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

pubs.acs.org/JPCA

Article

# Accurate Structures and Spectroscopic Parameters of Phenylalanine and Tyrosine in the Gas Phase: A Joint Venture of DFT and Composite Wave-Function Methods

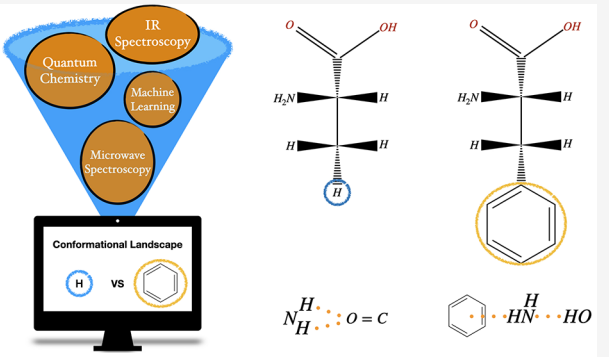
Vincenzo Barone* and Marco Fusè

Cite This: J. Phys. Chem. A 2023, 127, 3648−3657

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

**ABSTRACT:** A general strategy for the accurate computation of conformational and spectroscopic properties of flexible molecules in the gas phase is applied to two representative proteinogenic amino acids with aromatic side chains, namely, phenylalanine and tyrosine. The main features of all the most stable conformers predicted by this computational strategy closely match those of the species detected in microwave and infrared experiments. Together with their intrinsic interest, the accuracy of the results obtained with reasonable computer times paves the route for accurate investigations of other flexible bricks of life.






## INTRODUCTION

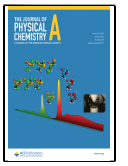
Amino acids represent central targets for accurate experimental and theoretical studies because their rich conformational landscape is tuned by the competition among different kinds of noncovalent interactions.[1,2] While zwitterionic forms of amino acids are found in crystals[3] and aqueous solutions,[4] neutral forms are observed in the gas phase[5−7] and in low-temperature inert matrices.[8,9] In the latter case, van der Waals forces between the sample and inner gas molecules have usually a negligible effect on spectroscopic parameters,[10] so high-resolution spectroscopic studies both in the gas phase (microwave, MW) and in inert matrices (infrared, IR) allow an unbiased disentanglement of intrinsic stereoelectronic effects without any strong perturbation from the environment.

Systematic investigations have revealed that the natural amino acids containing simple nonpolar side chains present two dominant conformers (usually referred to as type I and type II), stabilized by intrabackbone hydrogen bonds.[7,11,12] On the other hand, the conformational landscape of natural amino acids with polar side chains is much richer due to the synergy or competition between intrabackbone and backbone−(side chain) hydrogen bonds.[13−16] Here we will analyze in detail the situation for prototypical amino acids containing aromatic side chains (phenylalanine (Phe) and tyrosine (Tyr)), which can show an additional kind of noncovalent interaction between polar hydrogen atoms of the backbone and the $\pi$-system of the side chain.

Phenylalanine and tyrosine have been studied in the gas phase by high-resolution laser-induced fluorescence (LIF), hole-burning UV−UV, ion dip IR−UV, resonance-enhanced multi-photon ionization (REMPI) spectroscopy,[6,17−22] and, more recently, by microwave (MW) spectroscopy in combination with supersonic jets.[23,24] Phenylalanine has been studied also by matrix-isolation infrared spectroscopy.[10] These experiments have been interpreted in terms of a variable number of low-energy conformers due to the presence of effects (e.g., fast relaxation of some structures to more stable counterparts in the presence of low energy barriers and photodissociation of some products) playing a different role under different experimental conditions.[20,25,26]

Accurate quantum chemical (QC) computations can help to solve this kind of problems,[27−29] but the size of the systems and the number of different structures to be analyzed prevent a systematic use of very accurate but very expensive state-of-the-art QC methodologies.[30−32] Furthermore, an effective exploration of flat potential energy surfaces (PES) cannot be performed by straightforward systematic searches and/or local optimizations.[33,34] We have recently developed an integrated computational approach combining different QC methods driven by machine learning (ML) tools for the effective exploration of conformational PES and the successive refinement of the most significant stationary points.[35−37] In this framework, once a

Received: February 20, 2023
Revised: April 3, 2023
Published: April 13, 2023



suitable panel of low-energy minima has been defined, accurate structures are obtained by refining the optimized geometries by a linear regression approach[38,39] and accurate relative energies are computed by reduced-scaling composite methods.[40−45] The zero point energies (ZPEs) and thermal contributions leading from electronic to free energies are evaluated by parameter-free anharmonic approaches rooted in the second order vibrational perturbation theory (VPT2)[46−53] for sufficiently stiff degrees of freedom and proper treatment of hindered rotations.[54,55] Finally, accurate spectroscopic parameters of the energy minima with non negligible populations under the experimental conditions of interest are computed.[56]

This procedure has been recently validated for several amino acids containing polar side chains.[57−59] Here we tackle the problem of aromatic side chains analyzing both the rotational and vibrational spectra of the prototypical phenylalanine and tyrosine amino acids. In fact, previous studies of these amino acids employed QC methods of limited accuracy, payed marginal attention to the geometrical parameters and neglected vibrational corrections beyond the standard rigid rotor harmonic oscillator (RRHO) model. However, these limitations hamper any *a priori* prediction of the spectroscopic outcome, allowing at most its *a posteriori* interpretation in terms of the agreement between experimental and computed spectroscopic parameters for a predefined number of conformers. Our approach allows, instead, an unbiased comparison with spectroscopic results thanks to the accuracy of the computational results, which provide mean unsigned errors (MUEs) within 10 MHz for rotational constants and 10 $cm^{-1}$ for both relative energies and vibrational frequencies.[41,60]

## METHODS

A first characterization of low-energy conformers is performed at the B3LYP/jun-cc-pVDZ level,[61,62] also including Grimme's D3BJ dispersion corrections.[63] In the following this computational model is referred to simply as B3 and is used also for the computation of anharmonic contributions. The geometries and harmonic force fields of conformers lying within a predefined energy range are then refined by the revDSD-PBEP86-D3BJ/jun-cc-pv(T+d)Z model[64−67] (hereafter rDSD), which provides excellent conformational landscapes,[30,68,69] geometries,[39] dipole moments,[70] and spectroscopic parameters.[56,71] In the present context, we will analyze in detail only conformers not involved in fast relaxation processes and with a relative free energy below 620 $cm^{-1}$ at room temperature (i.e., a relative population of about 5% when only the considered conformer and the absolute free energy minimum are taken into account, since $kT/hc = 207$ $cm^{-1}$).

The typical MUEs of rDSD bond lengths and valence angles (0.3%)[39] are sufficient to obtain accurate relative electronic energies and vibrational spectra of different conformers, but the situation is different for the leading terms of MW spectra, namely, rotational constants of the vibrational ground state ($B_0^i$, where $i$ refers to the inertial axes $a$, $b$, $c$). These parameters include vibrational corrections ($\Delta B_{vib}^i$) in addition to equilibrium rotational constants ($B_e^i$).[72] In the framework of the VPT2 approximation,[73] the ground-state rotational constants can be expressed as

$$B_0^i = B_e^i + \Delta B_{vib}^i = B_e^i - \frac{1}{2}\sum_r \alpha_r^i \quad (1)$$

where the $\alpha_r$'s are the vibration−rotation interaction constants and the sum runs over all $r$ vibrational modes (for details, see refs 31, 74, and 75). With $\Delta B_{vib}^i$ contributions typically well below 1% of the corresponding $B_e^i$ rotational constants,[76] they can be safely determined at affordable levels of theory (B3 in the present context), which deliver typical errors within 10% (i.e., less than 0.1% of typical rotational constants).[31,77] On the other hand, inclusion of vibrational corrections is not warranted if the errors on the computed rotational constants are not much lower than 1% (10 MHz for a constant of 1000 MHz). Therefore, equilibrium rotational constants require very accurate geometrical parameters, which can be obtained only with state-of-the-art QC methods,[78−81] although reasonable relative errors (typically 0.4−0.5%) are obtained at the rDSD level.[39] We have recently shown that the systematic nature of the errors permits to improve significantly the rDSD geometrical parameters, and thus equilibrium rotational constants, by a linear regression approach (hereafter LRA).[39,82,83] In this model, the computed geometrical parameters ($r_{comp}$) are corrected for systematic errors by means of scaling factors ($a$) and offset values ($b$) depending on the nature of the involved atoms and determined once for ever from a large database of accurate semiexperimental (SE) equilibrium geometries:[39,59,83]

$$r_{LRA} = a \times r_{comp} + b \quad (2)$$

Since at least one $^{14}N$ quadrupolar nucleus is present in all amino acids, nuclear quadrupole coupling constants ($\chi_{ii}$, with $i$ referring to the inertia axis $a$, $b$, or $c$) are important for accurate predictions of rotational spectra because they determine a splitting of the rotational transitions, which generates the so-called hyperfine structure.[78,84] Furthermore, the components of dipole moments ($\mu_i$) determine the intensities of rotational transitions.[74,84] Several studies have shown that dipole moments and quadrupole coupling constants can be computed with sufficient accuracy at the rDSD level.[59,70] On the other hand, accurate electronic energies can be obtained by single-point energy evaluations at rDSD geometries using composite wave function methods rooted in the coupled cluster (CC) ansatz.[85] In particular, the CC model including single, double, and perturbative estimate of triple excitations (CCSD(T))[86] is employed here taking also into account complete basis set (CBS) extrapolation and core valence (CV) correlation. The starting point of the family of "cheap" schemes (ChS) developed in the last years[28,36,40,41] is a frozen core (fc) CCSD(T) computation in conjunction with a (partially augmented) triple-$\zeta$ basis set.[62,66,87] Then, in analogy with the correlation consistent composite approach (ccCA),[88,89] CBS and CV terms are computed with good accuracy and negligible additional cost by means of second order Møller−Plesset perturbation theory (MP2).[90] In particular, the CBS extrapolation by the standard $n^{-3}$ two-point formula[91] employs MP2/jun-cc-pV(X+d)Z energies with X = T and Q, whereas the CV contribution is incorporated as the difference between all-electron (ae) and fc MP2 calculations, both with the cc-pCVW(T+d)Z basis set.[92] Replacement of conventional methods by the explicitly correlated (F12) variants[93,94] leads to our current standard version of the approach, which is referred to as junChS-F12.[42,58,60] Comparison with the most accurate results available for a panel of representative non-covalent complexes provided an average absolute error smaller than 10 $cm^{-1}$.[41,42,60]

The relative stability of different conformers is determined by the corresponding relative enthalpy at 0 K ($\Delta H_0^\circ$) or free energy



https://doi.org/10.1021/acs.jpca.3c01174
J. Phys. Chem. A 2023, 127, 3648−3657

($\Delta G°$) at a temperature depending on the experimental conditions. The vibrational contributions to thermodynamic functions of stiff degrees of freedom are computed in the framework of second-order vibrational perturbation theory (VPT2),[46,48,50,73] employing rDSD harmonic and B3 anharmonic contributions. The same computations provide also anharmonic IR spectra.[53,95,96] Low-frequency contributions are, instead, taken into account by means of the quasi-harmonic (QH) approximation.[54,97]

All the computations have been performed with the Gaussian package[98] except the junChS-F12 and QH ones, which have been performed with the help of the Molpro[99] and GoodVibes[97] software, respectively.

## RESULTS AND DISCUSSION

The conformational landscape of Phe and Tyr is ruled by two soft degrees of freedom in the backbone ($\phi$ = H−N−C$^{\alpha}$−C′ and $\psi$ = N−C$^{\alpha}$−C′−O(H) dihedral angles) and two others in the side chain ($\chi_1$ = NC$^{\alpha}$−C$^{\beta}$−C$^{\gamma}$ and $\chi_2$ = C$^{\alpha}$−C$^{\beta}$−C$^{\gamma}$−C$^{\delta}$ dihedral angles) (see Figure 1). However, the nonplanarity of the $NH_2$

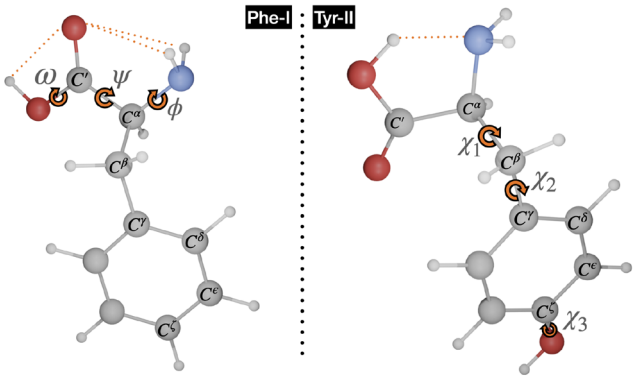



**Figure 1.** Structure and dihedral angles of phenylalanine (conformer of type I) and tyrosine (conformer of type II).

moiety suggests to replace the $\phi$ dihedral angle with $\phi'$ = LP−N−C$^{\alpha}$−C′ = $\phi$ + 120°, where LP is the nitrogen lone-pair perpendicular to the plane defined by the two amine hydrogens and the C$^{\alpha}$ atom. The customary *c*, *g*, *s*, *t* labels are then used to indicate the cis, gauche, skew, and trans conformations determined by $\phi'$, $\psi$, $\chi_1$, and $\chi_2$ dihedral angles.

Only planar conformations are allowed for the carboxylic moiety of both amino acids ($\omega$ = C$^{\alpha}$−C′−O−H = 0° or 180°) and the phenol group of tyrosine ($\chi_3$ = C$^{\epsilon}$−C$^{\zeta}$−O−H = 0° or 180°). In the former case $\omega$ = 0° is strongly preferred (due the formation of a weak hydrogen bridge with the carbonyl oxygen) except when the oxidryl hydrogen is involved in stronger hydrogen bonds with other acceptor groups. In the case of $\chi_3$ the two nonequivalent arrangements of the phenol OH moiety are labeled in the following as N and C with reference to the H(O) placement on the same side of the backbone N or C′ atom, respectively.

The exploration of the conformational landscape of phenylalanine provided several low-energy conformers stabilized by the formation of hydrogen bonds, classified as type I (bifurcated $NH_2$···O=C, $\phi' \approx 180°$, $\psi \approx 180°$, $\omega \approx 180°$), I′ (HNH···O=C, $\phi' \approx 90°$, $\psi \approx 180°$, $\omega \approx 180°$), and II (N···HO, $\phi' \approx 0°$, $\psi \approx 0°$, $\omega \approx 0°$).[100] Some conformers in which the O(H) oxygen of the carboxyl group forms a bifurcated hydrogen bond with the $NH_2$ moiety (type III, $\phi' \approx 180°$, $\psi \approx 0°$, $\omega \approx 180°$) have been also found, but they can easily relax to more stable I conformers by rotation of 180° around $\psi$. In fact, III conformers have been observed in MW spectra only in very special situations.[101] Also some conformers of type I relax to more stable structures (still of type I) through rotation around the $\chi_1$ dihedral angle. As a consequence, we are left with the six conformers shown in Figure 2, with the two I′ conformers corresponding to the nonequivalent g and g$^-$ orientations of the $\phi'$ dihedral angle (see Table 1).

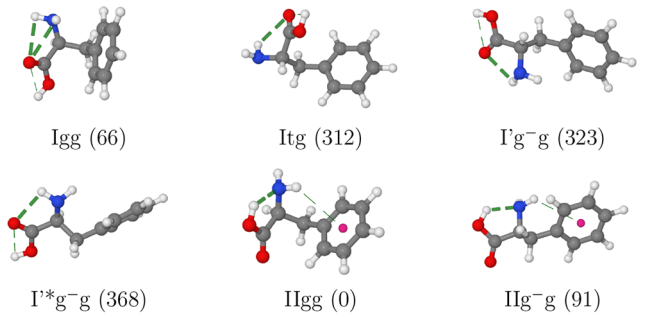



**Figure 2.** Conformers of phenylalanine with the computed relative free energies at room temperature (in cm$^{-1}$) given in parentheses. H-bonds are highlighted by dashed lines.

The structural and energetic characteristics of those six low-energy conformers are reported in Table 1, whereas the different contributions to the final electronic energies are given in Table S1 and the main spectroscopic parameters are given in Tables S2 and S3. The first noteworthy feature is the remarkable accuracy of rDSD conformational energies (MUE and maximum unsigned error (MAX) of 21.5 and 47.7 cm$^{-1}$ with respect to the junChS-F12 reference). Although rDSD electronic energies are not directly used to estimate relative stabilities, this finding gives further support to the use of rDSD geometries and harmonic force fields. Different staggered conformations are populated for $\chi_1$, which determines the relative position of the phenyl ring with respect to the amino acid backbone (see Figure 1). On the other hand, all low-energy conformers show a C$^{\alpha}$−C$^{\beta}$ bond nearly perpendicular to the plane of the phenyl ring ($\chi_2 \approx 90°$) and this feature has been confirmed experimentally by ENDOR spectroscopy for spin-labeled L-phenylalanine.[102] Interestingly, the same region of $\chi_2$ is dominantly populated and characteristic of Phe residues in proteins.[103]

Contrary to the usual situation for aliphatic amino acids,[59] conformers of type II are more stable than their types I and I′ counterparts in spite of a less favorable orientation of the OH group in the carboxyl moiety. The increased stability of these conformers is related to the existence of a "daisy chain" sequence of interactions involving aromatic $\pi$ electrons, together with the $NH_2$ and OH groups of the backbone.[19] Zero-point and thermal effects stabilize all conformers with respect to the IIgg absolute energy minimum, with this effect being particularly significant for conformers of types I and I′. As a consequence, conformer Igg becomes the second most stable form, with a relative free energy (with respect to IIgg) marginally lower than that of the IIg$^-$g conformer. Inclusion of anharmonic contributions in ZPEs is needed for obtaining quantitative results, but it does not alter the stability order of the different conformers. Finally, the main effect of the QH corrections is to reduce the overestimation of entropic contributions produced by the harmonic oscillator approximation, with the consequent excessive stabilization of all


https://doi.org/10.1021/acs.jpca.3c01174
J. Phys. Chem. A 2023, 127, 3648−3657

**Table 1. rDSD Relative Electronic Energies ($\Delta E_{rDSD}$), Harmonic Zero Point Energies ($\Delta ZPE_H$), Thermal Contributions ($\Delta Th_H$), and Quasi-Harmonic Corrections to Free Energies ($T\Delta S_{QH-H}$), Together with Differences between junChS-F12 and rDSD Electronic Energies ($\Delta ChS$) and B3 Anharmonic Corrections to ZPEs ($\Delta ZPE_{anh-H}$) for the Low-Lying Conformers of Phenylalanine[a]**

| label | $\Delta E_{rDSD}$ | $\Delta ChS$ | $\Delta ZPE_H$ | $\Delta Th_H$ | $\Delta ZPE_{anh-H}$ | $T\Delta S_{QH-H}$ | $\Delta G°$[b] | $\phi'$ | $\psi$ | $\omega$ | $\chi_1$ | $\chi_2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Igg | 317.3 | −4.6 | −152.1 | −142.2 | 49.4 | −2.2 | 65.6 | −178.4 | 174.3 | −178.9 | 62.6 | 89.6 |
| Itg | 596.9 | 36.6 | −180.8 | −236.2 | 41.3 | 54.2 | 312.0 | 173.5 | 129.7 | 177.4 | 179.9 | 81.4 |
| I′g⁻g | 660.4 | −11.0 | −190.7 | −206.1 | 43.1 | 26.9 | 322.6 | −74.4 | 151.9 | 176.2 | −60.4 | 99.1 |
| I′*g⁻g | 762.1 | −47.7 | −212.8 | −210.3 | 73.0 | 3.5 | 367.8 | 80.2 | −169.2 | −177.2 | −61.9 | 93.9 |
| IIgg | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | −26.1 | 12.2 | −2.6 | 52.7 | 81.0 |
| IIg⁻g | 176.9 | 7.5 | −31.6 | −92.8 | 31.6 | −0.9 | 90.7 | 33.8 | −18.0 | 4.0 | −63.2 | 103.3 |

[a]Best estimates of relative free energies at room temperature ($\Delta G°$) and dihedral angles optimized at the rDSD level ($\phi'$, $\psi$, $\omega$, $\chi_1$, and $\chi_2$) are also given. All the energetic quantities are in $cm^{-1}$ and the angles in degrees. [b]Sum of columns 2−7.

**Table 2. Ground-State Rotational Constants ($A_0$, $B_0$, and $C_0$ in MHz) and $^{14}N$-Nuclear Quadrupole Coupling Constants ($\chi$ in MHz) of the Two Most Stable Phenylalanine Conformers[a]**

| | IIgg | | | | IIg⁻g | | | |
|---|---|---|---|---|---|---|---|---|
| | $^{14}N$ | | $^{15}N$ | | $^{14}N$ | | $^{15}N$ | |
| param. | exp.[b] | calc.[c] | exp.[b] | calc.[c] | exp.[b] | calc.[c] | exp.[b] | calc.[c] |
| $A_0$ | 1666.0436(14) | 1658.3 | 1646.7381(17) | 1639.2 | 2457.05490(48) | 2456.6 | 2425.69(30) | 2425.7 |
| $B_0$ | 638.56314(12) | 641.3 | 636.56611(19) | 639.4 | 460.659722(79) | 461.1 | 459.64825(32) | 460.2 |
| $C_0$ | 568.76843(15) | 571.1 | 569.83617(20) | 568.3 | 424.74604(13) | 424.5 | 423.71892(25) | 423.6 |
| $\chi_{aa}$ | −0.283(16) | −0.403 | | | −0.777(22) | −0.505 | | |
| $\chi_{bb}$ | 1.275(55) | 1.479 | | | 0.0695(36) | −0.112 | | |
| $\chi_{cc}$ | −0.992(39) | −1.076 | | | 0.7075(14) | 0.617 | | |

[a]Rotational constants of the $^{15}N$ isotopomers are also reported. [b]From ref 23. [c]rDSD-LRA equilibrium geometries, rDSD properties and B3 vibrational corrections.

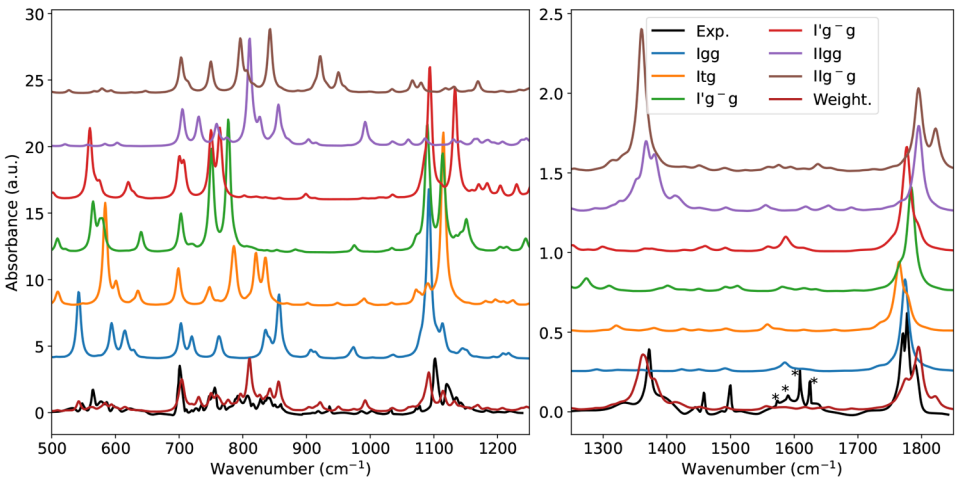



**Figure 3.** Comparison between the FT-IR spectrum of matrix-isolated phenylalanine (Exp) and the computed anharmonic spectra for the six most stable conformers and their Boltzmann average (Weight.) in the 500−1250 $cm^{-1}$ and 1250−1800 $cm^{-1}$ regions. See main text for details.

conformers with respect to the IIgg species. The final results suggest that Igg and IIg⁻g conformers have comparable populations (about 23% each at room temperature) only slightly lower than that of the IIgg conformer (about 33% at room temperature), whereas the other three conformers have equilibrium populations around 7% each at room temperature and their unequivocal characterization could be, therefore, quite difficult.

The rotational constants are very sensitive to the value of the $\chi_1$ dihedral angle, with g conformers showing values around 1650, 650, and 450 MHz, whereas g⁻ conformers show values around 2450, 450, and 400 MHz (see Tables 2 and S2). On the other hand, much larger dipole moments are computed for conformers of type II with respect to those of type I, irrespective of the orientation of the side chain (see Table S3). These general trends were used to support the assignment of experimental MW spectra[23] and are not very sensitive to the level of the QC computations. However, comparison of Table 2 (and Table S2) with the results of previous investigations[23] shows that our computational approach strongly improves the quantitative agreement between the computed and experimental rotational constants of the detected conformers concerning both the parent species and the $^{15}N$ isotopomers (MUE = 2.2 MHz and MAX = 7.7 MHz).

Despite the accessible relative energies of conformers of types I and I′′ (especially Igg), none of them has been detected in the


https://doi.org/10.1021/acs.jpca.3c01174
J. Phys. Chem. A 2023, 127, 3648−3657

MW study of ref 23, contrary to the conclusions issued from electronic and IR spectra.[10,19] As a matter of fact, the laser energy needed to obtain MW spectra is much higher than that required by other spectroscopic techniques. This leads to an increased photofragmentation of Phe, with this explaining the weakness of the observed spectra.[23] At the same time, both theory and experiments agree in indicating that conformers of type II have higher ionization energies than their I and I′ counterparts.[20] Thus, if the non observed conformers are preferentially ionized in the laser ablation process, their populations in the supersonic jet would be much lower than those estimated from a Boltzmann distribution.

Based on these premises, the IR spectra in inert matrices should show the signatures of conformers of types I, I′, II, and, possibly, III. However, in the most relevant spectral regions (mainly OH, C═O and C−O stretching together with COH bending regions) conformers of type III are predicted to absorb at frequencies very close to those of the corresponding type II structures. Therefore, they cannot be characterized by this technique too. Computed and experimental spectra in the mid-IR region are compared in Figure 3. The general agreement between theory and experiment confirms that different conformers are present in the experimental sample and that their relative populations are correctly predicted by our computations. The most distinctive features of the spectra concern the regions of C−O (1060−1160 $cm^{-1}$) and C═O (1750−1800 $cm^{-1}$) stretchings. The presence of multiple bands in the former region proves that several conformers of type I and I′ are present in the matrix. In the latter region, the bands originating from conformers of type I and I′ fall at lower wave numbers than those of type II conformers due to the involvement of the O═C group in a hydrogen bond with the amine hydrogens. However, this kind of hydrogen bond is not very strong, so that the wavenumber shift is quite limited.

The NH stretching vibrations are observed in the experimental spectrum as two broad bands of quite low intensity, centered at about 3340 and 3410 $cm^{-1}$, and the computed values are in good agreement with this finding for all low-lying conformers (see Table 3). In the case of I and I′ conformers these vibrations correspond to symmetric and antisymmetric combinations of the amine NH stretching modes.

**Table 3. Computed Wavenumbers for the NH and OH Stretching Vibrations of Phenylalanine and Tyrosyne Low-Energy Conformers[a]**

| phenylalanine | | | | |
|---|---|---|---|---|
| assign. | Igg | Itg | I′g⁻g | I′*g⁻g |
| $\nu NH_{sym}$ | 3325.7 | 3318.1 | 3340.8 | 3364.7 |
| $\nu NH_{asym}$ | 3403.9 | 3406.6 | 3421.7 | 3444.3 |
| $\nu OH$ | 3572.1 | 3563.6 | 3572.6 | 3579.3 |
| | IIgg | IIg⁻g | | |
| $\nu NH_{sym}$ + $\nu OH$ in phase | 3217.8 | 3268.8 | | |
| $\nu NH_{sym}$ + $\nu OH$ out of phase | 3339.7 | 3325.9 | | |
| $\nu NH_{asym}$ | 3427.6 | 3412.0 | | |
| tyrosine | | | | |
| assign. | IICgg | IINgg | IICg⁻g | IINg⁻g |
| $\nu NH_{sym}$ + $\nu OH$ in phase | 3208.5 | 3213.6 | 3243.9 | 3255.9 |
| $\nu NH_{sym}$ + $\nu OH$ out-of-phase | 3338.6 | 3337.5 | 3364.5 | 3364.5 |
| $\nu NH_{asym}$ | 3428.1 | 3426.6 | 3409.9 | 3408.9 |
| $\nu OH$ phenol | 3652.3 | 3661.0 | 3659.3 | 3661.1 |

[a]See main text for details.

However, for conformers of type II, the band above 3400 $cm^{-1}$ is due to the antisymmetric $NH_2$ stretching, but the band at about 3330 $cm^{-1}$ can be assigned to the out-of-phase combination of NH and OH stretchings. The broad experimental band centered at 3178 $cm^{-1}$ is assigned to the in phase combination of NH and OH stretchings of conformers of type II, which is computed slightly above 3200 $cm^{-1}$. Finally, the OH stretching of conformers of type I and I′ is computed between 3564 and 3579 $cm^{-1}$ and found experimentally between 3557 and 3567 $cm^{-1}$. In more general terms, a remarkable quantitative agreement with experiment is obtained for the positions of all the computed IR bands without employing any adjustable parameter.

The six Phe conformers discussed above give rise to 12 conformers of tyrosine (Tyr) due to the already mentioned presence of two nonequivalent orientations of the OH group in the para position of the phenyl ring. All these conformers are true energy minima, but, for the same reasons discussed above for Phe, we can expect that only species of type II are detected in MW experiments (see Figure 4). The structural and energetic characteristics of those four low-energy conformers are collected in Table 4, whereas the different contributions to the final electronic energies are given in Table S4. The situation envisaged by our computations was indeed confirmed by MW experiments, but, unfortunately, for conformers IINg⁻g and IICg⁻g too few lines could be assigned to allow an unequivocal characterization.[24]

We are thus left with the IINgg and IICgg conformers, whose computed spectroscopic parameters are compared to their experimental counterparts in Table 5 (see also Table S5). For purposes of completeness, the computed spectroscopic parameters of conformers IINg⁻g and IICg⁻g are given in the same tables. The agreement between experimental and computed rotational constants is remarkable, and even more importantly, all the trends and differences between the two conformers are reproduced quantitatively. The percentage MUE for the rotational constants is 0.38% and the percentage MAX 0.56%. These values are comparable to those obtained for the IIgg conformer of Phe and are more than three times smaller than those issued from the MP2 computations reported in ref 24 (percentage MUE = 1.42%, percentage MAX = 1.88%). Larger errors affect the quadrupolar coupling constants, but the agreement between computed and experimental values is largely sufficient to allow an unbiased assignment of different conformers.

The spectra of tyrosine in the mid-IR region computed taking into account all the type II conformers are shown in Figure 5. Although experimental spectra would probably show some characteristic signatures of conformers of types I and I′ (as discussed above for phenylalanine), in the present context we prefer to avoid any reference to species not well characterized experimentally.

The spectra are very similar to those of the corresponding conformers of phenylalanine except for the presence of a quite intense band at about 1200 $cm^{-1}$ (1197, 1193, 1198, and 1194 $cm^{-1}$ for conformer IICgg, IINgg, IICgg⁻, and IINgg⁻, respectively) assigned to the bending of the phenyl OH moiety. The same trend is observed also in the NH/OH stretching region, where an additional band due to the phenol OH stretching is found at about 3650 $cm^{-1}$ (see Table 3). A more detailed analysis is not pursued because it would be very similar to that given above for phenylalanine.



https://doi.org/10.1021/acs.jpca.3c01174
J. Phys. Chem. A 2023, 127, 3648−3657

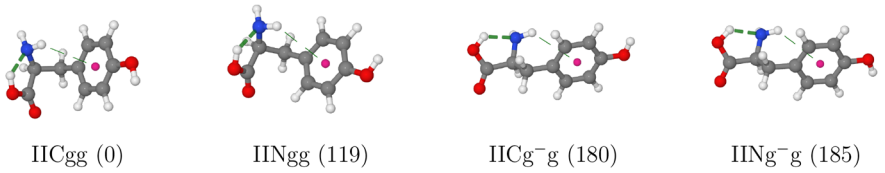



**Figure 4.** Conformers of tyrosine with the computed relative free energies at room temperature (in $cm^{-1}$) given in parentheses. H-bonds are highlighted by dashed lines.

**Table 4. rDSD Relative Electronic Energies ($\Delta E_{rDSD}$), Harmonic Zero Point Energies ($\Delta ZPE_H$), Thermal Contributions ($\Delta Th_H$), and Quasi-Harmonic Corrections to Free Energies ($T\Delta S_{QH-H}$), Together with Differences between junChS-F12 and rDSD Electronic Energies ($\Delta ChS$) and B3 Anharmonic Corrections to ZPEs ($\Delta ZPE_{anh-H}$) for the Low-Lying Conformers of Tyrosine[a]**

| label | $\Delta E_{rDSD}$ | $\Delta ChS$ | $\Delta ZPE_H$ | $\Delta Th_H$ | $\Delta ZPE_{anh-H}$ | $T\Delta S_{QH-H}$ | $\Delta G°$[b] | $\phi'$ | $\psi$ | $\omega$ | $\chi_1$ | $\chi_2$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| IICgg[c] | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 | −26.32 | 12.33 | −2.55 | 53.08 | 80.97 |
| IINgg[d] | 117.5 | 3.3 | −6.8 | −9.5 | 10.7 | 4.0 | 119.2 | −26.18 | 12.34 | −2.64 | 52.72 | 80.42 |
| IICg⁻g[c] | 278.0 | 11.5 | −32.9 | −101.9 | −10.4 | 35.6 | 179.9 | 33.92 | −18.25 | 4.09 | −62.64 | 104.25 |
| IINg⁻g[d] | 275.7 | 10.8 | −37.3 | −99.7 | −0.9 | 36.7 | 185.3 | 34.06 | −18.33 | 4.08 | −63.05 | 102.84 |

[a]Best estimates of relative free energies at room temperature ($\Delta G°$) and dihedral angles optimized at the rDSD level ($\phi'$, $\psi$, $\omega$, $\chi_1$, and $\chi_2$) are also given. All the energetic quantities are in $cm^{-1}$ and the angles in degrees. [b]Sum of columns 2−7. [c]$\chi_3 = 0°$. [d]$\chi_3 = 180°$.

**Table 5. Ground-State Rotational Constants ($A_0$, $B_0$, and $C_0$ in MHz), $^{14}N$-Nuclear Quadrupole Coupling Constants ($\chi$ in MHz), and Electric Dipole Moment Components ($\mu$ in debye) of the Two Observed Tyrosine Conformers[a]**

| parameter | $IINgg_{exp}$[b] | $IINgg_{calc}$[c] | $IICgg_{exp}$[b] | $IICgg_{calc}$[c] | $IINg^-g_{calc}$[c] | $IICg^-g_{calc}$[c] |
|---|---|---|---|---|---|---|
| $A_0$ | 1529.6791(40) | 1522.2 | 1525.2543(29) | 1517.6 | 2400.9 | 2404.3 |
| $B_0$ | 463.94021(32) | 464.5 | 465.48173(25) | 467.7 | 341.0 | 340.7 |
| $C_0$ | 425.76168(40) | 427.1 | 427.31023(27) | 429.4 | 321.1 | 321.6 |
| $\chi_{aa}$ | 0.709(14) | 0.5819 | 0.740(15) | 0.6111 | −0.5451 | −0.5301 |
| $\chi_{bb}$ | 0.236(88) | 0.3022 | 0.247(92) | 0.3134 | −0.1324 | −0.2353 |
| $\chi_{cc}$ | −0.945(88) | −0.8832 | −0.988(92) | −0.9245 | 0.6775 | 0.7654 |
| $\mu_a$ | | 2.0713 | | 1.4039 | 4.6739 | 4.7805 |
| $\mu_b$ | | 5.7026 | | 3.3171 | −0.7791 | −3.0797 |
| $\mu_c$ | | 0.9601 | | 0.4933 | 2.1722 | 1.2795 |

[a]The spectroscopic parameters computed for the IING⁻g and IING⁻g conformers are also reported. [b]From ref 24. [c]rDSD-LRA equilibrium geometries, rDSD properties and B3 vibrational corrections.

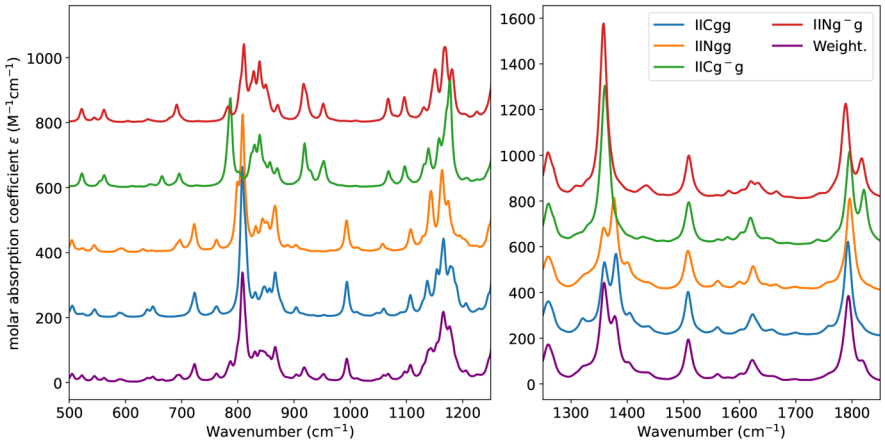



**Figure 5.** Computed anharmonic spectra for the four conformers and their Boltzmann average (Weight.) in the 500−1250 $cm^{-1}$ and 1250−1800 $cm^{-1}$ regions. See main text for details.

## CONCLUSIONS

In this paper, a general strategy aimed to the unbiased disentanglement of the conformational bath of flexible biomolecule building blocks has been applied to prototypical $\alpha$-amino acids containing aromatic side chains. Accurate structures and relative energies are obtained by the rDSD-LRA approach and the junChS-F12 composite method, respectively. Next, the spectroscopic parameters of sufficiently populated conformers can be safely computed at the rDSD level. The results obtained for phenylalanine and tyrosine are in full agreement with the available spectroscopic data, and permit their unbiased interpretation in terms of the cooperation or


https://doi.org/10.1021/acs.jpca.3c01174
J. Phys. Chem. A 2023, 127, 3648−3657

competition between intrabackbone and backbone-(side chain) hydrogen bonds.

Together with the intrinsic interest of the studied molecules, the results of the present investigation show that highly reliable analysis of structural and spectroscopic features is today possible for flexible building blocks of biomolecules in the gas phase.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.3c01174.

Additional computed data for phenylalanine and tyrosine, cartesian coordinates of rDSD optimized geometries for all the low-energy conformers of phenylalanine and tyrosine (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

### Author

**Marco Fusè** − *DMMT-sede Europa, Università di Brescia, 25121 Brescia, Italy;* orcid.org/0000-0003-0130-5175

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.3c01174


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

Funding from the Italian Ministry of University and Research (MUR, Grant 2017A4XRCA) and Italian Space Agency (ASI, "Life in Space" project no. 2019-3-U.0) 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. The Role of the Intramolecular Interactions in the Structural Behavior of Biomolecules: Insights from Rotational Spectroscopy. In *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) 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.

(6) Lee, Y.; Jung, J.; Kim, B.; Butz, P.; Snoek, L. C.; Kroemer, R. T.; Simons, J. P. Alanyl Side Chain Folding in Phenylalanine: Conformational Assignment Through Ultraviolet Rotational Band Contour Analysis. *J. Phys. Chem. A* **2004**, *108*, 69−73.

(7) Lesarri, A.; Sanchez, R.; Cocinero, E. J.; López, J. C.; Alonso, J. L. Coded Amino Acids in Gas Phase: the Shape of Isoleucine. *J. Am. Chem. Soc.* **2005**, *127*, 12952−12956.

(8) Lambie, B.; Ramaekers, R.; Maes, G. 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.

(9) Jarmelo, S.; Lapinski, L.; Nowak, M.; Carey, P. R.; Fausto, R. 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.

(10) Kaczor, A.; Reva, I. D.; 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.

(11) Lesarri, E. J.; Cocinero, J. C.; López, J. C.; Alonso, J. L. The Shape of Neutral Valine. *Angew. Chem., Int. Ed. Engl.* **2004**, *43*, 605−610.

(12) Cocinero, J. C.; Lesarri, E. J.; Grabow, J. U.; López, J. C.; Alonso, J. L. The Shape of Leucine in the Gas Phase. *ChemPhysChem* **2007**, *8*, 599−604.

(13) 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.

(14) Alonso, J. L.; Pérez, M.; Eugenia Sanz, M.; López, J. C.; Blanco, S. Seven Conformers of l-Threonine in the Gas Phase: a LA-MB-FTMW Study. *Phys. Chem. Chem. Phys.* **2009**, *11*, 617−627.

(15) Cabezas, C.; Varela, M.; Peña, I.; Mata, S.; López, J. C.; Alonso, J. L. The Conformational Locking of Asparagine. *Chem. Commun.* **2012**, *48*, 5934−5936.

(16) Sanz, M. E.; Cabezas, C.; Mata, S.; Alonso, J. L. Rotational Spectrum of Tryptophan. *J. Chem. Phys.* **2014**, *140*, 204308.

(17) Martinez, S. j.; Alfano, J. C.; Levy, D. H. The Electronic Spectroscopy of the Amino Acids Tyrosine and Phenylalanine in a Supersonic Jet. *J. Mol. Spectrosc.* **1992**, *156*, 421−430.

(18) Martinez, S. j.; Alfano, J. C.; Levy, D. H. The Electronic Spectroscopy of Tyrosine and Phenylalanine in a Supersonic Jet: Basic Analogs. *J. Mol. Spectrosc.* **1993**, *158*, 82−92.

(19) 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.

(20) 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.

(21) Inokuchi, Y.; Kobayashi, Y.; Ito, T.; Ebata, T. Conformation of L-Tyrosine Studied by Fluorescence-Detected UV-UV and IR-UV Double-Resonance Spectroscopy. *J. Phys. Chem. A* **2007**, *111*, 3209−3215.

(22) Shimozono, Y.; Yamada, K.; Ishiuchi, S.; Tsukiyama, K.; Fujii, M. Revised Conformational Assignments and Conformational Evolution of Tyrosine by Laser Desorption Supersonic Jet Laser Spectroscopy. *Phys. Chem. Chem. Phys.* **2013**, *15*, 5163−5175.

(23) 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.

(24) Perez, C.; Mata, S.; Cabezas, C.; Lopez, J. C.; Alonso, J. L. The Rotational Spectrum of Tyrosine. *J. Phys. Chem. A* **2015**, *119*, 3731−3735.

(25) 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.

(26) 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.

(27) Helgaker, T.; Klopper, W.; Tew, D. P. Quantitative Quantum Chemistry. *Mol. Phys.* **2008**, *106*, 2107−2143.

(28) 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.

(29) Karton, A. A Computational Chemist's Guide to Accurate Thermochemistry for Organic Molecules. *WIRES, Comp. Mol. Sci.* **2016**, *6*, 292−310.
(30) 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.
(31) 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.
(32) 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.
(33) 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*, 124110.
(34) 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.
(35) 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.
(36) 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.
(37) 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.
(38) 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.
(39) 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.
(40) 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.
(41) 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.
(42) 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.
(43) Gyevi-Nagy, L.; Kállay, 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.
(44) Kállay, M.; Horvath, R. A.; Gyevi-Nagy, L.; Nagy, P. R. Size-Consisten Explicitly Correlated Triple Excitation Correction. *J. Chem. Phys.* **2021**, *155*, 034107.
(45) 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.* **2022**, *120*, e2109526.
(46) Papousek, D.; Aliev, M. R. *Molecular Vibrational-Rotational Spectra*; Elsevier Scientific Publishing Company, 1982.
(47) Gaw, F.; Willetts, A.; Handy, N.; Green, W. In *Advances in Molecular Vibrations and Collision Dynamics*; Bowman, J. M., Ed.; JAI Press, 1992; Vol. *1*; pp 186−195.
(48) Clabo, D. A., Jr.; Allen, W. D.; Remington, R. B.; Yamaguchi, Y.; Schaefer, H. F., III A Systematic Study of Molecular Vibrational Anharmonicity and Vibration-Rotation Interaction by Self-Consistent-Field Higher-Derivative Methods. Asymmetric Top Molecules. *Chem. Phys.* **1988**, *123*, 187−239.
(49) Burcl, R.; Carter, S.; Handy, N. C. On the Representation of Potential Energy Surfaces of Polyatomic Molecules in Normal Coordinates: II. Parameterisation of the Force Field. *Chem. Phys. Lett.* **2003**, *373*, 357−365.
(50) Barone, V. Anharmonic Vibrational Properties by a Fully Automated Second Order Perturbative Approach. *J. Chem. Phys.* **2005**, *122*, 014108.
(51) Rosnik, A. M.; Polik, W. F. VPT2+K Spectroscopic Constants and Matrix Elements of the Transformed Vibrational Hamiltonian of a Polyatomic Molecule With Resonances Using Van Vleck Perturbation Theory. *Mol. Phys.* **2014**, *112*, 261−300.
(52) 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**, *125*, 1301−1324.
(53) Mendolicchio, M.; Bloino, J.; Barone, V. Perturb-Then-Diagonalize Vibrational Engine Exploiting Curvilinear Internal Coordinates. *J. Chem. Theory Comput.* **2022**, *18*, 7603−7619.
(54) Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. *Chemistry, Europ. J.* **2012**, *18*, 9955−9964.
(55) Li, S.-C.; Lin, Y.-C.; Li, Y.-P. Comparative Analysis of Uncoupled Mode Approximations for Molecular Thermochemistry and Kinetics. *J. Chem. Theory Comput.* **2022**, *18*, 6866−6877.
(56) 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*, 584203.
(57) León, I.; Fusè, M.; Alonso, E. R.; Mata, S.; Mancini, G.; Puzzarini, C.; Alonso, J. l.; Barone, V. Unbiased Disentanglement of Conformational Baths With the Help of Microwave Spectroscopy, Quantum Chemistry, and Artificial Intelligence: The Puzzling Case of Homocysteine. *J. Chem. Phys.* **2022**, *157*, 074107.
(58) Barone, V.; Fusè, M.; Aguado, R.; Potenti, S.; León, I.; Alonso, E. R.; Mata, S.; Lazzari, F.; Mancini, G.; Spada, L. Bringing Machine-Learning Enhanced Quantum Chemistry and Microwave Spectroscopy to Conformational Landscape Exploration: The Paradigmatic Case of 4-Fluoro-Threonine. *Chem. - Eur. J.* **2023**, e202203990.
(59) 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.
(60) Barone, V.; Di Grande, S.; Puzzarini, C. Toward Accurate Yet Effective Computations of Rotational Spectroscopy Parameters for Biomolecule Building Blocks. *Molecules* **2023**, *28*, 913.
(61) Becke, A. D. Density-Functional Exchange-Energy Approximation With Correct Asymptotic Behavior. *Phys. Rev. A* **1988**, *38*, 3098−3100.
(62) 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.
(63) 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*, 154104.
(64) 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.
(65) 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.


https://doi.org/10.1021/acs.jpca.3c01174
*J. Phys. Chem. A* 2023, 127, 3648−3657

(66) 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.
(67) 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.
(68) 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.
(69) 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.
(70) 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.
(71) Zapata Trujillo, J. C.; McKemmish, L. K. Model Chemistry Recommendations for Scaled Harmonic Frequency Calculations: A Benchmark Study. *J. Phys. Chem. A* **2023**, *127* (7), 1715−1735.
(72) 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.
(73) Mills, I. M. Vibration−Rotation Structure in Asymmetric- and Symmetric-Top Molecules. *Mol. Spectrosc.: Mod. Res.* **1972**, *1*, 115−140.
(74) Puzzarini, C.; Stanton, J. F.; Gauss, J. Quantum-Chemical Calculation of Spectroscopic Parameters for Rotational Spectroscopy. *Int. Rev. Phys. Chem.* **2010**, *29*, 273−367.
(75) 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*, 064119.
(76) Puzzarini, C.; Stanton, J. F. Connections Between the Accuracy of Rotational Constants and Equilibrium Molecular Structures. *Phys. Chem. Chem. Phys.* **2023**, *25*, 1421−1429.
(77) 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.
(78) Puzzarini, C.; Heckert, M.; Gauss, J. The Accuracy of Rotational Constants Predicted by High-Level Quantum-Chemical Calculations. I. Molecules Containing First-Row Atoms. *J. Chem. Phys.* **2008**, *128*, 194108.
(79) 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*, 104303.
(80) 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*, 119184.
(81) 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.
(82) 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.
(83) 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.
(84) Gordy, W.; Cook, R. L.; Weissberger, A. *Microwave Molecular Spectra*; Wiley: New York, 1984; Vol. *18*.
(85) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*; Cambridge University Press, 2009.
(86) 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.
(87) 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.
(88) DeYonker, N. J.; Cundari, T. R.; Wilson, A. K. The Correlation Consistent Composite Approach (ccCA): an Alternative to the Gaussian-n Methods. *J. Chem. Phys.* **2006**, *124*, 114104.
(89) DeYonker, N. J.; Wilson, B. R.; Pierpont, A. W.; Cundari, T. R.; Wilson, A. K. Towards the Intrinsic Error of the Correlation Consistent Composite Approach (ccCA). *Mol. Phys.* **2009**, *107*, 1107−1121.
(90) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618−622.
(91) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-Set Convergence of Correlated Calculations on Water. *J. Chem. Phys.* **1997**, *106*, 9639−9646.
(92) Hill, J. G.; Mazumder, S.; Peterson, K. A. Correlation Consistent Basis Sets for Molecular Core-Valence Effects With Explicitly Correlated Wave Functions: The Atoms B−Ne and Al−Ar. *J. Chem. Phys.* **2010**, *132*, 054108.
(93) Werner, H.-J.; Adler, T. B.; Manby, F. R. General Orbital Invariant MP2-F12 Theory. *J. Chem. Phys.* **2007**, *126*, 164102.
(94) Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 Methods: Theory and Benchmarks. *J. Chem. Phys.* **2009**, *130*, 054104.
(95) 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.
(96) Barone, V.; Festa, G.; Grandi, A.; Rega, N.; Sanna, N. Accurate Vibrational Spectra of Large Molecules by Density Functional Computations Beyond the Harmonic Approximation: the Case of Uracil and 2-thiouracil. *Chem. Phys. Lett.* **2004**, *388*, 279−283.
(97) Luchini, G.; Alegre-Requena, J.; Funes-Ardoiz, I.; Paton, R. GoodVibes: Automated Thermochemistry for Heterogeneous Computational Chemistry Data. *F1000Res.* **2020**, *9*, 291.
(98) 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.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16*, revision C.01; Gaussian, Inc.: Wallingford, CT, 2016.
(99) Werner, H.-J.; Knowles, P. J.; Manby, F. R.; Black, J. A.; Doll, K.; Heßelmann, A.; Kats, D.; Köhn, A.; Korona, T.; Kreplin, D. A.; et al. The Molpro Quantum Chemistry Package. *J. Chem. Phys.* **2020**, *152*, 144107.
(100) 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, 2015; pp 335−401.
(101) Jimenéz, A. I.; Vaquero, V.; Cabezas, C.; López, J. C.; Cativiela, C.; Alonso, J. L. The Singular Gas-Phase Structure of 1-Amino-cyclopropanecarboxylic Acid (Ac3c). *J. Am. Chem. Soc.* **2011**, *133*, 10621−10628.


https://doi.org/10.1021/acs.jpca.3c01174
J. Phys. Chem. A 2023, 127, 3648−3657

(102) Joela, H.; Mustafi, D.; Fair, C. C.; Makinen, M. W. Structure and Conformation of Spin-Labeled Methyl L-Phenylalaninate and L-Phenylalanine Determined by Electron Nuclear Double Resonance Spectroscopy. *J. Phys. Chem.* **1991**, *95*, 9135−9144.
(103) Duan, G. L.; Smith, V. H., Jr.; Weaver, D. F. Data Mining, Ab Initio, and Molecular Mechanics Study on Conformation of Phenylalanine and its Interaction with Neighboring Backbone Amide Groups in Proteins. *Int. J. Quantum Chem.* **2002**, *90*, 669−683.





https://doi.org/10.1021/acs.jpca.3c01174
*J. Phys. Chem. A* 2023, 127, 3648−3657