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

pubs.acs.org/JPCA

Article

Accurate Geometries of Large Molecules by Integration of the Pisa Composite Scheme and the Templating Synthon Approach

Published as part of The Journal of Physical Chemistry A virtual special issue "Gustavo Scuseria Festschrift".

Federico Lazzari,* Marco Mendolicchio, and Vincenzo Barone*

Cite This: *J. Phys. Chem. A* 2024, 128, 1385−1395

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: An effective yet reliable computational workflow is proposed, which permits the computation of accurate geometrical structures for large flexible molecules at an affordable cost thanks to the integration of machine learning tools and DFT models together with reduced scaling computations of vibrational averaging effects. After validation of the different components of the overall strategy, a panel of molecules of biological interest have been analyzed. The results confirm that very accurate geometrical parameters can be obtained at reasonable cost for molecules including up to about 50 atoms, which are the largest ones for which comparison with high-resolution rotational spectra is possible. Since the whole computational workflow can be followed employing standard electronic structure codes, accurate results for large-sized molecules can be obtained at DFT cost also by nonspecialists.

1. INTRODUCTION

Quantum chemical (QC) computations play an increasing role in complementing experimental studies since their accuracy and feasibility for large systems is constantly improving.[1−3] In this respect, reduced-cost QC methods are being developed for the accurate evaluation of electronic energies of large molecules.[4−9] Unfortunately, analytical energy derivatives are not yet available for those methods, so accurate computations of equilibrium geometries and vibrational frequencies for large molecules are still prohibitively expensive. Based on these premises, we have introduced and validated two models, which provide accurate molecular structures at the cost of geometry optimizations by a double hybrid functional in conjunction with a triple-ζ basis set followed by estimation of core−valence (CV) correlation effects by second-order Møller−Plesset perturbation theory (MP2)[10,11] or by a one-parameter empirical function.[12] The accuracy of the approach can be further increased by correcting the tendency of several methods rooted in the density functional theory (DFT) to overestimate delocalization by a simple function of Pauling bond orders[13] involving just one additional parameter.[14,15] The final model delivers an accuracy close to that of state-of-the-art composite wave function methods,[1,16−18] but can be routinely applied to much larger molecules. The purpose of the present paper is to further extend the model to molecules containing heavier elements[19] (in particular halogen atoms) and to large, flexible molecules for which the accuracy of the approach is not fully satisfactory. Some prototypical drugs, amino acids, and neurotransmitters are used to show that problematic moieties can be detected automatically, and the involved geometrical parameters can be corrected by the templating synthon approach described in the next section. Thanks to all of these improvements, the enhanced model introduced in the present paper can be safely (and automatically) employed to obtain accurate geometries and rotational constants.

2. METHODS AND COMPUTATIONAL TOOLS

While the molecular structures delivered by double-hybrid functionals in conjunction with suitable basis sets are generally quite accurate,[20−22] improved results can be obtained by the so-called linear regression approach (LRA), which employs bond-specific scaling factors[23] derived from a large database of accurate molecular structures.[21,24]

Whenever the LRA is not sufficiently accurate, one can resort to the templating molecule approach (TMA), which corrects different fragments of the target system with reference to molecules containing the same fragment, whose accurate geometries are already known.[25] While remarkable results have been obtained by combining LRA and TMA in the so-called Nano-LEGO model,[21,24,26] the use of a large number of parameters remains quite unsatisfactory and, above all, suitable templating molecules are not always available.[27,28]

Based on these premises, we have developed and validated the Pisa composite scheme (PCS),[10,12,27] which improves the

Received: December 23, 2023
Revised: January 24, 2024
Accepted: January 29, 2024
Published: February 13, 2024

https://doi.org/10.1021/acs.jpca.3c08382
J. Phys. Chem. A 2024, 128, 1385−1395

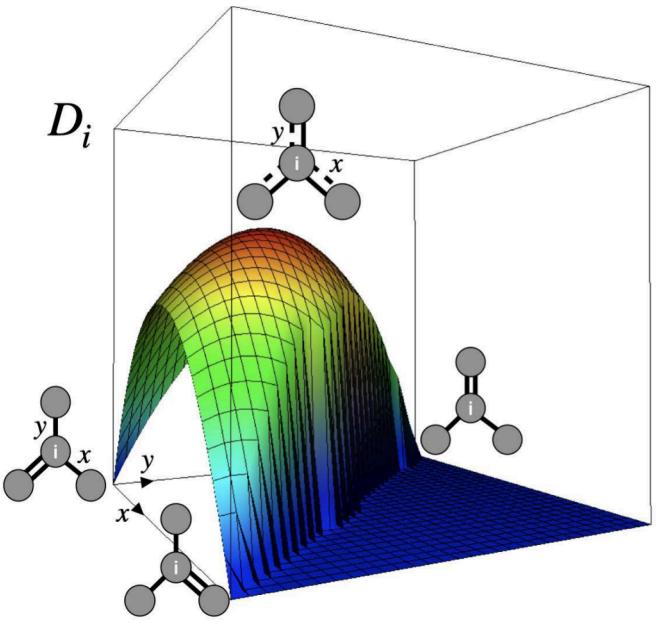

Figure 1. Delocalization feature space for a model assembly of atoms. In this case a constraint on the summation of bond orders is imposed so as to plot an analytical function. The key observation is that the highest value of the delocalization is obtained for a combination of noninteger bond orders around the central atom.

accuracy of current approaches for medium-sized molecules by nearly an order of magnitude, without any significant increase of computational cost, especially when an inexpensive one-parameter bond correction is employed (PCSB).[12,14,15]

Here, after extending the panel of molecules, which can be treated at the PCSB level, we propose to replace the TMA by the more general and robust templating synthon approach (TSA) in which for each bond the closest counterpart available in a reference database is employed to define automatically a suitable correction overcoming the problems related to hand selection or lack of a molecule containing an identical fragment.

Then, experimental rotational constants are employed to validate the computed geometries for large molecules, also thanks to the development of reduced cost procedures for the computation of vibrational corrections.

2.1. The Pisa Composite Scheme with Bond Corrections (PCSB). Previous studies[14,15] have shown that remarkably accurate geometrical parameters are delivered by the rev-DSD-PBEP86-D3BJ (hereafter rDSD) double-hybrid functional[20] in conjunction with a basis set obtained from the *s*, *p*, *d* functions of the cc-pVTZ-F12 basis set and the *f* functions of the cc-pVTZ basis set.[29−31] Furthermore, *d* functions can be neglected for H atoms. Following the same philosophy, for atoms of the fourth and fifth row of the periodic table, small-core pseudopotentials are used in conjunction with a basis set in which the *s*, *p*, *d* functions are taken from the cc-pVQZ-PP basis set and the *f* functions from the cc-pVTZ-PP basis set.[32,33] The final basis set (referred to as 3F12$^-$) is reported in the Supporting Information (SI) and has dimensions comparable with those of the jun-cc-pVTZ basis set[34] employed systematically in previous studies,[35] but delivers results much closer to those of augmented quadruple-ζ basis sets[10] which are, in turn, close to the complete basis set (CBS) limit.[36] Several studies have shown that at this level valence and dihedral angles are reproduced with remarkable accuracy, whereas bond lengths are slightly overestimated possibly due to the neglect of core−valence (CV) correlation.[10,12,14,37]

Starting from a set of atoms identified by atomic numbers and Cartesian coordinates ($\{Z_i, x_i, y_i, z_i\}$), Pauling bond orders (P_{ij})[13] can be determined from bond lengths (r_{ij}) and covalent radii (r^{cov})[38] according to

$$P_{ij} = \exp\left(\frac{r_{Zi}^{cov} + r_{Zj}^{cov} - r_{ij}}{0.3}\right) \quad (1)$$

and two atoms are considered bonded if P_{ij} is larger than 0.3. Then in the PCSB model, only bond lengths are further refined:

$$r_{ij}(\text{PCSB}) = r_{ij}(\text{rDSD/3F12}^-) + \Delta r_{ij}^b \quad (2)$$

It has been shown that Δr_{ij}^b can be expressed by a simple one-parameter function of the covalent radii and principal quantum numbers (n) of the involved atoms, together with the Pauling bond order (P_{ij}):[12,14]

$$\Delta r_{ij}^b = A_{ij}\sqrt{|P_{ij} - 2 + \Delta O(1 - P_{ij})|} \quad (3)$$

with

$$A_{ij} = -0.0011\sqrt{n_i' n_j' - 1}\,(r_{Zi}^{cov} + r_{Zj}^{cov}) \quad (4)$$

and $n_k'=max\{3,n_k\}$. An additional term can be introduced for correcting the underestimation of C−F bond hyperconjugation (Δ^{hyp}):

$$\Delta r_{ij}^{hyp} = -0.025(P_{ij} - 1)^2 \Delta F \quad (5)$$

In the above equations, $\Delta X = \delta_{iX} + \delta_{jX} - \delta_{iX}\delta_{jX}$ and δ_{xy} is a Kronecker δ. The PCSB model employs only two empirical parameters for describing any molecule at the same cost as the underlying DFT geometry optimization. All the rDSD/3F12$^-$ geometry optimizations have been performed with the Gaussian package,[39] whereas the computation of PCSB geometries has been implemented in a freely available Web site (https://www.skies-village.it/proxima/pcsbonds/), which, starting from a rDSD/3F12$^-$ optimized geometry, generates interactively the

https://doi.org/10.1021/acs.jpca.3c08382
J. Phys. Chem. A 2024, 128, 1385−1395

Figure 2. Trend of the rigidity feature for a central atom with unitary bond orders, so that the rigidity is exactly the bond order between the terminal atoms, i.e., the red line in the figure.

corresponding PCSB geometry and a 2D representation of the molecular structure.

2.2. From the Templating Molecule to the Templating Synthon Approach. The paradigm for molecular perception underlying the TSA is based on continuous rather than discrete indexes, providing a conceptual link between traditional chemical concepts and the quantum chemistry world. In particular, a new version of the PROXIMA tool[40] provides suitable continuous chemical descriptors (features) spanning different molecular topologies. In particular, a 4-dimension vector $\boldsymbol{A}$ is defined for each atom, which encodes a continuous 'atom type':

$$\mathbf{A}_i = \begin{bmatrix} D_i = \prod_k (1 + P_{ik}) \\ V_i = \sum_k \min\{1, P_{ik}\} \\ C_i = \dfrac{\sum_k Z_k P_{ik}}{\sum_k P_{ik}} \\ R_i = \sum_{m>n} P_{im} P_{in} P_{mn} \end{bmatrix} \quad (6)$$

The four elements of the $\boldsymbol{A}$ vector are the features describing different characteristics of atoms in molecules and are referred to as delocalization D_i, valence V_i, coordination C_i, and rigidity R_i, respectively. In particular, the delocalization feature has a maximum value when the bond orders of an atom are not integers (see Figure 1). The valence feature is a measure of the total number of covalent bonds in which the atom is involved, whereas the coordination feature is a weighted average of the atomic numbers of the atoms in the neighborhood of the examined center. Finally, the rigidity feature has been designed to account for situations where the atoms are packed very closely such as in the case of small tense rings and is expressed as a summation over all of the pairs of geminal atoms. As shown in Figure 2, the product between the bond orders hides an implicit dependence on the valence angles reaching very high values for small valence angles.

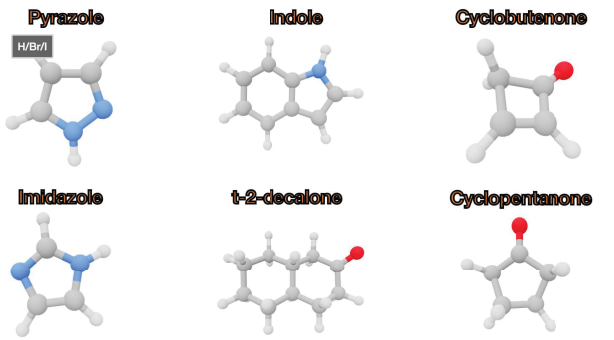

Figure 3. Prototypical semirigid molecules.

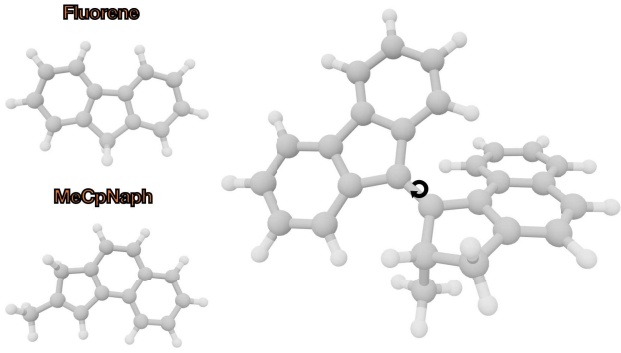

Figure 4. Molecular motor and its fragments.

The extension of such a continuous description of atom types to bond types is straightforward, taking into account that a bond is just a collection of two atoms. Thus, we can define a 9-dimension vector $\mathbf{S}_{ij}$ referred to as a synthon:

$$\mathbf{S}_{ij} = \begin{bmatrix} P_{ij} \\ \mathbf{A}_i \\ \mathbf{A}_j \end{bmatrix} \quad (7)$$

https://doi.org/10.1021/acs.jpca.3c08382
J. Phys. Chem. A 2024, 128, 1385−1395

where $\boldsymbol{A}_i$ and $\boldsymbol{A}_j$ are the 4-dimension vectors defined above and P_{ij} is the Pauling bond order. There is, of course, an ambiguity in the comparison of different bonds, unless a univocal atom ordering is defined. To this end, the traditional priority rules of stereochemistry are employed to establish a convention: the lowest priority atom is always the i atom.

When a new molecule is studied, the synthon of each bond is determined in real-time, and the corresponding templating synthon is selected as the most similar one in a reference database of high-quality molecular structures. Then, each bond length is corrected by the difference Δr_{ij}^{TS} between the PCSB and accurate bond length of the templating synthon:

$$r_{ij} = r_{ij}^{PCSB} + \Delta r_{ij}^{TS} \tag{8}$$

For the sake of simplicity, the Euclidean distance d_{tr} between two normalized synthons $\boldsymbol{s}_r$ and $\boldsymbol{s}_t$ is employed to quantify their "dissimilarity".

The main advantage of the TSA over its TMA predecessor is its fully unsupervised workflow, which does not require any molecule identical to a given fragment of the target system, Furthermore, the high quality of the PCSB starting structure permits the use of very simple search algorithms and a reference database of reduced dimensions. In particular, all the results reported in the present study have been obtained employing the following set of reference molecules, whose accurate semi-experimental structures are available in the SE127 database:[24] Alanine (conformers I and II), Uracil, Benzene, 2-Deoxyribose, Hydroxylamine, Imidazole, Phenol, Cyclohexane, HCN, Furane, Cis-Acrolein, Trans-Acrolein, Butadiene, Acetylbenzene, Benzonitrile, Pyridine, Pyrimidine, Pyrrole, HClC═O, CH_2Cl_2, CH_2F_2, CH_2Br_2, CH_3I, $CHBrF_2$, H_2C═O, H_2C═NH.

2.3. The Bottom-up Approach for the Validation of Molecular Structures. A direct comparison between experimental and computed geometrical parameters would allow a direct analysis of the performance of the PCSB and TSA models, but unfortunately, accurate experimental or semiexperimental (SE) equilibrium structures[41] are not available for the quite large molecules, which are the main targets of the present study. Under such circumstances, an unbiased evaluation of the quality of computed equilibrium geometries can be obtained by the so-called bottom-up approach, in which computed rotational constants are compared to their experimental counterparts.[42] However, the experimental rotational constants refer to the vibrational ground state (B_τ^0), so that proper comparison between theory and experiment requires the computation of electronic (ΔB_τ^{el}) and vibrational (ΔB_τ^{vib}) contributions:[41,43]

$$B_\tau^0 = B_\tau^{eq} + \Delta B_\tau^{el} + \Delta B_\tau^{vib} \tag{9}$$

with

$$\Delta B_\tau^{el} = -\frac{m_e}{m_p} g_{\tau\tau} B_\tau^{eq} \tag{10}$$

where the rotational $\boldsymbol{g}$ tensor[44] is expressed in units of the nuclear magneton, m_e is the electron mass, m_p is the proton mass, and τ is one of the principal axes of inertia. Since several tests have shown that ΔB_τ^{el} is smaller than the target accuracy of the proposed computational approach (at least for the molecules considered in the present study),[12,25] in the following this contribution will be neglected. On the other hand, the computation of ΔB_τ^{vib} requires harmonic frequencies (ω_i), Coriolis couplings ($\zeta_{ij,\tau}$),[45] and semidiagonal third derivatives of the energy with respect to normal modes (f_{iij}).[41,43,46] In the framework of vibrational second-order perturbation theory (VPT2),[47−55] ΔB_τ^{vib} can be expressed as

$$\Delta B_\tau^{vib} = \Delta B_\tau^{harm} + \Delta B_\tau^{Cor} + \Delta B_\tau^{anh} \tag{11}$$

In the above equation, ΔB_τ^{harm}, ΔB_τ^{Cor}, and ΔB_τ^{anh} represent the harmonic, Coriolis, and anharmonic contribution, respectively,

$$\Delta B_\tau^{harm} = -(B_\tau^{eq})^2 \sum_{\eta=x,y,z} \sum_{i=1}^{N} \frac{3(a_{i,\tau\eta})^2}{4\omega_i I_\eta^{eq}} \tag{12}$$

$$\Delta B_\tau^{Cor} = (B_\tau^{eq})^2 \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{(\zeta_{ij,\tau})^2(\omega_i - \omega_j)^2}{\omega_i\omega_j(\omega_i + \omega_j)} \tag{13}$$

$$\Delta B_\tau^{anh} = -(B_\tau^{eq})^2 \pi \sqrt{\frac{c}{h}} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{f_{iij} a_{j,\tau\tau}}{\omega_j^{3/2}} \tag{14}$$

where $a_{i,\tau\eta}$ is the derivative of the τ, η component of the inertia moment with respect to the i^{th} normal coordinate, c is the speed of light, h is the Planck constant, and I_η^{eq} is the principal moment of inertia along the axis η. Let us underline that at variance with the vibro-rotational interaction constants α_i, the expression of ΔB_τ^{vib} does not contain any resonant term.[56] The bottleneck of the whole computation is the ΔB^{anh} term, which contains semidiagonal cubic force constants (see eq 14) and represents indeed the largest contribution.[56] Since it is easier to increase than to decrease bond lengths and valence angles, vibrationally averaged inertia moments are usually larger than their equilibrium counterparts, and of course, just the opposite is true for rotational constants. From a practical point of view, the evaluation of VPT2 anharmonic frequencies and zero point energies for a nonlinear molecule containing N_a atoms requires analytical Hessians at $6N_a - 12$ geometries displaced along normal modes since the full cubic and semidiagonal quartic force field are required. However, if only vibrational corrections to rotational constants are needed, the semidiagonal cubic force field is sufficient, which can be obtained by a much faster approach based on $6N_a - 12$ analytical gradients at geometries displaced along the different normal modes since

$$f_{iij} = \frac{f_j(+\delta q_i) + f_j(-\delta q_j) - 2f_j(\mathbf{q}^{eq})}{\delta q_i^2} \tag{15}$$

where $\boldsymbol{q}^{eq}$ and q_i are respectively the vector collecting the dimensionless normal coordinates and its i^{th} element, δq_i is the displacement along the latter, and f_i is the corresponding analytical first-order energy derivative. In the present work, the use of this reduced-cost approach has been limited to the analysis of a specific molecule containing 47 atoms, whereas further details about the theoretical development and implementation in a new standalone code will be provided in a forthcoming paper.

In general terms, errors on rotational constants of the order of 0.1% (which correspond roughly to errors of the order of 0.001 Å for typical bond lengths and 0.05 degrees for typical valence angles) are considered a very satisfactory target even for small semirigid molecules.[57] Since vibrational corrections are about 2 orders of magnitude smaller than the corresponding equilibrium rotational constants, they cannot be neglected for reaching this accuracy. However, they can be safely computed by relatively cheap QC methods (e.g., the B3LYP/6-31G* model employed

https://doi.org/10.1021/acs.jpca.3c08382
J. Phys. Chem. A 2024, 128, 1385−1395

Table 1. Computed and Experimental Rotational Constants (in MHz) of Prototypical Semi-rigid Molecules

Molecule	Axis (τ)	B_τ^{eq}(rDSD)	ΔB_τ^{PCSB}	ΔB_τ^{vib}	B_τ^0(PCSB)	B_τ^0(Exp.)[a]
Pyrazole	*a*	9668.5	26.1	−75.6	9619.0	9625.6
	b	9462.3	29.4	−76.1	9415.6	9418.1
	c	4782.1	13.9	−39.8	4756.2	4758.7
4-Br-Pyrazole	*a*	9537.4	28.9	−79.9	9486.4	9481.1
	b	1262.5	4.4	−4.9	1262.0	1268.3
	c	1114.9	3.8	−5.2	1113.5	1118.4
4-I-Pyrazole	*a*	9550.9	29.1	−79.9	9500.1	9495.6
	b	955.2	3.4	−3.4	955.2	955.2
	c	868.3	3.1	−3.6	867.8	867.8
Cyclobutenone	*a*	12623.3	41.5	−109.8	12555.0	12548.3
	b	5313.9	20.5	−26.5	5307.9	5307.6
	c	3830.9	14.3	−24.2	3821.0	3820.2
Cyclopentanone	*a*	6662.4	24.0	−62.2	6624.2	6620.1
	b	3362.8	12.5	−25.2	3350.1	3351.5
	c	2420.3	9.1	−18.8	2410.6	2410.4
t-2-decalone	*a*	2237.0	8.4	−15.5	2229.9	2229.3
	b	763.1	2.8	−8.4	757.5	757.8
	c	609.1	2.3	−7.3	604.1	604.3
Imidazole	*a*	9769.3	28.7	−79.9	9718.1	9725.3
	b	9418.1	25.9	−77.2	9366.8	9374.0
	c	4795.2	13.7	−41.0	4767.9	4771.9
Indole	*a*	3893.3	11.6	−29.9	3875.0	3877.8
	b	1640.0	4.9	−10.0	1634.8	1636.0
	c	1153.9	3.4	−7.3	1150.0	1150.9
Fluorene	*a*	2186.0	6.7	−15.5	2177.2	2176.2
	b	588.0	1.8	−3.4	586.4	586.7
	c	464.7	1.4	−2.7	463.4	463.6
MeCpNaph[b]	*a*	1597.6	5.2	−10.5	1592.3	1591.1
	b	489.3	1.5	−3.0	487.8	488.1
	c	376.3	1.2	−2.2	375.3	375.4
MUE%		–	–	–	0.06	–
MAX%		–	–	–	0.50	–

[a]Experimental ground state rotational constants (taken from refs 58−66 for pyrazole, halogenated pyrazoles, cyclobutenone, cyclopentanone, t-2-decalone, imidazole, indole, fluorene, and 1-methyl-1*H*-cyclopenta[*a*]naphthalene, respectively) rounded to one decimal place. [b]1-Methyl-1*H*-cyclopenta[*a*]naphthalene.

in the present context and referred to in the following as B3/SVP), which deliver typical errors of about 5%.[25]

As we will see, equilibrium rotational constants obtained by the low-cost TSA/PCSB model described above in conjunction with B3/SVP vibrational corrections draw closer to the sought accuracy of 0.1% on the rotational constants of flexible molecules containing a few dozen atoms.

3. RESULTS AND DISCUSSION

A first aim of the present work is to extend the PCSB model to atoms of the fourth and fifth row of the periodic table employing pseudopotentials for replacing electrons of shells with a principal quantum number lower than $n - 1$. This is the reason why an upper bound of 3 is imposed on the principal quantum numbers in eq 4. Without any additional parameter, the carbon−halogen bond lengths in prototypical saturated compounds are very close

Table 2. Rotational Constants (in MHz) of the Molecular Motor Shown in Figure 4

Axis (τ)	B^{eq}(rDSD)	ΔB_τ^{PCSB}	ΔB_τ^{vib}	B_τ^0(PCSB)	B_τ^0(Exp.)[a]
a	308.266	0.967	−2.104	307.129	307.183
b	165.923	0.471	−1.129	165.265	164.951
c	122.793	0.363	−0.401	122.755	122.506
MUE%	0.392	–	–	0.137	–
MAX%	0.589	–	–	0.203	–

[a]Experimental ground state rotational constants (taken from ref 66) rounded to three decimal places.

https://doi.org/10.1021/acs.jpca.3c08382
J. Phys. Chem. A 2024, 128, 1385−1395

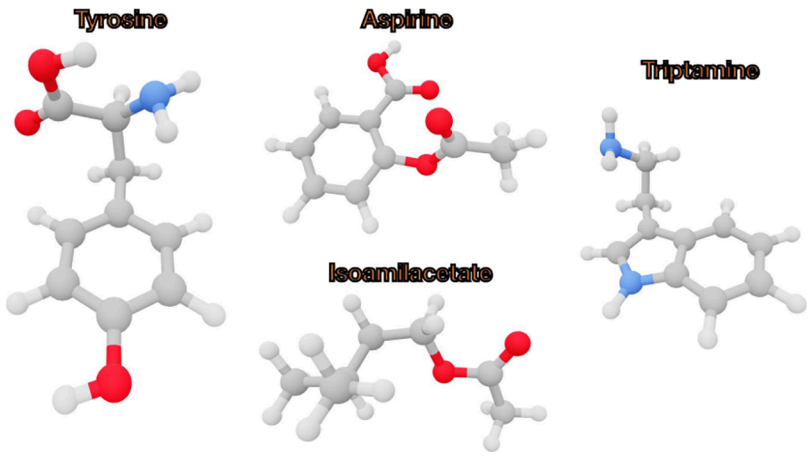

Figure 5. Molecular structures of tyrosine, aspirine, isoamilacetate, and triptamine.

Table 3. Rotational Constants (in MHz) of Prototypical Flexible Molecules

Molecule	Axis (τ)	B_τ^0(Exp.)[a]	ΔB_τ^{vib}	B_τ^0(rDSD)	B_τ^0(PCSB)	B_τ^0(TSA)
Creatinine-A1	*a*	3840.9	−29.3	3814.0	3826.1	3829.7
	b	1831.5	−14.0	1823.9	1830.4	1831.0
	c	1266.9	−11.9	1261.5	1265.8	1266.4
Creatinine-EI	*a*	3890.2	−41.1	3867.3	3878.5	3883.2
	b	1810.3	−11.9	1809.8	1809.8	1810.4
	c	1258.1	−11.2	1256.2	1256.2	1257.0
Creatinine-ZI	*a*	3842.4	−38.4	3811.6	3822.4	3827.2
	b	1825.4	−11.5	1817.3	1823.7	1824.5
	c	1261.0	−10.4	1254.7	1258.8	1259.6
Aspirine	*a*	1156.1	−7.6	1152.1	1155.6	1155.7
	b	762.6	−4.7	758.0	760.2	762.9
	c	509.0	−3.3	506.4	507.9	509.1
Isoamilacetate	*a*	3279.9	−28.9	3273.9	3287.7	3280.8
	b	711.6	−8.6	706.8	708.9	710.6
	c	689.7	−8.0	685.8	687.9	689.7
Histidine	*a*	1847.5	−13.9	1837.8	1845.2	1848.0
	b	831.7	−2.6	829.0	830.7	831.6
	c	745.9	−2.6	743.3	744.9	745.2
Tyrosine	*a*	1525.3	−13.7	1515.4	1521.4	1522.8
	b	465.5	−1.9	464.6	465.6	465.7
	c	427.3	−1.6	426.6	427.5	427.4
Triptamine	*a*	1730.2	−11.9	1722.2	1728.2	1730.2
	b	681.9	−4.9	680.0	681.6	681.8
	c	551.5	−3.8	549.9	551.3	551.4
MUE%		–	–	0.42	0.17	0.08
MAX%		–	–	0.80	0.52	0.40

[a]Experimental ground state rotational constants (taken from ref 67−72 for creatinine, aspirine, isoamilacetate, histidine, tyrosine, and triptamine, respectively) rounded to one decimal place.

to their semiexperimental counterparts taken from ref 22. For instance, the computed CBr bond lengths in CH_2Br_2 and $CHBrF_2$ are 1.9215 and 1.9278 Å to be compared to semiexperimental values of 1.9218 and 1.9279 Å. In the same vein, the computed CCl bond lengths in CH_2Cl_2 and the CI bond length in CH_3I are 1.7642 and 2.1323 Å to be compared with semiexperimental values of 1.7640 and 2.1336 Å.

The experimental and computed rotational constants for the prototypical molecules shown in Figures 3 and 4 are collected in Table 1, whereas the Cartesian coordinates of the rDSD/3F12$^-$

https://doi.org/10.1021/acs.jpca.3c08382
J. Phys. Chem. A 2024, 128, 1385−1395

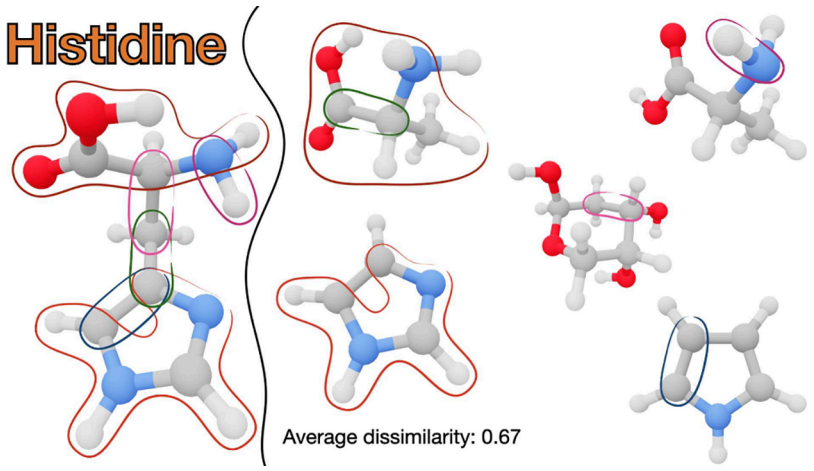

Figure 6. Templating syntyhon assignment for the bonds between non-hydrogen atoms of the histidine molecule. The average dissimilarity score of the highlighted atoms is also shown.

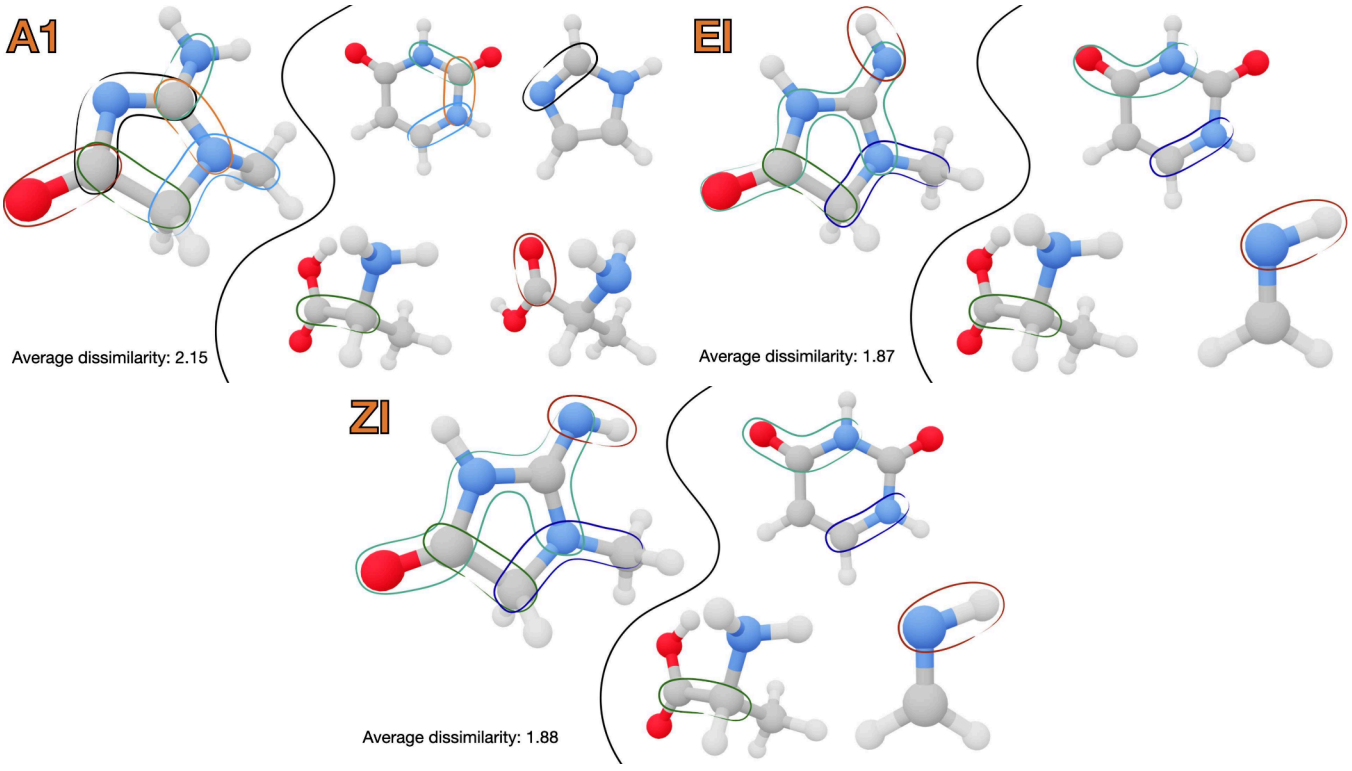

Figure 7. Templating synthon assignment for the bonds between non-hydrogen atoms and the imine NH bond in the three conformers of creatinine. The average dissimilarity score on the highlighted fragments is also shown.

optimized geometries are reported in the SI (the corresponding PCSB geometries can be obtained from their rDSD/3F12$^-$ counterparts by the free web utility quoted in section 2.1). As already mentioned, the corrections issuing from anharmonicity (leading contribution to ΔB^{vib}) and CV correlation (main contribution to ΔB^{B}) have opposite effects, with the former term being usually larger. As a consequence, the experimental ground state rotational constants (B^0) are smaller than their equilibrium counterparts (B^{eq}).

It is quite apparent that the PCSB model reaches the 0.1% target for halogen-substituted heteroaromatics (Br- and I-pyrazole[59]), molecules experiencing significant strain (cyclobutenone[60]) or involving large saturated cycles (t-2-decalone[62]), and typical building blocks of biomolecules (cyclopentanone,[61] pyrazole,[58] imidazole,[63] and indole[64]) or molecular motors (fluorene[65] and 1-methyl-1*H*-cyclopenta[*a*]naphthalene[66]). Taking into account also the results of previous studies,[14,15] it can be concluded that the structures of semirigid molecules containing up to about 20 atoms are described accurately at DFT cost.

As a first example of a larger molecule, we have considered the molecular motor shown in Figure 4, which contains 47 atoms, and whose microwave spectrum has been recently reported.[66] In this case, the computation of both the equilibrium geometry and the vibrational corrections to rotational constants by conventional approaches would be prohibitively expensive However, the PCSB model is orders of magnitude faster than composite wave function methods, and the new gradient procedure for

https://doi.org/10.1021/acs.jpca.3c08382
J. Phys. Chem. A 2024, 128, 1385−1395

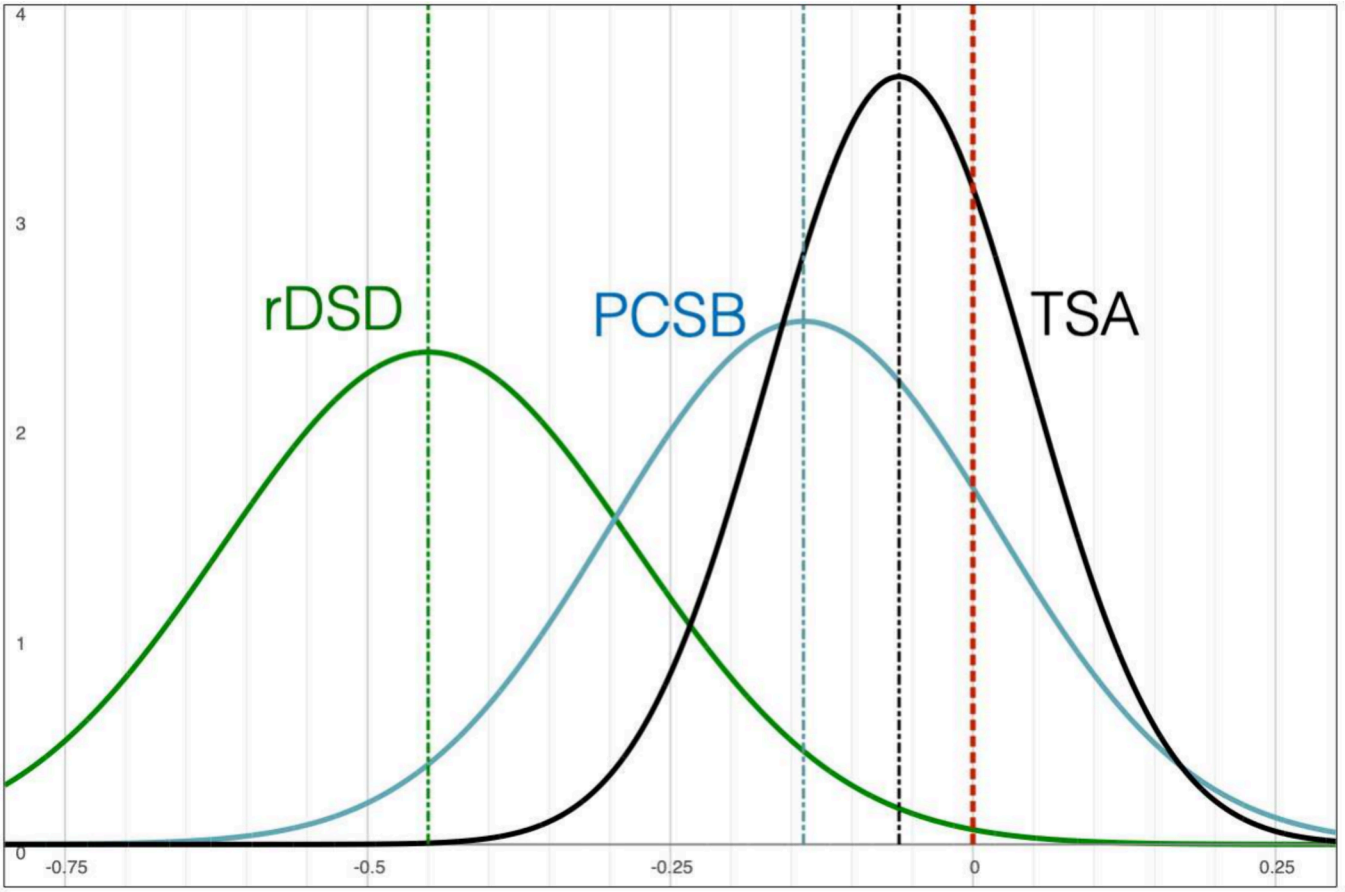

Figure 8. Gaussian error distributions for the rDSD, PCSB, and TSA models on the set of molecules of Table 3. The center of each Gaussian is evidenced by a vertical line and the zero-unsigned-error position is marked with a red vertical line.

vibrational corrections is about 2 orders of magnitude faster than its counterpart based on analytical Hessians. The results collected in Table 2 show that, thanks to the developments introduced in the present work, the sought accuracy of 0.1% is becoming possible for molecules containing about 50 atoms, which are the largest ones for which experimental rotational constants are available.

In general terms, the accuracy of computed rotational constants for non standard tautomeric forms or flexible moieties is lower than that of semirigid fragments.[15] In order to illustrate the performance of the TSA for such situations, we will make reference to the three stable tautomeric forms of creatinine (A1, EI, and ZI) and the most stable conformers of two drugs (aspirine and isoamilacetate), two α-amino acids (histidine and tyrosine), and a neurotransmitter (triptamine), which are shown in Figures 5, 6, and 7, and whose experimental rotational constants are available.[67−72] Also in this case, the Cartesian coordinates of all the rDSD/3F12$^-$ optimized structures are given in the SI.

Table 3 shows that the rDSD/3F12$^-$ results are already better than those delivered by the methods (B3LYP[73,74] or MP2[11]) currently available for large molecules and that the PCSB approach further improves the accuracy, although the 0.1% target is not always reached for the computed rotational constants. Then, we resort to the TSA, which produced, for instance, the synthons shown in Figures 6 and 7 for histidine and creatinine, respectively.

The TSA/PCSB rotational constants are collected in the last column of Table 3, whereas Figure 8 offers an intuitive picture of the performance of the different models in terms of normal distributions defined by

$$\rho(x) = N_c e^{-0.5\left(\frac{x-\Delta_{av}}{\Delta_{std}}\right)^2} \quad (16)$$

where Δ_{av} and Δ_{std} are the relative unsigned mean error and standard deviation, respectively, whereas N_c is a suitable normalization constant.

All of these results confirm that the unsupervised selection procedure underlying the TSA/PCSB model produces chemically significant templating synthons, which, in turn, further improve the PCSB results, leading to rotational constants satisfying the sought 0.1% accuracy.

4. CONCLUSIONS

The main targets of this study were the extension of the Pisa composite scheme to molecules containing heavy atoms and the introduction of the templating synthon approach for the description of problematic moieties. All of these developments and the implementation of the whole computational workflow in robust and user-friendly tools pave the way toward the unbiased assignment and interpretation of high-resolution spectra of large molecules at DFT cost by black-box quantum chemical computations of unprecedented accuracy.

https://doi.org/10.1021/acs.jpca.3c08382
J. Phys. Chem. A 2024, 128, 1385−1395

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.3c08382.

3F12$^-$ basis set in Gaussian format and Cartesian coordinates of all the molecular structures optimized at the rDSD/3F12$^-$ level. (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Federico Lazzari − *Scuola Normale Superiore, 56126 Pisa, Italy;* orcid.org/0000-0003-4506-3200; Email: federico.lazzari@sns.it

Vincenzo Barone − *INSTM, 50121 Firenze, Italy;* orcid.org/0000-0001-6420-4107; Email: vincebarone52@gmail.com

Author

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

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.3c08382

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Funding from Gaussian Inc. is gratefully acknowledged.

■ REFERENCES

(1) 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.

(2) Barone, V.; Alessandrini, S.; Biczysko, M.; Cheeseman, J. R.; Clary, D. C.; McCoy, A. B.; DiRisio, R. J.; Neese, F.; Melosso, M.; Puzzarini, C. Computational Molecular Spectroscopy. *Nature Rev. Meth. Prim.* **2021**, *1*, 38.

(3) Barone, V.; Di Grande, S.; Puzzarini, C. Toward Accurate Yet Effective Computations of Rotational Spectroscopy Parameters for Biomolecule Building Blocks. *Molecules* **2023**, *28*, 913.

(4) Janowski, T.; Pulay, P. Efficient Parallel Implementation of the CCSD External Exchange Operator and the Perturbative Triples (T) Energy Calculation. *J. Chem. Theory Comput.* **2008**, *4*, 1585−1592.

(5) Deumens, E.; Lotrich, V. F.; Perera, A.; Ponton, M. J.; Sanders, B. A.; Bartlett, R. J. Software Design of ACES III With the Super Instruction Architecture. *WIRES Comput. Mol. Sci.* **2011**, *1*, 895−901.

(6) Anisimov, V. M.; Bauer, G. H.; Chadalavada, K.; Olson, R. M.; Glenski, J. W.; Kramer, W. T. C.; Aprà, E.; Kowalski, K. Optimization of the Coupled Cluster Implementation in NWChem on Petascale Parallel Architectures. *J. Chem. Theory Comput.* **2014**, *10*, 4307−4316.

(7) Kaliman, I. A.; Krylov, A. New Algorithm for Tensor Contractions on Multi-Core CPUs, GPUs, and Accelerators Enables CCSD and EOM-CCSD Calculations With Over 1000 Basis Functions on a Single Compute Node. *J. Comput. Chem.* **2017**, *38*, 842−853.

(8) Kruse, H.; Šponer, J. Revisiting the Potential Energy Surface of the Stacked Cytosine Dimer: FNO-CCSD(T) Interaction Energies, SAPT Decompositions, and Benchmarking. *J. Phys. Chem. A* **2019**, *123*, 9209−9222.

(9) 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.

(10) 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.

(11) Møller, C.; Plesset, M. S. Note on an Approximation Treatment for Many-Electron Systems. *Phys. Rev.* **1934**, *46*, 618−622.

(12) Barone, V. PCS/Bonds: Pick your Molecule and Get Its Accurate Structure and Rotational Constants at DFT Cost. *J. Chem. Phys.* **2023**, *159*, No. 081102.

(13) Pauling, L. Atomic Radii and Interatomic Distances in Metals. *J. Am. Chem. Soc.* **1947**, *69*, 542−553.

(14) Barone, V.; Lazzari, F. Hunting for Complex Organic Molecules in the Interstellar Medium: the Role of Accurate Low-Cost Theoretical Geometries and Rotational Constants. *J. Phys. Chem. A* **2023**, *127*, 10517−10527.

(15) Barone, V. Quantum Chemistry Meets High-Resolution Spectroscopy for Characterizing the Molecular Bricks of Life in the Gas-Phase. *Phys. Chem. Chem. Phys.* **2024**, 25.

(16) 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 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.

(17) Warden, C. E.; Smith, D. G. A.; Burns, L. A.; Bozkaya, U.; Sherrill, C. D. Efficient and Automated Computation of Accurate Molecular Geometries Using Focal-Point Approximations to Large-Basis Coupled-Cluster Accuracy. *J. Chem. Phys.* **2020**, *152*, No. 124109.

(18) Watrous, A.; Westbrook, B. R.; Fortenberry, R. C. F12-TZ-cCR: A Methodology for Faster and Still Highly Accurate Quartic Force Fields. *J. Phys. Chem. A* **2021**, *125*, 10532−10540.

(19) Cosentino, U.; Villa, A.; Pitea, D.; Moro, G.; Barone, V.; Maiocchi, A. Conformational Characterization of Lanthanide(III) DOTA Complexes by ab Initio Investigation in Vacuo and in Aqueous Solution. *J. Am. Chem. Soc.* **2002**, *124*, 4901−4909.

(20) Santra, G.; Sylvetsky, N.; Martin, J. M. 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.

(21) 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.

(22) Ceselin, G.; Salta, Z.; Bloino, J.; Tasinato, N.; Barone, V. Accurate Quantum Chemical Spectroscopic Characterization of Glycolic Acid: a Route Toward its Astrophysical Detection. *J. Phys. Chem. A* **2022**, *126*, 2373−2387.

(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) Barone, V.; Ceselin, G.; Lazzari, F.; Tasinato, N. Toward Spectroscopic Accuracy for the Structure of Large Molecules at DFT Cost: Refinement and Extension of the Nano-LEGO Approach. *J. Phys. Chem. A* **2023**, *127*, 5183−5192.

(25) 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.

(26) Melli, A.; Barone, V.; Puzzarini, C. Unveiling Bifunctional Hydrogen Bonding with the Help of Quantum Chemistry: The Imidazole-Water Adduct as Test Case. *J. Phys. Chem. A* **2021**, *125*, 2989−2998.

(27) Barone, V.; Di Grande, S.; Lazzari, F.; Mendolicchio, M. Accurate Structures and Spectroscopic Parameters of Guanine Tautomers in the Gas Phase by the Pisa Conventional and Explicitly Correlated Composite Methods (PCS and PCS-F12). *J. Phys. Chem. A* **2023**, *127*, 6771−6778.

(28) Barone, V. DFT Meets Wave-Function Composite Methods for Characterizing Cytosine Tautomers in the Gas Phase. *J. Chem. Theory Comput.* **2023**, *19*, 4970−4981.

(29) 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.
(30) 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.
(31) 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.
(32) Peterson, K. A.; Figgen, D.; Goll, E.; Stoll, H.; Dolg, M. Systematically Convergent Basis Sets With Relativistic Pseudopotentials. II. Small-Core Pseudopotentials and Correlation Consistent Basis Sets for the Post-d Group 6−18 Elements. *J. Chem. Phys.* **2003**, *119*, 11113−11123.
(33) Peterson, K. A.; Shepler, B. C.; Figgen, D.; Stoll, H. On the Spectroscopic and Thermochemical Properties of ClO, BrO, IO, and Their Anions. *J. Phys. Chem. A* **2006**, *110*, 13877−13883.
(34) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, R. H.; Truhlar, G. D. Perspectives on Basis Sets Beautiful: Seasonal Plantings of Diffuse Basis Functions. *J. Chem. Theory Comput.* **2011**, *7*, 3027−3034.
(35) 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.
(36) Mehta, N.; Martin, J. M. L. Explicitly Correlated Double-Hybrid DFT: a Comprehensive Analysis of the Basis Set Convergence on the GMTKN55 Database. *J. Chem. Theory Comput.* **2022**, *18*, 5978−5991.
(37) 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.
(38) Cordero, B.; Gomez, V.; Platero-Prats, A. E.; Reves, M.; Echeverria, J.; Cremades, E.; Barragan, F.; Alvarez, S. Covalent Radii Revisited. *Dalton Trans* **2008**, 2832−2838.
(39) Frisch, M. J.; Trucks, W. G.; Schlegel, B. H.; Scuseria, E. G.; Robb, A. M.; Cheeseman, R. J.; Scalmani, G.; Barone, V.; Petersson, A. G.; Nakatsuji, H.; et al. *Gaussian 16 Revision C.01*; Gaussian Inc.: Wallingford, CT, 2016.
(40) 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.
(41) Demaison, J. Experimental, Semi-Experimental and Ab Initio Equilibrium Structures. *Mol. Phys.* **2007**, *105*, 3109−3138.
(42) Puzzarini, C.; Barone, V. Diving for Accurate Structures in the Ocean of Molecular Systems With the Help of Spectroscopy and Quantum Chemistry. *Acc. Chem. Res.* **2018**, *51*, 548−556.
(43) Mendolicchio, M.; Penocchio, E.; Licari, D.; Tasinato, N.; Barone, V. Development and Implementation of Advanced Fitting Methods for the Calculation of Accurate Molecular Structures. *J. Chem. Theory Comput.* **2017**, *13*, 3060−3075.
(44) Flygare, W. Magnetic Interactions in Molecules and an Analysis of Molecular Electronic Charge Distribution from Magnetic Parameters. *Chem. Rev.* **1974**, *74*, 653−687.
(45) Papousek, D.; Aliev, M. R. *Molecular vibrational-rotational spectra*; Elsevier Scientific Publishing Company: 1982.
(46) Csaszar, A. G.; Fabri, C.; Szidarovszky, T.; Matyus, E.; Furtenbacher, T.; Czako, G. The Fourth Age of Quantum Chemistry: Molecules in Motion. *Phys. Chem. Chem. Phys.* **2012**, *14*, 1085−1106.
(47) Van Vleck, J. H. On σ-Type Doubling and Electron Spin in the Spectra of Diatomic Molecules. *Phys. Rev.* **1929**, *33*, 467−506.
(48) Nielsen, H. H. The Vibration-Rotation Energies of Molecules. *Rev. Mod. Phys.* **1951**, *23*, 90−136.
(49) 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*; *Chapter SPECTRO - a Program for Derivation of Spectroscopic Constants from Provided Quartic Force Fields and Cubic Dipole Fields*, pp 186−195.
(50) Barone, V. Anharmonic Vibrational Properties by a Fully Automated Second Order Perturbative Approach. *J. Chem. Phys.* **2005**, *122*, No. 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) Yang, Q.; Mendolicchio, M.; Barone, V.; Bloino, J. Accuracy and Reliability in the Simulation of Vibrational Spectra: a Comprehensive Benchmark of Energies and Intensities Issuing from Generalized Vibrational Perturbation Theory to Second Order (GVPT2). *Front. Astron. Space Sci.* **2021**, *8*, No. 665232.
(53) Mendolicchio, M.; Bloino, J.; Barone, V. A 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**, *19*, 1759−1787.
(54) Mendolicchio, M.; Bloino, J.; Barone, V. Perturb-Then-Diagonalize Vibrational Engine Exploiting Curvilinear Internal Coordinates. *J. Chem. Theory Comput.* **2022**, *18*, 7603−7619.
(55) Mendolicchio, M. Harnessing the Power of Curvilinear Internal Coordinates: from Molecular Structure Prediction to Vibrational Spectroscopy. *Theor. Chem. Acc.* **2023**, *142*, 133.
(56) Puzzarini, C.; Stanton, J. F.; Gauss, J. Quantum-Chemical Calculation of Spectroscopic Parameters for Rotational Spectroscopy. *Int. Rev. Phys. Chem.* **2010**, *29*, 273−367.
(57) 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.
(58) Nygaard, L.; Christen, D.; Nielsen, D. J.; Pedersen, E. J.; Snerling, O.; Vestergaard, E.; Sorensen, G. O. Microwave Spectra of Isotopic Pyrazoles and Molecular Structure of Pyrazole. *J. Mol. Struct.* **1974**, *22*, 401−413.
(59) Cooper, G. A.; Medcraft, C.; Littlefair, J. D.; Penfold, T. J.; Walker, N. R. Halogen Bonding Properties of 4-Iodopyrazole and 4-Bromopyrazole Explored by Rotational Spectroscopy and Ab Initio Calculations. *J. Chem. Phys.* **2017**, *147*, No. 214303.
(60) Carrillo, M. J.; Marasinghe, D.; Feeley, E. B.; Sobie, K. M.; Zarzycki, R. J.; Carter-Fenk, K.; Fenk, C. J.; Tubergen, M. J. Theoretical and Microwave Spectroscopic Characterization of Cyclobutenone: Planar or Puckered? *J. Phys. Chem. A* **2023**, *127*, 9082−9087.
(61) Kim, H.; Gwinn, W. Ring Puckering in Five-Membered Rings. III. The Microwave Spectrum, Dipole Moment, and Structure of Cyclopentanone. *J. Chem. Phys.* **1969**, *51*, 1815−1819.
(62) Wachsmuth, D.; Jahn, M. K.; Blanco, S.; Gigosos, M. A.; Lesarri, A.; Grabow, J.-U. Inversion of Bicyclic Decanes: Rotational Spectra of the Trans and Double Cis Conformations of 2-Decalone. *ChemPhysChem* **2017**, *18*, 3620−3624.
(63) 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.; et al. Rotational Spectroscopy of Imidazole: Improved Rest Frequencies for Astrophysical Searches. *Astron. Astrophys.* **2019**, *628*, A53.
(64) Nesvadba, R.; Studecky, T.; Urban, S. Microwave Spectrum and Molecular Constants of Indole. *J. Mol. Spectrosc.* **2017**, *339*, 6−11.
(65) Thorwirth, S.; Theule, P.; Gottlieb, C. A.; McCarthy, M. C.; Thaddeus, P. Rotational Spectra of Small PAHs: Acenaphthene, Acenaphthylene, Azulene, and Fluorene. *Astrophys. J.* **2007**, *662*, 1309−1314.
(66) Domingos, S. R.; Cnossen, A.; Buma, W. J.; Browne, W. R.; Feringa, B. L.; Schnell, M. Cold Snapshot of a Molecular Rotary Motor Captured by High-Resolution Rotational Spectroscopy. *Angew. Chem., Int. Ed. Engl.* **2017**, *56*, 11209−11212.
(67) Léon, I.; Tasinato, N.; Spada, L.; Alonso, E. R.; Mata, S.; Balbi, A.; Puzzarini, C.; Alonso, J. L.; Barone, V. Looking for the Elusive Imine Tautomer of Creatinine: Different States of Aggregation Studied by Quantum Chemistry and Molecular Spectroscopy. *ChemPlusChem.* **2021**, *86*, 1374−1386.

https://doi.org/10.1021/acs.jpca.3c08382
J. Phys. Chem. A 2024, 128, 1385−1395

(68) Cabezas, C.; Alonso, J. L.; López, J. C.; Mata, S. Unveiling the Shape of Aspirin in the Gas Phase. *Angew. Chem., Int. Ed. Engl.* **2012**, *51*, 1375−1378.
(69) Sutikdja, L.; Jelisavac, D.; Stahl, W.; Kleiner, I. Structural Studies on Banana Oil, Isoamyl Acetate, by Means of Microwave Spectroscopy and Quantum Chemical Calculations. *Mol. Phys.* **2012**, *110*, 2883−2893.
(70) Bermúdez, C.; Mata, S.; Cabezas, C.; Alonso, J. L. Tautomerism in Neutral Histidine. *Angew. Chem., Int. Ed. Engl.* **2014**, *126*, 1195−1198.
(71) 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.
(72) Caminati, W. The Rotational Spectra of Conformers of Biomolecules: Tryptamine. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2806−2809.
(73) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785−789.
(74) Becke, A. D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

https://doi.org/10.1021/acs.jpca.3c08382
J. Phys. Chem. A 2024, 128, 1385−1395