Mem. S.A.It. Vol. 90, 467

Memorie della

The origin of organic chemistry on Earth: endogenous synthesis or exogenous delivery?

F. Vazart[1], N. Balucani[2], D. Skouteris[2], C. Ceccarelli[1], I. Shalayel[3], and Y. Vallée[3]

[1] IPAG, Université de Grenoble Alpes, 414 Rue de la Piscine, 38400 Saint-Martin-d'Hères, France, e-mail: fanny.vazart@univ-grenoble-alpes.fr
[2] Dipartimento di Chimica, Biologia e Biotecnologie, Università degli Studi di Perugia, via Elce di Sotto 8, 06123 Perugia, Italy
[3] Univ. Grenoble Alpes, DCM, CNRS, 301 rue de la Chimie, Saint-Martin-d'Hères, France

Abstract. How could life emerge on an originally inorganic Earth? This is the main question faced in the field of prebiotic chemistry. In order to solve this issue, two main, and not incompatible, hypotheses have been made. An endogenous synthesis one, based on the theory that building blocks of life were synthesized in the atmosphere of the early Earth and/or in primitive oceans, and an exogenous delivery one, according to which those building blocks were synthesized in space and then brought to Earth by comets and asteroids. In this contribution, based on state-of-the-art quantum chemistry computations, both hypotheses will be considered. First, possible gas-phase formation routes in the interstellar medium for two molecules of prebiotic interest will be investigated: glycolaldehyde ($HOCH_2CHO$), the simplest sugar-related molecule, able to ease the formation of more complex sugars and formamide (H_2NCHO), a possible precursor that might be able to link the appearances of metabolism and genetics. And then, thanks to a synergy between experimental and theoretical chemistry, a potential way of forming mercaptoacetaldehyde ($HSCH_2CHO$), a possible prebiotic precursor of cysteine, will be explored in the peculiar conditions of the early Earths ocean.

1. Introduction

How could life emerge on an originally inorganic Earth? This is the main question faced in the field of prebiotic chemistry. Indeed, life as we know is ruled by quite complex organic processes and the involved molecules were surely not present at the surface of the early Earth. In order to solve this issue, two main, and not incompatible, hypotheses have been proposed. The first one is based on an endogenous synthesis theory. According to this theory, the building blocks of life such as amino acids, nucleobases or simple sugars, for instance, are thought to have been synthesized in the atmosphere of the early Earth and/or in the primitive oceans, based on the very simple compounds that are believed to have been present at that time. These compounds include water, carbon dioxide or molecular nitrogen, for example, as shown in Fig. 1. The second hypothesis is based on an exogenous delivery theory, according to which those building blocks of life were synthesized in the interstellar medium (ISM), specifically the Solar Nebula, and then brought to Earth by comets and asteroids.

The endogenous synthesis theory is conceptually supported by the Miller-Urey experiment (Miller 1957) which consisted in the simulation of the assumed conditions of the early

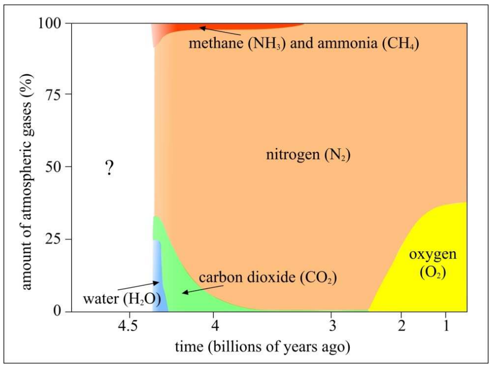

Fig. 1. Composition of the Earth's atmosphere as a function of time.

Earth in laboratory and the detection of several amino acids after a while. Although we now know that the early Earth atmosphere was different from that in the Miller-Urey experiment, the proof of concept remains. This explains why nowadays many chemists are looking at the atmosphere of Titan (the largest Saturn moon) because it is thought to be very similar to that of the early Earth.

As far as the exogenous delivery theory is concerned, the main support comes from the actual detection of potential building blocks of life on comets. For instance, glycine, the simplest amino acid was detected by the STARDUST mission in the Wild2 comet (Elsila et al. 2009) and then in the famous 67P/ChuryumovGerasimenko comet thanks to the ROSETTA mission (Altwegg et al. 2016).

As one can see in Fig. 2, around two hundred compounds have been detected so far in interstellar medium. They are mostly organic species composed of two to twelve atoms and among them, around sixty are interstellar Complex Organic Molecules (iCOMs) because they are carbon-bearing molecules of more than six atoms. These iCOMs are of a great interest for prebiotic chemistry because they might be considered as building blocks of life themselves, or at least as precursors or intermediates in their formations.

The major issues when it comes to reactivity in space are the harsh conditions. Indeed, the very low temperature leads to the absence of energy to help a reaction to occur, which brings to the fact that all considered reactions in space have to be (nearly) barrierless and exothermic. On the other hand, the very low pressure/density makes it already very difficult for two reactants to meet, so there is no possibility of considering three body collisions. To fulfill these requirements, two, again not incompatible, fields of chemistry can be applied. The first one regards surface reactions occurring at the top of dust grains (composed of a carbonaceous or silicate core, surrounded by ices of water, CO_2 or NH_3, for instance), that can help to reduce eventual barriers and ease the encounter between the reactants. However, this type of reaction involves physical processes, such as adsorption and desorption, that can be problematic in some regions where this amount of energy is not available, not to mention the difficulty of the reactions themselves (Rimola et al. 2018). The second field relies on gas-phase reactions, for which the reactants are free in gas-phase, encounter and react. The major issue in this field is that for multi-step reactions, we have to keep in mind that the only available energy comes from the reactants, so all the involved intermediates and transition states need to exhibit energies below that of the reactants.

2. Computational details

All the computations were carried out using Gaussian 09 and 16 suites of programs (Frisch et al. 2016). For all species, geometry optimizations were performed, followed by vibrational computations to make sure that all reactants, intermediates and products were true minima on the potential energy surface, and that all transition states exhibited a single imaginary frequency.

Regarding the support for endogenous synthesis, the B3LYP hybrid functional (Becke 1993) was used, in conjunction to the 6-31+G* double$-\xi$ basis set, including polarization and diffuse functions (Ditchfield et al. 1971). The solvent (water) was simulated thanks to the Polarizable Continuum Model (PCM) (Tomasi et al. 2005).

The calculations respecting the exogenous delivery theory were perfomed em-

2 atoms		3 atoms		4 atoms	5 atoms	6 atoms	7 atoms	8 atoms	9 atoms	10 atoms	11 atoms	12 atoms	>12 atoms
H_2	HD	C_3^+	AlNC	C_3H	C_5	C_5H	C_6H	CH_3C_3N	CH_3C_4H	CH_3C_5N	HC_9N	C_6H_6	$HC_{11}N$
AlF	FeO	C_2H	SiNC	C_3N	C_4H	H_2C_4	CH_2CHCN	$HCOOCH_3$	CH_3CH_2CN	$(CH_3)_2CO$	CH_3C_6H	C_3H_7CN	C_{60}
AlCl	O_2	C_2O	HCP	C_3O	C_4Si	C_2H_4	CH_3C_2H	CH_3COOH	$(CH_3)_2O$	$(CH_2OH)_2$	C_2H_5OCHO	$C_2H_5OCH_3$	C_{70}
C_2	CF^+	C_2S	CCP	C_3S	C_3H_2	CH_3CN	HC_5N	C_7H	CH_3CH_2OH	CH_3CH_2CHO	CH_3OCOCH_3		C_{60}^+
CH	SiH	CH_2	AlOH	C_2H_2	H_2CCN	CH_3NC	CH_3CHO	C_6H_2	HC_7N	CH_3CHCH_2O			
CH^+	PO	HCN	H_2O^+	NH_3	CH_4	CH_3OH	CH_3NH_2	CH_2OHCHO	C_8H				
CN	AlO	HCO	H_2Cl^+	HCCN	HC_3N	CH_3SH	C_2H_4O	HC_6H	CH_3CONH_2				
CO	OH^+	HCO^+	KCN	$HCNH^+$	HC_2NC	HC_3NH^+	H_2CCHOH	CH_2CHCHO	C_8H^-				
CO^+	CN^-	HCS^+	FeCN	HNCO	HCOOH	HC_2CHO	C_6H^-	CH_2CCHCN	C_3H_6				
CP	SH^+	HOC^+	HO_2	HNCS	H_2CNH	NH_2CHO	CH_3NCO	H_2NCH_2CN	CH_3CH_2SH				
SiC	SH	H_2O	TiO_2	$HOCO^+$	H_2C_2O	C_5N		CH_3CHNH					
HCl	HCl	H_2S	C_2N	H_2CO	H_2NCN	HC_4H							
KCl	TiO	HNC	Si_2C	H_2CN	HNC_3	HC_4N							
NH	ArH	HNO		H_2CS	SiH_4	H_2C_3O							
NO	N_2	MgCN		H_3O^+	H_2COH^+	H_2CCNH							
NS	NO^+	MgNC		SiC_3	C_4H^-	C_5N^-							
NaCl		N_2H^+		CH_3	HCOCN	HNCHCN							
OH		N_2O		C_3N^-	HNCNH								
PN		NaCN		PH_3	CH_3O								
SO		OCS		HCNO	NH_4^+								
SO^+		SO_2		HOCN	H_2NCO^+								
SiN		SiC_2		HSCN	$NCCNH^+$								
SiO		CO_2		H_2O_2									
SiS		NH_2		C_3H^+									
CS		H_3^+		HMgNC									
HF		SiCN		HCCO									

Fig. 2. List of detected species in ISM (2016), interstellar Complex Organic Molecules (iCOMs) in grey.

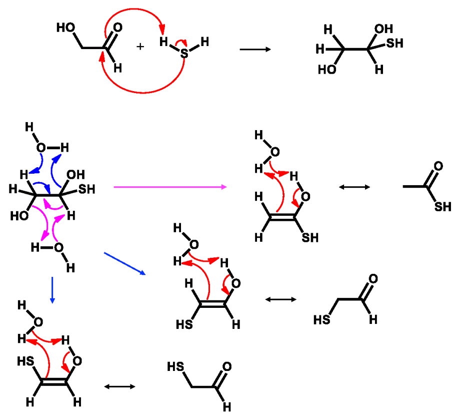

Fig. 3. Considered mechanism for the H_2S + $HOCH_2CHO$ reaction.

ploying the B2PLYP double-hybrid functional (Grimme 2006) in conjunction to the aug-cc-pVTZ triple−ξ basis set (Kendall et al. 1992). Semi-empirical dispersion effects were also included thanks to the GD3BJ model of Grimme (Grimme et al. 2011), leading to the so-called B2PLYP-D3 functional. Once the optimizations were completed, they were followed by energy reevaluations adopting the CCSD(T)/aug-cc-pVTZ level of theory (Scuseria et al. 1988; Grimme 2006).

3. Results and discussion

3.1. Endogenous synthesis

As mentioned before, according to the endogenous synthesis theory, the building blocks of life are thought to have been synthesized in the atmosphere of the early Earth and/or in the primitive oceans. In this context, mercaptoacetaldehyde ($HSCH_2CHO$) is an interesting species because it has been proposed to be an abiotic precursor of cysteine, an amino acid, in aqueous medium (Vallée et al. 2017). It is therefore interesting to wonder if its formation could be possible in the primitive oceans. However, mercaptoacetaldehyde is highly unstable and thus, difficult to manipulate in laboratory experiments. This is the reason why theoretical computations can help to study formation paths towards this molecule.

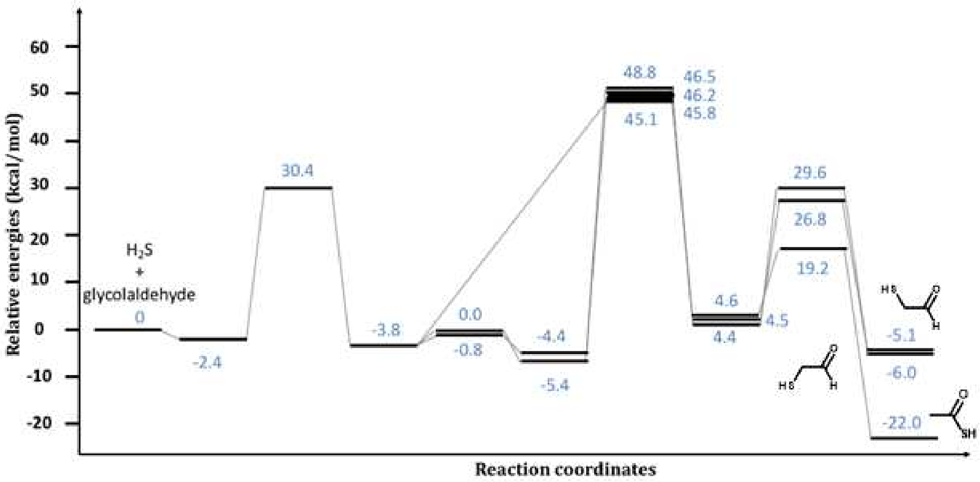

Fig. 4. Energetic profile of the H_2S + glycolaldehyde reaction, at the B3LYP/6-31+G* level of theory. Electronic energies are given in kcal/mol.

3.1.1. Considered reaction and mechanism

The reactants that were considered were glycolaldehyde and H_2S, which are believed to have been present in the primitive oceans. The considered mechanism for this reaction is shown in Fig. 3.

After an addition of the sulfur and hydrogen atoms of H_2S to the CO double bond of glycolaldehyde, one can observe a first intermediate $HOCH_2CHOHSH$. This intermediate can then undergo two types of dehydration, considering it has two different OH moieties, one on each carbon atom, with the help of an explicit water molecule:

1. Represented in pink: the leaving OH moiety is linked to the carbon atom different from the SH moiety, which leads to the CH_2COHSH enolic compound.
2. Represented in blue: the leaving OH moiety is linked to the carbon atom which is linked itself to the SH moiety, which can lead to both E and Z configurations of the enolic version of mercaptoacetaldehyde, depending on the original rotamer of the intermediate.

The three enolic species obtained after dehydration are then able to turn into both rotamers of mercaptoacetaldehyde or thioacetic acid (CH_3COSH), thanks to tautomeric rearrangements, again with the help of an explicit water molecule.

3.1.2. Energetic profile

To understand if this mechanism can be considered, a control of the energetic profile was needed and Fig. 4 shows this profile.

For the sake of consistency, the energy of one or two H_2O molecule(s) has been added to most intermediates because its explicit consideration helps to drop the activation energy of the dehydration and tautomeric steps by ca. 20 and 35 kcal/mol respectively. However, the explicit H_2O molecule in computations does not have a great impact on the intermediates (stabilization by 2 to 5 kcal/mol). As one can notice, the energies regarding activation and transition states are reasonable for the conditions of primitive oceans, so all three isomers should be obtained from this reaction.

3.2. Exogenous delivery

According to the exogenous delivery theory, the building blocks of life, or at least their precursors are thought to have been synthesized in space, and then brought to Earth by comets or asteroids. Here, two different molecules of prebiotic interest were considered.

The first one, formamide (H_2NCHO), can be a precursor in the abiotic amino acid syntheses (Pino et al. 2015) and perhaps also that of nucleobases (Saladino et al. 2012). This makes it a central compound that might be able to link both metabolism (transportation of energy, ruled by proteins) and genetics (transfer of information, ruled by DNA and RNA). It is detected in several regions of massive and solar-type star formation (López-Sepulcre et al. 2015).

The second one, glycolaldehyde ($HOCH_2CHO$), is important in prebiotic chemistry as the simplest sugar-related molecule. Indeed, it can be converted into amino acids or short dipeptides (Weber 1998) and ease the formation of more complex sugars (Cantillo et al. 2012). It has been detected in space, also towards Sgr B2 (Hollis et al. 2000), towards high- and low- mass star forming regions (Beltran et al. 2009; Jorgensen et al. 2012, 2016; Coutens et al. 2015; Taquet et al. 2015; De Simone et al. 2017) and more recently in shocked regions (Lefloch et al. 2017).

3.2.1. Formamide

The reactants NH_2 and formaldehyde (H_2CO) were considered to explain the formation of formamide. The corresponding reaction path is exhibited in Fig. 5 which describes the addition of NH_2 to H_2CO, followed by a hydrogen loss. This path is barrierless, exothermic and all the involved energies are below that of the reactants, which make it viable in ISM.

Formamide is the only product of this reaction but in order to figure out if it is formed in a non-negligible amount, a kinetics study was performed using RRKM theory. This study not only confirmed the formation of formamide but also gave rate constants of formation that, when included in an astrochemical model, provided abundances that matched the observed ones, which reassures the path.

In addition, the three types of deuterated formamide have also been studied using the

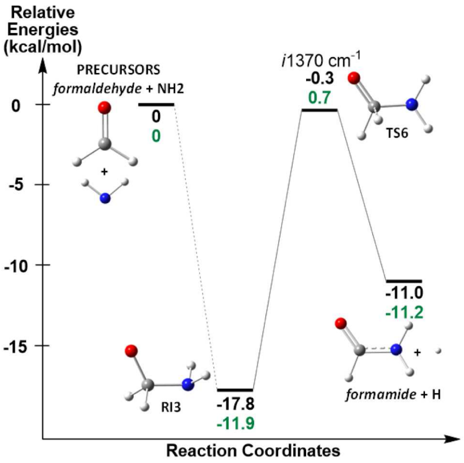

Fig. 5. Proposed reaction path for formamide formation.

same formation path, followed by their detection and quantification in ISM by Coutens et al. (2016). This led to the consideration of four reactions:

1. NH_2+HDCO→ $DCONH_2$+H
2. NH_2+HDCO→ $HCONH_2$+D
3. NHD+H_2CO→ *trans*-HCONHD+H
4. NHD+H_2CO→ *cis*-HCONHD+H

The obtained rate constants are shown in Fig. 6. As one can see, the undeuterated formamide is formed in smaller quantities through reaction (2) than all three others. This is easily understandable because this reaction requires the loss of a heavier deuterium atom compared to a hydrogen one.

When included in an astrochemical model, these rate constants again gave abundances that matched the observed ones. This contributes to a further reliability of the reaction path.

3.2.2. Glycolaldehyde

The formation of glycolaldehyde has already been explored, mostly using grain surface chemistry. The major theories involved radical-radical combination (Garrod et al. 2008) and irradiations by UV or particles (Woods et al. 2012; Maity et al. 2014). However, these studies were exclusively based on laboratory experiments, so we can question their reliability

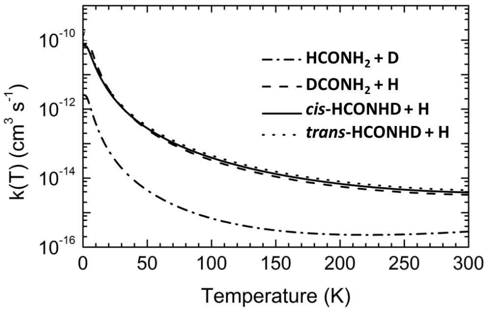

Fig. 6. Canonical rate constants for the four different deuteration reactions (with courtesy of Skouteris et al. 2017).

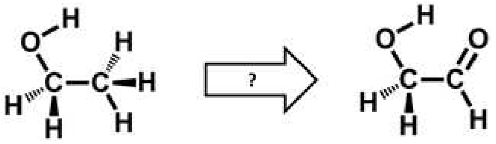

Fig. 7. Ethanol vs. Glycolaldehyde structures.

considering that quantum computations start to challenge the radical-radical combination theory (Rimola et al. 2018) and that the experimental irradiation fluxes are intensely higher than in ISM, by 108 times. Recently, a correlation has been noticed between the abundances of ethanol and glycolaldehyde in gas-phase (Jaber et al. 2014), which led to the assumption that the latter could be formed out of the first. If we take a look at both structures (see Fig. 7), it is noticeable that they are very similar. Indeed, if we remove a hydrogen atom of ethanol, we can obtain the radical $^{\bullet}CH_2CH_2$ OH, which is easily manageable in ISM, in presence of $Cl^{\bullet}$ or $OH^{\bullet}$. Then, if an oxygen atom adds to this radical and a second hydrogen one leaves, it should be possible to get glycolaldehyde.

This is why the studied reaction is $^{\bullet}CH_2CH_2OH$ + O(3P), and adding and oxygen atom to a radical is actually a common pattern since it has been used to explain the formation of methyl formate, a glycolaldehyde isomer, in space (Balucani et al. 2015).

The full reaction path following the addition of $O(^3P)$ to the radical $^{\bullet}CH_2CH_2OH$ is given in Fig. 8. This path is a lot more complicated than the previous one, mostly because the first intermediate is a lot more stable than the reactants, which give a wide range of energy for the further steps to occur. This steps mostly consist in rearrangements and dissociations and this reaction leads to many different products, including formaldehyde, epoxides, enols and glycolaldehyde. Once again, the first addition is barrierless and all the involved species exhibit energies below that of the reactants, which make the path feasible in ISM.

As we have seen, glycolaldehyde is a product of this reaction, but a kinetics study is again needed to figure out which is the main product of the reaction and if glycolaldehyde is formed in a non-negligible amount. This study indeed confirmed the formation of glycolaldehyde, although it is not the major product -formaldehyde is. The rate constants for the formation of glycolaldehyde were then included in an astrochemical model and gave abundances represented in red in Fig. 9. As one can see, there is a good agreement between the observed abundances in several sources and our simulations.

The lines represent different simulations and the dots represent observed sources.

4. Conclusions

As a conclusion, both endogenous synthesis and exogenous delivery can be supported by our theoretical investigations. Mercaptoacetaldehyde should be able to form in the aqueous medium of primitive oceans and then transformed into cysteine, as demonstrated by experiments. On the other hand, gas-phase chemistry is actually able to explain the formation of several iCOMs, but there are still a lot of compounds to investigate. NH_2 and formaldehyde seem to be efficient precursors for formamide in ISM, as supported by deuteration and astrochemical models, while ethanol might be an ancestor of glycolaldehyde, also supported by astrochemical modelling.

Acknowledgements. This project has received funding from the European Research Council (ERC)

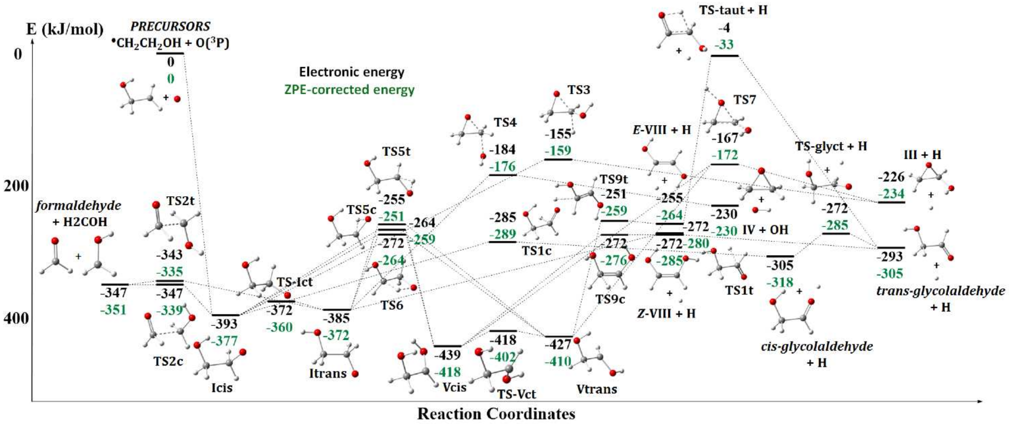

Fig. 8. Proposed reaction path for $CH_2CH_2OH + O(^3P)$.

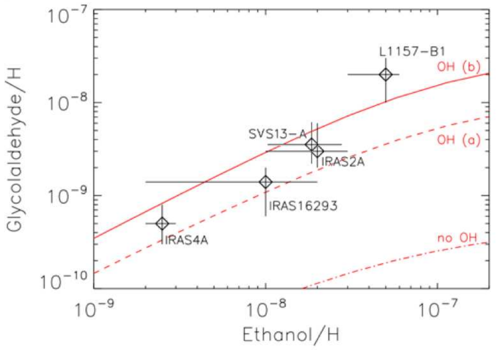

Fig. 9. Abundance of glycolaldehyde as a function of ethanol abundance in the gas phase. The lines represent different simulations and the dots represent observed sources.

under the European Union's Horizon 2020 research and innovation programme, for the Project "The Dawn of Organic Chemistry" (DOC), grant agreement No 741002, and from the French National Research Agency in the framework of the Investissements d'Avenir program (ANR-15-IDEX-02), through the funding of the "Origin of Life" project of the Univ. Grenoble-Alpes.

References

Altwegg, K., et al. 2016, Science Advances, 2, e1600285
Balucani, N., Ceccarelli, C. & Taquet, V. 2015, MNRAS, 449, L16
Becke, A. D. 1993, J. Chem. Phys., 98, 5648
Beltran, M. T., et al. 2009, ApJ, 690, L93
Cantillo, D., et al. 2012, Chem. Eur. J., 18, 8795
Coutens, A., et al. 2015, A&A, 576, 5
Coutens, A., et al 2016, A&A, 590, L6
De Simone, M., et al. 2017, A&A, 599, 121
Ditchfield, R., Hehre, W. J, & Pople, J. A. 1971, J. Chem. Phys., 54, 724
Elsila, J. E., Glavin, D. P & Dworkin, J. P. 2009, Meteoritics & Planetary Science, 44, 1323
Frisch, M. J., Trucks, G. W., Schlegel, H. B., et al. 2016, Gaussian 16 Revision B.01, gaussian Inc. Wallingford CT
Garrod, R. T., Weaver, S. L. W. & Herbst, E. 2008, ApJ, 682, 283
Grimme, S. 2006, J. Chem. Phys., 124, 034108
Grimme, S., Ehrlich, S. & Goerigk, L. 2011, J. Comp. Chem. 32, 1456
Hollis, J. M., Lovas, F. J. & Jewell, P. R. 2000, ApJ, 540, L107
Jaber, A., Ceccarelli, C., Kahane, C. & Caux, E. 2014, ApJ, 791, 29
Jorgensen, J. K., et al. 2012, ApJ, 757, L4
Jorgensen, J. K., et al. 2016, A&A, 595, 117
Kendall, R.A., Dunning, T. H. Jr. ,& Harrison, R. J. 1992, J. Chem. Phys., 96, 6796
Lefloch, B., et al. 2017, MNRAS, 469, L73
López-Sepulcre, A., et al. 2015, MNRAS, 449, 2438
Maity, S., Kaiser, R. & Jones, B. M. 2014, Faraday Discussions, 168, 485

Miller, S. 1957, Science, 117, 528
Pino, S., et al. 2015, Life, 5, 372
Rimola, A., et al. 2018, ACS Earth Space Chem., 2, 720
Saladino, R., et al. 2012, Phys. Life Rev, 9, 84
Scuseria, G. E., Janssen, C. L. & Schaefer, H. F. III 1988, J. Chem. Phys., 89, 7382
Skouteris, D., et al. 2017, MNRAS, 468, L1
Taquet, V., et al. 2015, ApJ, 804, 81
Tomasi, J., Mennucci, B. & Cammi, R. 2005, Chem. Rev., 105, 2999
Vallée, Y., et al. 2017, Int. J. Develop. Biol., 61, 471
Weber, A. L. 1998, Orig. Life Evol. Biosph., 28, 259
Woods, P. M., Kelly, G. & Viti, S. 2012, ApJ, 750, 19