

Glycolaldehyde formation mediated by interstellar amorphous ice: a computational study

M. A. M. Paiva¹,^{*} S. Pilling,² E. Mendoza³, B. R. L. Galvão⁴,^{*} and H. A. De Abreu¹,^{*}

¹*Departamento de Química, ICEx, Universidade Federal de Minas Gerais, Belo Horizonte-MG, 31270-901, Brazil*

²*Instituto de Pesquisa e Desenvolvimento (IP&D), Universidade do Vale do Paraíba (UNIVAP), São José dos Campos 12244-000, São Paulo, Brazil*

³*Departamento de Ciencias Integradas y Centro de Estudios Avanzados en Física, Matemáticas y Computación, Universidad de Huelva, 21071 Huelva, Spain*

⁴*Departamento de Química, Centro Federal de Educação Tecnológica de Minas Gerais, 30421-169, Belo Horizonte-MG, Brazil*

Accepted 2022 December 10. Received 2022 November 23; in original form 2022 August 1

ABSTRACT

Glycolaldehyde (HOCH₂CHO) is the most straightforward sugar detected in the Interstellar Medium (ISM) and participates in the formation pathways of molecules fundamental to life, red such as ribose and derivatives. Although detected in several regions of the ISM, its formation route is still debated and its abundance cannot be explained only by reactions in the gas phase. This work explores a new gas-phase formation mechanism for glycolaldehyde and compares the energy barrier reduction when the same route happens on the surface of amorphous ices. The first step of the mechanism involves the formation of a carbon–carbon bond between formaldehyde (H₂CO) and the formyl radical (HCO), with an energy barrier of 27 kJ mol⁻¹ (gas-phase). The second step consists of barrierless hydrogen addition. Density functional calculations under periodic boundary conditions were applied to study this reaction path on 10 different amorphous ice surfaces through an Eley–Rideal type mechanism. It was found that the energy barrier is reduced on average by 49 per cent, leading in some cases to a 100 per cent reduction. The calculated adsorption energy of glycolaldehyde suggests that it can be promptly desorbed to the gas phase after its formation. This work, thus contributes to explaining the detected relative abundances of glycolaldehyde and opens a new methodological framework for studying the formation routes for Complex Organic Molecules (COMs) in interstellar icy grains.

Key words: astrochemistry – molecular processes – ISM: molecules.

1 INTRODUCTION

Among the molecules detected in the interstellar medium (ISM), simple and complex organic molecules, SOM and COMs,¹ respectively, such as formaldehyde (H₂CO), formamide (CH₃NO), acetaldehyde (CH₃CHO), ethene (H₂CCH₂), glycolaldehyde (HOCH₂CHO), methyl and ethyl formate (HCOOCH₃, HCOOCH₂CH₃), have attracted much attention (Herbst & Van Dishoeck 2009; López-Sepulcre et al. 2015; Zamirri et al. 2019). In space, COMs have been found in interstellar clouds, hot corinos, hot molecular cores, molecular outflows, circumstellar envelopes around evolved stars, and in low and high mass star-forming regions in general (Coletta et al. 2020; Nazari et al. 2021). Detections of COMs from young stellar objects and protoplanetary discs suggest the participation of these molecules in the composition of comets and planets in formation, which increase the importance of COMs as the building blocks of the first organic molecules in primitive planets (Walsh et al. 2014; Fedoseev et al. 2015; Butscher et al. 2015).

The study of COMs has been an emerging topic in astrochemistry because of their potential role in prebiotic chemistry, particularly in the formation of amino acids and other fundamental species for the emergence of life (Bulak et al. 2021). The glycolaldehyde molecule is one of these molecules directly linked to the formation of complex sugars, like ribose and derivatives, essentials to genetic material throughout formose reaction. (Woods et al. 2013, 2012; Banfalvi 2021). From astronomical observations in the sub-mm domain, Hollis, Lovas & Jewell (2000) and Halfen et al. (2006) carried out a comprehensive study of glycolaldehyde in Sagittarius B2 discussing the role of the formose reaction to produce glycolaldehyde from formaldehyde. Recent theoretical and experimental works have investigated again the role of the formose reaction, also known as the Butlerov reaction, under interstellar conditions. Ahmad et al. (2020) studied the role of interstellar H₂CO to produce the C₂H₄O₂ isomers: glycolaldehyde, methyl formate, and acetic acid (CH₃COOH). Layssac et al. (2020) studied experimentally the chemistry of formaldehyde through formose-like reactions. They found that VUV processed interstellar ice analogues containing H₂CO can produce sugar-related molecules.

Glycolaldehyde was first detected in space in molecular clouds within Sagittarius B2(N) (Hollis et al. 2000), and later on it has been observed in several other different regions, such as the hot-molecular core G31.41+0.31 (Beltrán et al. 2008), the Class 0 protostellar binary IRAS 16293-2422 (Jørgensen et al. 2012), solar-type protostar NGC 1333 IRAS2A (Coutens et al. 2015), multiple

* E-mail: mamp@ufmg.br (MAMP); brenogalvao@gmail.com (BRLG); heitorabreu@ufmg.br (HADA)

¹As discussed in Herbst & van Dishoeck (2009), the adjective complex here stands for interstellar molecules containing an organic functional group, as the simple ones, but with more than 5 atoms.

sources of Perseus molecular cloud (De Simone et al. 2017), and in shock region L1157-B1 (Lefloch et al. 2017). Glycolaldehyde also has been detected in comets like Lovejoy (Biver et al. 2015) and 67P/Churyumov–Gerasimenko (Goesmann et al. 2015). Regarding bioessential sugars, Furukawa et al. (2019) detected sugar-related compounds in three carbonaceous chondrites demonstrating evidence of extraterrestrial ribose in primitive meteorites. These observations exemplify how widespread the glycolaldehyde molecule is in the ISM environment. Despite its importance, its chemical formation routes in space are still not well-understood (Woods et al. 2012, 2013; Vazart et al. 2018; Coutens et al. 2018).

The other two isomers of glycolaldehyde, the methyl formate and acetic acid have also been detected in several astrophysical environments (Mehring et al. 1997; Remijan et al. 2003, 2005; Shiao et al. 2010) being all involved in the synthesis of key biomolecules, such as glycine ($\text{NH}_2\text{CH}_2\text{COOH}$), the simplest amino acid (Sorrell 2001). Methyl formate is the most abundant of the three isomers detected in space, glycolaldehyde being the less abundant one. Indeed, among the three isomers, methyl formate was the first ever observed outside the Galaxy. Sewilo et al. (2018) reported its first extragalactic detection, along with dimethyl ether in the N113 star-forming region in the Large Magellanic Cloud. As discussed by Rachid, Faquine & Pilling (2017), such high abundances of methyl formate could be related to its high resistance to ionizing radiation field when in solid phase compared with other isomers. However, this species in the gas phase is highly sensitive to ionizing radiation (Fantuzzi et al. 2011). Both authors agreed that besides destruction by radiation, an explanation considering different efficiencies in both production and desorption routes to the gas phase might be invoked in an attempt to explain such observed abundances among isomers. According to Burke et al. (2015), the desorption of methyl formate from ice grain mantles occurs at lower temperatures than glycolaldehyde and acetic acid, and at shorter time-scale and therefore, the temperature of a region might be an important constrain for the abundances of these isomers (expected to be mainly produced in ices) to be detected in gas-phase.

In contrast, formaldehyde is a simple molecule found in many regions of space where glycolaldehyde is detected (Jørgensen et al. 2012; Leroux, Guillemin & Krim 2020). By modelling emission signals from the ^{18}O isotope transitions of the Sagittarius B2(N) region, an abundance of formaldehyde to glycolaldehyde is estimated in the ratio between 42 and 56 (Jørgensen et al. 2012). This abundance estimate provides the basis for a possible route of formation of glycolaldehyde from formaldehyde. However, the simple reaction between two formaldehyde molecules does not seem to be favourable even in the solid phase, as shown by Leroux et al. (2020), who concluded that the production of glyoxal followed by fragmentation would occur instead of glycolaldehyde production.

Shortly after the first detection of glycolaldehyde in Sagittarius B2(N), Sorrell (2001) proposed through theoretical astrophysical models that COMs such as glycolaldehyde could be formed in the bulk of icy grain mantles with the presence of radicals created by ultraviolet radiation. In this study, it was postulated that reactions between two formyl radical molecules and between formyl radical and methanol could lead to the formation of a glycolaldehyde molecule.

Through theoretical calculations, Jalbout (2007) proposes that gas-phase reactions between two molecules of formaldehyde with subsequent shock with H_3^+ have the potential for the final formation of glycolaldehyde and methyl formate. However, the chemical reaction proposed in this work leads to very slow reaction kinetics to produce the abundance of glycolaldehyde detected in the ISM.

Skouteris et al. (2018) proposed, through computational calculations, reaction pathways for the formation of glycolaldehyde through radical reactions between formaldehyde and hydroxymethyl radical in gas phase. Nevertheless, the reaction is promising with only two elementary steps, it comes up against an energy barrier that makes it clear that glycolaldehyde is not favourable for its formation in the gas phase, unless it develops through ethanol radical precursors.

In a nutshell, astrochemical models indicate that gas-phase routes would be inefficient for the formation of glycolaldehyde, and grain-surface formation are preferable candidates. Specifically, the mechanism with a formyl radical and methanol molecules with five reaction steps initially proposed by Sorrell (2001) is the best candidate for matching the observational estimates in low temperatures. However, mechanisms in two steps of two-body reactions involving HCO, H_2CO , and H look more feasible from a chemistry perspective Sorrell (2001), Beltrán et al. (2008), Bennett & Kaiser (2007), Woods et al. (2013).

The Bennett & Kaiser (2007) work shows that a hydrogen atom could add to a CO molecule, generating the formyl radical (HCO) for *a posteriori* reaction to generate the glycolaldehyde molecules and other isomers. In addition, the study developed by Pantaleone et al. (2020) shows that the HCO radical produced *in situ* by hydrogenation of CO remains adsorbed at the surface of water ice under ISM conditions. This radical will then be available for further collisions from incoming gas-phase molecules, such as formaldehyde.

Nevertheless, through astrochemical models and theoretical calculations Woods et al. (2013) showed that the recombination of two formyl radical molecules is not an efficient initial path for glycolaldehyde formation. In the same study, it was proposed that a surface reaction with an energy barrier of less than 100 K would be sufficient to match the observed abundance of glycolaldehyde in the ISM. Chuang et al. (2016) studied the formation of glycolaldehyde, ethylene glycol, and methyl formate in mixed CO, H_2CO , and CH_3OH ices. They highlighted the importance of radicals as HCO in the formation of COMs. A more recent study using simulations with a gas-grain chemical code for astrochemical modelling came to the same conclusions for recombining two formyl radical molecules (Coutens et al. 2018).

In this work, we perform computational predictions on the possibility of glycolaldehyde formation via reactions between formyl and formaldehyde, which are believed to be important precursors for this molecule (Beltrán et al. 2008; Woods et al. 2012). This is carried out both in the gas phase and in the presence of amorphous water ice using a new model.

2 COMPUTATIONAL METHODS

To investigate the influence of the ice in catalysing the reaction, the mechanism was initially explored in the gas phase and improved using benchmark calculations. Subsequently, we propose a model for the surface of an amorphous ice, and re-evaluate the reaction supported on it in several different ways.

2.1 Gas phase

Geometry optimizations and vibrational calculations were performed using density functional theory (DFT) with exchange–correlation functional M06-2x and aug-cc-pVQZ (AVQZ) basis set (Zhao & Truhlar 2008; Peterson, Adler & Werner 2008). In addition, Intrinsic Reaction Coordinate (IRC) (Maeda et al. 2015; Fukui 1970, 1981) and Nudged Elastic Band (NEB) (Henkelman, Uberuaga & Jónsson 2000) calculations ensured the correct link between the molecules

Table 1. Influence of decrease the energy barrier (EB) of the transition state by ice. Eley–Rideal reaction type starting with H₂CO or HCO molecules adsorbed in the surface (s).

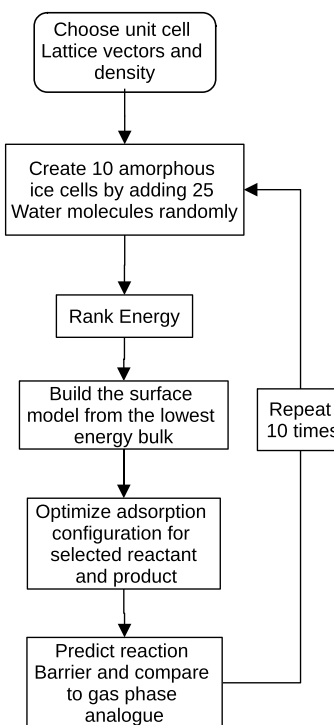
#Surface	Size of EB (kJ mol ⁻¹)		Decrease of EB (%)	
	s-H ₂ CO	s-HCO	s-H ₂ CO	s-HCO
1	12	0	14	100
2	13	15	5	0
3	14	30	1	0
4	0	17	100	0
5	10	18	24	0
6	0	15	100	0
7	4	14	68	0
8	15	15	0	0
9	1	6	91	52
10	2	27	83	0
Average	7	16	49	15

and the representation of the isomer with the lowest energy. To further improve the accuracy, single point energies were calculated at the M06-2X geometry using the explicitly correlated coupled cluster method with singles and doubles excitation with the third one obtained by perturbative method, with resolution of identity correction (CCSD(T)-F12/RI), using the cc-pVQZ-F12 (VQZ-F12) basis set (Knizia, Adler & Werner 2009). The T1 diagnostic of the various structures were evaluated and are presented in Table 1 of the Supplementary Material. The calculations are performed in the lowest spin state of each molecule. We have employed the ORCA program package 5.0.3 version in all gas phase calculations (Neese 2012, 2018; Neese et al. 2020).

2.2 Solid phase

The modelling of an amorphous ice surface is a non-trivial task, and there are several possible ways to tackle this problem. One of these is to use clusters of water molecules in the gas phase to mimic the disorganized pattern of a real grain (Al-Halabi et al. 2004; Ferrero et al. 2020). Alternatively, crystalline unit cells have also been used as an approximation to the amorphous ones (Karssemeijer, de Wijs & Cuppen 2014; Perrero et al. 2021; Molpeceres et al. 2019). Fewer works have employed amorphous unit cells for periodic calculations, which applied molecular dynamics on crystalline configurations to generate the random pattern (Andersson et al. 2006; Andersson & Van Dishoeck 2008; Oberg 2016).

In this work, we propose a faster and yet efficient approach to build representative amorphous unit cells. The main steps of this approach are summarized in Fig. 1. With a Fortran routine, we create a tetragonal box with dimensions of 10×10×7 Å. To match the mean density of ISM amorphous ice, 25 water molecules must be added inside such unit cell. The addition of each molecule is done by generating three random numbers, which are used to locate the centre of mass of the molecule at a random position inside the cell. Three other random numbers are employed for rotating the molecule around the three Euler angles. This is performed independently for each new molecule, and the only restriction to accept the position of a new one is that its distance to any previously added molecule must be greater than 1.6 Å. If this condition is not satisfied, the molecule is discarded and a new random position is tried. This process was repeated 10 times creating 10 different models for the amorphous ice. All obtained structures were optimized, and their total energies were ranked. The cells were classified according to their total energy

**Figure 1.** Flow chart of methodology steps used in solid-state phase.

and only the unit cell with the lowest energy of the group of 10 was selected for posterior surface creation.

Surfaces were created by increasing the height of the tetragonal unit cell by 13 Å resulting in a box of dimensions 10×10×20 Å. Another geometry relaxation was performed at the surface. As shown in Fig. 1, this process was repeated 10 times, yielding a diverse set of models for the surface of the amorphous ice grain.

All geometry optimizations were performed using DFT calculations with exchange functional PBE using plane waves (Perdew, Burke & Ernzerhof 1996) with Grimme D3 dispersion correction (Grimme et al. 2010) and the ultrasoft type pseudopotential (Garrity et al. 2014). The plane wave cutoff energy was set to 40 rydberg (Ry) with a mesh of k points set to 3×3×3 in the bulk and 2×2×1 in the surface models. The reactions involving radicals were of the spin-polarized calculation type, in which unpaired electrons in different orbitals are taken into account.

Subsequently, we proceed to study the adsorption and reaction of the molecules in each of the 10 surfaces generated, thus providing a sample of the many ways the reaction may occur in different ices. At each stage of the mechanism proposed, reagents, intermediate, and product molecules are positioned on the 10 surfaces randomly (following the same routine used for the water molecules in ice, producing random Cartesian coordinates near the surface) and structurally optimized. The Eley–Rideal reaction model was used for the mechanism proposed (Hama & Watanabe 2013), starting with one molecule adsorbed and the other approaching from the gas phase. Two initial routes were tested, one with HCO adsorbed while H₂CO approaches from the gas phase and the opposite.

The transition state search is performed using NEB type calculations with 10 images to reproduce the energy barrier expected by the approximation of the HCO and H₂CO molecules. Each image is separated by 0.5 Å from the other. To quantify the amount by which the reaction barrier is reduced due to the presence of the ice, a reference calculation employing the same methodology is

performed removing the water molecules and with a very large unit cell, thus representing the reaction in gas phase at this calculation level. Note that the benchmark calculations described in the previous section cannot be used as a reference to calculate the barrier height reduction here, as the methods for surface calculation are different.

It is important to mention that the energy barriers were not correctly described by the PBE functional. An in-depth analysis of this problem can be found in the Results section and more detailed in the Supplementary Material. It was shown that hybrid functionals are necessary to describe the barrier-height, and to circumvent this difficulty, single-point calculations were performed for each image obtained in the NEB calculations using the B3LYP functional with only the gamma point (Stephens et al. 1994). Therefore, all energy presented for the calculations including the ice surface are at the B3LYP level with PBE-optimized geometry. We have employed the QUANTUM ESPRESSO package 7.0 version in all solid phase calculations. (Giannozzi et al. 2009).

3 RESULTS

In this work, several possible reaction pathways for glycolaldehyde formation were initially screened, but most were deemed as potentially not important. The new and efficient route proposed here occurs in two parts: first a formaldehyde molecule reacts with a formyl radical, giving rise to an intermediate product radical, followed by a hydrogenation process leading to the glycolaldehyde formation, such as:



Reaction (1) is the decisive step and leads to the formation of the carbon–carbon bond. Unlike previous work, here we start with the premise that the HCO and H₂CO molecules are stable enough in the ice surface to react with another molecule (Pantaleone et al. 2020). First we consider HCO adsorbed in the surface and reacting with formaldehyde coming from the gas phase, and later the opposite situation.

Reaction (2) is expected to be fast due to its radical–radical character, and given the abundance of hydrogen atoms in the interstellar medium. The addition of a hydrogen atom promotes the formation of a very stable neutral and closed shell molecule. To understand the catalytic effect of the amorphous ice on this reaction, we first calculate its energy barrier in the gas phase, to later assess if it can be promoted on the surface of a grain.

3.1 Gas phase

The results obtained for gas-phase calculations are illustrated in Fig. 2. It is shown that the chemical reaction exemplified in equation (1) has an activation energy of 27 kJ mol⁻¹, probably due to electron repulsion from the molecular orbitals during the intermolecular approximation. As a result of carbon–carbon bond formation, the intermediate H₃C₂O₂ (I1) is produced, which is the glycolaldehyde species without a hydrogen atom, with relative energy of -6 kJ mol⁻¹ with respect to the reactants, suggesting that the formation of the carbon–carbon bond slightly stabilizes the composite system. The hydrogenation process that occurs in the sequence, as can be seen in Fig. 2, takes place without an energy barrier resulting in the target molecule glycolaldehyde, with relative energy of -455 kJ mol⁻¹. One can note that adding a hydrogen atom leads to a considerable stabilization of the molecule, as this addition

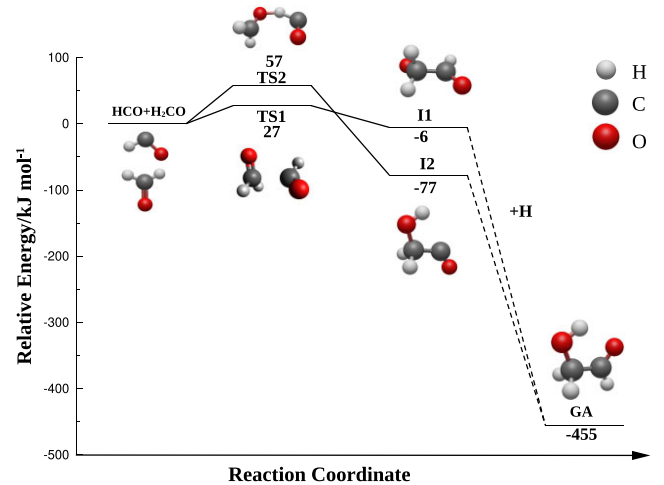


Figure 2. Gas phase mechanisms for Glycolaldehyde (GA) formation. The solid connections corresponds to reaction ((1)) while the dashed line corresponds to reaction ((2)). All energies are relative to the hypothetical H+HCO+H₂CO separated fragments and were obtained at the CCSD(T)-F12/VQZ-F12 level of theory with ZPE correction.

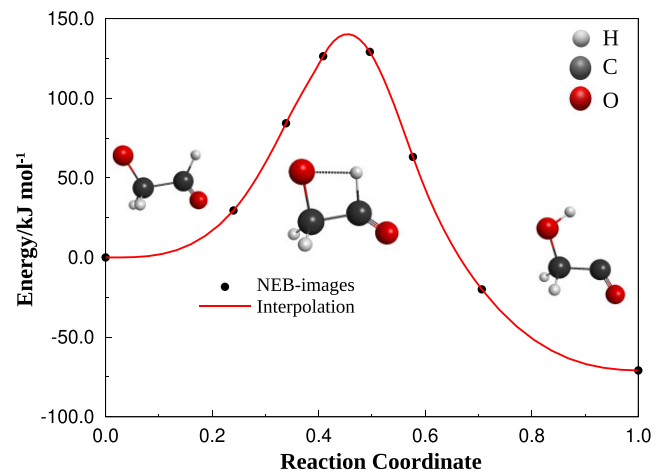


Figure 3. NEB calculation for the possible conversion of intermediate species I1 and I2. Gas phase calculation at M06-2X/aug-cc-pVTZ level.

leads to a non-radical diamagnetic system with 12 electrons in a closed-shell configuration.

An alternative mechanism of glycolaldehyde formation was also found, with a different direction of approximation between reactant molecules, as can be seen also in Fig. 2. In this mechanism, the H atom of the HCO radical migrates to the oxygen end of formaldehyde, leading to a transition state (TS2) that is higher in energy than TS1. In contrast, the intermediate formed in this way (I2) has lower energy than I1. The intermediate formation occurs with a hydrogen atom transfer, followed by the carbon–carbon bond formation. In sequence, the formation of the glycolaldehyde molecule occurs via a subsequent step in which the carbon (originally from the HCO molecule) is hydrogenated, and not the oxygen, as in the previous mechanism.

The interdependence of the mechanisms through TS1 and TS2 was investigated considering the isomerization of intermediates I1 and I2, as shown in Fig. 3. If a possible conversion of the intermediate I1 and I2 occurs, the reaction could go through TS1 to I1 and later isomerize via H-atom migration to I2 before formation of the

glycolaldehyde, since I2 has lower energy relative to the reagents. As shown in Fig. 3, this is not possible because the energy barrier for conversion between the intermediates is too high compared to their formation barriers. This result indicates that only I1 contributes to glycolaldehyde formation.

Complete energy data at DFT, CCSD(T), and ZPE correction are shown in Table 1 of the Supplementary Data.

3.2 Solid phase

3.2.1 Reaction mechanism

Fig. 4 shows one of the model surfaces of amorphous ice created in this work. The random pattern of water molecules with the presence of non-homogeneous cavities (as expected in an ISM ice) can be observed. The complete group of studied surfaces can be seen in Section 3 of the Supplementary Material. We emphasize here the importance of using a diverse sample of possible surfaces, because the structural arrangement of the site where the reaction takes place may vary substantially in a real amorphous ice grain.

The mechanism through TS1 (Fig. 2) was the most successful in the gas phase, and for this reason it was chosen to be the reference for the surface calculations. We have explored this reaction in two ways: one starting with an H₂CO molecule adsorbed at the surface (s-H₂CO) with HCO approaching from the gas phase, while the other starts with HCO adsorbed (s-HCO) and being approached by H₂CO. Examples of each reactant molecule adsorbed on the surface, as well as the intermediate I1 and the product glycolaldehyde can be seen in Fig. 5. Since we have 10 possible models for the amorphous ice surface, and the two reaction possibilities described above, this amounts to a total of 20 potential energy barriers (EB) predictions, which are presented in Table 1.

To investigate whether glycolaldehyde formation is in fact catalysed when carried out on the surface of amorphous ice, we calculate the decrease in EB provoked by the presence of the surface as $EB^{\text{gas}} - EB^{\text{surf}}$. It is not possible to carry out benchmark CCSD(T)-F12 with periodic conditions. The reference gas-phase EB was calculated at the same level of theory as the 20 calculations just described, which employ periodic boundary conditions with large unit cells containing only the HCO and H₂CO molecule at B3LYP level.

We should stress here that, to reproduce the benchmark gas-phase EB employing DFT calculations, it was necessary to use hybrid exchange–correlation functionals, since PBE was not able to predict a gas phase barrier. This is true both for localized basis sets and plane wave calculations. The complete exchange–correlation functional study for this system is shown in Section 1 of Supplementary Data.

The results of each different surface studied are shown in Table 1. As can be seen, the surface of amorphous ice catalyses the reaction of formation of glycolaldehyde at different levels with an average of 49 per cent of decrease of the energetic barrier in the s-H₂CO path. In some cases, the energetic barrier disappeared such as in surfaces number 4 and 6, leading to a reaction without barrier that should occur very rapidly.

It can also be seen in Table 1 that the path starting from the s-HCO structure is less efficient than that with s-H₂CO. In surfaces 1 and 9, there is a complete and partial decrease in the energy barrier. However, on the other surfaces, there is no decrease at all when compared to the gas phase reaction. It is assumed that this behaviour occurs due to the influence of the surface in stabilizing

the unpaired electron of the HCO radical, and thus decreasing its reactivity. In the s-H₂CO path on the other hand, the adsorption is likely to activate the otherwise stable formaldehyde molecule via an electron density transfer mechanism, and thus catalysing the process of carbon–carbon bond formation.

The reaction path over all surfaces starting with H₂CO adsorbed are illustrated in Fig. 6. This figure shows a tendency of energy decrease, initially due to the adsorption of the reactant molecules and followed by the chemical reaction itself, which presents more energetically stable products. This tendency is very similar to the gas phase reaction path, where the intermediate has a relative energy similar to the reactants and the final product presents very low energy. In contrast to the gas phase, in the catalysed route some reaction paths occur barrierlessly. The reaction paths starting with HCO adsorbed are illustrated in Fig. 7. As previously reported, the majority of cases did not lead to a decrease in energy barrier.

It is important to mention that the surface that presented the lower energy in the first step did not necessarily show the lower energy in the next step. In this way, it is challenging to point out a surface with a preferable topography for this chemical reaction to occur.

3.2.2 Adsorption of reactants and intermediates

It is important to know whether the newly formed products will be strongly bound to the surface and will remain on the grain for a long period, or if they are weakly bound and could be released to the gas phase using part of the energy released in the reaction. For this reason, we have also calculated the adsorption energy (E_{adsorp}) of each reactant, intermediate, and products on each of the 10 model surfaces used in this work as

$$E_{\text{adsorp}} = E_{\text{total}} - E_{\text{surface}} - E_{\text{molecule}}, \quad (3)$$

where E_{total} refers to the energy of the adsorbed configuration containing molecule and surface, while E_{surface} and E_{molecule} correspond to the energy of the clean surface and that of the isolated molecule, respectively. The adsorption energies for each chemical species are shown in Table 2.

In Table 2, we would like to highlight the results obtained for the formaldehyde and glycolaldehyde species. The adsorption energy for formaldehyde ranges from -3 to -49 kJ mol⁻¹, and the average value is -31 kJ mol⁻¹. These results are in excellent agreement with experimental results found in the literature. According to an experimental study performed by Noble et al. (2012), the adsorption energy of formaldehyde is estimated to be -27.1 kJ mol⁻¹, showing that our results present an average deviation of 4 kJ mol⁻¹ from the experimental one. It is possible to note that some surfaces, for example, surfaces 2 and 4, deviate in 2 and 3 kJ mol⁻¹ from the experimental result, and surfaces 6, 9, and 10 deviate from 12 kJ mol⁻¹, in the most.

For glycolaldehyde, we compare our results with the experiments performed by Burke et al. (2015) on the adsorption energy of important COMs on amorphous solid water ices using infrared and temperature programmed desorption techniques. This experiment obtained a value of -46 kJ mol⁻¹ for glycolaldehyde, which can be compared with the results of this work shown in the last column of Table 2. It can be seen that on average the deviation from the experiment was 18 kJ mol⁻¹. Several of our model surfaces showed deviations below 7 kJ mol⁻¹. Such deviations are expected for DFT calculations (Peverati & Truhlar 2011; Hammer, Hansen & Nørskov 1999), and suggests that our method for generating the

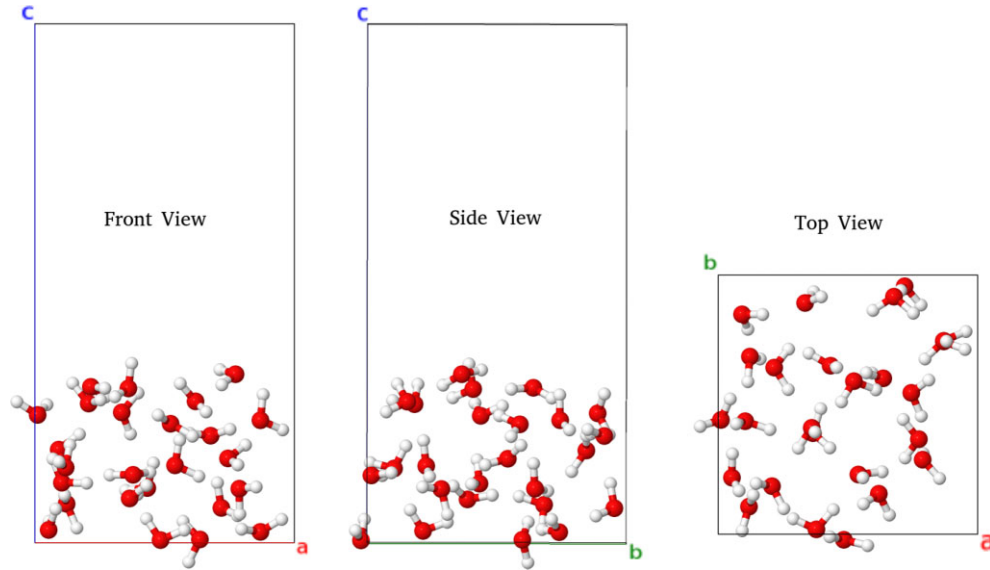


Figure 4. Front, side, and top view from one of the surfaces of amorphous ice created in this work.

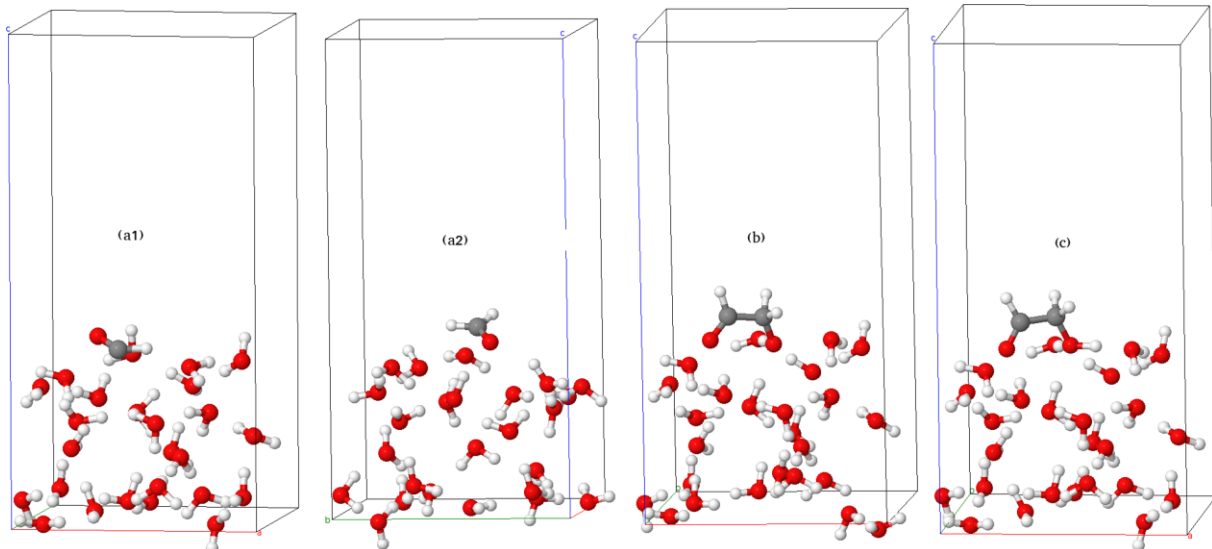


Figure 5. (a1) Adsorption of HCO radical, (a2) adsorption of H_2CO molecule, (b) intermediate molecule II, and (c) glycolaldehyde molecule adsorbed in surface of amorphous ice.

amorphous surfaces does not produce another significant source of error.

The results in Table 2 also suggest that the interaction between the reactants and the surface is strong enough to hold the reactants on the ice until the chemical reaction occurs.

Let us first analyse the mechanism starting with H_2CO adsorbed (*s*- H_2CO), which as argued in the previous section is the preferred one. In this case the intermediate formed is seen to have an adsorption energy of -37 kJ mol^{-1} on average. Recall that the formation of II is only slightly exothermic, and thus it would be formed on the grain with low vibrational energy content, which could also be dissipated to the grain. Given the low temperatures of the ISM, it is likely that this intermediate would stay on the grain long enough for the second step of hydrogenation to occur, finally leading to glycolaldehyde.

As for the less favourable route, which starts with HCO adsorbed and is attacked by a gas-phase H_2CO molecule, we see that the

average adsorption energy of the intermediate is much lower in magnitude. In fact, there were two cases (out of 10) with positive adsorption energy, which would mean that the nascent intermediate would be promptly desorbed. This low magnitude of the adsorption energy gives a further indication that this is not a viable mechanism, as the intermediate could in principle be desorbed right after formation, which would make it less prone for the subsequent step of hydrogen addition.

For the glycolaldehyde molecule, all surfaces presented negative adsorption energies (Table 2), indicating an attractive character to such molecule. However, the exothermicity of the second step of our proposed mechanism is large (on average releasing -751 kJ mol^{-1}), and for all of the 10 cases, the product glycolaldehyde molecule would in fact be formed with a vibrational energy content that is more than enough to overcome its adsorption energy and release it to the gas phase.

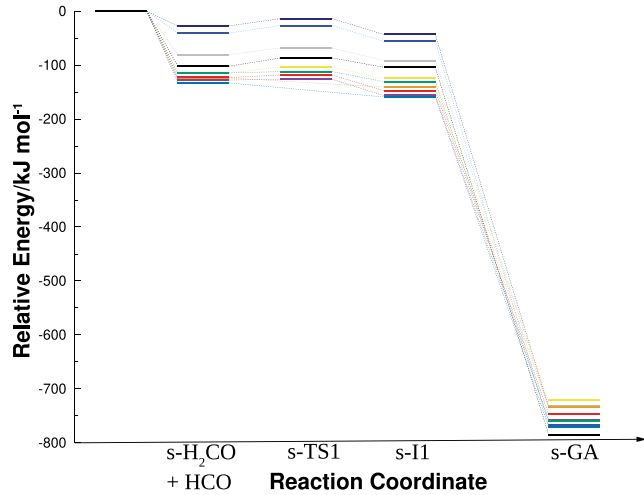


Figure 6. Formation of glycolaldehyde (GA) in the 10 model surfaces, each represented by a different colour. The first step corresponds to the adsorption of the H_2CO molecule at the surface, followed by HCO attack yielding the intermediate II and later the final product. The energy reference in this diagram is the sum of the ice surface energy and the separate reactants. The s- prefix indicates that the molecule is adsorbed on the surface.

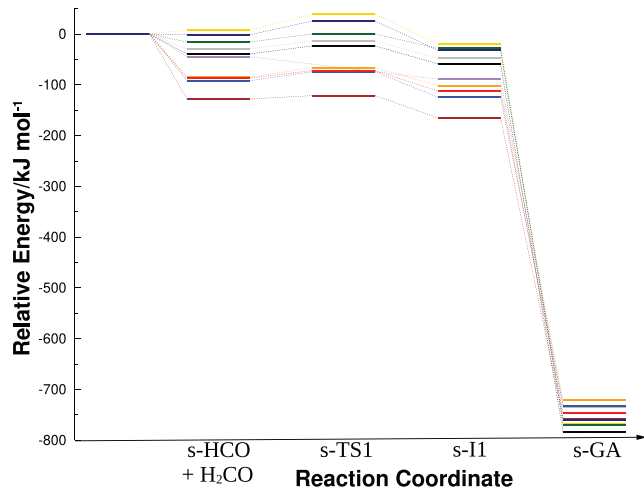


Figure 7. Formation of glycolaldehyde (GA) starting from an adsorbed HCO molecule.

Table 2. Adsorption energy of reactants, intermediates, and product in kJ mol^{-1} . Eley–Rideal reaction type starting with H_2CO and HCO molecule adsorbed in the surface (s).

# Surface	Reagents		Intermediates		Product GA
	s-HCO	s- H_2CO	s-HCO	s- H_2CO	
1	-114	-3	-44	-47	-52
2	-84	-25	-3	-60	-53
3	-117	-3	-1	-38	-64
4	-26	-24	-39	-59	-52
5	-114	-49	-9	-33	-41
6	-84	-35	-43	-41	-75
7	-114	-45	-16	-51	-67
8	-151	-49	38	-7	-100
9	-148	-34	-38	-29	-70
10	-113	-39	68	-4	-73
Average	-106	-31	-8	-37	-64

4 ASTROCHEMICAL IMPLICATIONS

In this study, two sides of a problem were investigated and significant results were obtained that improved our understanding of the formation of the glycolaldehyde molecule.

In the gas phase side, results show that the carbon–carbon bond can be achieved through neutral–radical reactions with a low energetic barrier, enabling this process to occur in many astronomical environments where glycolaldehyde is found. However, this association would require a radioactive energy release (radioactive association) or third-body collision for the stabilization of the intermediate complex. In fact, Woods et al. (2012) shows that the formation of glycolaldehyde in the gas phase is not sufficient to achieve the formation rates reported in the ISM.

At the surface of a grain, the reactants proposed in our mechanism (formyl radical and formaldehyde) can be formed in the ISM in several ways. For example, successive hydrogenation processes of CO can form the HCO and H_2CO , reactions already described in the KIDA data base² (Wakelam et al. 2015). Another process was recently described by Molpeceres et al. (2021), wherein the interstellar amorphous water ice catalyses the two-step $\text{C}+\text{H}_2\text{O}$ reaction to formaldehyde formation. The formaldehyde can be split in HCO radical and hydrogen atom in cosmic rays reactions, being a source of formation of this radical in the gas phase. Thus, our mechanism could be initiated also by a carbon atom on the amorphous ice, leading to our proposed reactants and culminating in the hydrogenation of the intermediate II on the ice surface as shown in Fig. 8.

When we directly compare the relative energies between the mechanism in the gas phase and the same applied to the surface of amorphous ice, as can be seen in Fig. 9, it is notable that the greatest influence occurs in the decrease of the energy barrier in the formation of the glycolaldehyde molecule. The energy barrier for this reaction reduces from a gas-phase value of 17 kJ mol^{-1} (2040 K) to only 7 kJ mol^{-1} (842 K) on average for the surface. This barrier can be easily overcome by local heating induced by collisions or cosmic ray impacts, and the reactivity may be enhanced by quantum tunnelling (Paudel 2014; Bialy 2020). This leads to the question of how many other chemical reactions take place in the ISM in a similar fashion.

In the case of the mechanism studied in this work, the exothermic chemical energy released in the formation of glycolaldehyde can lead to non-thermal surface desorption processes. With this desorption occurring glycolaldehyde would be released to the gas phase quickly after its formation. This hybrid mechanism between gas and solid phases would be the key to understanding the current abundances of glycolaldehyde in the ISM and other more complex COMs. Other reaction routes considering isomers of glycolaldehyde could also help to clarify this issue.

In comparison with other theoretical and experimental works, Layssac et al. (2020) discussed a general formation route for sugars and polyol compounds. They investigated the formation of COMs, such as glycerol and its saturated derivative glycerol, through VUV photolysis of interstellar ice analogues composed of H_2O and H_2CO . They proposed a formation mechanism for glycolaldehyde and polyoxomethylene (a formaldehyde polymer) based on a radical–non-radical reaction between HCO and H_2CO . Ahmad et al. (2020) found that reactions between two formaldehyde molecules exhibits a significant potential barrier, but due to quantum tunnelling these reactions can happen in the ISM. In addition to that, they also

²Kinetic Data base for Astrochemistry <https://kida.astrochem-tools.org/>

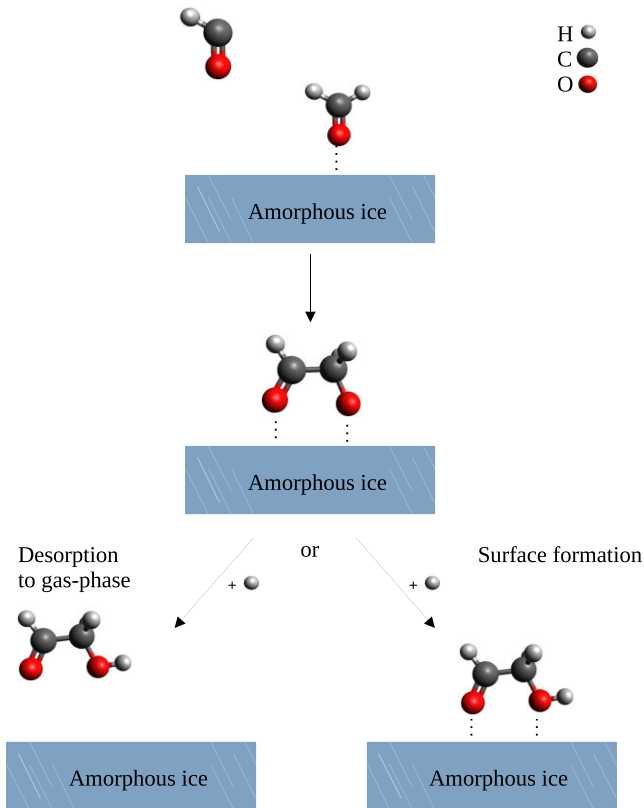


Figure 8. Most likely scenario of glycolaldehyde formation from the results obtained in this work. Dotted line means that the atom is bound to the surface through adsorption.

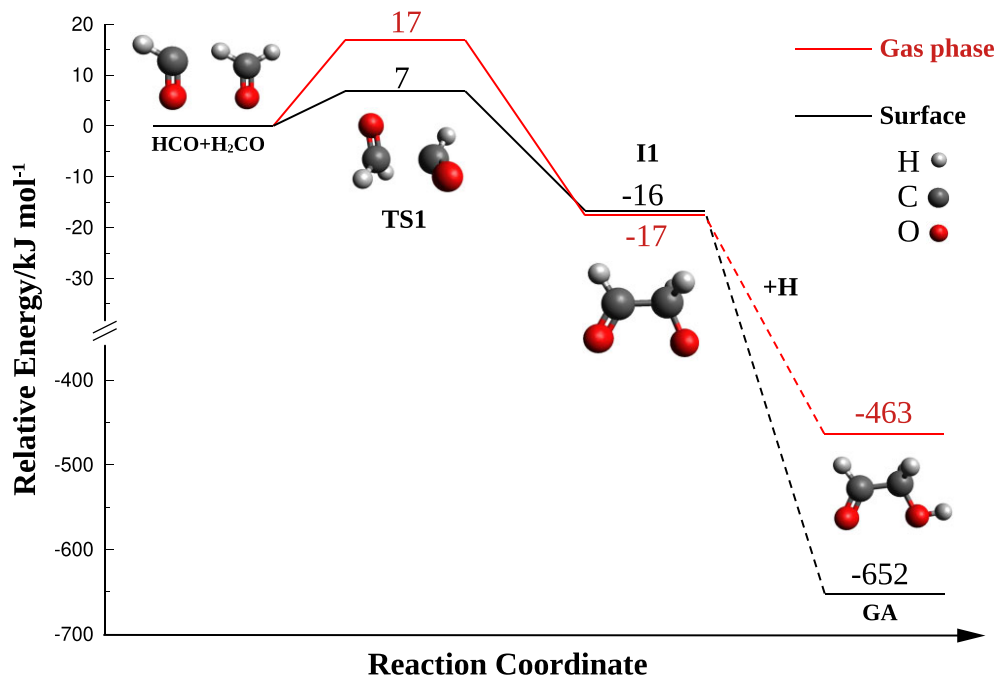


Figure 9. Comparison between the proposed mechanism in the gas phase and on the surface of amorphous ice. The displayed energy of the solid phase is the average of the energies of the 10 surfaces studied. All energies are at DFT level.

found that the chemical reaction is exothermic and capable of producing not only glycolaldehyde but also methyl formate and acetic acid.

In a future study, it would be interesting to investigate the thermodynamic influence of some non-polar amorphous surfaces with the presence of molecules such as CO and CO_2 in different proportions with water molecules. These studies should help quantify how much of the catalysis of chemical reactions would be directly linked to the surface composition.

It was shown in this work that small changes in the surface can lead to considerable changes in the final result. Therefore calculations with more details like explicit molecules, amorphous configuration, and periodic conditions have a significant importance for correct results.

5 CONCLUSIONS

In this work we computationally study the formation of glycolaldehyde from formaldehyde (H_2CO) and the formyl radical (HCO) over a two steps mechanism (reactions (1) and (2)), occurring both in the gas phase and catalysed by a surface (amorphous water ice).

In the gas phase, DFT and CCSD(T) calculations were used to construct the reaction path with the lowest possible energy. In the solid phase, periodic DFT calculations were used on a sample of 10 amorphous ice surfaces in order to study their influence on the energy barrier of an Eley–Rideal mechanism. A consistent decrease in the reaction barrier was obtained in almost all of the different amorphous water surfaces explored with an average reduction of 49 per cent (which amounts to an average barrier of 7 kJ mol^{-1}) and reaching a decrease of 100 per cent in some cases. Even though each grain will have its unique amorphous structure and reactive sites, our results suggest that the catalytic effect is the predominant scenario.

Given that the first step of the reaction is only slightly exothermic, our adsorption energy analysis indicates that the intermediate will remain on the grain for a subsequent hydrogenation step, leading to glycolaldehyde. The large amount of energy released in this step will be greater than the adsorption energy of the product molecule and sufficient to overcome its adsorption energy. Therefore, the mechanism culminates in the release of glycolaldehyde to the gas-phase in a non-thermal desorption process. The results on the energy barrier, as well as the adsorption energies indicate that the Eley–Rideal mechanism with an adsorbed H_2CO being attacked by HCO is much more efficient than the analogous one starting with an adsorbed HCO .

The mechanisms involving the surface (the average value) lead to a product that, in addition to having a lower activation barrier than in the gas phase, is also much more exothermic than the product in the gas phase. Which may suggest that the formation of glycolaldehyde in cold regions can be much more efficient than in gas. The proposed mechanism and the computational framework proposed here may be transferred to the study of several other COMs.

ACKNOWLEDGEMENTS

The authors would like to thank the financial support provided by the Coordenação de Aperfeiçoamento de Pes

soal de Nível Superior (CAPES), Brazil - Finance Code 001, and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), grants number 311402/2021-6 and 311508-2021-9.

REFERENCES

- Ahmad A., Shivani, Misra A., Tandon P., 2020, *Research in A&A*, 20, 014
- Al-Halabi A., Fraser H., Kroes G., Van Dishoeck E., 2004, *A&A*, 422, 777
- Andersson S., Al-Halabi A., Kroes G.-J., van Dishoeck E. F., 2006, *J. Chem. Phys.*, 124, 064715
- Andersson S., Van Dishoeck E., 2008, *A&A*, 491, 907
- Banfalvi G., 2021, *Int. J. Mol. Sci.*, 22, 3857
- Beltrán M. T., Codella C., Viti S., Neri R., Cesaroni R., 2008, *ApJ*, 690, L93
- Bennett C. J., Kaiser R. I., 2007, *ApJ*, 661, L899
- Bialy S., 2020, *Communications Physics*, 3, 1
- Biver N. et al., 2015, *Science Advances*, 1, e1500863
- Bulak M., Paardekooper D., Fedoseev G., Linnartz H., 2021, *A&A*, 647, 82
- Burke D. J., Puletti F., Brown W. A., Woods P. M., Viti S., Slater B., 2015, *MNRAS*, 447, 1444
- Butscher T., Duvernay F., Theule P., Danger G., Carissan Y., Hagebaum-Reignier D., Chiavassa T., 2015, *MNRAS*, 453, 1587
- Chuang K.-J., Fedoseev G., Ioppolo S., van Dishoeck E. F., Linnartz H., 2016, *MNRAS*, 455, 1702
- Coletta A., Fontani F., Rivilla V. M., Mininni C., Colzi L., Sánchez-Monge Á., Beltrán M. T., 2020, *A&A*, 641, 54
- Coutens A., Persson M. V., Jørgensen J. K., Wampfler S. F., Lykke J. M., 2015, *A&A*, 576, 5
- Coutens A., Viti S., Rawlings J., Beltrán M., Holdship J., Jiménez-Serra I., Quénard D., Rivilla V., 2018, *MNRAS*, 475, 2016
- De Simone M. et al., 2017, *A&A*, 599, 121
- Fantuzzi F., Pilling S., Santos A., Baptista L., Rocha A., Boechat-Roberty H., 2011, *MNRAS*, 417, 2631
- Fedoseev G., Cuppen H. M., Ioppolo S., Lamberts T., Linnartz H., 2015, *MNRAS*, 448, 1288
- Ferrero S., Zamirri L., Ceccarelli C., Witzel A., Rimola A., Ugliengo P., 2020, *ApJ*, 904, L11
- Fukui K., 1970, *J. Phys. Chem.*, 74, 4161
- Fukui K., 1981, *Accounts of chemical research*, 14, 363
- Furukawa Y., Chikaraishi Y., Ohkouchi N., Ogawa N. O., Glavin D. P., Dworkin J. P., Abe C., Nakamura T., 2019, *Proceedings of the National Academy of Science*, 116, 24440
- Garrity K. F., Bennett J. W., Rabe K. M., Vanderbilt D., 2014, *Comput. Mater. Sci.*, 81, 446
- Giannozzi P. et al., 2009, *J. Condens. Matter Phys.*, 21, 395502
- Goesmann F. et al., 2015, *Science*, 349, 6247
- Grimme S., Antony J., Ehrlich S., Krieg H., 2010, *J. Chem. Phys.*, 132, 154104
- Halfen D. T., Apponi A. J., Woolf N., Polt R., Ziurys L. M., 2006, *ApJ*, 639, L237
- Hama T., Watanabe N., 2013, *Chemical reviews*, 113, 8783
- Hammer B., Hansen L. B., Nørskov J. K., 1999, *Physical review B*, 59, 7413
- Henkelman G., Uberuaga B. P., Jónsson H., 2000, *J. Chem. Phys.*, 113, 9901
- Herbst E., Van Dishoeck E. F., 2009, *ARA&A*, 47, 427
- Herbst E., van Dishoeck E. F., 2009, *ARA&A*, 47, 427
- Hollis J. M., Lovas F. J., Jewell P. R., 2000, *ApJ*, 540, L107
- Jalbout A., 2007, *Molecular Physics*, 105, 941
- Jørgensen J. K., Favre C., Bisschop S. E., Bourke T. L., Van Dishoeck E. F., Schmalzl M., 2012, *ApJ*, 757, L4
- Karssemeijer L., de Wijs G., Cuppen H., 2014, *PCCP*, 16, 15630
- Knizia G., Adler T. B., Werner H.-J., 2009, *J. Chem. Phys.*, 130, 054104
- Layssac Y., Gutiérrez-Quintanilla A., Chiavassa T., Duvernay F., 2020, *MNRAS*, 496, 5292
- Lefloch B., Ceccarelli C., Codella C., Favre C., Podio L., Vastel C., Viti S., Bachiller R., 2017, *MNRAS*, 469, L73
- Leroux K., Guillemin J.-C., Krim L., 2020, *MNRAS*, 491, 289
- López-Sepulcre A., et al., 2015, *MNRAS*, 449, 2438
- Maeda S., Harabuchi Y., Ono Y., Taketsugu T., Morokuma K., 2015, *Int. J. Quantum Chem.*, 115, 258
- Mehringer D. M., Snyder L. E., Miao Y., Lovas F. J., 1997, *ApJ*, 480, L71
- Molpeceres G., Kastner J., Fedoseev G., Qasim D., Schomig R., Linnartz H., Lamberts T., 2021, *J. Phys. Chem. letters*, 12, 10854
- Molpeceres G., Rimola A., Ceccarelli C., Kästner J., Ugliengo P., Maté B., 2019, *MNRAS*, 482, 5389
- Nazari P., et al., 2021, *A&A*, 650, 150
- Neese F., 2012, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 2, 73
- Neese F., 2018, *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 8, e1327
- Neese F., Wennmohs F., Becker U., Riplinger C., 2020, *J. Chem. Phys.*, 152, 224108
- Noble J., Theule P., Mispelaer F., Duvernay F., Danger G., Congiu E., Dulieu F., Chiavassa T., 2012, *A&A*, 543, 5
- Oberg K. I., 2016, *Chemical Reviews*, 116, 9631
- Pantaleone S., Enrique-Romero J., Ceccarelli C., Ugliengo P., Balucani N., Rimola A., 2020, *ApJ*, 897, L56
- Paudel A., 2014, *Journal of Energy Research and Environmental Technology*, 1, 33
- Perdew J. P., Burke K., Ernzerhof M., 1996, *Phys. Rev. Lett.*, 77, 3865
- Perrero J., Rimola A., Corno M., Ugliengo P., 2021, *International Conference on Computational Science and Its Applications*. Springer, Cham, p. 608
- Peterson K. A., Adler T. B., Werner H.-J., 2008, *J. Chem. Phys.*, 128, 084102
- Peveatti R., Truhlar D. G., 2011, *The Journal of Physical Chemistry Letters*, 2, 2810
- Rachid M. G., Faquine K., Pilling S., 2017, *Planetary and Space Science*, 149, 83
- Remijan A. J., Wyrowski F., Friedel D. N., Meier D. S., Snyder L. E., 2005, *ApJ*, 626, L233
- Remijan A., Snyder L. E., Friedel D. N., Liu S.-Y., Shah R. Y., 2003, *ApJ*, 590, L314
- Sewilo M., et al., 2018, *ApJ*, 853, L19
- Shiao Y.-S. J., Looney L. W., Remijan A. J., Snyder L. E., Friedel D. N., 2010, *ApJ*, 716, L286

- Skouteris D., Balucani N., Ceccarelli C., Vazart F., Puzzarini C., Barone V., Codella C., Lefloch B., 2018, *ApJ*, 854, L135
- Sorrell W. H., 2001, *ApJ*, 555, L129
- Stephens P. J., Devlin F. J., Chabalowski C. F., Frisch M. J., 1994, *J. Phys. Chem.*, 98, 11623
- Vazart F., Skouteris D., Balucani N., Bianchi E., Ceccarelli C., Codella C., Lefloch B., 2018, in *International Conference on Computational Science and Its Applications*. Springer, Cham, p. 730
- Wakelam V. et al., 2015, *ApJS*, 217, 20
- Walsh C., Millar T. J., Nomura H., Herbst E., Widicus Weaver S., Aikawa Y., Laas J. C., Vasyunin A. I., 2014, *A&A*, 563, 33
- Woods P. M., Kelly G., Viti S., Slater B., Brown W. A., Puletti F., Burke D. J., Raza Z., 2012, *ApJ*, 750, L19
- Woods P. M., Slater B., Raza Z., Viti S., Brown W. A., Burke D. J., 2013, *ApJ*, 777, L90
- Zamirri L., Ugliengo P., Ceccarelli C., Rimola A., 2019, *ACS Earth and Space Chemistry*, 3, 1499

- Zhao Y., Truhlar D. G., 2008, *Accounts of chemical research*, 41, 157

SUPPORTING INFORMATION

Supplementary Data are available at [MNRAS](#) online.

SI-Glycolaldehyde_Formation_by_ISM_Ice.pdf

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.