

# Computational Study of the Reactions of Interstellar Molecules: CH<sub>2</sub> Reacting with HCNO and HNCO

Hongchen Ji<sup>1</sup>, Anita Rágyanszki<sup>1,2\*</sup>, and René A. Fournier<sup>1</sup>

<sup>1</sup>Department of Chemistry, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada

<sup>2</sup>Zuse Institute Berlin, Takustraße 7, 14195 Berlin, Germany

\*Corresponding author, e-mail address: ragyanszki@zib.de

---

**ABSTRACT:** Association reactions among small molecules known to exist in the interstellar medium are interesting for theories on the origins of life. A screening of thousands of reactions, using machine learning estimates of energy barriers, identified the reaction of CH<sub>2</sub> with HCNO and HNCO as particularly interesting. We report reaction mechanisms, including energies of transition states and products, computed with density functional theory and coupled cluster theory. The lowest energy pathway on the triplet ground state surface of CH<sub>2</sub> + HNCO has a barrier of 11 kcal/mol and produces CH<sub>2</sub>(CO)NH. Singlet CH<sub>2</sub> is 9 kcal/mol above the ground state. It can react with HCNO or HNCO without barrier giving four products: CH<sub>2</sub>NCHO, *N*-methyleneformamide, the thermodynamically favoured product; NHCHCHO; NHCHOCH; and (CH<sub>2</sub>OC)NH, oxiran-2-ylazanide. If triplet to singlet crossing occurs, an upper bound of roughly 10 kcal/mol is implied for the barrier to formation of these four products.

---

**Keywords:** *Interstellar medium • Reaction mechanism • Prebiotic chemistry • Methylene • Formamide • Imine acetaldehyde*

## 1. Introduction

The discovery of molecules in space and the study of their reactivity is of interest for their implications on the possible origins of life. The increasing number of organic molecules discovered through astronomical observations [1] suggests that prebiotic molecules may have formed in space and reached early Earth and other planets [2]. The spectroscopic signatures of more than 250 molecular species have been detected in the interstellar medium (ISM) and in circumstellar envelopes (CSE) [1,3]. It is difficult to establish the evidence for a molecular species in the ISM. It is therefore plausible that many more molecules present in the ISM are yet to be discovered. Recent observation of spectral lines assigned to the amino acid tryptophan in the interstellar gas of star-forming regions [4] points to the possibility of relatively complex and biologically important molecules.

Most molecules that have been observed in space have ten or fewer atoms. The mechanism for the formation of bigger molecules is not well known. Association reactions may take place in the gas phase or at the surface of carbonaceous grains [5]. These grains could be covered with ice and may catalyze some reactions. Given the low temperatures in the ISM (typically between 10 and 150 K) [6], it is likely that only reactions with a very low (or zero) energy barrier occur with a significant rate. Computations can guide the discovery of molecules and reactions in space. With quantum chemical methods, one can get transition states and compute energy barriers with an accuracy on the order of 1 kcal/mol [7]. Bimolecular reactions between the molecular species identified in the ISM are natural targets for this kind of study. But there are thousands of possibilities and accurate calculations for all of them is currently impossible.

The strategy adopted in the present work starts by considering all reactions with reactant molecules A and B chosen among those of Table 1. This makes 2,628 combinations. The molecules in Table 1 have all been observed in the ISM or CSE and contain no atom other than H, C, N and O. We consider association reactions A + B → P in which the product P is in a molecular database of 10,306 molecules taken from the reference [8]. This gives a set

of 136,081 chemical reactions. Then we use criteria to eliminate the vast majority of reactions: (1) the estimated energy barrier must be zero or very small; (2) the product molecule P should contain at least one atom each of H, C, N and O; (3) the proportion of the four elements (H:C:N:O) and functional groups in P should resemble those commonly found in molecules that form the basis of life. Finally, a very small subset of the reactions that are best according to these criteria are studied by methods of quantum chemistry.

Panda, Chiranjibi, Awasthi, and Anoop [9] studied the reaction mechanisms of HCN + H<sub>2</sub>CO and HNC + H<sub>2</sub>CO by density functional theory (DFT) methods and coupled cluster (CC) method and found barriers of reactions of 32 to 66 kcal/mol giving oxiran-2-ylazanide and (Z)-imine acetaldehyde as the main products. Several experimental and theoretical studies of reactions of radicals with HCNO have been conducted [10a–10i]. Theoretical studies of the reaction between NH and HCNO have been performed by a combination of DFT and quadratic configuration interaction (QCI) methods giving a reaction barrier of 2.9 kcal/mol for the kinetically favoured products, HCN and HON [10a]. Another study reported DFT, CC and QCI calculations predicting a zero barrier to the most feasible products, HCN and HNO [10b]. An experimental kinetics study of OH + HCNO reaction was conducted by Feng, Meyer, and Herschberger [10e] producing two major kinds of products CO + H<sub>2</sub>NO and HNO + HCO. A computational study of the same reaction by DFT and CC methods gave CO + NHOH as mainly favoured products with a zero barrier [10d]. The product channels of CH<sub>3</sub> + HCNO have been theoretically studied by DFT and CC methods: the favoured route produced Z and E-HC(CH<sub>3</sub>)NO with a reaction barrier of only 4 kcal/mol [10f]. There have been several experimental and theoretical studies of reactions of radicals and HNCO [10j–10n], including C<sub>3</sub>H<sub>3</sub> + HNCO [10j], OH + HNCO [10k], and CH<sub>2</sub>CH + HNCO [10m]. A theoretical study on the hydrogen transfer path of CH<sub>2</sub> + HNCO reaction has also been reported [10n].

**Table 1.** Molecule list of the ISM and CSE present in our molecule database<sup>a</sup>

Molecule Formula	SMILES	References	Molecule Formula	SMILES	References
CO	[C-]#[O+]	11, 12	CH <sub>3</sub> NCO	O=C=NC	47
H <sub>2</sub>	[HH]	13	HOCH <sub>2</sub> CN	OCC#N	48
N <sub>2</sub>	N#N	14	H <sub>3</sub> CNH <sub>2</sub>	CN	49
H <sub>2</sub> C	[CH2]	15	CH <sub>2</sub> CHCN	N#CC=C	23, 39
HCN	C#N	11, 16, 17	HCCCHNH	C#CC=N	50
CO <sub>2</sub>	C(=O)=O	18	CH <sub>3</sub> CHO	O=CC	11, 51
H <sub>2</sub> O	O	19	H <sub>2</sub> CHCOH	OC=C	52
N <sub>2</sub> O	N#[N+][O-]	20	c-C <sub>2</sub> H <sub>4</sub> O	O1CC1	51
H <sub>2</sub> CO	C=O	21	CH <sub>3</sub> C <sub>2</sub> H	CC#C	16
C <sub>2</sub> H <sub>2</sub>	C#C	22	NH <sub>2</sub> CH <sub>2</sub> CN	NCC#N	53
HNCO	N=C=O	23	(NH <sub>2</sub> ) <sub>2</sub> CO	C(=O)(N)N	54
HCNO	[O-][N+]#C	24	CH <sub>2</sub> CHCHO	O=CC=C	45
HCCN	N#C[CH+]	25	HCOOCH <sub>3</sub>	O=COC	23, 37
CNCN	[C-]#[N+]C#N	26	CH <sub>3</sub> COOH	CC(O)=O	55
HOCN	OC#N	27	CH <sub>3</sub> CHNH	CC=N	56
NH <sub>3</sub>	N	11, 28	HC <sub>3</sub> H <sub>2</sub> CN	C#CCC#N	57
CH <sub>4</sub>	C	29	H <sub>3</sub> CC <sub>2</sub> CN	CC#CC#N	58
H <sub>2</sub> CNH	C=N	30	CH <sub>2</sub> CCHCN	C=C=CC#N	45, 58
H <sub>2</sub> C <sub>2</sub> O	C=C=O	23	C <sub>2</sub> H <sub>3</sub> NH <sub>2</sub>	C=CN	59
HNCNH	C(=N)=N	31	CH <sub>3</sub> NHCHO	CNC=O	33
NH <sub>2</sub> CN	N#CN	32, 33	CH <sub>3</sub> CONH <sub>2</sub>	O=C(N)C	33, 42, 45
HCC-NC	C#C[N+]#[C-]	34	CH <sub>3</sub> OCH <sub>3</sub>	COC	60
HC(O)CN	C(=O)C#N	35	CH <sub>3</sub> CH <sub>2</sub> OH	OCC	61
HCOOH	C(=O)O	36, 37	CH <sub>3</sub> CH <sub>2</sub> CN	CCC#N	11, 23, 39
HC <sub>3</sub> N	C#CC#N	11, 16, 38	CH <sub>3</sub> CHCH <sub>2</sub>	C=CC	62
CH <sub>3</sub> OH	CO	23	CH <sub>3</sub> C <sub>4</sub> H	CC#CC#C	63
CH <sub>3</sub> CN	CC#N	23, 39	CH <sub>3</sub> CHCH <sub>2</sub> O	CC1CO1	64
C <sub>2</sub> H <sub>4</sub>	C=C	40	CH <sub>3</sub> CH <sub>2</sub> CHO	CCC=O	45
HC <sub>2</sub> CHO	C#CC=O	41	(CH <sub>3</sub> ) <sub>2</sub> CO	CC(=O)C	65
CH <sub>3</sub> NC	[C-]#[N+]C	39	NH <sub>2</sub> CH <sub>2</sub> CH <sub>2</sub> OH	NCCO	66
HC <sub>4</sub> N	[H]C#C[C]C#N	11	CH <sub>3</sub> COOCH <sub>3</sub>	O=C(OC)C	67
HCONH <sub>2</sub>	O=CN	42	C <sub>6</sub> H <sub>6</sub>	c1ccccc1	43

<i>l</i> -H <sub>2</sub> C <sub>4</sub>	C#CC#C	11, 43	(CH <sub>3</sub> ) <sub>2</sub> CHCN	CC(C)C#N	68
E-HNCHCN	[H]/N=C(C#N)\[H]	44	C <sub>3</sub> H <sub>7</sub> CN	CCCC#N	69
CH <sub>2</sub> CNH	C=C=N	45	C <sub>2</sub> H <sub>5</sub> NH <sub>2</sub>	CCN	59
<i>c</i> -H <sub>2</sub> C <sub>3</sub> O	C1=C(=O)=C1	46	C <sub>2</sub> H <sub>5</sub> OCH <sub>3</sub>	COCC	70

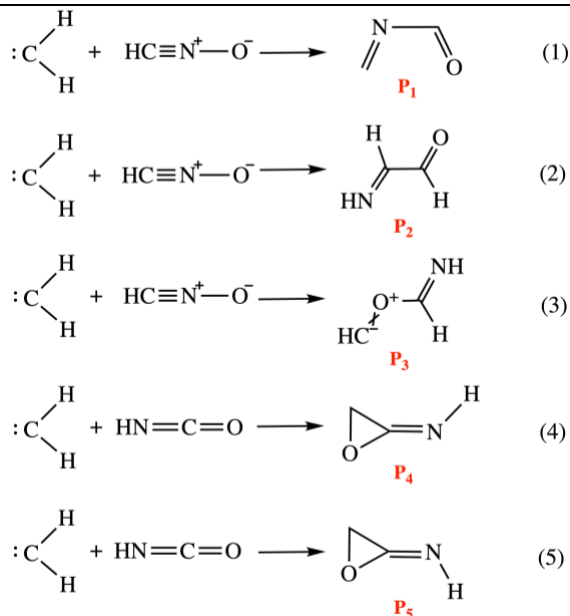
<sup>a</sup>The list of 72 molecules believed to be present in the ISM or CSE [1,11–70] that are also in our database [71]. The SMILES representations [72] and references are shown for each molecule.

Redondo *et al.* recently published high-level theoretical spectroscopic constants for hydroxy-azirine [73] and imine acetaldehyde [74]. These molecules were proposed as potential interstellar prebiotic molecules, but searches for their spectroscopic signature in the ISM have failed so far. Fourré *et al.* [75] suggested using the minimum energy principle (MEP) for screening potential ISM molecules. Interestingly, the screening of reactions with estimated energy barriers (present work) produced some of the same molecules proposed in reference [75].

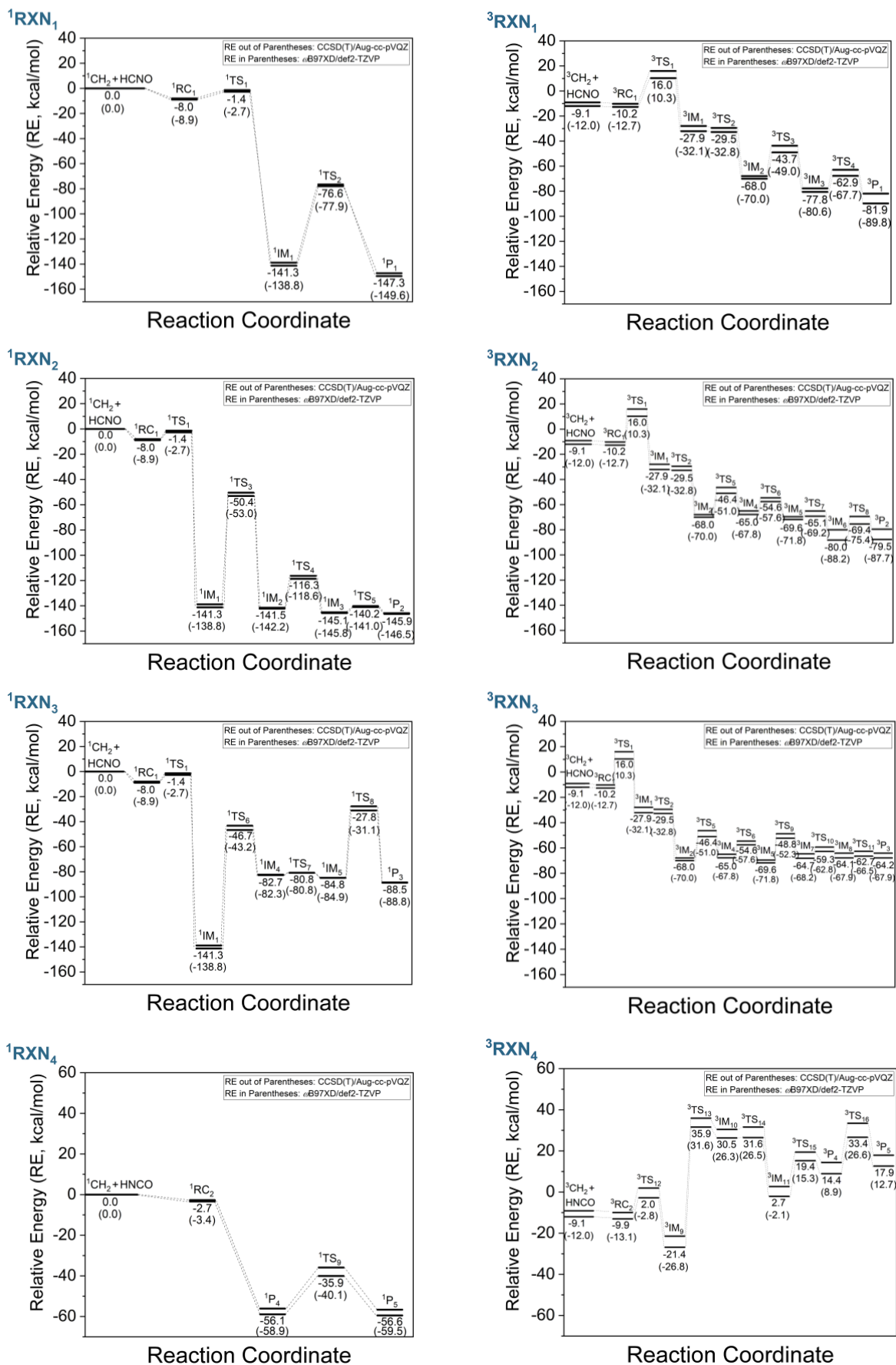
The dataset of molecules and reactions and the screening procedure are described in detail in section 2. The main result of the screening is the identification of CH<sub>2</sub> + HCNO/HNCO as a particularly promising system for study by accurate electronic structure calculation. The quantum chemical methods used to study them are described in section 3. Section 4 presents the main results of the study: a set of reaction mechanisms including transition states with energy barriers calculated with an accurate CCSD(T) method [76].

## 2. Chemical Reactions Dataset and Screening Procedure

One of the possible mechanisms for reactions in the ISM is gas-phase radiative association [5]:  $A + B \rightleftharpoons AB^* \rightarrow AB + h\nu$ . The temperature in the ISM varies roughly between 10 and 100 K and may exceed 100 K in “protostellar cores”. Therefore, reactions with low or zero energy barriers are much more likely to occur than those with high barriers. The formation of a relatively stable and long-lived reaction complex  $AB^*$  is also favourable [5]. Another important mechanism for ISM chemistry is the reaction of A and B at the surface of dust grains. The nature of these grains, and the mechanisms of reactions on them, are not well known and are difficult to model theoretically. For simplicity we assume that reactions with the smallest gas-phase energy barriers are also among those most likely to occur at the surface of grains.



**Scheme 1.** The five reactions of CH<sub>2</sub> with HCNO and HNCO and their products.



**Fig. 1.** Singlet and triplet potential energy diagrams showing the energies of reactant complexes (RCs), intermediates of reaction (IMs), transition states (TSs) and products (Ps). The relative energies calculated by CCSD(T) and DFT are shown, with DFT in parentheses.

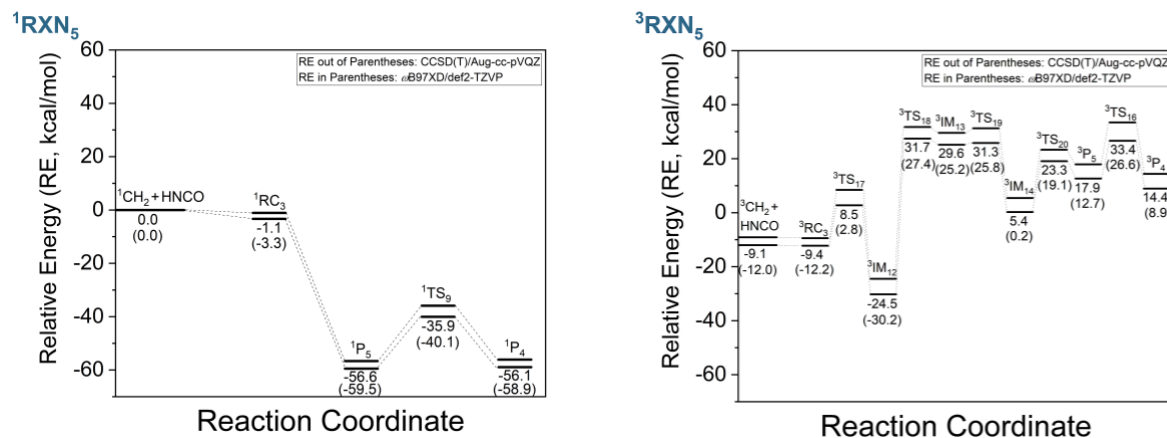


Fig. 1. (Continued).

A Kernel Ridge Regression (KRR) model described elsewhere [71] was used to estimate gas-phase reactions energy barriers. The procedure takes Simplified Molecular Input Line Entry Systems (SMILES) representations of molecules [72], converts them to 3D structures with Open Babel [77], and uses nuclear coordinates of reactants and products to calculate 300 features of a reaction. The features, which depend on interatomic distances and simple atomic properties like electronegativity and hardness [71], are fed to a KRR procedure with a Laplacian kernel. The KRR model was trained and tested on the dataset of Grambow *et al.* [8] which has 11,730 elementary reactions [71] and energy barriers computed at the  $\omega$ B97XD level of theory [78], with molecular species made of the elements H, C, N and O and no more than seven non-hydrogen atoms [71]. There is a total of 10,306 molecules in this dataset [8,71] of which 72, those in Table 1, were observed in the ISM or CSE [1,11–70]. We limit the present study to association reactions  $A + B \rightarrow P$  where A and B are from Table 1 because those are plausible ISM reactions and, for convenience, to cases where P is among the 10,306 molecules of the dataset [8,71]. The KRR model is computationally efficient and requires only twenty minutes for generating barrier estimates for 136,081 reactions with molecules A, B, and P that meet the above criteria. Although the KRR model was trained on elementary reactions it is used here to predict the barrier of the rate limiting step of reactions that may be elementary or complex. Despite this limitation, the KRR estimates provide a useful ranking of reactions in increasing order of estimated barriers. In particular, 2,214 reactions were predicted to have zero barriers or small barriers less than 6 kcal/mol (a root-mean square error of KRR [71]).

We apply a second round of screening to the 2,214 reactions with small energy barriers to bias for products (Ps) that look promising building blocks for larger molecules and prebiotic chemistry. For this we use two criteria: (i) there should be at least one atom of each element (H, C, N, O); (ii) the degree of unsaturation of the product divided by the number of non-H atoms  $n_u/n_x$  should fall within a range of [0.05, 0.50] so that it is similar to what is found in amino acids [79], mono and disaccharides [80], alpha-hydroxy acids [81], and many drug-like molecules [82]. Only a few reactions passed both rounds of screening.

Finally, five reactions stand out because they all involve the reaction of methylene ( $\text{CH}_2$ ) with fulminic acid (HCNO) or its isomer isocyanic acid (HNCO). These five reactions are shown in Scheme 1. The products of reactions exhibit an interesting variety of functional groups for such a small set.

### 3. Computational Details

All calculations were done with the Gaussian16 software [83] using GaussView [84] and Avogadro [85] as interfaces. The geometries of reactants, products, and intermediates of reactions were optimized by DFT with the  $\omega$ B97XD method for exchange-correlation and dispersion corrections [78] and a def2-TZVP basis set [86]. The search for transition states (TSs) was also done at the  $\omega$ B97XD/def2-TZVP level of theory [78,86]. All stationary points of the potential energy surface were characterized by normal mode analysis, those with one imaginary frequency being TS. The intrinsic reaction coordinate (IRC) [87] originating from every TS was followed to find the corresponding reactant and product. The energy profiles in Fig. 1 show sequences of minima (reactant, intermediate, or product) and the TS connecting each pair of minima. Single point (fixed geometry) calculations were performed with CCSD(T) (coupled cluster theory with single and double excitations and perturbative treatment of triple excitations) with a Aug-cc-pVQZ basis set [76,88] at the optimized minima and TS structures to get more accurate energy barriers and energies of reactions. Unless stated otherwise, the energies reported here are differences between 0 K energies at the CCSD(T)/Aug-cc-pVQZ [76,88] level with zero-point energy corrections and optimized geometries by  $\omega$ B97XD/def2-TZVP, a method denoted CCSD(T)/Aug-cc-pVQZ// $\omega$ B97XD/def2-

TZVP. Each of the five reactions of Scheme 1 was examined on the triplet and singlet potential energy surface. The ground states of HCNO and HNCO are both singlet. The ground state of CH<sub>2</sub> is triplet, with the lowest excited singlet state being 9.00 kcal/mol higher [89]. The CH<sub>2</sub> singlet excited state energy is calculated as +9.1 kcal/mol, very close to the experiment. In the molecular species symbols of the next section RC stands for reactant complex, IM for reaction intermediate, TS for transition state, P for product, the superscripts 1 and 3 represent singlet and triplet species, and subscripts are arbitrarily assigned to the structures (depicted in Fig. 2) in the order in which it appears in Fig. 1. In a few cases, following the IRC from different transition states produced slightly different reaction complex geometries with nearly identical energies. Only one version of each reaction complex is shown in Fig. 2.

#### 4. Results and Discussion

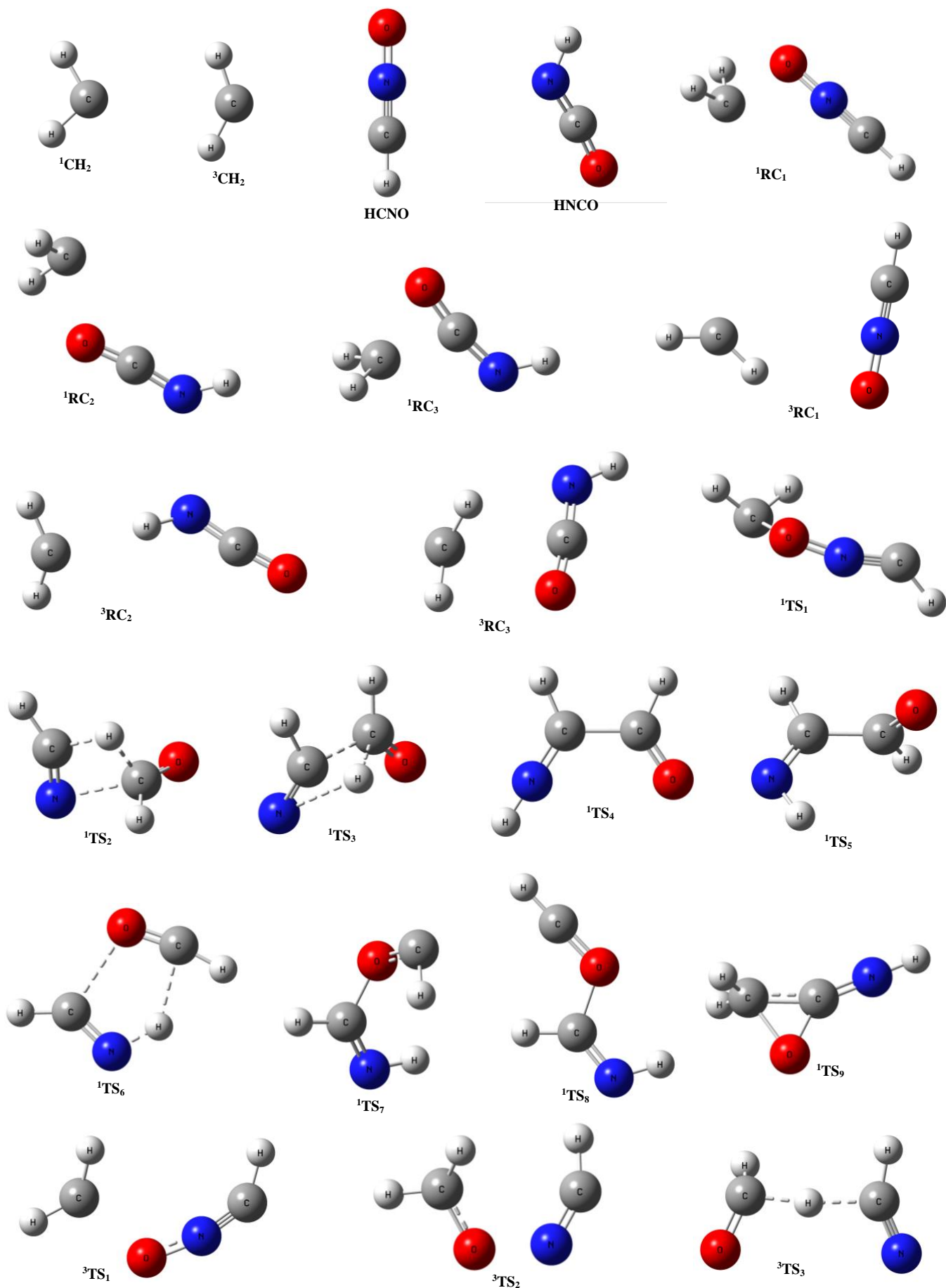
The products of reactions in Scheme 1 are *N*-methyleneformamide, imine acetaldehyde, ((λ<sup>3</sup>-methylene)-λ<sup>3</sup>-oxidanyl)methanimine, and two conformers of oxiran-2-ylazanide. None of them has been observed in the ISM or CSE [90,75]. Total energies including zero-point corrections (*E*s) are shown in Table 2 for all minima and transition states along with relative energies where the zero of energy is assigned to singlet CH<sub>2</sub> plus singlet HCNO or HNCO. The optimized geometries of reactants, reactant complexes, intermediates, transition states, and products are shown in Fig. 2 and the cartesian coordinates can be found in the Supplementary Information (SI). The basis set used here is not expected to give accurate results in weakly bound complexes. In particular, the geometries of <sup>1</sup>RC<sub>2</sub>, <sup>1</sup>RC<sub>3</sub>, and <sup>3</sup>RC<sub>3</sub> could not be fully optimized. Results on reactant complexes (RCs) should be considered only in a qualitative sense. Other relative energies also have numerical errors of ~2 kcal/mol, for instance, <sup>3</sup>IM<sub>1</sub> is found 1.6 kcal/mol higher than <sup>3</sup>TS<sub>2</sub> even though IRC following connects them. For instance, the energy of <sup>3</sup>IM<sub>1</sub>, without zero-point energy, is 0.22 kcal/mol lower than the <sup>3</sup>TS<sub>2</sub>. However, the zero-point correction pushes <sup>3</sup>IM<sub>1</sub> 1.6 kcal/mol higher than <sup>3</sup>TS<sub>2</sub>.

Reactions *k* (*k* = 1, 2, ..., 5) on the singlet and triplet energy surfaces are represented by <sup>1</sup>RXN<sub>*k*</sub> and <sup>3</sup>RXN<sub>*k*</sub>. Their energy profiles are shown in Fig. 1. In reaction <sup>1</sup>RXN<sub>1</sub>, singlet CH<sub>2</sub> and HCNO form a stable reactant complex, then CH<sub>2</sub> picks up the oxygen atom of HCNO in a barrierless process (see <sup>1</sup>TS<sub>1</sub> in Fig. 2 and <sup>1</sup>RXN<sub>1</sub> in Fig. 1). This leads to formation of <sup>1</sup>IM<sub>1</sub>, a complex of hydrogen cyanide (HCN) and formaldehyde (H<sub>2</sub>CO) both of which were identified in the ISM or CSE [11,16,17,21]. The complex of HCN and H<sub>2</sub>CO can go on to react in many ways, see <sup>1</sup>IM<sub>1</sub> and the transition states <sup>1</sup>TS<sub>2</sub>, <sup>1</sup>TS<sub>3</sub>, and <sup>1</sup>TS<sub>6</sub> in <sup>1</sup>RXN<sub>1</sub>, <sup>1</sup>RXN<sub>2</sub>, and <sup>1</sup>RXN<sub>3</sub>, respectively. These transition states are well below the <sup>1</sup>(CH<sub>2</sub> + HCNO) separation limit but well above HCN + H<sub>2</sub>CO (by 65 kcal/mol, 91 kcal/mol and 95 kcal/mol). The transition state <sup>1</sup>TS<sub>1</sub> is above triplet CH<sub>2</sub> plus HCNO by only 6.3 kcal/mol (9.0 - 2.7 = 6.3).

One conclusion we draw from energy profiles <sup>1</sup>RXN<sub>1</sub>, <sup>1</sup>RXN<sub>2</sub>, and <sup>1</sup>RXN<sub>3</sub> is that singlet CH<sub>2</sub>, it is present in the ISM, can react without barrier with HCNO. The thermodynamically favoured product of this reaction would be CH<sub>2</sub>NCHO (*N*-methyleneformamide, <sup>1</sup>P<sub>1</sub>), at -147.3 kcal/mol relative to reactants. Computational studies of *N*-methyleneformamide have reported vibrational spectra [91] and relative energies of conformers [92]. The conformer of *N*-methyleneformamide that we obtained is the *gauche*-conformation, which is reported as the most stable conformation [92]. The conformer <sup>1</sup>P<sub>2</sub> has a slightly higher energy, -145.9 kcal/mol. Some of the intermediates of reactions are also very close in energy: the HCN-H<sub>2</sub>CO complex <sup>1</sup>IM<sub>1</sub> (-141.3), and two conformers of imine acetaldehyde NHC(H)CHO, <sup>1</sup>IM<sub>2</sub> (-141.5) and <sup>1</sup>IM<sub>3</sub> (-145.1). However, it is much more likely that interstellar CH<sub>2</sub> is in its triplet ground state. Then, the formation of CH<sub>2</sub>NCHO would require approximately 6 to 9 kcal/mol excess energy and triplet-to-singlet surface crossing. Crossing might occur at some configurations of the CH<sub>2</sub> + HCNO reactant complex at an energetic cost of 9 kcal/mol (at large CH<sub>2</sub>-HCNO separation) or maybe less, at some hypothetical favourable configuration, but we were unable to locate such a configuration.

Recently Panda *et al.* [9] investigated in detail many reactions on the singlet energy surface of C<sub>2</sub>H<sub>3</sub>NO by a DFT method. They found many HCN + H<sub>2</sub>CO reaction pathways, but none of them with a barrier smaller than 43 kcal/mol. For the reaction of H<sub>2</sub>CO with HNC (another molecule observed in the ISM [90,93]) the lowest reaction energy barrier found was 15 kcal/mol [9], leading to the formation of <sup>1</sup>P<sub>5</sub> (shown in Fig. 2).

On the singlet surface, reactions 4 and 5 are barrierless and lead to the formation of two conformers of oxiran-2-ylazanide (<sup>1</sup>P<sub>4</sub> and <sup>1</sup>P<sub>5</sub>). On the triplet surface, reactions 4 and 5 have rate limiting barriers of 11.1 and 17.6 kcal/mol, respectively. The thermodynamically favoured triplet state product is <sup>3</sup>P<sub>1</sub> (see <sup>3</sup>RXN<sub>1</sub> in Fig. 1) with an energy of -82 kcal/mol (-73 kcal/mol relative to the ground state). It is 65 kcal/mol higher than singlet *N*-methyleneformamide. The kinetically favoured triplet product is <sup>3</sup>IM<sub>9</sub> (see Fig. 1, panel <sup>3</sup>RXN<sub>4</sub>). The barrier to forming <sup>3</sup>IM<sub>9</sub>, is 11.1 kcal/mol, making this reaction practically impossible in the gas phase in the ISM, but maybe possible in hot corinos regions or at the surface of dust grains. Molecules <sup>3</sup>P<sub>4</sub> and <sup>3</sup>P<sub>5</sub> are 27 and 23 kcal/mol higher than the reactants and are not likely products of these two triplet reactions.



**Fig. 2.** Optimized geometries of singlet and triplet reactants, RCs (with not fully optimized  $^1\text{RC}_2$ ,  $^1\text{RC}_3$ , and  $^3\text{RC}_3$ , see main text for details), TSs, IMs, and products of reactions (1), (2), (3), (4), and (5) at  $\omega\text{B97XD/def2-TZVP}$  level of theory. Cartesian coordinates of all studied species shown here are listed in SI.

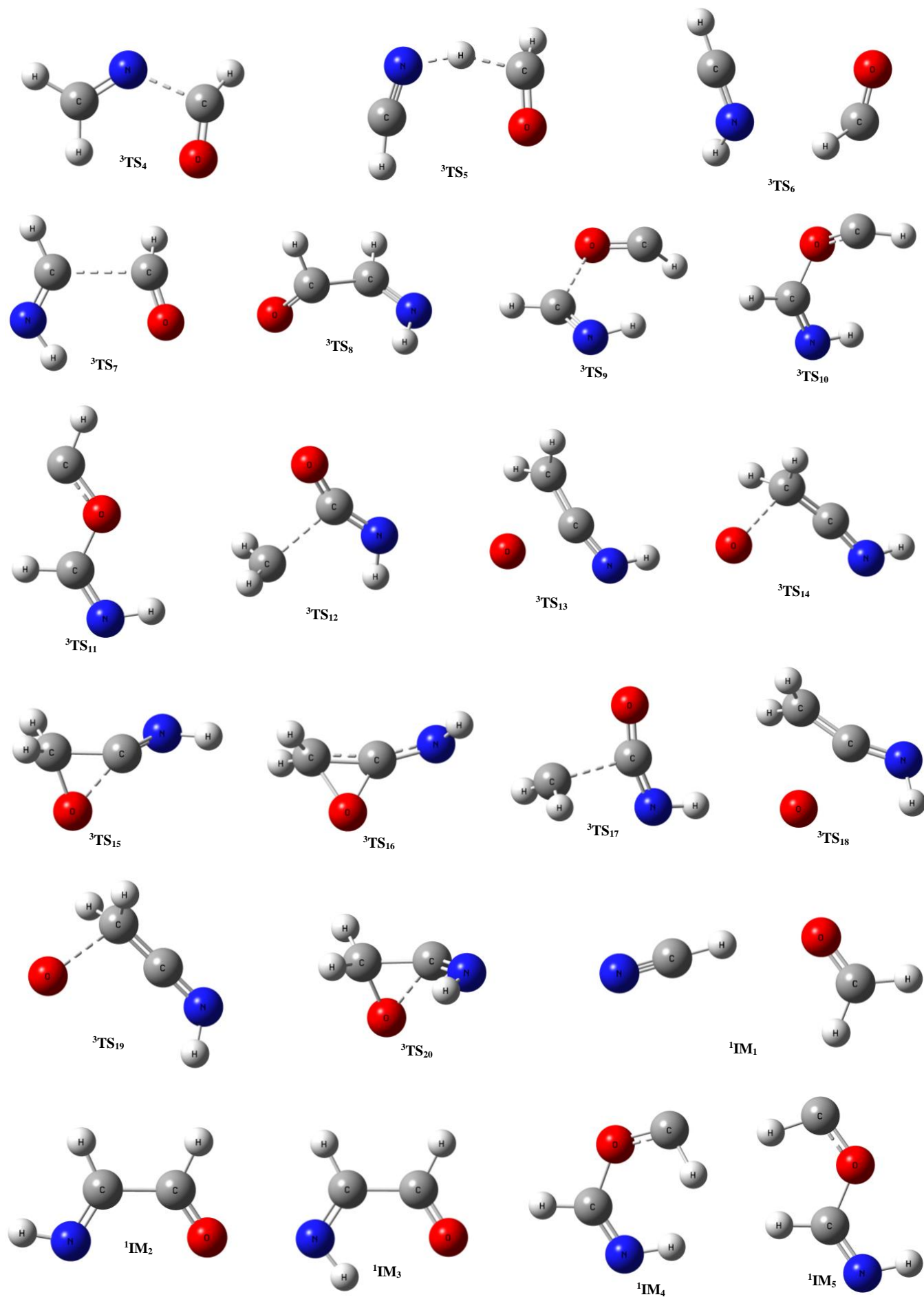


Fig. 2. (Continued).



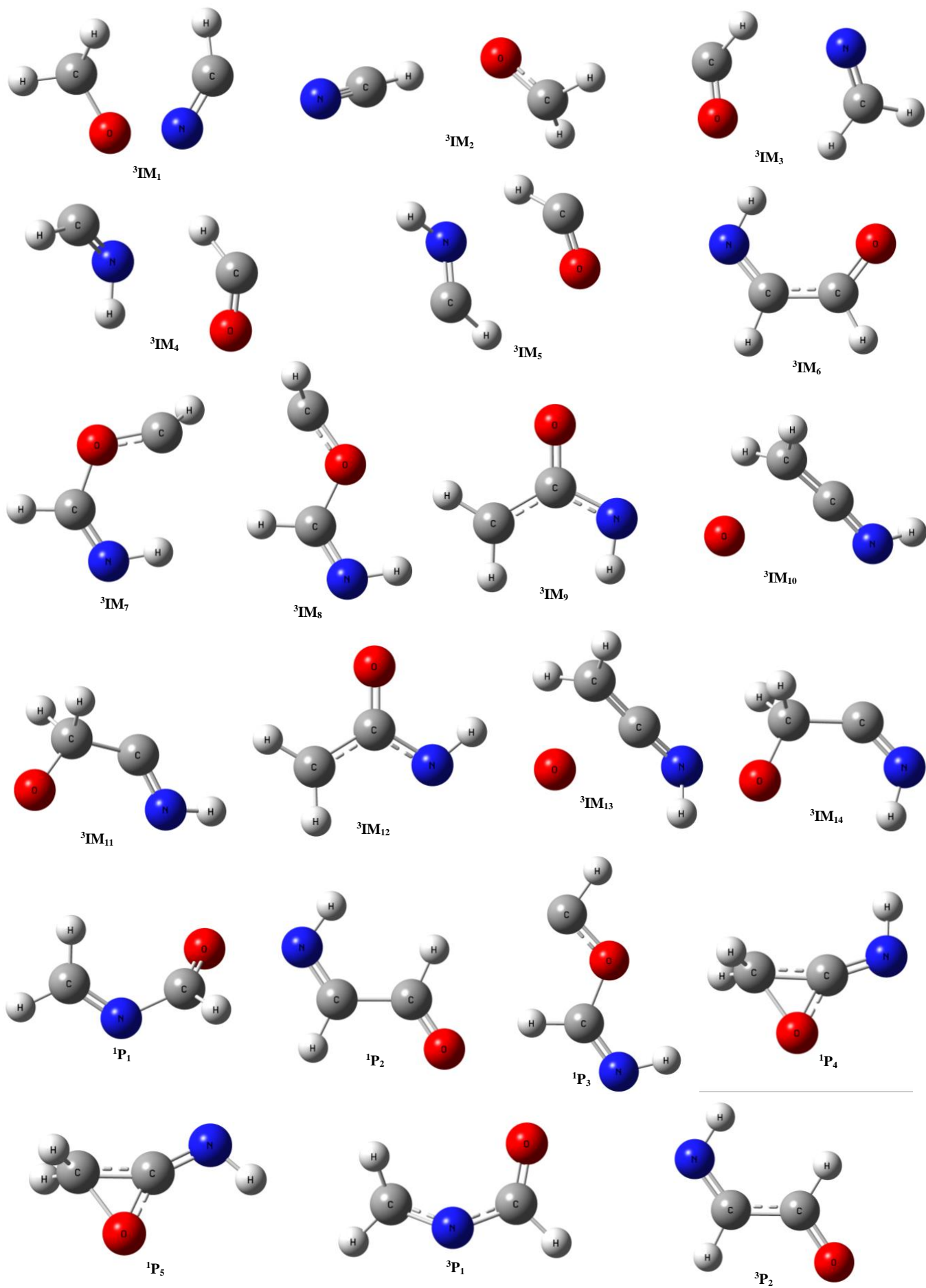


Fig. 2. (Continued).

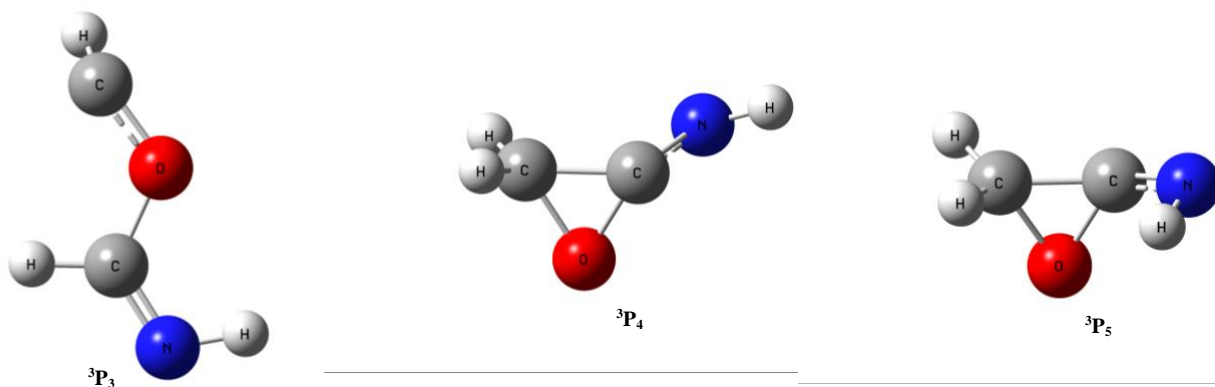


Fig. 2. (Continued).

**Table 2.** Energies ( $E$ s, a.u.), and relative energies (RE, kcal/mol) for the reactants, reactant complexes (RCs), transition states (TSs), intermediates (IMs), and products (Ps) for the reactions of  $\text{CH}_2$  with  $\text{HCNO}$  and  $\text{HNCO}$ , on singlet and triplet potential energy surfaces<sup>b</sup>

Singlet	$E$ (a.u.)	RE (kcal/mol)	Singlet	$E$ (a.u.)	RE (kcal/mol)
$^1\text{CH}_2 + \text{HCNO}$	-207.4127626 (-207.677513)	0.00 (0.00)	$^1\text{TS}_8$	-207.4570025 (-207.727097)	-27.76 (-31.11)
$^1\text{CH}_2 + \text{HNCO}$	-207.5233873 (-207.789059)	0.00 (0.00)	$^1\text{TS}_9$	-207.5805994 (-207.852908)	-35.90 (-40.07)
$^1\text{RC}_1$	-207.4255828 (-207.691713)	-8.04 (-8.91)	$^1\text{IM}_1$	-207.6378576 (-207.898754)	-141.25 (-138.83)
$^1\text{RC}_2^c$	-207.5276957 (-207.794436)	-2.70 (-3.37)	$^1\text{IM}_2$	-207.6383110 (-207.904179)	-141.53 (-142.24)
$^1\text{RC}_3^c$	-207.5251279 (-207.794367)	-1.09 (-3.33)	$^1\text{IM}_3$	-207.6440502 (-207.909825)	-145.14 (-145.78)
$^1\text{TS}_1$	-207.4149357 (-207.681737)	-1.36 (-2.65)	$^1\text{IM}_4$	-207.5445046 (-207.808732)	-82.67 (-82.34)
$^1\text{TS}_2$	-207.5347565 (-207.801660)	-76.55 (-77.90)	$^1\text{IM}_5$	-207.5479113 (-207.812760)	-84.81 (-84.87)
$^1\text{TS}_3$	-207.4930707 (-207.762031)	-50.39 (-53.04)	$^1\text{P}_1$	-207.6474740 (-207.915835)	-147.28 (-149.55)
$^1\text{TS}_4$	-207.5981231 (-207.866564)	-116.32 (-118.63)	$^1\text{P}_2$	-207.6451982 (-207.910901)	-145.86 (-146.45)
$^1\text{TS}_5$	-207.6361853 (-207.902231)	-140.20 (-141.01)	$^1\text{P}_3$	-207.5538158 (-207.818978)	-88.51 (-88.77)
$^1\text{TS}_6$	-207.4871563 (-207.746413)	-46.68 (-43.24)	$^1\text{P}_4$	-207.6127669 (-207.882985)	-56.09 (-58.94)
$^1\text{TS}_7$	-207.5414508 (-207.806329)	-80.75 (-80.83)	$^1\text{P}_5$	-207.6136523 (-207.883834)	-56.64 (-59.47)
Triplet	$E$ (a.u.)	RE (kcal/mol)	Triplet	$E$ (a.u.)	RE (kcal/mol)
$^3\text{CH}_2 + \text{HCNO}$	-207.4272761 (-207.696637)	-9.11 (-12.00)	$^3\text{TS}_{18}$	-207.4728192 (-207.745356)	31.73 (27.42)
$^3\text{CH}_2 + \text{HNCO}$	-207.5379008 (-207.808183)	-9.11 (-12.00)	$^3\text{TS}_{19}$	-207.4735590 (-207.747913)	31.27 (25.82)
$^3\text{RC}_1$	-207.4290474 (-207.697743)	-10.22 (-12.69)	$^3\text{TS}_{20}$	-207.4862105 (-207.758658)	23.33 (19.08)
$^3\text{RC}_2$	-207.5392324 (-207.809856)	-9.94 (-13.05)	$^3\text{IM}_1$	-207.4572629 (-207.728688)	-27.92 (-32.11)
$^3\text{RC}_3^c$	-207.5384130 (-207.808547)	-9.43 (-12.23)	$^3\text{IM}_2$	-207.5210788 (-207.789096)	-67.97 (-70.02)
$^3\text{TS}_1$	-207.3873336 (-207.661128)	15.96 (10.28)	$^3\text{IM}_3$	-207.5366754 (-207.805904)	-77.76 (-80.57)
$^3\text{TS}_2$	-207.4597489 (-207.729711)	-29.48 (-32.75)	$^3\text{IM}_4$	-207.5162937 (-207.785628)	-64.97 (-67.84)
$^3\text{TS}_3$	-207.4824223 (-207.755575)	-43.71 (-48.98)	$^3\text{IM}_5$	-207.5236526 (-207.791974)	-69.58 (-71.83)
$^3\text{TS}_4$	-207.5129872 (-207.785379)	-62.89 (-67.69)	$^3\text{IM}_6$	-207.5402356 (-207.818013)	-79.99 (-88.17)
$^3\text{TS}_5$	-207.4866958 (-207.758795)	-46.39 (-51.01)	$^3\text{IM}_7$	-207.5158644 (-207.786206)	-64.70 (-68.21)
$^3\text{TS}_6$	-207.4997159 (-207.769253)	-54.56 (-57.57)	$^3\text{IM}_8$	-207.5149810 (-207.785729)	-64.14 (-67.91)
$^3\text{TS}_7$	-207.5164750 (-207.787758)	-65.08 (-69.18)	$^3\text{IM}_9$	-207.5575397 (-207.831826)	-21.43 (-26.84)
$^3\text{TS}_8$	-207.5233604 (-207.797667)	-69.40 (-75.40)	$^3\text{IM}_{10}$	-207.4748374 (-207.747182)	30.47 (26.28)
$^3\text{TS}_9$	-207.4905591 (-207.760856)	-48.82 (-52.30)	$^3\text{IM}_{11}$	-207.5191279 (-207.792362)	2.67 (-2.07)
$^3\text{TS}_{10}$	-207.5073032 (-207.777536)	-59.33 (-62.77)	$^3\text{IM}_{12}$	-207.5623994 (-207.837195)	-24.48 (-30.21)
$^3\text{TS}_{11}$	-207.5126371 (-207.783471)	-62.67 (-66.49)	$^3\text{IM}_{13}$	-207.4762519 (-207.748969)	29.58 (25.16)

<sup>3</sup> TS <sub>12</sub>	-207.5202805 (-207.793478)	1.95 (-2.77)	<sup>3</sup> IM <sub>14</sub>	-207.5147488 (-207.788697)	5.42 (0.23)
<sup>3</sup> TS <sub>13</sub>	-207.4661591 (-207.738704)	35.91 (31.60)	<sup>3</sup> P <sub>1</sub>	-207.5433281 (-207.820537)	-81.93 (-89.75)
<sup>3</sup> TS <sub>14</sub>	-207.4730302 (-207.746845)	31.60 (26.49)	<sup>3</sup> P <sub>2</sub>	-207.5395170 (-207.817229)	-79.54 (-87.67)
<sup>3</sup> TS <sub>15</sub>	-207.4924691 (-207.764729)	19.40 (15.27)	<sup>3</sup> P <sub>3</sub>	-207.5149922 (-207.785738)	-64.15 (-67.91)
<sup>3</sup> TS <sub>16</sub>	-207.4701042 (-207.746624)	33.44 (26.63)	<sup>3</sup> P <sub>4</sub>	-207.5004641 (-207.774811)	14.38 (8.94)
<sup>3</sup> TS <sub>17</sub>	-207.5098942 (-207.784682)	8.47 (2.75)	<sup>3</sup> P <sub>5</sub>	-207.4948651 (-207.768897)	17.90 (12.65)

<sup>b</sup>Energies (*E*s, a.u.) and relative energies (kcal/mol) were calculated at the CCSD(T), and  $\omega$ B97XD (in parentheses), levels of theory.

## 5. Conclusion

The CH<sub>2</sub> + HCNO and CH<sub>2</sub> + HNCO reactions were identified as possible sources of as yet undetected interstellar prebiotic molecules by means of a screening procedure that combines energy barriers estimated by machine learning and empirical rules about chemical composition. These reactions were studied on the singlet and triplet energy surfaces by density functional theory at the  $\omega$ B97XD/def2-TZVP level for geometry optimization and normal mode frequencies and CCSD(T)/Aug-cc-pVQZ// $\omega$ B97XD/def2-TZVP for computing energy differences such as barriers to reaction. Singlet CH<sub>2</sub> reacts with HCNO without a barrier to produce *N*-methyleneformamide (<sup>1</sup>P<sub>1</sub>) and imine acetaldehyde (<sup>1</sup>P<sub>2</sub>), with the former being thermodynamically favoured by only ~1 kcal/mol. However, singlet CH<sub>2</sub> is 9.00 kcal/mol higher than ground state triplet CH<sub>2</sub> [89]. Therefore, reactions on the singlet energy surface would require triplet-to-singlet crossing and collision energies of ~9 kcal/mol, or less than 9 kcal/mol if smaller triplet-singlet energy gaps are present for some reactant complex geometries. Triplet CH<sub>2</sub> can react with HNCO with a barrier of 11.1 kcal/mol to produce <sup>3</sup>IM<sub>9</sub>, CH<sub>2</sub>C(=O)NH. These results do not give strong evidence for the formation of prebiotic molecules from reactions of triplet CH<sub>2</sub> with HCNO or HNCO in the ISM. However, the barriers to reaction relative to ground state reactants are found to be small for several reactions: ~9 kcal/mol for reactions 1 to 5 if triplet-to-singlet surface crossing occurs at large intermolecular distances, possibly less if crossing occurs at shorter distances; and on the triplet surface, ~11 kcal/mol for reaction 4, and ~18 kcal/mol for reaction 5. These results suggest three areas for future research: a more precise characterization of the triplet and singlet energy surfaces of weakly bound CH<sub>2</sub>-HCNO complexes, with a search for possible triplet-singlet crossing; the modeling of possible reactions of CH<sub>2</sub> with ice covered dust grains without and with coadsorbed HCNO; and a search for the spectroscopic signature of *N*-methyleneformamide in interstellar space.

## ASSOCIATED CONTENT

Element symbols with Cartesian coordinates of all singlet and triplet studied species are available in the Supplementary Information (SI) file.

## AUTHOR INFORMATION

### Corresponding Author

\*ragyanszki@zib.de

### Declaration of Competing Interest

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC). This research was enabled in part by support provided by Sharcnet (sharcnet.ca) and the Digital Research Alliance of Canada (alliancecan.ca).

## REFERENCES

- [1] M. Guélin, J. Cernicharo, Organic molecules in interstellar space: Latest advances, *Front. Astron. Space Sci.* 9 (2022). <https://doi.org/10.3389/fspas.2022.787567>.
- [2] a) C. F. Chyba, C. Sagan, Comets as a source of prebiotic organic molecules for the early earth, in: P. J. Thomas, C. F. Chyba, C. P. McKay (Eds.), *Comets and the origin and evolution of life*, Springer, New York, 1997, pp. 147–173; b) V. Barone, M. Biczysko, C. Puzzarini, Quantum chemistry meets spectroscopy for astrochemistry: Increasing complexity toward prebiotic molecules, *Acc. Chem. Res.* 48(5) (2015)

- 1413–1422. <https://doi.org/10.1021/ar5003285>; c) F. Tian, O. B. Toon, A. A. Pavlov, H. De Sterck, A hydrogen-rich early Earth atmosphere, *Science* 308 (2005) 1014–1017. <https://doi.org/10.1126/science.1106983>.
- [3] a) A. G. G. M. Tielens, S. B. Charnley, Circumstellar and interstellar synthesis of organic molecules, *Biophys. J.* 27 (1997) 23–51; b) D. Rehder, *The Interstellar Medium. Chemistry in Space: from Interstellar Matter to the Origin of Life*, Weinheim, Wiley-VCH, 2010.
- [4] S. Iglesias-Groth, A search for tryptophan in the gas of the IC 2348 star cluster of the Perseus molecular cloud, *Mon. Notices Royal Astron. Soc.* 523(2) (2023) 2876–2886. <https://doi.org/10.1093/mnras/stad1535>.
- [5] E. Herbst, Unusual chemical processes in interstellar chemistry: Past and present, *Front. Astron. Space Sci.* 8 (2021). <https://doi.org/10.3389/fspas.2021.776942>.
- [6] a) K. M. Ferrière, The interstellar environment of our galaxy, *Rev. Mod. Phys.* 73 (4) (2001) 1031–1066. <https://doi.org/10.1103/RevModPhys.73.1031>; b) L. M. Ziurys, The chemistry in circumstellar envelopes of evolved stars: following the origin of the elements to the origin of life, *Proc. Natl. Acad. Sci. U.S.A.* 103 (33) (2006) 12274–12279. <https://doi.org/10.1073/pnas.0602277103>.
- [7] a) B. J. Lynch, D. G. Truhlar, Small representative benchmarks for thermochemical calculations, *J. Phys. Chem. A* 107(42) (2003) 8996–8999. <https://doi.org/10.1021/jp035287b>; b) J. Zheng, Y. Zhao, D. G. Truhlar, The DBH24/08 database and its use to assess electronic structure model chemistries for chemical reaction barrier heights, *J. Chem. Theory Comput.* 5(4) (2009) 808–821. <https://doi.org/10.1021/ct800568m>.
- [8] a) C. A. Grambow, L. Pattanaik, W. H. Green, Reactants, products, and transition states of elementary chemical reactions based on quantum chemistry, *Sci. Data* 7(137) (2020). <https://doi.org/10.1038/s41597-020-0460-4>; b) [dataset] C. A. Grambow, L. Pattanaik, W. H. Green, Reactants, products, and transition states of elementary chemical reactions based on quantum chemistry, Zenodo, 1.0.1, 2020. <https://doi.org/10.5281/zenodo.3715478> (accessed 9 June 2021); c) L. Ruddigkeit, R. Van Deursen, L. C. Blum, J. L. Reymond, Enumeration of 166 billion organic small molecules in the chemical universe database GDB-17, *J. Chem. Inf. Model.* 52(11) (2012) 2864–2875. <https://doi.org/10.1021/ci300415d>.
- [9] S. Panda, A. Chiranjibi, D. Awasthi, A. Anoop, Exploring the Chemical Space of C<sub>3</sub>H<sub>2</sub>NO Isomers and Bimolecular Reactions with Hydrogen Cyanide and Formaldehyde: Insights into the Emergence of Life, *ChemRxiv*. (2023). <https://doi.org/10.26434/chemrxiv-2023-hfcl7-v2>. This content is a preprint and has not been peer-reviewed.
- [10] a) Y. Li, H.-L. Liu, Y.-B. Sun, Z. Li, X.-R. Huang, C.-C. Sun, Radical reaction HCNO + <sup>3</sup>NH: a mechanistic study, *Theor. Chem. Acc.* 124 (2009) 123–137. <https://doi.org/10.1007/s00214-009-0591-3>; b) Y. Gao, X.-J. Jia, S. Li, Y.-B. Yu, R.-S. Wang, X.-M. Pan, Theoretical study on the singlet and triplet potential energy surfaces of NH (X<sup>3</sup>Σ<sup>-</sup>) + HCNO reaction; c) J. A. Miller, S. J. Klippenstein, P. Glarborg, A kinetic issue in reburning: the fate of HCNO, *Combust. Flame* 135(3) (2003) 357–362. <https://doi.org/10.1016/j.combustflame.2003.07.002>; d) H. M. T. Nguyen, T. N. Nguyen, Calculations on the complex mechanism of the HCNO + OH reaction, *Chem. Phys. Lett.* 599 (2014) 15–22. <https://doi.org/10.1016/j.cplett.2014.03.001>; e) W. Feng, J. P. Meyer, J. F. Hershberger, Kinetics of the OH + HCNO Reaction, *J. Phys. Chem. A* 110 (2006) 4458–4464. <https://doi.org/10.1021/jp058305t>; f) T. N. Nguyen, H. M. T. Nguyen, Prediction of the product channels in the reaction of the methyl radical with fulminic acid, *Int. J. Chem. Kinet.* 53 (2021) 479–491. <https://doi.org/10.1002/kin.21458>; g) J.-L. Pang, H.-B. Xie, S.-W. Zhang, Y.-H. Ding, A.-Q. Tang, Theoretical study on reaction mechanism of fulminic acid HCNO with CN radical, *J. Phys. Chem. A* 112(23) (2008) 5251–5257. <https://doi.org/10.1021/jp709700u>; h) D. Mu, W.-X. Kong, S. Wang, A density functional investigation of the reaction mechanism of H<sub>2</sub>O + HCNO, *Int. J. Quantum Chem.* 111 (2011) 165–173. <https://doi.org/10.1002/qua.22745>; i) W. Feng, J. F. Hershberger, Theoretical study of the O(<sup>3</sup>P) + HCNO reaction, *Chem. Phys. Lett.* 457 (2008) 307–311. <https://doi.org/10.1016/j.cplett.2008.04.012>; j) T. V. Pham, A. V. Tran, Gas phase reaction of isocyanic acid: Kinetics, mechanisms, and formation of isopropyl aminocarbonyl, *ACS Omega* 6(50) (2021) 34661–34674. <https://doi.org/10.1021/acsomega.1c05063>; k) F. P. Tully, R. A. Perry, L. R. Thorne, M. D. Allendorf, Free-radical oxidation of isocyanic acid, *Symp. (Int.) Combust.* 22(1) (1989) 1101–1106. [https://doi.org/10.1016/S0082-0784\(89\)80120-1](https://doi.org/10.1016/S0082-0784(89)80120-1); l) W. Tsang, Chemical kinetic data base for propellant combustion. II. Reactions involving CN, NCO, and HNCO, *J. Phys. Chem. Ref. Data* 21 (1992) 753–791. <https://doi.org/10.1063/1.555914>; m) L.-C. Li, Y. Zheng, D. Zha, A.-M. Tian, M.-H. Xu, Theoretical studies on the reaction mechanism of CH<sub>2</sub>CH radical with HNCO, *Int. J. Quantum Chem.* 106 (2006) 1672–1682. <https://doi.org/10.1002/qua.20917>; n) Y.-Q. Ji, W.-L. Feng, Z.-F. Xu, M. Lei, M.-R. Hao, MP2 and QCISD Study of Hydrogen Transfer Reaction Path of the Reaction HNCO with Carbon-Hydrogen Radicals CH<sub>x</sub> (x = 1–3), *Acta Chim. Sinica* 60 (2002) 1167–1172.
- [11] L. M. Ziurys, The chemistry in circumstellar envelopes of evolved stars: Following the origin of the elements to the origin of life, *Proc. Natl. Acad. Sci. U.S.A.* 103(33) (2006) 12274–12279. <https://doi.org/10.1073/pnas.0602277103>.
- [12] W. R. F. Dent, M. C. Wyatt, A. Roberge, J. -C. Augereau, S. Casassus, S. Corder, J. S. Greaves, I. de Gregorio-Monsalvo, A. Hales, A. P. Jackson, A. Meredith Hughes, A. -M Lagrange, B. Matthews, D. Wilner, Molecular gas clumps from the destruction of icy bodies in the β Pictoris debris disk, *Science* 343(6178) (2014) 1490–1492. <https://doi.org/10.1126/science.1248726>.
- [13] G. R. Carruthers, Rocket observation of interstellar molecular hydrogen, *Astrophys. J.* 161 (1970) L81–L85. <http://doi.org/10.1086/180575>.
- [14] D. C. Knauth, B. -G. Andersson, S. R. McCandliss, H. W. Moos, The interstellar N<sub>2</sub> abundance towards HD 124314 from far-ultraviolet observations, *Nature* 429 (2004) 636–638. <https://doi.org/10.1038/nature02614>.
- [15] J. M. Hollis, P. R. Jewell, F. J. Lovas, Confirmation of interstellar methylene, *Astrophys. J.* 438 (1995) 259–264. <http://doi.org/10.1086/175070>.
- [16] A. Fuente, S. García-Burillo, M. Gerin, D. Teyssier, A. Usero, J. R. Rizzo, P. de Vicente, Photon-dominated chemistry in the nucleus of M82: Widespread HOC<sup>+</sup> emission in the inner 650 parsec disk, *Astrophys. J.* 619 (2005) L155–L158. <https://doi.org/10.1086/427990>.
- [17] L. E. Snyder, D. Buhl, Observations of radio emission from interstellar hydrogen cyanide, *Astrophys. J.* 163 (1971) L47–L52. <https://doi.org/10.1086/180664>.
- [18] D. C. B. Whittet, H. J. Walker, On the occurrence of carbon dioxide in interstellar grain mantles, *Mon. Notices Royal Astron. Soc.* 252(1) (1991) 63–67. <https://doi.org/10.1093/mnras/252.1.63>.
- [19] European Space Agency, Europe’s space telescope ISO finds water in distant places. [https://web.archive.org/web/20061222224048/http://www.iso.vilspa.esa.es/outreach/esa\\_pr/in9712.htm](https://web.archive.org/web/20061222224048/http://www.iso.vilspa.esa.es/outreach/esa_pr/in9712.htm) (accessed 7 March 2023).
- [20] L. M. Ziurys, A. J. Apponi, J. M. Hollis, L. E. Snyder, Detection of interstellar N<sub>2</sub>O: A new molecule containing an N-O bond, *Astrophys. J. Lett.* 436 (1994) L181–L184. <http://doi.org/10.1086/187662>.
- [21] L. E. Snyder, D. Buhl, B. Zuckerman, P. Palmer, Microwave detection of interstellar formaldehyde, *Phys. Rev. Lett.* 22(13) (1969) 679–681. <https://doi.org/10.1103/PhysRevLett.22.679>.
- [22] S. T. Ridgway, D. N. B. Hall, S. G. Kleinmann, D. A. Weinberger, R. S. Wojslaw, Circumstellar acetylene in the infrared spectrum of IRC +10° 216, *Nature* 264 (1976) 345–346. <https://doi.org/10.1038/264345a0>.
- [23] L. E. B. Johansson, C. Andersson, J. Ellidér, P. Friberg, Å. Hjalmarsen, B. Höglund, W. M. Irvine, H. Olofsson, G. Rydbeck, Spectral scan of Orion A and IRC + 10216 from 72 to 91 GHz, *Astron. Astrophys.* 130 (1984) 227–256.
- [24] N. Marcelino, J. Cernicharo, B. Tercero, E. Roueff, Discovery of fulminic acid, HCNO, in dark clouds, *Astrophys. J.* 690 (2009) L27–L30. <http://dx.doi.org/10.1088/0004-637X/690/1/L27>.
- [25] M. Guélin, J. Cernicharo, Astronomical detection of the HCCN radical. Toward a new family of carbon-chain molecules? *Astron. Astrophys.* 244 (1991) L21–L24.

- [26] M. Agúndez, N. Marcelino, J. Cernicharo, Discovery of interstellar isocyanogen (CNCN): Further evidence that dicyanopolynes are abundant in space, *Astrophys. J. Lett.* 861 (2018) L22. <https://doi.org/10.3847/2041-8213/aad089>.
- [27] S. Brünken, A. Belloche, S. Martín, L. Verheyen, K. M. Menten, Interstellar HOCN in the galactic center region, *Astron. Astrophys.* 516 (2010) A109. <https://doi.org/10.1051/0004-6361/200912456>.
- [28] N. Q. Rieu, D. Graham, V. Bujarrabal, Ammonia and cyanodiacetylene in the envelopes of CRL 2688 and IRC + 10216, *Astron. Astrophys.* 138 (1984) L5–L8.
- [29] J. H. Lacy, J. S. Carr, N. J. Evans II, F. Baas, J. M. Achtermann, J. F. Arens, Discovery of interstellar methane: Observations of gaseous and solid CH<sub>4</sub> absorption toward young stars in molecular clouds, *Astrophys. J.* 376 (1991) 556–560. <https://doi.org/10.1086/170304>.
- [30] J. E. Dickens, W. M. Irvine, C. H. DeVries, M. Ohishi, Hydrogenation of interstellar molecules: A survey for methylenimine (CH<sub>2</sub>NH), *Astrophys. J.* 479 (1997) 307–312. <https://doi.org/10.1086/303884>.
- [31] B. A. McGuire, R. A. Loomis, C. M. Charness, J. F. Corby, G. A. Blake, J. M. Hollis, F. J. Lovas, P. R. Jewell, A. J. Remijan, Interstellar carbodiimide (HNCNH): A new astronomical detection from the GBT PRIMOS survey via maser emission features, *Astrophys. J. Lett.* 758 (2012) L33. <http://doi.org/10.1088/2041-8205/758/2/L33>.
- [32] B. E. Turner, H. S. Liszt, N. Kaifu, A. G. Kisliakov, Microwave detection of interstellar cyanamide, *Astrophys. J.* 201 (1975) L149–L152. <https://doi.org/10.1086/181963>.
- [33] N. F. W. Ligterink, S. J. El-Abd, C. L. Brogan, T. R. Hunter, A. J. Remijan, R. T. Garrod, B. M. McGuire, The family of amide molecules toward NGC 6334I, *Astrophys. J.* 901 (2020) 37. <https://doi.org/10.3847/1538-4357/abad38>.
- [34] K. Kawaguchi, M. Ohishi, S. Ishikawa, Detection of isocyanoacetylene HCCNC in TMC-1, *Astrophys. J.* 386 (1992) L51–L53. <http://doi.org/10.1086/186290>.
- [35] A. J. Remijan, J. M. Hollis, F. J. Lovas, W. D. Stork, P. R. Jewell, D. S. Meier, Detection of interstellar cyanoformaldehyde (CNCHO), *Astrophys. J.* 675 (2008) L85–L88. <http://doi.org/10.1086/533529>.
- [36] B. Zuckerman, J. A. Ball, C. A. Gottlieb, Microwave detection of interstellar formic acid, *Astrophys. J.* 163 (1971) L41–L45. <https://doi.org/10.1086/180663>.
- [37] S. Liu, D. M. Mehringer, L. E. Snyder, Observations of formic acid in hot molecular cores, *Astrophys. J.* 552 (2001) 654–663. <http://doi.org/10.1086/320563>.
- [38] C. M. Walmsley, G. Winnewisser, F. Toelle, Cyanoacetylene and cyanodiacetylene in interstellar clouds, *Astron. Astrophys.* 81 (1980) 245–250.
- [39] A. J. Remijan, J. M. Hollis, F. J. Lovas, D. F. Plusquellic, P. R. Jewell, Interstellar isomers: The importance of bonding energy differences, *Astrophys. J.* 632 (2005) 333–339. <http://doi.org/10.1086/432908>.
- [40] A. L. Betz, Ethylene in IRC +10216, *Astrophys. J.* 244 (1981) L103–L105. <http://doi.org/10.1086/183490>.
- [41] W. M. Irvine, R. D. Brown, D. M. Cragg, P. Friberg, P. D. Godfrey, N. Kaifu, H. E. Matthews, M. Ohishi, H. Suzuki, A new interstellar polyatomic molecule: Detection of propynal in the cold cloud TMC-1, *Astrophys. J.* 335 (1988) L89–L93. <https://doi.org/10.1086/185346>.
- [42] J. M. Hollis, F. J. Lovas, A. J. Remijan, P. R. Jewell, V. V. Ilyushin, I. Kleiner, Detection of acetamide (CH<sub>3</sub>CONH<sub>2</sub>): The largest interstellar molecule with a peptide bond, *Astrophys. J.* 643 (2006) L25–L28. <https://doi.org/10.1086/505110>.
- [43] J. Cernicharo, A. M. Heras, A. G. G. M. Tielens, J. R. Pardo, F. Herpin, M. Guélin, L. B. F. M. Waters, Infrared space observatory's discovery of C<sub>4</sub>H<sub>2</sub>, C<sub>6</sub>H<sub>2</sub>, and benzene in CRL 618, *Astrophys. J.* 546 (2001) L123–L126. <https://doi.org/10.1086/318871>.
- [44] D. P. Zaleski, N. A. Seifert, A. L. Steber, M. T. Muckle, R. A. Loomis, J. F. Corby, O. Martinez Jr., K. N. Crabtree, P. R. Jewell, J. R. Hollis, Detection of E-cyanomethanimine toward Sagittarius B2(N) in the Green Bank Telescope PRIMOS survey, *Astrophys. J.* 765 (2013) L10. <https://doi.org/10.1088/2041-8205/765/1/L10>.
- [45] National Radio Astronomy Observatory, Researchers use NRAO telescope to study formation of chemical precursors to life. <https://www.nrao.edu/pr/2006/gbtmolecules/> (accessed 7 March 2023).
- [46] J. M. Hollis, A. J. Remijan, P. R. Jewell, F. J. Lovas, Cyclopropenone (c-H<sub>2</sub>C<sub>3</sub>O): A new interstellar ring molecule, *Astrophys. J.* 642 (2006) 933–939. <http://doi.org/10.1086/501121>.
- [47] D. T. Halfen, V. V. Ilyushin, L. M. Ziurys, Interstellar detection of methyl isocyanate CH<sub>3</sub>NCO in Sgr B2(N): A link from molecular clouds to comets, *Astrophys. J. Lett.* 812 (2015) L5. <http://doi.org/10.1088/2041-8205/812/1/L5>.
- [48] S. Zeng, D. Quénard, I. Jiménez-Serra, J. Martín-Pintado, V. M. Rivilla, L. Testi, R. Martín-Doménech, First detection of the pre-biotic molecule glycolonitrile (HOCH<sub>2</sub>CN) in the interstellar medium, *Mon. Not. R. Astron. Soc.: Lett.* 484(1) (2019) L43–L48. <https://doi.org/10.1093/mnras/lsz002>.
- [49] N. Kaifu, K. Takagi, T. Kojima, Excitation of interstellar methylamine, *Astrophys. J.* 198 (1975) L85–L88. <http://doi.org/10.1086/181818>.
- [50] L. Bizzocchi, D. Prudeniano, V. M. Rivilla, A. Pietropoli-Charmet, B. M. Giuliano, P. Caselli, J. Martín-Pintado, I. Jiménez-Serra, S. Martín, M. A. Requena-Torres, F. Rico-Villas, S. Zeng, J. -C. Guillemin, Propargylimine in the laboratory and in space: Millimetre-wave spectroscopy and its first detection in the ISM, *Astron. Astrophys.* 640 (2020) A98. <https://doi.org/10.1051/0004-6361/202038083>.
- [51] J. E. Dickens, W. M. Irvine, M. Ohishi, M. Ikeda, S. Ishikawa, A. Nummelin, Å. Hjalmarson, Detection of interstellar ethylene oxide (c-C<sub>2</sub>H<sub>4</sub>O), *Astrophys. J.* 489 (1997) 753–757. <https://doi.org/10.1086/304821>.
- [52] National Radio Astronomy Observatory, Scientists toast the discovery of vinyl alcohol in interstellar space. <https://www.nrao.edu/pr/2001/vinylalco/> (accessed 7 March 2023).
- [53] A. Belloche, K. M. Menten, C. Comito, H. S. P. Müller, P. Schilke, J. Ott, S. Thorwirth, C. Hieret, Detection of amino acetonitrile in Sgr B2(N), *Astron. Astrophys.* 482 (2008) 179–196. <https://doi.org/10.1051/0004-6361:20079203>.
- [54] A. J. Remijan, L. E. Snyder, B. A. McGuire, H. Kuo, L. W. Looney, D. N. Friedel, G. Y. Golubiatnikov, F. J. Lovas, V. V. Ilyushin, E. A. Alekseev, S. F. Dyubko, B. J. McCall, J. M. Hollis, Observational results of a multi-telescope campaign in search of interstellar urea [(NH<sub>2</sub>)<sub>2</sub>CO], *Astrophys. J.* 783(77) (2014). <http://doi.org/10.1088/0004-637X/783/2/77>.
- [55] D. M. Mehringer, L. E. Snyder, Y. Miao, Detection and confirmation of interstellar acetic acid, *Astrophys. J.* 480 (1997) L71–L74. <http://doi.org/10.1086/310612>.
- [56] R. A. Loomis, D. P. Zaleski, A. L. Steber, J. L. Neill, M. T. Muckle, B. J. Harris, J. M. Hollis, P. R. Jewell, V. Lattanzi, F. J. Lovas, The detection of interstellar ethanimine (CH<sub>3</sub>CHNH) from observations taken during the GBT PRIMOS survey, *Astrophys. J. Lett.* 765 (2013) L9. <http://doi.org/10.1088/2041-8205/765/1/L9>.
- [57] B. A. McGuire, A. M. Burkhardt, R. A. Loomis, C. N. Shingledecker, K. L. K. Lee, S. B. Charnley, M. A. Cordiner, E. Herbst, S. Kalenskii, E. Momjian, Early science from GOTHAM: Project overview, methods, and the detection of interstellar propargyl cyanide (HCCCH<sub>2</sub>CN) in TMC-1, *Astrophys. J. Lett.* 900 (2020) L10. <http://doi.org/10.3847/2041-8213/aba632>.
- [58] F. J. Lovas, A. J. Remijan, J. M. Hollis, P. R. Jewell, L. E. Snyder, Hyperfine structure identification of interstellar cyanoallene toward TMC-1, *Astrophys. J.* 637 (2006) L37–L40. <http://doi.org/10.1086/500431>.

- [59] S. Zeng, I. Jiménez-Serra, V. M. Rivilla, J. Martín-Pintado, L. F. Rodríguez-Almeida, B. Tercero, P. de Vicente, F. Rico-Villas, L. Colzi, S. Martín, M. A. Requena-Torres, Probing the chemical complexity of amines in the ISM: Detection of vinylamine (C<sub>2</sub>H<sub>5</sub>NH<sub>2</sub>) and tentative detection of ethylamine (C<sub>2</sub>H<sub>5</sub>NH<sub>2</sub>), *Astrophys. J. Lett.* 920 (2021) L27. <http://doi.org/10.3847/2041-8213/ac2c7e>.
- [60] L. E. Snyder, D. Buhl, P. R. Schwartz, F. O. Clark, D. R. Johnson, F. J. Lovas, P. T. Giguere, Radio detection of interstellar dimethyl ether, *Astrophys. J.* 191 (1974) L79–L82. <http://doi.org/10.1086/181554>.
- [61] B. Zuckerman, B. E. Turner, D. R. Johnson, F. O. Clark, F. J. Lovas, N. Fourikis, P. Palmer, M. Morris, A. E. Lilley, J. A. Ball, C. A. Gottlieb, M. M. Litvak, H. Penfield, Detection of interstellar trans-ethyl alcohol, *Astrophys. J.* 196 (1975) L99–L102. <http://doi.org/10.1086/181753>.
- [62] N. Marcelino, J. Cernicharo, M. Agúndez, E. Roueff, M. Gerin, J. Martín-Pintado, R. Mauersberger, C. Thum, Discovery of interstellar propylene (CH<sub>2</sub>CHCH<sub>3</sub>): Missing links in interstellar gas-phase chemistry, *Astrophys. J.* 665 (2007) L127–L130. <http://doi.org/10.1086/521398>.
- [63] A. J. Remijan, J. M. Hollis, L. E. Snyder, P. R. Jewell, F. J. Lovas, Methyltriacetylene (CH<sub>3</sub>C<sub>6</sub>H) toward TMC-1: The largest detected symmetric top, *Astrophys. J.* 643 (2006) L37–L40. <http://doi.org/10.1086/504918>.
- [64] B. A. McGuire, P. B. Carroll, R. A. Loomis, I. A. Finneran, P. R. Jewell, A. J. Remijan, G. A. Blake, Discovery of the interstellar chiral molecule propylene oxide (CH<sub>3</sub>CHCH<sub>2</sub>O), *Science* 352(6292) (2016) 1449–1452. <https://doi.org/10.1126/science.aae0328>.
- [65] L. E. Snyder, F. J. Lovas, D. M. Mehringer, N. Y. Miao, Y. Kuan, J. M. Hollis, P. R. Jewell, Confirmation of interstellar acetone, *Astrophys. J.* 578 (2002) 245–255. <https://doi.org/10.1086/342273>.
- [66] V. M. Rivilla, I. Jiménez-Serra, J. Martín-Pintado, C. Briones, L. F. Rodríguez-Almeida, F. Rico-Villas, B. Tercero, S. Zeng, L. Colzi, P. de Vicente, S. Martín, M. A. Requena-Torres, Discovery in space of ethanalamine, the simplest phospholipid head group, *Proc. Natl. Acad. Sci. U.S.A.* 118(22) (2021) e2101314118. <https://doi.org/10.1073/pnas.2101314118>.
- [67] B. Tercero, I. Kleiner, J. Cernicharo, H. V. L. Nguyen, A. López, G. M. Muñoz Caro, Discovery of methyl acetate and gauche ethyl formate in Orion, *Astrophys. J. Lett.* 770 (2013) L13. <https://doi.org/10.1088/2041-8205/770/1/L13>.
- [68] A. Belloche, R. T. Garrod, H. S. P. Müller, K. M. Menten, Detection of a branched alkyl molecule in the interstellar medium: Iso-propyl cyanide, *Science* 345(6204) (2014) 1584–1587. <https://doi.org/10.1126/science.1256678>.
- [69] A. Belloche, R. T. Garrod, H. S. P. Müller, K. M. Menten, C. Comito, P. Schilke, Increased complexity in interstellar chemistry: Detection and chemical modeling of ethyl formate and *n*-propyl cyanide in Sagittarius B2(N), *Astron. Astrophys.* 499 (2009) 215–232. <https://doi.org/10.1051/0004-6361/200811550>.
- [70] G. W. Fuchs, U. Fuchs, T. F. Giesen, F. Wyrowski, Trans-ethyl methyl ether in space: A new look at a complex molecule in selected hot core regions, *Astron. Astrophys.* 444 (2005) 521–530. <https://doi.org/10.1051/0004-6361:20053599>.
- [71] H. Ji, A. Rágyanszki, R. A. Fournier, Machine learning estimation of reaction energy barriers, [Manuscript submitted for publication] (2023). Available at SSRN: <https://ssrn.com/abstract=4535818> or <http://dx.doi.org/10.2139/ssrn.4535818>.
- [72] D. Weininger, SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules, *J. Chem. Inf. Comput. Sci.* 28 (1988) 31–36. <https://doi.org/10.1021/ci00057a005>.
- [73] P. Redondo, M. Sanz-Novo, C. Barrientos, A. Largo, Structure, spectroscopic signatures, and formation of hydroxy-azirine: a potential interstellar prebiotic molecule, *Mon. Notices Royal Astron. Soc.* 520(2) (2023) 1659–1671. <https://doi.org/10.1093/mnras/stad205>.
- [74] P. Redondo, A. Largo, C. Barrientos, Structure and spectroscopic properties of imine acetaldehyde: a possible interstellar molecule, *Mon. Notices Royal Astron. Soc.* 478(3) (2018) 3042–3048. <https://doi.org/10.1093/mnras/sty1175>.
- [75] I. Fourré, O. Matz, Y. Ellinger, J.-C. Guillemin, Relative thermodynamic stability of the [C,N,O] linkages as an indication of the most abundant structures in the ISM, *Astron. Astrophys.* 639 (2020) A16. <https://doi.org/10.1051/0004-6361/202037839>.
- [76] a) J. Čížek, On the use of the cluster expansion and the technique of diagrams in calculations of correlation effects in atoms and molecules, in: R. LeFebvre, C. Moser (Eds.), *Advances in chemical physics*, Wiley Interscience, New York, 1969, pp. 35–89; b) G. D. Purvis III, R. J. Bartlett, A full coupled-cluster singles and doubles model – the inclusion of disconnected triples, *J. Chem. Phys.* 76 (1982) 1910–1918. <https://doi.org/10.1063/1.443164>; c) G. E. Scuseria, C. L. Janssen, H. F. Schaefer III, An efficient reformulation of the closed-shell coupled cluster single and double excitation (CCSD) equations, *J. Chem. Phys.* 89 (1988) 7382–7387. <https://doi.org/10.1063/1.455269>; d) G. E. Scuseria, H. F. Schaefer III, Is coupled cluster singles and doubles (CCSD) more computationally intensive than quadratic configuration-interaction (QCISD)? *J. Chem. Phys.* 90 (1989) 3700–3703. <https://doi.org/10.1063/1.455827>; e) J. A. Pople, M. Head-Gordon, K. Raghavachari, Quadratic configuration interaction – a general technique for determining electron correlation energies, *J. Chem. Phys.* 87 (1987) 5968–5975. <https://doi.org/10.1063/1.453520>.
- [77] a) N. M. O’Boyle, M. Banck, C. A. James, C. Morley, V. Tim, G. R. Hutchison, Open Babel: an open chemical toolbox, *J. Cheminform.* 3 (2011) 1–14. <https://doi.org/10.1186/1758-2946-3-33>; b) N. Yoshikawa, G. R. Hutchison, Fast, efficient fragment-based coordinate generation for Open Babel, *J. Cheminform.* 11 (2019) 1–9. <https://doi.org/10.1186/s13321-019-0372-5>.
- [78] a) Y. S. Lin, G. D. Li, S. P. Mao, J. D. Chai, Long-range corrected hybrid density functionals with improved dispersion corrections, *J. Chem. Theory Comput.* 9 (2013) 263–272. <https://doi.org/10.1021/ct300715s>; b) J. D. Chai, M. Head-Gordon, Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections, *Phys. Chem. Chem. Phys.* 10 (2008) 6615–6620. <https://doi.org/10.1039/B810189B>; c) J. D. Chai, M. Head-Gordon, Systematic optimization of long-range corrected hybrid density functionals, *J. Chem. Phys.* 128(8) (2008) 084106. <https://doi.org/10.1063/1.2834918>.
- [79] a) T. MacKee, J. R. MacKee, Amino acids, peptides, and proteins, in: *Biochemistry: The Molecular Basis of Life*, 6<sup>th</sup> ed., Oxford University Press, New York, 2016, pp. 130–135; b) T. MacKee, J. R. MacKee, M. G. Sehorn, Amino acids, peptides, and proteins, in: *Student Study Guide/Solutions Manual for Use with Biochemistry: The Molecular Basis of Life*, 6<sup>th</sup> ed., Oxford University Press, New York, 2016, pp. 59–61; c) C. K. Mathews, K. E. Van Holde, D. R. Appling, S. J. Anthony-Cahill, Introduction to proteins: The primary level of proteins structure, in: *Biochemistry*, 4<sup>th</sup> ed., Pearson, Toronto, 2013, pp. 138–139; d) J. M. Otaki, M. Tsutsumi, T. Gotoh, H. Yamamoto, Secondary structure characterization based on amino acid composition and availability in proteins, *J. Chem. Inf. Model.* 50(4) (2010) 690–700. <https://doi.org/10.1021/ci900452z>.
- [80] C. K. Mathews, K. E. Van Holde, D. R. Appling, S. J. Anthony-Cahill, Carbohydrates: Sugars, saccharides, glycans, in: *Biochemistry*, 4<sup>th</sup> ed., Pearson, Toronto, 2013, pp. 309–311.
- [81] K. Chandru, N. Guttenberg, C. Giri, Y. Hongo, C. Butch, I. Mamajanov, H. J. Cleaves II, Simple prebiotic synthesis of high diversity dynamic combinatorial polyester libraries, *Commun. Chem.* 1(30) (2018). <https://doi.org/10.1038/s42004-018-0031-1>.
- [82] F. Mao, W. Ni, X. Xu, H. Wang, J. Wang, M. Ji, J. Li, Chemical structure-related drug-like criteria of global approved drugs, *Molecules* 21(75) (2016). <https://doi.org/10.3390/molecules21010075>.
- [83] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T.

- Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian 16, Revision C.01, Gaussian, Inc., Wallingford CT, 2016.
- [84] R. Dennington, T. A. Keith, J. M. Millam, GaussView, Version 6., Semichem Inc., Shawnee Mission, KS, 2016.
- [85] a) Avogadro: an open-source molecular builder and visualization tool (Version 1.2.0.), June 15, 2016. <http://avogadro.cc/> (accessed 21 April 2023); b) M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, G. R. Hutchison, Avogadro: an advanced semantic chemical editor, visualization, and analysis platform, *J. Cheminform.* 4 (2012) 1–17. <https://doi.org/10.1186/1758-2946-4-17>.
- [86] a) A. Schaefer, C. Huber, R. Ahlrichs, Fully optimized contracted gaussian-basis sets of triple zeta valence quality for atoms Li to Kr, *J. Chem. Phys.* 100 (1994) 5829–5835. <https://doi.org/10.1063/1.467146>; b) F. Weigend, R. Ahlrichs, Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: design and assessment of accuracy, *Phys. Chem. Chem. Phys.* 7 (2005) 3297–3305. <https://doi.org/10.1039/B508541A>; c) F. Weigend, Accurate coulomb-fitting basis sets for H to Rn, *Phys. Chem. Chem. Phys.* 8 (2006) 1057–1065. <https://doi.org/10.1039/B515623H>.
- [87] a) K. Fukui, The path of chemical reactions – The IRC approach, *Acc. Chem. Res.* 14(12) (1981) 363–368. <https://doi.org/10.1021/ar00072a001>; b) H. P. Hratchian, H. B. Schlegel, Finding minima, transition states, and following reaction pathways on *ab initio* potential energy surfaces, in: C. E. Dykstra, G. Frenking, K. S. Kim, G. Scuseria (Eds.), *Theory and applications of computational chemistry: The first 40 years*, Elsevier, Amsterdam, 2005, pp. 195–249.
- [88] a) R. A. Kendall, T. H. Dunning Jr., R. J. Harrison, Electron affinities of the first-row atoms revisited. Systematic basis sets and wave functions, *J. Chem. Phys.* 96 (1992) 6796–6806. <https://doi.org/10.1063/1.462569>; b) D. E. Woon, T. H. Dunning Jr., Gaussian-basis sets for use in correlated molecular calculations. III. The atoms Aluminum through Argon, *J. Chem. Phys.* 98 (1993) 1358–1371. <https://doi.org/10.1063/1.464303>; c) E. Papajak, J. Zheng, H. R. Leverentz, D. G. Truhlar, Perspectives on basis sets beautiful: Seasonal plantings of diffuse basis functions, *J. Chem. Theory and Comput.* 7 (2011) 3027–3034. <https://doi.org/10.1021/ct200106a>.
- [89] NIST Chemistry WebBook, NIST Standard Reference Database Number 69. <https://doi.org/10.18434/T4D303> (accessed 17 July 2023).
- [90] List of interstellar and circumstellar molecules. [https://en.wikipedia.org/wiki/List\\_of\\_interstellar\\_and\\_circumstellar\\_molecules#Molecules](https://en.wikipedia.org/wiki/List_of_interstellar_and_circumstellar_molecules#Molecules) (accessed 2 March 2023).
- [91] H. M. Badawi, DFT and MP2 vibrational spectra and assignments for *gauche* *N*-methyleneformamide CH<sub>2</sub>=N–CHO, *Theochem (J. Mol. Struct.)* 617 (2002) 9–15. [https://doi.org/10.1016/S0166-1280\(02\)00395-0](https://doi.org/10.1016/S0166-1280(02)00395-0).
- [92] D. Y. Afanas'yev, A. V. Prosyaniy, Why *N*-methyleneformamide, CH<sub>2</sub>=N–CHO, prefers *gauche*-conformation? A DFT/NBO study, *Theochem (J. Mol. Struct.)* 730 (2005) 45–49. <https://doi.org/10.1016/j.theochem.2005.05.001>.
- [93] a) P. Schilke, D. J. Benford, T. R. Hunter, D. C. Lis, T. G. Phillips, A line survey of Orion-KL from 607 to 725 GHz, *Astrophys. J.* 132 (2001) 281–364. <https://doi.org/10.1086/318951>; b) P. Schilke, C. Comito, S. Thorwirth, First detection of vibrationally excited HNC in space, *Astrophys. J.* 582 (2003) L101–L104. <https://doi.org/10.1086/367628>.