ChemComm

COMMUNICATION

View Article Online
View Journal | View Issue

Cite this: *Chem. Commun.*, 2024, **60**, 11283

Received 24th June 2024,
Accepted 10th September 2024

DOI: 10.1039/d4cc03074e

rsc.li/chemcomm

Unravelling the prebiotic origins of the simplest α-ketoacids in cometary ices: a computational investigation†

Soumya Ranjan Dash,‡[ab] Rinu Pandya,‡[ab] Geetika Singh,‡[a] Himanshu Sharma,‡[ab] Tamal Das,‡[ab] Hritwik Haldar,[c] Srinivas Hotha *[c] and Kumar Vanka *[ab]

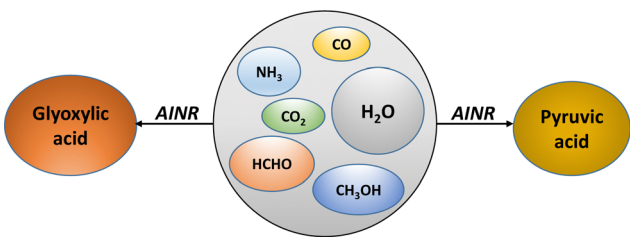
We have employed the *ab initio* nanoreactor (AINR) and DFT calculations to explore how the soft impact of comets entering early earth's dense atmosphere could induce chemical reactions in trapped interstellar ice components, leading to the origin of glyoxylic and pyruvic acids the simplest α-ketoacids essential for prebiotic metabolic cycles.

How life emerged remains a fundamental question within the realm of scientific exploration. Several hypotheses have been proposed to explain its origin, including the RNA world hypothesis and the metabolism-first principle; yet a definitive explanation continues to be elusive.[1] One approach focuses on asking how the building blocks of life arrived/originated on earth: did they form in earth's atmosphere, or on the ocean's surface, or in the depths of hydrothermal vents? Or can their origins be traced even further back to extra-terrestrial media, such as interstellar ice?[2] Meteorites and comets have also been proposed as space shuttles for transporting prebiotic molecules to earth.[3] Numerous investigations have rigorously debated whether these molecules could have survived the meteorites' entry and subsequent impact.[4,5] One theory suggests that the impact of meteorites on earth could induce shock compression of their contents – cometary ices – which, upon expansion to ambient conditions, might produce the building blocks of life.[4]

An alternative hypothesis suggests that the dense atmosphere of the early earth could have induced significant aero-braking on incoming comets, leading to a 'soft impact' scenario.[5] This process would decelerate the comet, reducing kinetic energy upon atmospheric entry, while still inducing compression of cometary ices without subjecting them to extreme shock pressures and temperatures.[5] This moderated impact environment could preserve the interstellar ice components trapped inside it, potentially triggering chemical reactions critical to the emergence of life. However, much of these researches have centered on the abiogenesis of amino acids, aldehydes, hydrocarbons, and the involvement of HCN in these processes.[6]

The advocates of the metabolism-first principle in prebiotic chemistry have focused greatly on the tricarboxylic acid (TCA) and the reverse-TCA (r-TCA) cycle.[7] The two simplest α-ketoacids, glyoxylic (Gly) and pyruvic (Pyr) acid, have been central not only to the prebiotic r-TCA cycle but also to the hydroketoglutarate-malonate cycle.[8] Moreover, investigations into the entirety of known metabolic pathways suggest a theoretical framework where these two α-ketoacids act as central junctions.[8,9] Therefore, it is imperative to explore the prebiotic origins of these two α-ketoacids (Scheme 1).

Despite a few attempts,[10] there is a notable lack of investigations into the non-radiative thermal processes involving the components of interstellar ices, which have been proposed to be reservoirs of essential prebiotic molecules.[11] These processes can occur during the soft impact induced compression of cometary ices as they enter earth's dense atmosphere.[5] One possible reason for this lack is the difficulty in replicating such conditions experimentally. This makes a computational approach to address this problem an attractive alternative.



Scheme 1 Components of interstellar ice taken in the AINR simulations.

[a] *Physical and Materials Chemistry Division, CSIR-National Chemical Laboratory, Pune 411008, India. E-mail: k.vanka@ncl.res.in*
[b] *Academy of Scientific and Innovative Research (AcSIR), Ghaziabad 201002, India*
[c] *Department of Chemistry, Indian Institute of Science Education and Research, Pune 411008, India. E-mail: s.hotha@iiserpune.ac.in*

† Electronic supplementary information (ESI) available. See DOI: https://doi.org/10.1039/d4cc03074e

‡ These authors contributed equally.

Here, we have employed the *ab initio* nanoreactor (AINR) in order to study how the soft impact-induced compression of interstellar ice components inside the comets could have led to the origin of Gly and Pyr through non-radiative, thermal pathways. The recently developed AINR method by the group of Martinez and co-workers has been employed in a variety of chemical systems to provide insights into reaction pathways and mechanisms without the need for experimental input.[12] These systems include prebiotic systems such as the Urey-Miller system, as well as HCN-water systems that could lead to the building blocks of life.[12] All the pathways obtained from the AINR (at the HF/3-21G level of theory, see ESI† for details) have been further investigated by conducting a comprehensive quantum chemical analysis using density functional theory (DFT) at the B3LYP-D3/6-311++G(d,p) (PCM = water) level of theory. As detailed below, our findings unveil interesting pathways for the formation of the two simplest α-ketoacids, as well as other prebiotically relevant molecules.

In the AINR simulations, a nearly homogenously mixed distribution of interstellar ice components was utilized as initial reactants. Initially, we started with a composition inspired by the research of Greenberg and colleagues,[4] featuring an ice mixture with a molar ratio of $H_2O:CH_3OH:NH_3:CO:CO_2$ of 2 : 1 : 1 : 1 : 1. During the AINR simulation (Fig. 1), we observed the formation of Gly and Pyr at 22 and 65 ps, respectively.

As illustrated in Fig. 2, the α-ketoacids are formed through a series of elementary steps. Initially, the carbene intermediate **8** transforms into intermediate **9** through a reaction that is highly exergonic (ΔG = −104.8 kcal mol^{-1}), due to the inherent instability of carbenes. Intermediate **9** then tautomerizes in the presence of water, acting as a catalyst, to form glycolaldehyde **10**, a key intermediate in the making of the target molecules. Additionally, we observed the formation of hydrogen peroxide, which facilitated the oxidation of glycolaldehyde **10** to intermediate **14**. This oxidation process continued through a two-step mechanism involving the formation of the peroxide intermediate **15** *via* a six-membered transition state with an energy barrier of 30.1 kcal mol^{-1}, ultimately resulting in the formation of Gly. The final product **16**, formed *via* a water-assisted six-membered transition state with an energy barrier of 31.1 kcal mol^{-1}. In the pathway leading to the formation of Pyr **13**, intermediate **5** reacts with glycolaldehyde to form intermediate **11**. This intermediate undergoes dehydration to yield the enol form of Pyr **12**, with an energy barrier of 46.5 kcal mol^{-1}. The enol form subsequently tautomerizes to produce the final product, Pyr **13**, with an energy barrier of 35.8 kcal mol^{-1}. It must be noted that most of these transition states are mediated by a proton-shuttling water catalyst. We note here that the proton-shuttling behaviour of water was not observed during the AINR simulations. However, this was included as a post-AINR refinement to the calculations. This is because the prominence of water among the interstellar ice components makes it a suitable catalyst to facilitate these reactions, and so it is likely that water would have mediated as a proton-shuttle catalyst in such interstellar ice systems.

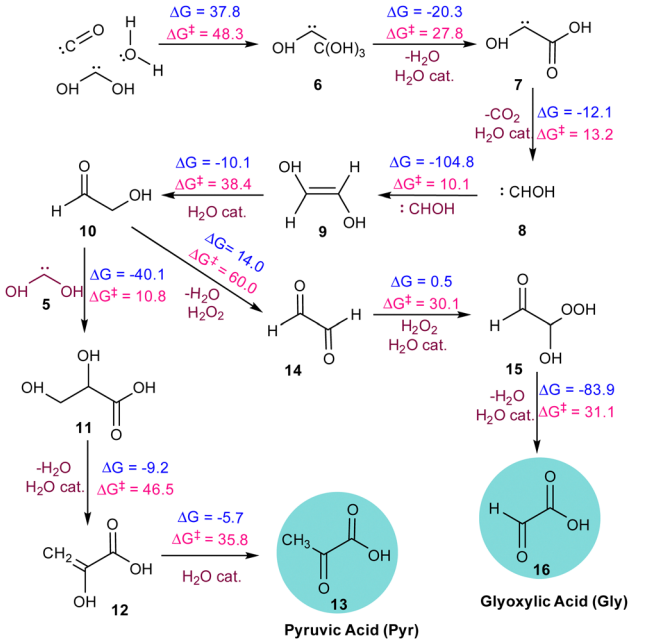


Fig. 2 The sequence of elementary reaction steps derived from the AINR. Molecules labelled "cat.", shown in brown, participate catalytically as proton shuttles. All the values are in kcal mol^{-1}.

Despite the prominence of H_2O, CO, CO_2, CH_3OH, and NH_3 in interstellar ice, there remains a lack of consensus on their precise ratios.[13] Therefore, we varied the molecular compositions in the simulations, as detailed in Table 1. In simulation AINR4, we included HCHO in the composition, following the report by Allamandola *et al.* that molecular clouds also contain ices composed of HCHO, along with other molecules.[14] However, in other AINR simulations (AINR2-5), only one of the acids was observed. Specifically, in simulations AINR3 and AINR4, glyoxylic acid (Gly) was identified at 29 and 31 ps, respectively. It was observed that in two AINR simulations, instead of Pyr, its

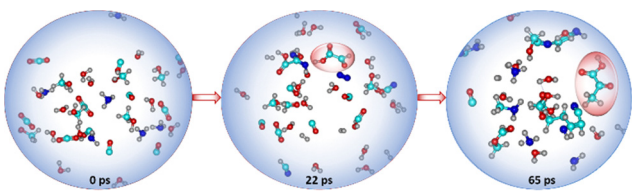


Fig. 1 Snapshots of AINR(1) simulation at different time scales revealing the formation of glyoxylic acid (22 ps) and pyruvic acid (65 ps). Color scheme: oxygen: red, carbon: teal, hydrogen: grey and nitrogen: blue.

Table 1 Different composition of the interstellar ice components taken in the AINR, along with their %content of carbon, nitrogen and oxygen. Time scales (in ps) where Pyr and Gly were formed are listed

AINR	Composition	%C	%O	%N	Gly	Pyr
1	$12H_2O + 6CH_3OH + 6NH_3 + 6CO + 6CO_2$	14.2	28.5	4.7	22	65
2	$6H_2O + 8CH_3OH + 12NH_3 + 7CO$	11.7	16.4	9.3	–	14
3	$6H_2O + 8CH_3OH + 12NH_3 + 7CO + 6CO_2$	14.3	17.8	8.2	29	–
4	$9H_2O + 9HCHO + 10NH_3 + 7CO_2$	12.9	25.8	8.0	31	–
5	$4H_2O + 8CH_3OH + 10NH_3 + 8CO$	13.8	17.2	8.6	–	29
6	$11CH_3OH + 12NH_3 + 8CO$	15.6	15.6	9.8	54	14

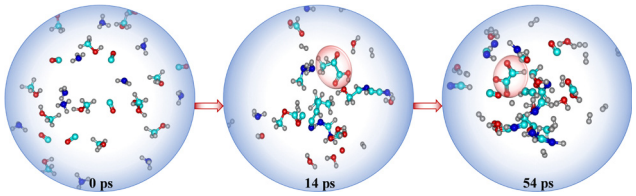


Fig. 3 Snapshots of AINR(6) simulation at different time scales revealing the formation of pyruvic acid (14 ps) and glyoxylic acid (54 ps). Colour scheme: oxygen: red, carbon: teal, hydrogen: grey and nitrogen: blue.

cyanide precursor, CH_3COCN, and the enol isomer of pyruvic acid, were formed.[15] In simulation AINR2, CH_3COCN was detected at 14 ps, which can undergo hydrolysis to form Pyr (see Fig. S2, ESI†). In simulation AINR5, both the enol form of Pyr and CH_3COCN were detected at 26 and 29 ps, respectively. The enol form further tautomerized to yield Pyr. The mechanistic pathways and their corresponding energy profiles are discussed in detail in the ESI.† We observed that in addition to simulation AINR1, simulation AINR6 also showed the formation of both targeted acids, as illustrated in Fig. 3. Gly **16** was detected at 54 ps, following a pathway similar to that in simulation AINR4 (see Fig. S4, ESI†). However, instead of the oxidation of HCHO **3** by H_2O_2 to form formic acid **44**, Gly was formed *via* the singlet carbene **5**. The formic acid **44**, then on further reaction with carbene **5**, provided intermediate **45**, which on dehydration gave Gly **16**, having a barrier of 27.0 kcal mol^{-1}.

Our investigation into the formation of Pyr in simulation AINR6 revealed an alternative pathway involving different intermediates. Initially, ethen-1-ol **21** was formed *via* the combination of two singlet carbenes **8** and **22** in a kinetically favourable reaction ($\Delta G = -144.2$ kcal mol^{-1}). Intermediate **21** then converted to ketene **23** with a barrier of 7.5 kcal mol^{-1}. The hydration of ketene **23** leads to the formation of intermediate **29**, also with a barrier of 38.9 kcal mol^{-1}. Both dehydrogenation and hydration processes were catalysed by water. Subsequently, intermediate **29** reacted with intermediate **19** to produce intermediate **30**, involving a six-membered transition state with a barrier of 45.6 kcal mol^{-1}. Intermediate **30** could dehydrate to form intermediate **31**, encountering a barrier of 26.2 kcal mol^{-1}, which upon further hydrolysis produced pyruvic acid (Fig. 4). Additionally, the hydrolysis of intermediate **30** yielded the diol form of pyruvic acid **35**. It is important to note that even though water was not included as an initial component in this composition (AINR6), it was generated *in situ* during the AINR collisions.

It was interesting to note that higher carbon concentrations (as shown in Table 1) favoured the formation of both glyoxylic and pyruvic acids, as observed in AINR1 and AINR6, with the exception of AINR3, even though the oxygen content was lowered. Additionally, increased nitrogen concentrations appeared to promote the formation of pyruvic acid, likely due to the formation of its cyanide precursor, as seen in AINR2 and AINR5. Furthermore, we have also observed that the pathways for the prebiotic formation of Gly and Pyr often involved transition states that are predominantly five- or six-membered rings, where

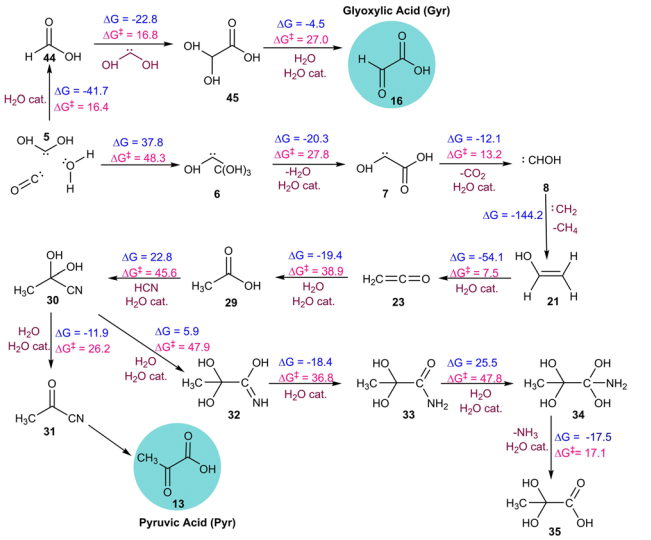


Fig. 4 The sequence of elementary reaction steps derived from the AINR. Molecules labelled "cat.", shown in brown, participate catalytically as proton shuttles. All the values are in kcal mol^{-1}.

water acts as a proton-shuttling catalyst in most of the reaction steps. Notably, dehydrogenation reactions, as observed in simulations AINR1, AINR3, and AINR5 exhibit high energy barriers, around 60.0 kcal mol^{-1} (Fig. 2 and Fig. S3, S5, ESI†). Additionally, reactions involving addition to carbonyl intermediates also present relatively high barriers compared to other reaction types.

Furthermore, our AINR simulations also revealed the formation of various singlet carbenes, which serve as crucial intermediates in the synthesis of the targeted molecules. The emergence of low-valent species such as carbenes (**5**, **6**, **7**, **8**, **22**, and **38**) significantly contributes to the relatively low to moderate energy barriers observed in the mechanistic pathways identified through the AINR approach.[16] Specifically, singlet carbenes such as **8**, **22**, and **38** facilitate the formation of formaldehyde **3**, ethen-1-ol **21**, and acetaldehyde **39**, respectively, *via* kinetically favourable reactions. These intermediates are essential in the subsequent pathways leading to the formation of glyoxylic and pyruvic acids.

Apart from the targeted α-ketoacids, we have also observed the formation of several key intermediates that are relevant in prebiotic chemistry, such as carbamic acid, glycine, isocyanic acid, formamide, glyceraldehyde, and acetylene (Ethyne), among others (see Fig. 5 and Fig. S7, ESI†).[17] In some pathways, these intermediates were found to contribute to the formation of Gly and Pyr, while in others, they form independently through distinct mechanisms.

In summary, we have demonstrated how the soft impact-induced compression experienced by comets entering early Earth's dense atmosphere could have triggered chemical reactions among the interstellar ice components that they carried. By employing a combined approach using the *ab initio* nanoreactor (AINR) and comprehensive quantum chemical static density functional theory calculations, our investigations have revealed non-radiative thermal pathways with accessible barriers for the

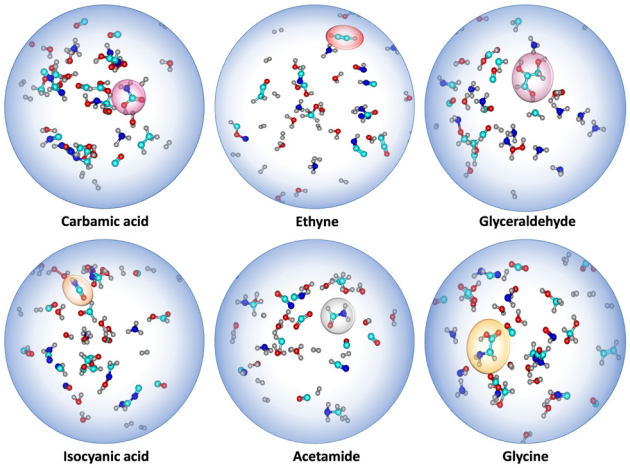


Fig. 5 Snapshots of different AINR simulations showing the formation of carbamic acid, ethyne, glyceraldehyde, isocyanic acid, acetamide and glycine. Colour scheme: oxygen: red, carbon: teal, hydrogen: grey and nitrogen: blue.

formation of not only the simplest α-ketoacids, essential for prebiotic metabolic pathways, but also other prebiotically relevant molecules. Additionally, we have shown how variations in the composition of interstellar ice components altered these reaction pathways. Overall, this study aims to provide a broader understanding of the reactions that could have led to the origin of life on the early earth.

K. V. is thankful for DST-SERB (CRG/2021/003255), GOI for providing financial assistance. K. V. and S. H. acknowledge the support and the resources provided by the 'PARAM Brahma Facility' under the National Supercomputing Mission, Government of India at IISER Pune.

Data availability

The data supporting this article have been included as part of the ESI.†

Conflicts of interest

There are no conflicts to declare.

Notes and references

1 T. Gánti, *The principles of life*, Oxford Univsersity Press, Oxford UK, 2003; F. Dyson, *Origins of life*, Cambridge University Press, Cambridge UK, 1999; D. Sutherland, *Nat. Rev.*, 2017, **1**, 1–7.
2 A. I. Oparin, *Proiskhozhedenie Zhizni*, Mosckovskii Rabochii, Moscow, 1924; Reprinted and translated in: J. D. Bernal, *The Origin of Life*, Weidenfeld and Nicolson, London, 1967; J. B. S. Haldane, *Ration. Ann.*, 1929, **148**, 3–10; K. Kvenvolden, *et al.*, *Nature*, 1970, **228**, 923–926; Y. Furukawa, T. Sekine and M. Oba, *Nat. Geosci.*, 2009, **2**, 62–66; S. V. Singh, V. Jayaram and J. K. Meka, *et al.*, *J. Indian Inst. Sci.*, 2023, **103**, 909–917.
3 J. Oro, *Nature*, 1961, **190**, 389–390; E. Pierazzo, D. A. Kring and H. J. Melosh, *J. Geophys. Res.*, 1998, **103**, 28607–28625; P. J. Thomas, R. D. Hicks, C. F. Chyba and C. P. McKay, *Comets and the Origin and Evolution of Life*, Springer, 2006.
4 N. Goldman, E. Reed and L. Fried, *Nat. Chem.*, 2010, **2**, 949–954; Z. Martins, M. Price and N. Goldman, *Nat. Geosci.*, 2013, **6**, 1045–1049; Y. Umeda, N. Fukunaga and T. Sekine, *J. Biol. Phys.*, 2016, **42**, 177–198; M. P. Kroonblawd, R. K. Lindsey and N. Goldman, *Chem. Sci.*, 2019, **10**, 6091–6098.
5 C. F. Chyba, P. J. Thomas, L. Brookshaw and C. Sagan, *Science*, 1990, **249**, 366–373; C. Chyba and C. Sagan, *Nature*, 1992, **355**, 125–132; J. G. Blank, G. H. Miller, M. J. Ahrens and R. E. Winans, *Origins Life Evol. Biospheres*, 2001, **31**, 15–51; G. Arney, S. D. Domagal-Goldman and V. S. Meadows, *Astrobiology*, 2018, **18**, 311–329.
6 G. Muñoz Caro, U. Meierhenrich and W. Schutte, *et al.*, *Nature*, 2002, **416**, 403–406; S. Panda and A. Anoop, *ACS Earth Space Chem.*, 2024, **8**, 348–360; H. Sandström and M. Rahm, *ACS Earth Space Chem.*, 2021, **5**, 2152–2159; M. Engel and S. Macko, *Nature*, 1997, **389**, 265–268.
7 E. Smith and H. J. Morowitz, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 13168–13173; K. B. Muchowska, S. J. Varma and J. Moran, *Chem. Rev.*, 2020, **120**, 7708–7744.
8 G. Springsteen, J. R. Yerabolu, J. Nelson, C. J. Rhea and R. Krishnamurthy, *Nat. Commun.*, 2018, **9**, 91; K. B. Muchowska, S. J. Varma and J. Moran, *Nature*, 2019, **569**, 104–107; R. T. Stubbs, M. Yadav and R. Krishnamurthy, *Nat. Chem.*, 2020, **12**, 1016–1022; S. A. Rauscher and J. Moran, *Angew. Chem., Int. Ed.*, 2022, **61**, e202212932.
9 J. E. Goldford, H. Hartman, T. F. Smith and D. Segrè, *Cell*, 2017, **168**, 1126–1134; N. Nogal, M. Sanz-Sánchez, S. Vela-Gallego, K. Ruiz-Mirazo and A. de la Escosura, *Chem. Soc. Rev.*, 2023, **52**, 7359–7388.
10 S. J. Varma, K. B. Muchowska and P. Chatelain, *J. Moran, Nat. Ecol. Evol.*, 2018, **2**, 1019–1024; A. K. Eckhardt, A. Bergantini, S. K. Singh, P. R. Schreiner and R. I. Kaiser, *Angew. Chem., Int. Ed.*, 2019, **58**, 5663; A. P. Clay, R. E. Cooke, R. Kumar, M. Yadav, R. Krishnamurthy and G. Springsteen, *Angew. Chem., Int. Ed.*, 2022, **61**, e202112572; G. Cooper, C. Reed, D. Nguyen, M. Carter and Y. Wang, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 14015–14020; F. S. Mohammed, K. Chen, M. Mojica, M. Conley, J. W. Napoline, C. Butch, P. Pollet, R. Krishnamurthy and C. L. Liotta, *Synlett*, 2017, 93–97; M. R. Marín-Yaseli, E. González-Toril, C. Mompeán and M. Ruiz-Bermejo, *Chem. Eur. J.*, 2016, **22**, 12785.
11 C. Puzzarini and S. Alessandrini, *ACS Cent. Sci.*, 2024, **10**, 13–15.
12 S. Seritan, C. Bannwarth and B. S. Fales, *et al.*, *WIREs Comput. Mol. Sci.*, 2021, **11**, e1494; L. P. Wang, A. Titov and R. McGibbon, *et al.*, *Nat. Chem.*, 2014, **6**, 1044–1048; T. Das, S. Ghule and K. Vanka, *ACS Cent. Sci.*, 2019, **5**, 1532–1540.
13 Y. Oba, Y. Takano and H. Naraoka, *Nat. Commun.*, 2019, **10**, 4413.
14 L. J. Allamandola, M. P. Bernstein, S. A. Sandford and R. L. Walker, *Space Sci. Rev.*, 1999, **90**, 219–232.
15 A. Danho, A. Mardyukov and P. R. Schreiner, *Chem. Commun.*, 2024, **60**, 5161–5164; R. Kakkar, P. Pathak and N. P. Radhika, *Org. Biomol. Chem.*, 2006, **45**, 886–895.
16 S. K. Mandal and H. W. Roesky, *Acc. Chem. Res.*, 2012, **45**, 298–307.
17 A. López-Sepulcre, N. Balucani, C. Ceccarelli, C. Codella, F. Dulieu and P. Theulé, *ACS Earth Space Chem.*, 2019, **3**, 2122–2137; R. Breslow, *Tetrahedron Lett.*, 1959, **1**, 22–26; A. Omran, C. Menor-Salvan, G. Springsteen and M. Pasek, *Life*, 2020, **10**, 125; J. H. Marks, J. Wang, B. J. Sun, M. McAnally, A. M. Turner, A. H. H. Chang and R. I. Kaiser, *ACS Cent. Sci.*, 2023, **9**, 2241–2250; E. O. Pentsak, M. S. Murga and V. P. Ananikov, *ACS Earth Space Chem.*, 2024, **8**, 798–856; R. S. Oremland and M. A. Voytek, *Astrobiology*, 2008, **8**, 45–58.

