

On the Plausibility of Pseudosugar Formation in Cometary Ices and Oxygen-rich Tholins

Nieves Lavado · Martín Ávalos · Reyes Babiano · Pedro Cintas · Mark E. Light · José Luis Jiménez · Juan C. Palacios

N. Lavado* · M. Ávalos · R. Babiano* · P. Cintas · J. L. Jiménez · J. C. Palacios
Departamento de Química Orgánica e Inorgánica, Facultad de Ciencias-UEX, Avenida de Elvas s/n, E-06006 Badajoz, Spain
E-mail: nlavador@gmail.com, reyes@unex.es

M. E. Light

Department of Chemistry, University of Southampton, Southampton SO17 1BJ, UK

Abstract We revisit herein the formation and structure of dihydroxy dioxanes, which can be obtained from prebiotically available precursors and can be regarded as primeval sugar surrogates. Previous studies dealing with the heterogeneous composition of interstellar bodies point to the existence of significant amounts of small polyalcohols along with oxygen-containing oligomers. Even though such derivatives did not give rise to nucleosides and oligonucleotides, nor they were incorporated into subsequent metabolic routes, molecular chimeras based on sugar-like species could be opportunistic scaffolds in pre-evolutionary scenarios. We could figure out that pseudosugars, assembled by hemiacetalic bonds from available precursors in both interstellar and terrestrial scenarios, were presumably more abundant than thought. Moreover, these species share some key features with naturally-occurring sugar rings, such as anomeric preferences, coordinating ability, and the prevalent occurrence of racemic compounds.

Keywords Prebiotic chemistry · Sugars · Comets · Tholins · Racemic compounds

Introduction

Unlike amino acids and other biomolecular signatures, which have long been scrutinized in meteorites and astronomical observations through spectroscopic emission lines, carbohydrates have received considerably less attention. Despite their inherent

lability, the putative formation of sugars and sugar-like structures may however be inferred from the large abundance of oxygen-containing molecules in the interstellar medium and bodies, and icy moons (Kaiser 2002; Williams and Viti 2013; Ruiz-Mirazo et al. 2014). It is generally agreed that the origin of carbohydrate structures, pentoses and hexoses in particular, is most likely related to the so-called formose reaction, the autocatalytic self-condensation of formaldehyde occurring in alkaline medium and the presence of some metal ions, which may be triggered under both thermal and radiation conditions (Harman et al. 2013; Gollihar et al. 2014). The overall transformation is a low-yielding, messy reaction leading to mixtures of pentoses and hexoses with different configurations. This extremely complex pathway arises from sequential aldol reactions accompanied by tautomerization and isomerization steps (Benner et al. 2010; Kim et al. 2011; Benner et al. 2012). Moreover, the process appears to be astrochemically relevant as HCHO and glycolaldehyde, the first aldol, are present and abundant in interstellar space and comets (Irvine 1999; Hollis et al. 2000; Meierhenrich 2015). Extracts isolated from Murchison and Murray meteorites contain variable amounts of ketone and polyol structures, such as dihydroxyacetone, glyceric acid, and 4C-6C alditols, among others, although the stereochemical bias of inherently chiral substances, if any, remains unknown (Cooper et al. 2001; Sephton 2001).

The formation of sugar and nucleosides has also been formulated by non-classical pathways that involve the in-situ generation and accumulation of heterocyclic precursors, in particular 2-aminooxazol as hypothesized by Sutherland and coworkers, thus overcoming the limitations of multistep formose cycles (Powner et al. 2009; Patel et al. 2015).

A plausible source of oxygen-containing precursors and heteropolymers should be comets as such organic compounds may be generated by UV photolysis or plasma discharge irradiation of condensed ices rich in $\text{H}_2\text{O}/\text{CO}/\text{CH}_3\text{OH}$ plus $\text{CH}_4/\text{C}_2\text{H}_6$ in rough mimicry of cometary chemistry (Meierhenrich 2015). The resulting ice tholins were analyzed by Sagan and associates through GC-MS, who found a wide range of polyalcohols, including ethylene glycol and glycerol, along with ethers, esters, carbonyls, and organic acids (McDonald et al. 1996). Polyoxymethylenes, generated by polymerization of formaldehyde, could not be detected, although the presence of ether linkages make them more volatile than higher alcohols containing OH groups bound by intermolecular hydrogen bonding. Further condensations of the above-mentioned molecules would also have given rise to more complex structures as conjectured from

UV and IR data. In this context, the term tholin appears to be appropriate as synonymous with complex organic matter, often intractable mixtures (Schwartz 2007). The term tholin, first coined by Sagan and Khare in the late 1970s, describes sticky brownish residues formed by photochemical or electrical discharges of cosmically significant gas mixtures (Sagan and Khare 1979). Accordingly, such residues are usually interpreted within the borderline of planetary simulations leading to aerosol or solid deposits of variable chemical composition, which may not reproduce the actual conditions present in solar moons like Titan (Cable et al. 2012). Obviously these premises do not exclude the prebiotic formation of similar structures in either watery environments or gas-rich atmospheres, even though the chemistry of dust grains and interstellar bodies seems to be largely dominated by surface and solid state reactions (Williams and Viti 2013). Furthermore, as reported recently, meteoritic samples (presumably containing tiny amounts of organics) can also catalyze the formation of numerous prebiotic molecules and analogs by irradiation with high-energy protons of formamide serving as versatile scaffold (Saladino et al. 2015).

Here we describe lab experiments that lead to dioxanes from glycols and carbonyls present in prebiotic scenarios, with solid-state features similar to those observed in oxygen-rich tholins, thus accounting for the putative formation of cyclic polyalcohols. Such oxygenated rings are also a structural motif shared by the simplest sugars, namely glycolaldehyde and glyceraldehyde in the solid state, though they exhibit complex dynamics in solution. The presence of OH functionalities is also responsible for ion coordination and dictates the fate of crystal packing. Finally, this study was motivated by the search of *inherently chiral dioxane structures*, which may help solving the puzzle of homochirogenesis in prebiotic sugar chemistry.

Materials and Methods

Materials All reagents (Sigma-Aldrich) and solvents (Scharlau) were purchased from commercial suppliers and used as received. Distilled water was produced in-house and used for all sample preparations and experiments.

Methods FT-IR spectra were recorded on a Thermo IR3000 spectrophotometer between 4000 and 600 cm⁻¹ using KBr pellets (spectrophotometric grade). NMR spectra were

recorded on a Bruker spectrometer at either 500 or 400 MHz for ^1H nuclei (125 or 100 MHz, respectively, for ^{13}C nuclei) in deuterated solvents. Chemical shifts (δ values) are in parts per million (ppm) relative to tetramethylsilane (TMS, δ 0.00 ppm) as internal standard. Coupling constants (J) are given in Hertz (Hz). Peak identification was facilitated by deuterium exchange (D_2O), double resonance experiments, and DEPT (*distortionless enhancement by polarization transfer*) experiments on decoupled ^{13}C NMR spectra. Low-resolution mass spectra (electrospray technique, positive mode, 0.1 M formic acid) were performed by the *Servicio de Apoyo a la Investigación* (SAIUEX) at the University of Extremadura.

Theoretical simulations to evaluate the stability of dioxane conformers in both gas phase and solution were performed at the M06-2X/6-311++G(d,p) level of theory (Zhao and Truhlar 2008) using the Gaussian09 package (Frisch et al. 2009).

Crystal Data Colorless block-shaped crystals of 2,3-dihydroxy-1,4-dioxane ($\text{C}_4\text{H}_8\text{O}_4$) were obtained from a saturated aqueous solution. A suitable crystal (0.26 x 0.14 x 0.12) was selected and mounted on a Mitgen holder in perfluoroether oil on a Rigaku AFC12 FRE-HF diffractometer equipped with an Oxford Cryosystems Cobra low-temperature apparatus. The crystal was kept at $T = 100(2)$ K during data collection. Using OLEX2 (Dolomanov et al. 2009), the structure was solved by Direct Methods using the ShelXT structure solution program (Sheldrick 2015). The model was refined with ShelXL (Sheldrick 2008) using least squares minimization. The structure was solved in space group Pbca (orthorhombic); $a = 9.5101 \text{ \AA}$, $b = 9.3995 \text{ \AA}$, $c = 12.0120 \text{ \AA}$, $\alpha = \beta = \gamma = 90^\circ$, $V = 1073.76(4) \text{ \AA}^3$, $T = 100(2) \text{ K}$, $Z = 8$, $Z' = 1$, $\mu(\text{MoK}\alpha) = 0.135$; 6659 reflections measured, 1748 unique ($R_{\text{int}} = 0.0184$) which were used in all calculations. The final wR_2 was 0.0834 (all data) and R_1 was 0.0289 ($I > 2(I)$). All non-hydrogen atoms were refined anisotropically. Hydrogen atom positions were calculated geometrically and refined using the riding model.

Dioxane Formation from Ethylene Glycol and Glyoxal The synthetic procedure involves the condensation of ethylene glycol (1 mL, 17.2 mmol) and glyoxal (40% aqueous solution: 2 mL, 17.5 mmol) under stirring at room temperature for 24 h. After evaporation at reduced pressure, the resulting 2,3-dihydroxy-1,4-dioxane crystallized on standing at room temperature, and the crystalline material was collected and washed successively with acetone and diethyl ether (1.95 g, 95%). This material (mp 102 °C)

was employed as intentional seeding in further experiments. The dioxane derivative (**1**) could also be obtained by heating the reaction mixture at *ca.* 60 °C for a week followed by water removal. Moreover, compound **1** could be isolated (67%, mp 105 °C) without water removal, after seeding with the crystalline material obtained above. In the absence of seeding, dioxane crystals were likewise obtained (50% yield), though the process took long (several weeks at 4 °C). Finally, the condensation was also conducted in benzene; the dioxane could be isolated from this solvent on seeding (16% yield, mp 104 °C).

Dioxane Formation from Glycerol and Glyoxal The crude product, 2,3-dihydroxy-5-hydroxymethyl-1,4-dioxane as inferred from IR and NMR data, was obtained by mixing glycerol (2.2 mL, 30.1 mmol) and glyoxal (40% aqueous solution: 3.5 mL, 30.6 mmol) under stirring at room temperature for 24 h, followed by water evaporation at reduced pressure. All attempts to purify the mixture by crystallization failed.

NMR Monitoring In order to verify the formation and stability of dihydroxy dioxanes in aqueous and dilute conditions, samples of either ethylene glycol or glycerol plus glyoxal (0.1 M each in D₂O) were prepared in NMR tubes. Proton NMR spectra were recorded for 4 days. The stability in acidic solution was checked by adding an equimolar amount of glacial acetic acid.

Boron Complexation Experiments To a suspension of borax (0.49 g, 1.3 mmol) in water (10 mL) was added 2,3-dihydroxy-1,4-dioxane (0.12 g, 1 mmol) and the mixture was stirred until complete dissolution. NMR monitoring was performed on D₂O solutions at different reaction times. After 16 h, however, the boron reagent precipitated in part. The complexation was repeated in methanol (10 mL), where both reagents remained in solution for a few weeks; borax precipitation was then observed. The crude residue of 2,3-dihydroxy-5-hydroxymethyl-1,4-dioxane (0.15 g) was likewise treated with borax (0.49 g) in water (5 mL) and kept at room temperature. Borax precipitation occurred after 3 weeks.

Results and Discussion

If polyalcohols were formed in oxygen-containing tholins, such molecules could have been delivered to the early Earth by cometary or meteoritic impacts (McDonald et al.

1996). Such polyalcohols, glycerol in particular, yet achiral structures, were once proposed as surrogates for pentoses in acyclic nucleotide analogs (Joyce et al. 1987, Joyce 1989, Schwartz 1993). Glyceronucleosides, however, show a poor ability to form duplex structures when they are incorporated into an oligonucleotide chain (Schneider and Benner 1990). Because of the coexistence of polyalcohols and carbonyl compounds in the condensed matter of interstellar bodies, it is quite plausible to anticipate the generation of their reaction products, which are on the one hand simpler to prepare than ribo- or deoxyribonucleosides and, on the other, less reactive than the raw precursors. Hydroxylated dioxanes are thus valuable candidates that can easily be formed from bifunctional molecules such as ethylene glycol and glyoxal. The former has been unequivocally identified in comets and the interstellar medium (Hollis et al. 2002; Crovisier et al. 2004). Likewise, glycerol is present in comets (Briggs et al. 1992) and its formation can also be simulated in ice analogs subjected to UV irradiation (Nuevo et al. 2010).

Such structures depicted in Fig. 1, 2,3-dihydroxy-1,4-dioxane (**1**) and 2,3-dihydroxy-5-hydroxymethyl-1,4-dioxane (**2**), show a close resemblance to glycolaldehyde and glyceraldehyde, known to be dimers in the solid state: 2,5-dihydroxy-1,4-dioxane (**3**) and 2,5-dihydroxy-3,6-dihydroxymethyl-1,4-dioxane (**4**), respectively. Glycolaldehyde itself can be obtained from irradiated ethylene glycol (Hudson et al. 2005). It is pertinent to highlight that, like **3** and **4**, structures **1** and **2** may be chiral species in the solid state (see last subsection, Results and Discussion), an extra motivation in the search for homochiral pseudosugars.

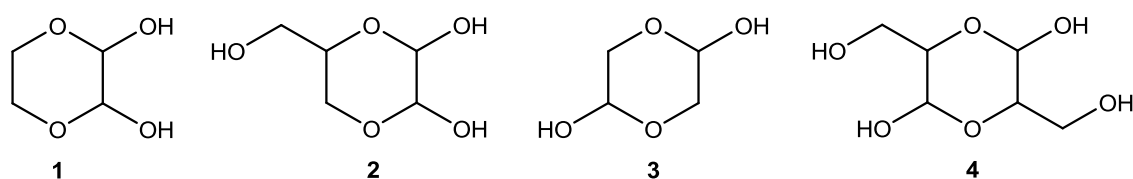


Fig. 1 Dioxanes generated from ethylene glycol (**1**) and glycerol (**2**) plus glyoxal, together with the dimeric structures of glycolaldehyde (**3**) and glyceraldehydes (**4**)

The survivability and stability of glyoxal in interstellar ices is however debatable. This species should reasonably be an intermediate *en route* to more stable derivatives and higher carbonaceous structures (Pavlovskaya and Telegina 1974).

Glyoxal remains stable in aqueous solution yielding monomeric hydrates (**5**, Fig. 2) at low molar concentrations, while dimeric and trimeric hydrates (**6**, **7**) form above 1M concentrations. According to a theoretical study, the dioxolane hydrate (**6**) appears to be kinetically favored over both hydration to the open dimer dihydrate and the dioxane dimer formed by ring closure (Kua et al. 2008).

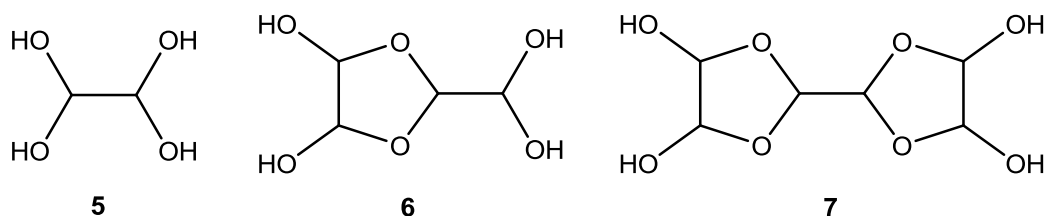


Fig. 2 Representative glyoxal hydrates in aqueous solution

As expected, the coupling of glycols and glyoxal to yield the corresponding adducts is deceptively simple and dates back to the early 1950s at least (Head 1955; Chitwood 1944). Product isolation may be tedious due to water removal, while the use of anhydrous glyoxal, prone to rapid polymerization, requires transformations to be conducted at 0 °C. In the 1980s, a useful variation was reported in anhydrous benzene at reflux with concomitant azeotropic removal of water (Venuti 1982). 2,3-Dihydroxy-1,4-dioxane was thus introduced as a synthetic equivalent of anhydrous glyoxal that can be released upon reaction with numerous organic partners. While these conditions are far from being prebiotic, we re-examined the condensation of neat reagents (glyoxal being *ca* 40% aqueous solution) and found that the dioxane adduct crystallized on standing after complete water evaporation, which is presumably akin to environmental conditions. Crystallization could also be observed in saturated aqueous solution, though this process may take long (several weeks at 4 °C, often cooling at 0 °C). Likewise, product crystallization took place in solution and open flasks after seeding, which gave rise to incipient crystallization in lower yields nevertheless,

All the samples, isolated by precipitation or re-crystallization, gave similar IR patterns. Notably, such spectra (Fig. 3) were quite similar to those of major and minor HPLC fractions separated from ice tholins (McDonald et al. 1996), thereby pointing out the dominance of hydroxylated compounds by photoirradiation. A very strong band at 3336 cm⁻¹ supports the existence of O–H bonds associated by hydrogen bonding. This broad band nearly overlaps C–H stretching bands between 3005 and 2888 cm⁻¹. The

strong bands located at 1148 and 1024 cm^{-1} are indicative of C–O and C–C vibrational modes. Unlike tholin mixtures, no significant bands appear in the range 1600–1700 cm^{-1} , which could be attributed to carbonyl and alkene absorptions.

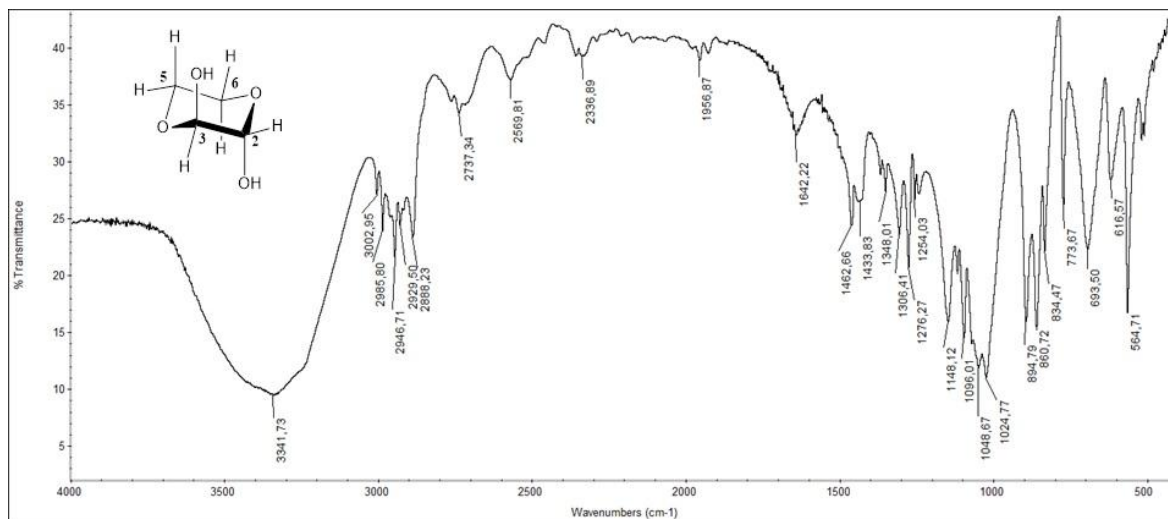


Fig. 3 FT-IR spectrum (KBr disc) of 2,3-dihydroxy-1,4-dioxane

The dihydroxylated dioxane can also exist as a mixture of four distinct stereoisomers and their corresponding conformers in solution (*vide infra*). It is well established that while glycolaldehyde exists essentially as stable dimer in the solid state (dioxane structure), a complex equilibrium between acyclic and cyclic forms (both five and six-membered rings) takes place in solution (Yaylayan et al. 1998). As shown in Fig. 4, four stereoisomeric structures are possible for dioxane **1** owing to the presence of two chiral centers. The *trans* isomer is a mixture of compounds (**1A**) and (**1C**) bearing an enantiomeric relationship, each existing in both diaxial and diequatorial conformers. For the sake of clarity, only one conformer is shown for a given enantiomer. The *cis* isomer is actually a *meso* diastereomer (**1B** = **1D**). Such forms are isoenergetic as they invariably possess one hydroxyl group in axial disposition and the other in equatorial orientation. Compounds **1B** and **1D** are identical albeit, as above, depicted in different conformation.

In principle, the anomeric effect should account for the greater stabilization of diaxial structures, yet counterbalancing steric hindrance. Most proton NMR spectra (as recorded in perdeuterated DMSO) show only one signal set that points to a single diastereomer; although a minor product (less than 1% on the NMR accuracy) could also be detected in a few crystallization batches (Fig. 5).

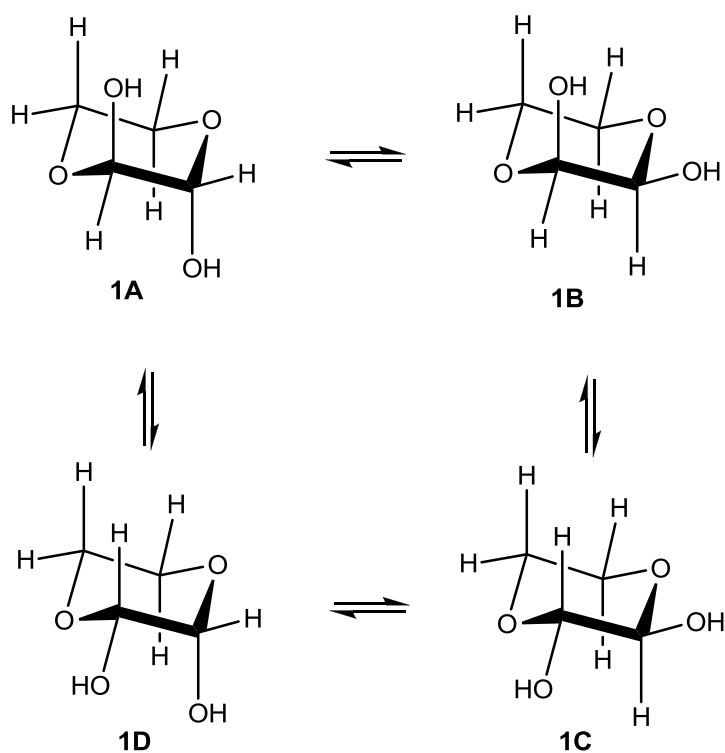


Fig. 4 Chair structures in equilibrium for 2,3-dihydroxy-1,4-dioxane

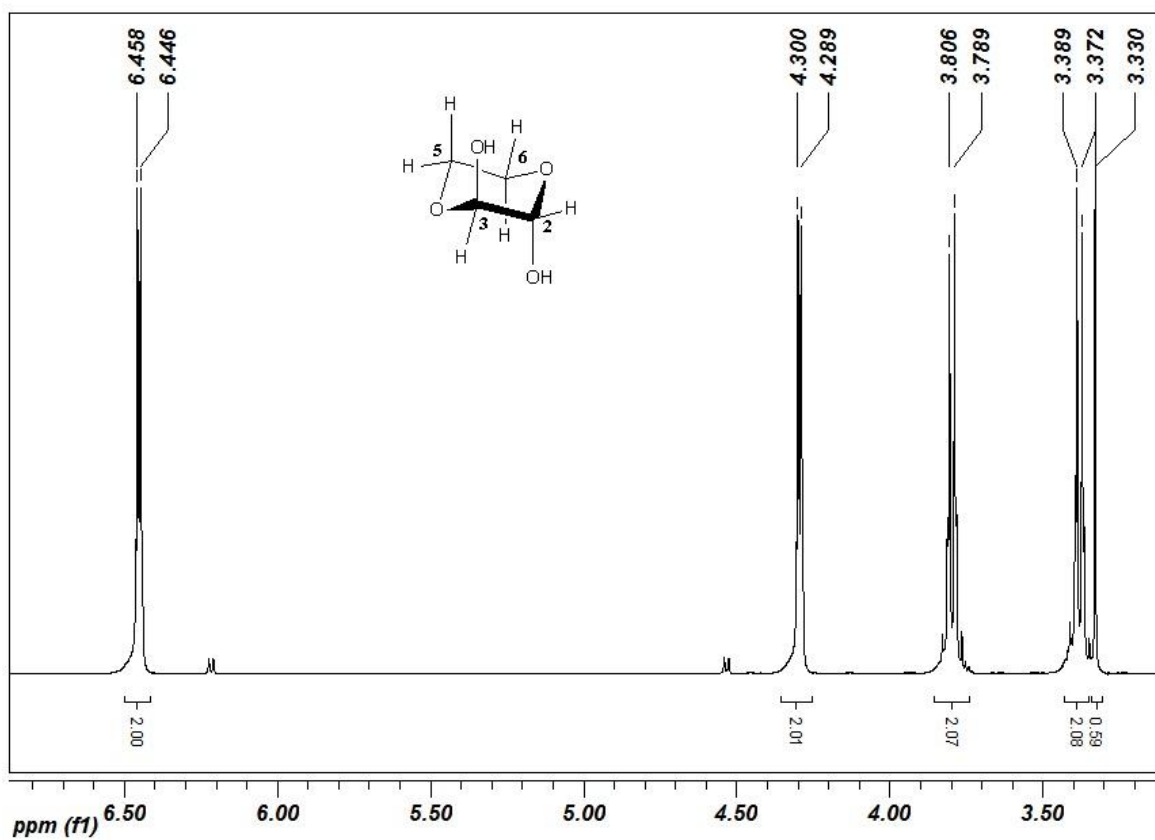


Fig. 5 ^1H NMR spectrum of dioxane **1** recorded in $\text{DMSO-}d_6$

The spectrum shows four non-equivalent (magnetically) proton sets, whose assignments were accomplished by selective irradiation plus deuterium exchange experiments. The OH groups resonate downfield at 6.45 ppm which exert electron-withdrawing effects on the vicinal hydrogens at C-2 and C-3 (4.29 ppm). Hydrogen atoms at C-5 and C-6 are much more shielded and resonate as two signal sets, which should reasonably be ascribed to either axial or equatorial arrangements. It is well known that axial protons are usually more deshielded than their equatorial counterparts in chair structures (Klod et al. 2002). As we shall see later, the diaxial structure is further supported by theoretical analysis. Accordingly, the minor isomer with proton signals at ~4.55 ppm would correspond to the diequatorial structure. In line with the preceding statement, those signals can be attributed to axial protons, whereas the major diaxial isomer places the same protons in equatorial positions at 4.29 ppm (upper field). It was often difficult to remove ethylene glycol from crystalline samples and the signal at 3.33 ppm does correspond to its methylene groups. This was further corroborated by adding deliberately ethylene glycol and recording the spectrum anew, which enhances the above-mentioned signal with the concomitant appearance of the OH resonance at 4.42 ppm. The apparent coupling constants measured on the coupled spectrum are $J_{\text{H,OH}} = 6$ Hz, $J_{\text{axial,equatorial}} = 2.5$ Hz, and $J_{\text{gem}} = 11$ Hz; the latter lies in the expected range for geminal couplings in chiral structures. The H/D exchange removes the OH signals and gives rise to a singlet (at 4.30 ppm) for the geminal protons. This collapse is also observed for the minor isomer, which supports the above assignments. The ^{13}C NMR spectrum is quite simple and shows only two resonances at 92.5 (C-2, C-3) and 60.6 ppm (C-5, C-6) as the diaxial isomer exhibits magnetic equivalence in both methylene and CH atoms; a fact that also rules out the existence of a structure with two OH groups in different orientations (axial and equatorial).

A preliminary assessment of stereoisomer stability, which corroborates in addition NMR data, comes from a high-level DFT computation with the MO6-2X hybrid functional and the extended basis 6-311++G(d,p) in both the gas phase and solution, the latter using the SMD method, as implemented in the Gaussian package (Frisch et al. 2009). On simulating structures **1A-1D** that differ by the spatial orientation of the OH functionality, the one showing both hydroxyl groups in axial dispositions (**1A**) corresponds to an energy minimum in water (Fig. 6, Table 1).

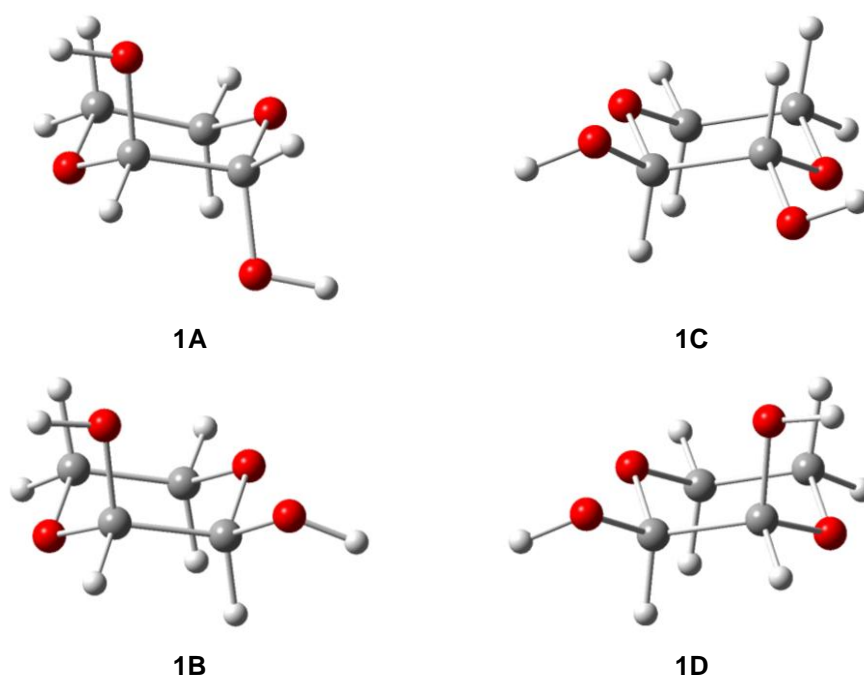


Fig. 6 Computation-based geometries for structures **1A-1D** of 2,3-dihydroxy-1,4-dioxane

Even though energy differences are less than 1 kcal/mol, the diaxial isomer appears to be the most stable structure. Structures **1B** and **1D** bear an enantiomeric relationship and their energies are identical, which ensures the validity of the computational approach. In other polar solvents, for instance DMSO where NMR data have been recorded, the diaxial isomer is again the most favored structure. This trend however changes in the gas phase (Table 2), where the axial-equatorial structure becomes the minimum along the conformational profile.

Table 1. Relative stabilities of dioxane **1** structures in water*

Conformer	ΔE (kcal/mol)	ΔH (kcal/mol)	ΔG (kcal/mol)
1A	0.00	0.00	0.00
1B	0.69	0.42	0.54
1C	0.83	0.10	0.07
1D	0.69	0.43	0.55

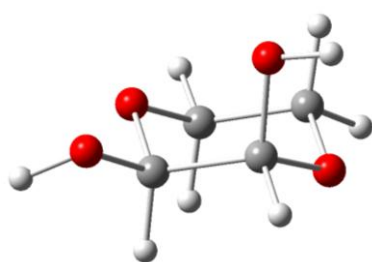
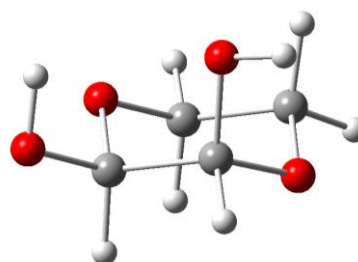
*At the M062X/6-311++G(d,p) level and the SMD model

Table 2. Relative stabilities of dioxane **1** structures in the gas phase*

Conformer	ΔE (kcal/mol)	ΔH (kcal/mol)	ΔG (kcal/mol)
1A	0.77	1.01	0.94
1B	0.00	0.00	0.00
1C	3.21	2.84	2.34
1D	0.00	0.00	0.00

*At the M062X/6-311++G(d,p) level

These variations can be altered significantly if one takes into account intramolecular hydrogen bonding present in these species. Table 3 gathers the relative energies of dioxane **1D** in DMSO, where the intramolecular bonding largely stabilizes the structure.

**1D-HB1****1D-HB2****Table 3.** Relative stabilities of HB-dioxane conformers **1D** in DMSO*

Conformer	ΔE (kcal/mol)	ΔH (kcal/mol)	ΔG (kcal/mol)
1D-HB1	2.79	2.76	2.59
1D-HB2	0.00	0.00	0.00

*At the M062X/6-311++G(d,p) level and the SMD model

At this stage one should necessarily move to dilute and aqueous conditions, which would be akin to prebiotic syntheses. We were still able to detect the formation of dioxane **1** under neutral conditions (from ethylene glycol and glyoxal, 0.1 M each in D₂O). Hydrolysis and concomitant release of ethylene glycol do indeed occur in water, but the system reaches an equilibrium state with little or no evolution after 24 h (peak integration remains essentially constant after 4 days).

Remarkably, the dioxane signals disappeared when NMR samples were heated at *ca.* 60 °C, thereby suggesting that ring hydrolysis is accelerated under thermal conditions. Notably, the presence of an acidic medium (by adding an equimolar amount of acetic acid) did promote the appearance of dioxane **1**. Although it is well-known that hemiacetals and ketals are sensitive to slightly acidic conditions, acid catalysis also favors their formation (Smith and March 2001). Likewise, an aqueous solution of dioxane **1** (0.1 M in D₂O) under acid conditions (equimolar amount of glacial acetic acid) gave rise to a steady equilibrium as evidenced by NMR monitoring.

In agreement with our theoretical predictions, ¹H NMR spectra of dioxane **1** in D₂O reveal the formation of the three structures (not only the diaxial one **1A** and occasionally **1C**), which were identified again through inspection of the proton resonances at C-2 and C-3. Such data together with an approximate estimation of stereoisomeric populations are gathered in Table 4.

Table 4. Proton resonances (H-2, H-3) for dioxane **1** structures in D₂O

Conformer	Proton shifts (ppm)	Population*
1A	4.62	62.5 %
1C	4.88	21.2 %
1B/1D	4.59, 4.85	16.3 %

*By integration of peak areas

The analogous condensation of neat reagents, glycerol and glyoxal, to afford the corresponding dioxane (**2**) by spontaneous crystallization was unsuccessful. There is little doubt about the formation of such adduct by IR inspection of the syrupy residue (Fig. 7). The spectrum resembles that of ice tholin II where glycerol and other alcohols formed upon irradiation (McDonald et al. 1996). The major peak is a strong and broad O-H stretching band centered at 3372 cm⁻¹. Medium-size C-H stretching bands at 2939 and 2884 cm⁻¹, indicative of aliphatic bonds, are accompanied by a strong band between 1000 and 1200 cm⁻¹ attributed to C-O-C linkages. The spectrum shows in addition a relatively strong band at 1636 cm⁻¹, which suggests a side product or unreacted glyoxal. However, its frequency is not consistent with a typical C=O stretching band; in fact the signal could not be corroborated through NMR spectroscopy either. Proton (Fig. 8) and

^{13}C NMR spectra were more complex than those of dioxane **1**, probably due to the existence of additional diastereomers in equilibrium (Fig. 9), each existing as a mixture of several conformers.

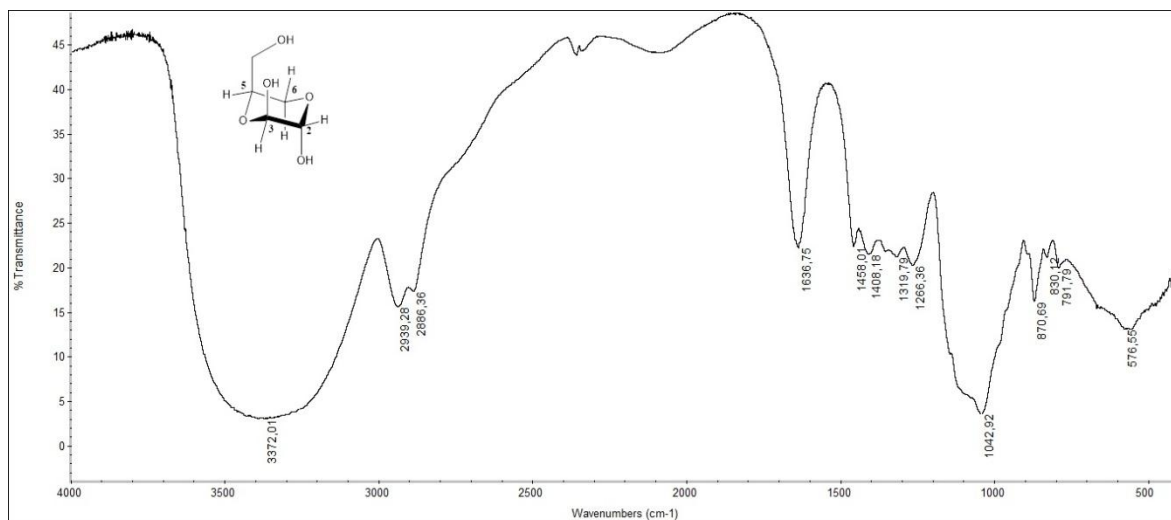


Fig. 7 FT-IR spectrum (KBr disc) of 2,3-dihydroxy-5-hydroxymethyl-1,4-dioxane (crude product)

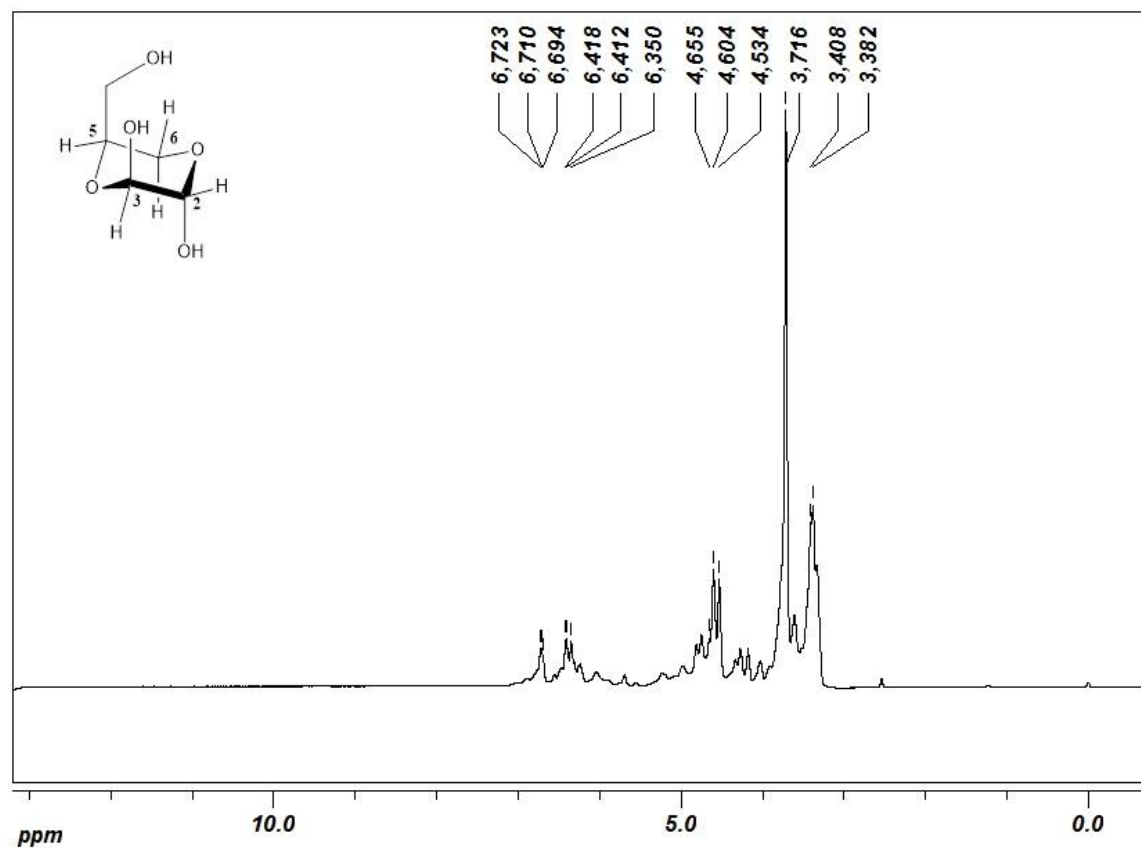


Fig. 8 ^1H NMR spectrum of dioxane **2** (crude residue) recorded in $\text{DMSO}-d_6$

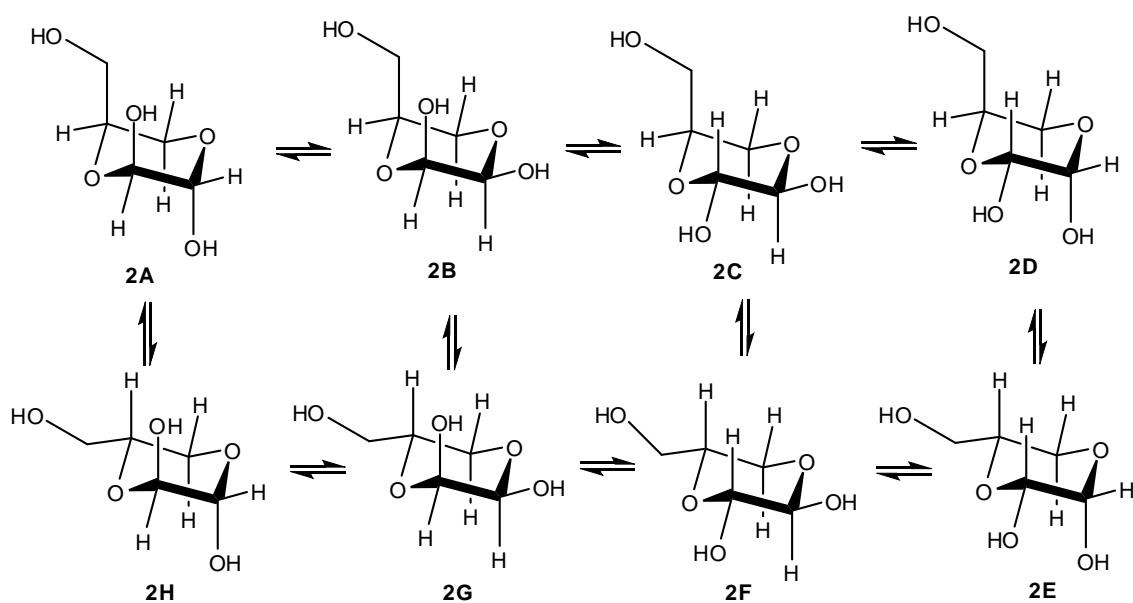
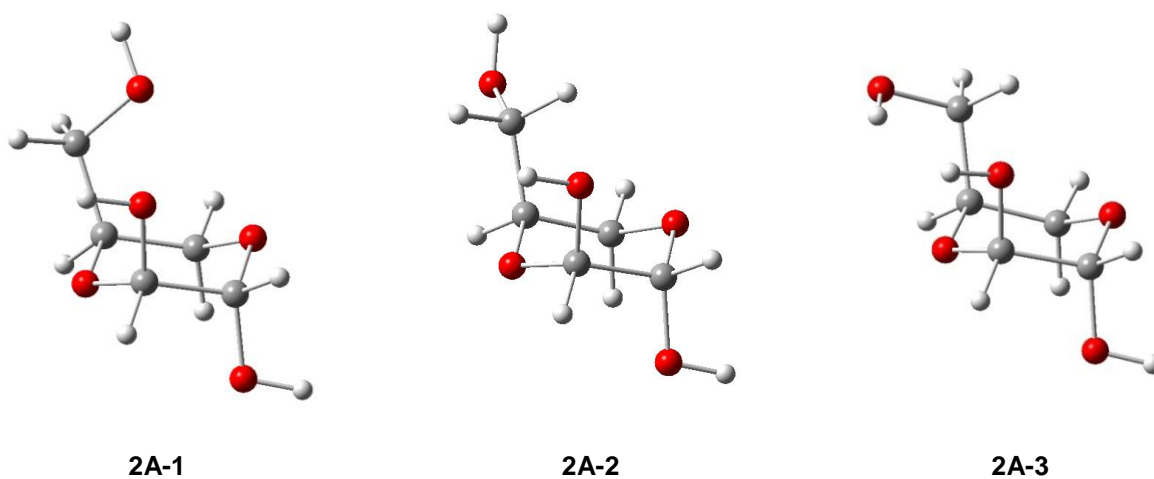
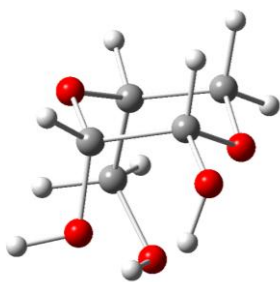


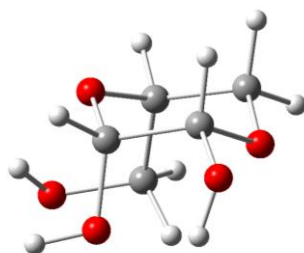
Fig 9 Diastereomeric equilibria for 2,3-dihydroxy-5-hydroxymethyl-1,4-dioxane

A further structural analysis could be accomplished for this dioxane derived from glycerol, although the iterative protocol is much more complex due to the potential existence of multiple conformers for structures **2A-2H**. Thus, series of conformers are depicted below (Fig. 10) by considering in every case the rotation around the hydroxymethyl group.

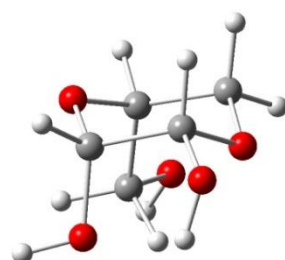




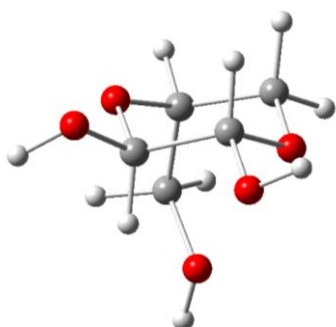
2B-1



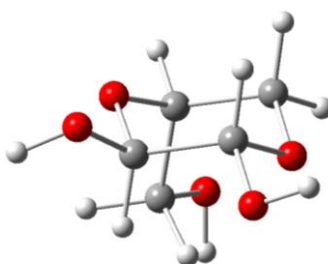
2B-1



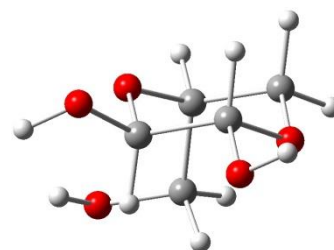
2B-3



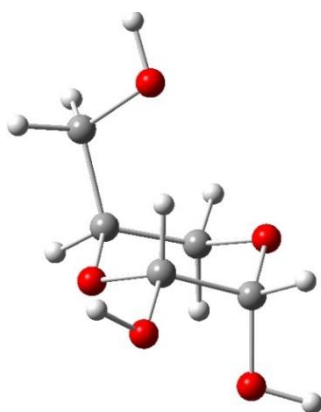
2C-1



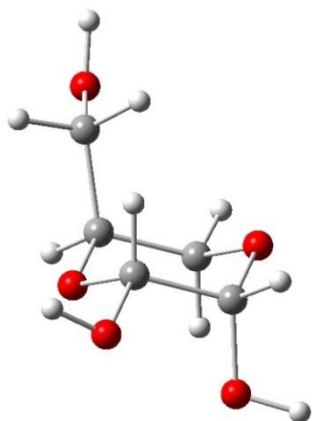
2C-2



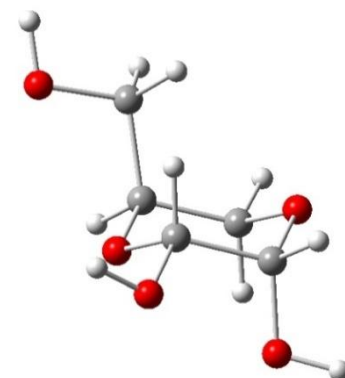
2C-3



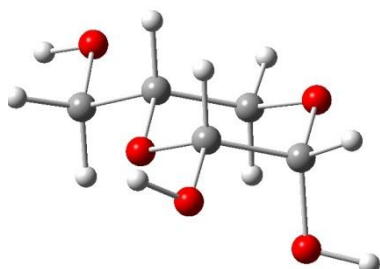
2D-1



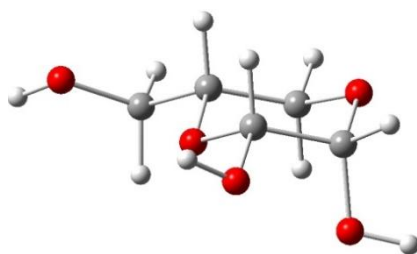
2D-2



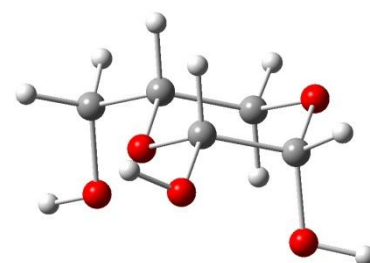
2D-3



2E-1



2E-2



2E-3

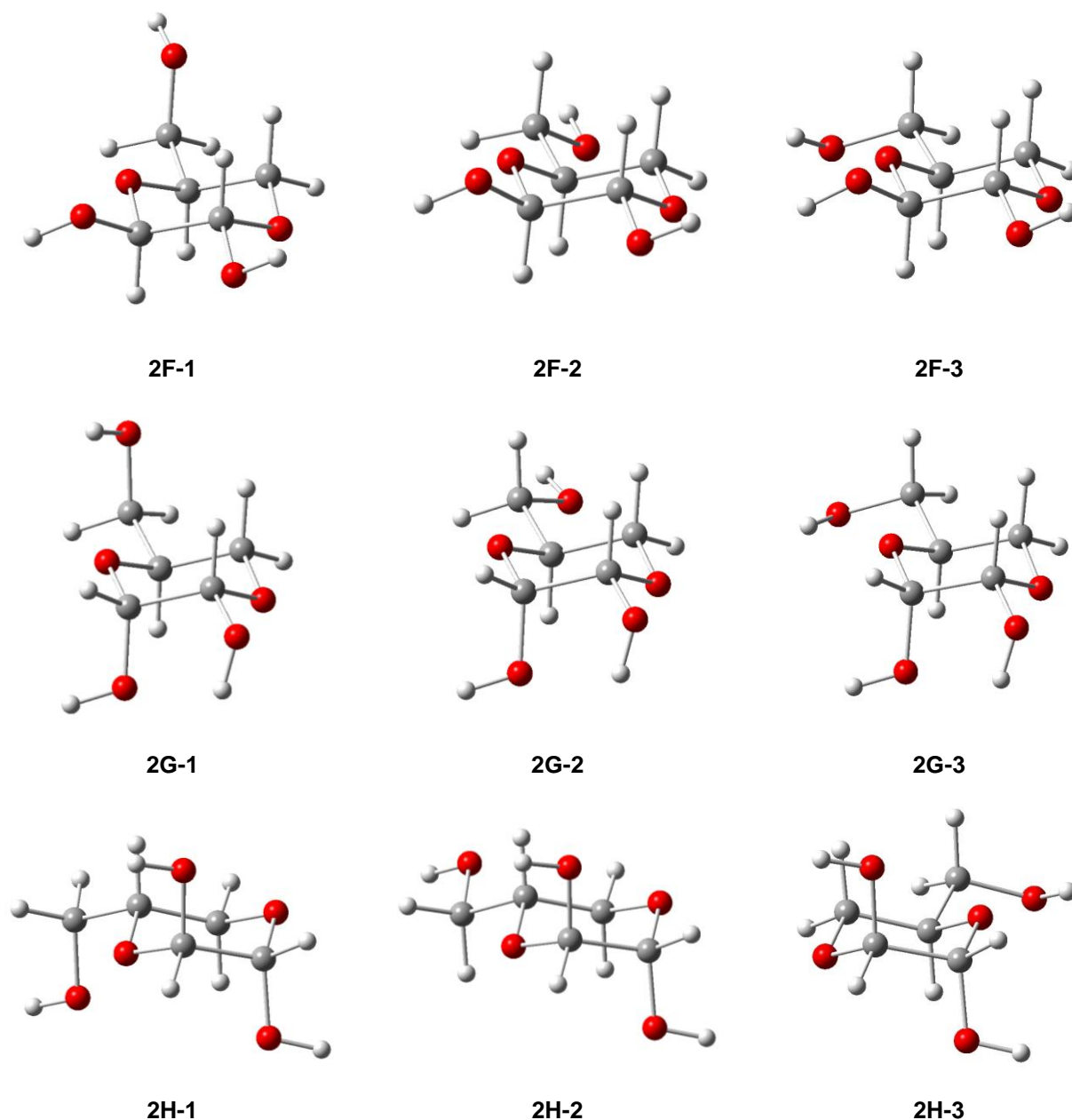


Fig. 10 Computation-based geometries for structures **2A-2H** of 2,3-dihydroxy-5-hydroxymethyl-1,4-dioxane

Table 5 shows the relative energies computed in water. The most stable structure (which is also obtained in DMSO and the gas phase) is the one having the hydroxyl groups in axial and equatorial dispositions and the hydroxymethyl substituent in equatorial position (**2G**). The energy differences increase in DMSO owing to its greater polarity, while decrease in water as a result of favorable hydrogen-bonding interactions.

Table 5. Relative stabilities of dioxane **2** structures in water*

Conformer	ΔE (kcal/mol)	ΔH (kcal/mol)	ΔG (kcal/mol)
2A-1	5.35	5.63	5.84
2A-2	2.81	2.98	3.16
2A-3	1.43	2.03	2.10
2B-1	0.90	1.28	2.38
2B-2	0.93	1.23	1.96
2B-3	2.51	2.41	2.36
2C-1	4.78	4.40	4.13
2C-2	3.64	3.15	2.88
2C-3	2.13	1.92	2.04
2D-1	4.80	4.76	4.96
2D-2	3.76	3.73	3.84
2D-3	3.20	3.32	3.43
2E-1	2.33	2.22	1.86
2E-2	1.60	1.68	1.69
2E-3	1.92	1.93	2.09
2F-1	1.99	1.51	1.33
2F-2	2.49	2.02	1.89
2F-3	1.04	0.62	0.53
2G-1	0.01	0.07	0.28
2G-2	1.41	1.51	1.69
2G-3	0.00	0.00	0.00
2H-1	1.03	1.28	1.48
2H-2	1.53	1.86	1.76
2H-3	0.07	0.42	0.48

*At the M062X/6-311++G(d,p) level and using the SMD method

As expected and like dioxane **1**, ^1H NMR spectra of glycerol and glyoxal mixtures (*ca.* 0.1 M each in D_2O) point to more complex configurational and conformational equilibria, which are consistent with the preceding computational assessment given the low energy differences. Such proton spectra are indeed complex, although on zooming the zone between 4.3 and 5.0 ppm the existence of multiple resonances with different coupling constants suggests the formation of at least four structures in freshly prepared solutions. The most abundant species, inferred from peak area integration, should most likely be structure **2F** (diequatorial isomer), which appears with two doublets resonating at 4.57 and 4.48 ppm ($J \sim 7$ Hz). In the aqueous medium progressive hydrolysis is observed, albeit *ca.* 10% of dioxane signals (relative to glycerol) remain unaffected after 4 days.

Finally we have recorded low-resolution mass spectra (electrospray technique, positive mode in 0.1 M formic acid solution) for dioxanes **1** and **2**. Although this technique in question appears to induce other *in-situ* transformations, significant peaks at m/z 117 for dioxane **1** (MW = 120) and m/z 147 for dioxane **2** (MW = 150) were observed. Such peaks would be consistent with the M+1 signals of the corresponding dilactones (i.e. generated by oxidation of the secondary hydroxyl groups).

Boron Coordination

As hydroxylated derivatives and, despite the fact that the OH function is a weak coordinating group, these dioxanes generated from polyalcohols are prone to undergo complexation with metal ions; actually a credible mechanism in environmental niches that channels subsequent evolutionary stages. There has been considerable controversy about the role and prebiotic abundance of certain metal species, although boron has emerged as the most promising candidate in selective sequestration of sugar-like molecules as chiefly proven by Benner and his coworkers (Kim et al. 2011; Ricardo et al. 2004; Scorei 2012).

It has long been established that boron-containing reagents coordinate with diols; an ability largely exploited in organic synthesis with boronic acids in organic solvents (Hall 2011). In general, the reactivity of boronates toward sugar diols follow the order *cis*-1,2-diol > 1,3-diol >> *trans*-1,3-diol. Although boron minerals are scarce, our preliminary test was conducted with dioxane **1** and borax, which is readily available as disodium salt, and monitored by proton NMR in D_2O . In this deuterated medium, borax shows a single resonance at 4.71 ppm. On recording the spectrum of a sample

containing equimolar amounts of **1** and borax, immediate changes were visible and strongly pointed to complexation. Thus, the two signal sets of axial and equatorial hydrogens collapse into a single signal at 3.57 ppm, while the hydrogen atoms at C-2 and C-3 resonate at 5.25 ppm (Fig. 11). As a result of rapid H/D exchange in D₂O, no signals for OH groups appear. Unfortunately, borax precipitated partly after 16 h, and the results should likely be interpreted in terms of complexation and rapid equilibration between the *trans* and *cis*-1,2-diol isomers.

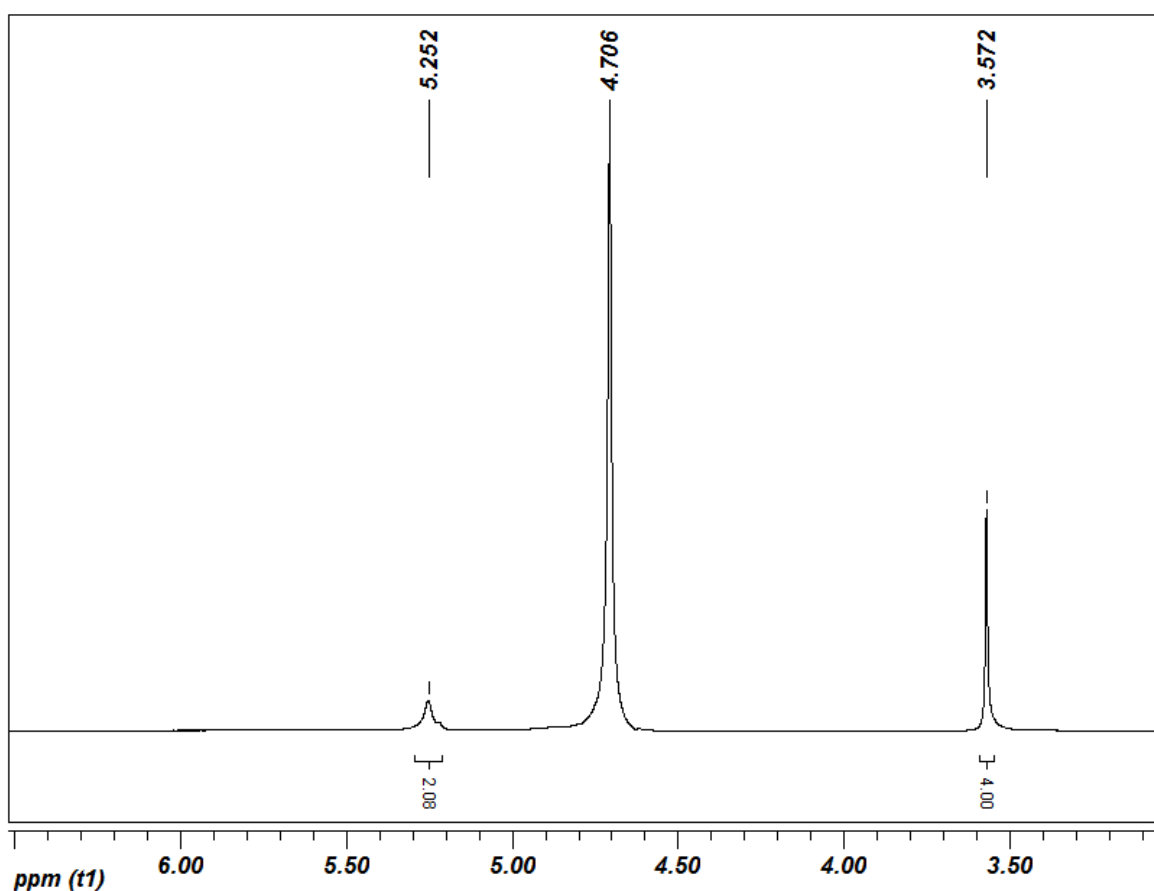


Fig. 11 ¹H NMR spectrum of borax and dioxane **1** in D₂O solution

Stereochemical Identity of 2,3-Dihydroxy-1,4-dioxane

Can dioxanes shed light into the stereochemical bias of pseudosugars in particular and sugars in general? As oxycarbacycles that mimic closely the structure of pyranoid sugars, dioxanes may certainly be prototypical models. Although natural sugars exhibit a distinctive D-configuration, the origin of this homochirality remains essentially unknown. One could envisage an *a priori scenario* where other chiral

molecules and auxiliaries would have deracemized the primeval sugars; on the other hand, in an *a posteriori* scenario, mirror symmetry breaking would have been unavoidable as a consequence of the increasing constitutional complexity, i.e. the number of oligonucleotide sequences from an inherently racemic mixture of components always exceeds a critical limit (Eschenmoser 2011).

Unprotected sugars do not form conglomeratic phases and, therefore spontaneous segregation by crystallization cannot be invoked. Pure enantiomers are difficult to obtain and, as paradigmatic example, D-ribose (for which at least two polymorphs exist) resisted all attempts to generate suitable crystals for X-ray diffraction studies until a few years ago (Šišak et al. 2010). Racemic DL-ribose crystallizes easily instead due to conformational preferences and H-bonding interactions (Patyk and Katrusiak 2014).

Having demonstrated that crystalline 2,3-dihydroxy-1,4-dioxane most likely corresponds to the *trans* isomer, which is inherently chiral (the *cis* isomer is obviously a *meso*, achiral isomer), our next task was the clear-cut identification of the racemate, i.e. either conglomerate (racemic mixture with separation of homochiral domains) or racemic compound possessing heterochiral domains within every crystal. Detection of conglomerates can usually be accomplished by methods that enable comparisons between enantiomers and racemates like IR, Raman, solid-state NMR spectroscopies, or X-ray powder diffraction (XRPD). Since pure enantiomers of 2,3-dihydroxy-1,4-dioxane were not available to us, we thought, rather naively, that further efforts could be paid to the search for conglomerates. To this end, the condensation of ethylene glycol and glyoxal was performed repeatedly under vigorous stirring looking for secondary nucleation conditions. Unfortunately, the solids collected showed no optical activity. To ascertain unequivocally the nature of such a racemate, we turned the attention to second harmonic generation (SHG), a non-linear method that is reliable enough for spotting conglomerates, based on signal detection at $\lambda/2$ wavelengths (Galland et al. 2009; Dupray 2012). Detection of SHG signals may be indicative of non-centrosymmetric crystals, which point to either conglomerates or racemic compounds crystallizing in non-centrosymmetric space groups. Still, false positives may also be encountered for centrosymmetric crystals exhibiting photoluminescence phenomena. However, the absence of a SHG signal most likely evidences the existence of a racemic compound (centrosymmetric space group), which turned to be the situation of dioxane **1**, thereby ruling out resolution by secondary nucleation. A few chiral space groups (*e.g.* $P4_12_12$ or $P4_32_12$), indeed rare among conglomerates, have been reported to be SHG inactive,

although the intensity of laser beams currently employed discard such false negatives too (Galland et al. 2009).

We were at last able to grow crystals of dioxane **1** suitable for X-ray diffraction after slow evaporation. It is relevant to note that, despite the structural simplicity of this dioxane, it has not yet been deposited with the Cambridge Structural Database (CSD). The ORTEP diagram and unit cell are shown in Fig. 12 (orthorhombic, space group *Pbca*, $Z = 8$), which exhibits a chair structure with axial OH groups, thus suggesting strong anomeric effects in the crystal lattice.¹

This structure is closely related to that of glycolaldehyde dimer as inferred from X-ray studies too (Mohaček-Grošek et al. 2013). As mentioned, the first chiral sugar, glyceraldehyde, may form a stable dimer in the solid state. *rac*-Glyceraldehyde adopts in its dimer form a chair symmetrical *p*-dioxane structure with all the hydroxyl and hydroxymethyl groups bound equatorially nevertheless (Senma et al. 1973; Lehmann and Luger, 1991). This arrangement, similar to that of β -D-glucopyranose, favors a heterochiral crystal packing (spatial group $P2_1/a$; $Z = 2$).

¹Crystallographic data have been deposited with the Cambridge Crystallographic Data Centre (registry number CCDC-1407505). Such data can be retrieved upon permission or from the authors

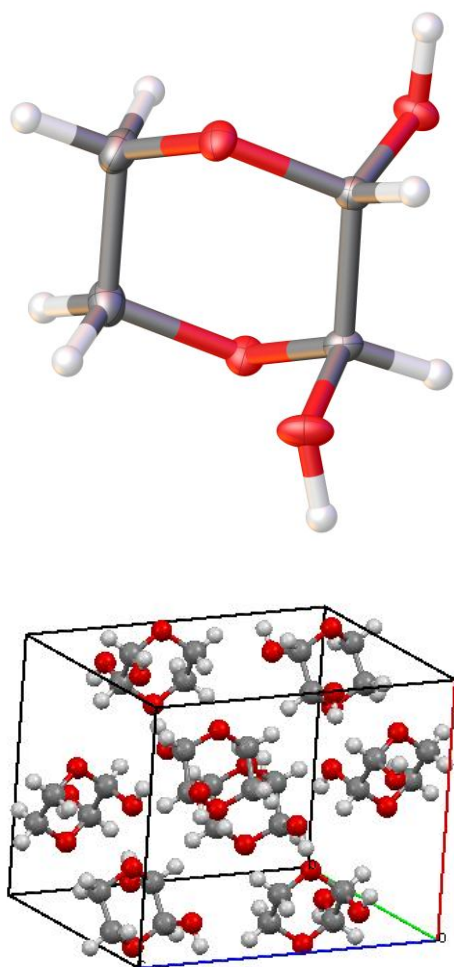


Fig. 12 Solid-state structure of 2,3-dihydroxy-1,4-dioxane (**1**). Thermal ellipsoids drawn at the 50 percent probability level. A representation of the unit cell indicates the number and orientation of dioxane molecules

In conclusion, formation of dioxane structures from polyalcohols and carbonyls present in both astrochemical and terrestrial arenas, should be plausible and prebiotically credible, especially in the solid state like the archetypal cases of glycolaldehyde and glyceraldehyde. This work documents in detail structural features that should be useful guidelines for future pursuits, as well as aqueous chemistry where formation of various stereoisomers might however have hampered selective routes of chemical biogenesis in favor of monohemiacetalic structures. The dioxane intermediates arguably complement the chemical space of sugar-like structures. Although the goal of achieving homochirality has not yet been fulfilled given the absence of chiral space groups in the solid state, at least for dioxane **1**, the formation of chiral dioxanes with

close resemblance to pyranoid rings suggests this effort is worthwhile. Similar condensations involving glycolaldehyde or glyceraldehyde, both in monomeric and dimeric form, with carbonyls or alcohols would afford additional routes within the puzzle of prebiotic carbohydrate chemistry. Such studies are under way in our laboratories.

Acknowledgments This research was funded by the Ministry of Economy and Competitiveness (Research Grant CTQ2013-44787-P). We also thank the financial support from the Autonomous Government of Extremadura and FEDER, aimed at the consolidation of research groups (Grant GRU15022). Last, but not least, we are grateful to Prof. Gérard Coquerel for assessing the racemic nature of 2,3-dihydroxy-1,4-dioxane by SHG measurements.

References

- Benner SA, Kim HJ, Kim MJ, Ricardo A (2010) Planetary organic chemistry and the origins of biomolecules. *Cold Spring Harbor Perspect Biol* 2:a003467
- Benner SA, Kim HJ, Yang Z (2012) Setting the stage: the history, chemistry, and geobiology behind RNA. *Cold Spring Harbor Perspect Biol* 4:a003541
- Briggs R, Ertem G, Ferris JP, Greenberg JM, McCain PJ, Mendoza-Gomez CX, Schutte W (1992) Comet Halley as an aggregate of interstellar dust and further evidence for the photochemical formation of organics in the interstellar medium. *Orig Life Evol Biosph* 22:287-307
- Cable ML, Hörst SM, Hodyss R, Beauchamp PM, Smith MA, Willis PA (2012) Titan tholins: simulating Titan organic chemistry in the Cassini-Huygens era. *Chem Rev* 112:1882-1909
- Chitwood HC (1944) Substituted dioxanes. US Patent 2,361,456
- Cooper G, Kimmich N, Belisle W, Sarinana J, Brabham K, Garrel L (2001) Carbonaceous meteorites as a source of sugar-related organic compounds for the early Earth. *Nature* 414:879-883

- Crovisier J, Bockelée-Morvan D, Biver N, Colom P, Despois D, Lis DC (2004) Ethylene glycol in comet C/1995 O1 (Hale-Bopp). *Astron Astrophys* 418:L35-L38
- Dolomanov OV, Bourhis LJ, Gildea RJ, Howard JAK, Puschmann H (2009). OLEX2: a complete structure solution, refinement and analysis program. *J Appl Cryst* 42:339-341
- Dupray V (2012) Recrystallization of enantiomers from conglomerates. In: Sztwiertnia K (ed) *Recrystallization*, Intech, Rijeka, pp 403-424
- Eschenmoser A (2011) Etiology of potentially primordial biomolecular structures: from vitamin B₁₂ to the nucleic acids and an inquiry into the chemistry of life's origin: a retrospective. *Angew Chem Int Ed* 50:12412-12472
- Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR et al (2009) *Gaussian09*, Revision B.01, Gaussian Inc, Wallingford
- Galland A, Dupray V, Berton B, Morin-Grognet S, Sanselme M, Atmani H, Coquerel G (2009) Spotting conglomerates by second harmonic generation. *Cryst Growth Des* 9:2713-2718
- Gollihar J, Levy M, Ellington AD (2014) *Biochemistry*. Many paths to the origin of life. *Science* 343:259-260
- Hall DG (2011) Structure, properties, and preparation of boronic acid derivatives. Overview of their reactions and applications. In: Hall DG (ed) *Boronic acids: preparation and applications in organic synthesis, medicine and materials*, 2nd edn. Wiley-VCH, Weinheim, pp 1-133
- Harman CE, Kasting JF, Wolf ET (2013) Atmospheric production of glycolaldehyde under hazy prebiotic conditions. *Orig Life Evol Biosph* 43:77-98
- Head FSH (1955) An addition compound of glyoxal and ethylene glycol. *J Chem Soc* 1036-1037
- Hollis JM, Lovas FJ, Jewell PR (2000) Interstellar glycolaldehyde: the first sugar. *Astrophys J* 540:L107-L110
- Hollis JM, Lovas FJ, Jewell PR, Coudert LH (2002) Interstellar antifreeze: ethylene glycol. *Astrophys J Lett* 571:L59-L62
- Hudson RL, Moore MH, Cook AM (2005) IR characterization and radiation chemistry of glycolaldehyde and ethylene glycol ices. *Adv Space Res* 36:184-189

Irvine WM (1999) The composition of interstellar molecular clouds. *Space Rev Sci* 90:203-218

Joyce GF (1989) RNA evolution and the origin of life. *Nature* 338:217-224

Joyce GF, Schwartz AW, Miller SL, Orgel LE (1987) The case for an ancestral genetic system involving simple analogues of the nucleotides. *Proc Natl Acad Sci USA* 84:4398-4402

Kaiser RI (2002) Experimental investigation on the formation of carbon-bearing molecules in the interstellar medium via neutral-neutral reactions. *Chem Rev* 102:1309-1358

Kim HJ, Ricardo A, Illangkoon HI, Kim MJ, Carrigan MA, Frye F, Benner SA (2011) Synthesis of carbohydrates in mineral-guided prebiotic cycles. *J Am Chem Soc* 133:9457-9468.

Kua J, Hanley SW, De Haan DO (2008) Thermodynamics and kinetics of glyoxal dimer formation: a computational study. *J Phys Chem A* 112:66-72

Klod S, Koch A, Kleinpeter E (2002) Ab initio quantum-mechanical GIAO calculation of the anisotropic effect of C-C and X-C single bonds—application to the ¹H NMR spectrum of cyclohexane. *J Chem Soc Perkin Trans 2* 1506-1509

Lehmann CW, Luger P (1991) Crystal structure redetermination of the d,l-glyceraldehyde dimer: a comparison of synchrotron and conventional X-ray diffraction data. *Z Kristallogr* 195:49-63

McDonald GD, Whited LJ, DeRuiter C, Khare BN, Patnaik A, Sagan C (1996) Production and chemical analysis of cometary ice tholins. *Icarus* 122:107-117

Meierhenrich U (2015) Comets and their origins. Wiley-VCH, Weinheim, pp 133-161

Mohaček-Grošek V, Prugovečki B, Prugovečki S, Strukan N (2013) Glycolaldehyde dimer in the stable crystal phase has axial OH groups: Raman, infrared and X-ray data analysis. *J Molec Struct* 1047:209-215

Nuevo M, Bredehöft JH, Meierhenrich UJ, d'Hendecourt L, Thiemann WHP (2010) Urea, glycolic acid, and glycerol in an organic residue produced by ultraviolet irradiation of interstellar/pre-cometary ice analogs. *Astrobiology* 10:245-256

- Patel BH, Percivalle C, Ritson DJ, Duffy CD, Sutherland JD (2015) Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nature Chem* 7:301-307
- Patyk E, Katrusiak A (2014) Conformational and H-bonding preferences for facile racemate crystallization of ribose. *Chirality* 26:806-810
- Pavlovskaya TE, Telegina TA (1974) Photochemical conversion of lower aldehydes in aqueous solutions and in fog. *Orig Life* 5:303-309
- Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* 459:239-242
- Ricardo A, Carrigan MA, Olcott AN, Benner SA (2004) Borate minerals stabilize ribose. *Science* 303:196
- Ruiz-Mirazo K, Briones C, De la Escosura A (2014) Prebiotic systems chemistry: new perspectives. *Chem Rev* 114:285-366
- Sagan C, Khare BN (1979) Tholins: organic chemistry of interstellar grains and gas. *Nature* 277:102-107
- Saladino R, Carota E, Botta G, Kapralov M, Timoshenko GN, Rozanov AY, Krasavin E, Di Mauro E (2015) Meteorite-catalyzed syntheses of nucleosides and of other prebiotic compounds from formamide under proton irradiation. *Proc Natl Acad Sci USA*. doi: 10.1073/pnas.1422225112
- Schneider KC, Benner SA (1990) Oligonucleotides containing flexible nucleoside analogues. *J Am Chem Soc* 112:453-455
- Schwartz AW (1993) Nucleotide analogs based on pentaerythritol—an hypothesis. *Orig Life Evol Biosph* 23:185-194
- Schwartz AW (2007) Intractable mixtures and the origin of life. *Chem Biodiv* 4:656-664.
- Scorei R (2012) Is boron a prebiotic element? A mini-review of the essentiality of boron for the appearance of life on earth. *Orig Life Evol Biosph* 42:3-17
- Senma M, Taira Z, Osaki K, Taga T (1973) Glyceraldehyde: an X-ray study of D,L-glyceraldehyde dimer. *J Chem Soc Chem Commun* 880-881
- Sephton MA (2001) Meteoritics: life's sweet beginnings? *Nature* 414:857-858

- Sheldrick GM (2008) A short history of SHELX. *Acta Cryst A* 64:112-122
- Sheldrick GM (2015) SHELXT—Integrated space-group and crystal-structure determination. *Acta Cryst A* 71:3-8
- Šišak D, McCusker LB, Zandomenighi G, Meier BH, Bläser D, Boese R, Scheizer WB, Gilmour R, Dunitz, JD (2010) The crystal structure of D-ribose—at last! *Angew Chem Int Ed* 49:4503-4505
- Smith MB, March J (2001) March's advanced organic chemistry, 5th ed. John Wiley & Sons Inc, New York, pp 1180-1181
- Venuti MC (1982) 2,3-Dihydroxy-1,4-dioxane: a stable synthetic equivalent of anhydrous glyoxal. *Synthesis* 61-63
- Williams D, Viti S (2013) Modelling interstellar physics and chemistry: implications for surface and solid-state processes. *Phil Trans R Soc A* 371:20110587
- Yaylayan VA, Harty-Majors S, Ismail AA (1998) Investigation of the mechanism of dissociation of glycolaldehyde dimer (2,5-dihydroxy-1,4-dioxane) by FTIR spectroscopy. *Carbohydr Res* 309:31-38
- Zhao Y, Truhlar DG (2008) The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor Chem Acc* 120:215-241