Semi-Experimental Equilibrium Structure Determinations by Employing B3LYP/SNSD Anharmonic Force Fields: Validation and Application to Semirigid Organic Molecules

Published Version:
Semi-Experimental Equilibrium Structure Determinations by Employing B3LYP/SNSD Anharmonic Force Fields: Validation and Application to Semirigid Organic Molecules / Piccardo, Matteo; Penocchio, Emanuele; Puzzarini, Cristina; Biczysko, Malgorzata; Barone, Vincenzo. - In: JOURNAL OF PHYSICAL CHEMISTRY. A, MOLECULES, SPECTROSCOPY, KINETICS, ENVIRONMENT, & GENERAL THEORY. - ISSN 1089-5639. - STAMPA. - 119:10(2015), pp. 2058-2082. [10.1021/jp511432m]

Availability:
This version is available at: https://hdl.handle.net/11585/523965 since: 2020-02-21

Published:
DOI: http://doi.org/10.1021/jp511432m

M. Piccardo, E. Penocchio, C. Puzzarini, M. Biczysko, V. Barone. "Semi-Experimental Equilibrium Structure Determinations by Employing B3LYP/SNSD Anharmonic Force Fields: Validation and Application to Semirigid Organic Molecules", J. Phys. Chem. A, **119** (2015) 2058

The final published version is available online at: DOI: 10.1021/jp511432m

Semi-Experimental Equilibrium Structure Determinations by Employing B3LYP/SNSD Anharmonic Force Fields: Validation and Application to Semirigid Organic Molecules

Matteo Piccardo,[†] Emanuele Penocchio,[‡,†] Cristina Puzzarini,[‡] Malgorzata Biczysko,[†,¶] and Vincenzo Barone[*,†]

Scuola Normale Superiore, Pisa, Italy, Dipartimento di Chimica “Giacomo Ciamician”, Universitá di Bologna, Via Selmi 2, I-40126 Bologna, Italy, and Consiglio Nazionale delle Ricerche, Istituto di Chimica dei Composti OrganoMetallici (ICCOM-CNR), Area della Ricerca CNR, UOS di Pisa, Via G. Moruzzi 1, I-56124 Pisa, Italy

E-mail: vincenzo.barone@sns.it.

Abstract

This work aims at extending the semi-experimental (SE) approach for deriving accurate equilibrium structures to large molecular systems of organic and biological interest. SE equilibrium structures are derived by a least-squares fit of the structural parameters to the experimental ground-state rotational constants of several isotopic species corrected by vibrational contributions computed by quantum mechanical (QM) methods. A systematic benchmark study on 21 small molecules (CCse set) is carried

---

[*] To whom correspondence should be addressed

[†] Scuola Normale Superiore

[‡] Dipartimento di Chimica “Giacomo Ciamician”

[¶] ICCOM-CNR

out to evaluate the performance of hybrid density functionals (in particular B3LYP) in the derivation of vibrational corrections to rotational constants. The resulting SE equilibrium structures show a very good agreement with the corresponding geometries obtained employing post-Hartree-Fock vibrational corrections. The use of B3LYP in conjunction with the double-ζ SNSD basis set strongly reduces the computational costs, thus allowing for the evaluation of accurate SE equilibrium structures for medium-sized molecular systems. On these grounds, an additional set of 26 SE equilibrium structures including the most common organic moieties has been set up by collecting the most accurate geometries available in the literature together with new determinations from the present work. The overall set of 47 SE equilibrium structures determined using B3LYP/SNSD vibrational corrections (B3se set) provides a high quality benchmark for validating the structural predictions of other experimental and/or computational approaches. Finally, we present a new strategy (referred to as template approach) to deal with the cases for which it is not possible to fit all geometrical parameters due to the lack of experimental data.

INTRODUCTION

The last decades have seen many efforts to determine accurate molecular structures for systems of increasing size and complexity.[1–14] Detailed knowledge of the equilibrium structures of isolated molecular systems of chemical, biological or technological interest is indeed a prerequisite for a deeper understanding of other physical-chemical properties, ranging from a precise evaluation of the electronic structure to the understanding and analysis of the dynamical and environmental effects affecting the molecular structures and properties.[1,2,15–17] Moreover, the availability of reference molecular structures allows one to test the accuracy of different quantum mechanical (QM) approaches,[2,18–22] and it is essential for a correct development of accurate force fields either of general applicability (e.g. for systems of biological interest)[23–26] or specifically tailored for individual systems.[27–30] Furthermore, robust and reliable computational approaches are of primary importance for conformational analysis and modeling of drugs and biomolecules,[18,31] as well as for a deeper understanding of chemical reactivity in terms of transition state structures,[32] which are not directly determinable from experiment. For a fruitful interplay of experiment and theory in the interpretation and quantification of molecular properties, and for validation purposes, it is hence desirable to have a large number of accurate equilibrium geometries at one disposal.

Nowadays, an increasing number of experimental data is available thanks to the growing interest in the field, but the structural parameters derived from experiment often depend on the chosen technique and can be biased by vibration and/or environmental conditions.[2,15] For example, the vibrationally averaged r_0 and substitution r_s structures are obtained from microwave and/or rotationally resolved infrared investigations through the analysis of the vibrational ground-state rotational constants for different isotopologues, but without an explicit consideration of vibrational effects.[33] The dependence of the results on experimental conditions complicates both the comparison of structures obtained with different experimental techniques and the subsequent use of these empirical structures in the computation of molecular properties. In addition, all vibrationally averaged structures (r_0, r_s, $r_{\alpha,T}$, $r_{g,T}$, etc.) depend on the isotopic species considered.[33,34]

A way to avoid all these problems is to resort to equilibrium structures (r_e), which are defined as the geometries at the minimum of the Born-Oppenheimer (BO) potential energy surface.[1,2] Although they are cumbersome to derive experimentally, and therefore generally available only for small molecules, this kind of structures is preferred as they exclude vibrational effects in a rigorous manner and, within the BO approximation,[35,36] are independent of the considered isotopic species. Moreover, depending solely on the electronic structure of the molecular system, r_e structures are directly comparable with the results from QM calculations.

Reference equilibrium structures can be obtained from high-level QM calculations, for instance making use of the coupled-cluster (CC) singles and doubles approximation augmented by a perturbative treatment of triple excitations, CCSD(T),[37] which is able to provide accurate structures, rivaling the best experimental results, provided that extrapolation to the complete basis-set limit and core correlation are taken into the proper account (see, for example, refs. 38–40). However, for medium-sized molecular systems such computations are still very challenging, due to the unfavourable scaling of highly correlated levels of theory with the number of basis functions.

An important step forward in this field has been provided by the introduction of the so-called semi-experimental (SE) equilibrium geometry (r_e^{SE}), which is obtained by a least-square fit of experimental rotational constants of different isotopologues corrected by computed vibrational contributions.[1,2,41] Introduced by Pulay *et al.*,[41] this method is nowadays considered the best approach to determine accurate equilibrium structures for isolated molecules.[42] Such an interplay of theory and experiment paves the route toward the extension of accurate structural studies to systems larger than those treatable by experimental and QM methods separately.

From a computational point of view, the bottleneck of the SE protocol is the calculation of the cubic force field at a level of theory sufficiently accurate to give reliable vibrational corrections to rotational constants.[42] Actually, CCSD(T) is considered the gold standard for this kind of determinations, but the computational cost restricts its applicability to systems of less than 10 atoms (see for example Refs. 43–45). Such a limitation needs to be overcome in order to set up a database of accurate molecular geometries to be used as references for benchmark QM calculations as well as for the validation of simpler models for larger systems, with special focus on biomolecule building blocks. Therefore, the setup and validation of a SE approach able to combine high accuracy and low computational cost is of great interest. In this view, we carried out a systematic study to demonstrate that the calculation of vibrational corrections from anharmonic force fields evaluated using the density functional theory (DFT) permits to obtain r_e^{SE} structures that agree well with the best equilibrium geometries reported in the literature, but with a significantly reduced computational effort. The key point is here the fact that the relevant quantity to correct equilibrium rotational constants is the sum of the vibration-rotation interaction constants α_i^β, and not the individual constants. The advantage is that resonance-free equations can be obtained and that, thanks to error cancellation, it is much easier to get sufficiently accurate values for the sum than for individual terms.

In previous studies, we showed that the B3LYP hybrid functional performs remarkably well for vibrational properties, when coupled to basis sets of at least polarized double-ζ quality including diffuse functions.[46–49] On these grounds, all DFT computations have been performed at the B3LYP level.

A first validation study was performed on 21 small molecules (hereafter CCse set) for which a sufficient number of experimental rotational constants is available and cubic CCSD(T) force fields with at least triple-ζ basis sets were computationally feasible or already known. These reference values were next compared with those issuing from B3LYP and MP2 cubic force fields. The remarkable accuracy of B3LYP/SNSD results allowed us to derive new SE equilibrium structures for an additional set of 26 medium-sized molecules characterized by the most representative bond patterns of organic systems, and including H, C, N, O, F, S and Cl atoms. The whole set of 47 SE equilibrium structures determined using B3LYP/SNSD vibrational corrections (hereafter referred to as B3se set) represents a high quality benchmark for structural studies and validation of computational models. In addition to the rigorous SE approach, theoretical and experimental data can also be combined in cases for which the lack of experimental information for a sufficient number of isotopologues prevents the derivation of a complete SE equilibrium structure. In these cases, fixing some geometrical parameters to reliable and accurate estimates allows for the determination of the remaining structural parameters for systems otherwise non-entirely characterizable (see for example Refs. 50–55). To this end, we introduced a new approach, denoted as template approach, that exploits the accurate SE results obtained for reference molecules in order to derive SE equilibrium structures for similar systems by avoiding highly expensive CC computations.

METHODOLOGY AND COMPUTATIONAL DETAILS

The so-called r_0 structure is obtained by a least-squares fit (LSF) of the molecular parameters to the experimental ground-state rotational constants $(B_0^\beta)^{\mathrm{EXP}}$ of a set of isotopologues, or their corresponding moments of inertia $(I_0^\beta)^{\mathrm{EXP}}$, where $\beta = x, y$ or z is one of the principal inertial axes in the molecule-fixed reference frame.

On the other hand, the mixed experimental-theoretical approach starts from the consideration that equilibrium rotational constants should be employed instead of the experimental ones, thus requiring vibrational and electronic contributions to be subtracted before the fitting procedure. The r_e^{SE} structure is then obtained by a LSF to the SE equilibrium rotational constants $(B_e^\beta)^{\mathrm{SE}}$, or their corresponding moments of inertia $(I_e^\beta)^{\mathrm{SE}}$, where $(B_e^\beta)^{\mathrm{SE}}$ are calculated from $(B_0^\beta)^{\mathrm{EXP}}$ as,

$$(B_e^\beta)^{\mathrm{SE}} = (B_0^\beta)^{\mathrm{EXP}} - (\Delta B_0^\beta)^{\mathrm{QM}} \tag{1}$$

$(\Delta B_0^\beta)^{\mathrm{QM}}$ is explicitly given by,

$$\begin{aligned}(\Delta B_0^\beta)^{\mathrm{QM}} &= \frac{m_e}{M_p} g^{\beta\beta} B_e^\beta - \sum_i \frac{\alpha_i^\beta d_i}{2} \\ &= \Delta B_{\mathrm{el}}^\beta + \Delta B_{\mathrm{vib}}^\beta \end{aligned} \tag{2}$$

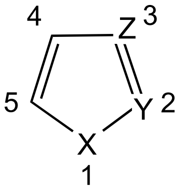
$\Delta B_{\mathrm{el}}^\beta$ is an electronic contribution, evaluated from the rotational **g** tensor and the ratio between electron (m_e) and proton (M_p) masses.[56–58] Although this term is often negligible, it will be systematically included in our computations for the sake of completeness. Within the BO approximation and enforcing Eckart-Sayvetz conditions,[35,36,59,60] the vibrational contribution $\Delta B_{\mathrm{vib}}^\beta$ is obtained by applying second-order vibrational perturbation theory (VPT2) to the molecular ro-vibrational Hamiltonian expressed in normal coordinates.[61–63] In the summation of eq. 2, α_i^β are the vibration-rotation interaction constants, explicitly given in the Supplementary Information (SI), and d_i is the degeneracy of the i-th vibrational normal mode.

In the present investigation, the cubic force fields required for the computation of the $\Delta B^{\beta}_{\mathrm{vib}}$ term have been evaluated at the CCSD(T),[37,64] second-order Møller-Plesset perturbation theory (MP2)[65,66] and DFT[67] levels. The correlation-consistent polarized cc-p(wC)VnZ basis sets[68–71] have mainly been used in CCSD(T) and MP2 calculations, with n = T, Q denoting the cardinal number of the corresponding basis set, shortly denoted as (wC)VnZ in the text. The frozen-core (fc) approximation has been adopted in conjunction with the VnZ sets, while all electrons (but 1s for second-row elements) have been correlated with wCVnZ. The hybrid B3LYP functional[72–74] has been used in conjunction with the SNSD basis set,[48,49,75] which represents an excellent compromise between accuracy and computational cost for vibrational studies.[46,48,49,76] Possible effects of basis set extension have also been investigated by employing the aug-cc-pVTZ (hereafter AVTZ) basis set.[69,77] In the specific case of CH_2CHF and *cis*-CHFCHCl, the cc-pVTZ (VTZ) basis set has been preferred, in view of the worsening in the vibrational frequencies coming from the inclusion of diffuse functions in the triple-ζ basis set for halo-ethylenes.[48]
The CFOUR program package[78] has been employed for MP2 and CCSD(T) computations, while DFT calculations have been performed with the GAUSSIAN suite of programs.[79] For all computational levels, the harmonic part has been obtained using analytic second derivatives, whereas the corresponding cubic force field has been determined in a normal-coordinate representation via numerical differentiation of the analytically computed harmonic force constants.[80–85] At the DFT level, the force field calculations have been carried out using very-tight criteria for the SCF and geometry optimization convergence, together with an ultra-fine grid for the numerical integration of the two-electron integrals and their derivatives. The numerical differentiations have been performed with the Gaussian default step of 0.01 Å.
The $\Delta B^{\beta}_{\mathrm{el}}$ contributions have been evaluated by calculating the $g^{\beta\beta}$ constants at the B3LYP/AVTZ level of theory.

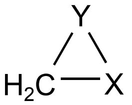
To obtain accurate reference equilibrium structures for pyridine, 2-fluoropyridine and 3-

fluoropyridine, we have performed geometry optimizations at the CCSD(T) level accounting for basis-set truncation errors and core-valence correlation effects by means of a composite approach.[38,39] The corresponding r_e is denoted as CCSD(T)/CBS+CV.

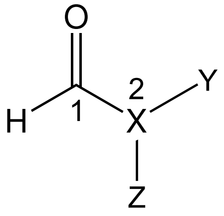
Figure 1: Sketch of the 36 asymmetric top molecules belonging to the B3se set (see also Tables 1, 4 and 5).



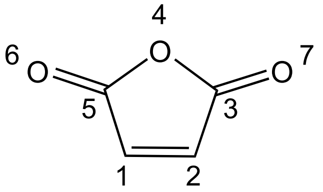
(a) (X,Y,Z)
(NH,CH,CH) pyrrole
(NH,N,CH) pyrazole
(NH,CH,N) imidazole
(O,CH,CH) furan
(S,CH,CH) thiophene



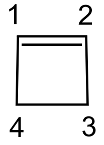
(b) (X,Y)
(NH,CH_2) aziridine
(O,CH_2) oxirane
(O,O) dioxirane



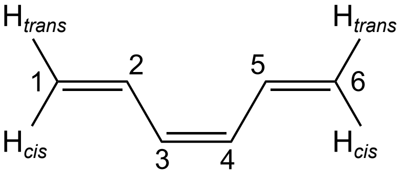
(c) (X,Y,Z)
(H,-,-) formaldehyde
(CH_2,OH,-) glycolaldehyde
(CH_2,CH_3,-) propanal
(C,CH_2,H) *cis*-acrolein
(C,H,CH_2) *trans*-acrolein
(O,H,-) *cis*-formic acid
(O,-,H) *trans*-formic acid
(O,CH_3,-) *cis*-methyl formate



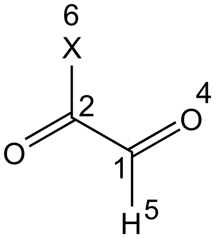
(d) maleic anhydride



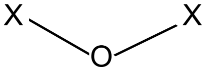
(e) cyclobutene



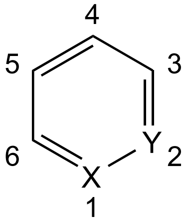
(f) *cis*-hexatriene



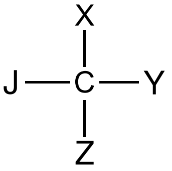
(g) X=H *trans*-glyoxal
X=OH glyoxylic acid



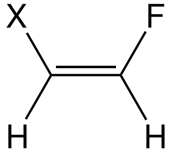
(h) X=H water
X=CH_3 dimethyl ether



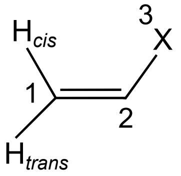
(i) (X,Y)
(N,CH) pyridine
(N,N) pyridazine



(j) (X,Y,Z,J)
(H,H,F,F) CH_2F_2
(Cl,Cl,F,F) CCl_2F_2
(H,H,Cl,Cl) CH_2Cl_2
(H,Cl,F,F) $CHClF_2$
(Cl,F,H,H) CH_2ClF



(k) X=H CH_2CHF
X=Cl *cis*-CHFCHCl



(l) X=H ethene
X=OH ethenol
X=CH_3 propene
X=$CHCH_2$ butadiene

RESULTS AND DISCUSSION

All the asymmetric top molecules of the B3se set, and when needed, their atom numbering are sketched in Figure 1.

Validation study: the performance of B3LYP force fields

As mentioned in the introduction, a set of 21 molecules, including linear (HCN, HNC, HCO^+, $HNCCN^+$, HCCH, HCCCCH), symmetric-top (H_2CCCH_2, SH_3^+, NH_3) and asymmetric-top (H_2O, H_2CO, CH_2ClF, CH_2CHF, *cis*-CHFCHCl, oxirane, dioxirane, cyclobutene, *trans*-glyoxal, *cis* and *trans*-acrolein, pyridazine) molecules, has been selected to investigate the performance of the B3LYP hybrid functional in the computation of the vibrational contributions to experimental vibrational ground-state rotational constants $(B_0^\beta)^{\mathrm{EXP}}$ subsequently used in the derivation of SE equilibrium geometries.

For all systems listed above, the experimental $(B_0^\beta)^{\mathrm{EXP}}$ constants and the $\Delta B_{\mathrm{vib}}^\beta$ contributions computed at the CCSD(T) and MP2 levels available in the literature have been collected, see Table 1 in the SI. When not available, MP2 and/or CCSD(T) vibrational contributions have been calculated in this work (see Table 1 for details), together with the $\Delta B_{\mathrm{vib}}^\beta$ contributions computed at the DFT level. The $\Delta B_{\mathrm{el}}^\beta$ contributions have also been taken into account. In particular, large $\Delta B_{\mathrm{el}}^\beta$ values are found for H_2O (about from 7.6 to 0.7% of $(\Delta B_0^\beta)^{\mathrm{QM}}$) and H_2CO (12.5-0.6%). Furthermore, the importance of taking into account the electronic contributions for *cis* and *trans*-acrolein and pyridazine is well known.[86–88] For both isomers of acrolein, ΔB_{el}^A, ΔB_{el}^B and ΔB_{el}^C are about 3.0-4.5%, 0.5-0.6% and 0.05-0.07% of $(\Delta B_0^\beta)^{\mathrm{QM}}$. For pyridazine, ΔB_{el}^A and ΔB_{el}^B are about 0.7-1.2% of $(\Delta B_0^\beta)^{\mathrm{QM}}$, while ΔB_{el}^C is about 0.3%. All $\Delta B_{\mathrm{vib}}^\beta$ and $\Delta B_{\mathrm{el}}^\beta$ contributions are given in Table 1 in the SI.

In Figure 2, for all the considered molecules, we have displayed the $(\Delta B_0^\beta)^{\mathrm{QM}}$ corrections for all isotopologues studied in terms of the percentage of the corresponding $(B_0^\beta)^{\mathrm{EXP}}$. From this Figure, it is apparent that, as expected, the $(\Delta B_0^\beta)^{\mathrm{QM}}$ are rather small contributions. They vary from 2-3% of $(B_0^\beta)^{\mathrm{EXP}}$, for systems like H_2O, to less than 1% in the case of $HNCCN^+$ and HCCCCH. Negative $(\Delta B_0^\beta)^{\mathrm{QM}}$ corrections are obtained for all molecules except for $(B_0^A)^{\mathrm{EXP}}$ of H_2O. The different computational levels used for the calculation of $(\Delta B_0^\beta)^{\mathrm{QM}}$ are discriminated by using different graphical symbols. This allows us to point out the very good agreement between the results obtained at the different levels of theory considered. Indeed, there is a generally good superposition of the graphical symbols for the various isotopologues: the discrepancies between the various methods and CCSD(T) are well within 1%. On these grounds, it is possible to estimate (see equation 13 of ref. 42) that the resulting SE geometrical parameters differ by at most 0.25% from those obtained with vibrational contributions at the CCSD(T) level.

In the following, the SE equilibrium structures derived using vibrational contributions from CCSD(T), MP2, B3LYP/SNSD and B3LYP/AVTZ force fields are referred to as CCSD(T) SE, MP2 SE, B3LYP/SNSD SE and B3LYP/AVTZ SE, respectively. The CCSD(T) and B3LYP/SNSD SE equilibrium structures are explicitly reported in Table 1, while the MP2 and B3LYP/AVTZ ones are given in Table 6 of the SI in terms of discrepancies with respect to the CCSD(T) SE geometrical parameters. In Table 1 the fully experimental r_0 structures are also collected together with the equilibrium geometries optimized at the B3LYP/SNSD level. The r_0 structures are reported to point out the non-negligible deviations that discourage their use even for establishing general trends. The B3LYP/SNSD equilibrium structures are shown in order to be easily accessible in view of the template approach presented later in the text. The root mean square (RMS) of the residuals in terms of equilibrium rotational constants (hereafter simply referred to as residuals), and for planar molecules, the mean inertial defects $\Delta_e = I^C - I^B - I^A$ are also given in Table 1 as indicators of the quality of the fits. Indeed, small values of the RMS residuals and Δ_e indicate a good quality of the fits and that $(\Delta B_0^\beta)^{\mathrm{QM}}$ corrections lead to good SE equilibrium rotational constants, respectively.

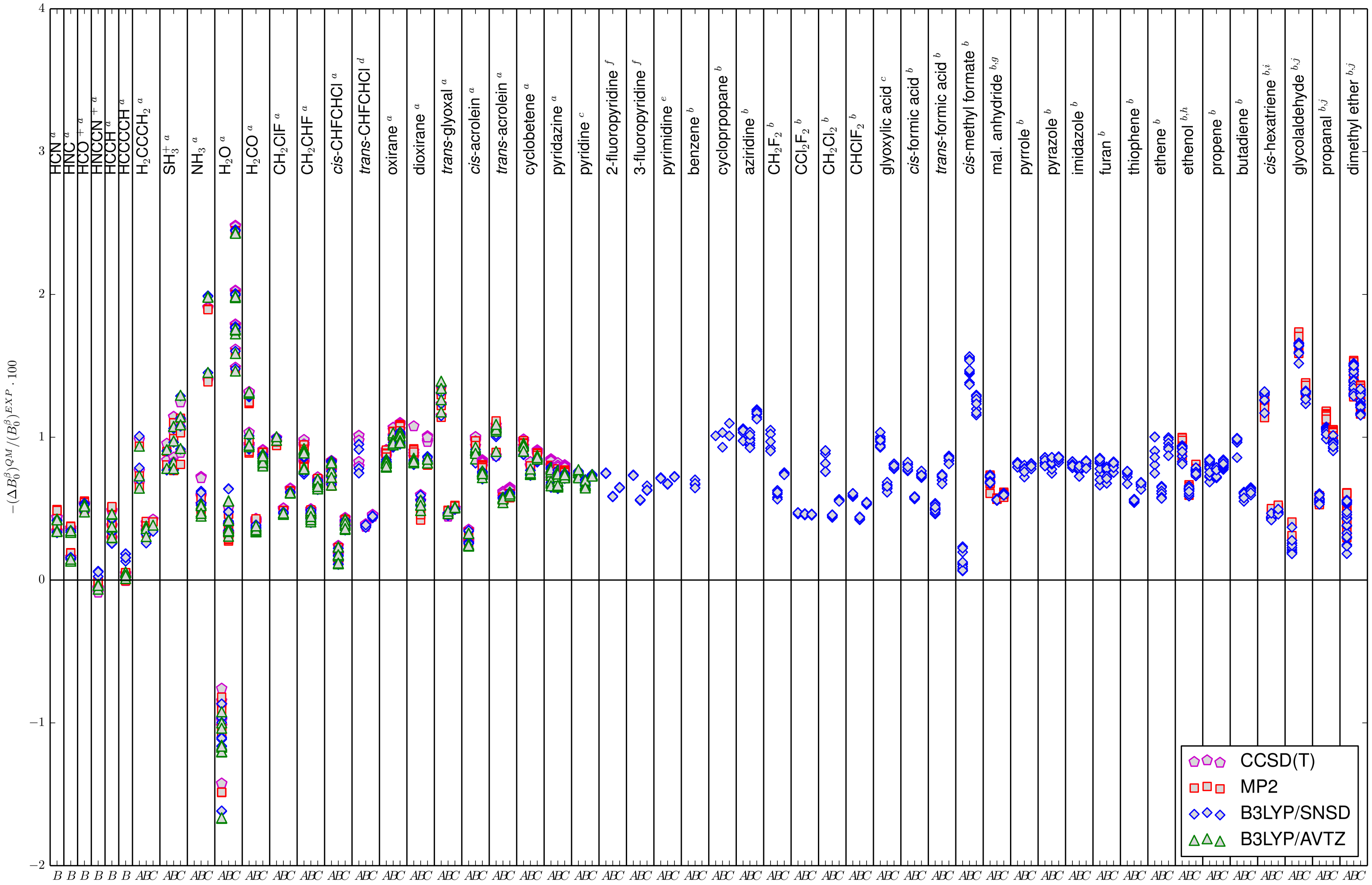
The differences in the geometrical parameters of the MP2, B3LYP/SNSD and B3LYP/AVTZ

SE equilibrium geometries with respect to the CCSD(T) SE equilibrium structures are graphically reported in Figure 3 (see also Table 6 in the SI). It is noteworthy that for the whole set of bond lengths the deviation of MP2 and B3LYP results from the CCSD(T) references never exceeds 0.0026 Å. The deviations show a nearly gaussian distribution with mean values close to zero and mean absolute errors (MAE) of 0.0004 Å, 0.0007 Å and 0.0005 Å for MP2, B3LYP/SNSD and B3LYP/AVTZ, respectively (see Table 2), thus pointing out the good accuracy of DFT vibrational contributions to rotational constants in evaluating SE equilibrium structures. The small standard deviations of MP2, B3LYP/SNSD and B3LYP/AVTZ can be considered fully satisfactory for geometrical parameter determinations. Focusing on specific bonds (see Figure 1 in the SI), it can be observed that the C-C bond lengths are those showing a quite large MAE with respect to the CH and CO bond lengths. A larger MAE (0.0056 Å) is obtained for r_0 structures, with a standard deviation of 0.0063 Å, a value significantly larger than the typical uncertainty affecting the SE methodology.

The deviations for angles are also very small, with MAE of 0.03, 0.06 and 0.06 degrees for MP2, B3LYP/SNSD and B3LYP/AVTZ, respectively. Similarly to bond lengths, these values correspond to accuracies comparable with the intrinsic errors of the SE fitting procedure. Also for angles, the deviations of r_0 structures are an order of magnitude larger than that of the various r_e^{SE}'s.

A linear least-square fit of the CCSD(T) r_e^{SE} values, expressed as a function of the corresponding MP2 and DFT ones (see Figure 2 in the SI), gives the parameters reported in Table 3. It is noteworthy that in all cases the angular coefficient is very close to 1 and the intercept never exceeds, in absolute value, 0.0025 Å for bond lengths and 0.06 degrees for angles. This confirms that using B3LYP corrections in the SE approach leads to results that reproduce very well the best SE equilibrium structures. In addition, the analysis of R^2 and standard deviation values of the linear regression does not point out any significant deviation from linearity.

Figure 2: $-(\Delta B_0^\beta)^{\mathrm{QM}}/(B_0^\beta)^{\mathrm{EXP}} \cdot 100$ for the isotopologues of all molecules considered in this work.



a) see Table 1 here and Table 1 in the SI. b) see Table 4 here and Table 2 in the SI. c) see Table 5 here and Table 2 in the SI. d) see Table 7 here and Table 4 in the SI. e) see Table 8 here and Table 5 in the SI. f) see Table 9 here and Table 3 in the SI.
g) MP2/VTZ $\Delta B_{\mathrm{vib}}^\beta$ from ref. 89. h) MP2/VQZ $\Delta B_{\mathrm{vib}}^\beta$ from ref. 90. i) MP2/VTZ $\Delta B_{\mathrm{vib}}^\beta$ from ref. 91. j) MP2/VTZ $\Delta B_{\mathrm{vib}}^\beta$ from ref. 90.

Table 1: r_0, r_e^{SE} and r_e geometries for the 21 molecules of the CCse set. Distances in Å, angles in degrees.

	$r_0{}^a$	$r_e^{\mathrm{SE}\,a}$		r_e
		CCSD(T)	B3LYP/SNSD	B3LYP/SNSD
		– linear molecules –		
HCN^b				
r(H-C)	1.0624(2)	$1.0651(1)^{\times,\dagger}$	$1.0645(1)^{\dagger}$	1.0707
r(C-N)	1.1568(1)	1.1533(1)	1.1536(1)	1.1551
Rms resid. [MHz]	0.0009	0.0001	0.0001	-
HNC^b				
r(H-N)	0.9863(2)	$0.9954(1)^{\times,\dagger}$	$0.9946(1)^{\dagger}$	1.0022
r(N-C)	1.1725(1)	1.1685(1)	1.1688(1)	1.1739
Rms resid. [MHz]	0.0010	0.0001	0.0001	-
HCO$^+$				
r(H-C)	1.0921(2)	$1.0919(1)^{\times,\dagger}$	$1.0916(1)^{\dagger}$	1.0995
r(C-O)	1.1091(1)	1.1057(1)	1.1057(1)	1.1075
Rms resid. [MHz]	0.0011	0.0001	0.0001	-
HNCCN^{+c}				
r(H-N)	1.0058(4)	$1.0133(1)^{\asymp,\dagger}$	$1.0138(2)^{\dagger}$	1.0191
r(N-C)	1.1400(8)	1.1406(1)	1.1392(4)	1.1455
r(C-C)	1.3762(9)	1.3724(1)	1.3735(4)	1.3686
r(C-N)	1.1584(7)	1.1634(1)	1.1607(3)	1.1628
Rms resid. [MHz]	0.0017	0.0002	0.0008	-
HCCH^d				
r(C≡C)	1.2084(1)	$1.2030(1)^{+,\dagger}$	$1.2036(1)^{\dagger}$	1.2060
r(C-H)	1.0572(2)	1.0617(1)	1.0611(1)	1.0676
Rms resid. [MHz]	0.0013	0.0001	0.0002	-
HCCCCH^e				
r(C≡C)	1.2079(3)	$1.2084(3)^{\times,\dagger}$	$1.2070(4)^{\dagger}$	1.2123
r(C-C)	1.3751(4)	1.3727(4)	1.3726(6)	1.3685
r(C-H)	1.0561(1)	1.0615(1)	1.0610(1)	1.0667
Rms resid. [MHz]	0.0008	0.0008	0.0012	-
		– symmetric top molecules –		
SH$_3^{+f}$				
r(S-H)	1.3563(2)	$1.3500(1)^{+,\dagger}$	$1.3502(1)^{\dagger}$	1.3683
a(H-S-H)	94.19(3)	94.15(1)	94.11(1)	94.30
Rms resid. [MHz]	0.0068	0.0010	0.0011	-
NH$_3{}^b$				
r(N-H)	1.0150(3)	$1.0110(2)^{\times,\dagger}$	$1.0111(2)^{\dagger}$	1.0176
a(H-N-H)	107.52(4)	106.94(2)	106.87(3)	106.59
Rms resid. [MHz]	0.0055	0.0033	0.0036	-
H$_2$**CCCH**$_2$				
r(C=C)	1.3096(4)	$1.3066(1)^{\div,\dagger}$	$1.3075(2)^{\dagger}$	1.3077
r(C-H)	1.0833(9)	1.0807(1)	1.0800(4)	1.0874
a(H-C-H)	118.56(11)	118.26(1)	118.37(5)	117.41
Rms resid. [MHz]	0.0169	0.0011	0.0081	-
		– asymmetric top molecules –		
H$_2$**O**^{b}				
r(O-H)	0.9567(1)	$0.9573(1)^{\times,\dagger}$	$0.9572(1)^{\dagger}$	0.9644
a(H-O-H)	104.93(2)	104.53(1)	104.47(1)	104.60
Rms resid. [MHz]	0.0014	0.0004	0.0004	-
Mean Δ_e [uÅ^2]	0.06277	0.00506	0.00603	-

H_2CO^b				
r(C-O)	1.2095(2)	1.2047(1)$^{\times,\dagger}$	1.2051(1)†	1.2052
r(C-H)	1.1064(4)	1.1003(1)	1.1002(1)	1.1099
a(H-C-O)	121.66(3)	121.65(1)	121.62(1)	121.84
Rms resid. [MHz]	0.0028	0.0003	0.0003	-
Mean Δ_e [uÅ^2]	0.06455	0.00297	0.00223	-
CH_2ClF^g				
r(C-H)	1.0891(13)	1.0840(1)‡,†	1.0842(1)†	1.0896
r(C-F)	1.3706(9)	1.3594(1)	1.3591(1)	1.3682
r(C-Cl)	1.7613(6)	1.7641(1)	1.7645(1)	1.7998
a(H-C-Cl)	109.34(38)	107.96(1)	107.93(1)	107.61
a(H-C-H)	110.21(14)	112.57(1)	112.55(1)	113.23
a(F-C-Cl)	110.18(99)	110.02(1)	110.02(2)	110.15
Rms resid. [MHz]	0.0008	0.0001	0.0001	-
CH_2CHF				
r(C1-F)	1.3574(29)	1.3424(2)$^{\div,\dagger}$	1.3412(5)†	1.3528
r(C1-H)	1.0921(18)	1.0792(1)	1.0784(4)	1.0856
r(C1-C2)	1.3169(31)	1.3213(2)	1.3234(6)	1.3247
r(C2-H_{trans})	1.0774(21)	1.0772(1)	1.0768(4)	1.0839
r(C2-H_{cis})	1.0854(14)	1.0785(1)	1.0782(3)	1.0848
a(F-C1-H)	107.70(62)	112.10(6)	112.36(19)	111.62
a(F-C1-C2)	121.57(3)	121.72(1)	121.68(1)	121.90
a(C1-C2-H_{trans})	118.83(21)	118.95(1)	118.94(4)	119.25
a(C1-C2-H_{cis})	121.02(19)	121.32(1)	121.29(3)	121.72
Rms resid. [MHz]	0.0023	0.0001	0.0004	-
Mean Δ_e [uÅ^2]	0.09356	0.00281	0.00162	-
***cis*-CHFCHCl^h**				
r(C1-Cl)	1.7271(19)	1.7129(2)$^{\div,\dagger}$	1.7124(14)†	1.7404
r(C1-H)	1.1109(19)	1.0795(2)	1.0795(14)	1.0818
r(C1=C2)	1.3162(26)	1.3244(2)	1.3266(19)	1.3278
r(C2-F)	1.3363(21)	1.3313(2)	1.3306(16)	1.3416
r(C2-H)	1.0858(16)	1.0796(1)	1.0776(13)	1.0849
a(Cl-C1=C2)	123.13(14)	123.08(1)	123.08(13)	123.74
a(H-C1=C2)	126.88(23)	121.08(2)	121.06(19)	120.91
a(F-C2=C1)	122.27(20)	122.56(2)	122.47(15)	123.10
a(H-C2=C1)	124.16(21)	123.49(2)	123.33(16)	123.45
Rms resid. [MHz]	0.0026	0.0002	0.0020	-
Mean Δ_e [uÅ^2]	0.19860	0.00712	0.01604	-
oxirane^i				
r(C-C)	1.4719(4)	1.4609(2)$^{\div,\dagger}$	1.4615(2)†	1.4674
r(C-O)	1.4357(2)	1.4274(1)	1.4281(1)	1.4324
r(C-H)	1.0823(3)	1.0816(2)	1.0814(2)	1.0889
a(C-O-C)	61.67(2)	61.56(1)	61.55(1)	61.63
a(H-C-H)	116.63(4)	116.25(2)	116.33(2)	115.75
a(H-C-O)	114.75(5)	114.87(3)	114.82(3)	115.04
Rms resid. [MHz]	0.0031	0.0015	0.0015	-
dioxirane^j				
r(C-O)	1.3914(4)	1.3846(5)$^{\div}$	1.3850(1)†	1.3901
r(O-O)	1.5192(1)	1.5133(5)	1.5140(1)	1.5006
r(C-H)	1.0837(11)	1.0853(15)	1.0850(1)	1.0919
a(H-C-H)	116.70(13)	117.03(20)	117.06(1)	116.96
Rms resid. [MHz]	0.0065	0.25	0.0006	-

***trans*-glyoxal**^{k}				
r(C=O)	1.2135(10)	1.2046(1)$^{\div,\dagger}$	1.2051(1)†	1.2069
r(C-C)	1.5155(15)	1.5157(1)	1.5149(2)	1.5262
r(C-H)	1.1031(6)	1.1006(1)	1.1006(1)	1.1093
a(H-C-C)	115.42(9)	115.23(1)	115.37(1)	115.14
a(O=C-H)	123.80(8)	123.60(1)	123.45(1)	123.45
Rms resid. [MHz]	0.0019	0.0001	0.0002	-
***cis*-acrolein**				
r(C1-C2)	1.4884(13)	1.4806(1)$^{\div,\dagger}$	1.4809(3)†	1.4840
r(C2-C3)	1.3389(12)	1.3350(1)	1.3368(2)	1.3377
r(C1-O)	1.2124(10)	1.2108(1)	1.2102(2)	1.2145
r(C1-H)	1.1047(10)	1.1024(1)	1.1021(2)	1.1113
r(C2-H)	1.0864(9)	1.0824(1)	1.0807(2)	1.0885
r(C3-H$_{cis}$)	1.0984(14)	1.0808(1)	1.0800(3)	1.0868
r(C3-H$_{trans}$)	1.0796(10)	1.0797(1)	1.0786(2)	1.0857
a(C1-C2-C3)	121.39(10)	121.21(1)	121.33(2)	122.28
a(O-C1-C2)	124.06(10)	123.96(1)	123.88(2)	124.66
a(C2-C1-H)	115.33(11)	115.83(1)	115.81(2)	115.24
a(C3-C2-H)	121.32(12)	121.57(1)	121.63(2)	121.21
a(C2-C3-H$_{cis}$)	118.59(10)	119.85(1)	119.86(2)	120.35
a(C2-C3-H$_{trans}$)	121.49(17)	121.61(1)	121.66(4)	121.73
Rms resid. [MHz]	0.0014	0.0001	0.0003	-
Mean Δ_e [uÅ^2]	-0.02202	0.01311	0.01346	-
***trans*-acrolein**				
r(C1-C2)	1.4803(10)	1.4702(1)$^{\div,\dagger}$	1.4703(1)†	1.4735
r(C2-C3)	1.3393(14)	1.3354(1)	1.3355(1)	1.3384
r(C1-O)	1.2122(12)	1.2103(1)	1.2109(1)	1.2142
r(C1-H)	1.1096(17)	1.1048(1)	1.1044(1)	1.1133
r(C2-H)	1.0809(14)	1.0814(1)	1.0817(1)	1.0876
r(C3-H$_{cis}$)	1.0872(15)	1.0825(1)	1.0826(1)	1.0883
r(C3-H$_{trans}$)	1.0833(14)	1.0795(1)	1.0792(1)	1.0856
a(C1-C2-C3)	120.18(8)	120.18(1)	120.21(1)	121.02
a(O-C1-C2)	123.66(13)	124.02(1)	123.97(1)	124.21
a(C2-C1-H)	114.70(12)	115.08(1)	115.11(1)	115.02
a(C3-C2-H)	122.79(15)	122.78(1)	122.85(1)	122.41
a(C2-C3-H$_{cis}$)	119.91(11)	120.46(1)	120.44(1)	120.89
a(C2-C3-H$_{trans}$)	121.71(18)	122.10(1)	122.07(1)	122.22
Rms resid. [MHz]	0.0013	0.0001	0.0001	-
Mean Δ_e [uÅ^2]	-0.01588	-0.00527	-0.00845	-
cyclobutene				
r(C1=C2)	1.3478(7)	1.3406(1)$^{\div,\dagger}$	1.3409(1)†	1.3420
r(C2-C3)	1.5210(2)	1.5141(1)	1.5149(1)	1.5193
r(C3-C4)	1.5727(15)	1.5639(1)	1.5646(2)	1.5736
r(C1-H)	1.0807(5)	1.0805(1)	1.0801(1)	1.0867
r(C3-H)	1.0923(3)	1.0894(1)	1.0892(1)	1.0961
r(C1-C2-C3)	94.24(4)	94.23(1)	94.23(1)	94.37
r(C1-C2-H)	133.59(7)	133.42(1)	133.47(1)	133.47
a(C4-C3-H)	114.57(5)	114.64(1)	114.60(1)	114.78
a(H-C3-H)	109.22(3)	109.09(1)	109.19(1)	108.59
Rms resid. [MHz]	0.0011	0.0001	0.0001	-

pyridazine^{l}				
r(N2-C3)	1.3395(100)	1.3302(12)$^{\aleph 0,\dagger}$	1.3324(24)†	1.3366
r(C3-C4)	1.3948(96)	1.3938(12)	1.3926(23)	1.3971
r(C4-C5)	1.3865(70)	1.3761(16)	1.3778(16)	1.3829
r(C4-H)	1.0797(23)	1.0802(4)	1.0791(5)	1.0858
r(C3-H)	1.0822(15)	1.0810(3)	1.0804(4)	1.0871
a(C3-C4-C5)	116.85(18)	116.85(3)	116.86(4)	116.88
a(N2-C3-C4)	123.91(22)	123.86(4)	123.87(5)	123.67
a(C4-C3-H)	121.50(46)	121.35(6)	121.39(11)	121.43
a(C5-C4-H)	122.26(29)	122.37(4)	122.32(7)	122.25
Rms resid. [MHz]	0.0054	-	0.0013	-
Mean Δ_e [uÅ^2]	0.03281	-	-0.00103	-

All computations have been performed in this work except where otherwise indicated.

a) Graphical symbols denote the basis sets used in the calculations of $\Delta B^{\beta}_{\mathrm{vib}}$ contributions: ÷ VTZ; ‡ CVTZ; ≍ VQZ; × CVQZ; + wCVQZ, ℵ0 ANO0. † denotes the inclusion of $\Delta B^{\beta}_{\mathrm{el}}$. For all the structures calculated in this work, the uncertainties on the geometrical parameters are reported within parentheses, rounded to $1 \cdot 10^{-4}$ Å for lengths and $1 \cdot 10^{-2}$ degrees for angles if smaller than these values.

$\Delta_e = I^C - I^B - I^A$ is the inertial defect.

b) CCSD(T) $\Delta B^{\beta}_{\mathrm{vib}}$ from ref. 92.

c) CCSD(T) $\Delta B^{\beta}_{\mathrm{vib}}$ from ref. 93.

d) CCSD(T) $\Delta B^{\beta}_{\mathrm{vib}}$ from ref. 94.

e) CCSD(T) $\Delta B^{\beta}_{\mathrm{vib}}$ from ref. 95.

f) CCSD(T) $\Delta B^{\beta}_{\mathrm{vib}}$ from ref. 96.

g) CCSD(T) $\Delta B^{\beta}_{\mathrm{vib}}$ from ref. 97.

h) CCSD(T) $\Delta B^{\beta}_{\mathrm{vib}}$ from ref. 98.

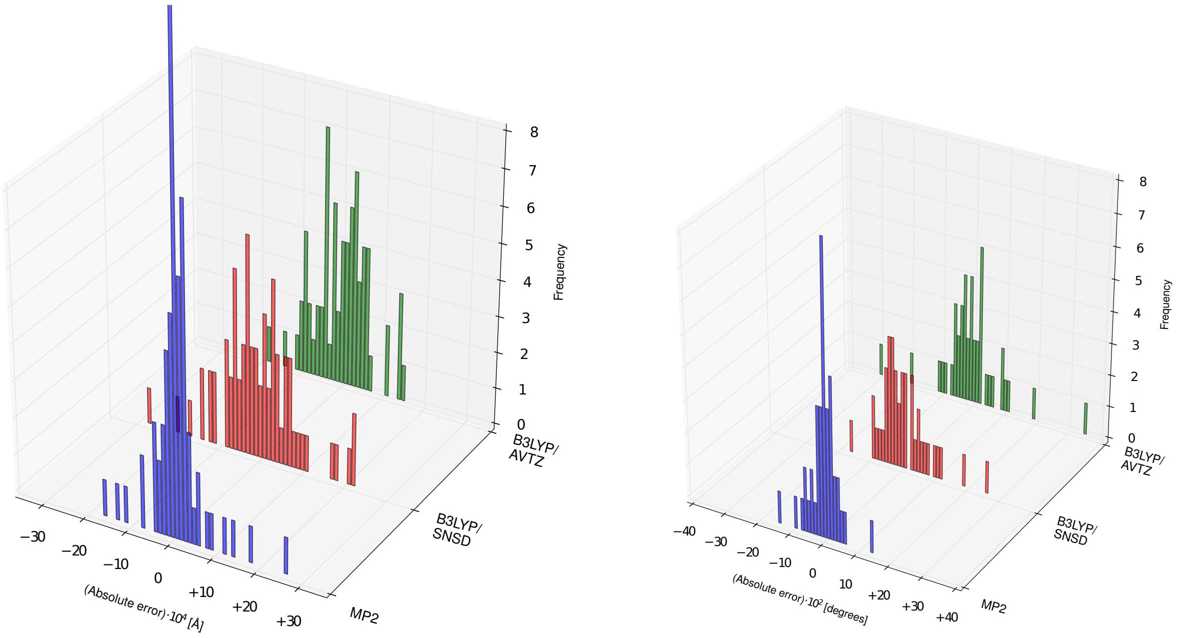
i) CCSD(T) $\Delta B^{\beta}_{\mathrm{vib}}$ from ref. 99.

j) CCSD(T) r_e^{SE} from ref. 85.

k) CCSD(T) $\Delta B^{\beta}_{\mathrm{vib}}$ from ref. 100.

l) CCSD(T) r_e^{SE} from ref. 88.

Figure 3: Statistical distributions of the MP2, B3LYP/SNSD and B3LYP/AVTZ deviations from CCSD(T) SE equilibrium parameters for the molecules belonging to the CCse set (see Table 1 and Table 6 in the SI).



(a) all bonds

(b) all angles

Table 2: Mean, standard deviation, and mean absolute error (MAE) for the MP2, B3LYP/SNSD and B3LYP/AVTZ deviations from CCSD(T) SE equilibrium parameters for the molecules belonging to the CCse set (see Table 1 and Table 6 in the SI). For the different types of bonds, only the sets having at least 7 items have been considered.

	MP2[a]	B3LYP/SNSD	B3LYP/AVTZ
All bonds (68 items)			
Mean	+0.0001	−0.0001	+0.0000
St. Dev.	0.0006	0.0009	0.0007
MAE	0.0004	0.0007	0.0005
CH bonds (27 items)			
Mean	+0.0002	−0.0005	−0.0005
St. Dev.	0.0003	0.0005	0.0005
MAE	0.0002	0.0006	0.0006
CC bonds (18 items)			
Mean	+0.0001	+0.0006	+0.0004
St. Dev.	0.0007	0.0010	0.0007
MAE	0.0005	0.0009	0.0007
CO bonds (7 items)			
Mean	+0.0001	+0.0004	+0.0003
St. Dev.	0.0002	0.0005	0.0002
MAE	0.0001	0.0005	0.0003
All angles (42 items)			
Mean	+0.00	+0.00	−0.01
St. Dev.	0.04	0.08	0.10
MAE	0.03	0.06	0.06

a) all MP2 calculations have been performed with basis sets of at least triple-ζ quality, Table 6 in the SI for details.

Table 3: Parameters for linear regressions of the CCSD(T) r_e^{SE} parameters versus the MP2, B3LYP/SNSD and B3LYP/AVTZ r_e^{SE} ones for the molecules belonging to the CCse set (see Table 1 and Table 6 in the SI).

	MP2[a]	B3LYP/SNSD	B3LYP/AVTZ
All bonds			
A	0.999740	0.998023	1.001581
B	0.000186	0.002475	−0.001953
R^2	0.999989	0.999976	0.999988
St. Dev.	0.000405	0.000582	0.000421
All angles			
A	1.000144	1.000004	1.000561
B	−0.012640	−0.005008	−0.058692
R^2	0.999985	0.999956	0.999928
St. Dev.	0.000609	0.001049	0.001347
$r_e^{\mathrm{SE}}(\mathrm{CCSD(T)}) = A \cdot r_e^{\mathrm{SE}}(\mathrm{MP2\ or\ B3LYP}) + B$			

a) all MP2 calculations have been performed with basis sets of at least triple-ζ quality, Table 6 in the SI for details.

From small to medium-large systems: B3LYP/SNSD semi-experimental structures

In the previous section we demonstrated that the SE equilibrium structures derived from $\Delta B^{\beta}_{\mathrm{vib}}$ contributions calculated at the B3LYP level have an accuracy comparable to that obtained when using CCSD(T) corrections. In view of these results, and aiming at increasing the number of geometrical patterns considered, in this section B3LYP SE equilibrium structures are presented for 26 organic molecules containing H, C, N, O, F, S, Cl atoms. For the systems considered, to the best of our knowledge, SE equilibrium structures derived from CCSD(T) vibrational contributions are not available, but a sufficient number of isotopologues has been characterized experimentally to allow for a reliable determination of all geometrical parameters without any constraint and/or assumption. Focusing on the B3LYP/SNSD quantum mechanical model, which permits to keep the computational costs low, the new SE equilibrium structures are compared with both the most accurate determinations available in the literature and vibrationally averaged r_0 geometries. Together with the 21 molecules previously considered, a high-quality benchmark set, including a total of 47 molecules (hereafter referred to as B3se set), has been set up for validating structural predictions from other experimental and/or computational approaches.

The geometrical parameters for CH_2F_2, CCl_2F_2, CH_2Cl_2, $CHClF_2$, ethene, ethenol, propene, butadiene, *cis*-hexatriene, cyclopropane, aziridine, benzene, pyrrole, pyrazole, imidazole, furan, thiophene, maleic anhydride, dimethyl ether, *cis* and *trans*-formic acid, *cis*-methyl formate, glycolaldehyde, and propanal are collected in Table 4 and compared with the best r_e^{SE} equilibrium structures available in the literature. In Table 5, we present the SE equilibrium structures for two additional molecules (glyoxylic acid and pyridine), which are then compared with the best theoretical r_e structures available. For all these molecules, the $(B_0^{\beta})^{\mathrm{EXP}}$, $\Delta B^{\beta}_{\mathrm{vib}}$ and $\Delta B^{\beta}_{\mathrm{el}}$ are summarized in Tables 2 of the SI. It is noteworthy that for most of the systems the RMS of the residuals for r_0 geometries is about one order of magnitude larger than the RMS of the residuals for SE equilibrium geometries. The small values for the latter, less than 7 kHz for all systems, demonstrate the good quality of the fits.

In analogy with the CCse set, equilibrium geometries obtained at the B3LYP/SNSD level are reported in Tables 4 and 5 because of their subsequent use within the template approach presented in the next section.

Halomethanes

A systematic evaluation of the SE equilibrium structure for a series of chlorinated and fluorinated methanes has been carried out recently.[101] In this work, in addition to the SE equilibrium structure of CH_2ClF reported in the previous section, CH_2F_2, CCl_2F_2, CH_2Cl_2, and $CHClF_2$ have been considered as models to investigate the C-X bond pattern, where X is an halogen atom and the C hybridization is sp^3. CH_2F_2, CCl_2F_2 and CH_2Cl_2 have C_{2v} symmetry and are completely characterized by 5 geometrical parameters, while $CHClF_2$ belongs to the C_s symmetry group and has 6 unique structural parameters. In ref. 101, the $\Delta B^{\beta}_{\mathrm{vib}}$ contributions have been calculated at the MP2/VTZ (with the modified V(T+d)Z for the chlorine atom[102]) level of theory for the halomethanes considered, except for CCl_2F_2 (B3LYP/6-311+G(3df,2pd)). For all these systems, there is a good agreement between the SE equilibrium structures obtained employing B3LYP and MP2 vibrational contributions. The C-F bond length shows the largest variation with the number of hydrogens bonded to the C atom, i.e., it increases from 1.3286 Å for CCl_2F_2 (no H atoms) to 1.3363 Å for $CHClF_2$ (one H atom) and to 1.3533 Å/1.3594 Å for CH_2F_2/CH_2ClF (two H atoms). A similar trend is shown by the CCl bond length, which changes from 1.7641/1.7642 Å for CH_2ClF/CH_2Cl_2 to 1.7558 Å for $CHClF_2$ and to 1.7519 Å for CCl_2F_2. On the contrary, the CH bond length is only marginally affected by the number of halogen atoms bonded to the C atom (1.0810, 1.0840, 1.0849 and 1.0867 Å for CH_2Cl_2, CH_2ClF, $CHClF_2$, and CH_2F_2, respectively).

Substituted alkene compounds

Together with CH_2CHF and *cis*-CHFCHCl (presented in Table 1), ethene, ethenol, butadiene, *cis*-hexatriene and propene have been studied as representatives of the Y-C=C-X bond pattern for non cyclic molecules, where C is sp^2 hybridized and X and Y are either halogens or C atoms.

Ethene and ethenol (or vinyl alcohol) are the simplest alkene and enol compound, respectively. A r_e^{SE} structure for ethene, which is defined by 3 internal parameters (D_{2h} symmetry), is available in the literature.[103] It corresponds to a weighted average of different r_e^{SE} geometries calculated by use of $\Delta B_{\mathrm{vib}}^{\beta}$ at the MP2 and B3LYP levels, in conjunction with basis sets of at least triple-ζ quality, and where scaled quadratic force fields have been coupled with unscaled cubic force fields in the vibrational corrections calculations. The uncertainties of 0.0010 Å and 0.10 degrees on the parameters of the structure of ref. 103 (see Table 4) include both the uncertainties related to the SE methodology and those estimated from the parameter differences found using the different QM models in the $\Delta B_{\mathrm{vib}}^{\beta}$ calculations. All the B3LYP/SNSD r_e^{SE} results, obtained by fitting the SE I_e^A and I_e^B moments of inertia, coincide with those of ref. 103 within the respective error bars.

The *syn* conformer of ethenol is fairly rigid and completely defined by 11 internal parameters (C_s symmetry). The SE equilibrium structure recently determined using $\Delta B_{\mathrm{vib}}^{\beta}$ computed at the MP2/VQZ level[90] is given in Table 4 together with the B3LYP/SNSD r_e^{SE} results, obtained by fitting the SE I_e^A and I_e^C moments of inertia. The agreement is extremely good and the small RMS residual and uncertainties on the fitted parameters indicate that the B3LYP/SNSD SE equilibrium geometry is also accurate.

Propene is the simplest mono-methyl internal rotor, and it has been largely studied by infrared and microwave spectroscopy (see Refs. 104,105 and references therein). As a consequence, experimental rotational constants are available for a large number of isotopologues (20). The molecular structure of propene (C_s symmetry with a synperplanar arrangement of the C1=C2-C3-H_{plane} moiety) is defined by 15 geometrical parameters, and has recently been evaluated by means of the SE approach using $\Delta B_{\mathrm{vib}}^{\beta}$ contributions at the MP2/VTZ(fc) level.[105] Some remarks on the fitting procedure need to be made. Due to large uncertainties affecting the $(B_0^A)^{\mathrm{EXP}}$ of some isotopologues[106] that lead to ill-conditioned results, the fit has been performed on the SE equilibrium moments of inertia corresponding to the $(B_0^B)^{\mathrm{EXP}}$ and $(B_0^C)^{\mathrm{EXP}}$ rotational constants. Moreover, the CHD_{cis}=$\mathrm{CDCH_3}$ and $\mathrm{CH_2}$=$^{13}\mathrm{CHCH_3}$ isotopologues have been excluded from the fit because of the corresponding large residuals affecting the equilibrium rotational constants. In this framework, B3LYP/SNSD vibrational contributions lead to residuals with a very small RMS. Some fitted geometrical parameters defining the methyl hydrogen atoms lying outside the molecular C-C-C plane (in particular the C3-H_{out} bond length (1.0895 Å) and the C1=C2-C3-H_{out} dihedral angle (120.47 degrees)) are significantly smaller than the corresponding values obtained using MP2 vibrational contributions (1.0949 Å and 121.08 degrees). The latter values are closer to their r_0 counterparts (1.1036 Å and 121.07 degrees) than our equilibrium ones. In contrast to the B3LYP trend, the MP2 SE C3-H_{out} bond length differs significantly from the other C-H bonds, which range between 1.0805 and 1.0862 Å.

Butadiene and *cis*-hexatriene are planar C_{2h} and C_{2v} molecules, respectively, belonging to the class of polyenes, which are of great importance in biology and organic electronics due to a π-electron delocalization that increases as the C=C chain gets longer. In analogy with ethene, a r_e^{SE} structure for butadiene was obtained by Craig and coworkers from the average of different MP2 and B3LYP r_e^{SE} geometries. The B3LYP/SNSD r_e^{SE} parameters, obtained by fitting the SE I_e^A and I_e^C moments of inertia, agree with those of ref. 103 within the respective error bars.

The B3LYP/SNSD SE equilibrium structure of *cis*-hexatriene is in good agreement with that obtained using a MP2/VTZ force field.[91] The largest discrepancy (about 0.0028 Å) is observed for the C3-H bond length. The r_0 structure shows consistently longer bond lengths, up to about 0.01 Å, again for the C3-H bond. Contrary to the molecules discussed above, in the case of *cis*-hexatriene the inclusion of the $(\Delta B_0^{\beta})^{\mathrm{QM}}$ terms in the fitting procedure only leads to a small reduction of the RMS residuals (1.7 kHz for r_e^{SE} with respect to 2.2 kHz for r_0).

It is noteworthy that the lengthening of the C=C double bonds as a consequence of π-electron delocalization is well reproduced: the C=C bond length is 1.3211/1.3245 Å in CH_2CHF/*cis*-CHFCHCl (a single C=C bonded to halogen atoms), 1.3326 Å in propene (a single C=C bond linked to a methyl group) and 1.3418/1.3509 Å in *cis*-hexatriene (three C=C groups). For the latter, the length of the central C=C bond (large conjugation) is 1.3509 Å and that of the terminal C=C one (lower conjugation) is 1.3418 Å, thus well reproducing the expected behaviour. It is also interesting to note that the C-C single bond length decreases from 1.4956 Å in propene ($sp^2 - sp^3$ type without any conjugation) to 1.4510 Å (conjugated $sp^2 - sp^2$ bond) in *cis*-hexatriene.

Cyclic and heterocyclic compounds

Cyclic and heterocyclic compounds are important building blocks of organic and biological molecules. Together with cyclobutene reported in Table 1, which is one of the smallest cycloalkenes, in this work we have studied cyclopropane and benzene (Table 4), which are among the simplest cycloalkanes and aromatic systems, and oxirane, dioxirane, pyridazine (Table 1), aziridine, pyrrole, pyrazole, imidazole, furan, thiophene, maleic anydride (Table 4) and pyridine (Table 5), as prototypical heterocyclic compounds.

Cyclopropane belongs to the D_{3h} symmetry group and it is completely defined by 3 geometrical parameters: the C-C and C-H distances and the HCH angle. The SE equilibrium structure has been previously determined by using a SDQ-MBPT(4)/VTZ cubic force field.[107] Though two rotational constants of the parent species (B_0^B and B_0^C), B_0^B of C_3D_6 and B_0^A, B_0^B, B_0^C of $C_3H_4D_2$ have been experimentally determined, the inclusion of all of them in the fitting procedure leads to large residuals, as also noticed in ref. 107. The geometry reported in Table 4 has been obtained by using the SE equilibrium moments of inertia of $C_3H_4D_2$, together with the SE I_e^B of the parent species. Thanks to its high symmetry (D_{6h}), the structure of benzene is defined by only 2 geometrical parameters: the C-H and C-C bond lengths. Its SE equilibrium structure has been determined for the first time by Stanton *et al.*[108] using vibrational contributions at the MP4(SDQ)/VTZ level. The B3LYP/SNSD SE equilibrium structures of cyclopropane and benzene show small uncertainties on the geometrical parameters and are in good agreement with the previous determinations.

Aziridine, also called ethylene imine, is one of the simplest non-aromatic N-heterocycles. Its equilibrium structure (C_s symmetry) is completely determined by 10 geometrical parameters, and is characterized by a high nitrogen inversion barrier. The rotational spectrum of aziridine has been studied in great detail because of its potential astrophysical interest.[109–111] Very recently, a SE equilibrium structure has been determined by combining the experimental ground-state rotational constants with $\Delta B^{\beta}_{\mathrm{vib}}$ contributions computed at the MP2/VTZ level.[112] The B3LYP/SNSD SE equilibrium structure has been derived by fitting SE equilibrium inertia moments, all equally weighted. The resulting SE equilibrium structure shows a very small RMS residual and a good agreement with the MP2/VTZ one.

Pyrazole and imidazole are two five-membered heteroaromatic rings, with adjacent and non-adjacent nitrogen atoms, respectively. Both molecules are completely characterized by 15 geometrical parameter, and have C_s symmetry. Pyrazole is used in the synthesis of many medical/organic molecules, while imidazole is present in important biological building-blocks, such as histidine and the related hormone histamine. The B3LYP/SNSD r_e^{SE} geometries shown in Table 4 have been obtained by fitting the SE I_e^B and I_e^C moments of inertia for pyrazole, and I_e^A and I_e^C for imidazole, all equally weighted.

For imidazole, the experimental rotational constants used were taken from ref. 113 without applying any corrections, while in ref. 112 the experimental values were corrected by the contribution of theoretical quartic distortion constants within the predicated method. In spite of these methodological differences, the B3LYP/SNSD r_e^{SE} parameters are in good agreement with the results of ref. 112.

Pyrrole, furan, and thiophene are three planar heterocyclic molecules belonging to the C_{2v}

point group, whose structures are completely defined by 9, 8 and 8 parameters, respectively. For pyrrole and furan, the best SE equilibrium structures reported in the literature were determined by correcting the vibrational ground-state rotational constants with $\Delta B^{\beta}_{\mathrm{vib}}$ contributions calculated at the MP2/wCVTZ and MP2/VTZ levels, respectively.[112,114] The best SE equilibrium geometry of thiophene was derived from a combined use of electron diffraction (ED), microwave spectroscopy (MW) and computed vibrational contributions at the B3LYP/6-311+G* level.[115] The B3LYP/SNSD SE equilibrium geometries of pyrrole and furane are in good agreement with those already available. On the contrary, the B3LYP/SNSD SE equilibrium parameters of thiophene collected in Table 4 show relevant differences with respect to those previously determined. For example, the B3LYP/SNSD SE value for r(CS) is 0.0087 Å longer than that of ref. 115, while the B3LYP/SNSD SE C=C bond length is 0.0095 Å shorter than the corresponding value of ref. 115.

It is interesting to note how the C-C bond lengths change when both the H atoms linked with the C atom in α-position with respect to the O atom of the ring are substituted with two O atoms, that is, when moving from furan to maleic anhydride. To the best of our knowledge, the most accurate SE equilibrium structure available for maleic anydride (C_{2v} symmetry) was derived using a MP2/VTZ cubic force field, also including the non negligible contribution due to $\Delta B^{\beta}_{\mathrm{el}}$.[89] As a matter of fact, for B3LYP calculations the inclusion of the latter contributions reduces the RMS residual from 8.5 kHz to 0.4 kHz and the SE equilibrium inertial defect from a mean value of -0.01634 uÅ^2 to -0.00779 uÅ^2. For this molecule, the MP2 and B3LYP SE equilibrium structures agree very well one another. From Table 4 we note a significant decrease of the C3-C4 bond length when moving from furan to maleic anhydride (1.4344 Å in furan with respect to r(C1-C2)=1.3320 Å in maleic anhydride, see Figure 1 d) and a contemporary increase of the C2-C3 bond length (1.3542 Å in furan with respect to 1.4857 Å in maleic anhydride), and of the ring C-O distance (from 1.3598 Å in furan to 1.3848 Å in maleic anhydride).

Thanks to the large number of isotopologues experimentally investigated[116,117] (see Table 2 in SI) and to the limited number of independent geometrical parameters (10), it is possible to determine a full SE equilibrium structure for pyridine (C_{2v} symmetry). The r_e^{SE} structure given in Table 5 has been obtained by fitting the SE I_e^A and I_e^C moments of inertia derived from the experimental rotational constants corrected by $\Delta B^{\beta}_{\mathrm{vib}}$ contributions calculated at the B3LYP/SNSD level. Even in this case, the inclusion of the $\Delta B^{\beta}_{\mathrm{el}}$ terms leads to a considerable improvement of the inertial defects. The CCSD(T)/CBS+CV and B3LYP/SNSD structures determined in this work are reported in Table 5 for comparison purposes and a subsequent use within the template approach (see next section). The CCSD(T)/CBS+CV and B3LYP SE equilibrium structures remarkably agree one to other and also with the r_e^{SE} determined in ref.[112] using B3LYP/6-311+G(3df,2pd) vibrational corrections and the so-called predicate approach. It is worthwhile noting that the SE equilibrium structures of benzene and pyridine derived using B3LYP/SNSD vibrational contributions show quite different C-C bond lengths: 1.3919 Å in benzene versus 1.3907 Å for the bond directly connected to the C-N bond and 1.3888 Å for the furthest bond in pyridine.

Ethers, aldehydes, esters and carboxylic acids

In addition to *trans*-glyoxal, *cis*- and *trans*-acrolein (see Table 1), dimethyl ether, glycolaldehyde, propanal, formic and glyoxylic acids as well as methyl formate have been investigated as models for the most significant oxygen-containing moieties.

Dimethyl ether, the simplest molecule with two internal rotors, has been studied in great detail as an interstellar molecule and because of the interest in its rotational-torsional spectrum.[118–120] Its equilibrium structure has C_{2v} symmetry (characterized by antiperiplanar arrangement of both the C-O-C-H_{plane} moieties) and is completely defined by 7 geometrical parameters. The SE equilibrium structure determined using B3LYP/SNSD vibrational contributions is in remarkable agreement with that obtained in ref.[90] using an MP2/VTZ cubic force field (see Table 4).

Formic acid (C_s symmetry) can be considered the prototype of carboxylic acids and presents two rotamers, the *cis* and *trans* forms. The SE equilibrium structures of both forms have been previously obtained by combining the experimental ground-state rotational constants of several isotopologues (11 and 7 for the *cis* and *trans* form, respectively) with $\Delta B^{\beta}_{\text{vib}}$ calculated from a MP2/VTZ cubic force field.[121] As shown in Table 4, for both conformers, the SE equilibrium structures issuing from B3LYP vibrational contributions are in very good agreement with the reference SE results.

cis-methyl formate is an important interstellar molecule and is considered the prototype system for studying the internal rotation of a methyl group.[122] At equilibrium, *cis*-methyl formate possesses a symmetry plane with one pair of equivalent out-of-plane hydrogen atoms (C_s symmetry) and d(C-O-C_m-H_{plane}) = 180.00 degrees, where H_{plane} is the methyl hydrogen on the symmetry plane. In ref. 122, $\Delta B^{\beta}_{\text{vib}}$ contributions derived from a MP2/VTZ cubic force field were combined with the available experimental rotational constants. The agreement between the MP2/VTZ and B3LYP/SNSD results is good for all parameters that are not related to the H_{plane} atom. In fact, quite large discrepancies are found for both the C_m-H_{plane} bond length (about 0.0052 Å) and the O-C_m-H_{plane} angles (about 0.69 Å). As noted for propene, the B3LYP/SNSD r_e^{SE} shows a smaller difference between the C_m-H_{plane} and C_m-H_{out} bond lengths than the MP2 SE equilibrium structure.

Glycolaldehyde can be considered the simplest sugar. Only the *syn* conformer, which is stabilized by an intramolecular hydrogen bond, has been observed by microwave spectroscopy.[123–127] It has C_s symmetry and is completely defined by 12 geometrical parameters. Recently, a SE equilibrium structure was determined by combining the ground-state experimental rotational constants with $\Delta B^{\beta}_{\text{vib}}$ contributions at the MP2/VTZ level.[90] The agreement with the new B3LYP/SNSD SE equilibrium structure is generally good, except for some small discrepancies on r(C1-C2) (1.5014 versus 1.5003 Å), r(O-H) (0.9618 versus 0.9593 Å), and a(C1-C2-H) (107.80 versus 108.11 degrees). It is noteworthy that the B3LYP/SNSD r_e^{SE} is in remarkable agreement with the high level fully theoretical r_e (referred to as r_e^{BO} in Table 6 of ref. 90): r(C1-C2)=1.5016 Å, r(O-H)=0.9653 Å and a(C1-C2-H)=107.794 Å.

The *syn* conformer of propanal, or propionaldehyde, which is significantly more stable than its *gauche* counterpart, has C$_s$ symmetry (with d(C1-C2-C-H$_{plane}$) = 180.00) degrees) and is completely defined by 15 geometrical parameters. Once again, Table 4 shows that MP2/VTZ [90] and B3LYP/SNSD cubic force fields lead to very similar SE equilibrium structures.

Finally, in Table 5 we report the first determination of the SE equilibrium structure of glyoxylic acid (C_s symmetry), which is the simplest α-oxoacid and is completely determined by 11 geometrical parameters. The B3LYP/SNSD SE equilibrium structure has been obtained by fitting the SE I_e^B and I_e^C moments of inertia of 8 out of 9 experimentally observed isotopologues (see Table 2), where the H^{13}COCOOH isotopologue has been excluded from the fit because of the large residuals shown by the fitted equilibrium rotational constants. In Table 5, the B3LYP SE equilibrium structure is compared with the theoretical r_e equilibrium geometry optimized at the CCSD(T)/VQZ level, taken from ref. 128. The agreement between these two structures is remarkable.

Table 4: r_0, r_e^{SE} and r_e geometries of CH_2F_2, CCl_2F_2, CH_2Cl_2, $CHClF_2$, ethene, ethenol, propene, butadiene, *cis*-hexatriene, cyclopropane, aziridine, benzene, pyrrole, pyrazole, imidazole, furan, thiophene, maleic anhydride, dimethyl ether, *cis* and *trans*-formic acid, *cis*-methyl formate, glycolaldehyde and propanal. Distances in Å, angles in degrees.

	$r_0{}^a$	$r_e^{\mathrm{SE}\,a}$		r_e
		Literature	B3LYP/SNSD	B3LYP/SNSD
		– halomethanes –		
CH_2F_2^{b}				
r(C-F)	1.3596(4)	1.35323(1)†	1.3533(1)†	1.3668
r(C-H)	1.0871(11)	1.08703(3)	1.0867(2)	1.0935
a(F-C-F)	108.06(3)	108.282(2)	108.29(1)	108.44
a(H-C-H)	113.42(13)	113.442(9)	113.48(2)	113.74
a(H-C-F)	108.81(5)	108.750(2)	108.74(1)	108.64
Rms resid. [MHz]	0.0082	-	0.0013	-
CCl_2F_2^{c}				
r(C-F)	1.3511(59)	1.3287(8)†	1.3286(7)†	1.3372
r(C-Cl)	1.7395(50)	1.7519(7)	1.7519(6)	1.7857
a(F-C-F)	105.57(47)	107.75(9)	107.77(6)	108.04
a(Cl-C-Cl)	113.22(35)	111.62(7)	111.61(4)	111.83
a(Cl-C-F)	109.44(20)	109.35(1)	109.34(3)	109.22
Rms resid. [MHz]	0.0060	-	0.0008	-
CH_2Cl_2^{d}				
r(C-H)	1.0743(15)	1.0816(2)†	1.0810(8)†	1.0862
r(C-Cl)	1.7711(4)	1.76425(3)	1.7642(2)	1.7956
a(H-C-H)	111.34(16)	111.772(4)	111.79(9)	112.48
a(Cl-C-Cl)	111.92(3)	112.166(3)	112.18(2)	112.85
a(Cl-C-H)	108.40(6)	108.237(10)	108.23(3)	107.90
Rms resid. [MHz]	0.0122	-	0.0069	-
$CHClF_2$^{d}				
r(C-H)	1.0953(6)	1.0850(11)	1.0849(2)†	1.0901
r(C-F)	1.3450(11)	1.3363(5)	1.3363(4)	1.3466
a(C-Cl)	1.7465(20)	1.7560(9)	1.7558(8)	1.7915
a(H-C-F)	109.14(11)	109.97(4)	110.02(4)	110.14
a(H-C-Cl)	110.41(23)	109.60(6)	109.45(9)	109.16
a(F-C-Cl)	110.25(6)	109.62(4)	109.63(2)	109.58
a(F-C-F)	107.59(9)	108.06(6)	108.06(3)	108.23
Rms resid. [MHz]	0.0017	-	0.0006	-
		– substituted alkene compounds –		
ethene^e				
r(C=C)	1.3373(3)	1.3305(10)	1.3317(1)†	1.3322
r(C-H)	1.0836(3)	1.0805(10)	1.0805(1)	1.0870
a(C=C-H)	121.26(2)	121.45(10)	121.40(1)	121.71
a(H-C-H)	117.49(3)	117.10(10)	117.19(1)	116.58
Rms resid. [MHz]	0.0024	-	0.0002	-
Mean Δ_e [uÅ^2]	0.06148	-	0.00119	-

ethenol[f]				
r(O-H)	0.9595(17)	0.9604(2)†	0.9605(1)†	0.9668
r(C2-O)	1.3708(20)	1.3594(8)	1.3598(1)	1.3638
r(C2-H)	1.0893(17)	1.0794(4)	1.0789(1)	1.0860
r(C1=C2)	1.3333(21)	1.3312(9)	1.3316(1)	1.3344
r(C1-H$_{cis}$)	1.0882(13)	1.0816(2)	1.0812(1)	1.0873
r(C1-H$_{trans}$)	1.0762(14)	1.0772(4)	1.0770(1)	1.0831
a(C2-O-H)	108.51(14)	108.81(4)	108.70(1)	109.41
a(C1=C2-H)	127.14(73)	122.65(32)	122.58(8)	122.74
a(C1=C2-O)	126.11(4)	126.297(5)	126.26(1)	126.84
a(C2-C1-H$_{cis}$)	121.37(13)	121.90(4)	121.87(1)	122.29
a(C2-C1-H$_{trans}$)	119.62(19)	119.59(2)	119.58(1)	119.86
Rms resid. [MHz]	0.0022	-	0.0001	-
Mean Δ_e [uÅ^2]	0.04867	-0.00424	-0.00206	-
propene[g]				
r(C1=C2)	1.3408(16)	1.3310(7)†	1.3326(2)†	1.3340
r(C2-C3)	1.5042(15)	1.4956(7)	1.4956(2)	1.5004
r(C1-H$_{cis}$)	1.0920(17)	1.0834(6)	1.0818(2)	1.0882
r(C1-H$_{trans}$)	1.0789(21)	1.0805(12)	1.0804(2)	1.0862
r(C2-H)	1.0874(13)	1.0857(4)	1.0841(2)	1.0909
r(C3-H$_{plane}$)	1.0823(30)	1.0862(8)	1.0880(4)	1.0947
r(C3-H$_{out}$)	1.1036(56)	1.0949(9)	1.0895(7)	1.0976
a(C1=C2-C3)	124.11(6)	124.47(2)	124.43(1)	125.29
a(C2-C1-H$_{cis}$)	120.42(13)	121.08(5)	121.13(2)	121.56
a(C2-C1-H$_{trans}$)	121.53(26)	121.55(14)	121.31(3)	121.59
a(C1=C2-H)	118.50(29)	118.75(13)	118.84(4)	118.68
a(C2=C3-H$_{plane}$)	111.24(12)	111.10(4)	111.07(2)	111.55
a(C2-C3-H$_{out}$)	110.00(54)	110.53(11)	111.02(7)	111.03
d(C1=C2-C3-H$_{out}$)	121.07(59)	121.08(14)	120.47(8)	120.80
Rms resid. [MHz]	0.0017	-	0.0002	-
butadiene[e]				
r(C1=C2)	1.3450(12)	1.3376(10)	1.3386(1)†	1.3411
r(C2-C3)	1.4603(17)	1.4539(10)	1.4543(2)	1.4561
r(C1-H$_{cis}$)	1.0847(8)	1.0819(10)	1.0815(1)	1.0876
r(C1-H$_{trans}$)	1.0822(10)	1.0793(10)	1.0793(1)	1.0854
r(C2-H)	1.0848(8)	1.0847(10)	1.0839(1)	1.0903
a(C1=C2-C3)	123.32(4)	123.62(10)	123.53(1)	124.29
a(C1=C2-H)	119.91(12)	119.91(10)	119.76(1)	119.34
a(C2=C1-H$_{cis}$)	120.49(6)	120.97(10)	120.94(1)	121.41
a(C2=C1-H$_{trans}$)	121.22(7)	121.47(10)	121.43(1)	121.63
Rms resid. [MHz]	0.0023	-	0.0002	-
Mean Δ_e [uÅ^2]	0.02434	-	-0.00790	-

***cis*-hexatriene**^{h}				
r(C1=C2)	1.3421(10)	1.33993(28)†	1.3418(8)†	1.3437
r(C2-C3)	1.4599(23)	1.45041(38)	1.4510(17)	1.4510
r(C3=C4)	1.3507(32)	1.34997(87)	1.3509(24)	1.3550
r(C1-H$_{cis}$)	1.0808(29)	1.08255(35)	1.0815(22)	1.0876
r(C1-H$_{trans}$)	1.0809(24)	1.07982(28)	1.0800(18)	1.0853
r(C2-H)	1.1107(25)	1.08788(37)	1.0866(19)	1.0881
r(C3-H)	1.0916(14)	1.08417(28)	1.0813(10)	1.0895
a(C1=C2-C3)	122.85(28)	122.755(38)	122.59(21)	123.61
a(C2-C3=C4)	126.21(9)	126.273(19)	126.23(6)	127.04
a(C2=C1-H$_{cis}$)	121.03(29)	121.017(34)	121.05(22)	121.42
a(C2=C1-H$_{trans}$)	121.43(30)	121.456(37)	121.38(23)	121.60
a(C1-C2-H)	120.38(29)	119.320(37)	119.51(22)	118.45
a(C4=C3-H)	119.67(18)	118.035(27)	117.48(14)	117.47
Rms resid. [MHz]	0.0022	-	0.0017	-
Mean Δ_e [uÅ^2]	-0.17124	-0.005	-0.01153	-
– cyclic and heterocyclic compounds –				
cyclopropane^i				
r(C-C)	1.5144(2)	1.5030(10)	1.5031(1)†	1.5094
r(C-H)	1.0790(4)	1.0786(10)	1.0787(2)	1.0856
a(H-C-H)	115.38(4)	114.97(10)	114.94(2)	114.25
Rms resid. [MHz]	0.0036	-	0.0022	-
aziridine^j				
r(C-N)	1.4800(4)	1.47013(6)†	1.4714(1)†	1.4740
r(C-C)	1.4800(4)	1.47703(8)	1.4777(1)	1.4845
r(N-H)	1.0118(8)	1.01279(13)	1.0126(1)	1.0180
r(C-H$_{cis}$)	1.0821(7)	1.08099(13)	1.0805(1)	1.0877
r(C-H$_{trans}$)	1.0803(7)	1.07971(13)	1.0791(1)	1.0866
a(C-N-H)	109.36(5)	109.376(9)	109.16(1)	110.03
a(C-N-C)	60.42(2)	60.311(6)	60.28(1)	60.47
a(N-C-C)	59.79(1)	59.845(3)	59.86(1)	59.77
a(N-C-H$_{cis}$)	118.47(9)	118.28(2)	118.19(1)	118.61
a(N-C-H$_{trans}$)	114.49(10)	114.46(2)	114.38(1)	114.66
a(C-C-H$_{cis}$)	117.52(9)	117.829(14)	117.81(1)	117.99
a(C-C-H$_{trans}$)	119.15(7)	119.538(14)	119.49(1)	119.85
Rms resid. [MHz]	0.0017	-	0.0001	-
benzene^k				
r(C-C)	1.3970(2)	1.3914(10)	1.3919(1)†	1.3961
r(C-H)	1.0807(12)	1.0802(20)	1.0795(1)	1.0865
Rms resid. [MHz]	0.0266	-	0.0016	-
Mean Δ_e [uÅ^2]	-	-	-	

pyrrole^l				
r(C-N)	1.3766(6)	1.36940(17)†	1.3694(1)†	1.3755
r(C2-C3)	1.3792(7)	1.3723(2)	1.3732(1)	1.3794
r(C3-C4)	1.4275(23)	1.4231(4)	1.4228(2)	1.4256
r(N-H)	0.9936(5)	1.00086(14)	1.0007(1)	1.0081
r(C2-H)	1.0741(4)	1.07532(13)	1.0744(1)	1.0801
a(C3-H)	1.0751(5)	1.07527(16)	1.0745(1)	1.0810
a(H-N-C2)	125.21(3)	125.096(8)	125.09(1)	125.08
a(C5-N-C2)	109.58(4)	109.809(16)	109.82(1)	109.85
a(N-C2-C3)	107.87(5)	107.762(15)	107.76(1)	107.66
a(C2-C3-C4)	107.34(8)	107.334(12)	107.33(1)	107.42
a(N-C2-H)	121.89(22)	120.99(7)	121.12(2)	121.16
a(C2-C3-H)	125.43(19)	125.94(6)	125.88(2)	125.70
Rms resid. [MHz]	0.0009	-	0.0001	-
Mean Δ_e [uÅ^2]	0.01545	-	0.00018	-
pyrazole^m				
r(N1-N2)	1.3530(40)	1.3431(6)†	1.3441(1)†	1.3486
r(N2=C3)	1.3302(46)	1.3286(7)	1.3289(1)	1.3329
r(C3-C4)	1.4190(571)	1.4093(6)	1.4090(11)	1.4144
r(C4=C5)	1.3794(49)	1.3771(8)	1.3765(1)	1.3817
r(C5-N1)	1.3611(33)	1.3523(6)	1.3519(1)	1.3587
r(N1-H)	0.9947(23)	1.0014(4)	1.0014(1)	1.0092
r(C3-H)	1.0768(23)	1.0755(4)	1.0757(1)	1.0817
r(C4-H)	1.0747(24)	1.0736(4)	1.0739(1)	1.0797
r(C5-H)	1.0754(35)	1.0740(5)	1.0745(1)	1.0805
a(N1-N2-C3)	103.92(18)	104.18(3)	104.11(1)	104.23
a(N2-C3-C4)	112.20(99)	111.90(5)	111.93(2)	111.88
a(C3-C4-C5)	104.38(78)	104.46(4)	104.46(2)	104.53
a(C4-C5-N1)	106.22(22)	106.23(4)	106.26(1)	106.19
a(C5-N1-N2)	113.27(1.60)	113.24(5)	113.24(3)	113.18
a(N2-N1-H)	121.17(1.09)	118.97(11)	118.95(2)	119.06
a(N2-C3-H)	123.12(1.41)	119.49(14)	119.51(3)	119.49
a(C3-C4-H)	126.94(1.45)	128.32(13)	128.18(3)	128.22
a(N1-C5-H)	121.03(1.09)	121.84(11)	121.75(2)	121.82
Rms resid. [MHz]	0.0042	-	0.0001	-
Mean Δ_e [uÅ^2]	0.03139	-	0.00120	-

imidazole[n]				
r(N1-C2)	1.3700(40)	1.3612(9)†	1.3616(7)†	1.3671
r(C2=N3)	1.3141(36)	1.3111(8)	1.3103(6)	1.3161
r(N3-C4)	1.3865(33)	1.3797(8)	1.3794(5)	1.3789
r(C4=C5)	1.3670(252)	1.3627(8)	1.3627(41)	1.3731
r(C5-N1)	1.3824(32)	1.3738(9)	1.3743(5)	1.3802
r(N1-H)	0.9942(22)	1.0008(5)	1.0011(3)	1.0096
r(C2-H)	1.0769(21)	1.0759(6)	1.0770(3)	1.0817
r(C4-H)	1.0768(25)	1.0747(6)	1.0752(4)	1.0809
r(C5-H)	1.0765(25)	1.0764(5)	1.0764(4)	1.0793
a(N1-C2-N3)	111.98(22)	111.91(6)	111.93(4)	111.56
a(C2-N3-C4)	105.07(26)	105.02(5)	105.03(4)	105.43
a(N3-C4-C5)	110.63(51)	110.60(6)	110.62(8)	110.61
a(C4-C5-N1)	105.50(1.27)	105.45(6)	105.43(21)	105.12
a(C5-N1-C2)	106.81(98)	107.02(5)	106.99(16)	107.29
a(C2-N1-H)	125.50(68)	126.23(16)	126.15(11)	126.41
a(N1-C2-H)	122.56(52)	122.53(12)	122.37(9)	122.43
a(N3-C4-H)	120.70(54)	121.51(11)	121.45(9)	121.40
a(N1-C5-H)	121.51(56)	121.92(12)	121.90(9)	122.22
Rms resid. [MHz]	0.0040	-	0.0006	-
Mean Δ_e [uÅ^2]	0.02737	-	0.00096	-
furan[o]				
r(C2-O)	1.3670(11)	1.3594(7)†	1.3598(4)†	1.3647
r(C2=C3)	1.3570(18)	1.3552(8)	1.3542(4)	1.3611
r(C3-C4)	1.4459(176)	1.432(2)	1.4344(19)	1.4357
r(C2-H)	1.0740(8)	1.0735(7)	1.0739(3)	1.0790
r(C3-H)	1.0729(7)	1.0753(6)	1.0743(3)	1.0806
a(C2-O-C5)	106.43(10)	106.63(6)	106.50(3)	106.89
a(O-C2-C3)	110.88(8)	110.66(9)	110.79(4)	110.42
a(C2-C3-C4)	105.91(59)	106.03(7)	105.96(6)	106.14
r(H-C2-O)	115.11(12)	115.88(6)	115.82(3)	115.85
a(H-C3-C4)	127.49(6)	127.66(5)	127.61(3)	127.46
Rms resid. [MHz]	0.0092	-	0.0010	-
Mean Δ_e [uÅ^2]	0.04728	-	0.00116	-
thiophene[p]				
r(S-C2)	1.7196(10)	1.704(2)	1.7127(5)†	1.7404
r(C2=C3)	1.3663(18)	1.372(3)	1.3625(9)	1.3678
r(C3-C4)	1.4307(44)	1.421(4)	1.4233(21)	1.4289
r(C2-H)	1.0763(10)	1.085(5)	1.0772(5)	1.0811
r(C3-H)	1.0783(6)	1.088	1.0792(3)	1.0844
a(C2-S-C5)	91.90(8)	92.4(2)	91.88(4)	91.32
a(S-C2-C3)	111.65(7)	111.6	111.66(3)	111.53
a(C2=C3-C4)	112.40(15)	112.2	112.40(7)	112.81
a(H-C2-S)	119.76(20)	119.9(3)	120.06(10)	119.84
a(H-C3-C4)	124.12(5)	124.4(4)	124.14(3)	123.92
Rms resid. [MHz]	0.0020	-	0.0010	-
Mean Δ_e [uÅ^2]	0.06604	-	0.00340	-

maleic anhydride[q]				
r(C1=C2)	1.3406(52)	1.3324(5)†	1.3320(10)†	1.3355
r(C2-C3)	1.4870(18)	1.4849(5)	1.4857(3)	1.4895
r(C3-O4)	1.3907(11)	1.3848(3)	1.3843(2)	1.3941
r(C3=O7)	1.1943(8)	1.1894(2)	1.1896(2)	1.1948
r(C1-H)	1.0747(7)	1.0765(2)	1.0761(1)	1.0822
a(C1=C2-C3)	107.88(17)	107.96(1)	107.94(3)	108.15
a(C2-C3=O7)	129.75(12)	129.67(3)	129.59(2)	129.85
a(C1=C2-H)	129.91(6)	129.90(1)	129.93(1)	129.84
a(C2-C3-O4)	108.01(8)	107.78(2)	107.79(2)	107.59
a(C3-O4-C5)	108.22(10)	108.52(3)	108.53(2)	108.53
a(O4-C5=O6)	122.24(15)	122.55(4)	122.61(3)	122.56
Rms resid. [MHz]	0.0023	-	0.0004	-
Mean Δ_e [uÅ^2]	-0.00800	-0.00057	-0.00779	-
– ethers, aldehydes, esters and carboxylic acids –				
dimethyl ether[r]				
r(C-O)	1.4160(1)	1.40660(2)†	1.4074(1)†	1.4139
r(C-H$_{plane}$)	1.0843(10)	1.0865(2)	1.0855(2)	1.0924
r(C-H$_{out}$)	1.0999(4)	1.09506(7)	1.0949(1)	1.1014
r(C-O-C)	111.71(2)	111.100(3)	111.06(1)	112.55
r(O-C-H$_{plane}$)	107.17(8)	107.515(14)	107.50(2)	107.34
r(O-C-H$_{out}$)	110.70(2)	111.191(3)	111.12(1)	111.40
d(C-O-C-H$_{out}$)	60.20(4)	60.542(6)	60.52(1)	60.67
Rms resid. [MHz]	0.0021	-	0.0005	-
***cis*-formic acid**[s]				
r(C-H)	1.1005(12)	1.0976(4)†	1.0985(3)†	1.1063
r(C=O)	1.1980(18)	1.1920(4)	1.1910(3)	1.1960
r(C-O)	1.3518(17)	1.3472(4)	1.3485(3)	1.3540
r(O-H)	0.9575(16)	0.9610(4)	0.9619(2)	0.9680
a(H-C=O)	121.51(88)	123.26(22)	124.21(18)	123.99
a(O-C=O)	122.23(3)	122.28(1)	122.30(1)	122.41
a(C-O-H)	109.29(15)	109.28(3)	109.00(2)	109.71
Rms resid. [MHz]	0.0023	-	0.0003	-
Mean Δ_e [uÅ^2]	0.06366	-	-0.00067	-
***trans*-formic acid**[s]				
r(C-H)	1.0944(7)	1.0920(1)†	1.0918(3)†	1.0000
r(C=O)	1.2048(15)	1.1980(1)	1.1973(5)	1.2026
r(C-O)	1.3442(16)	1.3406(1)	1.3417(5)	1.3469
r(O-H)	0.9712(12)	0.9662(1)	0.9656(4)	0.9729
a(H-C=O)	123.67(66)	125.04(1)	125.38(22)	125.15
a(O-C=O)	124.98(3)	124.83(1)	124.78(1)	125.05
a(C-O-H)	106.37(8)	106.97(1)	106.78(2)	107.37
Rms resid. [MHz]	0.0019	-	0.0005	-
Mean Δ_e [uÅ^2]	0.08253	-	0.00057	-

***cis*-methyl formate**[t]				
r(C$_m$-O)	1.4378(40)	1.4341(5)[†]	1.4358(16)[†]	1.4440
r(C-O)	1.3444(40)	1.3345(4)	1.3343(15)	1.3411
r(C$_m$-H$_{plane}$)	1.0526(63)	1.0793(10)	1.0845(30)	1.0893
r(C$_m$-H$_{out}$)	1.0935(20)	1.0871(3)	1.0875(8)	1.0924
r(C-H)	1.0952(46)	1.0930(5)	1.0925(18)	1.1006
r(C=O)	1.2051(46)	1.2005(5)	1.2001(17)	1.2049
a(C$_m$-O-C)	114.99(34)	114.32(4)	114.26(13)	115.77
a(O-C$_m$-H$_{plane}$)	108.01(98)	106.05(16)	105.35(42)	105.46
a(O-C$_m$-H$_{out}$)	109.76(14)	110.19(2)	110.07(5)	110.24
a(O-C-H)	109.74(44)	109.96(5)	109.54(17)	109.24
a(O-C=O)	125.37(43)	125.50(5)	125.50(16)	125.81
d(H$_{out}$-C-O-C)	58.98(21)	−60.28(3)	−60.36(8)	−60.37
Rms resid. [MHz]	0.0067	-	0.0025	-
glycolaldehyde[r]				
r(C1=O)	1.2138(23)	1.2086(4)[†]	1.2083(5)[†]	1.2115
r(C1-H)	1.0946(18)	1.1015(3)	1.1011(4)	1.1096
r(C1-C2)	1.5166(19)	1.5003(3)	1.5014(4)	1.5065
r(C2-H)	1.1029(11)	1.0969(2)	1.0964(2)	1.1033
r(C2-O)	1.3946(20)	1.3962(3)	1.3970(4)	1.4014
r(O-H)	1.0410(35)	0.9593(5)	0.9618(6)	0.9721
a(C2-C1=O)	121.22(19)	121.65[u]	121.68(4)	122.05
a(C2-C1-H)	116.80(24)	116.91[u]	116.85(5)	116.47
a(C1-C2-H)	107.68(13)	108.11(2)	107.80(3)	107.79
a(C1-C2-O)	112.47(17)	111.75(3)	111.72(3)	112.63
a(C2-O-H)	103.07(13)	106.28(2)	106.14(3)	106.53
d(H-C2-C1=O)	122.90(12)	122.35(2)	122.27(2)	122.82
Rms resid. [MHz]	0.0067	-	0.0006	-

propanal^r				
r(C1-C2)	1.5130(20)	1.5023(6)†	1.5037(4)†	1.5087
r(C2-C3)	1.5236(22)	1.5164(4)	1.5165(4)	1.5260
r(C3-H$_{plane}$)	1.0759(18)	1.0884(3)	1.0879(3)	1.0943
r(C3-H$_{out}$)	1.0944(10)	1.0883(2)	1.0888(2)	1.0938
r(C2-H)	1.0991(11)	1.0949(2)	1.0946(2)	1.1012
r(C1=O)	1.2093(22)	1.2074(4)	1.2075(4)	1.2099
r(C1-H)	1.1059(18)	1.1056(3)	1.1040(3)	1.1145
a(C2-C3-H$_{plane}$)	111.37(29)	110.66(4)	110.52(6)	110.62
a(C2-C3-H$_{out}$)	110.24(7)	110.72(2)	110.68(1)	111.02
a(C1-C2-C3)	113.89(17)	113.60(2)	113.65(3)	114.84
a(C1-C2-H)	107.11(14)	106.95(3)	106.75(3)	106.63
a(C2-C1-O)	124.24(19)	124.38(3)	124.34(4)	124.98
a(C2-C1-H)	115.16(23)	115.44(3)	115.34(4)	114.91
d(O-C1-C2-H)	123.91(12)	123.77(2)	123.78(2)	124.39
d(C1-C2-C3-H$_{out}$)	58.90(10)	59.46(2)	59.42(2)	59.64
Rms resid. [MHz]	0.0067	-	0.0004	-

a) All fits have been performed on moments of inertia. For all structures evaluated in this work, the uncertainties on the geometrical parameters are reported within parentheses, rounded to $1 \cdot 10^{-4}$ Å for lengths and $1 \cdot 10^{-2}$ degrees for angles if smaller than these values. $\Delta_e = I^C - I^B - I^A$ is the inertial defect. † denotes the inclusion of $\Delta B^{\beta}_{\mathrm{el}}$.
b) MP2/VTZ r_e^{SE} from ref. 101.
c) B3LYP/6-311+G(3df,2pd) r_e^{SE} from ref. 101.
d) MP2/V(T+d)Z r_e^{SE} from ref. 101.
e) literature r_e^{SE} obtained as average of different MP2 and B3LYP r_e^{SE}, with basis sets of at least triple-ζ quality, where the $\Delta B^{\beta}_{\mathrm{vib}}$ are derived coupling scaled quadratic force fields with unscaled cubic force fields, from ref. 103.
f) MP2/VQZ r_e^{SE} from ref. 90.
g) MP2/VTZ r_e^{SE} from ref. 105.
h) MP2/VTZ r_e^{SE} from ref. 91.
i) SDQ-MBPT(4)/VTZ r_e^{SE} from ref. 107.
j) MP2/VTZ r_e^{SE} from ref. 112.
k) SDQ-MBPT(4)/VTZ r_e^{SE} from ref. 108.
l) MP2(AE)/wCVTZ r_e^{SE} from ref. 112.
m) B3LYP/6-311+G(3df,2pd) r_e^{SE} from ref. 112.
n) B3LYP/6-311+G(3df,2pd) r_e^{SE} from ref. 112, where the experimental ground-state rotational constants were corrected within the predicates method.
o) MP2/VTZ r_e^{SE} from ref. 114.
p) SE structure ED + MW + vibSP(B3LYP/6-311+G* force field), see ref. 115.
q) MP2/VTZ r_e^{SE} from ref. 89.
r) MP2/VTZ r_e^{SE} from ref. 90.
s) MP2/VTZ r_e^{SE} from ref. 121.
t) MP2/VTZ r_e^{SE} from ref. 122.
u) values calculated as $121.65 = 180.00 - 58.35$ and $116.91 = 180.00 - 63.09$, where 58.35 and 63.09 are taken from Table 6 of ref. 90.

Table 5: r_0, r_e^{SE} and r_e geometries for glyoxylic acid and pyridine. Distances in Å, angles in degrees.

	$r_0{}^a$	$r_e^{\mathrm{SE}\,a}$	r_e	
		B3LYP/SNSD	CCSD(T)	B3LYP/SNSD
glyoxylic acid				
r(C1-C2)	1.5361(29)	1.5211(3)†	1.5256^b	1.5345
r(C1-H)	1.0959(28)	1.0964(3)	1.0963	1.1045
r(C1=O)	1.2063(30)	1.2067(3)	1.2087	1.2080
r(C2=O)	1.2039(30)	1.1994(3)	1.1977	1.2034
r(C2-O)	1.3310(33)	1.3325(3)	1.3317	1.3373
r(O-H)	0.9552(44)	0.9692(4)	0.9697	0.9764
r(C2-C1-H)	115.36(21)	115.59(2)	115.41	115.13
r(C2-C1=O)	120.54(27)	120.60(3)	120.66	121.03
r(C1-C2=O)	120.97(25)	121.95(3)	121.90	121.74
r(C1-C2-O)	113.95(29)	113.70(3)	113.35	113.78
r(C-O-H)	107.15(23)	106.84(2)	106.74	107.67
Rms resid. [MHz]	0.0030	0.0003	-	-
Mean Δ_e [uÅ^2]	0.06291	-0.01110	-	-
pyridine				
r(C2-C3)	1.3950(15)	1.3907(18)†	1.3898^c	1.3954
r(C3-C4)	1.3949(12)	1.3888(13)	1.3876	1.3930
r(N-C2)	1.3414(35)	1.3358(40)	1.3346	1.3391
r(C2-H)	1.0841(11)	1.0818(12)	1.0824	1.0887
r(C3-H)	1.0800(10)	1.0796(11)	1.0801	1.0858
r(C4-H)	1.0819(12)	1.0802(13)	1.0808	1.0865
a(C6-N-C2)	117.01(16)	116.93(18)	117.01	117.19
a(N-C2-C3)	123.75(18)	123.79(21)	123.73	123.64
a(C2-C3-C4)	118.50(9)	118.53(11)	118.54	118.50
a(C3-C4-C5)	118.49(9)	118.44(10)	118.44	118.54
a(N-C2-H)	115.49(24)	115.97(28)	115.97	116.00
a(C3-C2-H)	120.76(16)	120.25(18)	120.30	120.36
a(C2-C3-H)	120.73(16)	120.10(18)	120.14	120.22
a(C3-C4-H)	120.76(6)	120.78(7)	120.78	120.73
Rms resid. [MHz]	0.0028	0.0031	-	-
Mean Δ_e [uÅ^2]	0.03958	0.00392	-	-

a) The fits have been performed using SE I_e^B and I_e^C moments of inertia for glyoxylic acid and I_e^A and I_e^C ones for pyridine. The digits within parentheses are the uncertainties on the geometrical parameters. $\Delta_e = I^C - I^B - I^A$ is the inertial defect. † denotes the inclusion of $\Delta B_{\mathrm{el}}^{\beta}$.
b) r_e optimized at the CCSD(T)/VQZ level, from ref. 128.
c) r_e optimized at the CCSD(T)/CBS+CV level, from this work.

Figure 4: Sketch of the 4 molecules determined within the template approach.

(a) *trans*-CHFCHCl

(b) (X,Y)
(CH, N) pyrimidine
(CF,CH) 2-fluoropyridine
(CH,CF) 3-fluoropyridine

Toward larger systems: the template approach

In all cases presented above, a large number of experimental data, coupled with a limited number of molecular parameters, permitted the complete determination of the molecular structure. This is often not possible when the molecular size and topological complexity increase because of the large number of isotopologues then required. In these cases, the strategy widely adopted in the literature consists of fixing in the fitting procedure some parameters to the corresponding computed values,[50,52–54] or allowing some internal coordinates (called predicates) to vary from reference values within given uncertainties.[129]

Although, as shown above, $\Delta B^{\beta}_{\mathrm{vib}}$ and ΔB_{el} contributions calculated at the B3LYP level leads to very good r_e^{SE} results, the fixed parameters (or reference values within predicated methodology) need to be estimated at much higher levels of theory to achieve good accuracy. An example is provided by the case of glycine Ip, for which the differences between the CCSD(T)/VTZ and CCSD(T)/CBS+CV equilibrium structures are significant.[130] Since CCSD(T)/CBS+CV calculations for large systems are computationally very expensive, they are not always feasible.

In the following, we present a new approach to deal with medium-large systems and a series of test cases that allow us to point out its reliability. When one is interested in the determination of the r_e^{SE} structure of a molecule for which high-level computations are too expensive, it is possible to use a similar molecule (e.g., an isomer or substituted system), for which an accurate r_e^{SE} structure is available, as a template for deriving the parameters to be fixed in the fitting procedure. These parameters can be obtained as

$$r_e(\text{fixed}) = r_e + \Delta\mathrm{TM} \tag{3}$$

where ΔTM is defined as

$$\Delta\mathrm{TM} = r_e^{\mathrm{SE}}(\mathrm{TM}) - r_e(\mathrm{TM}) \tag{4}$$

r_e is the geometrical parameter of interest calculated at the same level for both the molecule under consideration and that chosen as reference, denoted as template-molecule (TM). In the following, some examples of application of this new approach are given and the reachable accuracy is addressed.

***trans*-1-chloro-2-fluoroethylene**

The first system analyzed is *trans*-1-chloro-2-fluoroethylene, for which the lack of experimental rotational constants for the ^{13}C-containing isotopologues does not allow the determination of the C=C bond length. As discussed above, a possible solution is to fix this parameter at a value obtained at a very high level of theory, as done in ref. 98. CCSD(T)/CBS+CV equilibrium geometries are available for both *cis* and *trans* isomers (see ref. 98 and Table 6) together with a complete r_e^{SE} equilibrium structure for the *cis* species (see Table 1). As a consequence, the *cis*-1-chloro-2-fluoroethylene can be used as a template for the estimation of the ΔTM correction for r(CC) of the *trans* isomer. According to eq. 4, the difference between the SE and theoretical values of r(CC) in *cis*-chlorofluoroethylene has been employed to correct the theoretical value of the C=C bond length of *trans*-chlorofluoroethylene, obtained at the same level of theory. Subsequently, the corrected r(CC) has been kept fixed in the fits performed, the corresponding results being reported in Table 7. In all cases, the $\Delta B_{\mathrm{vib}}^{\beta}$ terms have been calculated at the B3LYP/SNSD level, and the B3LYP/AVTZ $\Delta B_{\mathrm{el}}^{\beta}$ contributions have also been included. The fits have been performed on the SE I_e^A

and I_e^B moments of inertia, with 20 and 1 as relative weights, since I_e^B is about 20 times larger than I_e^A. The four fits differ for the level of theory used in the evaluation of r_e^{SE} of the template molecule and of r_e: r_e^{SE}(C=C) has been taken from SE equilibrium structure of *cis*-chlorofluoroethylene calculated with either CCSD(T)/VTZ $\Delta B_{\mathrm{vib}}^{\beta}$ contributions, fits 1 and 2, or B3LYP/SNSD $\Delta B_{\mathrm{vib}}^{\beta}$ contributions, fits 3 and 4, while the theoretical r_e(C=C) value at the CCSD(T)/CBS+CV level has been used for fits 1 and 3, and that at the B3LYP/SNSD level for fits 2 and 4. Table 7 shows that the results of fits 1 and 2 are similar to one another, and this is also the case for fits 3 and 4.

This suggests that the accuracy of the template approach is rather independent from the chosen theoretical method used in r_e estimation, and only limited by the accuracy of the template-molecule SE equilibrium structure considered. It is noteworthy that fit 4 allowed us to obtain a SE equilibrium structure completely independent from high-level (extremely expensive) computational results. This is particularly appealing in the treatment of medium-large systems, for which structural determinations at highly-correlated levels become computationally too expensive.

Table 6: r_e equilibrium geometries of cis- and *trans*-1-chloro-2-fluoroethylene. Distances in Å, angles in degrees.

	cis-CHFCHCl		*trans*-CHFCHCl	
	CBS+CV^a	B3LYP/SNSD	CBS+CV^a	B3LYP/SNSD
r(C1-Cl)	1.7107	1.7404	1.7163	1.7495
r(C1-H)	1.0764	1.0818	1.0775	1.0825
r(C1=C2)	1.3249	1.3278	1.3240	1.3266
r(C2-F)	1.3310	1.3416	1.3376	1.3499
r(C2-H)	1.0787	1.0849	1.0785	1.0840
a(Cl-C1=C2)	123.10	123.74	120.63	121.11
a(H-C1-C2)	120.43	120.91	122.95	123.62
a(F-C2=C1)	122.53	123.10	120.14	120.10
a(H-C2=C1)	123.43	123.45	125.82	126.50

a) CCSD(T)/CBS+CV r_e equilibrium geometry from ref. 98.

Table 7: r_e^{SE} equilibrium geometries of *trans*-1-chloro-2-fluoroethylene. Distances in Å, angles in degrees.

***trans*-chloro fluoroethylene**	$r_e^{\mathrm{SE}a}$			
	Fit 1^b	Fit 2^c	Fit 3^d	Fit 4^e
r(C1-Cl)	1.7188(9)	1.7190(9)	1.7178(9)	1.7179(9)
r(C1-H)	1.0775(21)	1.0774(21)	1.0780(21)	1.0779(21)
r(C1=C2)	1.3236$\surd$	1.3233$\surd$	1.3257$\surd$	1.3254$\surd$
r(C2-F)	1.3395(20)	1.3398(20)	1.3376(20)	1.3379(20)
r(C2-H)	1.0792(27)	1.0792(27)	1.0790(27)	1.0790(27)
a(Cl-C1=C2)	120.46(6)	120.45(6)	120.54(6)	120.53(6)
a(H-C1-C2)	122.60(15)	122.62(15)	122.47(15)	122.49(15)
a(F-C2=C1)	120.17(21)	120.16(21)	120.18(21)	120.18(21)
a(H-C2=C1)	125.86(5)	125.89(5)	125.69(5)	125.71(5)
Rms resid. [MHz]	0.0011	0.0011	0.0011	0.0011
Mean Δ_e [uÅ^2]	0.00308	0.00308	0.00308	0.00308

a) All fits have been performed on SE I_e^A and I_e^B moments of inertia, with 20 and 1 as weights, respectively, derived from $(B_0^\beta)^{\mathrm{exp}}$ constants corrected by B3LYP/SNSD $\Delta B_{\mathrm{vib}}^\beta$ and B3LYP/AVTZ $\Delta B_{\mathrm{el}}^\beta$ contributions. The digits within parentheses are the uncertainties on the geometrical parameters, while $\surd$ denotes the parameter kept fixed, obtained using *cis*-chlorofluoroethylene as TM (see text). $\Delta_e = I^C - I^B - I^A$ is the inertial defect.
b) r_e(fixed) = r_e(CBS+CV) + ΔTM; ΔTM = r_e^{SE}(CCSD(T)/VTZ) − r_e(CBS+CV).
c) r_e(fixed) = r_e(B3LYP/SNSD) + ΔTM; ΔTM = r_e^{SE}(CCSD(T)/VTZ) − r_e(B3LYP/SNSD).
d) r_e(fixed) = r_e(CBS+CV) + ΔTM; ΔTM = r_e^{SE}(B3LYP/SNSD) − r_e(CBS+CV).
e) r_e(fixed) = r_e(B3LYP/SNSD) + ΔTM; ΔTM = r_e^{SE}(B3LYP/SNSD) − r_e(B3LYP/SNSD).

Pyrimidine

Analogously to the case of *trans*-chlorofluoroethylene, the limited number of isotopologues experimentally investigated (7) makes the derivation of all geometrical parameters of pyrimidine not possible. In particular, no deutereted species have been studied experimentally, thus preventing the derivation of the C-H bond lengths and of the corresponding angles. In Table 8, three different fits for the SE equilibrium structure of pyrimidine are reported, all obtained by correcting the experimental rotational constants with $\Delta B_{\mathrm{vib}}^\beta$ contributions at the B3LYP/SNSD level and fitting on the I_e^A and I_e^C moments of inertia, with the inclusion of $\Delta B_{\mathrm{el}}^\beta$ (see Table 8). To evaluate the non-determinable parameters, pyridine has been used as TM (fit 1 in Table 8). The ΔTM corrections have been derived from the B3LYP/SNSD results of pyridine (Table 5). In particular, the N3-C4-H angle of pyrimidine has been estimated by using the ΔTM correction evaluated for the N-C2-H angle of pyridine; the ΔTM corrections for r(C2-H) and r(C4-H) have been based on the values for the C2-H distance of pyridine (these three parameters have in common the N-C-H pattern), while for the C5-H distance in pyrimidine the ΔTM correction has been calculated from r(C3-H) of pyridine (these two parameters shares a N-CH-C-H pattern). For the SE equilibrium structure of pyrimidine obtained following this procedure (fit 1 in Table 8), we expect an accuracy similar to that of a full SE equilibrium structure obtained with B3LYP/SNSD $\Delta B^{\beta}_{\mathrm{vib}}$ contributions.

Fit 2 and 3 in Table 8 show that it is possible to obtain very similar results using pyridazine (see Table 1) as TM instead of pyridine. This finding points out another interesting feature of the template approach, that is, the choice of TM is rather flexible: it is sufficient to find a molecule in which the parameter of interest, for example the C4-H bond length in pyrimidine, is present and involved in a similar bond pattern, a N-C-H bond chain for the case under consideration. The comparison of the results for fits 2 and 3 demonstrates that the SE equilibrium structure obtained with the TM approach does not change significantly if the r_e^{SE} (TM) parameter is taken from the best SE equilibrium structure available (CCSD(T)/ANO0 vibrational contributions in this case) or from the B3LYP/SNSD r_e^{SE}. A r_e^{SE} equilibrium structure of pyrimidine has been recently determined using a B3LYP/6-311+G(3df,2pd) cubic force field and the so-called predicate approach based on a CCSD(T) equilibrium geometry.[112] The remarkable agreement between the "template" and the "predicate" r_e^{SE} equilibrium geometries (see Table 8) gives further support to the template strategy, which has the advantage of avoiding any expensive CCSD(T) computation.

Table 8: r_e and r_e^{SE} equilibrium geometries for pyrimidine. Distances in Å, angles in degrees.

pyrimidine	$r_e^{\mathrm{SE}a}$				r_e	
	Fit 1[b]	Fit 2[c]	Fit 3[d]	Literature[e]	Literature[f]	B3LYP/SNSD
r(C2-N3)	1.3334(1)	1.3334(1)	1.3334(1)	1.3331(3)	1.3339	1.3374
r(N3-C4)	1.3355(1)	1.3355(1)	1.3358(1)	1.3349(6)	1.3349	1.3385
r(C4-C5)	1.3868(4)	1.3867(3)	1.3866(4)	1.3874(4)	1.3874	1.3921
r(C2-H)	1.0814✓	1.0822✓	1.0816✓	1.0820(23)	1.0819	1.0883
r(C4-H)	1.0819✓	1.0826✓	1.0820✓	1.0843(17)	1.0824	1.0887
r(C5-H)	1.0789✓	1.0795✓	1.0784✓	1.0799(23)	1.0793	1.0851
a(C2-N3-C4)	115.69(1)	115.69(1)	115.68(1)	115.712(21)	115.656	115.91
a(N3-C4-C5)	122.27(2)	122.27(2)	122.27(2)	122.276(19)	122.332	122.25
a(C4-C5-C6)	116.72(1)	116.72(1)	116.72(2)	116.661(36)	116.632	116.61
a(N1-C2-H)	116.31(1)	116.31(1)	116.31(1)	116.318(19)	116.303	116.46
a(N3-C4-H)	116.45✓	116.37✓	116.32✓	116.48(20)	116.461	116.48
a(C4-C5-H)	121.64(1)	121.64(1)	121.64(1)	121.670(18)	121.684	121.70
Rms resid. [MHz]	0.0004	0.0003	0.0004	-	-	-
Mean Δ_e [uÅ^2]	-0.00133	-0.00133	-0.00133	-	-	-

a) All fits have been performed on SE I_e^A and I_e^C moments of inertia, derived from the $(B_0^\beta)^{\mathrm{EXP}}$ constants corrected by B3LYP/SNSD $\Delta B_{\mathrm{vib}}^\beta$ and B3LYP/AVTZ $\Delta B_{\mathrm{el}}^\beta$ contributions. The digits within parentheses are the uncertainties on the geometrical parameters, while ✓ denotes the parameters kept fixed. $\Delta_e = I^C - I^B - I^A$ is the inertial defect.
b) r_e(fixed) $= r_e$(B3LYP/SNSD) $+ \Delta$TM; ΔTM $= r_e^{\mathrm{SE}}$(B3LYP/SNSD) $- r_e$(B3LYP/SNSD), with pyridine as TM: r(C2-H) and r(C4-H) from r(C2-H), r(C5-H) from r(C3-H), a(N3-C4-H) from a(N-C2-H), see text.
c) r_e(fixed) $= r_e$(B3LYP/SNSD) $+ \Delta$TM; ΔTM $= r_e^{\mathrm{SE}}$(CCSD(T)/ANO0) $- r_e$(B3LYP/SNSD), with pyridazine as TM: r(C2-H) and r(C4-H) from r(C3-H), r(C5-H) from r(C4-H), a(N3-C4-H) from a(N2-C3-H), see text.
d) r_e(fixed) $= r_e$(B3LYP/SNSD) $+ \Delta$TM; ΔTM $= r_e^{\mathrm{SE}}$(B3LYP/SNSD) $- r_e$(B3LYP/SNSD), with pyridazine as TM: r(C2-H) and r(C4-H) from r(C3-H), r(C5-H) from r(C4-H), a(N3-C4-H) from a(N2-C3-H), see text.
e) r_e^{SE} structure determined using B3LYP/6-311+G(3df,2pd) $\Delta B_{\mathrm{vib}}^\beta$ and the predicate approach, from ref. 112.
f) r_e^{BO}(II) in Table 9 of ref. 112.

Fluoropyridines

The first determinations of the SE equilibrium structure of 2- and 3-fluoropyridine are reported in Tables 9 and 10, respectively. Fluorine substitution reduces the molecular symmetry from C_{2v} to C_s, with the consequent increase of the number of unique internal parameters from 10 to 18. Because of the limited number of available experimental data, for these molecules it is not possible to evaluate all structural parameters. Therefore, some parameters have been fixed using the template approach.

Because of the lack of rotational data for deuterated species, only the parameters defining the C-C ring and the C-F bond length have been considered as free parameters for 2-fluoropyridine. On the other hand, for 3-fluoropyridine also the C-F bond length has been kept fixed in order to converge the fitting procedure.

Starting from the assumption that the substitution of a hydrogen atom with fluorine does not affect significantly the structure of the ring, we used pyridine as TM for 2-fluoropyridine. The best results have been obtained by fitting the SE I_e^A and I_e^C moments of inertia. In fit 1, the ΔTM corrections have been estimated by considering the CCSD(T)/CBS+CV level for r_e, while in fit 2, r_e has been calculated at the B3LYP/SNSD level. Even in this case, the values in Table 5 confirm that the resulting SE structures are negligibly affected by the level of theory chosen for r_e.

For 3-fluoropyridine, we proceeded analogously to 2-fluoropyridine for what concerns the C-H bonds, while for the C-F distance and the corresponding C3-C2-F angle, 2-fluoropyridine has been employed as TM. This is consistent with what discussed above for pyrimidine, namely that the choice of the TM molecule is quite flexible, thus allowing the simultaneous use of more than one TM in the determination of the parameters to be fixed. As in 2-fluoropyridine, we have used the B3LYP/SNSD r_e^{SE} of pyridine combined with both CCSD(T)/CBS+CV (fit 1) and B3LYP/SNSD (fit 2) structures as r_e in the calculation of ΔTM.

Table 9: r_e and r_e^{SE} equilibrium geometries for 2-fluoropyridine. Distances in Å, angles in degrees.

2-fluoropyridine	$r_e^{\mathrm{SE}a}$		r_e	
	Fit 1[b]	Fit 2[c]	CBS+CV	B3LYP/SNSD
r(N-C2)	1.3138(10)	1.3135(10)	1.3063	1.3120
r(N-C6)	1.3402(5)	1.3402(5)	1.3410	1.3438
r(C2-C3)	1.3838(14)	1.3840(14)	1.3898	1.3935
r(C3-C4)	1.3837(3)	1.3837(3)	1.3836	1.3898
r(C4-C5)	1.3949(4)	1.3948(4)	1.3933	1.3972
r(C5-C6)	1.3836(4)	1.3836(4)	1.3844	1.3909
r(C2-F)	1.3357(2)	1.3358(2)	1.3344	1.3483
r(C3-H)	1.0781√	1.0775√	1.0787	1.0839
r(C4-H)	1.0801√	1.0796√	1.0807	1.0860
r(C5-H)	1.0788√	1.0783√	1.0794	1.0848
r(C6-H)	1.0809√	1.0811√	1.0815	1.0876
a(C6-N-C2)	116.24(6)	116.24(6)	116.42	116.61
a(N-C2-C3)	126.17(7)	126.17(7)	126.17	126.08
a(C2-C3-C4)	116.73(3)	116.72(3)	116.59	116.53
a(C3-C4-C5)	119.06(1)	119.06(1)	119.09	119.21
a(C4-C5-C6)	118.37(0)	118.37(0)	118.39	118.33
a(C5-C6-N)	123.44(2)	123.44(2)	123.35	123.25
a(C3-C2-F)	118.43(12)	118.41(12)	117.87	117.98
a(C2-C3-H)	120.38√	120.48√	120.42	120.60
a(C3-C4-H)	120.33(1)	120.33(1)	120.34	120.10
a(C6-C5-H)	120.26√	120.25√	120.30	120.37
a(C5-C6-H)	120.86√	120.86√	120.91	120.97
Rms resid. [MHz]	0.0001	0.0001	-	-
Mean Δ_e [uÅ^2]	-0.00180	-0.00180	-	-

a) The fits have been performed on the SE I_e^A and I_e^C moments of inertia. The $(B_0^\beta)^{\mathrm{EXP}}$ constants have been corrected by B3LYP/SNSD $\Delta B_{\mathrm{vib}}^\beta$ and B3LYP/AVTZ $\Delta B_{\mathrm{el}}^\beta$ contributions. The digits within parentheses are the uncertainties on the geometrical parameters, while √ denotes the parameters kept fixed obtained using pyridine as TM, see text. $\Delta_e = I^C - I^B - I^A$ is the inertial defect.

b) r_e(fixed) = r_e(CBS+CV) + ΔTM; ΔTM = r_e^{SE}(B3LYP/SNSD) − r_e(CBS+CV).

c) r_e(fixed) = r_e(B3LYP/SNSD) + ΔTM; ΔTM = r_e^{SE}(B3LYP/SNSD) − r_e(B3LYP/SNSD).

Table 10: r_e and r_e^{SE} equilibrium geometries for 3-fluoropyridine. Distances in Å, angles in degrees.

3-fluoropyridine	$r_e^{\mathrm{SE}a}$		r_e	
	Fit 1^b	Fit 2^c	CBS+CV	B3LYP/SNSD
r(N-C2)	1.3346(9)	1.3352(9)	1.3323	1.3368
r(C2-C3)	1.3931(5)	1.3888(6)	1.3859	1.3904
r(C3-C4)	1.3714(6)	1.3757(6)	1.3786	1.3850
r(C4-C5)	1.3895(8)	1.3901(9)	1.3894	1.3933
r(C5-C6)	1.3940(5)	1.3926(5)	1.3888	1.3955
r(C6-N)	1.3314(9)	1.3326(10)	1.3337	1.3389
r(C2-H)	1.0811$\surd$	1.0813$\surd$	1.0817	1.0877
r(C3-F)	1.3406$\surd$	1.3398$\surd$	1.3393	1.3523
r(C4-H)	1.0791$\surd$	1.0787$\surd$	1.0797	1.0852
r(C5-H)	1.0794$\surd$	1.0791$\surd$	1.0800	1.0856
r(C6-H)	1.0808$\surd$	1.0812$\surd$	1.0814	1.0877
a(C6-N-C2)	117.68(6)	117.65(6)	117.73	117.88
a(N-C2-C3)	121.75(6)	121.85(6)	122.09	121.79
a(C2-C3-C4)	121.07(3)	121.07(3)	120.83	121.04
a(C3-C4-C5)	117.04(2)	116.94(2)	116.84	116.91
a(C4-C5-C6)	118.88(2)	118.91(2)	119.18	118.94
a(C5-C6-N)	123.58(5)	123.58(5)	123.33	123.43
a(C3-C2-H)	119.90$\surd$	120.12$\surd$	119.95	120.23
a(C4-C3-F)	120.59$\surd$	120.19$\surd$	119.14	119.76
a(C3-C4-H)	121.88(1)	121.97(1)	121.96	120.52
a(C6-C5-H)	120.24$\surd$	120.25$\surd$	120.28	120.37
a(C5-C6-H)	120.48$\surd$	120.36$\surd$	120.53	120.47
Rms resid. [MHz]	0.0004	0.0004	-	-
Mean Δ_e [uÅ^2]	-0.00180	-0.00180	-	-

a) The fits have been performed on the SE I_e^A and I_e^B moments of inertia. The $(B_0^\beta)^{\mathrm{EXP}}$ constants have been corrected by B3LYP/SNSD $\Delta B_{\mathrm{vib}}^\beta$ and B3LYP/AVTZ $\Delta B_{\mathrm{el}}^\beta$ contributions. The digits within parentheses are the uncertainties on the geometrical parameters, while $\surd$ denotes the parameters kept fixed obtained using pyridine and 2-fluoropyridine as TM, see text. $\Delta_e = I^C - I^B - I^A$ is the inertial defect.
b) $r_e(\text{fixed}) = r_e(\text{CBS+CV}) + \Delta\text{TM}$; $\Delta\text{TM} = r_e^{\mathrm{SE}}(\text{B3LYP/SNSD}) - r_e(\text{CBS+CV})$.
c) $r_e(\text{fixed}) = r_e(\text{B3LYP/SNSD}) + \Delta\text{TM}$; $\Delta\text{TM} = r_e^{\mathrm{SE}}(\text{B3LYP/SNSD}) - r_e(\text{B3LYP/SNSD})$.

CONCLUSIONS

The present paper is devoted to a thorough investigation on the determination of accurate equilibrium structures by means of a semi-experimental approach, avoiding as much as possible the use of expensive CC calculations. In the first part, 21 small molecules for which accurate SE structures determined using CCSD(T) vibrational contributions are available have been selected (CCse set) and used to demonstrate that the $\Delta B^{\beta}_{\mathrm{vib}}$ contributions derived from cubic force fields at the DFT level lead to results with an accuracy comparable to that obtainable at higher levels of theory (MP2 and, especially, CCSD(T)). In the second part, it has been shown that the B3LYP/SNSD model represents a very good compromise between accuracy and computational cost in the calculation of $\Delta B^{\beta}_{\mathrm{vib}}$ contributions, thus making the accurate determination of molecular structures for medium and large systems feasible. Within this context, new SE equilibrium structures have been determined for a set of 26 molecules, mostly including building blocks of biomolecules. These structures together with the SE equilibrium structures of the previous 21 molecules determined using B3LYP/SNSD vibrational contributions provide a set of 47 accurate equilibrium structures (referred to as the B3se set) which can be recommended as reference data for the investigation of molecular properties, as well as for parameterisations and validation of QM models. Finally, a new method, denoted template approach, has been proposed to deal with molecules for which there is a lack of experimental data and it is thus necessary to fix some geometrical parameters in the fitting procedure. This approach further extends the size of molecular systems amenable to highly accurate molecular structure determinations. The whole B3se set is available (in graphical interactive form) on our website dreams.sns.it.

Acknowledgement

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement N^{o} ERC-2012-AdG-320951-DREAMS. This work was also supported by Italian MIUR (PRIN 2012 "STAR: Spectroscopic and computational Techniques for Astrophysical and atmospheric Research" and PON01-01078/8) and by the University of Bologna (RFO funds). The high performance computer facilities of the DREAMS center (http://dreamshpc.sns.it) are acknowledged for providing computer resources. The support of COST CMTS-Action is also acknowledged.

Figure 5: TOC graphics

Supporting Information Available

Supporting Information Available: expressions of the vibration-rotation interaction constants; $(B_0^\beta)^{\mathrm{EXP}}$, $\Delta B_{\mathrm{vib}}^\beta$ and $\Delta B_{\mathrm{el}}^\beta$ results for all the molecules studied in the paper (Tables 1-5); the comparison between the r_0 and r_e^{SE} geometries estimated using $\Delta B_{\mathrm{vib}}^\beta$ from CCSD(T), MP2, B3LYP/SNSD and B3LYP/AVTZ cubic force fields (Table 6); the statistical distributions of the deviations from CCSD(T) SE equilibrium parameters for CH, CC and CO bonds of the molecules belonging to the CCse set (Figure 1); the plots of the CCSD(T) r_e^{SE} versus the MP2 and B3LYP ones for the molecules belonging to the CCse set (Figures 2).

This material is available free of charge via the Internet at `http://pubs.acs.org/`.

References

(1) Domenicano, A., Hargittai, I., Eds. *Accurate Molecular Structures. Their Determination and Importance*; Oxford University Press, 1992.

(2) Demaison, J., Boggs, J., Császár, A., Eds. *Equilibrium Molecular Structures: From Spectroscopy to Quantum Chemistry*; CRC Press, 2011.

(3) Bak, K. L.; Gauss, J.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Stanton, J. F. The accurate determination of molecular equilibrium structures. *J. Chem. Phys.* **2001**, *114*, 6548.

(4) Demaison, J. Experimental, semi-experimental and *ab initio* equilibrium structures. *Mol. Phys.* **2007**, *105*, 3109.

(5) Puzzarini, C.; Stanton, J. F.; Gauss, J. Quantum-chemical calculation of spectroscopic parameters for rotational spectroscopy. *Int. Rev. Phys. Chem.* **2010**, *29*, 273.

(6) Puzzarini, C.; Biczysko, M. In *Structure Elucidation in Organic Chemistry*; Cid, M., Ed.; Wiley-VCH Verlag GmbH & Co. KGaA, 2015; pp 27–64.

(7) Pérez, C.; Muckle, M. T.; Zaleski, D. P.; Seifert, N. A.; Temelso, B.; Shields, G. C.; Kisiel, Z.; Pate, B. H. Structures of Cage, Prism, and Book Isomers of Water Hexamer from Broadband Rotational Spectroscopy. *Science* **2012**, *336*, 897–901.

(8) Melandri, S.; Sanz, M. E.; Caminati, W.; Favero, P. G.; Kisiel, Z. The Hydrogen Bond between Water and Aromatic Bases of Biological Interest: An Experimental and Theoretical Study of the 1:1 Complex of Pyrimidine with Water. *J. Am. Chem. Soc.* **1998**, *120*, 11504–11509.

(9) Caminati, W. Nucleic Acid Bases in the Gas Phase. *Angew. Chem. Int. Edit.* **2009**, *48*, 9030–9033.

(10) Lovas, F. J.; McMahon, R. J.; Grabow, J.-U.; Schnell, M.; Mack, J.; Scott, L. T.; Kuczkowski, R. L. Interstellar Chemistry: A Strategy for Detecting Polycyclic Aromatic Hydrocarbons in Space. *J. Am. Chem. Soc.* **2005**, *127*, 4345–4349.

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

(12) Blanco, S.; Lesarri, A.; López, J. C.; Alonso, J. L. The Gas-Phase Structure of Alanine. *J. Am. Chem. Soc.* **2004**, *126*, 11675–11683.

(13) Peña, I.; Daly, A. M.; Cabezas, C.; Mata, S.; Bermúdez, C.; Niño, A.; López, J. C.; Grabow, J.-U.; Alonso, J. L. Disentangling the Puzzle of Hydrogen Bonding in Vitamin C. *J. Phys. Chem. Lett.* **2013**, *4*, 65–69.

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

(15) Barone, V., Ed. *Computational Strategies for Spectroscopy: from Small Molecules to Nano Systems*; Wiley, 2011.

(16) Grunenberg, J., Ed. *Computational Spectroscopy: Methods, Experiments and Applications*; Wiley, 2010.

(17) Quack, M., Merkt, F., Eds. *Handbook of High-resolution Spectroscopy*; John Wiley & Sons, Inc., 2011; p 2182.

(18) Grimme, S.; Steinmetz, M. Effects of London Dispersion Correction in Density Functional Theory on the Structures of Organic Molecules in the Gas Phase. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16031–16042.

(19) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. Benchmark database of accurate (MP2 and CCSD(T) complete basis set limit) interaction energies of small model complexes, DNA base pairs, and amino acid pairs. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

(20) Zhao, Y.; Truhlar, D. G. Density Functionals for Noncovalent Interaction Energies of Biological Importance. *J. Chem. Theory Comput.* **2007**, *3*, 289–300.

(21) Barone, V.; Biczysko, M.; Pavone, M. The role of dispersion correction to DFT for modelling weakly bound molecular complexes in the ground and excited electronic states. *Chem. Phys.* **2008**, *346*, 247–256.

(22) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. Accurate molecular structures and infrared spectra of trans-2,3-dideuterooxirane, methyloxirane, and trans-2,3-dimethyloxirane. *J. Chem. Phys.* **2014**, *141*, 034107.

(23) Senn, H. M.; Thiel, W. QM/MM Methods for Biomolecular Systems. *Angew. Chem. Int. Edit.* **2009**, *48*, 1198–1229.

(24) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.;

Won, Y.; Archontis, G.; Bartels, C.; Boresch, S. et al. CHARMM: The biomolecular simulation program. *J. Comp. Chem.* **2009**, *30*, 1545–1614.

(25) Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D. et al. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013**, *29*, 845–854.

(26) Grubisic, S.; Brancato, G.; Pedone, A.; Barone, V. Extension of the AMBER force field to cyclic α,α dialkylated peptides. *Phys. Chem. Chem. Phys.* **2012**, *14*, 15308–15320.

(27) Maple, J. R.; Dinur, U.; Hagler, A. T. Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 5350–5354.

(28) Dasgupta, S.; Yamasaki, T.; Goddard, W. A. The Hessian biased singular value decomposition method for optimization and analysis of force fields. *J. Chem. Phys.* **1996**, *104*, 2898–2920.

(29) Biczysko, M.; Bloino, J.; Brancato, G.; Cacelli, I.; Cappelli, C.; Ferretti, A.; Lami, A.; Monti, S.; Pedone, A.; Prampolini, G. et al. Integrated computational approaches for spectroscopic studies of molecular systems in the gas phase and in solution: pyrimidine as a test case. *Theo. Chem. Acc.* **2012**, *131*, 1201.

(30) Barone, V.; Cacelli, I.; De Mitri, N.; Licari, D.; Monti, S.; Prampolini, G. Joyce and Ulysses: integrated and user-friendly tools for the parameterization of intramolecular force fields from quantum mechanical data. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3736–3751.

(31) Risthaus, T.; Steinmetz, M.; Grimme, S. Implementation of nuclear gradients of range-separated hybrid density functionals and benchmarking on rotational constants for organic molecules. *J. Comp. Chem.* **2014**, *35*, 1509–1516.

(32) Cramer, C. *Essentials of Computational Chemistry: Theories and Models*; John Wiley & Sons Ltd., Chichester, UK, 2005.

(33) Császár, A. In *Equilibrium Molecular Structures: From Spectroscopy to Quantum Chemistry*; Demaison, J., Boggs, J., Császár, A., Eds.; CRC Press, 2011; pp 233–261.

(34) Kuchitsu, K. In *Accurate Molecular Structures. Their Determination and Importance*; Domenicano, A., Hargittai, I., Eds.; Oxford University Press, 1992; p 14.

(35) Born, M.; Oppenheimer, R. Zur Quantentheorie der Molekeln. *Ann. Phys.-Berlin* **1927**, *389*, 457–484.

(36) Jensen, P.; Bunker, P. In *Computational molecular spectroscopy*; Jensen, P., Bunker, P., Eds.; Wiley, 2000; pp 3–12.

(37) 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.

(38) Heckert, M.; Kállay, M.; Gauss, J. Molecular equilibrium geometries based on coupled-cluster calculations including quadruple excitations. *Mol. Phys.* **2005**, *103*, 2109.

(39) Heckert, M.; Kállay, M.; Tew, D. P.; Klopper, W.; Gauss, J. Basis-set extrapolation techniques for the accurate calculation of molecular equilibrium geometries using coupled-cluster theory. *J. Chem. Phys.* **2006**, *125*, 044108.

(40) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. The performance of composite schemes and hybrid CC/DFT model in predicting structure, thermodynamic and spectroscopic parameters: the challenge of the conformational equilibrium in glycine. *Phys. Chem. Chem. Phys.* **2013**, *15*, 10094–10111.

(41) 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.

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

(43) Senent, M. L.; Puzzarini, C.; Domínguez-Gómez, R.; Carvajal, M.; Hochlaf, M. Theoretical spectroscopic characterization at low temperatures of detectable sulfur-organic compounds: Ethyl mercaptan and dimethyl sulfide. *J. Chem. Phys.* **2014**, *140*, 124302.

(44) Ormond, T. K.; Scheer, A. M.; Nimlos, M. R.; Robichaud, D. J.; Daily, J. W.; Stanton, J. F.; Ellison, G. B. Polarized Matrix Infrared Spectra of Cyclopentadienone: Observations, Calculations, and Assignment for an Important Intermediate in Combustion and Biomass Pyrolysis. *J. Phys. Chem. A* **2014**, *118*, 708–718.

(45) Wang, X.; Huang, X.; Bowman, J. M.; Lee, T. J. Anharmonic rovibrational calculations of singlet cyclic C_4 using a new *ab initio* potential and a quartic force field. *J. Chem. Phys.* **2013**, *139*, 224302.

(46) Puzzarini, C.; Biczysko, M.; Barone, V. Accurate Harmonic/Anharmonic Vibrational Frequencies for Open-Shell Systems: Performances of the B3LYP/N07D Model for Semirigid Free Radicals Benchmarked by CCSD(T) Computations. *J. Chem. Theory Comput.* **2010**, *6*, 828.

(47) Bloino, J.; Biczysko, M.; Barone, V. General Perturbative Approach for Spectroscopy, Thermodynamics, and Kinetics: Methodological Background and Benchmark Studies. *J. Chem. Theory Comput.* **2012**, *8*, 1015.

(48) Carnimeo, I.; Puzzarini, C.; Tasinato, N.; Stoppa, P.; Pietropolli-Charmet, A.; Biczysko, M.; Cappelli, C.; Barone, V. Anharmonic theoretical simulations of infrared spectra of halogenated organic compounds. *J. Chem. Phys.* **2013**, *139*, 074310.

(49) Barone, V.; Biczysko, M.; Bloino, J. Fully anharmonic IR and Raman spectra of medium-size molecular systems: accuracy and interpretation. *Phys. Chem. Chem. Phys.* **2014**, *16*, 1759–1787.

(50) Demaison, J. F.; Craig, N. C. Semiexperimental Equilibrium Structure for cis,trans-1,4-Difluorobutadiene by the Mixed Estimation Method. *J. Phys. Chem. A* **2011**, *115*, 8049.

(51) 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.

(52) Demaison, J.; Craig, N. C.; Cocinero, E. J.; Grabow, J.-U.; Lesarri, A.; Rudolph, H. D. Semiexperimental Equilibrium Structures for the Equatorial Conformers of N-Methylpiperidone and Tropinone by the Mixed Estimation Method. *J. Phys. Chem. A* **2012**, *116*, 8684–8692.

(53) Demaison, J.; Craig, N. C.; Conrad, A. R.; Tubergen, M. J.; Rudolph, H. D. Semiexperimental Equilibrium Structure of the Lower Energy Conformer of Glycidol by the Mixed Estimation Method. *J. Phys. Chem. A* **2012**, *116*, 9116–9122.

(54) Demaison, J.; Jahn, M. K.; Cocinero, E. J.; Lesarri, A.; Grabow, J.-U.; Guillemin, J.-C.; Rudolph, H. D. Accurate Semiexperimental Structure of 1,3,4-Oxadiazole by the Mixed Estimation Method. *J. Phys. Chem. A* **2013**, *117*, 2278–2284.

(55) Barone, V.; Biczysko, M.; Bloino, J.; Egidi, F.; Puzzarini, C. Accurate structure, thermodynamics, and spectroscopy of medium-sized radicals by hybrid coupled clus-

ter/density functional theory approaches: The case of phenyl radical. *J. Chem. Phys.* **2013**, *138*, 234303.

(56) Papoušek, D.; Aliev, M. *Molecular vibrational-rotational spectra*; Elsevier, 1982.

(57) Gauss, J.; Ruud, K.; Helgaker, T. Perturbation-dependent atomic orbitals for the calculation of spin-rotation constants and rotational g tensors. *J. Chem. Phys.* **1996**, *105*, 2804–2812.

(58) Flygare, W. Magnetic interactions in molecules and an analysis of molecular electronic charge distribution from magnetic parameters. *Chem. Rev.* **1976**, *74*, 653–687.

(59) Sayvetz, A. The Kinetic Energy of Polyatomic Molecules. *J. Chem. Phys.* **1939**, *7*, 383–389.

(60) Eckart, C. Some Studies Concerning Rotating Axes and Polyatomic Molecules. *Phys. Rev.* **1935**, *47*, 552–558.

(61) Nielsen, H. H. The vibration-rotation energies of molecules. *Rew. Mod. Phys.* **1951**, *23*, 90–136.

(62) Watson, J. K. G. Simplification of the molecular vibration-rotation hamiltonian. *Mol. Phys.* **1968**, *15*, 479–490.

(63) Aliev, M.; Watson, J. K. G. *Molecular spectroscopy: modern research*; K. Narahari Rao, 1985; pp 1–67.

(64) Shavitt, I.; Bartlett, R. *Many-Body Methods in Chemistry and Physics*; Cambridge University Press, 2009.

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

(66) Harrison, R. J.; Fitzgerald, G. B.; Laidig, W. D.; Barteltt, R. J. Analytic MBPT(2) second derivatives. *Chem. Phys. Lett.* **1986**, *124*, 291.

(67) Stratmann, R. E.; Burant, J. C.; Scuseria, G. E.; Frisch, M. J. Improving harmonic vibrational frequencies calculations in density functional theory. *J. Chem. Phys.* **1997**, *106*, 10175–10183.

(68) Dunning Jr., 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.

(69) Woon, D. E.; Dunning Jr., T. H. Gaussian basis sets for use in correlated molecular calculations. III. The atoms aluminum through argon. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

(70) Woon, D. E.; Dunning Jr., T. H. Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *J. Chem. Phys.* **1995**, *103*, 4572–4585.

(71) Peterson, K. A.; Dunning Jr., 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.

(72) 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.

(73) Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652.

(74) Stephens, P. J.; Devlin, F. J.; Chabalowski, C. F.; Frisch, M. J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem.* **1994**, *98*, 11623.

(75) The SNSD basis set is available for download (accessed October 2014). `http://compchem.sns.it`.

(76) Barone, V.; Bloino, J.; Biczysko, M. Validation of the DFT/N07D computational model on the magnetic, vibrational and electronic properties of vinyl radical. *Phys. Chem. Chem. Phys.* **2010**, *12*, 1092.

(77) Kendall, R. A.; Dunning Jr., T. H.; Harrison, R. J. Electron affinities of the first row atoms revisited. Systematic basis sets and wave functions. *J. Chem. Phys.* **1992**, *96*, 6796.

(78) Stanton, J.; Gauss, J.; Harding, M.; Szalay, P.; Auer, A.; Bartlett, R.; Benedikt, U.; Berger, C.; Bernholdt, D.; Bomble, Y. et al. CFOUR, Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package. 2010; `http://www.cfour.de/`.

(79) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A. et al. Gaussian Development Version, Revision H.32. Gaussian Inc. Wallingford CT 2013.

(80) Schneider, W.; Thiel, W. Anharmonic force fields from analytic second derivatives: Method and application to methyl bromide. *Chem. Phys. Lett.* **1989**, *157*, 367.

(81) Stanton, J. F.; Gauss, J. Analytic second derivatives in high-order many-body perturbation and coupled-cluster theories: Computational considerations and applications. *Int. Rev. Phys. Chem.* **2000**, *19*, 61–95.

(82) Thiel, W.; Scuseria, G.; Schaefer, H. F. S.; Allen, W. D. The anharmonic force fields of HOF and F_2O. *J. Chem. Phys.* **1988**, *89*, 4965–4975.

(83) Barone, V. Anharmonic vibrational properties by a fully automated second-order perturbative approach. *J. Chem. Phys.* **2005**, *122*, 014108.

(84) Barone, V. Characterization of the potential energy surface of the HO_2 molecular system by a density functional approach. *J. Chem. Phys.* **1994**, *101*, 10666–10676.

(85) Stanton, J. F.; Lopreore, C. L.; Gauss, J. The equilibrium structure and fundamental vibrational frequencies of dioxirane. *J. Chem. Phys.* **1998**, *108*, 7190–7196.

(86) Benson, R. C.; Flygare, W. H. The molecular Zeeman effect of vinylene carbonate, maleic anhydride, acrolein and the benzene isomers, 3,4-dimethylenecyclobutene and fulvene. *J. Chem. Phys.* **1973**, *58*, 2366–2372.

(87) Blom, C. E.; Grassi, G.; Bauder, A. Molecular structure of s-cis- and s-trans-acrolein determined by microwave spectroscopy. *J. Am. Chem. Soc.* **1984**, *106*, 7427–7431.

(88) Esselman, B. J.; Amberger, B. K.; Shutter, J. D.; Daane, M. A.; Stanton, J. F.; Woods, R. C.; McMahon, R. J. Rotational spectroscopy of pyridazine and its isotopologs from 235-360 GHz: Equilibrium structure and vibrational satellites. *J. Chem. Phys.* **2013**, *139*, 224304.

(89) Vogt, N.; Demaison, J.; Rudolph, H. Equilibrium structure and spectroscopic constants of maleic anhydride. *Struct. Chem.* **2011**, *22*, 337–343.

(90) Vogt, N.; Demaison, J.; Vogt, J.; Rudolph, H. D. Why it is sometimes difficult to determine the accurate position of a hydrogen atom by the semiexperimental method: Structure of molecules containing the OH or the CH_3 group. *J. Comp. Chem.* **2014**, *35*, 2333–2342.

(91) Craig, N. C.; Chen, Y.; Fuson, H. A.; Tian, H.; van Besien, H.; Conrad, A. R.; Tubergen, M. J.; Rudolph, H. D.; Demaison, J. Microwave Spectra of the Deuterium Isotopologues of *cis*-Hexatriene and a Semiexperimental Equilibrium Structure. *J. Phys. Chem. A* **2013**, *117*, 9391–9400.

(92) 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.

(93) Puzzarini, C.; Cazzoli, G. Equilibrium structure of protonated cyanogen, $HNCCN^+$. *J. Mol. Spectrosc.* **2009**, *256*, 53.

(94) 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.

(95) Thorwirth, S.; Harding, M. E.; Muders, D.; Gauss, J. The empirical equilibrium structure of diacetylene. *J. Mol. Spectrosc.* **2008**, *251*, 220–223.

(96) Puzzarini, C. Ab initio anharmonic force field and equilibrium structure of the sulfonium ion. *J. Mol. Spectrosc.* **2007**, *242*, 70–75.

(97) Pietropolli-Charmet, A.; Stoppa, P.; Tasinato, N.; Giorgianni, S.; Barone, V.; Biczysko, M.; Bloino, J.; Cappelli, C.; Carnimeo, I.; Puzzarini, C. An integrated experimental and quantum-chemical investigation on the vibrational spectra of chlorofluoromethane. *J. Chem. Phys.* **2013**, *139*, 164302.

(98) Puzzarini, C.; Cazzoli, G.; Gambi, A.; Gauss, J. Rotational spectra of 1-chloro-2-fluoroethylene. II. Equilibrium structures of the cis and trans isomer. *J. Chem. Phys.* **2006**, *125*, 054307.

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

(100) Larsen, R. W.; Pawlowski, F.; Hegelund, F.; Jørgensen, P.; Gauss, J.; Nelander, B. The equilibrium structure of trans-glyoxal from experimental rotational constants and calculated vibration-rotation interaction constants. *Phys. Chem. Chem. Phys.* **2003**, *5*, 5031–5037.

(101) Vogt, N.; Demaison, J.; Rudolph, H. D. Accurate equilibrium structures of fluoro- and chloroderivatives of methane. *Mol. Phys.* **2014**, *112*, 2873–2883.

(102) Dunning Jr., 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.

(103) Craig, N. C.; Groner, P.; McKean, D. C. Equilibrium Structures for Butadiene and Ethylene: Compelling Evidence for Π-Electron Delocalization in Butadiene. *J. Phys. Chem. A* **2006**, *110*, 7461–7469, PMID: 16759136.

(104) Lide, D. R.; Mann, D. E. Microwave Spectra of Molecules Exhibiting Internal Rotation. I. Propylene. *J. Chem. Phys.* **1957**, *27*, 868–873.

(105) Demaison, J.; Rudolph, H. Ab initio anharmonic force field and equilibrium structure of propene. *J. Mol. Spectrosc.* **2008**, *248*, 66–76.

(106) Lide, D. R.; Christensen, D. Molecular Structure of Propylene. *J. Chem. Phys.* **1961**, *35*, 1374–1378.

(107) Gauss, J.; Cremer, D.; Stanton, J. F. The re Structure of Cyclopropane. *J. Phys. Chem. A* **2000**, *104*, 1319–1324.

(108) Gauss, J.; Stanton, J. F. The Equilibrium Structure of Benzene. *J. Phys. Chem. A* **2000**, *104*, 2865–2868.

(109) Thorwirth, S.; Müller, H. S. P.; Winnewisser, G. The Millimeter- and Submillimeter-Wave Spectrum and the Dipole Moment of Ethylenimine. *J. Mol. Spectrosc.* **2000**, *199*, 116–123.

(110) Thorwirth, S.; Gendriesch, R.; Müller, H. S. P.; Lewen, F.; Winnewisser, G. Pure Rotational Spectrum of Ethylenimine at 1.85 THz. *J. Mol. Spectrosc.* **2000**, *201*, 323–325.

(111) Bak, B.; Skaarup, S. The substitution structure of ethyleneimine. *J. Mol. Struct.* **1971**, *10*, 385–391.

(112) Császár, A. G.; Demaison, J.; Rudolph, H. D. Equilibrium Structures of 3-, 4-, 5-, 6-, and 7-Membered Unsaturated N-Containing Heterocycles. *J. Phys. Chem. A* **2014**, doi: 10.1021/jp5084168.

(113) Christen, D.; Griffiths, J. H.; Sheridan, J. The Microwave Spectrum of Imidazole; Complete Structure and the Electron Distribution from Nuclear Quadrupole Coupling Tensors and Dipole Moment Orientation. *Z. Naturforsch.* **1982**, *37a*, 1378.

(114) Demaison, J.; Császár, A. G.; Margulès, L. D.; Rudolph, H. D. Equilibrium Structures of Heterocyclic Molecules with Large Principal Axis Rotations upon Isotopic Substitution. *J. Phys. Chem. A* **2011**, *115*, 14078–14091.

(115) Kochikov, I. V.; Tarasov, Y. I.; Spiridonov, V. P.; Kuramshina, G. M.; Rankin, D. W. H.; Saakjan, A. S.; Yagola, A. G. The equilibrium structure of thiophene by the combined use of electron diffraction, vibrational spectroscopy and microwave spectroscopy guided by theoretical calculations. *J. Mol. Struct.* **2001**, *567*, 29–40.

(116) Mata, F.; Quintana, M. J.; Sørensen, G. O. Microwave spectra of pyridine and monodeuterated pyridines. Revised molecular structure of pyridine. *J. Mol. Struct.* **1977**, *42*, 1–5.

(117) Sørensen, G. O.; Mahler, L.; Rastrup-Andersen, N. Microwave spectra of [^{15}N] and [^{13}C] pyridines, quadrupole coupling constants, dipole moment and molecular structure of pyridine. *J. Mol. Struct.* **1974**, *20*, 119–126.

(118) Blukis, U.; Kasai, P. H.; Myers, R. J. Microwave Spectra and Structure of Dimethyl Ether. *J. Chem. Phys.* **1963**, *38*, 2753–2760.

(119) Groner, P.; Albert, S.; Herbst, E.; De Lucia, F. C. Dimethyl Ether: Laboratory Assignments and Predictions through 600 GHz. *Astrophys. J.* **1998**, *500*, 1059.

(120) Niide, Y.; Hayashi, M. Reinvestigation of microwave spectrum of dimethyl ether and rs structures of analogous molecules. *J. Mol. Spectrosc.* **2003**, *220*, 65–79.

(121) Demaison, J.; Herman, M.; Liévin, J. Anharmonic force field of *cis*- and *trans*-formic acid from high-level ab initio calculations, and analysis of resonance polyads. *J. Chem. Phys.* **2007**, *126*, 164305.

(122) Demaison, J.; Margulès, L.; Kleiner, I.; Császár, A. Equilibrium structure in the presence of internal rotation: A case study of cis-methyl formate. *J. Mol. Spectrosc.* **2010**, *259*, 70.

(123) Marstokk, K. M.; Møllendal, H. Microwave spectrum and dipole moment of glycolaldehyde. *J. Mol. Spectrosc.* **1970**, *5*, 205–213.

(124) Bouchez, A.; Margulès, L.; Motiyenko, R. A.; Guillemin, J.-C.; Walters, A.; Bottinelli, S.; Ceccarelli, C.; Kahane, C. The submillimeter spectrum of deuterated glycolaldehydes. *Astron. Astrophys.* **2012**, *540*, A51.

(125) Haykal, I.; Motiyenko, R. A.; Margulès, L.; Huet, T. R. Millimeter and submillimeter wave spectra of ^{13}C-glycolaldehydes. *Astron. Astrophys.* **2013**, *549*, A96.

(126) Carroll, P. B.; Drouin, B. J.; Widicus Weaver, S. L. The Submillimeter Spectrum of Glycolaldehyde. *Astrophys. J.* **2010**, *723*, 845.

(127) Carroll, P. B.; McGuire, B. A.; Zaleski, D. P.; Neill, J. L.; Pate, B. H.; Widicus Weaver, S. L. The pure rotational spectrum of glycolaldehyde isotopologues observed in natural abundance. *J. Mol. Spectrosc.* **2013**, *284*, 21–28.

(128) Bakri, B.; Demaison, J.; Margulès, L.; Møllendal, H. The Submillimeter-Wave Spectrum and Quantum Chemical Calculations of Glyoxylic Acid. *J. Mol. Spectrosc.* **2001**, *208*, 92–100.

(129) Demaison, J. In *Equilibrium Molecular Structures: From Spectroscopy to Quantum Chemistry*; Demaison, J., Boggs, J., Császár, A., Eds.; CRC Press, 2011; pp 29–52.

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