

Published in final edited form as:

*J Chem Theory Comput.* 2016 November 08; 12(11): 5385–5397. doi:10.1021/acs.jctc.6b00379.

## State-of-the-art thermochemical and kinetic computations for astrochemical complex organic molecules: formamide formation in cold interstellar clouds as a case study

Fanny Vazart<sup>#</sup>, Danilo Calderini<sup>#</sup>, Cristina Puzzarini<sup>§</sup>, Dimitrios Skouteris<sup>#</sup>, and Vincenzo Barone<sup>#</sup>

<sup>#</sup>Scuola Normale Superiore, piazza dei Cavalieri 7, 56125 Pisa, Italy

<sup>§</sup>Dipartimento di Chimica "Giacomo Ciamician", Via F. Selmi 2, 40126 Bologna, Italy

### Abstract

We propose an integrated computational strategy aimed at providing reliable thermochemical and kinetic information on the formation processes of astrochemical complex organic molecules. The approach involves state-of-the-art quantum-mechanical computations, second-order vibrational perturbation theory, and kinetic models based on capture and transition state theory together with the master equation approach. Notably, tunneling, quantum reflection, and leading anharmonic contributions are accounted for in our model. Formamide has been selected as a case study in view of its interest as a precursor in the abiotic amino acid synthesis. After validation of the level of theory chosen for describing the potential energy surface, we have investigated several pathways of the OH+CH<sub>2</sub>NH and NH<sub>2</sub>+HCHO reaction channels. Our results indicate that both reaction channels are essentially barrier-less (in the sense that all relevant transition states lie below or only marginally above the reactants) and can, therefore, occur under the low temperature conditions of interstellar objects provided that tunneling is taken into the proper account.

### Introduction

One of the central assumptions of contemporary astrochemistry and astrobiology is that molecular and/or supra-molecular complexity gradually evolves from simple atomic or diatomic species up to biological macromolecules, the building blocks of life. In this framework, characterization of the first steps toward chemical complexity is a mandatory starting point for understanding how life emerged on the Earth in the early past or could evolve in other scenarios. The presence of some species possessing a strong prebiotic potential (e.g. glycolaldehyde or formamide) among the nearly 200 molecules identified till now in the interstellar medium, poses a formidable challenge to physical chemistry. As a matter of fact the significant amount (up to a few percent of the total carbon content) of relatively complex organic molecules in extremely cold (down to 10 K) regions where the density is extremely low (of the order of 10<sup>4</sup> particles/cm<sup>3</sup>) implies efficient chemical routes in harsh conditions, a real challenge for the standard concept of chemical kinetics. In this

scenario, computational modeling plays an increasing role, provided that reliable yet feasible approaches are available for estimating accurate thermochemical and kinetic features.

The starting point of an effective computational strategy is the accurate computation of structural, energetic, and vibrational features of all the intermediates and transition states involved in the considered reaction scheme. In the last years we have developed and validated an integrated approach employing last-generation hybrid and double hybrid models rooted in the density functional theory (DFT) to obtain molecular structures at which accurate energies are evaluated by means of the so-defined “golden standard” CCSD(T) method, also accounting for extrapolation to the complete basis set limit and core correlation.<sup>1–4</sup> Furthermore, we also demonstrated that DFT can be used to evaluate the energy (and possibly property) derivatives needed by vibrational second-order perturbation theory (VPT2) to compute accurate anharmonic frequencies and intensities.<sup>5</sup> In the present paper, we introduce the integration of accurate kinetic calculations in the well-tested approach mentioned above. As exact dynamics is unfeasible for systems with more than five atoms, kinetic calculations based on the Rice-Ramsperger-Kassel-Markus (RRKM) scheme, possibly combined with capture theory and master equation resolution, become indispensable to meet the required accuracy. Also in this case, leading anharmonic, tunneling, and non-classical reflection effects are effectively accounted for by second-order perturbation theory.

In the present paper we applied this integrated approach to the investigation of the possible scenario for the formamide formation in cold interstellar clouds. As a matter of fact formamide, as the simplest amide, has attracted increasing attention in the field of prebiotic chemistry.<sup>6–9</sup> Indeed, taking also into account that it has the ability to act as a precursor in the abiotic amino acid synthesis and perhaps also in that of nucleic acid bases.<sup>10–12</sup> Formamide is, therefore, a central compound to connect metabolism (conversion of energy), which is ruled by proteins, and genetics (passage of information), ruled by RNA and DNA. Moreover, its key role in the interstellar medium appears to be also remarkable since it has been detected in the galactic center sources Sgr A and Sgr B2,<sup>13</sup> in the Orion-KL region (an active site of high-mass star formation embedded in OMC-1)<sup>14</sup> and more recently in a solar-type protostar.<sup>15</sup> These detections of centimeter- and millimeter-wave emissions due to rotational transitions, led prebiotic chemists to wonder how formamide had been formed in space. This is a major issue since several formation pathways have already been investigated for formamide, including ionic reactions or radical ones.<sup>16,17</sup> However, none of these reaction channels is likely to be open in space, *i.e.* in the gas phase at low temperature. Indeed, the first step might exhibit a barrier and would require a non-available energy, a third body collision could be needed which is not likely in the gas phase or the path could lead mainly to other products. Therefore, to solve this issue, we have undertaken a comprehensive quantum mechanical investigation, focusing on the addition of the  $\cdot\text{OH}$  radical to methanimine, as suggested in a study by *Akbar Ali and coworkers*.<sup>18</sup> When looking carefully to the system, one can envisage an additional plausible path involving the addition of the  $\cdot\text{NH}_2$  radical to formaldehyde, already suggested in a study by *Kahane and coworkers*.<sup>15</sup> In detail, we investigated the vibrational signatures and thermochemical properties of all stable molecules involved in these processes and the transition states connecting them (Figure 1) using state-of-the-art computational approaches. Then, reaction

rates have been computed using capture and Rice–Ramsperger–Kassel–Marcus (RRKM) theories within a general semi-classical approach including anharmonic contributions together with tunneling and non-classical reflection effects.

## Computational details

Computations were mainly performed with the double-hybrid B2PLYP functional,<sup>19</sup> in conjunction with the m-aug-cc-pVTZ basis set,<sup>20,21</sup> where *d* functions on hydrogens have been removed. Semiempirical dispersion contributions were also included into DFT computations by means of the D3BJ model of Grimme, thus leading to the so-called B2PLYP-D3 computational model.<sup>22,23</sup> The nature of all stationary points found on the potential energy surface (PES) (minimum or transition state) was checked by diagonalizing the corresponding Hessians. Cubic and semi-diagonal quartic force constants were subsequently computed in a normal mode representation by finite differences of analytical second derivatives of energy and used to obtain anharmonic frequencies within second-order vibrational perturbation theory (VPT2).<sup>24,25</sup> To take into the proper account possible resonances for frequencies<sup>26</sup> and IR intensities, including both mechanical and electrical anharmonicities, the generalized VPT2 (GVPT2) model was employed.<sup>27,28</sup>

For all stationary points, improved electronic energies were obtained by means of a composite approach based on coupled-cluster theory employing the singles and doubles approximation (CCSD) augmented by a perturbative treatment of triple excitations, CCSD(T),<sup>29,30</sup> and implemented in CFOUR.<sup>31,32</sup> The composite approach is based on the additivity approximation and accounts for the following contributions: the Hartree-Fock self-consistent-field (HF-SCF) energy extrapolated to the complete basis-set (CBS) limit, the valence correlation energy at the CCSD(T) level extrapolated to the CBS limit as well, and the core-valence correlation correction. The correlation-consistent cc-p(C)V $n$ Z ( $n = T-5$ ) basis sets<sup>21,33</sup> were used throughout. The so-called CCSD(T)/CBS+CV energy is therefore given by:

$$E_{\text{CBS+CV}} = E^{\text{CBS}}(\text{HF} - \text{SCF}) + \Delta E^{\text{CBS}}(\text{CCSD(T)}) + \Delta E(\text{CV}) \quad (1)$$

In this equation, the exponential extrapolation formula of Feller<sup>34</sup> for the HF-SCF energy and the  $n^{-3}$  extrapolation scheme for the CCSD(T) correlation contribution<sup>35</sup> were employed. In the formula above,  $n=T$ , Q and 5 were chosen for the HF-SCF extrapolation, whereas  $n = T$  and Q were used for CCSD(T). Core-valence correlation effects were included by adding the corresponding energy correction,  $E(\text{CV})$ , obtained as the difference between the all-electron and frozen-core CCSD(T) energies using the core-valence cc-pCVTZ basis set. For selected intermediates and transition states,  $n = Q$  and 5 were also used for the extrapolation to the CBS limit of the CCSD(T) correlation energy, and the cc-pCVQZ basis set for evaluating the  $E(\text{CV})$  correction. For the TS5 and TS6 transition states, to further inspect the small magnitude of the barrier with respect to formaldehyde and NH<sub>2</sub>, the effects due to a full treatment of triple (fT) and quadruple excitations (fQ) were also accounted for. The corresponding corrections were obtained as differences between CCSDT and CCSD(T) and between CCSDTQ and CCSDT calculations within the frozen-

core approximation employing the cc-pVTZ and cc-pVDZ basis sets, respectively. To this purpose, single-point CC singles, doubles, and triples (CCSDT)<sup>36,37</sup> as well as CC singles, doubles, triples, and quadruples (CCSDTQ)<sup>38</sup> energy calculations were carried out using the MRCC package by Kállay<sup>39</sup> interfaced to CFOUR.

At the VPT2 level, the vibrational energy of asymmetric top molecules can be written <sup>27</sup>

$$E_{\text{vib}} = E_{\text{el}} + \text{ZPE} + \sum_i \frac{h}{2\pi} \omega_i n_i + \sum_{i < j} X_{ij} \left( \frac{1}{2} n_i + \frac{1}{2} n_j + n_i n_j \right) \quad (2)$$

where  $E_{\text{el}}$  is the electronic energy, ZPE the zero point vibrational energy,  $h$  is the Planck constant,  $n_j$  and  $\omega_j$  are the vibrational quantum numbers and harmonic wavenumbers, respectively;  $\chi_{ij}$  is the anharmonic coupling constant between modes  $i$  and  $j$ , which is an explicit function of the cubic and semi-diagonal quartic force constants. For transition states (TS), all sums exclude the mode corresponding to the imaginary frequency. This expression presumes that no resonance (Fermi, Darling-Dennison etc.) involves the vibrational levels considered (the existence of resonances is not expected to affect the statistical ensemble of levels).

A hindered rotor model was used for those intermediates that involve large amplitude internal rotations. In such cases, the vibrational Hamiltonian has been assumed to be adiabatically separable in a set of coupled anharmonic oscillators and a Large Amplitude Motion (LAM), i.e. the hindered rotation. It is well known that this “forced” separation leads to a variation of the effective moment of inertia along the path.<sup>40</sup> On the other hand, recovering this kinetic coupling introduces a coordinate-dependent mass that makes the Hamiltonian non Hermitian<sup>41</sup> and this generates some theoretical and numerical ambiguities.<sup>42</sup> In order to include the effective mass variation, we used a generalized coordinate approach, which leads to a unitary reduced mass.<sup>43–45</sup> In practical terms, we performed a relaxed scan along the dihedral angle that defines the internal rotation, starting from an initial geometry, we minimized the translational and rotational interactions of the successive geometry and we computed the length of the curve as the distance in mass-weighted cartesian coordinates between pairs of successive geometries. Finally, we fitted the curve obtained with a cubic spline to obtain more points and then we diagonalized the Hamiltonian matrix obtained in a Discrete Variable Representation (DVR).<sup>46</sup>

The semi-classical definition of reaction probability proposed by Miller,<sup>47</sup> which includes both tunneling and anharmonicity, was used in the present paper for the evaluation of reaction rates. Full kinetic calculations were performed by an in-house code described in previous papers.<sup>48–50</sup> The initial bimolecular rate constant leading from the reactants to the intermediate has been evaluated using capture theory calculations, after fitting the energy values for the approaching reactants to a  $1/R^6$  functional form and assuming that each successful capture leads to the intermediate.

As far as dissociation back to reactants is concerned, we have used a detailed balance argument, whereby the unimolecular rate constant for back-dissociation is given by the equation

$$k_{\text{back}} = k_{\text{capt}} \frac{\rho(\text{R})}{\rho(\text{I})} \quad (3)$$

where  $k_{\text{capt}}$  is the capture rate constant,  $\rho(\text{R})$  is the density of states per unit volume for the reactants and  $\rho(\text{I})$  is the density of states for the intermediate. On the other hand, where, as opposed to the reactants, a well-defined transition state exists, we have performed a RRKM calculation. The microcanonical rate constant is calculated using the formula

$$k(E) = \frac{N(E)}{h\rho(E)} \quad (4)$$

where  $N(E)$  denotes the sum of states in the transition state at energy  $E$ ,  $\rho(E)$  is the reactant density of states at energy  $E$  and  $h$  is Planck's constant. At the harmonic level, simple algorithms are available for computing number and densities of states, whereas tunneling and quantum reflection have been treated by a simple Eckart barrier parameterized from the energies of reactants and products, together with energy and imaginary frequency of the transition state.

More generally,  $N(E)$  is given by

$$N(E) = \sum_{n_1=0} \sum_{n_2=0} \cdots \sum_{n_{F-1}=0} P_n(E) \quad (5)$$

with the semiclassical tunneling probability expressed by

$$P_n(E) = \frac{1}{1 + e^{2\vartheta(n,E)}} \quad (6)$$

In the above equations,  $F-1$  vibrational degrees of freedom ( $F$  being the number of internal degrees of freedom, i.e.  $3N-6$  where  $N$  is the number of atoms) are orthogonal to the reaction path. In the framework of VPT2,  $\vartheta(n,E)$  is an explicit function of the harmonic and anharmonic force constants.<sup>47</sup> Using the so-called Wang-Landau algorithm,<sup>51</sup> it becomes possible to evaluate effectively both  $\rho(E)$  and  $N(E)$  by separating the energy range in small bins (with width  $\delta E$ ), as shown explicitly for  $N(E)$  in eq. 7.

$$N(E) = \delta E \sum_{i=1}^{E/\delta E} \rho(E_i) \langle P(E_i) \rangle \quad (7)$$

where  $\langle P(E_i) \rangle$  is the average tunneling probability for the different vibrational quantum number choices leading to  $E \in [E_i - \delta E, E_i + \delta E]$ .

The corresponding canonical expression for the rate constant is

$$K(T) = \frac{\int_{-\infty}^{+\infty} N(E) e^{-\frac{E}{k_B T}} dE}{h Q_{re}(T)} \quad (8)$$

where  $Q_{re}(T)$  is the reactants partition function.

In agreement with the principle of detailed balance, both in the canonical and microcanonical ensemble, the equilibrium constant is given by the ratio of the two rate constants (direct and inverse). Both densities of states (reactant and transition state) were appropriately symmetrized with respect to the number of identical configurations of the reactants and/or transition state.

After all calculations were performed, for each energy the branching ratio among all possible products and back-dissociation was determined and the corresponding capture rate constant was multiplied by this ratio to give the rate constant for the formation of each product. Finally, the rate constants were Boltzmann-averaged in order to provide the rate constants as a function of temperature.

All electronic structure computations (but the coupled-cluster ones) and kinetic calculations (except those based on capture theory) were performed with a development version of the Gaussian suite of programs.<sup>52</sup>

## Results and Discussion

### Validation of structural and vibrational results

We have recently obtained reliable and accurate semi-experimental equilibrium structures ( $r_e^{SE}$ ) for a set of more than 50 molecules containing the most important building blocks of organic and biological systems.<sup>53</sup> From that compilation we can estimate the accuracy of equilibrium structures obtained by means of different computational methods. Table 1 shows that the B2PLYP double hybrid functional in comparison to the typical global hybrid ones halves the root mean square error (RMSE) with respect to  $r_e^{SE}$  and approaches the level of accuracy of the most sophisticated post-Hartree-Fock models. Errors comparable to those obtained at the CCSD(T) level are noted, unless extrapolation to the CBS limit and core-correlation corrections are accounted for in conjunction with the latter. These errors on geometries lead to negligible errors on computed energetic and thermodynamic properties, as for instance demonstrated by refs. 49–51.

A further check of the reliability of our computational model (*i.e.*, B2PLYP-D3/m-aug-cc-pVTZ) has been performed by comparing simulated and experimental infrared (IR) spectra of formamide (see Figure 2).<sup>51</sup>

As one can see, the simulated spectrum fits nicely the experimental one: such a good agreement gives us further confidence to investigate the present reaction path with the chosen level of theory. The infrared data are available in Supporting Information.

### Choice of methodology

In the paper by *Akbar Ali and coworkers*,<sup>18</sup> optimized geometries of several compounds at both the BHandHLYP/aug-cc-pVTZ and CCSD(T)/aug-cc-pVTZ levels of theory have been reported. Thanks to these data, a validation of our B2PLYP-D3/m-aug-cc-pVTZ geometries is possible. Indeed, as we will see later, the optimized geometries of the loose transition states strongly depend on the computational level, whereas those of the tight transition states and intermediates do not exhibit large changes when using different methods.

To further benchmark the accuracy and reliability of the B2PLYP-D3/m-aug-cc-pVTZ energies, the CCSD(T)/CBS+CV approach has been employed, which is known to provide gas-phase thermodynamic properties with a 1 kJ/mol accuracy when used in conjunction with DFT optimized geometries.<sup>54–56</sup> It has also been demonstrated that geometries optimized at the B3LYP/SNS,<sup>48</sup> B2PLYP/aug-cc-pVTZ<sup>56</sup> and CCSD(T)/cc-pVTZ<sup>55</sup> levels are suitable for fulfilling such an accuracy.

Figure 3 shows the comparisons of the optimized geometries at different levels of theory for the **RI0** and **TS0<sub>a</sub>** species (see Figure 1). The molecular structures will be discussed in detail later, but at this stage we note that **RI0** involves a hydrogen bond between OH<sup>•</sup> and methanimine, while **TS0<sub>a</sub>** is a weak Van der Waals complex. In the case of **RI0**, one can see that the optimized geometries obtained with the different levels of theory are quite similar one another. However, when looking at **TS0<sub>a</sub>**, the BHandHLYP functional underestimates the distance between OH<sup>•</sup> and methanimine, while B3LYP overestimates it compared to CCSD(T). Inclusion of semiempirical dispersion (D3) and of a fraction of MP2 contributions (B2PLYP-D3 functional) restores a remarkable agreement with CCSD(T).

In Figure 4, the comparison between the optimized geometries of **RI0b-vW** and **TS5**, which are both Van der Waals complexes, obtained with the BHandHLYP, B2PLYP-D3 and B3LYP functionals is depicted. Once again, it is apparent that the geometries of the intermediate optimized at the different levels of theory considered are quite similar, while the BHandHLYP geometry of the transition state underestimates the distance between the approaching partners (in this case NH<sub>2</sub> and formaldehyde). In this connection, it is noteworthy that also B3LYP provides inaccurate geometries for **TS5**, but in this case the distances between NH<sub>2</sub> and formaldehyde is overestimated. This trend is possibly related to the percentage of HF exchange included in the different global hybrid functionals.

These two verifications, together with the results discussed in the preceding section and several recent studies,<sup>57–59</sup> let us be fully confident about geometrical structures and non potential energy contributions (i.e., zero point energies, entropies, etc.) obtained at the B2PLYP-D3/m-aug-cc-pVTZ level. Then, as mentioned above, electronic energies have been improved by means of CCSD(T) computations (CCSD(T)/CBS+CV).

## Mechanistic study

**Approach OH' and methanimine**—In Figure 5, one can see plausible paths concerning the approach of OH' to methanimine, with the selected energies collected in Table 2. Starting from the separated reactants, a complex stabilized by the formation of a hydrogen bond (**RI0**), more stable by 30.4 kJ/mol than the reactants, is formed. If we add zero-point corrections, this complex is found to be more stable than the precursors by 22.2 kJ/mol. In the paper by Akbar Ali et al.<sup>18</sup>, CCSD(T)/aug-cc-pVTZ calculations gave a ZPE-corrected relative energy of -19.8 kJ/mol for this compound. Two transition states, **TS0<sub>a</sub>** or **TS0<sub>b</sub>**, with relative electronic energies of -4.5 and -4.1 kJ/mol (5.0 and -3.3 with ZPE corrections here, 2.8 and -2.8 in Ref 18), respectively, can follow this first step and govern hydrogen abstraction and addition of OH to the  $\pi$  bond of methanimine, leading to the intermediates **RI1** (**RI1z** or **RI1e**) and to H<sub>2</sub>CN + H<sub>2</sub>O, respectively.

Under such circumstances, capture theory seems to be more appropriate than transition state theory for describing this step.

**Proposed full OH + CH<sub>2</sub>NH reaction path**—Figure 6 depicts the possible paths of the CH<sub>2</sub>NH + 'OH reaction (skipping details about the approach step) and reports the relative electronic (at the CCSD(T)/CBS+CV level) and zero-point corrected energies (electronic CCSD(T)/CBS+CV energies + ZPE issuing from B2PLYP-D3/m-aug-cc-pVTZ anharmonic vibrational calculations) of all minima and transition states. Starting from CH<sub>2</sub>NH and 'OH, a hydrogen abstraction from CH<sub>2</sub>NH by 'OH can be observed, thus leading to H<sub>2</sub>CN' and H<sub>2</sub>O that are around 120 kJ/mol more stable than the reactants. An addition of 'OH to the  $\pi$  bond of CH<sub>2</sub>NH can also occur, thus giving two potential intermediates, **RI1e** or **RI1z**, which are respectively 131.4 and 128.0 kJ/mol more stable than the precursors and connected by a weak barrier (*ca.* 5 kJ/mol, through **TS-RI1ez**). Then, both **RI1** intermediates can lose a hydrogen atom, thus leading directly to the methanimidic acid isomers, through *ca.* 125 and 120 kJ/mol barriers for the *Z*- and *E*-isomer, respectively. In the case of **RI1e**, the hydrogen atom can migrate from oxygen to nitrogen, to give the intermediate **RI3**. This step exhibits a  $\sim$ 125 kJ/mol barrier. At this point, three reaction channels can be followed. The first one is a lengthening of the C-N bond that would lead to the formation of 'NH<sub>2</sub> and formaldehyde through a  $\sim$ 80 kJ/mol barrier. The second one involves the loss of the hydrogen atom linked to carbon, thus leading directly to formamide through a slightly lower barrier (by  $\sim$ 4 kJ/mol). As far as the third reaction path is concerned, the hydrogen atom linked to carbon migrates to the oxygen atom. This step exhibits a *ca.* 113 kJ/mol barrier and leads to the intermediate **RI5**. The latter is then able to lose the hydrogen atom bonded either to nitrogen or to the oxygen atom. The first possibility would give *E*- or *Z*-methanimidic acid, through *ca.* 138 and 155 kJ/mol barriers, respectively, and the second would lead to formamide, through a  $\sim$ 120 kJ/mol barrier. If we focus on the possible products, formamide + H is the most stable one, with a relative energy of -89.5 kJ/mol. The following ones are *E*-methanimidic acid + H and formaldehyde + NH<sub>2</sub>, at -44.8 and -43.5 kJ/mol respectively, and then the less stable product is *Z*-methanimidic acid + H that exhibits a relative energy of -30.1 kJ/mol.



It is noteworthy that all compounds involved in the reaction scheme of Figure 6 are lower in energy than the reactants, which makes the entire path viable in the interstellar medium. The mechanisms corresponding to these paths are depicted in Figure 7 and 8.

Let us now focus on the structural variations along the reaction paths. Table 3 reports the bond lengths and symmetry point group of all investigated compounds obtained from B2PLYP-D3/m-aug-cc-pVTZ geometry optimizations. One can first notice that the C-O and C-N bonds fluctuate significantly, while the O-H, N-H and C-H bonds remain close in distance (except for the transition states corresponding to H migrations and loss).

Let us first discuss the C-O bonds. In formaldehyde and formamide, the C-O distances, about 1.21 Å, are those typical of a double bond. In the methanimidic acid isomers, C-O bonds are single bonds, linked to a C-N double bond, thus exhibiting a length of 1.35 Å. If we analyze the intermediates and transition states, we can classify C-O bond lengths into three types. The first type, with C-O bond lengths around 1.25 Å, includes **TS5**, **TS6** and **TS7**. All three transition states lead directly to formaldehyde and formamide and exhibit therefore a formal C-O double bond, with a weak influence of the hydrogen atom. The second type, with C-O bond distances of about 1.35 Å, is present in **RI3**, both **TS3** transition states, **TS4** and both **TS8** transition states. The four **TS3** and **TS8** compounds evolve in methanimidic acid isomers and exhibit therefore similar C-O bond lengths. The C-O bond in **RI3** consists in a CO<sup>•</sup> radical moiety, and in **TS4** the hydrogen atom is migrating from C to the O atom. This explains why in these compounds the C-O distance is in between a single and a double bond. The last class, with C-O bond lengths around 1.40 Å, includes both **RI1** isomers, the transition state linking them (i.e. **TS-RI1ez**), **TS2** and **RI5**. In all these five compounds, an OH moiety is linked to a C-N single bond, therefore the C-O bond is a formal single bond as well.

Let us now consider how the C-N bond varies along the reaction. Once again, three types of bonds are observed. The first one, with the C-N bond lengths around 1.25 Å, is present in compounds that exhibit a formal double C-N bond: methanimine, both **TS3** transition states, both **TS8** transition states, H<sub>2</sub>CN and both methanimidic acid isomers. The second type is characterized by C-N bond lengths of about 1.37 Å. It is found in **TS4**, **TS6**, **TS7** and formamide, that is to say in compounds with a delocalization that makes the C-N bond lying somewhat in between a single and a double bond. The last class of C-N bond, which includes both **RI1** isomers, the connecting transition state (i.e. **TS-RI1ez**), **TS2**, **RI3** and **RI5**, exhibits a formal single bond of about 1.45 Å.

As far as the symmetry is concerned, it is noticeable that out of 14 intermediates and transition states, only 3 compounds have a *C<sub>S</sub>* group symmetry. However, all the precursors and products have relatively high symmetries (*C<sub>S</sub>*, *C<sub>∞v</sub>*, *C<sub>2v</sub>*).

**Approach NH<sub>2</sub><sup>•</sup> and formaldehyde**—Let us now consider an alternative path involving NH<sub>2</sub><sup>•</sup> and formaldehyde as precursors for formamide.

As in the case of the previous approach, CCSD(T)/CBS+CV relative energies are considered together with those at the B2PLYP-D3/m-aug-cc-pVTZ and CCSD(T)/m-aug-cc-pVTZ levels. The most relevant results are collected in Table 4.

Starting from the reactants either a weak van der Waals complex (**RI0b-vW**) or an intermediate stabilized by a hydrogen bond (**RI0b-Hbond**) can be formed. Both of them are followed by a low lying, first order saddle point (**TS5**) having an energy very close to that of reactants, which then leads to an intermediate (**RI3**). Once again, one can notice that this first step, leading to **RI3**, is nearly barrier-less, our best estimate for the transition state height with respect to reactants being 3.8 kJ/mol. Inclusion of the full-T (CBS+CV+fT) and full-Q (CBS+CV+fT+fQ) corrections further lowers this barrier to 2.05 and 1.67 kJ/mol, respectively.

Therefore, capture theory seems more appropriate than transition state theory to describe this step.

Figure 10 displays the possible path of the  $\text{CH}_2\text{O} + \cdot\text{NH}_2 \rightarrow \text{NH}_2\text{CHO} + \text{H}$  formation reaction (skipping again details about the approach step) and the relative electronic (CCSD(T)/CSB+CV) and zero-point corrected energies (ZPE at the B2PLYP-D3/m-aug-cc-pVTZ level) of all minima and transition states. In the case of the intermediate **RI3**, the hindered  $\text{NH}_2$  rotation needs to be carefully managed. To this end, all normal modes except that corresponding to the  $\text{NH}_2$  torsion have been included in a GVPT2 treatment, whereas a one-dimensional numerical approach has been used for the  $\text{NH}_2$  hindered rotation (HR). The overall zero point energy (ZPE) is then obtained by summing  $\text{ZPE}_{\text{GVPT2}}$  and  $\text{ZPE}_{\text{HR}}$ .

After the formation of the intermediate **RI3**, hydrogen loss can be observed, leading to formamide and the  $\cdot\text{H}$  radical through the transition state **TS6** that has a barrier of 73.2 kJ/mol. The rotation around the C-N bond in **RI3** can lead to another energy minimum, which forms again formamide through the same transition state **TS6**. The products were found 46 kJ/mol more stable than the reactants. The possible radical mechanism based on the proposed reaction path is given in Figure 11.

If we take a look at the structural variations along this reaction path (see Table 3, values in bold), it is evident that the C-O and C-N bonds fluctuate. Indeed, the C-O bonds in formaldehyde and formamide are formally double bonds, while in the first intermediate **RI3**, where the  $\text{NH}_2$  moiety is linked to the carbon atom, the CO bond becomes almost a single one. In the case of the transition state **TS6**, the presence of the H atom influences the C-O distance because it is a little bit longer than in formamide. As far as C-N is concerned, it is a single bond in the intermediate **RI3**, and then shortens once the H atom is removed, to reach in formamide a length in between the typical single and double bond. This observation can be explained by the delocalization that the amide moiety undergoes.

It is also noted that this reaction occurs keeping a relatively high symmetry during the process. Indeed, after an addition of two  $C_{2v}$  moieties, a compound of  $C_s$  symmetry is obtained, which proceeds toward the product molecule, of  $C_s$  symmetry as well, after the last transition state.

Let us now have a deeper insight into the hindered rotation occurring in the intermediate **RI3** that we mentioned above. Figure 12 shows the potential energy profile along the  $\varphi$  dihedral angle, which describes the  $\text{NH}_2$  rotation. The energies of minima and saddle points are given in Supporting Information.

In addition to the absolute energy minimum, the potential energy surface shows two additional equivalent local minima. All these minima correspond to staggered conformations and are connected by three transition states, corresponding to eclipsed conformations. The equivalent conformers **a** and **c** are separated by **b**, which is the highest transition state, with a barrier of *ca.* 3.8 kJ/mol. The most stable conformer (**e**) is found to be about 10.5 kJ/mol more stable than the other minima. The **d** and **f** equivalent transition states rule the **c** to **e** and **e** to **a** transformations and have a relative energy about 14.6 kJ/mol higher than the most stable conformer **e**.

### Kinetics of the OH + CH<sub>2</sub>NH path

The rate of back dissociation of the initial complex (assuming the validity of capture theory and using the principle of detailed balance) at a particular energy is proportional to the density of states (per unit volume) of the reactants. As the translational component of this density of states is proportional to the square root of the translational energy, at low energies, back dissociation of the initial complex is suppressed due to the reactant density of states tending to zero. As a result, at low energies the initial van der Waals complex is present with the chance to either rearrange to HOCH<sub>2</sub>NH (with subsequent possible formation of methanimidic acid) or eliminate a water molecule forming the H<sub>2</sub>CN radical. The other two possible products (formamide+H or formaldehyde+NH<sub>2</sub>) have negligible rates due to the fact that more than one steps are required to reach them from the original complex.

Both transition states corresponding to these channels lie *ca.* 4.5 kJ/mol below the reactant zero point energy. However, the TS leading to water elimination presents a higher sum of states than the second one at low energies, leading to a low-energy dominance of H<sub>2</sub>O elimination. This can be seen in Figures 13 and 14 as a constant ratio of around 8 between the rate constants of water elimination and methanimidic acid formation. It is to be noted that *Z*-methanimidic acid has a much lower rate of formation than its *E*-isomer. This is essentially due to the fact that the corresponding transition state is rather lower in energy. Moreover, as the two isomers of HOCH<sub>2</sub>NH (which dissociate to the corresponding isomers of methanimidic acid) equilibrate rapidly between themselves and the *E*-isomer is more stable, this is another factor contributing to this effect.

As the energy increases, the reactant density of states increases rapidly and this brings about a rapid increase in the rate of redissociation of the complex. Indeed, the two next highest rate constants diminish with increasing energy because of this effect (on the other hand, the rate constant of *Z*-methanimidic acid continues to increase slightly).

The same effects are also seen in the canonical rate constants, albeit smeared out due to Boltzmann averaging. Despite the predominance of redissociation, H<sub>2</sub>CN and *E*-methanimidic acid formation remain important processes up to 300K (with rate constants above  $10^{-11} \text{ cm}^3\text{s}^{-1}$ ).

In a previous paper,<sup>4</sup> a similar RRKM study was performed on the  $\text{NH}_2+\text{HCHO}$  reaction, based on calculations at the CBS-QB3 level of theory. This enabled us to assure that, even at low temperatures, the formation of formamide through that reaction was largely predominant over the back-dissociation into  $\text{NH}_2+\text{HCHO}$ . This result was confirmed by new RRKM calculations on this system using the CCSD(T)/CBS+CV energies, as one can see in Supporting Information. The rate constants are slightly lower than in the previous study (it drops from  $3.1 \times 10^{-10}$  to  $7.6 \times 10^{-11} \text{ cm}^3 \text{ s}^{-1}$  at 5 K), which is explained by a higher barrier for the transformation of **RI3** into formamide + H (represented by **TS6**), but the formamide formation still predominates at low temperatures since the reactants densities of states tend to zero. Moreover, when including the full-T (CBS+CV+fT) and full-Q (CBS+CV+fT+fQ) corrections, the **TS6** relative energy that was found at -1.3 kJ/mol drops to -2.3 and -2.8 kJ/mol respectively, which leads to slightly higher rate constants ( $9.9 \times 10^{-11} \text{ cm}^3 \text{ s}^{-1}$  at 5 K with CCSD(T)/CBS+CV+fT+fQ energies).

## Conclusion and perspectives

In this paper, we have provided new insights concerning the formation of formamide in the interstellar medium. Our computations allowed us to suggest two reaction paths combined to probable mechanisms concerning this formation. In both cases,  $\text{OH}+\text{CH}_2\text{NH}$  and  $\text{NH}_2+\text{HCHO}$ , the addition step does not involve significant barriers and can therefore occur in space. While RRKM calculations confirmed the effectiveness of the  $\text{NH}_2+\text{HCHO}$  reaction (since, once this first addition has been done, formamide formation largely predominates over back-dissociation at low energies), the  $\text{OH}+\text{CH}_2\text{NH}$  reaction leads preferentially to other products of prebiotic interest such as methanimidic acid and the  $\text{H}_2\text{CN}$  radical. Together with the specific interest of the studied system, the reliability, robustness, and reasonable computational cost of the proposed computational strategy paves the route toward the study of larger systems of interest for astrochemistry and astrobiology.

## Supplementary Information

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgement

The research leading to these results has received funding from the European Union's Seventh Framework Programme (FP/2007-2013) under the grant agreement no. ERC-2012-AdG-320951-DREAMS. This work was supported by Italian MIUR (PRIN 2012 "STAR: Spectroscopic and computational Techniques for Astrophysical and atmospheric Research") and by the University of Bologna (RFO funds). The authors thank Dr. Camille Latouche (Univ. Nantes) for useful discussions. The support of the COST CMTS Actions CM1405 (MOLIM: MOLEcules In Motion) and CM1404 (Our Astro-Chemical History) is also acknowledged.

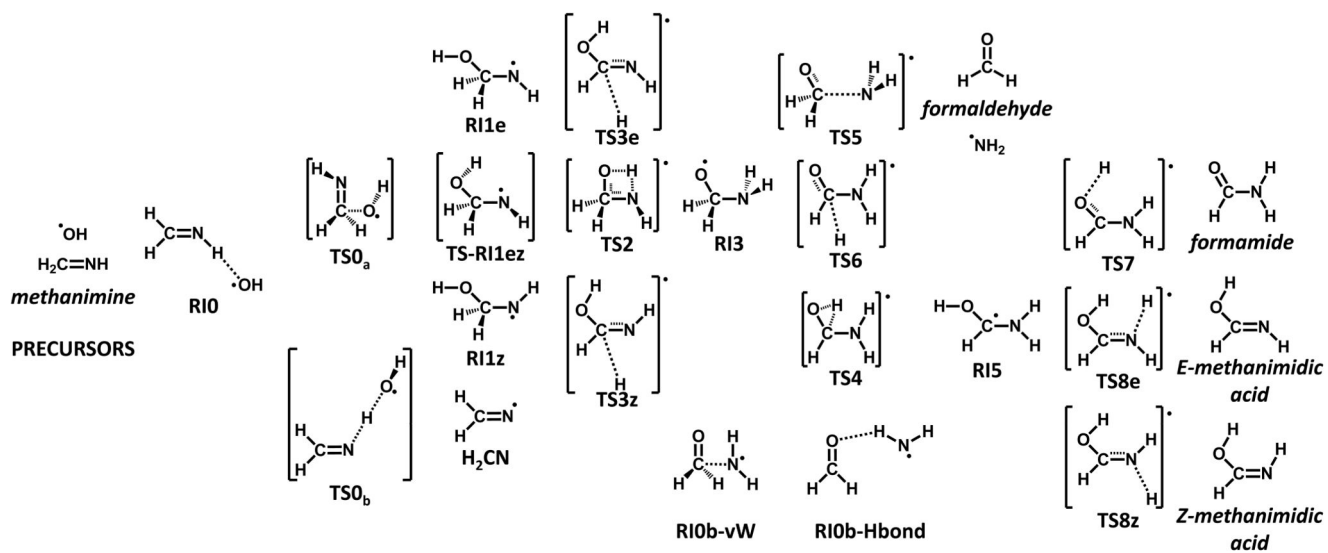
## References

- (1). Barone V, Biczysko M, Bloino J, Puzzarini C. The Performance of Composite Schemes and Hybrid CC/DFT Model in Predicting Structure, Thermodynamic and Spectroscopic Parameters: The Challenge of the Conformational Equilibrium in Glycine. *Phys Chem Chem Phys*. 2013; 15(25):10094–10111. [PubMed: 23599122]
- (2). Puzzarini C, Penocchio E, Biczysko M, Barone V. Molecular Structure and Spectroscopic Signatures of Acrolein: Theory Meets Experiment. *J Phys Chem A*. 2014; 118(33):6648–6656. [PubMed: 24842714]

- (3). Puzzarini C, Biczysko M, Barone V, Largo L, Peña I, Cabezas C, Alonso JL. 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(3):534–540. [PubMed: 26276605]
- (4). Barone V, Biczysko M, Bloino J, Cimino P, Penocchio E, Puzzarini C. CC/DFT Route toward Accurate Structures and Spectroscopic Features for Observed and Elusive Conformers of Flexible Molecules: Pyruvic Acid as a Case Study. *J Chem Theory Comput.* 2015; 11(9):4342–4363. [PubMed: 26575928]
- (5). 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(3): 1011–1018. [PubMed: 26765363]
- (6). Bada JL, Cleaves HJ. Ab Initio Simulations and the Miller Prebiotic Synthesis Experiment. *Proc Natl Acad Sci.* 2015; 112(4):E342. [PubMed: 25572968]
- (7). Jeilani YA, Nguyen HT, Cardelino BH, Nguyen MT. Free Radical Pathways for the Prebiotic Formation of Xanthine and Isoguanine from Formamide. *Chem Phys Lett.* 2014; 598(0):58–64.
- (8). Saladino R, Botta G, Pino S, Costanzo G, Di Mauro E. Genetics First or Metabolism First? The Formamide Clue. *Chem Soc Rev.* 2012; 41(16):5526. [PubMed: 22684046]
- (9). Barone V, Latouche C, Skouteris D, Vazart F, Balucani N, Ceccarelli C, Lefloch B. Gas Phase Formation of the Prebiotic Molecule Formamide: Insights from New Quantum Computations. *Mon Not R Astron Soc.* 2015; 453:L31–L35.
- (10). Halfen DT, Ilyushin V, Ziurys LM. Formation of Peptide Bonds in Space: A Comprehensive Study of Formamide and Acetamide in Sgr B2(N). *Astrophys J.* 2011; 743(2001):60.
- (11). Pino S, Sporer J, Costanzo G, Saladino R, Mauro E. From Formamide to RNA, the Path Is Tenuous but Continuous. *Life.* 2015; 5:372–384. [PubMed: 25647486]
- (12). Saladino R, Crestini C, Pino S, Costanzo G, Di Mauro E. Formamide and the Origin of Life. *Phys Life Rev.* 2012; 9(1):84–104. [PubMed: 22196896]
- (13). Gottlieb, Ca, Palmer, P., Rickard, L.J., Zuckerman, B. Studies of Interstellar Formamide. *Astrophys J.* 1973; 182:699.
- (14). Motiyenko RA, Tercero B, Cernicharo J, Margulès L. Rotational Spectrum of Formamide up to 1 THz and First ISM Detection of Its  $\nu_{12}$  Vibrational State. *A&A.* 2012; 548:A71.
- (15). Kahane C, Ceccarelli C, Faure a, Caux E. Detection of Formamide, the Simplest But Crucial Amide, in a Solar-Type Protostar. *Astrophys J.* 2013; 763:L38.
- (16). Redondo P, Barrientos C, Largo A. Some Insights Into Formamide Formation Through Gas-Phase Reactions in the Interstellar Medium. *Astrophys J.* 2014; 780(2011):181.
- (17). Jones BM, Bennett CJ, Kaiser RI. Mechanical Studies on the Production of Formamide ( $\text{H}_2\text{NCHO}$ ) Within Interstellar Ice Analogs. *Astrophys J.* 2011; 734:78.
- (18). Ali MA, Barker JR. Comparison of Three Isoelectronic Multiple-Well Reaction Systems:  $\text{OH} + \text{CH}_2\text{O}$ ,  $\text{OH} + \text{CH}_2\text{CH}_2$ , and  $\text{OH} + \text{CH}_2\text{NH}$ . *J Phys Chem A.* 2015; 119(28):7578–7592. [PubMed: 25839620]
- (19). Grimme S. Semiempirical Hybrid Density Functional with Perturbative Second-Order Correlation. *J Chem Phys.* 2006; 124(3):34108.
- (20). Papajak E, Leverentz HR, Zheng J, Truhlar DG. Efficient Diffuse Basis Sets: Cc-pVxZ+ and Maug-Cc-pVxZ. *J Chem Theory Comput.* 2009; 5(5):1197–1202. [PubMed: 26609710]
- (21). Dunning TH. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen. *J Chem Phys.* 1989; 90(2):1007.
- (22). Goerigk L, Grimme S. Efficient and Accurate Double-Hybrid-Meta-GGA Density Functionals—Evaluation with the Extended GMTKN30 Database for General Main Group Thermochemistry, Kinetics, and Noncovalent Interactions. *J Chem Theory Comput.* 2011; 7(2):291–309. [PubMed: 26596152]
- (23). Grimme S, Ehrlich S, Goerigk L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J Comput Chem.* 2011; 32(7):1456–1465. [PubMed: 21370243]
- (24). Mills, IM. 3.2 - Vibration–Rotation Structure in Asymmetric- and Symmetric-Top Molecules. *Molecular Spectroscopy.* Mathews, KNRW., editor. Academic Press; 1972. p. 115-140.

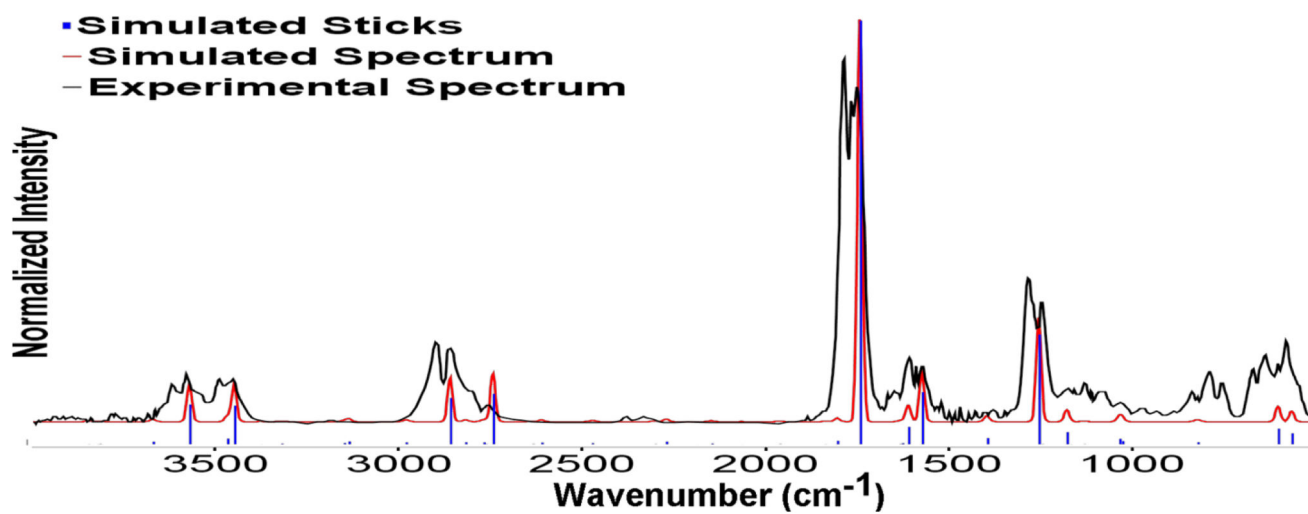
- (25). Hoy AR, Mills IM, Strey G. Anharmonic Force Constant Calculations. *Mol Phys.* 1972; 24(6): 1265–1290.
- (26). Barone V. Anharmonic Vibrational Properties by a Fully Automated Second-Order Perturbative Approach. *J Chem Phys.* 2005; 122
- (27). Bloino J, Biczysko M, Barone V. General Perturbative Approach for Spectroscopy, Thermodynamics, and Kinetics: Methodological Background and Benchmark Studies. *J Chem Theory Comput.* 2012; 8(3):1015–1036. [PubMed: 26593363]
- (28). 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(15):948–982. [PubMed: 26345131]
- (29). Raghavachari K, Trucks GW, Pople JA, Head-Gordon M. A Fifth-Order Perturbation Comparison of Electron Correlation Theories. *Chem Phys Lett.* 1989; 157(6):479–483.
- (30). Watts JD, Gauss J, Bartlett RJ. Coupled-cluster Methods with Noniterative Triple Excitations for Restricted Open-shell Hartree–Fock and Other General Single Determinant Reference Functions. Energies and Analytical Gradients. *J Chem Phys.* 1993; 98(11)
- (31). Stanton JF, Gauss J, Harding ME, Szalay PG. CFOUR A Quantum Chemical Program Package. 2011 with contributions from A. A. Auer, R. J. Bartlett, U. Benedikt, C. Berger, D. E. Bernholdt, Y. J. Bomble, O. Christiansen, M. Heckert, O. Heun, C. Huber, T.-C. Jagau, D. Jonsson, J. Jus\_elius, K. Klein, W. J. Lauderdale, F. Lipparini, D. Matthews, T. Metzroth, L. A. Mueck, D. P. O'Neill, D. R. Price, E. Prochnow, C. Puzzarini, K. Ruud, F. Schiffmann, W. Schwalbach, S. Stopkowitz, A. Tajti, J. Vazquez, F. Wang, J. D. Watts and the integral packages MOLECULE (J. Almloef and P. R. Taylor), PROPS (P. R. Taylor), ABACUS (T. Helgaker, H. J. Aa. Jensen, P. Jorgensen, and J. Olsen), and ECP routines by A. V. Mitin and C. van Wuelen. For the current version, see <http://www.cfour.de>.
- (32). Heckert M, Kállay M, Tew DP, Klopper W, Gauss J. Basis-Set Extrapolation Techniques for the Accurate Calculation of Molecular Equilibrium Geometries Using Coupled-Cluster Theory. *J Chem Phys.* 2006; 125(4)
- (33). Woon DE, Dunning TH. Gaussian Basis Sets for Use in Correlated Molecular Calculations. V. Core-valence Basis Sets for Boron through Neon. *J Chem Phys.* 1995; 103(11)
- (34). 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(9)
- (35). Helgaker T, Klopper W, Koch H, Noga J. Basis-Set Convergence of Correlated Calculations on Water. *J Chem Phys.* 1997; 106(23)
- (36). Noga J, Bartlett RJ. The Full CCSDT Model for Molecular Electronic Structure. *J Chem Phys.* 1987; 86(12)
- (37). Scuseria GE, Schaefer HF. A New Implementation of the Full CCSDT Model for Molecular Electronic Structure. *Chem Phys Lett.* 1988; 152(4):382–386.
- (38). Kállay M, Surján PR. Higher Excitations in Coupled-Cluster Theory. *J Chem Phys.* 2001; 115(7)
- (39). Kállay, M., Rolik, Z., Ladjanszki, I., Szegedy, L., Ladoczki, B., Csontos, J., Kornis, B. MRCC, a Quantum Chemical Program Suite. See Also Rolik Z, Kállay M. *J Chem Phys.* 2011; 135:104111. as well as [www.mrcc.hu](http://www.mrcc.hu). [PubMed: 21932880]
- (40). Aquino N, Campoy G, Yee-Madeira H. The Inversion Potential for NH<sub>3</sub> Using a DFT Approach. *Chem Phys Lett.* 1998; 296(1–2):111–116.
- (41). von Roos O. Position-Dependent Effective Masses in Semiconductor Theory. *Phys Rev B.* 1983; 27(12):7547–7552.
- (42). Bonatsos D, Georgoudis PE, Minkov N, Petrellis D, Quesne C. Bohr Hamiltonian with a Deformation-Dependent Mass Term for the Kratzer Potential. *Phys Rev C.* 2013; 88(3):34316.
- (43). Miller WH, Handy NC, Adams JE. Reaction Path Hamiltonian for Polyatomic Molecules. *J Chem Phys.* 1980; 72(1)
- (44). Barone V, Grand A, Minichino C, Subra R. Vibrational Modulation Effects on the Hyperfine Coupling Constants of Fluoromethyl Radicals. *J Chem Phys.* 1993; 99(9)
- (45). Minichino C, Barone V. From Concepts to Algorithms for the Characterization of Reaction Mechanisms. H<sub>2</sub>CS as a Case Study. *J Chem Phys.* 1994; 100(5)

- (46). Colbert DT, Miller WH. A Novel Discrete Variable Representation for Quantum Mechanical Reactive Scattering via the S-matrix Kohn Method. *J Chem Phys.* 1992; 96(3)
- (47). Miller WH, Hernandez R, Handy NC, Jayatilaka D, Willetts A. Ab Initio Calculation of Anharmonic Constants for a Transition State, with Application to Semiclassical Transition State Tunneling Probabilities. *Chem Phys Lett.* 1990; 172(1):62–68.
- (48). Leonori F, Skouteris D, Petrucci R, Casavecchia P, Rosi M, Balucani N. Combined Crossed Beam and Theoretical Studies of the C(<sup>1</sup>D) + CH<sub>4</sub> Reaction. *J Chem Phys.* 2013; 138:24311.
- (49). Leonori F, Petrucci R, Balucani N, Casavecchia P, Rosi M, Skouteris D, Berteloite C, Le Picard SD, Canosa A, Sims IR. Crossed-Beam Dynamics, Low-Temperature Kinetics, and Theoretical Studies of the Reaction S(<sup>1</sup>D)+C<sub>2</sub>H<sub>4</sub>. *J Phys Chem A.* 2009; 113:15328–15345. [PubMed: 19761231]
- (50). Vazart F, Latouche C, Skouteris D, Balucani N, Barone V. A Quantum Mechanical and RRKM Study on the Mechanisms and Rate Constants for the Formation of Cyanomethanimine Isomers in Cold Interstellar Clouds. *Astrophys J.* 2015
- (51). Nguyen TL, Barker JR. Sums and Densities of Fully Coupled Anharmonic Vibrational States: A Comparison of Three Practical Methods. *J Phys Chem A.* 2010; 114(10):3718–3730. [PubMed: 20170143]
- (52). Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, et al. Gaussian09 GDVI03. 2015
- (53). 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(10):2058–2082. [PubMed: 25648634]
- (54). Webook.nist.gov. IR-Formamide. May 20.2015
- (55). Puzzarini C. Isomerism of Cyanomethanimine: Accurate Structural, Energetic, and Spectroscopic Characterization. *J Phys Chem A.* 2015; 119(47):11614–11622. [PubMed: 26529434]
- (56). Biczysko M, Panek P, Scalmani G, Bloino J, Barone V. Harmonic and Anharmonic Vibrational Frequency Calculations with the Double-Hybrid B2PLYP Method. *J Chem Theory Comput.* 2010; 6(7):2115–2125. [PubMed: 26615939]
- (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. [PubMed: 26615939]
- (58). Penocchio E, Piccardo M, Barone V. Semiexperimental Equilibrium Structures for Building Blocks of Organic and Biological Molecules: The B2PLYP Route. *J Chem Theory Comput.* 2015; 11(10):4689–4707. [PubMed: 26574259]
- (59). Vazart F, Latouche C, Cimino P, Barone V. Accurate Infrared (IR) Spectra for Molecules Containing the C≡N Moiety by Anharmonic Computations with the Double Hybrid B2PLYP Density Functional. *J Chem Theory Comput.* 2015; 11(9):4364–4369. [PubMed: 26575929]

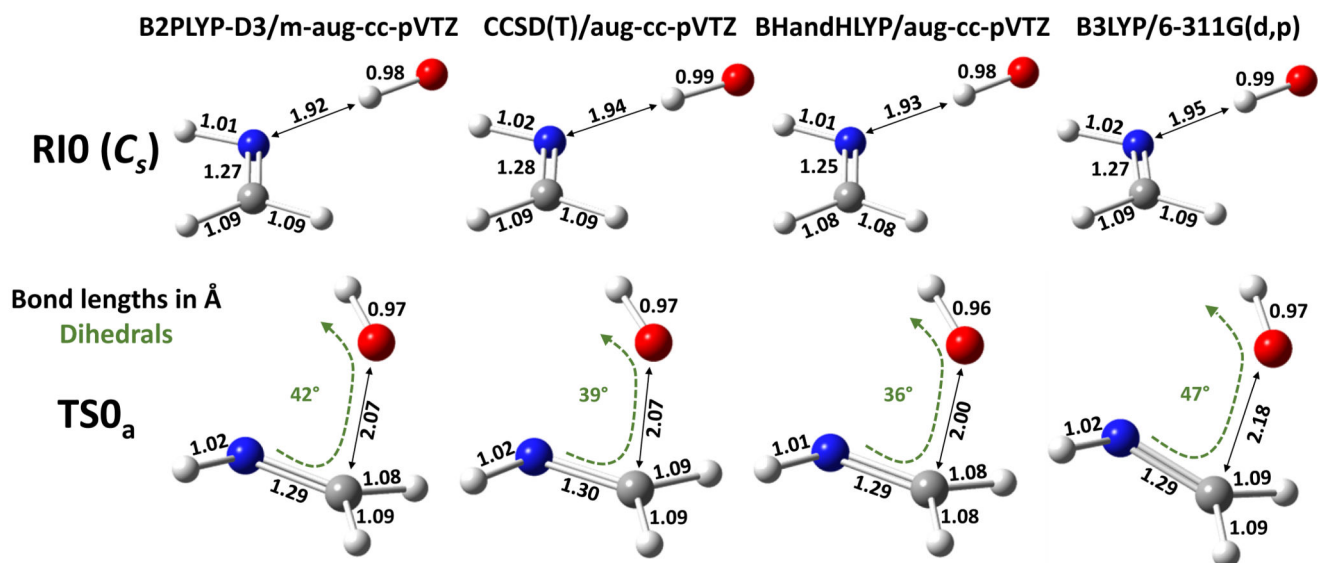


**Figure 1.**  
Investigated compounds

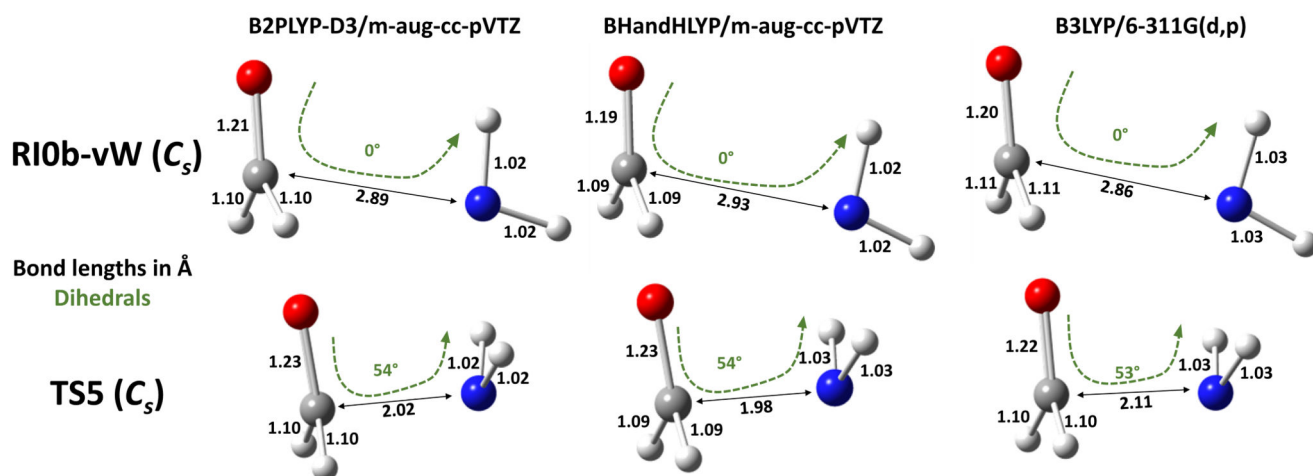




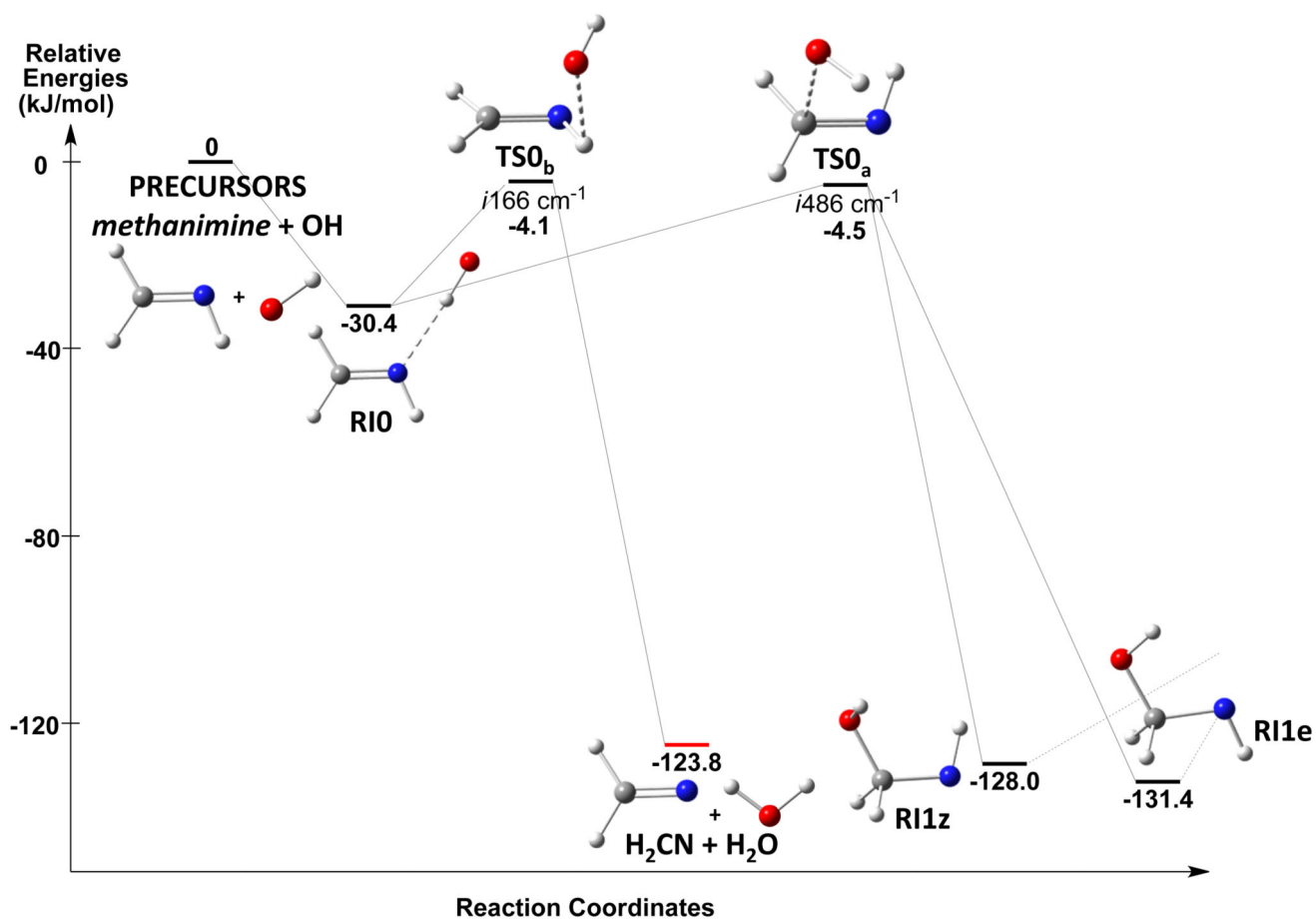
**Figure 2.** Simulated anharmonic (red, B2PLYP-D3/m-aug-cc-pVTZ) and Experimental (black) IR spectra of formamide in vacuum. Experimental spectrum taken from Ref. 54.

**Figure 3.**

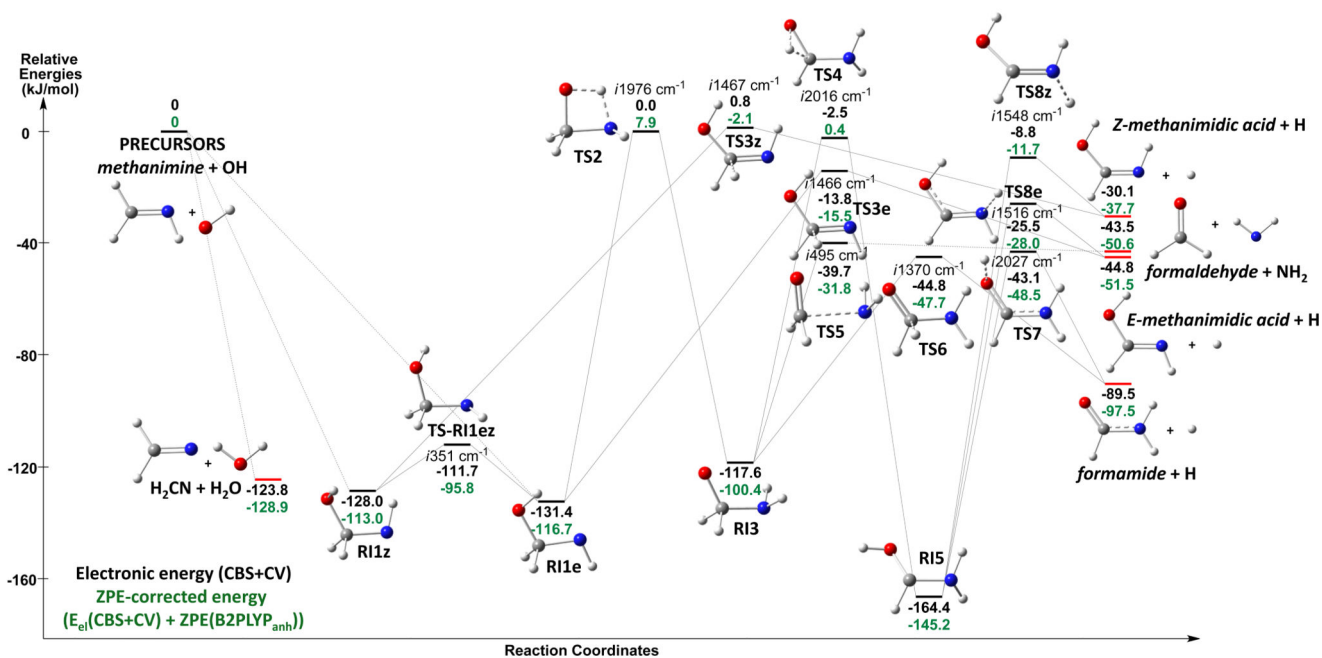
Comparisons between the optimized geometries of **RI0** and  **$TSO_a$**  obtained at the B2PLYP-D3/m-aug-cc-pVTZ, CCSD(T)/aug-cc-pVTZ, BHandHLYP/aug-cc-pVTZ and B3LYP/6-311G(d,p) levels of theory.



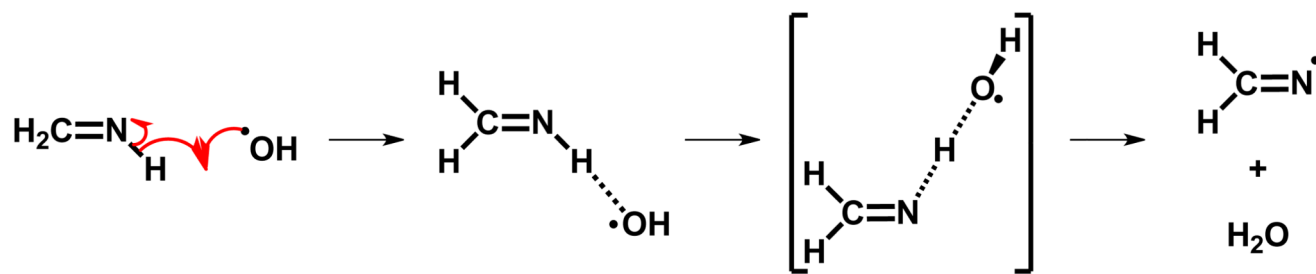
**Figure 4.** Comparisons between the optimized geometries of **RI0b-vW** and **TS5** obtained at the B2PLYP-D3/m-aug-cc-pVTZ, BHandHLYP/m-aug-cc-pVTZ and B3LYP/6-311G(d,p) levels of theory.



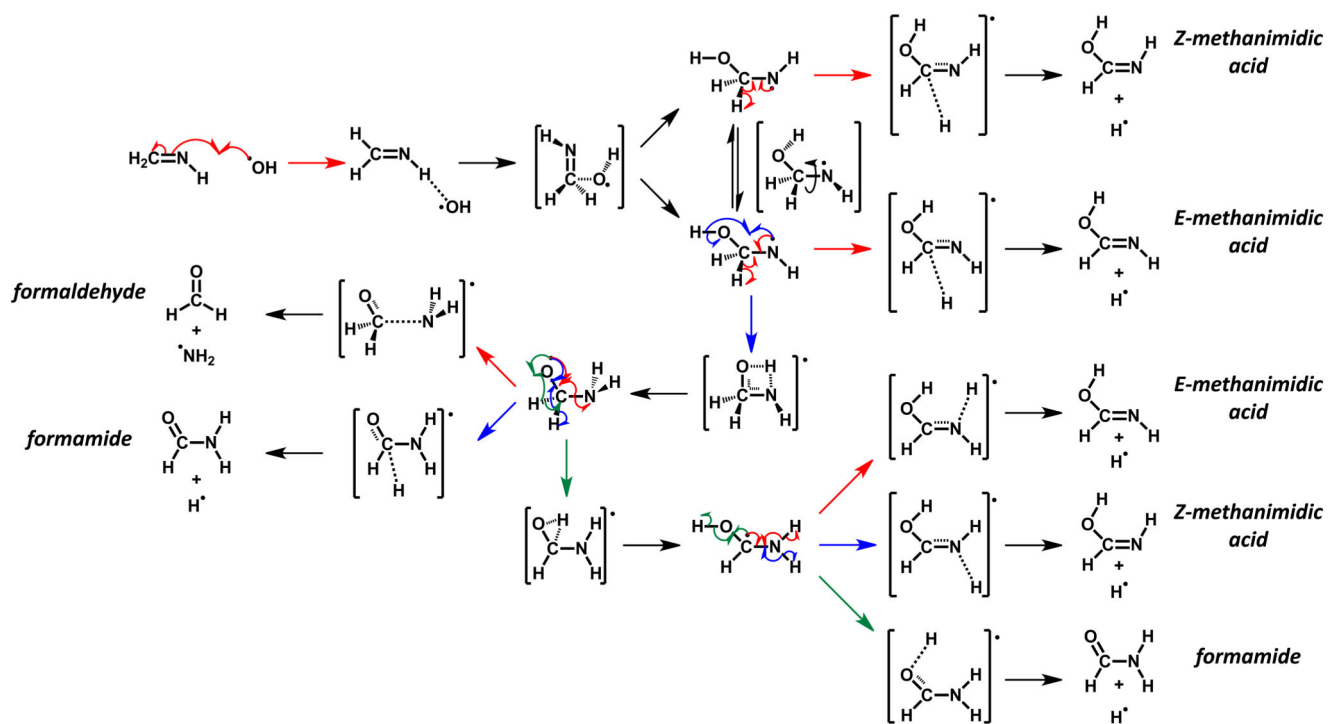
**Figure 5.**  
Proposed path for the approach of OH<sup>-</sup> and methanimine. Relative electronic CCSD(T)/CBS +CV energies in black. Energies in kJ/mol.



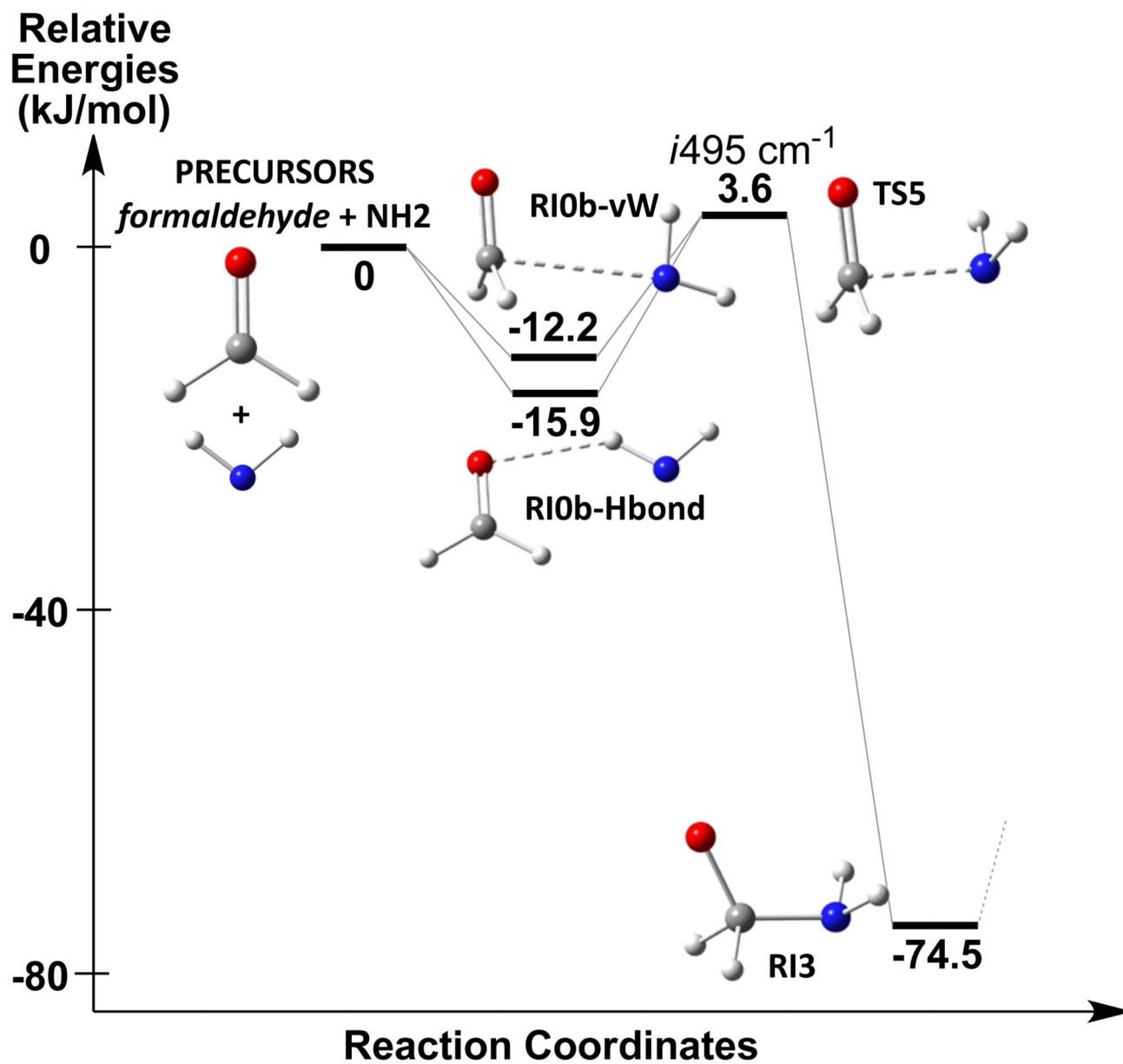
**Figure 6.** Proposed full reaction path. Electronic energies (black) are at the CCSD(T)/CBS+CV level and the ZPE corrected energies (green) are obtained by including the ZPE issuing from B2PLYP-D3/m-aug-cc-pVTZ anharmonic vibrational calculations. All energies are given in kJ/mol.



**Figure 7.**  
Possible radical mechanism for H abstraction.

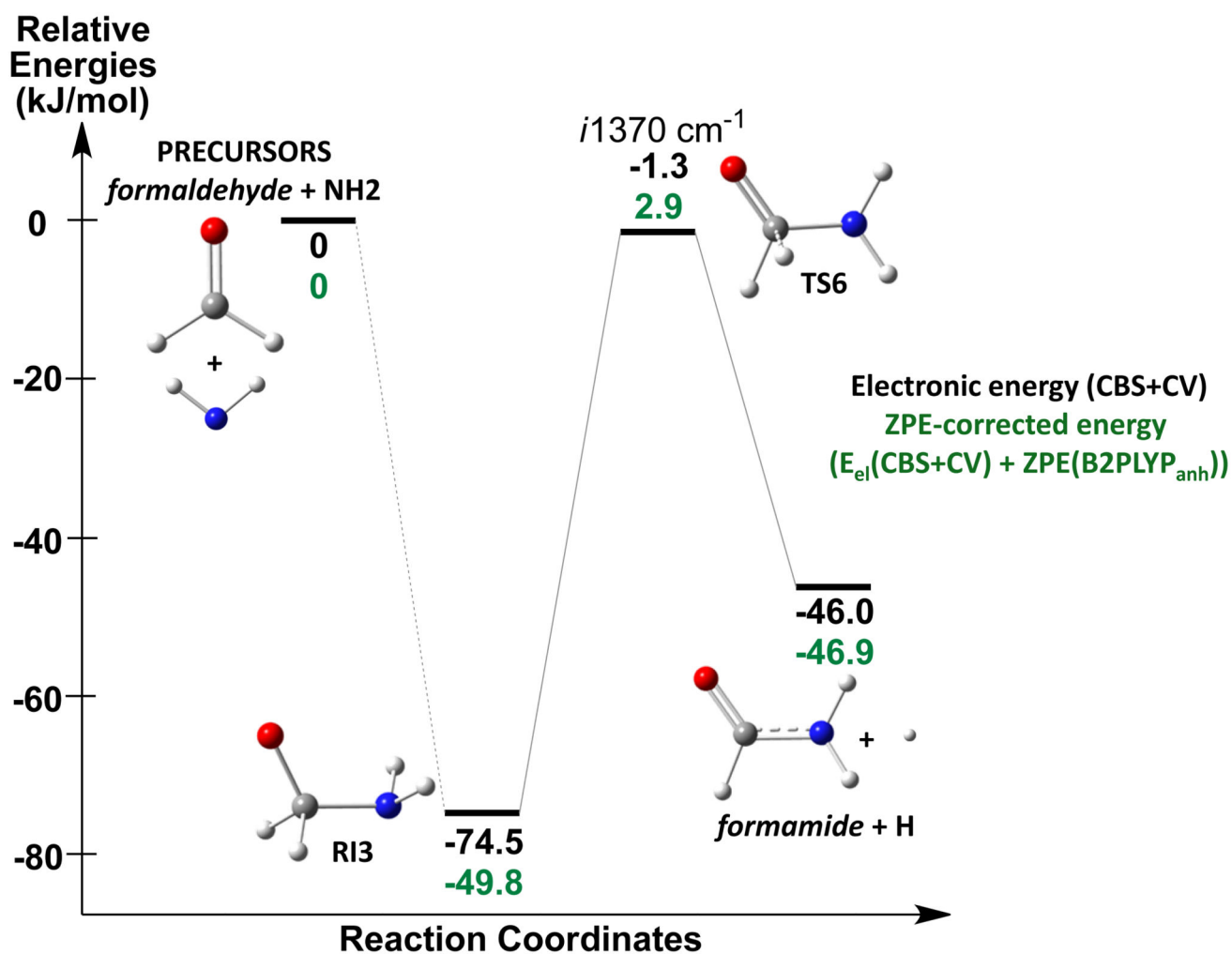


**Figure 8.**  
Possible radical mechanism for the addition of  $\text{OH}\cdot$  to methanimine.

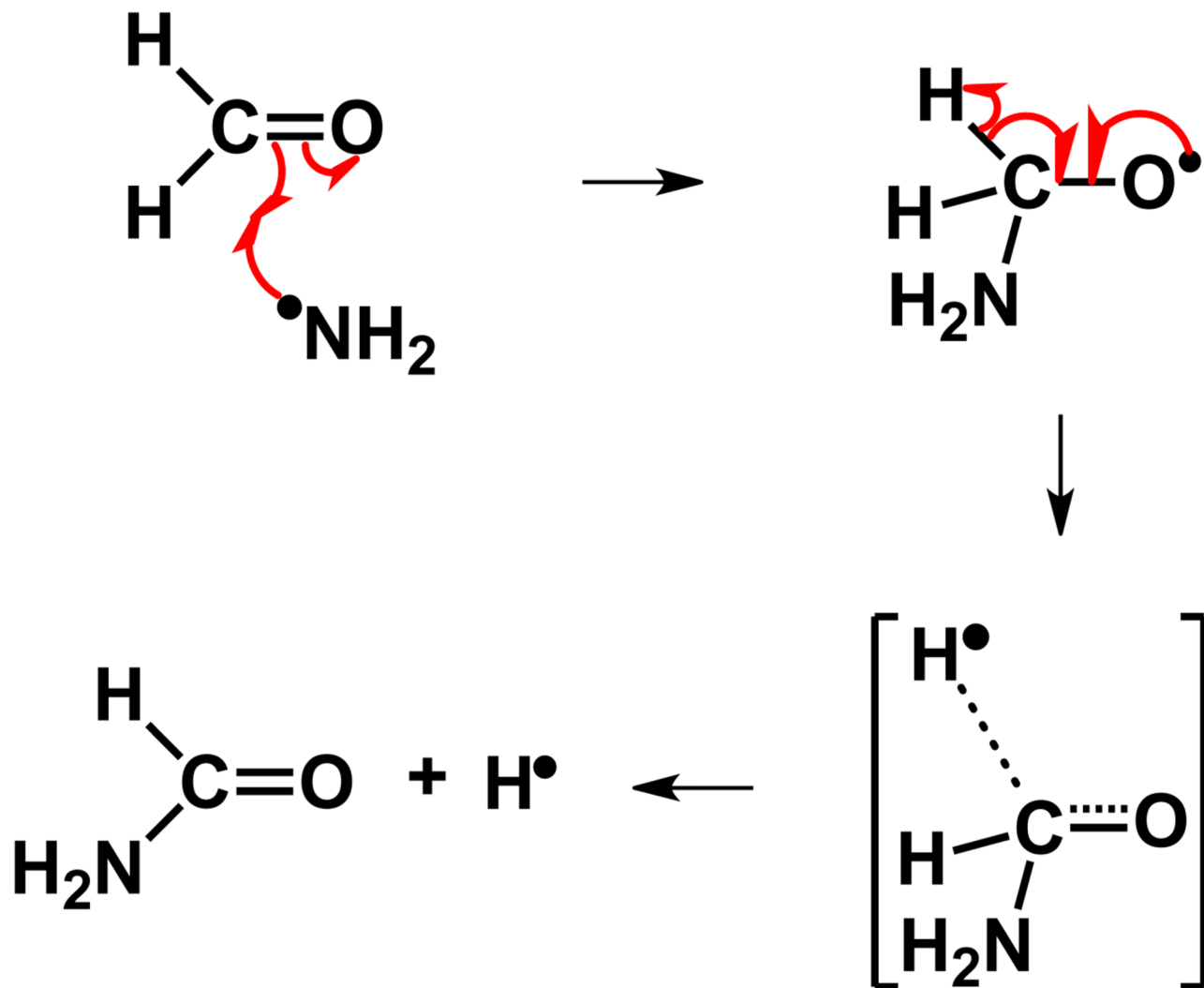


**Figure 9.**  
Proposed path for the approach of NH<sub>2</sub> and formaldehyde. Relative electronic CCSD(T)/CBS+CV (in kJ/mol) energies in black.

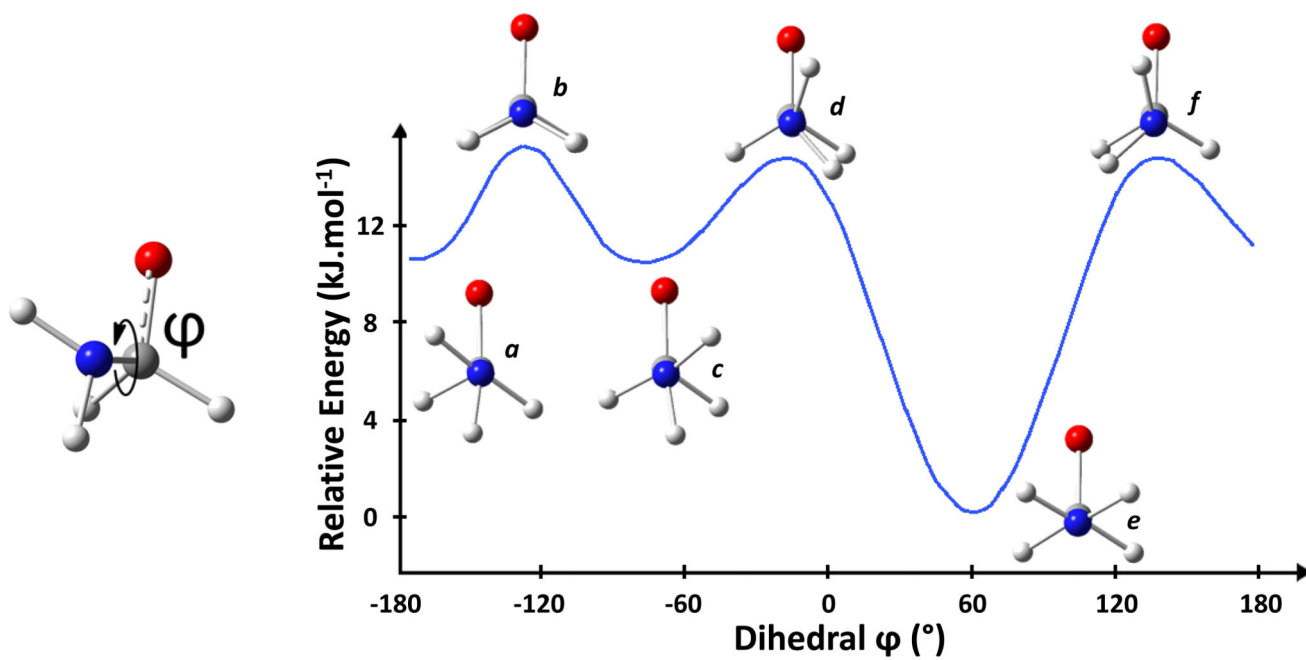




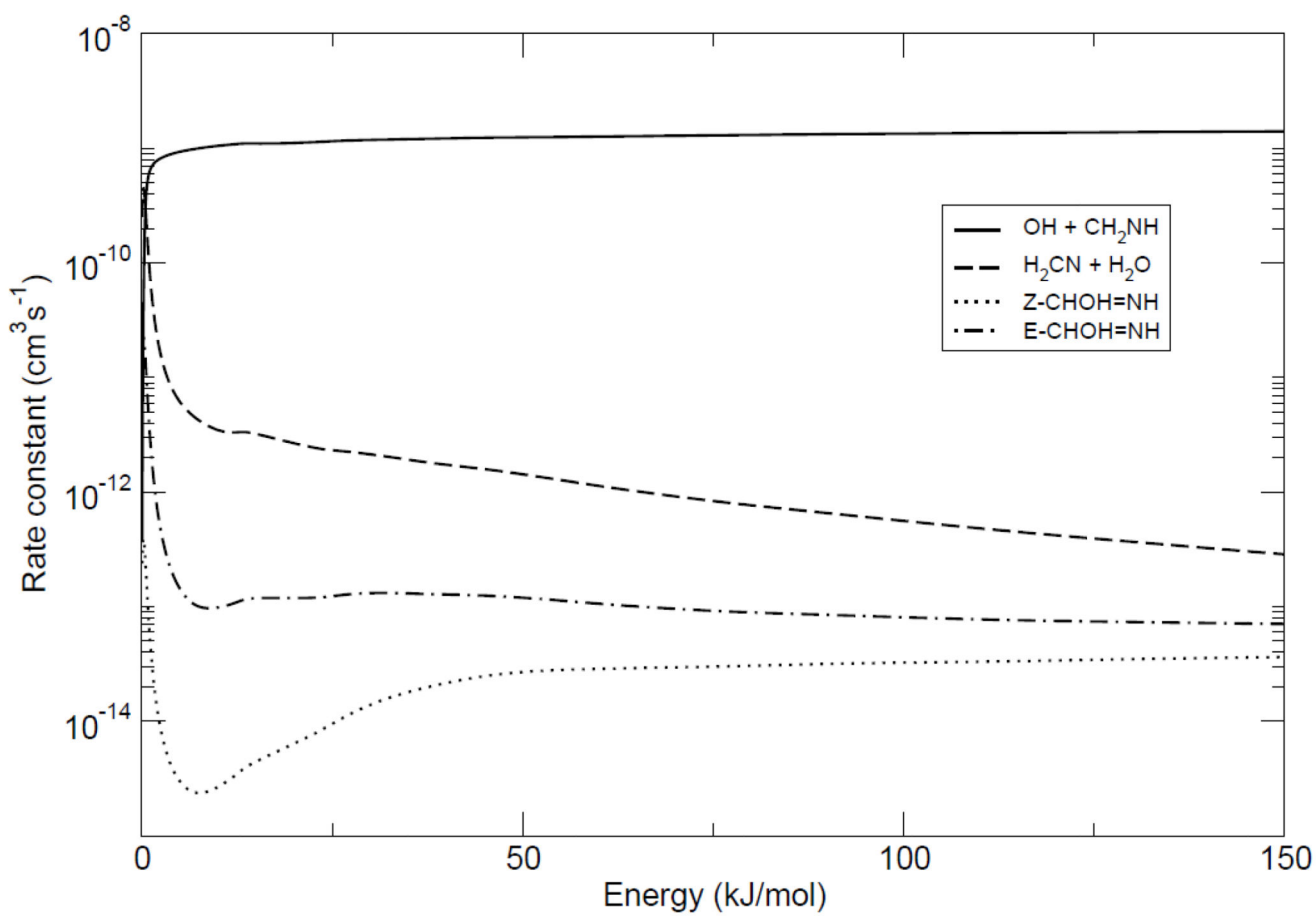
**Figure 10.** Proposed reaction path for formamide formation. Electronic energies (black) are at the CCSD(T)/CBS+CV level and the ZPE corrected energies (green) are obtained by including the ZPE issuing from B2PLYP-D3/m-aug-cc-pVTZ anharmonic vibrational calculations. All energies are given in kJ/mol.



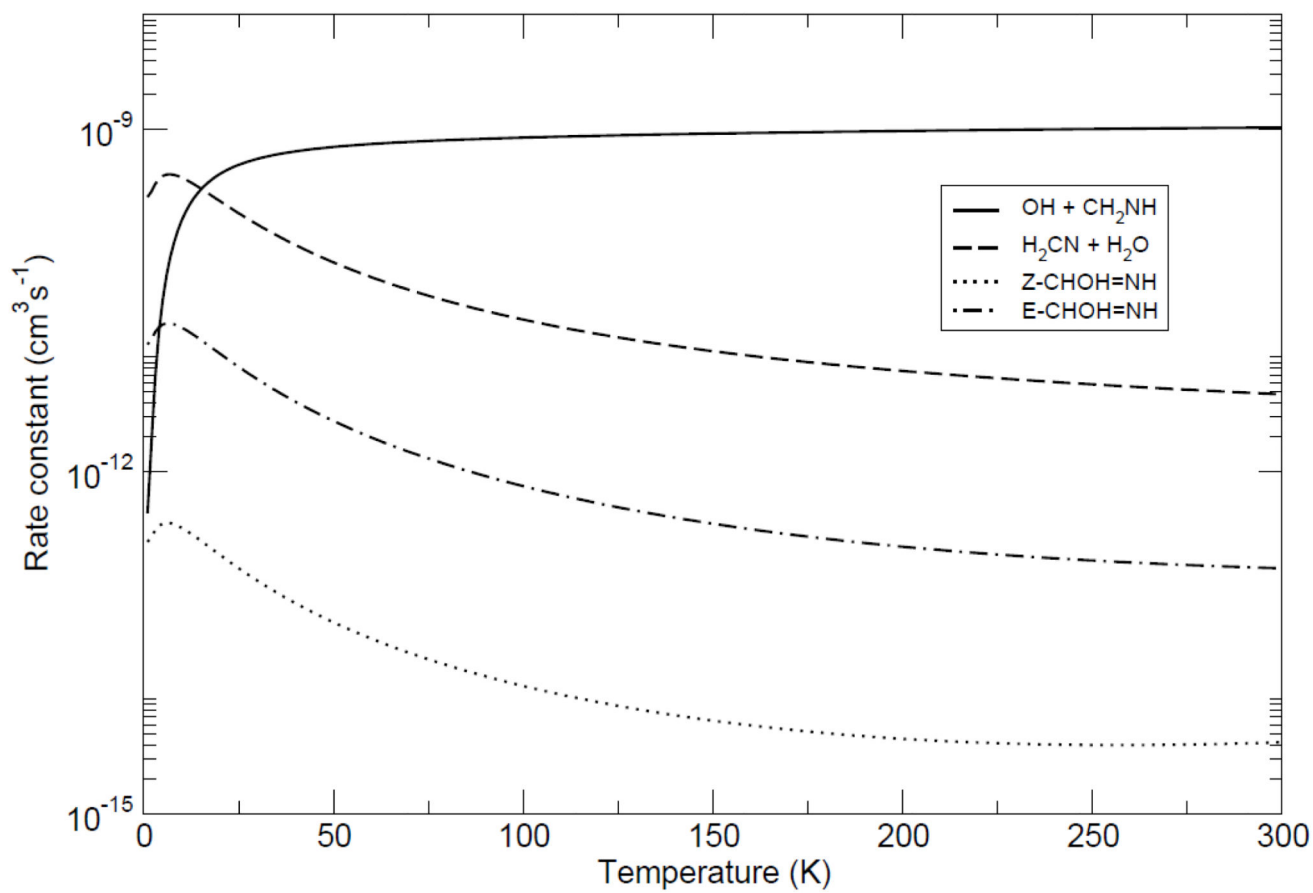
**Figure 11.**  
Possible radical mechanism for formamide formation through the NH<sub>2</sub> + CH<sub>2</sub>O reaction.



**Figure 12.** Potential energy profile along the  $\phi$  dihedral angle (NH<sub>2</sub> rotation) for compound RI3 obtained at the B2PLYP-D3/maug-cc-pVTZ level



**Figure 13.**  
Rate constant as a function of energy.



**Figure 14.**  
Rate constant as a function of temperature.

**Table 1**

Average root-mean-square errors of different geometrical parameters<sup>a</sup> with respect to the semi-experimental equilibrium values of the B3SE set of ref. 53.

	<b>B3LYP</b>	<b>B2PLYP</b>
C-O	0.006	0.003
C-N	0.005	0.002
C-H	0.007	0.001
N-H	0.007	0.001
Angles	0.41	0.26

<sup>a</sup>Distances in Å and angles in degrees

**Table 2**

Relative energies<sup>a</sup> (in kJ/mol, with respect to methanimine + OH) of **RI0**, **TS0<sub>a</sub>** and **TS0<sub>b</sub>** obtained at different levels of theory.

	B2PLYP-D3/m-aug-cc-pVTZ	CCSD(T)/m-aug-cc-pVTZ	CCSD(T)/CBS(TQ)+CV(CT) <sup>b</sup>	CCSD(T)/CBS(Q5)+CV(CQ) <sup>c</sup>
<b><i>Σ</i>precursors</b>	0	0	0	0
<b><i>RI0</i> (Cs)</b>	-32.09	-29.37	-30.38	-30.42
<b><i>TS0<sub>a</sub></i></b>	-7.53	-0.33	-4.52	-5.19
<b><i>TS0<sub>b</sub></i></b>	-8.28	-2.34	-4.14	-4.02
<b><i>RI1z</i></b>	-125.99	-	-128.03	-
<b><i>RI1e</i></b>	-128.87	-	-131.38	-

<sup>a</sup>Electronic energies computed at the corresponding B2PLYP/m-aug-cc-VTZ optimized geometries.

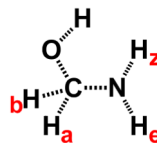
<sup>b</sup>The cc-pVTZ and cc-pVQZ basis sets were used for the extrapolation to the CBS limit and the cc-pCVTZ set for the corecorrelation correction.

<sup>c</sup>The cc-pVQZ and cc-pV5Z basis sets were used for the extrapolation to the CBS limit and the cc-pCVQZ set for the corecorrelation correction.

**Table 3**

Bond lengths (Å) and symmetry point groups of the investigated compounds (optimizations at the B2PLYP-D3/m-aug-cc-pVTZ level). The values related to the NH<sub>2</sub> + formaldehyde reaction are in bold.

	C-O	C-N	N-H <sub>c</sub>	N-H <sub>z</sub>	O-H	C-H <sub>a</sub>	C-H <sub>b</sub>	Sym.
<i>methanimine</i>	-	1.27	1.02	-	-	1.09	1.09	C <sub>S</sub>
<i>OH</i>	-	-	-	-	0.97	-	-	C <sub>∞v</sub>
<b>RI1z</b>	1.42	1.41	-	1.02	0.96	1.10	1.10	C <sub>S</sub>
<b>TS-RI1ez</b>	1.42	1.45	1.02	-	0.96	1.09	1.09	C <sub>I</sub>
<b>RI1e</b>	1.41	1.42	1.02	-	0.96	1.10	1.10	C <sub>S</sub>
<b>TS2</b>	1.42	1.47	1.02	1.27	1.24	1.09	1.09	C <sub>I</sub>
<b>RI3</b>	<b>1.37</b>	<b>1.44</b>	<b>1.01</b>	<b>1.01</b>	-	<b>1.10</b>	<b>1.10</b>	C <sub>S</sub>
<b>TS3z</b>	1.36	1.29	-	1.02	0.97	1.71	1.08	C <sub>I</sub>
<b>TS3e</b>	1.35	1.29	1.02	-	0.97	1.71	1.09	C <sub>I</sub>
<b>TS4</b>	1.38	1.38	1.01	1.01	1.22	1.29	1.09	C <sub>I</sub>
<b>TS5</b>	1.23	2.02	1.02	1.02	-	1.10	1.10	C <sub>S</sub>
<b>TS6</b>	<b>1.24</b>	<b>1.38</b>	<b>1.01</b>	<b>1.01</b>	-	<b>1.56</b>	<b>1.10</b>	C <sub>I</sub>
<b>RI5</b>	1.39	1.39	1.01	1.01	0.96	1.08	-	C <sub>I</sub>
<b>TS7</b>	1.25	1.35	1.00	1.01	1.38	1.09	-	C <sub>I</sub>
<b>TS8z</b>	1.36	1.27	1.64	1.02	0.96	1.08	-	C <sub>I</sub>
<b>TS8e</b>	1.34	1.28	1.01	1.62	0.97	1.08	-	C <sub>I</sub>
<i>H<sub>2</sub>CN</i>	-	1.24	-	-	1.09	1.09	-	C <sub>2v</sub>
<i>H<sub>2</sub>O</i>	-	-	-	-	0.96	-	-	C <sub>2v</sub>
<i>Z-methan. ac.</i>	1.36	1.26	-	1.02	0.97	1.08	-	C <sub>S</sub>
<i>E-methan. ac.</i>	1.35	1.26	1.01	-	0.97	1.09	-	C <sub>S</sub>
<i>formaldehyde</i>	<b>1.21</b>	-	-	-	-	<b>1.10</b>	<b>1.10</b>	C <sub>2v</sub>
<i>NH<sub>2</sub></i>	-	-	<b>1.03</b>	<b>1.03</b>	-	-	-	C <sub>2v</sub>
<i>formamide</i>	<b>1.21</b>	<b>1.36</b>	<b>1.00</b>	<b>1.00</b>	-	-	-	C <sub>S</sub>





**Table 4**

Relative energies<sup>a</sup> (in kJ/mol, with respect to formaldehyde+NH<sub>2</sub>) of **RI0b-vW**, **RI0b-Hbond**, **TS5** and **RI3** obtained at different levels of theory.

	B2PLYP-D3/m-aug-cc-pVTZ	CCSD(T)/m-aug-cc-pVTZ	CCSD(T)/CBS(TQ)+CV(CT) <sup>b</sup>	CCSD(T)/CBS(Q5)+CV(CQ) <sup>c</sup>
<b><i>Σ</i>precursors</b>	0	0	0	0
<b><i>RI0b-vW</i> (Cs)</b>	-12.72	-11.63	-12.18	-12.21
<b><i>RI0b-Hbond</i> (Cs)</b>	-15.69	-15.15	-15.90	-15.86
<b><i>TS5</i> (Cs)</b>	0.67	7.45	3.64	3.77
<b><i>RI3</i> (Cs)</b>	-71.55	-	-74.48	-

<sup>a</sup>Electronic energies computed at the corresponding B2PLYP/m-aug-cc-VTZ optimized geometries.

<sup>b</sup>The cc-pVTZ and cc-pVQZ basis sets were used for the extrapolation to the CBS limit and the cc-pCVTZ set for the core-correlation correction.

<sup>c</sup>The cc-pVQZ and cc-pV5Z basis sets were used for the extrapolation to the CBS limit and the cc-pCVQZ set for the core-correlation correction.