Published in final edited form as:

Can J Chem. 2016 December ; 94(12): 1065–1076. doi:10.1139/cjc-2016-0282.

Structural features of the carbon-sulfur chemical bond: a semi-experimental perspective

Emanuele Penocchio, Marco Mendolicchio, Nicola Tasinato, and Vincenzo Barone*

Scuola Normale Superiore, Pisa, Italy

Abstract

In this work semi-experimental and theoretical equilibrium geometries of 10 sulfur-containing organic molecules, as well as 4 oxygenated ones, are determined by means of a computational protocol based on density functional theory. The results collected in the present paper further enhance our online database of accurate semi-experimental equilibrium molecular geometries, adding 13 new molecules containing up to 8 atoms, for 12 of which the first semi-experimental equilibrium structure is reported, to the best of our knowledge. We focus in particular on sulfur-containing compounds, aiming both to provide new accurate data on some rather important chemical moieties, only marginally represented in the literature of the field, and to examine the structural features of carbon-sulfur bonds in the light of the previously presented linear regression approach. The structural changes issuing from substitution of oxygen by sulfur are discussed to get deeper insights on how modifications in electronic structure and nuclear potential can affect equilibrium geometries. With respect to our previous works, we perform non-linear constrained optimizations of equilibrium SE structures with a new general and user-friendly software under development in our group with updated definition of useful statistical indicators.

Keywords

Equilibrium Structure; DFT; Geometry Database; Nonlinear Regression

1 Introduction

In a recent perspective paper¹, George M. Whitesides, one of the most cited chemists² of the last decade, pointed out that one of the main concerns that chemistry has still to solve is the so-called “rational drug design”.^{3–5} Among the various experimental and theoretical steps one should overcome to reach such an ambitious target, those involving the prediction of the tertiary structure of a protein or a drug from their chemical *formulae* are still extremely challenging. Despite the large number of qualitative techniques developed nowadays, an accurate prediction of the three-dimensional structure continues to be a difficult task, even for medium-sized molecules. In order to overcome such limitations and go beyond our actual knowledge about the most peculiar features determining molecules’ tertiary structures, a systematic and accurate study of a large collection of molecules containing the relevant functional groups is mandatory. Another fascinating challenge of modern chemistry

*To whom correspondence should be addressed vincenzo.barone@sns.it.

is the search for additional experimental evidences about the origin of life and, as a strictly connected topic, extraterrestrial chemistry. 1,6,7 In this kind of research, computational microwave and infrared spectroscopies play a leading role in driving the interpretation of experimental data and elucidating the chemical composition of both planets' atmospheres and interstellar medium^{8–10}. On the other hand, the theoretical prediction of reliable spectroscopic properties relies on the computation of accurate equilibrium geometries of isolated molecular systems,^{11,12} that present the advantages of having a unique definition in the Born-Oppenheimer (B-O) picture and being independent from isotopic substitution and vibrational effects.^{11–26}

It is evident from the above discussion that a large database of accurate equilibrium molecular structures is of primary importance as it provides the required reference data for: (i) testing the precision of various quantum-mechanical (QM) approaches;^{14,27–31} (ii) parametrizing new density-functionals and force fields for bio-chemical applications;^{32–41} (iii) developing and validating new computational protocols for supporting and driving experimental observations^{11,42}.

Detailed knowledge of the equilibrium structures of isolated molecular systems, especially those of biological interest, is then a prerequisite for a deeper understanding and a more accurate modelling of paramount aspects of current chemical research. Unfortunately, from both experimental and theoretical points of view, the effort needed to obtain accurate equilibrium geometries becomes prohibitive except for very small systems.^{43–48} An important breakthrough, first introduced by Pulay and co-workers⁴⁹, has been the adoption of the so-called semi-experimental (SE) equilibrium geometry (r_e^{SE})^{50,51}, which can be obtained by a least-squares fit (LSF) of experimental rotational constants (B_0^β) for the ground vibrational state of different isotopologues corrected by vibrational contributions (ΔB_{vib}^β) computed theoretically. 13,14,49 This interplay between experimental measurements and theoretical calculations paved the route toward the extension of accurate structural studies to systems larger than those treatable by experimental and QM methods separately. 19,51–53

Thanks to the (never too) large set of available experimental data issuing from microwave spectroscopy, we employed DFT-based model chemistries to calculate ΔB_{vib}^β and afterwards r_e^{SE} for molecules containing up to 15 atoms, mostly including molecular building blocks of chemical, biological and technological interest. 54,55 The B3LYP functional^{56,57} in conjunction with a polarized double- ζ basis set has been shown to be an optimal compromise between accuracy and computational cost, thus paving the route to the extension of the SE approach to the structural analysis of larger systems.⁵⁴ In the few cases where the B3LYP-based scheme did not deliver high accuracy, the use of the B2PLYP functional^{58–60} coupled with a doubly polarized triple- ζ basis set was shown to represent a very accurate approach in the calculation of r_e^{SE} , without any significant loss of accuracy with respect to higher level post-Hartree-Fock calculations.^{55,61} The results we have so far obtained have been made available in a graphical interactive form on our web site, 62 where

r_e^{SE} of various molecules obtained using either B3LYP and B2PLYP functionals are collected in two databases (B3se and B2se set, respectively), together with other kinds of structures that may be useful, such as their fully theoretical counterparts (B3th and B2th sets) and, when feasible, r_e^{SE} obtained using coupled cluster (CC) techniques for the calculation of ΔB_{vib}^β (CCse set).⁶² These SE equilibrium structures not only represent a high quality benchmark for structural and spectroscopic studies, but, as already pointed out elsewhere^{54,55}, they can be seen as a set of “structural synthons” to be used, in the framework of the previously presented templating molecule and linear regression approaches (TMA and LRA, respectively), in order to build reliable equilibrium structures of larger molecules, such as amino acids, polypeptides and nucleobases.^{55,61}

The main aim of the present work is to introduce and discuss 10 new SE equilibrium structures of interest for modeling molecules containing one sulfur atom and hence of bio- and astro-chemical relevance. In addition, in the case of 4 molecules, we have also studied the SE equilibrium structure of the oxygenated analogues, in order to (i) investigate the structural effects of sulfur-oxygen substitution, thus gaining some insight on how the changes in electronic structure and nuclear potential could affect r_e^{SE} , and (ii) enlarge further our data set. The new structures add some important moieties to our collection, in which sulfur-containing functional groups were underrepresented until now, in spite of their important role in biological molecules^{63,64} as well as in prebiotic⁶⁵ and interstellar chemistry.^{66,67}

2 Methodology and Computational Details

In order to obtain the complete r_e^{SE} of a molecule, one has to perform a LSF on a set of molecular coordinates to the semi-experimental equilibrium moments of inertia, hereafter denoted as $(I_e^\beta)^{SE}$, of as many isotopologues as possible, where $\beta = a, b$ or c is one of the principal inertial axes in the molecule-fixed reference frame. The $(I_e^\beta)^{SE}$ can be obtained starting from the ground state experimental rotational constants, $(B_0^\beta)^{EXP}$, using the following equality

$$(I_e^\beta)^{SE} = \hbar / \left(4\pi (B_e^\beta)^{SE} \right) \quad (1)$$

where the equilibrium SE rotational constants $(B_e^\beta)^{SE}$, are given by

$$(B_e^\beta)^{SE} = (B_0^\beta)^{EXP} - (-\Delta B_0^\beta)^{QM} \quad (2)$$

In the previous equation $(\Delta B_0^\beta)^{\text{QM}}$ represents a correction computed from a QM approach and explicitly given by

$$\begin{aligned} (\Delta B_0^\beta)^{\text{QM}} &= \frac{m_e}{M_p} g^{\beta\beta} B_e^\beta - \sum_i \frac{\alpha_i^\beta d_i}{2} \\ &= \Delta B_{\text{el}}^\beta + \Delta B_{\text{vib}}^\beta \end{aligned} \quad (3)$$

$\Delta B_{\text{el}}^\beta$ is an electronic contribution, which derives from the rotational $\mathbf{g}$ tensor and the ratio between electron (m_e) and proton (M_p) masses^{68–70}. This term has been systematically included in our computations, even if often negligible. Within B-O approximation and Eckart-Sayvetz conditions^{71–74}, the expression of the vibrational contribution $\Delta B_{\text{vib}}^\beta$ can be found with the aid of second-order vibrational perturbation theory (VPT2) applied to the molecular ro-vibrational Hamiltonian expressed in normal coordinates^{75–77}. In the summation, α_i^β are the vibration-rotation interaction constants and d_i is the degeneracy of the i -th vibrational normal mode^{68,77,78}.

As can be inferred by looking at the explicit form of α_i^β -terms for asymmetric tops⁶⁸,

$$\alpha_i^\beta = -\frac{2(B_e^\beta)^2}{\omega_i} \left[\sum_{\gamma=x,y,z} \frac{3(a_{i,\beta\gamma})^2}{4I_e^\beta} + \sum_{j \neq i=1}^{3N-6} (\zeta_{ij,\beta})^2 \left[\frac{(\omega_i - \omega_j)^2}{\omega_i + \omega_j} + \frac{(\omega_i + \omega_j)^2}{\omega_i - \omega_j} \right] + \pi \left(\frac{c}{h} \right)^{1/2} \sum_{j=1}^{3N-6} K_{ijj} \frac{a_{j,\beta\beta} \omega_i}{\omega_j^{3/2}} \right] \quad (4)$$

the computational bottleneck of this procedure is the calculation of the elements K_{ijj} of the cubic force field. The other terms in the above equation are the inertial derivatives ($a_{i,\beta\gamma}$), the Coriolis coupling constants ($\zeta_{ij,\beta} = -\zeta_{ji,\beta}$) and the harmonic frequencies (ω_j); all the remaining symbols take their usual meaning. It is worth underlining how all possible terms involving accidental resonances ($\omega_i \approx \omega_j$) in equation 4 cancel once the summation in

equation 3 is performed, so that the $(\Delta B_0^\beta)^{\text{QM}}$ correction is devoid of the problems related to resonances. On the basis of previous validations^{54,55} all anharmonic calculations have been performed employing DFT techniques, using either the hybrid B3LYP^{56,57,79} and the double hybrid B2PLYP^{58–60} functionals. While the latter functional has always been coupled with a standard correlation-consistent polarized triple- ζ cc-pVTZ basis set^{80–83} (shortly denoted as VTZ), the B3LYP functional has been used in conjunction with the polarized double- ζ SNSD basis set^{84,85}, which was developed in our group starting from a standard 6-31G basis set and adding, with a subsequent optimization of orbital exponents, one contracted and one diffuse s -type function on each atom, together with diffuse p and d -types functions on second- and third-row atoms.⁸⁶ Both basis sets have been chosen because several studies have shown them to be an excellent compromise between accuracy and

computational cost for the calculation of several spectroscopic properties, in particular for anharmonic computational models.^{31,48,60,84,85,87–90} The $\Delta B_{\text{el}}^{\beta}$ contributions have been evaluated using the B3LYP functional in conjunction with the aug-cc-pVTZ (hereafter AVTZ) basis set^{81,91} in the calculation of the $g^{\beta\beta}$ quantities. All the calculations have been performed with the G09-D01 package⁹², computing harmonic contributions from analytical second derivatives and cubic force fields by numerical differentiation of analytical second derivatives of the electronic energy with respect to geometrical parameters, with the default step of 0.01 Å in the normal-coordinate representation^{93–98}. Very tight convergence criteria have been used for the SCF procedure and geometry optimization, together with an ultra-fine grid for the numerical integration connected to the exchange-correlation functional and its derivatives.

2.1 Least Squares Fitting and Error Analysis

Concerning the fitting procedure, all SE equilibrium geometries presented in this work have been obtained using a new software under development in our group, which allows one to perform non-linear constrained optimizations. One of the main features of the new tool is the possibility of computing the analytical Hessian matrix of the residual function (the squared norm of the residual vector) at the optimized geometry, thus allowing to characterize the critical point reached at the end of the minimization procedure. Its robustness has been proved by studying a representative set of different molecular systems, while the reliability of the results has been confirmed by comparison with the structural parameters computed by the *xrefit* module of the CFOUR program package⁹⁹. Although the full implementation of the new software will be presented in a forthcoming paper, we discuss here some details about the error estimate in order to avoid possible ambiguities arising from slightly different approximations employed in the available software, especially with respect to the *xrefit* code used in our previous work. As it is well known, the aim of an optimization is the determination of the vector $\mathbf{x}^*$ minimizing the residual function $R_f(\mathbf{x})$:

$$R_f(\mathbf{x}) = \sum_{i=1}^N r_i^2(\mathbf{x}) = \sum_{i=1}^N \left(y_i^{\text{SE}} - y_i^{\text{guess}}(\mathbf{x}) \right)^2 = \mathbf{R}(\mathbf{x})^T \mathbf{R}(\mathbf{x}) = \|\mathbf{R}(\mathbf{x})\|^2 \quad (5)$$

where $\mathbf{y}^{\text{SE}}$ and $\mathbf{y}^{\text{guess}}$ are N -dimensional vectors containing the semi-experimental and the guessed data, respectively, $\mathbf{R}$ represents the corresponding residual vector (with elements r_i), while $\mathbf{x} = (x_1, \dots, x_M)$ is the set of variables characterizing the optimization procedure. Due to the non-linearity of the residual function with respect to the variables $\mathbf{x}$, the optimization has to be performed through an iterative procedure, where at each iteration k the correction vector for the geometry can be found by solving a linear system of the type:

$$\begin{aligned} \mathbf{B}(\mathbf{x}_k) \Delta \mathbf{x}_k &= -\mathbf{g}(\mathbf{x}_k) \\ \mathbf{x}_{k+1} &= \mathbf{x}_k + \alpha_k \Delta \mathbf{x}_k \end{aligned} \quad (6)$$

where $\mathbf{B}(\mathbf{x}_k)$ is a general symmetric and positive definite matrix, $\mathbf{x}_k$ is the correction vector, $\mathbf{g}(\mathbf{x}_k)$ is the gradient of the residual function, and α_k is a scaling factor which ensures that

the value of the function decreases at each iteration¹⁰⁰. Once the optimized set of structural parameters $\mathbf{x}^*$ has been computed, the error-related quantities can be defined.

In our previous studies^{54,55,99} we adopted the following formulation of root-mean-square (RMS) error as an indicator of fits' quality

$$\overline{N} \quad (7)$$

In the present work this has been replaced by the residual standard deviation s defined by Demaison¹⁰¹ as

$$\sqrt{N-M} \quad (8)$$

in order to take into account also the number of degrees of freedom in discussing our results.

The estimates of the standard errors on the optimized variables are equal to the square roots of the diagonal entries of the variance-covariance matrix $C(\mathbf{x}^*)$:

$$C(\mathbf{x}^*) = s^2 \mathbf{H}^{-1}(\mathbf{x}) \quad (9)$$

where $\mathbf{H}(\mathbf{x}^*)$ is the Hessian matrix of the residual function evaluated at the optimum geometry. The standard deviation $s(x_j)$ of each parameter x_j can be defined as follows,

$$s(x_j) = \sqrt{[C(\mathbf{x}^*)]_{jj}} \quad (10)$$

which is the strategy adopted, for instance, in the *xrefit* code.

At variance, the uncertainties presented in the following have been defined through the evaluation of the confidence interval for each parameter x_j :

$$x_j - ts(x_j) \leq x_j \leq x_j + ts(x_j) \quad (11)$$

where t is given by the Student distribution. Eq. 11 can be easily extended to any function of the variables $\mathbf{x}$ through proper error propagation¹⁰¹.

In solving equations 6, it may happen that some parameters turn out to be very sensitive to small changes in the data and they suffer from large uncertainties so that they have poor significance. This is often due to the correlation of parameters that leads to an ill-conditioned optimization problem¹⁰¹. A reliable yardstick for assessing whether this kind of problems are present is provided by the condition number (κ)¹⁰²:

$$\kappa(\mathbf{J}) = \frac{\sqrt{\lambda_{\max}}}{\sqrt{\lambda_{\min}}} \quad (12)$$

in which $\lambda_{\max}$ and $\lambda_{\min}$ are the largest and the smallest eigenvalue of the square matrix $\mathbf{J}^T \mathbf{J}$, $\mathbf{J}$ being the Jacobian matrix whose columns have been normalized. As a rule of thumb, a large condition number ($\sim 10^3$ in the present case) has to be considered diagnostic for the presence of potentially harmful ill-conditioning^{101,102}.

3 Semi-Experimental Equilibrium Structures

The results for the molecular structure determination of the 14 molecules considered in this work are summarized in Tables 1, 2 and 3, with atoms labelled according to Figures 1 and 2, in the case of ambiguities. To the best of our knowledge, this is the first r_e^{SE} determination for thiirane, thiazole, 1,2,3-thiadiazole, 1,3,4-thiadiazole, 1,2,4-thiadiazole, 1,2,5-thiadiazole, isothiocyanic acid, thioformaldehyde, thioketene, propadienethione, isocyanic acid and propadienone. For all the molecules under study, the theoretical structure fully optimized at B3LYP/SNSD and B2PLYP/VTZ levels is also reported. Concerning r_e structures, the new results substantially reproduce our previous findings about the good performances of the B2PLYP/VTZ model chemistry⁵⁵, that shows a mean absolute error (MAE) with respect to SE equilibrium bond lengths lower than 0.005 Å (about half the value of 0.01 Å issuing from B3LYP/SNSD computations). The good accuracy of B2PLYP in predicting equilibrium structures has been recently extended by Adamo and co-workers to a larger panel of double-hybrid functionals³². In all cases our QM calculations systematically overestimate both $r_e(C-S)$ and $r_e(N-S)$, showing a mean error (ME), with respect to their r_e^{SE} counterparts, that is much larger than the previously mentioned averages, namely 0.0248 Å and 0.0623 Å for B3LYP/SNSD and 0.0097 Å and 0.0247 Å for B2PLYP/VTZ models. These values, about twice (for $r_e(C-S)$) and six times (for $r_e(N-S)$) larger than the average MAE on bond lengths involving only second row atoms, evidence the significant reduction in the accuracy of DFT-based methods for sulfur. We tried to optimize geometries with larger basis sets, obtained by adding various combinations of contracted d-functions on the sulfur atom, but the marginal improvement achieved suggests that the problem of reduced accuracy of structural parameters involving S atoms may be related to intrinsic limitations of the chosen functionals. In this respect, we optimized the geometries of selected molecules using M06-2X¹⁰³ and BHandH¹⁰⁴ two global hybrid functionals with a larger amount of Hartree-Fock exchange than B3LYP, using the VTZ basis set. In both cases (see Table 1 in Supporting Information) there is a shortening of almost all chemical bonds with respect to B3LYP/SNSD and B2PLYP/VTZ calculations, that in some instances provides a better description of the molecular structure around sulfur atoms, but, in general, does not improve the overall structural predictions. Therefore, we think that B3LYP and B2PLYP remain the methods of choice for the determination of semi-experimental equilibrium structures. For the sake of completeness we have also reported the fully empirical r_0 structure obtained by a least squares fit of bare experimental moments of inertia^{14,105}. Although this completely empirical structure is expected to be significantly less accurate than the SE one¹⁰⁶, its

differences with respect to SE equilibrium parameters can be roughly interpreted as a measure of molecular flexibility concerning both the whole structure and the individual bonds. In general terms our results discourage the use of the r_0 structure even for inferring qualitative considerations. The residual standard deviation (s), in terms of differences between observed and fitted moments of inertia, is always reported as an indicator of the quality of the fit and of the mutual coherence between experimental and theoretical data. Finally, it is worth to underline that the fits took an average number of 7 iterations in order to achieve convergence.

3.1 Thiirane, Thiazole, 1,2,3-thiadiazole, 1,2,4-thiadiazole, 1,2,5-thiadiazole and 1,3,4-thiadiazole

The chemical proprieties of heterocyclic organosulfur compounds are of considerable interest in modern research. In particular thiirane, thiazole and thiadiazoles have been found to have important biological¹⁰⁷, pharmacological^{108–110} and technological^{111–113} applications; furthermore they are believed to play a significant role in astrochemical processes^{114,115} and prebiotic chemistry.^{116,117} With the exception of thiirane, the planarity of these molecules, highlighted by an inertial defect ($\Delta I^c = I^c - I^b - I^a$) close to zero, leads to a dependence among the inertia moments, so it is necessary to select one of their possible couples in the fitting procedure¹⁴. For all the cases taken into account, the specific choice did not significantly alter the SE results, which means that we have dealt with problems without strong ill-conditioning. In the following we consider fits on I^a and I^b , because choosing the two smallest inertia moments usually gives the best results for the r_0 structure¹⁰⁶.

Thiirane, the simplest S-containing heterocyclic compound, is the sulfur analogue of oxirane (ethylene oxide) and it belongs to the C_{2v} symmetry point group. Its rotational and vibrational spectra are well known¹¹⁸, as well as those of its major isotopomers^{119–121}. The use of all the experimental moments of inertia measured till now permitted to fit the complete r_e^{SE} structure reported in Table 1, where the excellent agreement between B3LYP/SNSD and B2PLYP/VTZ SE schemes can be appreciated. The small magnitude of both condition numbers and uncertainties on resulting parameters confirm the good quality of both calculated and measured data.

Thiazole, or 1,3-thiazole, is a planar molecule (C_s symmetry) for which all mono-substituted isotopologues have been investigated¹²², thus allowing us to obtain its first complete SE equilibrium structure. As apparent from Table 1, this system suffers from larger uncertainties on the fitted parameters in comparison to the other molecules studied in this paper. In our opinion, this can be due to both possible inconsistencies in the experimental data, as suggested by a condition number about 50 times that of thiirane, and to the restricted number of degrees of freedom ($=5$) in the regression problem, the latter implicating a value of Student's t distribution of 2.57. In any case, it is worth noting how the inertial defect, that should vanish for planar systems, considerably decreases when experimental rotational constants are corrected by vibrational and electronic contributions, confirming the consistency of our computational schemes in the calculation of $(\Delta B_0^\beta)^{QM}$.

1,2,3-, 123 1,3,4-, 124 1,2,4-, 125 and 1,2,5-thiadiazole, 126 with C_s , $C_{2,v}$, C_s , and $C_{2,v}$ symmetry, respectively, represent all possible thiadiazoles and all of them have been extensively studied. For these molecules inspection of Table 1 shows satisfactory results, with low statistical uncertainties and differences between B3LYP/SNSD and B2PLYP/VTZ SE parameters that rarely exceed 0.001 Å for bond lengths or 0.1 degrees for bond angles. The only noteworthy discrepancy can be found in the $r(\text{C3-H})$ bond length of 1,2,4-thiadiazole (1.0829 Å and 1.0784 Å for B3LYP/SNSD and B2PLYP/VTZ vibrational corrections respectively), which shows large uncertainty together with $a(\text{N2-C3-H})$. This is probably due to the absence of experimental data concerning H→D isotopic substitution causing a scarce sensitivity of the fit to the H3 coordinate, that in turn affects both $r(\text{C3-H})$ and $a(\text{N2-C3-H})$. Despite this, the corresponding values obtained from B3LYP and B2PLYP computations are in agreement within their relatively large confidence intervals.

3.2 Isothiocyanic Acid, Thioformaldehyde, Thioketene, Propadienethione and their Oxygen Analogues

Sulfur and oxygen have a quite similar chemistry and can be found in analogous compounds. This parallelism can be exploited in order to understand, in a B-O picture, how changes in electronic structure and nuclear potential arising from the substitution of O with S can affect, qualitatively and quantitatively, the structural features of a molecule in its ground state equilibrium configuration. As an example of this kind of comparative analysis, we parallel cumulene-thiones up to the third member of the series (see Table 2) to their cumulenone counterparts (see Table 3), both families being of potential astrochemical and astrobiological interest.^{127–129} Also isothiocyanic and isocyanic acids are reported in the same tables. All the molecules discussed in this section are near-prolate asymmetric tops because of the linearity, or quasilinearity, of their heavy atom chains, which causes one eigenvalue of their inertia tensors (I^b) to be quite small, so that the corresponding rotational constant (B_0^a) is particularly difficult to measure or even estimate. For this reason the inertial defects of these molecules have only marginal significance and are not reported in Tables 2 and 3; as another consequence all the r_e^{SE} and r_0 structures have been obtained by fitting the I^b and I^c inertia moments.

Both isothiocyanic and isocyanic acids have been studied exploiting all the most common single isotopic substitutions^{130,131}, obtaining very satisfactory results from the points of view of uncertainties (s values) and agreement between B3LYP/SNSD and B2PLYP/VTZ r_e^{SE} . These two molecules are topologically very similar in their equilibrium structure and the only significant difference introduced by the replacement of oxygen with sulfur is an obvious lengthening of the $r(\text{C-S})$ parameter with respect to its $r(\text{C-O})$ counterpart, *i.e.* a very localized effect.

Almost the same occurs for thioformaldehyde¹³² and thioketene¹³³ in comparison to their oxygen analogues, formaldehyde^{54,55} (already present in our previous data sets) and ketene¹³⁴, whose new DFT-based r_e^{SE} structures are in excellent agreement with the most accurate equilibrium geometry previously reported.¹³⁴

Moving to the third members of cumulene-thiones and cumulenes series, larger uncertainties and unexpected structural differences are observed in both equilibrium and r_0 structures of propadienethione and propadienone. In fact the heavy atoms of propadienethione are collinear¹³⁵, whereas the oxygenated counterpart presents a strongly kinked structure^{136–138}. The uncertainties and the residual standard deviation for the B3LYP/SNSD r_e^{SE} of propadienethione are considerably larger than those at the B2PLYP/VTZ level. This

behaviour can be explained by noting that in this case the $\Delta B_{\text{vib}}^\beta$ contributions are very small (see Table 2 in Supporting Information) and the relative error of the B3LYP/SNSD method with respect to more accurate calculations becomes significant, leading to relevant changes in the result, also possibly related to the presence of ill-conditioning effects. For propadienone the situation is even worst, with condition numbers of the fits larger than 3500, a clear signal of severe ill-conditioning that, in this case, increases the variances of estimated parameters and can be responsible for important rounding errors¹⁰¹.

Moreover, propadienone is the only molecule, among those studied in this work, for which the conditioning of the problem worsens moving from experimental to SE inertia moments,

as indicated by the increase of κ by almost 400 units when $(I_e^\beta)^{\text{SE}}$ are fitted instead of $(I_0^\beta)^{\text{EXP}}$. The poor results obtained for the r_e^{SE} structures of propadienone may be due to the failure of VPT2 in accurately computing the vibrational contribution to the rotational constants of this rather flexible molecular system. In particular we believe that the problem concerns the calculation of the vibration-rotation interaction constant of the lowest frequency normal mode, along which the bending potential of propadienone presents a double minimum with a strong anharmonicity¹³⁹. For those reasons our results on propadienone have poor quantitative significance, but we decided to report their values in Table 3 in the light of their qualitative interest discussed in the following paragraph.

3.3 The Strange Case of Propadienethione and Propadienone

The bent structure of propadienone is of particular interest not only because it is unexpected on the basis of both chemical intuition and comparison with shorter cumulenes, but also because Hartree-Fock calculations predict it to have a wrong C_{2v} structure analogous to that of propadienethione¹³⁸. In fact, it has been shown how in a B-O framework the non-linearity of the CCCO chain is a surprisingly strong effect of electron correlation, that affects the structural features of the molecular system, correctly predicted only if electron correlation is properly taken into account¹⁴⁰. In particular MC-SCF calculations pinpointed the importance of the double in-plane $\pi - \pi^*$ electronic excitation in kinking the molecule at the β -carbon^{141,142}, with a resulting wave function compatible with a valence bond based hypothesis about the importance of the $H_2C = C^- - C \equiv O^+$ resonance structure for the stabilization of the system, thus justifying distortion from linearity. Electron correlation has been found to have such a decisive role in shaping, and hence in predicting, atom spatial arrangement also in longer cumulenes^{143,144} and it is certainly worth noting how DFT methods succeed in qualitatively reproducing the propadienone r_e structure. Regarding propadienethione, even though QM calculations at any level correctly reproduce the experimentally confirmed C_{2v} geometry, one would like to explain, in terms of reasonable

chemical arguments, the driving force that straightens this kind of system upon oxygen-sulfur exchange. Resorting to Lewis resonance structures does not help, because *a-priori* there are not important arguments that exclude a significant contribution of $H_2C=C^{\cdot-}-C\equiv S^+$. On the one hand, this is reinforced by the similar chemistry usually involving O and S atoms, on the other hand by CASSCF calculations that we performed with qualitative purposes at various levels, obtaining C_{2v} absolute minima with a multiconfigurational wave function very similar to that of the kinked propadienone in terms of the relative weights of corresponding excited determinants; therefore other important electronic and energetic aspects could affect these systems.

At this point it is useful to resort to the very elegant and effective model proposed in 1987 by Trinquier and Malrieu¹⁴⁵ that relates the occurrence of distortions from linearity in cumulene-like molecules to the singlet-triplet energy separation ($E_{S\rightarrow T}$) of the interacting fragments forming the multiple bond involved in the bending. Without entering the details of the model, well explained in the original work¹⁴⁵, one can consider propadienone ($H_2C=C=O$) as formed by the interaction between one vinylidene ($H_2C=C^{\cdot}$) and one carbonyl ($:C=O$) fragments, both being diradical species with a singlet ground state and a $E_{S\rightarrow T}$ of 46 and 139 kcal/mol, respectively.¹⁴⁶ As it is well known,^{147–149} the formation of a classical linear double bond ($\sigma + \pi$) requires the reaction between the triplet states of diradical fragments, so the higher $E_{S\rightarrow T}$ of the fragments (185 kcal/mol in the present case) with respect to the energy gain related to their linear double bond (estimated to be 146 kcal/mol for propadienone), the more the fragments tend to keep their singlet nature in a resulting bent structure, that anyway gives rise to some extent of conjugation, and thus represents the best compromise from an energetic point of view. The same considerations can be applied to propadienethione, for which the reduced $E_{S\rightarrow T}$ of the $:C=S$ fragment (79 kcal/mol), together with a quite similar energy gain for the linear double bond (144 kcal/mol), makes the linear C_{2v} structure the preferred one.

In the light of the above model, that works very well also in predicting the C_{2v} symmetry of shorter cumulenes and cumulene-thiones, it is clear that the substitution of the oxygen atom with a sulfur one causes the energy difference between singlet and triplet states of the involved diradical fragment to decrease sufficiently to stabilize the linear structure instead of the kinked one, thus justifying also the importance of electron correlation for a reliable *ab-initio* prediction of structural deformations.

As a concluding remark we would like to stress the fact that this example confirms the reasonableness of thinking at molecular systems as composed by fragments, or structural synthons, that tend to preserve their individualities and main structural features. This consideration further justifies the applicability and the successful results of our previously introduced TMA and LRA,^{54,55,61} both exploiting such common characteristics of the same moieties in different molecules.

4 Linear Regression Approach

As previously noted, B3LYP/SNSD and B2PLYP/VTZ predictions of $r(C-S)$ and $r(N-S)$ bond lengths suffer from large inaccuracy with respect to the reference SE parameters.

Following the same approach (LRA) adopted in a previous work⁵⁵, we looked for a linear relation between theoretical calculated parameters and SE ones by performing a linear regression of B2PLYP/VTZ r_e^{SE} r(C-S) bond lengths versus the corresponding r_e predictions (see Figure 3). In this way, defining B as the intercept for the linear regression and A as (slope-1), we can find out an empirical correction to improve the accuracy of r_e parameters:

$$r_e^{LRA} = r_e + \Delta r, \quad \text{with } \Delta r = A \cdot r_e + B \quad (13)$$

Even if the number of items is not very large, Figure 3 and Table 4 suggest the good quality of the linear regression for both B3LYP/SNSD and B2PLYP/VTZ, with an R^2 of 0.9991 and 0.9996 respectively. Further, it is remarkable that for the empirically corrected parameters (r_e^{LRA}) the MAE is reduced to 0.0018 Å for B3LYP/SNSD and to 0.0011 Å for B2PLYP/VTZ, with a root-mean-square error (RMSE) of 0.0024 Å and 0.0015 Å, respectively. In order to check the reliability of LRA empirical correction for r(C-S), we can exploit 2-thiouracil and thiophene as test cases, both of them being already present in our database. The best equilibrium value for r(C₂-S) in 2-thiouracil is 1.65020(3) Å⁹⁰, that was obtained by employing highly accurate coupled cluster *ab-initio* parameters in a partial SE equilibrium structure optimization; B3LYP/SNSD and B2PLYP/VTZ predict bond lengths of 1.6693 Å and 1.6567 Å, respectively⁵⁵. Applying eq. 13 to these values, we obtain a LRA empirical estimation for r(C₂-S) in 2-thiouracil of 1.6464 Å and 1.6477 Å for B3LYP/SNSD and B2PLYP/VTZ, thus reducing the absolute error to 0.0038 Å and 0.0025 Å, respectively. Concerning thiophene, our DFT calculations predict the r(C-S) distance to be quite longer than the best SE value of 1.7126 Å, with errors of 0.0278 Å (B3LYP/SNSD) and 0.0088 Å (B2PLYP/VTZ)^{54,55}. Again eq. 13 significantly improves the results, reducing the absolute error to 0.0015 Å for B3LYP/SNSD ($r_e^{LRA} = 1.7111$ Å) and to 0.0026 Å in the case of B2PLYP/VTZ ($r_e^{LRA} = 1.7100$ Å). These two examples clarify how LRA empirical corrections allow one to significantly improve the DFT theoretical predictions for r(C-S) distances, and thus can be useful in order to have a trustworthy guess of the equilibrium structure employing only non expensive computational calculations. Looking at interpolating straight lines in Figure 3, it is clear that the overestimation error of B3LYP/SNSD rapidly increases with the lengthening of SE equilibrium value of r(C-S), whereas B2PLYP/VTZ suffers of a small quasi-systematic error, the slope of the linear regression line being 0.96.

In the light of the large number of new SE equilibrium carbon-hydrogen bond lengths determined in this work, we report in Table 4 also the results for the linear regression of all B2PLYP/VTZ r_e^{SE} r(C-H) bond lengths versus both B3LYP/SNSD and B2PLYP/VTZ r_e predictions present in our online database. The very small changes in A and B parameters with respect to those previously found⁵⁵ are very satisfactory and, together with low MAE and RMSE of the LRA predictions, confirm the reliability of LRA approach for obtaining accurate estimates of r(C-H) bond lengths comparable with SE ones. It is worth to underline that, in the case of B3LYP/SNSD, r corrections reduce the MAE of 0.0062 Å and the

RMSE of 0.0064 Å to 0.0008 Å and 0.0011 Å, respectively, thus significantly improving the accuracy of the final results.

Another interesting remark concerns the fact that in all the four cases reported in Table 4 the error distribution of LRA parameters is well centred on zero, confirming the good accuracy of the LRA, while the small values of the standard deviation (SD), together with the significant size of the samples analysed, allows the definition of a quite narrow confidence interval for searching the optimal SE equilibrium value, for instance through the so called predicate method¹⁴, starting from a not too expensive QM calculation.

5 Conclusions

Beyond the significant enlargement of our online datasets of molecular geometries, this work vouches for the achieved ripeness of our computational protocol in dealing with the calculation of accurate SE equilibrium molecular structures of semi-rigid organic molecules containing the most common atoms and moieties of interest in astrochemistry as well as in life- and material-sciences. In fact, we have improved and further tested the accuracy and robustness of our fitting procedure by means of a new software able to perform non-linear constrained optimizations, which will be fully described in a forthcoming paper, once extensive testing and implementation of some additional features will be accomplished. The 68 molecules collected till now represent a valuable source of benchmark data to be used in parametrization and validation studies, as well as input information in many kind of analyses and computations. It is remarkable that the SE technique coupled with DFT anharmonic calculations allows one to obtain, with a very high level of accuracy, the equilibrium structure of molecules containing a third row atom like sulfur, which considerably increases the computational cost of high level post-Hartree-Fock methods because of their unfavourable scaling with the number of electrons.

Next these results can be exploited to find the best compromise between accuracy and computational cost in deriving equilibrium structure of molecular systems. In particular, in the context of the linear regression approach applied to carbon-hydrogen and carbon-sulfur bond lengths, it has been possible to define trustworthy empirical corrections, that permit one to make use of cost-effective, fully quantum mechanical, B3LYP/SNSD and B2PLYP/VTZ results to reach highly accurate estimates of equilibrium values. In this context, the large number of accurate SE parameters permits a profitable use of statistical estimates in order to extract reliable information about molecules' structural features and their evaluations by different approaches.

The reliable and robust performances offered by our protocol in dealing with medium-sized semi-rigid molecules, together with the dimension of the gathered dataset and the optimized tools developed till now, make us confident that it will be possible to obtain satisfactory results also for larger systems, which will be the focus of our future efforts because of their great interest in many areas of current fundamental and applied research.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgement

The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement n. [320951].

References

- (1). Whitesides G. *Angew Chem Int Ed*. 2015; 54:3196.
- (2). See for instance the Royal Society of Chemistry website. last update in December 2011 [http://www.rsc.org/images/H-index%20ranking%20of%20living%20chemists\(December%202011\).tcm18-211414.pdf](http://www.rsc.org/images/H-index%20ranking%20of%20living%20chemists(December%202011).tcm18-211414.pdf)
- (3). Mandal S, Moudgil M, Mandal SK. *Eur J Pharmacol*. 2009; 625:90. [PubMed: 19835861]
- (4). Verlinde CL, Hol WG. *Structure*. 1994; 2:577–587. [PubMed: 7922037]
- (5). Truhlar, D.Howe, W.Hopfinger, A., Dammkoehler, R., editors. *Rational Drug Design*. Springer-Verlag; New York, NY: 1999.
- (6). Shaw, AM. *Astrochemistry: From Astronomy to Astrobiology*. John Wiley & Sons, Inc.; New York, NY: 2006.
- (7). Herbst E. *Chem Soc Rev*. 2001; 30:168–176.
- (8). Puzzarini C. *Rend Fis Acc Lincei*. 22:165–172.
- (9). Winnewisser, G., Herbst, E., Ungerechts, H. *Spectroscopy of the earth's atmosphere and interstellar medium*. Rao, KN., Weber, A., editors. Academic Press; San Diego: 1992. p. 423-538. *Molecular Spectroscopy: Modern Research*
- (10). Tielens AGGM. *Rev Mod Phys*. 2013; 85:1021.
- (11). Barone, V., editor. *Computational strategies for spectroscopy: from small molecules to nano systems*. Wiley & Sons, Inc.; Hoboken, NJ: 2011.
- (12). Puzzarini C, Stanton JF, Gauss J. *Int Rev Phys Chem*. 2010; 29:273.
- (13). Domenicano, A., Hargittai, I., editors. *Accurate molecular structures. Their determination and importance*. Oxford University Press; New York, NY: 1992.
- (14). Demaison, J.Boggs, J., Császár, A., editors. *Equilibrium molecular structures: from spectroscopy to quantum chemistry*. CRC Press; Boca Raton, FL: 2011.
- (15). Grunenberg, J., editor. *Computational spectroscopy: methods, experiments and applications*. Wiley-VCH Verlag GmbH & Co. KGaA; Weinheim, DE: 2010.
- (16). Quack, M., Merkt, F., editors. *Handbook of high-resolution spectroscopy*. John Wiley & Sons, Inc; Weinheim, DE: 2011. p. 2182
- (17). Puzzarini, C., Biczysko, M. *Structure elucidation in organic chemistry*. Cid, M-M., editor. Wiley-VCH Verlag GmbH & Co. KGaA; Weinheim, DE: 2015. p. 27-64.
- (18). Bak KL, Gauss J, Jørgensen P, Olsen J, Helgaker T, Stanton JF. *J Chem Phys*. 2001; 114:6548.
- (19). Demaison J. *Mol Phys*. 2007; 105:3109.
- (20). Pérez C, Muckle MT, Zaleski DP, Seifert NA, Temelso B, Shields GC, Kisiel Z, Pate BH. *Science*. 2012; 336:897–901. [PubMed: 22605772]
- (21). Caminati W. *Angew Chem Int Ed*. 2009; 48:9030–9033.
- (22). Pietraperzia G, Pasquini M, Schiccheri N, Piani G, Becucci M, Castellucci E, Biczysko M, Bloino J, Barone V. *J Phys Chem A*. 2009; 113:14343–14351. [PubMed: 19813719]
- (23). Blanco S, Lesarri A, López JC, Alonso JL. *J Am Chem Soc*. 2004; 126:11675–11683. [PubMed: 15366915]
- (24). Lovas FJ, McMahon RJ, Grabow J-U, Schnell M, Mack J, Scott LT, Kuczkowski RL. *J Am Chem Soc*. 2005; 127:4345–4349. [PubMed: 15783216]
- (25). Peña I, Daly AM, Cabezas C, Mata S, Berúndez C, Niño A, López JC, Grabow J-U, Alonso JL. *J Phys Chem Lett*. 2013; 4:65–69. [PubMed: 26291213]
- (26). Puzzarini C, Biczysko MVB, Largo L, Peóla I, Cabezas C, Alonso JL. *J Phys Chem Lett*. 2014; 5:534–540. [PubMed: 26276605]
- (27). Grimme S, Steinmetz M. *Phys Chem Chem Phys*. 2013; 15:16031–16042. [PubMed: 23963317]

- (28). Jurecka P, Sponer J, Cerny J, Hobza P. *Phys Chem Chem Phys*. 2006; 8:1985–1993. [PubMed: 16633685]
- (29). Zhao Y, Truhlar DG. *J Chem Theory Comput*. 2007; 3:289–300. [PubMed: 26627172]
- (30). Barone V, Biczysko M, Pavone M. *Chem Phys*. 2008; 346:247–256.
- (31). Barone V, Biczysko M, Bloino J, Puzzarini C. *J Chem Phys*. 2014; 141:034107. [PubMed: 25053301]
- (32). Brémond É, Savarese M, Su NQ, Pérez-Jiménez ÁJ, Xu X, Sancho-García JC, Adamo C. *J Chem Theory Comput*. 2016; 12:459–465. [PubMed: 26730741]
- (33). Brémond É, Savares M, Sancho-García JC, Pérez-Jiménez ÁJ, Adamo C. *J Chem Phys*. 2016; 144:124104. [PubMed: 27036424]
- (34). Senn HM, Thiel W. *Angew Chem Int Ed*. 2009; 48:1198–1229.
- (35). Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, et al. *J Comput Chem*. 2009; 30:1545–1614. [PubMed: 19444816]
- (36). Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B, et al. *Bioinformatics*. 2013; 29:845–854. [PubMed: 23407358]
- (37). Grubisic S, Brancato G, Pedone A, Barone V. *Phys Chem Chem Phys*. 2012; 14:15308–15320. [PubMed: 23051698]
- (38). Maple JR, Dinur U, Hagler AT. *Proc Natl Acad Sci USA*. 1988; 85:5350–5354. [PubMed: 16593959]
- (39). Dasgupta S, Yamasaki T, Goddard WA. *J Chem Phys*. 1996; 104:2898–2920.
- (40). Biczysko M, Bloino J, Brancato G, Cacelli I, Cappelli C, Ferretti A, Lami A, Monti S, Pedone A, Prampolini G, Puzzarini C, et al. *Theor Chem Acc*. 2012; 131:1201.
- (41). Barone V, Cacelli I, De Mitri N, Licari D, Monti S, Prampolini G. *Phys Chem Chem Phys*. 2013; 15:3736–3751. [PubMed: 23389748]
- (42). Barone V, Biczysko M, Puzzarini C. *Acc Chem Res*. 2015; 48:1413–1422. [PubMed: 25894724]
- (43). Császár, A. *Equilibrium molecular structures: from spectroscopy to quantum chemistry*. Demaison, J, Boggs, J., Császár, A., editors. CRC Press; Boca Raton, FL: 2011. p. 233–261.
- (44). Kuchitsu, K. *Accurate molecular structures. Their determination and importance*. Domenicano, A., Hargittai, I., editors. Oxford University Press; New York, NY: 1992. p. 14
- (45). Raghavachari K, Trucks GW, Pople JA, Head-Gordon M. *Chem Phys Lett*. 1989; 157:479.
- (46). Heckert M, Kállay M, Gauss J. *Mol Phys*. 2005; 103:2109.
- (47). Heckert M, Kállay M, Tew DP, Klopper W, Gauss J. *J Chem Phys*. 2006; 125:044108.
- (48). Barone V, Biczysko M, Bloino J, Puzzarini C. *Phys Chem Chem Phys*. 2013; 15:10094–10111. [PubMed: 23599122]
- (49). Pulay P, Meyer W, Boggs JE. *J Chem Phys*. 1978; 68:5077.
- (50). Vázquez, J., Stanton, JF. *Equilibrium molecular structures: from spectroscopy to quantum chemistry*. Demaison, J, Boggs, J., Császár, A., editors. CRC Press; Boca Raton, FL: 2011. p. 53–87.
- (51). Pawłowski F, Jørgensen P, Olsen J, Hegelund F, Helgaker T, Gauss J, Bak K, Stanton J. *J Chem Phys*. 2002; 116:6482.
- (52). Puzzarini C, Barone V. *Phys Chem Chem Phys*. 2011; 13:7189. [PubMed: 21409277]
- (53). Puzzarini C, Penocchio E, Biczysko M, Barone V. *J Phys Chem A*. 2014; 118:6648–6656. [PubMed: 24842714]
- (54). Piccardo M, Penocchio E, Puzzarini C, Biczysko M, Barone V. *J Phys Chem A*. 2015; 119:2058–2082. [PubMed: 25648634]
- (55). Penocchio E, Piccardo M, Barone V. *J Chem Theory Comput*. 2015; 11:4689–4707. [PubMed: 26574259]
- (56). Lee C, Yang W, Parr RG. *Phys Rev B*. 1988; 37:785.
- (57). Becke AD. *J Chem Phys*. 1993; 98:5648–5652.
- (58). Grimme S. *J Chem Phys*. 2006; 124:034108. [PubMed: 16438568]
- (59). Neese F, Schwabe T, Grimme S. *J Chem Phys*. 2007; 126:124115. [PubMed: 17411116]

- (60). Biczysko M, Panek P, Scalmani G, Bloino J, Barone V. *J Chem Theory Comput.* 2010; 6:2115–2125. [PubMed: 26615939]
- (61). Barone V, Biczysko M, Bloino J, Penocchio E, Puzzarini C. *J Chem Theory Comput.* 11:4342.
- (62). [accessed February 2015] The CCse, B3se and B2se sets are available in Download section. <http://dreamslab.sns.it>
- (63). Patel OG, Mberu EK, Nzila AM, Macreadie IG. *Trends Paracitol.* 2004; 20:1–3.
- (64). Padmavathi V, Reddy GS, Padmaja A, Kondaiah P, Ali-Shazia. *Eur J Med Chem.* 2009; 44:2106–2112. [PubMed: 19036476]
- (65). Bhadra A, Ponnampuruma C. *Orig Life Evol Biosph.* 1986; 16:286–286.
- (66). Millar T, Herbst E. *Astron Astrophys.* 1990; 231:466.
- (67). Balucani N, Leonori F, Petrucci R, Casavecchia P, Skouteris D, Rosi M. *Mem S A It Suppl.* 2011; 16:91.
- (68). Papoušek, D., Aliev, M. *Molecular vibrational-rotational spectra.* Elsevier; New York, NY: 1982.
- (69). Gauss J, Ruud K, Helgaker T. *J Chem Phys.* 1996; 105:2804–2812.
- (70). Flygare W. *Chem Rev.* 1976; 74:653–687.
- (71). Born M, Oppenheimer R. *Ann Phys (Berlin, Ger).* 1927; 389:457–484.
- (72). Sayvetz A. *J Chem Phys.* 1939; 7:383–389.
- (73). Eckart C. *Phys Rev.* 1935; 47:552–558.
- (74). Jensen, P., Bunker, P. *Computational molecular spectroscopy.* Jensen, P., Bunker, P., editors. Wiley; New York, NY: 2000. p. 3-12.
- (75). Nielsen HH. *Rev Mod Phys.* 1951; 23:90–136.
- (76). Watson JKG. *Mol Phys.* 1968; 15:479–490.
- (77). Aliev, M., Watson, JKG. *Molecular spectroscopy: modern research.* Rao, KN., editor. Academic Press; New York, NY: 1985. p. 1-67.
- (78). Mills, IM. *Molecular spectroscopy: modern research.* K.N. Rao and C.W. Mathews; New York, NY: 1973. p. 115-140.
- (79). Stephens PJ, Devlin FJ, Chabalowski CF, Frisch MJ. *J Phys Chem.* 1994; 98:11623.
- (80). Dunning TH Jr. *J Chem Phys.* 1989; 90:1007–1023.
- (81). Woon DE, Dunning TH Jr. *J Chem Phys.* 1993; 98:1358–1371.
- (82). Woon DE, Dunning TH Jr. *J Chem Phys.* 1995; 103:4572–4585.
- (83). Peterson KA, Dunning TH Jr. *J Chem Phys.* 2002; 117:10548–10560.
- (84). Carnimeo I, Puzzarini C, Tasinato N, Stoppa P, Pietropolli-Charmet A, Biczysko M, Cappelli C, Barone V. *J Chem Phys.* 2013; 139:074310. [PubMed: 23968095]
- (85). Barone V, Biczysko M, Bloino J. *Phys Chem Chem Phys.* 2014; 16:1759–1787. [PubMed: 24346191]
- (86). The SNSD basis set is available in Download section of our website (accessed January 2015). <http://dreamslab.sns.it/sets/SNSD.gbs>
- (87). Puzzarini C, Biczysko M, Barone V. *J Chem Theory Comput.* 2010; 6:828. [PubMed: 26613310]
- (88). Barone V, Bloino J, Biczysko M. *Phys Chem Chem Phys.* 2010; 12:1092. [PubMed: 20094674]
- (89). Barone V, Biczysko M, Bloino J, Puzzarini C. *J Chem Theory Comput.* 2013; 9:1533. [PubMed: 26587615]
- (90). Puzzarini C, Biczysko M, Barone V, Peña I, Cabezas C, Alonso J. *Phys Chem Chem Phys.* 2013; 15:16965. [PubMed: 24002739]
- (91). Kendall RA, Dunning TH Jr, Harrison RJ. *J Chem Phys.* 1992; 96:6796.
- (92). Frisch, MJ., Trucks, GW., Schlegel, HB., Scuseria, GE., Robb, MA., Cheeseman, JR., Scalmani, G., Barone, V., Mennucci, B., Petersson, GA., Nakatsuji, H., et al. *Gaussian 09 Revision D.01.* Gaussian Inc; Wallingford CT: 2009.
- (93). Schneider W, Thiel W. *Chem Phys Lett.* 1989; 157:367.
- (94). Stanton JF, Gauss J. *Int Rev Phys Chem.* 2000; 19:61–95.
- (95). Thiel W, Scuseria G, Schaefer HFS, Allen WD. *J Chem Phys.* 1988; 89:4965–4975.
- (96). Barone V. *J Chem Phys.* 2005; 122:014108.

- (97). Barone V. *J Chem Phys.* 1994; 101:10666–10676.
- (98). Stanton JF, Lopreore CL, Gauss J. *J Chem Phys.* 1998; 108:7190–7196.
- (99). Stanton J, Gauss J, Harding M, Szalay P, Auer A, Bartlett R, Benedikt U, Berger C, Bernholdt D, Bomble Y, Cheng L, et al. CFOUR, Coupled-Cluster techniques for Computational Chemistry, a quantum-chemical program package. 2010<http://www.cfour.de/>
- (100). Bertsekas, DP. *Nonlinear Programming*. 2nd ed. Athena Scientific; Belmont, MA: 1999.
- (101). Demaison, J. *Equilibrium molecular structures: from spectroscopy to quantum chemistry*. Demaison, J, Boggs, J., Császár, A., editors. CRC Press; Boca Raton, FL: 2011. p. 29-52.
- (102). Belsley, DA., Kuh, E., Welsch, RE. *Regression Diagnostics*. 1st ed. John Wiley & Sons, Inc; Hoboken, NJ: 1980.
- (103). Y Z, Truhlar D. *TCA.* 2008; 120:215.
- (104). D BA. *J Chem Phys.* 1993; 98:1372.
- (105). Rudolph HD. *Struct Chem.* 1990; 2:581.
- (106). Rudolph, HD., Demaison, J. *Equilibrium molecular structures: from spectroscopy to quantum chemistry*. Demaison, J, Boggs, J., Császár, A., editors. CRC Press; Boca Raton, FL: 2011. p. 125-158.
- (107). Kashyap S, Garg V, Sharma P, Kumar N, Dudhe R, Gupta J. *Med Chem Res.* 2012; 21:2123–2132.
- (108). Li Y, Geng J, Liu Y, Yu S, Zhao G. *Chem Med Chem.* 2013; 8:27. [PubMed: 23208773]
- (109). Jain A, Sharma S, A V, Ravichandran V, Agrawal R. *Chem Biol Drug Des.* 2013; 81:557. [PubMed: 23452185]
- (110). Glossman M. *J Mol Struct(Theochem).* 1995; 330:385.
- (111). Nalwa, HS., editor. *Handbook of Organic Conductive Molecules and Polymers*. John Wiley & Sons, Inc; Hoboken, NJ: 1997.
- (112). Mullen, K., Wegner, G. *Electronic Materials: The Oligomer Approach*. John Wiley & Sons, Inc; Hoboken, NJ: 2008.
- (113). Glossman-Mitnik D, Márquez-Lucero A. *J Mol Struct(Theochem).* 2001; 548:153.
- (114). Saxena PP, Misra A. *Mon Not R Astron Soc.* 1995; 272:89–95.
- (115). Woon DE. *J Phys Chem A.* 2007; 111:11249–11253. [PubMed: 17536790]
- (116). Cole CA, Demarais NJ, Yang Z, Snow TP, Bierbaum VM. *AstroPhys J.* 2013; 779:181.
- (117). Balucani N. *Int J Mol Sci.* 2009; 10:2304. [PubMed: 19564951]
- (118). Evans CJ, Carter JP, Appadoo DR, Wong A, McNaughton D. *J Mol Spectrosc.* 2015; 316:32–37.
- (119). Hirao T, Okabayashi T, Tanimoto M. *J Mol Spectrosc.* 2001; 208:148–149. [PubMed: 11437564]
- (120). Allen WD, Bertie JE, Falk MV, Hess BA, Mast GB, Othen DA, Schaad LJ, Schaefer HF. *J Chem Phys.* 1986; 84:4211–4227.
- (121). Cunningham GL, Boyd AW, Myers RJ, Gwinn WD, Le Van WI. *J Chem Phys.* 1951; 19:676–685.
- (122). Nygaard L, Asmussen E, Høg JH, Maheshwari R, Nielsen C, Petersen IB, Rastrup-Andersen J, Sørensen G. *J Mol Struct.* 1971; 8:225–233.
- (123). Stiefvater OL. *Chem Phys.* 1976; 13:73–80.
- (124). Stiefvater OL. *Z Naturforsch.* 1989; 44a:29.
- (125). Stiefvater OL. *Z Naturforsch.* 1976; 31a:1681.
- (126). Stiefvater OL. *Z Naturforsch.* 1978; 33a:1511.
- (127). Ruiterkamp R, Charnley SB, Butner HM, Huang H-C, Rodgers SD, Kuan Y-J, Ehrenfreund P. *Astrophys Space Sci.* 2007; 310:181–188.
- (128). Leonori F, Petrucci R, Balucani N, Casavecchia P, Rosi M, Berteloite C, Le Picard SD, Canosa A, Sims IR. *Phys Chem Chem Phys.* 2009; 11:4701. [PubMed: 19492122]
- (129). Ehrenfreund P, Charnley SB. *ESA SP-496.* 2001:35–42.

- (130). Yamada K, Winnewisser M, Winnewisser G, Szalanski L, Gerry M. *J Mol Spectrosc.* 1980; 79:295–313.
- (131). Yamada K. *J Mol Spectrosc.* 1980; 79:323–344.
- (132). Johnson DR, Powell FX, Kirchhoff WH. *J Mol Spectrosc.* 1971; 39:136–145.
- (133). Bak B, Nielsen OJ, Svanholt H, Holm A, Toubro NH, Krantz A, Laureni J. *Acta Chem Scand.* 1979; 33a:161.
- (134). East ALL, Allen WD, Klippenstein SJ. *J Chem Phys.* 1995; 102:8506–8532.
- (135). Brown RD, Dyall KG, Elmes PS, Godfrey PD, McNaughton D. *J Am Chem Soc.* 1988; 110:789–792.
- (136). Brown RD, Godfrey PD, Champion R, McNaughton D. *J Am Chem Soc.* 1981; 103:5711–5715.
- (137). Brown RD, Champion R, Elmes PS, Godfrey PD. *J Am Chem Soc.* 1985; 107:4109–4112.
- (138). Brown R. *J Mol Struct.* 1983; 97:293–302.
- (139). Brown R, Godfrey PD, Champion R. *J Mol Spectrosc.* 1987; 123:93.
- (140). Farnell L, Radom L. *Chem Phys Lett.* 1982; 91:373–377.
- (141). Brown RD, Dittman RG. *Chem Phys.* 1984; 83:77–82.
- (142). Taylor PR. *J Comput Chem.* 1984; 5:589–597.
- (143). East ALL. *J Chem Phys.* 1998; 108:3574–3584.
- (144). Scott A, Radom L. *J Mol Struct.* 2000; 556:253–261.
- (145). Trinquier G, Malrieu JP. *J Am Chem Soc.* 1987; 109:5303–5315.
- (146). Davidson, ER. *Diradicals*. Borden, WT., editor. Wiley-Interscience; New York, NY: 1982. p. 73
- (147). Cheung LM, Sundberg KR, Ruedenberg K. *J Am Chem Soc.* 1978; 100:8024–8025.
- (148). Bernardi F, Robb MA. *Mol Phys.* 1983; 48:1345–1355.
- (149). Carter EA, G WA III. *J Phys Chem.* 1986; 90:998–1001.

Figure 1.
Graphical representation of the molecules in Table 1.

Figure 2.

Graphical representation of the molecules in Tables 2 and 3: (a) and (a') are isothiocyanic and isocyanic acids; (b) and (b') are thioformaldehyde and formaldehyde; (c) and (c') are thioketene and ketene; (d) and (d') are propadienethione and propadienone.

Figure 3.
Linear fits of SE $r(\text{CS})$ bond lengths with respect to their B3LYP/SNSD (left side) and B2PLYP/VTZ (right side) *ab-initio* predictions.

Table 1

B3LYP/SNSD (r_e , r_e^{SE}) and B2PLYP/VTZ (r_e , r_e^{SE}) and r_0 geometries of the five members eterocycles. Distances in Å, angles in degrees. Atom numbering as in Figure 1^a.

	r_e		r_e^{SE}		r_0
	B3LYP/SNSD	B2PLYP/VTZ	B3LYP/SNSD	B2PLYP/VTZ	exp.
– Sulfur eterocycles –					
thiirane					
r(S-C)	1.8476	1.8265	1.8115(1)	1.8111(1)	1.8187(4)
r(C-H)	1.0857	1.0799	1.0795(1)	1.0798(1)	1.0788(11)
a(C-S-C)	47.22	47.74	48.27(1)	48.25(1)	48.40(4)
a(S-C-H)	114.59	114.91	114.95(1)	114.98(1)	114.94(11)
a(H-C-H)	115.20	115.28	115.74(1)	115.66(1)	115.82(17)
s [u·Å ²]	-	-	0.0003	0.0001	0.0091
κ	-	-	11	11	11
thiazole					
r(S1-C2)	1.7527	1.7332	1.7211(17)	1.7213(14)	1.7275(61)
r(S1-C5)	1.7385	1.7200	1.7091(13)	1.7089(10)	1.7152(48)
r(C2-N3)	1.3002	1.3019	1.2995(25)	1.2986(20)	1.3035(90)
r(C4-C5)	1.3652	1.3643	1.3642(22)	1.3645(18)	1.3673(86)
r(C2-H)	1.0837	1.0779	1.0795(28)	1.0801(23)	1.0820(98)
r(C4-H)	1.0839	1.0781	1.0773(10)	1.0774(8)	1.0783(38)
r(C5-H)	1.0805	1.0751	1.0755(14)	1.0753(12)	1.0774(55)
a(C5-S1-C2)	88.48	88.97	89.25(6)	89.24(5)	89.31(23)
a(S1-C2-N3)	115.08	115.11	115.33(14)	115.33(6)	115.25(29)
a(C4-C5-S1)	109.73	109.66	109.66(7)	109.66(6)	109.65(28)
a(S1-C2-H)	120.56	120.75	120.42(77)	120.27(63)	119.51(241)
a(C5-C4-H)	124.88	124.78	125.05(14)	125.04(12)	124.94(54)
a(S1-C5-H)	121.32	121.64	121.79(25)	121.83(21)	121.05(96)
s [u·Å ²]	-	-	0.0007	0.0006	0.0028
Mean [u·Å ²]	-	-	0.00141	0.00088	0.07377
κ	-	-	577	570	542
1,2,3-thiadiazole					
r(S-N)	1.7722	1.7214	1.6856(6)	1.6845(4)	1.6927(23)
r(N-N)	1.2678	1.2825	1.2850(9)	1.2855(7)	1.2918(38)
r(N-C)	1.3729	1.3628	1.3646(7)	1.3645(5)	1.3721(29)
r(C-S)	1.7094	1.6928	1.6842(6)	1.6844(4)	1.6905(24)
r(C4-H)	1.0822	1.0768	1.0761(5)	1.0764(4)	1.0768(20)
r(C5-H)	1.0809	1.0754	1.0741(6)	1.0753(5)	1.0789(26)
a(C-S-N)	90.87	92.31	93.04(3)	93.05(2)	92.98(12)
a(N-N-C)	115.69	114.74	113.95(5)	113.90(4)	113.83(22)
a(S-N-N)	110.38	110.63	111.28(4)	111.29(3)	111.24(15)

	r_e		r_e^{SEa}		r_0
	B3LYP/SNSD	B2PLYP/VTZ	B3LYP/SNSD	B2PLYP/VTZ	exp.
a(N-C-H)	119.22	119.32	119.21(5)	119.15(4)	119.21(19)
a(S-C-H)	123.14	123.81	124.22(13)	124.01(9)	122.77(50)
s [$u \cdot \text{\AA}^2$]	-	-	0.0006	0.0005	0.0013
Mean [$u \cdot \text{\AA}^2$]	-	-	0.00273	0.00252	0.08892
κ	-	-	510	510	509
1,3,4-thiadiazole					
r(S-C)	1.7470	1.7282	1.7157(1)	1.7161(2)	1.7218(10)
r(C-N)	1.2995	1.3009	1.2997(2)	1.2987(3)	1.3033(13)
r(C-H)	1.0829	1.0770	1.0774(1)	1.0782(2)	1.0763(11)
a(C-S-C)	85.61	86.05	86.34(1)	86.31(1)	86.44(6)
a(S-C-N)	114.49	114.59	114.75(1)	114.74(1)	114.69(6)
a(S-C-H)	122.18	122.32	122.52(3)	122.25(5)	122.24(26)
s [$u \cdot \text{\AA}^2$]	-	-	0.0002	0.0004	0.0020
Mean [$u \cdot \text{\AA}^2$]	-	-	0.00303	0.00094	0.08556
κ	-	-	400	398	398
1,2,4-thiadiazole					
r(S1-N2)	1.6935	1.6627	1.6448(3)	1.6449(4)	1.6490(23)
r(N2-C3)	1.3129	1.3157	1.3144(5)	1.3146(8)	1.3196(42)
r(N4-C5)	1.3061	1.3090	1.3089(5)	1.3088(8)	1.3134(43)
r(C5-S1)	1.7354	1.7169	1.7034(4)	1.7031(6)	1.7077(32)
r(C3-H)	1.0854	1.0791	1.0829(27)	1.0784(45)	1.1046(243)
r(C5-H)	1.0842	1.0783	1.0785(3)	1.0782(5)	1.0766(24)
a(C5-S1-N2)	91.50	92.36	92.74(2)	92.74(2)	92.80(13)
a(S1-N2-C3)	106.87	106.92	107.11(1)	107.12(2)	107.08(13)
a(N4-C5-S1)	112.60	112.27	112.38(1)	112.37(2)	112.39(12)
a(N2-C3-H)	119.56	119.68	119.51(24)	119.54(40)	118.47(211)
a(S1-C5-H)	123.17	123.66	123.76(11)	123.90(19)	124.39(95)
s [$u \cdot \text{\AA}^2$]	-	-	0.0002	0.0003	0.0018
Mean [$u \cdot \text{\AA}^2$]	-	-	0.00245	0.00028	0.08454
κ	-	-	672	671	664
1,2,5-thiadiazole					
r(S-N)	1.6761	1.6447	1.6258(2)	1.6254(1)	1.6312(10)
r(N-C)	1.3179	1.3246	1.3238(2)	1.3237(2)	1.3280(15)
r(C-H)	1.0853	1.0791	1.0784(1)	1.0785(1)	1.0779(8)
a(N-S-N)	97.72	98.98	99.45(1)	99.44(1)	99.46(8)
a(S-N-C)	106.59	106.33	106.55(1)	106.56(1)	106.58(7)
a(N-C-H)	119.74	119.84	120.03(2)	120.06(2)	120.11(10)
s [$u \cdot \text{\AA}^2$]	-	-	0.0003	0.0003	0.0017
Mean [$u \cdot \text{\AA}^2$]	-	-	0.00199	0.00020	0.08284
κ	-	-	312	312	312

^{a)} All fits have been performed on moments of inertia. 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. $\Delta I^c = I^c - I^b - I^a$ is the inertial defect, s is the residual standard deviation (see Eq. 8), while κ is the condition number of the fit (see Eq. 12).

Table 2

B3LYP/SNSD (r_e , r_e^{SE}) and B2PLYP/VTZ (r_e , r_e^{SE}) and r_0 geometries of the sulfur containing molecules with heavy atoms chains. Distances in Å, angles in degrees. Atom numbering is reported in Figure 2.^a

	r_e		r_e^{SE}		r_0
	B3LYP/SNSD	B2PLYP/VTZ	B3LYP/SNSD	B2PLYP/VTZ	exp.
– Sulfur chains –					
isothiocyanic acid					
r(S-C)	1.5832	1.5744	1.5678(6)	1.5678(7)	1.5719(29)
r(C-N)	1.2091	1.2047	1.2049(9)	1.2047(9)	1.2074(38)
r(N-H)	1.0102	1.0035	1.0065(9)	1.0070(9)	1.0221(44)
a(S-C-N)	173.77	173.95	172.18(18)	172.23(18)	170.91(68)
a(H-N-C)	129.67	131.50	129.72(5)	129.75(5)	128.03(17)
s [u·Å ²]	-	-	0.0006	0.0006	0.0028
κ	-	-	270	270	270
thioformaldehyde					
r(S-C)	1.6227	1.6135	1.6096(1)	1.6093(1)	1.6137(10)
r(C-H)	1.0918	1.0859	1.0854(3)	1.0854(2)	1.0976(33)
a(H-C-S)	121.95	122.01	121.70(2)	121.73(2)	121.11(23)
s [u·Å ²]	-	-	0.0008	0.0007	0.0100
κ	-	-	7	7	7
thioketene					
r(S-C)	1.5724	1.5623	1.5557(4)	1.5556(4)	1.5583(55)
r(C-C)	1.3107	1.3082	1.3109(5)	1.3107(5)	1.3115(71)
r(C-H)	1.0863	1.0801	1.0805(2)	1.0806(1)	1.0981(21)
a(C-C-H)	120.66	120.46	120.13(1)	120.15(1)	119.47(16)
s [u·Å ²]	-	-	0.0005	0.0004	0.0065
κ	-	-	245	245	244
propadienethione					
r(S1-C2)	1.5876	1.5776	1.5733(61)	1.5715(15)	1.5704(226)
r(C2-C3)	1.2722	1.2691	1.2687(75)	1.2702(18)	1.2753(278)
r(C3-C4)	1.3223	1.3191	1.3246(36)	1.3228(9)	1.3182(133)
r(C4-H)	1.0905	1.0848	1.0825(13)	1.0842(3)	1.1086(46)
a(H-C4-C3)	121.80	121.68	121.22(8)	121.20(2)	120.42(30)
s [u·Å ²]	-	-	0.0051	0.0012	0.0190
κ	-	-	735	734	733

^{a)} All fits have been performed on moments of inertia. 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. s is the residual standard deviation (see Eq. 8), while κ is the condition number of the fit (see Eq. 12).

Table 3

B3LYP/SNSD (r_e , r_e^{SE}) and B2PLYP/VTZ (r_e , r_e^{SE}) and r_0 geometries of the oxygen containing molecules with heavy atoms chains. Distances in Å, angles in degrees. Atom numbering is reported in Figure 2.^a

	r_e		r_e^{SEa}		r_0
	B3LYP/SNSD	B2PLYP/VTZ	B3LYP/SNSD	B2PLYP/VTZ	exp.
– Oxygen Chains –					
isocyanic acid					
r(O-C)	1.1687	1.1683	1.1644(4)	1.1644(4)	1.1680(20)
r(C-N)	1.2196	1.2147	1.2149(4)	1.2147(4)	1.2186(21)
r(N-H)	1.0101	1.0037	1.0055(4)	1.0059(4)	1.0223(23)
a(O-C-N)	172.55	172.43	171.53(11)	171.51(10)	169.64(45)
a(H-N-C)	123.44	124.32	122.99(3)	123.02(2)	121.12(9)
s [u·Å ²]	-	-	0.0003	0.0003	0.0015
κ	-	-	190	190	190
formaldehyde					
r(O-C)	1.2052	1.2039	1.2050(1)	1.2047(1)	1.2084(8)
r(C-H)	1.1099	1.1030	1.1005(1)	1.1008(1)	1.1133(15)
a(H-C-O)	121.84	122.05	121.61(1)	121.61(1)	121.06(12)
s [u·Å ²]	-	-	0.0006	0.0007	0.0083
κ	-	-	8	8	8
ketene					
r(O-C)	1.1651	1.1638	1.1607(9)	1.1606(7)	1.1626(62)
r(C-C)	1.3140	1.3110	1.3126(9)	1.3124(8)	1.3146(65)
r(C-H)	1.0824	1.0761	1.0763(2)	1.0764(2)	1.0910(21)
a(C-C-H)	119.57	119.38	119.03(2)	119.05(2)	118.28(17)
s [u·Å ²]	-	-	0.0008	0.0007	0.0058
κ	-	-	194	194	194
propadienone^b					
r(O1-C2)	1.1711	1.1702	1.1659	1.1662	1.1635
r(C2-C3)	1.2998	1.3000	1.3056	1.3047	1.3197
r(C3-C4)	1.3277	1.3259	1.3292	1.3299	1.3225
r(C4-Hc)	1.0919	1.0859	1.0906	1.0883	1.1066
r(C4-Ht)	1.0901	1.0840	1.0865	1.0874	1.0928
a(O1-C2-C3)	169.73	169.19	168.34	168.31	167.85
a(C2-C3-C4)	146.90	145.11	142.71	142.46	144.18
a(Hc-C4-C3)	122.80	122.83	122.00	122.42	120.52
a(Ht-C4-C3)	121.32	121.16	120.57	120.62	117.52
s [u·Å ²]	-	-	0.0016	0.0016	0.0051
κ	-	-	3924	3936	3549

^{a)} All fits have been performed on moments of inertia. 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. s is the residual standard deviation (see Eq. 8), while κ is the condition number of the fit (see Eq. 12).

^{b)} SE data for this molecule are reported only for qualitative interest (see text).

Table 4 Statistical parameters for CS and CH bond lengths.^a

	r_e			r_e	
	B3LYP	B2PLYP		B3LYP	B2PLYP
CS bonds (10 items)			CH bonds (100 items)		
MAE (QM) [Å]	0.0248	0.0097	MAE (QM) [Å]	0.0062	0.0008
RMSE (QM) [Å]	0.0261	0.0103	RMSE (QM) [Å]	0.0064	0.0012
<i>A</i>	-0.088318	-0.038806	<i>A</i>	-0.086416	-0.068466
<i>B</i>	0.124429	0.055297	<i>B</i>	0.087847	0.073948
R^2	0.9991	0.9996	R^2	0.9840	0.9868
MAE (LRA) [Å]	0.0018	0.0011	MAE (LRA) [Å]	0.0008	0.0007
RMSE (LRA) [Å]	0.0024	0.0015	RMSE (LRA) [Å]	0.0011	0.0010
ME (LRA) [Å]	0.0000	0.0000	ME (LRA) [Å]	0.0000	0.0000
SD (LRA) [Å]	0.0025	0.0016	SD (LRA) [Å]	0.0011	0.0010

^a) B3LYP and B2PLYP in conjunction respectively with SNSD and VTZ basis sets. MAE and RMSE stand for mean absolute error and root-mean-square error, ME and SD stand for mean and standard deviation of the $(r_e^{LRA} - r_e^{SE})$ differences. *A* and *B* are (slope - 1) and the intercept for the linear regression of the fit of B2PLYP/VTZ r_e^{SE} parameters, taken as references, versus the corresponding B3LYP/SNSD and B2PLYP/VTZ r_e values; R^2 is the coefficient of determination.