

Interactions of Amino Acids and Aminoxazole Derivatives: Cocrystal Formation and Prebiotic Implications Enabled by Computational Analysis

Nieves Lavado¹  · Juan García de la Concepción¹ · Reyes Babiano¹ · Pedro Cintas¹ · Mark E. Light²

Received: 10 May 2019 / Accepted: 1 July 2019/

Published online: 20 July 2019

© Springer Nature B.V. 2019

Abstract

In line with the postulated intermediacy of aminoxazoles derived from small sugars toward the direct assembly of nucleoside precursors, we show here a potential prebiotic scenario where aminoxazolines might have also played further roles as complexing and/or sequestering agents of other primeval blocks, namely amino acids. To this end, a bis-aminoxazoline derivative, generated from dihydroxyacetone and cyanamide, gives rise to stable co-crystal forms with dicarboxylic amino acids (Asp and Glu), while ionic interactions owing to proton transfer are inferred from spectroscopic data in aqueous solution. The structure of a 1:2 aminoxazoline: aspartic acid complex, discussed in detail, was elucidated by X-ray diffractometry. Optimized geometries of such ionic structures with bulk aqueous solvation were assessed by DFT calculations, which disclose preferential arrangements that validate the experimental data. Peripherally, we were able to detect in a few cases amino acid dimerization (i.e. dipeptide formation) after prolonged incubation with the bis-aminoxazole derivative. A mechanistic simulation aided by computation provides some predictive conclusions for future explorations and catalytic design.

Keywords Prebiotic chemistry · Amino acids · Aminoxazole chemistry · Reaction mechanism · Dipeptide

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11084-019-09582-9>) contains supplementary material, which is available to authorized users.

✉ Nieves Lavado
nlavador@gmail.com

✉ Juan García de la Concepción
jugarco@unex.es

¹ Departamento de Química Orgánica e Inorgánica, Facultad de Ciencias-UEX, Avenida de Elvas s/n, E-06006 Badajoz, Spain

² Department of Chemistry, Faculty of Natural and Environmental Sciences, University of Southampton, Southampton SO17 1BJ, UK

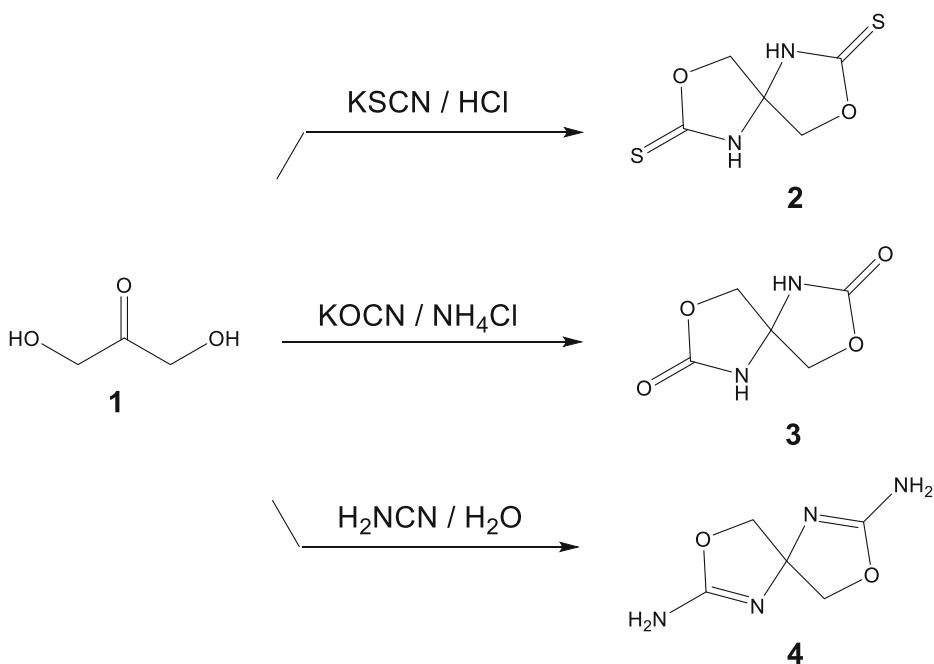
Introduction

Prebiotic chemistry deals with a vast repertoire of chemical reactions and physical processes aimed at elucidating the endogenous origin of essential building blocks and biomolecules on the primitive Earth, as well as their exogenous formation in space and subsequent delivery onto our planet by cometary or meteoritic impacts (Eschenmoser 2011; Ruiz-Mirazo et al. 2014). It is often assumed that the primeval soup consisted of a complex mixture where, in addition, several and disparate chemical routes coexisted and interacted each other. This inherent complexity, which often results in very low yields of individual components, contrasts with our conventional bench chemistry aimed at synthesizing a few compounds in high yields. The situation might not be so nightmarish whether some reactions and physical processes were able to favor selective transformations and sorting out over cascade and multistep scenarios (Schwartz 2007; Schwartz 2013; Budin and Szostak 2010; Pressman et al. 2015). Thus, *rac*-ribose, obtained in an extremely low yield (< 1%) from the messy formose reaction, would have instead been favored through the alkaline condensation of glycolaldehyde phosphate and formaldehyde as demonstrated by Eschenmoser et al. (Müller et al. 1990; Krishnamurthy et al. 1999). Higher sugars arose likely from C2 or C3 aldehydes via the formose reaction and, like ribose, their abundance would have been scarce.

A rather iconoclastic alternative to the RNA world, is Sutherland's aminoxazole-based chemistry by which a seemingly simple bottom-up construction of pyrimidine ribonucleotides can be achieved from purely prebiotic molecules like cyanamide, glycolaldehyde, glyceraldehyde, cyanoacetylene, and inorganic phosphate (Borsenberger et al. 2004; Anastasi et al. 2006; Powner et al. 2009; Powner et al. 2010; Powner et al. 2012; Ritson and Sutherland 2012). In subsequent work, Sutherland's group devised a proto-metabolic scenario where C2 and C3 sugars merge with HCN and its derivatives under photoredox catalysis (Patel et al. 2015). By focusing on dihydroxyacetone, which exists in equilibrium with glyceraldehyde (Yaylayan et al. 1999), extra products can be envisaged such as phosphorylated glycerols (precursors of lipid membranes) together with nucleosides, α -aminonitriles (precursors of amino acids) and nucleotides, thereby broadening the scope of synthetic prebiotic chemistry.

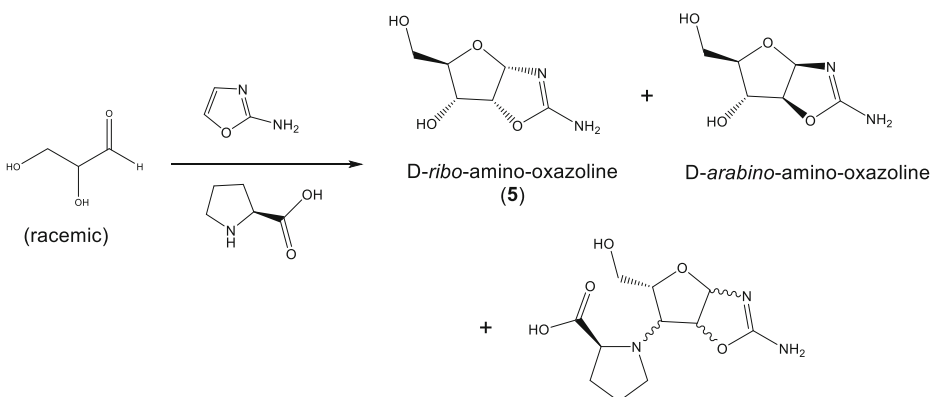
Simpler structural models might however throw light on non-covalent interactions and reactions involving different scaffolds. It is well established in carbohydrate chemistry that aldoses react with inorganic cyanates or thiocyanates affording glyco-oxazolidinones or glycol-oxazolidinethiones respectively, whereas ketoses often lead to cyclic and spiranic structures (Grouiller et al. 1988; Kovács et al. 1995). Thus, a prebiotic ketone, dihydroxyacetone (**1**), does react with KOCN, KSCN, or cyanamide giving rise to spiranic bicycles (Scheme 1) (Saul et al. 2000). The reaction with cyanamide is especially noticeable as the resulting bis(2-aminoxazoline), namely 2,7-diamino-1,6-dioxo-3,8-diazaspiro[4,4]nonan-2,7-diene (**4**), could also interact with other aldoses or primeval monomers, thereby fueling further chemistry. Compound **4** is thus structurally related to other aminoxazole derivatives like 2-amino-D-ribofuran[2,1-*d*]-2-oxazoline (**5**, Scheme 2), though less conformationally flexible than the latter. Moreover, and interestingly, such spiranes are chiral due to axial stereogenicity.

As mentioned, ribo-oxazoline **5** can likewise be obtained from D-ribose and cyanamide at basic pH according to previous protocols (Springsteen and Joyce 2004). This aminoxazoline derivative degrades more slowly than ribose itself (ca. 70-fold at pH ~10, $t_{1/2}$ > 1 week), thus being a masked form of the free sugar that could have boosted its stability and accumulation. In fact, the bicyclic adduct from ribose and cyanamide crystallizes spontaneously in aqueous solution, while the corresponding products derived from other hexoses and the rest of pentoses



Scheme 1 Condensation reactions of dihydroxyacetone with (thio)cyanates or cyanamide

do not (Springsteen and Joyce 2004). Compound **5** generated from enantiomerically pure D-ribose is obviously chiral. De novo synthesis from *rac*-glyceraldehyde and 2-aminoxazole, or alternatively from D- and L-ribose plus cyanamide, affords the racemic derivative, although the use of scalemic glyceraldehyde still enables enantioselective synthesis of **5** (Anastasi et al. 2006). In an elegant asymmetric variation, Blackmond and associates (Hein et al. 2011) showed that highly enantioenriched sugar aminoxazolines could be obtained from racemic starting materials with the handedness provided solely by a small enantiomeric imbalance of amino acids present in the reaction mixture (Scheme 2). Thus, a sample of L-proline (with an initial 1% *ee*) was added to a mixture of *rac*-glyceraldehyde and 2-aminoxazole and the



Scheme 2 Formation of enantioenriched RNA precursors from racemic glyceraldehyde in the presence of a chiral amino acid.

ensuing reaction afforded both D-ribo- and D-arabino-configured aminoxazolines in 20–80% *ee*. Cooling the mixture to 4 °C promoted crystallization of enantiopure D-ribo-aminoxazoline (**5**) crystals. The unnatural L-ribo-aminoxazoline was obtained similarly starting from a 1% *ee* of the D-proline enantiomer. Remarkably, L-glyceraldehyde could be sequestered by a side reaction joining L-proline and the aminoxazoline, (shown in Scheme 2), thereby enabling kinetic resolution of glyceraldehyde.

The present study shows the interactions of the above-mentioned aminoxazoline heterocycles, **4** in particular, with some proteinogenic amino acids, whose structures have been assessed by experimental and computational tools, which unravel less-explored mechanisms of accumulation and delivery of biomolecules in prebiotic niches (Schwartz 2007; Budin and Szostak 2010). Stable solid phases were obtained for dicarboxylic amino acids (Asp and Glu), and characterized in detail for the cocrystal of **4** and Asp. The serendipitous observation that amino acid dimerization took place in a few cases to a minor extent, prompted us to search for a mechanistic rationale for the role of such complexes as condensing agents.

Materials and Methods

All reagents were obtained from Aldrich-Sigma® and used as received without further purification. Aqueous solutions were prepared in bidistilled water. Compounds **4** (Saul et al. 2000) and **5** (Borsenberger et al. 2004) were prepared according to previously described methods.

IR spectra were recorded in the range of 4000–600 cm^{-1} on an FT-IR Thermo spectrophotometer using KBr (Merck) pellets. NMR spectra were recorded in D_2O solutions on a Bruker instrument operating at 500 MHz and 125 MHz for ^1H and ^{13}C nuclei, respectively. TMS was used as the internal standard ($\delta = 0.00$ ppm). Optical rotations were measured at 25 °C on a Perkin-Elmer 241 polarimeter at $\lambda = 589$ nm (D line, Na lamp), with all concentrations given in g/mL. Analytical microanalyses were determined on a Leco CHNS-932 analyzer.

Electrospray ionization mass spectra were analyzed using a 6520 Accurate-Mass Q-TOF LC/MS instrument with DUAL ESI interface (Agilent Technologies) coupled to an Agilent 1200 Liquid Chromatography (LC) apparatus. All samples were introduced in the system using injector and column switching valve directly to the interface. The mobile phase was milli-Q water containing 1.0% acetic acid with a flow at 0.3 mL/min. The acid imparts a positive charge on some functional groups; all samples were run in positive scan mode only. Mass spectra were recorded from 40 to 1000 m/z values in high-resolution mode (at 4 GHz ADC rate for data acquisition) and real-time processing on the mass peaks detected in each transient, weighting the apex data much more heavily than the shoulders. Conditions in the nebulizer were 350 °C (gas temperature), with a flow rate of 6 mL/min and 60 psi.

Complexation of 4 with DL-Asp (1:1) A mixture of bis(aminoxazoline) **4** (0.20 g, 1.28 mmol), DL-Asp (0.17 g, 1.28 mmol), and water (4 mL) was heated at ~50 °C until complete dissolution. Then, it was kept at 4 °C and MeOH was gradually added giving rise to a white precipitate. The first crop, obtained after addition of small volumes of MeOH (up to 5 mL), consisted essentially of DL-aspartic acid. The solid was filtered off and subsequent addition of MeOH (up to 18 mL) to the mother liquors afforded a crystalline solid (6, 0.098 g, 37%), mp = 155 °C (dec.), whose spectroscopic and X-ray diffraction data agree with formation of the title

complex (**6**) as a dihydrate. Anal. calcd for $C_{13}H_{22}N_6O_{10} \cdot 2H_2O$: C, 34.06; H, 5.72; N, 18.33; found: C, 33.77; H, 5.40; N, 18.30.

Complexation of 4 with L-Glu (1:2) A mixture of **4** (0.20 g, 1.28 mmol), L-Glu (0.38 g, 2.56 mmol), and water (8 mL) was heated at ~ 50 °C. The resulting solution was kept at 4 °C with successive additions of MeOH (small portions until a final volume of 18 mL), thus leading to the formation of a white solid (**7**, 0.38 g, 66%), mp = 191 °C, $[\alpha]_D = +2.5$ ($c = 0.008$, H_2O). Anal. calcd for $C_{15}H_{26}N_6O_{10} \cdot 2H_2O$: C, 37.04; H, 6.22; N, 17.28; found: C, 37.40; H, 5.90; N, 17.20.

Complexation of 4 with L-Glu (1:1) A mixture of **4** (0.20 g, 1.28 mmol), L-Glu (0.38 g, 2.56 mmol), and water (8 mL) was heated at ~ 50 °C. The resulting solution was kept at ca. 4 °C with successive additions of EtOH (small portions up to 16 mL), which afforded a white solid (**8**, 0.38 g, 98%), mp = 195 °C, $[\alpha]_D = +3.0$ ($c = 0.007$, H_2O). Anal. calcd for $C_{10}H_{17}N_5O_6 \cdot H_2O$: C, 37.38; H, 5.96; N, 21.80; found: C, 36.95; H, 5.90; N, 21.50.

Complexation of 4 with DL-Asp and DL-Glu (1:2:1) A mixture of **4** (0.20 g, 1.28 mmol), DL-Asp (0.17 g, 1.28 mmol), DL-Glu (0.19 g, 1.28 mmol), and water (6 mL) was heated until complete dissolution. Then, it was cooled (4 °C) and MeOH was added (4 mL) giving a white solid (**9**, 0.20 g, 54%), mp = 149 °C (dec.). Anal. calcd for $C_{18}H_{31}N_7O_{14} \cdot 3H_2O$: C, 34.67; H, 5.98; N, 15.72; found: C, 34.61; H, 5.54; N, 15.55.

Complexation of 5 with L-Asp (1:1) A mixture of **5** (0.20 g, 1.15 mmol), L-Asp (0.15 g, 1.15 mmol) and water (4 mL) was heated (~ 50 °C) until dissolution. It was then cooled at ca. 4 °C adding dropwise MeOH (up to 8 mL), which afforded an initial crop of L-aspartic acid (0.023 g, 0.17 mmol). After removal by filtration and subsequent addition of MeOH (14 mL), the second crop contained compound **10** (0.029 g, 12%), mp = 183 °C (dec.), $[\alpha]_D = +18.8$ ($c = 0.005$, H_2O). Anal. calcd for $C_{10}H_{17}N_3O_8 \cdot H_2O$: C, 36.93; H, 5.89; N, 12.92; found: C, 36.63; H, 5.40; N, 12.70.

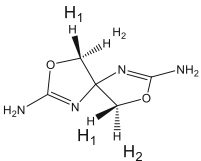
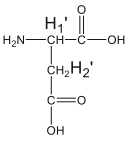
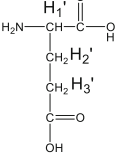
Dimerization of Amino Acids in the Presence of Compound 4 A mixture of **4** (0.25 g, 1.6 mmol) and the corresponding amino acid (8.0 mmol) was dissolved in water (20 mL) and heated at ~ 70 °C for approximately one month (670 h). Aliquots were regularly extracted (ranging from 24 h to one week), diluted to a final concentration of 10 ppm, and then subjected to ESI-MS analysis following the procedure described in the Electronic Supplementary Material (ESM, Tables S5 and S6).

Results and Discussion

Formation and Structural Features of Zwitterionic Complexes

As pointed out above the preparation of **4** can easily be conducted in water at room temperature from dihydroxyacetone and cyanamide (Saul et al. 2002). The resulting solid exhibits two methylene proton signals at δ 4.36 and 4.26 ppm, whose diastereotopicity is consistent with a chiral structure. Unfortunately, it crystallizes as racemic compound, not conglomeratic phase, which impedes spontaneous resolution by crystallization. Diastereomeric

Table 1 Proton resonances (δ) and proton shift variations ($\Delta\delta$) for compound **4**, Asp and Glu, and the resulting zwitterions in D₂O

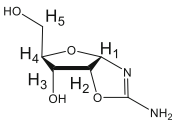
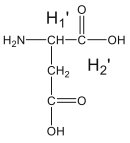
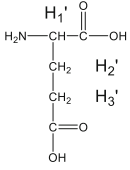
Comp.							
	H ₁	H ₂	H ₁ '	H ₂ ' _a	H ₂ ' _b	H ₃ ' _a	H ₃ ' _b
4	4.36	4.26					
DL-Asp			4.06	3.02	2.94		
1:1^a	4.71	4.66	3.89	2.81	2.67		
$\Delta\delta$	0.35	0.40	-0.17	-0.21	-0.27		
L-Asp			4.06	3.02	2.96		
1:1 Mix^a	4.72	4.67	3.89	2.81	2.67		
$\Delta\delta$	0.36	0.41	-0.17	-0.21	-0.29		
DL-Glu			3.80	2.14	2.14	2.54	2.54
1:1 Mix^a	4.66	4.61	3.74	2.12	2.05	2.34	2.34
$\Delta\delta$	0.30	0.35	-0.06	-0.02	-0.09	-0.20	-0.20
L-Glu			3.77	2.12	2.12	2.52	2.52
1:1 Mix^a	4.72	4.68	3.75	2.12	2.06	2.37	2.36
$\Delta\delta$	0.36	0.42	-0.02	0.00	-0.06	-0.15	-0.16

^a Oxazoline:amino acid ratio determined by proton integration (at 500 MHz)

resolution also failed for this spiranic substance. Homochiral **5** was prepared from D-ribose and cyanamide. We observe that, contrary to previous literature reporting its synthesis at basic pH (Borsenberger et al. 2004; Springsteen and Joyce 2004) crystallization of **5** also occurs under essentially neutral conditions (pH ~7.0) and mild heating (ca. 40 °C) as well.

In order to investigate whether **4** does actually interact with amino acids and exert a discriminating effect against the enantiomeric pair, 1:1 mixtures of both substrates in D₂O solutions (0.06 M each) were monitored by ¹H NMR. For comparison, both racemates (DL) and optically active (L-configured) amino acids were evaluated. In some cases the low aqueous solubility of the amino acid or parent heterocycle led to mixtures deviating from an equimolar ratio (molar relationships collected in Tables 1 and 2 were measured by integration). Chemical shift variations ($\Delta\delta$) refer to differences before and after interactions for a given proton signal: CH α (H₁') for the amino acid as well as the diastereotopic methylene protons H₁ and H₂ of aminoxazoline **4** and the bridged H₁ and H₂ signals of **5**. For most amino acids (Arg, Phe, Pro, Ala, Gly, Ser, Gln, Lys, His, Thr, or Asn), $\Delta\delta$ values were almost negligible (from 0.01 to 0.07 ppm for all proton resonances). However, dicarboxylic amino acids, such as Asp and Glu, showed significant variations for the oxazoline protons (from 0.22 to 0.47 ppm) and to a lesser extent for H₁' protons (see ESM). These observations point to interactions of the oxazoline moiety with the amino acid side chain, where the isoureido linkage of the former undergoes protonation, deshielding the vicinal hydrogen atoms, whereas the negative charge of the carboxylate group causes further shielding on the adjacent H-atoms.

Table 2 Proton resonances (δ) and proton shift variations ($\Delta\delta$) for compound **5**, amino acids (Asp, Glu), and the resulting zwitterions in D₂O

Comp.										
	H ₁	H ₂	H ₃	H ₄	H _{5a}	H _{5b}	H _{1'}	H _{2'a}	H _{2'b}	H _{3'}
5	5.79	4.98	4.12	3.61	3.91	3.72				
DL-Asp							4.06	3.02	2.96	
1:1 Mix^a	6.05	5.45	4.29	3.87	3.92	3.74	3.95	2.84	2.73	
$\Delta\delta$	0.26	0.47	0.17	0.26	0.01	0.02	-0.11	-0.18	-0.23	
L-Asp							4.06	3.02	2.96	
1:0.78 Mix^a	6.05	5.45	4.29	3.87	3.95	3.74	3.90	2.82	2.68	
$\Delta\delta$	0.22	0.47	0.17	0.26	0.04	0.02	-0.16	-0.20	-0.28	
DL-Glu							3.80	2.14	2.14	2.55
1:1 Mix^a	6.05	5.45	4.29	3.74	3.95	3.86	3.76	2.13	2.07	2.38
$\Delta\delta$	0.26	0.47	0.17	0.13	0.04	0.14	-0.04	-0.01	-0.07	-0.17
L-Glu							3.77	2.12	2.12	2.52
1:0.78 Mix^a	6.05	5.44	4.29	3.86	3.95	3.76	3.74	2.13	2.07	2.37
$\Delta\delta$	0.26	0.46	0.17	0.25	0.04	0.04	-0.03	0.01	-0.05	-0.15

^a Oxazoline:amino acid ratio determined by proton integration (at 500 MHz)

It is obvious that, in strict sense, the term *complexation* cannot be claimed in the present context, as the aforementioned results are clearly consistent with proton transfer from either the aspartic or glutamic acid side chain to the spiranic fragment of **4**. Such proton transfers could explain the expected effect on the chemical shifts of both substrates, amino acid and heterocycle. Accordingly, they may be solvated in the aqueous environment, thereby preventing complex formation per se. This could therefore be achieved by interactions that overcome the hydration of the binding partners or, alternatively by harnessing solvophobic effects, as demonstrated in anion recognition by peptide-like structures (Kubik 2017). In any case, the generation of zwitterionic aggregates triggered by simple proton transfer would have enabled the stabilization and accumulation of sensitive reagents. This proton transfer will also be noticeable if that interaction between acid and basic groups in reaction partners primarily drives the formation of cocrystals, thereby enhancing the separation process, as we shall discuss later.

Isolation and Characterization of Solid-State Structures

Crystallization screenings were performed for mixtures of **4** with Asp and Glu. The protocol was extended to other bifunctional amino acids like asparagine (Asn), though without success. Both reaction partners (in 1:1 and 1:2 M ratios) were dissolved in water at ca. 50 °C. Since no crystallization was observed, dropwise addition of MeOH or EtOH, caused the precipitation of

white solids. Unfortunately most trials were unsuccessful, leading to the selective precipitation of one of the starting materials. Gratifyingly, DL-Asp and L-Glu behaved differently and showed conversion to solid phases consistent with their cocrystallization with **4** (details included in the ESM, Table S3).

In addition to spectroscopic characterization, the complex obtained from **4** and DL-Asp could be unambiguously elucidated by single-crystal XRD. The solid **6** decomposes at ~ 155 °C and had elemental analysis that agrees with the formation of a hydrate derivative ($C_{13}H_{22}N_6O_{10} \cdot 2H_2O$), thus pointing to a 1:2:2 (**4**:DL-Asp:H₂O) stoichiometry. As determined by ¹H NMR analysis in D₂O, the 1:2 relationship between **4** and Asp is constant irrespective of the initial ratio of both substrates. This suggests that each amino group of **4** participates in hydrogen bonding with the carboxylate group of two different amino acid molecules. This would also support the involvement of the adjacent ring nitrogen in bonding, thus reinforcing the chelate structure (Fig. 1).

Water uptake can be inferred from the solid-state FT-IR spectrum, which shows broad and intense bands between 3500 and 3000 cm⁻¹ (Fig. S15). As mentioned, suitable crystals for X-ray diffraction allowed us to confirm the 3D-arrangement of reaction partners and water molecules in the unit cell, in agreement with the above surmises (Fig. 2). On the basis of a racemic amino acid, we also expected to see a centrosymmetric space group, which turned to be true as complex **6** crystallizes in the monoclinic *C2/c* group (*Z* = 4), where amino acid residues of opposite handedness coexist within the unit cell. This structure is similar to that of multicomponent crystals of some amino acids, namely L-Trp, with *N*-heterocyclic carboxylic acids, for which double zwitterions have been recently reported with change in the absolute configuration during cocrystallization (Das and Srivastava 2017). The term zwitterionic cocrystal is usually applied to systems consisting of a zwitterionic compound and a neutral molecule (Aakeröy et al. 2008), which show common features of supramolecular organization and hydrogen bonding (vide infra). Numerous zwitterionic cocrystals involve amino acids as the resulting solids are expected to have enhanced aqueous solubility, a property harnessed in the preparation of pharmaceutical cocrystals for oral delivery (Tilborg et al. 2014; Surov et al. 2018). Likewise, the spirane-amino acid structure **6** resembles that of guanidinium hosts with carboxylate anions (Berger and Schmidtchen 1999; Blondeau et al. 2007), though as noted above, their complexation was reported in aprotic media.

The asymmetric unit of compound **6** contains one Asp anion, half a dioxadiaza-spirane di-cation (on a 2-fold rotation axis) and one water molecule. Asp is in its 'straight' conformation with the torsion angle C4-C5-C6-C7 having a value of $-174.46(10)^\circ$ (the alternative bent conformation has torsion angles in the range ca. 50 – 70°). The carboxylate groups are in their resonance form C4-O2 = 1.2726(16) Å, C4-O3 = 1.2441(15) Å, a significant difference of 0.0285(22) Å and C7-O4 =

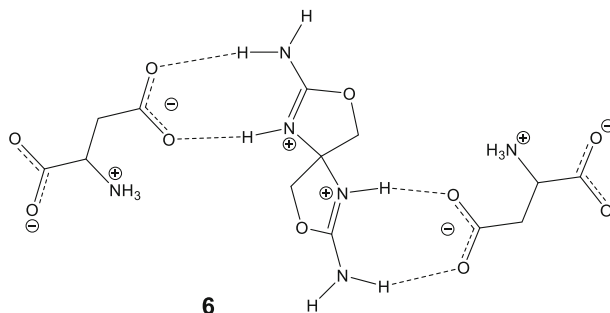


Fig. 1 Putative structure showing the complexation of **4** with DL-Asp. Water molecules could also be involved in H-bonding (not shown)

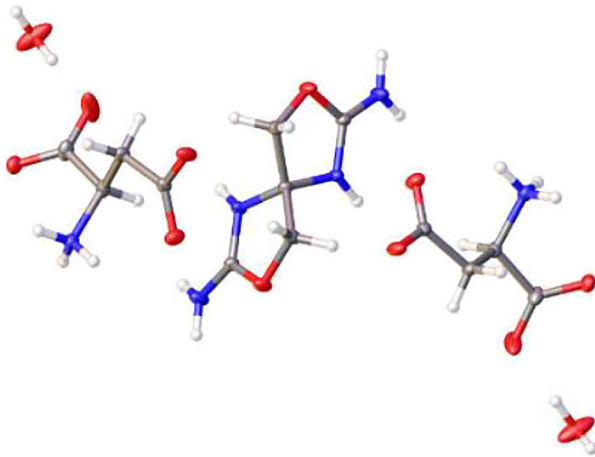


Fig. 2 X-Ray structure of complex **6** as a dihydrate, generated from **4** and DL-Asp. Ellipsoids are drawn at 50% probability

1.2454(18) Å, C7-O5 = 1.2513(16) Å, which are the same within error. Each five-membered ring of the spirane is in a twisted confirmation with Cremer-Pople puckering parameters of $Q(2) = 0.1760(12)$ Å and $\Phi(2) = 349.2(4)^\circ$ (Cremer and Pople 1975). The angle between the least squares planes is $87.08(8)^\circ$. The supramolecular structure is dominated by hydrogen bonding (all heteroatom hydrogens were clearly located in the difference map and freely refined) ultimately forming a 3D network. Figure 3 shows the unique hydrogen bond interactions with associated atom numbering.

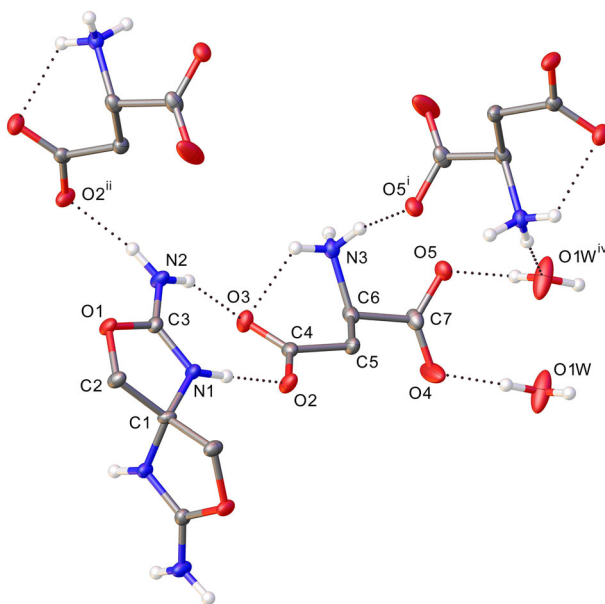


Fig. 3 Unique hydrogen bond interactions. Asymmetric unit and selected symmetry equivalent positions labelled, thermal ellipsoids drawn at the 50% probability level. Symmetry codes: (i) $3/2-x, -1/2 + y, 3/2-z$; (ii) $+x, -y, 1/2 + z$; (iv) $3/2-x, 3/2-y, 1-z$

In the asymmetric unit there are five acceptors (carboxylate groups and the water) and four donors with eight hydrogens (N-H, N-H2, N-H3 and H₂O), all are involved in hydrogen bond interactions. Taking the spirane as a central motif (Figs. 4 and 5) there are strong interactions to one of the carboxylate groups of the aspartate anion on either side described with the well-established graph-set terminology (Etter et al. 1990) as R2,2(8). The carboxylate at the other end of the anion makes hydrogen bonds to two water molecules and through a centre of symmetry forms a ring, R4,4,(12), with a symmetry related anion. A further ring motif, R4,4(24), is formed linking two spirane molecules through O2 of the carboxylate. The propagation, through the symmetry of the space-group, of these hydrogen bonded rings complete the 3D network. The corresponding symmetry codes (Hall and McMahon 2006) are included in Figs. 3 and 4.

For comparative purposes with Tables 1 and 2 where proton shift variations were measured for 1:1 mixtures, data gathered in Table 3 show ¹H NMR shifts in D₂O for **6** possessing a 1:2 stoichiometry (Fig. S16). The H₁ and H₂ signals of spiranic aminoxazoline exhibit larger $\Delta\delta$ variations than those of Tables 1 and 2. Alternatively, as expected for an increase in amino acid concentration, its proton signals are likewise less shielded in complex **6**. As emphasized earlier, the ionic species arising from H-transfer may be perfectly solvated in the aqueous solution.

Under similar experimental conditions, L-Asp failed to afford a cocrystal. Notably, the situation was exactly permuted for glutamic acid as a distinctive solid phase could only be obtained for the L-enantiomer. Starting from 1:2 (**4**:L-Glu) ratios, two different solids were isolated depending on the antisolvent employed for precipitation, either MeOH or EtOH. While solid-state FT-IR spectra were essentially identical for both complexes (Figs. S17 and S18), ¹H NMR spectra recorded in D₂O revealed two different stoichiometries (Figs. S19 and S20). Complex **7**, isolated by adding MeOH as precipitating agent, showed signals at $\delta = 4.34$ (H₂ of the spiranic unit, two equivalent protons) and $\delta = 3.73$ ppm in ca. 1:2 ratio, consistent with the formation of a 1:2 (**4**:L-Glu) complex, whereas integration of the same signals in complex **8**, isolated by adding EtOH as antisolvent, gave rise to ~1:1 relationship. Again, Fig. 6 displays putative structures for such complexes (see ESM: Table S1 for H-shift

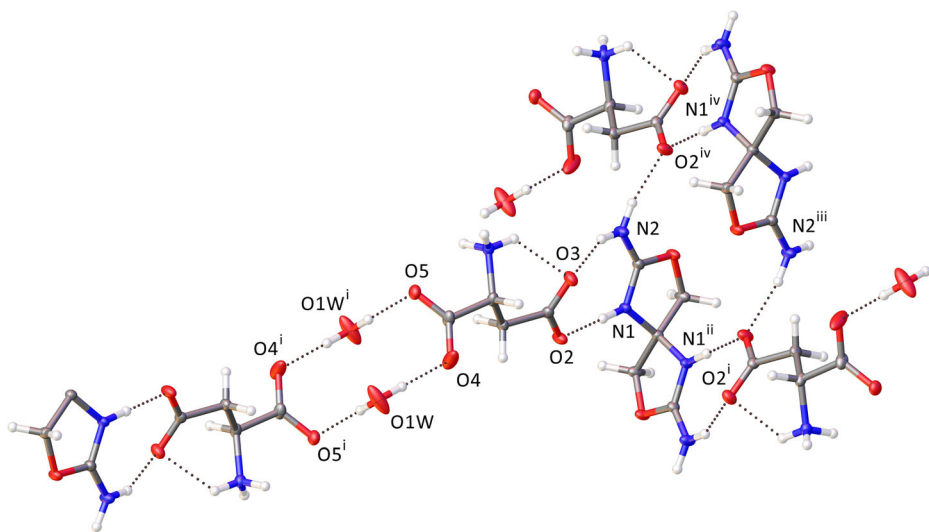


Fig. 4 The R4,4,(12), R2,2(8) and R4,4(24) ring motifs; symmetry codes: (i) $3/2-x, 3/2-y, 1-z$; (ii) $+x, -y, 1/2+z$; (iii) $1-x, 1-y, 2-z$; (iv) $1-x, +y, 3/2-z$

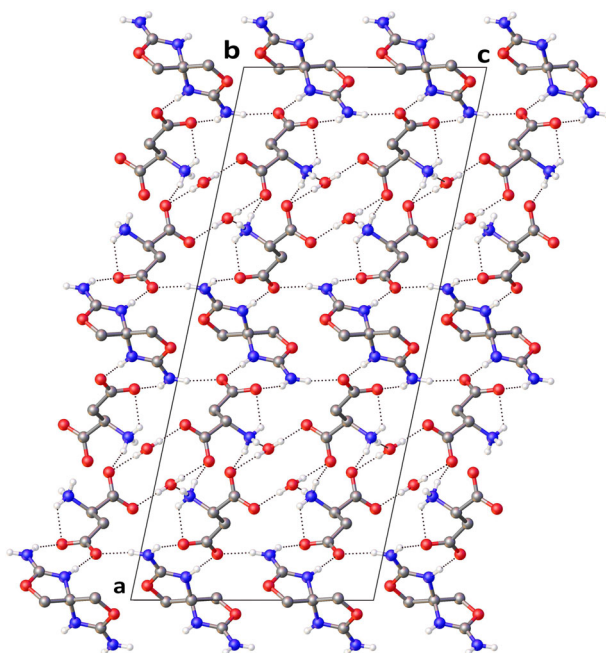


Fig. 5 View of the hydrogen bonding network of **6** along the *b*-axis

variations observed for **7** and **8** in D_2O). Only minor differences in chemical shifts could be observed and, as noted above for *rac*-Asp, the most significant changes are detected for the methylene group (H_3 signal) at the side chain, thus reflecting the interaction of that carboxylate group, rather than that of $C\alpha$, with the heterocyclic moiety. Elemental analyses measured for **7** and **8** were consistent with the formation of dihydrate and monohydrate phases, respectively. They had in addition different, yet close, melting points and optical rotations (see experimental procedures).

Given the bidentate character of **4**, we wondered if it might be possible to form the aminoxazoline derivative to two amino acids of mixed handedness. Accordingly, the assembly of **4** with DL-Asp and DL-Glu was investigated using equimolar amounts. Addition of MeOH gave a white precipitate whose spectral features were different from those of the starting materials (see ESM). The 1H NMR spectrum of the resulting complex (**9**) in D_2O shows clearly the coexistence of the three species (Fig. S22). Surprisingly, signal integrations unravels a ratio between the aminoxazoline and both amino acids that mismatches the initial stoichiometry, as the amount of Asp approximately doubles that of the others, giving rise to a

Table 3 Proton resonances (δ) and proton shift variations ($\Delta\delta$) for complex **6** (1:2 spirane:amino acid ratio) relative to **4** and DL-Asp in D_2O solution

Comp-	H_1	H_2	$H_{1'}$	$H_{2'a}$	$H_{2'b}$
6	4.91	4.85	3.92	2.85	2.73
4	4.37	4.26			
DL-Asp			4.06	3.02	2.96
$\Delta\delta$	0.54	0.59	-0.14	-0.17	-0.23

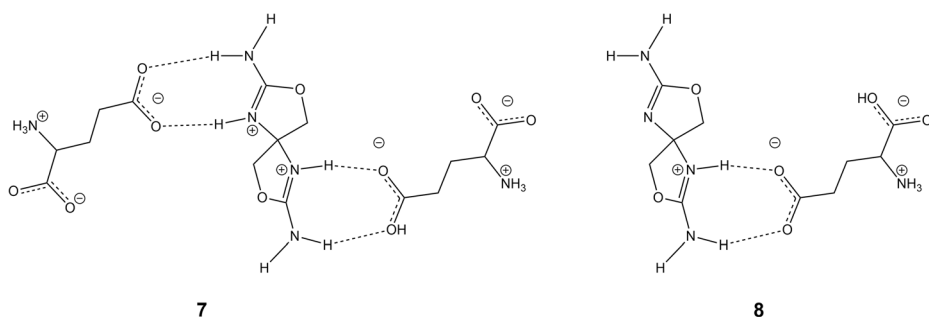


Fig. 6 Possible complexes **7** and **8** showing the coordination of **4** with L-Glu through carboxylate groups of the side chain. Water molecules could also be involved in H-bonding

1:2:1 (**4**:DL-Asp:DL-Glu) complex. No crystals suitable for X-ray diffraction could be obtained. Elemental analysis agreed with formation of a trihydrate, where water molecules might also stabilize a cage-like suprastructure through multiple hydrogen bonds. The solid-state FT-IR spectrum (Fig. S21) showed broad and intense peaks between 3500 and 3000 cm^{-1} accounting for the existence of bound OH bonds. Proton shifts collected in Table S2 (ESM) unveil once again the most salient interactions involving **4** with amino acids. The spiranic protons undergo a strong deshielding of ~ 0.6 ppm, which counterbalances the shielding effect on the amino acid units, especially for aspartic acid. The effect is also more pronounced on the hydrogen atoms of the side chain, thus pointing to prevalent coordination of the heterocycle with a terminal carboxylate group. Figure 7 shows a scheme accounting for the triple interaction of **4** with two molecules of Asp and one fragment of Glu.

Unlike **4**, compound **5** did not complex amino acids readily. Numerous failed attempts with Asp, Glu, and Asn led to precipitation of one or both parent materials. The sole exception was L-Asp, which coordinated with **5** in equimolar amounts affording a white solid in 12% yield from aqueous solution after successive additions of MeOH as antisolvent.

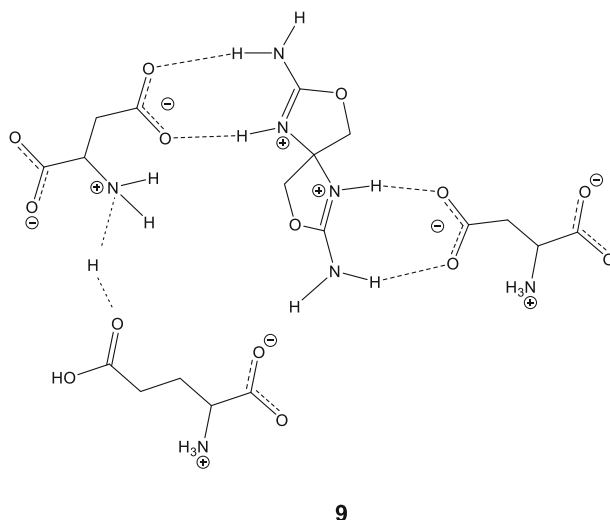


Fig. 7 Schematic proposal for the structure of complex **9** generated from aminoxazoline **4** with aspartic and glutamic acids. Water molecules that should be present in the coordinating sphere have been omitted for clarity

^1H NMR peak integration suggests formation of a 1:1 complex for **5** and L-Asp (Fig. S24). As expected the most deshielded signals were the bridged protons of the *ribo*-oxazoline bicycle ($\Delta\delta = 0.44$ and 0.23 ppm), whereas minor shielding was observed for H_α ($\Delta\delta = -0.15$ ppm) and diastereotopic CH_2 protons ($\Delta\delta = -0.18$ and -0.25 ppm) of the side chain. It seems reasonable to assume that the resulting complex (**10**) adopts a similar interaction to that of preceding solids involving the terminal carboxylate group and the aminoxazoline skeleton (Fig. 8).

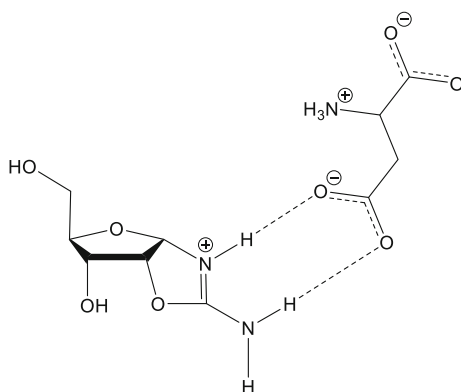
Geometry and Stability: Computation-Aided Models

To further assess the complexation of amino acids with **4**, the energy-minimized structures of some complexes have been calculated by DFT calculations at the M06-2X/6-31+G(d) level (Zhao and Truhlar 2008), (Ditchfield et al. 1971; Hariharan and Pople 1973; Gordon 1980; Francl et al. 1982) and modelling solvent effects (in water) by means of the continuum SMD model, (Marenich et al. 2009a; Marenich et al. 2009b; Ribeiro et al. 2010; Halim et al. 2010; Saielli 2010), as implemented in the Gaussian suite of programs (Frisch et al. 2009).

Both 1:1 and 1:2 aminoxazoline (**4**):amino acid complexes were computed, taking also into account for bifunctional amino acids the two possible coordination modes, i.e. either involving the carboxylate group adjacent to C_α (mode *a*) or the carboxylate group at the side chain (mode *b*) to ascertain the influence of the latter. Figures 9 and 10 show the resulting optimized structures for such complexes. Energy data are shown in Tables 4 and 5.

Although enthalpies for 1:1 complexes are thermodynamically favorable, the Gibbs energies are invariably positive as a result of the decreasing entropy upon complexation. Those values decrease (up to ca. $3 \text{ kcal}\cdot\text{mol}^{-1}$) for Asp and Glu acids which exhibit an intramolecular hydrogen bond. That interaction would also mimic coordination to water molecules (bearing in mind that the SMD method does not include explicit interaction with discrete water molecules, but merely simulates the solvation cage). For Glu, calculations predict a greater stability of the *b*-mode involving coordination to the side chain, the latter participating in non-covalent bonding with the aminoxazoline moiety. This result clearly mirrors the structural arrangement observed in the crystalline phase. Likewise, free energies are unfavorable for 1:2 complexes.

Fig. 8 Possible interactions of the complex of L-Asp with **5**



10

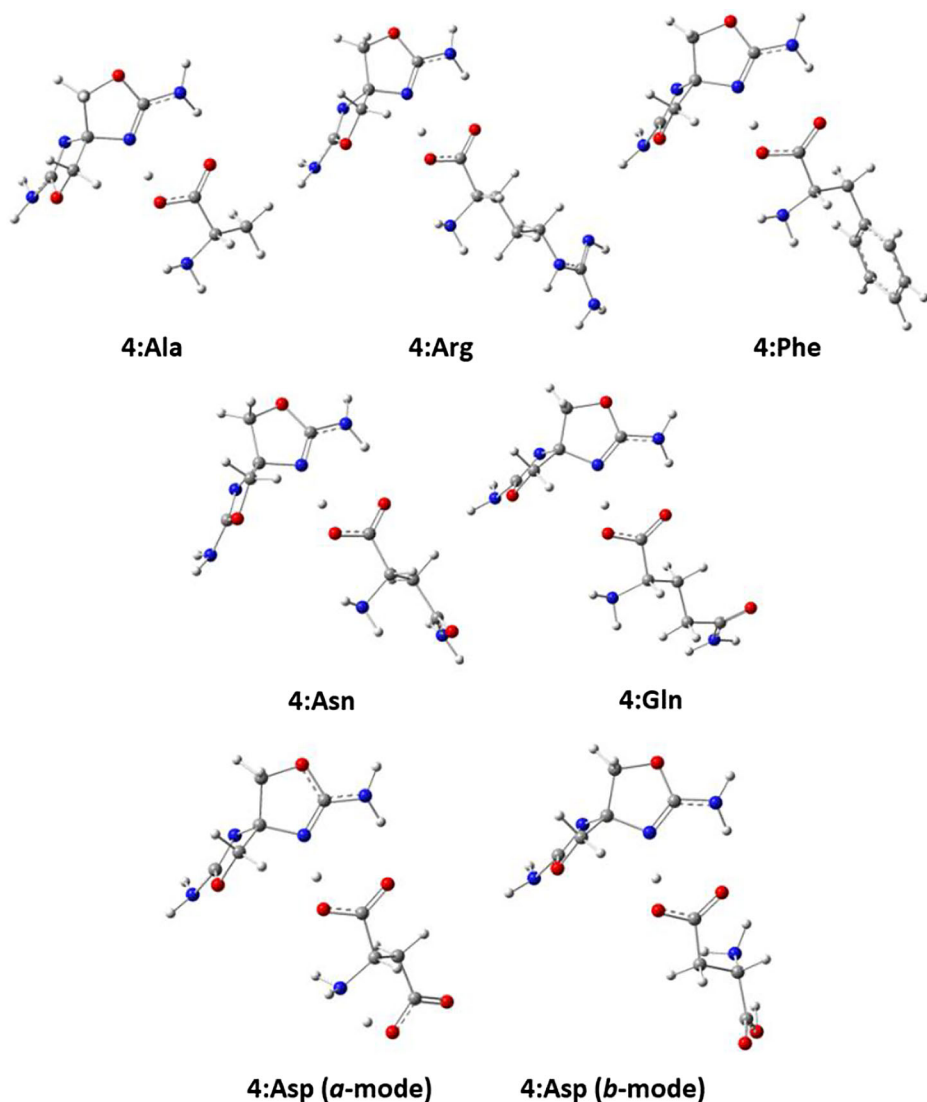


Fig. 9 Energy-minimized structures obtained for 1:1 (**4**:amino acid) complexes at the M06-2X/6-31 + G(d) level of theory, including solvent effects in water (SMD)

On the Catalytic Role of Aminoxazolines: A Computational Screening

When samples of Gly, Ala, or rac-Phe were left in neutral aqueous solution in the presence of **4** with prolonged heating for one week, and the mixture was then checked by mass spectra (MS) analysis (see ESM for details), peaks with m/z data consistent with linear or cyclic dipeptides (diketopiperazines, DKP) emerged. The *accidental discovery* prompted us to investigate theoretically how feasible that dimerization is.

The spiro compound **4** adopts a relatively rigid bowl-shaped structure whose nitrogen atoms are capable of interacting with amino acid carboxyl groups and may bring about a subsequent dimerization reaction. Compound **4**, obtained from dihydroxyacetone and

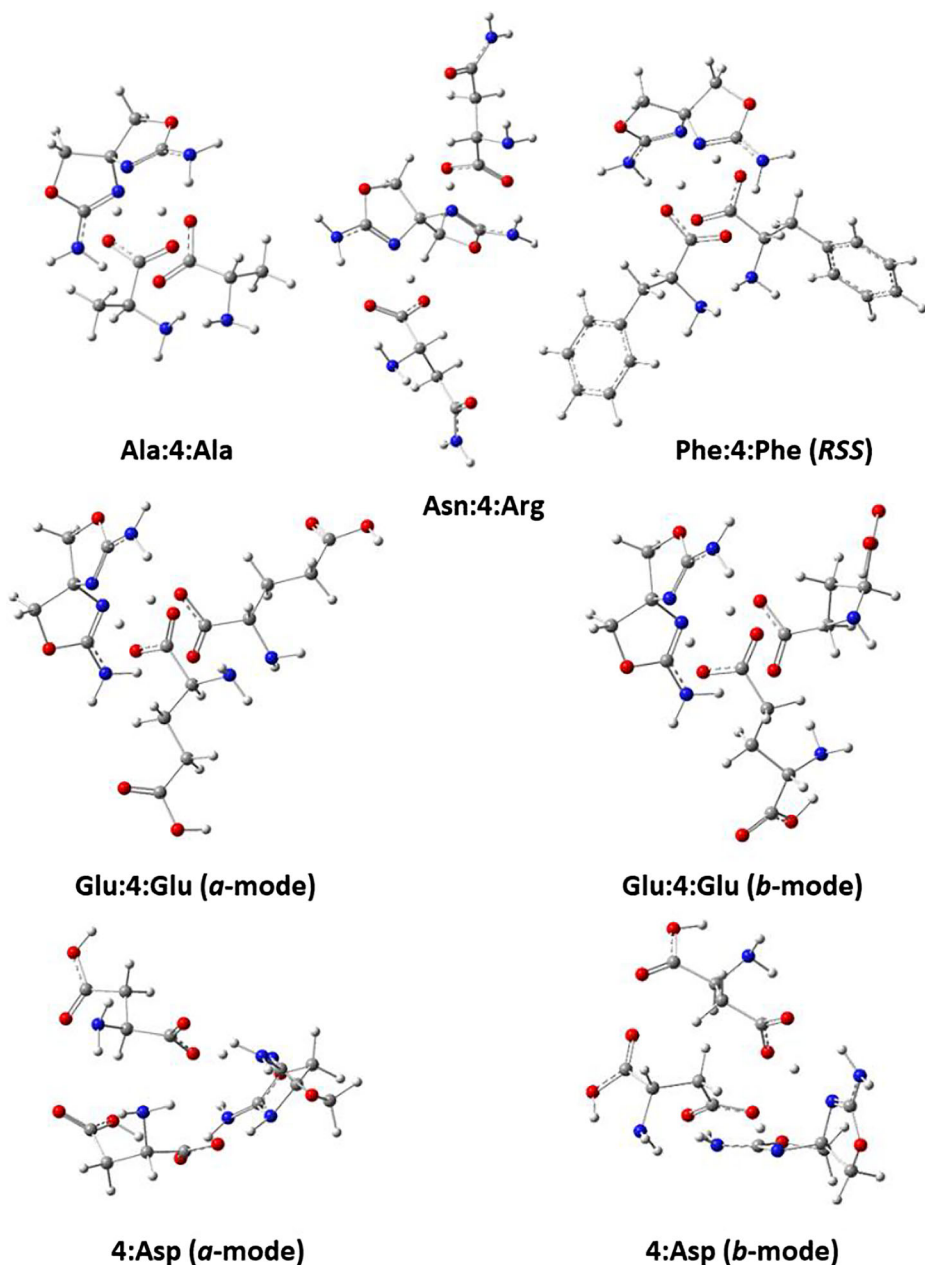


Fig. 10 Energy-minimized structures obtained for 1:2 (4:amino acid) complexes at the M06-2X/6-31 + G(d) level of theory, including solvent effects in water (SMD)

cyanamide, can be regarded as a masked surrogate of the latter. The condensing effect of cyanamide favoring the formation of dipeptides from a few amino acids was reported by Hawker and Oró in the early 1980s (Hawker and Oró 1981). Previously, Calvin and coworkers had suggested the use of cyanamide as activating/condensing agent for other biomolecules, although the group was unable to achieve the reaction of adenine and ribose in the presence of

Table 4 Stabilization energies (in kcal·mol⁻¹) for 1:1 aminoxazoline (**4**):amino acid complexes

Amino acid	ΔE	ΔH	ΔG
Ala	-0.21	-1.27	9.26
Arg	-0.52	-1.55	9.90
Asn	-0.29	-1.29	9.67
Asp (<i>a</i> -mode)	-4.91	-6.30	6.29
Asp (<i>b</i> -mode)	-3.47	-4.12	7.90
Gln	-0.77	-1.64	9.52
Glu (<i>a</i> -mode)	-0.96	-1.82	9.87
Glu (<i>b</i> -mode)	-3.66	-4.63	6.53
Phe	-0.74	-1.47	10.94

cyanamide (Steinman et al. 1964). Recently, dipeptides and diketopiperazines were detected during the incubation of amino acid solutions with cyanamide or dicyanamide in acid (Parker et al. 2014). The prebiotic generation of peptides challenging and numerous strategies thought to be prebiotically plausible have been postulated (Plasson et al. 2004; Pascal et al. 2005; Danger et al. 2012; Weissbuch et al. 2009; Weissbuch et al. 2011; Rode and Schwendinger 1990; Rode 1999; Reiner et al. 2006).

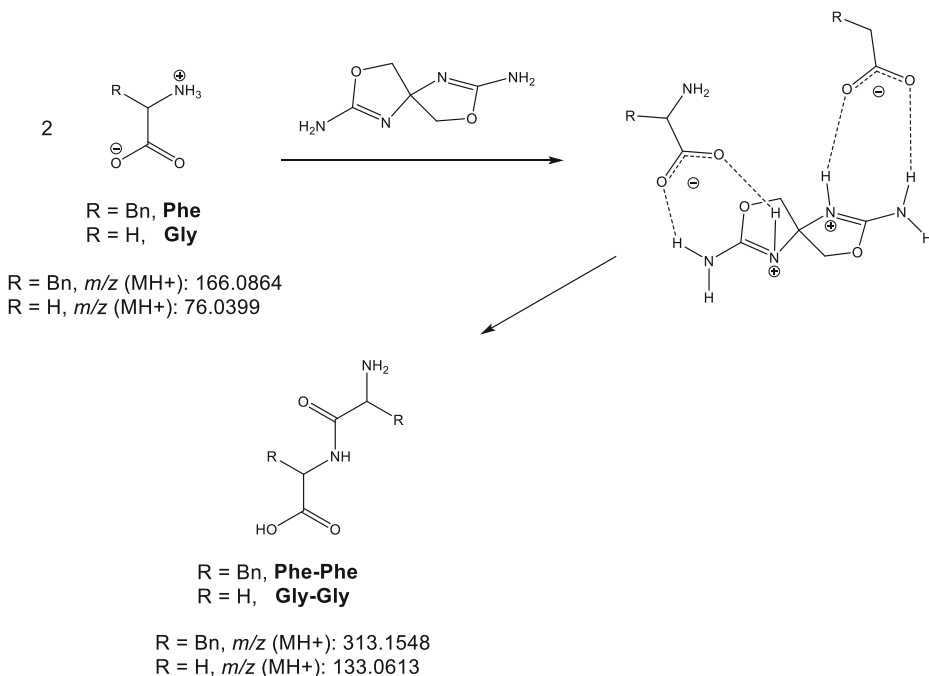
Scheme 3 shows the dimerization step which, if promoted by **4**, should occur via a ternary complex stabilized by electrostatic interactions and leading ultimately to dipeptides after intermolecular attack by another amino acid unit. Mass spectra revealed the formation of low-intensity peaks (< 2%), visible after ~100-h incubations, whose mass-to-charge values correspond to Gly-Gly and Phe-Phe. For Ala samples, peaks of higher intensity (up to 6%), consistent with formation of Ala-Ala were detected after ca. 330 h. Remarkably, those peaks could not be identified in the absence of **4**. To further corroborate the formation of the dipeptide, mass spectra were collected for reaction aliquots along with genuine samples of Phe and Phe-Phe at different voltages (60, 80 and 120 V) giving rise to $\Delta m/z$ less than 0.01 in all cases among the peaks in question. The intensity of Phe-Phe peaks increased as the reaction progressed, especially for samples measured at 60 V. After prolonged heating (ca. 650 h), a peak of low intensity and $m/z = 295.14$ Da $[M-H]^+$ emerges, which can be attributed to the corresponding DKP. That peak was not observed in the absence of **4**. In the cases of Gly-Gly and Ala-Ala, such dipeptides were better observed through ESI-MS experiments at 80 V.

Such results are rather disappointing in synthetic terms, even though primeval peptides would have likely been generated in tiny amounts by multiple mechanisms. We wondered whether amino acids other than Gly, Ala, and Phe, especially Asp and Glu, would eventually

Table 5 Stabilization energies (in kcal·mol⁻¹) for 1:2 aminoxazoline (**4**):amino acid complexes

Amino acid	ΔE	ΔH	ΔG
Ala	-6.66	-7.94	16.80
Asn	-1.49	-2.42	19.70
Asp (<i>a</i> -mode)	-12.41	-12.44	14.19
Asp (<i>b</i> -mode)	-11.11	-11.87	13.65
Glu (<i>a</i> -mode)	-4.78	-5.54	19.30
Glu (<i>b</i> -mode)	-12.24	-13.59	12.63
Phe (<i>RSS</i>) ^a	-8.45	-9.17	17.71

^aFor this amino acid, thermodynamic data of its complex refer to the (*RSS*)-configured diastereomer, which shows the highest stability (See next section for a computational rationale)



Scheme 3 Reaction scheme for dipeptide formation from Gly and Phe in the presence of **4** in water, via the intermediacy of a complex assembled through electrostatic interactions

produce peptides in the presence of **4**, which were unsuccessful. For Glu, peaks with $m/z = 259.0930$ and 277.1036 Da most likely correspond to dimeric structures. The peak at $m/z = 259$ is consistent with the DKP derivative, but also with a hybrid acyclic-cyclic structure, whereas the peak at $m/z = 277$ may include dipeptides involving the two carboxyl groups. As mentioned, an *in-silico* analysis could be instrumental in elucidating the factors accounting for the formation of Gly-Gly and Phe-Phe in the presence of **4** (*vide infra*), as aqueous solutions of those amino acids alone did not show evidence of dimerization. The choice of Gly and Phe obeys to their representative character as non-aromatic and aromatic amino acids, respectively. On the other hand, the lack of observable coupling for Asp and Glu may be ascribed to the strongest interaction of the aminoxazoline moiety with the distal carboxylate groups (*vide supra*: Table 5). That interaction also hampers dimerization, as the charged nitrogen atom at the α -carbon lacks enough nucleophilicity in aqueous medium.

The pK_as of Gly and Phe in water prevent their amino groups from acting as nucleophile as well. However, the experimental evidence accounting for the formation of Gly-Gly and Phe-Phe does actually suggest a role for the aminoxazole as promoter, which should remove a hydrogen from the ammonium groups, leading to a complex **CI** (Fig. 11) free amino groups could then be engaged in nucleophilic attack. That complex brings the two amino acid units in close proximity to allow their assembly in a peptide bond. Figure 11 depicts this catalytic scheme, where the first step involves the association of two molecules of the amino acid giving rise to **CI**, followed by intramolecular peptide bond formation. This sequence affords a second complex (**CII**), from which the dipeptide separates, enabling the heterocycle to initiate another catalytic cycle.

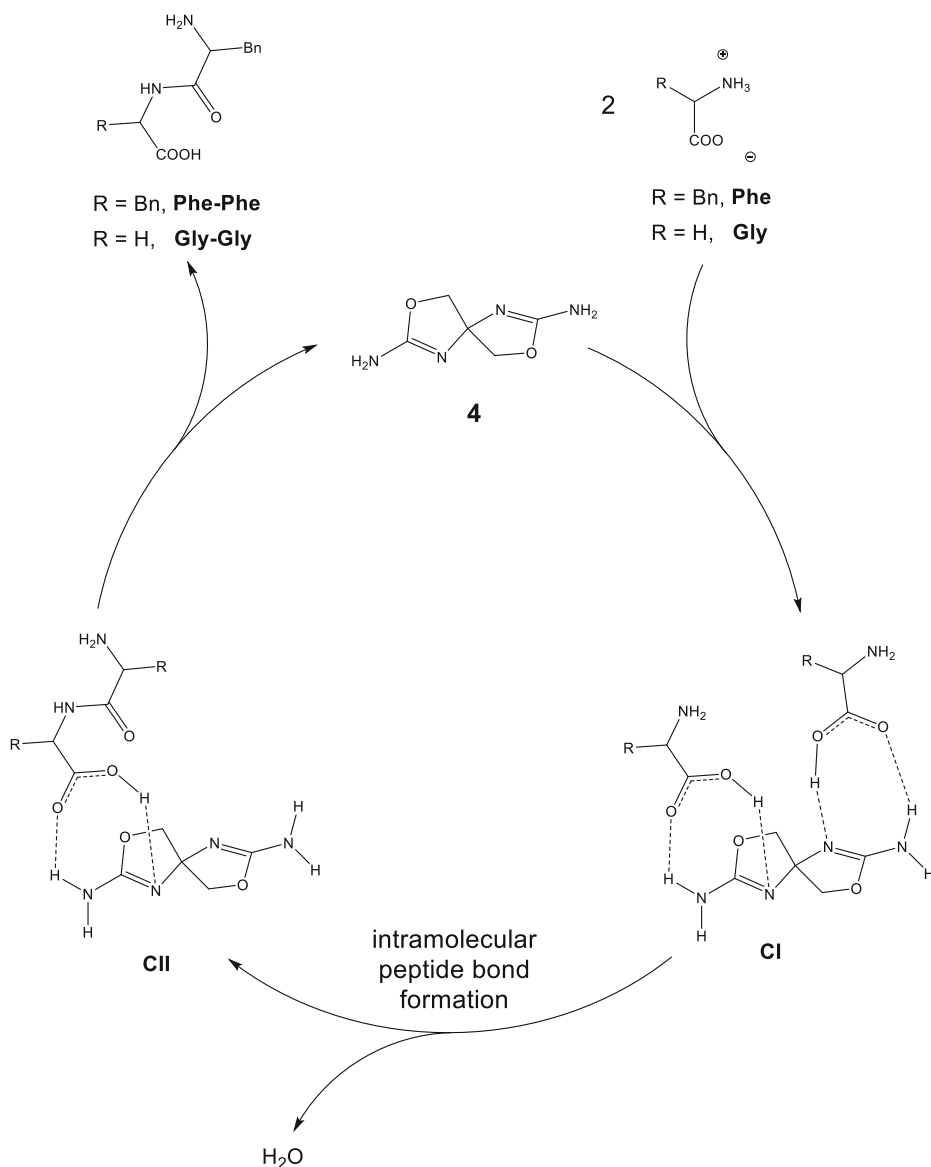


Fig. 11 Catalytic cycle proposed for the dimerization of Gly and Phe in the presence of **4** as catalyst

We have studied the complete reaction pathway yielding Gly-Gly from the initial complex CI_{Gly} at the M06-2X/6-31 + G(d) level of theory in water (SMD). Figure 12 shows the reaction pathways leading to Gly-Gly, whose initial step involves the generation of CI_{Gly} and the subsequent first saddle point (TS1_{Gly}), which corresponds to the nucleophilic attack of the amino group to the carboxylic carbon of the other Gly unit. The intermediate Int1_{Gly} , leads then to a second intermediate (Int2_{Gly}) through a second saddle point (TS2_{Gly}). The latter corresponds to proton transfer from the new ammonium group formed. The slightly higher relative free energy of Int2_{Gly} relative to TS2_{Gly} can be ascribed to their almost identical geometries, which translates into a flat energy profile

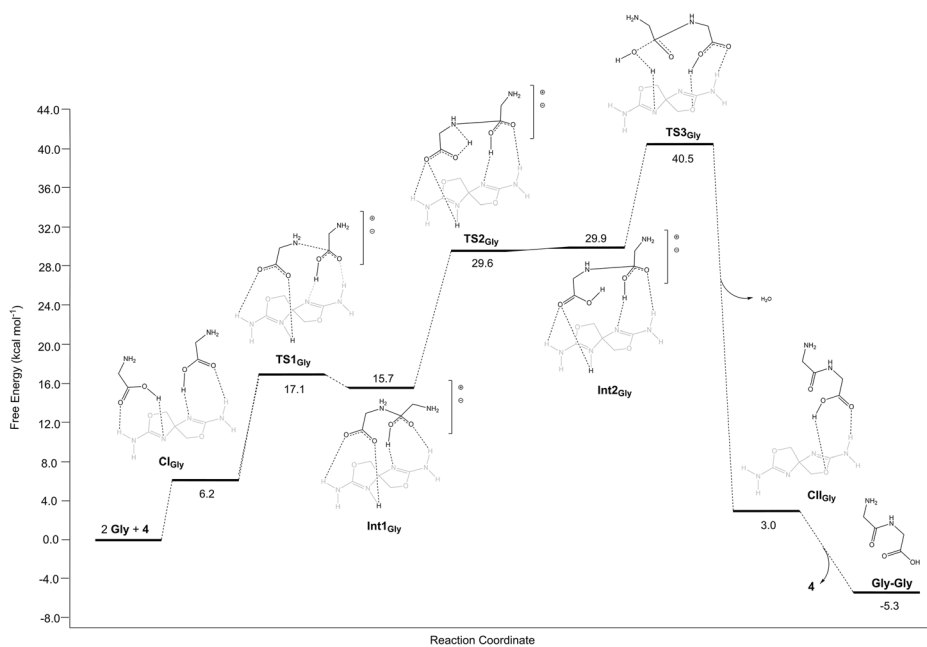


Fig. 12 Reaction pathways for peptide-bond formation between two glycine residues (**Gly**) catalyzed by spirane **4**. Free energy differences with respect to **Gly** and **4** are given in kcal mol⁻¹ at the M06-2X/6-31 + G(d) in water (SMD)

(see ESM for full energy data). In the last step, which is the rate-determining step (**TS3_{Gly}**), the newly generated tetrahedral carbon evolves into the amide bond where the hydroxyl plays the role of leaving group. This abstracts the hydrogen atom of the ammonium group of spirane **4** in a concerted fashion, thus liberating a water molecule. The second complex **CII_{Gly}** dissociates into **Gly-Gly** and spirane **4**, which is then available for a new catalytic cycle.

Since Gly is achiral information and the formation of Phe-Phe was also observed by mass spectrometry, we also simulated the reaction pathway leading to Phe-Phe in the presence of **4**. Here the chiral nature of the amino acid was taken into account. The *R* configuration of spirane **4** was arbitrarily chosen in order to ascertain whether an enantiomerically pure sample of the spiranic derivative would still exert chiral selection during the formation of the dimer. Accordingly, we computed all the reaction channels starting from the four diastereomers of the initial complex **CI**, i.e. *RRR*, *RRS*, *RSR* and *RSS*. For clarity, the first configuration refers to the chiral axis of the spiranic compound while the second and third labels allude to the stereogenic atoms of Phe residues. Figure 13 shows the energy landscape (barriers are given in Gibbs energies) leading to chiral Phe-Phe (i.e. *RR*, *SS*, *RS* and *SR* configurations). Mechanisms were identical to those of **Gly**, albeit the benzyl group slightly increases the energy gaps of some stationary points, particularly **TS2_{Phe}** and **Int2_{Phe}**. For an (*R*)-configured spirane, the most favored pathway yields the *RSS*-configured product, whose rate-determining step corresponds to the second saddle point [$\Delta G^\ddagger(\mathbf{TS2}_{\text{PheRSS}}) = 51.2$ kcal mol⁻¹]. This happens for the *RRS* path as well [$\Delta G^\ddagger(\mathbf{TS2}_{\text{PheRRS}}) = 52.1$ kcal mol⁻¹], while for the *RRR* and *RSR* pathways is the third one [$\Delta G^\ddagger(\mathbf{TS3}_{\text{PheRRR}}) = 52.7$ and $\Delta G^\ddagger(\mathbf{TS3}_{\text{PheRSR}}) = 55.0$ kcal mol⁻¹ respectively]. Although the difference in energy barriers lie in a few

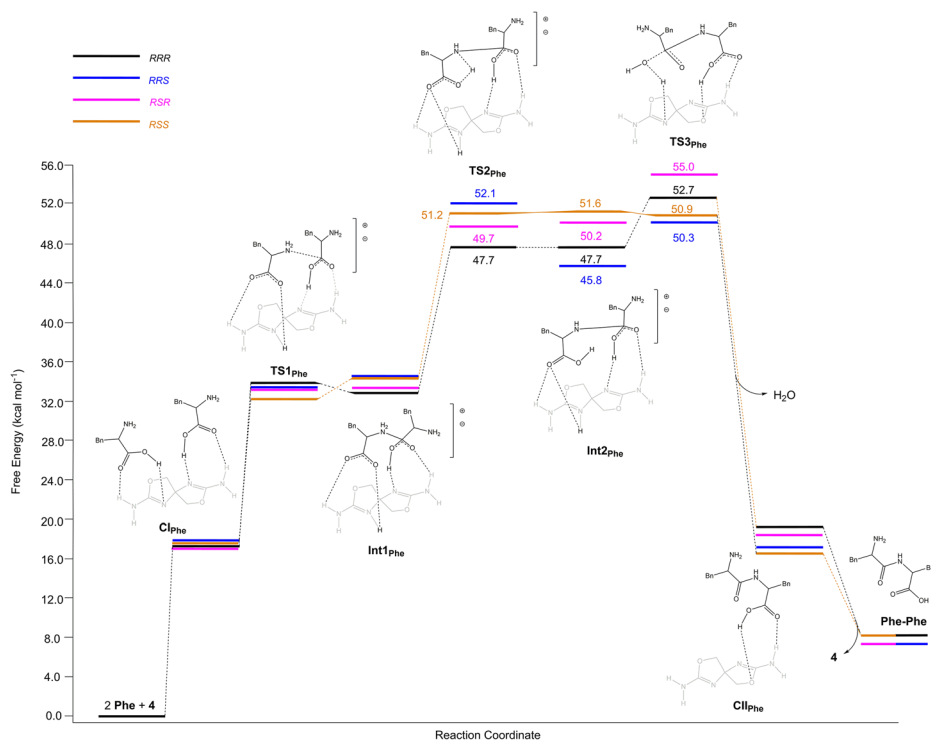


Fig. 13 Reaction pathways for the formation of peptide bonds between two phenylalanine residues (Phe) catalyzed by spirane **4**. Free energy differences are given with respect to Phe and **4** in kcal mol⁻¹ at the M06-2X/6-31+G(d) in water (SMD)

kcal mol⁻¹, the preference toward the homochiral (*S,S*)-dipeptide is noteworthy. The imbalance between homochiral and heterochiral aggregation, favoring the former, has also been observed for other amino acid clusters, for which an entropic advantage has been invoked. (Julian et al. 2005)

Conclusions

Aminoxazolines such as **4**, which can be readily generated from cyanamide and dihydroxyacetone, can serve as sequestering agents for amino acids, with strong interactions for dicarboxylic derivatives such as Glu and Asp. Although ¹H NMR data in solution provide some support, interaction-induced shifts due to proton transfer show small variations relative to unbound substrates, with the exception of bifunctional amino acids such as Glu and Asp, for which distinctive solid phases related to cocrystal formation could be isolated. The complex involving two molecules of Asp (as racemate) coordinated through the side carboxylate to the isoureido moiety of **4** could be elucidated by X-ray diffractometry. DFT calculations indicate that, despite unfavorable free energies owing to entropic loss after coordination in solution, both Glu and Asp exhibit less positive ΔG values when they coordinate to **4** through their carboxylate side chain. Optimized structures also reveal that aminoxazole derivatives are capable of holding the amino acids together, which is now of interest in prebiotic mechanisms

of selection and accumulation. In peripheral observations, formation of dipeptides in an aqueous soup containing **4** and Gly, Ala, or Phe could be inferred with confidence from ESI-MS measurements. Although the low yields of amino acid dimers preclude synthetic exploitation, this detection has value in a geochemical context, and adds further interest in the role that aminoxazole chemistry might have played in chemical evolution. In the search for a mechanistic rationale, DFT analyses support the proposal of a catalytic cycle between Gly and Phe with **4**. The computed hypothetical reaction of a racemic mixture of Phe and the *R*-enantiomer of **4** shows a diastereoselective bias towards the (*S*)-Phe-(*S*)-Phe isomer. Overall, the energy barriers to be overcome discard a facile route, albeit the computational analysis should serve as a preliminary toehold illustrating the potentiality of spiranic aminoxazoles as organocatalysts in future explorations.

Acknowledgements This work was supported by *Junta de Extremadura and Fondo Europeo de Desarrollo Regional* (Grants IB16167 and GR18015). We also gratefully acknowledge the Cénits/COMPUTAEX Foundation for providing computing time on the LUSITANIA Supercomputer.

References

- Aakeröy CB, Bahra GS, Brown CR, Hitchcock PB, Patell Y, Seddon K (2008) L-proline 2,5-dihydroxybenzoic acid (1/1): a zwitterion co-crystal. *Acta Chem Scand* 49:762–767. <https://doi.org/10.3891/acta.chem.scand.49-0762>
- Anastasi C, Crowe MA, Powner MW, Sutherland JD (2006) Direct assembly of nucleoside precursors from two- and three-carbon units. *Angew Chem Int Ed* 45:6176–6179. <https://doi.org/10.1002/anie.200601267>
- Berger M, Schmidchen FP (1999) Zwitterionic guanidinium compounds serve as electroneutral anion hosts. *J Am Chem Soc* 121:9986–9993. <https://doi.org/10.1021/ja992028k>
- Blondeau P, Segura M, Pérez-Fernández R, De Mendoza J (2007) Molecular recognition of oxoanions based on guanidinium receptors. *Chem Soc Rev* 36:198–210. <https://doi.org/10.1039/b603089k>
- Borsenberger V, Crowe MA, Lehbauer J, Raftery J, Helliwell M, Bhutia K, Cox T, Sutherland JD (2004) Exploratory studies to investigate a linked prebiotic origin of RNA and coded peptides. *Chem Biodivers* 1: 203–246. <https://doi.org/10.1002/cbdv.200490105>
- Budin I, Szostak JW (2010) Expanding roles for diverse physical phenomena during the origin of life. *Annu Rev Biophys* 39:245–263. <https://doi.org/10.1146/annurev.biophys.050708.133753>
- Cremer D, Pople JA (1975) A general definition of ring puckering coordinates. *J Am Chem Soc* 97:1354–1358. <https://doi.org/10.1021/ja00839a011>
- Danger G, Plasson R, Pascal R (2012) Pathways for the formation and evolution of peptides in prebiotic environments. *Chem Soc Rev* 41:5416–5429. <https://doi.org/10.1039/C2CS35064E>
- Das B, Srivastava HK (2017) Influence of the local chemical environment in the formation of multicomponent crystals of L-tryptophan with N-heterocyclic carboxylic acids: unusual formation of double zwitterions. *Cryst Growth Des* 17:3796–3805. <https://doi.org/10.1021/acs.cgd.7b00386>
- Ditchfield R, Hehre WJ, Pople JA (1971) Self-consistent molecular-orbital methods. IX. An extended Gaussian-type basis for molecular-orbital studies of organic molecules. *J Chem Phys* 54:724–728. <https://doi.org/10.1063/1.1674902>
- Eschenmoser A (2011) Etiology of potentially primordial biomolecular structures: from vitamin B 12 to the nucleic acids and an inquiry into the chemistry of life's origin: a retrospective. *Angew Chem Int Ed* 50: 12412–12472. <https://doi.org/10.1002/anie.201103672>
- Etter MC, MacDonald JC, Bernstein J (1990) Graph-set analysis of hydrogen-bond patterns in organic crystals. *Acta Crystal B* 46:256–262. <https://doi.org/10.1107/S0108768189012929>
- Francl MM, Pietro WJ, Hehre WJ, Binkley JS, DeFrees DJ, Pople JA, Gordon MS (1982) Self-consistent molecular orbital methods. XXIII A polarization-type basis set for second-row elements *J Chem Phys* 77: 3654–3665. <https://doi.org/10.1063/1.444267>
- Frisch MJ, Trucks GW, Schlegel HB et al. (2009) Gaussian 09. Revision D.01. Gaussian, Inc, Wallingford, CT, USA
- Gordon MS (1980) The isomers of silacyclopropane. *Chem Phys Lett* 76:163–168. [https://doi.org/10.1016/0009-2614\(80\)80628-2](https://doi.org/10.1016/0009-2614(80)80628-2)

- Grouiller A, Mackenzie G, Najib B, Shaw G, Ewing D (1988) A novel stereospecific synthesis of 5-amino-1- β -D-fructofuranosylimidazole-4-carboxamide. *J Chem Soc Chem Commun*:671–672. <https://doi.org/10.1039/C39880000671>
- Halim MA, Shaw DM, Poirier RA (2010) Medium effect on the equilibrium geometries, vibrational frequencies and solvation energies of sulfanilamide. *J Mol Struct (THEOCHEM)* 960:63–72. <https://doi.org/10.1016/j.theochem.2010.08.027>
- Hall SR, McMahon B (2006) International tables for crystallography vol G: definition and exchange of crystallographic data. International union of crystallography, Wiley, New York
- Hariharan PC, Pople JA (1973) The influence of polarization functions on molecular orbital hydrogenation energies. *Theor Chim Acta* 28:213–222. <https://doi.org/10.1007/BF00533485>
- Hawker JR Jr, Oró J (1981) Cyanamide mediated syntheses of peptides containing histidine and hydrophobic amino acids. *J Mol Evol* 17:285–294. <https://doi.org/10.1007/BF01795750>
- Hein JE, Tse E, Blackmond DG (2011) A route to enantiopure RNA precursors from nearly racemic starting materials. *Nat Chem* 3:704–706. <https://doi.org/10.1038/nchem.1108>
- Julian RR, Myung S, Clemmer DE (2005) Do homochiral aggregates have an entropic advantage? *J Phys Chem B* 109:440–444. <https://doi.org/10.1021/jp046478x>
- Kovács J, Pintér I, Köll P (1995) Direct transformation of D-idose and D-altrose with potassium cyanate into cyclic carbamates of derived glycosylamines. *Carbohydr Res* 272:255–262. [https://doi.org/10.1016/0008-6215\(95\)00009-1](https://doi.org/10.1016/0008-6215(95)00009-1)
- Krishnamurthy R, Arrhenius G, Eschenmoser A (1999) Formation of glycolaldehyde phosphate from glycolaldehyde in aqueous solution. *Orig Life Evol Biosph* 29:333–354. <https://doi.org/10.1023/A:1006698208873>
- Kubik S (2017) Anion recognition in aqueous media by cyclopeptides and other synthetic receptors. *Acc Chem Res* 50:2870–2878. <https://doi.org/10.1021/acs.accounts.7b00458>
- Marenich AV, Cramer CJ, Truhlar DG (2009a) Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *J Phys Chem B* 113:4538–4543. <https://doi.org/10.1021/jp810292n>
- Marenich AV, Cramer CJ, Truhlar DG (2009b) Performance of SM6, SM8, and SMD on the SAMPL1 test set for the prediction of small-molecule solvation free energies. *J Phys Chem B* 113:6378–6396. <https://doi.org/10.1021/jp809094y>
- Müller D, Pitsch S, Kittaka A, Wagner E, Wintner CE, Eschenmoser A, Ohloffgewidmet G (1990) Chemie von α -aminonitrilen. Aldomerisierung von glycolaldehyd-phosphat zu racemischen hexose-2,4,6-triphosphaten und (in gegenwart von formaldehyd) racemischen pentose-2,4-diphosphaten: rac-allose-2,4,6-triphosphat und rac-ribose-2,4-diphosphat sind die reaktionshauptprodukte. *Helv Chim Acta* 73:1410–1468. <https://doi.org/10.1002/hlca.19900730526>
- Parker ET, Zhou M, Burton AS, Glavin DP, Dworkin JP, Krishnamurthy R, Fernandez FM, Bada J (2014) A plausible simultaneous synthesis of amino acids and simple peptides on the primordial earth. *Angew Chem Int Ed* 53:8132–8136. <https://doi.org/10.1002/anie.201403683>
- Pascal R, Boiteau L, Commeyras A (2005) From the prebiotic synthesis of α -amino acids towards a primitive translation apparatus for the synthesis of peptides. *Prebiotic chemistry, Top Curr Chem* 259:69–122
- Patel BH, Percivalle C, Ritson DJ, Duffy CD, Sutherland JD (2015) Common origins of RNA, protein and lipid precursors in a cyanosulfidic protometabolism. *Nat Chem* 7:301–307. <https://doi.org/10.1038/nchem.2202>
- Plasson R, Bersini H, Commeyras A (2004) Recycling frank: spontaneous emergence of homochirality in noncatalytic systems. *Proc Natl Acad Sci U S A* 101:16733–16738. <https://doi.org/10.1073/pnas.0405293101>
- Powner MW, Gerland B, Sutherland JD (2009) Synthesis of activated pyrimidine ribonucleotides in prebiotically plausible conditions. *Nature* 459:239–242. <https://doi.org/10.1038/nature08013>
- Powner MW, Sutherland JD, Szostak JW (2010) Chemoselective multicomponent one-pot assembly of purine precursors in water. *J Am Chem Soc* 132:16677–16688. <https://doi.org/10.1021/ja108197s>
- Powner MW, Zheng SL, Szostak JW (2012) Multicomponent assembly of proposed DNA precursors in water. *J Am Chem Soc* 134:13889–13895. <https://doi.org/10.1021/ja306176n>
- Pressman A, Blanck C, Chen IA (2015) The RNA world as a model system to study the origin of life. *Curr Biol* 25:R953–R963
- Reiner H, Plankensteiner K, Fitz D, Rode BM (2006) The possible influence of L-histidine on the origin of the first peptides on the primordial earth. *Chem Biodivers* 3:611–621. <https://doi.org/10.1002/cbdv.200690064>
- Ribeiro RF, Marenich AV, Cramer CJ, Truhlar DG (2010) Prediction of SAMPL2 aqueous solvation free energies and tautomeric ratios using the SM8, SM8AD, and SMD solvation models. *J Comput Aided Mol Des* 24:317–333. <https://doi.org/10.1007/s10822-010-9333-9>
- Ritson D, Sutherland JD (2012) Prebiotic synthesis of simple sugars by photoredox systems chemistry. *Nat Chem* 4:895–899. <https://doi.org/10.1038/nchem.1467>

- Rode BM (1999) Peptides and the origin of life. *Peptides* 20:773–786
- Rode BM, Schwendinger MG (1990) Copper-catalyzed amino acid condensation in water—a simple possible way of prebiotic peptide formation. *Orig Life Evol Biosph* 20:401–410. <https://doi.org/10.1007/BF01808134>
- Ruiz-Mirazo K, Briones C, De La Escosura A (2014) Prebiotic systems chemistry: new perspectives for the origins of life. *Chem Rev* 114:285–366. <https://doi.org/10.1021/cr2004844>
- Saielli G (2010) Differential solvation free energies of oxonium and ammonium ions: insights from quantum chemical calculations. *J Phys Chem A* 114:7261–7265. <https://doi.org/10.1021/jp103783j>
- Saul R, Kern T, Kopf J, Pintér I, Köll P (2000) Reaction of 1,3-disubstituted acetone derivatives with pseudohalides: a simple approach to spiro[4.4]nonane-type bis-oxazolidines and -imidazolidines (bicyclic carbamates, thiocarbamates, ureas, and thioureas). *Eur J Org Chem*:205–209
- Schwartz AW (2007) Intractable mixtures and the origin of life. *Chem Biodivers* 4:656–664
- Schwartz AW (2013) Evaluating the plausibility of prebiotic multistage syntheses. *Astrobiology* 13:784–789. <https://doi.org/10.1089/ast.2013.1057>
- Springsteen G, Joyce GF (2004) Selective derivatization and sequestration of ribose from a prebiotic mix. *J Am Chem Soc* 126:9578–9583. <https://doi.org/10.1021/ja0483692>
- Steinman G, Lemmon RM, Calvin M (1964) Cyanamide: a possible key compound in chemical evolution. *Proc Natl Acad Sci U S A* 52:27–30. <https://doi.org/10.1073/pnas.52.1.27>
- Surov AO, Voronin AP, Vener MV, Churakov AV, Perlovich GL (2018) Specific features of supramolecular organisation and hydrogen bonding in proline cocrystals: a case study of fenamates and diclofenac. *CrystEngComm* 20:6970–6981. <https://doi.org/10.1039/c8ce01458b>
- Tilborg A, Norberg B, Wouters J (2014) Pharmaceutical salts and cocrystals involving amino acids: a brief structural overview of the state-of-art. *Eur J Med Chem* 74:411–426. <https://doi.org/10.1016/j.ejmech.2013.11.045>
- Weissbuch I, Illos RA, Bolbach G, Lahav M (2009) Racemic β -sheets as templates of relevance to the origin of homochirality of peptides: lessons from crystal chemistry. *Acc Chem Res* 42:1128–1140. <https://doi.org/10.1021/ar900033k>
- Weissbuch I, Leiserowitz L, Lahav M (2011) Achiral organic, inorganic, and metal crystals as auxiliaries for asymmetric transformations. *Isr J Chem* 51:1017–1033. <https://doi.org/10.1002/ijch.201100066>
- Yaylayan VA, Harty-Majors S, Ismail AA (1999) Investigation of DL-glyceraldehyde-dihydroxyacetone inter-conversion by FTIR spectroscopy. *Carbohydr Res* 318:20–25
- 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 function. *Theor Chem Accounts* 120:215–241. <https://doi.org/10.1007/s00214-007-0310-x>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.