THE JOURNAL OF PHYSICAL CHEMISTRY A
A JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

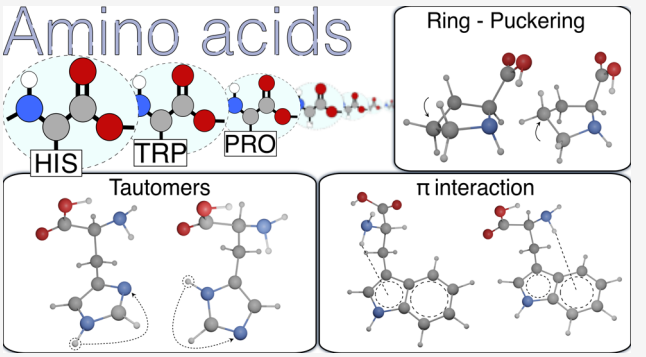
pubs.acs.org/JPCA
Article

DFT Meets Wave-Function Methods for Accurate Structures and Rotational Constants of Histidine, Tryptophan, and Proline

Vincenzo Barone,* Lina Marcela Uribe Grajales, Silvia Di Grande, Federico Lazzari, and Marco Mendolicchio

Cite This: J. Phys. Chem. A 2023, 127, 7534−7543
Read Online

ABSTRACT: A new computational strategy has been applied to the conformational and spectroscopic properties in the gas phase of amino acids with very distinctive features, ranging from different tautomeric forms (histidine) to ring puckering (proline), and heteroaromatic structures with non-equivalent rings (tryptophan). The integration of modern double-hybrid functionals and wave-function composite methods has allowed us to obtain accurate results for a large panel of conformers with reasonable computer times. The remarkable agreement between computations and microwave experiments allows an unbiased interpretation of the latter in terms of stereoelectronic effects.

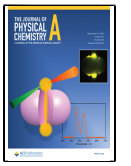


INTRODUCTION

Increasing attention has been paid in the last years to the conformational landscape of amino acids, which couple limited dimensions with a remarkable flexibility tuned by different kinds of non-covalent interactions.[1,2] Since environmental effects can strongly modify the characteristics of amino acids (for instance, zwitterionic forms are more stable in crystals[3] and aqueous solutions,[4] whereas neutral forms are exclusively found in the gas phase[5] or in inert matrixes[6]), an unbiased disentanglement of intrinsic stereoelectronic features requires preliminary studies in the gas phase. Thanks to the development of spectrometers coupling supersonic-jet expansion[7] and laser ablation,[8] thermolabile molecules with high melting points (like most amino acids) have become accessible to high-resolution spectroscopy studies. However, the interpretation of experimental spectra in structural and thermochemical terms is made difficult by the fast relaxation of some conformers to more stable counterparts whenever the corresponding energy barriers can be overcome under the specific experimental conditions.[9−11] Quantum chemical (QC) computations can help solve this kind of problems, provided that they couple accuracy[12−14] and feasibility for large numbers of different structures of medium-sized molecules.[15−17] Furthermore, an effective exploration of flat potential energy surfaces (PESs) requires more refined strategies[18,19] with respect to the systematic searches and/or local optimization techniques routinely employed for small semirigid molecules. In our opinion, the most suitable approach involves the synergistic use of QC methods of increasing sophistication in the different steps of an exploration/exploitation workflow driven by machine learning (ML) tools.[20−22] In this framework, after the preliminary discovery of conformers lying in a sufficiently large energy range by relatively cheap methods, the structures of the most stable conformers are refined by a double-hybrid functional and, possibly, further improved by a linear regression approach (LRA) involving a few empirical parameters in order to correct systematic errors.[23,24] Next, the transition states (TSs) ruling interconversion paths between pairs of conformers are found and energy minima connected to more stable species by low energy barriers are removed from the conformer list. Improved relative energies of the surviving conformers are evaluated by single-point computations with a wave-function composite method[25−30] and used, together with zero-point energies (ZPEs), computed by the double-hybrid functional mentioned above, to determine the final relative populations. Finally, spectroscopic parameters of the energy minima with non-negligible populations are computed.[16,31]

Systematic studies based on this strategy have confirmed that the backbone of natural amino acids containing simple non-polar side chains shows two main hydrogen bond patterns (usually referred to as type I and type II).[5,32−34] On the other hand, polar side chains give access to backbone−(side chain) hydrogen bonds, with this strongly increasing the number of

Received: June 23, 2023
Revised: August 12, 2023
Published: September 4, 2023



https://doi.org/10.1021/acs.jpca.3c04227
J. Phys. Chem. A 2023, 127, 7534−7543

low-energy conformers.[35−39] Additional interactions between backbone polar hydrogen atoms and side chain π-systems are possible for amino acids containing aromatic moieties.[40−42] This diversified landscape can be further enriched by a comprehensive analysis of amino acids showing additional features, like tautomerism (histidine, His), ring puckering (proline, Pro), or heteroaromatic structures with non-equivalent rings (tryptophan, Trp). Since all of these amino acids have been investigated in the gas phase by microwave (MW) spectroscopy,[38,43−45] the QC results must match those accurate experimental data. In this framework, the most distinctive feature of the present analysis with respect to previous studies is the coupling of feasibility and accuracy (relative mean unsigned errors (RMUEs) within 0.3% for rotational constants and 1% for quadrupolar coupling constants and relative energies), allowing the *a priori* prediction of experimental outcomes without any *ad hoc* assumption.

METHODS

As already mentioned in the Introduction, a preliminary exploration of the conformational PES by a fast semiempirical method[46] guided by ML algorithms[22,47] is followed by a characterization of low-energy conformers at the B3LYP-D3BJ/6-31+G* level.[48,49] The same combination of functional and basis set (hereafter B3/SVP) will be used also for the computation of anharmonic contributions (vide infra). The geometries of conformers lying within 1500 cm^{-1} above the absolute energy minimum are refined by the revDSD-PBEP86-D3BJ double-hybrid functional[50−53] (hereafter rDSD) in conjunction with the jun-cc-pVTZ basis set[52] (hereafter j3). While the rDSD/j3 model provides excellent conformational landscapes[15,54,55] and geometrical parameters,[24] further refinements are needed for structures lying below 1000 cm^{-1} and not connected to more stable energy minima by barriers lower than 400 cm^{-1}.[56,57] In fact, the main outcomes of experimental MW spectra are the ground-state rotational constants (B_τ^0, where τ refers to the inertial axes *a*, *b*, *c*), which include, together with equilibrium rotational constants (B_τ^{eq}), also electronic contributions (neglected in the following due to their very small values) and vibrational corrections (ΔB_τ^{vib}).[16,58,59] The leading (and most expensive) contribution to vibrational corrections comes from cubic force constants.[60] Fortunately, these terms can be obtained at affordable levels of theory (B3/SVP in the present context) since errors of the order of 10% on vibrational corrections correspond to errors lower than 0.1% on the overall rotational constants.[16,61] On the other hand, errors in the same range for equilibrium rotational constants can be obtained only by state-of-the-art QC methods.[62−65] As already mentioned in the Introduction, the systematic nature of the errors allows significant improvement of the rDSD/j3 results by the LRA;[24,66,67] however, the use of empirical parameters is not fully satisfactory. Therefore, an extensive benchmark of different basis sets and additional contributions was performed, which led to the selection of the cc-pVTZ-F12 basis set[68] (hereafter 3F12) and to the inclusion of core−valence (CV) correlation at the MP2 level in conjunction with the cc-pwCVTZ[69] (hereafter wC3) basis set. These choices define the new Pisa composite scheme (PCS),[70] in which each geometrical parameter (r) is obtained by combining the corresponding parameters optimized at different levels

$$r(\mathrm{PCS}) = r_{V2} + \Delta r_{CV2} \quad (1)$$

where r_{V2} is the geometrical parameter computed including an estimate of valence correlation energy with methods not exceeding the MP2 level (rDSD/3F12 in the present case) and Δr_{CV2} is the CV correction obtained from the difference between all-electron (ae) and frozen core (fc) MP2 computations in conjunction with the wC3 basis set. Several test computations have shown that PCS geometrical parameters are extremely accurate, and this will be further checked in the present context with reference to low-lying conformers of His, Trp, and Pro.

Since all amino acids contain at least one nitrogen atom, ^{14}N nuclear quadrupole coupling constants (χ_{ii}, with i referring to the inertia axis *a*, *b*, or *c*) play a non-negligible role in the accurate predictions of rotational spectra.[62,71] Furthermore, the intensities of the different MW transitions are determined by the components of dipole moments (μ_i).[58,71] While both dipole moments and quadrupole coupling constants can be computed with sufficient accuracy at the rDSD level,[39,72] accurate relative energies determining the conformer populations can be obtained by single-point energy evaluations on top of PCS geometries using composite wave-function methods rooted in the coupled cluster (CC) ansatz including single, double, and (perturbatively) triple excitations CCSD(T).[73,74] The final expression of the PCS energy is analogous to that of the PCS geometrical parameters, but now E_{V2} includes the CBS extrapolation and a further term (ΔE_V) is added to take into account valence correlation beyond the MP2 level. The CBS extrapolations appearing in both the E_{V2} and ΔE_V terms are performed by the standard n^{-3} two-point formula.[75] In order to allow the inclusion of those contributions, the E_{V2} term is evaluated at the MP2 level in place of the rDSD level employed for geometries

$$E(\mathrm{PCS}) = E_{V2} + \Delta E_V + \Delta E_{CV2} \quad (2)$$

where

$$E_{V2} = \frac{4^3 E(\mathrm{MP2/4F12}) - 3^3 E(\mathrm{MP2/3F12})}{4^3 - 3^3} \quad (3)$$

and

$$\Delta E_V = \frac{3^3 \Delta E(\mathrm{3F12}) - 2^3 \Delta E(\mathrm{2F12})}{3^3 - 2^3} \quad (4)$$

with

$$\Delta E(n\mathrm{F12}) = E(\mathrm{CCSD(T)}/n\mathrm{F12}) - E(\mathrm{MP2}/n\mathrm{F12}) \quad (5)$$

Finally

$$\Delta E_{CV2} = E(\mathrm{ae\text{-}MP2/wC3}) - E(\mathrm{MP2/wC3}) \quad (6)$$

In the equations above, all the energies are obtained within the fc approximation, unless the label ae (all-electron) is explicitly employed.

Finally, the ZPEs required for the computation of standard enthalpies at 0 K (ΔH_0°) are evaluated in the framework of vibrational perturbation theory to second order (VPT2),[76−78] employing rDSD/3F12 harmonic frequencies[31] and B3/SVP anharmonic contributions,[79,80] except in the case of tryptophan, where also harmonic contributions have been obtained at the B3/SVP level.

The aim of the PCS model is to approach the accuracy of CCSD(T)+CBS+CV computations at the cost of a triple-ζ

https://doi.org/10.1021/acs.jpca.3c04227
J. Phys. Chem. A 2023, 127, 7534−7543

Table 1. Equilibrium Rotational Constants and Vibrational Corrections for the Species Detected in MW Experiments Computed at the B3/SVP Level[a]

	B_a^{eq}	B_b^{eq}	B_c^{eq}	ΔB_a^{vib}	ΔB_b^{vib}	ΔB_c^{vib}
Imidazole	9698	9363	4764	−79.9	−77.2	−41.0
Indole	3872	1629	1146	−29.9	−10.0	−7.3
Histidine ϵIIgg$^-$	1838	822	740	−13.9	−2.6	−2.6
Tryptophan IIgg	1235	391	346	−11.3	−1.6	−1.5
Tryptophan IIg$^-$g	1293	332	287	−9.5	−2.5	−1.9
Proline IE$^-$ [b]	3926	1541	1341	−44.3	−16.5	−14.7
Proline IE$^+$ [b]	3995	1551	1269	−44.6	−13.6	−12.9
Proline IIE$^-$	3717	1651	1382	−46.9	−7.1	−6.0
Proline IIE$^+$ [b]	3981	1570	1256	−46.6	−13.6	−9.7

[a]All of the values are given in MHz. [b]The lowest frequency normal mode has been left harmonic.

Table 2. Comparison between the Experimental and Computed Rotational Constants of Imidazole and Indole[a]

Molecule	Parameter	Exp.[b]	rDSD/j3	LRA[c]	PCS[c]	B3/j3	MP2/j3
imidazole	B_a	9725.3	9755.6	9721.5	9723.7	9756.4	9755.9
	B_b	9374.0	9406.7	9370.3	9380.7	9463.6	9403.5
	B_c	4771.9	4789.0	4769.6	4772.8	4803.9	4788.2
	MAX		32.7	3.8	6.7	32.0	30.6
	MUE		26.7	3.3	3.1	24.5	25.5
	RMAX		0.36%	0.05%	0.07%	0.67%	0.34%
	RMUE		0.34%	0.04%	0.03%	0.37%	0.32%
indole	B_a	3877.8	3891.0	3876.2	3880.0	3907.0	3885.6
	B_b	1636.0	1638.9	1635.5	1636.7	1642.9	1642.7
	B_c	1150.9	1153.2	1150.5	1151.4	1156.6	1154.6
	MAX		13.2	1.6	2.2	29.2	7.8
	MUE		6.1	0.8	1.1	13.9	6.1
	RMAX		0.34%	0.04%	0.06%	0.75%	0.41%
	RMUE		0.24%	0.03%	0.05%	0.56%	0.31%

[a]All the values (except relative errors) are in MHz. [b]From ref 86 for imidazole and ref 87 for indole. [c]Includes B3/SVP vibrational corrections from Table 1.

CCSD(T) computation and without any empirical parameter thanks to the evaluation of CBS and CV contributions by the inexpensive MP2 model. While use of smaller basis sets requires the introduction of empirical factors,[81] the success of the "cheap" family of methods[26] witnesses that this goal can be reached starting from triple-ζ basis sets, and the PCS model further improves the results. Furthermore, single-point CCSD(T) energy evaluations with triple-ζ basis sets are feasible today for very large molecules thanks to the implementation of linear-scaling algorithms possibly employing local-correlation treatments.[82−84] The situation is different for gradient evaluations, where fast implementations are available only for DFT and MP2 (hence double-hybrids), which are, therefore, employed in the PCS model. In any case, since the dimensions of the studied molecules are small enough to allow the use of conventional approaches, all the computations have been performed with the Gaussian package.[85]

RESULTS AND DISCUSSION

The B3/SVP equilibrium rotational constants and vibrational corrections of all of the molecules studied in the present work are collected in Table 1. It is apparent that, as mentioned in the Introduction, vibrational corrections are of the order of 1% of the corresponding rotational constants. As a consequence, they cannot be neglected when aiming at unbiased comparisons with experiments.

Before analyzing the specific targets of this work, let us consider the semirigid molecules corresponding to the side chains of histidine and tryptophan, namely, imidazole and indole. Since MW spectra are available for both molecules, a first estimate of the reliability of different methods can be obtained in the absence of the additional challenges related to backbone flexibility. The maximum and mean unsigned errors (MAX and MUE, respectively) are used together with their relative values (RMAX and RMUE) to analyze the quality of the results delivered by different computational models.

The results collected in Table 2 confirm that at most qualitative trends can be obtained by the methods usually employed for the interpretation of MW spectra (B3LYP and MP2), whereas the LRA based on rDSD/j3 computations confirms its remarkable performance. However, the new PCS approach delivers comparable results without the need for any empirical parameter besides those already present in the underlying electronic structure method. At this level, both CV correlation and vibrational corrections need to be included since they play an opposite role, but the issuing error compensation is far from being perfect.

The "soft" dihedral angles governing the conformational landscape of His and Trp belong either to the backbone (ϕ = LP−N−C$^\alpha$−C′ and ψ = N−C$^\alpha$−C′−O(H) dihedral angles) or to the side chain (χ_1 = N−C$^\alpha$−C$^\beta$−C$^\gamma$ and χ_2 = C$^\alpha$−C$^\beta$−C$^\gamma$−N$^\delta$ for histidine or χ_2 = C$^\alpha$−C$^\beta$−C$^\gamma$−C$^\delta$ for tryptophan). LP is the nitrogen lone-pair perpendicular to the plane defined by the two amine hydrogens and the C$^\alpha$ atom (Figure 1). Only nearly planar conformations are allowed for the carboxy moiety of all amino acids (ω = C$^\alpha$−C′−O−H ≈ 0° or 180°), with ω ≈

https://doi.org/10.1021/acs.jpca.3c04227
J. Phys. Chem. A 2023, 127, 7534−7543

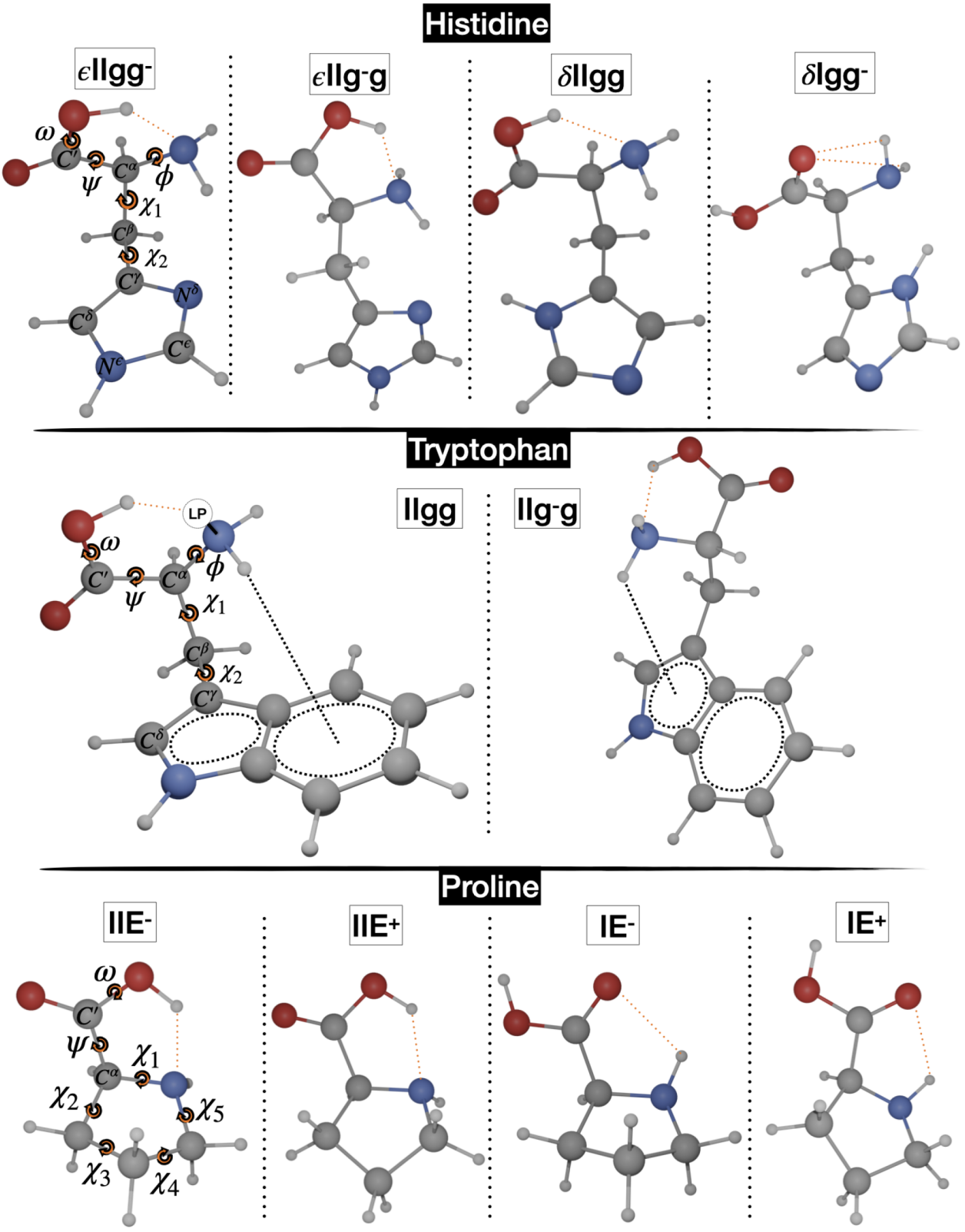


Figure 1. Low-energy minima of histidine, tryptophan, and proline.

0° being preferred, unless the oxidryl hydrogen is involved in strong hydrogen bonds with other electronegative atoms. The *c*, *g*, *s*, and *t* labels are then used to indicate the cis, gauche, skew, and trans conformations determined by the "soft" dihedral angles in the following order: ϕ, ψ, χ_1, and χ_2. In the case of Pro, the ψ and ω dihedral angles retain the same definitions, but the puckering of the pyrrolidine ring must be properly defined.[88] A full description of five-membered rings requires in principle two pseudorotation coordinates, the puckering amplitude (α) and the phase angle (τ),[89,90] which can be obtained from the endocyclic torsion angles χ_1, ..., χ_5 (see Figure 1):

https://doi.org/10.1021/acs.jpca.3c04227
J. Phys. Chem. A 2023, 127, 7534−7543

Table 3. rDSD/3F12 Relative Electronic Energies (ΔE_{rDSD}) and Harmonic Zero-Point Energies (ΔZPE_H), Together with Differences between PCS and rDSD/3F12 Electronic Energies (ΔPCS) and B3/SVP Anharmonic Corrections to ZPEs (ΔZPE_{A-H}) for the Low-Lying Tautomers and Conformers of Histidine[a]

Label	ΔE_{rDSD}	ΔPCS	ΔZPE_H[b]	$\Delta ZPE_{(A-H)}$[c]	ΔH_0°[d]	ϕ	ψ	ω	χ_1	χ_2
ϵIIgg$^-$	0	0	0	0	0	−8	9	−2	59	−72
ϵIIg$^-$g	395	18	5 (15)	31	449	13	−15	3	−67	65
δIIgg	165	−6	32 (37)	95	286	−17	20	−6	59	75
δIgg$^-$	1005	−10	−128 (−120)	88	955	82	−174	179	67	−49

[a]Best estimates of relative enthalpies at 0 K (ΔH_0°) and dihedral angles (Figure 1) optimized at the PCS level are also given. The angles are in degrees, whereas all of the energetic quantities are in cm^{-1}. [b]At the rDSD/3F12 level and (in parentheses) at the B3/SVP level. [c]At the B3/SVP level. [d]Sum of columns 2, 3, 4, and 5.

$$\alpha \cos(\tau) = \chi_1 \tag{7}$$

$$\alpha \sin(\tau) = \frac{\chi_2 - \chi_3 + \chi_4 - \chi_5}{-2[\sin(4\pi/5) + \sin(2\pi/5)]} \tag{8}$$

Two tautomeric forms are possible for histidine, depending on the presence of an acidic hydrogen on N$^\delta$ or N$^\epsilon$ (referred to as δ and ϵ in the following).

In agreement with previous computational studies,[91] the exploration of the conformational landscape of His provided several low-energy structures, most of which are stabilized by hydrogen bonds of type I (bifurcated $NH_2\cdots O{=}C$, $\phi \approx 180°$, $\psi \approx 180°$, $\omega \approx 180°$, *ttt*) and II (N···HO, $\phi \approx 0°$, $\psi \approx 0°$, $\omega \approx 0°$, *ccc*).[8]

As already reported for other amino acids[39] some low-energy conformers of type III ($\phi \approx 180°$, $\psi \approx 0°$, $\omega \approx 180°$, *tct*) have been found, but they can easily relax to more stable I conformers overcoming the very small energy barriers governing rotation around the ψ dihedral angle. Also the number of detectable conformers of type I is reduced by fast relaxations through nearly free rotation around the χ_1 dihedral angle. Therefore, our computations suggest that only four conformers might be detected in MW experiments. The corresponding structures are shown in Figure 1, and the main structural and energetic features are collected in Table 3 and Table S1 of the Supporting Information. Contrary to the usual situation for aliphatic amino acids,[39] conformers of type II are more stable than their type I counterparts (in spite of a less favorable orientation of the OH group in the carboxy moiety) since only the former conformers can establish a favorable interaction between the aromatic π-system of imidazole and one aminic hydrogen of the backbone.[92] The different intramolecular hydrogen bonding networks ruling the relative stability of type II structures have been recently analyzed,[91] and our results confirm the main conclusion of that work. The results collected in Table 3 show that ZPEs tune the relative stability of the type I and type II conformers.

Furthermore, the difference between rDSD/3F12 and PCS relative electronic energies does not exceed 20 cm^{-1} and that between B3/SVP and rDSD/3F12 relative ZPEs does not exceed 10 cm^{-1}. The former results confirm the reliability of the rDSD/3F12 model, whereas the second result permits confident use of B3/SVP ZPEs for larger molecules, like tryptophan.

Despite the accessible relative energy of the species δIIgg and ϵIIg$^-$g (δIIa and ϵIIb according to the nomenclature of ref 45) only the most stable species ϵIIgg$^-$ (ϵIIa according to the nomenclature of ref 45) has been detected in the experimental MW study.[45] The PCS spectroscopic parameters are given in Table 4 together with their experimental counterparts.

Table 4. Ground-State Rotational Constants (B_a^0, B_b^0, and B_c^0 in MHz) and ^{14}N-Nuclear Quadrupole Coupling Constants (χ in MHz) of the ϵIIgg$^-$ Structure (ϵIIa according to the Nomenclature of ref 45) of Histidine

Parameter	Experiment[a]	PCS[b]	MP2[a]
B_a^0	1847.5	1844	1839
B_b^0	831.7	834	859
B_c^0	745.9	748	770
χ_{aa}/N^δ	1.611	1.63	1.62
χ_{bb}/N^δ	−3.497	−3.61	−3.49
χ_{cc}/N^δ	1.886	1.97	1.87
χ_{aa}/N^ϵ	−0.179	−0.18	−0.18
χ_{bb}/N^ϵ	1.122	1.12	0.97
χ_{cc}/N^ϵ	−0.943	−0.93	−0.79
χ_{aa}/N^a	0.005	−0.02	0.04
χ_{bb}/N^a	2.098	2.31	2.10
χ_{cc}/N^a	−2.103	−2.29	−2.14

[a]From ref 45. [b]PCS equilibrium geometries, rDSD/3F12 properties, and B3/SVP vibrational corrections from Table 1.

Comparison of Table 4 with the results of previous investigations[45] shows that our computational approach reduces the RMUE on rotational constants by about 1 order of magnitude (0.27% at the PCS level and 2.32% from the MP2 computations reported in ref 45), which, in absolute terms, translates into a PCS maximum error of about 4 MHz (to be compared to 27 MHz at the MP2 level).

This finding shows that PCS computations are accurate enough to assign the detected conformer on the basis of rotational constants since the differences between the computed values of different structures (see Table 1 in ref 45) are larger than the maximum error. As already mentioned, the nitrogen atoms of His (N^a in the amino group of the backbone and N$^\delta$, N$^\epsilon$ in the imidazole ring of the side chain; see Figure 1) give rise to quadrupole coupling constants. The results collected in Table 4 show that the experimental and computed values are in remarkable agreement. It is also noteworthy that the PCS value of the HNC^{a}C′ dihedral angle (−21.8°) is very close to that estimated in ref 45 in order to minimize the difference between computed and experimental quadrupole coupling constants.

In agreement with previous studies,[38,93,94] the backbone of the most stable Trp conformers shows a hydrogen bond pattern of type II. Furthermore, the preferred conformation of the χ_2 dihedral angle (governing the position of the indole ring) is close to 90° (broadly referred to as *g*) and the most stable structures are characterized by the interaction of one aminic hydrogen of the backbone with the π-system of the phenyl ($\chi_1 \approx 60°$) or pyrrole ($\chi_1 \approx -60°$) ring of indole. The

https://doi.org/10.1021/acs.jpca.3c04227
J. Phys. Chem. A 2023, 127, 7534−7543

Table 5. rDSD/3F12 Relative Electronic Energies (ΔE_{rDSD}), Together with Differences between PCS and rDSD/3F12 Electronic Energies (ΔPCS), B3/SVP Harmonic Zero-Point Energies (ΔZPE_H), and Anharmonic Corrections to ZPEs (ΔZPE_{A-H}) for the Low-Energy Conformers of Tryptophan[a]

Label	ΔE_{rDSD}	ΔPCS	ΔZPE_H[b]	$\Delta ZPE_{(A-H)}$[b]	ΔH_0°[c]	ϕ	ψ	ω	χ_1	χ_2
IIgg	0	0	0	0	0	−12	14	−3	56	84
IIg$^-$g	353	37	−21	−7	362	15	−17	4	−62	109

[a]Best estimates of relative enthalpies at 0 K (ΔH_0°) and dihedral angles (see Figure 1) optimized at the PCS level are also given. The angles are in degrees, whereas all the energetic quantities are in cm^{-1}. [b]At the B3/SVP level. [c]Sum of columns 2, 3, 4, and 5.

Table 6. Experimental Ground-State Rotational Constants (B_a^0, B_b^0, and B_c^0 in MHz) and ^{14}N-Nuclear Quadrupole Coupling Constants (χ in MHz) of the Two Most Stable Tryptophan Conformers Compared with Computed Values[a]

	IIgg						IIg$^-$g			
	$^{14}N_i-^{14}N_a$		$^{15}N_i-^{14}N_a$		$^{15}N_i-^{15}N_a$		$^{15}N_i-^{14}N_a$		$^{15}N_i-^{15}N_a$	
Param.	Exp.[b]	Calc.[c]	Exp.[b]	Calc.[c]	Exp.[b]	Calc.[c]	Exp.[b]	Calc.[c]	Exp.[b]	Calc.[c]
B_a^0	1243.6	1237	1231.1	1225	1219.5	1213	1281.3	1286	1272.5	1273
B_b^0	392.5	395	392.2	394	391.3	394	333.7	344	332.4	333
B_c^0	346.9	349	345.7	348	344.3	346	287.1	288	286.3	287
χ_{aa}/N^a	0.31[d]	−0.1					−2.33	−2.34		
χ_{bb}/N^a	1.71	2.2					1.95	2.16		
χ_{cc}/N^a	−2.02	−2.1					0.38	0.18		
χ_{aa}/N^e	1.08	1.0								
χ_{bb}/N^e	1.30	1.4								
χ_{cc}/N^e	−2.38	−2.4								

[a]Rotational constants of the ^{15}N isotopomers are also reported. [b]From ref 38. [c]PCS equilibrium geometries, rDSD/3F12 properties, and B3/SVP vibrational corrections from Table 1. [d]Fixed in the fitting.

Table 7. rDSD/3F12 Relative Electronic Energies (ΔE_{rDSD}) and Harmonic Zero-Point Energies (ΔZPE_H), Together with Differences between PCS and rDSD/3F12 Electronic Energies (ΔPCS) and B3/SVP Anharmonic Corrections to ZPEs (ΔZPE_{A-H}) for the Low-Energy Structures of Proline[a]

Label	ΔE_{rDSD}	ΔPCS	ΔZPE_H[b]	$\Delta ZPE_{(A-H)}$[c]	ΔH_0°[d]	ψ	ω	τ
IIE$^-$	0	0	0	0	0	2	−1	−90
IIE$^+$	217	−17	−15 (−14)	9	194	1	2	88
IE$^-$	625	−64	−105 (−126)	9	465	169	179	−88
IE$^+$	605	−25	−139 (−157)	12	453	174	179	102

[a]Best estimates of relative enthalpies at 0 K (ΔH_0°) and dihedral angles (see Figure 1) optimized at the PCS level are also given. The angles are in degrees, whereas all the energetic quantities are in cm^{-1}. [b]At the rDSD/3F12 level and (in parentheses) at the B3/SVP level. [c]At the B3/SVP level. [d]Sum of columns 2, 3, 4, and 5.

structures of the two most stable conformers (IIgg and IIg$^-$g or IIb+ and IIc+ according to the nomenclature of ref 38) are shown in Figure 1, while their main features are collected in Table 5 and Table S2 of the Supporting Information. It is quite apparent that the stabilizing effect of interactions involving the phenyl ring (IIgg conformer) is considerably stronger, with respect to those in which the pyrrole ring is engaged (IIg$^-$g conformer). In analogy with the case of histidine, a single conformer (IIgg) was initially detected in MW studies. However, analysis of other isotopologues pointed out the presence of a second less abundant species. In fact, both ^{14}N and ^{15}N isotopologues could be observed for the nitrogen atoms of Trp (N^a in the amino group and N^e in the pyrrole ring; see Figure 1). Since our computations should be sufficiently accurate to discriminate between isotopologues, we compare in Table 6 the computed and experimental values. The agreement between theory and experiment is indeed remarkable, and the errors are once again about an order of magnitude smaller than those delivered by previous computations. As a consequence, the computed rotational constants allow the unbiased assignment of the detected species, with quadrupolar coupling constants further confirming the results.

Proline is the only natural amino acid whose side chain closes a (pyrrolidinic) cycle. Exhaustive conformational searches produced six low-energy conformers with two representatives each for the I, II, and III forms. However, the two species of type III are too unstable (more than 1000 cm^{-1} above the absolute energy minimum) to be detected in MW studies. The other four low-energy species adopt an envelope (E) arrangement with either *exo*- or *endo*-like placements of the carboxy moiety. The puckering amplitude α is always very close to 40°, and the phase angle τ is close either to 90° (E$^+$, *exo* COOH) or −90° (E$^-$, *endo* COOH)[88,95] (see Figure 1). All of those species have actually been detected in MW experiments,[43,44] and their structural and energetic parameters are given in Table 7 and Table S3 of the Supporting Information.

It is apparent that species of type II are significantly more stable than their counterparts of type I and that *exo* or *endo* placements of the carboxyl groups have comparable energies. The relative stabilities of the different species obtained at the rDSD/3F12 level are in fair agreement with their PCS

https://doi.org/10.1021/acs.jpca.3c04227
J. Phys. Chem. A 2023, 127, 7534−7543

Table 8. Experimental Ground-State Rotational Constants (B_a^0, B_b^0, and B_c^0 in MHz) and ^{14}N-Nuclear Quadrupole Coupling Constants (χ in MHz) of the Four Most Stable Proline Species Compared with Computed Values

	IE$^-$		IE$^+$		IIE$^-$		IIE$^+$	
Param.	Exp.[a]	Calc.[b]	Exp.[a]	Calc.[b]	Exp.[a]	Calc.[b]	Exp.[a]	Calc.[b]
B_a^0	3857.2	3879	4004.0	4017	3673.9	3684	3923.6	3966
B_b^0	1590.5	1573	1567.3	1556	1688.4	1682	1605.9	1583
B_c^0	1377.5	1365	1281.5	1272	1407.4	1405	1279.8	1266
χ_{aa}/N	2.47	2.71	1.14	1.24	0.88	0.97	0.04	0.16
χ_{bb}/N	1.83	1.81	2.32	2.47	−0.55	−0.54	−1.08	−1.11
χ_{cc}/N	−4.30	−4.52	−3.46	−3.71	−0.33	−0.43	1.04	0.95

[a]From ref 44. [b]PCS equilibrium geometries, rDSD/3F12 properties, and B3/SVP vibrational corrections from Table 1.

counterparts (MAX = 64 cm^{-1} and MUE = 35 cm^{-1}), which are, in turn, close (MAX = 28 cm^{-1} and MUE = 18 cm^{-1}) to the values obtained in ref 88 employing the so-called focal point analysis (FPA). Actually, the difference between rDSD/3F12 and PCS relative energies is often of the same order of magnitude as the corresponding difference between B3/SVP and rDSD/3F12 harmonic ZPEs or between harmonic and anharmonic ZPEs. This finding confirms that none of these contributions can be neglected in order to reach fully converged values.[96]

The experimental spectroscopic parameters are compared with the computed parameters in Table 8. It is quite apparent that the presence of large-amplitude puckering of the pyrrolidine ring increases the errors of the computed rotational constants with respect to those obtained for the other amino acids, with the effect being particularly strong for the B_a rotational constant of the IE$^+$ species. However, even in those circumstances, the agreement between computed and experimental rotational constants remains sufficiently good to permit the unequivocal assignment of the detected species, which is, anyway, further confirmed by quadrupolar coupling constants.

CONCLUSIONS

A general computational workflow aimed at the accurate description of the conformational landscape of flexible biomolecule building blocks has been applied to α-amino acids showing peculiar features such as tautomerism, ring puckering, or different aromatic rings. Accurate structures and relative energies are obtained by the new PCS model, which combines modern double-hybrid functionals and composite wave-function methods. In particular, geometries and spectroscopic parameters are obtained at the rDSD/3F12 level, whereas improved relative energies are computed by the PCS wave-function composite method. The agreement between computed and experimental results for histidine, tryptophan, and proline permits the unbiased interpretation of the latter in terms of well-defined stereoelectronic effects, with the synergism between intra-backbone and backbone−(side chain) non-covalent interactions playing a central role.

The above results, together with those of refs 39 and 42, provide a general panorama of natural α-amino acids. In more general terms, the reasonable cost and black-box implementation of the PCS model pave the way toward accurate studies of flexible prebiotic molecules containing a few dozen atoms also by nonspecialists.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.3c04227.

Different contributions to the relative energies and Cartesian coordinates of the PCS equilibrium geometries for all the low-energy structures discussed in the main text (PDF)

AUTHOR INFORMATION

Corresponding Author

Vincenzo Barone − *Scuola Normale Superiore di Pisa, 56126 Pisa, Italy;* orcid.org/0000-0001-6420-4107; Email: vincenzo.barone@sns.it

Authors

Lina Marcela Uribe Grajales − *Scuola Normale Superiore di Pisa, 56126 Pisa, Italy; Scuola Superiore Meridionale, 80138 Napoli, Italy*

Silvia Di Grande − *Scuola Normale Superiore di Pisa, 56126 Pisa, Italy; Scuola Superiore Meridionale, 80138 Napoli, Italy;* orcid.org/0000-0002-6550-0220

Federico Lazzari − *Scuola Normale Superiore di Pisa, 56126 Pisa, Italy;* orcid.org/0000-0003-4506-3200

Marco Mendolicchio − *Scuola Normale Superiore di Pisa, 56126 Pisa, Italy;* orcid.org/0000-0002-4504-853X

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.3c04227

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Funding from Gaussian, Inc., is gratefully acknowledged.

REFERENCES

(1) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. Accurate Structure, Thermochemistry 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.

(2) Alonso, E. R.; León, I.; Alonso, J. L. *Intra- and Intermolecular Interactions Between Non-Covalently Bonded Species*; Elsevier: 2020; pp 93−141.

(3) Simpson, H. J. J.; Marsh, R. E. The Crystal Structure of L-alanine. *Acta Crystallogr.* **1966**, *20*, 550−555.

(4) Freedman, T. B.; Diem, M.; Polavarapu, P. L.; Nafie, L. A. Vibrational Circular Dichroism in Amino Acids and Peptides. 6. Localized Molecular Orbital Calculations of the Amino Acids Carbon-

Hydrogen Stretching Vibrational Circular Dichroism in Deuterated Isotopomers of Alanine. *J. Am. Chem. Soc.* **1982**, *104*, 3343−3349.
(5) Lesarri, A.; Sanchez, R.; Cocinero, E. J.; Lopez, J. C.; Alonso, J. L. Coded Amino Acids in Gas Phase: the Shape of Isoleucine. *J. Am. Chem. Soc.* **2005**, *127*, 12952−12956.
(6) Jarmelo, S.; Lapinski, L.; Nowak, M.; Carey, P. R.; Fausto, R. Preferred Conformers and Photochemical (λ > 200 nm) Reactivity of Serine and 3,3-Dideutero-Serine in the Neutral Form. *J. Phys. Chem. A* **2005**, *109*, 5689−5707.
(7) Schols, G., Bassi, D., Buck, U., Eds. *Atomic and Molecular Beam Methods*; Oxford University Press: 1988.
(8) Alonso, J. L.; López, J. C. *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*; Springer: 2015; pp 335−401.
(9) Godfrey, P. D.; Brown, R. D. Proportions of Species Observed in Jet Spectroscopy-Vibrational Energy Effects: Histamine Tautomers and Conformers. *J. Am. Chem. Soc.* **1998**, *120*, 10724−10732.
(10) Florio, G. M.; Christie, R. A.; Jordan, K. D.; Zwier, T. S. Conformational Preferences of Jet-Cooled Melatonin: Probing trans- and cis-Amide Regions of the Potential Energy Surface. *J. Am. Chem. Soc.* **2002**, *124*, 10236−10247.
(11) Lee, K. T.; Sung, J.; Lee, K. J.; Park, Y. D.; Kim, S. K. Conformation-Dependent Ionization Energies of L-Phenylalanine. *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 4114−4117.
(12) Helgaker, T.; Klopper, W.; Tew, D. P. Quantitative Quantum Chemistry. *Mol. Phys.* **2008**, *106*, 2107−2143.
(13) 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.
(14) Karton, A. A Computational Chemist's Guide to Accurate Thermochemistry for Organic Molecules. *WIRES, Comp. Mol. Sci.* **2016**, *6*, 292−310.
(15) Kesharwani, M. K.; Karton, A.; Martin, J. M. Benchmark Ab Initio Conformational Energies for the Proteinogenic Amino Acids Through Explicitly Correlated Methods. Assessment of Density Functional Methods. *J. Chem. Theory Comput.* **2016**, *12*, 444−454.
(16) 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.
(17) Wang, P.; Shu, C.; Ye, H.; Biczysko, M. Structural and Energetic Properties of Amino Acids and Peptides Benchmarked by Accurate Theoretical and Experimental Data. *J. Phys. Chem. A* **2021**, *125*, 9826−9837.
(18) Mancini, G.; Fusè, M.; Lazzari, F.; Chandramouli, B.; Barone, V. Unsupervised Search of Low-Lying Conformers With Spectroscopic Accuracy: A Two-Step Algorithm Rooted Into the Island Model Evolutionary Algorithm. *J. Chem. Phys.* **2020**, *153*, No. 124110.
(19) Ferro-Costas, D.; Mosquera-Lois, I.; Fernandez-Ramos, A. Torsiflex: an Automatic Generator of Torsional Conformers. Application to the Twenty Proteinogenic. *Amino Acids. J. Cheminf.* **2021**, *13*, 100.
(20) Chandramouli, B.; Del Galdo, S.; Fusè, M.; Barone, V.; Mancini, G. Two-Level Stochastic Search of Low-Energy Conformers For Molecular Spectroscopy: Implementation and Validation of MM and QM Models. *Phys. Chem. Chem. Phys.* **2019**, *21*, 19921−19934.
(21) Barone, V.; Lupi, J.; Salta, Z.; Tasinato, N. Development and Validation of a Parameter-Free Model Chemistry for the Computation of Reliable Reaction Rates. *J. Chem. Theory Comput.* **2021**, *17*, 4913−4928.
(22) Mancini, G.; Fusè, M.; Lazzari, F.; Barone, V. Fast Exploration of Potential Energy Surfaces With a Joint Venture of Quantum Chemistry, Evolutionary Algorithms and Unsupervised Learning. *Digit. Discovery* **2022**, *1*, 790−805.
(23) Penocchio, E.; Piccardo, M.; Barone, V. Semiexperimental Equilibrium Structures for Building Blocks of Organic and Biological Molecules: The B2PLYP Route. *J. Chem. Theory Comput.* **2015**, *11*, 4689−4707.
(24) Ceselin, G.; Barone, V.; Tasinato, N. Accurate Biomolecular Structures by the Nano-LEGO Approach: Pick the Bricks and Build Your Geometry. *J. Chem. Theory Comput.* **2021**, *17*, 7290−7311.
(25) Puzzarini, C.; Biczysko, M.; Barone, V.; Largo, L.; Pena, I.; Cabezas, C.; Alonso, J. L. Accurate Characterization of the Peptide Linkage in the Gas Phase: a Joint Quantum-Chemistry and Rotational Spectroscopy Study of the Glycine Dipeptide Analogue. *J. Phys. Chem. Lett.* **2014**, *5*, 534−540.
(26) 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.
(27) Lupi, J.; Alessandrini, S.; Barone, V.; Puzzarini, C. junChS and junChS-F12 Models: Parameter-free Efficient yet Accurate Composite Schemes for Energies and Structures of Noncovalent Complexes. *J. Chem. Theory Comput.* **2021**, *17*, 6974−6992.
(28) Gyevi-Nagy, L.; Kallay, M.; Nagy, P. R. Accurate Reduced-Cost CCSD(T) Energies: Parallel Implementation, Benchmarks and Large-Scale Applications. *J. Chem. Theory Comput.* **2021**, *17*, 860−878.
(29) Kallay, M.; Horvath, R. A.; Gyevi-Nagy, L.; Nagy, P. R. Size-Consisten Explicitly Correlated Triple Excitation Correction. *J. Chem. Phys.* **2021**, *155*, No. 034107.
(30) Nagy, P. R.; Gyevi-Nagy, L.; Lorincz, B. D.; Kallay, M. Pursuing the Basis Set Limit of CCSD(T) Non-Covalent Interaction Energies for Medium-Sized Complexes: Case Study on the S66 Compilation. *Mol. Phys.* **2023**, *121*, No. e2109526.
(31) Barone, V.; Ceselin, G.; Fusè, M.; Tasinato, N. Accuracy Meets Interpretability for Computational Spectroscopy by Means of Hybrid and Double-Hybrid Functionals. *Front. Chem.* **2020**, *8*, No. 584203.
(32) Lesarri, E. J.; Cocinero, J. C.; Lòpez, J. C.; Alonso, J. L. The Shape of Neutral Valine. *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 605−610.
(33) Cocinero, J. C.; Lesarri, E. J.; Grabow, J. U.; Lòpez, J. C.; Alonso, J. L. The Shape of Leucine in the Gas Phase. *ChemPhysChem* **2007**, *8*, 599−604.
(34) Barone, V.; Di Grande, S.; Puzzarini, C. Toward Accurate Yet Effective Computations of Rotational Spectroscopy Parameters for Biomolecule Building Blocks. *Molecules* **2023**, *28*, 913.
(35) Blanco, S.; Sanz, M. E.; López, J. C.; Alonso, J. L. Revealing the Multiple Structures of Serine. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 20183−20188.
(36) Alonso, J. L.; Perez, C.; Eugenia Sanz, M.; Lopez, J. C.; Blanco, S. Seven Conformers of l-Threonine in the Gas Phase: a LA-MB-FTMW Study. *Phys. Chem. Chem. Phys.* **2009**, *11*, 617−627.
(37) Cabezas, C.; Varela, M.; Peña, I.; Mata, S.; Lòpez, J. C.; Alonso, J. L. The Conformational Locking of Asparagine. *Chem. Commun.* **2012**, *48*, 5934−5936.
(38) Sanz, M. E.; Cabezas, C.; Mata, S.; Alonso, J. L. Rotational Spectrum of Tryptophan. *J. Chem. Phys.* **2014**, *140*, No. 204308.
(39) Barone, V.; Fusè, M.; Lazzari, F.; Mancini, G. Benchmark Structures and Conformational Landscapes of Amino Acids in the Gas Phase: a Joint Venture of Machine Learning, Quantum Chemistry, and Rotational Spectroscopy. *J. Chem. Theory Comput.* **2023**, *19*, 1243−1260.
(40) Perez, C.; Mata, S.; Blanco, S.; Lopez, J. C.; Alonso, J. L. Jet-Cooled Rotational Spectrum of Laser-Ablated Phenylalanine. *J. Phys. Chem. A* **2011**, *115*, 9653−9657.
(41) Perez, C.; Mata, S.; Cabezas, C.; López, J. C.; Alonso, J. L. The Rotational Spectrum of Tyrosine. *J. Phys. Chem. A* **2015**, *119*, 3731−3735.
(42) Barone, V.; Fusè, M. Accurate Structures and Spectroscopic Parameters of Phenylalanine and Tyrosine in the Gas Phase: A Joint Venture of DFT and Composite Wave-Function Methods. *J. Phys. Chem. A* **2023**, *127*, 3648−3657.
(43) Lesarri, A.; Mata, S.; Cocinero, E. J.; Blanco, S.; López, J. C.; Alonso, J. L. The Structure of Neutral Proline. *Angew. Chem., Int. Ed. Engl.* **2002**, *41*, 4673−4676.
(44) Mata, S.; Vaquero, V.; Cabezas, C.; Pena, I.; Perez, C.; Lopez, J. C.; Alonso, J. L. Observation of Two New Conformers of Neutral Proline. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4141−4144.

https://doi.org/10.1021/acs.jpca.3c04227
J. Phys. Chem. A 2023, 127, 7534−7543

(45) Bermúdez, C.; Mata, S.; Cabezas, C.; Alonso, J. L. Tautomerism in Neutral Histidine. *Angew. Chem., Int. Ed. Engl.* **2014**, *126*, 11195−11198.
(46) Bannwarth, C.; Ehlert, S.; Grimme, S. GFN2-xTB, an Accurate and Broadly Parametrized Self-Consistent Tight-Binding Quantum Chemical Method With Multipole Electrostatics and Density-Dependent Dispersion Contributions. *J. Chem. Theory Comput.* **2019**, *15*, 1652−1671.
(47) Lazzari, F.; Salvadori, A.; Mancini, G.; Barone, V. Molecular Perception for Visualization and Computation: The Proxima Library. *J. Chem. Inf. Model.* **2020**, *60*, 2668−2672.
(48) Becke, A. D. Density-Functional Exchange-Energy Approximation With Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(49) 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*, No. 154104.
(50) Santra, G.; Sylvetsky, N.; Martin, J. M. L. Minimally Empirical Double-Hybrid Functionals Trained against the GMTKN55 Database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123*, 5129−5143.
(51) Dunning, T. H.; Peterson, K. A.; Wilson, A. K. Gaussian Basis Sets for Use in Correlated Molecular Calculations. X. The Atoms Aluminum Through Argon Revisited. *J. Chem. Phys.* **2001**, *114*, 9244−9253.
(52) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. Perspectives on Basis Sets Beautiful: Seasonal Plantings of Diffuse Basis Functions. *J. Chem. Theory Comput.* **2011**, *7*, 3027−3034.
(53) 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.
(54) Kang, Y. K.; Park, H. S. Assessment of CCSD(T), MP2, DFT-D, CBS-QB3, and G4(MP2) Methods for Conformational Study of Alanine and Proline Dipeptides. *Chem. Phys. Lett.* **2014**, *600*, 112−117.
(55) Kang, Y. K.; Park, H. S. Exploring Conformational Preferences of Alanine Tetrapeptide by CCSD(T), MP2, and Dispersion-Corrected DFT Methods. *Chem. Phys. Lett.* **2018**, *702*, 69−75.
(56) Godfrey, P. D.; Brown, R. D.; Rodgers, F. M. The Missing Conformers of Glycine and Alanine: Relaxation in Seeded Supersonic Jets. *J. Mol. Struct.* **1996**, *376*, 65−81.
(57) Ruoff, R. S.; Klots, T. D.; Emilsson, T.; Gutowsky, H. S. Relaxation of Conformers and Isomers in Seeded Supersonic Jets of Inert Gases. *J. Chem. Phys.* **1990**, *93*, 3142−3150.
(58) Puzzarini, C.; Stanton, J. F.; Gauss, J. Quantum-Chemical Calculation of Spectroscopic Parameters for Rotational Spectroscopy. *Int. Rev. Phys. Chem.* **2010**, *29*, 273−367.
(59) Liévin, J.; Demaison, J.; Herman, M.; Fayt, A.; Puzzarini, C. Comparison of the Experimental, Semi-Experimental and Ab Initio Equilibrium Structures of Acetylene: Influence of Relativisitic Effects and of the Diagonal Born−Oppenheimer Corrections. *J. Chem. Phys.* **2011**, *134*, No. 064119.
(60) Pulay, P.; Meyer, W.; Boggs, J. E. Cubic Force Constants and Equilibrium Geometry of Methane from Hartree-Fock and Correlated Wavefunctions. *J. Chem. Phys.* **1978**, *68*, 5077−5085.
(61) Piccardo, M.; Penocchio, E.; Puzzarini, C.; Biczysko, M.; Barone, V. Semi-Experimental Equilibrium Structure Determinations by Employing B3LYP/SNSD Anharmonic Force Fields: Validation and Application to Semirigid Organic Molecules. *J. Phys. Chem. A* **2015**, *119*, 2058−2082.
(62) Puzzarini, C.; Heckert, M.; Gauss, J. The Accuracy of Rotational Constants Predicted by High-Level Quantum-Chemical Calculations. I. Molecules Containing First-Row Atoms. *J. Chem. Phys.* **2008**, *128*, No. 194108.
(63) Heim, Z. N.; Amberger, B. K.; Esselman, B. J.; Stanton, J. F.; Woods, R. C.; McMahon, R. J. Molecular Structure Determination: Equilibrium Structure of Pyrimidine (m−$C_4H_4N_2$) From Rotational Spectroscopy (r_e^{SE}) and High-Level Ab Initio Calculation (r_e) Agree Within the Uncertainty of Experimental Measurement. *J. Chem. Phys.* **2020**, *152*, No. 104303.
(64) Gardner, M. B.; Westbrook, B. R.; Fortenberry, R. C.; Lee, T. J. Highly-Accurate Quartic Force Fields for the Prediction of Anharmonic Rotational Constants and Fundamental Vibrational Frequencies. *Spectrochim. Acta, Part A* **2021**, *248*, No. 119184.
(65) Watrous, A. G.; Westbrook, B. R.; Fortenberry, R. F12-TZ-cCR: a Methodology for Faster and Still Highly Accurate Quartic Force Fields. *J. Phys. Chem. A* **2021**, *125*, 10532−10540.
(66) Alonso, E. R.; Fusè, M.; Leòn, 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.
(67) Melli, A.; Tonolo, F.; Barone, V.; Puzzarini, C. Extending the Applicability of the Semi-experimental Approach by Means of Template Molecule and Linear Regression Models on Top of DFT Computations. *J. Phys. Chem. A* **2021**, *125*, 9904−9916.
(68) Peterson, K. A.; Adler, T. B.; Werner, H.-J. Systematically Convergent Basis Sets for Explicitly Correlated Wavefunctions: the Atoms H, He, B−Ne, and Al−Ar. *J. Chem. Phys.* **2008**, *128*, No. 084102.
(69) 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.
(70) Barone, V. Accuracy Meets Feasibility for the Structures and Rotational Constants of the Molecular Bricks of Life: a Joint Venture of DFT and Wave-Function Methods. *J. Phys. Chem. Lett.* **2023**, *14*, 5883−5890.
(71) Gordy, W.; Cook, R. L.; Weissberger, A. *Microwave Molecular Spectra*; Wiley: New York, 1984; Vol. *18*.
(72) Hait, D.; Head-Gordon, M. How Accurate is Density Functional Theory at Predicting Dipole Moments? An Assessment Using a New Database of 200 Benchmark Values. *J. Chem. Theory Comput.* **2018**, *14*, 1969−1981.
(73) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*; Cambridge University Press: 2009.
(74) 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.
(75) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-Set Convergence of Correlated Calculations on Water. *J. Chem. Phys.* **1997**, *106*, 9639−9646.
(76) Schuurman, M. S.; Allen, W. D.; Schaefer, H. F. The Ab Initio Limit Quartic Force Field of BH3. *J. Comput. Chem.* **2005**, *26*, 1106−1112.
(77) 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**, *17*, 4332−4358.
(78) Mendolicchio, M.; Bloino, J.; Barone, V. Perturb-Then-Diagonalize Vibrational Engine Exploiting Curvilinear Internal Coordinates. *J. Chem. Theory Comput.* **2022**, *18*, 7603−7619.
(79) Barone, V. Characterization of the Potential Energy Surface of the HO2Molecular System by a Density Functional Approach. *J. Chem. Phys.* **1994**, *101*, 10666−10676.
(80) Barone, V. Accurate Vibrational Spectra of Large Molecules by Density Functional Computations Beyond the Harmonic Approximation: the Case of Pyrrole and Furan. *Chem. Phys. Lett.* **2004**, *383*, 528−532.
(81) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gn theory. *WIRES Comput. Mol. Sci.* **2011**, *1*, 810−825.
(82) Ma, Q.; Werner, H. J. Explicitly Correlated Local Coupled-Cluster Methods Using Pair Natural Orbitals. *WIREs Comput. Mol. Sci.* **2018**, *8*, No. e1371.

https://doi.org/10.1021/acs.jpca.3c04227
J. Phys. Chem. A 2023, 127, 7534−7543

(83) Nagy, P. R.; Kallay, M. Approaching the basis set limit of CCSD(T) energies for large molecules with local natural orbital coupled-cluster methods. *J. Chem. Theory Comput.* **2019**, *15*, 5275−5298.
(84) Liakos, D. G.; Guo, Y.; Neese, F. Comprehensive benchmark results for the domain based local pair natural orbital coupled cluster method (DLPNO-CCSD(T)) for closed- and open-shell systems. *J. Phys. Chem. A* **2020**, *124*, 90−100.
(85) 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.01; Gaussian, Inc.: Wallingford, CT, 2016.
(86) Giuliano, B. M.; Bizzocchi, L.; Pietropolli Charmet, A.; Arenas, B. E.; Steber, A. L.; Schnell, M.; Caselli, P.; Harris, B. J.; Pate, B. H.; Guillemin, J. C.; Belloche, A. Rotational Spectroscopy of Imidazole: Improved Rest Frequencies for Astrophysical Searches. *Astron. Astrophys.* **2019**, *628*, A53.
(87) Nesvadba, R.; Studecky, T.; Uhlikova, T.; Urban, S. Microwave Spectrum and Molecular Constants of Indole. *J. Mol. Spectrosc.* **2017**, *339*, 6−11.
(88) Czinki, E.; Császár, A. G. Conformers of Gaseous Proline. *Chemistry, Eur. J.* **2003**, *9*, 1008−1019.
(89) Cremer, D.; Pople, J. A. General definition of ring puckering coordinates. *J. Am. Chem. Soc.* **1975**, *97*, 1354−1358.
(90) Paoloni, L.; Rampino, S.; Barone, V. Potential-Energy Surfaces for Ring-Puckering Motions of Flexible Cyclic Molecules through Cremer−Pople Coordinates: Computation, Analysis, and Fitting. *J. Chem. Theory Comput.* **2019**, *15*, 4280−4294.
(91) Yannacone, S. F.; Sethio, D.; Kraka, E. Quantitative Assessment of Intramolecular Hydrogen Bonds in Neutral Histidine. *Theor. Chem. Acc.* **2020**, *139*, 125.
(92) Snoek, L. C.; Robertson, E. G.; Kroemer, R. T.; Simons, J. P. Conformational Landscapes in Amino Acids: Infrared and Ultraviolet Ion-Dip Spectroscopy of Phenylalanine in the Gas Phase. *Chem. Phys. Lett.* **2000**, *321*, 49−56.
(93) Huang, Z.; Lin, Z. Detailed Ab Initio Studies of the Conformers and Conformational Distributions of Gaseous Tryptophan. *J. Phys. Chem. A* **2005**, *109*, 2656−2659.
(94) Kaczor, A.; Reva, I. D.; Proniewicz, L. M.; Fausto, R. Matrix-Isolated Monomeric Tryptophan: Electrostatic Interactions as Non-trivial Factors Stabilizing Conformers. *J. Phys. Chem. A* **2007**, *111*, 2957−2965.
(95) Allen, W. D.; Czinki, E.; Császár, A. G. Molecular Structure of Proline. *Chemistry, Eur. J.* **2004**, *10*, 4512−4517.
(96) Barone, V.; Lupi, J.; Salta, Z.; Tasinato, N. Reliable Gas Phase Reaction Rates at Affordable Cost by Means of the Parameter-Free JunChS-F12 Model Chemistry. *J. Chem. Theory Comput.* **2023**, *19*, 3526−3537.

https://doi.org/10.1021/acs.jpca.3c04227
J. Phys. Chem. A 2023, 127, 7534−7543