Computational strategies for the accurate thermochemistry and kinetics of gas-phase reactions

Settore Scientifico Disciplinare CHIM/02

Candidato
dr. Jacopo Lupi

Relatore

Prof. Vincenzo Barone

Computational strategies for the accurate thermochemistry and kinetics of gas-phase reactions

by

Jacopo Lupi

Abstract

This PhD thesis focuses on the theoretical and computational modeling of gas phase chemical reactions, with a particular emphasis on astrophysical and atmospherical ones. The ability to accurately determine the rate coefficients of key elementary reactions is deeply connected to the accurate determination of geometrical parameters, vibrational frequencies and, even more importantly, electronic energies and zero-point energy corrections of reactants, transition states, intermediates and products involved in the chemical reaction, together with a suitable choice of the statistical approach for the rate computation (i.e. the proper transition state theory model). The main factor limiting the accuracy of this process is the computational time requested to reach meaningful results (i.e. reaching subchemical accuracy below $1\,\mathrm{kJ\,mol^{-1}}$), which increases dramatically with the the size of the system under investigation. For small-sized systems, several nonempirical procedures have been developed and presented in the literature. However, for larger systems the well-known model chemistries are far from being parameter-free since they include some empirical parameters and employ geometries which are not fully reliable for transition states and noncovalent complexes possibly ruling the entrance channels. Based on these premises, this dissertation has been focused on the development of new "cheap" composite schemes, entirely based on the frozen core coupled cluster ansatz including single, double, and (perturbative) triple excitation calculations in conjunction with a triple-zeta quality basis set, including the contributions due to the extrapolation to the complete basis set limit and core-valence effects using second-order Møller-Plesset perturbation theory. For the first time the "cheap" scheme has been extended to explicitly-correlated methods, which have an improved performance with respect to their conventional counterparts. Benchmarks with different sets of state of the art energy barriers, interaction energies and geometrical parameters spanning a wide range of values show that, in the absence of strong multireference contributions, the proposed models outperforms the most well-known model chemistries, reaching a subchemical accuracy without any empirical parameter and with affordable computer times. Besides the composite schemes development efforts, a robust protocol for disclosing the thermochemistry and kinetics of reactions of atmospheric and astrophysical interest, rooted in the so-called ab initio-transition-state-theory-based master equation approach have been thoroughly investigated and validated.

Publications

Main author

1. J. Lupi, S. Alessandrini, C. Puzzarini, and V. Barone. junChS and junChS-F12 models: parameter-free efficient yet accurate composite schemes for energies and structures of noncovalent complexes. *J. Chem. Theory Comput.*, 17(11): 6974–6992, 2021

2. J. Lupi, C. Puzzarini, and V. Barone. Methanimine as a key precursor of imines in the interstellar medium: the case of propargylimine. *Astrophys. J. Lett.*, 903(2):L35, 2020

3. J. Lupi, C. Puzzarini, C. Cavallotti, and V. Barone. State-of-the-art quantum chemistry meets variable reaction coordinate transition state theory to solve the puzzling case of the $H_2S + Cl$ system. *J. Chem. Theory Comput.*, 16(8): 5090–5104, 2020

4. J. Lupi, M. Martino, A. Salvadori, S. Rampino, G. Mancini, and V. Barone. Virtual reality tools for advanced modeling. *AIP Conf. Proc.*, 2145(1):020001, 2019

Co-author

1. C. Baiano, J. Lupi, N. Tasinato, and V. Barone. Gliding on ice in search of accurate and cost-effective computational methods for astrochemistry on grains: the puzzling case of the HCN isomerization. *J. Chem. Theory Comput.*, XX(XX):XXXX, 2022

2. V. Barone, J. Lupi, Z. Salta, and N. Tasinato. Development and validation of a parameter-free model chemistry for the computation of reliable reaction rates. *J. Chem. Theory Comput.*, 17(8):4913–4928, 2021

3. F. Tonolo, J. Lupi, C. Puzzarini, and V. Barone. The quest for a plausible formation route of formyl cyanide in the interstellar medium: a state-of-the-art quantum-chemical and kinetic approach. *Astrophys. J.*, 900(1):85, 2020

4. C. Baiano, J. Lupi, N. Tasinato, C. Puzzarini, and V. Barone. The role of state-of-the-art quantum-chemical calculations in astrochemistry: formation route and spectroscopy of ethanimine as a paradigmatic case. *Molecules*, 25 (12):2873, 2020

5. C. Puzzarini, Z. Salta, N. Tasinato, J. Lupi, C. Cavallotti, and V. Barone. A twist on the reaction of the CN radical with methylamine in the interstellar medium: new hints from a state-of-the-art quantum-chemical study. *Mon. Not. R. Astron. Soc.*, 496(4):4298–4310, 2020

6. Z. Salta, N. Tasinato, J. Lupi, R. Boussessi, A. Balbi, C. Puzzarini, and V. Barone. Exploring the maze of $C_2N_2H_5$ radicals and their fragments in the interstellar medium with the help of quantum-chemical computations. *ACS Earth Space Chem.*, 4(5):774–782, 2020

7. Z. Salta, J. Lupi, V. Barone, and O. N. Ventura. H-abstraction from dimethyl sulfide in the presence of an excess of hydroxyl radicals. A quantum chemical evaluation of thermochemical and kinetic parameters unveils an alternative pathway to dimethyl sulfoxide. *ACS Earth Space Chem.*, 4(3):403–419, 2020

8. Z. Salta, J. Lupi, N. Tasinato, V. Barone, and O. N. Ventura. Unraveling the role of additional OH-radicals in the H–abstraction from dimethyl sulfide using quantum chemical computations. *Chem. Phys. Lett.*, 739:136963, 2020

THIS PAGE INTENTIONALLY LEFT BLANK

Contents

THIS PAGE INTENTIONALLY LEFT BLANK

Introduction

Gas phase reactions play a major role in many different environments. Understanding the reaction pathways that lead to the formation and dissociation of species present in planetary systems, with Earth being a special case and in the interstellar medium (ISM), is of paramount importance.

The comprehension of processes ruling the evolution of our atmosphere is essential since they are related, for example, to global climate changes or to the build-up of chemical complexity in space. The interest for atmospheric chemistry has grown considerably in the last years because it is now recognized that chemistry is at the heart of a wide range of atmospheric phenomena including acid rain, ozone depletion, air pollution, the atmospheric transport, conversion and deposition of pollutants such as mercury and perfluorinated species, greenhouse gas budgets, determining the impact of biofuels and biogenic emissions, and cloud formation. [13]

Shifting the focus on ISM and exoplanetary atmospheres, since the early Sixties the discovery of new molecules in these environments has continued at a nearly steady pace. However, in the last decade, the rate of detection of new interstellar complex organic molecules (iCOMs) has been further accelerating. The importance of these species (containing at least six atoms including carbon) lies in the evidence, accumulated in the last decades, that iCOMs play the key role of precursors of biochemical building blocks, either in outer space or in planetary atmospheres resembling that of the primordial Earth.

The accurate computation of the formation rates of these species at reduced computational cost is crucial in order to use them within global kinetic schemes. Among the theoretical tools commonly used to do that, we recall quantum dynamics calculations,[14] Multi-Configuration Time-Dependent Hartree (MCTDH),[15] Ring Poly-

mer Molecular Dynamics (RPMD)[16] and Quasi-Classical Trajectory calculations (QCT).[17,18] The main drawback of these methods is the high computational time requested already with medium sized systems (more than 3-4 atoms for quantum dynamics and more than 10-12 atoms for the others) since they require the knowledge of the whole reactive potential energy surface (PES). Because of their limited applicability, nowadays one of the most powerful tools of chemical kinetics is the transition state theory (TST). The broad success of the theory can be understood looking at its simplicity and efficiency. Indeed TST only requires the knowledge of few critical points of the PES, making it applicable also to larger chemical systems. Moreover, its widespread use arises also from the fact that electronic structure calculations can nowadays be carried out by the majority of commercial and free codes in an almost black-box manner. For all the PES stationary points some key properties are needed: equilibrium geometries, electronic energies, harmonic (or anharmonic) frequencies and zero-point energies. Ultimately, the quality of the rate constant depends on two fundamental factors: the accuracy of the electronic structure calculations and the sophistication of the transition state theory.

Let us start by discussing the former point. The most crucial aspect is the computation of accurate values for all the energy barriers ruling the different elementary steps of the PES. It is worthwhile to recall that chemists usually aim to the so called chemical ($4\,\mathrm{kJ\,mol^{-1}}$) or even sub-chemical accuracy. To do that it is essential to recover electron correlation energy, i.e. to go beyond Hartree-Fock electronic energy. Following Lowdin,[19] the electron correlation energy is defined as *the difference between the exact eigenvalue of the Hamiltonian and its expectation value in the Hartree-Fock approximation for the state under consideration*, in the non relativistic approximation. Several procedures have been developed and employed to recover the maximum possible amount of electron correlation in order to generate accurate thermochemical data, which for small systems come close to the full configuration interaction (FCI) complete basis set (CBS) limit.[20] Among the most successful approaches are the Weizmann-n series (with the most accurate being W4),[21] the focal point analysis (FPA),[22,23] the Feller–Dixon–Peterson model (FDP),[24] and the extrapolated ab initio thermochemistry (HEAT) protocol.[25–27] A simplified version of

the HEAT protocol is obtained by retaining only the extrapolation to the CBS limit at the CCSD(T) level and incorporating the core-valence corrections, thus leading to the model referred to usually as CBS-CV. This approach is rather well tested in the literature and was shown to provide results with an accuracy well within $2\,\mathrm{kJ\,mol^{-1}}$ for systems not showing excessive multireference character. Recently, alternative protocols have been proposed, which employ explicitly correlated approaches:[20,28,29] thanks to the faster convergence to the complete basis set limit, these approaches allow some computer time saving, but the rate-determining step remains the evaluation of higher-level contributions. For larger molecular systems, more approximate composite methods are unavoidable, which aim at reaching the chemical accuracy. The most well-known among these so-called model chemistries are the last versions of the Gn[30] (G4[31]) and CBS-x[32] (CBS-QB3[33]) families. However, all of these models include some empirical parameters and employ geometries, which are not fully reliable for transition states and noncovalent complexes ruling the entrance channels of most reactions of astrochemical and atmospheric interest. As a matter of fact, the most reliable protocols (e.g., HEAT) push geometry optimizations to the limit to obtain accurate energetics, whereas, at the other extreme, Gn and CBS-x schemes employ B3LYP geometries, whose accuracy is often unsatisfactory.[34]

On the other hand, the choice of the most suitable transition state theory to describe the system under examination is crucial. The transition state concept was first proposed in 1935 by Eyring[35] and Evans and Polanyi[36] and relies on the idea that there exists a transition state (TS) or "activated complex" that separates the reactants from the products. The main idea behind TST is that the reactants will move in the configurational space under the appropriate laws of motion, but the reaction will take place only if the system reaches the TS geometry. In fact, this configuration is characterized by a maximum along the reaction coordinate, but it is a minimum along all the other motions orthogonal to it, and therefore it is a first order saddle point along the multidimensional PES. A necessary condition for the successful application of TS is that the reaction must be elementary, in the sense specified by the IUPAC Goldbook[37] i.e. *a reaction for which no reaction*

intermediates have been detected or need to be postulated in order to describe the chemical reaction on a molecular scale. Almost every chemical reaction of practical interest consists of a network of elementary steps, each giving its own contribution to the global kinetics, where each reactant and product of each step is the product and the reactant of another elementary reaction forming a chemical pathway.

It can be proved that the standard TST reactive flux is an *upper bound* for the true classical reactive flux.[38] This means that the variational principle can be employed in order to search for the best estimate of the rate constant. Within the validity range of the so called *non re-crossing assumption*, which means that all the trajectories in the phase space which pass through the diving surface never come back, and the Boltzmann distribution of the reactants, TST provides the exact classical rate constant. Unfortunately these hypotheses are often violated and a bunch of additional problematics turn out. In particular the non re-crossing assumption is the most limiting one, as it is not always verified. All the methodologies which aim to overcome this limitation fall into the broad class of *variational transition state theory* (VTST).[39] The thermal reactant distribution assumption is also critical. Indeed, in some fast reactions, energized molecules can react more rapidly than the rate at which high energy states are thermalized by collision with the environment. Moreover the assumption that the dynamics proceeds on a single Born-Oppenheimer surface leaves out all the spin forbidden reactions, e.g. all the singlet to triplet ones, while the classic treatment of nuclei dynamics excludes all those reactions characterized by quantum effects, such as tunneling and non classical reflection. Other limitations, instead, come from the definition of the reactants and TS partition functions and in particular on how we decide to treat the internal modes of vibration of the molecules. Commonly, within the TST approach, the internal motions of the molecule are treated with the harmonic oscillator approximation (HO), which is not always legitimate. For example, in case of torsional vibrations, characterized by very low frequencies, the so called hindered rotors cannot be treated as harmonic oscillators.

There are actually two types of possible transition states: loose and tight. A loose transition state is made up of fragments that rotate freely or nearly freely in

more than one dimension with regard to one another. It should be noticed that the existence of torsions, which are one-dimensional internal rotations, is not enough to make a transition state loose. A typical example is provided by bond fission reactions without intrinsic barrier (i.e., when the potential energy is monotonically uphill or monotonically downhill in the reverse association), which exhibit typical loose transition states. The second class includes tight TSs, which rule reactions with non-negligible barriers. Furthermore, if a saddle point exists but is too low in energy (e.g., below the downhill direction reactants), the dynamical bottleneck might be a tight transition state around the saddle point, or a loose transition state leading to a well between the reactants and the saddle point. In fact, there are most likely two dynamical bottlenecks that should be treated with *ad hoc* theories. It should be clear that for loose transition states the standard methodologies cannot be applied, since it is not possible to identify a TS structure. A variety of different models to overcome this limitation have been proposed, from simplest ones, e.g. Phase Space Theory (PST),[40,41] to much more sophisticated ones, such as Variable Reaction Coordinate Transition State Theory (VRC-TST).[42–44]

Lastly, TST, as it has been sketched here does not take into account any pressure dependence of the rate coefficient. However, it has been demonstrated that for many gas-phase reactions, especially those involving small molecules, the rate coefficient can show a strong dependence on pressure. Finally, in order to successfully describe the global kinetics of multistep reactions (taking into account also the pressure dependence) the Master Equation (ME) approach came out as one of the most effective model and it will be used and discussed throughout this thesis.

Aim and strategy

This dissertation aims to develop a computational strategy for the accurate characterization of gas-phase reactive chemical systems. The strategy is based on two main pillars: the first one is the accurate calculation of the molecular structures and the thermodynamics of the species involved, i.e. the characterization of their reactive potential energy surfaces, while the second one is the kinetics calculation. As for

the former, the goal is to be able to obtain accurate results with contained computational times. This is done by developing composite schemes which can rivalry in accuracy with more resource demanding ones. As regards the latter, the calculation of reliable rate constants is done through the so called ab initio-transition-state-theory-based master equation approach (AITSTME). It is *ab initio* since it takes as input molecular properties calculated from first principles. These information are thus used in conjunction with TST coupled with a master equation approach in order to obtain phenomenological rate constants.

Hereafter the objectives of this thesis are defined and a related strategy to pursue them is sketched. In the introduction of each Chapter the proper objectives will be recalled.

O1. Define a methodology with moderate computational time in order to accurately describe geometries of covalent and non-covalent molecules. The robustness of the so developed methodology will be tested benchmarking its results with high level reference data.

O2. Extend the methodology to the calculation of activation energies. Again, this will be benchmarked against high level reference data.

O3. Define a simple yet reliable protocol to calculate reaction rate constants, taking as input molecular properties calculated with the aforementioned methodology, based on AITSTME. The protocol will be validated against experimental or high level theoretical results.

O4. Apply the computational strategy to paradigmatic test case reactions in order to disclose their thermochemistry and kinetics, thus getting insights on their role in atmospheric or astrophysical contexts.

Thesis outline

In this doctoral thesis, new developments of composite schemes for the accurate calculation of reactive PES will be presented. A general and robust strategy to

successfully calculate accurate rate constants for challenging gas-phase reactions will be also proposed. This thesis is structured as follows.

1. In Chapter 1 definitions and discussions of theoretical methodologies employed in this dissertation are given.

2. In Chapter 2 a new composite scheme based on explicitly-correlated approaches for the description of non-covalent complexes has been defined starting from the well-tested "cheap" version employing conventional methods. The results are presented in the paper "junChS and junChS-F12 Models: Parameter-free Efficient yet Accurate Composite Schemes for Energies and Structures of Non-covalent Complexes" (published, *Journal of Chemical Theory and Computation*).[1]

3. In Chapter 3 a slightly modified version of the junChS composite scheme is proposed and validated employing as benchmark different sets of state-of-the-art energy barriers. The result is a robust computational scheme which, combined with ME calculations, provides reliable reaction rates at reduced computational cost. The results are presented in the paper "Development and Validation of a Parameter-Free Model Chemistry for the Computation of Reliable Reaction Rates" (published, *Journal of Chemical Theory and Computation*).[6]

4. In Chapter 4 the general strategy outlined in the previous chapters is applied to the challenging reaction of hydrogen sulfide with chlorine radical. The investigation allows to present a robust approach for disclosing the thermochemistry and kinetics of reactions of atmospheric and astrophysical interest. The results are presented in the paper "State-of-the-Art Quantum Chemistry Meets Variable Reaction Coordinate Transition State Theory to Solve the Puzzling Case of the $H_2S + Cl$ System" (published, *Journal of Chemical Theory and Computation*).[3]

5. In Chapter 5 the same general strategy is applied to a challenging astrophysical test case: the formation of 2-propyn-1-imine (propargylimine) in the ISM.

The idea routes a general mechanism for the formation of complex imines in the ISM starting from methanimine as precursor. The results are presented in the paper "Methanimine as a Key Precursor of Imines in the Interstellar Medium: The Case of Propargylimine" (published, *The Astrophysical Journal Letters*).[2]

6. In Chapter 6, the focus is shifted to gas-grain reactivity on interstellar ices. In particular, a detailed study of the HCN/HNC isomerization by state-of-the-art quantum chemical methods and realistic cluster models is presented. The results are collected in the paper "Gliding on Ice in Search of Accurate and Cost-Effective Computational Methods for Astrochemistry on Grains: The Puzzling Case of the HCN Isomerization " (published, *Journal of Chemical Theory and Computation*).[5]

7. Finally, in Chapter 7, the conclusions are presented, together with possible future improvements to the models and computational methods developed in the thesis. Along with them, some preliminary results of undergoing works are shown.

All the chemical kinetics and explicitly correlated calculations have been done by the present author. Moreover, electronic structure calculations by means of Gaussian code, have been carried out in conjunction with co-authors, except in Chapter 6, where the author mainly focused on the kinetic modeling.

Chapter 1

Theoretical background

This Chapter introduces the theoretical fundamentals about electronic structure theory and methods for the determination of rate constants. While many of the discussed methodologies can be found in standard quantum chemistry[45–48] and chemical kinetics[49,50] textbooks, this is the opportunity to give a concise and coherent framework which helps to contextualise the computational approaches used in this dissertation. Hereafter, some basics of wavefunction theory and density functional theory are presented. An introduction to composite schemes is also given. Regarding chemical kinetics, transition state theory and the Rice-Ramsperger-Kassel-Marcus theory are discussed along with master equation approach to solve chemical reactions involving multiple, interconnected potential wells.

The molecular Hamiltonian Chemical reactions can only be deeply understood if we know what occurs to the electronic structure during the process. Quantum mechanics is the fundamental theory of physics that can explain the microscopical behaviour of matter. The starting point is the Schrödinger equation

$$\hat{\mathcal{H}} \left| \Psi(\mathbf{r}, t) \right\rangle = i\hbar \frac{\partial}{\partial t} \left| \Psi(\mathbf{r}, t) \right\rangle, \tag{1.1}$$

where $\hbar = \frac{h}{2\pi}$, with h Planck's constant, $\hat{\mathcal{H}}$ the Hamiltonian operator and $\left| \Psi \right\rangle$ the wavefunction which fully describes a state of the system. This linear partial differential equation describes the time evolution of a N particles system, with coordinates

$\mathbf{r} = \{\mathbf{r}_i\}$, $i = 1, \ldots, N$.

The wavefunction itself has no physical meaning, but $|\Psi(\mathbf{r}, t)|^2$ is the probability density of finding each particle at a given point and at a given time. Other physical properties can also be obtained from the wavefunction by applying operators, since $|\Psi\rangle$ contains all the information of the system. In the framework of quantum mechanics, any measurable property has an associated linear operator $\hat{A}$, which is Hermitian, i.e. $\hat{A}^\dagger = \hat{A}$. For example, in the case of N charged particles, $\hat{\mathcal{H}}$ becomes:

$$\hat{\mathcal{H}} = -\sum_{i=1}^{N} \frac{\hbar^2}{2m_i} \nabla_i^2 + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{q_i q_j}{4\pi\varepsilon_0 |\mathbf{r}_i - \mathbf{r}_j|}, \tag{1.2}$$

where the first term is the kinetic energy $\hat{T}$, with m_i being the mass of the i-th particle and ∇^2 being the Laplacian operator acting on the coordinates of the i-th particle, and the second one is the potential energy $\hat{V}$, including the sum of two-particle Coulomb's interactions.

Since Eq. 1.2 has no explicit dependence on time, the wavefunction can be factorized as $|\Psi(\mathbf{x}_i, t)\rangle = |\Psi(t)\rangle |\Psi_x(\{\mathbf{x}_i\})\rangle$, where $|\Psi_t(t)\rangle = A\exp(-iEt)$, i.e. it is a phase and therefore it carries no information on the physics of the system. Therefore one can keep only the space-dependent part, reaching the time-independent (non-relativistic) Schrödinger equation:

$$\hat{\mathcal{H}} |\Psi(\{\mathbf{x}_i\})\rangle = E |\Psi(\{\mathbf{x}_i\})\rangle, \tag{1.3}$$

from which the energy can be obtained as $E = \langle \Psi | \hat{\mathcal{H}} | \Psi \rangle$. Chemical systems consist of 2 different types of particles: electrons and nuclei. For a system with n electrons and N nuclei, the kinetic energy contributions can be separated into the electronic and the nuclear terms, $\hat{T}_e$ and $\hat{T}_n$. Regarding the potential energy term, there are crossed terms: electron-electron, nucleus-nucleus and electron-nucleus potential energy terms ($\hat{V}_{ee}, \hat{V}_{nn}, \hat{V}_{en}$). The last term couples electrons and nuclei, making the system not separable into electron-dependent and nuclei-dependent wavefunctions.

Thus, the molecular Hamiltonian (in atomic units) is:

$$\hat{\mathcal{H}} = -\sum_{i=1}^{n} \frac{1}{2}\nabla_i^2 - \sum_{k=1}^{N} \frac{1}{2M_k}\nabla_k^2 - \sum_{i=1}^{n}\sum_{k=1}^{N} \frac{Z_k}{|\mathbf{r}_i - \mathbf{R}_k|} + \sum_{i<j}^{n} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{k<l}^{N} \frac{Z_k Z_l}{|\mathbf{r}_k - \mathbf{R}_l|}.$$

(1.4)

where Z_k are the atomic numbers and $\mathbf{r}$ and $\mathbf{R}$ are the electronic and nuclear coordinates, respectively. The non-separability of the wavefunction into an electronic and a nuclear part limits the applicability of Eq. 1.4, mainly for hydrogenoid atoms (i.e. atoms composed by nuclei and one electron). However, molecules are multielectronic systems and some additional approximations have to be introduced.

The Born-Oppenheimer approximation   The Born-Oppenheimer approximation, sometimes called adiabatic approximation, is crucial in quantum chemistry. It finds its roots in to the fact that nuclei are much heavier than electrons (about 1800 times more) and therefore they move more slowly. In this scenario, the kinetic energy of the nuclei can be neglected and the nuclei-nuclei repulsion term can be considered to be a constant. The remaining terms form the so called electronic Hamiltonian, which describes the motion of N electron in the field of M clamped point charges

$$\hat{\mathcal{H}}_{elec} = -\sum_{i=1}^{n} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{n}\sum_{k=1}^{N} \frac{Z_k}{r_{ik}} + \sum_{i<j}^{n} \frac{1}{r_{ij}},$$

(1.5)

where $r_{ik} = |\mathbf{r}_i - \mathbf{R}_k|$ and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$. Within the approximation, the electronic dynamics can be considered separated by the nuclear one. The latter only appears as a parameter inside the electronic wavefunction, as can be seen by the following equation

$$\hat{\mathcal{H}}_{elec} |\psi(\mathbf{r};\mathbf{R})\rangle = E_{elec} |\psi(\mathbf{r};\mathbf{R})\rangle.$$

(1.6)

The $|\psi(\mathbf{r};\mathbf{R})\rangle$ are eigenfunctions of $\hat{\mathcal{H}}_{elec}$ and therefore they form an orthonormal basis set in the electronic Hilbert space, i.e.

$$\langle\psi_q(\mathbf{r};\mathbf{R})|\psi_s(\mathbf{r};\mathbf{R})\rangle = \delta_{qs}.$$

Assuming to be able to solve Eq. 1.6 for each nuclear configuration $\{\mathbf{R}\}$, the

total molecular wavefunction $|\Psi(\mathbf{r}; \mathbf{R})\rangle$ can be expanded on the basis of $|\psi_q(\mathbf{r}; \mathbf{R})\rangle$

$$|\Psi(\mathbf{r}; \mathbf{R})\rangle = \sum_q \chi_q(\mathbf{R}) |\psi_q(\mathbf{r}; \mathbf{R})\rangle . \qquad (1.7)$$

Inserting Eq. 1.7 into Eq. 1.6 and projecting onto the ψ_q basis set (integrating only on electronic coordinates), we can evaluate how the different terms of $\hat{\mathcal{H}}$ act on $\chi_q(\mathbf{R})$ and $\psi_q(\mathbf{r}; \mathbf{R})$. In particular the kinetic energy operator of nuclei $-\sum_k \frac{\hbar^2}{2M_k}\nabla_k^2$ is not diagonal on the $|\psi_q(\mathbf{r}; \mathbf{R})\rangle$ basis. The Born-Oppenheimer approximation now consists of neglecting the off diagonal terms of this operator and therefore decoupling the electronic dynamics from the nuclear one. In general, this assumption is an extremely mild one, and it is entirely justified in most cases. It is worthwhile emphasizing that this approximation has very profound consequences from a conceptual point of view: without the Born–Oppenheimer approximation, for example, we would lack the concept of a potential energy surface.

Pauli exclusion principle   Electrons are non-distinguishable particles. That is, any observable derived from the wavefunction, like their probability density, must be invariant given any permutation of the electronic coordinates:

$$|\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_n)|^2 = |\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_j, \ldots, \mathbf{r}_i, \ldots, \mathbf{r}_n)|^2$$

Introducing the permutation operator $\hat{P}$ of parity p, which permutes two electrons in the wavefunction, we can formalize the following: $\hat{P}|\Psi\rangle = (-1)^p|\Psi\rangle$. This is the Pauli exclusion principle, which states that a n-electrons wavefunction wavefunction must be antisymmetric with respect to the exchange of any two electrons. Any electronic (and more generally fermionic) wavefunction must obey this principle.

Variational principle and Hartree-Fock method   The variational principle is an important theorem in quantum chemistry, in which many computational techniques are rooted. Given a normalized trial wavefunction $|\Phi\rangle$ that satisfies the appropriate boundary conditions (usually that the wavefunction goes to zero at infinity), then the expectation value of the Hamiltonian is an upper bound to the

exact ground state energy. That is, if $\langle\Phi|\Phi\rangle = 1$ then $\langle\Phi|\hat{\mathcal{H}}|\Phi\rangle \geq E_0$. The equality holds when $|\Phi\rangle$ is identical to the ground state wavefunction $|\Psi_0\rangle$.

Now the question is how we can find a good trial wavefunction. At first, one could think about building-up the wavefunction as product of non interacting single electron spin-orbitals $|\chi(\mathbf{x})\rangle = |\Psi(\mathbf{r})\rangle\,|\omega(\sigma)\rangle$, where the spin function $|\omega(\sigma)\rangle$ can assume two values, $|\alpha\rangle$ or $|\beta\rangle$. However, this form does not respect the Pauli principle. A better choice is provided by the so called Slater determinant:

$$\Psi_{SD}(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} = |\chi_1\chi_2\ldots\chi_N|. \tag{1.8}$$

This is the starting point of the Hartree-Fock method (HF). Indeed, we want to calculate the quantity $\langle\Psi_{SD}|\hat{\mathcal{H}}|\Psi_{SD}\rangle$, that is

$$V_{NN} + \sum_{i=0}^{N_{elec}} h_i + \sum_{i=0}^{N_{elec}} \sum_{j>i}^{N_{elec}} J_{ij} - K_{ij}, \tag{1.9}$$

where h_i collects one-electron terms and $J_{ij} - K_{ij}$ two-electrons ones. These terms are defined as $h_i = \langle\chi_i|\hat{T}_e + \hat{V}_{en}|\chi_i\rangle$, $J_{ij} = \langle\chi_i\chi_j|r_{ij}^{-1}|\chi_i\chi_j\rangle$ and $K_{ij} = \langle\chi_i\chi_j|r_{ij}^{-1}|\chi_j\chi_i\rangle$. The two latter terms come from $\hat{V}_{ee}$ acting on the wavefunction, where r_{ij} is the distance between the i-th and j-th electrons. Here J_{ij} is the Coulomb integral and represent the electrostatic interaction between two charge densities $|\chi_i|^2$ and $|\chi_j|^2$ while K_{ij} , called the exchange integral, arises from the antisymmetric nature of the Slater determinant. Both J_{ij} and K_{ij} are positive, and therefore, if $i = j$ they vanish, preventing any self-interaction spurious effect which would come from J_{ii}. The core of the Hartree-Fock theory is to apply the variational method to this independent particle model. By proceeding this way, J_{ij} and K_{ij} integrals can be expressed as single-electron operators, defined by the spin-orbitals themselves:

$$J_j\chi_i(\mathbf{x}) = \chi_i(\mathbf{x}) \int \frac{|\chi_j(\mathbf{x}')|^2}{|\mathbf{x} - \mathbf{x}'|} d\mathbf{x}', \tag{1.10}$$

$$K_j \chi_i(\mathbf{x}) = \chi_j(\mathbf{x}) \int \frac{\chi_j^*(\mathbf{x}')\chi_i(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} d\mathbf{x}'. \tag{1.11}$$

In other words, it turns out that Hartree-Fock theory is a mean-field theory, i.e. each electron feels the averaged effect of the rest by means of these one-electron operators. Collecting these two terms plus h_i, one obtains the Fock operator, which may be defined for a given spin direction σ, either up or down (α, β), as:

$$\hat{F}^\sigma \left| \chi_i^\sigma \right\rangle = \left[\hat{h} + \sum_{j,\sigma'} \hat{J}_j^{\sigma'} - \delta_{\sigma\sigma'} K_j^\sigma \right] \left| \chi_i^\sigma \right\rangle = \varepsilon_i^\sigma \left| \chi_i^\sigma \right\rangle \tag{1.12}$$

where ε_i^σ are the energies associated with the i-th spin-orbtial with spin σ. These are the unrestricted HF (UHF) equations, in which each electron is allowed to occupy a different spatial orbital thanks to the separation into the electron's spin. Nevertheless the two sets of orbitals are still coupled by the Coulomb's term. For a closed-shell molecule, i.e. same number of spin α and β electrons, the previous equation becomes

$$\hat{F} \left| \chi_i \right\rangle = \left[\hat{h} + \sum_j \hat{J}_j - K_j \right] \left| \chi_i \right\rangle = \varepsilon_i \left| \chi_i \right\rangle , \tag{1.13}$$

in which each spin-orbital is filled by two electrons, one spin up and one spin down. This gives rise to the restricted HF (RHF) theory. One could also take a midway by forcing orbitals to behave as RHF as far as possible and the remaining ones as UHF, which is the restricted open-shell HF (ROHF) method. The total RHF and ROFH wavefunctions are also eigenfunctions of the total squared spin operator $\hat{S}^2$, while the UHF ones are not.

Electron correlation However, the Hartree-Fock method is far from giving the exact values of energy. HF misses some "electronic correlation" energy, as each electron only feels the averaged effect of the rest. This lack in electronic correlation energy is the energy required to reach the exact value:

$$E_{corr} = E_{exact} - E_{HF}. \tag{1.14}$$

Electron correlation can be classified in two ways: as Coulomb's and Fermi's correlation, or as static and dynamic correlation. The first classification accounts for the different interaction of electrons with parallel (Fermi correlation) or antiparallel (Coulomb correlation) spins. Taking RHF theory, due to the Pauli principle, a given spin-orbital can only be populated by two electrons with antiparallel spins. Therefore Coulomb correlation can take place between electrons in both the same spatial orbital and different ones. On the other hand, Fermi correlation can only happen between electrons in different orbitals, and consequently, it is smaller. The second classification (into dynamic and static correlation) is analogous to the previous one, but has some conceptual differences. Dynamic correlation takes into account the instantaneous correlation of electrons while the static one is associated to near-degenerate states (spin-orbitals of similar energy). The dynamic correlation is more important for electrons in the same orbital, while static one is more important between electrons in different orbitals. In this sense, UHF can retrieve more static correlation than RHF.

Møller-Plesset perturbation theory   One of the possible ways to recover electronic correlation is via perturbation theory, i.e. considering correlation as a small perturbation $\hat{\mathcal{H}}'$ to the reference HF Hamiltonian $\hat{\mathcal{H}}^{(0)}$. The perburbed Hamiltonian is

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}^{(0)} + \lambda \hat{\mathcal{H}}',$$

where λ is a dimensionless parameter and $\hat{\mathcal{H}}^{(0)} = \sum_i \hat{F}_i$. The perturbed Schrödinger equation to solve is then

$$(\hat{\mathcal{H}}^{(0)} + \lambda \hat{\mathcal{H}}') \left| \Psi \right\rangle = E \left| \Psi \right\rangle. \tag{1.15}$$

The eigenvalues and eigenfunctions of this equation are a continuous functions of λ and therefore can be expanded in series:

$$E(\lambda) = \sum_{k=0}^{\infty} \lambda^k E^{(k)}, \tag{1.16}$$

$$\Psi(\lambda) = \sum_{k=0}^{\infty} \lambda^k \left|\Psi^{(k)}\right\rangle, \tag{1.17}$$

where each term of the expansion is an order of the correction, and the terms of the wavefunction are orthogonal between them. Substituting Eq. 1.16 and 1.17 into Eq. 1.15 one obtains:

$$(\hat{\mathcal{H}}^{(0)} + \lambda\hat{\mathcal{H}}')(\left|\Psi^{(0)}\right\rangle + \lambda\left|\Psi^{(1)}\right\rangle + \lambda^2\left|\Psi^{(2)}\right\rangle + \ldots) = (E^{(0)} + \lambda E^{(1)} + \lambda^2 E^{(2)}$$
$$+ \ldots)(\left|\Psi^{(0)}\right\rangle + \lambda\left|\Psi^{(1)}\right\rangle + \lambda^2\left|\Psi^{(2)}\right\rangle + \ldots).$$

Then, collecting the terms by orders of λ:

$$\hat{\mathcal{H}}^{(0)}\left|\Psi^{(0)}\right\rangle = E^{(0)}\left|\Psi^{(0)}\right\rangle$$

$$\hat{\mathcal{H}}^{(0)}\left|\Psi^{(1)}\right\rangle + \hat{\mathcal{H}}'\left|\Psi^{(0)}\right\rangle = E^{(0)}\left|\Psi^{(1)}\right\rangle + E^{(1)}\left|\Psi^{(0)}\right\rangle,$$

$$\hat{\mathcal{H}}^{(0)}\left|\Psi^{(2)}\right\rangle + \hat{\mathcal{H}}'\left|\Psi^{(1)}\right\rangle = E^{(0)}\left|\Psi^{(2)}\right\rangle + E^{(1)}\left|\Psi^{(1)}\right\rangle + E^{(2)}\left|\Psi^{(0)}\right\rangle$$

and so on. This approach is commonly known as Møller-Plesset (MP) perturbation theory. In order to compute the successive corrections one may use the knowledge of the previous ones by using the energy of the unperturbed (zeroth order) equation. In the MP formalism it is just the sum of the energies of the occupied spin-orbitals (which counts the electron-electron twice). It can be shown that the n-th correction to the energy is $E^{(n)} = \langle\Psi^{(0)}|\hat{\mathcal{H}}'|\Psi^{(n-1)}\rangle$. Then, the first order correction removes the electron-electron interaction once, and therefore the sum of zeroth and first order is just the HF energy. Hence, MP does not recover any correlation until second order. Higher order solutions adds excited Slater determinants. However the expansion in excited determinants is normally truncated to a certain order, at least second. Møller-Plesset methods are labeled as MPn where n is the order of the correction where the expansion is truncated. The usual ones are MP2, MP3 and MP4. Since this family of methods are not variational, energy is not an upper bound to the real one, and therefore, it may actually be lower than the actual one.

Coupled cluster theory Coupled cluster (CC) theory was originally formulated for the quantum-chemical treatment of nuclear matter. After its introduction into electronic structure theory it became one of the most powerful schemes in quantum chemistry for the electron-correlation treatment and for high-accuracy computations.

CC theory uses an exponential ansatz for the wavefunction

$$|\Psi_{CC}\rangle = \exp(\hat{T})\,|\Psi_{HF}\rangle \tag{1.18}$$

where $\hat{T}$ is the so called cluster operator, which is an excitation operator and consists of the weighted sum of all excitations,

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \ldots \hat{T}_{N_{elec}}. \tag{1.19}$$

$\hat{T}_1 + \hat{T}_2 + \ldots$ denote the weighted sums of single, double, etc., excitations with the unknown parameters given by the weighting coefficients that are usually referred to as amplitudes. The chosen exponential ansatz in Eq. 1.18 ensures size-consistency and size-extensivity of the electron-correlation treatment even within a truncated scheme that does not include all excitations. CC theory, therefore, is, by construction, a size extensive approach. Let us discuss briefly the meaning of the two aforementioned terms. Size consistency of a method means that the energy of two molecules separated by an infinite distance is the same as the sum of the energies individually calculated on each, in formula $E_{AB} = E_A + E_B$ for $r \longrightarrow \infty$. Instead, size extensivity means that correlation energy scales correctly (linearly) with the size of the system.

Because of the exponential ansatz, the CC wavefunction is typically not determined via the variational principle. Instead, one uses a projection approach in which the CC wavefunction is inserted into the Schrödinger equation; the latter is then multiplied from the left with $\exp(-\hat{T})$, and an expression for the energy is obtained by projection onto the reference determinant

$$E = \langle\Psi_{HF}|\exp(-\hat{T})\hat{\mathcal{H}}\exp(\hat{T})|\Psi_{HF}\rangle \tag{1.20}$$

and nonlinear equations for the amplitudes are obtained by projection onto the excited determinants $|\Psi_Q\rangle$

$$0 = \langle\Psi_Q|\exp(-\hat{T})\hat{\mathcal{H}}\exp(\hat{T})|\Psi_{HF}\rangle\,, \tag{1.21}$$

which has to be solved for every possible $|\Psi_Q\rangle$. CC theory demonstrates its advantages only when used with a truncated cluster operator. The usual choices are $\hat{T} = \hat{T}_1 + \hat{T}_2$ (CC singles and doubles (CCSD)), $\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3$ (CC singles, doubles, triples (CCSDT)), and $\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \hat{T}_4$ (CC singles, doubles, triples, quadruples (CCSDTQ)), etc.

CC higher than CCSD are computationally very expensive, but third order excitations can be added perturbationally giving rise to the CCSD(T) method, often referred as the "gold standard" in quantum chemistry. Indeed, the triples contribution is computed separately through fourth order perturbation theory and added to the CCSD results. In addition, a term coming from the fifth order perturbation that describes the coupling between singles and triples is also included.

Explicitly correlated methods Electron correlation methods generally suffer from the problem of slow basis set convergence, and very large basis sets are often required to obtain converged results. The basis set problem arises from fact that the expansion in products of limited 1-electron functions (orbitals) does not describe well the shape of the wave function at small to intermediate values of interelectron distance r_{12}, which should satisfy the electronic wave function cusp condition:[51,52]:

$$\left.\frac{\partial\Psi}{\partial r_{12}}\right|_{r_{12}=0} = \frac{1}{2}\Psi(r_{12}=0). \tag{1.22}$$

This problem can be avoided by including terms in the wave function that depend explicitly on r_{12} and thus can describe the cusp properly. Early implementations used a linear R_{12} correlation factor, and a number of so-called R_{12}-methods were developed. It was found later that a Slater-type function $F_{12} \sim \exp(-\gamma r_{12})$ yields much better basis set convergence and numerical stability. This family of methods is called the F12 methods, and has been used widely in this dissertation. The γ

exponent is a parameter usually chosen in the input files. Various tests have shown that an optimal value for γ is especially dependent on the basis set size and varies for the most cases from 0.8 to 1.4. The value 1 has been found to be a good choice for most cases. In general it holds that larger values of γ are better for bigger basis sets.[53]

The general expression of the wavefunctions, that holds both for standard and explicitly correlated MP2 and CCSD, is

$$\Psi_{\mathrm{MP2(-F12)}} = (1 + \hat{T}_2)\Psi_{\mathrm{HF}} \tag{1.23}$$

$$\Psi_{\mathrm{CCSD(-F12)}} = e^{\hat{T}_1 + \hat{T}_2}\Psi_{\mathrm{HF}} \tag{1.24}$$

For explicitly correlated methods the double excitation operator $\hat{T}_2$ is extended by an additional term with the explicitly correlated amplitudes $\mathcal{T}_{\alpha\beta}^{ij}$. The single excitations, only present in Eq. 1.23, are not expanded because the explicit correlation acts on electron pairs:

$$\hat{T}_1 = \sum_{ia} t_a^i \hat{E}_i^a$$

$$\hat{T}_2 = \frac{1}{2}\left(\sum_{ijab} T_{ab}^{ij}\hat{E}_{ij}^{ab} + \sum_{ij\alpha\beta} \mathcal{T}_{\alpha\beta}^{ij}\hat{E}_{ij}^{\alpha\beta}\right)$$

where the explicitly correlated amplitudes are build by the projector $\hat{Q}_{12}$ and the correlation factor F_{12} as

$$\mathcal{T}_{\alpha\beta}^{ij} = \sum_{mn} T_{mn}^{ij} \langle mn|F_{12}\hat{Q}_{12}|\alpha\beta\rangle .$$

The indexes $i, j, k, l, \ldots$ identify occupied orbitals, a, b the virtual orbitals while α, β the complete virtual orbital basis. The $\hat{Q}_{12}$ operator ensures orthogonality between the explicitly correlated part and the reference one. In order not to overburden the mathematical notation, we refer readers to an elegant and clear review by Liguo and coworkers for the F12 equations derivation.[29]

The meaning of some F12 related keywords often encountered in this thesis is hereafter clarified. In order to calculate some blocks of the Fock matrix and exchange

operators in the MP2-F12 method, it is common to rely on density fitting (DF) which allows to reduce the associated computational cost. The DF approach needs the use of an additional basis set. Resolution of Identity (RI) is used to approximate three- and four-electron integrals in terms of only one- and two-electron integrals. This approximation becomes exact if an infinite basis set is used. Since in practice any finite basis used will not be complete, the RI becomes an approximation. The original formulations of the RI approach utilized the orbital basis as the expansion basis. Such a choice was found to be inefficient, as these bases are generally optimized to best describe the occupied subspace; the virtual subspace often has significant overlap with the orthogonal complement of the space spanned by the orbital basis. As such, auxiliary bases were introduced specifically for the RI approximation. The most commonly used approach is that of the complementary auxiliary basis set (CABS), which is formed so as to span the orthogonal complement of the space spanned by the orbital basis. The CABS is then constructed as the union of the orbital and auxiliary bases. DF and RI auxiliary basis sets have the same functional form as the orbital basis sets, but different exponents, and they usually are largely uncontracted. Like the orbital basis sets, they are taken from basis set libraries.

Both MP2-F12 and CCSD(T)-F12 methods have a large number of ansatz and approximations.[54–56] The so called diagonal fixed amplitudes approximation (FIX) is the recommended one for MP2-F12: in that way the method becomes size consistent and orbital invariant and the use of localized orbitals is no longer a necessity. The amplitudes of the explicitly correlated part do not have to be calculated. The error introduced by this approximation depends on the system and the property calculated but is in general rather small. For CCSD-F12 two main approximations exist, namely the F12a and F12b. From a practical point of view, it has been noted that CCSD-F12a tends to overestimate the correlation energy, while CCSD-F12b is systematically below the basis set limit.

A final remark on perturbative triples correction is deserved, indeed explicit correlation has not been implemented for them yet. Therefore a trick is usually employed to speed-up the convergence in the (T) energy contribution with respect to the basis set size (for example in Molpro software package). The ratio of the

MP2-F12 and MP2 correlation energies is used to estimate a scaling factor for the perturbative triples correction, as follows

$$\Delta E_{(T^*)} = \Delta E_{(T)} \frac{E_{\text{MP2}-\text{F12}}^{\text{corr}}}{E_{\text{MP2}}^{\text{corr}}}.$$

This is based on the assumption that explicit correlation affects the triples energy in the same way that the MP2 correlation contribution does. This appears to be valid when considering that the same assumption is made in the basis set extrapolation, where the entire correlation contribution is extrapolated in the same way. The good results of the CCSD(T*)-F12a, which are very close to the CCSD(T) basis set limit, support this.

Density functional theory All the methods discussed above are wavefunction based methods, which rely on a very complex entity (Ψ) that depends on 4N variables (3N coordinates and N spins). Density functional theory (DFT), on the other hand, relies on a much simpler quantity: the electron density, which just depends on three spatial variables and may be derived from a wavefunction:

$$\rho(\mathbf{r}) = N \int \dots \int d\mathbf{x}_1 d\mathbf{x}_2 \dots \mathbf{x}_N \Psi^*(\mathbf{x}_1, \dots, \mathbf{x}_N) \Psi(\mathbf{x}_1, \dots, \mathbf{x}_N). \qquad (1.25)$$

DFT formalizes the electronic energy of a molecular system as a functional of the density

$$E[\rho] = V_{NN} + E_e[\rho] = V_{NN} + V_{eN}[\rho] + J_e[\rho] + T[\rho] + E_{xc}[\rho], \qquad (1.26)$$

where $V_{eN}[\rho] = -\sum_{k=0}^{N} Z_k \int \rho(\mathbf{r}) r_{ik}^{-1}$ is the electron-nucleus interaction, $J_e[\rho] = \frac{1}{2} \int \int \rho(\mathbf{r}_1)\rho(\mathbf{r}_2) r_{12}^{-1} d\mathbf{r}_1 d\mathbf{r}_2$. It may seem somewhat unnatural that the molecular Hamiltonian can be fully specified from a quantity that does not describe two-particle distributions but only one-particle ones. However, it can be intuitively shown that the system can be completely defined by the electron density in its ground state. The Hamiltonian is completely determined by the number of electrons and the potential V_{Ne}, usually called "external" potential. Both quantities can be derived from the electron density: its integral over the whole volume is the number

of electrons, the cusps of the density are the positions of nuclei, and their heights define the nuclear charges, these two, in turn, define V_{Ne}. Finally, the Hamiltonian determines the energy, the wave-function and the associated properties. Therefore, there must be a one-to-one correspondence of the electron density and the energy of the system. This was mathematically proved in the first Hohenberg and Kohn theorem, which ensures the existence of a universal energy functional of the electronic density. Additionally they showed in their second theorem that the energy can be obtained by means of the variational principle, starting with a trial density, $\tilde{\rho}$, so that $E[\tilde{\rho}] \geq E[\rho]$.[57] Furthermore, the unrestricted scheme can be introduced to DFT by simply splitting the total electron density into the electron density for α and β electrons ($\rho = \rho_\alpha + \rho_\beta$). The only problem is that, although $E[\rho]$ is proved to exist and to be universal, it is unknown. The reason lies in the $T[\rho]$ and $J_e[\rho]$ functionals that cannot be directly written in terms of $\rho(\mathbf{r})$. If they were known, DFT would return the exact ground-state energy. To overcome this, Kohn and Sham ingeniously considered a fictitious systems, of non-interacting electrons moving within an "external" potential (V_{eN}).[58] Such a system can be exactly described by a Slater determinant made up by auxiliary Kohn-Sham (KS) spin-orbitals. As a constraint, it is then required that the electron density derived from these auxiliary functions is the same as the one of DFT i.e., $\rho_{KS}(\mathbf{r}) = \sum_i \int |\chi_i(\mathbf{x})|^2 d\sigma = \rho(\mathbf{r})$. The (small) difference in kinetic energy between the real and the fictitious ($T_s[\rho]$) system is added to the exchange-correlation term, which becomes $E_{xc}[\rho] = E'_{xc}[\rho] + T[\rho] - T_s[\rho]$. This $E_{xc}[\rho]$ is now the only unknown term. The design of different $E_{xc}[\rho]$ functionals has led to the rise of several hundreds of methods, which make possible the realization of DFT.

It is customary to separate $E_{xc}[\rho]$ into a term for exchange and another for correlation, $E_{xc}[\rho] = E_x[\rho] + E_c[\rho]$. Then each DFT method relies on certain approximations to compute each part. There are hundreds of DFT methods, but no systematic way to improve functionals. According to the fundamental ingredients in each method, Perdew and Schmidt proposed a classification into five families.[59] The resulting Jacob's ladder of DFT is a classification that connects the "hell" of non-interacting electrons to the "heaven" of chemical accuracy. The farther up the

ladder one goes, the more accurate the results are, but at the cost of additional computational effort. The ladder's steps are as follows, in order of increasing difficulty: (i) the local spin density approximation (LSDA) assumes that the electronic density is a slowly changing function of $\mathbf{r}$, so that the uniform electron gas model is able to describe the non-uniform density system locally, (ii) the generalized gradient approximation (GGA), in which corrections accounting for the non uniformity of ρ through the gradient of the electron density are introduced, (iii) the meta-GGA family, which adds higher order corrections based on higher order derivatives of the electronic density (and in some cases the kinetic energy from the auxiliary orbitals), (iv) the hybrid DFT methods, which improve the highly local nature of the previous families (they care about the electronic density at specific points and their close vicinity) by mixing some "exact" exchange energy (computed following the HF method with the KS auxiliary spin-orbitals), which is non local, to the GGA and meta-GGA rungs, and finally (v) one can also include some (dynamical) correlation estimated by perturbative methods like MP2, with this leading to double-hybrid methods, at the expenses of a higher computational cost.

A final remark is deserved regarding dispersion corrections. It is well known that LDA and GGA approximations predict exponential falloffs instead of the expected asymptotic r^{-6} behavior. Hybrids do not improve this picture, as the inclusion of HF-like exact exchange may end up in repulsive forces. Double-hybrids instead do improve the picture by the addition of correlation from the perturbative correction. However, usually DFT functionals are corrected in order to properly describe London dispersion forces. The most popular strategy is due to Grimme, and relies on its simplicity negligible additional computational cost. Indeed, the so called D3[60,61] and D3(BJ)[62] corrections, improve the descriptions of medium range interactions with respect to previous generation corrections, are less empirical and incorporate dependence on the chemical environment of each atom by accounting for the number of directly bonded atoms. The D3(BJ) version includes the Becke-Johnson damping function, which controls the overlap between short and long range interactions, as

the former is described by DFT. The equation for D3 correction is given by

$$\Delta E_{\text{disp}}^{\text{D3}} = -\frac{1}{2} \sum_{n=6,8} \sum_{A \neq B} s_n \frac{C_n^{AB}}{R_{AB}^n} f_{\text{damp}}(R_{AB}),$$

where s_n is a scaling factor which depends on the chosen functional, C_n^{AB} are the nth order dispersion parameters for each AB atom pair, R_{AB} are the AB inter-nuclear distances and

$$f_{\text{damp}}(R_{AB}) = \frac{1}{1 + \exp\left(-\gamma\left(\frac{R_{AB}}{s_{r,n}R_0^{AB}} - 1\right)\right)}.$$

are damping functions where R_{AB} are cut-off radii, $s_{r,6}$ are DFT functional dependent scaling factors, $s_{r,8}$ is set to 1 for all functionals and γ are constants, set to 14 for $n = 6$ and 16 for $n = 8$. Regarding D3(BJ) correction, the equation becomes

$$\Delta E_{\text{disp}}^{\text{D3(BJ)}} = -\frac{1}{2} \sum_{n=6,8} \sum_{A \neq B} s_n \frac{C_n^{AB}}{R_{AB}^n + f_{\text{damp}}^n(R_{AB}^0)}.$$

Here, $f(R_{AB}^0) = a_1 R_{AB}^0 + a_2$ where a_i are fit parameters (the BJ parameters) and $R_{AB}^0 = \sqrt{\frac{C_8^{AB}}{C_6^{AB}}}$. For either D3 and D3(BJ), the fit parameters ($s_{r,6}$ and s_8 for the former, a_1, s_8 and a_2 for the latter) are determined in least-squares fit to a set of 130 dispersion interaction energies.

Composite schemes   As already stated in the Introduction, composite schemes (also known as composite models or model chemistries) aim for high accuracy (at least $4\,\text{kJ}\,\text{mol}^{-1}$) by combining the results of several calculations. Usually, they combine methods with a high level of theory and a small basis set with methods that use lower levels of theory with larger basis sets. In this way it is possible to set up less computationally intensive calculations, concerning both CPU and disk storage. In this paragraph the composite schemes used in the dissertation will be formally presented and discussed. They are the so called HEAT scheme and the "cheap" family schemes, namely junChS and junChS-F12, developed and validated in this thesis. Along with them, also the "Best" scheme, again developed in the thesis, will be sketched.

HEAT   The HEAT scheme (which stands for High-accuracy Extrapolated Ab initio Thermochemistry) defines the total molecular energy in the following way: [27,63,64]

$$E_{\text{HEAT}} = E_{\text{HF}}^{\infty} + \Delta E_{\text{CCSD(T)}}^{\infty} + \Delta E_{\text{CCSDT}} + \Delta E_{\text{CCSDTQ}} + \Delta E_{\text{rel}} + \Delta E_{\text{ZPE}} + \Delta E_{\text{DBOC}} + \Delta E_{\text{SO}}.$$
(1.27)

First of all, the geometries of species are taken from optimizations carried out at the CCSD(T) level of theory with the correlation-consistent cc-pVQZ basis sets and ZPE are evaluated at the same level of theory (ΔE_{ZPE}). Then, HF-SCF and CCSD(T) correlation energies are obtained in a hierarchical series of basis sets, and then extrapolated separately to obtain estimates of the corresponding basis set limits. For the HF-SCF energy, calculations are carried out using the augmented correlation consistent basis sets aug-cc-pCVnZ, where $n = T, Q, 5$, which are designed to treat core correlation effects properly. These three energies were then extrapolated with the Feller's equation:

$$E_{\text{HF}}^{(n)} = E_{\text{HF}}^{\infty} + \frac{a}{e^{bn}}.$$

The parameters, a, b, and the extrapolated HF-SCF energy E_{HF}^{n} are determined uniquely from the three energies. The coupled cluster correlation energy is obtained through the following two-point extrapolation formula

$$\Delta E_{\text{CCSD(T)}}^{(n)} = \Delta E_{\text{CCSD(T)}}^{\infty} + \frac{a}{n^3},$$

employing the aug-cc-pCVQZ and aug-cc-pCV5Z basis sets. In order to push the theory beyond the CCSD(T) limit, higher level correlation effects have to be included, namely triples and quadruples excitations. Full triples contribution is computed as following

$$\Delta E_{\text{CCSDT}} = E_{\text{CCSDT}}^{\text{TQ}}(\text{fc}) - E_{\text{CCSD(T)}}^{\text{TQ}}(\text{fc}).$$

Here TQ denotes that the corresponding contribution has been obtained by the correlation energy two-points extrapolation formula using the frozen-core CCSDT and CCSD(T) energies obtained with the cc-pVTZ and cc-pVQZ basis sets. The

effects of full quadruples excitations are computed as

$$\Delta E_{\text{CCSDTQ}} = E_{\text{CCSDTQ}}(\text{fc}) - E_{\text{CCSDT}}(\text{fc})$$

employing a cc-pVDZ basis set.

Now, if more accuracy is sought, diagonal Born–Oppenheimer corrections have to be included. Indeed, electronic energy, as obtained from traditional electronic structure calculations, is not equivalent to the expectation value of the molecular Hamiltonian over the electronic (clamped nuclei) wave function. The difference lies in the contribution of the nuclear kinetic energy operator $\hat{T}_n$, which can be viewed as a first-order correction to the usual electronic energy. This correction is calculated by the expectation value

$$\Delta E_{\text{DBOC}} = \langle \Psi(\mathbf{r}; \mathbf{R}) | \hat{T}_n | \Psi(\mathbf{r}; \mathbf{R}) \rangle$$

at the HF-SCF level in conjunction with aug-cc-pVTZ basis set.

Spin-orbit correction ΔE_{SO} are included with a spin–orbit configuration interaction procedure: the core electrons are described by relativistic effective core potentials including spin–orbit terms that allow a straightforward calculation of the spin–orbit interaction integrals. Lastly, the so-called scalar relativistic contributions ΔE_{rel} are included by contracting the one-particle density matrix obtained at the CCSD(T)/ aug-cc-pCVTZ level with Darwin and mass velocity terms.

Because of its high computational cost, many cheaper schemes have been proposed by the authors. One of them, the so called HEAT-like (sometimes in this thesis referred as CBS-CVH), has been widely used in this dissertation and its definition can be found in Chapters 3 and 4.

junChS and junChS-F12   All the model chemistries belonging to the family of "cheap" schemes, are based on the same general idea, i.e. CCSD(T) with a triple ζ quality basis set in conjunction with CBS limit and additive CV corrections computed at MP2 level of theory. Both junChS and junChS-F12 rely on the use of partially augmented basis sets. Indeed, the starting point is the evaluation of molec-

ular geometries at the DFT level, employing the revDSD-PBEP86-D3(BJ) double hybrid functional in conjunction with the jun-cc-pVTZ basis set. The anharmonic ZPE is then evaluated through generalized second order vibrational perturbation theory. Concerning the electronic energy, junChS leads to the following expression

$$E_{\text{jChS}} = E_{\text{CCSD(T)}} + \Delta E_{\text{MP2}}^{\infty} + \Delta E_{\text{CV}}, \tag{1.28}$$

while junChS-F12 to

$$E_{\text{jChS}-\text{F12}} = E_{\text{CCSD(T)}-\text{F12a}} + \Delta E_{\text{MP2}-\text{F12}}^{\infty} + \Delta E_{\text{CV}-\text{F12}}. \tag{1.29}$$

$E_{\text{CCSD(T)}}$ and $E_{\text{CCSD(T)}-\text{F12}}$ are computed by jun-cc-pVTZ basis sets. Corrections for the CBS limit are computed respectively at the MP2 and MP2-F12 level of theory using the usual two-point extrapolation formula, here reported again for the sake of convenience

$$\Delta E_{\text{MP2}(-\text{F12})}^{(n)} = \Delta E_{\text{MP2}(-\text{F12})}^{\infty} + \frac{a}{n^3},$$

employing jun-cc-pVnZ basis sets, with $n = T, Q$. Finally, the core-valence correction is recovered at MP2(-F12) level of theory as the difference between all electron and frozen core calculations, i.e.

$$\Delta E_{\text{CV}(-\text{F12})} = E_{\text{MP2}(-\text{F12})}(\text{ae}) - E_{\text{MP2}(-\text{F12})}(\text{fc}), \tag{1.30}$$

using the cc-pwCVTZ basis set. The workflows for junChS and junChS-F12 schemes are reported in Fig. 1-1.
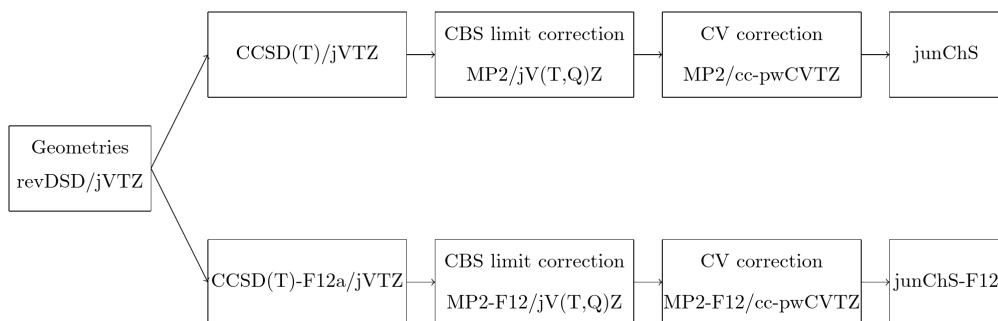

Figure 1-1: junChS and junChS-F12 workflows.

"Best" scheme In case more accuracy is sought, as stated in Chapter 2 and 3, the starting point are molecular geometries at F12 level, obtained as

$$r_{\text{F12}} = r_{\text{CCSD(T)-F12b}} + \Delta r_{\text{CV}}, \tag{1.31}$$

where $r_{\text{CCSD(T)-F12b}}$ are computed with jun-cc-pVTZ basis sets and Δr_{CV} is computed at the MP2-F12 level of theory in conjunction with cc-pwCVTZ. It can be noted that it is the analogue of the junChS-F12 scheme, discarding the MP2-F12 extrapolation term. On top of these geometries, electronic energies are usually calculated as

$$E_{\text{Best}} = E_{\text{CCSD(T)-F12b}} + \Delta E_{\text{CV-F12}} + \Delta E_{\text{CCSDT}} + \Delta E_{\text{CCSDT(Q)}} + \Delta E_{\text{rel}} + \Delta E_{\text{DBOC}}. \tag{1.32}$$

$E_{\text{CC-F12}}$ stands for CCSD(T)-F12b/aug-cc-pVQZ-F12. The core valence correction $\Delta E_{\text{CV-F12}}$ is computed as the CCSD(T)-F12b energy difference between all electron and frozen core calculations employing the cc-pwCVTZ basis set. The diagonal Born-Oppenheimer correction ΔE_{DBOC} and the scalar relativistic contribution to the energy ΔE_{rel} are computed at the HF-SCF/aug-cc-pVDZ and CCSD(T)/aug-cc-pCVDZ level. Finally, ΔE_{CCSDT} and $\Delta E_{\text{CCSDT(Q)}}$ are computed, within the fc approximation, as energy differences between CCSDT and CCSD(T) and between CCSDT(Q) and CCSDT calculations employing the aug-cc-pVTZ and aug-cc-pVDZ basis set, respectively.

Transition state theory For a single step reaction $\text{A} + \text{B} \longrightarrow \text{P}$, the rate of change of the abundances of either reactants and products is called the reaction rate law and its given by

$$\text{rate} = \frac{d[P]}{dt} = k[A][B], \tag{1.33}$$

where $[\ldots]$ are the abundances of species involved in the reaction and k is the reaction rate constant. The experimental relation for k is given by the Arrhenius equation:

$$k = A \exp\left(-\frac{E_a}{k_B T}\right) \tag{1.34}$$

where A is the pre-exponential or frequency factor (which may involve a small dependence on temperature) and E_a is the activation energy. If a reaction obeys the Arrhenius equation, then the Arrhenius plot ($\ln k$ versus $\frac{1}{T}$) should be a straight line with the slope and the intercept being $-\frac{E_a}{R}$ and A, respectively. The activation energy can be very roughly interpreted as the minimum energy (kinetic plus potential, relative to the lowest state of reactants) that reactants must have to form products and the pre-exponential factor is a measure of the rate at which collisions occur. A more precise interpretation of E_a was provided by Tolman[65,66] who showed that the Arrhenius energy of activation is the average total energy (relative translation plus internal) of all reacting pairs of reactants minus the average total energy of all pairs of reactants, including non-reactive pairs.

TST provides an expression of k in terms of statistical mechanics quantities that can be calculated for every molecular species. The main result of TST is the following expression for the thermal rate coefficient:

$$k = \frac{k_B T}{h} \frac{Z^{\ddagger}}{Z_R} \exp\left(-\frac{E_a}{k_B T}\right) \qquad (1.35)$$

where $Z^{\ddagger}$ and Z_R are the partition functions of the transition state and reactants, respectively. TST is based on some key assumptions:

1. there is chemical equilibrium between reactants and the "activated complex" (the high energy species formed corresponding to the saddle point, the TS)

2. once a molecule has reached the TS it may either come back to reactants or evolve into products

3. if it crosses the TS, the system can only evolve into products.

The canonical transition state theory can be applied also to reactions that present a loose transition state. In this case, as the molecular configuration of the transition state cannot be univocally identified, it is necessary to apply the variational principle to the computation of the rate coefficient. Indeed, the transition state corresponds in terms of reactive flux to the minimum flux of reactive molecules that pass from reactants to products per unit of time. This principle allows us to locate

the transition state, choosing the position along the reaction coordinate that minimizes the canonical rate coefficient. This procedure requires only a little more effort than the standard TST, as one repeats the calculation of k for a number of different transition states, until the position giving the minimum value for k is determined.

Rice-Ramsperger-Kassel-Marcus theory As we have seen in the previous paragraph, a molecule that actually undergoes reaction must have a high internal energy and at least it has to exceed the barrier existing between reactant and products. Molecules in the gas phase can increase their internal energy through excitation induced by collision with other molecules in the bath gas. Hence, a reaction is the combination of two steps: (i) the addition or removal of energy by collision, i.e. $A + M \longleftrightarrow A^* + M$, and (ii) the actual reaction event $A^* \longrightarrow$ products (where A is the reactant, M is the bath gas and * indicates internal excitation). During the twenties of the last century, Lindemann and Ramsperger supposed and verified experimentally that such a mechanism were able to explain the dynamics of gas-phase unimolecular reactions. The experimental evidence showed that $k(T)$ has a non-negligible dependence on pressure. Indeed, at high pressures, collisional activation and deactivation are very fast, hence the rate-determining step of the mechanism is the reaction event and as such, the overall rate coefficient will correspond to the rate coefficient of the reaction event. Since this does not involve the bath gas, the overall rate coefficient will be independent of pressure. At low pressure, the situation is the opposite. In fact, the collisional activation and deactivation is the rate-determining step and the overall rate coefficient is proportional to the bath gas pressure. In the intermediate region, the k follows a falloff regime.

In this description, the molecule A and its excited "version" A^* are considered as if they wouldn't have any internal states, which is obviously not true. In order to improve the model, the microscopic rates of collisional energy transfer from energy level E_j to another level with energy E_i, i.e. $R(E_i, E_j)$, and the microscopic rate of reaction $k(E_i)$ from a level with E_i have to be considered. Obviously, only molecules that reach an energy level higher than the activation energy, can react. The time

evolution of population $x_i(t)$ of reactant molecule can therefore be written as:

$$\frac{dx_i(t)}{dt} = [M] \sum_{j=0}^{\infty}(R_{ij}x_j - R_{ji}x_i) - k_i x_i(t). \tag{1.36}$$

This is the most basic form of the master equation. Solving this equation it is possible to extract information about the phenomenological rate coefficient, in other words the measurable rate constant of the considered reaction as a function of temperature and pressure. Nevertheless, the solution of this equation can be attempted only if information about $k(E_i)$ and the collisional energy transfer R_{ij} are known. Collisional energy transfer can be determined theoretically from classical trajectories simulations. This type of calculation, however, requires a significant computational effort and provides more information than what is actually needed.

In general, $k(E_i)$ is a function of not just the energy E, but also of the angular momentum J. This parameter, as for the collisional energy transfer can be obtained exactly from complete classical trajectories calculations. Nonetheless, the most commonly employed and widely accepted approximate description of the dynamics of the reaction step is Rice-Ramsperger-Kassel-Marcus (RRKM) theory, which gives the following fundamental equation for the rate constant

$$k(E) = \frac{N^{\ddagger}(E - E_0)}{h\rho(E)} = \frac{\int_0^{E-E_0} d\varepsilon^{\ddagger}\rho^{\ddagger}(E - E_0 - \varepsilon^{\ddagger})}{h\rho(E)}, \tag{1.37}$$

where $N^{\ddagger}(E - E_0)$ is the sum of states from 0 to $E - E_0$ for the TS (computed by excluding the normal mode with imaginary frequency under the assumption that the motion along the reaction coordinates is separable from that of the other modes), $\rho(E)$ is the density of states of reactants, E_0 is the energy barrier containing the vibrational ZPEs of reactants and the TS. The sum of states of the TS can be expressed as an integral of its density of states over the translational energy of the reaction coordinate $\varepsilon^{\ddagger}$.

Tunneling effects can be conveniently included by introducing the so called trans-

mission probability $T(\varepsilon^\ddagger)$

$$k(E) = \frac{\int_{-E_0}^{E-E_0} d\varepsilon^\ddagger T(\varepsilon^\ddagger) \rho^\ddagger (E - E_0 - \varepsilon^\ddagger)}{h\rho(E)}. \tag{1.38}$$

Note that now the integral is extended also to energies below the reaction threshold E_0.

Analytical expressions for $T(E)$ can be obtained by using model potential shapes. In this dissertation one of these, the unsymmetric Eckart potential, has been widely used.[67] In this case, the potential (i.e. the barrier shape) is modeled by the equation:

$$V(x) = \frac{Ae^{\tilde{a}x}}{1 + e^{\tilde{a}x}} + \frac{Be^{\tilde{a}x}}{(1 + e^{\tilde{a}x})^2}, \tag{1.39}$$

where $\frac{1}{\tilde{a}}$ is the length scale of the barrier, $A = V_1 - V_2$ and $B = (V^{\frac{1}{2}} + V^{\frac{1}{2}})^2$, with V_1 and V_2 the barriers from reactants to the TS and from products to the TS, so that $A \leq 0$ and $B > 0$. The transmission probability for such a barrier is given by:

$$T(E) = \frac{\cosh(a + b) - \cosh(a - b)}{\cosh(a + b) + \cosh(\sqrt{4\alpha_1\alpha_2 - \pi^2})}, \tag{1.40}$$

where

$$a = 2\frac{\sqrt{\alpha_1 \zeta}}{\frac{1}{\sqrt{\alpha_1}} + \frac{1}{\sqrt{\alpha_2}}},$$

$$b = 2\frac{\sqrt{(\zeta - 1)\alpha_1 + \alpha_2)}}{\frac{1}{\sqrt{\alpha_1}} + \frac{1}{\sqrt{\alpha_2}}},$$

$$\alpha_1 = \frac{2\pi V_1}{h\nu^\ddagger}$$

$$\alpha_2 = \frac{2\pi V_2}{h\nu^\ddagger}.$$

Here, $\zeta = \frac{E}{V_1}$, V_1 is the reactants to saddle point barrier, V_2 that of products to the saddle point one and $\nu^\ddagger$ the frequency of the transition state in absolute value.

Master equation   Let us take a deeper look into the master equation, already introduced with Eq. 1.36. A number of chemical reactions occur with non equilibrium energy distributions (non-Boltzmann), in other words the chemical reaction

time-scale is comparable to the collision time-scale. A successful approach that has been developed to model this class of reactions is the time dependent, multiple well master equation, in which a coarse-grained description of the energy levels accessible to the molecules is adopted. Considering the representation of a generic reactive
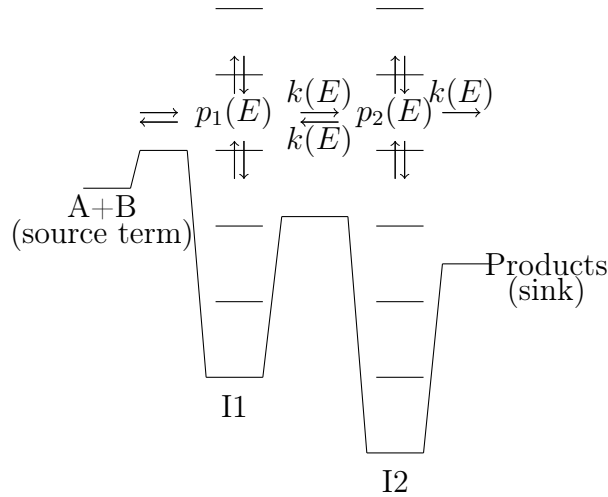

Figure 1-2: Representation of an energy grained master equation model for association reaction with two wells (intermediates), I1 and I2, and an irreversible product channel.

system illustrated in Figure 1-2, it is possible to formulate the master equation for a generic isomer m on a reactive PES, as:

$$
\begin{aligned}
\frac{dp_m(t)}{dt} =& \omega \int_0^\infty P(E|E')p_m(E')dE' - \omega p_m(E) + \sum_{n\neq m}^{M} k_{m,n}(E)p_n(E) - \sum_{n\neq m}^{M} k_{n,m}(E)p_m(E) \\
& - k_{S,m}(E)p_m(E) + K_{R,m}^{eq}k_{R,m}(E)\frac{\rho(E)e^{-\frac{E}{k_B T}}}{Z_m(T)}n_A p_B - k_{R,m}(E)p_m(E).
\end{aligned}
$$

$$(1.41)$$

Here, $p_m(E)$ is the rovibrational population density within a particular energy grain E, ω is the Lennard-Jones collision frequency, and $P(E|E')$ is the probability that collision with bath gas will result in a transition from a grain with energy E' to a grain with energy E. The right side of the equation has seven terms, three are positive and correspond to the flux into $p_m(E)$, while the other four terms are negative and correspond to outward fluxes. The first two terms are correlated to the gain or loss of the population density due to collisional energy transfer phenomena with

the bath gas. The second two terms are the inward flux coming from an adjacent intermediate n and the outward flux due to the conversion of the isomer m into another isomer n. The fifth term describes the reaction into the products and is considered if intermediate m is directly connected to the products. Finally, the last two terms are related to the bimolecular association source term and apply only to those isomers that are populated through a bimolecular association reaction, in which it is assumed that reactant A is in significant excess compared to reactant B. How this master equation is written, it does not represent a closed system of differential equations since p_B is not specified. Thus, if a bimolecular reaction is included in the reaction network, then it is necessary to include an additional equation:

$$\frac{dp_B}{dt} = \sum_{m=1}^{M} \int_0^\infty k_{R,m}(E)dE - n_A p_B \sum_{m=1}^{M} K_{R,m}^{eq} \int_0^\infty k_{R,m}(E)\frac{\rho(E)e^{-\frac{E}{k_B T}}}{Q_m(T)}dE. \quad (1.42)$$

In order to solve this set of coupled ordinary differential equations, it is possible to use essentially two different strategies: stochastic approaches or matrix diagonalization techniques. In the first family of master equation solvers, the Multiwell Program Suite of Barker and coworkers that employs Gillespie's exact stochastic method can be found.[68] The second one, includes, for example, the MESS code of Klippenstein and coworkers, which has been used throughout this thesis.[69]

How MESS works   The master equation calculation begins with the construction of the global relaxation matrix, which describes both chemical transformations (bimolecular to complex reactions and internal isomerizations) and collisional energy relaxation, followed by the determination of the relaxation matrix's eigenstates, and finally the extraction of the full set of phenomenological chemical rate coefficients from those. The calculations are done at microcanonical level. Although alternative expressions for the reactive flux can be incorporated through read options, the chemical transformations are generally described using RRKM theory. For evaluating the densities or numbers of states for barriers, wells, and bimolecular products, a variety of models are available. Similarly, a variety of model energy transfer kernels can be used to describe collisional energy relaxation. The one used in all the calculations

of this dissertation is the single exponential-down model,[70] i.e.

$$P(E|E') = C_N(E')^{-1} \exp\left(-\frac{E' - E}{\Delta E_{\text{down}}}\right),$$

for $E \leq E'$. $C_N(E')$ is a normalization factor and ΔE_{down} is the average energy transferred in a deactivating collision. In order to recover also the probability for upwards energy transfer, microscopic reversibility condition has to be imposed

$$f(E')P(E|E') = f(E)P(E'|E),$$

where $f(E)$ is the Boltzmann distribution function. Klippenstein and Miller's chemically significant eigenstate (CSE) approach provides the foundation for relating the eigenstates of the transition matrix to the phenomenological rate coefficients.[71,72] As described by Georgievskii *et al.*,[69] the particular formalism used involves an irreversible rather than reversible treatment of bimolecular species. The bimolecular products and reactants are treated on an equal footing in this method. The concentrations of sources and sinks are not included as variables in the global phase space. The global relaxation matrix is constructed for each temperature and pressure using appropriate discretization, and the chemically significant eigenstates are found, analyzed, and used to evaluate all of the phenomenological rate coefficients. More info about the code, along with source code and a manual, can be found at https://github.com/Auto-Mech/MESS.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 2

Composite schemes for molecular energies and geometries

The junChS-F12 model is proposed in this Chapter as an extension of the junChS approach[73] to explicitly-correlated F12 methods. A comprehensive benchmark investigation spanning a wide spectrum of intermolecular interactions shows that junChS-F12 is an effective, accurate, and parameter-free scheme for calculating accurate interaction energies. The suggested model rivals the accuracy of state-of-the-art composite approaches while having a much lower computational cost and scaling with the system dimensions. In fact, "ChS" stands for "cheap scheme" in the acronyms "junChS" and "junChS-F12," with "cheap" emphasizing the low computational cost. A reliable, accurate, and cost-effective reference geometry is obtained using the double-hybrid revDSD-PBEP86-D3(BJ) (hereafter revDSD) functional in combination with a triple-zeta basis set, which is an essential part of the junChS-F12 scheme. The application of composite schemes to highly accurate structural determinations of molecular complexes is also explored in this study. The accuracy of the traditional junChS model was demonstrated through a thorough validation procedure. When using F12 methods, the most efficient approach is to apply MP2-F12 core-valence correlation corrections to fc-CCSD(T)-F12/jun-cc-pVTZ geometries without recovering the basis set superposition or truncation errors. As a result of the validation and benchmarking process, the junChS and junChS-F12 databases are also presented. In this Chapter, objective O1 is pursued.

JCTC
Journal of Chemical Theory and Computation

Article

junChS and junChS-F12 Models: Parameter-free Efficient yet Accurate Composite Schemes for Energies and Structures of Noncovalent Complexes

Jacopo Lupi,[§] Silvia Alessandrini,[§] Cristina Puzzarini,* and Vincenzo Barone*

Read Online

ACCESS | 📊 Metrics & More | 📰 Article Recommendations | 🆂🅸 Supporting Information

ABSTRACT: A recently developed model chemistry (denoted as junChS [Alessandrini, S.; et al. *J. Chem. Theory Comput.* **2020**, *16*, 988−1006]) has been extended to the employment of explicitly correlated (F12) methods. This led us to propose a family of effective, reliable, and parameter-free schemes for the computation of accurate interaction energies of molecular complexes ruled by noncovalent interactions. A thorough benchmark based on a wide range of interactions showed that the so-called junChS-F12 model, which employs cost-effective revDSD-PBEP86-D3(BJ) reference geometries, has an improved performance with respect to its conventional counterpart and outperforms well-known model chemistries. Without employing any empirical parameter and at an affordable computational cost, junChS-F12 reaches subchemical accuracy. Accurate characterizations of molecular complexes are usually limited to energetics. To take a step forward, the conventional and F12 composite schemes developed for interaction energies have been extended to structural determinations. A benchmark study demonstrated that the most effective option is to add MP2-F12 core−valence correlation corrections to fc-CCSD(T)-F12/jun-cc-pVTZ geometries without the need of recovering the basis set superposition error and the extrapolation to the complete basis set.
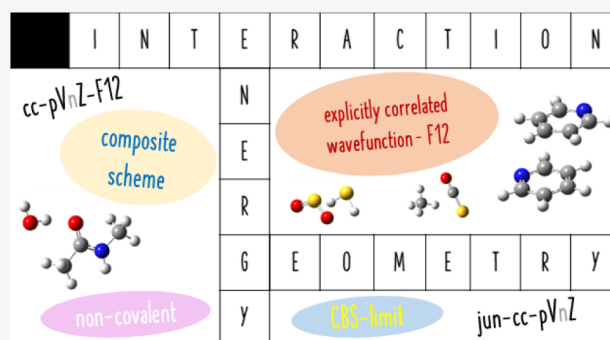
INTRODUCTION

The quantitative description of noncovalent interactions is of paramount relevance in widely different fields, ranging from biochemistry to nanoscience, going through chemical reactivity in the Earth atmosphere and in the interstellar space.[1−6] The reason lies in the fine tuning of geometric and energetic properties issuing from the formation of noncovalent bonds, which is due to the synergistic role played by electrostatic, induction, and dispersion contributions. Concerning biological model systems, noncovalent interactions are—for example—at the basis of DNA bases pairing and complexation with solvent molecules. Moving to reactivity, in the fields of astrochemistry and combustion chemistry, prereactive complexes are crucial points, whose properties are ruled by weak interactions. Furthermore, loose transition states also belong to the class of noncovalently bonded compounds.

Aiming at the definition of a general protocol able to accurately characterize noncovalent complexes at a non-prohibitive computational cost, the focus of this work is mainly on the interaction energy at the equilibrium geometry for medium-sized molecular systems containing up to about 20 atoms belonging to the first three rows of the periodic table. For intermolecular complexes that do not show strong static correlation effects, the most successful strategy is based on the

coupled cluster (CC) ansatz. In particular, the CC model that accounts for the full treatment of single and double excitations and the perturbative inclusion of triple excitations, CCSD(T),[7] has become the so-called "gold standard" in contemporary computational chemistry. As a matter of fact, a fortuitous but quite general error compensation makes CCSD(T) performances often better than those delivered by the model including the full treatment of triple excitations (CCSDT),[8] so that improvement beyond CCSD(T) results requires, together with the full treatment of triple excitations, at least perturbative inclusion of quadruple excitations (CCSDT(Q)).[9]

In order to establish a protocol able to combine accuracy and efficiency, the additivity approximation is adopted and leads to the definition of composite schemes, which aim at minimizing (i) the errors due to the truncation of the basis set and (ii) the *N*-electron error associated to the CCSD(T) method. While CCSD(T) can be considered to be affected by a small *N*-

electron error, the complete basis set (CBS) limit can be accurately estimated by means of extrapolation procedures based on hierarchies of basis sets.[10,11] These being computationally demanding already for medium-sized systems, the frozen-core (fc) approximation is usually employed with an additive incorporation of the core−valence (CV) correlation effects. Although CV contributions are usually small for interaction energies of molecules containing only second-row atoms, this is not the case for geometries and, when third-row atoms are present, also for energies.[12]

Focusing on molecular complexes, the accurate description of noncovalent interactions requires the presence of diffuse functions in the basis set (basis sets denoted as "augmented") because these interactions are particularly sensitive to the tails of the wave functions of the partners. In the last decade, a number of systematic investigations have shown that CCSD(T) computations in conjunction with "augmented" triple- and quadruple-ζ basis sets and followed by extrapolation to the CBS limit provide, in most cases, highly accurate results.[13−15] However, composite schemes entirely based on CCSD(T) calculations are computationally expensive and not affordable for larger systems, such as those of biological or prebiotic relevance. Among the approaches proposed to reduce the computational cost, we have introduced the so-called "cheap" scheme (hereafter ChS) that, on top of fc-CCSD(T) calculations in conjunction with a triple-ζ-quality basis set, includes the contributions due to the extrapolation to the CBS limit and CV effects using second-order Møller−Plesset perturbation theory (MP2).[16] This scheme, whose denomination is due to the strong reduction of the computational cost with respect to an approach entirely based on CC computations, is well tested for semirigid and rather flexible systems[17−22] and has been recently extended to incorporate partially augmented basis sets (junChS) in order to treat molecular complexes.[23]

To extend the range of applicability of the ChS approach, one can exploit the fact that the extrapolation to the CBS limit can benefit from explicitly correlated methods and resolution of identity, as well illustrated in the literature.[10,24] The rate-determining step of the junChS model is the CCSD(T) energy evaluation in conjunction with the jun-cc-pVTZ basis set. Since the use of a double-ζ-quality basis set fairly reduces the accuracy of conventional approaches, a possible solution is offered by the CCSD(T)-F12 method.[25] In this approach, the CCSD part benefits from the F12 implementation, thus rapidly converging to the CBS limit,[26] with the computational bottleneck represented by the conventional evaluation of the (T) contribution. These considerations prompted us to develop an "F12 version" of the junChS model (junChS-F12) that, on top of CCSD(T)-F12 calculations, employs MP2-F12[27] for the extrapolation to the CBS limit. In this work, we investigate whether explicitly correlated F12 approaches can improve the reliability and/or reduce the computational cost of the junChS model by reducing basis set dimensions. In this regard, both basis sets specifically developed for explicitly correlated computations (the cc-pVnZ-F12 and aug-cc-pVnZ-F12 families with n = 2−4)[28] and traditional basis sets (the jun-cc-pVnZ family with n = 3 and 4)[29,30] will be investigated in conjunction with the F12 methods.

While accurate characterizations of molecular complexes are usually limited to their energetics, structural information is likewise important because the energetics marginally depends on the structure only when reliable reference geometries are employed.[31,32] In this respect, benchmark studies are missing or

very limited. A significant example is provided by model systems to study biological interactions, where the geometry has sometimes a huge impact on the type and the strength of interactions established. Another example is offered by the investigation of reactive potential energy surfaces, which are often characterized by crucial stationary points (either intermediates or transition states) that are molecular complexes.[33] However, the most compelling example is the spectroscopic characterization of molecular complexes. Spectroscopic techniques are very powerful, also because they are noninvasive. The interpretation of all spectroscopic studies requires preliminary accurate structural evaluations because there is a strong relationship between the experimental outcome and the electronic structure of the system under investigation.[34] Since the accurate determination of equilibrium geometries of noncovalent complexes is a quite unexplored field, in this study, we have decided to also investigate the performance of conventional and explicitly correlated composite schemes for this purpose. While the conventional ChS is well tested for semirigid and flexible molecules,[17−20,35] its use in structural characterization of molecular complexes is limited to a few studies.[21,22,36] A more systematic investigation is carried out here together with the extension of different composite schemes to the explicitly correlated counterparts.

To conclude, we can summarize the goals of this study as follows:

- Definition and validation of explicitly correlated ChSs for energy evaluation.
- Definition and validation of conventional and explicitly correlated composite schemes for structural determination.
- Definition of the junChS and junChS-F12 databases.

The manuscript is organized as follows. In the next section, the methodology is described in detail. In the following section, the performance of the explicitly correlated composite schemes is discussed for interaction energy first and for structural determination then. Finally, the junChS and junChS-F12 databases are introduced and analyzed.

■ METHODOLOGY

In the following, the computational strategies employed in our work are thoroughly described. We start by introducing the Computational Details and the set of intermolecular complexes selected. Then, the ChSs based on F12 explicitly correlated methods are presented. This requires, first of all, the introduction of a sort of glossary to be used along with the manuscript. Finally, the application of the conventional and F12 ChS to structural determinations is described.

Computational Details. The Gaussian[37] and Molpro[38,39] suites of programs have been employed for conventional and explicitly correlated computations, respectively. Density fitting (DF) and resolution of identity approximations have been used throughout in F12 calculations. In all F12 calculations, the Slater-type geminal exponent has been set to $\gamma = 1.0\ a_0^{-1}$.

Concerning CCSD(T)-F12 computations, the F12b approximation[40] has been used in combination with the extrapolation to the CBS limit, while the F12a variant[41] has been employed in the other cases. For MP2-F12 computations, the default approximation (3C with EBC) has always been employed.[27] As already noted, in the CCSD(T)-F12 method, the perturbative treatment of triples is the same as in the conventional CCSD(T) counterpart. Indeed, a full F12
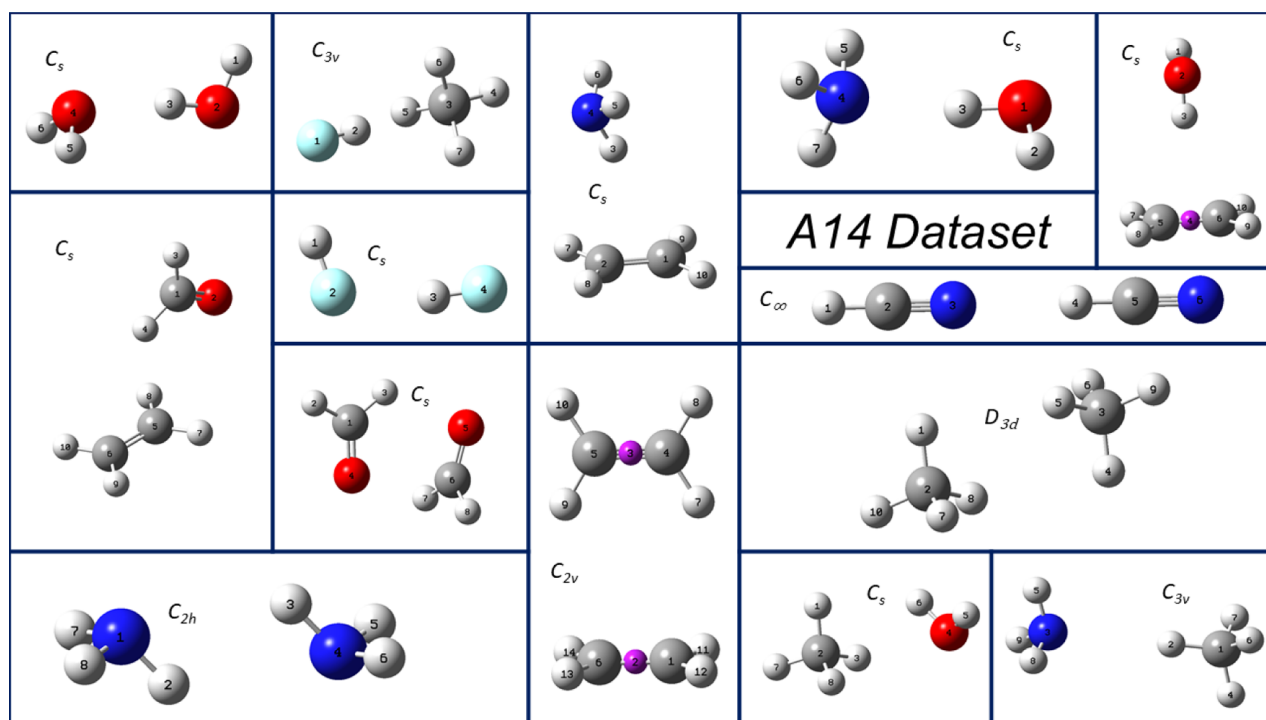
Figure 1. Total of 14 small-sized noncovalent molecular complexes of the A24 dataset used in the present work. The symmetry point group is also provided.

correction to the term accounting for triple excitations is not yet available in commercial codes. Although this contribution can be estimated from the difference between MP2 and MP2-F12 energies or from pair contributions,[25,42] we preferred to avoid any empirical term in the ChS-F12 family of composite schemes.

As mentioned in the Introduction, specific (aug-)cc-pVnZ-F12 basis sets and standard aug-cc-pVnZ and "seasonal" basis sets have been employed in combination with F12 calculations. Whenever aug-cc-pVnZ-F12 or cc-pVnZ-F12 orbital basis sets are used, suitable DF basis and resolution of the identity (RI) basis sets are automatically chosen using the Molpro program.[38] In all the other cases, these basis sets should be carefully selected. The general recipe adopted in this paper is to choose DF and RI basis sets corresponding to the augmented versions of the orbital ones. It is worthwhile recalling that the generic cc-pVnZ-F12 basis set actually corresponds to the may-cc-pV(n+1)Z set, where the highest angular momentum functions are neglected for all atoms and an additional set of p-functions is used for nonhydrogen atoms. As a consequence, jun-cc-pVTZ and aug-cc-pVDZ-F12 basis sets have comparable dimensions. Finally, it should be noted that for third-row atoms, the variants including additional tight d-functions were employed for the Dunning-like basis sets, for example, cc-pV(n+d)Z or jun-cc-pV(n+d)Z.[43]

Aiming at setting up an approach affordable also for quite large systems, the reference geometry should be evaluated at an effective level of theory. As done in the definition of the junChS model,[23] we rely on density functional theory with dispersion corrections incorporated using Grimme's D3 formulation. The double-hybrid revDSD-PBEP86-D3(BJ) functional[44−46] in conjunction with the jun-cc-pVTZ basis set (hereafter referred to as revDSD) has been employed. Although in ref 23, we have employed the B2PLYP functional,[47] the revDSD level is expected to provide improved results at the same computational cost.[32,48]

Focusing on structural determinations, the conventional junChS model (see ref 23 for its definition) together with different ChS-F12 and full CCSD(T)-F12 composite schemes will be tested, with their performance evaluated against the geometries reported by Hobza in ref 49. For the geometry optimizations carried out with Molpro, the RMS of the displacements and gradients has been set to 4×10^{-6} and 1×10^{-6} au, respectively. These thresholds correspond to the "very tight" criteria used in the geometry optimizations performed with Gaussian.

ChS Models: Dataset and Reference Energies. The selection of the systems to be considered in this study has been based on the A24 dataset,[49] which is composed by 24 small noncovalent complexes covering strong interactions, such as hydrogen bonds, but also weaker contacts dominated by dispersion terms. Our interest mainly being the development of efficient strategies for the accurate description of large systems with a biological character, a subset of 14 complexes (hereafter referred to as the A14 set) has been selected from the A24 dataset mentioned above. The systems containing either a noble gas, borane, ethyne, or ethane have been removed, as previously done in the definition of the junChS approach.[23] The chosen systems are reported in Figure 1, and they are as follows:

- Homodimers: $H_2O\cdots H_2O$, $NH_3\cdots NH_3$, $CH_2O\cdots CH_2O$, $HF\cdots HF$, $HCN\cdots HCN$, $CH_4\cdots CH_4$, and $C_2H_4\cdots C_2H_4$;
- Heterodimers: $CH_4\cdots H_2O$, $NH_3\cdots H_2O$, $CH_4\cdots NH_3$, $CH_4\cdots HF$, $C_2H_4\cdots H_2O$, $C_2H_4\cdots NH_3$, and $C_2H_4\cdots CH_2O$.

For these systems, accurate interaction energies are available from the A24 dataset.[49] They have been obtained using a CCSD(T)-based composite scheme that accounts for the extrapolation to the CBS limit within the fc approximation together with the extrapolation to the CBS limit of the CV correlation energy.[50] Therefore, this level of theory can be
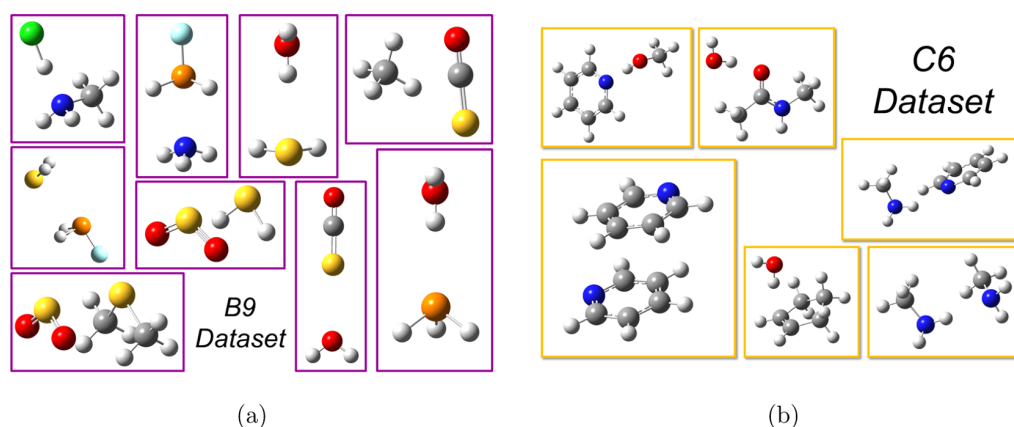
(a)   (b)

Figure 2. Additional noncovalent complexes considered. (a) B9 dataset: small systems containing atoms of the first three rows of the periodic table. (b) C6 dataset: larger systems containing atoms of the first two rows of the periodic table.

Table 1. Glossary of Conventional and F12 Composite Schemes: Acronyms and Their Explanation

conventional methods		F12 methods	
acronym	explanation	acronym	explanation
CBS	CCSD(T) extrapolated to the CBS limit (cc-pVnZ)	CBS-F12	CCSD(T)-F12 extrapolated to the CBS limit (cc-pVnZ-F12)
junCBS	CCSD(T) extrapolated to the CBS limit using jun-cc-pVnZ basis sets	junCBS-F12	CCSD(T)-F12 extrapolated to the CBS limit using jun-cc-pVnZ basis sets
CBS+CV	CCSD(T) extrapolated to the CBS limit and CCSD(T)/CV correction (cc-pwCVTZ basis)	CBS+CV-F12	CCSD(T)-F12 extrapolated to the CBS limit and CCSD(T)-F12/CV correction (cc-pCVnZ-F12 basis)
junCBS+CV	CCSD(T) extrapolated to the CBS limit using jun-cc-pVnZ basis sets and CCSD(T)/CV corr. (cc-pwCVTZ basis)	junCBS+CV-F12	CCSD(T)-F12 extrapolated to the CBS limit using jun-cc-pVnZ basis sets and CCSD(T)-F12/CV corr. (cc-pwCVTZ basis)
ChS	"cheap" scheme with conventional methods (cc-pVnZ/cc-pwCVTZ)	ChS-F12	"cheap" scheme using CCSD(T)-F12 and MP2-F12 (cc-pVnZ-F12/cc-pCVnZ-F12)
mayChS	ChS using may-cc-pVnZ basis sets	mayChS-F12	ChS-F12 using may-cc-pVnZ basis sets
junChS	ChS using jun-cc-pVnZ basis sets	junChS-F12	ChS-F12 using jun-cc-pVnZ basis sets
augChS	ChS using aug-cc-pVnZ basis sets	augChS-F12	ChS-F12 using aug-cc-pVnZ-F12 basis sets
augF12ChS	ChS using aug-cc-pVnZ-F12 basis sets		
ChS+aug$\Delta\alpha$	ChS plus additive contribution for the diffuse function using aug-cc-pVTZ		
ChS+jun$\Delta\alpha$	ChS plus additive contribution for the diffuse function using jun-cc-pVTZ		

denoted as CCSD(T)/CBS+CV or more simply as CBS+CV. While more accurate values incorporating up to quadruple excitations in the cluster expansion and relativistic effects are available,[50] one aim of this work is to provide an effective strategy to recover the error associated to the truncation of the basis set while accounting for the contribution of all electrons on top of fc-CCSD(T) calculations.[51] Therefore, the CBS+CV level will also be designed as "reference" or more shortly as "ref", in the following. Instead, we will denote the CCSD(T)/CBS level,[49] that is, the scheme not accounting for CV corrections, as "ref-CBS".

While our conventional and explicitly correlated ChSs use revDSD geometries, in order to investigate possible structural effects, the reference geometries employed in the "ref" CBS+CV calculations presented above will also be exploited. We recall that they were obtained by means of a composite scheme that starts from the HF-SCF/aug-cc-pVQZ level of theory and incorporates the extrapolation to the CBS limit at the MP2 level using aug-cc-pVnZ basis sets (n=T,Q) and the contribution of triple excitations via CCSD(T) in combination with the aug-cc-pVDZ basis set. To rule out the basis set superposition error (BSSE) effects, the counterpoise (CP) correction[52] was included at each iteration step. The resulting geometry will be denoted as "CBS-georef" in the following.

Since the A14 (and also A24) dataset only includes species with atoms of the first two rows of the periodic table, a series of small complexes bearing at least one third-row atom (referred to as the B9 set) has been added to the set of molecular complexes employed in this work. Furthermore, to test the new approach for larger noncovalent complexes, we have also built the C6 dataset. While a summary of these systems is provided here below, their graphical representation is reported in Figure 2a,b:

- B9 dataset: $FH_2P\cdots NH_3$, $FH_2P\cdots H_2S$, $H_2O\cdots H_2S$, $H_2O\cdots PH_3$, $OCS\cdots H_2O$, $SO_2\cdots H_2S$, $OCS\cdots CH_4$, $CH_3NH_2\cdots HCl$, and $(CH_3)_2S\cdots SO_2$.
- C6 dataset: c-$C_5H_8\cdots H_2O$ (Z isomer), $CH_3NH_2\cdots C_5H_5N$, $CH_3OH\cdots C_5H_5N$, $H_2O\cdots C_3H_7NO$, $C_5H_5N\cdots C_5H_5N$, and $CH_3NH_2\cdots CH_3NH_2$.

A last comment is deserved about the sought accuracy for computed interaction energies. First, it should be taken into account that they span a quite large range: from the relatively strong hydrogen bonds (10−30 kJ mol^{-1}) to the much weaker dispersion interactions (1−10 kJ mol^{-1}). As a consequence, the so-called gold-standard absolute error of 0.2 kJ mol^{-1} suggested in ref 53 is fully satisfactory for the first class of interactions, but it is questionable for the weak ones. Therefore, we prefer to retain the target accuracy introduced in ref 23: an average

relative error within 1% without any outlier above 3% and an average absolute deviation below 0.2 kJ mol^{-1}.

ChS Models: the Glossary. To guide the reader through the plethora of ChS and ChS-F12 models that will be addressed in this manuscript, Table 1 summarizes and explains the corresponding acronyms. These can be seen as formed by three different pieces. The central part refers to the type of composite scheme. In the final part, the indication whether F12 methods are used is given, while in the initial part, information on the basis set is provided. Whenever the type of basis set is not specified, the use of basis sets specifically developed for F12 computations (e.g., cc-pVnZ-F12) in combination with F12 methods and the use of the standard Dunning basis sets (e.g., cc-pVnZ) together with conventional methods are implied. If not otherwise stated, the extrapolation to the CBS limit is performed using triple- and quadruple-ζ basis sets when conventional basis sets are employed and double-ζ and triple-ζ whenever F12 sets are used. In combination with the latter, the CV term is evaluated using the cc-pCVDZ-F12 basis set,[54] while for schemes employing conventional basis sets, the CV contribution is computed with the cc-pwCVTZ basis set.[55] Finally, the +jun$\Delta\alpha$ and +aug$\Delta\alpha$ labels indicate the additive inclusion of the contribution due to diffuse functions. In the corresponding models, the ChS outlined above is augmented by a term evaluated as the difference between two MP2 computations, one using the aug-cc-pVTZ (or jun-cc-pVTZ) basis set and the other performed with cc-pVTZ.

ChS Models: Explicitly Correlated F12 Variants. While the conventional ChS approaches have already been introduced and applied,[5,17,20,23,56,57] the explicitly correlated ChS models are defined here for the first time. As already noted for the conventional ChS approaches, different variants can be defined for the ChS-F12 model, with the general expression given by the following

$$E_{\text{ChS-F12}} = E_{\text{CC-F12}} + \Delta E_{\text{CBS}}(\text{MP2-F12})$$
$$+ \Delta E_{\text{CV}}(\text{MP2-F12}) \tag{1}$$

The starting fc-CCSD(T)-F12 energy, $E_{\text{CC-F12}}$, is computed in conjunction with different basis sets, which are also employed in the CBS term. According to the basis sets used in the CCSD(T)-F12 and MP2-F12/CBS calculations, different schemes are obtained, as summarized in Table 1. Let us see in detail the expressions for the ChS-F12 and junChS-F12 models. For the former, the CCSD(T)-F12 term is computed in combination with the cc-pVDZ-F12 basis set, while in the latter, it is evaluated using the jun-cc-pVTZ set.

The extrapolation to the CBS limit is performed using the n^{-3} extrapolation formula[11]

$$\Delta E_{\text{CBS}}(\text{MP2-F12})$$
$$= \frac{n^3 E(\text{MP2-F12}/n\text{Z}) - (n-1)^3 E(\text{MP2-F12}/(n-1)\text{Z})}{n^3 - (n-1)^3} \tag{2}$$

where "nZ" and "$(n-1)$Z" stand for jun-cc-pVQZ and jun-cc-pVTZ, respectively, in the case of the junChS-F12 model, and for cc-pVTZ-F12 and cc-pVDZ-F12, respectively, when considering the ChS-F12. Noted is that "$(n-1)$Z" is the basis set used in the corresponding fc-CCSD(T)-F12 calculation.

The CV contribution is then computed as a difference between all-electron (ae) and fc MP2-F12 energy calculations

$$\Delta E_{\text{CV}}(\text{MP2-F12}) = E(\text{aeMP2-F12}/\text{C}n\text{Z})$$
$$- E(\text{fcMP2-F12}/\text{C}n\text{Z}) \tag{3}$$

where "CnZ" stands for cc-pwCVTZ for all schemes not employing F12-basis sets (e.g., junChS-F12), while for the latter, the cc-pCVDZ-F12 basis set is used.

A special note is deserved for the ChS-F12 model. In fact, while the standard model employs F12 double- and triple-ζ basis sets, the variant using F12 triple- and quadruple-ζ sets is also defined. In the latter case, the CCSD(T)-F12 term is computed with the cc-pVTZ-F12 set, the MP2-F12 extrapolation to the CBS limit is computed using cc-pVTZ-F12 and cc-pVQZ-F12, and the MP2-F12 CV contribution is computed with cc-pCVTZ-F12.

In addition to the different ChS-F12 models generically defined above and summarized in Table 1, composite schemes entirely based on CCSD(T)-F12 calculations and accounting for the extrapolation to the CBS limit and CV correction have also been defined and tested. According to Table 1, these are defined as CBS+CV-F12. However, we need to specify the basis sets employed. If the cc-pVDZ-F12 and cc-pVTZ-F12 sets are used for the extrapolation and cc-pCVDZ-F12 is used for the CV term, then, we obtain the CBS2+CV-F12. Otherwise, if the extrapolation to the CBS limit is performed using cc-pVTZ-F12 and cc-pVQZ-F12, and cc-pCVTZ-F12 is employed for the CV correction, the scheme is denoted as CBS3+CV-F12. If not explicitly specified, CBS+CV-F12 refers to CBS2+CV-F12. Finally, if the may-cc-pVnZ and jun-cc-pVnZ sets are used, we refer to it as mayCBS+CV-F12 and junCBS+CV-F12, respectively. In these latter cases, the CV contribution is evaluated using the cc-pwCVTZ basis set. If the CV term is neglected, the model is defined as CBS (conventional) and CBS-F12 (either CBS2-F12 or CBS3-F12; if not specified, we refer to the former).

ChS Models: Structural Determination. While the conventional ChSs have already been employed in structural determinations, their application to geometry optimizations of molecular complexes has not been introduced yet. As mentioned in the Introduction, this is accomplished in the present work for both conventional and explicitly correlated ChSs. The resulting composite approaches are denoted as "geometry" schemes and require geometry optimizations at different levels of theory. All the defined schemes do not account for the CP correction in view of its very small contribution, especially for explicitly correlated approaches, as will be demonstrated in the Results and Discussion section.

The additivity of the various terms is exploited at a parameter level. For a generic structural parameter R, its expression is given by

$$R(\text{ChS[-F12]}) = R(\text{CCSD}(T)\text{[-F12]})$$
$$+ \Delta R(\text{MP2[-F12]/CBS})$$
$$+ \Delta R(\text{MP2[-F12]/CV}) \tag{4}$$

where the first term on the right-hand side is the structural parameter R optimized using CCSD(T) (or CCSD(T)-F12 for explicitly correlated schemes) in conjunction with the specific basis set that defines the type of approach. To give a couple of examples, for the junChS model, this first term is the result of the fc-CCSD(T)/jun-cc-pVTZ geometry optimization, while it is the fc-CCSD(T)-F12/cc-pVDZ-F12 optimized parameter in the case of ChS-F12. The second term is the contribution due to

the extrapolation to the CBS limit, which is derived from the extrapolation of the geometrical parameters obtained from fc-MP2 or MP2-F12 optimizations with two members of the hierarchical series of bases defining the scheme. This means jun-cc-pVTZ and jun-cc-pVQZ for the junChS model and cc-pVDZ-F12 and cc-pVTZ-F12 for the ChS-F12 approach. As done for energetics, the extrapolation is performed using the n^{-3} expression as follows

$$
\Delta R(\text{CBS}) = \frac{n^3 R_{\text{MP2[-F12]}/nZ} - (n-1)^3 R_{\text{MP2[-F12]}/(n-1)Z}}{n^3 - (n-1)^3}
$$
$$
- R_{\text{MP2[-F12]}/(n-1)Z} \tag{5}
$$

where $n = 4$ (i.e., QZ), with the only exception of the ChS-F12 variant. The last term in eq 4 introduces the effects on the geometry arising from the correlation of core electrons. This contribution is usually non-negligible because equilibrium structures are particularly sensitive to the correlation of the inner electrons. The $\Delta R(\text{CV})$ term is given by the difference between the structural parameter R resulting from a geometry optimization with all electrons correlated and that within the fc approximation. The basis sets used are (i) the cc-pwCVTZ basis set for all conventional models and for the F12 approaches not employing F12 basis sets and (ii) the cc-pCVnZ-F12 basis sets, with $n = $ D or T, for those methods exploiting the F12 family of basis sets.

Starting from the ChS model, two additional schemes have been defined by incorporating the effect of diffuse functions via an additive $\Delta\alpha$ term, $\Delta R(\text{MP2/diff})$, which can be computed using the jun-cc-pVTZ or aug-cc-pVTZ basis set. Such an addition leads to the definition of the ChS+jun$\Delta\alpha$ and ChS+aug$\Delta\alpha$. The $\Delta R(\text{MP2/diff})$ term is evaluated as a difference in the structural parameters optimized within the fc approximation. For the jun$\Delta\alpha$ correction, we have

$$
\Delta R(\text{MP2/diff}) = R(\text{fc-MP2/junTZ}) - R(\text{fc-MP2/TZ}) \tag{6}
$$

The CBS+CV scheme, entirely based on CC calculations, has also been used with the additivity being again applied at the geometry parameter level. In addition to the CBS2+CV-F12 and CBS3+CV-F12 models, entirely similar in the definition to the corresponding energy schemes, the junCBS+CV-F12 and junCBS+CV models have been introduced

$$
R(\text{junCBS+CV[-F12]})
$$
$$
= R(\text{CCSD(T)[-F12]/CBS}) + \Delta R(\text{CCSD(T)[-F12]} /\text{CTZ}) \tag{7}
$$

where, in both schemes, the jun-cc-pVTZ and jun-cc-pVQZ sets are used for the extrapolation to the CBS limit and cc-pwCVTZ is used for the CV term.

RESULTS AND DISCUSSION

In the following section, the performance of the various ChS-F12 models, based on the A14 dataset, for both interaction energy evaluations and structural determinations will be presented with the aim of defining the most effective scheme. As a byproduct of this investigation, a database of accurate structures and energies for the A14 + B9 + C6 molecular complexes will be defined. In passing, the performance of the revDSD level of theory for both geometries and energies will be discussed.

Performance of the ChS-F12 Models: Interaction Energy. Together with the various ChS-F12 models (see Table 1), the conventional junChS has also been considered for comparison purposes, and the role of the different contributions has been investigated. As mentioned in the Methodology section, while the revDSD structures are the reference geometries for all our schemes, the geometry effect has been studied by resorting to the "CBS-georef" reference geometries. Noted is that in addition to CP-corrected interaction energies, half-corrected (half-CP) and noncorrected CP(NCP) results will also be discussed.

In Table 2, the results for the A14 dataset are summarized in terms of absolute and relative energies with respect to the "ref"

Table 2. Relative and Absolute Errors[a,b] of the revDSD Level, the Conventional junChS and augF12ChS Models, and all ChS-F12 Variants[c]

model		relative error (%)	absolute error
revDSD[c]	CP	6.34	0.49
	NCP	3.59	0.34
	half-CP	3.11	0.24
junChS	CP	1.38	0.12
	NCP	2.06	0.22
	half-CP	1.68	1.68
augF12ChS	CP	1.12	0.14
	NCP	9.70	0.79
	half-CP	4.34	0.33
mayChS-F12	CP	1.23	0.10
	NCP	2.81	0.19
	half-CP	1.93	0.14
junChS-F12	CP	0.68	0.05
	NCP	1.10	0.08
	half-CP	0.88	0.06
jun-(d,f)H_ChS-F12	CP	0.65	0.05
	NCP	1.07	0.08
	half-CP	0.84	0.06
ChS-F12[d]	CP	2.17 (0.93)	0.18 (0.09)
	NCP	1.20 (0.86)	0.12 (0.09)
	half-CP	1.07 (0.82)	0.08 (0.09)
augChS-F12	CP	0.67	0.06
	NCP	2.12	0.21
	half-CP	1.01	0.10

[a]Absolute errors in kJ mol^{-1}. [b]Errors evaluated with respect to CBS+CV "ref" reference energies. See text. [c]For all schemes: revDSD reference geometries. [d]In parentheses, the results for the scheme employing the F12 triple- and quadruple-ζ basis sets are given. See text.

values. The most promising F12 results are also graphically presented in Figure 3, where their relative errors are compared with those of the junChS counterpart. All F12 models of Table 1 are considered together with the conventional junChS and augF12ChS, with the latter—despite being a conventional model—employing the aug-cc-pVDZ-F12 and aug-cc-pVTZ-F12 basis sets. It is noted that the revDSD model provides large relative errors, these ranging from ~3 to ~6%. However, in absolute terms, the deviations are smaller than 0.5 kJ mol^{-1}. The smallest mean absolute error (MAE) is delivered by the junChS-F12 model. Interestingly, reducing the dimension of the basis set by removing d and f polarization functions on hydrogen atoms (for triple- and quadruple-ζ basis sets, respectively) and thus resorting to the jun-(d,f)H_ChS-F12
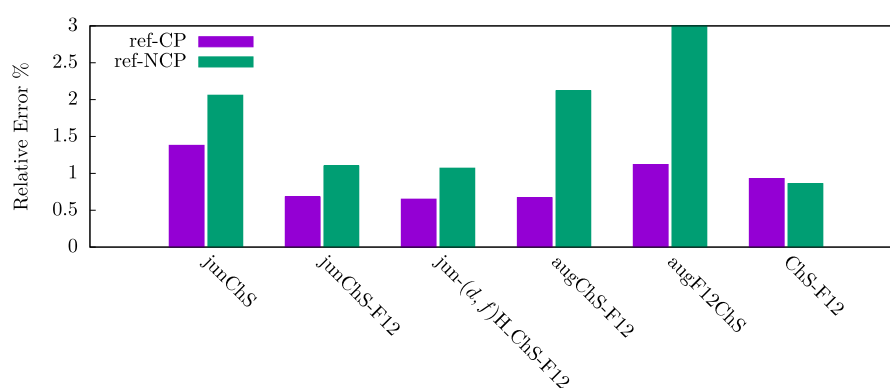
Figure 3. CP and NCP energy differences with respect to reference, "ref", using revDSD geometries. ChS-F12 employes cc-pVTZ-F12 and cc-pVQZ-F12 basis sets (cc-pCVTZ-F12 for the CV term).

model even slightly improves the MAE. As far as the comparison with conventional schemes is concerned, it is apparent that moving from the conventional junChS approach to the junChS-F12 counterpart nearly halves the MAE. However, the absolute errors point out that the differences between the two models, that is, only fractions of kJ mol^{-1}, are much less relevant with respect to what the MAEs tend to suggest. This is because half of the molecular complexes of the A14 set are characterized by small interaction energies.

Back to Table 2, the standard ChS-F12, which employs the F12 double- and triple-ζ basis sets, gives disappointing results, and the same applies to the mayChS-F12 model. Indeed, d and f diffuse functions on nonhydrogen atoms (not included at these levels) play a non-negligible role. While the extension of the ChS-F12 to the cc-pVTZ-F12 and cc-pVQZ-F12 sets (see numbers in parentheses and Figure 3) leads to improved results, this comes at the expense of a huge increase in the computational cost. Instead, the augChS-F12 model represents a much more effective alternative. Indeed, with the jun-cc-pVTZ and aug-cc-pVDZ-F12 basis sets having comparable dimensions, it is somewhat natural that they lead to similar results.

In Table 3, the results of a statistical analysis for some CBS +CV schemes entirely based on the CCSD(T)-F12 method are reported. In addition to the CBS+CV-F12, mayCBS+CV-F12,

and junCBS+CV-F12 models, the approach employing the jun-cc-pVnZ-(d,f)H basis sets has also been considered. From the comparison of Tables 2 and 3, it is apparent that if we consider CP results, the ChS models perform even better than the CBS +CV schemes, the only exception being the approaches involving the may-cc-pVnZ basis sets. The situation is completely different if we analyze the NCP values. In fact, with the exception of the approaches involving the F12 basis sets, the best performance is obtained using the CBS+CV schemes. This behavior can be explained by a better extrapolation to the CBS limit in the schemes entirely based on the CCSD(T)-F12 method. Indeed, on a pure theoretical basis, CP and NCP values should be identical at the CBS limit because the BSSE should vanish. Therefore, CP-correcting the energies should worsen the results, what actually occurs in the case of CBS+CV schemes. Since the MP2-F12 extrapolation is less effective, the CP correction within a ChS approach is still required and leads to improved results. Furthermore, CP correction in ChSs is computationally more convenient than in a CBS+CV scheme, the overall conclusion being that the junChS-F12 and jun-(d,f)H_ChS-F12 models are the best-performing schemes among the F12 approaches considered in this work.

Extrapolation to the CBS Limit. In view of the discussion given above, it is interesting to investigate the convergence to the CBS limit, thereby resorting to the "ref-CBS" interaction energy as a reference. In Tables 4 and 5, the extrapolation to the CBS limit within CBS+CV-F12 and ChS-F12 is analyzed, respectively, with the cc-pVnZ-F12, may-cc-pVnZ, jun-cc-pVnZ, and jun-cc-pVnZ-(d,f)H basis sets being considered. Furthermore, the exponent m of the n^{-m} extrapolation has been tested. Among different preliminary investigations, we selected the $m = 5$ exponent together with the standard $m = 3$ exponent. From the CCSD(T)-F12/CBS results, summarized in terms of absolute and relative errors, it is noted that the MAEs of the NCP-mayCBS-F12 and NCP-junCBS-F12 interaction energies drop from 2.13 and 1.21% to 1.23 and 1.18%, respectively, when moving from n^{-3} to n^{-5}. The reduction of the MAE is even more marked when considering the CP interaction energies; indeed, they reduce from 2.80 and 1.64% to 0.81 and 0.76%, respectively, going from n^{-3} to n^{-5}. Analogous reductions are noted also for the jun-cc-pVnZ-(d,f)H and cc-pVnZ-F12 families of basis sets when considering the (T,Q) combination. For the (D,T) combination, the CP energies show instead the opposite behavior, possibly due to an error compensation related to the small basis set.

Table 3. Relative and Absolute Errors^{a,b} of Different CBS +CV Composite Schemes^{c}

model		relative error (%)	absolute error
mayCBS+CV-F12	CP	0.87	0.08
	NCP	1.18	0.09
	half-CP	0.92	0.08
junCBS+CV-F12	CP	0.79	0.07
	NCP	0.92	0.09
	half-CP	0.75	0.07
jun-(d,f)H_CBS+CV-F12	CP	0.79	0.07
	NCP	0.99	0.10
	half-CP	0.79	0.08
CBS+CV-F12^{d}	CP	1.36	0.10
	NCP	1.46	0.16
	half-CP	1.00	0.08

^{a}Absolute errors in kJ mol^{-1}. ^{b}Errors evaluated with respect to CBS +CV "ref" reference energies. See text. ^{c}For all schemes: revDSD reference geometries. ^{d}The extrapolation to the CBS limit using the n^{-5} formula, with $n = 3, 4$.

Table 4. Relative and Absolute Errors[a,b] of the CCSD(T)-F12/CBS Approach in Combination with Different Basis Sets on Top of the revDSD Geometries

model		n^{-3}		n^{-5}	
		relative error (%)	absolute error	relative error (%)	absolute error
cc-pV(D,T)Z-F12	CP	0.86	0.07	1.68	0.11
	NCP	1.77	0.19	1.51	0.16
	half-CP	1.17	0.12	1.07	0.09
cc-pV(T,Q)Z-F12	CP	1.22	0.13	0.77	0.08
	NCP	0.93	0.11	0.85	0.11
	half-CP	1.03	0.12	0.81	0.09
may-cc-pV(T,Q)Z	CP	2.80	0.24	0.81	0.08
	NCP	2.13	0.14	1.23	0.10
	half-CP	1.77	0.15	0.93	0.08
jun-cc-pV(T,Q)Z	CP	1.64	0.16	0.76	0.08
	NCP	1.21	0.11	1.18	0.11
	half-CP	1.25	0.13	0.91	0.09
jun-cc-pV(T,Q)Z-d,fH	CP	1.80	0.17	0.77	0.08
	NCP	1.36	0.13	1.25	0.13
	half-CP	1.44	0.15	0.95	0.10

[a]Absolute errors in kJ mol^{-1}. [b]Errors evaluated with respect to "ref-CBS" reference energies. See text.

Table 5. Relative and Absolute Errors[a,b] of the Extrapolation to the CBS Limit within ChS Approaches: $E_{CC\text{-}F12} + \Delta E_{CBS}(MP2\text{-}F12)$

model		n^{-3}		n^{-5}	
		relative error (%)	absolute error	relative error	absolute error
cc-pV(D,T)Z-F12	CP	2.27	0.18	2.88	0.23
	NCP	1.10	0.10	1.35	0.18
	half-CP	1.30	0.09	1.62	0.15
cc-pV(T,Q)Z-F12	CP	0.73	0.07	0.70	0.09
	NCP	0.90	0.09	0.79	0.06
	half-CP	0.75	0.07	0.71	0.07
may-cc-pV(T,Q)Z	CP	1.25	0.10	1.90	0.14
	NCP	2.87	0.19	1.59	0.09
	half-CP	2.00	0.14	1.65	0.10
jun-cc-pV(T,Q)Z	CP	0.66	0.05	1.03	0.07
	NCP	1.05	0.07	1.17	0.10
	half-CP	0.84	0.06	0.77	0.05
jun-cc-pV(T,Q)Z-d,fH	CP	0.72	0.06	1.12	0.09
	NCP	1.02	0.07	1.09	0.08
	half-CP	0.82	0.06	0.76	0.05

[a]Absolute errors in kJ mol^{-1}. [b]Errors evaluated with respect to "ref-CBS" reference energies. See text.

Table 6. junChS-F12 CP Energies (kJ mol^{-1}): the Various Contributions for the A14 Complexes

complex	"ref"	CC/junTZ	$\Delta MP2^{\infty}$/jun(T,Q)Z	$\Delta MP2$-CV/wCTZ	total	rel. error (%)	abs. error
$H_2O\cdots H_2O$	−21.0832	−20.8822	0.0344	−0.1512	−20.9990	−0.40	0.08
$NH_3\cdots NH_3$	−13.2131	−12.9057	−0.2132	−0.0807	−13.1997	−0.10	0.01
$HF\cdots HF$	−19.2213	−19.1430	0.0014	−0.1078	−19.2494	0.15	−0.03
$CH_2O\cdots CH_2O$	−18.9284	−18.5310	−0.3690	−0.0552	−18.9553	0.14	−0.03
$HCN\cdots HCN$	−19.9828	−19.7537	−0.0085	−0.0787	−19.8410	−0.71	0.14
$C_2H_4\cdots C_2H_4$	−4.6024	−4.3114	−0.3000	−0.0493	−4.6607	1.27	−0.06
$CH_4\cdots CH_4$	−2.2301	−1.9832	−0.2083	−0.0055	−2.1970	−1.48	0.03
$H_2O\cdots NH_3$	−27.3759	−27.1443	−0.1294	−0.2003	−27.4740	0.36	−0.10
$H_2O\cdots C_2H_4$	−10.7696	−10.4381	−0.2219	−0.1092	−10.7692	−0.004	0.0004
$C_2H_4\cdots CH_2O$	−6.7948	−6.4974	−0.2677	−0.0620	−6.8271	0.48	−0.03
$NH_3\cdots C_2H_4$	−5.7865	−5.5505	−0.2083	−0.0603	−5.8190	0.56	−0.03
$HF\cdots CH_4$	−6.9162	−6.7403	−0.1943	−0.1072	−7.0418	1.82	−0.13
$H_2O\cdots CH_4$	−2.8242	−2.6431	−0.0979	−0.0335	−2.7746	−1.76	0.05
$NH_3\cdots CH_4$	−3.2175	−3.0678	−0.1160	−0.0428	−3.2266	0.28	−0.01
MAE						0.68	0.05

Table 7. jun-(d,f)H_ChS-F12 CP Energies (kJ mol^{-1}): the Various Contributions for the A14 Complexes

	"ref"	CC/junTZ-(d,f)H	ΔMP2$^{\infty}$/jun(T,Q)Z-(d,f)H	ΔMP2-CV/wCTZ	total	rel. error (%)	abs. error
$H_2O{\cdots}H_2O$	−21.0832	−20.8347	0.0352	−0.1512	−20.9506	0.63	−0.13
$NH_3{\cdots}NH_3$	−13.2131	−12.8501	−0.2421	−0.0807	−13.1729	0.30	−0.04
$HF{\cdots}HF$	−19.2213	−19.0976	−0.0035	−0.1078	−19.2089	0.06	−0.01
$HCN{\cdots}HCN$	−19.9828	−19.7414	−0.0045	−0.0787	−19.8246	0.79	−0.16
$CH_4{\cdots}CH_4$	−2.2301	−1.9667	−0.2252	−0.0055	−2.1974	1.47	−0.03
$CH_2O{\cdots}CH_2O$	−18.9284	−18.4796	−0.3971	−0.0552	−18.9319	0.02	0.00
$C_2H_4{\cdots}C_2H_4$	−4.6024	−4.2954	−0.3218	−0.0493	−4.6665	1.39	0.06
$H_2O{\cdots}C_2H_4$	−10.7696	−10.4093	−0.2320	−0.1091	−10.7505	0.18	−0.02
$H_2O{\cdots}CH_4$	−2.8242	−2.6338	−0.0999	−0.0339	−2.7676	2.00	−0.06
$H_2O{\cdots}NH_3$	−27.3759	−27.0710	−0.1545	−0.2002	−27.4257	0.18	0.05
$NH_3{\cdots}CH_4$	−3.2175	−3.0571	−0.1161	−0.0423	−3.2155	0.06	0.00
$NH_3{\cdots}C_2H_4$	−5.7865	−5.5344	−0.2189	−0.0602	−5.8136	0.47	0.03
$HF{\cdots}CH_4$	−6.9162	−6.6717	−0.2239	−0.1072	−7.0029	1.25	0.09
$C_2H_4{\cdots}CH_2O$	−6.7948	−6.4739	−0.2820	−0.0600	−6.8159	0.31	0.02
MAE						0.65	0.05

Table 8. junCBS+CV-F12 CP Energies (kJ mol^{-1}): the Various Contributions for the A14 Complexes

	"ref"	CC-CBS/jun(T,Q)Z	ΔCC-CV/wCTZ	total	rel. error (%)	abs. error
$H_2O{\cdots}H_2O$	−21.0832	−21.0524	−0.1200	−21.1724	0.42	−0.09
$NH_3{\cdots}NH_3$	−13.2131	−13.1601	−0.0605	−13.2206	0.06	−0.01
$HF{\cdots}HF$	−19.2213	−19.3738	−0.0818	−19.4556	1.22	−0.23
$CH_2O{\cdots}CH_2O$	−18.9284	−18.9566	0.0382	−18.9184	−0.05	0.01
$HCN{\cdots}HCN$	−19.9828	−19.8441	−0.0811	−19.9253	−0.29	0.06
$C_2H_4{\cdots}C_2H_4$	−4.6024	−4.5555	−0.0068	−4.5623	−0.87	0.04
$CH_4{\cdots}CH_4$	−2.2301	−2.2058	0.0049	−2.2009	−1.31	0.03
$H_2O{\cdots}NH_3$	−27.3759	−27.4233	−0.1633	−27.5865	0.77	−0.21
$H_2O{\cdots}C_2H_4$	−10.7696	−10.6645	−0.0532	−10.7177	−0.48	0.05
$C_2H_4{\cdots}CH_2O$	−6.7948	−6.7604	−0.0010	−6.7704	−0.36	0.02
$NH_3{\cdots}C_2H_4$	−5.7865	−5.7393	−0.0243	−5.7636	−0.40	0.02
$HF{\cdots}CH_4$	−6.9162	−7.0176	−0.0749	−7.0925	2.55	−0.18
$H_2O{\cdots}CH_4$	−2.8242	−2.7537	−0.0206	−2.7743	−1.77	0.05
$NH_3{\cdots}CH_4$	−3.2175	−3.2065	−0.0274	−3.2338	0.51	−0.02
MAE					0.79	0.07

From the inspection of Tables 4 and 5, it appears that the n^{-5} formula performs better than the n^{-3} counterpart for the extrapolation of CCSD(T)-F12 energies. Instead, concerning MP2-F12 energies within the ChS-F12 models, we note that the conventional n^{-3} extrapolation formula performs better than the n^{-5} form, the only exception being the (T,Q) combination for the cc-pVnZ-F12 basis sets. Therefore, n^{-3} has been retained for ChS-F12 models, while n^{-5} has been chosen as the extrapolation of CCSD(T)-F12 energies.

Role of the Various ChS-F12 Contributions. To discuss the different contributions within a given composite scheme, the best ChS-F12 variants have been selected, namely, junChS-F12 and jun-(d,f)H_ChS-F12. For comparison purposes, the junCBS+CV-F12 model has also been considered. The results are collected in Table 6 for junChS-F12, Table 7 for jun-(d,f)H_ChS-F12, and Table 8 for junCBS+CV-F12. In all cases, CP-corrected energies are reported, the NCP counterparts being available in the Supporting Information (Tables S1−S3). In all tables, CC stands for fc-CCSD(T)-F12 and, when present, MP2 stands for MP2-F12.

From Tables 6 and 7, we note that—in absolute terms—the CCSD(T)-F12/jun-cc-pVTZ (or CCSD(T)-F12/jun-cc-pVTZ-(d,f)H) terms underestimate the "ref" values by a quantity ranging from ∼0.1 to ∼0.4 kJ mol^{-1}. In almost all cases, about 70% of these differences are recovered by the CBS

correction term evaluated at the MP2-F12 level. This contribution is in almost all cases negative, which means that it increases—in absolute terms—the interaction energies obtained at the CCSD(T)-F12 level in conjunction with the triple-ζ basis set. Furthermore, if the comparison is extended to Table 8, we note that the CCSD(T)-F12 terms augmented by the MP2-F12/CBS corrections (Tables 6 and 7) are very close to the corresponding junCBS-F12 values. This means that the approximation of recovering the extrapolation to the CCSD(T)-F12 CBS limit at the MP2-F12 level is valid. Moving to the CV contribution, it is again noted that CCSD(T)-F12 and MP2-F12 provide similar corrections. In almost all cases, the CV corrections are negative, thus further increasing the computed interaction energies. Overall, the CV term is not negligible at all, indeed this being in almost all cases larger (up to 1 order of magnitude) than the absolute error associated to the model considered.

Finally, it is interesting to observe that if NCP interaction energies are considered (see the Supporting Information), all the abovementioned observations remain valid, the only difference being that for the junChS-F12 and jun-(d,f)H_ChS-F12, the MAE increases from 0.68 to 1.10% and from 0.65 to 1.07%, respectively. Analogously, for the junCBS+CV-F12 model, the MAE increases from 0.79 to 0.92%. However, despite these significant differences in the CP- and NCP-MAEs,

38

in absolute terms, they are very small, indeed pointing out that—owing to the extrapolation to the CBS limit—the CP corrections are small. If the comparison is extended to the conventional junChS model (see the Supporting Information, Tables S4 and S5), we note that the CP−NCP energy differences are slightly more pronounced for schemes involving conventional methods when comparing models employing the same basis sets (see Figure 3). For example, the difference is larger for junChS than for junChS-F12. This is even more pronounced when comparing augF12ChS and augChS-F12 (both employing the aug-cc-pVnZ-F12 sets; see Table 1). To conclude this subsection, it is worthwhile noting once again that the junChS-F12 and jun-(df)H_ChS-F12 approaches considered in the detailed discussion given above perform better than the junCBS+CV-F12 scheme despite their reduced computational cost.

Geometry Effects on Interaction Energies. In order to investigate the effect of the reference structures and possible errors associated to their choice, a comparison has been carried out between NCP and CP interaction energies of the A14 complexes obtained on top of the "CBS-georef" and revDSD structures. In passing, we note that all revDSD-optimized structures belong to the same symmetry point group as those reported by Hobza,[49] with the revDSD structural parameters of the A14 dataset gathered in the Supporting Information (Table S6). The results to be compared are those of Table 2 (revDSD geometry) and Table 9 (CBS-georef geometry).

Table 9. Relative and Absolute Errors[a,b] of the revDSD Level and the Conventional junChS and the junChS-F12 Variants Employing "CBS-georef" Reference Geometries

model		relative error (%)	absolute error
revDSD	CP	6.74	0.54
	NCP	3.31	0.29
	half-CP	3.45	0.28
junChS	CP	0.93	0.10
	NCP	1.35	0.14
	half-CP	1.03	0.11
junChS-F12	CP	0.67	0.073
	NCP	0.98	0.072
	half-CP	0.82	0.071
jun-(d,f)H_ChS-F12	CP	0.74	0.083
	NCP	0.95	0.066
	half-CP	0.83	0.073

[a]Absolute errors in kJ mol^{-1}. [b]Errors evaluated with respect to CBS+CV "ref" reference energies. See text.

We limit the discussion to the three most promising composite methods, namely, junChS, junChS-F12, and junChS-F12-(d,f)H. For them, when CP interaction energies are considered, the MAEs are 1.38, 0.68, and 0.65% when using revDSD geometries and 0.93, 0.67, and 0.74% when employing the "CBS-georef" structures, respectively. Similar variations are noted for NCP interaction energies. We can thus conclude that improving the reference structure leaves essentially unchanged the performance of the jun-ChS-F12, while that of the conventional junChS is significantly improved. A slight worsening is noted for the junChS-F12-(d,f)H model. However, in absolute terms, these variations lead to negligible energy differences, well below 0.05 kJ mol^{-1}.

Performance of the ChS-F12 Models: Structural Determination. The performance of the geometry schemes

introduced in the Methodology section has been tested using the molecular complexes of the A14 dataset. The most critical and sensitive parameter to analyze in the case of noncovalent complexes is the intermolecular distance ruling the interaction, and indeed, this will be the quantity discussed in detail in the following. For the sake of completeness, all parameters describing the complexes are reported, for all the approaches considered, in the Supporting Information (Table S6). The main comment on the intramolecular distances is that they appear to be similar at the different levels of theory considered, with differences of few mÅ. The same applies to the intramolecular angles, which are quite insensitive to the level of theory, and only changes smaller than 1° are observed. For example, in the dimer of water, the two intramolecular H−O−H angles are predicted to be 104.98 and 104.88° using the junCBS+CV-F12 model. The same angles are 104.90 and 104.84°, respectively, at the junChS level and 104.72 and 104.92°, respectively, from the junChS-F12 approach. The intermolecular O4−H3−O2 angle spans from 171.15 to 172.68°, the maximum variation between different composite schemes thus being 1.5°. Since the PES along an intermolecular angle is usually quite flat, one can conclude that such small changes lie within the uncertainty of the structural determination and that the overall energetics of the system is not influenced by small changes in intra-/intermolecular angles. Another example to support the conclusion that intramolecular distances and angles and intermolecular angles are less affected by the model considered is offered by the CH$_4$···NH$_3$ adduct. This is characterized by a H−C−H intramolecular angle, which is found to be 109.7° by all the schemes employed. As before, the intramolecular distances are very similar at all levels of theory, with differences well within 4 mÅ. At the junCBS+CV-F12 level, the N−H and C−H bond lengths are 1.011 and 1.087 Å, respectively. To monitor different types of complexes, we can also consider a system dominated by dispersion interactions, namely, the methane dimer. In this case, the C−H intramolecular distances obtained by exploiting different composite schemes are very similar (see the Supporting Information for values), and the same applies to the intermolecular angle, whose predicted value is always close to 70.4°.

Let us now focus the discussion on intermolecular distances, which are instead particularly sensitive to the computational approach considered. The first question to address is the reliability of models employing an additive term for incorporating the effects of diffuse functions (the ChS+aug$\Delta\alpha$ or ChS+jun$\Delta\alpha$ schemes; see Table 1). Let us consider a couple of examples. For the rather rigid NH$_3$···H$_2$O complex, the N···O distance is 1.957 Å at the ChS+aug$\Delta\alpha$ level, to be compared with the value of about ~1.975 Å delivered by the junChS and junCBS+CV-F12 models. The difference being close to 20 mÅ leads to the conclusion that the ChS+aug$\Delta\alpha$ approach is not suitable for evaluating geometries of noncovalent complexes. The N···O distance issuing from ChS+jun$\Delta\alpha$ geometry optimizations is 1.960 Å, thus reducing the difference with respect to junCBS+CV-F12 to 15 mÅ, which is—however—an unsatisfactory value. Moving to a more flexible system, that is, the CH$_4$···NH$_3$ complex, the inadequacy of the ChS+aug$\Delta\alpha$ and ChS+jun$\Delta\alpha$ schemes is even more evident: these models provide a N···H distance of about 0.2 Å longer than what was predicted by the other composite approaches, irrespective of the considered family of basis sets. In conclusion, the additive approximation for the effect of diffuse functions is not suitable

Table 10. Comparison of CP- and NCP-Corrected Geometries for Some Paradigmatic Cases (atom labeling in Figure 1)

		junChS		junChS-F12		junCBS+CV		junCBS+CV-F12[a]		CC-F12/VDZ-F12[b]
		CP	NCP	CP	NCP	CP	NCP	CP	NCP	NCP
H$_2$O···H$_2$S	r(O1–H2)	0.9603	0.9603	0.9605	0.9606	0.9606	0.9612	0.9608 (0.9603)	0.9608 (0.9603)	0.9651
	r(O1–H3)	0.9560	0.9561	0.9563	0.9563	0.9563	0.9570	0.9566 (0.9561)	0.9566 (0.9561)	0.9608
	θ(H2–O1–H3)	104.67	104.69	104.81	104.91	104.82	104.66	104.73 (104.74)	104.73 (104.75)	104.40
	r(H3···S4)	3.4821	3.4758	3.4656	3.4716	3.4738	3.4799	3.4847 (3.4805)	3.4836 (3.4850)	3.4722
	θ(O1–H3–S4)	116.22	116.69	119.32	117.78	116.61	116.48	116.27 (116.27)	116.33 (116.47)	116.83
	r(H5/H6–S4)	1.3363	1.3363	1.3367	1.3344	1.3369	1.3368	1.3366 (1.3365)	1.3366 (1.3364)	1.3391
	θ(H5–S4–H3)	84.29	83.78	81.85	83.07	84.31	84.05	84.36 (84.38)	84.34 (84.21)	82.82
	φ(H5–S4–H3–O1)	133.55	133.50	133.23	133.39	133.49	133.49	133.48 (133.48)	133.48 (133.47)	133.41
	φ(H6–S4–H3–O1)	−133.55	−133.50	−133.23	−133.39	−133.49	−133.49	−133.48 (−133.48)	−133.48 (−133.47)	−133.41
	θ(H5–S4–H6)	92.46	92.29	92.42	92.33	92.43	92.38	92.45 (92.46)	92.44 (92.45)	92.23
H$_2$O···H$_2$O	r(O2–H1)	0.9548	0.9554	0.9550	0.9556	0.9563	0.9563	0.9559 (0.9554)	0.9559 (0.9554)	0.9602
	r(O2–H3)	0.9625	0.9626	0.9631	0.9629	0.9634	0.9635	0.9631 (0.9627)	0.9632 (0.9626)	0.9672
	θ(H3–O2–H1)	105.38	104.84	105.48	104.92	104.80	104.81	104.89 (104.89)	104.88 (104.89)	104.49
	r(H3···O4)	2.1860	1.9507	2.1614	1.9487	1.9571	1.9513	1.9533 (1.9494)	1.9513 (1.9512)	1.9600
	θ(O4–H3–O2)	168.35	171.15	169.18	171.24	171.78	171.41	171.79 (171.93)	171.82 (171.75)	171.73
	r(H5/H6–O4)	0.9554	0.9570	0.9567	0.9572	0.9579	0.9579	0.9575 (0.9570)	0.9575 (0.9570)	0.9618
	θ(H5–O4–H3)	107.70	110.74	109.96	111.85	111.59	110.89	111.63 (111.92)	111.54 (111.66)	109.79
	φ(H5–O4–H3–O2)	55.63	57.89	57.32	58.67	58.50	58.00	58.59 (58.78)	58.46 (58.54)	57.24
	φ(H6–O4–H3–O2)	−55.63	−57.89	−57.32	−58.67	−58.50	−58.00	−58.59 (−58.78)	−58.46 (−58.54)	−57.24
	θ(H5–O4–H6)	104.25	104.90	103.98	104.72	104.89	104.88	104.98 (105.01)	104.98 (104.99)	104.61
CH$_4$···CH$_4$	r(C2–H1)	1.0857	1.0857	1.0857	1.0857	1.0864	1.0864	1.0860 (1.0858)	1.0860 (1.0858)	1.0883
	r(C2···C3)	3.6233	3.6231	3.6397	3.6373	3.6435	3.6388	3.6403 (3.6251)	3.6385 (3.6296)	3.6526
	θ(C3–C2–H1)	70.44	70.44	70.44	70.45	70.44	70.44	70.44 (70.44)	70.44 (70.44)	70.44
	r(C3–H9)	1.0857	1.0857	1.0857	1.0857	1.0864	1.0864	1.0860 (1.0858)	1.0860 (1.0858)	1.0883
HCN···HCN	r(H1–C2)	1.0658	1.0656	1.0659	1.0658	1.0662	1.0661	1.0660 (1.0660)	1.0659 (1.0658)	1.0666
	r(C2–N3)	1.1505	1.1505	1.1500	1.1501	1.1518	1.1519	1.1507 (1.1500)	1.1508 (1.1501)	1.1560
	r(N3···H4)	2.2159	2.2164	2.2160	2.2162	2.2192	2.2157	2.2156 (2.2129)	2.2146 (2.2153)	2.2204
	r(H4–C5)	1.0715	1.0714	1.0717	1.0716	1.0719	1.0719	1.0717 (1.0717)	1.0717 (1.0717)	1.0720
	r(C5–N6)	1.1526	1.1528	1.1521	1.1522	1.1540	1.1540	1.1529 (1.1522)	1.1529 (1.1523)	1.1581

[a] Extrapolations using the n^{-3} formula are reported in parentheses. [b] The acronym stands for fc-CCSD(T)-F12/cc-pVDZ-F12.

not only for interaction energies[23] but also for structural determinations.

Having excluded the $\Delta\alpha$ composite schemes, we can now move to address the reliability of the other models. In general terms, the performances of the junChS and junChS-F12 variants are comparable, while the junCBS+CV-F12 approach gives typically different bond distances, which can be either longer or shorter than the junChS and junChS-F12 counterparts. Analyzing in detail the trends, one may observe that for some species, the convergence to the CBS limit of the CCSD(T)-F12 method is not what we expected. Indeed, for $NH_3\cdots H_2O$, $H_2O\cdots C_2H_4$, $HCN\cdots HCN$, and $CH_4\cdots HF$, the intermolecular bond distance increases when going from the jun-cc-pVTZ to the jun-cc-pVQZ basis set, a behavior which is opposite to that systematically observed for the conventional CCSD(T) and MP2 methods and for MP2-F12. From our analysis, it can be concluded that CCSD(T)-F12 geometries do not seem to benefit from extrapolation to the CBS limit, either with the n^{-3} or n^{-5} formula, and actually, the extrapolation can lead to unreliable corrections. The former point confirms the fact that F12 methodologies approach in a steep manner the CBS limit and that the convergence is achieved with relatively small basis sets. This is also confirmed by the fact that the MP2-F12 corrective term for incorporating the CBS limit in ChS-F12 often leads to corrections smaller than 0.1 mÅ. In this respect, we can mention the $CH_4\cdots H_2O$ and $C_2H_4\cdots NH_3$ complexes. Interestingly, the comparison between CCSD(T)-F12/jun-cc-pVTZ and junCBS+CV-F12 geometries points out that the difference is negligible or can be attributed to the CV term. This means that the simplest and "safest" route to obtain accurate structural parameters when relying on CCSD(T)-F12 calculations is to add CV corrections, evaluated at either the MP2-F12 or CCSD(T)-F12 level, to CCSD(T)-F12/jun-cc-pVTZ geometries.

From the inspection of Table S6 of the Supporting Information, it is observed that the junChS and junChS-F12 models show the best performances, with the CBS-georef structures used as references. For intermolecular distances, the average deviation from the latter is −0.01 Å. Half of such a discrepancy can be attributed to the CV term, which is not included in the CBS-georef reference structures and, for the junChS and junChS-F12 approaches, indeed amounts—on average—to −0.005 Å. Even if not accounting for the diffuse function effects, the ChS shows a performance only negligibly degraded with respect to that of junChS and junChS-F12. The CBS and CV contributions on the geometrical parameters within the ChS, junChS, and junChS-F12 models are collected in Table S7 of the Supporting Information. From the inspection of this table, we note that—on average—the CBS correction to the intermolecular distances is larger for ChS and junChS than for junChS-F12. While for the two schemes involving conventional methods, the CBS contribution is often in the order of a few hundredths of Å, for junChS-F12, this term is generally smaller than 0.01 Å. Furthermore, the CBS contribution is larger for intermolecular distances than for the intramolecular counterparts. Moving to the CV term, there is not a noticeable difference between intra- and intermolecular parameters. In the case of distances, the correction is always in the order of a few mÅ. Interestingly, the revDSD geometries show a good accuracy, thus supporting the choice made in the definition of our ChS/ChS-F12 approaches. The average deviation, in absolute terms, from the CBS-georef structures is 0.016 Å, with maximum discrepancies of up to 0.03 Å. While there is a

systematic trend in the deviations of the junChS, junChS-F12, and ChS models from CBS-georef, the same does not apply to revDSD, the differences being either positive or negative.

While the geometry schemes discussed above do not incorporate the CP corrections, four paradigmatic complexes (well representing different types of interactions) have been selected to address (i) the difference between CP- and NCP-corrected geometries and (ii) the difference between the conventional and explicitly correlated (F12) CBS+CV schemes. The results are collected in Table 10. As far as the comparison between CP and NCP geometries is concerned, the differences are very small, that is, in the order of few mÅ. This gives support to the choice of defining geometry schemes without CP corrections. The conclusion is that the computational cost due to the incorporation of CP corrections is not necessary owing to the extrapolation to the CBS limit performed in all ChS/ChS-F12 approaches. From Table 10, it is also noted that the differences between conventional and F12 CBS+CV schemes are small as well, these also being in the order of few mÅ. The last comment concerns the fc-CCSD(T)-F12/cc-pVDZ-F12 level. Its performance is nearly as good as that of composite approaches, with deviations of about 0.01 Å for intermolecular distances.

In summary, the extrapolation to the CBS limit of the CCSD(T)-F12 geometries is non-trivial, and the analysis of the convergence of such methods must be carefully checked. The extrapolation formula employed, either n^{-3} or n^{-5}, introduces an error of 3−5 mÅ, which is close to the error associated with the use of CP or NCP geometries. While F12 basis sets have been purposely tailored for F12 methodologies, it seems that there is no convenience to employ them in comparison with the "seasonal" jun-cc-pVnZ sets. Indeed, for the latter family, triple- and quadruple-ζ basis sets have sizes comparable to those of the double- and triple-ζ sets of the F12 family. However, the "seasonal" basis sets guarantee a solid extrapolation to the CBS limit, also including in the most consistent way diffuse functions. Furthermore, this equivalence between the two families of basis sets is confirmed by the comparison between CCSD(T)-F12/jun-cc-pVTZ (see Table S6) and CCSD(T)-F12/cc-pVDZ-F12 (Table 10) geometries. These give results with an agreement better than 6 mÅ, with the cc-pVDZ-F12 basis set giving typically shorter bonds. A final note concerns the full CCSD(T) schemes (conventional and F12) using the aug-cc-pVnZ-F12 basis sets (see Table S8). It is apparent that their performance is very similar to that of junCBS+CV-F12, however, at an increased computational cost.

In conclusion, our suggestion is to employ conventional methods within composite schemes to account for the extrapolation to the CBS limit and the CV correction, such as the junChS model, and to employ the F12 explicitly correlated methods within simple geometry optimizations, possibly improved by the inclusion of the CV term in an additive manner. In view of the incorporation of the CV correction and the use of a larger basis set for the CCSD(T) term, we can consider that the junChS model provides improved results with respect to the CBS-georef counterparts.

Semi-experimental Equilibrium Intermolecular Parameters. Owing to the direct connection between rotational constants and the molecular structure, microwave spectroscopy can be effectively exploited to obtain accurate geometrical parameters for molecular systems with a nonvanishing dipole moment and sufficiently stable in the gas phase.[60] While a pure experimental approach cannot be exploited in the majority of

Table 11. Structural Parameters[a] for the FA···H₂O Complex (Distances in Å and Angles in Degrees)

HCONH₂	revDSD	junChS	junChS-F12	r_{SE}[b]
r(C−H)	1.1028	1.0994	1.0995	1.097(3)
r(C=O)	1.2144	1.2088	1.2084	1.212(2)
θ(NCH)	112.56	112.69	112.73	112.4(13)
r(C−N)	1.3582	1.3539	1.3533	1.354(2)
θ(NCO)	124.73	124.56	124.55	124.2(5)
r(NH$_{trans}$)	1.0035	1.0004	1.0003	1.017(2)
θ(CNH$_{trans}$)	121.11	121.11	121.09	120.5(2)
r(NH$_{cis}$)	1.0061	1.0030	1.0030	1.008(2)
θ(CNH$_{cis}$)	119.37	119.21	119.19	119.9(1)
H₂O	revDSD	junChS	junChS-F12	r_{SE}[c]
r(H−O)	0.9610	0.9563	0.9565	0.9573(1)
θ(HOH)	104.46	104.49	104.55	104.53(1)

| | | | | r_{SE} | | | |
| | | | | | TM | | |
HCONH₂···H₂O	revDSD	junChS	junChS-F12	fixed[d]	revDSD	junChS	junChS-F12
r(C=O)	1.2246	1.2192	1.2190		1.2223	1.2224	1.2226
r(N−C)	1.3474	1.3432	1.3430		1.3431	1.3433	1.3437
r(C−H)	1.1002	1.0971	1.0970		1.0945	1.0947	1.0945
r(N−H$_{trans}$)	1.0035	1.0004	1.0000		1.0171	1.0170	1.0167
r(N−H$_{cis}$)	1.0121	1.0094	1.0090		1.0020	1.0143	1.0140
r(Ow−H$_{cis}$)*	2.0522	2.0328	2.0300	2.128(3)	2.136(3)	2.122(3)	2.122(3)
r(O−H$_c$)	0.9746	0.9694	0.9700		0.9708	0.9704	0.9708
r(O−H$_b$)	0.9592	0.9547	0.9550		0.9554	0.9557	0.9587
θ(NCO)	125.09	124.93	124.92		124.57	124.57	124.57
θ(HCN)	113.43	113.51	113.54		113.27	113.22	113.21
θ(H$_{trans}$NC)	120.17	120.13	120.12		119.56	119.52	119.52
θ(H$_{cis}$NC)	119.66	119.53	119.54		120.19	120.22	119.55
θ(O$_w$H$_{cis}$N)*	136.10	135.48	136.39	133.6(1)	133.6(1)	134.0(1)	134.0(1)
θ(HOH)	106.68	106.76	106.79		106.75	106.81	106.78
θ(H$_c$O$_{wat}$H$_{cis}$)*	82.44	83.31	83.13	71.8(4)	69.5(4)	69.2(5)	69.2(5)
σ				6.4×10^{-3}	6.83×10^{-3}	7.96×10^{-3}	7.96×10^{-3}

[a]Intermolecular parameters are denoted with asterisks. [b]Taken from ref 58. [c]Taken form ref 59. [d]The parameters not reported are fixed at the values of the corresponding r_e^{SE} values of the isolated monomers.

cases,[35] the so-called semi-experimental (SE) approach[61] allows the determination of equilibrium structures of experimental quality, the so-called SE equilibrium structures, r_e^{SE}. These are obtained from a least-squares fit of the SE equilibrium rotational constants, which are, in turn, derived from the experimental ground-state counterparts by subtracting vibrational corrections. The latter terms can be very effectively computed by hybrid density functionals in conjunction with double-/triple-ζ basis sets (B3LYP-D3(BJ)/SNSD[62,63]) in the present case).[64] Whenever experimental information is available for a sufficient number of isotopologues, a complete structural determination is possible, but this is rarely the case for molecular complexes. As a consequence, the application of the SE approach in such cases implies fixing the intramolecular parameters at those of the isolated fragments and then fitting the most significant intermolecular parameters.

In the framework of this work, it is worthwhile testing whether ChS/ChS-F12 models are able to deliver accurate geometrical parameters to be employed in the determination of r_e^{SE}s of molecular complexes. To this end, we have selected two representative complexes, namely, formamide−water (FA···H₂O) and dimethylsulfide-sulfur dioxide (DMS···SO₂). Let us start from the structures of the isolated molecules. For all of them, r_e^{SE} determinations are feasible, the results being collected in Tables 11 and 12 for FA···H₂O and DMS···SO₂, respectively,

with the atom labeling being provided in Figure 4. In these tables, the optimized geometries at the revDSD, junChS, and junChS-F12 levels are also reported. From the inspection of Tables 11 and 12, it is evident that the junChS and junChS-F12 models provide results very similarly to the r_e^{SE} counterparts, with discrepancies of a few mÅ for bond lengths and a few tenths of degree for valence angles. The close agreement between the conventional and explicitly correlated versions of the junChS composite scheme is also retained for the intermolecular parameters of both the studied complexes, with, however, increased absolute deviations from the r_e^{SE}. In addition, these tables further confirm the reliability of the revDSD model.

Before further proceeding with the discussion of the SE equilibrium structures, it is necessary to explain how they have been determined. As mentioned above, when there is a lack of experimental data, the intramolecular parameters can be fixed to those of the isolated monomers. This is the case of the results of Tables 11 and 12 denoted as "fixed", with the r_e^{SE} values of the monomers being employed. In the other cases, we resorted to the template molecule (TM) approach[64] to evaluate the parameter values to be kept fixed (r_e(fixed))

$$r_e(\text{fixed}) = r_e^{\text{level}} + \Delta r_e(\text{TM}) \qquad (8)$$

where r_e^{level} is the value of a generic intramolecular parameter of the molecular complex evaluated at a given level of theory

Table 12. Structural Parameters[a] for the DMS···SO₂ Complex (Distances in Å and Angles in Degrees)

$(CH_3)_2S$	revDSD	junChS	junChS-F12	r_{SE}[b]
$r(C-S)$	1.8058	1.7984	1.7977	1.79863(13)
$\theta(CSC)$	98.69	98.43	98.49	98.58000(81)
$r(H_2-C)$	1.0900	1.0867	1.0870	1.08857(38)
$\theta(H_2CS)$	107.38	107.45	107.46	107.4196(69)
$r(H_1-C)$	1.0910	1.0879	1.088	1.08972(47)
$\theta(H_1CS)$	110.87	110.75	110.79	110.688(29)
$\phi(H_1CSH_2)$	−118.88	−118.98	−118.96	−119.053(44)
$\phi(CSXX)$	130.66	130.78	130.76	129.15
$\phi(H_1CSC)$	−61.12	−61.02	−61.04	−60.95

SO_2	revDSD	junChS	junChS-F12	r_{SE}[c]
$r(S-O)$	1.4421	1.4298	1.4288	1.4307858(15)
$\theta(OSO)$	119.28	119.29	119.24	119.329872(81)

| | | | | r_{SE} | | | |
| | | | | | | TM | |
$(CH_3)_2S\cdots SO_2$	revDSD	junChS	junChS-F12	fixed[d]	revDSD	junChS	junChS-F12
$r(S1-S2)$*	2.9288	2.9570	2.9632	2.944(2)	2.943(3)	2.945(2)	2.944(2)
$r(S1-O)$	1.4500	1.4266	1.4358		1.4387	1.4276	1.4378
$\theta(OS1S2)$*	94.92	94.18	94.21	95.5(1)	95.4(1)	95.1(1)	95.4(1)
$\phi(O-S1-S2-X)$	−121.07	−120.76	−121.08		−121.01	−120.11	−120.81
$r(C-S2)$	1.8029	1.7928	1.7958		1.7957	1.7930	1.7968
$\theta(CS2S1)$*	91.52	91.16	91.1	91.5(1)	91.6(1)	91.6(1)	91.6(1)
$\phi(C-S2-S1-X)$*	130.17	130.30	130.32	130.7(3)	130.5(3)	130.7(3)	130.6(3)
$r(H1-C)$	1.0912	1.0886	1.0887		1.0898	1.0904	1.0904
$\theta(H1CS3)$	110.60	110.66	110.63		110.42	110.60	110.53
$\phi(H1CS3C)$	−63.89	−64.34	−64.00		−63.72	−64.26	−63.90
$r(H2-C)$	1.0895	1.0866	1.0865		1.0881	1.0885	1.0881
$\theta(H2CS3)$	107.25	107.35	107.35		107.29	107.32	107.31
$\phi(H2CS2H1)$	−118.69	−118.74	−118.73		−118.85	−118.81	−118.82
$r(H3C)$	1.0901	1.0873	1.0872		1.0887	1.0891	1.0889
$\theta(H3CS2)$	109.93	109.94	109.94		109.75	109.89	109.84
$\phi(H3CS2C)$	58.563	58.05	58.36		58.39	57.98	58.27
σ				4.8×10^{-3}	6.3×10^{-3}	4.4×10^{-3}	5.1×10^{-3}

[a]Intermolecular parameters are denoted with asterisks. [b]Taken from ref 65. [c]Taken from ref 66. [d]The parameters not reported are fixed at the values of the corresponding r_e^{SE} values of the isolated monomers.
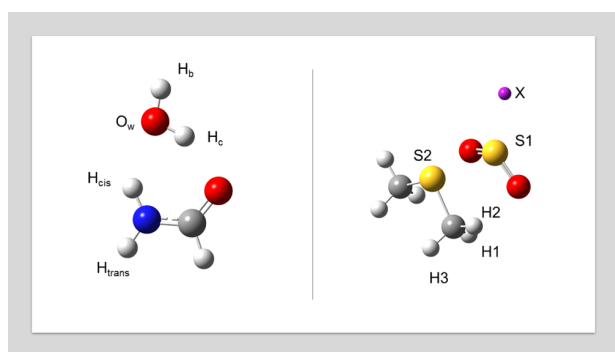

Figure 4. FA···H₂O and DMS···SO₂ complexes: structure and atom labeling.

(revDSD, junChS, or junChS-F12, in the present case) and $\Delta r_e(TM)$ is correction to be applied based on the corresponding value within the isolated monomer

$$\Delta r_e(TM) = r_e^{SE} - r_e^{level} \qquad (9)$$

Focusing on the various SE equilibrium structures reported in Tables 11 and 12, it is interesting to note that using the revDSD, junChS, or junChS-F12 level in the TM approach leads to very

similar results. This suggests that whenever exploiting this methodology for accurate structural determinations of molecular complexes, there is no need of composite schemes, but instead, a very cost-effective level of theory such as revDSD is largely sufficient.

junChS and junChS-F12 Databases. The previous sections have demonstrated the reliability and the adequate accuracy of the revDSD-optimized geometries as reference structures, which have thus been retained for setting up a database of accurate interaction energies. The CP, NCP, and half-CP interaction energies for the A14 complexes using both the junChS and junChS-F12 are collected in Table 13. While the statistics for these two approaches have already been discussed, from this table, it is apparent that in absolute terms, the differences between the two models are only fractions of kJ mol⁻¹. It is interesting to note that CP and NCP results are always very similar, with differences in almost all cases well within 0.1 kJ mol⁻¹, thus pointing out the effectiveness of the extrapolation to the CBS limit in recovering the BSSE. However, the CP−NCP differences are systematically smaller for the junChS-F12.

To extend the representativeness of both our benchmark study and the database to be set up, we have selected some molecules including third-row atoms either taken from our

Table 13. Interaction Energies of the A14 Complexes with revDSD Reference Geometries[a]

	junChS			junChS-F12		
	CP	NCP	half-CP	CP	NCP	half-CP
$H_2O\cdots H_2O$	−21.10	−21.37	−21.24	−21.00	−21.11	−21.05
$NH_3\cdots NH_3$	−13.30	−13.34	−13.32	−13.20	−13.19	−13.20
$HF\cdots HF$	−19.45	−19.59	−19.52	−19.25	−19.41	−19.33
$HCN\cdots HCN$	−19.88	−19.70	−19.79	−19.84	−19.99	−19.91
$CH_4\cdots CH_4$	−2.25	−2.22	−2.24	−2.20	−2.16	−2.18
$CH_2O\cdots CH_2O$	−19.23	−19.43	−19.33	−18.96	−19.10	−19.03
$C_2H_4\cdots C_2H_4$	−4.75	−4.78	−4.77	−4.66	−4.70	−4.68
$H_2O\cdots C_2H_4$	−10.86	−11.01	−10.93	−10.77	−10.80	−10.79
$H_2O\cdots CH_4$	−2.78	−2.79	−2.79	−2.77	−2.75	−2.76
$H_2O\cdots NH_3$	−27.57	−27.69	−27.63	−27.47	−27.52	−27.50
$NH_3\cdots CH_4$	−3.24	−3.24	−3.24	−3.23	−3.22	−3.22
$NH_3\cdots C_2H_4$	−5.89	−5.99	−5.94	−5.82	−5.84	−5.83
$HF\cdots CH_4$	−7.13	−7.14	−7.13	−7.04	−7.06	−7.05
$C_2H_4\cdots CH_2O$	−6.94	−7.07	−7.01	−6.83	−6.91	−6.87

[a]Energies in kJ mol^{-1}.

Table 14. junChS and junChS-F12 Interaction Energies (revDSD Reference Geometry) for the B9 and C6 Datasets[a]

	junChS			junChS-F12			literature
	CP	NCP	half-CP	CP	NCP	half-CP	
B9 Dataset							
$FH_2P\cdots H_2S$	−14.92	−15.42	−15.17	−14.80	−14.89	−14.85	−14.94[b]
$FH_2P\cdots NH_3$	−29.46	−30.03	−29.74	−29.08	−29.24	−29.16	−29.54[b]
$H_2O\cdots H_2S$	−12.18	−12.41	−12.30	−12.23	−12.25	−12.24	
$H_2O\cdots PH_3$	−10.77	−10.95	−10.86	−10.76	−10.77	−10.76	
$CH_3NH_2\cdots HCl$	−52.27	−52.81	−52.54	−52.32	−52.41	−52.37	
$OCS\cdots CH_4$	−4.36	−4.45	−4.41	−4.25	−4.29	−4.27	
$OCS\cdots H_2O$	−7.96	−8.09	−8.02	−7.87	−7.89	−7.88	
$SO_2\cdots H_2S$	−12.22	−12.45	−12.34	−12.11	−12.20	−12.15	
$DMS\cdots SO_2$	−33.53	−34.40	−33.97	−32.97	−33.36	−33.17	−34.04[b]
C6 Dataset							
cyclopentene$\cdots H_2O$	−16.59	−16.71	−16.65	−16.37	−16.44	−16.40	−12.8[c]
$H_2O\cdots$peptide	−35.16	−35.54	−35.35	−34.91	−35.05	−34.98	−33.89[d]; −32.31[e]
$CH_3NH_2\cdots CH_3NH_2$	−17.82	−17.94	−17.88	−17.64	−17.67	−17.66	−17.41[d]; −15.23[e]
$CH_3NH_2\cdots$pyridine	−20.17	−20.27	−20.22	−19.94	−19.99	−19.96	−16.61[d]; −13.19[e]
$CH_3OH\cdots$pyridine	−28.18	−28.65	−28.42	−28.08	−28.21	−28.15	−31.00[d]; −28.57[e]
pyridine$\cdots$pyridine	−16.32	−16.58	−16.45	−16.12	−16.26	−16.19	−16.32[d]; −9.97[e]

[a]Values in kJ mol^{-1}. [b]Ref 23: at the junChS level. [c]Vibrational ground-state dissociation energy taken from ref 35: ChS CP-corrected electronic energy augmented by harmonic zero-point energy at the B2PLYP-D3/maug-cc-pVTZ-*d*H level. [d]Ref 67: see text for the description of the level of theory. [e]Ref 68: CCSD/CBS energy computed as the sum of HF/aug-cc-pVQZ energy and PNO-CCSD-F12/aug-cc-pVTZ energy.

previous work[23] or purposely selected for this study (B9 dataset). Furthermore, some additional larger molecular complexes (C6 dataset) have been chosen, most of them from the S66 dataset,[67] in order to improve the coverage for the H, C, N, and O atoms.[67] The results are collected in Table 14, where the junChS and junChS-F12 models (both employing revDSD reference geometries) are compared, with CP, NCP, and half-CP interaction energies being considered. In Table 14, previously available data are also provided. In this respect, it should be noted that the junChS results for $FH_2P\cdots H_2S$, $FH_2P\cdots NH_3$, and $DMS\cdots SO_2$ were already reported in ref 23; however, small discrepancies (<0.1 kJ mol^{-1}) can be noted due to the use of a different reference geometry (B2PLYP-D3/maug-cc-pVTZ in ref 23). The situation is different for the complexes taken from the S66 dataset, for which differences as large as 3 kJ mol^{-1} can be observed. Actually, this is somewhat expected, the level of theory employed in the S66 dataset being less accurate than our junChS and junChS-F12 approaches. Indeed, in ref 67, starting

from HF/aug-cc-pVQZ energies, the CBS contribution is evaluated at the MP2 level using the aug-cc-pVTZ and aug-cc-pVQZ basis sets, and the effect of triple excitations is incorporated via the CCSD(T)-MP2 energy difference computed using the aug-cc-pVDZ basis set. Furthermore, the CV correction is not included. The cyclopentene−water complex was previously investigated at the ChS level;[36] thus, the improvement is due to the balanced incorporation of the diffuse function effects.

Focusing on the comparison between the junChS and junChS-F12, inspection of Table 14 shows that the two approaches provide very similar results, with a maximum difference of 0.6 kJ mol^{-1} for CP-corrected interaction energies and 0.8 kJ mol^{-1} for the NCP counterparts. Furthermore, it is noted that in absolute terms, the CP-corrected junChS interaction energies are usually larger than the junChS-F12 counterparts by about 0.1−0.2 kJ mol^{-1}. As expected in view of the extrapolation to the CBS limit, the CP corrections are small.

As noted for the A14 dataset, the CP−NCP energy differences are slightly more pronounced for schemes involving conventional methods. Indeed, the CP correction to junChS-F12 interaction energies is usually as low as 0.1 kJ mol⁻¹, the exception being the DMS−SO₂ complex, for which it is about 0.4 kJ mol⁻¹. This suggests that the CP correction can be safely neglected when employing the junChS-F12 model, with remarkably saving computer time. The average CP correction, ∼0.3 kJ mol⁻¹, is quite small also for the conventional junChS model. However, it is comparable (if not larger) to the expected error for this approach; therefore, its neglect is not recommended. The CP correction of the interaction energy of DMS−SO₂ is particularly large, that is, 0.9 kJ mol⁻¹, also when considering the conventional model.

Another remarkable aspect of the results for the B9 dataset is that the trends for molecules containing third-row atoms are extremely similar to those only bearing first/second-row elements, with the only exception being the DMS−SO₂ complex. For the systems of Table 14 previously investigated,[23,67] since the junChS-F12 model does not change the trends in any significant manner, we only recall that the present results are the most accurate currently available for both the B9 and C6 dataset.

In conclusion, based on the results of this study and those of ref 23, the junChS and junChS-F12 models allowed us to establish a database of accurate interaction energies at a limited computational cost, the actual version of this database consisting of the A14, B9, and C6 datasets. The present work has demonstrated the applicability of the junChS and junChS-F12 approaches to molecular complexes involving third-row atoms without any degradation in the accuracy and the possibility of its application to large systems. As a part of the database, the structural parameters of the A14 dataset are also provided. According to the discussion of a previous section, only the junChS and CCSD(T)-F12/jun-cc-pVTZ+CV(MP2/cc-pwCVTZ) levels have been retained. The database will be extended in order to incorporate more information for both interaction energies and equilibrium structures.

■ CONCLUDING REMARKS

Composite schemes based on explicitly correlated (F12) approaches for the description of noncovalent complexes have been defined, starting from the well-tested "cheap" methodology for conventional methods.[23] The performance has been analyzed in detail for systems ruled by different intermolecular interactions employing a representative dataset including atoms belonging to the first two rows of the periodic table. While these composite schemes have been first introduced for the accurate evaluation of interaction energies, they have then been extended to structural determinations.

Among the different approaches investigated, the so-called junChS-F12 model together with its conventional counterpart, junChS, and a modified version obtained by removing d and f polarization functions on hydrogen atoms, the jun-(d,f)H_ChS-F12 model, have been found to be the best-performing schemes. The junChS-F12 and jun-(d,f)H_ChS-F12 approaches, tested on the dataset denoted as A14, showed average relative errors of 0.68 and 0.65% (for CP-corrected interaction energies), respectively, which means—in absolute terms—average deviations of 0.05 kJ mol⁻¹. Interestingly, if the CP correction is neglected, thus further reducing the computational cost, the performances of the different "cheap" variants show a limited worsening, with the average relative errors of junChS-F12 and

jun-(d,f)H_ChS-F12 models increasing to 1.10 and 1.07%, respectively.

The present study confirms the outcomes of our previous work that led to the definition of the junChS:[23] (i) The inclusion of up to d diffuse functions on nonhydrogen atoms is mandatory in the CCSD(T) computational step for both conventional and explicitly correlated approaches. For this purpose, the aug-cc-pVDZ-F12 and jun-cc-pVTZ basis sets (which, as explained in the text, have comparable dimensions) are the smallest options for accurate results and, actually, deliver comparable results. (ii) The extrapolation to the CBS limit at the MP2(-F12) level improves significantly the accuracy of the results without any relevant increase in the computational cost for both conventional and explicitly correlated approaches. (iii) The CV correlation correction always represents a quite small contribution that cannot, however, be neglected in view of the very small absolute errors.

The junChS-F12 approach slightly outperforms its conventional junChS counterpart, the latter showing a relative error of 1.38%. For quite small systems, single-point junChS(-F12) computations require no more than twice the time of the underlying coupled-cluster step and are one order of magnitude faster than the CBS+CV counterparts. Enlarging the dimension of the systems increases the effectiveness of the junChS(-F12) model because of the favorable scaling of MP2(-F12) computations with respect to CCSD(T)(-F12), with the performance of MP2(-F12) being further enhanced using the resolution of identity and other acceleration techniques. Concerning the coupled-cluster step, inclusion of explicit correlation increases the computer time by 20% at most, a burden which is accompanied by a non-negligible improvement in accuracy. Replacing MP2-F12 with CCSD(T)-F12 in the extrapolation step (thus leading to the junCBS+CV-F12 model) increases the computational cost by at least one order of magnitude without improving the performance of the composite scheme. Indeed, for CP-corrected interaction energies, the average relative error increases from 0.68 to 0.79%, and the average absolute error increases from 0.05 to 0.07 kJ mol⁻¹.

Moving to the structural investigation, an important result is that the geometries optimized employing the recent rev-DSD-PBEP86 double-hybrid functional augmented by D3(BJ) empirical dispersion correction and in conjunction with the jun-cc-pVTZ basis set show a reasonably good accuracy in most situations. If improved accuracy is sought, the most effective option is to add the core−valence correlation correction, evaluated at the MP2-F12 level, to fc-CCSD(T)-F12/jun-cc-pVTZ geometries without accounting for the extrapolation to the CBS limit and the basis set superposition error. Indeed, F12 explicitly correlated methods deliver geometrical parameters close to the basis set convergence, thus leading to negligible basis set superposition errors.

Although the study of larger systems requires further developments, possibly related to the use of local-correlation treatments (e.g., PNO or DLPNO possibly including explicit correlation[68,69]), the different variants of the so-called "cheap" composite scheme described in the present paper already represent reliable and effective tools for the investigation of intermolecular interactions between biomolecule building blocks. In particular, both junChS and junChS-F12 models have been successfully applied to the C6 dataset, which involves rather large systems such as the pyridine dimer.

■ ASSOCIATED CONTENT

ⓈⒾ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c00869.

Statistics for the following composite schemes applied on top of the A14 dataset: junChS-F12 NCP energies, jun-(d,f)H_ChS-F12 NCP energies, junCBS+CV-F12 NCP energies, and junChS CP and NCP energies, full account on the geometrical parameters obtained for the A14 dataset for all the computational approaches considered, detailed analysis of the CBS and CV terms for the ChS, junChS, and junChS-F12 models, and CP and NCP geometries for the $H_2O\cdots H_2S$, $H_2O\cdots H_2O$, $HCN\cdots HCN$, and $CH_4\cdots CH_4$ complexes, obtained using the augF12CBS+CV and augF12CBS+CV-F12 composite approaches (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Cristina Puzzarini — *Dipartimento di Chimica "Giacomo Ciamician", Università di Bologna, I-40126 Bologna, Italy;* ⓞ orcid.org/0000-0002-2395-8532; Email: cristina.puzzarini@unibo.it

Vincenzo Barone — *Scuola Normale Superiore, I-56126 Pisa, Italy;* ⓞ orcid.org/0000-0001-6420-4107; Email: vincenzo.barone@sns.it

Authors

Jacopo Lupi — *Scuola Normale Superiore, I-56126 Pisa, Italy;* ⓞ orcid.org/0000-0001-6522-9947

Silvia Alessandrini — *Scuola Normale Superiore, I-56126 Pisa, Italy; Dipartimento di Chimica "Giacomo Ciamician", Università di Bologna, I-40126 Bologna, Italy*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.1c00869

Author Contributions

§J.L. and S.A. contributed equally to this work.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Hansen, J. C.; Francisco, J. S. Radical–Molecule Complexes: Changing Our Perspective on the Molecular Mechanisms of Radical–Molecule Reactions and their Impact on Atmospheric Chemistry. *ChemPhysChem* **2002**, *3*, 833–840.

(2) Klemperer, W.; Vaida, V. Molecular complexes in close and far away. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10584–10588.

(3) Riley, K. E.; Hobza, P. Noncovalent interactions in biochemistry. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2011**, *1*, 3–17.

(4) Tang, W.; Johnston, S.; Iggs, J. A.; Berry, N. G.; Phelan, M.; Lian, L.; Bacsa, J.; Xiao, J. Cooperative catalysis through noncovalent interactions. *Angew. Chem., Int. Ed.* **2013**, *52*, 1668–1672.

(5) Puzzarini, C.; Spada, L.; Alessandrini, S.; Barone, V. The challenge of non-covalent interactions: theory meets experiment for reconciling accuracy and interpretation. *J. Phys.: Condens. Matter* **2020**, *32*, 343002.

(6) Umadevi, D.; Panigrahi, S.; Sastry, G. N. Noncovalent Interaction of Carbon Nanostructures. *Acc. Chem. Res.* **2014**, *47*, 2574–2581.

(7) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

(8) Noga, J.; Bartlett, R. J. The Full CCSDT Model for Molecular Electronic Structure. *J. Chem. Phys.* **1987**, *86*, 7041–7050.

(9) Kállay, M.; Gauss, J. Approximate treatment of higher excitations in coupled-cluster theory. *J. Chem. Phys.* **2005**, *123*, 214105.

(10) Zhang, J.; Valeev, E. F. Prediction of Reaction Barriers and Thermochemical Properties with Explicitly Correlated Coupled-Cluster Methods: A Basis Set Assessment. *J. Chem. Theory Comput.* **2012**, *8*, 3175–3186.

(11) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639–9646.

(12) Martin-Drumel, M.-A.; Baraban, J. H.; Changala, P. B.; Stanton, J. F.; McCarthy, M. C. The hunt for elusive molecules: insights from joint theoretical and experimental investigations. *J. Chem. Phys.* **2019**, *25*, 7243–7258.

(13) Karton, A.; Martin, J. M. L. Prototypical $\pi - \pi$ dimers re-examined by means of high-level CCSDT(Q) composite ab initio methods. *J. Chem. Phys.* **2021**, *154*, 124117.

(14) Řezáč, J. Non-Covalent Interactions Atlas Benchmark Data Sets: Hydrogen Bonding. *J. Chem. Theory Comput.* **2020**, *16*, 2355–2368.

(15) Kříž, K.; Nováček, M.; Řezáč, J. Non-Covalent Interactions Atlas Benchmark Data Sets 3: Repulsive Contacts. *J. Chem. Theory Comput.* **2021**, *17*, 1548–1561.

(16) Møller, C.; Plesset, M. S. Note on an approximation treatment for many-electron systems. *Phys. Rev.* **1934**, *46*, 618.

(17) Puzzarini, C.; Barone, V. Extending the molecular size in accurate quantum-chemical calculations: The equilibrium structure and spectroscopic properties of uracil. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7189–7197.

(18) Puzzarini, C.; Biczysko, M.; Barone, V.; Peña, I.; Cabezas, C.; Alonso, J. L. Accurate molecular structure and spectroscopic properties of nucleobases: a combined computational−microwave investigation of 2-thiouracil as a case study. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16965−16975.

(19) Barone, V.; Biczysko, M.; Bloino, J.; Puzzarini, C. Accurate structure, thermodynamic and spectroscopic parameters from CC and CC/DFT schemes: the challenge of the conformational equilibrium in glycine. *Phys. Chem. Chem. Phys.* **2013**, *15*, 10094.

(20) Puzzarini, C.; Biczysko, M.; Barone, V.; Largo, L.; Peña, I.; Cabezas, C.; Alonso, J. L. Accurate characterization of the peptide linkage in the gas phase: A joint quantum-chemical and rotational spectroscopy study of the glycine dipeptide analogue. *J. Phys. Chem. Lett.* **2014**, *5*, 534−540.

(21) Spada, L.; Tasinato, N.; Vazart, F.; Barone, V.; Caminati, W.; Puzzarini, C. Noncovalent Interactions and Internal Dynamics in Pyridine−Ammonia: A Combined Quantum-Chemical and Microwave Spectroscopy Study. *Chem.—Eur. J.* **2017**, *23*, 4876−4883.

(22) Obenchain, D. A.; Spada, L.; Alessandrini, S.; Rampino, S.; Herbers, S.; Tasinato, N.; Mendolicchio, M.; Kraus, P.; Gauss, J.; Puzzarini, C.; Grabow, J. U.; Barone, V. Unveiling the sulfur-sulfur bridge: accurate structural and energetic characterization of a homo chalcogen inter-molecular bond. *Angew. Chem., Int. Ed.* **2018**, *57*, 15822−15826.

(23) Alessandrini, S.; Barone, V.; Puzzarini, C. Extension of the "Cheap" Composite Approach to Noncovalent Interactions: The jun-ChS Scheme. *J. Chem. Theory Comput.* **2020**, *16*, 988−1006.

(24) Kong, L.; Bischoff, F. A.; Valeev, E. F. Explicitly Correlated R12/F12 Methods for Electronic Structure. *Chem. Rev.* **2012**, *112*, 75−107.

(25) Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 methods: Theory and benchmarks. *J. Chem. Phys.* **2009**, *130*, 054104.

(26) Kesharwani, M. K.; Sylvetsky, N.; Köhn, A.; Tew, D. P.; Martin, J. M. L. Do CCSD and approximate CCSD-F12 variants converge to the same basis set limits? The case of atomization energies. *J. Chem. Phys.* **2018**, *149*, 154109.

(27) Werner, H.-J.; Adler, T. B.; Manby, F. R. General orbital invariant MP2-F12 theory. *J. Chem. Phys.* **2007**, *126*, 164102.

(28) Peterson, K. A.; Adler, T. B.; Werner, H.-J. Systematically convergent basis sets for explicitly correlated wavefunctions: The atoms H, He, B−Ne, and Al−Ar. *J. Chem. Phys.* **2008**, *128*, 084102.

(29) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(30) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, G. D. Perspectives on Basis Sets Beautiful: Seasonal Plantings of Diffuse Basis Functions. *J. Chem. Theory Comput.* **2011**, *7*, 3027−3034.

(31) Spackman, P. R.; Jayatilaka, D.; Karton, A. Basis set convergence of CCSD(T) equilibrium geometries using a large and diverse set of molecular structures. *J. Chem. Phys.* **2016**, *145*, 104101.

(32) Karton, A.; Spackman, P. R. Evaluation of density functional theory for a large and diverse set of organic and inorganic equilibrium structures. *J. Comp. Chem.* **2021**, *42*, 1590−1601.

(33) Barone, V.; Lupi, J.; Salta, Z.; Tasinato, N. Development and Validation of a Parameter-Free Model Chemistry for the Computation of Reliable Reaction Rates. *J. Chem. Theory Comput.* **2021**, *17*, 4913−4928.

(34) Puzzarini, C.; Bloino, J.; Tasinato, N.; Barone, V. Accuracy and interpretability: the devil and the holy grail. New routes across old boundaries in computational spectroscopy. *Chem. Rev.* **2019**, *119*, 8131−8191.

(35) Puzzarini, C.; Barone, V. Diving for accurate structures in the ocean of molecular systems with the help of spectroscopy and quantum chemistry. *Acc. Chem. Res.* **2018**, *51*, 548−556.

(36) Wang, J.; Spada, L.; Chen, J.; Gao, S.; Alessandrini, S.; Feng, G.; Puzzarini, C.; Gou, Q.; Grabow, J. U.; Barone, V. The Unexplored World of Cycloalkene-Water Complexes: Primary and Assisting Interactions Unraveled by Experimental and Computational Spectroscopy. *Angew. Chem., Int. Ed.* **2019**, *58*, 13935−13941.

(37) Frisch, M. J.; et al. *Gaussian 16*, Revision C.01.; Gaussian Inc.: Wallingford CT, 2016.

(38) Werner, H.-J.; et al. The Molpro quantum chemistry package. *J. Chem. Phys.* **2020**, *152*, 144107.

(39) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 242−253.

(40) Feller, D.; Peterson, K. A. An expanded calibration study of the explicitly correlated CCSD(T)-F12b method using large basis set standard CCSD(T) atomization energies. *J. Chem. Phys.* **2013**, *139*, 084110.

(41) Adler, T. B.; Knizia, G.; Werner, H.-J. A simple and efficient CCSD(T)-F12 approximation. *J. Chem. Phys.* **2007**, *127*, 221106.

(42) Kállay, M.; Horváth, R. A.; Nagy, P. R. Size-consistent explicitly correlated triple excitation correction. *J. Chem. Phys.* **2021**, *155*, 034107.

(43) Dunning, T. H.; Peterson, K. A.; Wilson, A. K. Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited. *J. Chem. Phys.* **2001**, *114*, 9244−9253.

(44) Santra, G.; Sylvetsky, N.; Martin, J. M. L. Minimally Empirical Double-Hybrid Functionals Trained against the GMTKN55 Database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123*, 5129−5143.

(45) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(46) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.

(47) Grimme, S. Semiempirical Hybrid Density Functional with Perturbative Second-order Correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(48) Barone, V.; Ceselin, G.; Fusè, M.; Tasinato, N. Accuracy Meets Interpretability for Computational Spectroscopy by Means of Hybrid and Double-Hybrid Functionals. *Front. Chem.* **2020**, *8*, 584203.

(49) Řezáč, J.; Hobza, P. Describing Noncovalent Interactions beyond the Common Approximations: How Accurate Is the "Gold Standard," CCSD(T) at the Complete Basis Set Limit? *J. Chem. Theory Comput.* **2013**, *9*, 2151−2155.

(50) Řezáč, J.; Dubecký, M.; Jurečka, P.; Hobza, P. Extensions and applications of the A24 data set of accurate interaction energies. *Phys. Chem. Chem. Phys.* **2015**, *17*, 19268−19277.

(51) Kesharwani, M. K.; Karton, A.; Sylvetsky, N.; Martin, J. M. L. The S66 noncovalent interactions benchmark reconsidered using explicitly correlated methods near the basis set limit. *Aust. J. Chem.* **2018**, *71*, 238−248.

(52) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19*, 553−566.

(53) Řezáč, J.; Hobza, P. Benchmark calculations of interaction energies of noncovalent complexes and their applications. *Chem. Rev.* **2016**, *116*, 5038−5071.

(54) Hill, J. G.; Mazumder, S.; Peterson, K. A. Correlation consistent basis sets for molecular core-valence effects with explicitly correlated wave functions: The atoms B−Ne and Al−Ar. *J. Chem. Phys.* **2010**, *132*, 054108.

(55) Peterson, K. A.; Dunning, T. H. Accurate correlation consistent basis sets for molecular core−valence correlation effects: The second row atoms Al−Ar, and the first row atoms B−Ne revisited. *J. Chem. Phys.* **2002**, *117*, 10548−10560.

(56) Li, W.; Spada, L.; Tasinato, N.; Rampino, S.; Evangelisti, L.; Gualandi, A.; Cozzi, P. G.; Melandri, S.; Barone, V.; Puzzarini, C. Theory meets experiment for noncovalent complexes: The puzzling case of pnicogen interactions. *Angew. Chem., Int. Ed.* **2018**, *57*, 13853−13857.

(57) Alonso, E. R.; Fusè, M.; León, I; Puzzarini, C.; Alonso, J. L.; Barone, V. Exploring the maze of cycloserine conformers in the gas phase guided by microwave spectroscopy and quantum chemistry. *J. Phys. Chem. A* **2021**, *125*, 2121−2129.

(58) Alessandrini, S.; Puzzarini, C. Structural and Energetic Characterization of Prebiotic Molecules: The Case Study of Formamide and Its Dimer. *J. Phys. Chem. A* **2016**, *120*, 5257−5263.

(59) Piccardo, M.; Penocchio, E.; Puzzarini, C.; Biczysko, M.; Barone, V. Semi-Experimental Equilibrium Structure Determinations by Employing B3LYP/SNSD Anharmonic Force Fields: Validation and Application to Semirigid Organic Molecules. *J. Phys. Chem. A* **2015**, *119*, 2058−2082.

(60) Gordy, W.; Cook, R. L. *Microwave Molecular Spectra*; Wiley, 1984.

(61) Pulay, P.; Meyer, W.; Boggs, J. E. Cubic force constants and equilibrium geometry of methane from Hartree−Fock and correlated wavefunctions. *J. Chem. Phys.* **1978**, *68*, 5077−5085.

(62) Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A: At., Mol., Opt. Phys.* **1988**, *38*, 3098.

(63) Barone, V.; Cimino, P.; Stendardo, E. Development and validation of the B3LYP/N07D computational model for structural parameter and magnetic tensors of large free radicals. *J. Chem. Theory Comput.* **2008**, *4*, 751−764.

(64) Piccardo, M.; Penocchio, E.; Puzzarini, C.; Biczysko, M.; Barone, V. Semi-experimental equilibrium structure determinations by employing B3LYP/SNSD anharmonic force fields: Validation and application to semirigid organic molecules. *J. Phys. Chem. A* **2015**, *119*, 2058−2082.

(65) Demaison, J.; Margulès, L.; Rudolph, H. D. Accurate determination of an equilibrium structure in the presence of a small coordinate: The case of dimethylsulfide. *J. Mol. Struct.* **2010**, *978*, 229−233.

(66) Demaison, J.; Liévin, J. Accuracy of the equilibrium structure of sulphur dioxide. *Mol. Phys.* **2021**, No. e1950857. in press

(67) Řezáč, J.; Riley, K. E.; Hobza, P. S66: a well balanced database of benchmark interaction energies relevant to biomolecular structures. *J. Chem. Theory Comput.* **2011**, *7*, 2427−2438.

(68) Schmitz, G.; Hättig, C.; Tew, D. P. Explicitly correlated PNO-MP2 and PNO-CCSD and their application to the S66 set and large molecular systems. *Phys. Chem. Chem. Phys.* **2014**, *16*, 22167−22178.

(69) Sandler, I.; Chen, J.; Taylor, M.; Sharma, S.; Ho, J. Accuracy of DLPNO-CCSD(T): Effect of Basis Set and System Size. *J. Phys. Chem. A* **2021**, *125*, 1553−1563.

Chapter 3

Reliable reaction rates at affordable computational cost

In this Chapter a slightly modified version of the original junChS[73] model chemistry is presented. It is proposed as an effective, reliable and parameter-free scheme for the computation of accurate reaction rates with special reference to astrochemical and atmospheric processes. Benchmarks with different sets of state-of-the-art energy barriers spanning a wide range of values show that, in the absence of strong multi-reference contributions, the proposed model outperforms the most well-known (and commonly used) model chemistries, reaching a sub-chemical accuracy without any empirical parameter and with non-prohibitive computer times. Some test cases show that geometries, energy barriers, zero point energies and thermal contributions computed at this level can be used in the framework of the AITSTME for obtaining accurate reaction rates. In this Chapter, objectives O1, O2, O3 are pursued.

Article

Development and Validation of a Parameter-Free Model Chemistry for the Computation of Reliable Reaction Rates

Vincenzo Barone,* Jacopo Lupi, Zoi Salta, and Nicola Tasinato

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: A recently developed model chemistry (jun-Cheap) has been slightly modified and proposed as an effective, reliable, and parameter-free scheme for the computation of accurate reaction rates with special reference to astrochemical and atmospheric processes. Benchmarks with different sets of state-of-the-art energy barriers spanning a wide range of values show that, in the absence of strong multireference contributions, the proposed model outperforms the most well-known model chemistries, reaching a subchemical accuracy without any empirical parameter and with affordable computer times. Some test cases show that geometries, energy barriers, zero point energies, and thermal contributions computed at this level can be used in the framework of the master equation approach based on the ab initio transition-state theory for obtaining accurate reaction rates.
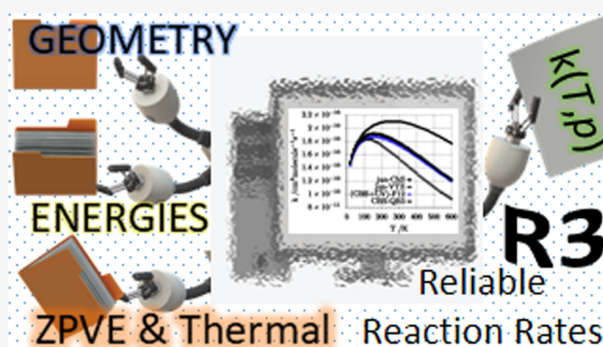
INTRODUCTION

For many years, scientists were skeptical about the presence of molecular systems in the interstellar space due to the harsh physical conditions (low temperature and pressure in the presence of high-energy radiation fields) characterizing this environment. However, contrary to these expectations, more than 200 molecules have now been identified in the interstellar and circumstellar medium (ISM),[1] including several so-called interstellar complex organic molecules (iCOMs), namely, molecules containing carbon and a total of more than six atoms.[2] Most of the observed species should have a very short lifetime according to Earth-based standards, but the intermolecular processes leading to thermodynamic equilibrium are not effective in the ISM due to its extreme physical parameters.[3,4] This situation calls for a strong interplay among observations, laboratory studies, and computational approaches to understand the chemical evolution in these regions and to explain the observed abundances of different species.

Astrochemical models are virtual laboratories including thousands of reactions and whose main goal is to reproduce the observational data to the best possible extent. Although the available astrochemical models show widely different degrees of sophistication,[5] all of them share the same basic ingredients:[6] a set of initial conditions (total density, temperature, etc.) and a panel of chemical reactions characterized by their respective temperature-dependent rate constants and most likely exit channels. To improve the current predictions provided by these models, the reactions responsible for the largest uncertainties on the abundances

must be studied in more detail by laboratory experiments and/or theoretical methods to provide improved rate constants and branching ratios.

Chemical kinetics plays a fundamental role also in the different but related context of atmospheric models that try to reproduce and interpret the large number of chemical processes occurring in the troposphere and stratosphere. Reaction rate coefficients and product yields have been either traditionally obtained by means of suitable experimental techniques[7] or estimated using structure−activity relationships.[8] The massive number of organic compounds released in the atmosphere and the corresponding huge number of possible reactions ruling their oxidation/degradation pathways make experimental measurements of even a small fraction of key processes a daunting task. In recent years, computational chemistry has begun to contribute substantially to a better understanding of several important reaction sequences in the atmosphere.[9] These contributions have, at their heart, the use of electronic structure calculations to determine the energies and other characteristics (mainly geometries and vibrational frequencies) of stable species, reactive complexes, and transition states, which are then used in theoretical frameworks

to determine rate coefficients. The main factor limiting the accuracy of this process is the computation of accurate values for all of the energy barriers ruling the different elementary steps. Next, zero point energies (ZPEs) and finite temperature contributions (FTCs) come into play, whose contributions may become non-negligible already for medium-sized systems.

Several nonempirical procedures have been developed and employed for the generation of accurate thermochemical data, which for small systems come close to the full configuration interaction (FCI) complete basis set (CBS) limit.[10] Among the most successful approaches are the Weizmann-n series (with the most accurate being W4[11]), the focal point analysis (FPA),[12,13] the Feller−Dixon−Peterson model (FDP),[14] and the extrapolated ab initio thermochemistry (HEAT) protocol.[15−17] A simplified version of the HEAT protocol is obtained by retaining only the extrapolation to the CBS limit at the CCSD(T) level and incorporating the core-valence corrections, thus leading to the model referred to in the following as CBS-CV. This approach is rather well tested in the literature and was shown to provide results with an accuracy well within 0.5 kcal mol^{-1}. Recently, alternative protocols have been proposed, which employ explicitly correlated approaches:[10,18] thanks to the faster convergence to the complete basis set limit, these approaches allow some computer time saving, but the rate-determining step remains the evaluation of higher-level contributions.

For larger molecular systems, more approximate composite methods are unavoidable, which aim at reaching the so-called chemical accuracy (1 kcal mol^{-1}). The most well-known among these so-called model chemistries are the last versions of the Gn[19] (G4[20]) and CBS-x[21] (CBS-QB3[22]) families. However, all of these models include some empirical parameters and employ geometries, which are not fully reliable for transition states and noncovalent complexes ruling the entrance channels of most reactions of astrochemical and atmospheric interest. As a matter of fact, the most reliable protocols (e.g., HEAT) push geometry optimizations to the limit to obtain accurate energetics, whereas, at the other extreme, Gn and CBS-x schemes employ B3LYP geometries, whose accuracy is often unsatisfactory.[23]

In the last few years, a reliable and accurate computational protocol, referred to as the cheap scheme (ChS) and devoid of any empirical parameter, has been developed and tested with remarkable success for structural and energetic data.[24−26] In conjunction with geometries and harmonic frequencies issuing from double-hybrid functionals, ChS has given promising results also for the activation energies of some reactions of astrochemical interest.[27−31] More recently, an improved variant (referred to as the jun-Cheap scheme, jChS) has been introduced, which, thanks to the use of the "june" partially augmented basis set of the "calendar" family,[32] provides very accurate results also for noncovalent interactions.[33,34] On these grounds, in this paper, we provide a comprehensive benchmark of the jChS model chemistry for several classes of reactions for which accurate reference results are available or have been purposely computed. Together with electronic energies, we analyze also zero point energies, thermal contributions to enthalpies and entropies, and overall reaction rates computed for elementary reactions in the framework of the master equation (ME) approach based on the ab initio transition-state theory (AITSTME).[35−37]

The paper is organized as follows. In the first part, we validate the jChS model chemistry with reference to some well-known databases: (i) the 24 energy barriers available in the latest updated version of the DBH24 database,[38] (ii) the 52 barriers of Truhlar's HTBH38[39] and NHTBH38[40] databases not included in DBH24, and (iii) seven representative reactions from Karton's BH28 database.[41] When needed, the reference values are updated by new computations performed with a composite method closely resembling the W3.2 model.[42]

Next, the reliability of the jChS model chemistry for zero point energies and thermal contributions to enthalpies and entropies is assessed with respect to the new databases THCS21 and THOS10 containing accurate reference values for closed- and open-shell systems, respectively.

Finally, the role of different contributions in determining the overall accuracy of computed reaction rates is analyzed by means of some simple elementary reactions and two more complex reaction networks relevant for astrochemistry and atmospheric chemistry. Conclusions and perspectives are given in the last section.

■ COMPUTATIONAL DETAILS

All of the composite schemes employed in the present work extrapolate single-point energies computed at suitable geometries (see next sections) using the cc-pV(n + d)Z (hereafter nZ)[43] or jun-cc-pV(n + d)Z (hereafter jnZ)[32] families of basis sets. The coupled cluster (CC) ansatz including single, double, and (perturbative) triple excitations (CCSD(T))[44] within the frozen-core approximation and in conjunction with 3Z or j3Z basis sets is always employed in the first step. Next, CBS extrapolation and core-valence correlation (CV) are added using either MP2[45] (leading, in conjunction with jnZ basis sets, to our standard jChS model) or CCSD(T). In the latter case, inclusion of higher-level terms (diagonal Born−Oppenheimer,[46−49] scalar relativistic,[50,51] full triple and perturbative quadruple excitations[52−54]) and systematic use of nZ basis sets lead to the CBS-CVH scheme.

The effect of spin−orbit coupling is added to the energies of the O, OH, SH, and Cl radicals, lowering their electronic energies by 0.22, 0.20, 0.54, and 0.84 kcal mol^{-1}, respectively.[55]

Vibrational contributions are always obtained by the rev-DSDPBEP86-D3(BJ) double-hybrid functional,[56] in conjunction with the j3Z basis set (hereafter rev-DSD). Harmonic frequencies are computed by analytical second derivatives[57] and anharmonic corrections, when needed, by the generalized second-order vibrational perturbation theory (GVPT2) employing third- and semidiagonal fourth derivatives obtained by numerical differentiation of second derivatives implemented by one of the present authors in Gaussian software.[58−60]

All of the computations have been performed with the Gaussian code,[60] except for CCSD(T) geometry optimizations that have been carried out with the Molpro package,[61] CCSDT or CCSDT(Q) energy evaluations with the MRCC program,[62] and DBOC together with relativistic computations with the CFOUR code.[63]

jChS Model Chemistry. The jChS total electronic energies are obtained by single-point computations at rev-DSD geometries

$$E_{jChS} = E(CCSD(T)/j3Z) + \Delta E_{MP2}^{CBS} + \Delta E_{CV} \qquad (1)$$

where the CBS term is

Table 1. Theoretical Values of Barrier Heights in the DBH24/08 Data Set Obtained at Different Levels of Theory[e]

	reactions	forward/reverse barrier height			
		CCSD(T)	jChS	jChS[a]	ref[b]
	Heavy-Atom Transfer				
a1[c]	H• + N₂O → OH• + N₂	17.89/84.96	17.53/83.25	17.58/83.27	17.13/82.47
a2	H• + ClH → HCl + H•	18.89/18.89	17.31/17.31	17.33/17.33	18.00/18.00
a3[c]	CH₃• + FCl → CH₃F + Cl•	7.21/62.20	7.16/60.37	7.05/60.28	6.75/60.00
	Nucleophilic Substitution				
a4	Cl⁻···CH₃Cl → ClCH₃···Cl⁻	13.56/13.56	13.26/13.26	13.28/13.28	13.41/13.41
a5	F⁻···CH₃Cl → FCH₃···Cl⁻	3.52/29.47	3.39/29.09	3.41/29.09	3.44/29.42
a6	OH⁻ + CH₃F → HOCH₃ + F⁻	−2.39/17.78	−2.48/17.36	−2.51/17.35	−2.44/17.66
	Unimolecular and Association				
a7	H• + N₂ → HN₂•	15.23/11.01	14.34/11.12	14.36/11.09	14.36/10.61
a8	H• + C₂H₄ → C₂H₅•	2.43/42.59	1.9/42.19	1.92/42.21	1.72/41.75
a9	HCN ↔ HNC	47.45/32.77	47.98/33.24	48.02/33.28	48.07/32.82
	Hydrogen Transfer				
a10[d]	OH• + CH₄ → CH₃• + H₂O	7.05/19.05	6.63/20.04	6.52/19.94	6.71/19.60
a11[c,d]	H• + OH• →H₂ + ³O	10.38/14.62	11.51/13.77	11.42/13.78	10.71/13.12
a12[c]	H• + H₂S → H₂ + HS•	4.23/19.23	3.7/17.94	3.69/17.96	3.62/17.33
	MAX	2.49	0.80	0.80	
	MUE	0.71	0.36	0.35	
	RMSD	0.97	0.44	0.43	

[a]At QCISD/MG3 geometries. [b]Ref 38. [c]Spin−orbit contributions on the reverse reaction barrier. [d]Spin−orbit contributions on the forward reaction barrier. [e]All of the values (exclusive of ZPE) are in kcal mol⁻¹.

$$\Delta E_{\mathrm{MP2}}^{\mathrm{CBS}} = \frac{4^3 E(\mathrm{MP2/j4Z}) - 3^3 E(\mathrm{MP2/j3Z})}{4^3 - 3^3} - E(\mathrm{MP2/j3Z}) \tag{2}$$

and the core-valence correction ΔE_{CV} is the MP2 energy difference between all electron (ae) and frozen-core (fc) calculations employing the cc-pwCVTZ basis set.[64] At this level, the extrapolation of Hartree−Fock (HF) and correlation contributions is performed with the same equation and basis sets since several tests have shown that this simplified recipe has a negligible impact on the overall accuracy of the results. Furthermore, scalar relativistic effects are neglected, which is not a serious approximation since the heaviest element involved in this study is Cl.

CBS-CVH Composite Scheme. The CBS-CVH total electronic energies are obtained from single-point computations at geometries optimized by the jChS composite method described above for energies

$$E_{\mathrm{tot}} = E_{\mathrm{HF}}^{\mathrm{CBS}} + \Delta E_{\mathrm{CCSD(T)}}^{\mathrm{CBS}} + \Delta E_{\mathrm{CV}} + \Delta E_{\mathrm{fT}} + \Delta E_{\mathrm{pQ}}$$
$$+ \Delta E_{\mathrm{rel}} + \Delta E_{\mathrm{DBOC}} \tag{3}$$

In this case, HF and correlation energies are extrapolated separately. In particular, the HF CBS limit is estimated using Feller's exponential formula[65]

$$E_{\mathrm{HF}}(n) = E_{\mathrm{HF}}^{\mathrm{CBS}} + B \exp(-Cn) \tag{4}$$

whereas the CBS limit of the correlation energy is obtained by the n^{-3} formula proposed by Helgaker and co-workers[66]

$$\Delta E_{\mathrm{corr}}(n) = \Delta E_{\mathrm{corr}}^{\mathrm{CBS}} + An^{-3} \tag{5}$$

The three-point extrapolation of HF energies employs 3Z, 4Z, and 5Z basis sets, whereas the two smaller basis sets are used in the two-point extrapolation of correlation energies. The core-valence correction ΔE_{CV} is computed as the CCSD(T) energy difference between all electron and frozen-core calculations employing the cc-pCVTZ basis set.[64]

The diagonal Born−Oppenheimer correction ΔE_{DBOC}[46−49] and the scalar relativistic contribution to the energy ΔE_{rel}[50,51] are computed at the HF-SCF/aug-cc-pVDZ and CCSD(T)/aug-cc-pCVDZ levels, after having checked their convergence with respect to contributions calculated with triple-ζ basis sets for a few stationary points.

Finally, the corrections due to full treatment of triple (ΔE_{fT}) and perturbative treatment of quadruple (ΔE_{pQ}) excitations are computed, within the fc approximation, as energy differences between CCSDT and CCSD(T) and between CCSDT(Q) and CCSDT calculations employing the cc-pVTZ and cc-pVDZ basis sets, respectively.

Kinetic Models. Global and channel-specific rate constants were computed solving the multiwell one-dimensional master equation using the chemically significant eigenvalue (CSE) method within the Rice−Ramsperger−Kassel−Marcus (RRKM) approximation.[67] The collisional energy-transfer probability is described using the exponential down model[68] with a temperature-dependent ΔE_{down} of $260 \times (T/298)^{0.875}$ cm⁻¹ in an argon bath gas.

For channels ruled by a distinct saddle point, rate coefficients are determined by the conventional transition-state theory (TST) within the rigid-rotor harmonic-oscillator (RRHO) approximation[69] and including tunneling as well as nonclassical reflection effects using the Eckart model.[70] Instead, rate constants for barrierless elementary reactions are computed employing the phase space theory (PST),[71,72] again within the RRHO approximation. The isotropic attractive potential V_{eff} entering the PST is described by a $\frac{C}{R^6}$ power law, whose C coefficient is obtained by fitting rev-DSD energies computed at various long-range distances of fragments. We obtained the following C coefficients for the PST calculations of barrierless channels: 230 $a_0^6 E_{\mathrm{h}}$ for the H₂S + Cl entrance channel, 64.2 $a_0^6 E_{\mathrm{h}}$ for the CH₃NH₂ + CN entrance channel on the methyl side, and 94.4 $a_0^6 E_{\mathrm{h}}$ for the CH₃NH₂ + CN entrance channel on the nitrogen side.

Table 2. Theoretical Values of Barrier Heights for the Forward and Reverse Reactions in the NHTBH38/08 Data Set Not Included in the DBH24 Selection[d]

	reaction		forward/reverse barrier height		
		jChS	jChS[a]	ref 38	
NHT1	$H^\bullet + FH \rightarrow HF + H^\bullet$	41.99/41.99	42.02/42.02	42.18/42.18	
NHT2	$H^\bullet + FCH_3 \rightarrow HF + CH_3^\bullet$	30.31/57.54	30.31/57.54	30.38/57.02	
NHT3*	$H^\bullet + F_2 \rightarrow HF + F^\bullet$	3.50/107.18[b]	1.49/105.25	2.27/105.80	
NHT4	$F^- + CH_3F \rightarrow FCH_3 + F^-$	−0.70/−0.70	−0.71/-0.71	−0.34/−0.34	
NHT5	$F^-\cdots CH_3F \rightarrow FCH_3\cdots F^-$	13.21/13.21	13.20/13.20	13.38/13.38	
NHT6	$Cl^- + CH_3Cl \rightarrow ClCH_3 + Cl^-$	2.27/2.27	2.33/2.33	3.10/3.10	
NHT7	$F^- + CH_3Cl \rightarrow FCH_3 + Cl^-$	−12.32/19.29	−12.31/19.31	−12.54/20.11	
NHT8	$OH^-\cdots CH_3F \rightarrow HOCH_3\cdots F^-$	11.14/47.38	11.14/47.38	10.96/47.20	
NHT9	$H^\bullet + CO \rightarrow HCO^\bullet$	3.22/22.87	3.19/22.82	3.17/22.68	
NHT10	$CH_3^\bullet + C_2H_4 \rightarrow CH_3CH_2CH_2^\bullet$	6.37/32.77	6.35/32.74	6.85/32.97	
	MAX[c]	0.83	0.80		
	MUE[c]	0.33	0.32		
	RMSD[c]	0.42	0.40		

[a]jChS on QCISD/MG3 geometry. [b]Employing restricted open-shell geometry; the values using the unrestricted geometry are 4.46/108.14. [c]Neglecting the problematic reaction NHT3 (marked with an asterisk; see the text for discussion). [d]All of the values (exclusive of ZPE) are in kcal mol^{-1}.

The rate constants of the overall reactions evaluated in different temperature ranges are fitted by the three-parameter modified Arrhenius equation proposed by Kooij[73,74]

$$k(T) = A \left(\frac{T}{300} \right)^n \exp\left(-\frac{E}{RT} \right) \tag{6}$$

where A, n, and E are the fitting parameters and R is the universal gas constant.

■ RESULTS AND DISCUSSION

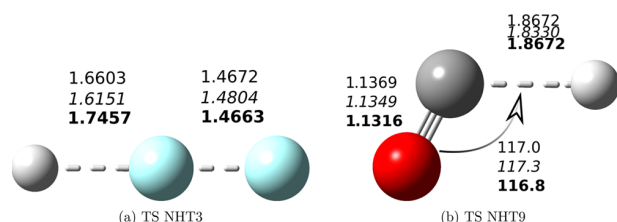In the original jChS model, geometries and force fields were computed with the B2PLYP double-hybrid functional[75]

Figure 1. Sketch of the structures of the transition states ruling the reactions $H^\bullet + F_2 \rightarrow HF + F^\bullet$ (NHT3) and $H^\bullet + CO \rightarrow HCO^\bullet$ (NHT9). The key geometrical parameters issuing from rev-DSD, QCISD/MG3 (italics), and jChS (bold) geometry optimizations are also reported. Bond distances are in Å, and angles are in degrees. The following colors are used for the different atom types: white, H; black, C; red, O; and light blue, F.

augmented by empirical dispersion contributions (namely, the D3(BJ) model)[76,77] in conjunction with partially augmented triple-zeta basis sets.[33] However, the recently developed rev-DSD model[56] delivers improved descriptions of noncovalent interactions and activation energies.[78,79] Therefore, we benchmarked the performances of this functional (still in conjunction with partially augmented triple-zeta basis sets) for geometrical parameters and vibrational frequencies, obtaining results close to those delivered by the CCSD(T) ansatz in conjunction with comparable basis sets, but at a much reduced computational cost.[80] As a consequence, the jChS

model chemistry now uses by default rev-DSD geometries and force fields.

If the spin contamination from higher spin states is large, the potential energy surfaces computed by unrestricted wave functions can be significantly distorted, showing, for example, anomalously high reaction barriers.[81] This means that UMP2 estimates of CBS and CV contributions in the jChS model could become problematic. On the other hand, CCSD fully eliminates the $S + 1$ contaminant[82] and CCSD(T) reduces also the $S + 2$ contaminant[83] so that calculations at the CCSD(T) level are usually relatively insensitive to the choice of (restricted or unrestricted) orbitals.[84] However, in cases where higher spin contaminants become important, CCSD(T) can also fail.[83] On these grounds, all of the jChS and CBS-CVH energies have been computed by the restricted open-shell approach.

Concerning density-functional theory (DFT) methods, it is well-known that the extent of spin contamination in unrestricted versions of hybrid density functionals increases with the amount of HF exchange.[85] However, Menon and Radom[86] showed that in unrestricted double-hybrid procedures, the opposing behavior of UHF and UMP2 with respect to spin contamination leads to smaller differences between the energies predicted by unrestricted and restricted open-shell variants. Although rev-DSD energies are not used in the present context, spin contamination can have an effect also on gradients and Hessians. We have, therefore, checked systematically the spin contamination and found that its effect is always negligible (within the target accuracy of the jChS model) except for the CN radical and the transition state ruling the reaction $H^\bullet + F_2 \rightarrow HF + F^\bullet$, which will be analyzed in detail in a following section.

Reaction Barriers. The most well-known database of accurate reaction barriers is the DBH24 compilation[38,87] containing results mostly obtained at the CCSDTQ5/CBS level via the W4 theory[88] for a statistically representative set including three prototypes for each of the following classes of reactions: heavy-atom transfer, nucleophilic substitution, unimolecular and association reactions, and hydrogen-transfer reactions.

Table 3. Theoretical Values of Barrier Heights for the Forward and Reverse Reactions in the HTBH38/08 Data Set Not Included in the DBH24 Selection[e]

	reaction		forward/reverse barrier height		
		jChS	jChS[a]	ref 38
HT1*[b]	H• + HCl → H₂ + Cl•	4.97/7.80	5.57/7.95	5.49/7.42
HT2[c]	OH• + H₂ → H₂O + H•	5.67/21.76	5.58/21.69	5.10/21.20
HT3	CH3• + H₂ → CH₄ + H•	11.96/14.64	11.95/14.64	12.10/15.30
HT4	H• + H₂ → H₂ + H•	9.58/9.58	9.58/9.58	9.60/9.60
HT5*[c]	OH• + NH₃ → H₂O + NH₂•	4.13/14.41	3.55/13.85	3.20/12.70
HT6[b]	HCl + CH₃• → Cl• + CH₄	1.69/7.19	1.70/7.61	1.70/7.90
HT7[c]	OH• + C₂H₆ → H₂O + C₂H₅•	4.00/20.91	3.84/20.75	3.40/19.90
HT8	F• + H₂ → HF + H•	1.69/33.90	1.77/34.00	1.80/33.40
HT9*[b,c]	³O + CH₄ → OH• + CH₃•.	14.77/9.83	14.87/9.82	13.70/8.10
HT10*	H• + PH₃ → PH₂• + H₂	2.85/25.09	2.82/25.05	3.10/23.20
HT11*[b,c]	³O + HCl → OH• + Cl•	10.81/11.38	10.85/11.70	9.80/10.40
HT12*	NH₂• + CH₃• → CH₄ + NH	9.49/22.09	9.50/22.11	8.00/22.40
HT13*	NH₂• + C₂H₅ → NH + C₂H₆	9.97/19.08	10.39/19.51	7.50/18.30
HT14	NH₂• + C₂H₆ → NH₃ + C₂H₅•	11.24/17.85	11.18/17.80	10.40/17.40
HT15	NH₂• + CH₄ → NH₃ + CH₃•	13.82/16.94	13.80/16.92	14.50/17.80
HT16*	s−*trans cis*-C₅H₈ → same	39.66/39.66	39.63/39.63	38.40/38.40
	MAX[d]	1.01	0.88	
	MUE[d]	0.48	0.42	
	RMSD[d]	0.58	0.52	

[a]jChS on QCISD/MG3 geometry. [b]Spin−orbit corrections on the reverse reaction barrier. [c]Spin−orbit corrections on the forward reaction barrier. [d]Neglecting the problematic reactions (marked with an asterisk). [e]All of the values (exclusive of ZPE) are in kcal mol⁻¹.
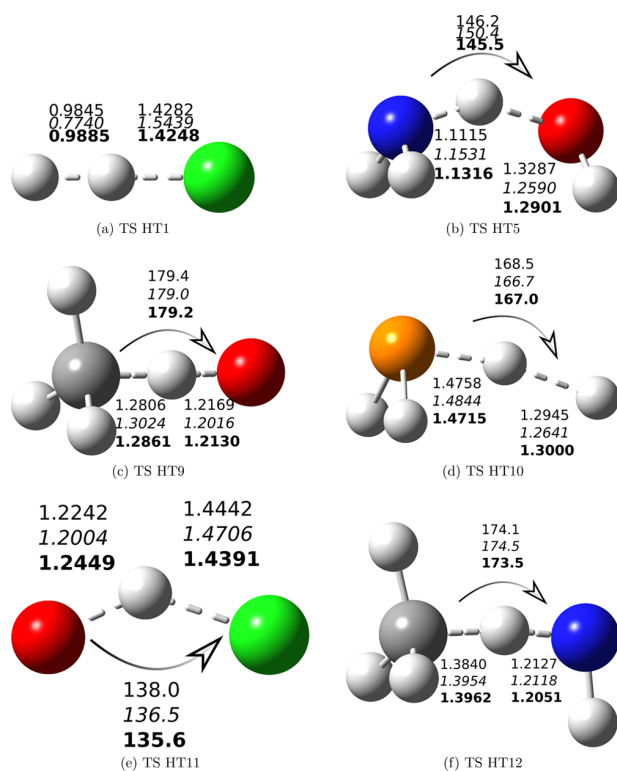

Figure 2. Sketch of the structures of the transition states ruling the reactions collected in Table 4. The key geometrical parameters issuing from rev-DSD, QCISD/MG3 (italics), and jChS (bold) geometry optimizations are also reported. Bond distances are in Å, and angles are in degrees. The following colors are used for the different atom types: white, H; black, C; blue, N; red, O; orange, P; and green, Cl.

Table 1 compares the reaction barriers computed at CCSD(T) and jChS levels to the reference values of ref 38. The arithmetic (mean unsigned error—MUE) and geometric (root-mean-square deviation—RMSD) average errors show that the jChS model chemistry fulfills the target of subchemical accuracy without any outlier above 1 kcal mol⁻¹ (max error = 0.80 kcal mol⁻¹). It is also remarkable that estimation of CBS and CV contributions by inexpensive MP2 computations and without any empirical parameter halves the error of the underlying CCSD(T) computation. To investigate the role of geometries on the computed barriers, we repeated the computations using the QCISD/MG3 structures employed in the original compilation.[38] It is quite apparent that in this case the results are only marginally affected by geometry optimizations at different computational levels. We will come back to this aspect in the following since the situation could be different for more complex transition structures and/or the noncovalent complexes ruling the entrance channels of barrierless reactions. In this connection, further support to the reliability of rev-DSD structures is provided by the respectable MUE and RMSD (1.7 and 2.4 kcal mol⁻¹, respectively) of the energy barriers computed at this level.

Zhang and co-workers[18] have shown that, for the same set of reactions, inclusion of explicit correlation (F12) in CCSD(T) computations[89] reduces the mean and maximum unsigned errors of the conventional CCSD(T) approach (0.66 and 1.77 kcal mol⁻¹) to 0.29 and 0.85 kcal mol⁻¹ when using basis sets slightly larger than j3Z (including also f diffuse functions on non-hydrogen atoms). As shown in Table 1, this improvement is close to that obtained when going from CCSD(T)/j3Z (0.71 and 2.49 kcal mol⁻¹) to jChS (0.36 and 0.80 kcal mol⁻¹). These trends suggest that inclusion of explicit correlation or two-point extrapolation at the MP2 level is an effective route for improving significantly the accuracy of computed energy barriers, without introducing additional computational bottlenecks with respect to the underlying CCSD(T)/j3Z reference.

Table 4. Theoretical Values of the Forward and Reverse Barriers Ruling the "Challenging" HTBH38/08 Reactions[d]

geometry		QCISD		rev-DSD	jChS		
	forward/reverse barrier	ref 38	jChS[a]	jChS	jChS	CBS-CV	CBS-CVH
HT1[a]	$H^\bullet + HCl \rightarrow H_2 + Cl^\bullet$	5.49/7.42	5.57/7.95	4.97/7.80	4.97/7.85	5.25/8.23	5.41/8.19
HT5[b]	$OH^\bullet + NH_3 \rightarrow H_2O + NH_2^\bullet$	3.20/12.70	3.55/13.85	4.13/14.41	4.41/14.60	4.40/14.34	4.39/13.63
HT9[a,b]	$^3O + CH_4 \rightarrow OH^\bullet + CH_3^\bullet$	13.70/8.10	14.87/9.82	14.77/9.83	14.93/9.76	14.70/9.37	14.64/9.30
HT10	$H^\bullet + PH_3 \rightarrow PH_2^\bullet + H_2$	3.10/23.20	2.82/25.05	2.85/25.09	2.85/25.12	2.87/24.48	2.89/24.52
HT11	$^3O + HCl \rightarrow OH^\bullet + Cl^\bullet$	9.80/10.40	10.85/11.70	10.81/11.38	10.67/11.43	10.93/11.34	10.27/10.98
HT12	$NH_2^\bullet + CH_3^\bullet \rightarrow CH_4 + NH$	8.00/22.40	9.50/22.11	9.49/22.09	8.94/21.84	8.87/21.89	9.24/22.26
	MAX[c]	1.32	0.84	0.57	0.60	0.66	
	MUE[c]	0.74	0.39	0.34	0.34	0.20	
	RMSD[c]	0.87	0.46	0.38	0.38	0.28	

[a]Spin−orbit corrections on the reverse reaction barrier. [b]Spin−orbit corrections on the forward reaction barrier. [c]Neglecting the reverse barrier of reaction HT5. [d]All of the values (exclusive of ZPE) are in kcal mol^{-1}.

(a) b1 BHPERI1   (b) b2 CRBH1   (c) b3 CRBH4

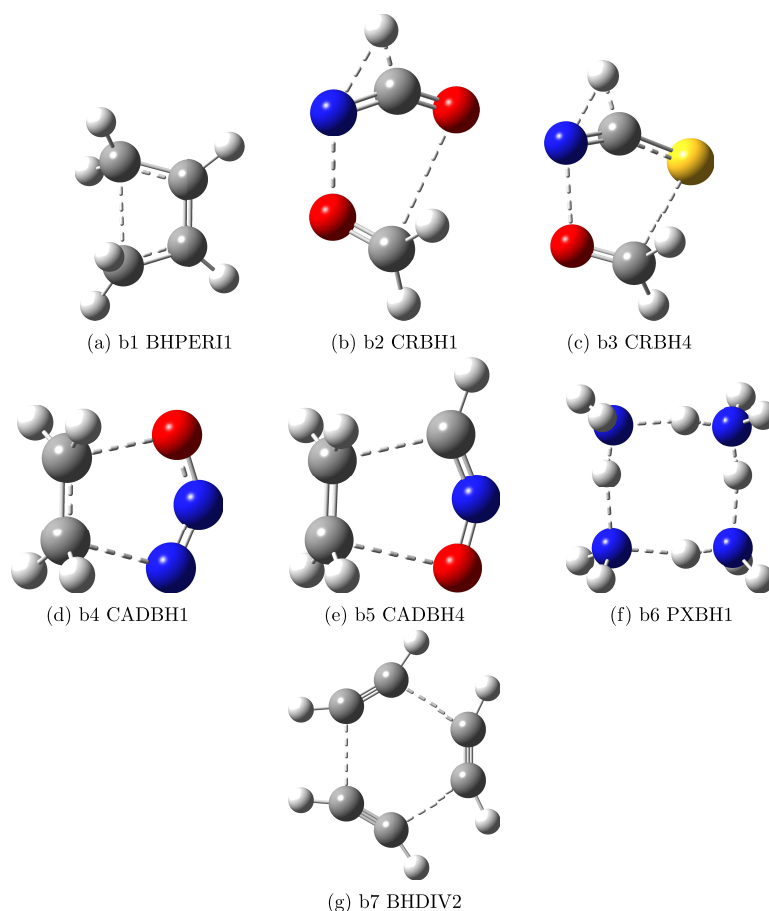(d) b4 CADBH1   (e) b5 CADBH4   (f) b6 PXBH1

(g) b7 BHDIV2

Figure 3. Sketch of the structures of the transition states ruling the reactions of Table 5. The following colors are used for the different atom types: white, H; black, C; blue, N; red, O; and yellow, S.

As a matter of fact, already for reactions involving two heavy atoms (e.g., A7, A8, A9, A10 in Table 1), single-point jChS computations require no more than twice the time of the CCSD(T)/jun-cc-pVTZ step and are an order of magnitude faster than the CBS-CV counterparts. On increasing the dimensions of the systems, the effectiveness of the jChS model increases because of the favorable scaling of MP2 computations with respect to CCSD(T) ones, which can be further enhanced by approaches employing resolution of identity and other acceleration techniques. Furthermore, jChS computations can be performed also with the widely diffused electronic

structure codes lacking explicitly correlated approaches (e.g., Gaussian or CFOUR), and the accuracy of the results surpasses that of all of the model chemistries considered by Zheng et al.[38]

Two larger databases of prototypical reactions are also available for barriers related to transfers of hydrogen and non-hydrogen atoms (HTBH38[39] and NHTBH38,[40] respectively). However, the reaction barriers not already included in the DBH24 set have been obtained at a lower computational level (W1). We have thus decided to compute at the jChS level all of the reactions of the above two sets not contained in the

Article

Table 5. Theoretical Values of Barrier Heights for the Forward and Reverse Reactions in the BH14 Data Set, Obtained at Different Levels of Theory[h]

	label in BH28	forward/reverse barrier height	forward reaction barrier height		
		jChS	jChS[a]	W3lite-F12[b]	CCSDT(Q)-CCSD(T)
b1	BHPERI1[c]	35.07/95.13		35.01	−0.17
b2	CRBH1[d]	47.24 (78.59)	47.01	46.15	−1.10
b3	CRBH4[d]	46.54 (64.14)	46.12	44.89	−1.60
b4	CADBH1[e]	27.26/36.08		27.56	0.00
b5	CADBH4[e]	11.57/57.52		11.64	−0.24
b6	PXBH1[f]	48.59/48.59		48.45	−0.12
b7	BHDIV2[g]	51.15/201.05		50.10	−0.14
	MAX	1.65			
	MUE	0.62			
	RMSD	0.86			

[a]Using the geometries of ref 41. [b]Ref 41. [c]Refs 90, 91. [d]Ref 92. [e]Refs 91, 93. [f]Ref 94. [g]Ref 95. [h]All of the values (exclusive of ZPE) are in kcal mol^{-1}.

Table 6. ThCS21 Database: ZPEs in kcal mol^{-1} and Absolute Entropies at 298.15 K and 1 atm in cal (mol K)$^{-1}$

molecule	ZPE_{harm}[a]	ZPE_{anh}[a,b]	ZPE_{exp}[c]	S_{harm}[a]	S_{exp}[d]
HF	5.89	5.84	5.86	41.46	41.50
HCl	4.30	4.27	4.24	44.57	44.64
H_2	6.36 (6.24)	6.30	6.23	31.13	31.20
N_2	3.33	3.32	3.36	45.77	45.77
F_2	1.42	1.41	1.30	48.33	48.44
CO	3.09	3.08	3.09	47.24	47.21
Cl_2	0.81	0.81	0.80	53.18	53.29
CO_2	7.26	7.23	7.30	51.09	51.07
CS_2	4.36	4.35	4.34	56.78	56.85
H_2O	13.45 (13.21)	13.24	13.26	45.09	45.10
H_2S	9.59	9.47	9.48	49.12	49.16
HOF	8.77 (8.65)	8.64	8.65	54.11	54.17
HOCl	8.31 (8.19)	8.19	8.19	56.47	56.49
N_2O	6.84	6.80	6.77	52.51	52.54
HCN	10.03 (9.91)	9.95	10.00	48.16	48.21
SO_2	4.33	4.31	4.41	59.35	59.30
C_2H_2	16.72 (16.48)	16.56	16.49	47.91	47.99
H_2CO	16.76 (16.52)	16.54	16.52[e]	52.23	52.30
NH_3	21.63 (21.27)	21.26	21.20	45.98	46.04
CH_4	28.20 (27.72)	27.79	27.71	44.48	44.48
C_2H_4	32.06 (31.58)	31.67	31.46[f]	52.35	52.39
MUE	0.15 (0.04)	0.05		0.05	
RMSD	0.22 (0.06)	0.07		0.06	

[a]rev-DSD-PBEP86-D3(BJ)/jun-cc-pV(T + d)Z. [b]HDCPT2 model. [c]From ref 96. [d]From ref 104. The original values have been lowered by 0.03 cal (mol K)$^{-1}$ to take into account the passage from 1 bar (0.1 MPa) to 1 atm (0.10135 MPa) references. [e]From accurate diffusion Monte Carlo computations[105] since the value of 16.10 reported in ref 96 is affected by an estimated error of 0.51 kcal mol^{-1}. [f]From the accurate computations of ref 106 since the value of 30.70 reported in ref 96 is affected by an estimated error of 0.40 kcal mol^{-1}.

original DBH24 compilation using both rev-DSD and the original QCISD/MG3 geometries. Whenever significant discrepancies were found, the reactions were recomputed also at the CBS-CVH level.

The reactions from the NHTBH38 set not included in the DBH24 selection are collected in Table 2. It is noteworthy that

Table 7. Absolute Entropies at 298.15 K and 1 atm in cal (mol K)$^{-1}$

molecule	S_{harm}[a]	S_{HR}[a,b]	S_{exp}
CH_3CH_3	54.38	54.70	54.79[c,d]
CH_3OH	57.00	57.36	57.29[c,d]
CH_3SH	60.57	60.99	60.96[c,d]
CH_3CHO	62.66	63.11	63.06[c,d]
CHOCHO	64.93	65.09	65.10[c,e]

[a]rev-DSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z. [b]Including HR correction. [c]The original values have been lowered by 0.03 cal (mol K)$^{-1}$ to take into account the passage from 1 bar (0.1 MPa) to 1 atm (0.10135 MPa) references. [d]From ref 111. [e]From ref 112.

rev-DSD energy barriers, although not directly used in the jChS model chemistry, show MUEs smaller than 2.0 kcal mol^{-1}, thus suggesting that the corresponding geometries should be sufficiently accurate for single-point energy evaluations at higher computational levels. This is confirmed by the finding that only for reaction NHT3, QCISD and rev-DSD geometries lead to significantly different results (cf. columns 2 and 3 of Table 2). Geometry optimization at the jChS level provides results far from both values (Figure 1a). However, as mentioned in a previous section, unrestricted rev-DSD computations show a strong spin contamination for the TS ruling reaction NHT3 ($\langle S^2 \rangle$ = 1.03 in place of the correct value of 0.75). We have, therefore, reoptimized the geometry of this TS employing the restricted open-shell approach in conjunction with numerical energy derivatives. The issuing geometrical parameters (rHF = 1.6603, rFF = 1.4672 Å) are closer to the jChS counterparts (rHF = 1.7457, rFF = 1.4663 Å) than the unrestricted values (rHF = 1.5700, rFF = 1.4021 Å) and, indeed, even better than the QCISD/MG3 values of ref 38 (rHF = 1.6151, rFF = 1.4804 Å), thus giving further support to the accuracy of rev-DSD geometries. To check the accuracy of computed energies irrespective of geometry effects, we have recomputed the forward and reverse barriers of reaction NHT3 at the CBS-CVH level employing QCISD/MG3 geometries. The results (1.57 and 104.84 kcal mol^{-1}) are much closer to the jChS values (1.49 and 105.25, MUE = 0.25 kcal mol^{-1}) than to the results of ref 38 (2.27 and 105.80, MUE = 0.83 kcal mol^{-1}), thus confirming the reliability and robustness of the jChS model chemistry. However, in this case, fully reliable results can be obtained only employing more accurate geometries: as a matter of fact, the forward and reverse barriers obtained from single-point CBS-CVH computations at jChS geometries are 2.59 and 105.77 kcal mol^{-1}, respectively. The seemingly good agreement with the results of ref 38 is due to a fortuitous error compensation between poor geometry and limited accuracy of the electronic energy. With the exception of this reaction, the agreement between jChS energies and the reference values is satisfactory, suggesting that for this kind of reaction the jChS errors are in line with those discussed above for the DBH24 database.

The reactions from the HTBH38 set not included in the DBH24 selection are collected in Table 3. Once again, it is noteworthy that the rev-DSD energy barriers, although not directly used in the jChS model chemistry, do not show any unrealistic outlier. Only for reactions HT1 and HT5, QCISD and rev-DSD geometries lead to significantly different results (cf. columns 2 and 3 of Table 3). Geometry optimization at the jChS level provides results close to rev-DSD (Figure 2a) for HT1 and intermediate between rev-DSD and QCISD for

Table 8. ThOS10 Database: ZPEs and Nonpotential Energy Terms for Representative Open-Shell Species at 298.15 K and 1 atm in cal (mol K)$^{-1}$

molecule	ZPE$_{calc}$[a,b]	ZPE$_{exp}$[b]	S_{calc}[a,c]	S_{exp}[c,d]	$H-H^0_{calc}$[a,b]	$H-H^0_{exp}$[b,e]
OH($^2\pi$)	5.25 (5.33)	5.29[e]	43.95	43.88	2.07	2.11
SH($^2\pi$)	3.88 (3.88)	3.82	47.27	46.76	2.07	2.07
CN($^2\sum^{+}$,[f])	2.83 (2.83)	2.95	48.35	48.43	2.07	2.07
NO($^2\pi$)	2.80 (2.77)	2.71	50.42	50.34	2.07	2.07
NH$_2$(^{2}B$_1$)	11.89 (11.83)	11.52[g]	46.49	46.54	2.37	2.37
HCO(^{2}A')	8.06 (8.09)	8.09[h]	53.58	53.66	2.39	2.39[h]
HO$_2$(^{2}A'')	8.85 (8.87)	8.78[i]	54.67	54.76	2.39	2.39
CH$_3$(^{2}A$_2$'')	18.62 (18.42)	18.48[i]	46.26	46.38	2.46	2.45
t-HOCO(^{2}A')	13.00 (13.07)	13.10[i]	60.08		2.61	
CH$_3$CO(^{2}A')	26.82 (26.85)	26.69[i]	64.23	63.92	2.98	2.96

[a]rev-DSD-PBEP86-D3(BJ)/jun-cc-pV(T + d)Z HRHO model and (in parenthesis) HRHO model including a correction to ZPEs of −0.12 for each CH, NH, or OH bond. [b]In kcal mol^{-1}. [c]In cal (mol K)$^{-1}$. [d]From ref 113. When needed, entropy values have been lowered by 0.03 cal (mol K)$^{-1}$ to take into account the passage from 1 bar (0.1 MPa) to 1 atm (0.10135 MPa) references. [e]From ref 96. [f]Restricted open shell with an equilibrium bond length of 1.179 Å;; the unrestricted result is 3.43 kcal mol^{-1} with S^2 = 0.854 and an equilibrium bond length of 1.159 Å. [g]CBS-CV results from ref 114. [h]CBS-CV results from ref 115. [i]Diffusion Monte Carlo results from ref 105.
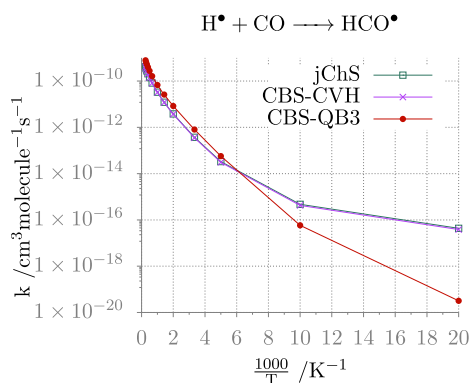
Figure 4. Temperature dependence of the H$^{\bullet}$ + CO reaction rate constants calculated at various levels of theory in the high-pressure limit.

HT5 (Figure 2b). The agreement between jChS energies and the reference values is generally worse than for the NHTBH38 set and particularly disappointing for reactions HT1, HT5, HT9, HT10, HT11, HT12, HT13, and HT16. To have a first check of the accuracy of the jChS results irrespective of geometry effects, the forward and reverse barriers of two reactions in this group (HT1 and HT12) have been recomputed at the CBS-CVH level on top of QCISD/MG3 geometries. In the first case, the CBS-CVH values (5.95 and 8.73 kcal mol^{-1}) are quite close to the results of both ref 38 and the jChS counterparts for the forward barrier and much closer to the jChS result for the reverse barrier. The situation is reversed for reaction HT12, where the CBS-CVH results (9.35 and 22.37 kcal mol^{-1}) confirm the similar results of jChS and ref 38 for the reverse barrier but are much closer to the jChS ones for the forward barrier. Once again, the jChS model chemistry does not show any outlier above the threshold of

chemical accuracy, whereas this is not the case for the original reference values of ref 38. For the forward and reverse barriers of the remaining eight reactions, the deviations of the jChS results from those of ref 38 are well within subchemical accuracy (MUE around 0.5 kcal mol^{-1}).

We then selected six "challenging reactions" among those mentioned above for further investigation. To this end, we report in Table 4 the results obtained at different geometries together with new reference values obtained at the CBS-CVH level on top of jChS geometries. A first general remark is that some of the new reference values differ by more than 1 kcal mol^{-1} from those reported in ref 38 (cf. columns 1 and 6 of Table 4). Furthermore, the only barrier showing significant contributions by higher-order terms (mainly full triple and perturbative quadruple excitations) is the reverse barrier of reaction HT5 (cf. columns 5 and 6 in Table 4). Neglecting this barrier, the results of ref 38 show an MUE of 0.82 kcal mol^{-1} and a maximum error of 1.61 kcal mol^{-1}, whereas the jChS approach has an MUE lower than 0.40 kcal mol^{-1} without any absolute error larger than 1 kcal mol^{-1}, irrespective of the level of geometry optimizations. As a matter of fact, the relatively cheap rev-DSD geometries can be confidently employed for reaching subchemical accuracy and the use of more accurate structures does not really improve the results. The CBS-CV approach reduces significantly the MUE, but at the price of employing more accurate (and costly) geometries together with CCSD(T) computations performed with partially augmented 4Z basis sets. In conclusion, the jChS model chemistry can be confidently employed for evaluating reaction barriers of all of the reactions included in the HTBH38 and NHTBH38 data sets with subchemical accuracy without any outlier above 1 kcal mol^{-1}.

To extend the benchmark to larger and more complex systems, we resorted to the BH28 set of ref 41, which includes accurate (W3lite-F12) energy barriers for several pericyclic

Table 9. Arrhenius−Kooij Parameters for the H$^{\bullet}$ + CO Reaction

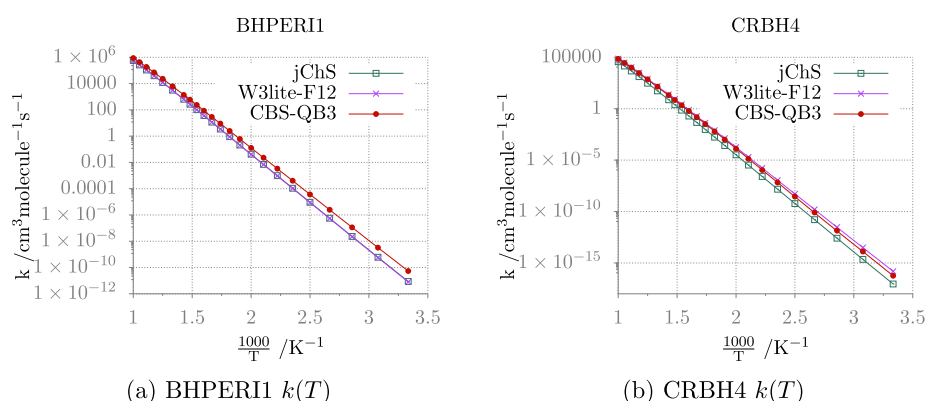| forward/reverse	ref 116	jChS	CBS-CVH	CBS-QB3
A (cm^3 molecule^{-1} s^{-1})	2.98 × 10^{-11}/1.37 × 10^{13}	3.86 × 10^{-11}/1.47 × 10^{13}	3.87 × 10^{-11}/1.46 × 10^{13}	7.96 × 10^{-11}/ 3.81 × 10^{13}
n	1.03/1.06	1.07/1.20	1.06/1.20	1.02/1.04
E (kcal mol^{-1})	2.64/17.79	2.86/18.14	2.89/18.08	2.76/17.95
rms		4× 10^{-14}/ 3.20 × 10^{-2}	3.9 × 10^{-14}/3.21× 10^{-2}	1.91 × 10^{-14}/ 2.56 × 10^{-2}

(a) BHPERI1 $k(T)$ (b) CRBH4 $k(T)$

Figure 5. Rate constant temperature-dependence plots of the BHPERI1 and CRBH4 reactions from the BH14 data set for a pressure of 1 atm.

Table 10. Arrhenius−Kooij Parameters for BHPERI1 and CRBH4 Reactions from the BH14 Data Set

	BHPERI1			CRBH4		
	jChS	W3lite-F12	CBS-QB3	jChS	W3lite-F12	CBS-QB3
A (cm^3 molecule^{-1} s^{-1})	7.43×10^{13}	8.18×10^{13}	1.26×10^{14}	1.24×10^{16}	2.58×10^{16}	1.14×10^{17}
n	-1.05	-1.12	-1.44	-2.81	-3.37	-4.02
E (kcal mol^{-1})	34.31	34.41	33.53	45.60	44.35	45.83
rms	7.67×10^{-2}	7.80×10^{-2}	1.06×10^{-1}	1.05×10^{-1}	1.08×10^{-1}	1.06×10^{-1}
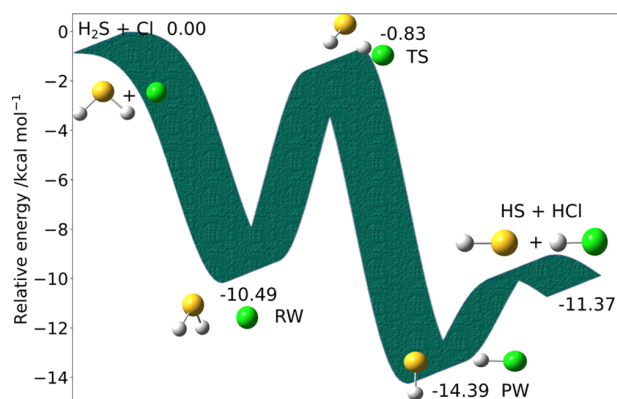
Figure 6. $H_2S + Cl$ reaction mechanism. Electronic energies are computed at the jChS level.
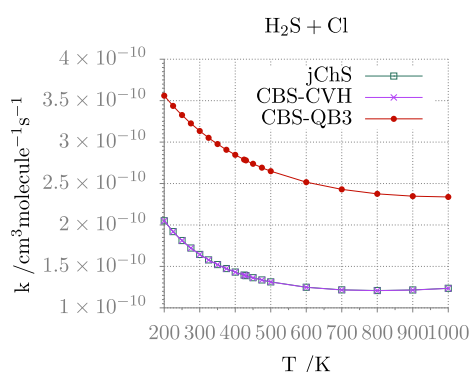
Figure 7. Temperature-dependence plots of the $H_2S + Cl$ reaction rate constants calculated at various levels of theory for a pressure of 1 atm.

(BHPERI), bipolar cycloaddition (CADBH), cycloreversion (CRBH), multiple proton exchange (PXBH), and different

(BHDIV) reactions. For each of these five classes of reactions, we selected no more than two representative cases. The structures of the seven selected transition states are shown in Figure 3, and the corresponding forward and reverse reaction barriers (BH14 set) are collected in Table 5.

The average and maximum errors are larger than those of the DBH24 set; however, a closer inspection of the results shows that, as already pointed out in ref 41, the role of full triple and quadruple excitations is non-negligible for CRBH reactions. This effect cannot be captured, of course, by the jChS model and leads to errors well above 1 kcal mol^{-1}. In all other cases, the errors are below the target of the jChS model chemistry. As a matter of fact, excluding the contribution of triple and quadruple excitations (last column in Table 5) reduces the MUE of jChS results to 0.24 kcal mol^{-1}. Furthermore, the error related to the difference between rev-DSD and reference geometries is lower than 0.3 kcal mol^{-1} even in the worst cases.

Zero Point Energy and Finite Temperature Contributions. Accurate determination of thermochemical and kinetic parameters by quantum chemical methods requires, in addition to electronic energies, also zero point and finite temperature contributions (FTCs), which are usually obtained within the RRHO approximation, possibly employing empirical scaling factors.[96] However, it is well-known that the scaling factors are intrinsically different for zero point energies (ZPEs) and vibrational frequencies, with the results for the latter quantities often being not sufficiently accurate.[97] One effective strategy devoid of any empirical parameter is offered by the generalized second-order vibrational perturbation theory in conjunction with a separate treatment of large-amplitude motions.[58,98] In fact, a resonance-free expression for ZPEs of energy minima and transition states,[99,100] an unsupervised smoothing procedure (HDCPT2) for fundamental frequencies,[101] and a fully automatic detection and treatment of torsional motions (hindered rotor, HR, approximation)[102] have been implemented in the Gaussian code[60] and

Table 11. Arrhenius–Kooij Parameters for the H_2S + Cl Reaction

	jChS	CBS-CVH	CBS-QB3
A (cm^3 molecule^{-1} s^{-1})	9.12×10^{-11}	9.11×10^{-11}	2.63×10^{-10}
n	7.65×10^{-2}	7.67×10^{-2}	1.60×10^{-1}
E (kcal mol^{-1})	-3.42×10^{-1}	-3.42×10^{-1}	-9.94×10^{-2}
rms	2.51×10^{-12}	2.51×10^{-12}	2.87×10^{-12}
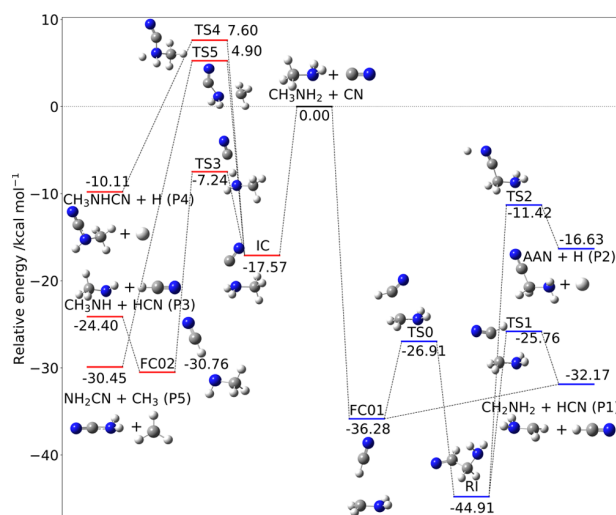
Figure 8. CH_3NH_2 + CN reaction mechanism. The pathway for the attack of CN to the N moiety of methylamine is marked in red and for the abstraction of H from the methyl group by CN in blue. Electronic energies are computed at the jChS level.

validated.[59] As a consequence, a fully black-box procedure is available for taking into account all of these contributions.

Next, the so-called simple perturbation theory (SPT)[103] can be applied for computing partition functions without the need for performing explicit (or stochastic) summations of individual energy levels. In fact, the SPT retains the formal expression of the harmonic partition function but employing the anharmonic ZPE and fundamental levels (Δ_i) issuing from HDCPT2 and HR computations.

$$Q_{vib} = \frac{\exp\left(-\frac{ZPE}{KT}\right)}{\prod_i \left[1 - \exp\left(-\frac{\Delta_i}{KT}\right)\right]} \tag{7}$$

This approximation provides results in remarkable agreement with accurate reference values and leads to analytical expressions for the different thermodynamic functions.[103]

On these grounds, we will now analyze the performances of the jChS model chemistry in dealing with these terms starting from a benchmark of the RRHO approximation with reference to accurate quantum chemical results and then proceeding to take into account anharmonic contributions. For illustration purposes, we will focus our attention on ZPEs and absolute entropies (S), which are especially sensitive to high and low frequencies, respectively.

To this end, a new database has been built (ThCS21), which contains accurate experimental values for the ZPEs and absolute entropies of 21 semirigid closed-shell molecules, whose estimated errors are below 0.1 kcal mol^{-1} and 0.05 cal (mol K)$^{-1}$, respectively. The results collected in Table 6 show that already at the harmonic level, the errors are well within the level of accuracy expected from the jChS model chemistry and

the anharmonic results can be confidently used in conjunction with the most sophisticated models (e.g., CBS-CVH). Actually, the harmonic frequencies obtained at this level do not require any empirical correction to compensate for method and/or basis set deficiency but only for genuine anharmonic effects, which, in turn, give significant contributions to ZPEs only for some XH bonds (X = C, N, O). As a consequence, an empirical correction of 0.12 kcal mol^{-1} for each bond of this kind provides results very close to the anharmonic counterparts (see results in parenthesis in the first column of Table 6).

Accurate entropy values are also available for the same set of molecules, and harmonic computations perform a remarkable job in reproducing the experimental values. However, entropy is exquisitely sensitive to low-frequency vibrations, so that a set of flexible molecules is collected in Table 7. It is apparent that the HRHO model (which does not add any computational burden with respect to the underlying RRHO model) performs a remarkable job for systems containing a single torsion. The situation is more involved for larger flexible systems due to the presence of several low-energy minima contributing to the overall thermodynamic functions. Although this aspect goes beyond the main topic of the present contribution, we point out that several strategies are being proposed, following systematic search,[107] stochastic,[108] and, more recently, machine learning[109] approaches. Other kinds of large-amplitude motions can be taken into account by means of one-dimensional variational or quasi-variational approaches[110] followed by SPT or direct count of energy levels.[98]

Another issue is represented by open-shell species, which are of paramount importance in both astrochemistry and atmospheric chemistry. In this case, experimental zero point energies are available only for diatomic species and accurate determinations are quite limited also for the other thermodynamic functions. The jChS results collected in Table 8 for a few representative systems (ThOS10 database) suggest that (in the absence of strong multireference effects) the expected accuracy is close to that reached for closed-shell systems.

Reaction Rates. In this section, we analyze the impact on reaction rates of the different ingredients discussed in the previous section, comparing the results issuing from different model chemistries including CBS-QB3, jChS, and CBS-CVH. Starting from simple elementary mechanisms, we proceed to more complex potential energy surfaces including several intermediates and transition states, possibly leading to different products.

The first test case is the high-pressure limit of the reaction $H^\bullet$ + CO, which has been recently investigated by Vichietti et al.[116] This reaction belongs to the HTBH38 set, whose jChS results have been discussed in the section devoted to energy barriers. For purposes of comparison, we have computed also the barriers at the CBS-CVH level on top of jChS geometries obtaining values (3.26 and 22.86 kcal mol^{-1}) for the forward and reverse barriers very close to the jChS counterparts at rev-DSD geometries (3.22 and 22.87 kcal mol^{-1}). Although the presence of a van der Waals prereactive complex has been
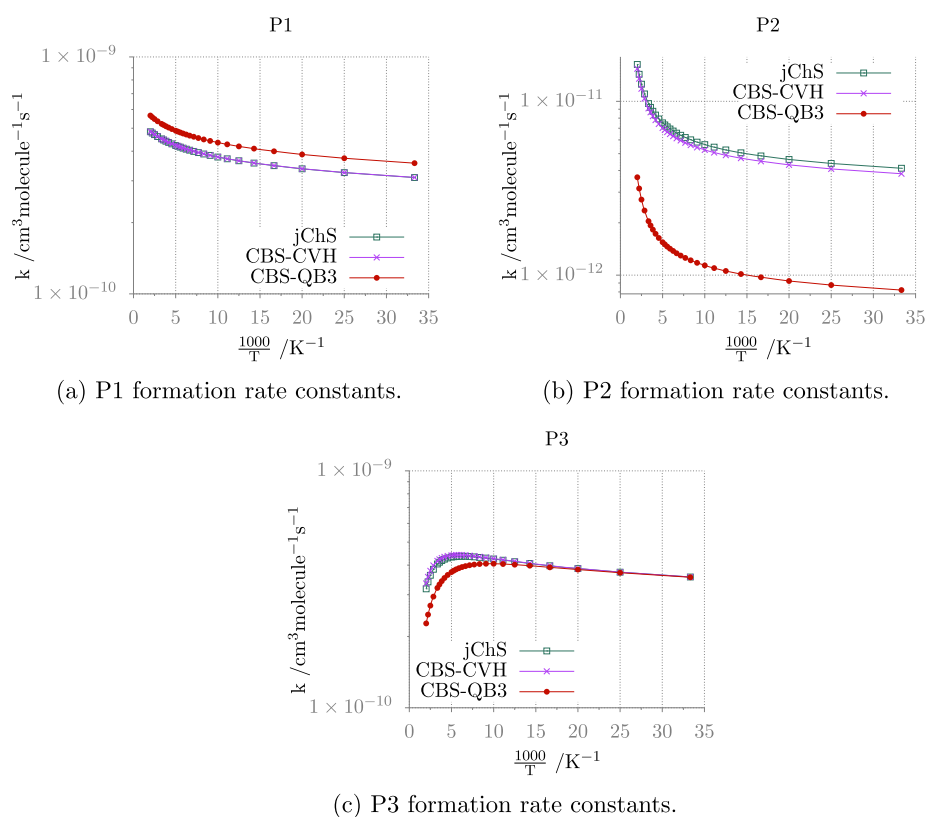
(a) P1 formation rate constants.

(b) P2 formation rate constants.

(c) P3 formation rate constants.

Figure 9. Temperature-dependence plots of the CH_3NH_2 + CN reaction rate constants calculated at various levels of theory for a pressure of 10^{-8} atm.

suggested, its stability (if any) is so small that its impact on the computed reaction rates is negligible.

The reaction rates computed in the 50−4000 K temperature interval are shown in Figure 4, and the parameters of the corresponding Arrhenius−Kooij fits obtained by different electronic structure methods are collected in Table 9. The non-Arrhenius behavior of the reaction is quite apparent, but the small errors of all of the fits show that the Arrhenius−Kooij model captures the essence of the deviation. Furthermore, the jChS results are close to the reference values of ref 116, whereas this is not the case for the largely employed CBS-QB3 approach at least at low temperatures.

We next consider the BHPERI1 and CRBH4 reactions discussed in the section on the energy barriers (see Figure 3a,c). The rates computed in the 300−1000 K temperature interval by different electronic structure methods are shown in Figure 5a,b, whereas the parameters of the corresponding Arrhenius−Kooij fits are collected in Table 10. Both reactions are characterized by quite high energy barriers, and their rates show a clear Arrhenius behavior. In these circumstances, the different electronic structure methods deliver comparable results over the whole temperature range.

The next example is the reactive potential energy surface for H_2S + Cl (see Figure 6), which involves a van der Waals prereactive complex (RW) followed by the transition state TS leading to a productlike van der Waals complex (PW) and then to the products, i.e., HS + HCl. Since this reaction has been recently investigated at the CBS-CVH level,[28] it represents a meaningful test for the jChS model chemistry. Once again, the largest deviation from the reference values for all of the

stationary points is lower than 0.3 kcal mol^{-1}, to be compared to errors larger than 1 kcal mol^{-1} especially for transition states at the CBS-QB3 level. Errors of this magnitude can lead to unreliable rate constants, especially for reactions like this where the dynamical bottleneck is located at the inner transition state, as already pointed out in ref 28.

The reaction rates issuing from jChS computations are compared in Figure 7 to the CBS-QB3 and CBS-CVH counterparts, whereas the parameters of the corresponding Arrhenius−Kooij fittings (see eq 6) are collected in Table 11. The root-mean-square deviations reported in Table 11 demonstrate that the data are indeed well fitted by the Arrhenius−Kooij expression with a negative activation energy (E) at 0 K. The results issuing from jChS and CBS-CVH computations are virtually indistinguishable, whereas significantly larger rates are obtained at the CBS-QB3 level.

The last example is the quite complex reactive potential energy surface ruling the addition of CN to CH_3NH_2 shown in Figure 8 together with the jChS energies of all of the stationary points. The experimental reaction rate at different temperatures[117] has been recently well reproduced employing CBS-CVH electronic structure computations within a master equation treatment similar to that employed in the present paper.[118] This system represents, therefore, a challenging test for the jChS model.

The attack of CN on the nitrogen side of methylamine proceeds via a potential well associated with a prereactive complex, NC⋯NH$_2$CH$_3$ (IC), which evolves in an inner (submerged) transition state (TS3) that, passing through an NCH⋯NHCH$_3$ intermediate (FC02), forms the HCN +

Table 12. Arrhenius−Kooij Parameters for the CH$_3$NH$_2$ + CN Reaction

	jChS			CBS-CVH			CBS-QB3		
	P1	P2	P3	P1	P2	P3	P1	P2	P3
A (cm^3 molecule^{-1} s^{-1})	4.51×10^{-10}	8.95×10^{-12}	4.28×10^{-10}	4.52×10^{-10}	8.38×10^{-12}	4.38×10^{-10}	5.22×10^{-10}	1.86×10^{-12}	3.54×10^{-10}
n	1.50×10^{-1}	8.70×10^{-1}	-2.43×10^{-1}	1.51×10^{-1}	8.84×10^{-1}	-2.09×10^{-1}	1.63×10^{-1}	9.78×10^{-1}	-4.45×10^{-1}
E (kcal mol^{-1})	2.10×10^{-3}	-8.31×10^{-2}	4.99×10^{-2}	2.02×10^{-3}	-8.48×10^{-2}	4.63×10^{-2}	5.28×10^{-4}	-9.72×10^{-2}	6.71×10^{-2}
rms	9.70×10^{-13}	4.54×10^{-13}	1.75×10^{-11}	9.40×10^{-13}	4.33×10^{-13}	1.69×10^{-11}	3.31×10^{-13}	1.11×10^{-13}	1.73×10^{-11}

NHCH$_3$ products (P3). Alternative channels, and in particular that leading to NH$_2$CN + CH$_3$, are ruled by nonsubmerged transition states and are, therefore, closed under the ISM conditions. The attack on the methyl side forms directly the FC01 complex, which, in turn, leads to HCN + NH$_2$CH$_2$ (P1) without any potential energy barrier. In this case, the alternative two-step mechanism (TS0-RI-TS2-P2) leading to aminoacetonitrile + H is open since it involves only submerged transition states, but it is less favorable.

Comparison with the CBS-CVH results of ref 118 shows MAE of 0.26 kcal mol^{-1} and a maximum deviation of −0.55 kcal mol^{-1} for the relative energies of all of the stationary points. The errors of the CBS-QB3[22] model are again larger than 1 kcal mol^{-1}, in agreement with the estimates of previous studies.[119] The reaction rates issuing from jChS computations are compared in Figure 9 with the CBS-QB3 and CBS-CVH counterparts, whereas the parameters of the corresponding Arrhenius−Kooij fits (see eq 6) are collected in Table 12. It is to be noted that pressure does not influence the reaction rate, as the reactants always proceed to form the products without experiencing significant collisional stabilization in the investigated pressure range (0.001−1 bar).

A curved Arrhenius plot is obtained when the activation energy depends on the temperature, and this behavior is captured by the Arrhenius−Kooij equation when this dependence is linear. The root-mean-square deviations reported in Table 12 demonstrate that the data for the different products are indeed well fitted by the Arrhenius−Kooij expression. Within this model, E represents the activation energy at 0 K, and the activation energy at a generic temperature T is given by $E + n\left(\frac{RT}{300}\right)$. In the present case, the activation energy is positive for P1 and P3 as a result of both the capture rate constant and the subsequent energy barriers for the unimolecular steps. The value is instead negative for P2, but in this case, the Arrhenius plot is essentially linear. The n parameter (the first derivative of the activation energy with respect to temperature) is positive for the P1 and P2 products, thus reflecting an increase of the activation energy with temperature, while the opposite behavior (n negative) is obtained for P3. Finally, the value of the pre-exponential factor A is typical for this kind of reaction and rules the branching ratio between the different channels.

■ CONCLUSIONS

Astrochemistry and atmospheric chemistry require accurate kinetic data for processes occurring at low to moderate temperatures and involving barrier heights spanning a large range of values. Furthermore, the chemical species involved in these processes can contain more than 10 non-hydrogen atoms and noncovalent interactions may often rule the entrance channels. We have, therefore, developed and validated a new general model rooted in the master equation formalism employing the ab initio transition-state theory for computing the reaction rates of elementary processes. To this end, we have slightly modified and validated the recently proposed jChS model chemistry with reference to very accurate energetic and kinetic data. The results obtained for a large panel of systems and reaction channels show an average error within 0.3 kcal mol^{-1} without the need for any empirical parameter, which allows the evaluation of accurate branching ratios and leads to errors within 20% for reaction rates.

The computational bottleneck of the proposed model chemistry is the CCSD(T)/jun-cc-pVTZ step and, in this connection, recent localized treatments of correlation (e.g., using local pair natural orbitals[120,121]) will be investigated to further increase the dimension of molecular systems amenable to accurate computations with reasonable computer requirements. Furthermore, the performances of the jChS model for other classes of reactions of particular interest for astrochemistry and/or atmospheric chemistry (e.g., those involving ozone and Criegee intermediate) need be investigated in deeper detail. In the same vein, further refinements and validations are surely needed for specific situations (e.g., non-negligible static correlation effects or intersystem crossing). However, even taking these caveats into account, we think that the results reported in the present paper pave the route for the accurate study of chemical processes under widely different temperature and pressure conditions.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c00406.

> Extended table of results for the HTBH38/08 data set, Cartesian coordinates of all of the stationary points optimized in this work, and MESS input files (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Vincenzo Barone − *SMART Laboratory, Scuola Normale Superiore di Pisa, 56125 Pisa, Italy;* orcid.org/0000-0001-6420-4107; Email: vincenzo.barone@sns.it

Authors

Jacopo Lupi − *SMART Laboratory, Scuola Normale Superiore di Pisa, 56125 Pisa, Italy;* orcid.org/0000-0001-6522-9947

Zoi Salta − *SMART Laboratory, Scuola Normale Superiore di Pisa, 56125 Pisa, Italy;* orcid.org/0000-0002-7826-0182

Nicola Tasinato − *SMART Laboratory, Scuola Normale Superiore di Pisa, 56125 Pisa, Italy;* orcid.org/0000-0003-1755-7238

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.1c00406

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Woon, D. E. The Astrochymist. http://www.astrochymist.org/ (accessed Dec 5, 2018).

(2) Herbst, E.; van Dishoeck, F. E. Complex Organic Interstellar Molecules. *Annu. Rev. Astron. Astrophys.* **2009**, *47*, 427−480.

(3) Snow, T. P.; McCall, B. J. Diffuse Atomic and Molecular Clouds. *Annu. Rev. Astron. Astrophys.* **2006**, *44*, 367−414.

(4) Yamamoto, S. *Introduction to Astrochemistry*; Springer, 2017.

(5) Dobrijevic, M.; Loison, J. C.; Hickson, K. M.; Gronoff, G. 1D-coupled photochemical model of neutrals, cations and anions in the atmosphere of Titan. *Icarus* **2016**, *268*, 313−339.

(6) Carbo, R.; Ginebreda, A. Interstellar chemistry. *J. Chem. Educ.* **1985**, *62*, 832−836.

(7) Atkinson, R.; Baulch, D. L.; Cox, R. A.; Crowley, J. N.; Hampson, R. F.; Hynes, R. G.; Jenkin, M. E.; Rossi, M. J.; Troe, J.; IUPAC Subcommittee. Evaluated kinetic and photochemical data for atmospheric chemistry: Volume II - gas phase reactions of organic species. *Atmos. Chem. Phys.* **2006**, *6*, 3625−4055.

(8) Kwok, E. S.; Atkinson, R. Estimation of hydroxyl radical reaction rate constants for gas-phase organic compounds using a structure-reactivity relationship: An update. *Atmos. Environ.* **1995**, *29*, 1685−1695.

(9) Vereecken, L.; Glowacki, D. R.; Pilling, M. J. Theoretical chemical kinetics in tropospheric chemistry: methodologies and applications. *Chem. Rev.* **2015**, *115*, 4063−4114.

(10) Karton, A. A computational chemist's guide to accurate thermochemistry for organic molecules. *WIREs Comput. Mol. Sci.* **2016**, *6*, 292−310.

(11) Karton, A.; Martin, J. M. L. Explicitly correlated Wn theory: W1-F12 and W2-F12. *J. Chem. Phys.* **2012**, *136*, No. 124114.

(12) Laane, J.; Dakkouri, M.; van der Veken, B.; Oberhammer, H. *Structures and Conformations of Non-Rigid Molecules*; Springer: Netherlands, 1993; Vol. *410*.

(13) Császár, A. G.; Allen, W. D.; Schaefer, H. F. In pursuit of the ab initio limit for conformational energy prototypes. *J. Chem. Phys.* **1998**, *108*, 9751−9764.

(14) Peterson, K. A.; Feller, D.; Dixon, D. A. Chemical accuracy in ab initio thermochemistry and spectroscopy: Current strategies and future challenges. *Theor. Chem. Acc.* **2012**, *131*, 1−50.

(15) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. HEAT: High accuracy extrapolated ab initio thermochemistry. *J. Chem. Phys.* **2004**, *121*, 11599−11613.

(16) Bomble, Y. J.; Vázquez, J.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J.; Stanton, J. F. High-accuracy extrapolated ab initio thermochemistry. II. Minor improvements to the protocol and a vital simplification. *J. Chem. Phys.* **2006**, *125*, No. 064108.

(17) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. High-accuracy extrapolated ab initio thermochemistry. III. Additional improvements and overview. *J. Chem. Phys.* **2008**, *128*, No. 114111.

(18) Zhang, J.; Valeev, E. F. Prediction of reaction barriers and thermochemical properties with explicitly correlated coupled-cluster methods: a basis set assesment. *J. Chem. Theory Comput.* **2012**, *8*, 3175−3186.

(19) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Gn theory. *WIREs Comput. Mol. Sci.* **2011**, *1*, 810−825.

(20) Curtiss, A. L.; Redfern, C. P.; Raghavachari, K. Gaussian-4 theory. *J. Chem. Phys.* **2007**, *126*, No. 084108.

(21) Montgomery, J. A.; Ochterski, J. W.; Petersson, G. CBS. *J. Chem. Phys.* **1994**, *101*, 5900−5909.

(22) Montgomery, J. A.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. A complete basis set model chemistry. VII. Use of the minimum population localization method. *J. Chem. Phys.* **2000**, *112*, 6532−6542.

(23) Curtiss, L. A.; Redfern, P. C.; Raghavachari, K. Assessment of Gaussian-4 theory for energy barriers. *Chem. Phys. Lett.* **2010**, *499*, 168−172.

(24) Puzzarini, C.; Barone, V. Extending the molecular size in accurate quantum-chemical calculations: the equilibrium structure and spectroscopic properties of uracil. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7189−7197.

(25) Puzzarini, C.; Biczysko, M.; Barone, V.; Peña, I.; Cabezas, C.; Alonso, J. L. Accurate molecular structure and spectroscopic

62

properties of nucleobases: a combined computational-microwave investigation of 2-thiouracil as a case study. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16965−16975.

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

(27) Salta, Z.; Tasinato, N.; Lupi, J.; Boussessi, R.; Balbi, A.; Puzzarini, C.; Barone, V. Exploring the Maze of $C_2N_2H_5$ Radicals and Their Fragments in the Interstellar Medium with the Help of Quantum-Chemical Computations. *ACS Earth Space Chem.* **2020**, *4*, 774−782.

(28) Lupi, J.; Puzzarini, C.; Cavallotti, C.; Barone, V. State-of-the-Art Quantum Chemistry Meets Variable Reaction Coordinate Transition State Theory to Solve the Puzzling Case of the H_2S + Cl System. *J. Chem. Theory Comput.* **2020**, *16*, 5090−5104.

(29) Lupi, J.; Puzzarini, C.; Barone, V. Methanimine as a Key Precursor of Imines in the Interstellar Medium: The Case of Propargylimine. *Astrophys. J. Lett.* **2020**, *903*, No. L35.

(30) Baiano, C.; Lupi, J.; Tasinato, N.; Puzzarini, C.; Barone, V. The Role of State-of-the-Art Quantum-Chemical Calculations in Astrochemistry: Formation Route and Spectroscopy of Ethanimine as a Paradigmatic Case. *Molecules* **2020**, *25*, No. 2873.

(31) Ballotta, B.; Nandi, S.; Barone, V.; Rampino, S. Gas-Phase Formation and Isomerization Reactions of Cyanoacetaldehyde, a Prebiotic Molecule of Astrochemical Interest. *ACS Earth Space Chem.* **2021**, *5*, 1071−1082.

(32) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, R. H.; Truhlar, G. D. Perspectives on Basis Sets Beautiful: Seasonal Plantings of Diffuse Basis Functions. *J. Chem. Theory Comput.* **2011**, *7*, 3027−3034.

(33) Alessandrini, S.; Barone, V.; Puzzarini, C. Extension of the "Cheap" Composite Approach to Noncovalent Interactions: The jun-ChS Scheme. *J. Chem. Theory Comput.* **2020**, *16*, 988−1006.

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

(35) Bao, J. L.; Truhlar, D. G. Variational transition state theory: theoretical framework and recent developments. *Chem. Soc. Rev.* **2017**, *46*, 7548−7596.

(36) Miller, A. J.; Klippenstein, J. S. Master equation methods in gas phase chemical kinetics. *J. Phys. Chem. A* **2006**, *110*, 10528−10544.

(37) Klippenstein, S. J. From Theoretical Reaction Dynamics to Chemical Modeling of Combustion. *Proc. Combust. Inst.* **2017**, *36*, 77−111.

(38) Zheng, J.; Zhao, Y.; Truhlar, D. G. The DBH24/08 Database and Its Use to Assess Electronic Structure Model Chemistries for Chemical Reaction Barrier Heights. *J. Chem. Theory Comput.* **2009**, *5*, 808−821.

(39) Zhao, Y.; Lynch, B. J.; Truhlar, D. G. Multi-coefficient extrapolated density functional theory for thermochemistry and thermochemical kinetics. *Phys. Chem. Chem. Phys.* **2005**, *7*, 43−52.

(40) Zhao, Y.; González-García, N.; Truhlar, D. G. Benchmark Database of Barrier Heights for Heavy Atom Transfer, Nucleophilic Substitution, Association, and Unimolecular Reactions and Its Use to Test Theoretical Methods. *J. Phys. Chem. A* **2005**, *109*, 2012−2018.

(41) Karton, A. Highly Accurate CCSDT(Q)/CBS Reaction Barrier Heights for a Diverse Set of Transition Structures: Basis Set Convergence and Cost-Effective Approaches for Estimating Post-CCSD(T) Contributions. *J. Phys. Chem. A* **2019**, *123*, 6720−6732.

(42) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. W4 theory for computational thermochemistry: In pursuit of confident sub-kJ/mol predictions. *J. Chem. Phys.* **2006**, *125*, No. 144108.

(43) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

(44) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479−483.

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

(46) Sellers, H.; Pulay, P. The adiabatic correction to molecular potential surfaces in the SCF approximation. *Chem. Phys. Lett.* **1984**, *103*, 463−465.

(47) Handy, N. C.; Yamaguchi, Y.; Schaefer, H. F. The diagonal correction to the Born-Oppenheimer approximation: Its effect on the singlet-triplet splitting of CH_2 and other molecular effects. *J. Chem. Phys.* **1986**, *84*, 4481−4484.

(48) Handy, N. C.; Lee, A. M. The adiabatic approximation. *Chem. Phys. Lett.* **1996**, *252*, 425−430.

(49) Kutzelnigg, W. The adiabatic approximation I. The physical background of the Born-Handy ansatz. *Mol. Phys.* **1997**, *90*, 909−916.

(50) Cowan, R. D.; Griffin, M. Approximate relativistic corrections to atomic radial wave functions. *J. Opt. Soc. Am.* **1976**, *66*, 1010−1014.

(51) Martin, R. L. All-electron relativistic calculations on silver hydride. An investigation of the Cowan-Griffin operator in a molecular species. *J. Phys. Chem. A* **1983**, *87*, 750−754.

(52) Bomble, Y. J.; Stanton, J. F.; Kállay, M.; Gauss, J. Coupled-cluster methods including noniterative corrections for quadruple excitations. *J. Chem. Phys.* **2005**, *123*, No. 054101.

(53) Kállay, M.; Gauss, J. Approximate treatment of higher excitations in coupled-cluster theory. *J. Chem. Phys.* **2005**, *123*, No. 214105.

(54) Kállay, M.; Gauss, J. Approximate treatment of higher excitations in coupled-cluster theory. II. Extension to general single-determinant reference functions and improved approaches for the canonical Hartree−Fock case. *J. Chem. Phys.* **2008**, *129*, No. 144101.

(55) Moore, C. E. Atomic Energy Levels; Natl. Bur. Stand. (US) Circ., 1949.

(56) Santra, G.; Sylvetsky, N.; Martin, J. M. Minimally empirical double-hybrid functionals trained against the GMTKN55 database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123*, 5129−5143.

(57) Biczysko, M.; Panek, P.; Scalmani, G.; Bloino, J.; Barone, V. Harmonic and Anharmonic Vibrational Frequency Calculations with the Double-Hybrid B2PLYP Method: Analytic Second Derivatives and Benchmark Studies. *J. Chem. Theory Comput.* **2010**, *6*, 2115−2125.

(58) Barone, V. Anharmonic vibrational properties by a fully automated second order perturbative approach. *J. Chem. Phys.* **2005**, *122*, No. 014108.

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

(60) Frisch, M. J. et al.*Gaussian 16*, revision C.01; Gaussian Inc.: Wallingford CT, 2016.

(61) Werner, H.-J.; et al. The Molpro quantum chemistry package. *J. Chem. Phys.* **2020**, *152*, No. 144107.

(62) Kállay, M. et al. MRCC, a quantum chemical program suite, 2018.

(63) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G. A quantum chemical program package. 2016; with contributions from Auer, A. A.; Bartlett, R. J.; Benedikt, U.; Berger, C.; Bernholdt, D. E.; Bomble, Y. J.; Christiansen, O.; Engel, F.; Heckert, M.; Heun, O.; Huber, C.; Jagau, T.-C.; Jonsson, D.; Jusèlius, J.; Klein, K.; Lauderdale, W. J.; Lipparini, F.; Matthews, D.; Metzroth, T.; Mück, L. A.; O'Neill, D. P.; Price, D. R.; Prochnow, E.; Puzzarini, C.; Ruud, K.; Schiffmann, F.; Schwalbach, W.; Stopkowicz, S.; Tajti, A.; Vázquez, J.; Wang, F.; Watts, J. D., and the integral packages MOLECULE (J. Almlöf and P. R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen. For the current version, see http://www.cfour.de.

(64) Woon, D. E.; Dunning, T. H., Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. V. Core-Valence Basis Sets for Boron through Neon. *J. Chem. Phys.* **1995**, *103*, 4572−4585.

(65) Feller, D. Application of systematic sequences of wave functions to the water dimer. *J. Chem. Phys.* **1992**, *96*, 6104−6114.

(66) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-set convergence of correlated calculations on water. *J. Chem. Phys.* **1997**, *106*, 9639−9646.

(67) Georgievskii, Y.; Miller, A. J.; Burke, P. M.; Klippenstein, J. S. Reformulation and solution of the master equation for multiple-well chemical reactions. *J. Phys. Chem. A* **2013**, *117*, 12146−12154.

(68) Tardy, D. C.; Rabinovitch, B. Collisional Energy Transfer. Thermal Unimolecular Systems in the Low-Pressure Region. *J. Chem. Phys.* **1966**, *45*, 3720−3730.

(69) Fernández-Ramos, A.; Miller, J. A.; Klippenstein, S. J.; Truhlar, D. G. Modeling the Kinetics of Bimolecular Reactions. *Chem. Rev.* **2006**, *106*, 4518−4584.

(70) Eckart, C. The penetration of a potential barrier by electrons. *Phys. Rev.* **1930**, *35*, 1303−1309.

(71) Hunter, M.; Reid, S. A.; Robie, D. C.; Reisler, H. The Monoenergetic Unimolecular Reaction of Expansion-Cooled NO$_2$: NO Product State Distributions at Excess Energies 0−3000 cm^{-1}. *J. Chem. Phys.* **1993**, *99*, 1093−1108.

(72) Skouteris, D.; Balucani, N.; Ceccarelli, C.; Vazart, F.; Puzzarini, C.; Barone, V.; Codella, C.; Lefloch, B. The Genealogical Tree of Ethanol: Gas-phase Formation of Glycolaldehyde, Acetic Acid, and Formic Acid. *Astrophys. J.* **2018**, *854*, No. 135.

(73) Kooij, D. M. Z. *Phys. Chem.* **1893**, *12*, 155−161.

(74) Laidler, K. J. A glossary of terms used in chemical kinetics, including reaction dynamics (IUPAC recommendations 1996). *Pure Appl. Chem.* **1996**, *68*, 149−192.

(75) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, No. 034108.

(76) Grimme, S. Density functional theory with London dispersion corrections. *WIREs Comput. Mol. Sci.* **2011**, *1*, 211−228.

(77) Caldeweyher, E.; Bannwarth, C.; Grimme, S. Extension of the D3 dispersion coefficient model. *J. Chem. Phys.* **2017**, *147*, No. 034112.

(78) Xie, F.; Fusè, M.; Hazrah, A. S.; Jaeger, W.; Barone, V.; Xu, Y. Discovering the elusive global minimum in a ternary chiral cluster: rotational spectra of propylene oxide trimer. *Angew. Chem., Int. Ed.* **2020**, *59*, 22427−22430.

(79) Alonso, E. R.; Fusè, M.; León, I.; Puzzarini, C.; Alonso, J. L.; Barone, V. Exploring the maze of cycloserine conformers in the gas phase guided by microwave spectroscopy and quantum chemistry. *J. Phys. Chem. A* **2021**, *125*, 2121−2129.

(80) Barone, V.; Ceselin, G.; Fusè, M.; Tasinato, N. Accuracy Meets Interpretability for Computational Spectroscopy by Means of Hybrid and Double-Hybrid Functionals. *Front. Chem.* **2020**, *8*, 584203-1−584203-14.

(81) Sosa, C.; Schlegel, H. B. Ab initio calculations on the barrier height for the hydrogen addition to ethylene and formaldehyde. The importance of spin projection. *Int. J. Quantum Chem.* **1986**, *29*, 1001−1015.

(82) Schlegel, H. B. Møller-Plesset Perturbation Theory with Spin Projection. *J. Phys. Chem. A* **1988**, *92*, 3075−3378.

(83) Yuan, H.; Cremer, D. The expectation value of the spin operator S^2 as a diagnostic tool in coupled cluster theory. The advantages of using UHF-CCSD theory for the description of homolytic dissociation. *Chem. Phys. Lett.* **2000**, *324*, 389−402.

(84) Stanton, J. F. On the extent of spin contamination in open-shell coupled-cluster wave functions. *J. Chem. Phys.* **1994**, *101*, 371−374.

(85) Cohen, A. J.; Tozer, D. J.; Handy, N. C. Evaluation of S^2 in density functional theory. *J. Chem. Phys.* **2007**, *126*, No. 214104.

(86) Menon, A. S.; Radom, L. Consequences of Spin Contamination in Unrestricted Calculations on Open-Shell Species: Effect of Hartree Fock and Møller Plesset Contributions in Hybrid and Double-Hybrid

Density Functional Theory Approaches. *J. Phys. Chem. A* **2008**, *112*, 13225−13230.

(87) Karton, A.; Tarnopolsky, A.; Lamère, J.; Schatz, G. C.; Martin, J. M. L. Highly Accurate First-Principles Benchmark Data Sets for the Parametrization and Validation of Density Functional and Other Approximate Methods. Derivation of a Robust, Generally Applicable, Double-Hybrid Functional for Thermochemistry and Thermochemical Kinetics. *J. Phys. Chem. A* **2008**, *112*, 12868−12886.

(88) Karton, A.; Rabinovich, E.; Martin, J. M. L.; Ruscic, B. W4 Theory for computational thermochemistry: in pursuit of confident sub-kJ/mol predictions. *J. Chem. Phys.* **2011**, *125*, No. 144108.

(89) Kong, L.; Bischoff, F. A.; Valeev, E. F. Explicitly Correlated R12/F12 Methods for Electronic Structure. *Chem. Rev.* **2012**, *112*, 75−107.

(90) Guner, V.; Khuong, K. S.; Leach, A. G.; Lee, P. S.; Bartberger, M. D.; Houk, K. N. A Standard Set of Pericyclic Reactions of Hydrocarbons for the Benchmarking of Computational Methods: The Performance of ab Initio, Density Functional, CASSCF, CASPT2, and CBS-QB3 Methods for the Prediction of Activation Barriers, Reaction Energetics, and Transition State Geometries. *J. Phys. Chem. A* **2003**, *107*, 11445−11459.

(91) Karton, A.; Goerigk, L. Accurate reaction barrier heights of pericyclic reactions: Surprisingly large deviations for the CBS-QB3 composite method and their consequences in DFT benchmark studies. *J. Comput. Chem* **2015**, *36*, 622−632.

(92) Yu, L.-J.; Sarrami, F.; O'Reilly, R. J.; Karton, A. Reaction barrier heights for cycloreversion of heterocyclic rings: An Achilles' heel for DFT and standard ab initio procedures. *Chem. Phys.* **2015**, *458*, 1−8.

(93) Ess, D. H.; Houk, K. N. Activation Energies of Pericyclic Reactions: Performance of DFT, MP2, and CBS-QB3 Methods for the Prediction of Activation Barriers and Reaction Energetics of 1,3-Dipolar Cycloadditions, and Revised Activation Enthalpies for a Standard Set of Hydrocarbon Pericyclic Reactions. *J. Phys. Chem. A* **2005**, *109*, 9542−9553.

(94) Karton, A.; O'Reilly, R. J.; Chan, B.; Radom, L. Determination of Barrier Heights for Proton Exchange in Small Water, Ammonia, and Hydrogen Fluoride Clusters with G4(MP2)-Type, MPn, and SCS-MPn Procedures−A Caveat. *J. Chem. Theory Comput.* **2012**, *8*, 3128−3136.

(95) Goerigk, L.; Hansen, A.; Bauer, C.; Ehrlich, S.; Najibi, A.; Grimme, S. A look at the density functional theory zoo with the advanced GMTKN55 database for general main group thermochemistry, kinetics and noncovalent interactions. *Phys. Chem. Chem. Phys.* **2017**, *19*, 32184−32215.

(96) Irikura, K. K.; Johnson, R.; Kacker, R. N.; Kessel, R. Uncertainties in scaling factors for ab initio vibrational zero-point energies. *J. Chem. Phys.* **2009**, *130*, No. 114102.

(97) Kesharwani, M. K.; Brauer, B.; Martin, J. M. L. Frequency and Zero-Point Vibrational Energy Scale Factors for Double-Hybrid Density Functionals (and Other Selected Methods): Can Anharmonic Force Fields Be Avoided? *J. Phys. Chem. A* **2015**, *119*, 1701−1714.

(98) Skouteris, D.; Calderini, D.; Barone, V. Methods for calculating partition functions of molecules involving large amplitude and/or anharmonic motions. *J. Chem. Theory Comput.* **2016**, *12*, 1011−1018.

(99) Schuurman, M. S.; Allen, W. D.; von Raguè Schleyer, P.; Schaefer, H. F. I. The highly anharmonic BH5 potential energy surface characterized in the ab initio limit. *J. Chem. Phys.* **2005**, *122*, No. 104302.

(100) Piccardo, M.; Bloino, J.; Barone, V. Generalized vibrational perturbation theory for rotovibrational energies of linear, symmetric and asymmetric tops: Theory, approximations, and automated approaches to deal with medium-to-large molecular systems. *Int. J. Quantum Chem.* **2015**, *115*, 948−982.

(101) Bloino, J.; Biczysko, M.; Barone, V. General perturbative approach for spectroscopy, thermodynamics and kinetics: methodological background and benchmark studies. *J. Chem. Theory Comput.* **2012**, *8*, 1015−1036.

(102) Ayala, P. Y.; Schlegel, H. B. Identification and treatment of internal rotation in normal mode vibrational analysis. *J. Phys. Chem. A* **1998**, *108*, 2314−2325.

(103) Truhlar, D. G.; Isaacson, A. D. Simple perturbation theory estimates of equilibrium constants from force fields. *J. Chem. Phys.* **1991**, *94*, 357−359.

(104) Afeefy, H. Y.; Liebman, J. F.; Stein, S. E. In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J.; Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD, 2021; Chapter Neutral Thermochemical Data.

(105) Harding, L. B.; Georgievskii, Y.; Klippenstein, S. J. Accurate anharmonic zero-point energies for some combustion-related species from diffusion Monte Carlo. *J. Phys. Chem. A* **2017**, *121*, 4334−4340.

(106) Avila, G.; Carrington, T. J. Using a pruned basis, a non-product quadrature grid, and the exact Watson normal-coordinate kinetic energy operator to solve the vibrational Schrödinger equation for C2H4. *J. Chem. Phys.* **2011**, *135*, No. 064101.

(107) Li, Y.-P.; Bell, A. T.; Head-Gordon, M. Thermodynamics of anharmonic systems: uncoupled mode approximations for molecules. *J. Chem. Theory Comput.* **2016**, *12*, 2861−2870.

(108) Chandramouli, B.; Del Galdo, S.; Fuse, M.; Barone, V.; Mancini, G. Two-level strochastic search of low-energy conformers for molecular spectroscopy: implementation and validation of MM and QM models. *Phys. Chem. Chem. Phys.* **2019**, *21*, 19921−19934.

(109) Mancini, G.; Fuse, M.; Lazzari, F.; Chandramouli, B.; Barone, V. Unsupervised search of low-lying conformers with spectroscopic accuracy: a two-step algorithm rooted into the island model evolutionary algorithm. *J. Chem. Phys.* **2020**, *153*, No. 124110.

(110) Puzzarini, C.; Bloino, J.; Tasinato, N.; Barone, V. Accuracy and interpretability: The devil and the holy grail. New routes across old boundaries in computational spectroscopy. *Chem. Rev.* **2019**, *119*, 8131−8191.

(111) Chao, J.; Hall, K. R.; Marsh, K. N.; Wilhoit, R. C. Thermodynamic Properties of Key Organic Oxygen Compounds in the Carbon Range C1 to C4. Part 2. Ideal Gas Properties. *J. Phys. Chem. Ref. Data* **1986**, *15*, 1369−1436.

(112) Dorofeeva, O.; Novikov, V. P.; Neumann, D. B. NIST-JANAF Thermochemical Tables. I. Ten Organic Molecules Related to Atmospheric Chemistry. *J. Phys. Chem. Ref. Data* **2001**, *30*, 475−513.

(113) Ruscic, B.; Boggs, J. E.; Burcart, A.; Császár, A.; Demaison, J.; Janoshek, R.; Martin, J. M. l.; Morton, M. L.; Rossi, M. J.; Stanton, J. F.; Szalas, P. G.; Westmoreland, P. R.; Zabel, F.; Berces, T. IUPAC critical evaluation of thermodynamical properties of selected radicals. Part I. *J. Phys. Chem. Ref. Data* **2005**, *34*, 573−656.

(114) Demaison, J.; Margulès, L.; Boggs, J. E. Equilibrium structure and force field of NH_2. *Phys. Chem. Chem. Phys.* **2003**, *5*, 3359−3363.

(115) Marenich, A. V.; Boggs, J. E. Coupled cluster CCSD(T) calculations of equilibrium geometries, anharmonic force fields, and thermodynamic properties of formyl (HCO) and isoformyl (HOC) radical species. *J. Phys. Chem. A* **2003**, *107*, 2343−2350.

(116) Vichietti, R. M.; Machado, F. B.; Haiduke, L. A. Accurate rate constants for the forward and reverso H• + CO → HCO• reactions at the high-pressure limit. *ACS Omega* **2020**, *5*, 23975−23982.

(117) Sleiman, C.; El Dib, G.; Rosi, M.; Skouteris, D.; Balucani, N.; Canosa, A. Low temperature kinetics and theoretical studies of the reaction CN + CH3NH2: a potential source of cyanamide and methyl cyanamide in the interstellar medium. *Phys. Chem. Chem. Phys.* **2018**, *20*, 5478−5489.

(118) Puzzarini, C.; Salta, Z.; Tasinato, N.; Lupi, J.; Cavallotti, C.; Barone, V. A twist on the reaction of the CN radical with methylamine in the interstellar medium: new hints from a state-of-the-art quantum-chemical study. *Mon. Not. R. Astron. Soc.* **2020**, *496*, 4298−4310.

(119) Simmie, J. M.; Somers, K. P. Benchmarking Compound Methods (CBS-QB3, CBS-APNO, G3, G4, W1BD) against the Active Thermochemical Tables: A Litmus Test for Cost-Effective Molecular Formation Enthalpies. *J. Phys. Chem. A* **2015**, *119*, 7235−7246.

(120) Nagy, P. R.; Kallay, M. Approaching the basis set limit of CCSD(T) energies for large molecules with local natural orbital coupled-cluster methods. *J. Chem. Theory Comput.* **2019**, *15*, 5275−5298.

(121) Liakos, D. G.; Guo, Y.; Neese, F. Comprehensive benchmark results for the domain based local pair natural orbital coupled cluster method (DLPNO-CCSD(T)) for closed- and open-shell systems. *J. Phys. Chem. A* **2020**, *124*, 90−100.

THIS PAGE INTENTIONALLY LEFT BLANK

Chapter 4

A challenging case study: hydrogen sulfide reaction with chlorine radical

In this Chapter the gas-phase reaction of H_2S with Cl has been re-investigated both for its intrinsic interest in atmospheric chemistry and astrochemistry and in order to check if, as previously suggested, only explicit dynamical computations can lead to an accurate evaluation of the reaction rate because of strong recrossing effects and the breakdown of the variational extension of transition state theory. State-of-the-art electronic computations including up to quadruple excitations together with CBS extrapolation, spin-orbit contributions and diagonal Born-Oppenheimer corrections provide energetic quantities with a sub-$kJ\,mol^{-1}$ accuracy. Next, VTST for tight transition state steps and its variable coordinate extension for barrierless steps show that, contrary to the conclusions of previous less refined computations, both steps play a role in determining the overall reaction rate. Feeding of these data in a one dimensional master equation model leads to results in remarkable agreement with experiment in a wide range of temperatures thus unambiguously showing that state-of-the-art transition state theory based on accurate electronic structure computations is perfectly adequate to study this system. At the same time, various versions of the ChS models show remarkably accurate energetic quantities. Moreover, revDSD functional capabilities to accurately describe long-range interactions are shown, promoting it as a strong candidate for the density functional treatment of barrierless channels. In this Chapter, objectives O3 and O4 are pursued.

67

Article

State-of-the-Art Quantum Chemistry Meets Variable Reaction Coordinate Transition State Theory to Solve the Puzzling Case of the H₂S + Cl System

Jacopo Lupi, Cristina Puzzarini, Carlo Cavallotti, and Vincenzo Barone*

Cite This: *J. Chem. Theory Comput.* 2020, 16, 5090−5104

Read Online

ACCESS | 📊 Metrics & More | 📰 Article Recommendations | 🆂🆠 Supporting Information

ABSTRACT: The atmospheric reaction of H₂S with Cl has been reinvestigated to check if, as previously suggested, only explicit dynamical computations can lead to an accurate evaluation of the reaction rate because of strong recrossing effects and the breakdown of the variational extension of transition state theory. For this reason, the corresponding potential energy surface has been thoroughly investigated, thus leading to an accurate characterization of all stationary points, whose energetics has been computed at the state of the art. To this end, coupled-cluster theory including up to quadruple excitations has been employed, together with the extrapolation to the complete basis set limit and also incorporating core−valence correlation, spin−orbit, and scalar relativistic effects as well as diagonal Born−Oppenheimer corrections. This highly accurate composite scheme has also been paralleled by less expensive yet promising computational approaches. Moving to kinetics, variational transition state theory and its variable reaction coordinate extension for barrierless steps have been exploited, thus obtaining a reaction rate constant (8.16×10^{-11} cm³ molecule⁻¹ s⁻¹ at 300 K and 1 atm) in remarkable agreement with the experimental counterpart. Therefore, contrary to previous claims, there is no need to invoke any failure of the transition state theory, provided that sufficiently accurate quantum-chemical computations are performed. The investigation of the puzzling case of the H₂S + Cl system allowed us to present a robust approach for disclosing the thermochemistry and kinetics of reactions of atmospheric and astrophysical interest.
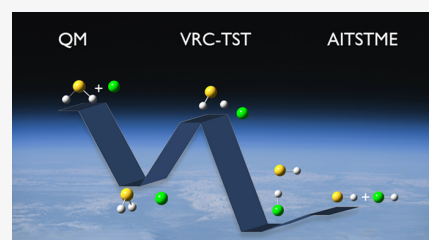
1. INTRODUCTION

Transition state theory (TST) and its variational extension are accepted to be the "workhorses" in computational kinetics for several fields ranging from combustion to atmospheric chemistry and astrochemistry (see, e.g., refs 1−3). It is thus fundamental to understand the features of a reactive potential energy surface (PES) that can introduce a breakdown of such theory. In other words, it is important to rationalize its possible limitations. In this respect, some cases of breakdown of TST have been reported in the literature. For instance, Hase and co-workers pointed out that many recrossings can take place in the Cl⁻ + CH₃Cl gas-phase reaction,[4] and the same behavior was observed for the unimolecular isomerizations of NCCN and CH₃CN.[5] Another interesting example is offered by the roaming mechanism in H₂CO photodissociation.[6,7] The unsatisfactory results obtained for the H₂S + Cl reaction by applying canonical variational transition state theory (CVTST) have been interpreted as another failure of TST.[8] However, the quantum-chemical approach employed there calls for a deeper reinvestigation of the reaction.

To provide a definitive elucidation of the mechanism of the H₂S + Cl reaction, in the present work, we have investigated its reactive PES by means of state-of-the-art quantum-chemical computations and coupled them with sophisticated kinetic models still rooted in the TST. In the framework of an ab initio-transition-state-theory-based master equation (AITSTME) treatment, different approaches have been employed to deal with barrierless reactions. The investigation of this specific reaction will lead us to the definition of an accurate protocol for the investigation of the thermochemistry and kinetics of challenging reactions.

In addition to offer a puzzling case study, the H₂S + Cl reaction plays an important role in atmospheric chemistry and might be of relevance also for the investigation of other planetary atmospheres. Indeed, the atmospheric sulfur cycle has been the subject of intensive investigations for a long time, mostly because of the need of continuously assessing and monitoring the contribution of anthropogenic sulfur compounds to problems such as acid rain, visibility reduction, and climate modification.[9] In particular, reduced sulfur-containing species are important in the chemistry of the atmosphere; among them, hydrogen sulfide (H₂S) is one of the simplest, but yet it plays an important role in earth's and planetary
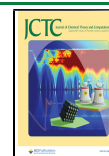
atmospheres. Concerning the terrestrial environment, its concentration in the atmosphere is significantly due to the decomposition of organic matter and volcanic eruptions, which can inject H_2S directly into the stratosphere.[10,11] For example, measurements of H_2S concentration by UV spectroscopy at volcanic sites in Italy have shown that this quantity can be on the order of 100 ppm (i.e., much larger than its average atmospheric concentration), H_2S thus being 2−3 times more abundant than SO_2.[12,13] However, the anthropogenic emission of H_2S should not be overlooked.[14,15]

In earth's atmosphere, H_2S is mainly removed by the hydroxyl radical (OH) by means of the gas-phase reaction[16]

$$H_2S + OH \rightarrow HS + H_2O \tag{1}$$

However, in some marine remote boundary layers and coastal urban areas, the concentration of the chlorine radical (Cl) is larger than that of OH.[17] Therefore, the reaction of H_2S with Cl, namely

$$H_2S + Cl \rightarrow HS + HCl \tag{2}$$

is also important. Concerning planetary systems, reaction 2 can play a role in Venus' atmosphere, the latter being rich in H_2S.[18−20] In addition to its importance in atmospheric processes, reaction 2 is of great interest as a prototype for heavy−light−heavy atom reactive systems,[21] and, furthermore, it leads to the production of vibrationally excited HCl molecules, which can be used in infrared chemiluminescence and laser-induced fluorescence studies.[22]

Reaction 2 has been studied since 1980, both experimentally and theoretically. Nevertheless, there is still some uncertainty concerning its detailed mechanism. At room temperature, the experimental rate constant spans from 3.7×10^{-11} to 10.5×10^{-11} cm^3 $molecule^{-1}$ s^{-1}.[23−32] According to the review study by Atkinson et al.,[33] the most reliable results are those by Nicovich et al.,[30] who reported an extensive investigation over a wide range of experimental conditions. In that work, the value of the rate constant at room temperature was found to be pressure-independent over a wide range, the latter being 33−800 mbar (i.e., 25−600 Torr).

One proposed mechanism is the direct hydrogen abstraction, the reaction thus proceeding through a transition state. Another possibility is offered by an addition/elimination mechanism, which has been discussed by various groups and nowadays seems to be widely accepted. Despite this, to the best of our knowledge, no theoretical work was able to correctly reproduce and interpret the experimental data. Indeed, the computational work by Wilson and Hirst pointed out the existence of a $H_2S \cdots Cl$ adduct, but without being able to connect it with the products, i.e., HS and HCl.[34] In ref 34, the rate constant for the direct abstraction was computed using conventional transition state theory (cTST), which led to a value, at room temperature, of 2.8×10^{-12} cm^3 $molecule^{-1}$ s^{-1}. However, such rate constant is 1 order of magnitude smaller than the experimental datum. Resende et al. instead investigated the addition/elimination mechanism.[8] They obtained a rate constant of 1.2×10^{-9} cm^3 $molecule^{-1}$ s^{-1}, 1 order of magnitude larger than the experimental value, and, as already mentioned above, they ascribed this discrepancy to a "breakdown of transition state theory".

To solve this puzzle, we have undertaken a comprehensive analysis of the whole reaction mechanism using state-of-the-art electronic structure and kinetic models. The paper is organized as follows. In the next section, the computational methodology

is described in some detail, thus introducing the different approaches employed for the electronic structure and kinetic calculations. Then, the results will be reported and discussed: first, the characterization of the reactive PES will be provided, followed by the accurate evaluation of its thermochemistry; then, the reaction rate constants will be addressed. Finally, the major outcomes of this work will be summarized in the concluding remarks.

2. COMPUTATIONAL METHODOLOGY

In this section, the methodology employed for the characterization of the $H_2S + Cl$ PES and its energetics will be first of all introduced. Then, we will move to the definition of the models used for the accurate interpretation of the kinetic aspects of the title reaction.

2.1. Electronic Structure Calculations. Several works have shown that double-hybrid functionals in conjunction with basis sets of at least triple-ζ quality represent a remarkable compromise between accuracy and computational cost.[35−38] For the calculation of equilibrium geometries and vibrational frequencies, the B2PLYP functional often approaches, and in some cases even overcomes, the accuracy of the much more computationally expensive CCSD(T) method, when used in conjunction with comparable basis sets (see, e.g., refs 39 and 40). CCSD(T), often denoted as the "gold standard" for accurate calculations, stands for the coupled-cluster (CC) method including a full account of single and double excitations, CCSD,[41] and a perturbative estimate of triple excitations (CCSD(T)).[42] In this respect, the recent work by Martin's group has led to the development of the revDSD-PBEP86 functional.[43] This represents a significant improvement with respect to B2PLYP, especially for transition states and noncovalent interactions, also showing very good performances (as B2PLYP) for equilibrium geometries and, especially, vibrational frequencies. Although the very recent D4 model for dispersion contributions[44] provides some improvement on energy evaluations, the D3(BJ) model[45,46] is already remarkably accurate. Therefore, since a full analytical implementation of second derivatives of energy is available for the latter,[47] we have decided to rely on the D3(BJ) scheme for incorporating dispersion effects. For all of these reasons, in this paper, we have characterized all stationary points of the reactive PES under consideration with the revDSD-PBEP86-D3(BJ) functional in conjunction with the jun-cc-pV(T+d)Z basis set.[48−50]

Subsequently, the energetics of all stationary points was accurately determined by exploiting the composite scheme denoted "HEAT-like", which is a state-of-the-art approach and will be described in detail in the following. Using this model as reference, the performance of different variants of the so-called "cheap" composite scheme[51,52] (described in the following as well), denoted as ChS, will be investigated. For comparison purposes, the CBS-QB3 model[53,54] will also be considered because it is extensively used in the evaluation of the thermochemistry of reactive systems.

For all levels of theory, the spin−orbit (SO) corrections for the Cl radical as well as for all open-shell species have been computed, within the state-interacting approach implemented in the MOLPRO program,[55−57] at the complete active space self-consistent field (CASSCF)[58,59] level in conjunction with the aug-cc-pVTZ basis set[60,61] and the full valence as active space. For Cl, calculations of the SO corrections have also been carried out using the multireference configuration interaction

(MRCI) method[62−64] in conjunction with the aug-cc-pVnZ sets, with n = T, Q, and 5. The computations for Cl have been performed to calibrate the level of theory to be used for the other radicals. Since it has been noted that the CASSCF values are almost independent of the basis set used (from 274.5 cm^{-1} with aug-cc-pVTZ to 275.7 cm^{-1} with aug-cc-pV5Z) and very close to the results at the MRCI level (276.1 cm^{-1} with aug-cc-pVTZ and 279.5 cm^{-1} with aug-cc-pVQZ), the cheapest level of theory has been chosen for all computations. In passing, we note that only at the MRCI/aug-cc-pV5Z level a value of 295.2 cm^{-1}, in very good agreement with the experimental result of 293.663 cm^{-1},[65] was obtained. In conclusion, the CASSCF/aug-cc-pVTZ level is expected to provide SO corrections affected by uncertainties not exceeding 0.2 kJ mol^{-1}.

Finally, electronic energies need to be corrected for the zero-point vibrational energy (ZPE) contribution. These corrections have been obtained both within the harmonic approximation and at the anharmonic level. In both cases, they have been computed using the double-hybrid revDSD-PBEP86-D3(BJ) functional in conjunction with the jun-cc-pV(T+d)Z basis set. Second-order vibrational perturbation theory (VPT2)[66] has been exploited for the evaluation of anharmonic ZPEs. Density functional theory (DFT) geometry optimizations and force field computations have been performed with Gaussian 16 quantum-chemical software.[67]

2.1.1. Reference Structures. The first issue that needs to be addressed is the effect of the reference geometries on the energetics. Indeed, a reactive PES can be very complicated and characterized by several alternative mechanisms. Therefore, the search of the stationary points of a reactive PES and their geometry optimizations might become the computational rate-determining step of the investigation. In this view, it is important to rely on a suitable and reliable level of theory for this purpose. The revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z level seems indeed to meet the requirements. However, further calculations to check the accuracy obtainable in structural determinations and their suitability for energetics have been performed. Concerning the latter point, the ChS approach applied to geometry optimizations has been employed, with the computational details being provided in the specific section.

Focusing on the reactant-well (RW) adduct (see Figure 1), as a test case (a radical species with a noncovalent bond), the
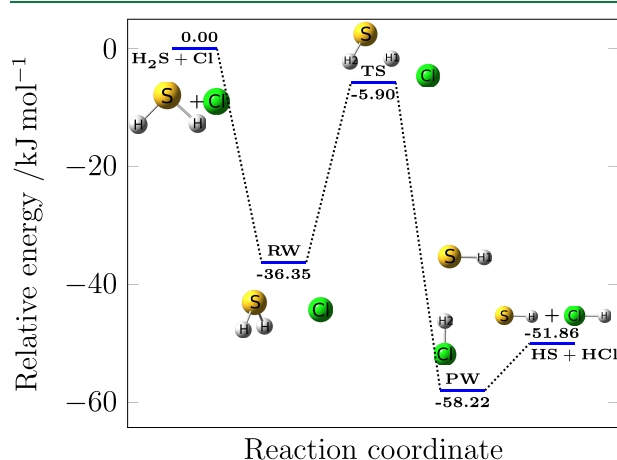

Figure 1. Reaction mechanism for the H$_2$S + Cl reaction. SO- and ZPE-corrected HEAT-like energies are reported.

equilibrium structure has been accurately evaluated by means of a combination of gradient and geometry approaches entirely based on coupled-cluster (CC) theory.[41] First of all, the so-called CCSD(T)/CBS+CV equilibrium structure has been obtained by minimizing the following gradient

$$\frac{dE_{CBS}}{dx} = \frac{dE^{CBS}(HF-SCF)}{dx} + \frac{d\Delta E^{CBS}(CCSD(T))}{dx} + \frac{d\Delta E_{CV}}{dx} \tag{3}$$

where the first two terms on the right-hand side are the energy gradients for the extrapolation to the complete basis set (CBS) limit and the last term incorporates the effect of core−valence (CV) correlation. The exponential formula introduced by Feller[68] and the two-point n^{-3} expression by Helgaker et al.[69] are used for the extrapolation to the CBS limit of the Hartree−Fock self-consistent field (HF-SCF) energy gradient and the CCSD(T) correlation contribution, respectively. The cc-pVnZ basis sets[49,60,70,71] have been employed, with n = Q, 5, and 6 being chosen for the HF-SCF extrapolation and n = Q and 5 for CCSD(T). Since the extrapolation to the CBS limit is performed within the frozen-core (fc) approximation, CV correlation effects have been considered by adding the corresponding correction, $d\Delta E_{CV}/dx$, where the all-electron−frozen-core energy difference is evaluated employing the cc-pCVQZ basis set.[72,73] It is to be noted that the CCSD(T)/CBS+CV equilibrium structure employing the aug-cc-pVnZ sets (n = T, Q, 5 for HF-SCF and n = T, Q for CCSD(T)) for the extrapolation to the CBS limit and the cc-pCVTZ basis set for the CV contribution has been obtained for all minima: H$_2$S, HS, HCl, and the reactant-well (RW) and product-well (PW) adducts.

The contributions due to the full treatment of triple ($\Delta r(fT)$) and quadruple ($\Delta r(fQ)$) excitations have been obtained at the "geometry" level, by adding the following differences to the CCSD(T)/CBS+CV geometrical parameters

$$\Delta r(fT) = r(CCSDT) - r(CCSD(T))$$
$$\Delta r(fQ) = r(CCSDTQ) - r(CCSDT) \tag{4}$$

where r denotes a generic structural parameter. The cc-pVTZ basis set has been used for the fT correction and the cc-pVDZ set for the fQ contribution. This implies that geometry optimizations at the fc-CCSDT[74−76]/cc-pVTZ, fc-CCSD(T)/cc-pVTZ, fc-CCSDTQ[77]/cc-pVDZ, and fc-CCSDT/cc-pVDZ levels have been performed.

2.1.2. HEAT-like Approach. The CC-based approach that has been denoted as HEAT-like takes the HEAT protocol[78−80] as reference and starts from the CCSD(T) method. In detail, the scheme (that will be denoted as "CBS+CV+DBOC+rel+fT +fQ" in the following) can be summarized as follows

$$E_{tot} = E^{CBS}_{HF-SCF} + \Delta E^{CBS}_{CCSD(T)} + \Delta E_{CV} + \Delta E_{rel} + \Delta E_{DBOC} + \Delta E_{fT} + \Delta E_{fQ} \tag{5}$$

In the expression above, "CBS" means that CCSD(T) energies, obtained within the fc approximation, have been extrapolated to the CBS limit. Analogously to what has been done for the CCSD(T)/CBS+CV gradient scheme, the extrapolation to the CBS limit has been performed in two steps, i.e., HF-SCF and the CCSD(T) correlation energies have been extrapolated

separately. The HF-SCF CBS limit has been evaluated by exploiting the exponential expression introduced by Feller[68]

$$E_{SCF}(n) = E_{SCF}^{CBS} + B\exp(-Cn) \tag{6}$$

For the CCSD(T) correlation contribution, the extrapolation to the CBS limit has been carried out using the n^{-3} formula by Helgaker and co-workers[69]

$$\Delta E_{corr}(n) = \Delta E_{corr}^{CBS} + An^{-3} \tag{7}$$

The cc-pVQZ and cc-pV5Z basis sets have been employed for the CCSD(T) correlation energy, whereas the cc-pVnZ sets, with n = Q, 5, and 6, have been used for HF-SCF.

By making use of the additivity approximation, the CV effects have been taken into account by means of the following expression

$$\Delta E_{CV} = E_{core+val} - E_{val} \tag{8}$$

thus incorporating the CCSD(T) energy difference obtained from all electrons (ae) and fc calculations, both in the cc-pCVQZ basis set.[72,73]

The diagonal Born–Oppenheimer correction, ΔE_{DBOC},[81–84] and the scalar relativistic contribution to the energy, ΔE_{rel}, have been computed at the HF-SCF/aug-cc-pVTZ and MP2/unc-cc-pCVQZ (correlating all electrons) levels, respectively, where MP2 stands for Møller–Plesset theory to second order[85] and "unc" denotes the use of the uncontracted basis set. The contributions due to relativistic effects have been evaluated using the lowest-order direct perturbation theory (second order in $1/c$, DPT2).[86]

In analogy to eq 4, corrections due to a full treatment of triples, ΔE_{fT}, and of quadruples, ΔE_{fQ}, have computed, within the fc approximation, as energy differences between CCSDT and CCSD(T) and between CCSDTQ and CCSDT calculations employing the cc-pVTZ and cc-pVDZ basis sets, respectively.

In addition to the scheme described above, a variant including the less expensive CCSDT(Q) method[87–89] has also been considered

$$E_{tot} = E_{HF-SCF}^{CBS} + \Delta E_{CCSD(T)}^{CBS} + \Delta E_{CV} + \Delta E_{rel} + \Delta E_{DBOC}$$
$$+ \Delta E_{fT} + \Delta E_{pQ} \tag{9}$$

where the only difference with respect to the former approach is the use of CCSDT(Q), which incorporates the quadruple excitations in a perturbative manner, instead of CCSDTQ. This approach will be denoted as "CBS+CV+DBOC+rel+fT+pQ".

It should be noted that the HEAT-like schemes also provide all possible intermediate approaches, such as CCSD(T)/CBS, CCSD(T)/CBS+CV, and CCSD(T)/CBS+CV+DBOC+rel, together with the results for the single-basis calculations like fc-CCSD(T)/cc-pVQZ and fc-CCSD(T)/cc-pV5Z.

All computations within the HEAT-like schemes have been carried out using the quantum-chemical CFOUR program package,[90] with the MRCC code[91] being interfaced to CFOUR to perform the calculations including quadruple excitations.

For all schemes described above, the geometries of the stationary points have been optimized using the double-hybrid revDSD-PBEP86-D3(BJ) in conjunction with the jun-cc-pV(T+d)Z basis set.

2.1.3. ChS Variants. The so-called cheap geometry approach (as already mentioned, denoted as ChS) was initially

developed for accurate molecular structure determinations[51,92] and then extended to energetic evaluations[52] for medium-sized systems. Recently, this scheme has been improved to accurately describe intermolecular, noncovalent interactions (thus leading to the definition of the jun-ChS approach).[93]

The starting point of this composite scheme is the CCSD(T) method in conjunction with a basis set of triple-ζ quality within the fc approximation. To improve the accuracy of this level of theory, the ChS model requires the incorporation of the extrapolation to complete basis set (CBS) and the core–valence (CV) correlation contribution, with all of them taken into account using MP2. Therefore, the general ChS model can be described by the following expression

$$E_{ChS} = E(CCSD(T)/TZ) + \Delta E^{CBS}(MP2) + \Delta E_{CV}(MP2) \tag{10}$$

In all cases, the extrapolation to the CBS limit has been performed:

(a) In two steps, as done in the framework of the HEAT-like approach. For HF-SCF, n = T, Q, and 5 have been used, while n = T and Q have been considered for MP2.

(b) In one step, according to

$$\Delta E^{CBS}(MP2) = \frac{4^3 E(MP2/QZ) - 3^3 E(MP2/TZ)}{4^3 - 3^3}$$
$$- E(MP2/TZ) \tag{11}$$

thus employing triple- and quadruple-ζ basis sets.

In analogy to the HEAT-like approach, the CV contribution, ΔE_{CV}, has been evaluated as energy (at the MP2 level) difference between ae and fc calculations. A basis set of triple-ζ quality has been employed for all ChS models.

The ChS variants employed in the present study can be classified as follows:

1. Standard, based on standard CCSD(T) and MP2 methods: ChS[51,52] and jun-ChS.[93]
2. F12, based on explicitly correlated F12 methods: ChS-F12.

For both standard ChS approaches, d-augmented basis sets have been employed for CCSD(T) calculations as well as for the extrapolation to the CBS limit: the cc-pV(n+d)Z[49,50,60] basis sets for ChS and jun-cc-pV(n+d)Z[48–50] for jun-ChS. The CV correlation correction has been evaluated employing the weighted-core–valence cc-pwCVTZ basis set.[73]

While the standard ChS approaches have already been largely applied (the reader is referred to the cited references for more details), the ChS-F12 model has been introduced for the first time. In detail, the general expression of eq 10 becomes the following

$$E_{ChS-F12} = E(CCSD(T)-F12/TZ) + \Delta E^{CBS}(MP2-F12)$$
$$+ \Delta E_{CV}(MP2-F12) \tag{12}$$

For evaluating the CC term (using CCSD(T)-F12[94–96] within the fc approximation), the cc-pVTZ-F12 basis set[97] has been used. As in the case of ChS and jun-ChS, the extrapolation to the CBS limit has been performed both in one and two steps. In the case of the two-step procedure, the HF/CBS contribution has been taken from ChS, while for the extrapolation of the correlation contribution with the MP2-F12 method,[98] two sets of basis sets have been considered:[97]

cc-pVDZ-F12/cc-pVTZ-F12 (extrapolation D,T) and cc-pVTZ-F12/cc-pVQZ-F12 (extrapolation T,Q). For the one-step extrapolation to the CBS limit, the two sets of bases have been used as well. In the case of the D,T extrapolation, formula 11 has been modified accordingly. The CV contribution has been computed using the cc-pCVTZ-F12 basis set.[99] Furthermore, both the F12a and F12b variants[94] have been employed, and, as described in ref 100, the geminal exponent (β) was set to 1.0 for all the F12 basis sets considered.

The final remark concerns the reference geometries used in the energy evaluations. In the case of the standard ChS approaches, the molecular structures employed have been obtained at the same level of theory, the only exception being the tests made on the effect of the reference geometries. This means that an expression analogous to eq 10 has been exploited to obtain the geometries of the stationary points. The only difference lies in the fact that, for structural determinations, the extrapolation to the CBS limit has been performed only in one step, according to eq 11. Conversely, for ChS-F12, the reference geometries have been determined at the fc-CCSD(T)-F12/cc-pVDZ-F12 level. As mentioned above, to investigate the structural effects on energetics, the jun-ChS model has also been applied to revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z geometries.

For the ChS approaches, the geometry optimizations and energy evaluations have been performed with the MOLPRO quantum-chemistry package.[55,56]

2.2. Kinetic Models. The reactive PES for the H_2S + Cl reaction is summarized in Figure 1. As evident from this figure, the Cl atom reacts with H_2S, thus leading to the $H_2S\cdots Cl$ intermediate well, denoted as RW. From here onward, the only available channel is, at least under atmospheric conditions, the addition/elimination reaction: RW isomerizes via the transition state, TS, to the H-bonded $HS\cdots HCl$ product well, denoted as PW, which then evolves to the products, i.e., HS + HCl. This process can be described in the framework of the AITSTME approach through a three-channel, two-well master equation. Depending on the examined temperature and pressure conditions, the global reaction rate is controlled either by the rate of conversion of RW into PW or by both its rate and that of formation of RW. The rate of decomposition of PW is generally fast with respect to the two other reaction channels and thus does not impact the global reaction rate.

The rate of formation of RW, which proceeds over a barrierless PES, has been determined at three different levels of theory. At the lowest level, the rate constant has been determined using phase space theory (PST),[2] assuming a $\frac{C}{R^6}$ attractive potential, with the coefficient C obtained by fitting the energies computed at various long-range distances (R) of the fragments using the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z level of theory. Alternatively, the rate constant has been determined using variational transition state theory (VTST) and variable reaction coordinate transition state theory (VRC-TST),[2] which, among the considered theoretical approaches, is the most accurate to describe properly the large amplitude motions that characterize this reaction channel.

VTST calculations have been performed within the rigid-rotor harmonic-oscillator (RRHO) approximation over a PES scanned as a function of the distance between Cl and H_2S in the 3.0−4.6 Å interval using a 0.2 Å step. Structures and vibrational frequencies of each point along the PES have been determined at the CASPT2/cc-pVDZ[101−103] level using a

(9e,7o) active space composed of the four H−S σ bonding and antibonding orbitals (4e,4o) and of the three p valence orbitals of chlorine (5e,3o). Higher-level energies have been determined at the CASPT2/aug-cc-pVTZ level over a (21e,13o) active space, consisting of the same active space used for the geometry optimization with the addition of the remaining valence electrons of Cl and S and of the three 2p orbitals and electrons of S. All CASPT2 calculations have been performed by state averaging over three states using a 0.2 level shift, the only exception being made for high-level calculations, which used a 0.25 IPEA shift. The IPEA shift was chosen as it allows for reproducing with relatively good accuracy the HEAT-like electronic energy of the entrance van der Waals well RW: −43.7 vs −41.89 kJ mol^{-1}.

VRC-TST calculations have been performed sampling the PES over multifaceted dividing surfaces constructed using three pivot points, positioned as schematized in Figure 2. Two
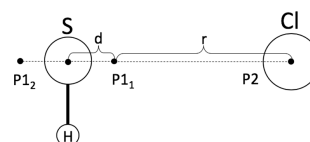

Figure 2. Scheme describing the position of the pivot points used to construct the multifaceted dividing surface for VRC-TST calculations. The H_2S plane is perpendicular to the plane of the picture, and the second H atom is hidden behind the S atom.

pivot points ($P1_1$ and $P1_2$) were placed in proximity of the S atom, symmetrically displaced along the axis perpendicular to the H_2S plane and passing from S by a distance d, while the third pivot point was centered on Cl. The multifaceted dividing surface is constructed varying the distance r between Cl and the pivot points, as described by Georgievskii and Klippenstein in ref 104, between 5 and 11 a_0 with 1.0 a_0 step. Calculations were repeated changing the position of the P1 pivot points varying d between 0.01 and 0.6 a_0, using 0.1 a_0 steps. Reactive fluxes were computed through Monte Carlo sampling using a 5% convergence threshold.

A minimum of 200 sampling points have been taken for each dividing surface. The calculated reactive fluxes have then been multiplied by a flat 0.9 factor to correct for recrossing dynamical effects. The 0.9 correction factor comes from the comparison of VRC-TST with trajectory calculations, which showed that VRC-TST total rate coefficients are generally about 10% greater than those obtained from related trajectory simulations,[105] when VRC-TST is applied at the level of theory used in the present study.[106] Energies have been computed at the CASPT2/cc-pVDZ level using a (5e,3o) active space consisting of the p orbitals of Cl, keeping the structures of H_2S frozen in its minimum-energy configuration and state averaging over three states. The sampled energies have been corrected, using a one-dimensional potential function of the distance between the S and Cl atoms, for geometry relaxation, energy accuracy, and SO effects. The correction energy for geometry relaxation (ΔE_{GEOM}) has been determined at the level of theory used to optimize geometries for VTST calculations, while that for the energy accuracy (ΔE_{HL}) has been estimated at the higher level used for VTST calculations (CASPT2/aug-cc-pVTZ level over a (21e,13o) active space). Spin−orbit corrections along the PES (ΔE_{SO}) have been computed using the state interacting method with a Breit−Pauli Hamiltonian

Table 1. Structural Parameters of the Stationary Points of the H₂S + Cl Reaction at Different Levels of Theory[g]

		ChS	jun-ChS	CC-F12[a]	DSD-D3[b]	CBS+CV[c]	CBS+CV+fT+fQ[d]	QCISD[e]	experiment[f]
H_2S	r(H−S)	1.336	1.336	1.337 (1.338)	1.338	1.335 (1.335)	1.335		1.3356
	θ(H−S−H)	92.2	92.2	92.2 (92.3)	92.5	92.3 (92.3)	92.3		92.11
$H_2S{\cdots}Cl$	r(H−S)	1.337	1.336	1.338 (1.339)	1.339	1.336 (1.336)	1.336	1.337	
	r(S−Cl)	2.582	2.586	2.585 (2.584)	2.595	2.567 (2.585)	2.568	2.670	
	θ(H−S−H)	93.0	92.8	92.6 (92.8)	92.9	92.9 (92.8)	92.9		
	θ(H−S−Cl)	87.4	87.6	87.5 (87.5)	88.0	87.4 (87.5)	87.4	88.2	
TS	r(H1−Cl)	1.645	1.644	1.630 (1.642)	1.633	1.642 (1.635)		1.614	
	r(H1−S)	1.470	1.469	1.476 (1.471)	1.472	1.468 (1.469)		1.478	
	r(H2−S)	1.339	1.339	1.340 (1.341)	1.341	1.340 (1.338)		1.339	
	θ(Cl−H1−S)	127.0	128.2	129.8 (128.0)	129.6	127.2 (127.4)		137.0	
	θ(H1−S−H2)	90.7	90.7	90.6 (90.8)	91.3	91.0 (90.8)			
	φ(Cl−H1−S−H2)	281.0	281.0	281.4 (281.0)	280.4	280.9 (281.1)			
HS⋯HCl	r(H2−S)	2.508	2.506	2.484 (2.506)	2.481	2.492 (2.492)	2.492		
	r(H1−S)	1.341	1.341	1.343 (1.343)	1.343	1.341 (1.341)	1.341		
	r(H2−Cl)	1.284	1.284	1.285 (1.286)	1.288	1.284 (1.284)	1.284		
	θ(H1−S−H2)	92.4	92.3	91.8 (91.8)	92.6	92.0 (91.9)	92.0		
	θ(Cl−H2−S)	176.5	176.1	175.8 (176.0)	176.6	175.8 (175.8)	175.8		
HS	r(H−S)	1.340	1.340	1.341 (1.342)	1.342	1.340 (1.340)	1.340		1.3406194(3)
HCl	r(H−Cl)	1.274	1.274	1.274 (1.276)	1.276	1.274 (1.274)	1.274		1.274565598(53)

[a]CC-F12 stands for fc-CCSD(T)-F12 in conjunction with the cc-pVDZ-F12 basis set. Values within parentheses have been obtained in conjunction with the cc-pVTZ-F12 basis set. [b]DSD stands for revDSD-PBEP86-D3(BJ) in conjunction with the jun-cc-pV(T+d)Z basis set. [c]CBS+CV stands for CCSD(T)/CBS+CV, with the aug-cc-pVnZ sets (n = T, Q) used for the extrapolation to the CBS limit and cc-pCVTZ for the evaluation of the CV contribution. Within parentheses are given the results for cc-pVQZ and cc-pV5Z being used for CBS and cc-pCVQZ for CV. See text. [d]CBS+CV+fT+fQ stands for CCSD(T)/CBS+CV augmented by fT and fQ contributions. See text. [e]In conjunction with the cc-pV(T+d)Z basis set. Values are taken from ref 8. [f]H₂S: ref 113; HCl: ref 114; HS: ref 115. [g]Distances are in angstroms, and angles are in degrees.

and a CASSCF wave function determined with a (9e,7o) active space state and six electronic states. The VRC-TST energy is thus computed as

$$E_{VRC-TST} = E(CASPT2(5e, 3o)/cc-pVDZ) + \Delta E_{GEOM}$$
$$+ \Delta E_{HL} + \Delta E_{SO} \qquad (13)$$

The rate of conversion of RW into PW has been calculated using both conventional and variational TST, as the reaction proceeds through a well-defined saddle point. Energies and structures of the stationary points (wells and saddle point) have been evaluated as described in Section 2.1. VTST calculations have been performed computing frequencies and structures along a minimum-energy path (MEP) determined through intrinsic reaction coordinate (IRC) calculations, at the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z level, taking 10 steps of 0.03 a_0 in both directions. Frequencies along the MEP have been computed both in Cartesian and in internal curvilinear coordinates.[107] Tunneling corrections have been taken into account by means of both small curvature theory (SCT)[108] and the Eckart model,[109] using an asymmetric potential.

VRC-TST calculations have been performed using the VaReCoF software,[106,110] while master equation simulations have been carried out using MESS.[111] VTST calculations (in Cartesian and internal curvilinear coordinates) as well as the determination of stationary points along the MEP to determine the VRC-TST potential and the construction of the input files for VaReCoF computations have been performed using a modified version of EStokTP.[112] All CASPT2, CASSCF, and SO calculations have been carried out using the MOLPRO program.[56] For all TST variants, where required, the energies

Table 2. Relative Electronic Energies[a] for the H₂S + Cl Reaction[g]

	reactants	RW	TS	PW	products
	H₂S + Cl	H₂S···Cl		HS···HCl	HS + HCl
CCSD(T)/VTZ	0.00	−28.29	7.74	−56.10	−44.90
CCSD(T)/VQZ	0.00	−37.57	1.30	−58.23	−46.35
CCSD(T)/V5Z	0.00	−41.32	−0.84	−59.11	−46.95
CCSD(T)/CBS	0.00	−44.84	−3.10	−60.10	−47.51
CBS+CV	0.00	−45.09	−3.42	−60.30	−47.65
CBS+CV+DBOC	0.00	−45.09	−2.05	−60.15	−47.64
CBS+CV+DBOC+rel	0.00	−45.11	−2.31	−59.98	−47.45
CBS+CV+DBOC+rel+fT+pQ	0.00	−45.14	−3.41	−60.09	−47.45
CBS+CV+DBOC+rel+fT+fQ	0.00	−45.11	−3.33	−60.08	−47.44
ChS[b]	0.00	−42.81	−2.49	−60.29	−47.94
		[−42.94]	[−2.54]	[−60.41]	[−47.44]
jun-ChS[b]	0.00	−43.89	−3.46	−60.19	−47.58
		(−43.81)	(−3.45)	(−60.18)	(−47.59)
		[−43.89]	[−3.46]	[−60.19]	[−47.58]
ChS-F12a/F12b CBS(D,T)[c]	0.00	−44.64/−44.58	−1.97/−1.93	−59.80/−59.80	−47.11/−47.15
		[−43.97/−43.91]	[−2.74/−2.70]	[−60.10/−60.10]	[−47.06/−47.09]
ChS-F12a/F12b CBS(T,Q)[c]	0.00	−43.87/−43.81	−3.32/−3.29	−60.33/−60.33	−47.42/−47.46
		[−43.54/−43.48]	[−3.39/−3.35]	[−60.27/−60.27]	[−47.28/−47.31]
CBS-QB3[d]	0.00	−48.16	−8.12	−61.13	−53.43
revDSD-PBEP86-D3(BJ)[e]	0.00	−47.82	−3.85	−63.09	−46.11
		(−45.50)	(−1.80)	(−59.43)	(−46.00)
QCISD/cc-pV(T+d)Z[f]	0.00	−23.14	20.17		−43.76
PMP2/CBS[f]	0.00	−46.69	−9.08		−52.72
ae-CCSD(T)/CBS[f]	0.00	−46.48	−7.57		−48.16

[a]Unless otherwise stated, reference geometries at the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z level. [b]Reference geometries at the same level as the energy evaluation. Extrapolation to the CBS limit in one step. Values within parentheses: revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z geometries as reference. Values within square brackets: extrapolation to the CBS limit in two steps. [c]Reference geometries at the fc-CCSD(T)-F12/cc-pVDZ-F12 level. Extrapolation to the CBS limit in one step. Values within square brackets: Extrapolation to the CBS limit in two steps [d]Reference geometries at the B3LYP/6-31G(d) level. [e]Values obtained in conjunction with the jun-cc-pV(Q+d)Z basis set. Within parentheses: results for the jun-cc-pV(T+d)Z basis set. In both cases: reference structures at the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z level. [f]Results from ref 8. [g]Values in kJ mol⁻¹.

of the stationary points of the PES have been considered at the CBS+CV+DBOC+rel+fT+fQ level, also including SO corrections and anharmonic ZPE contributions.

3. RESULTS AND DISCUSSION

The reaction mechanism for the H₂S + Cl reaction is shown in Figure 1 and, as mentioned in the Introduction, none of the previous computational works provided a complete picture for it. In particular, the presence of both H₂S···Cl and HS···HCl has been thoroughly investigated only in the present study. In the following, the molecular structures of the stationary points are first presented. Then, the associated thermochemistry and kinetics are reported and discussed.

3.1. Molecular Structures. The structural parameters of the stationary points located on the H₂S + Cl reactive PES, optimized at different levels of theory, are collected in Table 1. From the inspection of this table, the first conclusion that can be drawn is that for covalently bonded compounds, i.e., H₂S, HS, and HCl, there is a perfect agreement among ChS, jun-ChS, CCSD(T)/CBS+CV, and revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z results. For the intermediate adducts, i.e., RW and PW, we note again that the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z covalent bond lengths agree very well with those determined by means of composite schemes. The discrepancies are evident only for the noncovalent distances, i.e., S···Cl in RW and S···H in PW. However, the differences are well within 0.01−0.02 Å, meaning that their impact on energetics is

expected to be negligible (as will be demonstrated in the next section).

To further test the performance of different ChS approaches and of the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z level, as mentioned in the Section 2, the CCSD(T)/CBS+CV+fT+fQ scheme has been exploited for the RW adduct. This allowed us to confirm beyond all doubts the accuracy of the ChS models and the suitability of the revDSD-PBEP86-D3(BJ) functional for the characterization of noncovalently bonded systems. Moreover, the results of Table 1 point out the very good performance of the CCSD(T)-F12/cc-pVDZ-F12 level of theory, which is only marginally improved by the use of the cc-pVTZ-F12 basis set. Indeed, also in the case of fc-CCSD(T)-F12 calculations, a very good agreement with the structural parameters obtained from composite approaches can be noted.

Finally, it is apparent that the results at the QCISD/cc-pV(T+d)Z level from ref 8 show remarkable deficiencies in the description of the geometrical parameters. In particular, the θ(Cl−H1−S) angle in the transition state is about 10° larger than the corresponding best-estimated value. Another important deviation is observed for the RW adduct, the S···Cl distance being too long by ∼0.1 Å and the θ(H−S−Cl) too large by ∼1°.

3.2. H₂S + Cl Reaction: Previous Works. In the following, a critical analysis of the previous computational investigations on the H₂S + Cl reaction is reported.

In the work by Wilson et al.,[34] the RW and TS stationary points were characterized by evaluating the electronic energies at the ae-MP4/6-311+G(2df,p) level of theory,[116,117] using ae-MP2/6-311G** reference structures. Moreover, the addition and abstraction pathways were treated as separate processes instead of combining them in an addition/elimination mechanism, as done in this work. As a consequence, the rate constant was calculated with standard TST considering only the abstraction step. However, a value 1 order of magnitude lower than the experimental one is mainly explained by an incorrect evaluation of the barrier height.

In a more recent work, in which the PW adduct was not considered as well, Resende et al.[8] optimized the geometries of the stationary points at the QCISD/cc-pV(T+d)Z level of theory.[118,119] Using these reference structures, the energies were subsequently evaluated at the PMP2 level of theory[120−123] in conjunction with cc-pV(n+d) basis sets (with n = D, T, Q), extrapolated to the CBS limit using the exponential expression introduced by Feller, and finally added to the corresponding ae-CCSD(T)/cc-pV(T+d)Z energies. Resende et al. thus employed for the energetics a composite scheme similar to our ChS. Using the aforementioned geometries and energies, the rate constant was subsequently calculated with VTST, thereby obtaining a value that was 1 order of magnitude higher than the experimental datum. To justify their overestimated result, the authors advocated a failure of transition state theory. The role of explicit dynamical effects (e.g., recrossing) was analyzed by performing some explicit trajectory computations. Although the results obtained were not conclusive, they induced the authors to advocate the inadequacy of TST in reproducing the significant role of vibrationally excited H_2S in stabilizing the $H_2S\cdots Cl$ adduct.

3.3. H_2S + Cl Reaction: Thermochemistry. The relative electronic energies, obtained at various computational levels, are detailed in Table 2. In particular, fc-CCSD(T) results in conjunction with basis sets of increasing size up to the CBS limit are collected together with the CBS+CV, CBS+CV+DBOC, CBS+CV+DBOC+rel, and CBS+CV+DBOC+rel+fT+fQ values. These allow us to inspect the trend of the relative energies as a function of the basis set as well as the role of the various contributions.

For all stationary points, it is noted that even for a basis set as large as cc-pV5Z the results are not yet converged, with values being quantitatively accurate only at the CBS limit. A particular remark is warranted for the fc-CCSD(T)/cc-pVTZ level because it is often used in the investigation of reactive PESs, in particular for astrophysical purposes. This level predicts the RW adduct to lie ∼28 kJ mol^{-1} below the reactants, which means less stable by about 17 kJ mol^{-1} with respect to what evaluated by more accurate calculations. Furthermore, at this level, the transition state is predicted to emerge above the reactants by more than 7 kJ mol^{-1}. By enlarging the basis set, we note that TS is still emerged with the quadruple-ζ set and becomes barely submerged only when the quintuple-ζ basis is used.

Moving from the fc-CCSD(T)/CBS level on, it is observed that the CV corrections are small, these being, on average, of the order of 0.2 kJ mol^{-1}. Roughly of the same order of magnitude are the combined DBOC and scalar relativistic contributions, with the only exception being the transition state, for which the correction amounts to about 1 kJ mol^{-1}. Analogous is the situation for the fT and fQ contributions, with the corresponding correction for TS being, however, in the

opposite direction. The overall conclusion is that the CBS+CV level provides results in very good agreement with the CBS+CV+DBOC+rel+fT+fQ model. Furthermore, it is noted that when basis sets more suitable for describing the third-row elements are used (i.e., cc-pV(n+d)Z), for the CBS+CV model, smaller basis sets can be employed. In another test, the extrapolation to the CBS limit has been applied to ae-CCSD(T)/cc-pCVnZ (n = Q, 5) energies. Since the differences with the CBS+CV level lie within 0.2 kJ mol^{-1}, the validity of the additivity approximation for the CV contribution has been confirmed. From Table 2, it is also evident that the perturbative (instead of full) treatment of quadruple excitations marginally affects the relative energies.

Moving to less expensive approaches, the very good performance of all ChS variants deserves to be highlighted. Indeed, ChS, jun-ChS, and ChS-F12 provide very similar results, which deviate, in the cases, by about 2 kJ mol^{-1} from our best estimates (i.e., CBS+CV+DBOC+rel+fT+fQ). The jun-ChS values obtained using also the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z reference geometries are reported, within parentheses, in Table 2. A perfect agreement between the two sets of jun-ChS relative energies is observed, the differences being well below 0.1 kJ mol^{-1}. Such a comparison confirms the accuracy of the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z structures for thermochemical studies. In addition, the revDSD-PBEP86-D3(BJ) functional provides accurate results also for the energetics, as clear from the results collected in Table 2. We note that this functional performs better when combined with the jun-cc-pV(T+d)Z basis set for all minima, while the jun-cc-pV(Q+d)Z set seems to be required for correctly evaluating the relative energy of the transition state.

In Table 2, for all ChS variants, relative energies based on extrapolations to the CBS limit performed in both one and two steps are reported. It is worth noting that the two approaches provide very similar results. This is an important outcome because the extrapolation in one step avoids the problems related to basis sets of quintuple-ζ quality for which convergence of the HF-SCF energy can be troublesome, especially for open-shell species. Furthermore, the computation of the corresponding integrals might become particularly expensive for large systems.

As mentioned in Section 2, both F12a and F12b variants have been employed in the ChS-F12 model. From the results of Table 2, it is evident that the two approximations provide nearly coincident results. For this scheme, the extrapolation to the CBS limit using double- and triple-ζ basis sets has also been tested. It is apparent that even in this case the results are very good, showing deviations from the best-estimated values well within 2 kJ mol^{-1}. However, this ChS variant is the only one presenting some difference (about 1 kJ mol^{-1}) between the one- and two-step procedures for the extrapolation to the CBS limit.

As already pointed out in the literature (see, e.g., refs 124 and 125), despite its widespread use, the CBS-QB3 model provides disappointing results for energetics and, in particular, for barrier heights. In fact, the transition state lies too low in energy by about 5 kJ mol^{-1} with respect to the best estimate.

Finally, the comparison of our results with those from ref 8 is deserved. First, it is noted that the QCISD/cc-pV(T+d)Z level, whose unsuitability in the determination of reference structures has been previously pointed out, provides unreliable energetics, the transition state being predicted ∼20 kJ mol^{-1}

Table 3. Best-Estimated Relative Electronic Energies (Including Spin−Orbit) together with ZPE and Thermochemical Corrections[d]

	reactants	RW	TS	PW	products
	$H_2S + Cl$	$H_2S\cdots Cl$		$HS\cdots HCl$	$HS + HCl$
CBS+CV+DBOC+rel+fT+fQ[a]	0.00	−41.89	−0.05	−56.79	−46.25
	(3.28)	(0.06)	(0.01)	(0.00)	(2.10)
anharm-ZPE[b]	0.00	5.54	−5.85	−1.43	−5.61
harm-ZPE[b]	0.00	5.86	−5.27	−1.55	−5.86
$\Delta H° - \Delta H_0°$ [c]	0.00	−2.78	−3.44	−0.58	0.94

[a]Within parentheses, the SO corrections (at the CASSCF/aug-cc-pVTZ level) are given. [b]Relative ZPE corrections at the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z level. [c]Standard state: 1 atm, 298 K; at the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z level. [d]Values in kJ mol^{-1}.

above the reactants. The situation improves on moving to the PMP2 level and, in particular, to the level denoted in Table 2 as ae-CCSD(T)/CBS, which corresponds, as previously described, to a composite scheme similar to the ChS model. Despite this improvement, the barrier height is underestimated by about 3 kJ mol^{-1}.

To complete the thermochemistry of the $H_2S + Cl$ reaction, the relative energies need to be corrected for SO coupling (which was neglected in ref 8). Doing so for our best level of theory leads to the values collected in the first row of Table 3. The major consequence is that the transition state is no longer submerged and lies above the reactants by 0.12 kJ mol^{-1}. However, a further correction should be taken into consideration. Once ZPE is incorporated, the transition state is again submerged by about 5 kJ mol^{-1}. Furthermore, it has to be noted that ZPE corrections evaluated within the harmonic approximation are in very good agreement with the anharmonic values.

The computed standard formation enthalpies (298.15 K, 1 atm) of H_2S and products (HS and HCl), obtained from our electronic structure calculations and the experimental formation enthalpies of the H, S, and Cl atomic species,[126] are in remarkable agreement with the most accurate experimental data taken from ref 126, except for that of HS,[127] namely, 141.86 vs 141.87, −91.82 vs −92.17, and −20.34 vs −20.50 kJ mol^{-1} for HS, HCl, and H2S, respectively. All details are provided in the Supporting Information (SI) (see Table S1). From these results, we obtain a sub-kJ mol^{-1} accuracy also for the reaction enthalpy of the title reaction (50.92 vs 51.24 kJ mol^{-1}).

3.4. $H_2S + Cl$ Reaction: Rate Constants. The rate constant for the H abstraction from H_2S by Cl has been computed at different levels of theory, as described in detail in Section 2.2, by solving the multiwell one-dimensional master equation using the chemically significant eigenvalue (CSE) method within the Rice−Ramsperger−Kassel−Marcus (RRKM) approximation, as detailed by Miller and Klippenstein in ref 128.

Although the use of a master equation model is generally not necessary to evaluate the rate constant of an abstraction reaction, there are several reasons that warrant its employment, as described by Cavallotti et al. in ref 112. In fact, it allows one to properly limit the contributions to the reactive flux from energy states below the asymptotic energies of the fragment, as well as to account for limitations to the reaction fluxes determined by an eventually slow rate of formation of the precursor complex. This is indeed the situation for the present system, for which the flux through the outer transition state, leading to the formation of the precursor complex, and that through the inner transition state, leading to H abstraction and

the formation of the product complex, have comparable values. The consequence is that part of the flux passing through the outer transition state is reflected from the inner transition state.

As mentioned in Section 2.2, reactive fluxes through the outer transition state, which occur on a barrierless MEP, have been computed at three different levels of theory: VRC-TST, VTST, and PST. The Cl−H_2S interaction potential used to determine the VRC-TST flux uncorrected for geometry relaxation, active-space size and basis-set size, and at different levels of corrections is reported in Figure 3. It is interesting to
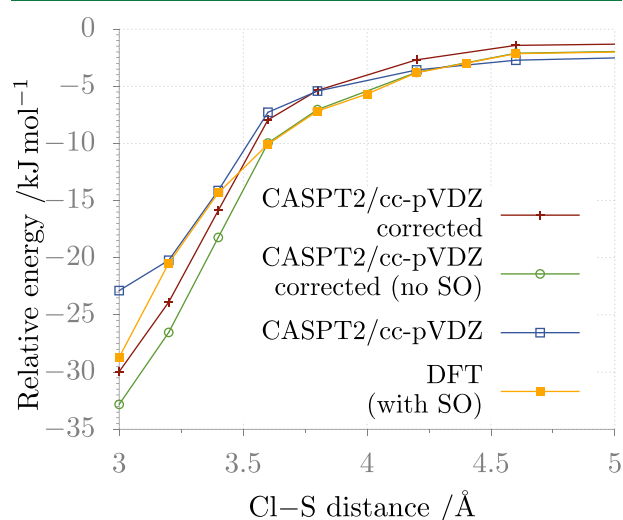

Figure 3. Interaction potential between Cl and H_2S calculated at different levels of theory: the uncorrected CASPT2/cc-pVDZ level; the CASPT2/cc-pVDZ level corrected for geometry relaxation, high-level energy contributions, and SO effects (see eq 13); the CASPT2/cc-pVDZ level corrected for geometry relaxation and high-level energy contributions; and the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z level corrected for SO effects.

observe that the corrected CASPT2 potential used in VRC-TST calculations is similar, though not exactly overlapped with that determined at the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z level of theory. This suggests that this functional may be used to perform VRC-TST calculations for open-shell systems for which it may be cumbersome, e.g., for difficulties in converging to a proper active space, to determine the interaction potential at the CASPT2 level.

As already mentioned, master equation simulations have been performed using MESS software. The collisional energy transfer probability has been described by means of the single exponential down model[129] with a temperature dependence

$\langle \Delta E \rangle_{down}$ of $260(T/298)^{0.875}$ cm^{-1} in an argon bath gas. Different models have been employed to compute reaction fluxes through the inner and the outer transition states. The highest-level simulations have been obtained using VRC-TST for the outer TS and VTST in curvilinear internal coordinates for the inner TS (referred as VTSTin), the results being reported in Table 4 and compared with experimental data in Figure 4.

Table 4. Rate Coefficients for the H$_2$S + Cl Reaction at Various Temperatures (Pressure = 1 atm)[a,b]

T (K)	VRC-VTSTin	VTST-VTSTin	PST-VTSTin
200	1.01×10^{-10}	1.02×10^{-10}	1.31×10^{-10}
225	9.49×10^{-11}	9.53×10^{-11}	1.22×10^{-10}
250	8.94×10^{-11}	9.03×10^{-11}	1.14×10^{-10}
275	8.51×10^{-11}	8.64×10^{-11}	1.08×10^{-10}
300	8.16×10^{-11}	8.33×10^{-11}	1.03×10^{-10}
325	7.89×10^{-11}	8.07×10^{-11}	9.84×10^{-11}
350	7.63×10^{-11}	7.87×10^{-11}	9.50×10^{-11}
375	7.44×10^{-11}	7.69×10^{-11}	9.21×10^{-11}
400	7.26×10^{-11}	7.56×10^{-11}	8.98×10^{-11}
425	7.13×10^{-11}	7.45×10^{-11}	8.78×10^{-11}
450	6.99×10^{-11}	7.36×10^{-11}	8.62×10^{-11}
475	6.89×10^{-11}	7.29×10^{-11}	8.49×10^{-11}
500	6.80×10^{-11}	7.24×10^{-11}	8.38×10^{-11}
600	6.56×10^{-11}	7.17×10^{-11}	8.16×10^{-11}
700	6.45×10^{-11}	7.24×10^{-11}	8.16×10^{-11}
800	6.41×10^{-11}	7.41×10^{-11}	8.31×10^{-11}
900	6.44×10^{-11}	7.65×10^{-11}	8.57×10^{-11}

[a]The various prefixes stand for the theoretical methods used to handle the barrierless entrance channel, while the VTSTin suffix means that the inner TS is handled with VTST in curvilinear internal coordinates. The barrierless exit channel is always treated with PST. [b]Values in cm^3 molecule^{-1} s^{-1}.

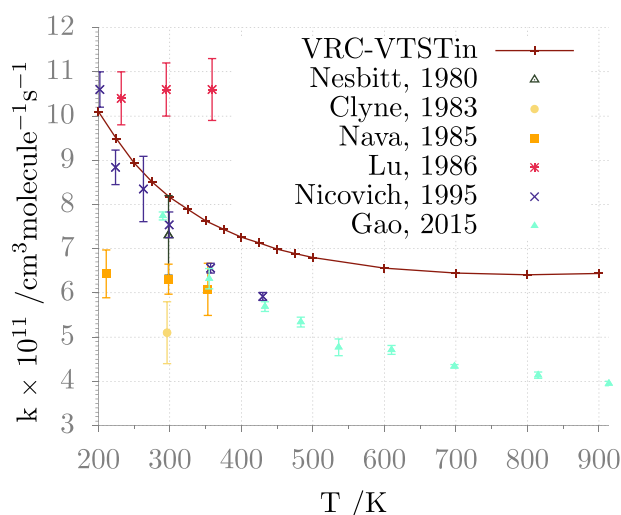**Figure 4.** H$_2$S + Cl global rate constant: comparison between computed (VRC-TST theory for the outer TS, VTST with vibrational frequencies evaluated using curvilinear internal coordinates for the inner TS, small curvature theory for tunneling) and experimental data.

The comparison between calculated and experimental data shows an excellent agreement at 300 K, the temperature at which most of the measurements were made. Indeed, the calculated 7.76×10^{-11} cm^3 molecule^{-1} s^{-1} value is in quite

good agreement with the 7.4×10^{-11} cm^3 molecule^{-1} s^{-1} datum recommended by Atkinson et al.[33] on the basis of an extensive review. Furthermore, the calculated rate is in excellent agreement with the rate constant measured in the 200−433 K temperature range by Nicovich et al.,[30] from which it differs by about 12% at most. The calculated global rate constant is almost pressure-independent in the considered conditions, an outcome that agrees with the experimental data that show that there is no measurable collisional stabilization of the entrance well.[32] However, the temperature trend is not perfectly reproduced, as the experimental data are slightly underestimated at low temperatures and slightly overestimated at high temperatures. The discrepancy is more evident for the recent measurements by Gao et al.,[32] which are overestimated by a factor of 1.5 at 900 K. Even if such a disagreement is relatively small and it has been observed only with respect to a single set of experimental data, it is anyway useful to try to understand its origin. It is first of all noted that the structure of the saddle point of the inner TS has an optical isomer, which thus effectively doubles the density of states (DOS) of the TS and, consequently, the rate constant. The two optical isomers are separated by a second-order saddle point with a barrier of about 10 kJ mol^{-1}. Therefore, it is likely that, as the temperature increases, the two isomers interconvert among themselves. If this is the case, then the DOS of the TS is overestimated by up to a factor of 2. To investigate whether this can be the case, the one-dimensional (1D) PES for the conversion between the two isomers has been determined as a function of the Cl−H−S−H dihedral angle and the partition function of the corresponding vibrational internal motion has been replaced with a 1D hindered rotor model. The rate constants calculated at different temperatures are compared with the experimental temperature-dependent data in Figure 5. The computed results are now in quantitative agreement with experiments at high temperature and differ at most by a factor of 1.38 at 200 K. This outcome confirms, as it is well known in the literature, that one of the key aspects in the estimation of an accurate rate constant using TST is the proper description
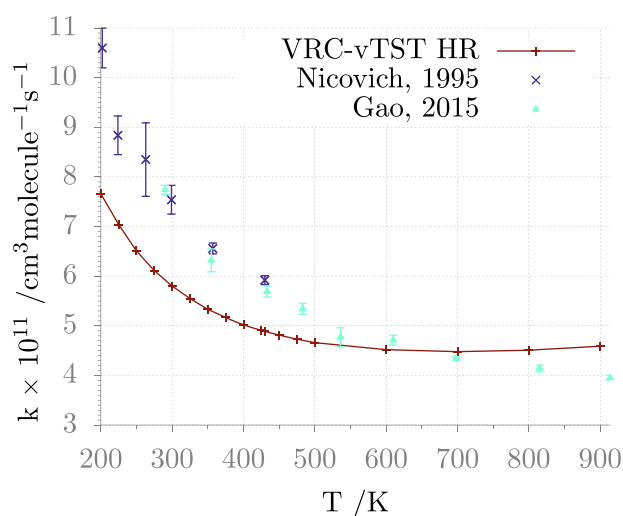

Figure 5. H$_2$S + Cl global rate constant: comparison between computed (same level as in Figure 4, but modeling of the internal motion for the interconversion between the optical isomers of the TS with a 1D hindered rotor) and experimental data.

of anharmonic internal motions and that it is sometimes necessary to use different models depending on the investigated temperature and pressure conditions.

To compare the contribution of the entrance and inner channels to the global rate constant, it is interesting to report the rates of each channel computed solving the ME fictitiously enhancing the rate of the other channel (Figure 6). As it can be
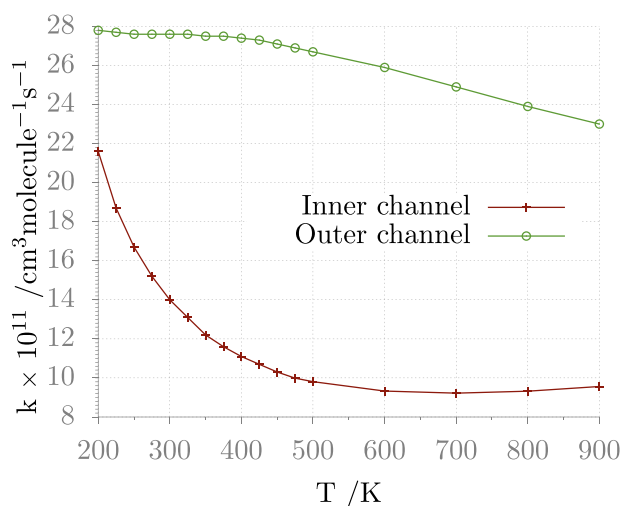

Figure 6. Rate constants of the outer and inner channels computed using VRC-TST and VTST, respectively.

observed, both rates exhibit a negative activation energy, in agreement with experimental observations, and their values are comparable, though the rate of the inner channel is smaller, and it thus impacts more significantly the global reaction flux.

The impact of the level of theory chosen to compute the inner and outer TS fluxes on the global rate constant is analyzed in Figure 7. It can be observed that, for the outer channel, VRC-TST and VTST give similar results that differ at most by a factor of 1.1, while PST predictions deviate by up to a factor of 1.2. Despite this, it is interesting to notice that there is a slight but significant qualitative difference in the temperature dependence between the rate constants computed using VRC-TST, in better agreement with the experimental trend, and those determined at the other theoretical levels. It should also be recalled that the global rate constant is mostly controlled by the rate of the inner TS, so that the differences between the levels of theory used for the outer TSs are mitigated. The analysis of the impact of the chosen theoretical level for the inner TS shows that variational effects have a minor impact, though the rate constant computed using the internal coordinate model is in better agreement with experimental data. The most relevant effect on the rate constant, as commented above, is given by the use of the 1D hindered rotor model. Finally, though not shown, it has been found that using the Eckart model rather than small curvature theory to compute the tunneling contributions has a negligible impact on the rate constant evaluation. The reason is that tunneling corrections are small for this system because the RW adduct is not significantly collisionally stabilized in the examined temperature and pressure conditions and the energy barrier is submerged with respect to reactants.

It can be concluded that the elementary processes that contribute to the reactive fluxes change depending on the
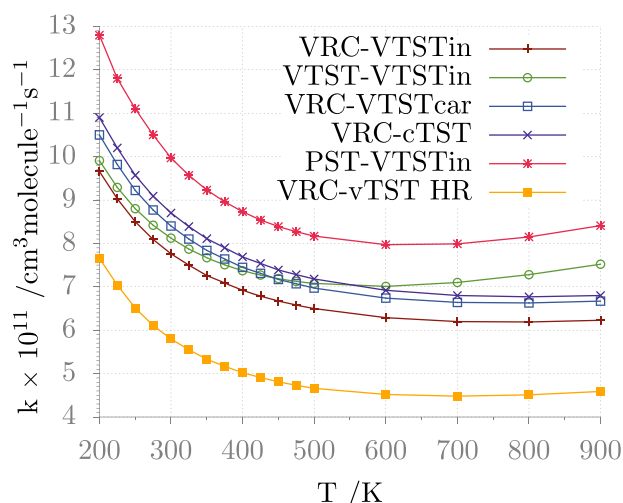

Figure 7. Global rate constants computed using different theoretical approaches to determine the fluxes through the inner and outer transition states. The nomenclature is "outer TS−inner TS": (1) the outer TS flux: VRC-TST (VRC), VTST (VTST), or PST (PST); (2) the inner TS flux: the VTST level with vibrational frequencies computed using internal curvilinear coordinates (VTSTin) or Cartesian coordinates (VTSTcar), conventional TST (cTST), VTST with one internal mode modeled as a 1D hindered rotor (vTST HR).

ranges of temperatures and pressures that are investigated and that their rate must be determined at a suitable level of theory to obtain quantitative agreement with experimental data. For example, canonical VTST is not apt to study the entrance channel for this system, not much because it assumes a thermal distribution (thus allowing for computing a canonical $k(T)$ rate constant instead of the microcanonical E, or E,J resolved rates, $k(E)$ or $k(E,J)$), but rather because, as it is often implemented in the literature, it uses the harmonic approximation to evaluate reactive fluxes along the minimum-energy path, which is improper for loosely interacting fragments. A more "proper" evaluation of the reactive fluxes is that given by VRC-TST.

4. CONCLUSIONS

The kinetics of radical−molecule reactions is of remarkable interest in several fields including, inter alia, atmospheric- and astrochemistry. However, obtaining quantitative rate constants for such reactions by means of theoretical methods is challenging because of the difficulties that can be faced in the accurate description of some stationary points (intermediates and/or transition states). Indeed, they might show strong correlation effects, the situation being more involved when third-row atoms are present. Furthermore, SO coupling might be relevant for open-shell species. On these grounds, the first aim of this paper was to investigate a prototypical reaction of this kind, namely, the H_2S + Cl addition/elimination reaction, beyond the usual "gold standard" of quantum-chemical calculations, represented by CCSD(T) possibly including the extrapolation to the CBS limit. To this aim, a HEAT-like approach, which includes the full treatment of triple and quadruple excitations together with diagonal Born−Oppenheimer corrections and relativistic effects, combined with a proper treatment of the SO coupling has been employed. This level of theory, in conjunction with anharmonic ZPE corrections evaluated using the double-

hybrid revDSD-PBEP86-D3(BJ) functional in the framework of the VPT2 model, is expected to fulfill a sub-kJ mol^{-1} accuracy, thus allowing an unbiased analysis of the ability of different kinetic models in reproducing the experimental reaction rates. In the present work, different approaches of increasing accuracy have been employed to describe the barrierless entrance channel of the reaction, whose role (in evaluating the global reaction rate) depends on the examined temperature and pressure conditions.

In this connection, even the quite simple PST leads to results within a factor of 2 with respect to their experimental counterparts, whereas the more refined VTST and, especially, VRC-TST models lead to results in quantitative agreement with experiment. These outcomes show unambiguously that this reaction can be well described by models based on the transition state theory, provided that the underlying electronic structure computations are sufficiently accurate and that barrierless channels are properly described.

Furthermore, in view of extending the accuracy reached by our approach to reactive PESs involving larger systems, we have tested the performance of computationally less expensive composite schemes, which would become indeed unavoidable in such cases. Our conclusion is that different variants of the so-called cheap approach perform remarkably well. At the same time, last-generation double-hybrid functionals can be profitably used to optimize geometries and evaluate vibrational contributions. In summary, in our opinion, a promising route for computing reaction rates in semiquantitative agreement with experiment (i.e., well within a factor of 2) for quite large molecular systems can be based on ChS energy evaluations of the relevant stationary points coupled to TST for the activated steps and to PST for barrierless steps. All of these ingredients must be finally introduced in a master equation model of the overall reaction network, which must include all of the elementary processes that can contribute to the reactive fluxes, with rates computed at a suitable level of theory.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.0c00354.

> Input file (example) for the RRKM-ME calculation of the rate constant for the H$_2$S + Cl reaction using the MESS package; thermochemistry results (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Vincenzo Barone − *Scuola Normale Superiore, I-56126 Pisa, Italy;* ⓘ orcid.org/0000-0001-6420-4107; Email: vincenzo.barone@sns.it

Authors

Jacopo Lupi − *Scuola Normale Superiore, I-56126 Pisa, Italy;* ⓘ orcid.org/0000-0001-6522-9947

Cristina Puzzarini − *Department of Chemistry "Giacomo Ciamician", University of Bologna, I-40126 Bologna, Italy;* ⓘ orcid.org/0000-0002-2395-8532

Carlo Cavallotti − *Department of Chemistry, Materials, and Chemical Engineering "G. Natta", Politecnico di Milano, I-20131 Milano, Italy;* ⓘ orcid.org/0000-0002-9229-1401

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.0c00354

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Allen, J. W.; Goldsmith, C. F.; Green, W. H. Automatic estimation of pressure-dependent rate coefficients. *Phys. Chem. Chem. Phys.* **2012**, *14*, 1131−1155.

(2) Bao, J. L.; Truhlar, D. G. Variational transition state theory: theoretical framework and recent developments. *Chem. Soc. Rev.* **2017**, *46*, 7548−7596.

(3) Klippenstein, S. J.; Cavallotti, C. Mathematical Modelling of Gas-Phase Complex Reaction Systems: Pyrolysis and Combustion. In *Computer Aided Chemical Engineering*; Faravelli, T.; Manenti, F.; Ranzi, E., Eds.; Elsevier, 2019; Vol. 45, pp 115−167.

(4) Sun, L.; Hase, W. L.; Song, K. Trajectory Studies of S$_N$2 Nucleophilic Substitution. 8. Central Barrier Dynamics for Gas Phase Cl$^-$ + CH$_3$Cl. *J. Am. Chem. Soc.* **2001**, *123*, 5753−5756.

(5) Pritchard, H. O. Recrossings and Transition-State Theory. *J. Phys. Chem. A* **2005**, *109*, 1400−1404.

(6) Bowman, J. M.; Zhang, X. New insights on reaction dynamics from formaldehyde photodissociation. *Phys. Chem. Chem. Phys.* **2006**, *8*, 321−332.

(7) Townsend, D.; Lahankar, S. A.; Lee, S. K.; Chambreau, S. D.; Suits, A. G.; Zhang, X.; Rheinecker, J.; Harding, L. B.; Bowman, J. M. The Roaming Atom: Straying from the Reaction Path in Formaldehyde Decomposition. *Science* **2004**, *306*, 1158−1161.

(8) Resende, S. M.; Pliego, J. R., Jr.; Vandresen, S. Ab initio study of the Cl + H$_2$S atmospheric reaction: is there a breakdown of the transition state theory. *Mol. Phys.* **2008**, *106*, 841−848.

(9) Möller, D. Estimation of the global man-made sulphur emission. *Atmos. Environ. (1967)* **1984**, *18*, 19−27.

(10) Seinfeld, J. H.; Pandis, S. N. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*; John Wiley & Sons, 2016.

(11) Kotra, J. P.; Finnegan, D. L.; Zoller, W. H.; Hart, M. A.; Moyers, J. L. El Chichón: Composition of plume gases and particles. *Science* **1983**, *222*, 1018−1021.

(12) O'Dwyer, M.; Padgett, M.; McGonigle, A.; Oppenheimer, C.; Inguaggiato, S. Real-time measurement of volcanic H$_2$S and SO$_2$ concentrations by UV spectroscopy. *Geophys. Res. Lett.* **2003**, *30*, No. 1652.

(13) Inguaggiato, S.; Diliberto, I. S.; Federico, C.; Paonita, A.; Vita, F. Review of the evolution of geochemical monitoring, networks and methodologies applied to the volcanoes of the Aeolian Arc (Italy). *Earth-Sci. Rev.* **2018**, *176*, 241−276.

(14) Fioletov, V. E.; McLinden, C. A.; Krotkov, N.; Li, C.; Joiner, J.; Theys, N.; Carn, S.; Moran, M. D. A global catalogue of large SO$_2$ sources and emissions derived from the Ozone Monitoring Instrument. *Atmos. Chem. Phys.* **2016**, *16*, 11497−11519.

(15) Klimont, Z.; Smith, S. J.; Cofala, J. The last decade of global anthropogenic sulfur dioxide: 2000−2011 emissions. *Environ. Res. Lett.* **2013**, *8*, No. 014003.

(16) Finlayson-Pitts, B. J.; Pitts, J. N., Jr. *Chemistry of the Upper and Lower Atmosphere: Theory, Experiments, and Applications*; Elsevier, 1999.

(17) Spicer, C.; Chapman, E.; Finlayson-Pitts, B.; Plastridge, R.; Hubbe, J.; Fast, J.; Berkowitz, C. Unexpectedly high concentrations of molecular chlorine in coastal air. *Nature* **1998**, *394*, 353−356.

(18) Lewis, J. S. Venus: Atmospheric and lithospheric composition. *Earth Planet. Sci. Lett.* **1970**, *10*, 73−80.

(19) Hoffman, J.; Hodges, R.; Donahue, T.; McElroy, M. Composition of the Venus lower atmosphere from the Pioneer Venus mass spectrometer. *J. Geophys. Res.* **1980**, *85*, 7882−7890.

(20) Yung, Y. L.; DeMore, W. Photochemistry of the stratosphere of Venus: Implications for atmospheric evolution. *Icarus* **1982**, *51*, 199−247.

(21) Polanyi, J. C. Concepts in reaction dynamics. *Acc. Chem. Res.* **1972**, *5*, 161−168.

(22) Agrawalla, B.; Setser, D. Infrared chemiluminescence and laser-induced fluorescence studies of energy disposal by reactions of fluorine and chlorine atoms with hydrogen sulfide, deuterium sulfide, hydrogen selenide, water, water-d2, and methanol. *J. Phys. Chem. A.* **1986**, *90*, 2450−2462.

(23) Braithwaite, M.; Leone, S. R. Laser-initiated chemical reactions: $Cl + H_2S \rightarrow HCl + HS$: Rate constant, product energy distribution, and direct detection of a chain mechanism. *J. Chem. Phys.* **1978**, *69*, 839−845.

(24) Nesbitt, D. J.; Leone, S. R. Laser-initiated chemical chain reactions. *J. Chem. Phys.* **1980**, *72*, 1722−1732.

(25) Clyne, M. A.; Ono, Y. Determination of the rate constant of reaction of ground-state Cl and H atoms with H_2S using resonance fluorescence in a discharge flow. *Chem. Phys. Lett.* **1983**, *94*, 597−602.

(26) Clyne, M. A.; MacRobert, A. J.; Murrells, T. P.; Stief, L. J. Kinetics of the reactions of atomic chlorine with H_2S, HS and OCS. *J. Chem. Soc., Faraday Trans. 2* **1984**, *80*, 877−886.

(27) Nava, D. F.; Brobst, W. D.; Stief, L. J. Temperature study of the rates of the reactions of atomic chlorine with hydrogen sulfide and C_2H_4S. *J. Phys. Chem. B* **1985**, *89*, 4703−4707.

(28) Lu, E. C. C.; Iyer, R. S.; Rowland, F. S. Reaction rates for thermal chlorine atoms with hydrogen sulfide from 232 to 359 K by a radiochemical technique. *J. Phys. Chem. C.* **1986**, *90*, 1988−1990.

(29) Hossenlopp, J. M.; Hershberger, J. F.; Flynn, G. W. Kinetics and product vibrational energy disposal dynamics in the reaction of chlorine atoms with hydrogen sulfide-d2. *J. Phys. Chem. D.* **1990**, *94*, 1346−1351.

(30) Nicovich, J.; Wang, S.; Wine, P. Kinetics of the reactions of atomic chlorine with H_2S, D_2S, CH_3SH, and CD_3SD. *Int. J. Chem. Kinet.* **1995**, *27*, 359−368.

(31) Chen, K.-S.; Cheng, S.-S.; Lee, Y.-P. Reaction dynamics of Cl + H_2S: Rotational and vibrational distribution of HCl probed with time-resolved Fourier-transform spectroscopy. *J. Chem. Phys.* **2003**, *119*, 4229−4236.

(32) Gao, Y.; Alecu, I.; Goumri, A.; Marshall, P. High-temperature kinetics of the reaction between chlorine atoms and hydrogen sulfide. *Chem. Phys. Lett.* **2015**, *624*, 83−86.

(33) Atkinson, R.; Baulch, D.; Cox, R.; Crowley, J.; Hampson, R.; Hynes, R.; Jenkin, M.; Rossi, M.; Troe, J. Evaluated kinetic and photochemical data for atmospheric chemistry: Volume I-gas phase reactions of Ox, HOx, NOx and SOx species. *Atmos. Chem. Phys.* **2004**, *4*, 1461−1738.

(34) Wilson, C.; Hirst, D. M. Ab initio study of the reaction of chlorine atoms with H_2S, CH_3SH, CH_3SCH_3 and CS_2. *J. Chem. Soc., Faraday Trans.* **1997**, *93*, 2831−2837.

(35) Graham, D. C.; Menon, A. S.; Goerigk, L.; Grimme, S.; Radom, L. Optimization and basis-set dependence of a restricted-open-shell form of B2-PLYP double-hybrid density functional theory. *J. Phys. Chem. A* **2009**, *113*, 9861−9873.

(36) Sancho-García, J. C.; Adamo, C. Double-hybrid density functionals: merging wavefunction and density approaches to get the best of both worlds. *Phys. Chem. Chem. Phys.* **2013**, *15*, 14581−14594.

(37) Goerigk, L.; Grimme, S. Double-hybrid density functionals. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2014**, *4*, 576−600.

(38) Mehta, N.; Casanova-Páez, M.; Goerigk, L. Semi-empirical or non-empirical double-hybrid density functionals: which are more robust. *Phys. Chem. Chem. Phys.* **2018**, *20*, 23175−23194.

(39) Penocchio, E.; Piccardo, M.; Barone, V. Semi-Experimental Equilibrium Structures for Building Blocks of Organic and Biological

Molecules: the B2PLYP Route. *J. Chem. Theory Comput.* **2015**, *11*, 4689−4707.

(40) Puzzarini, C.; Bloino, J.; Tasinato, N.; Barone, V. Accuracy and Interpretability: The Devil and the Holy Grail. New Routes across Old Boundaries in Computational Spectroscopy. *Chem. Rev.* **2019**, *119*, 8131−8191.

(41) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics*, 1st ed.; Cambridge University Press, 2009.

(42) Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479−483.

(43) Santra, G.; Sylvetsky, N.; Martin, J. M. Minimally Empirical Double-Hybrid Functionals Trained against the GMTKN55 Database: revDSD-PBEP86-D4, revDOD-PBE-D4, and DOD-SCAN-D4. *J. Phys. Chem. A* **2019**, *123*, 5129−5143.

(44) Caldeweyher, E.; Ehlert, S.; Hansen, A.; Neugebauer, H.; Spicher, S.; Bannwarth, C.; Grimme, S. A generally applicable atomic-charge dependent London dispersion correction. *J. Chem. Phys.* **2019**, *150*, No. 154122.

(45) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, No. 154104.

(46) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456−1465.

(47) Biczysko, M.; Panek, P.; Scalmani, G.; Bloino, J.; Barone, V. Harmonic and Anharmonic Vibrational Frequency Calculations with the Double-Hybrid B2PLYP Method: Analytic Second Derivatives and Benchmark Studies. *J. Chem. Theory Comput.* **2010**, *6*, 2115−2125.

(48) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. Perspectives on Basis Sets Beautiful: Seasonal Plantings of Diffuse Basis Functions. *J. Chem. Theory Comput.* **2011**, *7*, 3027−3034.

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

(50) Dunning, T. H.; Peterson, K. A.; Wilson, A. K. Gaussian basis sets for use in correlated molecular calculations. X. The atoms aluminum through argon revisited. *J. Chem. Phys.* **2001**, *114*, 9244−9253.

(51) Puzzarini, C.; Barone, V. Extending the Molecular Size in Accurate Quantum-Chemical Calculations: the Equilibrium Structure and Spectroscopic Properties of Uracil. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7189−7197.

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

(53) Montgomery, J. A., Jr.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. A complete basis set model chemistry. VI. Use of density functional geometries and frequencies. *J. Chem. Phys.* **1999**, *110*, 2822−2827.

(54) Montgomery, J. A., Jr.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. A. A Complete Basis Set Model Chemistry. VII. Use of the Minimum Population Localization Method. *J. Chem. Phys.* **2000**, *112*, 6532−6542.

(55) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 242−253.

(56) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. et al. *MOLPRO, A Package of Ab Initio Programs*, version 2019.2; TTI GmbH Stuttgart, 2019.

(57) Berning, A.; Schweizer, M.; Werner, H.-J.; Knowles, P. J.; Palmieri, P. Spin-orbit matrix elements for internally contracted multireference configuration interaction wavefunctions. *Mol. Phys.* **2000**, *98*, 1823−1833.

(58) Werner, H.; Knowles, P. J. A second order multiconfiguration SCF procedure with optimum convergence. *J. Chem. Phys.* **1985**, *82*, 5053−5063.

(59) Knowles, P. J.; Werner, H.-J. An efficient second-order MCSCF method for long configuration expansions. *Chem. Phys. Lett.* **1985**, *115*, 259−267.

(60) Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.

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

(62) Werner, H.; Knowles, P. J. An efficient internally contracted multiconfiguration-reference configuration interaction method. *J. Chem. Phys.* **1988**, *89*, 5803−5814.

(63) Knowles, P. J.; Werner, H.-J. An efficient method for the evaluation of coupling coefficients in configuration interaction calculations. *Chem. Phys. Lett.* **1988**, *145*, 514−522.

(64) Shamasundar, K. R.; Knizia, G.; Werner, H.-J. A new internally contracted multi-reference configuration interaction method. *J. Chem. Phys.* **2011**, *135*, No. 054101.

(65) Moore, C. E. *Atomic Energy Levels*, NSRDS-NBS 35; Office of Standard Reference Data, National Bureau of Standards: Washington, DC, 1971.

(66) Bloino, J.; Biczysko, M.; Barone, V. General perturbative approach for spectroscopy, thermodynamics, and kinetics: Methodological background and benchmark studies. *J. Chem. Theory Comput.* **2012**, *8*, 1015−1036.

(67) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16*, revision C.01; Gaussian Inc.: Wallingford, CT, 2016.

(68) Feller, D. The Use of Systematic Sequences of Wave Functions for Estimating the Complete Basis Set, Full Configuration Interaction Limit in Water. *J. Chem. Phys.* **1993**, *98*, 7059.

(69) Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. Basis-Set Convergence of Correlated Calculations on Water. *J. Chem. Phys.* **1997**, *106*, 9639.

(70) Van Mourik, T.; Dunning, T. H., Jr. Gaussian basis sets for use in correlated molecular calculations. VIII. Standard and augmented sextuple zeta correlation consistent basis sets for aluminum through argon. *Int. J. Quantum Chem.* **2000**, *76*, 205−221.

(71) Wilson, A. K.; van Mourik, T.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. VI. Sextuple zeta correlation consistent basis sets for boron through neon. *J. Mol. Struct.: THEOCHEM* **1996**, *388*, 339−349.

(72) Woon, D. E.; Dunning, T. H., Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. V. Core-Valence Basis Sets for Boron through Neon. *J. Chem. Phys.* **1995**, *103*, 4572−4585.

(73) Peterson, K. A.; Dunning, T. H., Jr. Accurate correlation consistent basis sets for molecular core-valence correlation effects: The second row atoms Al-Ar, and the first row atoms B-Ne revisited. *J. Chem. Phys.* **2002**, *117*, 10548−10560.

(74) Noga, J.; Bartlett, R. J. The Full CCSDT Model for Molecular Electronic Structure. *J. Chem. Phys.* **1987**, *86*, 7041−7050.

(75) Scuseria, G. E.; Schaefer, H. F., III A New Implementation of the Full CCSDT Model for Molecular Electronic-Structure. *Chem. Phys. Lett.* **1988**, *152*, 382−386.

(76) Watts, J. D.; Bartlett, R. J. The Coupled-Cluster Single, Double, and Triple Excitation Model for Open-Shell Single Reference Functions. *J. Chem. Phys.* **1990**, *93*, 6104−6105.

(77) Kállay, M.; Surján, P. R. Higher Excitations in Coupled-Cluster Theory. *J. Chem. Phys.* **2001**, *115*, 2945−2954.

(78) Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. HEAT: High accuracy extrapolated ab initio thermochemistry. *J. Chem. Phys.* **2004**, *121*, 11599−11613.

(79) Bomble, Y. J.; Vázquez, J.; Kállay, M.; Michauk, C.; Szalay, P. G.; Császár, A. G.; Gauss, J.; Stanton, J. F. High-accuracy extrapolated ab initio thermochemistry. II. Minor improvements to the protocol and a vital simplification. *J. Chem. Phys.* **2006**, *125*, No. 064108.

(80) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. High-accuracy extrapolated ab initio thermochemistry. III. Additional improvements and overview. *J. Chem. Phys.* **2008**, *128*, No. 114111.

(81) Sellers, H.; Pulay, P. The adiabatic correction to molecular potential surfaces in the SCF approximation. *Chem. Phys. Lett.* **1984**, *103*, 463.

(82) Handy, N. C.; Yamaguchi, Y.; Schaefer, H. F. The diagonal correction to the Born-Oppenheimer approximation: Its effect on the singlet-triplet splitting of CH_2 and other molecular effects. *J. Chem. Phys.* **1986**, *84*, 4481.

(83) Handy, N. C.; Lee, A. M. The adiabatic approximation. *Chem. Phys. Lett.* **1996**, *252*, 425.

(84) Kutzelnigg, W. The adiabatic approximation I. The physical background of the Born-Handy ansatz. *Mol. Phys.* **1997**, *90*, 909.

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

(86) Kutzelnigg, W. In *Relativistic Electronic Structure Theory. Part I. Fundamentals*; Schwerdtfeger, P., Ed.; Elsevier: Amsterdam, 2002.

(87) Bomble, Y. J.; Stanton, J. F.; Kállay, M.; Gauss, J. Coupled-cluster methods including noniterative corrections for quadruple excitations. *J. Chem. Phys.* **2005**, *123*, No. 054101.

(88) Kállay, M.; Gauss, J. Approximate treatment of higher excitations in coupled-cluster theory. *J. Chem. Phys.* **2005**, *123*, No. 214105.

(89) Kállay, M.; Gauss, J. Approximate treatment of higher excitations in coupled-cluster theory. II. Extension to general single-determinant reference functions and improved approaches for the canonical Hartree-Fock case. *J. Chem. Phys.* **2008**, *129*, No. 144101.

(90) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G. CFOUR. A Quantum Chemical Program Package, 2016, with contributions from Auer, A. A., Bartlett, R. J., Benedikt, U., Berger, C., Bernholdt, D. E., Bomble, Y. J., Christiansen, O., Engel, F., Heckert, M., Heun, O., Huber, C., Jagau, T.-C., Jonsson, D., Jusélius, J., Klein, K., Lauderdale, W. J., Lipparini, F., Matthews, D., Metzroth, T., Mück, L. A., O'Neill, D. P., Price, D. R., Prochnow, E., Puzzarini, C., Ruud, K., Schiffmann, F., Schwalbach, W., Stopkowicz, S., Tajti, A., Vázquez, J., Wang, F., Watts, J. D., and the integral packages MOLECULE (Almlöf, J., Taylor, P. R.), PROPS (Taylor, P. R.), ABACUS (Helgaker, T., Jensen, H. J. A., Jørgensen, P., Olsen, J.), and ECP routines by Mitin, A. V., van Wüllen, C. For the current version, see: http://www.cfour. de.

(91) Kállay, M.; Nagy, P. R.; Rolik, Z.; Mester, D.; Samu, G.; Csontos, J.; Csóka, J.; Szabó, B. P.; Gyevi-Nagy, L.; Ladjánszki, I.; Szegedy, L.; Ladóczki, B.; Petrov, K.; Farkas, M.; Mezei, P. D.; Hégely, B. *MRCC, A Quantum Chemical Program Suite*, 2018. For the current version, see: http://www.mrcc.hu.

(92) Puzzarini, C.; Biczysko, M.; Barone, V.; Peña, I.; Cabezas, C.; Alonso, J. L. Accurate molecular structure and spectroscopic properties of nucleobases: a combined computational-microwave investigation of 2-thiouracil as a case study. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16965−16975.

(93) Alessandrini, S.; Barone, V.; Puzzarini, C. Extension of the "Cheap" Composite Approach to Noncovalent Interactions: The jun-ChS Scheme. *J. Chem. Theory Comput.* **2020**, *16*, 988−1006.

(94) Adler, T. B.; Knizia, G.; Werner, H.-J. A simple and efficient CCSD(T)-F12 approximation. *J. Chem. Phys.* **2007**, *127*, No. 221106.

(95) Werner, H.-J.; Knizia, G.; Manby, F. R. Explicitly correlated coupled cluster methods with pair-specific geminals. *Mol. Phys.* **2011**, *109*, 407−417.

(96) Knizia, G.; Adler, T. B.; Werner, H.-J. Simplified CCSD(T)-F12 methods: Theory and benchmarks. *J. Chem. Phys.* **2009**, *130*, No. 054104.

(97) Peterson, K. A.; Adler, T. B.; Werner, H.-J. Systematically convergent basis sets for explicitly correlated wavefunctions: The atoms H, He, B-Ne, and Al-Ar. *J. Chem. Phys.* **2008**, *128*, No. 084102.

(98) Werner, H.-J.; Adler, T. B.; Manby, F. R. General orbital invariant MP2-F12 theory. *J. Chem. Phys.* **2007**, *126*, No. 164102.

(99) Hill, J. G.; Mazumder, S.; Peterson, K. A. Correlation consistent basis sets for molecular core-valence effects with explicitly correlated wave functions: The atoms B-Ne and Al-Ar. *J. Chem. Phys.* **2010**, *132*, No. 054108.

(100) Hill, J. G.; Peterson, K. A.; Knizia, G.; Werner, H.-J. Extrapolating MP2 and CCSD explicitly correlated correlation energies to the complete basis set limit with first and second row correlation consistent basis sets. *J. Chem. Phys.* **2009**, *131*, No. 194105.

(101) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218−1226.

(102) Dyall, K. G. The choice of a zeroth-order Hamiltonian for second-order perturbation theory with a complete active space self-consistent-field reference function. *J. Chem. Phys.* **1995**, *102*, 4909−4918.

(103) Celani, P.; Werner, H.-J. Multireference perturbation theory for large restricted and selected active space reference wave functions. *J. Chem. Phys.* **2000**, *112*, 5546−5557.

(104) Georgievskii, Y.; Klippenstein, S. J. Transition State Theory for Multichannel Addition Reactions: Multifaceted Dividing Surfaces. *J. Phys. Chem. A* **2003**, *107*, 9776−9781.

(105) Georgievskii, Y.; Klippenstein, S. J. Variable reaction coordinate transition state theory: Analytic results and application to the C_2H_3 + H → C_2H_4 reaction. *J. Chem. Phys.* **2003**, *118*, 5442−5455.

(106) Harding, L. B.; Georgievskii, Y.; Klippenstein, S. J. Predictive Theory for Hydrogen Atom-Hydrocarbon Radical Association Kinetics. *J. Phys. Chem. A* **2005**, *109*, 4646−4656.

(107) Jackels, C. F.; Gu, Z.; Truhlar, D. G. Reaction-path potential and vibrational frequencies in terms of curvilinear internal coordinates. *J. Chem. Phys.* **1995**, *102*, 3188−3201.

(108) Liu, Y. P.; Lynch, G. C.; Truong, T. N.; Lu, D. H.; Truhlar, D. G.; Garrett, B. C. Molecular modeling of the kinetic isotope effect for the [1,5]-sigmatropic rearrangement of cis-1,3-pentadiene. *J. Am. Chem. Soc.* **1993**, *115*, 2408−2415.

(109) Eckart, C. The penetration of a potential barrier by electrons. *Phys. Rev.* **1930**, *35*, 1303.

(110) Georgievskii, Y.; Harding, L.; Klippenstein, S. *VaReCoF*, 2016.3.23.

(111) Georgievskii, Y.; Miller, J. A.; Burke, M. P.; Klippenstein, S. J. Reformulation and solution of the master equation for multiple-well chemical reactions. *J. Phys. Chem. A* **2013**, *117*, 12146−12154.

(112) Cavallotti, C.; Pelucchi, M.; Georgievskii, Y.; Klippenstein, S. EStokTP: Electronic Structure to Temperature-and Pressure-Dependent Rate Constants−A Code for Automatically Predicting the Thermal Kinetics of Reactions. *J. Chem. Theory Comput.* **2019**, *15*, 1122−1145.

(113) Cook, R. L.; De Lucia, F. C.; Helminger, P. Molecular force field and structure of hydrogen sulfide: recent microwave results. *J. Mol. Struct.* **1975**, *28*, 237−246.

(114) Klaus, T.; Belov, S.; Winnewisser, G. Precise Measurement of the Pure Rotational Submillimeter-Wave Spectrum of HCl and DCl in Their v= 0, 1 States. *J. Mol. Spectrosc.* **1998**, *187*, 109−117.

(115) Martin-Drumel, M. A.; Eliet, S.; Pirali, O.; Guinet, M.; Hindle, F.; Mouret, G.; Cuisset, A. New investigation on THz spectra of OH and SH radicals ($X^2\Pi_i$). *Chem. Phys. Lett.* **2012**, *550*, 8−14.

(116) Krishnan, R.; Pople, J. A. Approximate fourth-order perturbation theory of the electron correlation energy. *Int. J. Quantum Chem.* **1978**, *14*, 91−100.

(117) Krishnan, R.; Frisch, M.; Pople, J. Contribution of triple substitutions to the electron correlation energy in fourth order perturbation theory. *J. Chem. Phys.* **1980**, *72*, 4244−4245.

(118) Gauss, J.; Cremer, D. Analytical evaluation of energy gradients in quadratic configuration interaction theory. *Chem. Phys. Lett.* **1988**, *150*, 280−286.

(119) Salter, E.; Trucks, G. W.; Bartlett, R. J. Analytic energy derivatives in many-body methods. I. First derivatives. *J. Chem. Phys.* **1989**, *90*, 1752−1766.

(120) Schlegel, H. B. Møller-Plesset perturbation theory with spin projection. *J. Phys. Chem. E.* **1988**, *92*, 3075−3078.

(121) Schlegel, H. B. Potential energy curves using unrestricted Møller-Plesset perturbation theory with spin annihilation. *J. Chem. Phys.* **1986**, *84*, 4530−4534.

(122) Knowles, P. J.; Handy, N. C. Projected unrestricted Møller-Plesset second-order energies. *J. Chem. Phys.* **1988**, *88*, 6991−6998.

(123) Chen, W.; Schlegel, H. B. Evaluation of S^2 for correlated wave functions and spin projection of unrestricted Møller-Plesset perturbation theory. *J. Chem. Phys.* **1994**, *101*, 5957−5968.

(124) Vazart, F.; Calderini, D.; Puzzarini, C.; Skouteris, D.; Barone, V. State-of-the-Art Thermochemical and Kinetic Computations for Astrochemical Complex Organic Molecules: Formamide Formation in Cold Interstellar Clouds as a Case Study. *J. Chem. Theory Comput.* **2016**, *12*, 5385−5397.

(125) Puzzarini, C.; Barone, V. The challenging playground of astrochemistry: an integrated rotational spectroscopy - quantum chemistry strategy. *Phys. Chem. Chem. Phys.* **2020**, *22*, 6507−6523.

(126) Cox, J. D.; Wagman, G. D.; Medvedev, D. A. *CODATA Key Values for Thermodynamics*; Hemisphere Publishing Corporation, 1989; p 1.

(127) Császár, A.; Leininger, M. L.; Burcat, A. Enthalpy of formation of HS. *J. Phys. Chem. A* **2003**, *107*, 2061−2065.

(128) Miller, J. A.; Klippenstein, S. J. Master equation methods in gas phase chemical kinetics. *J. Phys. Chem. A* **2006**, *110*, 10528−10544.

(129) Tardy, D. C.; Rabinovitch, B. Collisional Energy Transfer. Thermal Unimolecular Systems in the Low-Pressure Region. *J. Chem. Phys.* **1966**, *45*, 3720−3730.

Chapter 5

Methanimine as prebiotic precursor: the case of 2-propyn-1-imine

The detection of nitrogen-bearing complex organic molecules in the interstellar space is a topic of remarkable interest because most of them represent key intermediates toward the main building-blocks of biomolecules, namely aminoacids and nucleobases. Within this class, several imines have been detected in the last years and, in particular the Z-isomer of propargylamine (2-propyn-1-imine) has been identified in 2020 in the quiescent G+0693-0.027 molecular cloud. The formation of imines has been mainly ascribed to gas-phase isomerizations of simple nitriles or to partial hydrogenations on dust grain surfaces. However, addition/elimination by reactive radicals on methanimine can lead to more complex imines through reaction channels possibly open also in the harsh conditions of the interstellar medium. In this contribution we report state-of-the-art quantum chemical and kinetic computations providing strong evidence that addition of ethynyl radical to methanimine (both of which have been detected in the same region where propargylamine was identified) can lead to propargylamine through a mechanism involving only submerged transition states and with nearly constant low-pressure (10^{-12} bar) reaction rate (around 10^{-10} cm^3 molecule^{-1} s^{-1}) in the 20-30 K temperature range. Together the specific interest of propargylamine, the proposed reaction mechanism represents a quite general route for obtaining complex imines by attack of reative radicals on the methanimine precursor. In this Chapter, objective O4 is pursued.

Methanimine as a Key Precursor of Imines in the Interstellar Medium: The Case of Propargylimine

Jacopo Lupi[1] , Cristina Puzzarini[2] , and Vincenzo Barone[1]

[1] Scuola Normale Superiore, Piazza dei Cavalieri 7, Pisa, I-56126, Italy; vincenzo.barone@sns.it

[2] Department of Chemistry "Giacomo Ciamician," University of Bologna, Via F. Selmi 2, Bologna, I-40126, Italy; cristina.puzzarini@unibo.it

Received 2020 August 11; revised 2020 October 5; accepted 2020 October 17; published 2020 November 9

Abstract

A gas-phase formation route is proposed for the recently detected propargylimine molecule. In analogy to other imines, such as cyanomethanimine, the addition of a reactive radical (C_2H in the present case) to methanimine (CH_2NH) leads to reaction channels open also in the harsh conditions of the interstellar medium. Three possible isomers can be formed in the $CH_2NH + C_2H$ reaction: Z- and E-propargylimine (Z-,E-PGIM) as well as N-ethynyl-methanimine (N-EMIM). For both PGIM species, the computed global rate coefficient is nearly constant in the 20–300 K temperature range, and of the order of 2–3×10^{-10} cm^3 molecule^{-1} s^{-1}, while that for N-EMIM is about two orders of magnitude smaller. Assuming equal destruction rates for the two isomers, these results imply an abundance ratio for PGIM of [Z]/[E] ~ 1.5, which is only slightly underestimated with respect to the observational datum.

Unified Astronomy Thesaurus concepts: Interstellar molecules (849); Interstellar medium (847); Interstellar abundances (832); Astrochemistry (75)

1. Introduction

Currently, the number of molecules detected in the interstellar medium (ISM) thanks to their rotational signatures far exceeds 200 (McGuire 2018). Among them, more than 70 species belong to the class of the so-called interstellar complex organic molecules (iCOMs), namely molecules containing at least one carbon atom and a total of more than six atoms (Herbst & van Dishoeck 2009). Nitrogen-bearing iCOMs are particularly interesting because of their prebiotic character; indeed, they represent key intermediates toward the main building blocks of biomolecules, like amino acids and nucleobases. Within this class of iCOMs, six members of the imine family have been detected so far in the ISM, namely methanimine (CH_2NH, Godfrey et al. 1973; Dickens et al. 1997), ethanimine (CH_3CHNH, Loomis et al. 2013), keteni-mine (CH_2CNH, Lovas et al. 2006), 3-imino-1,2-propadieny-lidene (CCCNH, Kawaguchi et al. 1992), C-cyanomethanimine (NCCHNH, Zaleski et al. 2013; Rivilla et al. 2018), and—very recently—Z-propargylimine (2-propyn-1-imine, HC≡C-CH=NH, Bizzocchi et al. 2020).

The main hypotheses on their formation mechanisms in astrophysical environments involve either tautomerization of simple nitriles (Lovas et al. 2006) or their partial hydrogenation on dust-grain surface (Theule et al. 2011; Krim et al. 2019). However, for C-cyanomethanimine, a gas-phase formation route has been recently proposed that involves addition of the cyano radical (CN) to methanimine (Vazart et al. 2015). It is thus quite natural to hypothesize that methanimine can play a role in the formation of other imines upon addition/elimination of reactive radicals already detected in the ISM, like CH_3, C_2H or OH. Indeed, the reaction of the hydroxyl radical with methanimine is proven to effectively lead to the formation of formamide in the gas phase (Vazart et al. 2016; Codella et al. 2017).

The focus of the present Letter is the possible formation pathway of propargylimine (PGIM), whose Z-isomer has been very recently identified in the quiescent G+0.693-0.027

molecular cloud with an estimated column density of $0.24 \pm 0.02 \times 10^{14}$ cm^{-2} (Bizzocchi et al. 2020). In the same study, an upper limit of 0.9×10^{-10} was retrieved for the fractional abundance (w.r.t. H_2) of the higher-energy E isomer (which means a column density <0.13 10^{14} cm^{-2}), which instead was not observed. After the spectroscopic characterization of this imine and its astronomical detection, Bizzocchi et al. (2020) put forward some speculations about feasible formation routes based on the relative abundances of a number of possible precursors in the G+0.693-0.027 molecular cloud. However, in spite of the detection of CH_2NH (Zeng et al. 2018), this has not been taken into consideration, notwith-standing the authors reported, among the others, a large fractional abundance for the ethynyl radical (C_2H, $^2\Sigma^+$), i.e., 3.91×10^{-8}.

Based on these premises, we decided to perform a state-of-the-art quantum-chemical (QC) characterization of the stationary points on the doublet reactive $C_2H + CH_2NH$ potential energy surface (PES) followed by kinetic computations in the framework of a master equation model rooted in generalized transition state estimates of the elementary reaction rates. From a theoretical point of view, the reactions between the ethynyl radical and several substrates have been recently investigated by state-of-the-art QC approaches (Bowman et al. 2020), but addition/elimination reactions with unsaturated substrates have not yet been explored.

2. Computational Methodology

The starting point for the study of the formation pathway of PGIM is the identification of the potential reactants and the analysis of the corresponding reactive PES, which implies the accurate characterization of all stationary points from both a structural and energetic point of view. This first step then requires to be completed by kinetic calculations. In the derivation of a feasible reaction mechanism, one has to take into account the extreme conditions of the ISM: low temperatures (10–100 K) and low number density (10–10^7 cm^{-3}). By translating density in

terms of pressure, a number density of 10^4 cm^{-3} corresponds to a pressure of 3.8×10^{-10} Pa ($\sim 3.8 \times 10^{-15}$ atm).

2.1. Reactive Potential Energy Surface

We have followed the general computational strategy validated in several recent studies (Baiano et al. 2020; Lupi et al. 2020; Puzzarini et al. 2020; Salta et al. 2020; Tonolo et al. 2020), which involves the following steps.

1. The stationary points have been located and characterized using the double-hybrid B2PLYP functional (Grimme 2006), combined with D3(BJ) corrections (to incorporate dispersion effects; Grimme et al. 2010, 2011) and in conjunction with the jun-cc-pVTZ "seasonal" basis set (Papajak et al. 2011).
2. Single-point energy calculations, at the B2PLYP-D3 (BJ)/jun-cc-pVTZ geometries, have been performed by means of the so-called "cheap" composite scheme (ChS; Puzzarini & Barone 2011; Puzzarini et al. 2014), which starts from the coupled-cluster theory including single and double excitations augmented by a perturbative estimate of triples (CCSD(T); Raghavachari et al. 1989) in conjunction with a triple-zeta basis set (cc-pVTZ; Dunning 1989) and within the frozen-core (fc) approximation. To improve this level of theory, the ChS model considers the extrapolation to the complete basis set (CBS) limit and the effect of core-valence (CV) correlation using Møller-Plesset theory to second order (MP2; Møller & Plesset 1934). Concerning the former contribution, the fc-MP2 energy is extrapolated to the CBS limit using the n^{-3} expression (Helgaker et al. 1997) in conjunction with the cc-pVTZ and cc-pVQZ basis sets. The CV correlation correction is, instead, the difference between the MP2 energy evaluated correlating all electrons and that computed within the fc approximation, both in conjunction with the cc-pCVTZ basis set (Woon & Dunning 1995).
3. ChS energies have been combined with anharmonic zero-point energy (ZPE) corrections evaluated at the B2PLYP-D3(BJ)/jun-cc-pVTZ level within hybrid degeneracy-corrected second-order vibrational perturbation theory (HDCPT2; Bloino et al. 2012; Puzzarini et al. 2019).

All calculations have been performed with the Gaussian software (Frisch et al. 2016).

2.2. Kinetic Models

Global rate constants have been calculated by using a master equation (ME) approach based on ab initio transition state theory (AITSTME), thereby employing the MESS software as master equation solver (Georgievskii et al. 2013). For elementary reactions involving a transition state, rate constants have been computed using transition state theory (TST), while for barrierless elementary reactions, they have been evaluated by means of phase space theory (PST; Pechukas & Light 1965; Chesnavich 1986). The basic assumption of PST is that the interaction between two reacting fragments is isotropic (following a $\frac{C_6}{R^6}$ power law) and does not affect the internal fragment motions (Fernández-Ramos et al. 2006). This approximation is generally valid for low-temperature phenomena, as those occurring in the ISM. To be more precise, in order to obtain the C_6 parameter for the PST calculation, we

performed a scan of the HCC-CH$_2$NH and HCC-NHCH$_2$ distances for the C- and N-end attack, respectively. Then, the corresponding minimum energy paths have been fitted to a $f(x) = f_0 - \frac{C_6}{x^6}$ function, thus obtaining a C_6 value of 131.96 $a_0{}^6 E_h$ for the former attack and of 180.59 $a_0{}^6 E_h$ for the latter. In all cases, tunneling has been accounted for using the Eckart model (Eckart 1930).

The rate constants of the overall reactions leading to the C$_3$H$_3$N imine isomers (namely, the E-,Z-PGIM species and N-ethynyl-methanimine, N-EMIM) have been evaluated in the 20–500 K temperature range. To model their temperature dependence, the rate constants at different temperatures have been fitted to a three-parameter modified Arrhenius equation, namely the Arrhenius–Kooij expression (Kooij 1893; Laidler 1996):

$$k(T) = A\left(\frac{T}{300}\right)^n \exp\left(-\frac{E}{RT}\right) \qquad (1)$$

where A, n, and E are the fitting parameters, R being the universal gas constant.

3. Results and Discussion

3.1. Reactivity and Energetics

A recent re-investigation of the reaction channel starting from the attack of the cyano radical to the C-end of methanimine (Puzzarini & Barone 2020) has shown that, for all stationary points, the ChS model has a maximum absolute deviation of 3 kJ mol^{-1} and an average absolute deviation of 1.1 kJ mol^{-1} with respect to a reference composite scheme, which is able to reach sub-kJ accuracy energetics. These errors are much smaller than those issuing from widely employed composite schemes (e.g., CBS-QB3, Montgomery et al. 2000, or G4, Curtiss et al. 2007) and well sufficient for obtaining quantitative estimates of reaction rates and branching ratios (Baiano et al. 2020; Lupi et al. 2020; Puzzarini et al. 2020; Salta et al. 2020; Tonolo et al. 2020). On these grounds, we have performed a full characterization of the doublet PES for the addition-elimination reactions of both CN and CCH radicals to methanimine at the ChS level.

As far as the reaction mechanism is concerned, hydrogen abstraction could be competitive with addition/elimination (Bowman et al. 2020), but test computations showed that the former reaction channel is at least one order of magnitude slower than the latter one. As a consequence, only the addition/elimination reaction channel is analyzed in detail in the following. The reaction mechanism proposed in the present Letter for the formation of N-EMIM and the PGIM isomers is sketched in Figure 1 and the relative energies of all the stationary points, with respect to reactants, are collected in Table 1 together with the corresponding results for the CH$_2$NH+CN reaction. There are three possible initial adducts, corresponding to the attack of the ethynyl radical to the C or N ends and to the π-system of the imine double bond. However, the cyclic adduct resulting from the third option (CYCLO-1) is significantly less stable and easily interconverts to one of the corresponding open-chain minima (1Z or 1N). For both the CN and CCH radicals, the intermediate obtained upon attack to the N moiety is slightly more stable, but the reaction channels originating from it are ruled by transition states significantly less stable (albeit always submerged) than those ruling the corresponding channels issuing from 1Z or 1E. Noted is that
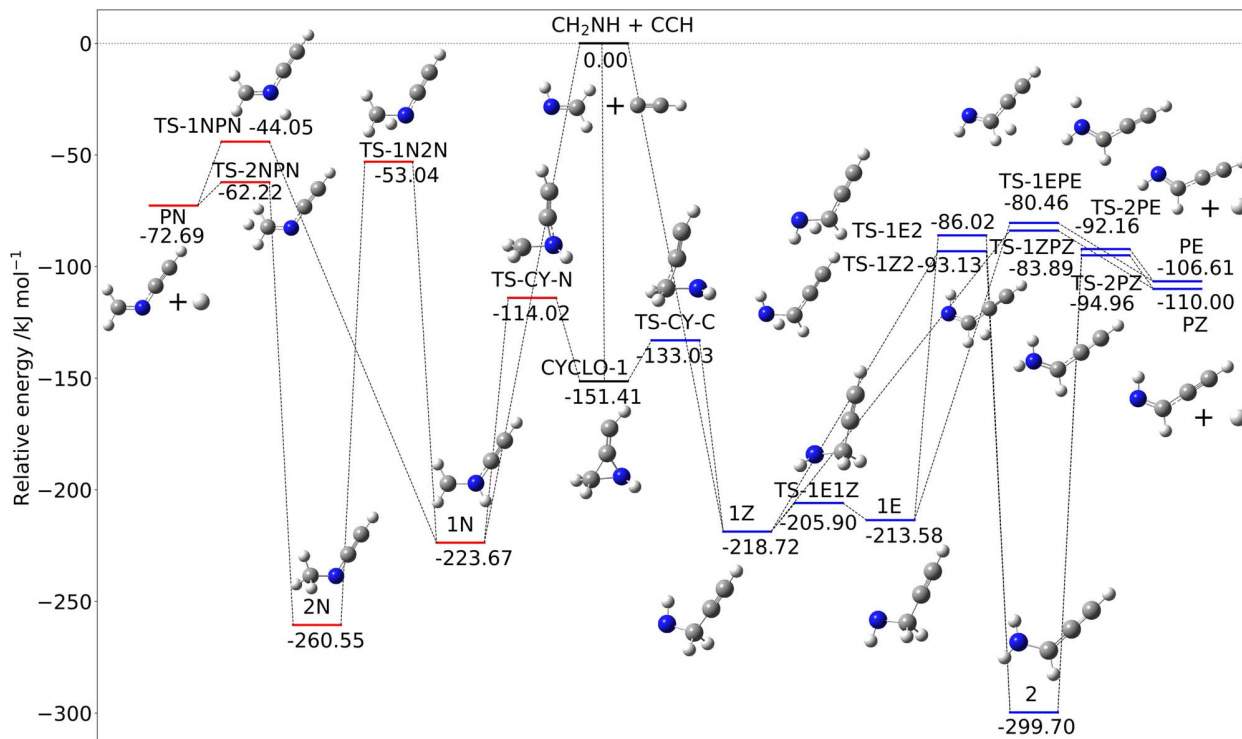
Figure 1. Formation route of N-EMIM and the PGIM isomers: ChS energies augmented by anharmonic B2PLYP-D3(BJ) ZPE corrections.

Table 1
ChS Relative Electronic Energies (ΔE_{el}) and Corresponding Standard Enthalpies at 0 K (ΔH_0°) for the Stationary Points of the CH$_2$NH+X Reaction

	X = C$_2$H		X = CN	
	ΔE_{el}	ΔH_0°	ΔE_{el}	ΔH_0°
CH$_2$NH + X	0.00	0.00	0.00	0.00
1Z	−229.36	−218.72	−203.59	−198.67
TS-1E1Z	−217.65	−205.90	−192.39	−183.24
1E	−225.75	−213.58	−201.48	−191.95
TS-1Z2	−96.36	−93.13	−63.55	−62.81
TS-1E2	−88.80	−86.02	−57.47	−57.18
2	−314.60	−299.70	−284.59	−271.29
TS-2PZ	−88.09	−94.96	−48.37	−58.27
TS-2PE	−84.99	−92.16	−46.22	−56.40
TS-1ZPZ	−77.81	−83.89	−40.62	−49.89
TS-1EPE	−73.86	−80.46	−37.97	−47.70
Z-IM + H (PZ)	−99.21	−110.00	−60.89	−74.73
E-IM + H (PE)	−95.40	−106.61	−58.55	−72.78
CYCLO-1	−167.50	−151.41	−108.93	−96.69
TS-CY-C	−144.35	−133.03	−97.62	−88.32
TS-CY-N	−122.80	−114.02	−88.77	−80.94
1N	−233.04	−223.67	−208.44	−201.33
TS-1N2N	−53.16	−53.04	−25.66	−27.16
2N	−272.23	−260.55	−223.19	−211.86
TS-2NPN	−52.98	−62.22	−22.76	−33.56
TS-1NPN	−35.54	−44.05	−3.85	−14.68
N-IM + H (PN)	−59.22	−72.69	−30.64	−45.80

**Note.** Values in kJ mol^{-1}.

the PES for the CH$_2$NH+CN reaction is, in any detail, analogous to that of the C$_2$H radical.

Starting from the very stable 1Z (or 1E) pre-reactive complex, one might observe a loss of the hydrogen radical,

leading directly to the Z (or E) isomer of PGIM. This step has an exit barrier of about ∼135 (or ∼133) kJ mol^{-1}. On the other hand, considering the presence of the stabilizing C$_2$H moiety on the carbon atom, hydrogen migration might be observed in order to localize the unpaired electron on this atom. This migration occurs through the submerged transition state TS-1Z2 (TS-1E2 for the E-PGIM), which lies 125.6 kJ mol^{-1} above 1Z (127.6 kJ mol^{-1} above 1E for the E-route), thus forming the most stable intermediate of the whole PES, namely 2, which is nearly 300 kJ mol^{-1} below the reactants. Next, loss of hydrogen leads again to the Z (or E) form of PGIM through the submerged transition state TS-2PZ (TS-2PE), lying about 95 kJ mol^{-1} (92 kJ mol^{-1} for the E species) below the reactants (exit barrier of about 205 and 208 kJ mol^{-1}, respectively). The comparison with the analogous reaction paths for the gas-phase production of C-cyanomethanimine (Puzzarini & Barone 2020) shows that the formation of PGIM is characterized by greater exothermicity (−108 versus −60 kJ mol^{-1} for the average of Z and E isomers) and lower exit barriers (126 versus 140 kJ mol^{-1} for the average of TS-1Z2 and TS-1E2 and 206 versus 238 kJ mol^{-1} for the average of TS-2PZ and TS-2PE). Furthermore, the stability of the pre-reactive complex 1Z or 1E (ruling the barrierless entrance channel) and that of the intermediate 2 (involved in the two-step mechanism) are greater in the case of the addition of C$_2$H than for CN (−218.7 versus −203.6 kJ mol^{-1} for the average of 1Z,1E and −299.7 versus −284.6 kJ mol^{-1} for 2).

Moving to the attack to the N-end of methanimine, from the inspection of Figure 1, it is evident that the two possible paths originating from the 1N pre-reactive complex are similar to those described above for the C-end attack, as already noted for the CH$_2$NH+CN reaction (Vazart et al. 2015). With the only exception of 1N, which lies lower in energy than 1Z and 1E, all intermediates and transition states of these paths are less stable
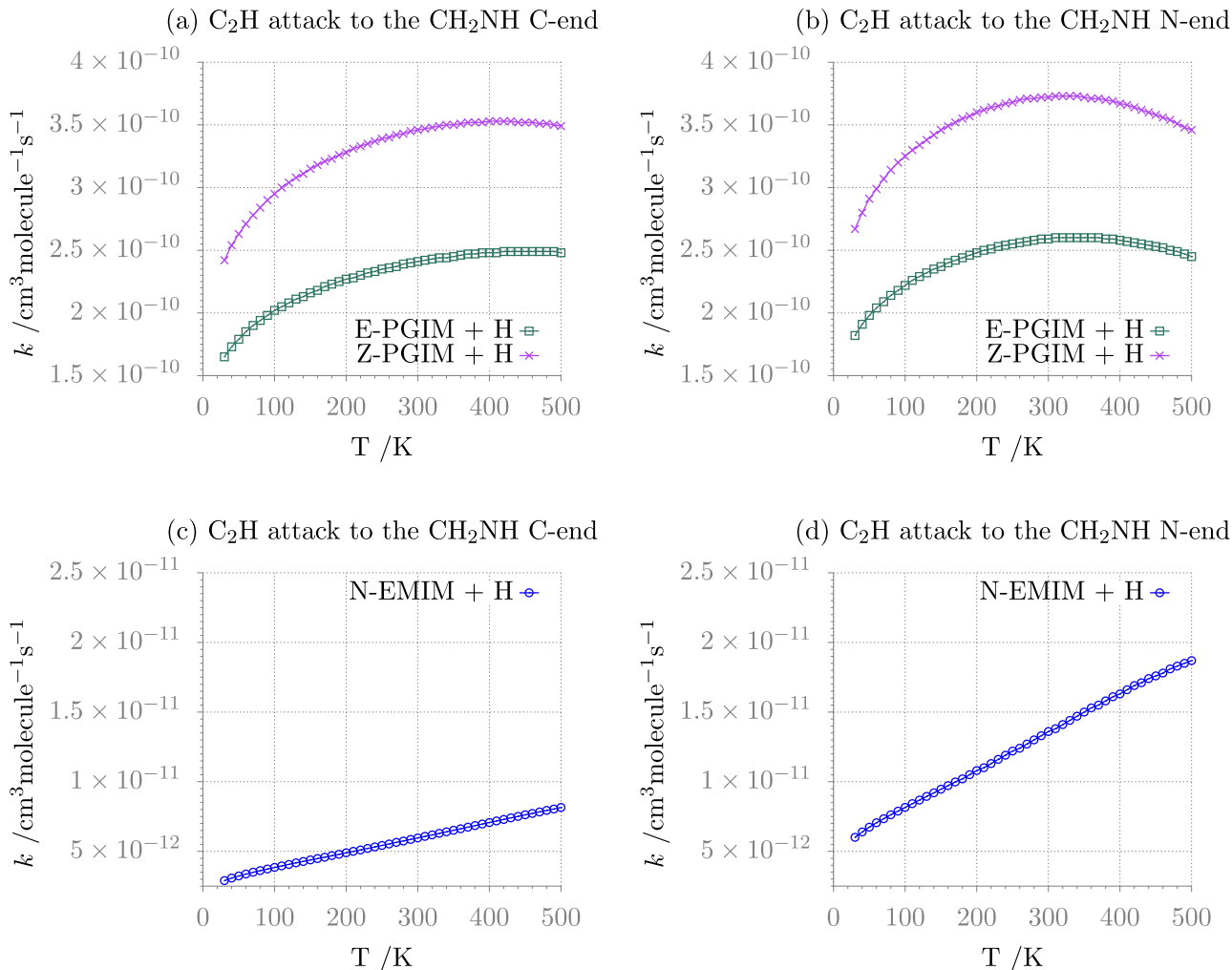
Figure 2. Temperature-dependence plots of the CH_2NH+C_2H reaction rate constants.

with respect to the C-end counterparts. The product itself, i.e., N-EMIM + H (PN), lies at higher energy: $-72.7 \, kJ \, mol^{-1}$, to be compared with $-106.6 \, kJ \, mol^{-1}$ for E-PGIM + H (PE) and $-110.0 \, kJ \, mol^{-1}$ for Z-PGIM + H (PZ).

Studies for reactions of radical species with molecules containing a double bond have shown that the reactivity depends on the type of system. For the C = C bond, addition/ elimination is barrierless and strongly favored over hydrogen elimination (e.g., Bouwman et al. 2012), whereas for C = O bonds, only H elimination is barrierless, whereas both the C- and O- addition/eliminations involve small barriers (e.g., Dong et al. 2005). Preliminary computations for the addition of other radicals (e.g., CP, OH, and CH_3) to methanimine show that the mechanism described in the previous paragraphs for the reaction with C_2H or CN represents a quite general route to the formation of complex imines, although in a few cases (e.g., CH_3) some transition states are not submerged with respect to reactants.

3.2. Rate Constants

To definitely confirm the effectiveness of the mechanism proposed, kinetic computations are required. The product specific rate constants as a function of temperature are shown in Figure 2 for the reaction of methanimine with C_2H and in

Figure 3 for the reaction with CN, whereas the parameters of the Arrhenius–Kooij fits are given in Table 2. These have been obtained by fitting the global rate constants computed in the 20–500 K range. In more detail, for each figure, four panels are provided: those on the left refer to the C-end attack (panels (a) and (c)), while those on the right to the N-end attack (panels (b) and (d)). In both figures, the top panels show the temperature profiles of rate constants for the formation of the "C-isomers" (namely, Z-/E-PGIM and Z-/E-C-cyanomethanimine, CMIM), while the bottom panels refer to the formation of the "N-isomers" (namely, N-EMIM and N-cyanomethanimine, N-CMIM).

Focusing on the C-end reaction paths, the prevalence of the Z-product is related to the slightly lower energy of the corresponding transition states compared to those leading to the E isomer. Back-dissociation into reactants is negligible in the whole temperature range considered, whereas the overall rate constant for the PGIM formation raises by increasing the temperature, also showing progressive deviations from the Arrhenius behavior. The overall rate constant, which is of the order of 2–3 $\times$ $10^{-10} \, cm^3 \, molecule^{-1} \, s^{-1}$, is mainly ruled by the one-step mechanism leading to products from the 1Z/1E pre-reactive complex through the TS-1ZPZ/TS-1EPE transition state. However, this is always true for Z-PGIM, while for the E
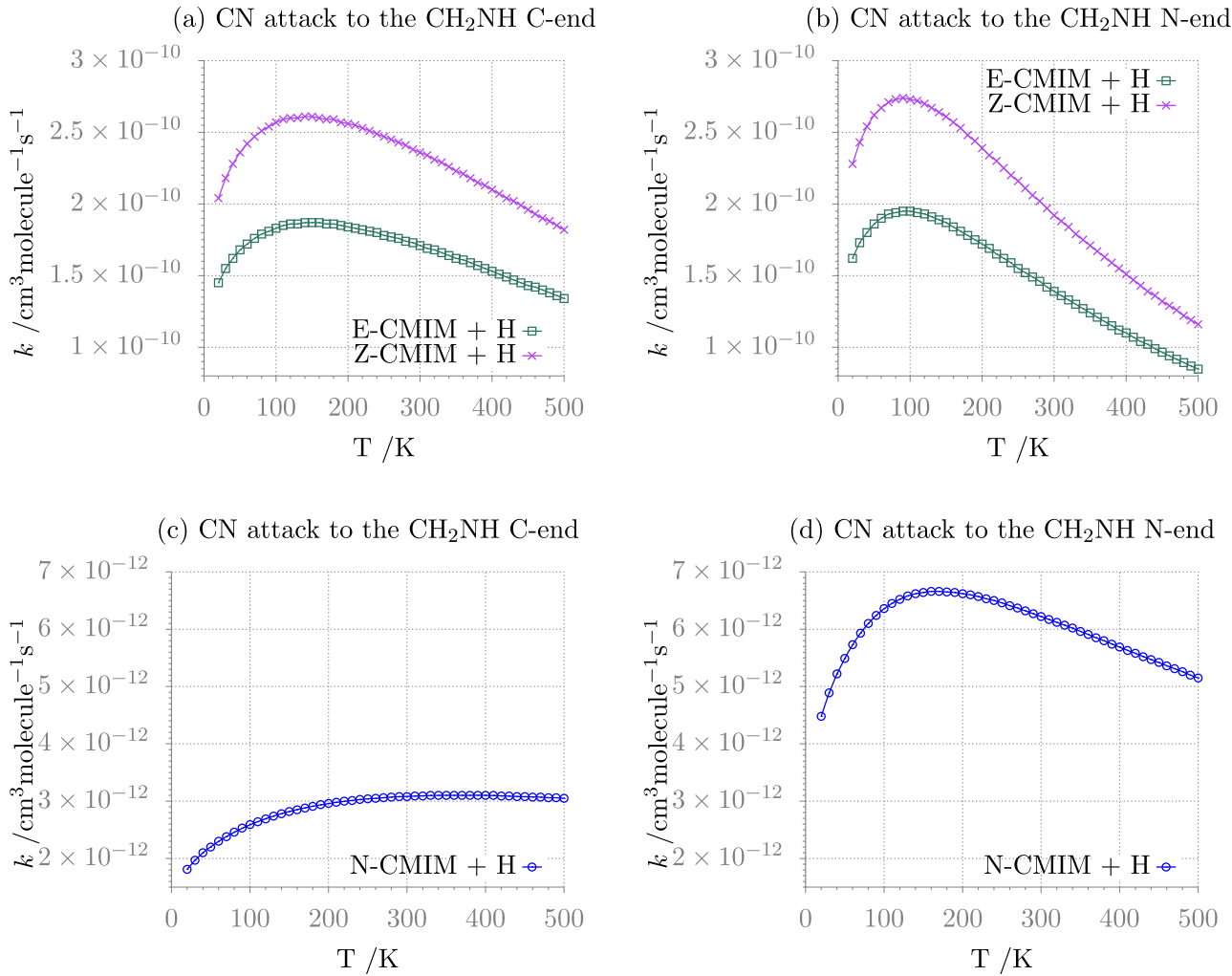
Figure 3. Temperature-dependence plots of the CH$_2$NH+CN reaction rate constants.

isomer the two-step mechanism seems to be the rate determining one above $\sim$350 K. The derived branching ratio is of the order of 1.5, smaller than the observational result ($\geqslant$1.9), as already noted in the case of C-cyanomethanimine. However, as in the latter case, the isomer abundance ratio obtained from astronomical observations is affected by a large uncertainty (the fractional abundance of E being indirectly derived), and it is, instead, close to the value obtained from a thermodynamic estimate based on the relative stability of the E and Z isomers. The considerations above on the computed isomers ratio assume a similar destruction rate for both E and Z species. In the case of C-cyanomethanimine, it has been suggested that the strong difference between the dipole moments of the E and Z forms (4.2 D and 1.4 D, respectively, from Vazart et al. 2015) leads to significantly different destruction rates (see Shingledecker et al. 2020). More specifically, Shingledecker et al. (2020) proposed a general rule-of-thumb for estimating the abundances of isomers based on their dipole moments, which has been denoted as "relative dipole principle." According to this, for propargylimine, whose isomers have very similar dipole moments ($\sim$2 D, see Bizzocchi et al. 2020), the assumption that they have similar destruction rates seems to be reasonable. At the same time, different reaction rates with H radicals cannot be invoked since, for both cyanomethanimine and propargylimine, the

corresponding reactions are ruled by non-submerged transition states. Further investigation of alternative mechanisms would be surely warranted, but it is out of the scope of the present Letter. As already noted for thermochemistry, the general kinetic features for the C$_2$H and CN additions to methanimine are very similar, thus giving further support to the plausibility and generality of the proposed mechanism. At 100 K, for PGIM, the overall rate constants for Z and E species (in cm^3 molecule^{-1} s^{-1}) are 3.25×10^{-10} and 2.22×10^{-10}, respectively, to be compared to 2.73×10^{-10} and 1.95×10^{-10} for the two corresponding isomers of C-cyanomethanimine.

As far as the formation of the N-species is concerned, it is interesting to note that this process would be a little bit favored over formation of the C-species if the attacks to the two ends of the imino group would be independent (as actually is in the case of the CN addition to the CH$_3$ or NH$_2$ moiety of methylamine; see Puzzarini et al. 2020, with a rate constant of 4–5×10^{-10} cm^3 molecule^{-1} s^{-1}). However, Figure 1 shows that the two channels are connected by a low-lying cyclic intermediate. Under these circumstances (also valid for the attack of the CN radical), the formation of C-products becomes faster by two orders of magnitude with respect to formation of the N-product, with the rate constant of the latter process slightly increasing with the temperature. To provide a graphical
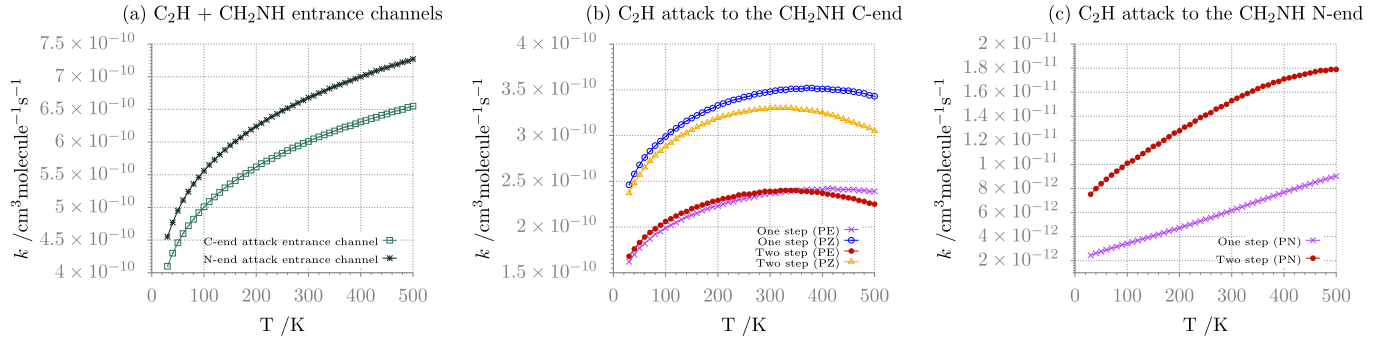
Figure 4. Temperature dependence of the rate constants for the elementary steps of the overall CH_2NH+C_2H reaction, namely barrierless entrance (panel (a)), and one- or two-step paths leading to Z-/E-PGIM (panel (b)) and N-EMIM (panel (c)).

Table 2
The Arrhenius–Kooij Parameters for the CH_2NH+X Reaction

	C-end Attack			N-end Attack		
X = C_2H	E	Z	N	E	Z	N
A/cm^3 molecule^{-1} s^{-1}	2.43×10^{-10}	3.51×10^{-10}	5.54×10^{-12}	2.66×10^{-10}	3.83×10^{-10}	1.26×10^{-11}
n	7.58×10^{-2}	3.86×10^{-2}	6.33×10^{-1}	-6.10×10^{-2}	-9.22×10^{-2}	6.59×10^{-1}
E/kJ mol^{-1}	6.74×10^{-2}	8.72×10^{-2}	-2.32×10^{-1}	1.62×10^{-1}	1.77×10^{-1}	-2.15×10^{-1}
rms[a]	4.37×10^{-12}	7.12×10^{-12}	1.15×10^{-13}	1.02×10^{-11}	1.54×10^{-11}	1.97×10^{-13}
X = CN	E	Z	N	E	Z	N
A/cm^3 molecule^{-1} s^{-1}	1.75×10^{-10}	2.42×10^{-10}	3.12×10^{-12}	1.46×10^{-10}	2.03×10^{-10}	6.51×10^{-12}
n	-3.20×10^{-1}	-3.40×10^{-1}	2.68×10^{-2}	-6.56×10^{-1}	-6.68×10^{-1}	-2.72×10^{-1}
E/ kJ mol^{-1}	2.17×10^{-1}	2.24×10^{-1}	1.03×10^{-1}	3.37×10^{-1}	3.39×10^{-1}	2.33×10^{-1}
rms[a]	1.09×10^{-11}	1.55×10^{-11}	8.39×10^{-14}	1.47×10^{-11}	2.07×10^{-11}	3.29×10^{-13}

Note.
[a] rms stands for root-mean-square deviation of the fit.

explanation of the behavior of the global constant with temperature, the contributions of some specific reaction channels are shown in Figure 4. These are the two barrierless (C- and N-end) entrance channels, the one- and two-step processes leading to Z-/E-PGIM for the C-end attack and the corresponding channel leading to N-EMIM for the attack to the N end of methanimine. It is noted that, even if the entrance channel flux for the N-end attack is faster than the C-end attack one, the subsequent high barriers of the N-EMIM formation path slow down the flux, thus resulting in the preferential formation of the E,Z-PGIM, which presents lower lying barriers. In this picture, an important role is played by the TS-CY-N transition state linking 1N to the cyclic pre-reactive complex, CYCLO-1. In fact, this interconversion is the elementary step characterized by the lowest barrier for the N-end side of the overall CH_2NH+C_2H reaction. Similar arguments also apply to the reaction involving CN.

It is noteworthy that the behavior discussed above for both types of radical attack to methanimine is specific of the low pressure limit (see computational details). In fact, moving to a pressure of 1 atm (of limited astrophysical interest, but of potential relevance in planetary atmospheres), the N-EMIM formation remains unfavorable with respect to E,Z-PGIM but all formation rate constants become slower. This trend is due to the stabilization of the entrance channel wells (namely 1N and 1Z) by collisions (that occurs at pressure values as high as 1 atm), thus leading to an increase of the effective reaction

barriers with the consequent decrease of the overall rate constant, which shows a monotonic increase with temperature.

A curved Arrhenius plot is obtained when the activation energy depends on the temperature and this behavior is captured by the Arrhenius–Kooij formula (see Equation (1)) when this dependence is linear. The rms deviations reported in Table 2 demonstrate that the data for the C_3H_3N imine isomers are indeed well fitted by the Arrhenius–Kooij expression. Within this model, E represents the activation energy at 0 K and the activation energy at a generic temperature T is given by $E + n\left(\frac{RT}{300}\right)$. In the present case, the activation energy is always positive, with the exception of N-EMIM, as a result of both the capture rate constant and the subsequent energy barriers for the unimolecular steps. The n parameter (the first derivative of the activation energy with respect to temperature) is always positive for the C-end attack, while it is negative for the PGIM isomers when the N-end attack takes place. Finally, the values of the pre-exponential factor A are typical for this kind of reactions and rule the branching ratio between the Z and E PGIM isomers. Indeed, the ratio of the A factors is 1.44 and the branching ratio ranges between 1.43 and 1.47 in the whole temperature range (20–500 K).

4. Concluding Remarks

In this Letter, we have proposed a gas-phase formation route for the recently detected Z-PGIM molecule. In analogy to the addition of the CN radical to methanimine leading to cyanomethanimine,

Lupi, Puzzarini, & Barone

addition of the isoelectronic ethynyl radical easily leads to PGIM through a similar reaction mechanism, which involves the formation of a stable pre-reactive complex and its successive evolution by means of submerged transition states. Since the level of the QC and kinetic computations carried out gives strong supports to the quantitative accuracy of our results, search for PGIM isomers in the other regions of the ISM where methanimine and the ethynyl radical have been both detected could be attempted to further validate the proposed reaction mechanism.

In a more general perspective, the results of our state-of-the-art computations provide convincing evidences about the feasibility of a general addition/elimination mechanism for the formation of complex imines. This starts from methanimine as a precursor and involves reactive radicals abundantly present in the interstellar space.

ORCID iDs

Jacopo Lupi ⓘ https://orcid.org/0000-0001-6522-9947
Cristina Puzzarini ⓘ https://orcid.org/0000-0002-2395-8532
Vincenzo Barone ⓘ https://orcid.org/0000-0001-6420-4107

References

Baiano, C., Lupi, J., Tasinato, N., Puzzarini, C., & Barone, V. 2020, Molecules, 25, 2873
Bizzocchi, L., Prudenzano, D., Rivilla, V. M., et al. 2020, A&A, 640, A98
Bloino, J., Biczysko, M., & Barone, V. 2012, J. Chem. Theory Comp., 8, 1015
Bouwman, J., Goulay, F., Leone, S. R., & Wilson, K. R. 2012, JPCA, 116, 3907
Bowman, M. C., Burke, A. D., Turney, J. M., & Schaefer, H. F., III 2020, MolPh, 118, e1769214
Chesnavich, W. J. 1986, JChPh, 84, 2615
Codella, C., Ceccarelli, C., Caselli, P., et al. 2017, A&A, 605, L3
Curtiss, L. A., Redfern, P. C., & Raghavachari, K. 2007, JChPh, 126, 084108
Dickens, J. E., Irvine, W. M., DeVries, C. H., & Ohishi, M. 1997, ApJ, 479, 307
Dong, H., Ding, Y.-h., & Sun, C.-c. 2005, JChPh, 122, 204321
Dunning, T. H., Jr. 1989, JChPh, 90, 1007
Eckart, C. 1930, PhRv, 35, 1303
Fernández-Ramos, A., Miller, J. A., Klippenstein, S. J., & Truhlar, D. G. 2006, ChRv, 106, 4518
Frisch, M. J., Trucks, G. W., Schlegel, H. B., et al. 2016, Gaussian 16 Revision C.01
Georgievskii, Y., Miller, J. A., Burke, M. P., & Klippenstein, S. J. 2013, JPCA, 117, 12146
Godfrey, P. D., Brown, R. D., Robinson, B. J., & Sinclair, M. W. 1973, ApL, 13, 119
Grimme, S. 2006, JChPh, 124, 034108
Grimme, S., Antony, J., Ehrlich, S., & Krieg, H. 2010, JChPh, 132, 154104
Grimme, S., Ehrlich, S., & Goerigk, L. 2011, JCoCh, 32, 1456
Helgaker, T., Klopper, W., Koch, H., & Noga, J. 1997, JChPh, 106, 9639
Herbst, E., & van Dishoeck, E. F. 2009, ARA&A, 47, 427
Kawaguchi, K., Takano, S., Ohishi, M., et al. 1992, ApJL, 396, L49
Kooij, D. 1893, Zeitschr. Phys. Chem., 12, 155
Krim, L., Guillemin, J. C., & Woon, D. E. 2019, MNRAS, 485, 5210
Laidler, K. J. 1996, PApCh, 68, 149
Loomis, R. A., Zaleski, D. P., Steber, A. L., et al. 2013, ApJL, 765, L9
Lovas, F. J., Hollis, J. M., Remijan, A. J., & Jewell, P. R. 2006, ApJL, 645, L137
Lupi, J., Puzzarini, C., Cavallotti, C., & Barone, V. 2020, J. Chem. Theory Comp., 16, 5090
McGuire, B. A. 2018, ApJS, 239, 17
Møller, C., & Plesset, M. S. 1934, PhRv, 46, 618
Montgomery, J. A., Frisch, M. J., Ochterski, J. W., & Petersson, G. A. 2000, JChPh, 112, 6532
Papajak, E., Zheng, J., Xu, X., Leverentz, H. R., & Truhlar, D. G. 2011, J. Chem. Theory Comp., 7, 3027
Pechukas, P., & Light, J. C. 1965, JChPh, 42, 3281
Puzzarini, C., & Barone, V. 2011, PCCP, 13, 7180
Puzzarini, C., & Barone, V. 2020, PCCP, 22, 6507
Puzzarini, C., Biczysko, M., Barone, V., et al. 2014, J. Phys. Chem. Lett., 5, 534
Puzzarini, C., Bloino, J., Tasinato, N., & Barone, V. 2019, ChRv, 119, 8131
Puzzarini, C., Salta, Z., Tasinato, N., et al. 2020, MNRAS, 496, 4298
Raghavachari, K., Trucks, G. W., Pople, J. A., & Head-Gordon, M. 1989, CPL, 157, 479
Rivilla, V. M., Martín-Pintado, J., Jiménez-Serra, I., et al. 2018, MNRAS, 483, L114
Salta, Z., Tasinato, N., Lupi, J., et al. 2020, ESC, 4, 774
Shingledecker, C. N., Molpeceres, G., Rivilla, V. M., Majumdar, L., & Kästner, J. 2020, ApJ, 897, 158
Theule, P., Borget, F., & Mispelaer, F. e. a. 2011, A&A, 534, A64
Tonolo, F., Lupi, J., Puzzarini, C., & Barone, V. 2020, ApJ, 900, 85
Vazart, F., Calderini, D., Puzzarini, C., Skouteris, D., & Barone, V. 2016, J. Chem. Theory Comp., 12, 5385
Vazart, F., Latouche, C., Skouteris, D., Balucani, N., & Barone, V. 2015, ApJ, 810, 111
Woon, D. E., & Dunning, T. H., Jr. 1995, JChPh, 103, 4572
Zaleski, D. P., Seifert, N. A., Steber, A. L., et al. 2013, ApJL, 765, L10
Zeng, S., Jiménez-Serra, I., Rivilla, V. M., et al. 2018, MNRAS, 478, 2962

Chapter 6

From gas-phase to gas-grain chemistry: hydrogen cyanide isomerization on ices

In this Chapter a simple but paradigmatic test case of gas-grain chemistry in the ISM is presented. Isomerization of hydrogen cyanide to hydrogen isocyanide on icy grain surfaces is investigated by different composite scheme. The main aim of this work was the implementation and validation of a general computational strategy for the study of the thermochemistry and kinetics of chemical processes taking place on interstellar icy-grains. To this end composite methods rooted in the coupled cluster ansatz have been combined with hybrid and double hybrid functionals together with molecular mechanics force field to characterize the stationary points ruling the reactive potential energy surfaces on model clusters sufficiently large to minimize spurious boundary effects. Next powerful master equation/RRKM models have been employed to compute reaction rates including tunneling effects. As a demanding test case we have selected the HCN/HNC reactions for which the previous available computational results are not fully satisfactory. In this Chapter, objectives O1, O2 and O4 are pursued.

JCTC
Journal of Chemical Theory and Computation

Article

Gliding on Ice in Search of Accurate and Cost-Effective Computational Methods for Astrochemistry on Grains: The Puzzling Case of the HCN Isomerization

Carmen Baiano, Jacopo Lupi, Vincenzo Barone,* and Nicola Tasinato*

Cite This: https://doi.org/10.1021/acs.jctc.1c01252

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: The isomerization of hydrogen cyanide to hydrogen isocyanide on icy grain surfaces is investigated by an accurate composite method (jun-Cheap) rooted in the coupled cluster ansatz and by density functional approaches. After benchmarking density functional predictions of both geometries and reaction energies against jun-Cheap results for the relatively small model system $HCN\cdots(H_2O)_2$, the best performing DFT methods are selected. A large cluster containing 20 water molecules is then employed within a QM/QM′ approach to include a realistic environment mimicking the surface of icy grains. Our results indicate that four water molecules are directly involved in a proton relay mechanism, which strongly reduces the activation energy with respect to the direct hydrogen transfer occurring in the isolated molecule. Further extension of the size of the cluster up to 192 water molecules in the framework of a three-layer QM/QM′/MM model has a negligible effect on the energy barrier ruling the isomerization. Computation of reaction rates by the transition state theory indicates that on icy surfaces, the isomerization of HNC to HCN could occur quite easily even at low temperatures thanks to the reduced activation energy that can be effectively overcome by tunneling.
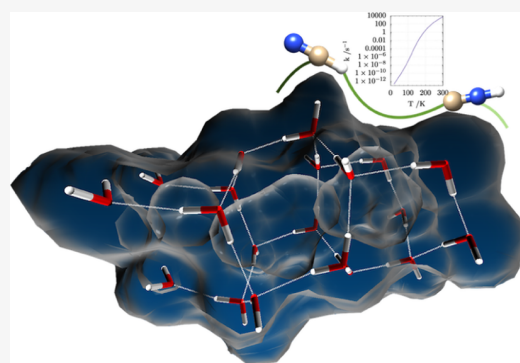
1. INTRODUCTION

In 2018, McGuire published a census of Interstellar, Circumstellar, Extragalactic, Protoplanetary Disks, and Exoplanetary Molecules[1] including more than 200 molecules (containing from 2 to 70 atoms) and this number is steadily increasing thanks to the modern technologies of new observatory telescopes.[2] The identification of many interstellar complex organic molecules (iCOMs) defeated the old and general idea that the interstellar medium (ISM) was an empty vial where chemical reactivity could not operate. Questions about the formation of iCOMs in such extreme conditions and the evolution of molecular complexity fueled the curiosity of astrochemists all over the world.[3] While gas-phase reactions seemed the obvious choice to explore the formation pathways of molecular systems in such rarefied environments, the ubiquitous presence of dust and grains and the mismatch between some observations and the molecular abundances predicted by gas-phase models have boosted the role of solid-state chemistry.[4,5] Since the discovery of the catalytic role of grains for H_2 formation,[6,7] astrochemists and physicists have struggled looking for gas-grain models that could provide a comprehensive picture of chemical processes in the ISM. At the low temperatures of molecular clouds, molecules in the gas phase accrete icy mantles freezing out onto grain surfaces[8,9] and leading to porous and amorphous icy surfaces,[10−12] which can host local reactants triggering a molecular reactivity not

feasible in the gas phase. The composition and morphological features make the simulation of these icy structures a great challenge in this field.[9,13]

The difficulty of performing experimental studies for systems capable of mimicking the harsh conditions of the ISM, calls for computational simulations of periodic surfaces and/or suitable model clusters able to take into the proper account the main structural features responsible for the chemistry at the interface.[14,15] This translates into the necessity of simulating extended systems, thus making the computational burden prohibitive for the accurate state-of-the-art methods developed for isolated molecules.[16] Because water is the main component of polar icy mantles,[17,18] a lot of efforts have been devoted to the investigation of the adsorption and formation of iCOMs on water clusters used to mimic interstellar ices. The structures of H_2O clusters containing up to 22 atoms have been worked out from molecular dynamics simulations and made available in online databases.[19] Some years ago, Rimola et al. studied iCOM formation pathways on clusters including up to 33

ACS Publications

water molecules obtained by combining two $(H_2O)_{18}$ clusters taken from the (010) surface of ice-XI[20] and removing three molecules to facilitate the construction of the final cluster.[21] Furthermore, attempts to include the structural modifications induced by UV and cosmic rays photo-processing have been made by means of small radical and ionized water clusters.[22] More recently, molecular dynamics has been used to model amorphous water ices[23] and to simulate mixed CO/H_2O ices.[24] Adsorption energies on clusters of larger size have been evaluated by a two-layer our own N-layered integrated molecular orbital molecular mechanics (ONIOM) model, with the higher-level layer treated by means of density functional theory (DFT), and the lower-level one described through molecular mechanics (MM)[25,26] or semiempirical quantum chemical methods.[27]

While coupled cluster theory including full treatment of single and double excitations together with perturbative estimation of triple excitations [CCSD(T)], possibly in conjunction with composite schemes to estimate the complete basis set (CBS) limit, is considered the gold-standard for accurate predictions,[28] the size of the systems to be dealt with in the case of ice-mediated chemistry makes density functional theory the only viable route in terms of accuracy to computational cost trade-off. As is well known, the reliability of DFT strongly depends on the specific system and properties at hand and on the choice of the density functional (DF) among an ever increasing number of possible formulations. In this respect, benchmark is a fundamental step for ranking the reliability of DFT model chemistries, also in connection with the computational cost, and hence it represents a very active field of research.

Concerning the specific topic of adsorption and reactivity of iCOMs on interstellar ice analogues, to the best of our knowledge, systematic benchmark studies are still lacking. In this connection, Enrique-Romero et al.[29] performed a calibration analysis of radical−water interactions and activation energy for NH_2 + HCO and CH_3 + HCO reactions in the presence of one and two water molecules. They tested the accuracy of B3LYP and BHLYP functionals (both with and without dispersion corrections) in conjunction with the 6-311++G(2df,2pd) basis set taking CASPT2/cc-pVTZ and CCSD(T)/aug-cc-pVTZ levels of theory as reference. That analysis was focused on the interaction and activation energies, while recent works have highlighted that reliable geometries are fundamental prerequisites for accurate thermochemistry and kinetics.[16] In this respect, the B3LYP functional can be unable to predict correct structures for van der Waals complexes[30] and transition states.[31] Furthermore, the use of CCSD(T)/triple-ζ energies cannot be recommended as a reference in benchmark studies because basis set truncation and lack of core−valence correlation limit the accuracy, thus introducing a bias in the reference values. This issue can be overcome by resorting to composite methods that aim at minimizing the errors relying on well-tested additive approximations.[16,32]

In this work, we assess the performances of several DFT model chemistries in evaluating the structural and energetic aspects of ice-mediated interstellar reactions employing the HCN ⇌ HNC isomerization catalyzed by water molecules as a paradigmatic process. On the one side, this can be considered a model for more complex reactions mediated by ice surfaces and, on the other side, the chosen system is small enough to allow the exploitation of state-of-the-art composite methods to generate accurate reference values for both geometries and reaction energies. The HCN ⇌ HNC isomerization has been widely studied because the observed HNC/HCN ratio in the ISM cannot be predicted on the basis of the proposed gas-phase mechanisms. Moreover, both HCN and HNC can be involved in the formation of amino acid precursors in the Strecker synthesis of glycine.[33,34] Gardebien and Sevin investigated the process for the isolated molecule and with explicit inclusion of two to four water molecules[35] finding that the most favorable mechanism consists of a one-step path involving a proton relay mediated by the water cluster. Koch et al. employed a more realistic model including seven additional water molecules to simulate the local environment of the icy surface and employing the polarizable continuum model (PCM) to account for bulk effects.[36] According to the available data, the water cluster acts as a catalyst lowering the energy barrier with respect to the gas phase, an effect that progressively smooths increasing the number of H_2O molecules. Intermolecular hydrogen bond drives both the interaction of HCN and HNC with the ice surface and the isomerization process. This represents the most common mechanism through which molecules adsorb and react on ISM polar ices.

On these grounds, we decided to perform a detailed study of the HCN ⇌ HNC isomerization by state-of-the-art quantum chemical methods and realistic cluster models. The work is organized as follows: the computational methods are described in Section 2, while the outcomes of the benchmark are detailed in Section 3 concerning both geometries and energies, thus leading to the identification of the best performing DFT model chemistries in terms of the trade-off between accuracy and computational cost. Despite the fact that the benchmark is carried out on a simplified model, the outcomes are expected to be of general validity, especially with respect to the relative performances of the tested methods which can then be transferred to larger H_2O clusters. With this in mind, at the end of Section 3, the best performing methods are employed to simulate the HCN ⇌ HNC isomerization catalyzed by a cluster of 20 water molecules and then further embedded in a 172 water slab described through MM. Finally, reaction rates are computed in the framework of the transition state theory (TST) including tunneling.

2. COMPUTATIONAL METHODOLOGY

For the benchmark study, we selected 10 DFs belonging to different families: two hybrids (B3LYP, BHLYP),[37−39] a long-range corrected DF (ωB97X-D),[40] three meta-hybrids (PW6B95,[41] BMK,[42] and M06-2X[43]), one meta-NGA (MN15[44]), the B2PLYP,[45] and the two spin-component-scaled (DSD-PBEP86 and revDSD-PBEP86)[46,47] double hybrids. To test the accuracy to computational cost trade-off, for each functional, six basis sets have been considered. In particular, we selected Dunning's aug-cc-pVnZ basis sets (n = D, T)[48,49] as well as the corresponding jun- and jul-modifications from Truhlar's calendar family.[50] All the DFT calculations include empirical dispersion corrections according to the DFT-D3 scheme proposed by Grimme[51] with the Becke-Johnson damping function (BJ),[52,53] which are fundamental for the correct prediction of van der Waals complexes,[54−56] transition states,[57] and surface processes.[58,59] Accurate reference geometries and energies for the benchmark were generated by using the Cheap composite scheme (ChS)[60,61] and its recent jun-Cheap revision (jun-ChS),[16,32] with the latter appearing

the best option because of the increased reliability for noncovalent interactions and the better description of the water dimer structure. Indeed, for $(H_2O)_2$, ChS and jun-ChS geometries were first compared to highly accurate CCSD(T)-F12b/CBS + fT + fQ + CV + REL + DBOC values.[62] The results, reported in Table S1 of the Supporting Information, show that bond lengths and valence angles are reproduced very accurately, with maximum errors of −0.003 Å and −0.2°, while there is a deviation of 3° for the angle defining the orientation of the C_2 axis of the acceptor water molecule with respect to the O−O axis. On the basis of the reliable geometry delivered by jun-ChS, this method was used as reference for both equilibrium geometries and electronic energies.

Preliminary B3LYP-D3/aug-cc-pVTZ computations of the HCN ⇌ HNC reactive PES were refined at the jun-ChS level. The nature of the identified stationary points (minima or saddle points) was checked through frequency calculations performed at each level of theory. All calculations have been carried out with the Gaussian software,[63] except the geometry optimizations at the ChS and jun-ChS levels, which have been performed using the CFOUR package.[64,65] Because revDSD-PBEP86 is not among the Gaussian built-in functionals, it has been defined by setting proper IOP flags on top of the DSD-PBEP86 functional.

Full geometry optimizations were performed for the complexes containing 2−4 H_2O molecules, whereas for the 20 water model cut from the ice XI (010) surface, 8 molecules belonging to the cluster edge (see Figure S2 of the Supporting Information) were kept frozen at their positions in the crystal in order to prevent geometrical distortions causing a nonphysical breakdown of the crystalline pattern. The best-performing methods were employed within a QM/QM′ strategy for simulating the HCN ⇌ HNC isomerization on this cluster in order to evaluate the catalytic effect of the ice surface. For the purpose, we employed the ONIOM method[66] treating the reaction center (i.e., the adsorbate and four water molecules) at a higher level of theory (i.e., a double-hybrid DF or even jun-ChS), whereas a less computationally demanding method (i.e., a meta-hybrid DF) was used for the remaining molecules of the cluster. A much larger cluster containing 192 water molecules was also investigated by means of a three-layer (QM/QM′/MM) ONIOM approach enforcing the so-called mechanical embedding and employing the Amber force field.[67] In this case, the structural degrees of freedom of the adsorbate and the first 20 H_2O molecules were optimized while freezing the coordinates of the remaining 172 water molecules to those of the regular (010) surface of ice XI. Test computations with the more refined electrostatic embedding showed negligible differences on the relative energies.

Rate constants were computed solving the multiwell one-dimensional master equation using the chemically significant eigenvalues method.[68] Rate coefficients were determined using conventional TST within the rigid-rotor harmonic-oscillator approximation,[69] also incorporating tunneling and nonclassical reflection effects by means of the Eckart model.[70] The rates evaluated at different temperatures were fitted by a simple Arrhenius equation or by the three-parameter modified Arrhenius equation proposed by Kooij[71,72]

$$k(T) = A\left(\frac{T}{300}\right)^n \exp\left(-\frac{E_a}{RT}\right) \qquad (1)$$

where A, n, and E_a are the fitting parameters, R is the universal gas constant, and the limiting Arrhenius behavior is recovered when $n = 0$. All the kinetic computations were performed with the MESS code.[68]

3. RESULTS AND DISCUSSION

As widely discussed in the Introduction, the reliable modeling of interstellar ices is an extremely complex task, requiring the assessment of DFT methods for geometry and energy predictions that offer the proper balance between accuracy and computational burden. The lack of systematic studies addressing this issue for solid-state astrochemical processes calls for a dedicated benchmark. While small-size clusters cannot be fully representative of an extended substrate, the interaction of small molecules with water ice surfaces is generally guided by hydrogen bonds between the polar functional groups of the molecule and the exposed H and O atoms of the ice surface, which are already present in the smallest cluster models. Therefore, while the thermochemistry computed by using clusters composed of a small number of H_2O molecules is not representative of real icy-grain chemistry, the outcomes of the benchmark are safely transferable to larger clusters. In the following subsections we report the results of our benchmark study, concerning first geometries and then reaction and activation energies. Finally, to scale-up to a more realistic water ice model, we report a full characterization of the PES of the HCN ⇌ HNC isomerization on clusters composed by either 20 or 192 H_2O molecules.

3.1. Geometry Snow-Board. The HCN → HNC isomerization is an endothermic process involving a high activation energy and the jun-ChS results are close to the current best estimates[73] for both the reaction (61.6 vs 63.8 kJ/mol) and activation (198.5 vs 201.1 kJ/mol) energy. The addition of two water molecules leads to the formation of a hydrogen-bonded van der Waals adduct featuring the interactions between the H atom of HCN and the oxygen of one water molecule and between the N atom and one hydrogen of the second water molecule. Then, the reaction proceeds through a transition state for the $(H_2O)_2$-mediated proton transfer reaching, in this way, a post-reactive complex in which carbon is engaged in a weak H-bond with a hydrogen of the first H_2O molecule, while the H atom of HNC interacts with the oxygen of the second water molecule. The structures of all the stationary points ruling the reactive PES are sketched in Figure 1 together with selected geometrical parameters obtained at the jun-ChS level.

The accuracy of the considered DFT model chemistries has been evaluated with respect to jun-ChS values and the overall mean absolute errors (MAEs) and mean absolute relative errors (REs) have been evaluated over all the bond lengths, valence, and dihedral angles of the species involved in the PES. The full list of data can be found in Table S2 and Figure S1 of the Supporting Information. As a rule of thumb (with some exceptions for dihedral angles), triple-ζ basis sets show smaller errors than the corresponding double-ζ ones, with the improvement being less pronounced along the jun-, jul-, and aug-series. In general, the tested hybrid and meta-hybrid DFs on the one side, and the double-hybrids on the other, give similar trends for the MAEs, with the notable exception of BHLYP-D3 in conjunction with the jul-cc-pVDZ basis set, that strongly overshoots and the ωB97X-D functional that shows larger deviations from the jun-ChS reference values, especially for valence and dihedral angles. In the case of the BHLYP-D3/
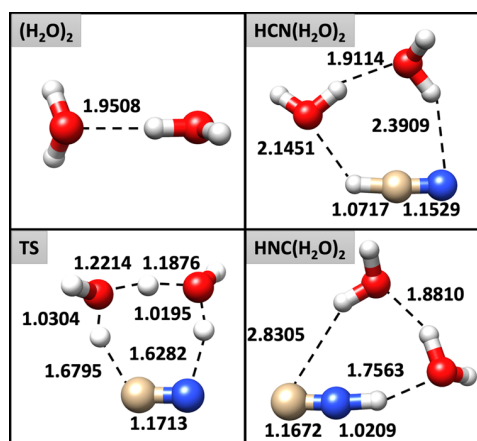
94

pubs.acs.org/JCTC

Figure 1. Stationary points on the reactive PES of the HCN ⇌ HNC isomerization catalyzed by two water molecules. Representative bond lengths (Å) obtained at the jun-ChS level are reported.

jul-cc-pVDZ model, MAEs as large as 0.09 Å, 6°and 10°were observed for bond lengths, valence, and dihedral angles, respectively. These results are related to the inability of reproducing a tight structure for the post-reactive complex. Specifically, one H-bond in CNH···(H$_2$O)$_2$ (see Figure 1) is broken and the product collapses into an open structure. All in all, it can be observed that the most promising (meta-)hybrid DFs are PW6B95-D3, BMK-D3, M06-2X, and MN15 coupled to triple-ζ basis sets (or, at least, the jul-cc-pVDZ one). Concerning the double-hybrid functionals, the best structural predictions are delivered by DSD-PBEP86-D3 and revDSD-PBEP86-D3 that show comparable accuracy. In order to have a clearer picture of the performance of the different model chemistries in the prediction of the geometries involved in the HCN ⇌ HNC isomerization assisted by two water molecules, Figure 2 reports the overall REs of each method, evaluated by averaging the REs of the geometrical parameters of all the species on the reactive PES. Inspection of this figure reveals that, among the (meta-)hybrid DFs, the best results for double-ζ basis sets are delivered by PW6B95-D3 and BMK-
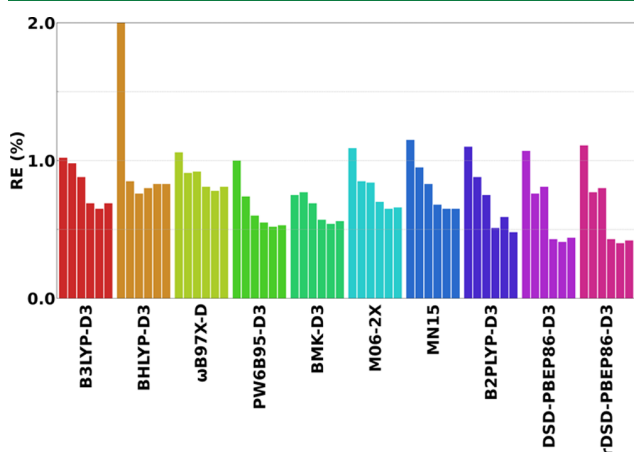

Figure 2. Total REs (%) of the geometries of the species on the PES of the HCN ⇌ HNC isomerization assisted by two water molecules for the investigated DFT methods with respect to jun-ChS reference values. For each functional, the different basis sets are reported in the following order: jun-DZ, jul-DZ, aug-DZ, jun-TZ, jul-TZ, and aug-TZ.

D3. In particular, PW6B95-D3/jul-cc-pVDZ, BMK-D3/aug-cc-pVDZ, and PW6B95-D3/aug-cc-pVDZ score REs in the 0.60−0.74% range. The PW6B95-D3 and BMK-D3 DFs are the best performers also in conjunction with triple-ζ basis sets showing REs around 0.55%. Concerning the double-hybrid functionals, it is apparent that their use in conjunction with a double-ζ basis set does not justify the computational overload in comparison with hybrid functionals; however, both DSD-PBEP86-D3 and its revision predict improved geometries when employed in conjunction with triple-ζ basis sets, reaching a RE of only 0.4% for the jul-cc-pVTZ basis set. In passing, it is interesting to point out that, these functionals have also demonstrated to be excellent performers in predicting structural and spectroscopic properties of gas-phase molecules.[74,75]

3.2. Skiing on Adsorption, Reaction, and Activation Energies. The functional/basis set combinations with the optimal accuracy/cost trade-off for geometry predictions have been identified in the previous section. Reactivity studies require the calculation of accurate formation and activation energies for the subsequent kinetic analysis. For this reason, some of the DFT methods delivering the best geometrical predictions have been selected and their accuracy for computing adsorption, activation, and reaction (electronic) energies explored using again jun-ChS results as references. In a first step, the impact of the geometry on the energetics has been assessed, by evaluating jun-ChS electronic energies for the different DFT structures. In a second step, the formation energies stemming from full DFT computations (for both geometries and energies) have been analyzed.

Electronic energies obtained at the jun-ChS level on top of selected DFT geometries are reported in Table 1, while the corresponding error analysis is presented in Figure 3.

Table 1. jun-ChS Formation Energies (kJ/mol) with Respect to Isolated HCN and (H$_2$O)$_2$ for Each Species along the HCN/HNC···(H$_2$O)$_2$ Isomerization PES Evaluated on Top of DFT Geometries

level of theory for geometry	HCN···(H$_2$O)$_2$	TS	CNH···(H$_2$O)$_2$	HNC + (H$_2$O)$_2$
PW6B95-D3/jul-DZ	−33.38	99.37	15.23	62.31
BHLYP-D3/aug-DZ	−33.40	99.21	14.87	61.88
PW6B95-D3/aug-DZ	−33.44	99.40	15.17	62.28
BMK-D3/aug-DZ	−33.12	99.13	15.04	62.26
M06-2X/aug-DZ	−33.32	99.67	15.38	62.28
MN15/aug-DZ	−33.39	99.36	15.61	62.63
PW6B95-D3/jul-TZ	−33.49	99.15	14.73	61.89
BMK-D3/jul-TZ	−33.47	99.18	14.59	61.66
M06-2X/jul-TZ	−33.38	99.49	14.91	61.85
MN15/jul-TZ	−33.51	99.33	14.98	62.03
DSD-PBEP86-D3/jul-TZ	−33.52	99.35	15.06	62.22
revDSD-PBEP86-D3/jul-TZ	−33.54	99.31	15.02	62.22
jun-ChS	−33.42	99.26	15.03	62.27

It is quite apparent that the energetic results obtained employing geometries optimized with all the tested methods are in remarkable agreement with the jun-ChS reference, with deviations smaller than 0.6 kJ/mol, even though some of them provide an unbalanced description of the different elementary processes. For example, the MN15/aug-cc-pVDZ and BMK-D3/jul-cc-pVTZ models yield excellent predictions of both the interaction energy of hydrogen cyanide with (H$_2$O)$_2$ and the
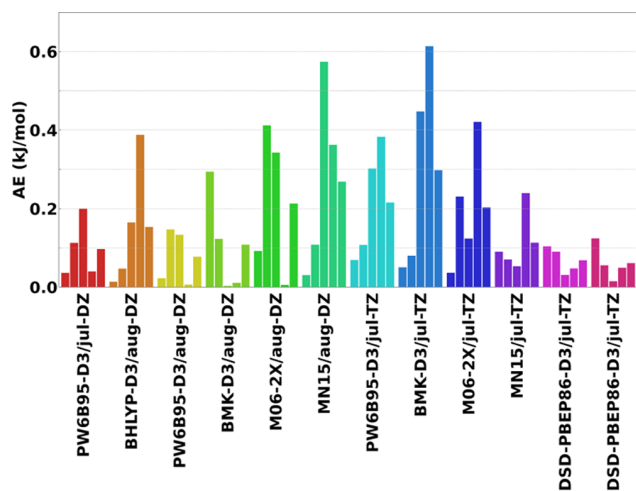
Figure 3. Error analysis for jun-ChS formation energy (kJ/mol) obtained on top of DFT geometries in comparison with full (both energies and geometries) jun-ChS results. Each color corresponds to a DFT model chemistry and collects absolute errors for the formation energy of each species along the PES with respect to isolated reactants: (1) pre-reactive complex; (2) transition state; (3) post-reactive complex; (4) products; and (5) MAE over all of the steps along the PES.

transition state energy, with errors around 0.05 and 0.1 kJ/mol, respectively; however, the computed HNC formation energy (at the BMK-D3 level) and its interaction energy with the water dimer (at the MN15 level) show significantly larger errors. Among the (meta-)hybrid functionals, the best and most consistent energetic description is given by PW6B95-D3 in conjunction with jul- or aug-cc-pVDZ basis sets, which reaches an overall MAE (evaluated by considering the relative electronic energies of all the stationary points ruling the PES) close to 0.1 kJ/mol and a maximum deviation of 0.2 kJ/mol.

Moving to the double-hybrid DFs, the DSD-PBEP86-D3 and revDSD-PBEP86-D3 models in conjunction with the jul-cc-pVTZ basis set yield excellent performances, scoring a MAE of about 0.06 kJ/mol and reproducing the formation energies of all the elementary steps with a maximum deviation of 0.12 kJ/mol for the pre-reactive complex at the revDSD-PBEP86-D3/jul-cc-pVTZ level.

The relative electronic energies of all the stationary points fully evaluated at different DFT levels (i.e., energies and geometries) are collected in Table 2 and the MAEs from the jun-ChS computations are shown in Figure 4. In general terms, the results mirror those obtained for jun-ChS energies evaluated on top of DFT geometries, with the only difference being the much larger deviations, which now span the 5−29 kJ/mol range. Furthermore, the relative stability of CNH⋯$(H_2O)_2$ is always strongly underestimated (becoming even negative with BHLYP-D3, BMK-D3, and MN15 functionals) except at the PW6B95-D3 and, especially, DSD-PBEP86-D3 and revDSD-PBEP86-D3 levels in conjunction with the jul-cc-pVTZ basis set. All the (meta-)hybrid DFs show MAEs larger than 10 kJ/mol, with the exception of PW6B95-D3, which is the only functional that reaches a MAE around 6 kJ/mol in conjunction with the jul- and aug-cc-pVDZ basis set and of 3.8 kJ/mol employing the jul-cc-pVTZ basis. The DSD-PBEP86-D3 and revDSD-PBEP86-D3 functionals confirm their good performances in conjunction with the jul-cc-pVTZ basis set, with MAE around 3 kJ/mol and maximum deviations of 10.4

Table 2. DFT Formation Energies (kJ/mol) with Respect to Isolated HCN and Water Dimer $(H_2O)_2$ for Each Species along the HCN⋯$(H_2O)_2$ Isomerization PES

level of theory[a]	HCN⋯$(H_2O)_2$	TS	CNH⋯$(H_2O)_2$	HNC + $(H_2O)_2$
PW6B95-D3/jul-DZ	−34.12	86.96	8.34	56.74
BHLYP-D3/aug-DZ	−38.42	89.54	−1.35	51.12
PW6B95-D3/aug-DZ	−33.99	88.45	7.70	56.03
BMK-D3/aug-DZ	−36.30	81.76	−6.60	42.58
M06-2X/aug-DZ	−37.46	70.46	1.14	52.78
MN15/aug-DZ	−36.67	76.94	−2.82	47.31
PW6B95-D3/jul-TZ	−32.67	94.10	11.31	56.81
BMK-D3/jul-TZ	−35.97	88.16	0.11	48.23
M06-2X/jul-TZ	−37.16	76.00	4.64	54.06
MN15/jul-TZ	−35.26	85.53	1.92	48.88
DSD-PBEP86-D3/jul-TZ	−35.08	88.86	14.86	64.93
revDSD-PBEP86-D3/jul-TZ	−33.97	93.50	16.21	64.85
jun-ChS	−33.42	99.26	15.03	62.27

[a]For both energy and geometry.
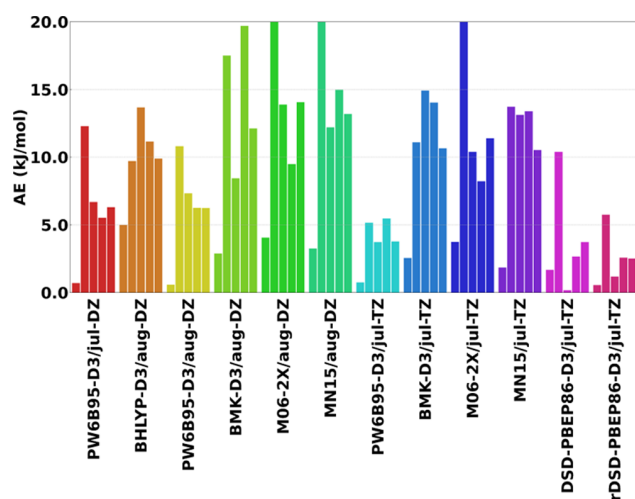
Figure 4. Error analysis for DFT formation energies (kJ/mol) in comparison with jun-ChS values. Each color corresponds to a DFT model chemistry (used for both geometry and energy) and collects absolute errors for the formation energy of each species along the PES with respect to isolated reactants: (1) pre-reactive complex; (2) transition state; (3) post-reactive complex; (4) products; and (5) MAE over all of the steps along the PES.

kJ/mol. Hence, the model chemistries with the optimal accuracy for structural parameters are also the best choices for thermochemistry.

These results confirm the conclusions of recent benchmarks about the quality of PW6B95-D3/jul-cc-pVDZ and DSD-PBEP86-D3/jul-cc-pVTZ models for geometries, vibrational frequencies, and other spectroscopic parameters.[74,75] Noted is that the core−valence correlation has not been included for double hybrid functionals because it was not taken into account in their original parametrization and its contribution is anyway within the expected error bar at least for molecular systems containing only hydrogen and second-row atoms (see Table S3 of the Supporting Information for CV contributions in jun-ChS results). Furthermore, some test computations performed with quadruple-ζ basis sets showed that complete basis set extrapolation has a negligible effect on all the trends discussed above. For example, the relative electronic energies

of the stationary points obtained by using the DSD-PBEP86-D3 functional in conjunction with the aug-cc-pVQZ basis set ($\Delta E = -35.04$, 89.51, 14.83, and 64.78 kJ/mol) differ from the counterparts obtained employing the jul-cc-pVTZ basis set by 0.65 kJ/mol at most (for the TS). Finally, although triple-ζ basis sets possibly deliver more robust results for hybrid functionals, this computational level will be used in the following only to describe the environmental effects in the framework of QM/QM′ computations where the increased computational cost with respect to double-ζ results is not justified, in our opinion, by the marginally improved robustness.

3.3. Scaling-Up toward Extended Systems: Best Performers at Work. The benchmark performed for both geometries and energies permits the identification of the best candidates for setting up a QM/QM′ ONIOM strategy for the study of the HCN $\rightleftharpoons$ HNC isomerization on large clusters capable of providing a more realistic modeling of the icy-grain and of the molecule−surface interactions.

At first, a cluster composed by 20 water molecules (shown in Figures 5 and S2 of the Supporting Information) has been
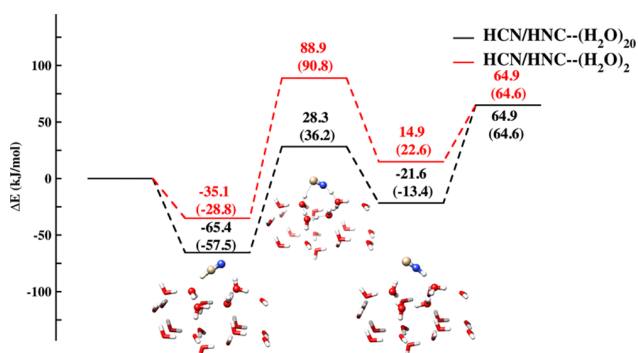
Figure 5. Potential energy profile for HCN $\rightleftharpoons$ HNC isomerization mediated by the $(H_2O)_{20}$ cluster and the $(H_2O)_2$ dimer. Red lines refer to the HCN isomerization catalyzed by $(H_2O)_2$ and both geometries and ΔE have been computed at the DSD-PBEP86-D3/jul-cc-pVDZ level. Black lines refer to the ONIOM results for the reaction catalyzed by $(H_2O)_{20}$. The ball and stick representation is used for atoms of the highest QM level (DSD-PBEP86-D3/jul-cc-pVDZ), while the tube representation is used for the atoms belonging to the QM′ (PW6B95-D3/jul-cc-pVDZ) portion. ΔE corrected for ZPVE are reported in parenthesis with ZPVEs calculated at the same level of theory as the corresponding energies and geometries.

used, in which the pattern of exposed water molecules is suitable for a H-relay mechanism mediated by four water molecules. It should be noted that in ref 36 a proton-relay mechanism mediated by three water molecules, in turn solvated by seven additional waters, was used. In the present work, the four H_2O molecules involved in the hydrogen transfer and the adsorbed species have been considered as the reaction center of the process under study; hence, they constitute the higher-level QM portion of the system. Following the outcomes of the benchmark study, the DSD-PBEP86-D3 functional in conjunction with the jul-cc-pVTZ basis set has been used for the purpose, while the remaining part of the cluster, treated at a lower QM′ level, has been described by the PW6B95-D3 DF in conjunction with the jul-cc-pVDZ basis set. The energetic profile of the HCN $\rightleftharpoons$ HNC isomerization occurring on the $(H_2O)_{20}$ cluster is reported in Figure 5 where it is also compared with that for the $(H_2O)_2$-

mediated process. Going from the process assisted by two waters to that assisted by four water molecules in the $(H_2O)_{20}$ cluster lowers the energy of all the species present in the reactive PES. The most remarkable effect is the reduction of the energy barrier ruling the isomerization when considering the 20 water cluster in place of just two water molecules involved in the simplest possible relay mechanism.

The dependence of the energy profile on the number of water molecules involved in the relay mechanism was already pointed out.[35,36] However, only few water molecules were considered and no attempt to simulate the effect of ice bulk has been reported beyond the PCM level, whose reliability is, however, questionable for hydrogen-bonding solids. For comparison, Table 3 lists the relative energies (corrected for the zero point vibrational energies, ZPVEs of the elementary steps obtained by Koch et al.,[36] and in the present work) (further details are given in Table S5 of the Supporting Information). As it can be seen, the relative energy for HCN interacting with the water cluster is only marginally affected by the cluster size, but there is a huge effect on the activation barrier. While an overall fair agreement between the present results and those obtained in ref 36, can be noted, there is a difference of about 18 kJ/mol for the energy of the transition state. This can be explained by considering that Koch et al.[36] investigated the role of the crystalline environment by optimizing for the different stationary points the positions of seven water molecules around the HNC···$(H_2O)_3$ complex without any constraint related to the arrangement of water molecules in icy structures. The importance of the morphological pattern in ice is highlighted by the present results: indeed, using a $(H_2O)_{20}$ cluster with the same molecular arrangement as in ice XI rules out the possibility of a process catalyzed by two or three water molecules. Rather, the molecular arrangement at the surface permits a process assisted by four water molecules (see Figure 5).

Further support to the reliability of the results is provided by the comparable barrier obtained by another ONIOM computation in which the high-level part of the system [HCN···$(H_2O)_4$] is treated at the jun-ChS instead of DSD-PBEP86-D3 level without any additional geometry optimization (last line of Table 3). What is even more gratifying is that the differences between the results obtained for the smallest HCN···$(H_2O)_2$ model and the larger model clusters (values in parenthesis in Table 3) obtained at the DSD-PBEP86 level are in quantitative agreement with the jun-ChS counterparts. This paves the route toward the computation of very reliable parameters for reactions occurring on icy grains by combining jun-ChS results for small models and ONIOM(DSD-PBEP86:PW6B95-D3) values for large model clusters.

This approach can be further extended to very large models by employing a three-layer QM/QM′/MM ONIOM model. In order to also analyze this aspect, we have embedded the HCN···$(H_2O)_{20}$ cluster in a large model of ice-XI containing 172 water molecules described by the Amber force field (see Figure 6). The results collected in Table 3 show that inclusion of the MM layer further stabilizes the HCN isomer with respect to the HNC counterpart by about 4 kJ/mol, but has a negligible effect on the energy barrier (less than 0.4 kJ/mol). Taking into account the estimated error bar of the overall computational approach (about 4 kJ/mol), the results obtained for the HCN···$(H_2O)_{20}$ model can be considered essentially converged with respect to further extension of the ice substrate.

Table 3. Relative Ground-State Energies (kJ/mol) with Respect to HNC···(H₂O)ₙ Post Reactive Complex and Comparison with the Results of ref36[a]

	total H₂O (n)	relay H₂O (n_R)	TS[b]	HCN··· (H₂O)ₙ[b]
ref 36	2	2	74.1	−42.3
	3	3	43.9 (−30.2)	−41.4 (−0.9)
	10	3	13.8 (−60.3)	−41.4 (−0.9)
B3LYP[c]	2	2	73.8	−39.8
	3	3	44.1 (−29.7)	−41.7 (−1.9)
PW6B95-D3[d]	2	2	70.9	−43.6
	3	3	52.4 (−18.5)	−42.0 (−1.6)
	4	4	49.3 (−21.6)	−42.7 (−0.9)
	192	4	36.6[e] (−33.4)	−40.0[e] (−3.6)
DSD-PBEP86-D3[f]	2	2	68.1	−51.5
	3	3	48.3 (−19.8)	−49.6 (−1.9).
	4	4	46.1 (−22.0)	−47.8 (−3.7)
	20	4	32.3[g] (−35.8)	−43.7[g] (−7.8)
			32.1[h] (−36.0)	−41.7[h] (−9.8)
	192	4	32.4[i] (−35.7)	−45.7[i] (−6.8)
jun-ChS	2	2	78.3	−50.0
	3	3	58.5 (−19.8)[j]	−48.0 (−2.0)[j]
	20	4	44.1 (−34.2)[k]	−40.8 (−9.2)[k]

[a]Both the total number of water molecules (n) and the number of water molecules directly involved in the relay mechanism (n_R) are indicated. All values include ZPVEs. [b]In parentheses is the difference with respect to (H₂O)₂ results. [c]6-31+G(d,p) basis set as in ref 36. [d]jul-cc-pVDZ basis set. [e]QM/MM energies and ZPVEs. 20 water molecules treated at the PW6B95-D3 level, the remaining molecules described by the Amber force field. [f]jul-cc-pVTZ basis set. [g]ONIOM geometries and ZPVE. DSD-PBEP86/jul-cc-pVTZ for adsorbate and molecules involved in the relay mechanism, PW6B96-D3/jul-cc-pVDZ for the water molecules not involved in the relay mechanism. [h]DSD-PBEP86/jul-cc-pVTZ energies with geometries and ZPVE at the PW6B95-D3/jul-cc-pVDZ level. [i]DSD-PBEP86:PW6B95-D3:Amber energies on PW6B95-D3:Amber geometries. ZPEs at PW6B95-D3:Amber level. [j]jun-ChS electronic energy, PW6B95-D3/jul-cc-pVDZ geometry, and ZPVE. [k]jun-ChS:PW6B95 electronic energy, PW6B95-D3/jul-cc-pVDZ geometry, and ZPVE.

3.4. Reaction Rates.
In an astrochemical context, HNC can either isomerize to HCN or diffuse on ice surfaces and then react with another molecule (e.g., CH₂NH to produce acetonitrile) at the low temperatures typical of the ISM.

The reaction rates computed for the HNC ⇌ HCN isomerization with the methodology described in Section 2 are shown in Figure 7. It is apparent that the rates computed for the HNC···(H₂O)₂ system [Figure 7, panels (c,d)] are very small irrespective of the inclusion or not of tunneling. The situation is completely different for the HNC···(H₂O)₂₀ model, where the rate not including tunneling (corresponding to the one used by Koch and coworkers[36]) remains very small at low temperatures (Figure 7b), but inclusion of tunneling (Figure 7a) permits an effective reaction even at temperatures
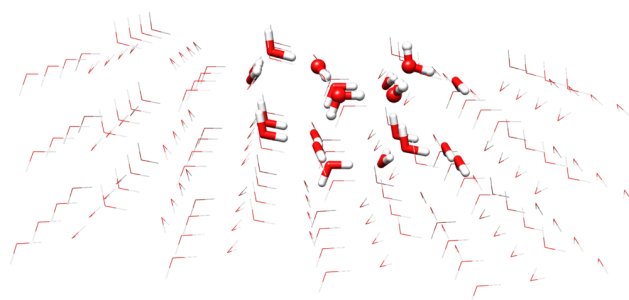
Figure 6. Structural model for the (H₂O)₁₉₂ cluster treated by three-layer ONIOM DSD-PBEP86:PW6B95:Amber strategy (geometry at PW6B95:Amber level). Ball and stick and tubular representation for the QM sections treated at the DSD-PBEP86/jul-cc-pVTZ and PW6B95-D3/jul-cc-pVDZ levels, respectively.
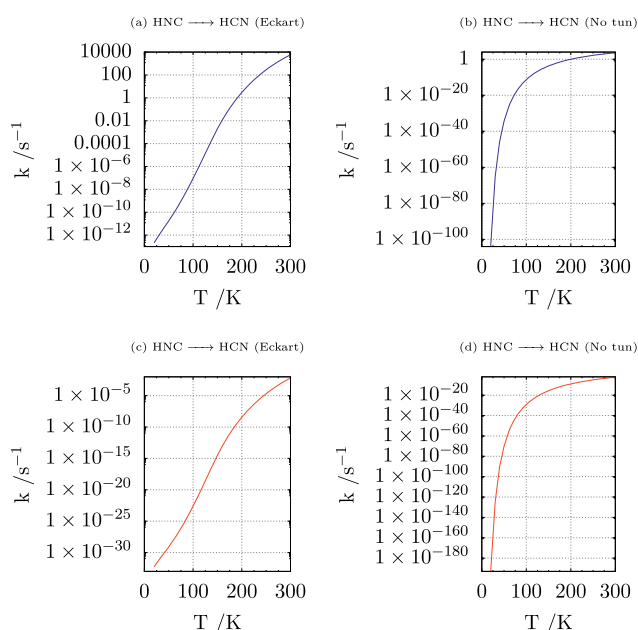
Figure 7. Reaction rates for the HNC ⇌ HCN isomerization including (Eckart) or excluding (no tun) tunneling. Panels (a,b) refer to the HNC···(H₂O)₂₀ model, whereas panels (c,d) refer to the HNC···(H₂O)₂ model.

characteristic of the ISM. Noted is that the rates computed taking tunneling into account show a clear bimodal shape and cannot be fitted by a simple Arrhenius (or Kooij) function.[71,72]

Unfortunately, the diffusion coefficients of HNC (or even HCN) on ice have not yet been reported.[76] According to a recent classification of ice adsorbates,[77] HCN (hence probably HNC) is assigned to the intermediate class, which induces some deformation of the surface, but does not form hydrates nor penetrates rapidly into the ice bulk. An upper limit to the surface diffusion coefficient can be estimated with reference to the guess of 4×10^{-11} cm² s⁻¹ at 130 K provided by Livingston et al. for SO₂,[78] which corresponds to a mean distance of 1260 Å in 1 s. Because the computed isomerization rate at 130 K is about 10^{-4} s⁻¹ (which lowers to 1×10^{-10} s⁻¹ at 50 K), the average diffusion of HNC before isomerization can reach 100 Å at 130 K (1 cm at 50 K).

Therefore, if the formation of aminoacetonitrile is faster than the isomerization to HCN when HNC and CH₂NH are nearest neighbors,[34] our results suggest that diffusion of HNC

along significant distances could permit the formation of aminoacetonitrile on icy grains containing CH_2NH even at low concentrations.

4. CONCLUSIONS AND OUTLOOK

The main aim of this work was the implementation and validation of a general computational strategy for the study of the thermochemistry and kinetics of chemical processes taking place on interstellar icy-grains. To this end, composite methods rooted in the coupled cluster ansatz have been combined with hybrid and double hybrid functionals together with molecular mechanics force field to characterize the stationary points ruling the reactive potential energy surfaces on model clusters sufficiently large to minimize spurious boundary effects. Next powerful master equation/TST models have been employed to compute reaction rates including tunneling effects. As a demanding test case, we have selected the HCN/HNC reactions for which the available computational results are not fully satisfactory.

Ten different (meta-)hybrid and double-hybrid density functionals have been considered in conjunction with the jun-, jul-, and aug-cc-pVnZ basis sets of double- and triple-ζ quality, and their accuracy in predicting geometries together with thermochemical and kinetic data (adsorption, activation, and reaction energies) has been assessed in comparison to reference values computed using the jun-ChS composite method. This benchmark has led to the conclusion that, among (meta-)hybrid functionals, BMK-D3 and PW6B95-D3 in conjunction with partially augmented double- and triple-ζ basis sets yield the most reliable description of geometries, with the optimal trade-off between accuracy and computational cost being offered by the PW6B95-D3/jul-cc-pVDZ model chemistry. Concerning double-hybrids, DSD-PBEP86-D3 and revDSD-PBEP86-D3 in conjunction with the jul-cc-pVTZ basis set deliver accurate predictions of both geometries and reaction energies. Next, these outcomes have been used to investigate the effect of cluster size and ice surface on the isomerization process of HCN. In particular, a cluster containing 20 water molecules has been cut from the (010) surface of ice XI and used in a multiscale ONIOM calculation, in which the reaction center has been modeled at the DSD-PBEP86-D3/jul-cc-pVTZ level, while for the remaining portion of the $(H_2O)_{20}$ cluster, the PW6B95-D3 functional has been employed in conjunction with the jul-cc-pVDZ basis set. This approach has allowed the proper modeling of the surface with an accurate yet cost-effective strategy. The pivotal role of the structural arrangement of surface molecules in driving the evolution of catalytic processes has been pointed out. The accuracy of the results has been further improved by combining jun-ChS results for small models to QM/QM′ (DSD-PBEP86:PW6B95-D3) values for medium-size model clusters and/or three-layer QM/QM′/MM computations for very large clusters.

On top of these computations, reaction rates have been computed by methods rooted in the transition state theory including tunneling which plays the dominant role at low temperature for processes involving the motion of light atoms. At variance with previous investigations, our results show that the isomerization is ruled by a proton relay mechanism directly involving four water molecules, but tuned by relatively distant waters belonging to the model cluster employed to mimic the ice surface. The resulting activation energy is strongly reduced with respect to that governing the isomerization of the bare HCN molecule, but only tunneling allows for effective isomerization of HNC in the harsh conditions characterizing astrochemical processes.

Together with the intrinsic interest of the studied system, the results of the present work have allowed to define the best strategy for future modeling of iCOMs-ices interactions in the framework of a QM/QM′/MM approach. This also represents the starting point for hybrid QM/QM′/periodic approaches, in which the outcome of the multiscale (QM/QM′) description of the cluster is corrected for environmental effects obtained by simulating the surface using periodic boundary conditions.[79] However, the crystalline water ice surfaces usually employed to simulate icy dust grains could be inadequate to describe their amorphous structure. Work in this and related directions is under way in our laboratory in order to achieve a more realistic modeling of chemical processes occurring on icy mantles of interstellar grains.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.1c01252.

> Main structural parameters for the water dimer obtained at the ChS and jun-ChS levels; error analysis for the structural parameters of the $HCN\cdots(H_2O)_2$ system; contributions to jun-ChS electronic energies; error analysis for the formation energy of the species involved in HNC ⇌ HCN isomerization assisted by the $(H_2O)_2$ cluster; relative ground-state electronic energies for the stationary points on the HCN ⇌ HNC isomerization PES with respect to isolated HCN and $(H_2O)_n$ for n = 2, 20; Cartesian coordinates of the stationary points on the HCN ⇌ HNC isomerization PES on a $(H_2O)_n$ cluster for n = 2, 3, and 4 optimized at the DSDPBEP86-D3/jul-cc-pVTZ level of theory; Cartesian coordinates of the stationary points on HCN ⇌ HNC isomerization PES catalyzed by the $(H_2O)_{20}$ cluster optimized at the DSDPBEP86-D3/jul-cc-pVTZ:PW6B95-D3/jul-cc-pVDZ level of theory (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Vincenzo Barone — *Scuola Normale Superiore, I-56126 Pisa, Italy;* ⓞ orcid.org/0000-0001-6420-4107; Email: vincenzo.barone@sns.it

Nicola Tasinato — *Scuola Normale Superiore, I-56126 Pisa, Italy;* ⓞ orcid.org/0000-0003-1755-7238; Email: nicola.tasinato@sns.it

Authors

Carmen Baiano — *Scuola Normale Superiore, I-56126 Pisa, Italy*

Jacopo Lupi — *Scuola Normale Superiore, I-56126 Pisa, Italy;* ⓞ orcid.org/0000-0001-6522-9947

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jctc.1c01252

Notes

The authors declare no competing financial interest.

Journal of Chemical Theory and Computation pubs.acs.org/JCTC Article

■ REFERENCES

(1) McGuire, B. A. 2018 Census of Interstellar, Circumstellar, Extragalactic, Protoplanetary Disk, and Exoplanetary Molecules. *Astrophys. J. Suppl.* **2018**, *239*, 17.

(2) van Dishoeck, E. F. Astrochemistry: overview and challenges. *Proc. Int. Astron. Union* **2017**, *13*, 3.

(3) Puzzarini, C. Grand Challenges in Astrochemistry. *Front. Astron. Space Sci.* **2020**, *7*, 19.

(4) Garrod, R. T.; Weaver, S. L. W.; Herbst, E. Complex Chemistry in Star-forming Regions: An Expanded Gas-Grain Warm-up Chemical Model. *Astrophys. J.* **2008**, *682*, 283.

(5) Herbst, E.; van Dishoeck, E. F. Complex Organic Interstellar Molecules. *Annu. Rev. Astron. Astrophys.* **2009**, *47*, 427.

(6) Hollenbach, D. J.; Werner, M. W.; Salpeter, E. E. Molecular Hydrogen in H I Regions. *Astrophys. J.* **1971**, *163*, 165.

(7) Wakelam, V.; Bron, E.; Cazaux, S.; Dulieu, F.; Gry, C.; Guillard, P.; Habart, E.; Hornekær, L.; Morisset, S.; Nyman, G.; Pirronello, V.; Price, S. D.; Valdivia, V.; Vidali, G.; Watanabe, N. H_2 formation on interstellar dust grains: The viewpoints of theory, experiments, models and observations. *Mol. Astrophys.* **2017**, *9*, 1.

(8) Allamandola, L. J.; Bernstein, M. P.; Sandford, S. A.; Walker, R. L. Evolution of Interstellar Ices. *Composition and Origin of Cometary Materials*; Springer Science & Business Media: Dordrecht, 1999; pp 219−232.

(9) Burke, D. J.; Brown, W. A. Ice in space: surface science investigations of the thermal desorption of model interstellar ices on dust grain analogue surfaces. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5947.

(10) Watanabe, N.; Kouchi, A. Ice surface reactions: A key to chemical evolution in space. *Prog. Surf. Sci.* **2008**, *83*, 439.

(11) Hama, T.; Watanabe, N. Surface Processes on Interstellar Amorphous Solid Water: Adsorption, Diffusion, Tunneling Reactions, and Nuclear-Spin Conversion. *Chem. Rev.* **2013**, *113*, 8783.

(12) Boogert, A. C. A.; Gerakines, P. A.; Whittet, D. C. B. Observations of the Icy Universe. *Annu. Rev. Astron. Astrophys.* **2015**, *53*, 541.

(13) Cuppen, H. M.; Walsh, C.; Lamberts, T.; Semenov, D.; Garrod, R. T.; Penteado, E. M.; Ioppolo, S. Grain Surface Models and Data for Astrochemistry. *Space Sci. Rev.* **2017**, *212*, 1.

(14) Tasinato, N.; Ceselin, G.; Stoppa, P.; Pietropolli Charmet, A.; Giorgianni, S. A Bit of Sugar on TiO_2: Quantum Chemical Insights on the Interfacial Interaction of Glycolaldehyde over Titanium Dioxide. *J. Phys. Chem. C* **2018**, *122*, 6041.

(15) Rimola, A.; Ferrero, S.; Germain, A.; Corno, M.; Ugliengo, P. Computational Surface Modelling of Ices and Minerals of Interstellar Interest—Insights and Perspectives. *Minerals* **2021**, *11*, 26.

(16) Barone, V.; Lupi, J.; Salta, Z.; Tasinato, N. Development and Validation of a Parameter-Free Model Chemistry for the Computation of Reliable Reaction Rates. *J. Chem. Theory Comput.* **2021**, *17*, 4913.

(17) Ehrenfreund, P.; Schutte, W. A. ISO observations of interstellar ices: Implications for the pristinity of comets. *Adv. Space Res.* **2000**, *25*, 2177.

(18) Gibb, E. L.; Whittet, D. C. B.; Boogert, A. C. A.; Tielens, A. G. M. Interstellar Ice: The Infrared Space Observatory Legacy. *Astrophys. J. Suppl.* **2004**, *151*, 35.

(19) Maheshwary, S.; Patel, N.; Sathyamurthy, N.; Kulkarni, A. D.; Gadre, S. R. Structure and Stability of Water Clusters $(H_2O)_n$, n = 8—20: An Ab Initio Investigation. *J. Phys. Chem. A* **2001**, *105*, 10525.

(20) Rimola, A.; Sodupe, M.; Ugliengo, P. Deep-space glycine formation via Strecker-type reactions activated by ice water dust mantles. A computational approach. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5285.

(21) Rimola, A.; Skouteris, D.; Balucani, N.; Ceccarelli, C.; Enrique-Romero, J.; Taquet, V.; Ugliengo, P. Can Formamide Be Formed on Interstellar Ice? An Atomistic Perspective. *ACS Earth Space Chem.* **2018**, *2*, 720.

(22) Rimola, A.; Sodupe, M.; Ugliengo, P. Computational study of interstellar glycine formation occurring at radical surfaces of water-ice dust particles. *Astrophys. J.* **2012**, *754*, 24.

(23) Shimonishi, T.; Nakatani, N.; Furuya, K.; Hama, T. Adsorption Energies of Carbon, Nitrogen, and Oxygen Atoms on the Low-temperature Amorphous Water Ice: A Systematic Estimation from Quantum Chemistry Calculations. *Astrophys. J.* **2018**, *855*, 27.

(24) Zamirri, L.; Casassa, S.; Rimola, A.; Segado-Centellas, M.; Ceccarelli, C.; Ugliengo, P. IR spectral fingerprint of carbon monoxide in interstellar water—ice models. *Mon. Not. R. Astron. Soc.* **2018**, *480*, 1427.

(25) Sameera, W. M. C.; Senevirathne, B.; Andersson, S.; Maseras, F.; Nyman, G. ONIOM(QM:AMOEBA09) Study on Binding Energies and Binding Preference of OH, HCO, and CH3 Radicals on Hexagonal Water Ice (Ih). *J. Phys. Chem. C* **2017**, *121*, 15223.

(26) Sameera, W. M. C.; Senevirathne, B.; Andersson, S.; Al-lbadi, M.; Hidaka, H.; Kouchi, A.; Nyman, G.; Watanabe, N. CH_3O Radical Binding on Hexagonal Water Ice and Amorphous Solid Water. *J. Phys. Chem. A* **2021**, *125*, 387.

(27) Duflot, D.; Toubin, C.; Monnerville, M. Theoretical Determination of Binding Energies of Small Molecules on Interstellar Ice Surfaces. *Front. Astron. Space Sci.* **2021**, *8*, 645243.

(28) Puzzarini, C.; Bloino, J.; Tasinato, N.; Barone, V. Accuracy and interpretability: The devil and the holy grail. New routes across old boundaries in computational spectroscopy. *Chem. Rev.* **2019**, *119*, 8131.

(29) Enrique-Romero, J.; Rimola, A.; Ceccarelli, C.; Ugliengo, P.; Balucani, N.; Skouteris, D. Reactivity of HCO with CH_3 and NH_2 on Water Ice Surfaces. A Comprehensive Accurate Quantum Chemistry Study. *ACS Earth Space Chem.* **2019**, *3*, 2158.

(30) Spada, L.; Tasinato, N.; Bosi, G.; Vazart, F.; Barone, V.; Puzzarini, C. On the competition between weak OH···F and CH···F hydrogen bonds, in cooperation with CH···O contacts, in the difluoromethane − tert-butyl alcohol cluster. *J. Mol. Spectrosc.* **2017**, *337*, 90.

(31) Puzzarini, C.; Salta, Z.; Tasinato, N.; Lupi, J.; Cavallotti, C.; Barone, V. A twist on the reaction of the CN radical with methylamine in the interstellar medium: new hints from a state-of-the-art quantum-chemical study. *Mon. Not. R. Astron. Soc.* **2020**, *496*, 4298.

(32) Alessandrini, S.; Barone, V.; Puzzarini, C. Extension of the "Cheap" Composite Approach to Noncovalent Interactions: The jun-ChS Scheme. *J. Chem. Theory Comput.* **2020**, *16*, 988.

(33) Woon, D. E. Ab Initio Quantum Chemical Studies of Reactions in Astrophysical Ices 3. Reactions of $HOCH_2NH_2$ Formed in H_2CO/NH_3/H_2O Ices. *J. Phys. Chem. A* **2001**, *105*, 9478.

(34) Koch, D. M.; Toubin, C.; Peslherbe, G. H.; Hynes, J. T. A Theoretical Study of the Formation of the Aminoacetonitrile Precursor of Glycine on Icy Grain Mantles in the Interstellar Medium. *J. Phys. Chem. C* **2008**, *112*, 2972.

(35) Gardebien, F.; Sevin, A. Catalytic Model Reactions for the HCN Isomerization. I. Theoretical Characterization of Some Water-Catalyzed Mechanisms. *J. Phys. Chem. A* **2003**, *107*, 3925.

(36) Koch, D. M.; Toubin, C.; Xu, S.; Peslherbe, G. H.; Hynes, J. T. Concerted Proton-Transfer Mechanism and Solvation Effects in the HNC/HCN Isomerization on the Surface of Icy Grain Mantles in the Interstellar Medium. *J. Phys. Chem. C* **2007**, *111*, 15026.

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

(38) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1988**, *37*, 785.

100

(39) Vosko, S. H.; Wilk, L.; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.* **1980**, *58*, 1200.

(40) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615.

(41) Zhao, Y.; Truhlar, D. G. Design of Density Functionals That Are Broadly Accurate for Thermochemistry, Thermochemical Kinetics, and Nonbonded Interactions. *J. Phys. Chem. A* **2005**, *109*, 5656.

(42) Boese, A. D.; Martin, J. M. L. Development of density functionals for thermochemical kinetics. *J. Chem. Phys.* **2004**, *121*, 3405.

(43) Zhao, Y.; Truhlar, D. G. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, non-covalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06 functionals and 12 other functionals. *Theor. Chem. Acc.* **2008**, *120*, 215.

(44) Yu, H. S.; He, X.; Li, S. L.; Truhlar, D. G. MN15: A Kohn–Sham global-hybrid exchange–correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. *Chem. Sci.* **2016**, *7*, 5032.

(45) Grimme, S. Semiempirical hybrid density functional with perturbative second-order correlation. *J. Chem. Phys.* **2006**, *124*, 034108.

(46) Kozuch, S.; Martin, J. M. L. DSD-PBEP86: in search of the best double-hybrid DFT with spin-component scaled MP2 and dispersion corrections. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20104.

(47) Kozuch, S.; Martin, J. M. L. Spin-component-scaled double hybrids: An extensive search for the best fifth-rung functionals blending DFT and perturbation theory. *J. Comput. Chem.* **2013**, *34*, 2327.

(48) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007.

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

(50) Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. Perspectives on Basis Sets Beautiful: Seasonal Plantings of Diffuse Basis Functions. *J. Chem. Theory Comput.* **2011**, *7*, 3027.

(51) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **2010**, *132*, 154104.

(52) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.* **2011**, *32*, 1456.

(53) Smith, D. G. A.; Burns, L. A.; Patkowski, K.; Sherrill, C. D. Revised Damping Parameters for the D3 Dispersion Correction to Density Functional Theory. *J. Phys. Chem. Lett.* **2016**, *7*, 2197.

(54) Burns, L. A.; Mayagoitia, Á. V.; Sumpter, B. G.; Sherrill, C. D. Density-functional approaches to noncovalent interactions: A comparison of dispersion corrections (DFT-D), exchange-hole dipole moment (XDM) theory, and specialized functionals. *J. Chem. Phys.* **2011**, *134*, 084107.

(55) Klimeš, J.; Michaelides, A. Perspective: Advances and challenges in treating van der Waals dispersion forces in density functional theory. *J. Chem. Phys.* **2012**, *137*, 120901.

(56) Tasinato, N.; Grimme, S. Unveiling the non-covalent interactions of molecular homodimers by dispersion-corrected DFT calculations and collision-induced broadening of ro-vibrational transitions: Application to $(CH_2F_2)_2$ and $(SO_2)_2$. *Phys. Chem. Chem. Phys.* **2015**, *17*, 5659.

(57) Goerigk, L.; Kruse, H.; Grimme, S. Benchmarking Density Functional Methods against the S66 and S66x8 Datasets for Non-Covalent Interactions. *Phys. Chem. Chem. Phys.* **2011**, *12*, 3421.

(58) Piane, M. D.; Corno, M.; Ugliengo, P. Does Dispersion Dominate over H-Bonds in Drug-Surface Interactions? The Case of Silica-Based Materials As Excipients and Drug-Delivery Agents. *J. Chem. Theory Comput.* **2013**, *9*, 2404.

(59) Tasinato, N.; Moro, D.; Stoppa, P.; Pietropolli Charmet, A.; Toninello, P.; Giorgianni, S. Adsorption of F_2CCFCl on TiO_2 nano-powder: Structures, energetics and vibrational properties from DRIFT spectroscopy and periodic quantum chemical calculations. *Appl. Surf. Sci.* **2015**, *353*, 986.

(60) Puzzarini, C.; Barone, V. Extending the molecular size in accurate quantum-chemical calculations: the equilibrium structure and spectroscopic properties of uracil. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7189.

(61) Puzzarini, C.; Biczysko, M.; Barone, V.; Peña, I.; Cabezas, C.; Alonso, J. L. Accurate molecular structure and spectroscopic properties of nucleobases: a combined computational–microwave investigation of 2-thiouracil as a case study. *Phys. Chem. Chem. Phys.* **2013**, *15*, 16965.

(62) Lane, J. R. CCSDTQ Optimized Geometry of Water Dimer. *J. Chem. Theory Comput.* **2013**, *9*, 316.

(63) Frisch, M. J. et al. *Gaussian 16*, Revision C. 01; Gaussian Inc.: Wallingford, CT, 2016.

(64) Stanton, J. F.; Gauss, J.; Harding, M. E.; Szalay, P. G. CFOUR. A quantum chemical program package; with contributions from A. A. Auer, R. J. Bartlett, U. Benedikt, C. Berger, D. E. Bernholdt, Y. J. Bomble, O. Christiansen, F. Engel, M. Heckert, O. Heun, C. Huber, T.-C. Jagau, D. Jonsson, J. Jusélius, K. Klein, W. J. Lauderdale, F. Lipparini, D. Matthews, T. Metzroth, L. A. Mück, D. P. O'Neill, D. R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, S. Stopkowicz, A. Tajti, J. Vázquez, F. Wang, J. D. Watts and the integral packages MOLECULE (J. Almlöf and P. R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jørgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wüllen, 2016. For the current version, see http://www.cfour.de.

(65) Matthews, D. A.; Cheng, L.; Harding, M. E.; Lipparini, F.; Stopkowicz, S.; Jagau, T.-C.; Szalay, P. G.; Gauss, J.; Stanton, J. F. Coupled-cluster techniques for computational chemistry: The CFOUR program package. *J. Chem. Phys.* **2020**, *152*, 214108.

(66) Vreven, T.; Morokuma, K. Chapter 3 Hybrid Methods: ONIOM(QM:MM) and QM/MM. *Annu. Rep. Comput. Chem.* **2006**, *2*, 35.

(67) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179.

(68) Georgievskii, Y.; Miller, J. A.; Burke, M. P.; Klippenstein, S. J. Reformulation and solution of the master equation for multiple-well chemical reactions. *J. Phys. Chem. A* **2013**, *117*, 12146.

(69) Fernández-Ramos, A.; Miller, J. A.; Klippenstein, S. J.; Truhlar, D. G. Modeling the Kinetics of Bimolecular Reactions. *Chem. Rev.* **2006**, *106*, 4518.

(70) Eckart, C. The penetration of a potential barrier by electrons. *Phys. Rev.* **1930**, *35*, 1303.

(71) Kooij, D. M. Über die Zersetzung des gasförmigen Phosphorwasserstoffs. *Z. Phys. Chem.* **1893**, *12*, 155.

(72) Laidler, K. J. A glossary of terms used in chemical kinetics, including reaction dynamics (IUPAC Recommendations 1996). *Pure Appl. Chem.* **1996**, *68*, 149.

(73) Zheng, J.; Zhao, Y.; Truhlar, D. G. The DBH24/08 Database and Its Use to Assess Electronic Structure Model Chemistries for Chemical Reaction Barrier Heights. *J. Chem. Theory Comput.* **2009**, *5*, 808.

(74) Barone, V.; Ceselin, G.; Fusè, M.; Tasinato, N. Accuracy Meets Interpretability for Computational Spectroscopy by Means of Hybrid and Double-Hybrid Functionals. *Front. Chem.* **2020**, *8*, 584203–584211.

(75) Ceselin, G.; Barone, V.; Tasinato, N. Accurate Biomolecular Structures by the Nano-LEGO Approach: Pick the Bricks and Build Your Geometry. *J. Chem. Theory Comput.* **2021**, *17*, 7290.

(76) Huthwelker, T.; Ammann, M.; Peter, T. The uptake of acidic gases on ice. *Chem. Rev.* **2006**, *106*, 1375.

(77) Devlin, J. P.; Buch, V. Vibrational Spectroscopy and Modeling of the Surface and Subsurface of Ice and of Ice Adsorbate Interactions. *J. Phys. Chem. B* **1997**, *101*, 6095.

(78) Livingston, F. E.; Smith, J. A.; George, S. M. General Trends for Bulk Diffusion in Ice and Surface Diffusion on Ice. *J. Phys. Chem. A* **2002**, *106*, 6309.

(79) Sauer, J. Ab Initio Calculations for Molecule−Surface Interactions with Chemical Accuracy. *Acc. Chem. Res.* **2019**, *52*, 3502.

Chapter 7

Conclusions and outlook

In this dissertation, novel composite schemes for the calculation of accurate thermo-chemistry and kinetics of gas-phase reactions have been developed and presented. The starting point of all these schemes is the use of spin-component scaled double hybrid density functionals including dispersion corrections (namely revDSD-PBEP86-D3(BJ)) for the calculation of molecular geometries and vibrational frequencies, in conjunction with basis sets of triple-ζ quality including diffuse functions. It has been demonstrated that this model chemistry can efficiently evaluate precise zero-point energies and thermal contributions. This approach permits to obtain geometrical parameters with a reasonably good accuracy in most situations. If improved accuracy is sought, the most effective option is to add the core-valence correlation correction, evaluated at the MP2-F12 level, to CCSD(T)-F12/jun-cc-pVTZ geometries. On top of these, the so called "cheap" schemes are employed in order to obtain accurate electronic energies. All the variants of this scheme are rooted in the coupled cluster ansatz, to which complete basis set and core-valence corrections are added, both evaluated at the MP2 level of theory, with the aim of reducing computational cost without sacrificing accuracy. Indeed the aforementioned schemes have been subsequently benchmarked against a large panel of reference reaction barriers and molecular structures. The results show that the average error is within $1.25\,\mathrm{kJ\,mol^{-1}}$. If more accurate electronic energies are required, the starting molecular geometries are those obtained at the F12 level, as detailed before. On top of these, CCSD(T)-F12 single point computations in conjunction with augmented basis sets of quadruple-ζ

quality are already at the CBS limit. Core correlation can be easily computed at the CCSD(T)-F12 level with the cc-pwCVTZ basis set. Full triple and perturbative quadruple corrections can be computed with triple-ζ and double-ζ quality basis sets, respectively. Finally, in order to push the electronic structure computations to the HEAT-like limit, diagonal Born-Oppenheimer corrections and scalar relativistic contributions are needed, the former computed at the HF-SCF/aug-cc-pVDZ level and the latter at the CCSD(T)/aug-cc-pCVDZ level. It has been shown that the resulting composite scheme can represent an accurate reference for the benchmarking of reaction barriers.

Once the electronic energies have been obtained, they have to be inserted along with the geometries and the vibrational frequencies into master equation models, within the so called ab initio transition-state theory master equation approach. In order to prove the robustness of this approach, some challenging reactions of atmospherical and astrophysical interest have been chosen and carefully investigated with the developed methods. In particular, H_2S oxidation by Cl radical and CH_2NH reaction with CN, C_2H radicals have been reported and discussed. The resulting systems' reactive potential energy surfaces and reaction rates are among the most precise currently available (especially for H_2S + Cl). Moreover, CH_2NH computations have permitted to propose a general model for the formation of more complex imines in the interstellar medium. The main conclusion is that models based on the transition state theory can accurately represent the reaction if the underlying electronic structure computations are accurate enough and barrierless channels are adequately characterized. Last but not least, the isomerization of HCN on icy grain surfaces has been investigated. The accuracy of several DFT methods for predicting structural and energetic properties of the system has been analysed, finding out that among (meta-)hybrid functionals, BMK-D3(BJ) and PW6B95-D3(BJ) in conjunction with partially augmented double- and triple-ζ basis sets yield the most reliable description of geometries, with the optimal trade-off between accuracy and computational cost being offered by the PW6B95-D3(BJ)/jul-cc-pVDZ model chemistry. Concerning double-hybrid functionals, DSD-PBEP86-D3(BJ) and revDSD-PBEP86-D3(BJ) in conjunction with the jul-cc-pVTZ basis set deliver accurate predictions for both

geometries and reaction energies.

The objectives outlined in the Introduction have been reached satisfactorily. Clearly the proposed computational strategy for the investigation of gas-phase reactions can be further improved. As far as electronic structure calculations are concerned, it would be desirable to apply composite schemes to larger molecular systems, both covalent and non-covalent ones. One possible way to go in this direction is to use local-correlation treatments (e.g., PNO or DLPNO possibly including explicit correlation).[74,75] Regarding kinetics calculations, the automation of the use of more sophisticated theories is necessary, also in view to insert them into black-box codes that can be routinely used. In particular the accurate description of the barrierless channels with VRC-TST (in a DFT formulation) and the calculation of anharmonic densities of the states,[76] along with the implementation of semi-classical tunnelling,[77] could represent possible ways of improving the protocol.

Undergoing work In this final paragraph, some of the undergoing works' preliminary results are collected and discussed. These results should be intended as partial and subject to future refinements.

HTBH38/08 and NHTBH38/08 databases by Truhlar's group have been further investigated with junChS and junChS-F12 composite schemes in order to check the role of MP2 CBS extrapolation contribution. To do this, a new reference energy was defined as "Best" (already defined in Chapter 1), which surpasses in accuracy the old ones. As can be seen in Figure 7-1, junChS-F12 outperforms junChS when MP2-F12 and MP2 CBS corrections are excluded, because the MP2 extrapolation is essential in order to overcome the basis set incompleteness error of conventional CCSD(T). This means that CCSD(T)-F12 employing a triple ζ quality basis set is already close to the CBS limit. Various contributions to the total Best energy are reported in Tables 7.1 and 7.2.

Moreover, the different performances of three composite schemes, namely CBS-QB3, junChS-F12 and Best, in computing reaction rates within a RRKM-ME approach have been investigated. In particular two single step reactions, i.e. H + PH$_3$ $\longrightarrow$ PH$_2$ + H$_2$ (HT11) and CH$_4$ + OH $\longrightarrow$ CH$_3$ + H$_2$O (HT4), two isomer-
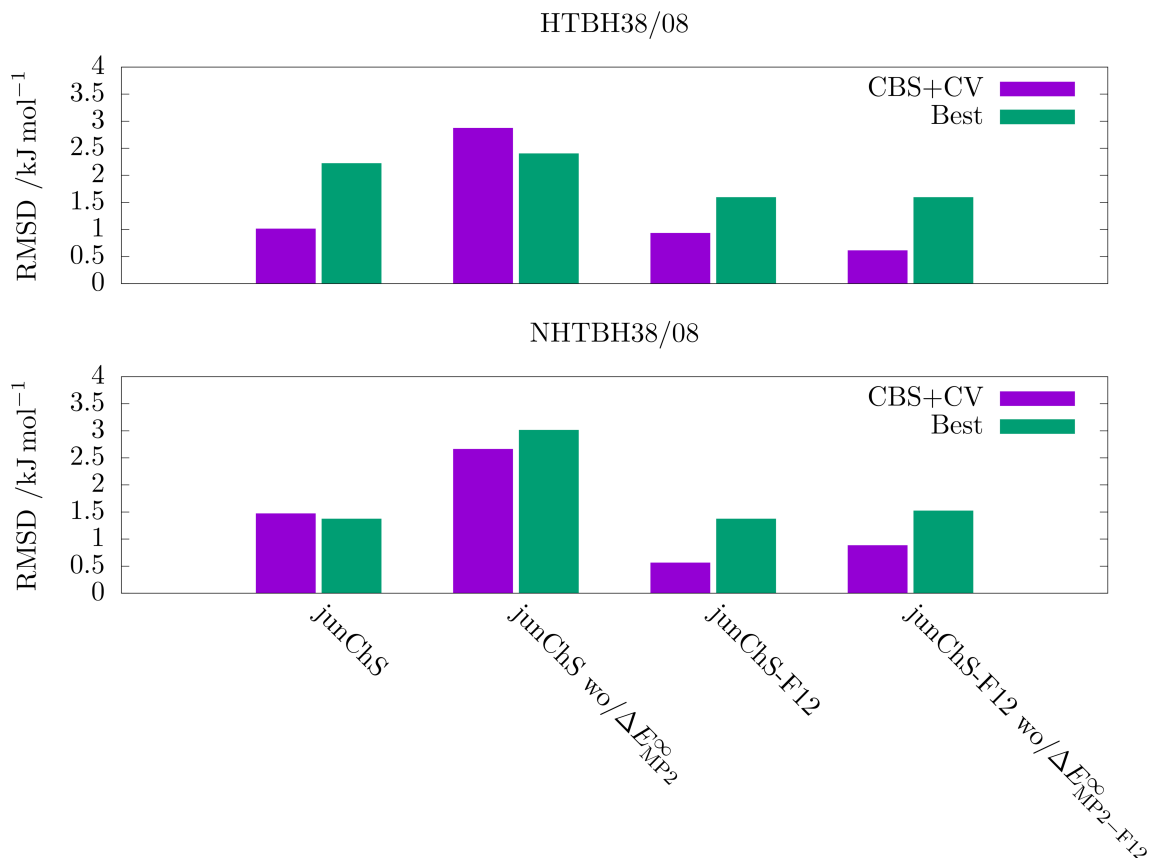
Figure 7-1: Root-mean-square deviations of HTBH38/08 and NHTBH38/08 databases' reactions calculated with different cheap schemes, both with and without MP2 CBS correction.

ization reactions, i.e. HCN $\longrightarrow$ HNC (NHT19) and Cl$^-\cdots$CH$_3$Cl $\longrightarrow$ ClCH$_3\cdots$Cl$^-$ (NHT10), from HTBH38/08 and NHTBH38/08 datasets have been selected. In order to take into account also the effect of entrance and exit van der Waals wells, two multistep reactions of astrophysical and atmospherical relevance have been investigated, i.e. cyano radical (CN) addition to ethylene (C$_2$H$_4$) and Criegee intermediate (CI) water reaction (CH$_2$OO + H$_2$O). In Figure 7-2 the high-pressure limit rate constants of HT11 and HT4 are collected. For HT11 both junChS-F12 and Best schemes are in good accordance with the temperature trend of experimental results,[78,79] while CBS-QB3 tends to overestimate the rate at high temperatures, while underestimates it at lower ones. For HT4 reaction, the accordance between the schemes is worse, likely because of the hindered rotation present in the transition state. In Figure 7-3 high-pressure limit rate constant of NHT19 and reported. Both reactions are characterized by quite high energy barriers, and their rates show

Table 7.1: Theoretical values of the barrier heights in the HTBH38/08 dataset obtained at different levels of theory. All the values are in $kJ\,mol^{-1}$.

		ΔZPE-H	ΔGEOM	CV-F12ᵃ	fT	pQ	DBOC	rel	ΔSO
HT1	H• + HCl ⟶ H₂ + Cl•	-3.31/-11.97	-0.04/1.63	0.10/-0.30 (0.13/-0.32)	-0.31/0.24	-0.08/-0.20	1.60/0.95	-0.53/0.83	0.00/3.51
HT2	OH• + H₂ ⟶ H₂O + H•	2.09/-5.10	0.00/0.00	-0.01/0.80 (-0.03/0.83)	-0.53/-1.08	-0.33/0.29	0.22/1.15	0.03/-0.56	0.84/0.00
HT3	CH₃• + H₂ ⟶ CH₄ + H•	7.03/-5.73	0.00/0.00	-0.04/0.60 (-0.06/0.61)	-0.25/-0.41	-0.14/-0.05	1.19/1.88	0.00/-0.09	-
HT4	OH• + CH₄ ⟶ CH₃• + H₂O	-6.36/-0.75	0.04/0.04	0.25/0.43 (0.26/0.45)	-0.07/-0.46	-0.66/-0.13	0.41/0.65	0.04/-0.47	0.84/0.00
HT5	H• + H₂ ⟶ H₂ + H•	-3.05/-3.05	-0.08/-0.08	0.00/0.00 (0.00/0.00)	-0.31/-0.31	-0.80/-0.80	1.71/1.71	-0.02/-0.02	-
HT6	OH• + NH₃ ⟶ H₂O + NH₂•	-2.55/3.60	-0.17/0.75	0.33/-0.05 (0.36/-0.02)	-0.89/-0.96	-1.22/-0.82	1.67/1.54	-0.02/-0.20	0.84/0.00
HT7	HCl + CH₃• ⟶ Cl• + CH₄	4.06/-17.36	-0.04/1.67	0.01/0.26 (0.11/0.33)	0.01/0.41	-0.31/-0.34	0.29/0.33	-0.04/1.23	0.00/3.51
HT8	OH• + C₂H₆ ⟶ H₂O + C₂H₅•	-6.65/-0.42	0.04/0.04	0.16/0.60 (0.17/0.62)	-0.24/-0.44	-0.62/-0.13	-0.33/0.27	-0.18/-0.69	0.84/0.00
HT9	F• + H₂ ⟶ HF + H•	-2.80/-0.84	0.00/0.04	0.07/0.67 (0.05/0.70)	-0.52/-0.90	-0.19/0.61	0.11/0.90	0.01/-0.77	-
HT10	³O + CH₄ ⟶ OH• + CH₃•	-13.81/3.10	-0.38/-1.26	0.43/0.27 (0.41/0.26)	-0.22/-0.06	-0.22/0.16	0.14/-0.05	0.07/-0.33	0.92/0.84
HT11	H• + PH₃ ⟶ PH₂• + H₂	-1.63/-0.33	-0.04/-0.04	-0.06/-0.3 (-0.04/-0.42)	-0.47/-0.37	-0.08/-0.21	0.86/0.20	-0.25/0.54	-
HT12	H• + OH• ⟶ H₂ + ³O	-3.47/-7.61	-0.59/0.33	0.45/-0.03 (0.45/-0.07)	-0.62/-0.61	0.05/-0.42	1.46/0.95	-0.40/0.09	0.84/0.92
HT13	H• + H₂S ⟶ H₂ + HS•	-2.43/-5.19	0.00/0.00	0.00/-0.41 (0.02/-0.45)	-0.38/-0.13	-0.10/-0.32	0.95/0.29	-0.32/0.69	0.00/2.26
HT14	³O + HCl ⟶ OH• + Cl•	-8.37/-12.84	-0.67/0.13	0.08/0.16 (0.37/0.44)	-2.37/-1.82	-1.26/-0.91	1.44/1.31	-0.43/0.44	0.92/4.35
HT15	NH₂• + CH₃• ⟶ CH₄ + NH	1.51/-7.32	-2.47/-1.17	0.36/0.26 (0.36/0.25)	0.04/-0.01	-0.30/-0.59	2.03/2.10	-0.19/0.07	-
HT16	NH₂• + C₂H₅• ⟶ NH + C₂H₆	0.71/-8.79	-2.55/-1.30	0.48/0.11 (0.48/0.10)	0.32/0.09	-0.41/-0.66	1.14/0.84	-0.20/0.06	-
HT17	NH₂• + C₂H₆ ⟶ NH₃ + C₂H₅•	-3.35/-3.26	0.00/0.04	0.06/0.89 (0.06/0.88)	0.46/0.32	-0.62/-0.53	0.72/1.46	0.03/-0.30	-
HT18	NH₂• + CH₄ ⟶ NH₃ + CH₃•	-1.46/-2.05	0.00/0.00	0.14/0.70 (0.13/0.69)	0.49/0.18	-0.57/-0.45	0.80/1.18	0.05/-0.29	-
HT19	s-trans cis−C₅H₈ ⟶ same	-9.46/-9.46	0.08/0.08	0.61/0.61 (0.62/0.62)	0.37/0.37	-1.68/-1.68	0.33/0.33	-0.03/-0.03	-

ᵃF12a (F12b)

Table 7.2: Theoretical values of the barrier heights in the NHTBH38/08 dataset obtained at different levels of theory. All the values are in $kJ\,mol^{-1}$.

		ΔZPE-H	ΔZPE(A-H)	ΔGEOM	CV-F12ᵃ	fT	pQ	DBOC	rel	ΔSO
NHT1	H• + N₂O ⟶ OH• + N₂	4.06/-3.68	-0.75/-0.75	0.17/-0.88	0.62/0.14 (0.60/0.11)	-0.61/1.91	-0.78/-5.10	1.13/0.39	-0.27/0.44	0.00/0.84
NHT2	H• + FH ⟶ HF + H•	-7.36/-7.36	-1.26/-1.26	0.08/0.08	0.52/0.52 (0.53/0.53)	-0.12/-0.12	-0.36/-0.36	2.29/2.29	-0.69/-0.69	-
NHT3	H• + ClH ⟶ HCl + H•	-4.10/-4.10	8.58/8.58	-0.08/-0.08	0.24/0.24 (0.24/0.24)	-0.38/-0.38	-0.17/-0.17	1.46/1.46	-0.83/-0.83	-
NHT4	H• + FCH₃ ⟶ HF + CH₃•	-1.30/-0.67	-1.84/-2.47	-0.08/0.00	0.58/1.34 (0.54/1.32)	-0.69/-0.31	-0.74/-0.86	0.68/0.17	-0.66/-0.69	-
NHT5	H• + F₂ ⟶ HF + F•	4.44/-14.27	-1.34/-1.21	0.84/0.84	0.09/1.64 (0.07/1.60)	-0.21/0.34	0.27/-2.81	0.74/0.34	-0.14/-1.01	-
NHT6	CH₃• + FCl ⟶ CH₃F + Cl•	8.24/-12.30	0.88/-0.08	0.00/0.00	0.34/1.12 (0.26/0.97)	-0.12/-0.04	-0.94/-1.69	0.50/0.61	-0.33/-0.70	0.00/3.51
NHT7	F⁻ + CH₃F ⟶ FCH₃ + F⁻	-0.75/-0.75	-0.42/-0.42	0.00/0.00	1.47/1.47 (1.51/1.51)	-0.47/-0.47	-0.60/-0.60	0.04/0.04	-0.23/-0.23	-
NHT8	F⁻···CH₃F ⟶ FCH₃···F⁻	-1.76/-1.76	-0.25/-0.25	0.00/0.00	1.12/1.12 (1.13/1.13)	-0.32/-0.32	-0.46/-0.46	0.06/0.06	-0.23/-0.23	-
NHT9	Cl⁻ + CH₃Cl ⟶ ClCH₃ + Cl⁻	-1.42/-1.42	0.00/0.00	0.00/0.00	1.19/1.19 (1.37/1.37)	-0.56/-0.56	-0.53/-0.53	0.00/0.00	-0.24/-0.24	-
NHT10	Cl⁻···CH₃Cl ⟶ ClCH₃···Cl⁻	-2.18/-2.18	0.13/0.13	0.00/0.00	1.07/1.07 (1.19/1.19)	-0.53/-0.53	-0.46/-0.46	0.03/0.03	-0.50/-0.50	-
NHT11	F⁻ + CH₃Cl ⟶ FCH₃ + Cl⁻	0.08/-3.89	-0.17/-0.08	0.00/0.00	1.44/0.76 (1.49/0.92)	-0.63/-0.46	-0.37/-0.93	0.02/-0.01	-0.20/-0.10	-
NHT12	F⁻···CH₃Cl ⟶ FCH₃···Cl⁻	-0.46/-4.98	0.04/0.00	0.00/0.00	0.98/0.63 (1.00/0.74)	-0.37/-0.44	0.25/-0.36	0.01/0.04	-0.18/-0.35	-
NHT13	OH⁻ + CH₃F ⟶ HOCH₃ + F⁻	3.51/-5.65	-0.38/0.00	0.00/0.00	1.29/1.96 (1.32/1.99)	-0.27/-0.52	-1.04/-0.92	0.01/0.12	-0.14/-0.48	-
NHT14	OH⁻···CH₃F ⟶ HOCH₃···F⁻	0.29/-3.64	-0.25/1.76	0.00/0.00	1.06/1.76 (1.07/1.76)	-0.22/-0.65	-0.86/-0.60	0.06/0.14	-0.17/-0.45	-
NHT15	H• + N₂ ⟶ HN₂•	3.97/-18.07	-0.38/0.13	0.00/-0.04	0.26/0.43 (0.24/0.38)	-0.51/0.73	1.01/-0.06	1.34/0.44	0.10/-0.38	-
NHT16	H• + CO ⟶ HCO•	2.05/-19.37	-0.17/0.42	0.00/0.00	0.06/1.02 (0.07/1.03)	-0.43/-0.09	-0.16/-0.10	0.78/-0.10	0.03/-0.43	-
NHT17	H• + C₂H₄ ⟶ C₂H₅•	5.44/-17.11	-0.33/-0.25	0.00/0.17	0.19/0.06 (0.21/0.03)	-0.50/0.16	-0.03/-0.76	0.75/-0.02	-0.02/-0.18	-
NHT18	CH₃• + C₂H₄ ⟶ CH₃CH₂CH₂•	10.79/-9.96	-0.79/-0.08	-0.13/-0.08	0.55/0.59 (0.59/0.56)	-0.82/-1.05	-0.49/-0.80	0.03/-0.01	0.02/-0.21	-
NHT19	HCN ⟶ HNC	-14.02/-12.84	0.08/0.08	0.13/0.13	1.64/0.80 (1.67/0.80)	-0.64/0.17	0.39/-0.80	0.11/0.20	-0.27/-0.43	-

ᵃF12a (F12b)

a clear Arrhenius behavior in the medium to high temperature range. Going down in temperature, in the HCN system tunneling effect starts dominating and therefore the trend deviates from the Arrhenius one. A similar trend is seen for NHT10, with CBS-QB3 greatly mismatching junChS-F12 and Best. The last examples are the reactive potential energy surfaces of C_2H_4 + CN and CH_2OO + H_2O, collected in Figure 7-4. Regarding the former, the CN addition rate constant is essentially flat in the whole temperature range, while for the latter, CBS-QB3 deviates from the junChS-F12 trend. Best calculations are still undergoing, but they are expected to behave similarly to junChS-F12. Additional hints about the reliability of junChS-F12 scheme with respect to CBS-QB3 come from the Arrhenius-Kooij fit parameters, which are reported in Table 7.3.

(a) $k(T)$ H + PH$_3$ (b) $k(T)$ CH$_4$ + OH
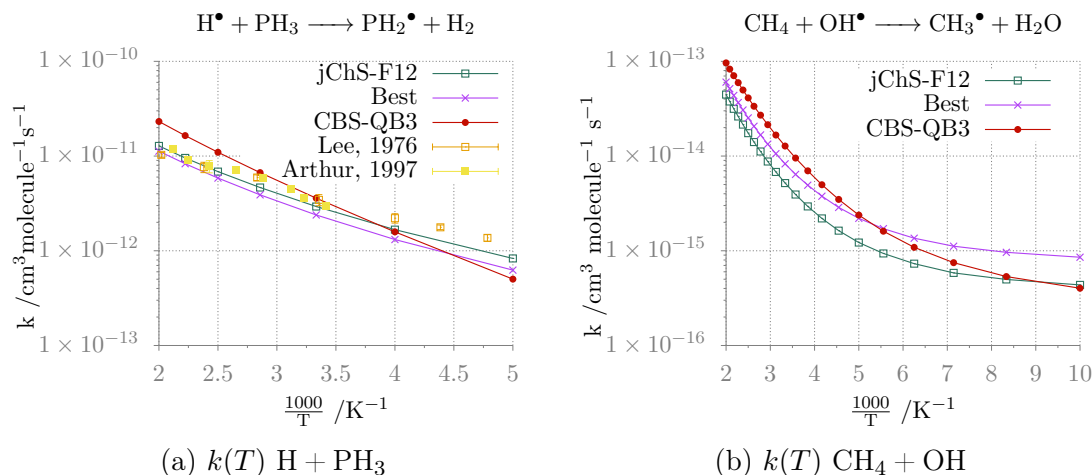
Figure 7-2: High-pressure limit rate constant for bimolecular one-step reactions.

(a) $k(T)$ HCN $\longrightarrow$ HNC   (b) $k(T)$ Cl$^-$···CH$_3$Cl $\longrightarrow$ ClCH$_3$···Cl$^-$
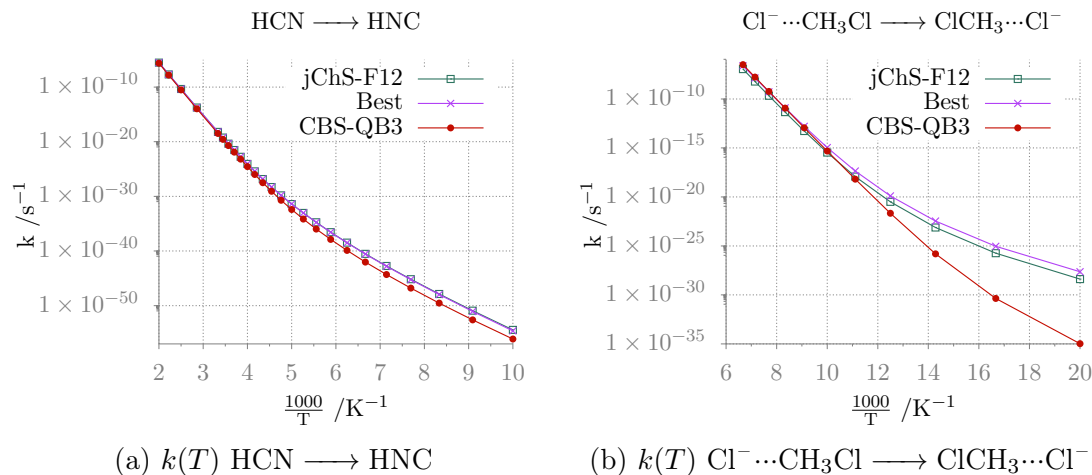
Figure 7-3: High-pressure limit rate constant for unimolecular reactions.

Lastly, development efforts of new composite schemes are currently undergoing. Indeed, it would be desirable to have a composite scheme where one can dispense with extrapolating to CBS. Thus, the idea is to build "special" basis sets combining $n + 1$ s, p functions with n or even $n - 1$ d, f functions. In this way, the s, p space functions can be considered already at the CBS convergence (considering $n = 4$), while polarization ones (d, f) can be pushed towards CBS limit exploiting the F12 convergence. Geometries of some selected molecules taken from the recent SE100 database[80] by Ceselin *et al.* and some others works, as detailed in Table 7.4, have been computed by means of composite schemes employing these modified basis sets. The core of these composite schemes is the same as the ones developed in the disser-
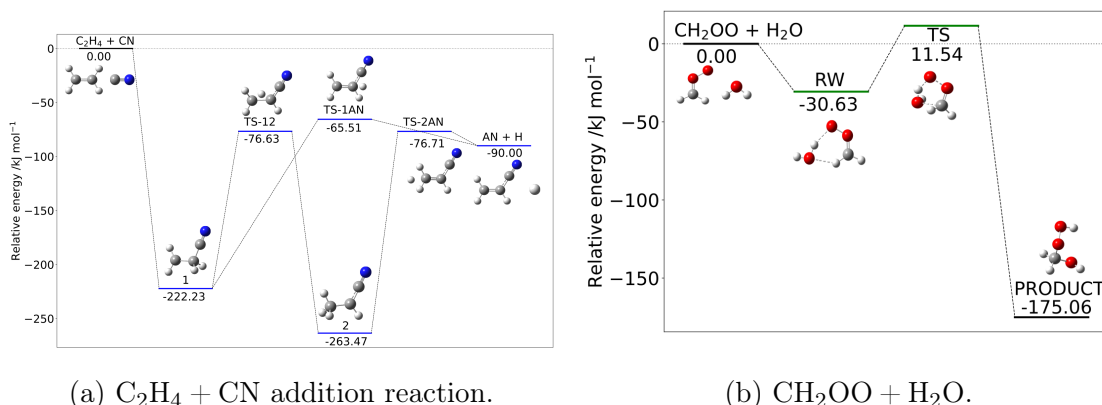
(a) $C_2H_4 + CN$ addition reaction.

(b) $CH_2OO + H_2O$.

Figure 7-4: Reactive potential energy surfaces. Electronic energies at junChS-F12 level of theory augmented by revDSD-PBEP86-D3(BJ) anharmonic ZPE.
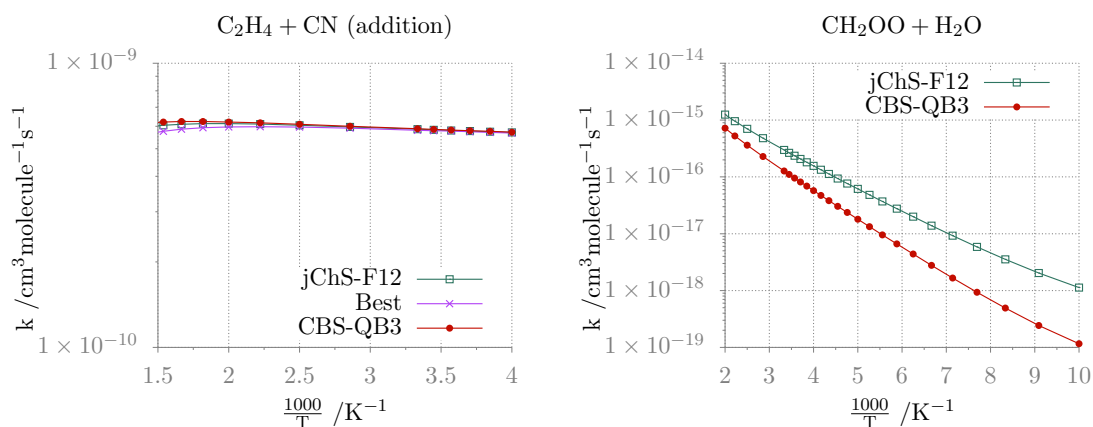
(a) $k(T)$ $C_2H_4 + CN$ addition reaction.

(b) $k(T)$ $CH_2OO + H_2O$ reaction.

Figure 7-5: Rate constants for bimolecular multiwell reactions. Pressure is set at 1 atm.
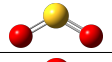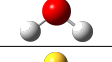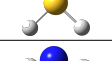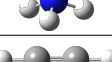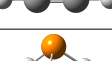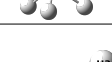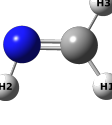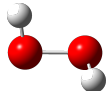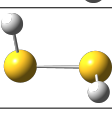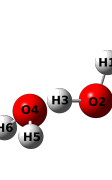
tation, i.e. a CCSD(T)-F12b term and a core-valence additive correction calculated at the MP2-F12 level. Now, the difference is that jun-cc-pVTZ and cc-pwCVTZ basis sets are replaced by their modified versions. In particular, in the (CC+CV)-F12mod scheme, "mod" appendix means that jun-cc-pVnZ and cc-pwCVnZ basis sets are built combining s, p functions from the $(n + 1)$ basis sets with d, f functions from the n ones. Preliminary results are reported in Table 7.4. (CC+CV)-F12 and (CC+CV)-F12mod geometries are very close to each other. In order to have common reference schemes, CCSD(T)-F12b/cc-pVDZ-F12 and CCSD(T)-F12b/cc-pVTZ-F12 including CV at MP2-F12/cc-pwCVTZ geometries have been calculated and indicated as VDZ-F12+CV and VTZ-F12+CV, along with junChS ones. The

Table 7.3: Arrhenius-Kooij parameters of selected reactions.

		junChS-F12	Best	CBS-QB3
$H + PH_3 \longrightarrow PH_2 + H_2$	A /cm^3 molecule^{-1} s^{-1}	9.06×10^{-12}	8.96×10^{-12}	4.84×10^{-11}
	n	2.02	2.03	1.60
	E / kJ mol^{-1}	2.85	3.34	6.47
$CH_4 + OH \longrightarrow CH_3 + H_2O$	A /cm^3 molecule^{-1} s^{-1}	3.76×10^{-14}	3.77×10^{-14}	8.19×10^{-14}
	n	2.86	2.86	2.56
	E / kJ mol^{-1}	5.31	4.07	4.74
$HCN \longrightarrow HNC$	A /s^{-1}	1.13×10^{14}	8.61×10^{13}	1.78×10^{13}
	n	6.42×10^{-2}	1.93×10^{-1}	1.01
	E / kJ mol^{-1}	1.88×10^{2}	1.87×10^{2}	1.82×10^{2}
$Cl^- \cdots CH_3Cl \longrightarrow ClCH_3 \cdots Cl^-$	A /s^{-1}	1.19×10^{11}	1.19×10^{11}	2.34×10^{11}
	n	1.05	1.06	7.56×10^{-1}
	E / kJ mol^{-1}	5.10×10^{1}	5.00×10^{1}	5.07×10^{1}
$C_2H_4 + CN$	A /cm^3 molecule^{-1} s^{-1}	5.89×10^{-10}	5.86×10^{-10}	5.90×10^{-10}
	n	8.55×10^{-2}	5.05×10^{-2}	1.05×10^{-1}
	E / kJ mol^{-1}	5.45×10^{-2}	7.73×10^{-2}	4.18×10^{-2}
$CH_2OO + H_2O$	A /cm^3 molecule^{-1} s^{-1}	2.83×10^{-15}	-	2.36×10^{-15}
	n	1.05	-	1.12
	E / kJ mol^{-1}	5.63	-	7.30

good agreement with semi-experimental geometries confirms the reliability of the proposed (CC+CV)-F12mod scheme, which is essentially the same of the standard (CC+CV)-F12.

Table 7.4: Geometries of selected molecules from SE100 database.[80]All distances in Å, all angles in degrees. All reference geometries are taken from Ref. 80 except CH_2NH[81], H_2S[82], H_2S_2[83], H_2O_2[84] and $H_2O\cdots H_2O$.[85]

			SE	(CC+CV)-F12mod	(CC+CV)-F12	VDZ-F12+CV[a]	VTZ-F12+CV	junChS
SO_2		r(S-O)	1.4308	1.4305	1.4305	1.4305 (1.4332)	1.4303	1.4299
		θ(S-O-S)	119.31	119.34	119.35	119.41 (119.40)	119.34	119.29
H_2O		r(O-H)	0.9573	0.9573	0.9576	0.9576 (0.9585)	0.9577	0.9563
		θ(H-O-H)	104.53	104.54	104.57	104.47 (104.34)	104.53	104.49
H_2S		r(S-H)	1.3356	1.3351	1.3353	1.3343 (1.3365)	1.3359	1.3356
		θ(H-S-H)	92.11	92.24	92.28	92.13 (92.19)	92.22	92.21
NH_3		r(N-H)	1.0110	1.0106	1.0108	1.0109 (1.0122)	1.0110	1.0100
		θ(H-N-H)	106.94	106.82	106.79	106.77 (106.58)	106.79	106.82
C_2H_2		r(C-H)	1.0617	1.0618	1.0620	1.0620 (1.0633)	1.0619	1.0617
		r(C-C)	1.2030	1.2031	1.2032	1.2033 (1.2061)	1.2027	1.2023
PH_3		r(P-H)	1.4117	1.4110	1.4113	1.4109 (1.4140)	1.4118	1.4116
		θ(H-P-H)	93.42	93.43	93.46	93.36 (93.48)	93.41	93.42
CH_2NH		r(H1-C)	1.0919	1.0898	1.0899	1.0900 (1.0915)	1.0900	1.0894
		r(C-N)	1.2709	1.2710	1.2711	1.2713 (1.2741)	1.2708	1.2702
		r(N-H2)	1.0195	1.0185	1.0186	1.0187 (1.0200)	1.0188	1.0177
		r(C-H3)	1.0839	1.0856	1.0856	1.0857 (1.0872)	1.0858	1.0852
		θ(H1-C-N)	123.72	124.29	124.28	124.30 (124.32)	124.28	124.29
		θ(C-N-H2)	110.35	110.29	110.31	110.14 (109.98)	110.30	110.34
		θ(N-C-H3)	119.25	118.70	118.70	118.70 (118.67)	118.68	118.71
BH_2OH		r(O-B)	1.3498	1.3502	1.3502	1.3501 (1.3534)	1.3499	1.3485
		r(O-H1)	0.9558	0.9579	0.9582	0.9581 (0.9590)	0.9582	0.9567
		r(B-H2)	1.1957	1.1941	1.1942	1.1943 (1.1967)	1.1943	1.1938
		r(B-H3)	1.1899	1.1886	1.1886	1.1887 (1.1911)	1.1887	1.1882
		θ(B-O-H1)	112.90	112.97	112.98	112.86 (112.73)	113.00	112.96
		θ(H2-B-O)	119.80	120.41	120.42	120.42 (120.41)	120.41	120.42
		θ(H3-B-O)	-	116.82	116.82	116.84 (116.80)	116.80	116.82
H_2O_2		r(O-O)	1.4524	1.4480	1.4480	1.4496 (1.4519)	1.4488	1.4482
		r(O-H)	0.9617	0.9623	0.9626	0.9625 (0.9634)	0.9626	0.9611
		θ(H-O-H)	99.76	100.17	100.18	100.12 (100.04)	100.17	100.22
		φ(H-O-O-H)	113.6	113.22	113.29	113.28 (113.10)	112.85	112.08
H_2S_2		r(S-S)	2.0513	2.0518	2.0525	2.0505 (2.0548)	2.0515	2.0518
		r(S-H)	1.3403	1.3391	1.3394	1.3383 (1.3405)	1.3399	1.3396
		θ(H-S-H)	98.13	98.14	98.16	98.14 (98.16)	98.15	98.14
		φ(H-S-S-H)	90.72	90.78	90.70	90.77 (90.72)	90.64	90.66
$H_2O\cdots H_2O$		r(H1-O2)	0.9569	0.9565	0.9568	0.9568 (0.9576)	0.9569	0.9554
		r(H3-O2)	0.9641	0.9637	0.9640	0.9638 (0.9646)	0.9641	0.9626
		r(O2$\cdots$O4)	2.9092	2.9109	2.9077	2.9138 (2.9170)	2.9099	2.9056
		r(H5/H6-O4)	0.9584	0.9581	0.9584	0.9584 (0.9592)	0.9585	0.9570
		θ(H1-O2-H3)	104.85	104.85	104.88	104.75 (104.62)	104.86	104.82
		θ(H5-O4-H6)	104.95	104.94	104.89	104.80 (104.78)	104.86	104.78
		θ(O4-O2-H3)	5.69	5.24	5.19	5.01 (5.10)	5.59	5.93
		φ(H6-O4-O2-H1)	123.46	122.12	122.26	122.36 (122.53)	122.69	123.00

[a]In parentheses values without CV contribution.

Bibliography

[1] J. Lupi, S. Alessandrini, C. Puzzarini, and V. Barone. junChS and junChS-F12 models: parameter-free efficient yet accurate composite schemes for energies and structures of noncovalent complexes. *J. Chem. Theory Comput.*, 17(11): 6974–6992, 2021.

[2] J. Lupi, C. Puzzarini, and V. Barone. Methanimine as a key precursor of imines in the interstellar medium: the case of propargylimine. *Astrophys. J. Lett.*, 903 (2):L35, 2020.

[3] J. Lupi, C. Puzzarini, C. Cavallotti, and V. Barone. State-of-the-art quantum chemistry meets variable reaction coordinate transition state theory to solve the puzzling case of the $H_2S + Cl$ system. *J. Chem. Theory Comput.*, 16(8): 5090–5104, 2020.

[4] J. Lupi, M. Martino, A. Salvadori, S. Rampino, G. Mancini, and V. Barone. Virtual reality tools for advanced modeling. *AIP Conf. Proc.*, 2145(1):020001, 2019.

[5] C. Baiano, J. Lupi, N. Tasinato, and V. Barone. Gliding on ice in search of accurate and cost-effective computational methods for astrochemistry on grains: the puzzling case of the HCN isomerization. *J. Chem. Theory Comput.*, XX (XX):XXXX, 2022.

[6] V. Barone, J. Lupi, Z. Salta, and N. Tasinato. Development and validation of a parameter-free model chemistry for the computation of reliable reaction rates. *J. Chem. Theory Comput.*, 17(8):4913–4928, 2021.

[7] F. Tonolo, J. Lupi, C. Puzzarini, and V. Barone. The quest for a plausible formation route of formyl cyanide in the interstellar medium: a state-of-the-art quantum-chemical and kinetic approach. *Astrophys. J.*, 900(1):85, 2020.

[8] C. Baiano, J. Lupi, N. Tasinato, C. Puzzarini, and V. Barone. The role of state-of-the-art quantum-chemical calculations in astrochemistry: formation route and spectroscopy of ethanimine as a paradigmatic case. *Molecules*, 25(12): 2873, 2020.

[9] C. Puzzarini, Z. Salta, N. Tasinato, J. Lupi, C. Cavallotti, and V. Barone. A twist on the reaction of the CN radical with methylamine in the interstellar medium: new hints from a state-of-the-art quantum-chemical study. *Mon. Not. R. Astron. Soc.*, 496(4):4298–4310, 2020.

[10] Z. Salta, N. Tasinato, J. Lupi, R. Boussessi, A. Balbi, C. Puzzarini, and V. Barone. Exploring the maze of $C_2N_2H_5$ radicals and their fragments in the interstellar medium with the help of quantum-chemical computations. *ACS Earth Space Chem.*, 4(5):774–782, 2020.

[11] Z. Salta, J. Lupi, V. Barone, and O. N. Ventura. H-abstraction from dimethyl sulfide in the presence of an excess of hydroxyl radicals. A quantum chemical evaluation of thermochemical and kinetic parameters unveils an alternative pathway to dimethyl sulfoxide. *ACS Earth Space Chem.*, 4(3):403–419, 2020.

[12] Z. Salta, J. Lupi, N. Tasinato, V. Barone, and O. N. Ventura. Unraveling the role of additional OH-radicals in the H–abstraction from dimethyl sulfide using quantum chemical computations. *Chem. Phys. Lett.*, 739:136963, 2020.

[13] G. Ritchie. *Atmospheric chemistry: from the surface to the stratosphere*. World Scientific Publishing Company, 2017.

[14] S. C. Althorpe and D. C. Clary. Quantum scattering calculations on chemical reactions. *Ann. Rev. Phys. Chem.*, 54(1):493–529, 2003.

[15] M. H. Beck, A. Jäckle, G. A. Worth, and H.-D. Meyer. The multiconfiguration time-dependent Hartree (MCTDH) method: a highly efficient algorithm for propagating wavepackets. *Phys. Rep.*, 324(1):1–105, 2000.

[16] I. R. Craig and D. E. Manolopoulos. Chemical reaction rates from ring polymer molecular dynamics. *The Journal of Chemical Physics*, 122(8):084106, 2005.

[17] L. M. Raff and D. L. Thompson. *Theory of chemical reaction dynamics*, volume III, pages 1–122. M. Baer, CRC Press, 1985.

[18] G. H. Peslherbe, H. Wang, and W. L. Hase. *Monte Carlo sampling for classical trajectory simulations*, pages 171–201. John Wiley & Sons, Ltd, 1999.

[19] P.-O. Löwdin. *Correlation problem in many-electron quantum mechanics I. Review of different approaches and discussion of some current ideas*, pages 207–322. John Wiley & Sons, Ltd, 1958.

[20] A. Karton. A computational chemist's guide to accurate thermochemistry for organic molecules. *WIRES Comput. Mol. Sci.*, 6:292–310, 2016.

[21] A. Karton and J. M. L. Martin. Explicitly correlated Wn theory: W1-F12 and W2-F12. *J. Chem. Phys.*, 136:124114, 2012.

[22] J. Laane, M. Dakkouri, B. van der Veken, and H. Oberhammer. *Structures and conformations of non-rigid molecules*, volume 410. Springer Netherlands, 1993.

[23] A. G. Császár, W. D. Allen, and H. F. Schaefer. In pursuit of the ab initio limit for conformational energy prototypes. *J. Chem. Phys.*, 108:9751–9764, 1998.

[24] K. A. Peterson, D. Feller, and D. A. Dixon. Chemical accuracy in ab initio thermochemistry and spectroscopy: current strategies and future challenges. *Theor. Chem. Acc.*, 131:1–50, 2012.

[25] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton. HEAT: high accuracy extrapolated ab initio thermochemistry. *J. Chem. Phys.*, 121:11599–11613, 2004.

[26] Y. J. Bomble, J. Vázquez, M. Kállay, C. Michauk, P. G. Szalay, A. G. Császár, J. Gauss, and J. F. Stanton. High-accuracy extrapolated ab initio thermochemistry. II. Minor improvements to the protocol and a vital simplification. *J. Chem. Phys.*, 125:064108, 2006.

[27] M. E. Harding, J. Vázquez, B. Ruscic, A. K. Wilson, J. Gauss, and J. F. Stanton. High-accuracy extrapolated ab initio thermochemistry. III. Additional improvements and overview. *J. Chem. Phys.*, 128:114111, 2008.

[28] J. Zhang and E. F. Valeev. Prediction of reaction barriers and thermochemical properties with explicitly correlated coupled-cluster methods: a basis set assessment. *J. Chem. Theory Comput.*, 8:3175–3186, 2012.

[29] L. Kong, F. A. Bischoff, and E. F. Valeev. Explicitly correlated R12/F12 methods for electronic structure. *Chem. Rev.*, 112:75–107, 2012.

[30] L. A. Curtiss, P. C. Redfern, and K. Raghavachari. Gn theory. *WIRES Comput. Mol. Sci.*, 1:810–825, 2011.

[31] L. A. Curtiss, P. C. Redfern, and K. Raghavachari. Gaussian-4 theory. *J. Chem. Phys.*, 126(8):084108, 2007.

[32] J. A. Montgomery, J. W. Ochterski, and G.A. Petersson. A complete basis set model chemistry. IV. An improved atomic pair natural orbital method. *J. Chem. Phys.*, 101:5900–5909, 1994.

[33] J. A. Montgomery Jr., M. J. Frisch, J. W. Ochterski, and G. A. Petersson. A complete basis set model chemistry. VII. Use of the minimum population localization method. *J. Chem. Phys.*, 112:6532–6542, 2000.

[34] L. A. Curtiss, P. C. Redfern, and K. Raghavachari. Assessment of Gaussian-4 theory for energy barriers. *Chem. Phys. Lett.*, 499:168–172, 2010.

[35] H. Eyring. The activated complex in chemical reactions. *J. Chem. Phys.*, 3(2): 107–115, 1935.

[36] M. G. Evans and M. Polanyi. Some applications of the transition state method to the calculation of reaction velocities, especially in solution. *Trans. Faraday Soc.*, 31:875–894, 1935.

[37] A. D. McNaught and A. Wilkinson. *IUPAC. Compendium of chemical terminology*. Blackwell Scientific Publications, Oxford, 1997.

[38] M. J. Pilling and P. W. Seakins. *Reaction kinetics*. Oxford University Press, 1996.

[39] J. L. Bao and D. G. Truhlar. Variational transition state theory: theoretical framework and recent developments. *Chem. Soc. Rev.*, 46:7548–7596, 2017.

[40] P. Pechukas and J. C. Light. On detailed balancing and statistical theories of chemical kinetics. *J. Chem. Phys.*, 42(9):3281–3291, 1965.

[41] W. J. Chesnavich. Multiple transition states in unimolecular reactions. *J. Chem. Phys.*, 84(5):2615–2619, 1986.

[42] S. J. Klippenstein. Variational optimizations in the Rice–Ramsperger–Kassel–Marcus theory calculations for unimolecular dissociations with no reverse barrier. *J. Chem. Phys.*, 96(1):367–371, 1992.

[43] Y. Georgievskii and S. J. Klippenstein. Variable reaction coordinate transition state theory: analytic results and application to the C_2H_3 + H $\longrightarrow$ C_2H_4 reaction. *J. Chem. Phys.*, 118(12):5442–5455, 2003.

[44] Y. Georgievskii and S. J. Klippenstein. Transition state theory for multichannel addition reactions: multifaceted dividing surfaces. *J. Phys. Chem. A*, 107(46): 9776–9781, 2003.

[45] T. Helgaker, P. Jørgensen, and J. Olsen. *Molecular electronic-structure theory.* John Wiley & Sons, 2000.

[46] A. Szabo and N. S. Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory.* Dover Publications, 1996.

[47] R. G. Parr and W. Yang. *Density-functional theory of atoms and molecules.* Oxford University Press, 1989.

[48] F. Jensen. *Introduction to computational chemistry.* John Wiley & Sons, 2017.

[49] W. Forst. *Unimolecular reactions: a concise introduction.* Cambridge University Press, 2003.

[50] R. G. Gilbert and S. C. Smith. *Theory of unimolecular and recombination reactions.* Blackwell Scientific Publishing, 1990.

[51] T. Kato. On the eigenfunctions of many-particle systems in quantum mechanics. *Comm. Pure Appl. Math.*, 10(2):151–177, 1957.

[52] R. T. Pack and W. B. Brown. Cusp conditions for molecular wavefunctions. *J. Chem. Phys*, 45(2):556–559, 1966.

[53] O. Marchetti and H.-J. Werner. Accurate calculations of intermolecular interaction energies using explicitly correlated wave functions. *Phys. Chem. Chem. Phys.*, 10:3400–3409, 2008.

[54] H.-J. Werner, T. B. Adler, and F. R. Manby. General orbital invariant MP2-F12 theory. *J. Chem. Phys.*, 126(16):164102, 2007.

[55] G. Knizia, T. B. Adler, and H.-J. Werner. Simplified CCSD(T)-F12 methods: theory and benchmarks. *J. Chem. Phys.*, 130:054104, 2009.

[56] T. B. Adler, G. Knizia, and H. J. Werner. A simple and efficient CCSD(T)-F12 approximation. *J. Chem. Phys.*, 127:221106, 2007.

[57] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136: B864–B871, 1964.

[58] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

[59] J. P. Perdew and K. Schmidt. Jacob's ladder of density functional approximations for the exchange-correlation energy. *AIP Conf. Proc.*, 577(1):1–20, 2001.

[60] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.*, 132(15), 2010.

[61] D. G. A. Smith, L. A. Burns, K. Patkowski, and C. D. Sherrill. Revised damping parameters for the D3 dispersion correction to density functional theory. *J. Phys. Chem. Lett.*, 7(12):2197–2203, 2016.

[62] S. Grimme, S. Ehrlich, and L. Goerigk. Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.*, 32(7):1456–1465, 2011.

[63] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez, and J. F. Stanton. HEAT: high accuracy extrapolated ab initio thermochemistry. *J. Chem. Phys.*, 121:11599–11613, 2004.

[64] Y. J. Bomble, J. Vázquez, M. Kállay, C. Michauk, P. G. Szalay, A. G. Császár, J. Gauss, and J. F. Stanton. High-accuracy extrapolated ab initio thermochemistry. II. Minor improvements to the protocol and a vital simplification. *J. Chem. Phys.*, 125:064108, 2006.

[65] R. C. Tolman. *Statistical mechanics with applications to physics and chemistry*. The Chemical Catalog Company, Inc., 1927.

[66] D. G. Truhlar. Interpretation of the activation energy. *J. Chem. Educ.*, 55(5): 309, 1978.

[67] C. Eckart. The penetration of a potential barrier by electrons. *Phys. Rev.*, 35 (11):1303–1309, 1930.

[68] J. R. Barker. Multiple-well, multiple-path unimolecular reaction systems. I. MultiWell computer program suite. *Int. J. Chem. Kinet.*, 33(4):232–245, 2001.

[69] Y. Georgievskii, J. A. Miller, M. P. Burke, and S. J. Klippenstein. Reformulation and solution of the master equation for multiple-well chemical reactions. *J. Phys. Chem. A*, 117(46):12146–12154, 2013.

[70] D. C. Tardy and B. S. Rabinovitch. Collisional energy transfer. Thermal uni-molecular systems in the low-pressure region. *J. Chem. Phys.*, 45(10):3720–3730, 1966.

[71] S. J. Klippenstein and J. A. Miller. From the time-dependent, multiple-well master equation to phenomenological rate coefficients. *J. Phys. Chem. A*, 106 (40):9267–9277, 2002.

[72] J. A. Miller and S. J. Klippenstein. Master equation methods in gas phase chemical kinetics. *J. Phys. Chem. A*, 110(36):10528–10544, 2006.

[73] S. Alessandrini, V. Barone, and C. Puzzarini. Extension of the "cheap" composite approach to noncovalent interactions: the jun-ChS scheme. *J. Chem. Theory Comput.*, 16:988–1006, 2020.

[74] P. R. Nagy and M. Kallay. Approaching the basis set limit of CCSD(T) energies for large molecules with local natural orbital coupled-cluster methods. *J. Chem. Theory Comput.*, 15:5275–5298, 2019.

[75] D. G. Liakos, Y. Guo, and F. Neese. Comprehensive benchmark results for the domain based local pair natural orbital coupled cluster method (DLPNO-CCSD(T)) for closed- and open-shell systems. *J. Phys. Chem. A*, 124:90–100, 2020.

[76] T. L. Nguyen and J. R. Barker. Sums and densities of fully coupled anharmonic vibrational states: a comparison of three practical methods. *J. Phys. Chem. A*, 114(10):3718–3730, 2010.

[77] W. H. Miller, R. Hernandez, N. C. Handy, D. Jayatilaka, and A. Willets. Ab initio calculation of anharmonic constants for a transition state, with application to semiclassical transition state tunneling probabilities. *Chem. Phys. Lett.*, 172:62–68, 1990.

[78] J. H. Lee, J. V. Michael, W. A. Payne, D. A. Whytock, and L. J. Stief. Absolute rate constant for the reaction of atomic hydrogen with phosphine over the temperature range 209 to 495 K. *J. Chem. Phys.*, 65(8):3280–3283, 1976.

[79] N. L. Arthur and I. A. Cooper. Arrhenius parameters for the reactions of H atoms with PH_3 and AsH_3. *J. Chem. Soc., Faraday Trans.*, 93:521–524, 1997.

[80] G. Ceselin, V. Barone, and N. Tasinato. Accurate biomolecular structures by the nano-LEGO approach: pick the bricks and build your geometry. *J. Chem. Theory Comput.*, 17(11):7290–7311, 2021.

[81] L. Margulès, J. Demaison, P. B. Sreeja, and J.-C. Guillemin. Submillimeterwave spectrum of CH_2PH and equilibrium structures of CH_2PH and CH_2NH. *J. Mol. Spectrosc.*, 238(2):234–240, 2006.

[82] R. L. Cook, F. C. De Lucia, and P. Helminger. Molecular force field and structure of hydrogen sulfide: recent microwave results. *J. Mol. Struct.*, 28(2): 237–246, 1975.

[83] H. Ye, M. Mendolicchio, H. Kruse, C. Puzzarini, M. Biczysko, and V. Barone. The challenging equilibrium structure of HSSH: another success of the rotational spectroscopy/quantum chemistry synergism. *J. Mol. Struct.*, 1211: 127933, 2020.

[84] Joshua H. Baraban, P. Bryan Changala, and John F. Stanton. The equilibrium structure of hydrogen peroxide. *J. Mol. Struct.*, 343:92–95, 2018.

[85] J. Řezáč and P. Hobza. Describing noncovalent interactions beyond the common approximations: how accurate is the "gold standard," CCSD(T) at the complete basis set limit? *J. Chem. Theory Comput.*, 9(5):2151–2155, 2013.

[86] A. Császár, M. L. Leininger, and A. Burcat. Enthalpy of formation of HS. *J. Phys. Chem. A*, 107:2061–2065, 2003.

[87] J. D. Cox, G. D. Wagman, and D. A. Medvedev. *CODATA key values for thermodynamics*. Hemisphere Publishing Corp., New York, 1989.

[88] J. R. Lane. CCSDTQ optimized geometry of water dimer. *J. Chem. Theory Comput.*, 9(1):316, 2013.

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix A

Supporting data for Chapter 2

Table S1: junChS-F12 NCP-energies ($\text{kJ}\,\text{mol}^{-1}$): the various contributions for the A14 complexes.

	"ref"	CC/junTZ	ΔMP2$^\infty$/jun(T,Q)Z	MP2-CV/wCVTZ	Total	Rel. Error (%)	Abs. Error
$H_2O\cdots H_2O$	-21.0832	-21.4317	0.4631	-0.1384	-21.1070	0.11	-0.02
$NH_3\cdots NH_3$	-13.2131	-13.2529	0.1188	-0.0605	-13.1946	-0.14	0.02
$HF\cdots HF$	-19.2213	-19.6172	0.3155	-0.1038	-19.4055	0.96	-0.18
$CH_2O\cdots CH_2O$	-18.9284	-19.1210	0.0745	-0.0506	-19.0971	0.89	-0.17
$HCN\cdots HCN$	-19.9828	-20.1677	0.3029	-0.1230	-19.9878	0.02	0.00
$C_2H_4\cdots C_2H_4$	-4.6024	-4.5937	-0.0544	-0.0482	-4.6962	2.04	-0.09
$CH_4\cdots CH_4$	-2.23007	-2.1810	0.0298	-0.0086	-2.1598	-3.15	0.07
$H_2O\cdots NH_3$	-27.3759	-27.7949	0.4503	-0.1766	-27.5212	0.53	-0.15
$H_2O\cdots C_2H_4$	-10.7696	-10.8837	0.1783	-0.0994	-10.8049	0.33	-0.04
$C_2H_4\cdots CH_2O$	-6.79482	-6.8380	-0.0105	-0.0577	-6.9061	1.64	-0.11
$NH_3\cdots C_2H_4$	-5.78647	-5.8409	0.0541	-0.0532	-5.8399	0.92	-0.05
$HF\cdots CH_4$	-6.91615	-7.1018	0.1460	-0.0992	-7.0550	2.01	-0.14
$H_2O\cdots CH_4$	-2.8242	-2.8971	0.1713	-0.0261	-2.7518	-2.56	0.07
$NH_3\cdots CH_4$	-3.2175	-3.3536	0.1629	-0.0279	-3.2186	0.03	-0.001
MAE						0.26	-0.06

Table S2: jun-(d, f)H_ChS-F12 NCP-energies ($\text{kJ}\,\text{mol}^{-1}$): the various contributions for the A14 complexes.

	"ref"	CC/junTZ-dH	ΔMP2$^\infty$/jun(T,Q)Z-dH	MP2-CV/wCVTZ	Total	Rel. Error (%)	Abs. Error
$H_2O\cdots H_2O$	-21.0832	-21.4194	0.4617	-0.1384	-21.0961	0.06	0.01
$NH_3\cdots NH_3$	-13.2131	-13.2161	0.0941	-0.0605	-13.1826	0.23	-0.03
$HF\cdots HF$	-19.2213	-19.5602	0.2534	-0.1038	-19.4105	0.98	0.19
$HCN\cdots HCN$	-19.9828	-20.1691	0.3200	-0.1230	-19.9721	0.05	-0.01
$CH_4\cdots CH_4$	-2.2301	-2.1747	0.0325	-0.0086	-2.1508	3.56	-0.08
$CH_2O\cdots CH_2O$	-18.9284	-19.0957	0.0899	-0.0506	-19.0563	0.68	0.13
$C_2H_4\cdots C_2H_4$	-4.6024	-4.5783	-0.0628	-0.0482	-4.6893	1.89	0.09
$H_2O\cdots C_2H_4$	-10.7696	-10.8734	0.1665	-0.0993	-10.8062	0.34	0.04
$H_2O\cdots CH_4$	-2.8242	-2.8937	0.1691	-0.0265	-2.7512	2.59	-0.07
$H_2O\cdots NH_3$	-27.3759	-27.7545	0.4001	-0.1765	-27.5309	0.57	0.16
$NH_3\cdots CH_4$	-3.2175	-3.3580	0.1714	-0.0274	-3.2140	0.11	0.00
$NH_3\cdots C_2H_4$	-5.7865	-5.8372	0.0511	-0.0530	-5.8391	0.91	0.05
$HF\cdots CH_4$	-6.9162	-7.0291	0.0898	-0.0992	-7.0386	1.77	0.12
$C_2H_4\cdots CH_2O$	-6.7948	-6.8289	0.0016	-0.0567	-6.8840	1.31	0.09
MAE						1.08	0.05

Table S3: junCBS+CV-F12 NCP-energies (kJ mol^{-1}): the various contributions for the A14 complexes.

	"ref"	CC-CBS/jun(T,Q)Z	CC-CV/wCVTZ	Total	Rel. Error (%)	Abs. Error
$H_2O\cdots H_2O$	-21.0832	-21.1486	-0.1024	-21.2510	0.80	-0.17
$NH_3\cdots NH_3$	-13.2131	-13.2052	-0.0397	-13.2449	0.24	-0.03
$HF\cdots HF$	-19.2213	-19.4387	-0.0678	-19.5065	1.48	-0.29
$CH_2O\cdots CH_2O$	-18.9284	-19.0591	0.0476	-19.0116	0.44	-0.08
$HCN\cdots HCN$	-19.9828	-19.8982	-0.1201	-20.0183	0.18	-0.04
$C_2H_4\cdots C_2H_4$	-4.6024	-4.6129	-0.0007	-4.6136	0.24	-0.01
$CH_4\cdots CH_4$	-2.2301	-2.1778	0.0015	-2.1763	-2.41	0.05
$H_2O\cdots NH_3$	-27.3759	-27.5016	-0.1367	-27.6383	0.96	-0.26
$H_2O\cdots C_2H_4$	-10.7696	-10.7484	-0.0378	-10.7862	0.15	-0.02
$C_2H_4\cdots CH_2O$	-6.7948	-6.8230	0.0001	-6.8228	0.41	-0.03
$NH_3\cdots C_2H_4$	-5.7865	-5.8003	-0.0143	-5.8146	0.49	-0.03
$HF\cdots CH_4$	-6.9162	-7.0312	-0.0614	-7.0926	2.55	-0.18
$H_2O\cdots CH_4$	-2.8242	-2.7652	-0.0135	-2.7787	-1.61	0.05
$NH_3\cdots CH_4$	-3.2175	-3.2335	-0.0116	-3.2451	0.86	-0.03
MAE					0.92	0.09

Table S4: junChS CP-energies (kJ mol^{-1}): the various contributions for the A14 complexes.

	"ref"	CC/junTZ	$\Delta MP2^{\infty}$/jun(T,Q)Z	MP2-CV/wCVTZ	Total	Rel. Error (%)	Abs. Error
$H_2O\cdots H_2O$	-21.0832	-19.3953	-1.5582	-0.1461	-21.0996	0.08	-0.02
$NH_3\cdots NH_3$	-13.2131	-11.8951	-1.3311	-0.0784	-13.3046	0.69	-0.09
$HF\cdots HF$	-19.2213	-17.5377	-1.8012	-0.1118	-19.4507	1.19	-0.23
$CH_2O\cdots CH_2O$	-18.9284	-16.2045	-2.9410	-0.0850	-19.2305	1.60	-0.30
$HCN\cdots HCN$	-19.9828	-18.7592	-1.0467	-0.0700	-19.8759	-0.54	0.11
$C_2H_4\cdots C_2H_4$	-4.6024	-3.7738	-0.9277	-0.0484	-4.7499	3.20	-0.15
$CH_4\cdots CH_4$	-2.2301	-1.6577	-0.5897	-0.0064	-2.2538	1.06	-0.02
$H_2O\cdots NH_3$	-27.3759	-25.4086	-1.9709	-0.1872	-27.5667	0.70	-0.19
$H_2O\cdots C_2H_4$	-10.7696	-9.4665	-1.2918	-0.1003	-10.8586	0.83	-0.09
$C_2H_4\cdots CH_2O$	-6.7948	-5.7002	-1.1790	-0.0636	-6.9427	2.18	-0.15
$NH_3\cdots C_2H_4$	-5.7865	-4.9559	-0.8829	-0.0556	-5.8943	1.86	-0.11
$HF\cdots CH_4$	-6.9162	-5.9096	-1.1172	-0.1006	-7.1274	3.05	-0.21
$H_2O\cdots CH_4$	-2.8242	-2.3175	-0.4339	-0.0319	-2.7833	-1.45	0.04
$NH_3\cdots CH_4$	-3.2175	-2.6927	-0.5117	-0.0399	-3.2443	0.83	-0.03
MAE						1.38	0.12

Table S5: junChS NCP-energies (kJ mol^{-1}): the various contributions for the A14 complexes.

	"ref"	CC/junTZ	$\Delta MP2^{\infty}$/jun(T,Q)Z	MP2-CV/wCVTZ	Total	Rel. Error (%)	Abs. Error
$H_2O\cdots H_2O$	-21.0832	-21.0233	-0.2079	-0.1429	-21.3741	1.38	-0.29
$NH_3\cdots NH_3$	-13.2131	-12.7819	-0.4828	-0.0784	-13.3432	0.98	-0.13
$HF\cdots HF$	-19.2213	-19.3800	-0.0959	-0.1140	-19.5899	1.92	-0.37
$CH_2O\cdots CH_2O$	-18.9284	-18.2645	-1.1814	0.0140	-19.4319	2.66	-0.50
$HCN\cdots HCN$	-19.9828	-20.5372	0.9151	-0.0829	-19.7049	1.39	0.28
$C_2H_4\cdots C_2H_4$	-4.6024	-4.7291	-0.0019	-0.0502	-4.7813	3.89	-0.18
$CH_4\cdots CH_4$	-2.2301	-2.0900	-0.1193	-0.0084	-2.2176	0.56	0.01
$H_2O\cdots NH_3$	-27.3759	-27.2308	-0.2795	-0.1782	-27.6885	1.14	-0.31
$H_2O\cdots C_2H_4$	-10.7696	-10.7880	-0.1205	-0.1008	-11.0094	2.23	-0.24
$C_2H_4\cdots CH_2O$	-6.7948	-6.8647	-0.1498	-0.0587	-7.0732	4.10	-0.28
$NH_3\cdots C_2H_4$	-5.7865	-5.8092	-0.1219	-0.0556	-5.9867	3.46	-0.20
$HF\cdots CH_4$	-6.9162	-7.0660	0.0317	-0.1056	-7.1399	3.23	-0.22
$H_2O\cdots CH_4$	-2.8242	-2.8609	0.1007	-0.0297	-2.7900	1.21	0.03
$NH_3\cdots CH_4$	-3.2175	-3.2828	0.0736	-0.0297	-3.2388	0.66	-0.02
MAE						2.06	0.22

Table S6: Geometries of the A14 complexes at different levels of theory. The inter-molecular parameters are highlighted in bold.

		CC/junVTZ	CC-F12/junTZ	CC-F12/junTZ+MP2-CV(wCVTZ)	ChS	ChS+augΔα	ChS+junΔα	junChS	junCBS+CV-F12	junChS-F12	revDSD	Hobza
HCN···HCN	r(H1-C2)	1.0688	1.0675	1.0664	1.0658	1.0663	1.0667	1.0656	1.0659	1.0658	1.0682	1.0687
	r(C2-N3)	1.1586	1.1553	1.1525	1.1504	1.1506	1.1506	1.1505	1.1508	1.1501	1.1556	1.1550
	r(N3-H4)	2.2167	2.2143	2.2114	2.2149	2.2059	2.2167	2.2164	2.2146	2.2162	2.2015	2.2253
	r(H4-C5)	1.0740	1.0729	1.0718	1.0718	1.0719	1.0729	1.0714	1.0717	1.0716	1.0740	1.0743
	r(C5-N6)	1.1609	1.1576	1.1548	1.1526	1.1529	1.1531	1.1528	1.1529	1.1522	1.1580	1.1572
HF···HF	r(H1-F2)	0.9231	0.9217	0.9212	0.9181	0.9211	0.9204	0.9179	0.9187	0.9184	0.9240	0.9199
	r(F2-H3)	1.8304	1.8280	1.8257	1.8244	1.8419	1.8468	1.8176	1.8212	1.8225	1.8261	1.8327
	r(H3-F4)	0.9257	0.9244	0.9239	0.9209	0.9249	0.9240	0.9208	0.9216	0.9213	0.9273	0.9227
	θ(H3-F2-F1)	115.18	114.70	115.05	115.68	120.32	121.17	114.79	115.16	115.10	114.03	115.02
	θ(F4-H3-F2)	170.63	169.41	169.67	170.71	174.39	174.86	169.81	169.78	168.23	169.53	169.56
H₂O···H₂O	r(O2-H1)	0.9603	0.9586	0.9576	0.9554	0.9576	0.9572	0.9554	0.9559	0.9556	0.9601	0.9577
	r(O2-H3)	0.9671	0.9658	0.9649	0.9625	0.9654	0.9649	0.9626	0.9632	0.9629	0.9680	0.9648
	θ(H3-O2-H1)	104.42	104.71	104.86	104.86	105.67	105.58	104.84	104.88	104.92	104.77	104.73
	r(H3-O4)	1.9600	1.9541	1.9503	1.9477	1.9472	1.9535	1.9507	1.9513	1.9487	1.9557	1.9618
	θ(O4-H3-O2)	172.62	171.88	171.95	172.68	171.45	172.36	171.15	171.82	171.24	171.98	171.72
	r(H5-O4)	0.9618	0.9602	0.9593	0.9570	0.9587	0.9582	0.9570	0.9575	0.9572	0.9618	0.9593
	θ(H5-O4-H3)	111.67	110.92	111.10	110.72	114.69	115.75	110.74	111.54	111.85	110.47	114.44
	φ(H5-O4-H3-O2)	58.32	58.01	58.17	57.83	60.23	60.89	57.89	58.46	58.67	57.78	58.36
	θ(H5-O4-H6)	104.53	104.79	104.94	104.95	105.43	105.39	104.90	104.98	103.98	104.86	104.83
NH₃···H₂O	r(O1-H2)	0.9601	0.9584	0.9574	0.9552	0.9574	0.9570	0.9552	0.9558	0.9555	0.9599	0.9576
	r(O1-H3)	0.9727	0.9714	0.9706	0.9680	0.9722	0.9717	0.9679	0.9688	0.9684	0.9742	0.9702
	θ(H3-O1-H2)	104.65	104.93	105.09	105.07	106.12	106.03	105.05	105.16	105.14	104.99	104.96
	r(H3-N4)	1.9784	1.9733	1.9696	1.9731	1.9569	1.9604	1.9789	1.9738	1.9753	1.9707	1.9850
	θ(N4-H3-O1)	170.87	170.53	170.45	170.77	166.01	166.23	169.86	169.91	169.84	170.81	169.98
	r(H5-N4)	1.0148	1.0133	1.0119	1.0107	1.0117	1.0112	1.0107	1.0110	1.0107	1.0142	1.0132
	θ(H5-N4-H3)	103.53	102.96	102.76	102.36	101.02	101.18	102.22	102.40	102.47	103.12	102.63
	r(H6-N4)	1.0144	1.0128	1.0114	1.0102	1.0111	1.0107	1.0102	1.0105	1.0102	1.0137	1.0127
	θ(H6-N4-H5)	106.54	106.72	106.90	106.88	107.26	107.21	106.89	106.91	106.90	106.88	106.70
	φ(H6-N4-H5-H3)	123.18	122.98	122.79	122.81	122.45	122.50	122.79	122.78	122.80	122.81	123.00

Table S6 Continued.

		CC/junVTZ	CC-F12/junTZ	CC-F12/junTZ+ MP2-CV(wCVTZ)	ChS	ChS+augΔα	ChS+junΔα	junChS	junCBS+CV-F12	junChS-F12	revDSD	Hobza
CH$_2$O···CH$_2$O	r(C1-H2)	1.1015	1.1005	1.0990	1.0983	1.0981	1.0979	1.0985	1.099	1.099	1.1025	1.1011
	r(C1-H3)	1.1016	1.1008	1.0994	1.0989	1.0995	1.0992	1.0990	1.099	1.099	1.1023	1.1015
	θ(H3-C1-H2)	117.52	117.64	117.66	117.62	117.91	117.87	117.60	117.60	117.58	117.42	117.60
	r(O4-C1)	1.2146	1.2117	1.2096	1.2068	1.2090	1.2087	1.2065	1.2065	1.2060	1.2104	1.2106
	θ(O4-C1-H3)	121.12	121.07	121.06	121.10	120.96	120.97	121.10	121.09	121.11	121.41	121.31
	r(**O5-H3**)	2.4136	2.3949	2.3870	2.3776	2.3516	2.3787	2.3707	2.3839	2.3781	2.4143	2.3827
	θ(**O5-H3-C1**)	121.62	121.35	121.43	121.41	120.98	121.07	121.29	121.33	121.25	120.86	121.55
	r(C6-O5)	1.2144	1.2116	1.2096	1.2070	1.2098	1.2093	1.2066	1.2064	1.2060	1.2106	1.2106
	θ(**C6-O5-H3**)	98.30	98.37	98.46	98.38	98.41	98.53	98.37	98.41	98.50	98.09	98.34
	r(H7-C6)	1.1010	1.1000	1.0984	1.0977	1.0974	1.0972	1.0980	1.0983	1.0982	1.1018	1.1006
	θ(H7-C6-O5)	121.79	121.77	121.75	121.78	121.59	121.56	121.82	121.80	121.80	121.88	121.79
	φ(**H7-C6-O5-H3**)	90.32	90.36	90.35	90.43	90.49	90.40	90.44	90.38	90.37	90.45	90.39
CH$_4$···NH$_3$	r(C1-H2)	1.0897	1.0885	1.0870	1.0857	1.0863	1.0856	1.0862	1.0865	1.0862	1.0898	1.0887
	r(**H2-N3**)	2.8166	2.8003	2.7948	2.8220	2.9307	2.9892	2.7713	2.7818	2.7751	2.7964	2.7933
	r(C1-H4)	1.0902	1.0888	1.0872	1.0861	1.0868	1.0862	1.0863	1.0866	1.0863	1.0901	1.0889
	θ(H4-C1-H2)	109.72	109.72	109.72	109.70	109.63	109.62	109.71	109.71	109.71	109.72	109.71
	r(H5-N3)	1.0145	1.0129	1.0114	1.0103	1.0113	1.0108	1.0102	1.0105	1.0102	1.0138	1.0128
	θ(**H5-N3-H2**)	112.44	112.27	112.07	112.08	111.69	111.73	112.06	112.05	112.06	112.11	112.26
	θ(H6-C1-H4)	109.22	109.22	109.23	109.24	109.32	109.32	109.23	109.23	109.23	109.22	109.23
	θ(H5-N3-H8)	106.34	106.54	106.75	106.74	107.16	107.12	106.77	106.77	106.76	106.71	106.54
CH$_4$···HF	r(H2-F1)	0.9223	0.9211	0.9206	0.9177	0.9220	0.9213	0.9173	0.9184	0.9179	0.9235	0.9193
	r(**H2-C3**)	2.3138	2.3022	2.2980	2.3191	2.3323	2.3444	2.3103	2.3077	2.3133	2.3079	2.3195
	r(H4-C3)	1.0886	1.0873	1.0857	1.0847	1.0856	1.0851	1.0850	1.0852	1.0849	1.0886	1.0873
	r(H5-C3)	1.0910	1.0897	1.0881	1.0870	1.0879	1.0873	1.0872	1.0876	1.0873	1.0909	1.0898
	θ(H5-C3-H4)	108.74	108.72	108.70	108.74	108.66	108.63	108.75	108.74	108.76	108.70	108.76
NH$_3$···NH$_3$	r(N1-H2)	1.0159	1.0145	1.0131	1.0119	1.0127	1.0121	1.0120	1.0123	1.0120	1.0154	1.0145
	r(**H2-N4**)	2.1886	2.1760	2.1703	2.1627	2.2044	2.2250	2.1598	2.1612	2.1651	2.1802	2.1717
	θ(H3-H2-N1)	98.14	98.01	98.08	98.40	98.47	98.46	98.29	98.30	98.16	98.15	98.17
	r(H5-N4)	1.0144	1.0128	1.0114	1.0102	1.0110	1.0105	1.0102	1.0105	1.0102	1.0137	1.0128
	θ(H5-N4-H3)	106.70	106.88	107.09	107.07	107.73	107.64	107.15	107.13	107.14	107.05	106.89
	φ(**H5-N4-H3-H2**)	123.39	123.19	122.98	122.98	122.38	122.46	122.93	122.96	122.94	123.01	123.18

Table S6 Continued.

		CC/junVTZ	CC-F12/junTZ	CC-F12/junTZ+MP2-CV(wCVTZ)	ChS	ChS+augΔα	ChS+junΔα	junChS	junCBS+CV-F12	junChS-F12	revDSD	Hobza
CH₄···CH₄	r(C2-H1)	1.0895	1.0882	1.0865	1.0856	1.0865	1.0897	1.0857	1.0860	1.0857	1.0894	1.0883
	r(**C2-C3**)	3.6953	3.6659	3.6605	3.6244	3.5438	3.6041	3.6231	3.6385	3.6373	3.6642	3.6380
	θ(**C3-C2-H1**)	70.47	70.46	70.46	70.44	70.43	70.45	70.44	70.44	70.45	70.46	70.45
	r(C3-H9)	1.0895	1.0882	1.0866	1.0855	1.0863	1.0895	1.0857	1.0860	1.0857	1.0894	1.0882
C₂H₄···C₂H₄	r(C1-C6)	1.3383	1.3350	1.3318	1.3302	1.3312	1.3312	1.3302	1.3335	1.3299	1.3319	1.3346
	r(**X2-X3**)	3.8497	3.8347	3.8232	3.7915	3.7102	3.7795	3.7816	3.8142	3.7994	3.8216	3.8116
	r(C4-C5)	1.3381	1.3348	1.3316	1.3301	1.3313	1.3313	1.3298	1.3333	1.3295	1.3316	1.3342
	r(C4-H7)	1.0836	1.0826	1.0812	1.0806	1.0813	1.0809	1.0807	1.0824	1.0807	1.0837	1.0832
	θ(H7-C4-C5)	121.07	121.03	121.01	121.01	120.93	121.03	121.01	121.01	121.00	121.06	121.03
	r(H8-C4)	1.0837	1.0826	1.0811	1.0806	1.0812	1.0808	1.0806	1.0823	1.0806	1.0837	1.0832
	θ(H8-C4-C5)	121.53	121.52	121.52	121.54	121.56	121.53	121.54	121.52	121.53	121.60	121.52
	θ(**C1-C6-C5**)	90.00	89.93	90.00	90.00	90.00	90.00	90.00	90.00	89.92	90.00	90.00
	r(H12-C1)	1.0837	1.0826	1.0811	1.0805	1.0812	1.0808	1.0806	1.0823	1.0806	1.0837	1.0831
	θ(H12-C1-C6)	121.42	121.40	121.40	121.41	121.39	121.41	121.42	121.40	121.40	121.46	121.41
	φ(**H12-C1-C4-C5**)	-121.42	-121.40	-121.40	-121.41	-121.39	-121.41	-121.42	-121.40	-121.40	-121.46	-121.41
C₂H₄···CH₂O	r(C1=O2)	1.2126	1.2096	1.2075	1.2047	1.2069	1.2053	1.2044	1.2044	1.2039	1.2084	1.2084
	r(H3-C1)	1.1031	1.1022	1.1006	1.1000	1.0998	1.1024	1.1002	1.1005	1.1004	1.1043	1.1028
	θ(H3-C1-O2)	121.45	121.42	121.40	121.44	121.28	121.29	121.44	121.45	121.73	121.52	121.43
	r(H4-C1)	1.1022	1.1014	1.0999	1.0994	1.0995	1.1021	1.0998	1.0997	1.1106	1.1032	1.1021
	θ(H4-C1-O2)	121.50	121.46	121.44	121.47	121.30	121.39	121.48	121.49	121.79	121.57	121.48
	r(**H4-C5**)	2.8807	2.8650	2.8572	2.8344	2.7848	2.8293	2.8472	2.8488	2.8748	2.8706	2.8607
	θ(**C5-H4-C1**)	114.22	114.08	114.11	115.21	116.54	117.54	114.38	114.13	109.28	114.10	114.19
	r(C5-C6)	1.3388	1.3355	1.3323	1.3307	1.3319	1.3366	1.3306	1.3303	1.3435	1.3324	1.3350
	θ(**C6-C5-C1**)	94.11	94.12	93.56	94.63	91.20	96.75	94.34	93.12	93.24	109.76	108.45
	r(H7-C5)	1.0833	1.0822	1.0808	1.0801	1.0808	1.0833	1.0805	1.0803	1.0909	1.0833	1.0827
	θ(H7-C5-C6)	121.72	121.72	121.72	121.73	121.69	121.83	121.72	121.71	121.74	121.76	121.71
	φ(**H7-C5-C6-H4**)	89.78	89.75	89.74	89.82	89.78	89.83	89.73	89.72	89.73	89.76	89.72
	r(H9-C6)	1.0838	1.0827	1.0812	1.0806	1.0812	1.0836	1.0809	1.0807	1.0917	1.0838	1.0833
	θ(H9-C6-C5)	121.48	121.46	121.46	121.47	121.46	121.59	121.46	121.46	121.48	121.52	121.47
	φ(**H9-C6-C5-H4**)	-90.16	-90.17	-90.18	-90.17	-90.21	-90.19	-90.18	-90.20	-90.18	-90.17	-90.18

125

Table S6 Continued.

	CC/junVTZ	CC-F12/junTZ	CC-F12/junTZ+ MP2-CV(wCVTZ)	ChS	ChS+aug Δα	ChS+jun Δα	junChS	junCBS+CV-F12	junChS-F12	revDSD	Hobza
r(H1-O2)	0.9609	0.9592	0.9583	0.9562	0.9582	0.9577	0.9561	0.9566	0.9563	0.9608	0.9584
r(H3-O2)	0.9641	0.9626	0.9617	0.9594	0.9625	0.9621	0.9593	0.9599	0.9595	0.9646	0.9615
θ(H3-O2-H1)	104.25	104.53	104.69	104.68	105.72	105.63	104.68	104.76	104.74	104.63	104.64
r(**H3-X4**)	2.4206	2.4168	2.4105	2.4199	2.4157	2.4272	2.4252	2.4242	2.4283	2.4006	2.4224
θ(**O2-H3-X4**)	173.41	173.41	173.27	177.41	163.64	163.99	170.75	172.87	172.72	166.39	161.14
r(C5-C6)	1.3397	1.3364	1.3332	1.3317	1.3327	1.3328	1.3316	1.3317	1.3313	1.3336	1.3361
r(C5-H8)	1.0836	1.0826	1.0811	1.0804	1.0811	1.0807	1.0805	1.0808	1.0805	1.0837	1.0832
θ(H8-C5-C6)	121.45	121.44	121.43	121.43	121.39	121.41	121.45	121.44	121.44	121.49	121.42
r(H7-C5)	1.0836	1.0826	1.0811	1.0804	1.0810	1.0807	1.0805	1.0808	1.0805	1.0837	1.0831
θ(H7-C5-C6)	121.45	121.44	121.43	121.45	121.44	121.43	121.46	121.44	121.44	121.47	121.41
φ(**C5-H3-O2-H1**)	-112.51	-112.53	-112.95	-112.99	-160.01	-159.49	-117.80	-114.32	-114.81	-130.26	-139.53
φ(**H8-C5-X4-H3**)	-90.10	-90.10	-90.11	-85.09	-79.53	-80.52	-87.57	-90.11	-90.12	-77.76	-72.92
r(C1-C2)	1.3388	1.3356	1.3324	1.3308	1.3320	1.3366	1.3307	1.3308	1.3304	1.3325	1.3351
r(**C2-H3**)	2.6878	2.6778	2.6698	2.6953	2.6880	2.7529	2.6616	2.6661	2.6694	2.6706	2.6835
θ(**C1-C2-H3**)	88.55	88.88	88.28	82.88	73.54	77.77	87.38	86.99	86.73	87.02	86.93
r(H3-N4)	1.0150	1.0136	1.0122	1.0111	1.0122	1.0137	1.0111	1.0112	1.0110	1.0146	1.0135
θ(**C2-H3-N4**)	146.07	145.67	146.51	157.89	173.55	170.12	148.10	148.23	148.46	147.48	147.84
r(N4-H5)	1.0145	1.0129	1.0114	1.0101	1.0107	1.0129	1.0102	1.0104	1.0102	1.0138	1.0126
θ(H3-N4-H5)	106.53	106.73	106.94	106.89	107.74	107.28	107.00	106.99	106.99	106.90	104.75
φ(**C2-H3-N4-H5**)	123.42	123.23	123.01	123.05	122.24	122.68	122.95	122.96	122.96	123.05	123.22
r(C2-H7)	1.0835	1.0824	1.0810	1.0803	1.0811	1.0835	1.0804	1.0806	1.0804	1.0835	1.0828
θ(C1-C2-H7)	121.53	121.52	121.51	121.48	121.38	121.52	121.53	121.50	121.50	121.57	121.52
φ(**H7-C2-C1-H3**)	89.87	89.85	89.84	89.91	90.03	90.02	89.85	89.84	89.85	89.87	89.85
r(C1-H9)	1.0837	1.0826	1.0811	1.0804	1.0810	1.0835	1.0805	1.0808	1.0806	1.0837	1.0830
θ(C2-C1-H9)	121.46	121.44	121.44	121.45	121.425	121.55	121.46	121.44	121.44	121.50	121.45
φ(**H9-C1-C2-H3**)	-90.12	-90.13	-90.14	-90.16	-90.17	-90.16	-90.15	-90.14	-90.15	-90.13	-90.14

$C_2H_4\cdots H_2O$

$C_2H_4\cdots NH_3$

Table S6 Continued.

		CC/junVTZ	CC-F12/junTZ	CC-F12/junTZ+MP2-CV(wCVTZ)	ChS	ChS+augΔα	ChS+junΔα	junChS	junCBS+CV-F12	junChS-F12	revDSD	Hobza
	r(C2-H3)	1.0888	1.0880	1.0864	1.0857	1.0869	1.0862	1.0852	1.0859	1.0857	1.0893	1.0880
	r(**H3-O4**)	2.6659	2.6270	2.6232	2.8410	2.7831	2.8169	2.6434	2.6238	2.6229	2.6445	2.6279
	r(H1-C2)	1.0898	1.0886	1.0870	1.0863	1.0869	1.0863	1.0860	1.0864	1.0861	1.0900	1.0887
	r(H7-C2)	1.0892	1.0887	1.0870	1.0856	1.0867	1.0862	1.0854	1.0865	1.0861	1.0899	1.0887
	r(H5-O4)	0.9601	0.9596	0.9586	0.9565	0.9582	0.9577	0.9553	0.9569	0.9566	0.9612	0.9583
	θ(**O4-H3-C2**)	161.98	179.96	179.87	141.28	157.40	158.17	161.25	179.74	179.60	179.41	179.09
	θ(H1-C2-H3)	109.94	109.67	109.68	109.73	109.49	109.44	109.93	109.66	109.67	109.70	109.65
	θ(H7-C2-H3)	109.55	109.29	109.66	109.58	109.69	109.70	109.55	109.67	109.30	109.66	109.68
	θ(**H5-O4-H3**)	80.35	127.77	127.30	77.91	118.12	125.06	80.26	127.88	127.30	123.52	127.69
$H_2O\cdots CH_4$	φ(**H5-O4-H3-C2**)	52.82	89.94	82.77	53.89	68.17	91.56	51.26	89.96	82.74	71.54	87.36
	φ(**H8-C2-H3-O4**)	-120.09	-120.02	-120.03	-120.01	-119.94	-119.93	-120.08	-120.02	-120.03	-120.03	-120.02

Table S7: ChS, junChS, and junChS-F12 models: CBS and CV contributions for the A14 complexes. The inter-molecular parameters are highlighted in bold.

		ΔR(CBS)			ΔR(CV)
		ChS	junChS	junChS-F12	
HCN···HCN	r(H1-C2)	-0.0009	-0.0021	-0.0006	-0.0012
	r(C2-N3)	-0.0055	-0.0055	-0.0024	-0.0025
	r(N3-H4)	0.0035	0.0021	0.0048	-0.0024
	r(**H4-C5**)	0.0002	-0.0014	-0.0001	-0.0012
	r(C5-N6)	-0.0055	-0.0057	-0.0025	-0.0024
HF···HF	r(H1-F2)	-0.0022	-0.0046	-0.0027	-0.0005
	r(**F2-H3**)	0.0146	-0.0110	-0.0032	-0.0018
	r(H3-F4)	-0.0012	-0.0043	0.0462	-0.0005
	θ(**H3-F2-F1**)	5.53	-0.60	-0.003	0.21
	θ(**F4-H3-F2**)	3.88	-0.93	-1.44	0.11
H_2O···H_2O	r(O2-H1)	-0.0024	-0.0042	-0.0020	-0.0008
	r(O2-H3)	-0.0016	-0.0038	-0.0020	-0.0008
	θ(H3-O2-H1)	1.04	0.31	0.05	0.11
	r(**H3-O4**)	-0.0055	-0.0066	-0.0016	-0.0027
	θ(**O4-H3-O2**)	-0.55	-1.59	-0.71	0.12
	r(H5-O4)	-0.0029	-0.0038	-0.0021	-0.0008
	θ(**H5-O4-H3**)	3.23	-1.12	0.74	0.21
	φ(**H5-O4-H3-O2**)	2.01	-0.59	0.50	0.16
	θ(H5-O4-H6)	0.73	0.26	0.04	0.12

Table: S7 Continued.

		ΔR(CBS)			ΔR(CV)
		ChS	junChS	junChS-F12	
$NH_3{\cdots}H_2O$	r(O1-H2)	-0.0026	-0.0040	-0.0020	-0.0008
	r(O1-H3)	-0.0007	-0.0042	-0.0023	-0.0007
	θ(H3-O1-H2)	1.23	0.28	0.05	0.12
	r(**H3-N4**)	-0.0154	0.0028	0.0057	-0.0025
	θ(**N4-H3-O1**)	-4.38	-1.00	-0.61	-0.02
	r(H5-N4)	-0.0024	-0.0029	-0.0012	-0.0012
	θ(**H5-N4-H3**)	-2.08	-1.16	-0.29	-0.14
	r(H6-N4)	-0.0024	-0.0029	-0.0012	-0.0012
	θ(H6-N4-H5)	0.55	0.21	-0.001	0.14
	φ(**H6-N4-H5-H3**)	-0.54	-0.24	0.001	-0.15
$CH_2O{\cdots}CH_2O$	r(C1-H2)	-0.0022	-0.0017	-0.0001	-0.0013
	r(C1-H3)	-0.0010	-0.0012	-0.0001	-0.0013
	θ(H3-C1-H2)	0.35	0.07	-0.04	0.001
	r(O4-C1)	-0.0040	-0.0061	-0.0015	-0.0020
	θ(O4-C1-H3)	-0.17	-0.03	0.02	0.003
	r(**O5-H3**)	-0.0370	-0.0396	-0.0038	-0.0032
	θ(**O5-H3-C1**)	-0.57	-0.36	-0.07	0.03
	r(C6-O5)	-0.0035	-0.0059	-0.0015	-0.0020
	θ(**C6-O5-H3**)	0.16	0.03	0.02	0.04
	r(H7-C6)	-0.0024	-0.0016	-0.0001	-0.0013
	θ(H7-C6-O5)	-0.21	0.05	0.02	-0.002
	φ(**H7-C6-O5-H3**)	0.10	0.12	0.01	-0.01

Table: S7 Continued.

		ΔR(CBS)			ΔR(CV)
		ChS	junChS	junChS-F12	
$CH_4 \cdots NH_3$	r(C1-H2)	-0.0024	-0.0021	-0.0008	-0.0014
	r**(H2-N3)**	0.1489	-0.0426	-0.0198	-0.0027
	r(C1-H4)	-0.0024	-0.0024	-0.0009	-0.0014
	θ(H4-C1-H2)	-0.10	-0.02	-0.002	-0.004
	r(H5-N3)	-0.0026	-0.0031	-0.0012	-0.0012
	θ**(H5-N3-H2)**	-0.55	-0.24	-0.01	-0.15
	θ(H6-C1-H4)	0.10	0.02	0.002	0.004
	θ(H5-N3-H8)	0.61	0.24	0.01	0.17
$CH_4 \cdots HF$	r(H2-F1)	-0.0007	-0.0045	-0.0027	-0.0005
	r**(H2-C3)**	0.0296	0.0007	0.0153	-0.0041
	r(H4-C3)	-0.0019	-0.0022	-0.0008	-0.0014
	r(H5-C3)	-0.0021	-0.0022	-0.0008	-0.0015
	θ(H5-C3-H4)	-0.10	0.03	0.05	-0.01
$NH_3 \cdots NH_3$	r(N1-H2)	-0.0025	-0.0027	-0.0012	-0.0012
	r**(H2-N4)**	0.0373	-0.0251	-0.0052	-0.0038
	θ(H3-H2-N1)	0.09	0.06	0.07	0.10
	r(H5-N4)	-0.0027	-0.003	-0.0012	-0.0012
	θ(H5-N4-H3)	0.76	0.29	0.05	0.16
	φ**(H5-N4-H3-H2)**	-0.73	-0.29	-0.04	-0.16

Table: S7 Continued.

		ΔR(CBS)			ΔR(CV)
		ChS	junChS	junChS-F12	
CH$_4$···CH$_4$	r(C2-H1)	-0.0020	-0.0024	-0.0008	-0.0014
	r**(C2-C3)**	-0.0746	-0.0696	-0.0232	-0.0026
	θ**(C3-C2-H1)**	-0.01	-0.03	-0.02	-0.001
	r(C3-H9)	-0.0022	-0.0024	-0.0009	-0.0014
C$_2$H$_4$···C$_2$H$_4$	r(C1-C6)	-0.0045	-0.0053	-0.0019	-0.0028
	r**(X2-X3)**	-0.0685	-0.0611	-0.0239	-0.0070
	r(C4-C5)	-0.0042	-0.0055	-0.0021	-0.0028
	r(C4-H7)	-0.0014	-0.0016	-0.0004	-0.0013
	θ(H7-C4-C5)	-0.04	-0.06	-0.01	0.003
	r(H8-C4)	-0.0016	-0.0018	-0.0005	-0.0013
	θ(H8-C4-C5)	-0.01	0.001	0.01	0.01
	θ**(C1-C6-C5)**	0.001	0.001	-0.01	0.00
	r(H12-C1)	-0.0016	-0.0018	-0.0005	-0.0013
	θ(H12-C1-C6)	0.02	0.01	-0.002	-0.01
	φ**(H12-C1-C4-C5)**	-0.02	-0.01	0.002	0.01

Table: S7 Continued.

		ΔR(CBS)			ΔR(CV)
		ChS	junChS	junChS-F12	
	$r(\text{C}_1\!=\!\text{O}_2)$	-0.0042	-0.0063	-0.0037	-0.0020
	$r(\text{H3-C1})$	-0.0021	-0.0016	-0.0003	-0.0013
	$\theta(\text{H3-C1-O2})$	-0.19	0.003	0.05	-0.002
	$r(\text{H4-C1})$	-0.0016	-0.0014	-0.0002	-0.0013
	$\theta(\text{H4-C1-O2})$	-0.16	0.001	0.05	-0.01
	$r(\textbf{H4-C5})$	-0.0511	-0.0375	-0.0085	-0.0043
	$\theta(\textbf{C5-H4-C1})$	2.23	-0.24	0.01	0.03
$\text{C}_2\text{H}_4\cdots\text{CH}_2\text{O}$	$r(\text{C5-C6})$	-0.0042	-0.0053	-0.002	-0.0028
	$\theta(\textbf{C6-C5-C1})$	-1.95	-1.29	-0.44	-0.18
	$r(\text{H7-C5})$	-0.0015	-0.0018	-0.0005	-0.0013
	$\theta(\text{H7-C5-C6})$	-0.04	-0.01	-0.01	0.01
	$\varphi(\textbf{H7-C5-C6-H4})$	0.02	-0.03	-0.02	-0.001
	$r(\text{H9-C6})$	-0.0016	-0.0018	-0.0005	-0.0013
	$\theta(\text{H9-C6-C5})$	-0.02	-0.01	0.002	0.01
	$\varphi(\textbf{H9-C6-C5-H4})$	-0.02	-0.02	-0.01	-0.004

Table: S7 Continued.

		ΔR(CBS)			ΔR(CV)
		ChS	junChS	junChS-F12	
	r(H1-O2)	-0.0025	-0.004	-0.0020	-0.0008
	r(H3-O2)	-0.0013	-0.004	-0.0022	-0.0008
	θ(H3-O2-H1)	1.19	0.31	0.05	0.12
	r(**H3-X4**)	0.0049	0.0084	0.0179	-0.0038
	θ(**O2-H3-X4**)	-9.22	-2.71	-0.56	0.05
$C_2H_4 \cdots H_2O$	r(C5-C6)	-0.0042	-0.0053	-0.0019	-0.0028
	r(C5-H8)	-0.0016	-0.0018	-0.0006	-0.0013
	θ(H8-C5-C6)	-0.03	-0.02	0.003	0.01
	r(H7-C5)	-0.0016	-0.0018	-0.0006	-0.0013
	θ(H7-C5-C6)	-0.02	-0.004	0.003	0.01
	φ(**C5-H3-O2-H1**)	-45.31	-5.12	-1.76	0.17
	φ(**H8-C5-X4-H3**)	4.99	2.48	-0.01	0.05

Table: S7 Continued.

		ΔR(CBS)			ΔR(CV)
		ChS	junChS	junChS-F12	
	r(C1-C2)	-0.0042	-0.0053	-0.0019	-0.0028
	r**(C2-H3)**	0.0164	-0.0211	-0.0004	-0.0052
	θ**(C1-C2-H3)**	-14.60	-0.74	-1.55	-0.43
	r(H3-N4)	-0.0023	-0.0028	-0.0011	-0.0012
	θ**(C2-H3-N4)**	27.28	1.28	1.95	0.75
	r(N4-H5)	-0.0029	-0.003	-0.0012	-0.0012
	θ(H3-N4-H5)	0.89	0.30	0.05	0.17
$C_2H_4\cdots NH_3$	φ**(C2-H3-N4-H5)**	-0.86	-0.30	-0.05	-0.17
	r(C2-H7)	-0.0014	-0.0018	-0.0005	-0.0013
	θ(C1-C2-H7)	-0.16	-0.01	-0.01	0.01
	φ**(H7-C2-C1-H3)**	0.18	-0.01	0.003	0.002
	r(C1-H9)	-0.0017	-0.0018	-0.0005	-0.0013
	θ(C2-C1-H9)	-0.03	-0.01	0.004	0.01
	φ**(H9-C1-C2-H3)**	-0.03	-0.02	-0.02	-0.003

Table: S7 Continued.

		ΔR(CBS)			ΔR(CV)
		ChS	junChS	junChS-F12	
	r(C2-H3)	-0.0017	-0.0022	-0.0008	-0.0014
	r(**H3-O4**)	0.1777	-0.0199	-0.0003	-0.0026
	r(H1-C2)	-0.0021	-0.0024	-0.0009	-0.0014
	r(H7-C2)	-0.0022	-0.0024	-0.0009	-0.0014
	r(H5-O4)	-0.0028	-0.004	-0.0020	-0.0008
$H_2O\cdots CH_4$	θ(**O4-H3-C2**)	-20.56	-0.59	-0.28	-0.14
	θ(H1-C2-H3)	-0.21	-0.02	-0.01	0.002
	θ(H7-C2-H3)	0.03	-0.003	-0.004	-0.002
	θ(**H5-O4-H3**)	-2.49	-0.14	-0.003	0.05
	φ(**H5-O4-H3-C2**)	1.02	-1.61	0.003	0.05
	φ(**H8-C2-H3-O4**)	0.08	0.004	0.004	-0.002

Table S8: Comparison of CP and NCP corrected geometries for some paradigmatic cases at the augF12CBS+CV and augF12CBS+CV-F12 levels.

		augF12CBS+CV[a]	augF12CBS+CV-F12[b]
		NCP	NCP
$H_2O\cdots H_2S$	$r(\text{O1-H2})$	0.9641	0.9623
	$r(\text{O1-H3})$	0.9598	0.9581
	$\theta(\text{H2-O1-H3})$	104.52	104.69
	$r(\text{H3}\cdots\text{S4})$	3.4639	3.4783
	$\theta(\text{O1-H3-S4})$	116.65	116.61
	$r(\text{H5/H6-S4})$	1.3395	1.3379
	$\theta(\text{H5-S4-H3})$	83.02	83.85
	$\varphi(\text{H5-S4-H3-O1})$	133.49	133.48
	$\varphi(\text{H6-S4-H3-O1})$	-133.49	-133.48
	$\theta(\text{H5-S4-H6})$	92.15	92.35
$H_2O\cdots H_2O$	$r(\text{O2-H1})$	0.9592	0.9574
	$r(\text{O2-H3})$	0.9664	0.9647
	$\theta(\text{H3-O2-H1})$	104.67	104.83
	$r(\text{H3}\cdots\text{O4})$	1.9442	1.9494
	$\theta(\text{O4-H3-O2})$	171.54	171.72
	$r(\text{H5/H6-O4})$	0.9607	0.9590
	$\theta(\text{H5-O4-H3})$	111.01	111.51
	$\varphi(\text{H5-O4-H3-O2})$	58.04	58.43
	$\varphi(\text{H6-O4-H3-O2})$	-58.04	-58.43
	$\theta(\text{H5-O4-H6})$	104.75	104.92
$CH_4\cdots CH_4$	$r(\text{C2-H1})$	1.0889	1.0870
	$r(\text{C2}\cdots\text{C3})$	3.6294	3.6470
	$\theta(\text{C3-C2-H1})$	70.44	70.45
	$r(\text{C3-H9})$	1.0889	1.0870
$HCN\cdots HCN$	$r(\text{H1-C2})$	1.0679	1.0664
	$r(\text{C2-N3})$	1.1545	1.1522
	$r(\text{N3}\cdots\text{H4})$	2.2108	2.2120
	$r(\text{H4-C5})$	1.0737	1.0723
	$r(\text{C5-N6})$	1.1567	1.1544

[a] Conventional CBS+CV scheme using the aug-cc-pVnZ-F12 basis sets.
[b] CBS+CV-F12 scheme using the aug-cc-pVnZ-F12 basis sets.

Appendix B

Supporting data for Chapter 3

HTBH38/08 dataset extended table results

Table S1: Theoretical values of the barrier heights for the forward and reverse reactions in the HTBH38/08 and NHTBH38/08 datasets, obtained at different levels of theory. All values (not including zero-point energy and spin-orbit corrections) are in kcal mol^{-1}.

	Reaction	forward/reverse barrier height					
		revDSD[a]	CCSD(T)/revDSD[b]	CCSD(T)/QCISD[c]	jChS/revDSD[d]	jChS/QCISD[e]	"best"[f]
c1	$H^\bullet + HCl \longrightarrow H_2 + Cl^\bullet$	1.27/7.54	5.63/9.07	5.90/9.34	4.97/7.80	5.57/7.95	5.49/7.42[g]
c2	$OH^\bullet + H_2 \longrightarrow H_2O + H^\bullet$	5.88/20.92	6.04/20.92	5.97/20.84	5.47/21.76	5.38/21.69	5.10/21.20
c3	$CH_3^\bullet + H_2 \longrightarrow CH_4 + H^\bullet$	11.32/14.36	12.30/14.99	12.29/14.97	11.96/14.64	11.95/14.64	12.10/15.30
c4	$H^\bullet + H_2 \longrightarrow H_2 + H^\bullet$	8.96/8.96	10.02/10.02	10.01/10.01	9.58/9.58	9.58/9.58	9.60/9.60
c5	$OH^\bullet + NH_3 \longrightarrow H_2O + NH_2^\bullet$	4.67/14.37	4.21/14.29	4.18/14.27	3.93/14.41	3.35/13.85	3.20/12.70
c6	$HCl + CH_3^\bullet \longrightarrow Cl^\bullet + CH_4$	1.09/10.40	2.27/8.39	2.27/8.39	1.69/6.35	1.70/6.77	1.70/7.90
c7	$OH^\bullet + C_2H_6 \longrightarrow H_2O + C_2H_5^\bullet$	4.47/19.97	4.17/19.71	4.03/19.57	3.80/20.91	3.64/20.75	3.40/19.90
c8	$F^\bullet + H_2 \longrightarrow HF + H^\bullet$	2.77/32.22	2.26/32.76	2.26/32.75	1.69/33.90	1.77/34.00	1.80/33.40
c9	$^3O + CH_4 \longrightarrow OH^\bullet + CH_3^\bullet$	14.60/8.24	15.33/8.42	15.36/8.45	14.55/9.63	14.65/9.62	13.70/8.10
c10	$H^\bullet + PH_3 \longrightarrow PH_2^\bullet + H_2$	2.61/24.50	3.33/25.90	3.30/25.88	2.85/25.09	2.82/25.05	3.10/23.20
c11	$O + HCl \longrightarrow OH^\bullet + Cl^\bullet$	7.60/10.56	11.67/10.89	11.61/10.81	10.59/10.34	10.63/10.66	9.80/10.40
c12	$NH_2^\bullet + CH_3^\bullet \longrightarrow CH_4 + NH$	8.94/21.97	8.69/22.99	8.69/23.01	9.49/22.09	9.50/22.11	8.00/22.40
c13	$NH_2^\bullet + C_2H_5 \longrightarrow NH + C_2H_5$	9.28/18.81	9.01/19.98	9.43/20.40	9.97/19.08	10.39/19.51	7.50/18.30
c14	$NH_2^\bullet + C_2H_6 \longrightarrow NH_3 + C_2H_5^\bullet$	11.62/17.43	11.85/17.31	11.83/17.28	11.24/17.85	11.18/17.80	10.40/17.40
c15	$NH_2^\bullet + CH_4 \longrightarrow NH_3 + CH_3^\bullet$	14.11/46.42	14.47/16.59	14.47/16.58	13.82/16.94	13.80/16.92	14.50/17.80
c16	s-trans cis-$C_5H_8 \longrightarrow$ same	39.42/39.42	39.87/39.87	39.87/39.87	39.66/39.66	39.63/39.63	38.40/38.40
c17	$H^\bullet + FH \longrightarrow HF + H^\bullet$	40.68/40.68	42.64/42.64	42.64/42.64	41.99/41.99	42.02/42.02	42.18/42.18
c18	$H^\bullet + FCH_3 \longrightarrow HF + CH_3^\bullet$	30.97/55.29	30.51/57.43	30.47/57.40	30.31/57.54	30.31/57.54	30.38/57.02
c19	$H^\bullet + F_2 \longrightarrow HF + F^\bullet$	9.32/111.60	5.07/108.93	1.85/105.78	2.38/106.07[h]	1.49/105.25	2.59/105.77[i]
c20	$F^- + CH_3F \longrightarrow FCH_3 + F^-$	-0.71/-0.71	-0.02/-0.02	-0.04/-0.04	-0.70/-0.70	-0.71/-0.71	-0.34/-0.34
c21	$F^- \cdots CH_3F \longrightarrow FCH_3 \cdots F^-$	12.72/12.72	13.53/13.53	13.50/13.50	13.21/13.21	13.20/13.20	13.38/13.38
c22	$Cl^- + CH_3Cl \longrightarrow ClCH_3 + Cl^-$	1.86/1.86	3.11/3.11	3.12/3.12	2.27/2.27	2.33/2.33	3.10/3.10
c23	$F^- + CH_3Cl \longrightarrow FCH_3 + Cl^-$	-12.48/19.02	-11.73/20.09	-11.73/20.08	-12.32/19.29	-12.31/19.31	-12.54/20.11
c24	$OH^- \cdots CH_3F \longrightarrow HOCH_3 \cdots F^-$	10.42/47.21	11.18/47.80	11.15/47.76	11.14/47.38	11.14/47.38	10.96/47.20
c25	$H^\bullet + CO \longrightarrow HCO^\bullet$	3.52/23.65	3.74/22.52	3.71/22.50	3.22/22.87	3.19/22.82	3.17/22.68
c26	$CH_3^\bullet + C_2H_4 \longrightarrow CH_3CH_2CH_2^\bullet$	7.91/35.57	6.74/32.64	6.61/32.51	6.37/32.77	6.35/32.74	6.85/32.97
	MUE	1.27/2.18 (0.89/2.25)	0.73/0.71 (0.61/0.60)	0.64/0.63 (0.61/0.63)	0.70/0.88 (0.50/0.58)	0.52/0.66 (0.50/0.59)	-
	RMSD	1.91/5.84 (1.02/6.49)	0.99/1.05 (0.82/0.89)	0.85/0.90 (0.84/0.93)	1.04/1.23 (0.74/0.73)	0.80/0.80 (0.81/0.74)	-

[a] revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z; [b] CCSD(T)/jun-cc-pV(T+d)Z//revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z; [c] CCSD(T)/jun-cc-pV(T+d)Z; [d] jChS on revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z optimized geometry; [e] jChS on QCISD/MG3 optimized geometry; [f] Zheng et al. *J. Chem. Theory Comput.* 2009, 5, 808–821; [g] CBS-CVH result (this work); [h] jChS geometry; the values at the rev-DSD geometry were 4.46/108.14; [i] CBS-CVH result (this work): the values at the rev-DSD geometry were 2.27/105.80. the original values were 5.49/7.42; the original values were 2.27/105.80

138

Appendix C

Supporting data for Chapter 4

Thermochemistry

Using the electronic structure calculations carried out and literature data, the heats of formation (at 298 K, 1 atm) of H_2S, HS, and HCl have been derived. This only required additional computations for the sulfur atom (^3P).

Table S1: Thermochemistry.[a]

	E_{el}[b]	ΔE_{el}[b]	ZPE[c]	$H° - H°_\circ$	SO[d]	$\Delta_f H°$ (calc)	$\Delta_f H°$ (exp)
	(Hartree)	(kJ mol^{-1})	(kJ mol^{-1})	(kJ mol^{-1})	(kJ mol^{-1})	(kJ mol^{-1})	(kJ mol^{-1})
H	-0.5	0.0	0.0	6.20[e]	0.0	–	218.00[e]
S	-399.1715515	0.0	0.0	6.66[e]	-2.16	–	277.17[e]
Cl	-461.5429218	0.0	0.0	6.20	-3.28	–	121.30[f]
HS	-399.8107595	-365.46	16.14	8.81	-2.10	141.86	141.87[e]
HCl	-462.2138181	-448.66	17.88	8.78	0.0	-91.82	-92.31[f]
H_2S	-400.4635863	-766.68	39.63	10.44	0.0	-20.34	-20.50[f]

[a] Standard state: 1 atm, 298 K. [b] At the CBS+CV+DBOC+rel+fT+fQ level.
[c] At the the revDSD-PBEP86-D3(BJ)/jun-cc-pV(T+d)Z level.
[d] At the CASSCF/aug-cc-pVTZ level. [e] Ref. 86. [f] Ref. 87.

MESS input files

Given the length of the input files, in order not to overload the thesis excessively, please refer directly to the SI published here https://pubs.acs.org/doi/10.1021/acs.jctc.0c00354

THIS PAGE INTENTIONALLY LEFT BLANK

Appendix D

Supporting data for Chapter 6

Cartesian coordinates of the stationary points can be found in the SI published here

https://pubs.acs.org/doi/10.1021/acs.jctc.1c01252.

Table S1: Main structural results for the water dimer obtained at the ChS and junChS level are compared with the CCSD(T)-F12b/CBS+fT+fQ+CV+REL+DBOC values of ref.[88] θ_a and θ_d are the H-O-H valence angles for the acceptor and the donor respectively, α gives a measure of the deviation from a linear hydrogen bond and β gives the orientation of the C2 axis of the acceptor with respect to the O-O axis. Bond lengths in Å and angles in degrees.

	ChS	junChS	Ref. value[88]
r(O-O)	2.9049	2.9058	2.9092
θ_a	104.78	104.77	104.95
θ_d	104.87	104.84	104.85
α	4.91	5.91	5.69
β	126	127	123.46

Table S2: MAE and RE for bond lengths (r), angles (θ) and dihedrals (ϕ) of the HCN$\cdots$(H$_2$O)$_2$ system. Averaged values over all the species along the PES are collected. Last column collect total RE obtained averaging over all the structural parameters. MAE for bond lengths in Å and in degrees for angles and dihedrals. RE are in %.

		MAE			RE			
		r	θ	ϕ	r	θ	ϕ	Total
B3LYP-D3	jun-DZ	0.028	1.72	2.21	1.78	1.31	2.43	1.02
	jul-DZ	0.023	1.45	3.56	1.47	1.23	3.65	0.98
	aug-DZ	0.021	1.24	2.53	1.36	1.08	2.61	0.88
	jun-TZ	0.013	1.07	2.69	0.87	0.89	3.08	0.69
	jul-TZ	0.013	1.03	2.44	0.86	0.86	2.69	0.65
	aug-TZ	0.013	1.09	2.51	0.89	0.92	2.82	0.69
BHLYP-D3	jun-DZ	0.092	7.34	9.21	3.50	6.17	10.47	4.82
	jul-DZ	0.016	1.64	2.52	0.87	1.39	2.38	0.85
	aug-DZ	0.014	1.46	1.64	0.79	1.25	1.81	0.76
	jun-TZ	0.014	1.28	1.59	0.96	1.07	1.65	0.80
	jul-TZ	0.014	1.34	1.54	0.95	1.10	1.82	0.83
	aug-TZ	0.015	1.34	1.48	0.97	1.11	1.72	0.83
ωB97X-D	jun-DZ	0.030	2.36	4.01	1.65	1.90	3.28	1.06
	jul-DZ	0.023	1.82	5.01	1.25	1.57	4.86	0.91
	aug-DZ	0.021	1.70	5.15	1.14	1.48	4.89	0.92
	jun-TZ	0.017	1.63	5.06	0.88	1.37	5.23	0.81
	jul-TZ	0.017	1.59	4.87	0.88	1.34	4.95	0.78
	aug-TZ	0.018	1.62	5.00	0.93	1.38	5.17	0.81
PW6B95-D3	jun-DZ	0.022	2.03	2.47	1.23	1.50	2.48	1.00
	jul-DZ	0.015	1.34	3.11	0.80	1.09	3.05	0.74
	aug-DZ	0.013	1.05	1.96	0.73	0.90	1.65	0.60
	jun-TZ	0.011	0.96	2.56	0.55	0.79	2.59	0.55
	jul-TZ	0.010	0.94	2.31	0.53	0.76	2.29	0.52
	aug-TZ	0.010	0.95	2.34	0.52	0.78	2.36	0.53
BMK-D3	jun-DZ	0.015	1.87	3.32	1.00	1.36	2.71	0.75

	jul-DZ	0.013	1.67	4.09	0.84	1.38	3.84	0.77
	aug-DZ	0.012	1.60	2.84	0.77	1.34	2.47	0.69
	jun-TZ	0.008	1.03	2.32	0.52	0.84	2.04	0.57
	jul-TZ	0.008	1.00	2.00	0.52	0.81	1.61	0.54
	aug-TZ	0.008	1.03	2.02	0.51	0.84	1.68	0.56
M06-2X	jun-DZ	0.014	1.95	2.65	1.05	1.37	2.99	1.09
	jul-DZ	0.013	1.16	2.96	0.91	0.96	2.72	0.85
	aug-DZ	0.014	1.18	2.62	0.90	0.96	2.74	0.84
	jun-TZ	0.012	1.05	2.36	0.72	0.82	2.06	0.70
	jul-TZ	0.012	1.00	1.90	0.74	0.79	1.39	0.65
	aug-TZ	0.012	1.04	1.86	0.75	0.83	1.41	0.66
MN15	jun-DZ	0.017	2.15	3.08	1.18	1.59	2.62	1.15
	jul-DZ	0.013	1.49	3.70	0.94	1.17	3.23	0.95
	aug-DZ	0.010	1.20	2.80	0.82	0.99	2.87	0.83
	jun-TZ	0.009	1.03	3.13	0.63	0.82	2.48	0.68
	jul-TZ	0.008	1.04	2.69	0.60	0.82	1.99	0.65
	aug-TZ	0.008	1.01	2.60	0.58	0.80	1.93	0.65
B2PLYP-D3	jun-DZ	0.025	1.88	2.36	1.58	1.39	2.89	1.10
	jul-DZ	0.020	1.23	2.71	1.29	1.03	2.62	0.88
	aug-DZ	0.018	1.00	1.74	1.21	0.88	1.65	0.75
	jun-TZ	0.010	0.77	1.65	0.63	0.63	1.77	0.51
	jul-TZ	0.009	0.69	2.13	0.58	0.57	2.72	0.59
	aug-TZ	0.009	0.74	1.50	0.60	0.63	1.60	0.48
DSD-PBEP86-D3	jun-DZ	0.021	1.78	2.48	1.44	1.28	3.02	1.07
	jul-DZ	0.017	0.88	2.63	1.18	0.74	2.60	0.76
	aug-DZ	0.016	0.96	2.39	1.15	0.84	2.64	0.81
	jun-TZ	0.008	0.54	1.50	0.60	0.44	1.69	0.43
	jul-TZ	0.007	0.46	1.39	0.57	0.37	1.55	0.41
	aug-TZ	0.007	0.59	1.53	0.58	0.49	1.73	0.44
	jun-DZ	0.023	1.80	2.53	1.49	1.31	2.98	1.11
	jul-DZ	0.017	0.89	2.64	1.15	0.75	2.58	0.77

rDSD-PBEP86-D3

143

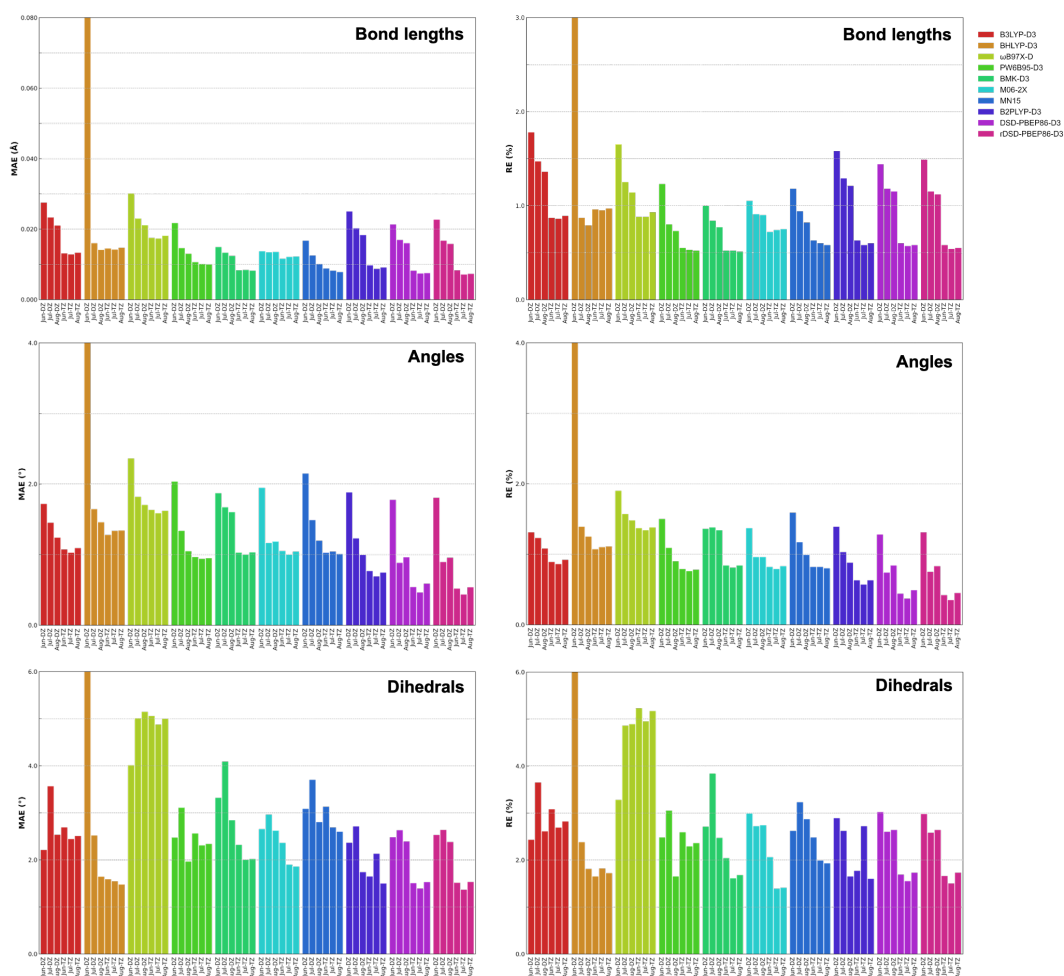|							
aug-DZ	0.016	0.96	2.38	1.12	0.83	2.64	0.80
jun-TZ	0.008	0.52	1.51	0.58	0.42	1.66	0.43
jul-TZ	0.007	0.43	1.36	0.54	0.35	1.50	0.40
aug-TZ	0.007	0.54	1.53	0.55	0.45	1.73	0.42

Figure S1: Mean Absolute Errors (MAE) and Relative Errors (RE) for bond lengths, valence and dihedral angles. The values are obtained by averaging absolute and relative errors of structural parameters over all the species involved in the reactive PES for the HCN⇌HNC isomerization assisted by two water molecules.

	HCN	$(H_2O)_2$	HCN$\cdots$$(H_2O)_2$	TS	HNC$\cdots$$(H_2O)_2$	HNC
E(CC)	-93.2774	-152.6856	-245.9753	-245.9254	-245.9574	-93.2540
ΔE_{MP2}^{∞}	-0.0446	-0.0818	-0.1268	-0.1262	-0.1264	-0.0444
ΔE_{MP2}^{CV}	-0.0944	-0.1029	-0.1974	-0.1974	-0.1973	-0.0943
E(junChS)	-93.4165	-152.8703	-246.2995	-246.2490	-246.2810	-93.3927

Table S3: junChS contributions (in Hartree) to energy evaluated on junChS geometries. E(CC) is the CCSD(T) energy computed with the jun-cc-pVTZ basis set; ΔE_{MP2}^{∞} is the difference between the fc-MP2/jun-cc-pVTZ energy and the corresponding extrapolated value estimated by the jun-cc-pVnZ basis sets with $n = $ T and Q. ΔE_{MP2}^{CV} accounts for core-valence correlation and is obtained as difference between ae- and fc- MP2 calculations with cc-pwCVTZ basis set. The final junChS energy is reported in the last row.

Table S4: AE and MAE of the formation energies computed at the junChS level on top of different optimized geometries with respect to full junChS results. The formation energy of each species (in kJ mol^{-1}) is calculated with respect to isolated HCN + $(H_2O)_2$. RC stands for reactant complex, i. e. NCH$\cdots$$(H_2O)_2$, TS for transition state and PC for product complex, i. e. CNH$\cdots$$(H_2O)_2$.

|	RC	TS	PC	HCN + $(H_2O)_2$	MAE
PW6B95-D3/jul-DZ	0.04	0.11	0.20	0.04	0.10
BHLYP-D3/aug-DZ	0.01	0.05	0.16	0.39	0.15
PW6B95-D3/aug-DZ	0.02	0.15	0.13	0.01	0.08
BMK-D3/aug-DZ	0.29	0.12	0.00	0.01	0.11
M06-2X/aug-DZ	0.09	0.41	0.34	0.01	0.21
MN15/aug-DZ	0.03	0.11	0.57	0.36	0.27
PW6B95-D3/jul-TZ	0.07	0.11	0.30	0.38	0.22
BMK-D3/jul-TZ	0.05	0.08	0.45	0.61	0.30
M06-2X/jul-TZ	0.04	0.23	0.12	0.42	0.20
MN15/jul-TZ	0.09	0.07	0.05	0.24	0.11
DSD-PBEP86-D3/jul-TZ	0.10	0.09	0.03	0.05	0.07
rDSD-PBEP86-D3/jul-TZ	0.12	0.06	0.02	0.05	0.06
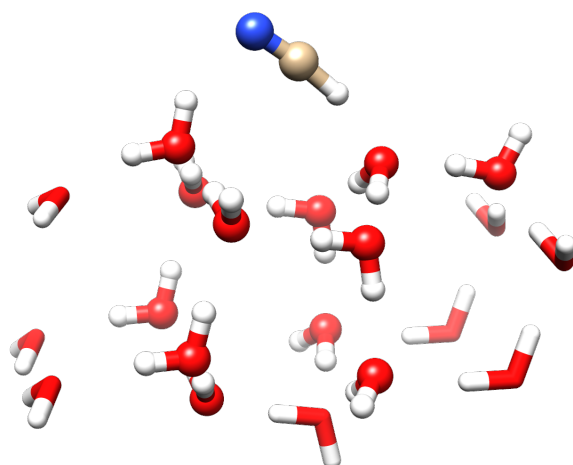
Figure S2: HCN···$(H_2O)_{20}$ reported in order to highlight molecules frozen during the optimization. Ball and stick representation used for atoms free to move while tubular representation for molecules kept frozen in order to prevent structural distortion of the cluster.

		HCN$\cdots$(H$_2$O)$_n$	TS	HNC$\cdots$(H$_2$O)$_n$	HNC + (H$_2$O)$_n$
junChS	n=2 [2]	-33.4	99.3	15.0	62.3
		6.3	**1.9**	**7.8**	**-1.3**
		(-27.2)	(101.2)	(22.8)	(61.0)
	n = 4 [20]	-61.4	43.4	-18.5	62.3
		10.6	**-9.3**	**8.6**	**-2.0**
		(-50.7)	(34.2)	(-9.9)	(60.2)
PW6B95	n=4 [20]	-64.1	27.6	-27.8	56.7
		10.6	**-9.3**	**8.6**	**-2.0**
		(-53.5)	(18.4)	(-19.2)	(54.7)
DSD-PBEP86	n=4 [20]	-65.4	28.3	-21.6	64.9
		7.9	**-8.9**	**8.1**	**-1.3**
		(-57.5)	(19.4)	(-13.4)	(63.6)

Table S5: Ground state electronic energies with respect to isolated HCN and (H$_2$O)$_n$. The number of water molecules involved in the relay mechanism is explicitly indicated together with the total number of water molecules included in the model system (in parenthesis). ZPVEs are in bold while energy corrected for ZPVEs are in parenthesis. Values in kJ mol^{-1}. junChS energy and geometries and ZPVEs at DSD-PBEP86-D3/jul-cc-pVTZ for $n = 2$. junChS energies for the adsorbate and the water molecules involved in the proton relay and PW6B95D3/jul-cc-pVDZ for the remaining water molecules in the cluster for $n = 4$ [20]. PW6B95-D3/jul-cc-pVDZ for PW6B95 results. DSD-PBEP86 results refers to ONIOM(DSD-PBEP86/jul-cc-pVTZ:PW6B95D3/jul-cc-pVDZ), both geometries and ZPVEs at this level.