MEMÓRIAS
DA
ACADEMIA DAS CIÊNCIAS
DE
LISBOA

CLASSE DE CIÊNCIAS

TOMO XLVIII

Ab initio potentials: From CBS extrapolation to globalness to riddles in the chemistry of small carbon clusters

A.J.C. Varandas

ACADEMIA DAS CIÊNCIAS
DE LISBOA

LISBOA • 2022

Ab initio potentials: From CBS extrapolation to globalness to riddles in the chemistry of small carbon clusters

A.J.C. Varandas[2]

António J. C. Varandas obtained a degree in Chemical Engineering from University of Porto (1971), and a Ph.D. in Theoretical Chemistry from University of Sussex (1976). After graduation, he joined the Department of Chemistry of Universidade de Coimbra, where he is Full Professor since 1988. Distinguished Professor at the Department of Physics of Qufu Normal University (China), he is also Professor at the Department of Physics of Universidade Federal do Espírito Santo (Brazil). With an impact factor of h=56 (Google Scholar), he published more than 430 papers, co-authored the first monograph in Molecular Potential Energy Functions (Wiley 1984), and wrote 2 other books in portuguese. He received awards: Prize Artur Malheiros for Physics and Chemistry of Lisbon Academy of Sciences (1985), Prize Ferreira da Silva of Portuguese Chemical Society (2001), and Prize Stimulus to Excellence of Portuguese Ministry of Science, Innovation and High Studies (2004). He serves amongst others the editorial board of Journal of Theoretical and Computational Chemistry where recurrently acts as Editor. He received honorific recognitions abroad, the most recent in China where he has been awarded in 2018 the title of Shandong Provincial Distinguished Foreign Expert. He is corresponding member of the Academia de Ciências de Lisboa since 2006, and member of the EU Academy of Sciences since 2014.

The importance of *ab initio* electronic structure calculations in quantum molecular science has prompted this short overview with emphasis on carbon compounds at the Periodic Table celebratory sessions in the Academy of Sciences of Lisbon. Aiming at accuracy, the issue of extrapolating the calculated raw energies to the complete one-electron basis-set (CBS) limit is first examined. For brevity, only the electron correlation contribution to the total energy is considered since it is the most difficult to converge. With the uniform-singlet-and-triplet-extrapolation scheme at the focal point, the emphasis is on recent updates. Still, rather than an even survey, the discussion centers on applications to pure carbon clusters and related carbon-hydrogen compounds, with references given here and there to other material that is left uncovered. Aiming at spectroscopic and reaction dynamics studies, the representation of global potential energy surfaces is then briefly addressed by concentrating on methods developed over the years in the author's Group. Because the purist route to the calculation and modeling of

[2] Department of Physics, Qufu Normal University, China; Department of Physics, Universidade Federal do Espírito Santo, 29075-910 Vitória, Brazil and Department of Chemistry, and Chemistry Centre, University of Coimbra, 3004-535 Coimbra, Portugal

global potentials for large-sized clusters from first-principles appears unaffordable at present, a predictive scheme is suggested to prompt first guesses for more complete *ab initio* work ahead. Prospects for future work conclude the overview.

I. INTRODUCTION

Approximations are unavoidable in molecular physics, with that due to Born and Oppenheimer[1] (BO) being most fundamental. Owing to the mass disparity of nuclei and electrons (the former are at least 1837 times heavier than the latter), BO proposed that their motions could be treated separately. In fact, an even smaller mass ratio may justify such an approximation: by considering four equally charged fermions, two positive and two negative, we have shown[2,3] that a mass ratio between the heavier (assumed the positive ones) and lighter (negative) fermions of 200 was enough to validate the separability of their motions up to ≃ 80 %. As a result, the electronuclear Schrödinger equation splits into two: one for the electrons moving at a fixed arrangement of the nuclei (electronic Schrödinger equation, eSE), the other for the nuclei moving on the potential energy surface (PES or potential) created by the electrons (nSE). The nuclei are said to move adiabatically governed by the PES. Only the eSE is of concern in the present work, leaving aside the nSE which is key in reaction dynamics[4] and where classical[5] (and references therein) approaches are often validated due to the large masses of the nuclei.

Two major difficulties (explosions) arise in computational quantum chemistry based solely in first principles (*ab initio*):[6] a) the ≃ X^{3N-6} explosion signals the number of times that the eSE needs to be solved pointwise to map the PES of a N-atom species (X is a typical number required per dimension); b) the X^{12} explosion which indicates how the cost per point raises with the cardinality of the basis (X is its cardinal number).[7] Added to such explosions is the need for PESs with chemical accuracy (≤ 1 cal mol^{-1}) in reaction dynamics, and spectroscopic accuracy (≤ 1 cm^{-1}) if rovibrational calculations are at stake. Both essentially imply that the PES is at least calculated at the one-electron complete basis set (CBS) limit. Extrapolation is then required, which may use purely mathematical methods or be based on a physically motivated asymptotic theory as is the case here.

The utility of CBS extrapolation gets enhanced when combined with fragment-based methods in which a large molecule is made tractable by explicitly considering all parts into which it can be fragmented, thence as a many-body expansion[8] (MBE) development. If the eSE eigenenergies are first split into HF- and correlation-type contributions, as usually done in modern *ab initio* theory, the approach is known as double-MBE[9,10] (DMBE).

Appearing in the second row of the periodic table, the carbon atom has four bonding electrons in its valence shell, and hence it can form four bonds with other atoms. In particular, C atoms can bond together forming C-clusters. Given the flexibility of carbon to form bonds with most elements, we focus on applications to clusters and elementary reactions where it is involved, in particular with hydrogen. Naturally, the focus will be on relatively small clusters and molecules, since they are key to underpin the properties of larger ones.

Following common practice in the literature, bond lengths are given in bohr (a_0=0.529 × 10^{-10} m, and energies in hartree, kcal mol^{-1} or kJ mol^{-1}; 1 Eh = 627.510 kcal mol^{-1} = 2625.5 kJ mol^{-1}).

II. ELECTRONIC STRUCTURE METHODS: A SYNOPSIS

Methods for solving the eSE are of utmost importance in computational molecular science.[11] The simplest is Hartree-Fock (HF), a mean-field theory where electron correlation is ignored. The error due to its disregard is significant, and hence more sophisticated single-reference (SR) MO-based ones emerged: variational (configuration interaction, CI), perturbative Møller-Plesset (many-body perturbation theory like MP2), and couple-cluster (CC). Of these, the CC singles and doubles with perturbative triples method, CCSD(T), is commonly viewed as the golden rule of quantum chemists.

Because the electron-electron repulsion operator has a singularity at $r_{12} = 0$, the exact wave function must have a discontinuous derivative as implied by Kato's[12] cusp condition. Because the conventional methods fail to satisfy it, this largely explains their very slow convergence. An enormous progress has recently been done toward the solution of this problem through the development of so-called explicitly correlated (R12 and F12) electron correlation methods since they allow to accelerate the basis set convergence of the wavefunction.[13-15] In fact, studies of thermochemistry,[16-19] noncovalent interactions,[20-24] and vibrational frequencies[25,26] have reported gains of at least two[27] angular momentum increments on their conventional counterparts. Yet, such methods involve approximations: benchmark runs with the CCSD-F12a variant of CCSD show a slight overestimation of correlation, while the CCSD-F12b variant favours a slight underestimation. Moreover, their convergence to CBS limit is often nonmonotonic. Because integration of the eSE with accuracy at demand is expensive, mostly unreachable, the alternative is to systematize the error of conventional methods and extrapolate to predict the inherent error.

Two ways stand therefore to obtain accurate energies: solution of the eSE after explicit introduction of correlation in the wave function,[15]and exploitation of the convergence of hierarchized correlation consistent basis sets toward the CBS limit. Despite a fast convergence (often reported $\propto X^{-7}$), explicitly correlated (R12-type) methods appear to perform inefficiently with small basis sets[15]. Additionally, conventional and R12 methods are known to converge to the same asymptotic energy, with CBS extrapolation even outperforming in overcoming noncompleteness of the one-electron basis set, a merit recognized[28] by the number of CBS schemes vying the R12 techniques (see elsewhere[29] for an extrapolation calculator developed for some popular schemes).

Suffice it to add at this point that the HF energy converges exponentially, while being computationally less demanding.[30] The focal point here will then be at the correlation energy,[6,31] with the reader addressed elsewhere[32,33] for HF/CBS or CASSCF/CBS [the latter involves only static (nondynamical) correlation; see later] extrapolation schemes.

Since SR methods still miss the nondynamical correlation, this must be recovered at the multireference (MR) level, typically with complete-active-space-self-consistent-field (CASSCF; particularly popular is the so-called full-valence CASSCF or FVCAS variant) and MRCI wave functions, the latter accounting for the dynamical correlation by inclusion of singles and doubles excitations (MRCISD), often also with inclusion of Davidson's correction for quadruple excitations, MRCI+Q. In this case too, extrapolation to the one-electron CBS limit plays an extremely useful role.[32] Although extrapolation to the $\mathcal{N}$-electron basis set limit has been investigated,[34] its application has been less common in the literature.[35-37]

Another popular a priori electronic structure approach is density functional theory (DFT). By far the leading method used in computing the electronic structure properties of medium and large-sized molecules, the Kohn-Sham[38] (KS) DFT variant is its mainstream. It is an exact formulation of quantum mechanical electronic structure theory but relies on approximate exchange-correlation functionals.[39] As a result, there is a proliferation of DFT functionals, with the best for one application being often not the best for another.[40] Recently,[41,42] we have shown that second-order Møller-Plesset perturbation theoretic results extrapolated from the first steps of the hierarchical staircase[6,31,43] to the CBS limit can rival DFT/M06-2X[39] (and references therein) both on time and accuracy. Such a performance actually extends to other popular DFT functionals.[41,42]

The surge of DFT methods in fields like cluster chemistry and organometallic catalysis to find the many existing stationary points and even reaction pathways comes therefore as no surprise given its cost-effectiveness. In fact, although chemical intuition and comparison with similar reactions can help on the endeavour, the number of such topographical features makes it a formidable task which, most importantly, remains prone to overlooks. To overcome drawbacks, the development of automated procedures to find intermediate species is pivotal.[44-50] Some of these techniques combine geometrical approaches to identify the stationary point with dynamics simulations, with the minima obtained by tracing the intrinsic reaction coordinate paths from the transition states.

III. CBS EXTRAPOLATION: ELECTRON CORRELATION

From a partial-wave expansion for two-electron atoms, it has been established that:[51] a) for natural-parity singlet states, the leading contribution to the energy at second-order of perturbation theory is $\propto (\ell + 1/2)^{-4}$ with no odd-terms either $\propto (\ell + 1/2)^{-5}$ or $\propto (\ell + 1/2)^{-7}$; b) for triplet states, the leading term is or $\propto (\ell + 1/2)^{-6}$; These findings remain essentially unaltered for atoms with any number of electrons.[51,52] If $\Delta E_l \propto \sum_{m=4} a_m (l + 1/2)^{-m}$, the convergence error when $\ell \geq L$ assumes then the form:

$$\Delta E = \sum_{m=4} A_{m-1} (L+1)^{-m+1} \qquad (1)$$

with A_3 and A_{35} being the first two leading coefficients; considering just the first can be accuracy--limiting.[52]

Largely motivated by the possibility of CBS extrapolation, modern basis sets are commonly built according to a principal expansion. Among them are the popular Gaussian-type orbital (GTO) correlation-consistent basis sets[53-55] (cc-pV*X*Z or V*X*Z), diffuse augmented ones (aug-cc-pV*X*Z or AV*X*Z), etc., where the cardinal number X (= 2: *D*, 3: *T*, 4: *Q*...) is identifiable with n and $L + 1$.[56,57] The above slow convergence must be compounded with further scalings[58] at MP2, CCSD, and CCSD(T) calculation levels, namely $N^5 N_b^4$, $N^6 N_b^4$, and $N^7 N_b^4$, where N_b is the number of basis functions per atom; $N_b \simeq X^3$ for a V*X*Z basis.[7,57] Because correlated calculations beyond *Q*Z are often unaffordable for many interesting systems, the raw energies may then be left too far for a safe extrapolation by some popular schemes.

A. The USTE scheme: update survey

CBS extrapolation of conventional electronic energies is best performed by extrapolating separately its HF and correlation components. The latter, the only of concern in the present work, scales $\propto X^{-3}$ for opposite spin electron pairs and $\propto X^{-5}$ for pairs with the same spin. The USTE[59] scheme accounts for both as

$$E_X^{corr} = E_\infty^{corr} + \frac{A_3}{(X+\alpha)^3}\left[1 + \frac{\tau_{53}}{(X+\alpha)^2}\right] \quad (2)$$

where $\tau_{53} = (A_5^o/A_5 + cA_3^{m-1}$; α, A_5^o, and c are universal *ab-initio*-based parameters. Empirical-free and showing the correct asymptotic behavior,[6] it is dual-level giving a prediction as accurate as one possibly can get when extrapolating from raw energies for the two highest affordable X values. Yet, use of (D, T) at most is key for larger systems. This is the goal of GUSTE,[60] where τ_{53} has been suggested to be treated as invariant over configuration space once determined for one geometry with $X \geq Q$ raw data. However, even this is there out of reach.

Eq. (2) may be rewritten[61] as $E_X^{corr} = E_\infty^{corr} + Ax^{-3}$ where the hierarchical number $x \equiv \tilde{X}$ is defined by

$$\tilde{X} = (X+\alpha)\left[1 + \frac{A_5/A_3}{(X+\alpha)^2}\right]^{-1/3} \quad (3)$$

The novel concept is that the basis is educated to account for deficiencies on its composition according to the recovered correlation energy. Stated differently, a hierarchical staircase as straight as possible in X^{-3} is envisaged to enhance reliability when extrapolating from any two steps.[61] Although more than one possibility exists for reassignment,[61] we suggested[43] to obtain the new hierarchical numbers (x) as statistical averages of the values obtained from the condition that the $X \leq 6$ values fall on the straight line obtained by fitting USTE(5,6) correlation energies.[59] The method kept the original acronym[43] but specifies the hierarchical number-pair used for the extrapolation: USTE(x_1, x_2). The novel hierarchical numbers $x = d, t, q, p, h, \ldots$, are real positive but still universal as they apply to any correlation-consistent--type basis sets.[62] Since the correspondence for sub-minimal [sM, thence smaller than *D*Z, which are minimal (M); larger ones are extended (E)] basis sets may not be obvious,[63] the basis may alternatively be indicated.

The most recent version of USTE assumes the form:[63]

$$E_X^{corr} = \eta\, E_\infty^{corr} + \frac{A_3}{x^3} \quad (4)$$

where η is a tolerance factor, and

$$x = X - p_0 - \frac{1}{2}\tanh\left(\frac{X-X_0}{p_1}\right) \quad (5)$$

where X_0, p_1 and p_2 are universal parameters. Named USTE$_a$(x-1, x), where *a* stands for analytic, this protocol[63] yields high quality results while allowing to extrapolate from any pair of x values, thence any basis sets. Its reliability has actually been checked against the best available estimates which have also been employed as reference to scrutinize raw energies obtained from MP2-F12 and CCSD(T)-F12 calculations.[64] To enhance agreement and delve into subchemical accuracy (<1 kcal mol^{-1}), a tiny scaling (fixed at $\eta \simeq 1 \pm 0.001$) has been allowed. This tiny scaling helps to level off the effect of having used CBS(V*5*Z, V*6*Z) energies as reference, and the fact that the calculations were not at optimized geometries but at all-electron CCSD(T)/CVTZ ones.[43] Indeed, it enhances agreement of the USTE$_a$ predictions with the reference raw F12 energies:[64] rmsd of 0.180 and 0.086 kcal mol^{-1} for MP2/CBS and CCSD(T)/CBS, respectively. To go beyond this (*i.e.*, to attain spectroscopic accuracy) would imply including other corrections such as core and core-valence effects, perturbative contributions for connected quadruple excitations, and relativistic effects that lie outside the scope of the model.

Given its high reliability, USTE$_a$(x-1, x) may also be used to assign a cardinal number to any arbitrary basis.[63] With it, basis sets from subminimal Pople's STO-*n*G and Huzinaga's MINI ones to the most advanced extended *VXZ*-F12 anstazes have been ranked from their ability to recover the correlation energy.

USTE has also been extensively applied, and much of the work recently reviewed.[6] Most recently, it has been used[41] to assess how correlated MO calculations perform versus Kohn-Sham DFT by testing the performance of both methods on the calculation of 38 hydrogen transfer barrier heights and 38 non-hydrogen transfer barrier heights/isomerizations extracted from extensive databases, in addition to four 2p isomerization reactions and six others for large organic molecules.[39] All KS DFT calculations employed the popular M06-2X functional, while the correlated MO-based ones used MP2 and CCSD(T) with the raw MO energies subsequently CBS(*d*, *t*) extrapolated. MP2/CBS(*d*, *t*) was found[41] to be as cost effective as DFT/M06-2X while showing a satisfactory accuracy when compared with the reference data. A similar performance was observed for even-numbered carbon clusters.[65]

Another illustration that is claimed by the paper's title involves an organic molecule with as many as 45 isomers, since such molecules are known to pose a challenging problem to DFT.[66-69] Specifically considered is C_8H_8 since accurate isomerization energies are available[70] for comparison from *ab initio* calculations and the W1-F12 thermochemical protocol.[71] Moreover, a whole range of hydrocarbon functional groups [these include (linear and cyclic) polyacetylene, polyyne, and cumulene moieties, as well as aromatic, anti-aromatic, and highly-strained rings] is involved, while results are available also from composite semiempirical procedures and a panoply of DFT functionals.[70]

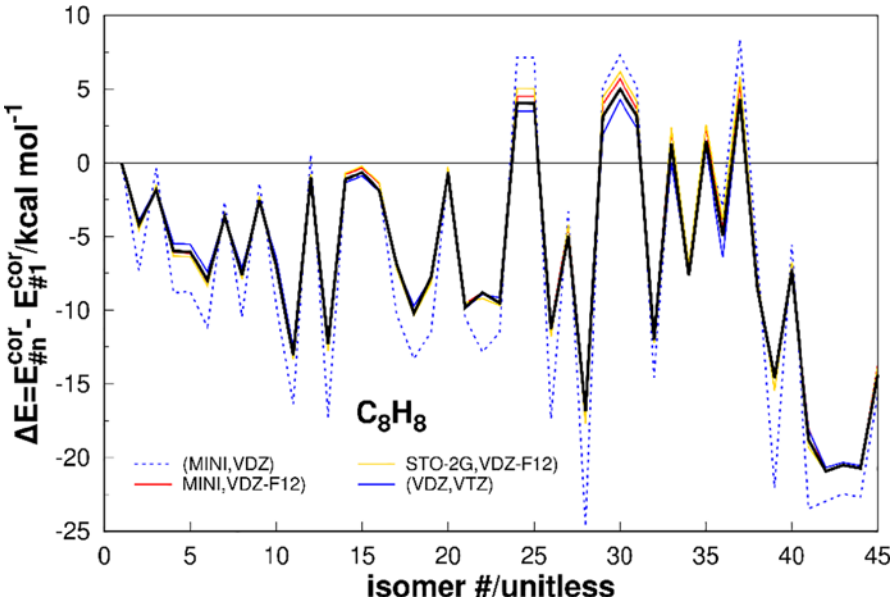

Figure 1
Energy separations of the C_8H_8 isomers at CCSD(T) level of theory; see elsewhere[70] for their names and B3LYP geometries. Shown[70] in solid black is CBS(VDZ-F12,VTZ-F12). Adapted from Ref. 63.

Figure 1 compares the CCSD(T)/CBS correlation energies for C_8H_8 with the best available results.[63] Suffice here to note that the trends observed with MP2 are similar but at a drastically smaller cost.[63] Indeed, the MP2/CBS energetics may be enhanced at zero-cost to approximate couple-cluster quality via spin scaling[72] using variable-scaling opposite spin[73] (VOS) theory. The following results are highlighted:[63] a) use of a (sM, M) basis-set pair is enough to mimic the CCSD(T) reference data[71] with high accuracy [rmsd of 3.01, 0.42, and 0.64 kcal mol^{-1} from (MINI, VDZ), (MINI, VDZ-F12), and (STO-2G, VDZ-F12), respectively], which compares with 0.49 kcal mol^{-1} from our recently[41] recommended CBS(*d, t*) scheme; b) the wall-times are generally much smaller than the references,[70] and up to a fiftyfold factor than CCSD(T)/CBS(V*D*Z-F12, V*T*Z-F12) for the cheapest extrapolation pair; c) the above results outperform DFT/M06-2X by up to 2.8 kcal mol^{-1}, which performs itself similar to MP2/CBS(MINI, VDZ-F12); d) CBS(sM, M) schemes are pseudo single-level;[6,74] e) CBS(sM, sM) extrapolations show somewhat modest performances, but at drastically smaller costs while occasionally performing at an accuracy comparable to some DFT functionals. Reasons for such a performance were advanced based on the so-called closeness criterion.[63] In summary, while giving a reliable prediction of the ups and downs in the evolution of the isomerization energy, such educated predictions contrast with the mismatched pattern observed at raw ab initio level with sM bases, an ordering also difficult to get with DFT.[70]

IV. A FURTHER GLIMPSE ON CARBON COMPOUNDS

Continuing on carbon compounds, suffice it to note that much of their diversity and complexity stems from the capacity of C atoms to bond with one another in various chain and ring structures

and 3*D* conformations, as well as for linking with other atoms. Indeed, they are probably as many as the different types of living organisms, thus justifying the specialized field of organic chemistry. Although organic molecules may contain other elements, it is the carbon-hydrogen bond that defines them as organic and organic chemistry as chemistry of life. Having addressed some intricacies of C_8H_8 in section IIIA, we turn in this section to pure carbon clusters and C_3H, an akin carbon--hydrogen compound.

A. Small pure carbon clusters

Carbon clusters have long attracted both chemists and physicists alike. The small ones are key in the chemistry of carbon rich stars, comets and interstellar molecular clouds, while acting as building blocks in the formation of complex C-containing species. Besides the panoply of astrophysical significance, they are important in the formation of fullerenes, nanotubes, and carbon-rich thin films, while predominant in terrestrial shooting flames.[75-83] All this due to the exceptional properties of C in forming single, double, and triple (eventually quadruple in the dimer[84]) bonds. Clearly, the elucidation of possible mechanisms leading to formation and growth of such C-clusters requires that the properties of small ones are understood.[80] It turns out that distinct but nearly isoenergetic isomers can be formed in a high-density of low-lying singlet and triplet states, which makes their study most challenging.[81,83] In fact, C-clusters in the small size range have been described in a variety of mass spectrometric observations,[65,75-77,93-99] while both spin states have been extensively studied with *ab initio* and DFT[100,101] calculations.

Although KS DFT has been vastly used in studying C-clusters, we have recently[41] shown that MP2/CBS(*d*, *t*) energies[6,31,43] rival DFT calculations with the popular M06-2X functional[39] (and references therein) both in time and accuracy. Such a performance extends to other functionals: MP2/CBS(*d*, *t*) outperforms DFT/B3LYP-D3[102] for the same V*DZ* basis set by showing energy errors at least twice smaller for the same test set. Suffice to add that the MP2/CBS(*d*, *t*) energetics could be enhanced at zero-cost to approximate couple-cluster quality by using variable-scaling opposite spin[73] (VOS) theory.

Carbon clusters also offer a fertile ground to test methods for automated location of the many stationary points occurring on their potentials, and even full reaction pathways. Even if no claim is made of a fully elaborated tool,[65] we have suggested a simple scheme based solely on *ab initio* calculations that may be used to locate their most relevant intermediates. The approach consists of inducing an adiabatic breakup of a bond (preferably at a minimum), which is then stimulated to follow until dissociation using *ab initio* techniques. Because it is essentially a generalization of our own optimized reaction coordinate[61,103] (ORC) method, it was named ORC for stimulated evolution (ORCSE[65]) where all but the inactive degrees of freedom (DOF) are optimized. Briefly, the following three-point premise is accepted in ORCSE: (1) all intermediates are well approximated at MP2/CBS(*d*, *t*) level of theory; (2) all are accessible through a reaction coordinate that involves the stretch of a bond, a twist, or even any specially designed combination of stretches and twists, once all other DOF are fully optimized; (3) given the limitations of the optimization process, alternative paths may potentially be induced in unveiling other (unknown) stationary points. Although full optimization of all DOF but the inactive coordinate warrants in principle completeness, this cannot be ensured

due to difficulties in covering the full configuration space and the fact that most algorithms converge to the closest stationary point.

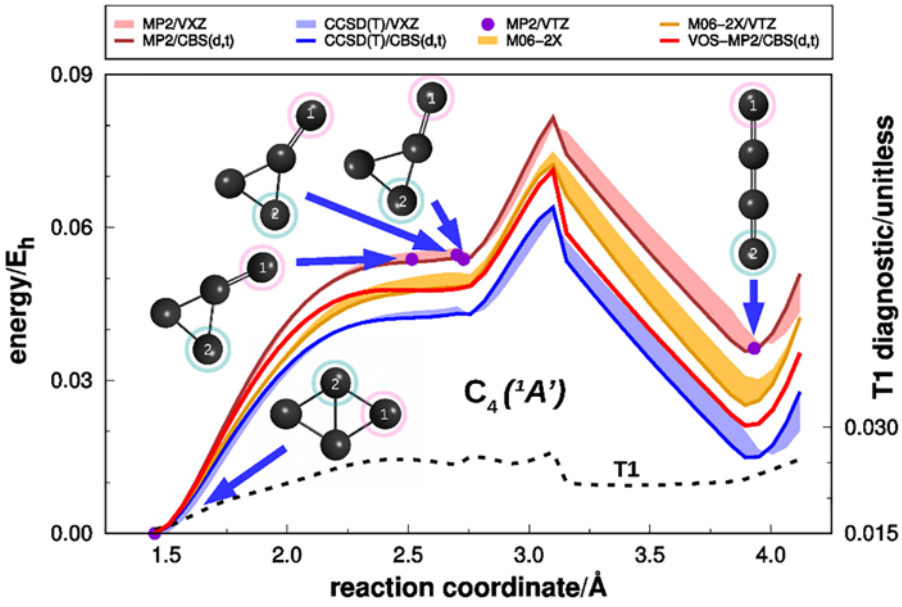

Figure 2
ORCSE path showing all structural isomers of $C_4(^1A')$ obtained by varying the distance between atoms 1 (circled in pink) and 2 (cyan). With R_{12} the inactive coordinate, all other DOF have been fully optimized, and the energies taken relative to the starting geometry. The shaded areas indicate the range of energies covered from *DZ* to CBS(*d*, *t*), with the latter indicated by the solid colored line. A similar procedure is adopted for DFT, except for the line that refers now to M06-2X/*VTZ*. The dots indicate fully optimized MP2/*VTZ* energies, and the black dashed line the T1 diagnostic[106] for validity of single--reference methods. For illustration, two other structures (not necessarily stationary points up to a tight convergence) before and after the C_{2v} TS, are also shown. Adapted from Ref. 65.

In the following, we illustrate ORCSE for singlet C_4 (see elsewhere[65] for other clusters), which aimed originally at validation of the approach since it had already been studied at high levels of *ab initio* theory.[83] In fact, linear triplet C_4 has also been studied spectroscopically, although cyclic ^1A_g eluded identification thus far. Being a small cluster, the *VTZ* basis set could be utilized,[104,105] with all raw energies subsequently CBS extrapolated from the two lowest steps of the hierarchical staircase for correlation consistent basis sets, $x = d$ and t. Figure 2 shows the ORCSE path[65] for $C_4(^1A')$. The overall evolution process is seen to occur stepwise in a plane: starting from the ^1A_g global minimum (rhombic), the system attains a C_s monocyclic ring form (distorted kite) via a ring-opening process in which a single peripheral bond is broken, crosses the C_{2v} transition state (TS), visits the other equivalent kite structure, and finally attains linearity after passing a peak of high energy associated with a *L*-shaped structure. Along such a path, single-point MP2/V*DZ*, VOS-MP2/V*XZ*, and CCSD(T)/V*XZ* (*X*=*D*, *T*) calculations were next performed, and subsequently CBS(*d*, *t*) extrapolated.[65] Notably, the height of the C_{2v} transition state in C_4 shows good agreement with our own CASDC/CBS(*T*,*Q*) estimate[83] of 29.73 kcal mol^{-1} using AV*XZ* basis sets (see elsewhere[107] for the CASDC method). The corresponding result with DFT/M06-2X and a V*DZ* basis is 33.1 kcal mol^{-1}, thence

similar to MP2/CBS(*d*, *t*) value of 33.9 kcal mol^{-1}. No comparison has been possible for the V*T*Z basis since M06-2X does not predict such a saddle point, apparently not an uncommon finding.[108] Of course, the very good agreement[50] observed may have been somewhat accidental as the CASDC/CBS(*T*, *Q*) estimate[83] for the relative stability of the cyclic *vs* linear forms places the latter 6.14 kcal mol^{-1} above the former, which is nearly 50 % smaller than VOS-MP2/CBS(*d*, *t*) but yet a fourfold factor smaller than with DFT/M06-2X.[108]

Of crucial interest for reaction dynamics is the availability of a global PES, preferably in analytical form. Among the most reliable approaches, the MBE[8,109] and DMBE[9,110,11] methods play a prominent role having acquired popularity. By developing the total interaction energy as an expansion of the energies of all involved atomic subclusters,[8] such methods provide an accurate description of valence interactions while accounting for the correct asymptotic behavior of all *n*-body terms in the series. In fact, all dissociation limits (as well as long- range interactions in DMBE) are naturally warranted. In fact, once the potentials of all fragments have been obtained, a truncated series may even be used to predict an approximate version of the PES for the target polyatomic[8,109] (see later). A word of caution is mandatory though: even if converging rapidly, chemical accuracy is generally attainable only when including up to the highest-order non-pairwise-additive terms.[8,9]

Regarding *l*-$C_4(^3\Sigma_g^-)$, the Wigner-Witmer spin-spatial correlation rules[112,113] show that it dissociates adiabatically as follows:

$$l\text{-}C_4(^3\Sigma_g^-) \rightarrow C_2(a^3\Pi_u) + C_2(a^3\Pi_u) \quad (6)$$

$$\rightarrow l\text{-}C_3(\tilde{X}^1\Sigma_g^+) + C(^3P) \quad (7)$$

$$\rightarrow C_2(a^3\Pi_u) + 2C(^3P) \quad (8)$$

$$\rightarrow 4C(^3P) \quad (9)$$

So, like $C_3(\tilde{X}^1\Sigma_g^+)$,[8,114-116] *l*-$C_4(^3\Sigma_g^-)$ does not dissociate to ground-state C_2 fragments. Indeed, channel (6) lies[117] 17.2 kJ mol^{-1} above $C_2(X^1\Sigma_g^+) + C_2(X^1\Sigma_g^+)$ which is spin-forbidden for *l*-$C_4(^3\Sigma_g^-)$; it correlates[118] with *l*-$C_4(^1\Sigma_g^-)$. Dissociation into $C_2(a^3\Pi_u)$ fragments is also found to occur with an endothermicity of 607.0 kJ mol^{-1}, while dissociation into C_3 and C fragments gives both in their ground states; for the energetics, see left panel of Figure 3. The above collinear reaction path is actually the lowest for formation of *l*-$C_4(^3\Sigma_g^-)$ being exothermic by 503.7 kJ mol^{-1}. In fact, the first excited asymptotic channel $C_3(\tilde{a}^3\Pi_u) + C(^3P)$ lies[119,120] 202.8 kJ mol^{-1} above the asymptote in Eq. (7), thus correlating with higher excited states like *l*-$C_4(^3\Pi_u)$[121] and *l*-$C_4(^1\Sigma_g^+)$. Atomization occurs via dissociation of $C_3(\tilde{X}^1\Sigma_g^+)$ into $C_2(a^3\Pi_u)$ + $C(^3P)$, followed by fragmentation of the diatomic into $C(^3P) + C(^3P)$.[116]

Using the above, the DMBE[9,110,111] PES of C_4 truncated at three-body terms assumes the form

$$V_{C_4}^{(2+3)}(\mathbf{R}) = V_{C_aC_b}^{(2)}(R_1)+V_{C_aC_c}^{(2)}(R_2)+V_{C_aC_d}^{(2)}(R_3)+V_{C_cC_d}^{(2)}(R_4)+V_{C_bC_d}^{(2)}(R_5)+V_{C_bC_c}^{(2)}(R_6)$$
$$+V_{C_aC_bC_c}^{(3)}(R_1,R_2,R_6)+V_{C_aC_cC_d}^{(3)}(R_2,R_3,R_4)+V_{C_aC_bC_d}^{(3)}(R_1,R_3,R_5)+V_{C_bC_cC_d}^{(3)}(R_4,R_5,R_6), \quad (10)$$

where **R**={*Ri*} is a collective variable of all interparticle distances. According to the DMBE[9,110,111] formalism, each *n*-body term is then partitioned into its extended Hartree-Fock (EHF) and dynamical correlation (dc) contributions; see elsewhere[116,122] for details. Suffice to add that all the fragments dissociate into ground-state C atoms, hence one-body terms are not required as they are set as reference.

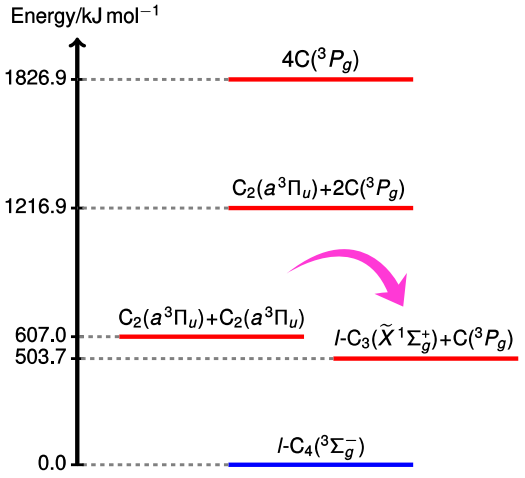

Figure 3
Left: Energetics of the various asymptotic channels of 1-$C_4(^3\Sigma_g^-)$; the arrow signals the $C_2 + C_2 \rightarrow C_3 + C$ exothermic reaction, not studied thus far. Right: Optimized reaction path for interconversion between 1-$C_4(^3\Sigma_g^-)$, and r-$C_4(^3B_{2g})$, via *lr*-$C_4(^3A'')$. The inactive coordinate corresponds to the peripheral bond length of the r-$C_4(^3B_{2g})$ structure (R), with all remaining DOF optimized at each grid point. Stationary structures obtained at CAS(8,12)/AVTZ, CASDC/CBS and CASDC/CBS//CAS(8,12)/AVTZ levels are symbolized by circles, diamonds and triangles, respectively. Energies are relative to *l*-$C_4(^3\Sigma_g^-)$. See also text. Adapted from Ref. 83.

Regarding the potential functions used for two- and three-body terms, they have been taken from *ab initio* potentials previously reported. For enhanced reliability, some have been fine-tuned from available spectroscopic data. Specifically, for the trimers, a simplified version of the multiple energy switching (ES) scheme[123,124] has been employed, with the functions so obtained showing rmsds of some cm^{-1} for rovibrational energies up to about 4000 cm^{-1} above zero-point energy (ZPE); see the original publications for details.

Even if repeatedly noted, high accuracy can only be expected when adding up to the highest-order terms in the MBE, and hence a four-body term is needed. To enhance the PES accuracy, an approximate four-body term has therefore been added to DMBE(2+3), which is denoted DMBE/ES-SS-(2+3) in Ref. 83. For this, the *n*-body distributed polynomial method[125] was employed but with Gaussian functions centered at convenient geometries. It assumes the form[83]

$$V_{C_4}^{(4)}(\mathbf{R}) = \sum_{i=1}^{7} P_i^{(4)}(\mathbf{\Gamma}) G_i(\mathbf{\Gamma}), \tag{11}$$

where $P_i^{(4)}(\mathbf{\Gamma})$ are cubic polynomials,

$$P_i^{(4)}(c_0 + c_1\Gamma_1 + c_2\Gamma_1^2 + c_3\Gamma_2 + c_4\Gamma_3 + c_5\Gamma_1^3 + c_6\Gamma_1\Gamma_2 + c_7\Gamma_1\Gamma_3 + c_8\Gamma_4 + c_9\Gamma_5 + c_{10}\Gamma_6) \tag{12}$$

and

$$G_i(\mathbf{\Gamma}) = \exp[-\gamma_i(\Gamma_1)^2] \tag{13}$$

are range-determining factors. In turn, Γ_i (i=1-6) are totally symmetric combinations of displacement coordinates[8,126]

$$\Gamma_1 = Q_1 \tag{14}$$

$$\Gamma_2 = Q_2^2 + Q_3^2 + Q_4^2 \tag{15}$$

$$\Gamma_3 = Q_5^2 + Q_6^2 \tag{16}$$

$$\Gamma_4 = Q_2 Q_3 Q_4 \tag{17}$$

$$\Gamma_5 = Q_6^3 + 3Q_6Q_5^2 \tag{18}$$

$$\Gamma_6 = Q_6(2Q_2^2 - Q_3^2 - Q_4^2) + \sqrt{3}Q_5(Q_3^2 - Q_4^2) \tag{19}$$

where Q_i (i=1-6) are symmetrized displacements from a T_d reference of bond length R_0:[8,126]

$$\begin{pmatrix} Q_1 \\ Q_2 \\ Q_3 \\ Q_4 \\ Q_5 \\ Q_6 \end{pmatrix} = \begin{pmatrix} \sqrt{1/6} & \sqrt{1/6} & \sqrt{1/6} & \sqrt{1/6} & \sqrt{1/6} & \sqrt{1/6} \\ \sqrt{1/2} & 0 & 0 & -\sqrt{1/2} & 0 & 0 \\ 0 & \sqrt{1/2} & 0 & 0 & -\sqrt{1/2} & 0 \\ 0 & 0 & \sqrt{1/2} & 0 & 0 & -\sqrt{1/2} \\ 0 & 1/2 & -1/2 & 0 & 1/2 & -1/2 \\ \sqrt{1/3} & -\sqrt{1/12} & -\sqrt{1/12} & \sqrt{1/3} & -\sqrt{1/12} & -\sqrt{1/12} \end{pmatrix} \begin{pmatrix} R_1\text{-}R_0 \\ R_2\text{-}R_0 \\ R_3\text{-}R_0 \\ R_4\text{-}R_0 \\ R_5\text{-}R_0 \\ R_6\text{-}R_0 \end{pmatrix}. \tag{20}$$

Thence, in the T_d symmetry point group, Q_1 transforms as A_1, (Q_2, Q_3, Q_4) as the triply-degenerate T_2 irreducible representation and (Q_5, Q_6) as the double-denegerate E mode.

To calibrate the above four-body term, 663 *ab initio* points have been calculated for C_4. Of them, 53 refer to constrained optimized geometries for collinear approximations of C_3 + C and C_2 + C_2 at CASDC/CBS level, using C_{2v} and D_{2h} symmetries, respectively. The four-body interaction energies were then obtained from the requirement that they should vanish at all dissociation limits once subtracting DMBE(2+3) from the total interaction energies. Additionally, 76 arrangements related to the ORC path in Figure 3 have been included. To acquire CASDC/CBS quality, the actually computed MRCI(Q)-8/AV*T*Z energies were finally scaled to reproduce the CASDC/CBS splitting between the *l*-$C_4(^3\Sigma_g^-)$ and r-$C_4(^3B_{2g})$ forms (ΔE_{lr} = 0.0555 E_h) using

$$\mathcal{F} = \frac{[E_3^{\text{MRCI(Q)-8}}(\mathbf{R}_r) - E_3^{\text{MRCI(Q)-8}}(\mathbf{R}_l)]}{\Delta E_{lr}}, \tag{21}$$

where $E_3^{\text{MRCI(Q)-8}}(\mathbf{R}_r)$ and $E_3^{\text{MRCI(Q)-8}}(\mathbf{R}_l)$ denote MRCI(Q)-8/AV*T*Z energies of the rhombic and linear isomers at FVCAS/AV*T*Z optimized geometries, yielding $\mathcal{F}$=1.7010. From Eq. (21), the total interaction energy of any arbitrary structure x with respect to *l*-$C_4(^3\Sigma_g^-)$ was then obtained as

$$E_x(\mathbf{R}) = \frac{[E_3^{\text{MRCI(Q)-8}}(\mathbf{R}_x) - E_3^{\text{MRCI(Q)-8}}(\mathbf{R}_l)]}{\mathcal{F}} + E_l(\mathbf{R}), \tag{22}$$

where $E_l(\mathbf{R})$ = -0.7127 E_h is the total interaction energy of the linear global minima (with respect to the infinitely separated atoms) predicted from CASDC/CBS calculations, a result in excellent agreement with the experimental estimate of -0.6958 E_h.[98,116,127]

Figures 4 and 5 illustrate salient attributes of the final DMBE(2+3+4) PES so obtained for ground-state triplet C_4.[83] Specifically, Figure 4 shows a contour plot for C moving around C_3. The notable feature is the T-shaped structure at (x, y)=(0.00, 2.69) a_0, which resembles the top of a barrier connecting the two symmetry equivalent l-$C_4(^3\Sigma_g^-)$ structures. It turns out not to be a true transition state in the 6*D* configuration space of C_4 but a saddle point of index 3, as actually predicted from the MRCI(Q) calculations. Another feature is the T-shaped long-range structure at (x, y) = (0.00, 4.02) a_0 where C_4 assumes an equilateral triangular form with D_{3h} symmetry possibly a symmetry-imposed conical intersection due to the involved C_3 fragments where such a topological feature is present. It turns out that preliminary CAS(8,12)/AV*T*Z calculations and the sign-reversal property of the wave-function[128] have not confirmed such a prediction but revealed the presence of a high-density of close-in-energy states in that region. Note that C-clusters are fertile ground for the appearance of such topological features, with the interested reader referred elsewhere[83,129,130] (and references therein) for details on the underlying theory. The in-plane attack of C_2 by another C_2 is shown in Figure 5.[83] A notable feature is a C_{2v} structure at (x, y)=(0.00, 4.05) a_0. Apparently a minimum, it is actually a saddle point in 6*D*. Related to the degenerate isomerization of symmetrically equivalent l-$C_4(^3\Sigma_g^-)$ structures, such a feature has been confirmed both at MRCI(Q)-8/AV*T*Z and CCSD(T)/AV*T*Z levels of theory.[83] Instead, such a C_3-monocyclic form is found to be a minimum in DFT calculations.[108,131] Also visible are two additional extrema on the DMBE potential: a distorted C_2 capped triangle at (x, y)=(4.58, 2.64) a_0 and a quasi-rhomboedric form at (x, y)=(5.08, 0.67) a_0. Although predicted as minimum and transition state (respectively) in 2*D*, their rigorous assignment could not be done as it would require high-level FVCAS/MRCI frequency calculations, a task unaffordable at present.

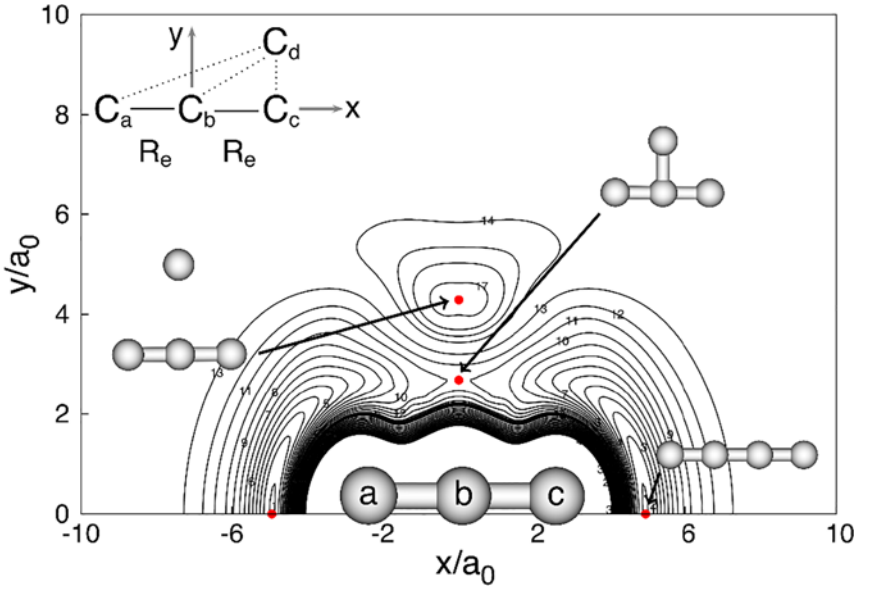

Figure 4

Partially relaxed contour plot of DMBE(2+3+4) PES [DMBE/ES-SS-(2+3+4) in the original work] for C moving around C_3 which lies along the x axis with the origin fixed at the central carbon atom. Contours are equally spaced by 0.015 E_h, starting at -0.7312 E_h. Adapted from Ref. 83.

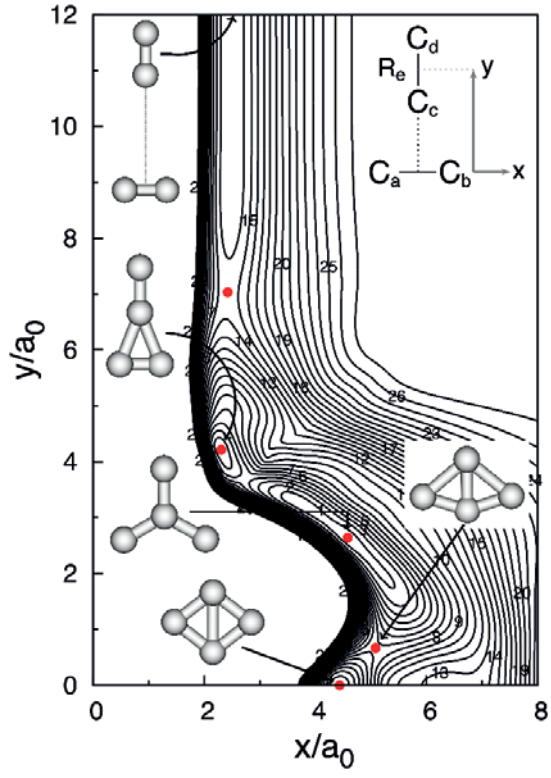

Figure 5
Partially relaxed contour plot for C2v insertion of C2 into another C2 obtained from the DMBE(2+3+4) [ES-SS-(2+3+4) in the original work] PES. Contours are equally spaced by 0.015 Eh, starting at -0.6770 Eh. Adapted from Ref. 83.

B. A key carbon-hydrogen molecule: C_3H

Ubiquitous in the interstellar medium (ISM),[132] small carbon-bearing species like C_n and C_nH(n = 1-3) are conspicuous in driving C-chemistry[133] in cold dense clouds[134-136] and circumstellar envelopes of evolved C-rich stars.[137-139] In ISM,[140] C_3H is believed to play a major role, reaching high fractional abundances ($\simeq 10^{-9}$) when compared to H_2.[133,141] Both its cyclic[142] (c-CH_3, cyclopropynylidyne) and linear[143] (ℓ-C_3H; propynylidyne) isomers are deemed as formed there either via dissociative electron recombination of[144] c,ℓ-$C_3H_2^+/C_3H_3^+$ or through the $C+C_2H_2$ neutral pathway[144-147] which is key in the formation of C-chains in space.[145-147] This prompted further surmises[148] on the role of c-C_3H as intermediate (via c-C_3H_2 formation) in the synthesis of interstellar polycyclic aromatic hydrocarbons which are recognized as potential carriers of unidentified IR bands.[149] The following summarizes the current status of the title radical with emphasis on the work done at the author's Group.[150]

Starting with ℓ-C_3H, it has a $^2\Pi$ ground-state and two bending modes, ν_4(C-C-H) and ν_5(C-C-C), which are perturbed by Renner-Teller (RT) and spin-orbit effects.[143] Discovered by Gottlieb *et al.*[151] who measured its microwave spectra in both $^2\Pi_{1/2}$ and $^2\Pi_{3/2}$ (ground) vibronic states, it was further studied by Yamamoto[152] and Kanada *et al.*[153] who recorded pure rotational lines in ν_4 ($^2\Sigma^\mu$) while finding ℓ-C_3H to have an extremely low vibrationally excited state ($\approx$ 27 cm^{-1} above $^2\Pi_{1/2}$) due to strong RT effects in ν_4.[152] Subsequent work focused on improving[154] the spectroscopic constants of $^2\Pi_r$ and ν_4 ($^2\Sigma^\mu$) and

extending the range of rotational transitions in the $^2\Sigma$ vibrationally excited manifold.[155] IR vibrational band centers for stretching modes ν_1, ν_2 and ν_3 were provided both in Ar matrices,[156] and gas phase.[157]

Regarding *c*-C_3H, it has a $^2B_2$[142] ground electronic state as first detected by Yamamoto *et al.*[142,158] using microwave spectroscopy. Based on the predicted rotational constants, they reported the molecular structure of *c*-C_3H, while confirming its C_{2v} symmetry.[158]

Theoretically, a wealth of *ab initio* calculations were reported for C_3H. Early ones[142,153,156,159-164] were mainly devoted to elucidate discrepancies between the predicted symmetries of the ℓ- and *c*-C_3H forms and those inferred from microwave spectroscopy,[153,158] with the best estimate[164] placing *c*-C_3H ≈ 14 kJ mol^{-1}more stable than ℓ-C_3H. Such studies suggest that the calculated optimum structures may be transition states and that a slightly distorted C_s form may prevail. Such a symmetry breaking issue was first considered for the cyclic isomer by Stanton and coworkers.[165,166] Using the equation-of-motion CCSD method for ionized states, they emphasized the basis set role on the proper determination of its C_{2v} symmetry and increase of X $^2B_2 \rightarrow$ A $^2A_1$ excitation energy. They further noted that the pseudo-Jahn-Teller effect between such states (previously considered[158] as responsible for its C_s equilibrium structure) is weakened when the size of the one-electron basis set is enhanced and CBS extrapolation done. Similar conclusions (regarding also the *N*-electron basis) were drawn at the MRCI+Q/V*X*Z (*X* = *D* – *Q*) level of theory,[104,167] with Halvick[168] noting that the stiffness of the PES along the C-C asymmetric w_4 increases with correlation enhancement. In turn, Bassett and Fortenberry[169] reported a quartic force field (QFF) for *c*-C_3H from a composite scheme based on accurate CBS extrapolated CCSD(T)/AV*X*Z (*X* = *T*-5) energies that were additively corrected for core correlation and scalar relativistic effects.[169] From a QFF local form so obtained and using second-order vibrational perturbation theory, the authors further reported[169] rotational constants, structural parameters and anharmonic vibrational frequencies for the ground-state (X $^2B_2$) that are likely the highest-level *ab initio* estimates thus far. Most recently, Bennedjai *et al.*[170] performed a spectroscopic characterization of the C_3H isomers at CCSD(T)-F12/AV*T*Z level of theory. Yet, the most comprehensive theoretical study to date on ℓ-C_3H was carried out by Perić *et al.*[171] who provided local MRCI+Q/V*T*Z forms (including relativistic effects) for both ^{2}A′ and ^{2}A″ electronic states correlating with the $^2\Pi$ term. They further employed their PESs to compute the vibronic and spin-orbit structure of the ℓ-C_3H spectrum using a variational approach.[171] Moreover, they noted the extremely flat nature of the CCC-H bending potential curve (^{2}A′) and, like others,[153,162] did not rule out the possibility of its quasilinearity. Despite the fact that their local forms assume $C_{\infty v}$ equilibrium geometries, the various spectroscopic parameters have shown excellent agreement with the experimental results. Indeed, Ding *et al.*[172] have shown that the minimum structure of ℓ-C_3H tends to move with basis set size enhancement (from AV*T*Z to AV*Q*Z) from a bent to a linear geometry at both CASSCF and CCSD(T) levels of theory. In turn, from CCSD(T)/6-311+G(3df,2p) calculations on the ground-state C_3H, Mebel and Kaiser[173] pointed out that the ℓ- and *c*-forms forms rearrange into each other through a ring-opening step via an asymmetric transition state (ℓc-C_3H) with a barrier of about 115 kJ mol^{-1}. They further note that besides C(^3P_j) + $C_2H(X^1\Sigma_g^+)$, the barrierless reactions CH(X $^2\Pi$) + $C_2(X^1\Sigma_g^+)$ and C(^{3}P) + $C_2H(X\ ^2\Sigma^+)$ may be facile neutral-neutral exothermic pathways to yield carbon trimer in cold interstellar environments. Figure 6 underpins[150] most prominent features of the C_3H PES, clearly making it a unique and challenging species both from the chemical and astrophysical viewpoints.

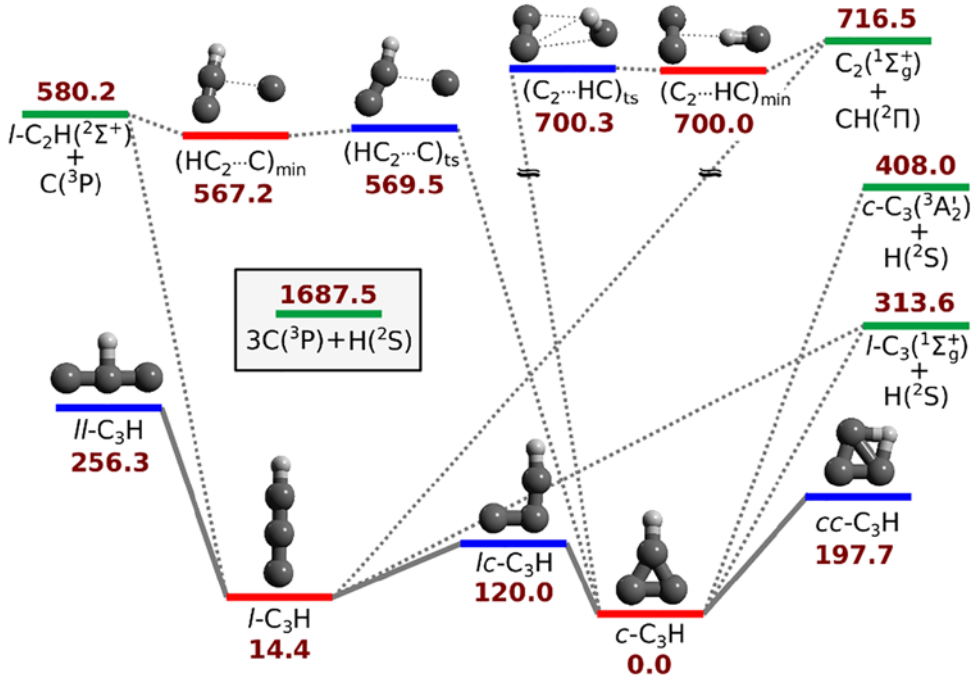

Figure 6
Pathways connecting the stationary points in CHIPR PES. Energies (in kJ mol^{-1}) refer to the global c-C3H minimum. Dotted lines connect dissociative channels (in green), while solid ones connect isomeric structures (red for minima, blue for transition states). Adapted from Ref. 150.

Despite the prevalence of C_3H in ISM and its relevance in C-chain formation having stimulated considerable experimental[142,151-158,172] and theoretical[159-166,168-171,173,174] work to understand its intricate chemistry, most studies focused on energetics, symmetry and spectroscopy of its isomeric forms, and hence on local potential functions. Most recently, we have reported[150] the first global PES for ground state C_3H based on the CHIPR[175-177] method. It correlates with the $^2\Pi$ state at linear geometries and $^2B_2$ at cyclic ones, while describing correctly all fragmentation channels (see Figure 6). Indeed, it has already been successfully employed to carry out the first dynamics study of the H + C_3 reaction.[150] The analytical form is based on the matrix

$$\mathcal{H}_e = \begin{pmatrix} \mathcal{V}_{11}^{(2+3)}(\mathbf{R}) & \epsilon(\mathbf{R}) \\ \epsilon(\mathbf{R}) & \mathcal{V}_{22}^{(2+3)}(\mathbf{R}) \end{pmatrix} \tag{23}$$

where

$$\begin{aligned} \mathcal{V}_{11}^{(2+3)}(\mathbf{R}) = & \sum_{i=1}^{3} V^{(2)}_{\mathrm{CH}(^2\Pi)}(R_i) + \sum_{i=4}^{6} V^{(2)}_{\mathrm{C_2}(^3\Pi_u)}(R_i) + V^{(3)}_{\mathrm{C_2H}(^2\mathrm{A'})}(R_2,R_3,R_6) + V^{(3)}_{\mathrm{C_2H}(^2\mathrm{A'})}(R_1,R_3,R_5) \\ & + V^{(3)}_{\mathrm{C_2H}(^2\mathrm{A'})}(R_2,R_3,R_6) + V^{(3)}_{\mathrm{C_3}(^1\mathrm{A'})}(R_4,R_5,R_6) \end{aligned} \tag{24}$$

$$\mathcal{V}_{22}^{(2+3)}(\mathbf{R}) = \sum_{i=1}^{3} V^{(2)}_{\mathrm{CH}(^2\Pi)}(R_i) + \sum_{i=4}^{6} V^{(2)}_{\mathrm{C_2}(^1\Sigma_g^+)}(R_i) + V^{(3)}_{\mathrm{C_2H}(^2\mathrm{A'})}(R_2,R_3,R_6) + V^{(3)}_{\mathrm{C_2H}(^2\mathrm{A'})}(R_1,R_3,R_5) + V^{(3)}_{\mathrm{C_2H}(^2\mathrm{A'})}(R_2,R_3,R_6) + V^{(3)}_{\mathrm{C_3}(^3\mathrm{A'})}(R_4,R_5,R_6) \quad (25)$$

Because CHIPR assumes the MBE[8] form, with the reference being the infinitely separated ground-state C and H atoms, only a few remarks are required on its two-body and three-body terms. Following Ref. 178, $V^{(2+3)}(\mathbf{R})$ has been first defined as the lowest PES arising from diagonalization of a 2×2 pseudo-diabatic matrix, with the diagonal terms constructed from previously reported two- and three-body (ground-state) potentials for $C_2H(^2A')$,[179] $C_3(^1A')$[180] and $C_3(^3A')$[178] (see the original work[178-180] for details). In turn,

$$\epsilon^{(4)}(\mathbf{R}) = E^{\mathrm{CC/CBS}(d,\,t)}(\mathbf{R}) - V^{(2+3)}(\mathbf{R}) \quad (26)$$

is a coupling term chosen to warrant that the eigenvalues of $\mathcal{H}_e$ are continuous everywhere. The energetics of the various dissociation channels predicted from $V^{(2+3)}(\mathbf{R})$ compare well[150] with CC/CBS(d, t) calculations as well as experiment.[127,179,180] In fact, the agreement is quite good, with just a slight discrepancy for $\mathcal{L}$-$C_3(^1\Sigma_g^+)$ + $H(^2S)$, which is attributed to the more attractive nature of the $\mathcal{L}$-$C_3(^1\Sigma_g^+)$ three-body term. Note that this channel has the largest experimental uncertainty, while an accurate estimate of the $\mathcal{L}$-C_3 atomization energy still awaits determination. Because four-body energies vanish at all asymptotic channels, the dissociation energies in $V^{(2+3+4)}(\mathbf{R})$ remain as in $V^{(2+3)}(\mathbf{R})$.

An effective four-body term $V^{(4)}(\mathbf{R})$ must next be added,[150] namelly

$$V^{(4)}(\mathbf{R}) = \sum_{i,j,k,l,m,n=0}^{L} C_{i,j,k,l,m,n} \left[\sum_{g \in G} \mathcal{P}_g^{(i,j,k,l,m,n)} (y_1^i y_2^j y_3^k y_4^l y_5^m y_6^n) \right] \quad (27)$$

In fact, according to the CHIPR formalism the interaction energy, $\epsilon^{(4)}(\mathbf{R})$ can be conveniently modelled[175,176] using a L^{th}-degree polynomial ($i + j + k + l + m + n \le L$), where $C_{i,j,k,l,m,n}$ are expansion coefficients and y_p (p = 1, 2, ..., 6) are transformed coordinates (see below). Note that only C coefficients referring to excitations of at least four modes (thence $i \le j \le k \le l \le m \le n$) can be included, thus satisfying the constraints that avoid inclusion of two- or three-body contributions. Note further that the second summation runs over all permutation elements g ϵ G, where G is a subgroup of the $\mathcal{S}_4$ symmetric group.[182] For an AB_3-type molecule like the title one, $P_g^{(i,j,k,l,m,n)}$ operators that reflect the action of the particle permutations[182] of $\mathcal{S}_4$ onto the exponent set {i, j, k, l, m, n} brought by the first summation in Eq. (27). They will generate the required symmetrized sums of monomials that make $V^{(4)}(\mathbf{R})$ invariant to all permutations of identical atoms.[175-177]

Being a key point in CHIPR, every y_p is then expanded as a distributed-origin contracted basis[175-177]

$$y_p = \sum_{\alpha=1}^{M} c_\alpha \phi_{p,\alpha}, \quad (28)$$

with $\phi_{p,\alpha}$ expressed either as

$$\phi_{p,\alpha}^{[1]} = \mathrm{sech}^{\eta}(\gamma_{p,\alpha}\rho_{p,\alpha}) \tag{29}$$

or

$$\phi_{p,\alpha}^{[2]} = \left[\frac{\tanh(\beta R_p)}{R_p}\right]^{\sigma} \mathrm{sech}^{\eta}\left(\gamma_{p,\alpha}\rho_{p,\alpha}\right) \tag{30}$$

where $\rho_{p,\alpha} = R_p - R_{p,\alpha}^{\mathrm{ref}}$ defines the displacement coordinate from the origin of the α^{th} primitive, $\gamma_{p,\alpha}$, are non-linea$_\alpha$ parameters, and $\eta = 1$, $\sigma = 6$ and $\beta = 1/5$ are constants.[175-177] As usual,[10] both $\phi_{p,\alpha}^{[1]}$ and $\phi_{p,\alpha}^{[2]}$ are employed,[150] with the latter appearing only once as the last term in the summation. Additionally, all distributed origins are related by[175-177]

$$R_{p,\alpha}^{\mathrm{ref}} = \zeta(R_p^{\mathrm{ref}})^{\alpha-1} \tag{31}$$

where ζ and R_p^{ref} are adjustable parameters.

Multiple cuts on the 6*D* configurational space of the CHIPR form have shown that it to reproduces reliably all its major topographical attributes.[150] For example, the left panel of Figure 7 illustrates the isomerization of the symmetry-equivalent *c*-C_3H structures which occurs via the C_{2v} transition state *cc*-C_3H, yet unreported at the date of our publication.[150] Located 197.7 kJ mol^{-1} above *c*-C_3H, it shows an imaginary frequency of 1693.9 cm^{-1} for the H wagging motion through the *c*-C_3 center-of-mass. In turn, classical barrier heights calculated[178] from CC/CBS(*d*, *t*) and MRCI+Q/CBS(*D*, *T*) protocols have shown good agreement with the CHIPR predictions, namely 194.2 and 188.8 kJ mol^{-1}, in the same order. Indeed, a close look at inset (b) shows that the CHIPR form reproduces accurately the entire minimum energy path (MEP) calculated at the CC/CBS(*d*, *t*) level of theory.

Besides *c*-C_3H, the CHIPR form[150] predicts a linear ℓ-C_3H isomer which shows as a minimum 14.4 kJ mol^{-1} above the cyclic form. For comparison, the corresponding energy differences predicted at CC/CBS(*d*, *t*), CC-F12/AV*T*Z,[170] and MRCI+Q/CBS(*D*, *T*) levels of theory are 11.8, 15.4 and 5.3 kJ mol^{-1}, respectively. Regarding ℓ-C_3H, the structural parameters predicted from the CHIPR form[150] also agree well with the experimental (vibrationally averaged) r_0 values,[153] with CHIPR placing ℓ-C_3H ~ 14.5 kJ mol^{-1} above *c*-C_3H. However, discrepancies may be expected in the description of C-C-H and C-C-C degenerate bending modes whose vibrational angular momenta is known[153,154] to couple with the $^2\Pi$ (total) electronic angular momentum. Recall[153,171] that the strong RT effect makes the lowest ^{2}A′ PES very flat along w_4, with appearance of a quasi-linear C_3H molecule. Yet, in accordance with recent CC-F12 calculations[170] and Ref. 172, CBS extrapolations tend to favor the highest symmetry $C_{\infty v}$ species. Suffice to add that the predicted structural parameters and harmonic frequencies for ℓ-C_3H are close to those predicted from the local PES of Perić *et al.*[171].

To conclude this section, the right hand side panel of Figure 7 shows a contour plot for CH moving around C_2. Clearly visible is the presence of a T-shaped (C_2v) transition state, $\ell\ell$-C_3H, which is responsible for the degenerate isomerization of the symmetry-equivalent ℓ-C_3H structures. Its

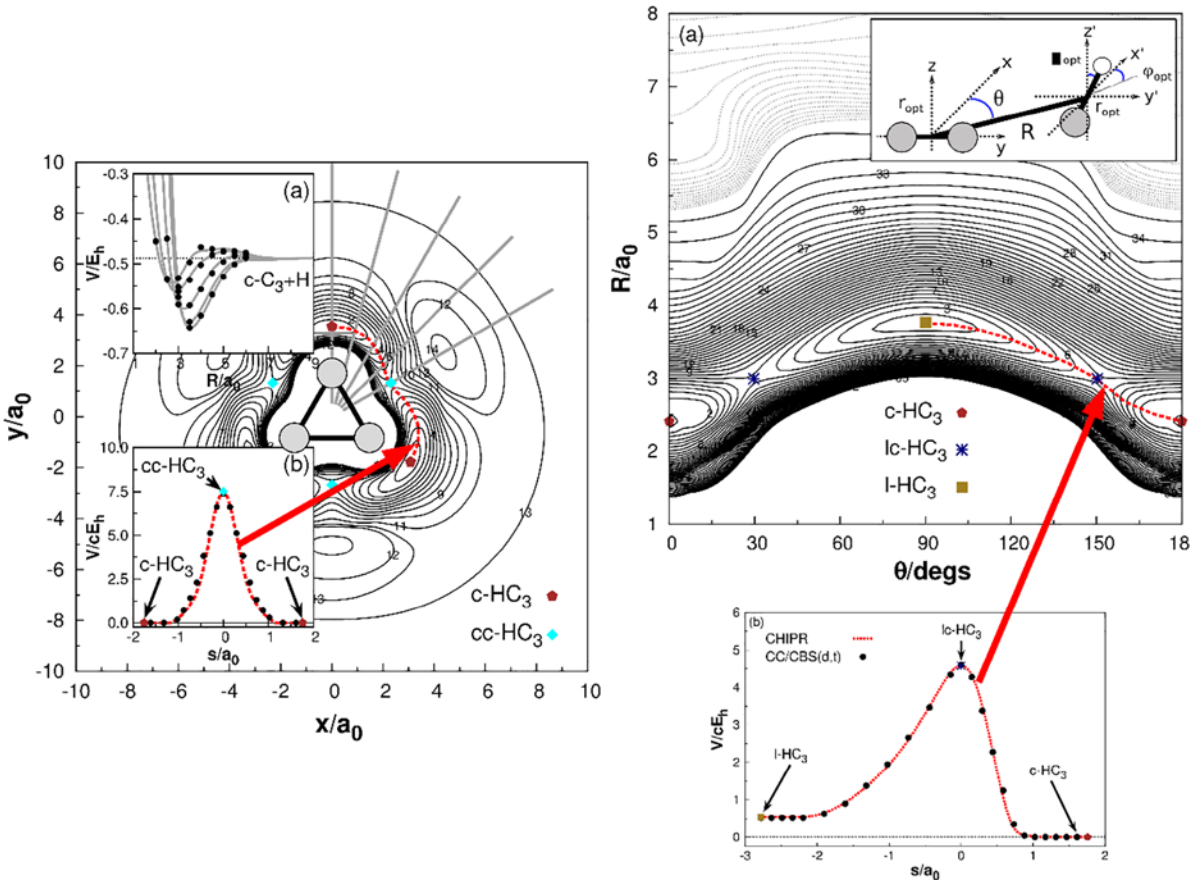

Figure 7

Left: contours for H atom moving around partially relaxed *c*-C_3 with center-of-mass fixed at the origin. Contours are equally spaced by 0.0135 E_h starting at -0:65 E_h. The reference energy is that of the infinitely separated atoms. Insets: (a) optimized 1D cuts; (b) minimum energy path for isomerization process (s is the reaction coordinate, with the reference energy being that of *c*-C_3H). Solid dots indicate ab initio CC/CBS(d, t) points while the dotted lines represent the dissociation predicted from $V^{(2+3)}(\mathbf{R})$. Right: (a) contours for CH moving around a C_2 diatomic with center-of-mass fixed at the origin. All DOF but R and θ are partially relaxed. Black solid and gray dotted contours are equally spaced by 0.0075 and 0.00075Eh starting at -0.65 and -0.385 E_h. (b). MEP (s is the reaction coordinate in mass-scaled atomic units) for the isomerization *c*-C_3H ⇋ *ℓ*-C_3H via transition state *ℓc*-C_3H. In both contour plots, the zero of energy refers to the infinitely separated atoms while for the right hand side (b) plot it corresponds to *c*-C_3H. Adapted from Ref. 150.

calculated imaginary frequency amounts to 1111.7 cm^{-1} and points toward the H wagging motion through the *ℓ*-C_3 center-of-mass. Such a feature, firstly reported in our work,[150] lies about 241.9 and 256.3 kJ mol^{-1} above the *ℓ*-C_3H and *c*-C_3H (respectively), with the corresponding isomerization barriers predicted from CC/CBS(*d*, *t*) and MRCI+Q/CBS(*D*, *T*) being 246.3 and 245.8 kJ mol^{-1}. Moreover, the bottom right-hand-side panel of Figure 7 evinces the reliability of the CHIPR form in reproducing the CC/CBS(*d*, *t*) MEP for this process. Note that the isomerization between *c*-C_3H and *ℓ*-C_3H occurs via the *ℓc*-C_3H transition state[144,173] whose imaginary frequency (795.1 cm^{-1}) points along the C-C bond breaking/forming process. Indeed, its classical barrier height of 120.0 kJ mol^{-1} relative to *c*-C_3H compares well with both CC/CBS(*d*, *t*) and MRCI+Q/CBS(*D*, *T*) estimates of 122.1 and 123.8 kJ mol^{-1}, respectively.

C. Large-size carbon clusters: are global potentials affordable?

Although a clear yes might be the expected answer, some remarks are in order. First, to recall that all reactions take place on a PES, although in a very large molecule (always assumed along the paper to be in vacuum) where thousands of coordinates may be involved (a well known case is protein folding,[183] even if in vacuum) there will be a difference relative to a simple chemical reaction: the conformational freedom is much larger than in the case of a small molecule, and hence the possibility arises that entropy is more important than in most small-molecule reactions. Because the point here is to discuss PESs rather than free-energy surfaces, there is the need to impose a limiting size to what is meant by large molecule: it must be such that reaction is expected to be governed by the potential energy (rather than the free energy) but still big enough for stimulating the posed query and some helpful analysis.

As a second remark, the terms full-dimensionality and globalness (often used in the literature as equivalent) deserve clarification. A local (versus global) potential function may be full-dimensional when involving all (3*N*-6) DOF, but a global potential must be full-dimensional, while implying that all possible channels associated to the involved *N*-atom molecule are in principle assessable. Thence, the latter should in principle be applicable in any dynamics study, from intramolecular to scattering, non-reactive or reactive.

Now, if global, the PES may be obtained by numerical interpolation [although the literature is vast[184] (and references therein), the most traditional approach – not implying cost-effectiveness or feasibility – employs cubic splines[185]] or by least-squares fitting of some functional form. While a truly global numerical potential is hardly conceivable (suffice it to recall the fact that the eSE may not have a converged solution at some regions of configuration space), analytical forms are commonly cheaper to use but suffer from being hard to formulate in hugely dimensional configuration spaces. All this without having in mind the computational cost of such a numerical potential, an issue already raised in the Introduction. Entering the realm of C-clusters, one should now recall that all intricacies encountered in the PESs of small clusters (and even more subtle ones that are likely to arise with increasing cluster size) will be present in the hugely dimensional configuration spaces of large clusters (174*D* for C_{60}).

From the above paragraphs, it is plausible to conclude that obtaining an accurate *ab initio*-based global PES for such large C-clusters is currently little less than unaffordable. Approximate methods are therefore at stake, with the problem demanding approximations both in the selection of the *ab initio* method for solution of the eSE and approach in fitting the myriad of points that must be calculated.

Rather than focusing on the purist approach of solving the eSE and fitting of the calculated points in a blind manner, we survey a recent suggestion to predict the structures and energetics of such C-clusters[186] by employing a truncated MBE[8] or its DMBE congener.[9,110,111] Accordingly, the PES of the large cluster is first approximated through a MBE development in terms of the potentials of the involved subclusters up to four-body ones.[8] This will not only warrant a correct description of the asymptotic behavior of every *n*-body term in the MBE, but also the involved dissociation limits (as well as long--range interactions in the case of using the DMBE method) are naturally described. Once such building--blocks are available, the truncated MBE or DMBE so obtained will enable a first prediction of the PES for the target polyatomic.[8,109] Of course, and emphasizing the point once more, high accuracy is only attainable when including up to the highest-order (in the limit, all non-pairwise-additive) terms.[8,9] Still,

a major advantage over popular semiempirical valence-bond theories such as the DIM method[187-189] is that simpler and more flexible functional forms can be employed to fit *ab initio* and/or experimental information, and hence obtain reliable high-dimensional global potentials.

Regarding C-clusters, we follow Pitzer and Clementi[190] who first recognized the ℓ-C_x forms as low energy isomeric structures on their corresponding PESs. Thence, odd- and even-numbered chains (for $x > 2$) are assumed to correlate with their $^1\Sigma_g^+$ and $^3\Sigma_g^-$ ground electronic states, respectively. This is to say that all comparisons with the predictions to be made follow such an assumption.

Indeed, despite the fact *ab initio* calculations tend to support that monocyclic (singlet) isomers are nearly isoenergetic or even more stable than linear (triplet) arrangements,[191] the current predictions seem to indicate (in absence of cyclic global minima) that the fragments on which the corresponding $V^{(2+3+4)}$ DMBE forms are built correlate with triplet states of the C_x (x = 4, 6, 8, 10) clusters. This is rationalized from the fact that the PESs of ℓ-$C_3(\tilde{X}\,^1\Sigma_g^+)$ and ℓ-$C_4(^3\Sigma_g^-)$ dissociate as follows:

$$\ell\text{-}C_3(\tilde{X}\,^1\Sigma_g^+) \rightarrow C_2(a\,^3\Pi_u)+C(^3P) \quad (32)$$

$$\rightarrow 3C(^3P) \quad (33)$$

and

$$\ell\text{-}C_4(^3\Sigma_g^-) \rightarrow C_2(a\,^3\Pi_u)+C_2(a^3\Pi_u) \quad (34)$$

$$\rightarrow \ell\text{-}C_3(\tilde{X}^1\Sigma_g^+) + C(^3P) \quad (35)$$

$$\rightarrow C_2(a\,^3\Pi_u)+2C(^3P) \quad (36)$$

$$\rightarrow 4C(^3P) \quad (37)$$

thus restricting any dissociation of the C_x (x = 5-10) PESs to occur according to the above.

We next survey some virtues of the approach[186] by examining the first estimates of global DMBE PESs for C_x (x = 5-10) clusters obtained from the reported potentials for C_2, C_3 and C_4.[83,116] First, Figure 9 illustrates the predictive capability of $V^{(2+3+4)}$ in providing estimates of global minima and their associated thermochemical properties.[186] Clearly, the good quality of the results give important insights into the structure-determining nature of the (2+3+4)-body terms. Indeed, they allow to judge the methodology as providing at least satisfactory first guesses of the true C_x potentials. Furthermore, it opens the way to construction of global reliable forms for large C-clusters (and possibly even other atomic clusters). Of course, the present approach may be open to improvement, *e.g.* by employing simple functional forms for higher-order terms (say, n=5 or 6) while assuming that the MBE/DMBE approximately converges beyond them. Unfortunately, this is difficult to judge since such terms may be minor but are numerous, *e.g.* for C_{10} there are 210 and 120 six- and seven-body terms, respectively. Still, for large clusters, it is fair to say that a large number of such terms approximately vanish since many of the atoms are far from the others.

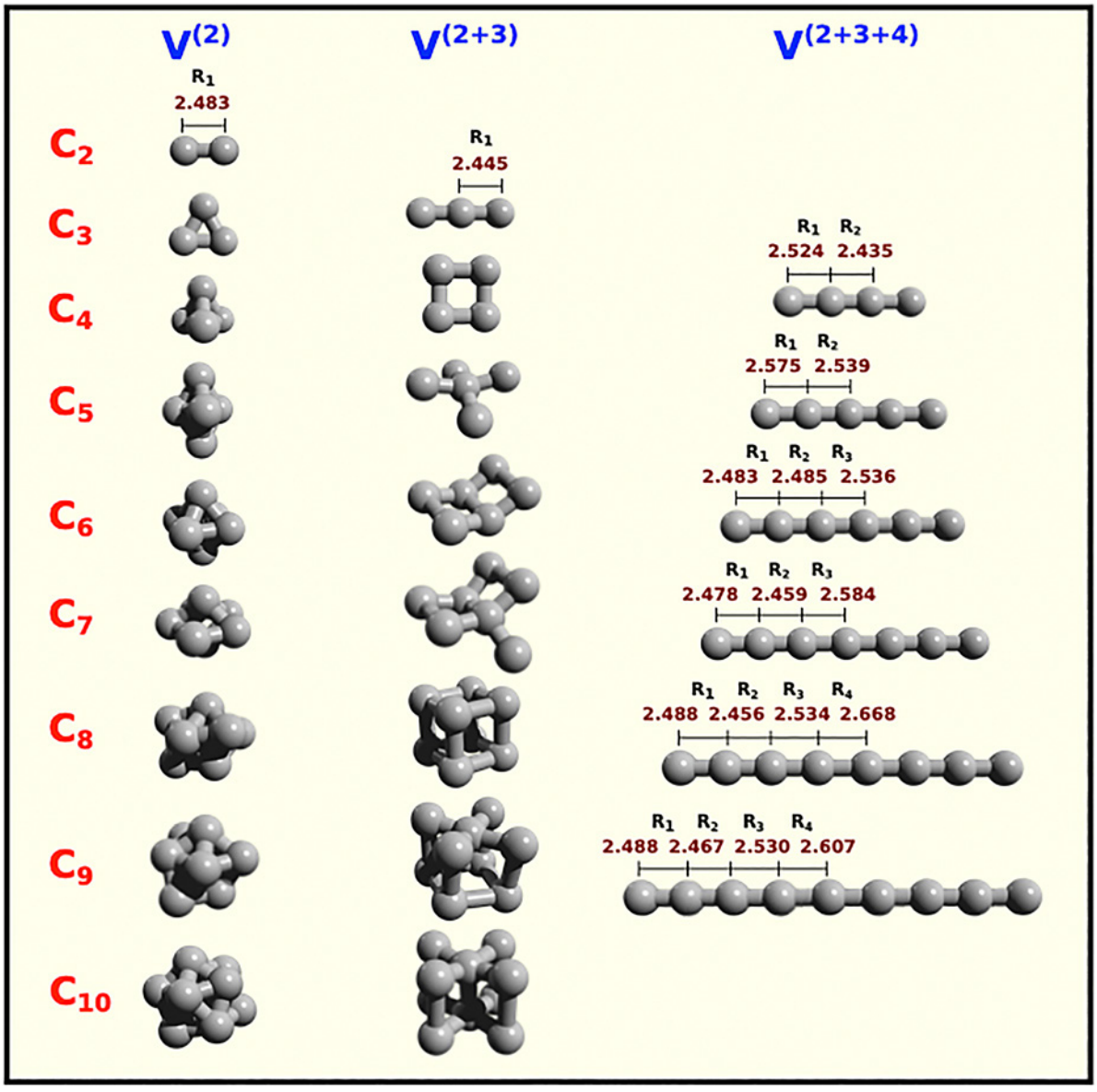

Figure 8
C_x (x=2-9) global minima as predicted from a truncated $V^{(2+3+4)}$ DMBE expansion. Also shown are structures obtained when PESs with only two- [$V^{(2)}$] or two- plus three-body potentials [$V^{(2+3)}$] are employed to get structural parameters, harmonic vibrational frequencies and dissociation energies for the ℓ-C_x clusters. Adapted from Ref. 186.

Figure 9 shows contour plots for the collinear reactions ℓ-$C_{(x-2)}$ + C_2 → ℓ-$C_{(x-1)}$+ C, thence for the lowest-energy path yielding ℓ-C_x as predicted from the truncated DMBE forms.[186] Clearly, the predictive nature and cost-effectiveness of the method suggest that such truncated DMBE potentials may even be employed in evaluating gradients and hessians for large species, which may then be used as first guesses for stationary point searches in actual electronic structure calculations. Indeed, the approach has proven useful for C_4.[83]

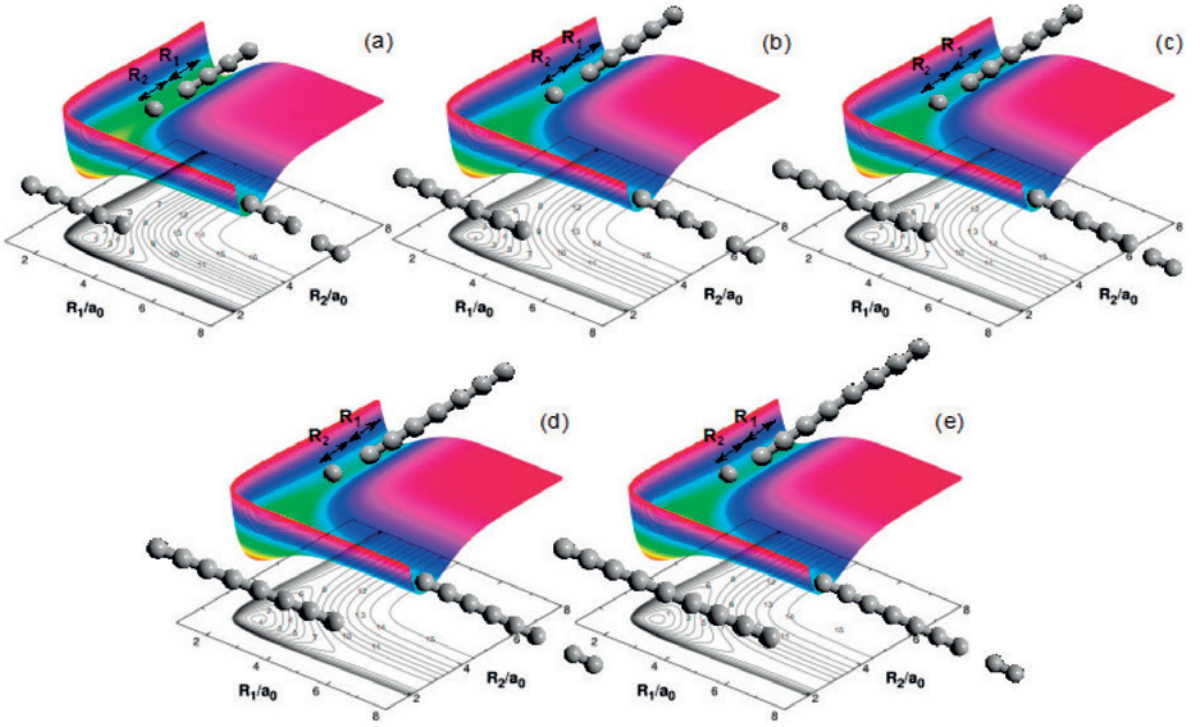

Figure 9
Partially relaxed contour plots for the collinear reactions (a) ℓ-C_3 + C_2 → ℓ-C_4 + C. Contours are equally spaced by 0.02 E_h, starting at -1.1 E_h; (b) ℓ-C4 + C2 → ℓ-C_5 + C. Contours are equally spaced by 0.03 E_h, starting at -1.3 E_h; (c) ℓ-C_5+C_2 → ℓ-C_6+C. Contours are equally spaced by 0.02 E_h, starting at -1.1 E_h; (d) ℓ-C_6 + C_2 → ℓ-C_7 + C. Contours are equally spaced by 0.02 E_h, starting at -1.1 E_h; (e) ℓ-C_7 + C_2 → ℓ-C_8 + C. Contours are equally spaced by 0.02 E_h, starting at -1.1 E_h. Adapted from Ref. 186.

V. CONCLUDING REMARKS

Saturating a basis in electronic structure calculations is key but mostly unaffordable for medium- and large-sized molecules. CBS extrapolation offers a cost-effective and reliable way out, and we have shown[6,63] how to obtain highly accurate correlation energies from various methods and basis sets by CBS extrapolating the raw energies with the USTE scheme. The method is applicable to any basis set family even if not of the correlation consistent type, which we have shown how to hierarchize from the recovered correlation energy. Because the joint use of a SM and an E basis sets costs as much as a single--point calculation just with the latter, its $USTE_a$ variant has become, as efficient and perhaps even more reliable than any genuine single-level scheme.[6] Indeed, due to the low-cost and reliability of MP2/CBS(sM, M) and MP2/CBS(M, M) methods *vs* KS DFT, a wealth of topics are open to revisitation, ranging from CBS extrapolations in large systems to explicitly correlated calculations and on-the-fly dynamics. Left unreviewed in this work was progress on optimal basis sets[192,193] for direct extrapolation of the correlation energy. Specifically built for CBS extrapolating the correlation energy at the same cost as Dunning's V*X*Z and AV*X*Z ansatzes from which they have been derived, the novel optimized basis sets typically outperform the latter by factors of three- to fivefold. Although extrapolation of the HF energy is also key when accuracy is at demand, the topic could not be discussed for brevity, with the reader being addressed elsewhere.[6,31]

Using MRCI or CCSD(T) calculations, we have also shown how to model the global potential energy surfaces of any triatomic or even tetratomic systems in their full dimensionality by using the DMBE expansion and CHIPR methods. Indeed, a general computer code has recently been made available to generate triatomic CHIPR forms, while another of the same family will soon be reported for any tetratomic molecule jointly with an option for refining CHIPR two-body *ab initio* curves to spectroscopic accuracy from a fit to available vibrational levels. Both DMBE and CHIPR methods have also been illustrated by considering ground-state C_4 and the akin C_3H radical. It is hoped that this work may stimulate further theoretical and experimental studies on such species which are of utmost relevance in carbon chemistry both in terrestrial and interstellar media, just to mention a few areas.

ACKNOWLEDGMENTS

I wish to thank all my coworkers for their valuable contributions to the topics covered in the present work. The support from Fundação para a Ciência e a Tecnologia and Coimbra Chemistry Centre, Portugal, through the project UI0313/QUI/2013, co-funded by FEDER/COMPETE 2020-EU, and from China's Shandong Province "Double-Hundred Talent Plan" (2018) is also gratefully acknowledged.

REFERENCES

1. M. Born and J. R. Oppenheimer, Ann. Phys. **84**, 457 (1927).
2. A. J. C. Varandas, J. P. da Providência, M. Brajczewska, and J. da Providência, Chem. Phys. Lett. **610**, 167 (2014).
3. A. J. C. Varandas, J. da Providência, and J. P. da Providência, Can. J. Phys. **98**, 379 (2020).
4. S. Ghosh, R. Sharma, S. Adhikari, and A. J. C. Varandas, Phys Chem Chem Phys **21**, 20166 (2019).
5. A. J. C. Varandas, Chem. Phys. Lett. **439**, 386 (2007)
6. A. J. C. Varandas, Ann. Rev. Phys. Chem. **69**, 177 (2018).
7. A. J. C. Varandas, J. Chem. Phys. **113**, 8880 (2000).
8. J. N. Murrell, S. Carter, S. C. Farantos, P. Huxley, and A. J. C. Varandas, *Molecular Potential Energy Functions* (Wiley, Chichester, 1984).
9. A. J. C. Varandas, Adv. Chem. Phys. **74**, 255 (1988).
10. A. J. C. Varandas, in *Reaction Rate Constant Computations: Theories and Applications*, edited by K. Han and T. Chu (The Royal Society of Chemistry, 2013), chap. 17, pp. 408–445.
11. T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, Chichester, 2000).
12. T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).
13. S. Ten-no and J. Noga, WIREs Comput Mol Sci **2**, 114–125 (2012).
14. L. Kong, F. A. Bischoff, and E. F. Valeev, Chem. Rev. **112**, 75 (2012).
15. C. Hättig, W. Klopper, A. Köhn, and D. P. Tew, Chem. Rev. **112**, 4 (2012).
16. D. Vogiatzis, R. Haunschild, and W. Klopper, Theor Chem Acc **133**, 1446 (2014).
17. W. Klopper, B. Ruscic, D. P. Tew, F. A. Bischoff, and S. Wolfsegger, Chem. Phys. **356**, 14–24 (2009).
18. N. Sylvetsky, K. A. Peterson, A. Karton, and J. M. L. Martin, J. Chem. Phys. **144**, 214101 (2016).
19. M. K. Kesharwani, N. Sylvetsky, A. Köhn, D. P. Tew, and J. M. L. Martin, J. Chem. Phys. **149**, 154109 (2018).
20. L. A. Burns, M. S. Marshall, and C. D. Sherrill, J. Chem. Phys. **141**, 234111 (2014).
21. D. A. Sirianni, L. A. Burns, and C. D. Sherrill, J. Chem. Theory Comput. **13**, 86-99 (2017).
22. M. K. Kesharwani, A. Karton, and J. M. L. Martin, J. Chem. Theory Comput. **12**, 444-454 (2016).
23. B. Brauer, M. K. Kesharwani, S. Kozuch, and J. M. L. Martin, Phys. Chem. Chem. Phys. **18**, 20905 (2016).
24. D. Manna, M. K. Kesharwani, N. Sylvetsky, and J. M. L. Martin, J. Chem. Theory Comput. **13**, 3136 (2017).
25. D. P. Tew, W. Klopper, M. Heckert, and J. Gauss, J. Phys. Chem. A **111**, 11242 (2007).
26. J. M. L. Martin and M. K. Kesharwani, J. Chem. Theory Comput. **10**, 2085 (2014).

27. D. P. Tew, W. Klopper, C. Neissb, and C. Hättig, Phys. Chem. Chem. Phys. **9**, 1921–1930 (2007).
28. D. Feller, K. A. Peterson, and T. D. Crawford, J. Chem. Phys. **124**, 054107 (2006).
29. V. Vasilyev, Computational and Theoretical Chemistry **1115**, 1 (2017).
30. W. Kutzelnigg, Int. J. Quantum Chem. **51**, 447 (1994).
31. F. N. N. Pansini, A. C. Neto, and A. J. C. Varandas, Theor. Chem. Acc. **135**, 261 (2016).
32. A. J. C. Varandas, J. Chem. Phys. **126**, 244105 (2007).
33. A. J. C. Varandas, Phys. Chem. Chem. Phys. **21**, 8022 (2019).
34. L. Bytautas and K. Ruedenberg, J. Chem. Phys. **122**, 154110 (2005).
35. L. Bytautas and K. Ruedenberg, J. Chem. Phys. **128**, 214308 (2008).
36. L. Bytautas and K. Ruedenberg, J. Chem. Phys. **130**, 204101 (2009).
37. L. Bytautas and K. Ruedenberg, J. Chem. Phys. **132**, 074109 (2010).
38. W. Kohn and L. Sham, Phys. Rev. A **140**, 1133 (1965).
39. R. Peverati and D. G. Truhlar, Phil. Trans. R. Soc. A **372**, 20120476 (2014).
40. N. Mardirossian and M. Head-Gordon, J. Chem. Theory Comput. **12**, 4303 (2016).
41. A. J. C. Varandas, M. M. González, L. A. Montero-Cabrera, and J. M. G. de la Vega., Chem. Eur. J. **23**, 9122 (2017).
42. M. M. González, F. G. D. Xavier, J. Li, L. A. Montero-Cabrera, J. M. G. de la Vega, and A. J. C. Varandas, J. Phys. Chem. A **124**, 126 (2020).
43. A. J. C. Varandas and F. N. N. Pansini, J. Chem. Phys. **141**, 224113 (2014).
44. K. Bondensgard and F. Jensen, J. Chem. Phys. **104**, 8025 (1996).
45. W. Quapp, M. Hirsch, O. Imig, and D. Heidrich, J. Comp. Chem. **19**, 1087 (1998).
46. K. Irikura and R. Johnson, J. Phys. Chem. A **104**, 2191 (2000).
47. E. Muller, A. de Meijere, and H. Grubmuller, J. Chem. Phys. **116**, 897 (2002).
48. D. J. Wales and J. P. K. Doye, J. Phys. Chem. A **101**, 5111 (1997).
49. S. Habershon, J. Chem. Theor. Comput. **12**, 1786 (2016).
50. S. Maeda, S. Komagawa, M. Uchiyama, and K. Morokuma, Angew. Chem. Int. Ed. **50**, 644 (2011).
51. W. Kutzelnigg and J. D. Morgan, J. Chem. Phys. **96**, 4484 (1992).
52. J. R. Flores, R. Slupski, and K. Jankowski, J. Chem. Phys. **124**, 104107 (2006).
53. T. H. Dunning Jr., K. A. Peterson, and D. E. Woon, in *Encyclopedia of Computational Chemistry*, edited by P. v. R. Schleyer, N. L. Allinger, T. Clark, J. Gasteiger, P. A. Kolman, and H. F. Schaefer III (Wiley, Chichester, 1998), p. 88.
54. T. H. Dunning Jr., J. Phys. Chem. A **104**, 9062 (2000).
55. F. Jensen, Advanced Review p. 1 (WIREs Comput. Mol. Sci.) (2012).
56. T. Helgaker, W. Klopper, H. Koch, and J. Noga, J. Chem. Phys. **106**, 9639 (1997)
57. A. Halkier, T. Helgaker, P. Jørgensen, W. Klopper, H. Koch, J. Olsen, and A. K. Wilson, Chem. Phys. Lett. **286**, 243 (1998).
58. K. Raghavachari and J. B. Anderson, J. Phys. Chem. **100**, 12960 (1996).
59. A. J. C. Varandas, J. Chem. Phys. **126**, 244105 (2007).
60. A. J. C. Varandas, J. Phys. Chem. A **112**, 1841 (2008).
61. A. J. C. Varandas, J. Phys. Chem. A **114**, 8505 (2010).
62. F. N. N. Pansini, A. C. Neto, and A. J. C. Varandas, Chem. Phys. Lett. **641**, 90 (2015).
63. A. J. C. Varandas, Phys. Chem. Chem. Phys. **20**, 22084 (2018).
64. J. G. Hill, K. A. Peterson, G. Knizia, and H.-J. Werner, J. Chem. Phys. **131**, 194105 (2015).
65. A. J. C. Varandas, The European Physical Journal D **72**, 180 (2018).
66. S. Grimme, Angew. Chem. Int. Ed. **45**, 4460 (2006).
67. P. R. Schreiner, A. A. Fokin, R. A. Pascal Jr, and A. de Meijere, Org. Lett. **8**, 3635 (2006).
68. P. R. Schreiner, Angew. Chem. Int. Ed. **46**, 4217 (2007).
69. M. S. S. Grimme and M. Korth, J. Org. Chem. **72**, 2118 (2007).
70. A. Karton and J. M. Martin, Mol. Phys. **110**, 2477 (2012).
71. A. Karton and J. M. L. Martin, J. Chem. Phys. **136**, 124114 (2012).
72. S. Grimme, J. Chem. Phys. **118**, 9095 (2003).
73. A. J. C. Varandas, J. Chem. Phys. **133**, 064104 (2010).
74. F. N. N. Pansini and A. J. C. Varandas, Chem. Phys. Lett. **631-632**, 70 (2015).
75. S. Arulmozhiraja and T. Ohno, J. Chem. Phys. **128**, 114301 (2008).

76. L. Belau, S. E. Wheeler, B. W. Ticknor, M. Ahmed, S. R. Leone, W. D. Allen, H. F. Schaefer III, and M. A. Duncan, J. Am. Chem. Soc. **129** (2007).
77. K. E. Yousaf and P. R. Taylor, Chem. Phys. **349**, 58 (2008).
78. N. Sakai and S. Yamamoto, Chem. Rev. **113**, 8981 (2013).
79. N. L. J. Cox and F. Patat, A&A **565**, A61 (2014).
80. T. W. Yen and S. K. Lai, J. Chem. Phys. **142**, 084313 (2015).
81. C. M. Rocha and A. J. C. Varandas, J. Chem. Phys. **144**, 064309 (2016).
82. S. K. Lai, I. Setiyawati, T. W. Yen, and Y. H. Tang, Theor. Chem. Acc. **136**, 20 (2017).
83. A. J. C. Varandas and C. M. Rocha, Phil. Trans. Roy. Soc. A **376**, 20170145 (2018).
84. S. Shaik, D. Danovich, W. Wu, P. Su, H. S. Rzepa, and P. C. Hiberty, Nat Chem. **4**, 195 (2012).
85. E. A. Rohlfing, D. M. Cox, and A. Kaldor, J. Chem. Phys. **81**, 3322 (1984).
86. H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley, Nature **318**, 162 (1985).
87. M. Y. Hahn, E. C. Honea, A. J. Paguia, K. E. Schriver, A. M. Camarena, and R. L. Whetten, Chem. Phys. Lett. **130**, 12 (1986).
88. E. A. Rohlfing, J. Chem. Phys. **93**, 7851 (1990).
89. T. Moriwaki, K. Kobayashi, M. Osaka, M. Ohara, H. Shiromaru, and Y. Achiba, J. Chem. Phys. **107**, 8927 (1997).
90. Y. K. Choi, H. S. Im, and K. K. W. Jung, Int. J. Mass Spectrom. **189**, 115 (1999).
91. M. E. Geusic, M. F. Jarrold, T. J. McIlrath, R. R. Freeman, and W. L. Brown, J. Chem. Phys. **86**, 3862 (1987).
92. C. H. Bae and S. M. Park, J. Chem. Phys. **117**, 5347 (2002).
93. J. D. Watts and R. J. Bartlett, Chem. Phys. Lett. **190**, 19 (1992).
94. J. Hutter and H. P. Lüthi, J. Chem. Phys. **101**, 2213 (1994).
95. J. M. L. Martin and P. R. Taylor, J. Phys. Chem. **100**, 6047 (1996).
96. R. O. Jones, J. Chem. Phys. **110**, 5189 (1999).
97. Y. Shlyakhter, S. Sokolova, A. Lüchow, and J. B. Anderson, J. Chem. Phys. **110**, 10725 (1999).
98. Karton, A. Tarnopolsky, and J. M. Martin, Mol. Phys. **107**, 977 (2009).
99. A. J. C. Varandas, European Physical Journal D **72**, 134 (2018).
100. N. A. Poklonski, S. V. Ratkevich, and S. A. Vyrko, J. Phys. Chem. A **119**, 9133 (2015).
101. C. Mauney, M. B. Nardelli, and D. Lazzati, ApJ **800**, 1 (2015).
102. S. Grimme, J. Comput. Chem. **27**, 1787 (2006).
103. A. J. C. Varandas, Phys. Chem. Chem. Phys. **16**, 16997 (2014).
104. T. H. Dunning Jr., J. Chem. Phys. **90**, 1007 (1989).
105. T. H. Dunning, Jr, K. A. Peterson, and A. K. Wilson, J. Chem. Phys. **114**, 9244 (2001).
106. T. J. Lee, Chem. Phys. Lett. **372**, 362 (2003).
107. A. J. C. Varandas, J. Phys. Chem. A **117**, 7393 (2013b).
108. A.C.Ngandjong, Z.Mezei, J. Mougenot, A.Michau, K.Hassouni, G.Lombardi, M.Seydou, and F.Maurela, Comp. Theor. Chem. **1102**, 105 (2017).
109. A. J. C. Varandas and J. N. Murrell, Faraday Discuss. Chem. Soc. **62**, 92 (1977).
110. A. J. C. Varandas, Mol. Phys. **53**, 1303 (1984).
111. A. J. C. Varandas, J. Mol. Struct. Theochem. **21**, 401 (1985).
112. E. P. Wigner and E. E. Witmer, Z. Physik. **51**, 859 (1928).
113. G. Herzberg, *Molecular Spectra and Molecular Structure. III Electronic Spectra and Electronic Structure of Polyatomic Molecules* (Van Nostrand, New York, 1966).
114. S. Carter, I. M. Mills, and J. N. Murrell, J. Mol. Spectrosc. **81**, 110 (1980).
115. S. Carter, I. M. Mills, and R. N. Dixon, J. Mol. Spectrosc. **106**, 411 (1984).
116. C. M. R. Rocha and A. J. C. Varandas, J. Chem. Phys. **143**, 074302 (2015).
117. M. Martin, J. Photochem. Photobiol. A **66**, 263 (1992)
118. J. P. Ritchie, H. F. King, and W. S. Young, J. Chem. Phys. **85**, 5175 (1986).
119. W. Weltner and D. McLeod, J. Chem. Phys. **40**, 1305 (1964).
120. Y. Sych, P. Bornhauser, G. Knopp, Y. Liu, T. Gerber, R. Marquardt, and P. P. Radi, J. Chem. Phys. **139**, 154203 (2013).
121. V. Wakelam, J.-C. Loison, E. Herbst, D. Talbi, D. Quan, and F. Caralp, A&A **495**, 513 (2009).
122. C. M. R. Rocha and A. J. C. Varandas, J. Chem. Phys. **144**, 064309 (2016).
123. A. J. C. Varandas, J. Chem. Phys. **105**, 3524 (1996).

124. A. J. C. Varandas, J. Chem. Phys. **119**, 2596 (2003).
125. E. Martínez-Núñez and A. J. C. Varandas, J. Phys. Chem. A **105**, 5923 (2001).
126. A. J. C. Varandas and J. N. Murrell, Chem. Phys. Lett. **84**, 440 (1982).
127. K. A. Gingerich, H. C. Finkbeiner, and R. W. Schmude, J. Am. Chem. Soc. **116**, 3884 (1994).
128. A. J. C. Varandas, J. Tennyson, and J. N. Murrell, Chem. Phys. Lett. **61**, 431 (1979).
129. A. J. C. Varandas, Chem. Phys. Lett. **487**, 139 (2010).
130. A. J. C. Varandas and B. Sarkar, Phys. Chem. Chem. Phys. **13**, 8131 (2011).
131. S. J. Blanksby, D. Schroder, S. Dua, J. H. Bowie, and H. Schwarz, J. Am. Chem. Soc. **2000**, 7105 (2000).
132. B. A. McGuire, ApJS **239**, 17 (2018).
133. R. I. Kaiser, Chem. Rev. **102**, 1309 (2002).
134. I. W. M. Smith, E. Herbst, and Q. Chang, MNRAS **350**, 323 (2004).
135. I. W. M. Smith, A. M. Sage, N. M. Donahue, E. Herbst, and D. Quan, Faraday Discuss. **133**, 137 (2006).
136. J.-C. Loison, V. Wakelam, K. M. Hickson, A. Bergeat, and R. Mereau, MNRAS **437**, 930 (2014).
137. T. J. Millar and E. Herbst, A&A **288**, 561 (1994).
138. R. J. Hargreaves, K. Hinkle, and P. F. Bernath, MNRAS **444**, 3721 (2014).
139. M. Agúndez, J. Cernicharo, G. Quintana-Lacaci, A. Castro-Carrizo, L. Velilla Prieto, N. Marcelino, M. Guélin, C. Joblin, J. A. Martín-Gago, C. A. Gottlieb, et al., A&A **601**, A4 (2017).
140. *NIST standard reference database number 101* (2013), release 16a. Editor: Russell D. Johnson III, URL http://cccbdb.nist.gov/.
141. M. Agúndez and V. Wakelam, Chem. Rev. **113**, 8710 (2013).
142. S. Yamamoto, S. Saito, M. Ohishi, H. Suzuki, S.-I. Ishikawa, N. Kaifu, and A. Murakami, ApJ **322**, L55 (1987).
143. P. Thaddeus, C. A. Gottlieb, A. Hjalmarson, L. E. B. Johansson, W. M. Irvine, P. Friberg, and R. A. Linke, ApJ **294**, L49 (1985).
144. J.-C. Loison, M. Agúndez, V. Wakelam, E. Roueff, P. Gratier, N. Marcelino, D. N. Reyes, J. Cernicharo, and M. Gerin, MNRAS **470**, 4075 (2017).
145. K. M. Hickson, J.-C. Loison, and V. Wakelam, Chem. Phys. Lett. **659**, 70 (2016).
146. D. C. Clary, E. Buonomo, I. R. Sims, I. W. M. Smith, W. D. Geppert, C. Naulin, M. Costes, L. Cartechini, and P. Casavecchia, J. Phys. Chem. A **106**, 5541 (2002).
147. R. I. Kaiser, C. Ochsenfeld, M. Head-Gordon, Y. T. Lee, and A. G. Suits, Science **274**, 1508 (1996).
148. T. Furtenbacher, I. Szabó, A. G. Császár, P. F. Bernath, S. N. Yurchenko, and J. Tennyson, ApJS **224**, 44 (2016).
149. A. G. G. M. Tielens, Annu. Rev. Astron. Astrophys. **46**, 289 (2008).
150. C. M. R. Rocha and A. J. C. Varandas, Phys. Chem. Chem. Phys. **21**, 24406 (2019).
151. C. A. Gottlieb, J. M. Vrtilek, E. W. Gottlieb, P. Thaddeus, and A. Hjalmarson, ApJ **294**, L55 (1985).
152. S. Yamamoto, S. Saito, H. Suzuki, S. Deguchi, N. Kaifu, S.-I. Ishikawa, and M. Ohishi, ApJ **348**, 363 (1990).
153. M. Kanada, S. Yamamoto, S. Saito, and Y. Osamura, J. Phys. Chem. **104**, 2192 (1996).
154. M. Caris, T. Giesen, C. Duan, H. Muller, S. Schlemmer, and K. Yamada, J. Mol. Spectrosc. **253**, 99 (2009).
155. M. C. McCarthy, K. N. Crabtree, M.-A. Martin-Drumel, O. Martinez, B. A. McGuire, and C. A. Gottlieb, ApJS **217**, 10 (2015).
156. Q. Jiang, C. M. L. Rittby, and W. R. M. Graham, J. Chem. Phys. **99**, 3194 (1993).
157. S. M. Sheehan, B. F. Parsons, J. Zhou, E. Garand, T. A. Yen, D. T. Moore, and D. M. Neumark, J. Chem. Phys. **128**, 034301 (2008).
158. S. Yamamoto and S. Saito, J. Chem. Phys. **101**, 5484 (1994).
159. H. Yamagishi, H. Taiko, S. Shimogawara, A. Murakami, T. Noro, and K. Tanaka, Chem. Phys. Lett. **250**, 165 (1996).
160. J. Takahashi and K. Yamashita, J. Chem. Phys. **104**, 6613 (1996).
161. K. Aoki, S. Ikuta, and A. Murakami, J. Mol. Struct.: THEOCHEM **365**, 103 (1996).
162. C. Ochsenfeld, R. I. Kaiser, Y. T. Lee, A. G. Suits, and M. Head-Gordon, J. Chem. Phys. **106**, 4141 (1997).
163. R. K. Chaudhuri, S. Majumder, and K. F. Freed, J. Chem. Phys. **112**, 9301 (2000).
164. J. C. Sancho-García and A. J. Pérez-Jiménez, J. Phys. B At. Mol. Opt. **35**, 3689 (2002).
165. J. F. Stanton, Chem. Phys. Lett. **237**, 20 (1995).
166. J. C. Saeh and J. F. Stanton, J. Chem. Phys. **111**, 8275 (1999).
167. R. A. Kendall, T. H. Dunning Jr., and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).
168. P. Halvick, Chem. Phys. **340**, 79 (2007).
169. M. K. Bassett and R. C. Fortenberry, J. Chem. Phys. **146**, 224303 (2017).
170. S. C. Bennedjai, D. Hammoutène, and M. L. Senent, ApJ **871**, 255 (2019).

171. M. Perić, M. Mladenović, K. Tomić, and C. M. Marian, J. Chem. Phys. **118**, 4444 (2003)
172. H. Ding, T. Pino, F. Güthe, and J. P. Maier, J. Chem. Phys. **115**, 6913 (2001).
173. A. M. Mebel and R. I. Kaiser, Chem. Phys. Lett. **360**, 139 (2002).
174. Y. Wang, B. J. Braams, and J. M. Bowman, J. Phys. Chem. A **111**, 4056 (2007).
175. A. J. C. Varandas, J. Chem. Phys. **138**, 054120 (2013).
176. A. J. C. Varandas, J. Chem. Phys. **138**, 134117 (2013).
177. A. J. C. Varandas, in *Reaction Rate Constant Computations: Theories and Applications* (The Royal Society of Chemistry, U.K., 2014), vol. 17, chap. Putting together the pieces: global description of valence and long-range forces via combined hyperbolic inverse-power representation of the potential energy surface for use in reaction dynamics, pp. 408–445.
178. C. M. R. Rocha and A. J. C. Varandas, J. Phys. Chem. A **123**, 8154 (2019).
179. S. Joseph and A. J. C. Varandas, J. Phys. Chem. A **114**, 2655 (2010).
180. C. M. R. Rocha and A. J. C. Varandas, Chem. Phys. Lett. **700**, 36 (2018).
181. R. S. Urdahl, Y. Bao, and W. M. Jackson, Chem. Phys. Lett. **178**, 425 (1991).
182. B. E. Sagan, *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions* (Springer, New York, 2001).
183. M. Karplus, Folding & Design **2**, S69 (1997).
184. H. Wang and R. P. A. Bettens, Phys. Chem. Chem. Phys. **21**, 4513 (2019).
185. M. Patrício, J. L. Santos, F. Patrício, and A. J. C. Varandas, J. Math. Chem. **51**, 1729 (2013).
186. C. M. Rocha, J. Li, and A. J. C. Varandas, J. Phys. Chem. A **123**, 3121 (2019).
187. F. O. Ellison, J. Am. Chem. Soc. **85**, 3540 (1963).
188. J. C. Tully, Adv. Chem. Phys. **42**, 63 (1980).
189. P. J. Kuntz, in *Atom-Molecule Collision Theory*, edited by R. Bernstein (Plenum, New York, 1979), p. 79.
190. K. S. Pitzer and E. Clementi, J. Am. Chem. Soc. **81**, 4477 (1959).
191. A. Van Orden and R. J. Saykally, Chem. Rev. **98**, 2313 (1998).
192. A. J. C. Varandas and F. N. N. Pansini, J Chem. Phys. **150**, 154106 (2018).
193. A. J. C. Varandas and F. N. N. Pansini, International Journal of Quantum Chemistry **120**, e26135 (2020)
194. C. M. Rocha and A. J. C. Varandas, Comp. Phys. Comm. **247**, 106913 (2020).