while adding diffuse functions almost halves these values.

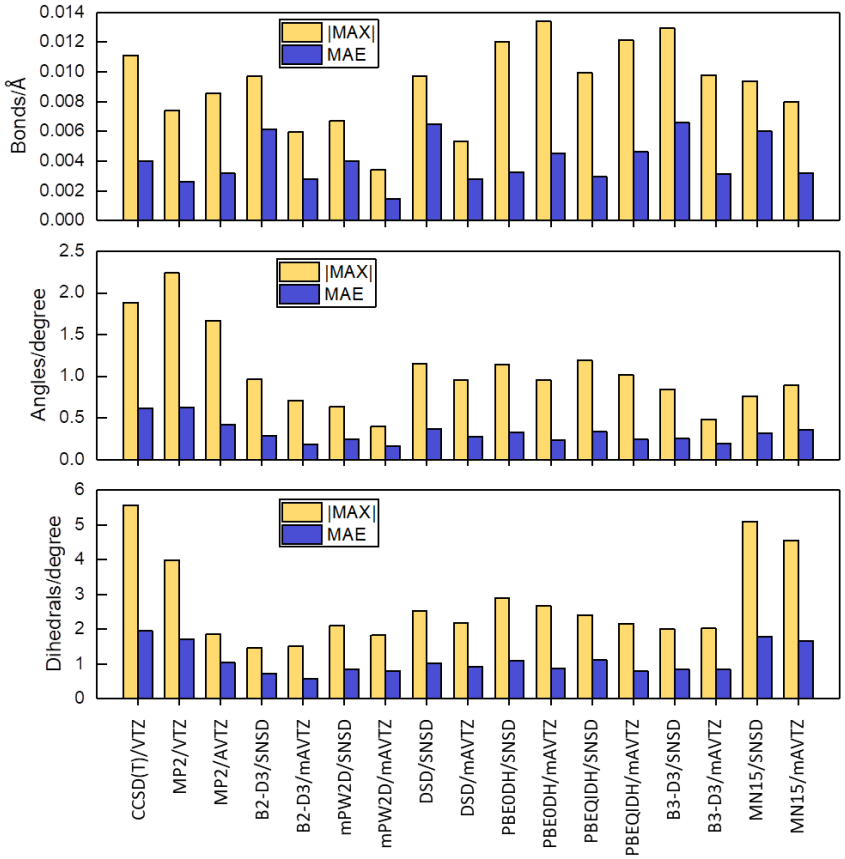


Figure 3. Maximum (|MAX|) and mean (MAE) absolute deviations for equilibrium bond lengths (in Å), angles and dihedral angles (in degrees) computed at different levels of theory. Data are reported as differences with respect to the CCSD(T)/(CBS+CV)^{MP2} best estimates.

Interestingly, B2PLYP-D3, mPW2PLYPD and DSDPBEP86, in conjunction with maug-cc-pVTZ basis set, yield bond lengths in very close agreement with the reference structure, with absolute maximum and mean errors below 0.006 Å and 0.003 Å, respectively. From an inspection of DFT results, it is evident that all computations performed with a triple-ζ basis set show maximum and mean errors of 0.003-0.008 Å and 0.003 Å, respectively, while using the double-ζ basis set nearly doubles the average errors. Interestingly, the PBE0DH and PBEQIDH functionals show the opposite trend, but they clearly perform worse than the other considered DFT models. For bond angles, all DFT functionals (with either double- or triple-ζ basis sets) show a better agreement with the CCSD(T)/(CBS+CV)^{MP2} level than MP2 (maximum errors below 1.2 deg and MAE below 0.4 deg). Among them, B2PLYP-D3

and mPW2PLYPD when used in conjunction with the maug-cc-pVTZ basis set yield the best results (|MAX| of 0.7 deg and 0.4 deg, and MAE 0.18 deg and 0.15 deg, respectively). For dihedral angles, the largest errors (MAX ranging from 3.5 to 5 deg and MAE from 1.4 to 2 deg) are found for MN15 and double-hybrid functionals from the PBE family. The mean absolute errors of the best-performing methods, i.e. B2PLYP-D3/maug-cc-pVTZ, mPW2PLYPD/maug-cc-pVTZ and DSDPBEP86/maug-cc-pVTZ, are below 1 deg for both Ser I and Ser II. Overall, we can conclude that for dihedral angles, B2PLYP-D3, mPW2PLYPD and DSDPBEP86 perform better than CCSD(T) and MP2 when used in combination with limited basis sets, and that inclusion of diffuse functions and triple excitations are important for correlated methods.

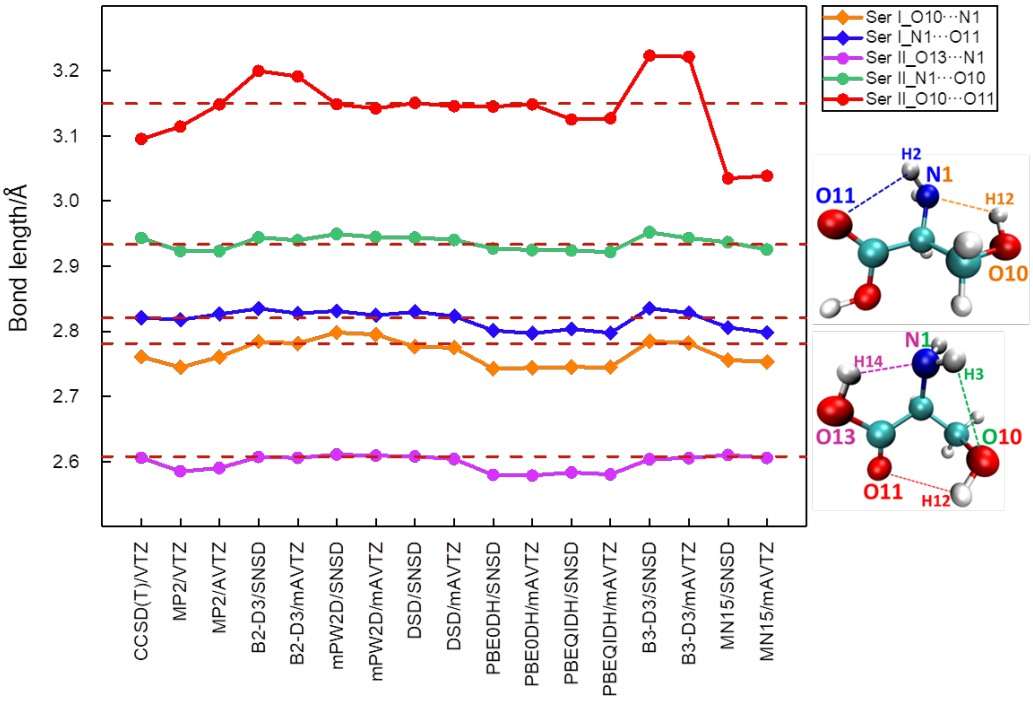


Figure 4. HB donor-acceptor distances calculated using different methodologies. Data are reported as deviations with respect to the CCSD(T)/(CBS+CV)^{MP2} best estimates (marked as dashed lines).

An inspection of the structural parameters defining the weak interactions further confirms the general findings discussed above. However, it can be also observed that

the accuracy of the D···A distances varies for different hydrogen bonds. In all four cases involving both N and O atoms, the error variation over all methods is relatively small, with differences within 0.04 Å, 0.02 Å, 0.03 Å and 0.02 Å for the O10···N1 (Ser I), N1···O11 (Ser I), O13···N1 (Ser II), and N1···O10 (Ser II) HBs, respectively. On the contrary, for O10···O11 (Ser II), much larger deviations can be observed; in particular, the MN15 functional combined with both basis sets yields distances shorter by more than 0.1 Å and B3LYP-D3 longer by 0.07 Å with respect to the reference value. Other DFT models are within ±0.05 Å from the best estimate. The O10···O11 (Ser II) distance shorter by about 0.05 Å than the best-estimated value, is also observed for the CCSD(T) and MP2 methods in conjunction with the cc-pVTZ basis set. In line with the NCI analysis, the O10···O11 HB is characterized by the longest donor-acceptor distance, thus suggesting the weakest hydrogen-bond interaction.

Considering all structural parameters, we conclude that both the basis set incompleteness and lack of triple excitations clearly affects the results of wave-function-based methodologies. Indeed, computations at the CCSD(T)/cc-pVTZ and MP2/(aug)-cc-pVTZ levels show noticeable deviations from the best estimates. Furthermore, it has been further demonstrated that the double-hybrid functionals combined with augmented triple-ζ basis sets provide rather accurate structural parameters at a greatly reduced computational cost. From the present analysis, it is evident that the B2PLYP-D3, mPW2PLYPD and DSDPBEP86 functionals in conjunction with the maug-cc-pVTZ basis set perform very well for equilibrium structure determinations.

Finally, a comment regarding the maug-cc-pVTZ basis set is deserved. Recently, an extensive structural study of the glycine conformers pointed out problems in the correct characterization of the IIIp/tct structure for computations using the maug-cc-pVTZ basis set [90]. This failure has been related to the flat energy profile along the NH_2 out-of-plane bending vibration, but a similar issue has not been encountered in this work for the studied serine conformers. Our conclusion is that, despite promising results, this basis set cannot be further recommended for flexible molecules.

3.2.2 DFT Conformational Energies

The conformational energies computed at different levels of theory (evaluated on top of the corresponding optimized equilibrium structures) are graphically reported in Figure 5, where the reference CCSD(T)/CBS+CV value is also made evident. In Figure 5, the absolute deviations from the latter are also displayed. From this figure, we first of all note that all DFT functionals predict the correct order of the serine conformers, Ser II being more stable than Ser I. While the B2PLYP-D3/maug-cc-pVTZ and mPW2PLYPD/maug-cc-pVTZ levels of theory underestimate the energy difference between the two conformers by about 1 kJ/mol, the MN15 functional (with both basis sets) overestimates it by more than 2 kJ/mol. Even if overestimating the population of Ser I, B2PLYP-D3/maug-cc-pVTZ and mPW2PLYPD/maug-cc-pVTZ correctly predict that both conformers are sufficiently stable to be observed in experiments; on the contrary, MN15/SNSD and MN15/maug-cc-pVTZ wrongly suggest instead that Ser I should be excluded from the analysis. Except for these two extreme cases, the other methods predict reliable electronic energies, with differences with respect to the best estimate within 0.5 kJ/mol.

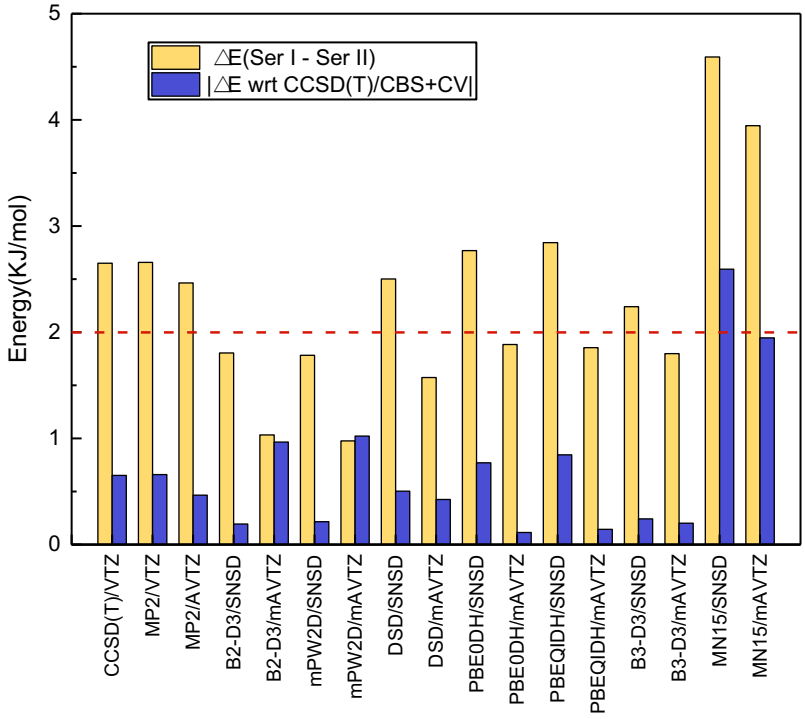


Figure 5. Conformational energy difference between the two conformers, ΔE(SerI-SerII), and absolute error, |ΔE| (absolute deviation with respect to CCSD(T)/CBS+CV) calculated at different levels of theory. The reference CCSD(T)/CBS+CV value is reported as dashed line.

3.2.3 DFT Harmonic Wavenumbers

The accuracy of DFT harmonic wavenumbers (the complete list can be found in Table S5) is assessed from their comparison with the CCSD(T)/(CBS+CV)^{MP2} best estimates. All functionals perform in a very similar way for both conformers, with the overall mean and maximum absolute deviations being graphically reported in Figure 6. The largest discrepancies are mainly positive for the Ser I conformer, while they are either positive or negative for Ser II, as shown in Table S5. Furthermore, the largest deviations are related to the O-H and N-H stretching modes.

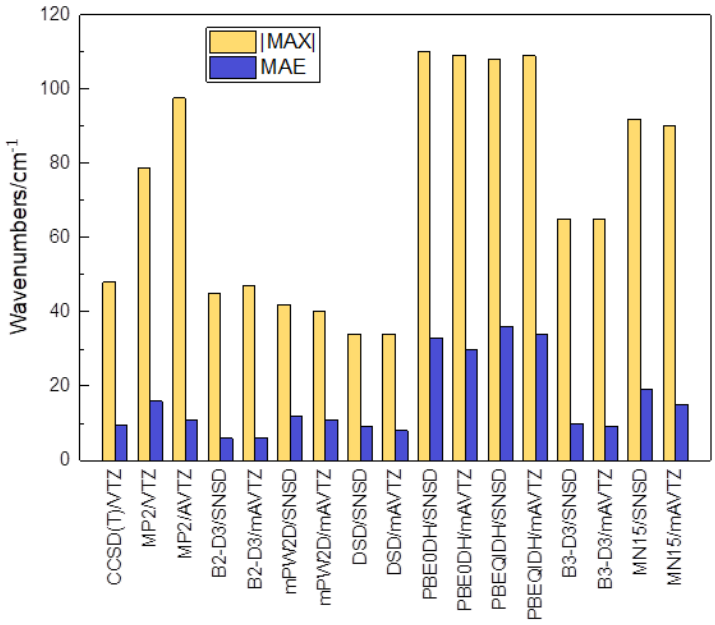


Figure 6. Harmonic vibrational wavenumbers (in cm^{-1}) of both conformers computed at different levels of theory: maximum (|MAX|) and mean (MAE) absolute deviations with respect to the $CCSD(T)/(CBS+CV)^{MP2}$ best estimates.

In analogy to molecular structures, DFT computations employing the double-hybrid B2PLYP-D3 and DSDPBEP86 functionals perform very well, better than or similarly to CCSD(T)/cc-pVTZ, with the mean absolute error for both conformers being smaller than 10 cm^{-1}, and the maximum discrepancies well within 50 cm^{-1}. The maximum deviations are negative and associated to the symmetric $\nu(NH_2)$ vibration of Ser II. A MAE of about 10 cm^{-1} is also found for mPW2PLYPD and B3LYP-D3, but for the former, the overall good performance is contaminated by negative discrepancies up to about -70 cm^{-1} associated to the ν(O-H) and ν(N-H) vibrations. The MN15 functional shows the largest basis-set effect, with MAE of about 19 and 15 cm^{-1} for the SNSD and maug-cc-pVTZ, respectively, and large (either positive or negative) errors up to 60 cm^{-1}. Significantly larger errors, with mean and maximum greater than 30 cm^{-1} and 100 cm^{-1}, are found for both the PBE0DH and PBEQIDH functionals.

Previous studies on glycine [16] and pyruvic acid [66] pointed out the problem of accurately describing the O-H stretching vibrations, in particular when involved in HB interactions. It is thus interesting to analyse this aspect in more detail for both serine conformers. Figure 7 depicts the harmonic vibrational wavenumbers of the four OH stretching vibrations in comparison with the best estimates. Among DFT functionals, B2PLYP-D3 (10 cm^{-1}) and DSDPBEP86 (20 cm^{-1}) provide the O-H stretching values that are in the best agreement with the reference data. Larger discrepancies (up to 40 cm^{-1}) are instead observed for mPW2PLYPD. The PBE0DH, PBEQIDH and MN15 (with both basis sets) functionals yield wavenumbers significantly higher than the reference values, with deviations that can be as large as 110 cm^{-1}, 110 cm^{-1} and 60 cm^{-1}, respectively.

Since -as expected from the literature data [16][66][106]- the B2PLYP-D3 and DSDPBEP86 functionals show the best performance, they have been selected for the anharmonic computations. The mPW2PLYPD has been also considered because of its overall good performance.

Focusing on the ZPE correction, errors exceeding 500 cm^{-1} are found for both PBE0DH and PBEQIDH functional, with the best prediction being provided by B2PLYP-D3 (errors below 80 cm^{-1}, i.e. ~1 kJ/mol). A similar good accuracy is noted for the MN15 functional and is likely due to errors compensation. Nevertheless, all methodologies provide ZPE differences between two conformers that well agree with the best estimate within ~25cm^{-1}. Therefore, on general grounds, the use of DFT functionals for the evaluation of the ZPE correction is not expected to introduce significant errors in the evaluation of conformational energies.

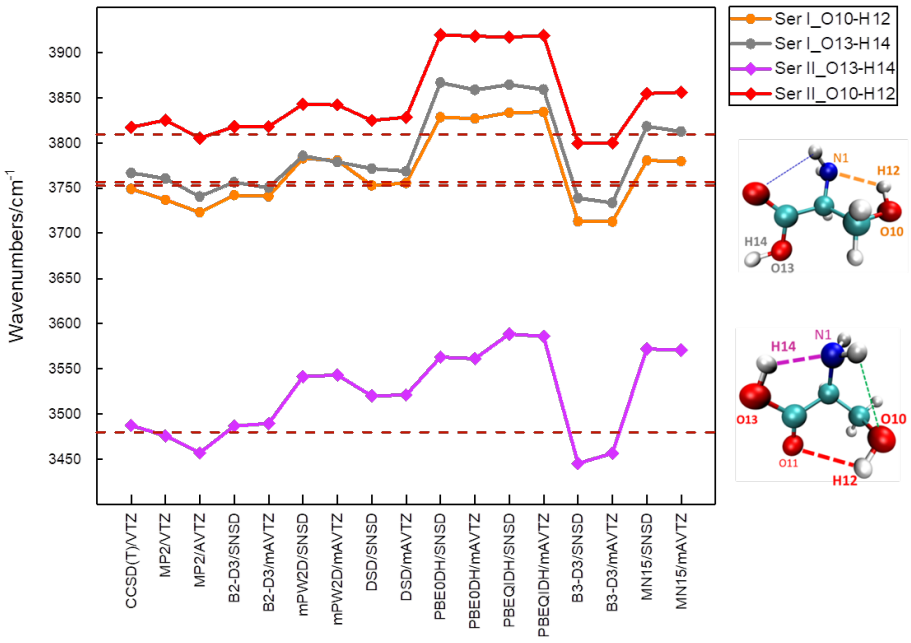


Figure 7. O-H stretching harmonic wavenumbers (in cm^{-1}) calculated at different levels of theory. CCSD(T)/(CBS+CV)^{MP2} best estimates are marked as dashed lines.

3.3 DFT and Hybrid CC/DFT Anharmonic Force Field

Assessing the accuracy of computed wavenumbers by comparison with experimental results [64][110][111] requires going beyond the harmonic approximation. Table 5 collects the computed DFT and hybrid CC/DFT anharmonic wavenumbers for selected normal modes, along with the available experimental results and the most accurate computations available prior to this work [18]. The latter employed MP2/aug-cc-pVTZ harmonic wavenumbers combined with anharmonic corrections at the MP2/6-31G* level. Furthermore, Table 5 reports the statistical analysis performed using all normal modes for both conformers. The whole set of vibrational wavenumbers is provided in the SI, Table S6.

For the Ser I conformer, a very good agreement is obtained for all DFT functionals, with errors halved in comparison to the previous theoretical predictions at the MP2 level. Moreover, inclusion of CCSD(T)/(CBS+CV)^{MP2} harmonic wavenumbers in the hybrid CC/DFT model further improves accuracy, indeed leading

to MAE values of about 6 cm^{-1} and maximum deviations within 13 cm^{-1} for anharmonic corrections at the B2PLYP-D3/maug-cc-pVTZ level. This good agreement further supports the original assignment of spectra reported in reference [64]. The situation is instead more involved for the Ser II conformer, with all computations predicting wavenumbers blue shifted by 50-120 cm^{-1} for the $\nu(O\text{-}H)_{bb}$ vibration involved in the strongest HB, even if best-estimated harmonic wavenumbers are used. These discrepancies are attributed to the too simplified description of the potential energy curve along the proton stretching coordinate [145], which would require going beyond both the quartic expansion of the potential and the VPT2 treatment, thus exploiting -for instance- variational solutions of the nuclear Schrodinger equation [146][147]. If one excludes this wavenumber from the statistical analysis, the best overall agreement with experiment is again obtained by combining $CCSD(T)/(CBS+CV)^{MP2}$ harmonic wavenumbers with B2PLYP-D3/maug-cc-pVTZ anharmonic corrections (CC/B2-D3) within GVPT2 computations, with 8 cm^{-1} and 23 cm^{-1} being the mean and maximum absolute errors, respectively. Furthermore, DFT computations, with MAE and MAX of 10 cm^{-1} and 36 cm^{-1}, respectively, show an improvement with respect the MP2 results, which show MAE and MAX values of 16 cm^{-1} and 52 cm^{-1}, respectively.

In the case of serine, a direct vis-à-vis comparison of experimental and computed IR spectra for only the two most stable conformers is not feasible because of the concomitant presence of other conformers in the experimental mixture [64]. Still, anharmonic computations allow the prediction of the spectra of each conformer, as shown in Figure 8 for CC/B2-D3 computations. It can be observed that the spectra patterns of Ser I and Ser II are quite different in terms of both band positions and band intensities, thus they can be easily distinguished by infrared features.

Table 5. Anharmonic vibrational wavenumbers (in cm^{-1}) for selected modes of Ser I and Ser II.

Mode	Assignment	B2PLYP-D3	CC/B2-D3	mPW2PLYPD	CC/mPW2D	DSDPBEP86	CC/DSD	MP2[18]	Experiment[64][110][111]
					Ser I				
1	$\nu(O\text{-}H)_{bb}$	3565	3571	3603	3577	3579	3566	3574	3579.9
2	$\nu(O\text{-}H)_{sc}$	3551	3565	3588	3560	3567	3564	3546	3561.1
3	$\nu_{asym}(NH_2)$	3427	3420	3447	3421	3436	3425	3442	3414.2
4	$\nu_{sym}(NH_2)$	3359	3349	3378	3350	3363	3351	3392	3344.0
5	$\nu_{asym}(C_\beta H_2)$	3009	2990	3010	2991	3001	2992	3016	2983.0
6	$\nu(C_\alpha\text{-}H)$	2969	2956	2977	2954	2970	2955	2991	2945.0
7	$\nu_{sym}(C_\beta H_2)$	2864	2853	2874	2862	2863	2853	2877	2865.0
8	$\nu(C{=}O)$	1759	1767	1781	1765	1775	1764	1787	1773.0
11	$\delta(COH)_{sc}$	1407	1419	1419	1421	1412	1418	1411	1409.7
14	$\delta(COH)_{bb}$	1323	1319	1332	1318	1327	1316	1348	1328.1
18	$\nu(C\text{-}O)_{bb}$	1136	1144	1147	1147	1149	1146	1151	1140.9
19	$\nu(N\text{-}C_\alpha)$	1094	1106	1104	1104	1106	1103	1105	1105.2
20	$\nu(C\text{-}O)_{sc}$	1057	1065	1072	1065	1072	1062	1076	1065.7
24	$\delta_{wag}(NH_2)$	819	814	824	809	828	813	848	820.2
	|MAX|[a]	25.5	13.9	34.1	17.9	25.2	13.8	51.7	
	MAE[a]	9.2	6.8	13.2	7.1	8.1	7.3	17.8	
					Ser II				
1	$\nu(O\text{-}H)_{sc}$	3638	3629	3664	3631	3648	3629	3644	3642
2	$\nu_{asym}(NH_2)$	3434	3432	3454	3427	3440	3429	3450	3417
3	$\nu_{sym}(NH_2)$	3370	3395	3385	3362	3380	3378	3375	3404
4	$\nu(O\text{-}H)_{bb}$	3223	3231	3287	3278	3237	3250	3224	3190
5	$\nu_{asym}(C_\beta H_2)$	2985	2978	3010	2976	2995	2978	3013	2983
6	$\nu_{sym}(C_\beta H_2)$	2986	2973	2993	2978	2987	2974	2969	2950
7	$\nu(C_\alpha\text{-}H)$	2935	2924	2915	2925	2915	2924	2940	
8	$\nu(C{=}O)$	1786	1784	1801	1783	1798	1782	1804	1785
11	$\delta(COH)_{bb}$	1398	1416	1411	1419	1398	1411	1409	1400
16	$\delta(COH)_{sc}$	1209	1212	1217	1211	1215	1211	1215	1211
17	$\nu(C\text{-}O)_{bb}$	1177	1181	1190	1183	1188	1179	1195	1190
21	$\nu(C\text{-}O)_{sc}$	970	973	978	973	978	972	976	974
22	$\delta_{wag}(NH_2)$	912	911	917	909	915	911	923	914
23	$\tau(O\text{-}H)_{bb}$	879	875	877	865	884	870	876	873
24	$\nu(N\text{-}C_\alpha)$	838	831	852	845	853	843	840	846
	|MAX|[a]	36.1	41.8	97.1	88.5	47.5	60.5	38.9	
	MAE[a]	10.5	10.2	18.2	13.8	13.3	12.1	15.6	
					Ser I & Ser II[b]				
	|MAX|[a]	36.1	22.7	43.4	42.4	36.5	32.0	51.7	
	MAE[a]	9.4	7.9	14.1	8.9	10.0	8.7	16.3	

[a] Maximum |MAX| and mean (MAE) absolute deviations with respect to experiment.

[b] Statistical analysis excluding $\nu(O\text{-}H)_{bb}$ of Ser II.

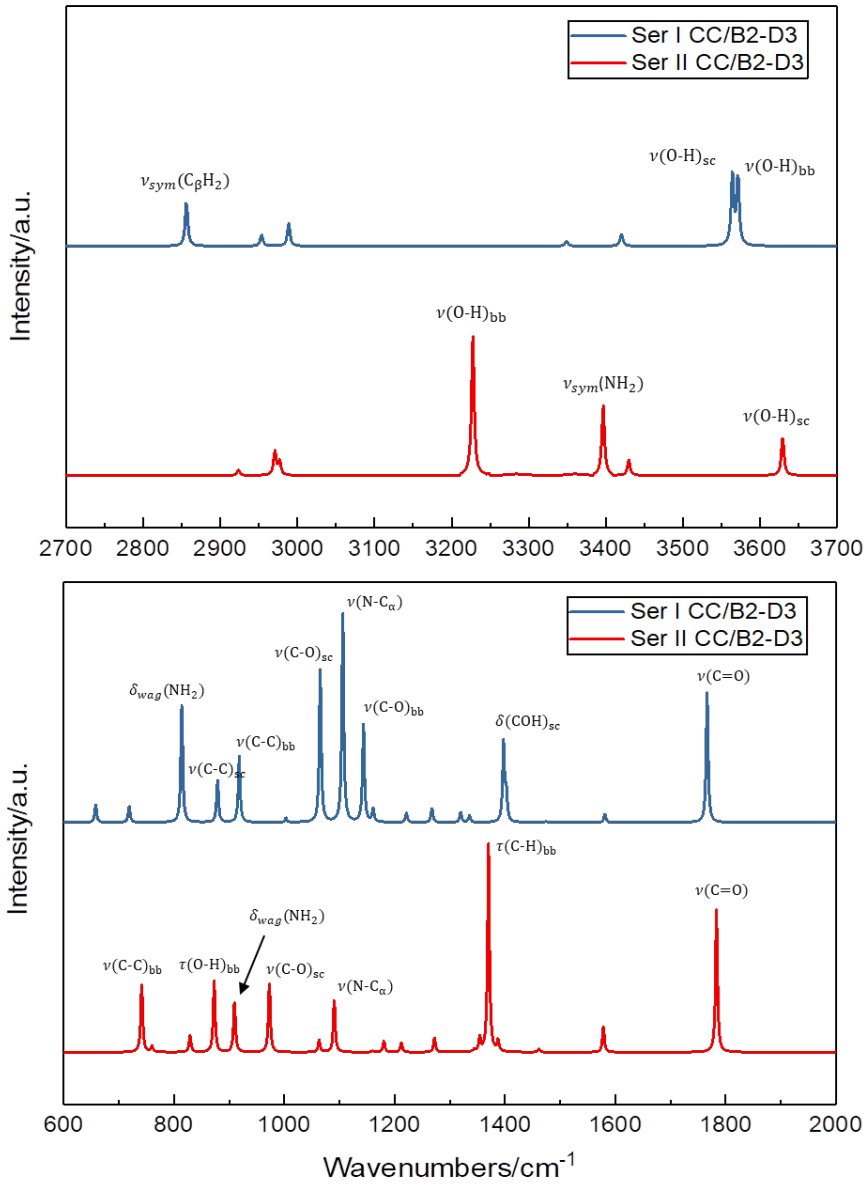


Figure 8. Anharmonic IR spectra of Ser I and Ser II computed at the CC/B2-D3 GVPT2 level.

4. Conclusions

The structures, relative stabilities, and vibrational wavenumbers of the two lowest energy conformers of serine have been characterized using state-of-the-art composite schemes. More precisely, the so-called "cheap" approach (CCSD(T)/(CBS+CV)^{MP2}) has been employed for structural and harmonic wavenumbers determinations. The great accuracy of CCSD(T)/(CBS+CV)^{MP2} equilibrium structures, even for flexible system, is well-known and, in this study, is further confirmed based on the comparison with experiment. Indeed, the corresponding equilibrium rotational constants only deviate 0.2-0.3% from their semi-experimental counterparts. It has been shown that the extrapolation to the CBS limit gives the largest contribution to all structural parameters, with those involved in hydrogen bonds being the most significantly affected by going beyond the CCSD(T)/cc-pVTZ level. Moving to harmonic vibrational wavenumbers, the present work is the first application of the "cheap" scheme to a system as large as serine. Inspection of the various contributions points out that the corrections on top of CCSD(T)/cc-pVTZ are not negligible, these being ~10 cm^{-1} on average, and are mainly related to basis set effects. Interestingly, for the various modes, the overall contribution varies both in sign and magnitude, and therefore it cannot be simply recovered by using scaling factors. The electronic energies have been obtained at the CCSD(T)/CBS+CV level (i.e. using a composite scheme entirely based on CCSD(T) calculations) on top of CCSD(T)/(CBS+CV)^{MP2} geometries and they confirm that the conformer II is more stable the Ser I. While the energy difference is 2 kJ/mol at the equilibrium, this reduces to ~0.5 kJ/mol upon inclusion of the zero-point-vibrational corrections, without altering the stability order.

In the second part of this study, the accurate results issued from composite schemes have been employed as references for benchmarking selected DFT functionals, thereby focusing on the hybrid and double-hybrid dispersion-corrected ones. It is demonstrated that double-hybrid functionals, in particular B2PLYP-D3, mPW2PLYPD and DSDPBEP86, combined with augmented triple-ζ basis sets, provide rather accurate structural parameters and vibrational properties. Their

accuracy is better than or similar to that of CCSD(T)/cc-pVTZ, but obtained at a much reduced computational cost. Most DFT functionals reliably predict the electronic energy difference between the two conformers, with differences within 0.5 kJ/mol with respect to the best estimate. Only the B2PLYP-D3/maug-cc-pVTZ and mPW2PLYPD/maug-cc-pVTZ underestimate the energy difference between the two conformers by about 1 kJ/mol, while the MN15 functional (with both basis sets) overestimates it by more than 2 kJ/mol.

Finally, hybrid CC/DFT anharmonic vibrational computations within the GVPT2 treatment allowed for simulating IR spectra. The best overall agreement with experiment is obtained by combining CCSD(T)/(CBS+CV)^{MP2} harmonic wavenumbers with B2PLYP-D3/maug-cc-pVTZ anharmonic corrections, with mean and maximum absolute errors in the mid-infrared region (excluding ν(O-H)$_{bb}$ of Ser II) of 8 cm^{-1} and 23 cm^{-1}, respectively. Interestingly, anharmonic computations employing B2PLYP-D3 and DSDPBEP86 functionals also perform very well, indeed showing MAE and MAX of 10 cm^{-1} and 36 cm^{-1}, respectively. Our study allowed us to confirm that the B2PLYP-D3 and DSDPBEP86 functionals employed in conjunction with triple-ζ basis sets augmented by diffuse funcions are reliable and cost effective methodologies that can be recommended for structural and spectroscopic studies of medium-sized flexible biomolecules showing intramolecular hydrogen bonds.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/xxxxx. *(i)* Best-estimated equilibrium geometries: structural parameters and Cartesian coordinates; *(ii)* Best-estimated harmonic and anharmonic vibrational wavenumbers; *(iii)* Equilibrium geometries computed at different levels of theory; *(iv)* Harmonic and anharmonic vibrational wavenumbers computed at different levels of theory.

Author Information

Corresponding Author: Malgorzata Biczysko

International Centre for Quantum and Molecular Structures, Physics Department, Shanghai University, 99 Shangda Road, Shanghai, 200444 China

email: biczysko@shu.edu.cn

ORCID ID: 0000-0002-6439-5561

Corresponding Author: Cristina Puzzarini

Department of Chemistry "Giacomo Ciamician", University of Bologna, Via F. Selmi 2, 40126 Bologna, Italy

email: cristina.puzzarini@unibo.it

ORCID ID: 0000-0002-2395-8532

Author Contributions

M. B. and C. P. designed the research. M. S. and F. S. performed the calculations. M. S. and M. B. analyzed the data. M. S., M. B. and C. P. wrote the paper.

Notes

The authors declare no competing financial interest.

Acknowledgements

This work was enabled by the financial support from the National Natural Science Foundation of China (Grant No. 31870738). In Bologna, this work has been supported by RFO funds (university of Bologna).

References:

[1]. Chin, W.; Piuzzi, F.; Dimicoli, I.; Mons, M. Probing the Competition between Secondary Structures and Local Preferences in Gas Phase Isolated Peptide Backbones. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1033-1048.

[2]. de Vries, M. S.; Hobza, P. Gas-Phase Spectroscopy of Biomolecular Building Blocks. *Annu. Rev. Phys. Chem.* **2007,** *58*, 585-612.

[3]. Lee, J. J.; Albrecht, M.; Rice, C. A.; Suhm, M. A. Adaptive Aggregation of Peptide Model Systems. *J. Phys. Chem. A* **2013**, *117*, 7050-7063.

[4]. Roy, T. K.; Kopysov, V.; Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V.; Gerber, R. B. Conformational Structures of a Decapeptide Validated by First Principles Calculations and Cold Ion Spectroscopy. *Chem. Phys. Chem.* **2015**, *16*, 1374-1378.

[5]. Alonso, J. L.; López, J. C. Microwave Spectroscopy of Biomolecular Building Blocks, In Gas-Phase IR Spectroscopy and Structure of Biological Molecules. Rijs, A. M.; Oomens, J., Eds. *Springer Berlin Heidelberg* **2015**, *364*, 335-401.

[6]. Rijs, A. M.; Oomens, J. Gas-Phase IR Spectroscopy and Structure of Biological Molecules Preface, In Gas-Phase IR Spectroscopy and Structure of Biological Molecules. Rijs, A. M.; Oomens, J., Eds. **2015**, *364*, V-VII.

[7]. Teschmit, N.; Horke, D. A.; Küpper, J. Spatially Separating the Conformers of a Dipeptide. *Angew. Chem. Int. Ed.* **2018**, *57*, 13775.

[8]. Gloaguen, E.; Mons, M.; Schwing, K.; Gerhards, M. Neutral Peptides in the Gas Phase: Conformation and Aggregation Issues. *Chem. Rev*. **2020**, *120*, 12490-12562.

[9]. Bakels, S.; Gaigeot, M. -P.; Rijs, A. M. Gas-Phase Infrared Spectroscopy of Neutral Peptides: Insights from the Far-IR and THz Domain. *Chem. Rev.* **2020**, *120*, 3233-3260.

[10]. Alonso, E. R.; Fusè, M.; León, I.; Puzzarini, C.; Alonso, J. L.; Barone, V. Exploring the Maze of Cycloserine Conformers in the Gas Phase Guided by Microwave Spectroscopy and Quantum Chemistry. *J. Phys. Chem. A* **2021**, *125*, 2121-2129.

[11]. Czinki, E.; Császár, A. G. Conformers of Gaseous Proline. *Chem.-Eur. J.* **2003**, *9*, 1008-1019.

[12]. Allen, W. D.; Czinki, E.; Császár, A. G. Molecular Structure of Proline. *Chem.-Eur. J.* **2004**, *10*, 4512-4517.

[13]. Pollreisz, F.; Gomory, I.; Schlosser, G.; Vékey, K.; Solt, I.; Császár, A. G. Mass Spectrometric and Quantum-Chemical Study on the Structure, Stability, and Chirality of Protonated Serine Dimers. *Chem. -Eur. J.* **2005**, *11*, 5908-5916.

[14]. Wilke, J. J.; Lind, M. C.; Schaefer, H. F.; Császár, A. G.; Allen, W. D. Conformers of Gaseous Cysteine. *J. Chem. Theory Comput.* **2009**, *5*, 1511-1523.

[15]. Jaeger, H. M.; Schaefer, H. F.; Demaison, J.; Császár, A. G.; Allen, W. D. Lowest-Lying Conformers of Alanine: Pushing Theory to Ascertain Precise Energetics and Semiexperimental Re Structures. *J. Chem. Theory Comput.* **2010**, *6*, 3066-3078.

[16]. Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. Accurate Structure, Thermodynamic and Spectroscopic Parameters from CC and CC/DFT Schemes: The Challenge of the Conformational Equilibrium in Glycine. *Phys. Chem. Chem. Phys.* **2013**, *15*, 10094-10111.

[17]. Puzzarini, C.; Biczysko, M.; Barone, V.; Largo, L.; Peña, I.; Cabezas, C.; Alonso, J. L. Accurate Characterization of the Peptide Linkage in the Gas Phase: A Joint Quantum-chemical and Rotational Spectroscopy Study of the Glycine Dipeptide Analogue. *J. Phys. Chem. Lett.* **2014**, *5*, 534-540.

[18]. He, K.; Allen, W. D. Conformers of Gaseous Serine. *J. Chem. Theory Comput.* **2016**, *12*, 3571-3582.

[19]. Roy, T. K.; Sharmac, R.; Gerber, R. B. First-Principles Anharmonic Quantum Calculations for Peptide Spectroscopy: VSCF Calculations and Comparison with Experiments. *Phys. Chem. Chem. Phys*. **2016**, *18*, 1607-1614.

[20]. Bakels, S.; Porskamp, S. B. A.; Rijs, A. M. Formation of Neutral Peptide Aggregates as Studied by Mass-Selective IR Action Spectroscopy. *Angew. Chem. Int. Edit.* **2019**, *58*, 10537-10541.

[21]. Bakels, S.; Meijer, E. M.; Greuell, M.; Porskamp, S. B. A.; Rouwhorst, G.; Mahé, J.; Gaigeot, M. -P.; Rijs, A. M. Interactions of Aggregating Peptides Probed by IR-UV Action Spectroscopy. *Faraday Discuss*. **2019**, *217*,

322-341.

[22]. Blodgett, K. N.; Jang, G.; Kim, S.; Kim, M. K.; Choi, S. H.; Zwier, T. S. Coexistence of Left- and Right-Handed 12/10-Mixed Helices in Cyclically Constrained β-Peptides and Directed Formation of Single-Handed Helices upon Site-Specific Methylation. *J. Phys. Chem. A* **2020**, *124*, 5856-5870.

[23]. Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L.; Improved Peptide and Protein Torsional energetics with the OPLS-AA force field. *J. Chem. Theory Comput*. **2015**, *11*, 3499-3509.

[24]. Grubisic, S.; Brancato, G.; Barone, V. An Improved AMBER Force Field for - Dialkylated Peptides: Intrinsic and Solvent-Induced Conformational Preferences of Model Systems. *Phys. Chem. Chem. Phys*. **2013**, *15*, 7395-17407.

[25]. Jiang, Z.; Biczysko, M.; Moriarty, N. W. Accurate Geometries for "Mountain Pass" Regions of the Ramachandran Plot Using Quantum Chemical Calculations. *Proteins: Structure, Function, and Bioinformatics*, **2018**, *86*, 273-278.

[26]. Prasad, V. K.; Otero-de-la-Roza, A.; DiLabio, G. A. PEPCONF, a Diverse Data Set of Peptide Conformational Energies. *Sci. Data* **2019**, *6*, 180310.

[27]. Kulik, H. J.; Luehr, N.; Ufimtsev, I. S.; Martinez, T. J. Ab Initio Quantum Chemistry for Protein Structures. *J. Phys. Chem. B*, **2012**, *116*, 12501-12509.

[28]. Schmitz, S.; Seibert, J.; Ostermeir, K.; Hansen, A.; Göller, A. H.; Grimme, S. Quantum Chemical Calculation of Molecular and Periodic Peptide and Protein Structures. *J. Phys. Chem. B*, **2020**, *124*, 3636-3646.

[29]. Zheng, M.; Biczysko, M.; Xu, Y.; Moriarty, N. W.; Kruse, H.; Urzhumtsev, A.; Waller, M. P.; Afonine, P. V. Including Crystallographic Symmetry in Quantum-based Refinement: Q|R#2. *Acta Crystallographica Section D-Structural Biology* **2020**, *76*, 41-50.

[30]. Frenkel-Pinter, M.; Samanta, M.; Ashkenasy, G.; Leman, L. J. Prebiotic Peptides: Molecular Hubs in the Origin of Life. *Chem. Rev.* **2020**, *120*, 4707-4765.

[31]. Comets and the Origin and Evolution of Life. Thomas, P. J.; Hicks, R. D.; Chyba, C. F.; McKay, C. P., Eds. 2nd ed. Adv. Astrobiol. Biogeophys XVII *Springer Verlag* **2006**.

[32]. Bernstein, M. P.; Dworkin, J. P.; Sandford, S. A.; Cooper, G. W.; Allamandola, L. J. Racemic Amino Acids from the Ultraviolet Photolysis of Interstellar Ice Analogues. *Nature* **2002**, *416*, 401-403.

[33]. Bada, J. L.; Glavin, D. P.; McDonald, G. D.; Becker, L. A Search for Endogenous Amino Acids in Martian Meteorite ALH84001. *Science* **1998**, *279*, 362-365.

[34]. Kvenvolden, K.; Lawless, J.; Pering, K.; Peterson, E.; Flores, J.; Ponnamperuma, C.; Kaplan, I. R.; Moore, C. Evidence for Extraterrestrial Amino-Acids and Hydrocarbons in the Murchison Meteorite. *Nature* **1970,** *228*, 923-926.

[35]. Elsila, J. E.; Glavin, D. P.; Dworkin, J. P. Cometary Glycine Detected in Samples Returned by Stardust. *Meteorit. Planet. Sci.* **2009**, *44*, 1323-30.

[36]. Altwegg, K.; Balsiger, H.; Bar-Nun, A.; Berthelier, J. J.; Bieler, A.; Bochsler, P.; Briois, C.; Calmonte, U.; Combi, M. R.; Cottin, H., et al. Prebiotic Chemicals — Amino Acid and Phosphorus — in the Coma of Comet 67P/Churyumov-Gerasimenko. *Sci. Adv.* **2016,** *2*, E1600285.

[37]. Burton, A. S.; Stern, J. C.; Elsila, J. E.; Glavin, D. P.; Dworkin, J. P. Understanding Prebiotic Chemistry through the Analysis of Extraterrestrial Amino Acids and Nucleobases in Meteorites. *Chem. Soc. Rev.* **2012**, *41*, 5459-5472.

[38]. Elsila, J. E.; Aponte, J. C.; Blackmond, D. G.; Burton, A. S.; Dworkin, J. P.; Glavin, D. P. Meteoritic Amino Acids: Diversity in Compositions Reflects Parent Body Histories. *ACS Cent. Sci*. **2016**, *2*, 370-379.

[39]. Morse, A. D.; Chan, Q. H. S. Observations of Cometary Organics: A Post Rosetta Review. *ACS Earth Space Chem.* **2019,** *3*, 1773-1791.

[40]. Elsila, J. E.; Dworkin, J. P.; Bernstein, M. P.; Martin, M. P.; Sandford, S. A. Mechanisms of Amino Acid Formation in Interstellar Ice Analogs. *Astrophys. J.* **2007**, *660*, 911-918.

[41]. Sandford, S. A.; Nuevo, M.; Bera, P. P.; Lee, T. J. Prebiotic Astrochemistry and the Formation of Molecules of Astrobiological Interest in Interstellar Clouds and Protostellar Disks. *Chem. Rev.* **2020**, *120*, 4616-4659.

[42]. Shivani; Singh, A.; Gupta, V.; Misra1, A.; Tandon, P. Quantum-Chemical Approach to Serine Formation in the Interstellar Medium: A Possible Reaction Pathway. *A&A* **2014**, *563*, A55.

[43]. Ehrenfreund, P.; Irvine, W.; Becker, L.; Blank, J.; Brucato, J. R.; Colangeli, L.; Derenne, S.; Despois, D.; Dutrey, A.; Fraaije, H., et al. Astrophysical and Astrochemical Insights into the Origin of Life. *Rep. Prog. Phys.* **2002**, *65*, 1427-1487.

[44]. Patel, B. H.; Percivalle, C.; Ritson, D. J.; Duffy, C. D.; Sutherland, J. D. Common Origins of RNA, Protein and Lipid Precursors in a Cyanosulfidic Protometabolism. *Nat. Chem.* **2015**, *7*, 301-307.

[45]. Puzzarini, C.; Barone, V. A Never-Ending Story in the Sky: The Secrets of Chemical Evolution. *Phys. Life Rev.* **2020,** *32*, 59-94.

[46]. Öberg, K. I. Photochemistry and Astrochemistry: Photochemical Pathways to Interstellar Complex Organic Molecules, *Chem. Rev.* **2016,** *116*, 9631-9663.

[47]. Puzzarini, C.; Barone, V. The Challenging Playground of Astrochemistry: An Integrated Rotational Spectroscopy - Quantum Chemistry Strategy. *Phys. Chem. Chem. Phys.* **2020**, *22*, 6507-6523.

[48]. Garrod, R. T.; Widicus Weaver, S. L. Simulations of Hot-Core Chemistry. *Chem. Rev.* **2013**, *113*, 8939-8960.

[49]. Caminati, W.; Grabow, J. U. Microwave Spectroscopy-Chapter 15: Molecular Systems In Frontiers of Molecular Spectroscopy. *Elsevier*, **2009**, 455-552.

[50]. Cabezas, C.; Varela, M.; Cortijo, V.; Jimenez, A. I.; Peña, I. Daly, A. M.; López, J. C.; Cativiela, C.; Alonso, J. L. The Alanine Model Dipeptide Ac-Ala-NH2 Exists as a Mixture of C7(eq) and C5 Conformers. *Phys. Chem. Chem. Phys*. **2013**, *15*, 2580-2585.

[51]. Eugenia, S. M.; Cabezas, C.; Mata, S.; Alonso, J. L. Rotational Spectrum of Tryptophan. *J. Chem. Phys.* **2014**, *140*.

[52]. Leon, I.; Alonso, E. R.; Mata, S.; Cabezas, C.; Alonso, J. L. Unveiling the Neutral Forms of Glutamine. *Angew. Chem. Int. Edit.* **2019**, *58*, 16002-16007.

[53]. Perez, G.; Mata, S.; Cabezas, C.; López, J. C.; Alonso, J. L. The Rotational Spectrum of Tyrosine. *J. Phys. Chem. A* **2015**, *119*, 3731-3735.

[54]. Cabezas, C.; Robben, M. A. T.; Rijs, A. M.; Peña, I.; Alonso, J. L. Fourier Transform Microwave Spectroscopy of Ac-Ser-NH2: The Role of Side Chain Interactions in Peptide Folding. *Phys. Chem. Chem. Phys.* **2015**, *17*, 20274-20280.

[55]. 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.

[56]. Kaczor, A.; Reva, I.; Proniewicz, L. M.; Fausto, R. Importance of Entropy in the Conformational Equilibrium of Phenylalanine: A Matrix-Isolation Infrared Spectroscopy and Density Functional Theory Study. *J. Phys. Chem. A*, **2006,** *110*, 2360-2370.

[57]. Pohl, G.; Perczel, A.; Vass, E.; Magyarfalvid, G.; Tarczay, G. A Matrix Isolation Study on Ac-Gly-NHMe and Ac-l-Ala-NHMe, The Simplest Chiral and Achiral Building Blocks of Peptides and Proteins. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4698-4708.

[58]. Legrady, B.; Vass, E.; Tarczay, G. Matrix-Isolation Vibrational Circular Dichroism Spectroscopy in Structural Studies of Peptides: Conformational Landscape of the Ac(-Ala)1-4-OMe Depsipeptide Series. *J. Mol. Spectrosc.* **2018**, *351*, 29-38.

[59]. Balabin, R. M. Conformational Equilibrium in Glycine: Experimental Jet-Cooled Raman Spectrum. *J. Phys. Chem. Lett.* **2010**, *1*, 20-23.

[60]. Unterberg, C.; Gerlach, A.; Schrader, T.; Gerhards, M. Structure of the Protected Dipeptide Ac-Val-Phe-OMe in the Gas Phase: Towards a β-Sheet Model System. *J. Chem. Phys.* **2003**, *118*, 8296-8300.

[61]. Burke, N. L.; DeBlase, A. F.; Redwine, J. G.; Hopkins, J. R.; McLuckey, S. A.; Zwier, T. S. Gas-Phase Folding of a Prototypical Protonated Pentapeptide: Spectroscopic Evidence for Formation of a Charge-Stabilized Beta-Hairpin. *J. Am. Chem. Soc.* **2016**, *138*, 2849-2857.

[62]. Bazsó, G.; Magyarfalvi, G.; Tarczay, G. Tunneling Lifetime of the ttc/VIvip Conformer of Glycine in Low-Temperature Matrices. *J. Phys. Chem. A* **2012**, *116*, 10539-10547.

[63]. Bazsó, G.; Najbauer, E. E.; Magyarfalvi, G.; Tarczay, G. Near-Infrared Laser Induced Conformational Change of Alanine in Low-Temperature Matrixes and the Tunneling Lifetime of Its Conformer VI. *J. Phys. Chem. A* **2013**, *117*, 1952-1962.

[64]. Najbauer, E. E.; Bazsó, G.; Apóstolo, R.; Fausto, R.; Biczysko, M.; Barone, V.; Tarczay, G. Identification of Serine Conformers by Matrix-Isolation IR Spectroscopy Aided by Near-Infrared Laser-Induced Conformational Change, 2D Correlation Analysis, and Quantum Mechanical Anharmonic Computations. J. *Phys. Chem. B* **2015**, *119*, 10496-10510.

[65]. Barone, V.; Alessandrini, S.; Biczysko, M.; Cheeseman, J. R.; Clary, D. C.; McCoy, A. B.; DiRisio, R. J.; Neese, F.; Melosso, M.; Puzzarini, C. Computational Molecular Spectroscopy. *Nat. Rev. Methods Primers* **2021**, *1*, 38.

[66]. Barone, V.; Biczysko, M.; Bloino, J.; Cimino, P.; Penocchio, E.; Puzzarini, C. CC/DFT Route toward Accurate Structures and Spectroscopic Features for Observed and Elusive Conformers of Flexible Molecules: Pyruvic Acid as a Case Study. *J. Chem. Theory Comput.* **2015**, *11*, 4342-4363.

[67]. Puzzarini, C.; Bloino, J.; Tasinato, N.; Barone, V. Accuracy and Interpretability: The Devil and the Holy Grail. New Routes across Old Boundaries in Computational Spectroscopy. *Chem. Rev.* **2019**, *119*, 8131-8191.

[68]. Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-Set Convergence of Correlated Calculations on Water. *J. Chem. Phys.* **1997**, *106*, 9639-9646.

[69]. Puzzarini, C. Extrapolation to the Complete Basis Set Limit of Structural Parameters: Comparison of Different Approaches. *J. Phys. Chem. A*, **2009**, *113*, 14530-14535.

[70]. Császár, A. G.; Allen, W. D.; Schaefer, H. F. III. In Pursuit of the Ab Initio Limit for Conformational Energy Prototypes. *J. Chem. Phys*. **1998**, *108*, 9751-9764.

[71]. 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.

[72]. Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical Accuracy in Ab Initio Thermochemistry and Spectroscopy: Current Strategies and Future Challenges. *Theor. Chem. Acc.* **2012**, *131*, 1079-1-20.

[73]. Puzzarini, C.; Heckert, J.; 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.

[74]. 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.

[75]. Heckert, M.; Kállay, M.; Gauss, J. Molecular Equilibrium Geometries Based on Coupled-Cluster Calculations Including Quadruple Excitations. *Mol. Phys.* **2005**, *103*, 2109-2115.

[76]. 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.

[77]. Raghavachari, K. Historical Perspective on: A Fifth-Order Perturbation Comparison of Electron Correlation Theories [Volume 157, Issue 6, 26 May 1989, Pages 479-483]. *Chem. Phys. Lett.* **2013**, 589, 35-36.

[78]. Alessandrini, S.; Barone, V.; Puzzarini, C. Extension of the "Cheap" Composite Approach to Noncovalent Interactions: The jun-ChS Scheme. *J. Chem. Theory Comput.* **2020**, *16*, 988-1006.

[79]. Puzzarini, C.; Barone, V. Diving for Accurate Structures in the Ocean of Molecular Systems with the Help of Spectroscopy and Quantum Chemistry. *Acc. Chem. Res.* **2018**, *51*, 548-556.

[80]. Kruse, H.; Šponer, J. Highly Accurate Equilibrium Structure of the C2h Symmetric N1-to-O2 Hydrogen-Bonded Uracil-Dimer. *Int. J. Quantum Chem.* **2018,** *118*, e25624.

[81]. Cerny, J.; Pitonak, M.; Riley, K. E.; Hobza, P. Complete Basis Set Extrapolation and Hybrid Schemes for Geometry Gradients of Noncovalent Complexes. *J. Chem. Theory Comput.* **2011**, *7*, 3924-3934.

[82]. 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.

[83]. Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618-622.

[84]. Puzzarini, C.; Biczysko, M.; Bloino, J.; Barone, V. Accurate Spectroscopic Characterization of Oxirane: A Valuable Route to its Identification in Titan's Atmosphere and the Assignment of Unidentified Infrared Bands. *Astrophys. J.* **2014**, *785*, 107.

[85]. Spada, L.; Tasinato, N.; Vazart, F.; Barone, V.; Caminati, W.; Puzzarini, C. Noncovalent Interactions and Internal Dynamics in Pyridine-Ammonia: A Combined Quantum-Chemical and Microwave Spectroscopy Study. *Chem. - Eur. J.* **2017**, *23*, 4876-4883.

[86]. Morgan, W. J.; Matthews, D. A.; Ringholm, M.; Agarwal, J.; Gong, J. Z.; Ruud, K.; Allen, W. D.; Stanton, J. F.; Schaefer, H. F. Geometric Energy Derivatives at the Complete Basis Set Limit: Application to the Equilibrium Structure and Molecular Force Field of Formaldehyde. *J. Chem. Theory Comput.* **2018**, *14*, 1333-1350.

[87]. 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.

[88]. 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.

[89]. Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini. C. Accurate Molecular Structures and Infrared Spectra of Trans-2,3-Dideuterooxirane, Methyloxirane, and Trans-2,3-Dimethyloxirane. *J. Chem. Phys.* **2014**, *141*, 034107.

[90]. Shu, C.; Jiang, Z.; Biczysko, M. Toward Accurate Prediction of Amino Acid Derivatives Structure and Energetics from DFT: Glycine Conformers and Their Interconversions. *J. Mol. Model*. **2020**, *26*, 129.

[91]. Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. Characterization of the Elusive Conformers of Glycine from State-of-the-Art Structural, Thermodynamic, and Spectroscopic Computations: Theory Complements Experiment. *J. Chem. Theory Comput.* **2013**, *9*, 1533-1547.

[92]. Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. Glycine Conformers: A Never-Ending Story? *Phys. Chem. Chem. Phys*. **2013**, *15*, 1358-1363.

[93]. Puzzarini, C.; Biczysko, M.; Barone, V.; Peña, I.; Cabezas, C.; Alonso, J. L. Accurate Molecular Structure and Spectroscopic Properties of Nucleobases: A Combined Computational-Microwave Investigation of

2-Thiouracil as a Case Study. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16965-16975.

[94]. 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.

[95]. Carbonniere, P.; Lucca, T.; Pouchan, C.; Rega, N.; Barone, V. Vibrational Computations beyond the Harmonic Approximation: Performances of the B3LYP Density Functional for Semirigid Molecules. *J. Comput. Chem.* **2005**, *26*, 384-388.

[96]. 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.

[97]. Fogarasi, G.; Pulay, P. Ab Initio Vibrational Force Fields. *Ann. Rev. Phys. Chem.* **1984**, *35*, 191-213.

[98]. Handy, N. C.; Gaw, J. F.; Simandiras, E. D. Accurate Ab Initio Prediction of Molecular Geometries and Spectroscopic Constants, Using SCF and MP2 Energy Derivatives. *J. Chem. Soc., Faraday Trans. 2*, **1987**, *83*, 1577-1593.

[99]. Biczysko, M.; Bloino, J.; Brancato, G.; Cacelli, I.; Cappelli, C.; Ferretti, A.; Lami, A.; Monti, S.; Pedone, A.; Prampolini, G., et al. 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.

[100]. Dressler, S.; Thiel, W. Anharmonic Force Fields from Density Functional Theory. *Chem. Phys. Lett.* **1997**, *273*, 71-78.

[101]. Lee C.; Yang W.; Parr R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B*, **1988**, *37*, 785-789.

[102]. Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648-5652.

[103]. Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

[104]. 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.

[105]. Barone, V.; Biczysko, M.; Bloino, J. Fully Anharmonic IR and Raman Spectra of Medium-Size Molecular Systems: Accuracy and Interpretation. *Phys. Chem. Chem. Phys.* **2014**, *16*, 1759.

[106]. Barone, V. The Virtual Multifrequency Spectrometer: A New Paradigm for Spectroscopy. *Wires. Comput. Mol. Sci.* **2016**, *6*, 86-110.

[107]. Chaudret, R.; de Courcy, B.; Contreras-García, J.; Gloaguen, E.; Zehnacker-Rentien, A.; Mons, M.; Piquemal, J. -P. Unraveling Non-Covalent Interactions within Flexible Biomolecules: from Electron Density Topology to Gas Phase Spectroscopy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9876.

[108]. Řezáč, J. Non-Covalent Interactions Atlas Benchmark Data Sets: Hydrogen Bonding. *J. Chem. Theory Comput*. **2020**, *16*, 2355-2368.

[109]. Blanco, S.; Sanz, M. E.; López, J. C.; Alonso, J. L. Revealing the multiple structures of serine. *Proc. Natl. Acad. Sci.* **2007**, *104*, 20183-20188.

[110]. Jarmelo, S.; Lapinski, L.; Nowak, M.; Carey, P.; Fausto, R. J. Preferred Conformers and Photochemical ($\lambda > 200$ nm) Reactivity of Serine and 3,3-Dideutero-Serine In the Neutral Form. *J. Phys. Chem. A* **2005**, 109, 5689-5707.

[111]. Lambie, B.; Ramaekers, R.; Maes, G. J. Conformational Behavior of Serine: An Experimental Matrix-Isolation FT-IR and Theoretical DFT(B3LYP)/6-31++G** Study. *J. Phys. Chem. A* **2004**, *108*, 10426-10433.

[112]. Jarmelo, S.; Fausto, R. Entropy Effects in Conformational Distribution and Conformationally Dependent UV-Induced Photolysis of Serine Monomer Isolated in Solid Argon. *J. Mol. Struct.* **2006**, *786*, 175-181.

[113]. Gauss, J.; Stanton, J. Analytic CCSD(T) Second Derivatives. *Chem. Phys. Lett.* 1997, 276, 70-77.

[114]. Dunning, T. H. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007-1023.

[115]. Woon, D. E.; Dunning, T. H. 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.

[116]. Peterson, K. A.; Dunning, T. H. Accurate Correlation Consistent Basis Sets for Molecular Core-Valence Correlation Effects: The Second Row Atoms Al-Ar, and the First Row Atoms B-Ne Revisited. *J. Chem. Phys*. **2002**, *117*, 10548-10560.

[117]. Puzzarini, C. Accurate Molecular Structures of Small- and Medium-Sized Molecules. *Int. J. Quantum Chem.* **2016**, *116*, 1513-1519.

[118]. Kendall, R. A.; Dunning, Jr. T. H.; Harrison, R. J. Electron Affinities of the First-Row Atoms Revisited. Systematic Basis Sets and Wave Functions. *J. Chem. Phys.* **1992**, *96*, 6796-6806.

[119]. 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.

[120]. Stanton, J. F.; Gauss, J.; Cheng, L.; Harding, M. E.; Matthews, D. A.; Szalay, P. G.; Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C., et al. *CFOUR*, Revision V.2; **2019**. For the current version, see http://www.cfour.de.

[121]. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H., et al. *Gaussian 16,* Revision C.1; Gaussian, Inc.: Wallingford, CT, **2019**.

[122]. Burke, K. Perspective on Density Functional Theory. *J. Chem. Phys.* **2012**, *136*, 150901.

[123]. Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn-Sham Global-Hybrid Exchange-Correlation Density Functional with Broad Accuracy for Multi-Reference and Single-Reference Systems and Noncovalent Interactions. Chem. Sci. **2016**, *7*, 5032-5051.

[124]. Schwabe, T.; Grimme, S. Double-Hybrid Density Functionals with Long-Range Dispersion Corrections: Higher Accuracy and Extended Applicability. *Phys. Chem, Chem. Phys.* **2007**, *9*, 3397-3406.

[125]. Kozuch, S.; Martin, J. M. L. DSD-PBEP86: in Search of the Best Double-Hybrid DFT with Spin-Component Scaled MP2 and Dispersion Corrections. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104-20107.

[126]. Brémond. E.; Adamo, C. Seeking for Parameter-Free Double-Hybrid Functionals: The PBE0-DH Model. *J. Chem. Phys*. **2011**, *135*, 024106.

[127]. Brémond, E.; Sancho-Garcia, J. C.; Perez-Jimenez, A. J.; Adamo, C. Communication: Double-Hybrid Functionals from Adiabatic-Connection: The QIDH Model. *J. Chem. Phys.* **2014**, *141*, 031101.

[128]. Carnimeo, I.; Puzzarini, C.; Tasinato, N.; Stoppa, P.; Charmet, A. P.; Biczysko, M.; Cappelli, C.; Barone, V. Anharmonic Theoretical Simulations of Infrared Spectra of Halogenated Organic Compounds. *J. Chem. Phys.* **2013,** *139*, 074310.

[129]. Papajak, E.; Leverentz, H. R.; Zheng, J.; Truhlar, D. G. Efficient Diffuse Basis Sets: cc-pVxZ plus and maug-cc-pVxZ. *J. Chem. Theory Comput.* **2009**, *5*, 1197-1202.

[130]. Grimme, S. Density Functional Theory with London Dispersion Corrections. *Wiley Interdisciplinary Reviews-Computational Molecular Science.* **2011**, *1*, 211-228.

[131]. Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate ab initio Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys*. **2010**, *132*.

[132]. Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J. Comput. Chem.* **2011**, *32,* 1456-1465.

[133]. Duschinsky, F. The Importance of the Electron Spectrum in Multi Atomic Molecules. Concerning the Franck-Condon principle. *Acta Physicochim. URSS*, **1937**, *7*, 551.

[134]. Bloino, J.; Baiardi, A.; Biczysko, M. Aiming at an Accurate Prediction of Vibrational and Electronic Spectra for Mediu-to-Large Molecules: An Overview. *Int. J. Quantum Chem.* **2016**, *116*, 1543-1574.

[135]. Fornaro, T.; Carnimeo, I.; Biczysko, M. Toward Feasible and Comprehensive Computational Protocol for Simulation of the Spectroscopic Properties of Large Molecular Systems: The Anharmonic Infrared Spectrum of Uracil in the Solid State by the Reduced Dimensionality/Hybrid VPT2 Approach. *J. Phys. Chem. A* **2015**, *119*, 5313-5326.

[136]. Nielsen, H. H. The Vibration-Rotation Energies of Molecules, *Rev. Mod. Phys.* **1951**, *23*, 90-136.

[137]. Franke, P. R.; Stanton, J. F.; Douberly, G. E. How to VPT2: Accurate and Intuitive Simulations of CH Stretching Infrared Spectra Using VPT2+K with Large Effective Hamiltonian Resonance Treatments. *J. Phys. Chem. A* **2021**.

[138]. Barone, V. Anharmonic Vibrational Properties by a Fully Automated Second-Order Perturbative Approach. *J. Chem. Phys.* **2005**, *122*, 14108.

[139]. 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.

[140]. Bloino, J.; Biczysko, M.; Barone, V. Anharmonic Effects on Vibrational Spectra Intensities: Infrared, Raman, Vibrational Circular Dichroism, and Raman Optical Activity. *J. Phys. Chem. A* **2015**, *119*, 11862-11874.

[141]. Mendolicchio, M.; Bloino, J.; Barone, V. General Perturb-Then-Diagonalize Model for the Vibrational Frequencies and Intensities of Molecules Belonging to Abelian and Non-Abelian Symmetry Groups. *J. Chem. Theory. Comput.* **2021**.

[142]. Johnson, E. R.; Keinan, S.; Mori-Sánchez, P.; Contreras-García, J.; Cohen, A. J.; Yang, W. Revealing Noncovalent Interactions. *J. Am. Chem. Soc.* **2010**, *132*, 6498-6506.

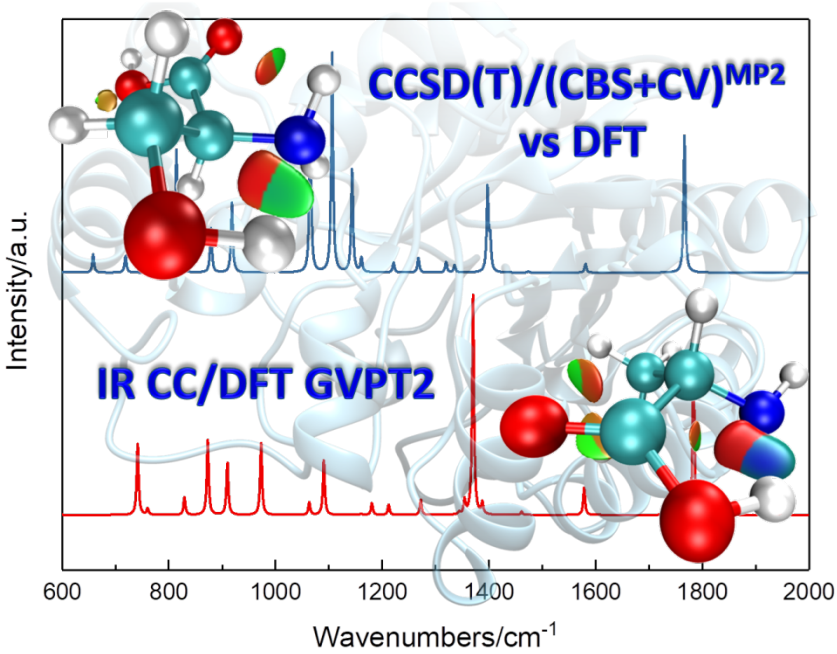
[143]. Lu, T.; Chen, F. Multiwfn: A multifunctional wavefunction analyzer. *J. Comput. Chem.* **2012**, *33*, 580-592.

[144]. Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual Molecular Dynamics. *J. Mol. Graphics* **1996**, *14*, 33-38.

[145]. Biczysko, M.; Latajka, Z. Accuracy of Theoretical Potential Energy Profiles along Proton-Transfer Coordinate for XH−NH_3 (X = F, Cl, Br) Hydrogen-Bonded Complexes. *J. Phys. Chem. A* **2002**, *106*, 3197-3201.

[146]. Baiardi, A.; Reiher, M. The Density Matrix Renormalization Group in Chemistry and Molecular Physics: Recent Developments and New Challenges. *J. Chem. Phys*. **2020**, *152*, 040903.

[147]. Baiardi, A.; Stein, C. J.; Barone, V.; Reiher, M. Vibrational Density Matrix Renormalization Group. *J. Chem. Theory Comput.* **2017**, *13*, 3764-3777.



TOC Graphics