# Comprehensive Search of Stable Isomers of Alanine and Alanine Precursors in Prebiotic Syntheses

Mitsuo Shoji, Natsuki Watanabe, Yuta Hori, Kenji Furuya, Masayuki Umemura, Mauro Boero, Yasuteru Shigeta

**Comprehensive Search of Stable Isomers of Alanine and Alanine Precursors in Prebiotic Syntheses**

Mitsuo Shoji,*[,1,2] Natsuki Watanabe,[3] Yuta Hori,[1] Kenji Furuya,[4] Masayuki Umemura,[1] Mauro Boero,[5] Yasuteru Shigeta[1]

[1] Center for Computational Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8577, Japan.

[2] JST-PRESTO, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

3 Graduate School of Pure and Applied Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8571, Japan.

[4] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

[5] University of Strasbourg, Institut de Physique et Chimie des Matériaux de Strasbourg, CNRS, UMR 7504, 23 rue du Loess, F-67034 France.

* corresponding author.

E-mail address: mshoji@ccs.tsukuba.ac.jp (M. Shoji)

Abstract

Enantiomeric excesses of $_L$-amino acids were detected in meteorites; however, their molecular mechanism and prebiotic syntheses are still a matter of debate. To shed some light on the origin of the homochirality, alanine and the chiral precursors formed in the prebiotic processes were investigated in terms of their stabilities among their isomers by resorting on the minimum energy principle, namely the abundancy of a molecule in the interstellar medium is directly correlated to the stability among isomers. To facilitate the search for possible isomers, we developed a new isomer search algorithm, the random connection (RC) method, and performed a thorough search of all the stable isomers within a given chemical formula. We found that alanine and most of its precursors are located at higher energy by more than by 5.7 kcal $mol^{-1}$, with respect to the most stable isomer consisting in a linear-chain structure, while only the 2-aminopropanenitrile is the most stable isomer along all the possible ones. The inherent stability of the $\alpha$-amino nitrile suggests that the 2-aminopropanenitrile is the dominant

contribution in the formation of the common enantiomeric excess over α-amino acids. (184/200)

Key words. Minimum energy principle–Chirality–Alanine precursor–isomer–2-aminopropanenitrile

## 1. Introduction

Amino acids are the fundamental building blocks of proteins, and only twenty kinds of them with the L-form of the molecular chirality are exploited by all known living organisms on Earth. Although it is still unclear why L-amino acids are selected, a large enantiomeric excess of L-amino acids was found in carbonaceous meteorites, such as the Murchison meteorite. This discovery suggests an extraterrestrial origin of L-amino acids (Glavin *et al.* 2020; Elsila *et al.* 2016). In the Murchison meteorite, a wide variety of organic compounds, including insoluble and soluble organic ones, have been identified (Koga *et al.*, 2017). For the amino acid component, the meteoritic amino acids possess α, β, γ, and δ amino structures presenting the carbon moiety $C_2$-$C_9$. The most abundant amino acid in the Murchison meteorite is glycine (~3 ppm), whereas other relatively abundant amino acids are alanine, β-alanine and isovaline, all detected in a concentration of ~1.7 ppm. The diverse composition of the meteoritic amino acids suggests a prebiotic synthesis of amino acids in meteorites' parent body and/or the interstellar medium (ISM). The Rosetta mission was able to detect, via mass spectrometry, the presence of volatile glycine and precursor molecules in the coma of comet 67P/Churyumov-Gerasimenko in 2014 (Altwegg *et al*, 2016; Grady *et al.,* 2018). However, mass spectrometry cannot quantitatively distinguish between structural isomers in the same molecular weight. Therefore, the mechanisms of the chemical evolution of glycine and to amino acids still remain unsolved.

The proposed formation pathways for the amino acids are summarized in Figure 1. The Strecker synthesis in the parent body of meteorites is a typical mechanism (Elsila *et al.,* 2016). The Strecker mechanism has long been postulated for the formation of α-amino acids by the hydrocyanation of imines (RCH=NH) and a subsequent hydrolyzation via the formations of amino nitriles (Figure 1). Similar to the Strecker mechanism, the cyanohydrin reaction can explain the formation of α-hydroxy acids, also detected in the Murchison meteorite, and are distributed similarly to the α-amino acids. Another proposed mechanism worthy of attention is the reaction between the imine and carbomyl anion to form an amide intermediate (Koga *et al.*, 2017). This scheme is consistent with the formations of a variety of meteoritic amino acids, soluble organic compounds and alkylated pyridines present in

the Murchison meteorite.

An alternative reaction pathway for the synthesis of amino acids is a radical coupling mechanism on grain surfaces in the ISM. Radical species are formed upon interstellar UV photolysis and cosmic rays irradiation, and their reaction barriers are low enough to proceed in the cold molecular cloud ($T$ ~ 10K). The suggested radical coupling mechanisms for the glycine formation can be realized according to three different paths (Garrod 2013; Singh *et al.,* 2013; Sato *et al.,* 2018);

$$CH_2NH_2 + CHO \rightarrow NH_2CH_2CHO$$

$$NH_2CH_2CHO + 2OH \rightarrow NH_2CH_2COOH + H_2O, \quad (1)$$

$$CH_2NH_2 + COOH \rightarrow NH_2CH_2COOH, \quad (2)$$

$$NH_2 + CH_2COOH \rightarrow NH_2CH_2COOH. \quad (3)$$

According to the chemical kinetic simulation in star-forming regions (Garrod 2013), the dominant reaction processes are (1), (2) and (3) in the low (40K < $T$ < 55K), middle (55K < $T$ < 75K), and high (75K < $T$ < 120K) temperature ranges, respectively. The simulation has evidenced that the production of glycinal ($NH_2CH_2CHO$) is a key event for the glycine formation. For the formation of other amino acids, the radical coupling reactions of the type $NH_2CHCN + R \rightarrow NH_2CHRCN$ have been suggested. These reactions proceed through the formation of aminoacetonitrile ($NH_2CH_2CN$).

Based on these α-amino acid formation mechanisms, the origin of the molecular chirality can be traced back to their nitrile, aldehyde, and amino compounds with chiral centers, highlighted in Figure 1 by the color frames. Even though the generation of amino acids requires these multistep reactions, it is still unclear which formation pathways are dominant compared to other molecular formation reaction channels. If the radical coupling reactions are dominant in these chemical reactions, the molecular mechanism of homochirality is expected to be directly related to the formation processes.

The Minimum Energy Principle (MEP) has been used to predict the presence of interstellar molecules in ISM (Lattelais *et al.*, 2009; Lattelais *et al.*, 2010). This principle is very efficient for many interstellar molecules with a small number of atoms (from two to six), although a few exceptional cases exist (Fourré *et al.,* 2016; Fourré *et al.,* 2020). MEP assumes that the most abundant molecule/isomer in ISM corresponds the most stable one. Although the number of isomers dramatically increase with the number of atoms, in the present study, we did a computational effort in a comprehensive way to evaluate the stability of alanine and chiral alanine precursors within the MEP. The stability of related chiral molecules, namely α-hydroxy compounds, as well as glycine were also investigated for the purposes of comparison and test of the present theoretical procedure. We

implemented a new algorithm to generate isomers for a given chemical composition, and we shall refer to this approach as the random connection (RC) method. The validity of this method was carefully assessed for glycine, whose isomers and isomerization channels were recently explored by K.Ohno and coworkers using their scaled hypersphere search-anharmonic downward distortion following (SHS-ADDF) method (Ohno *et al.*, 2021). The group of Fourré developed and benchmarked a software tool able to determine all possible isomers of a given composition (Fourré *et al.,* 2016; Fourré *et al.,* 2020). Their approach uses a one-to-one correspondence between atoms with bonds and weighted graph with vertexes within the graph theory framework. A thorough inspection of a specified vertex in a graph allows to extract all the possible isomers identified as unique graphs. Such procedure is well suited for small-size molecules if the number of conformational degrees of freedoms is negligible. However, for larger molecules counting of all the possible graphs implies the risk of including unrealistic structures and connections, thus, thus requiring a more careful handling of the conformational degrees of freedom. For this reason, our RC method represents a viable way to overcome these difficulties by introducing random samplings for both bond connections and initial atomic geometries and also by introducing a universal molecular force field which do not require to be adjusted according to the molecules targeted. The detailed computational procedure is described in section 2. Advantages and disadvantages of our RC method compared to the previous approach by Fourré *et al* are summarized in the supporting information (Table S1). In section 3, the results of stable isomers for seven molecules, specifically glycine (**A**), alanine (**B**), alanineamide (**C**), 2-aminopropanal (**D**), 2-aminopropanenitrile (**E**), lactic acid (**F**) and lactonitrile (**G**) are presented. Their molecular structures are sketched in Figure 2, and we can observe that **C**-**E** are the chiral precursors of **B**, whereas **G** is a unique chiral precursor of α-hydroxy acid **F**. These compounds are all strictly related to the promising prebiotic syntheses of **B** and **F**. Discussions concerning the origin of homochirality and limitation of the present theoretical approach, in the context of the results obtained here, are given in sections 3.8 and 3.9. Finally in section 4, we summarized discussions as conclusion.

## 2. Method

The computational procedures of the RC method are detailed in this section. The overall flowchart is sketched in Figure 3 and can be summarized as follows:

1. A target molecule is initially identified. The number of atoms and their chemical types are then counted by virtually eliminating all their covalent bonds. For example, the molecule shown in

Figure 3 contains (1) one oxygen atom with two covalent bonds, (2) one carbon atom with four covalent bonds, (3) one nitrogen atom with three covalent bonds, and (4) five hydrogen atoms with one covalent bond. This virtual splitting allows to identify all the main building blocks of related isomers.

2. To construct new isomers, the isolated atoms are redistributed uniformly in a cubic box of 10 Å side length. Despite looking a bit arbitrary, this procedure is necessary to sample different conformers. The atoms are then newly bond-connected to other atoms with unconnected bonds. Initially, heavy atoms are connected. The bond-connections are randomly selected from all the available possibilities. Then, unconnected bonds are saturated with hydrogen atoms. By this procedure, the bonds are drawn to form a single molecule saturating the number of covalent bonds of each atom, avoiding spurious formation of separate molecules. If the atom types are not unique, all the possible atom-type-sets are searched separately. For instance, the O atom in a carboxylic acid has two possibilities, either the double-bonded carbonyl oxygen (=O) or the single-bonded hydroxyl/ether oxygen (-O-).
3. After the molecule has the complete information of bond connection, the molecular geometry is optimized at a force field level. Our specific force field consists of a harmonic and a van der Waals component, accounting for the bonding ($E_{\mathrm{R}}$) and non-bonding ($E_{\mathrm{vdW}}$) interactions, respectively

$$E = E_{\mathrm{R}} + E_{\mathrm{vdW}}$$

$$E_{\mathrm{R}} = \sum_{ij} \frac{1}{2} k_{ij} \left(r - r_{\mathrm{ij}}\right)^2$$

$$E_{\mathrm{vdW}} = \sum_{kl} D_{kl} \{-2(r_{kl}/r)^6 + (r_{kl}/r)^{12}\}$$

The indices *i*, *j* run over all the bonded atom pairs and the indices *k*, *l* run over the non-bonded atom pairs. This force field is a simplified form of the Universal Force Field (UFF) (Rappe *et al.*, 1992). The other terms of the original UFF, i.e., bond angle bending ($E_{\mathrm{q}}$), dihedral angle torsion ($E_{\mathrm{q}}$) and electrostatic terms ($E_{\mathrm{el}}$), are not considered in the present study, because an exploratory rough optimization is needed in this stage. Further and more accurate optimizations within electronic structure calculations are carried out in the next optimization step, where the bond information is not valid under the calculation. All the parameters used in the present study ($r_{\mathrm{ij}}$, $D_{\mathrm{ij}}$) are reported in the supporting information (Table S2).

4. The accurate structural optimization within higher-level theoretical approaches is performed at this step. We stress the fact that in practical applications of this protocol, steps 2-4 should be

carried out many times for sampling. It is not always granted that the former step succeeds in the generation of new bond-connections and different conformations of the isomer. It is recommended carefully check the sets of the generated molecular structures before performing higher-level theoretical calculations, which is the most time-consuming. In our case, we adapted a procedure making use of two different level of electronic structure calculation. On a first instance, we used a density functional theory (DFT) approach at the theoretical levels of B3LYP/6-31G and B3LYP/6-311++G(d,p). The DFT can properly reproduce the stable molecular structures and has shown to be a viable way applicable to a high number of samplings. Zero-point vibrational energy (ZPE) corrections were evaluated at the B3LYP/6-311++G(d,p) theoretical level by performing frequency calculations. In the RC method, random samplings were performed for both connections and geometries. If the number of samplings is insufficient, a complete search for isomers is not pursued. Duplications resulting into an identical molecule were used to check whether or not the sampling was sufficient. We found that 1000 samplings were sufficient to generate stable isomers for the present molecules (**A**-**G**). On a second instance, for the stable isomers, a more accurate *ab initio* method, CCSD(T)/cc-pVTZ, was used for a more accurate energy evaluation. ZPE corrections for the CCSD(T) were obtained from the B3LYP/6-311++G(d,p) calculations.

All the electronic structure calculations were performed using the Gaussian16 program package (Frisch *et al.*, 2016). Molecular structures shown in the figures were drawn using the visual molecular dynamics program (Humphrey *et al.*, 1996).

## 3. Results and Discussion

### 3.1 Glycine isomers ($C_2H_5NO_2$, **A**)

Glycine is the simplest amino acid and is a fundamental constituent of the main chain of proteins. Moreover, glycine can be synthesized in laboratory and has been found in meteorites. Theoretical studies have been performed to inspect the formation and molecular properties of glycine, which special attention to conformers and isomers. This makes glycine the best reference molecule to check the validity of our RC method. Ohno and coworkers have shown that there are two isomers lower in energy, hence more stable, than glycine by using their SHS-ADDF method (Ohno *et al.*, 2021). The most stable isomer was identified as $CH_3NHCOOCH$ (*N*-methylcarbamic acid, **A1**) with four conformations and the second stable isomer was $CH_3OCONH_2$ (methyl carbamate, **A2**) with one

conformer, more stable than glycine (**A3**). Glycine has two low-lying conformers (10.0, 10.7 kcal mol$^{-1}$), and $NH_2COCH_2OH$ (glycolamide, **A4**) with two conformers (10.3, 10.9 kcal mol$^{-1}$) turned out to be energetically stable and comparable to the glycine isomers.

The RC method was able to reproduce the reported results, and three stable conformers of glycine were obtained (Table 1, Figure 4). These results suggest that glycine is not the most stable isomer in $C_2H_5NO_2$, and its energy is higher by 10.0 kcal mol$^{-1}$ relative to the most stable isomer. The formation of the $C_\alpha$ carbon, -NH-$C_\alpha$-CO-, requires a higher in energy compared to the formation of the direct carbonyl-amino groups linkage, i.e., the peptide bond (-NH-CO-). All these stable isomers contain the carbonyl group (-C=O-), and the other unsaturated bonds of the ethylene group (C=C) and imine group (C=N) are not relatively stable. $NH_2CH_2OCHO$ (aminomethyl formate, **A5**) is the next stable isomer with respect to **A4**.

### 3.2 Alanine isomers ($C_3H_7NO_2$, **B**)

Given the reliability of the RC method in generating the stable isomers of glycine, we used this approach to search for alanine isomers ($C_3H_7NO_2$). After 1000 samplings, we obtained four isomers that were more stable than alanine: *N*-ethylcarbamic acid ($CH_3CH_2NHCOOH$, **B1**), ethylcarbamate ($NH_2COOCH_2CH_3$, urethane, **B2**), *N*-dimethylcarbamic acid ($(CH_3)_2NCOOH$, **B3**), and 2-hydroxypropanamide ($NH_2COCH(OH)CH_3$, **B4**). The molecular structures are shown in Figure 5 and their relative energies are summarized in Table 2. Alanine (**B5**) is energetically located at 10.3 kcal mol$^{-1}$ above the most stable molecule, **B1**. Compared to the isomers of glycine, the stable isomers of alanine can be rationalized in terms of methyl-group substitutions to the glycine isomers. For example, the methyl group substitutions of **A1** at the methyl group hydrogen, amide group hydrogen and carboxyl group hydrogen generate the isomers **B1**, **B3** and **B8** (**A1** → (**B1**, **B3**, **B8)**). In an analogous way, other stable isomers of alanine can be realized as follows: **A2** → **B2**, **A3** → **B5**, **A4**→ (**B4**, **B9**) and **A5** → **B10**, where the arrow "→" indicates a methyl group substitution. According to the relation of the methyl group substitution, the higher energy state of alanine (10.3 kcal mol$^{-1}$) is related to the higher energy state of glycine (10.0 kcal mol$^{-1}$), because **A3** and **B5** carry the same α-amino acid structure and the most stable isomers **A1** and **B1** share the same molecular framework with a carbamic acid group (-NHCOOH). It should be mentioned that there are no stable isomers containing the ethylene group and imine in the stable isomers of alanine. It is also expected that all the monomer forms of the amino acids are characterized by an instability in comparison with the isomers, because

the main chain part of the amino acid, $NH_2$-CH-COOH, is energetically located at a higher value with respect to the methylcarbamic acid part (-$CH_2$-NH-COOH).

### 3.3 Alanineamide isomers ($C_3H_8N_2O$, **C**)

As alanine is not the most stable isomer in $C_3H_8N_2O$, we turned our investigation to the alanine derivative, alanineamide. For this system, Figure 6 and Table 3 show the molecular structures and summarize their relative energies. We found that only the *N*-ethylurea (**C1**) is more stable than alanineamide (**C2**). Indeed, C2 is energetically located 5.7 kcal $mol^{-1}$ above **C1**. **C1** can assume four stable conformations differing in the conformation of the amino and ethyl groups. The isomers whose stability is comparable to **C2** are 3-amino propan amide (**C3**), *N*,*N'*-dimethylurea (**C4**), *N*,*N*-dimethylurea (**C5**) and *N*-(aminomethyl)acetamide (**C6**). 2-Amino-*N*-methylacetamide (**C7**) and *N*-(1-aminoethyl)formamide (**C8**) are still existing isomers, however, they are located at higher relative energies than **C2** (11.7, 12.2 kcal $mol^{-1}$).

**C1**, **C4** and **C5** are all urea derivatives. From the viewpoint of group substitution, **C1** is an aminated compound of **B1** at the carboxyl hydroxy group, and **C1** is also related to an aminated compound of **B2** at the carboxyl ether group. Similarly, **C2** is related to both **B5** and **B4**. The relatively higher stability of **C2** compared to **B5** (10.3 kcal $mol^{-1}$) can be explained by the substitution relation of **B4** → **C2** (5.7 kcal $mol^{-1}$). Another remarkable feature in the alanineamide isomers (**C**) is that the β-amino amide (**C3**) becomes the next stable isomer to the α-amino acid (**C2**). This feature was observed for the alanine isomers (**B7**).

### 3.4 2-Aminopropanal isomers ($C_3H_7NO$, **D**)

In the ensemble of systems having the $C_3H_7NO$ composition, we found that there are eight isomers more stable than 2-aminopropanal (**D9**) (Figure 7). These stable isomers are propanamide (**D1**), *N*-methylacetamide (**D2**), *N*-ethylformamide (**D3**), propanimidic acid (**D4**), *N*,*N*-dimethylformamide (**D5**), 1-aminopropan-2-one (**D6**), *N*-methylethanimidic acid (**D7**), and ethylmethanimidic acid (**D8**). At variance with these eight stable structures, **D9** is higher in energy than the most stable isomer **D1** by 23.9 kcal $mol^{-1}$ (Table 4). The number of conformations of the stable isomers is 2, 3, 2, 4, 1, 4, 2, 1 for all the systems from **D1** to **D8** in this order. The stable isomers of **D1**, **D2**, **D3** and **D5** are all amide compounds, and various chemical species can be present, not only with the carbonyl group (C=O) but also with the hydroxyl group (**D4**) and with the ether group (**D10**) in the energy range around the value

of **D9**. Concerning the nitrogen compound, the imine group (C=N) also confers a relatively stability to the compounds **D4, D7**, **D8**, **D10** next to the amide compounds. Other stable isomers above **D10** are 3-aminopropanal (**D11**) and methyl methylmethanimidate (**D12**). For the group substitution, the analogy with alanine can be seen by replacing a hydroxy group with a hydrogen atom. By substitution at the carboxy group, the following correspondence can be inferred **B1** → **D3**, **B3** → **D5**, **B5** → **D9** and **B7** → **D11**. Similarly, by substitution at the hydrocarbon, the correspondence becomes **B4**→ **D1**, **B6** → **D1** and **B9** → **D2**. The latter substitution contributes to stabilize the isomers. Another correspondence worthy of note is the tautomerization of amide (-NH-CO-) into imidic acid (-N=C(OH)-). The analogies that arise in this case are **D1**→ **D4**, **D2** → **D7** and **D3** → **D8**. Due to the relative abundance of more stable isomers, 2-aminopropanal (**D9**) cannot be regarded as a stable isomer for the $C_3H_7NO$ composition.

### 3.5 2-Aminopropanenitrile isomers ($C_3H_6N_2$, **E**)

As a further step in our investigation, we focused on the isomers of amid nitril, $NH_2CH(CH_3)CN$. We found that the most stable form is 2-aminopropanenitrile (**E1**) (Figure 8). The second stable isomer is 3-aminopropanenitrile (**E2**) and the relative energy with respect to **E1** is only 0.1 kcal $mol^{-1}$ (Table 5). The energy difference becomes 0.9 kcal $mol^{-1}$ at the more accurate theoretical level, CCSD(T)/cc-pVTZ (Table 5). Other stable isomers are propanimidamide (**E3**), ethylcyanamide (**E4**), imidazoline (dihydro imidazole, **E5**), *N*-ethylmethanimidamide (**E6**), propane-1,2-diimine (**E7**) and (methylamino)acetonitrile (**E8**), and they are also comparable in stability to **E1**, having relative energies of 3.5, 4.1, 5.3, 6.3, 7.3 and 7.9 kcal $mol^{-1}$, respectively. Imidazoline (**E5**) is characterized by a five-membered ring structure and the molecular framework is different from other isomers. Amide compounds (-CN) confer stability to the isomers **E1**, **E2** and **E4**, and other isomers having the imine group (C=N) accompanies an unsaturated bond or ring formation. The reason of higher stability of **E1** can be ascribed to the fact that the nitril group does not become a linker but acts only a terminal block and prefers to form a chemical bond with C atom in comparison with N atom of e.g. **E4**. Hence, in this case the most stable isomer turns out to be α-amid nitrile (**E1**), followed by β-amid nitrile (**E2**).

### 3.6 Lactic acid isomers ($C_3H_6O_3$, **F**)

Isomers of lactic acid, 2-hydroxypropanoic acid ($HOCH(CH_3)COH$), were our next target. We found that only etabonic acid (**F1**) is more stable than lactic acid (**F2**) by 3.2 kcal $mol^{-1}$(Table 6, Figure

9). **F1** has four stable conformations energetically lower than **F2**. The higher energy isomers are 3-hydroxypropanoic acid (**F3**), hydroxymethyl acetate (**F4**) and 1-hydroxyethyl formate (**F5**), which is still located on the energy scale at a value lower than 10 kcal $mol^{-1}$. Compared to the alanine isomers (**B**), group substitution from the amide group (-NH-) to the oxygen atom (-O-) shows close correspondence, specifically (**B1**, **B2) → F1**, (**B4**, **B5) → F2**, (**B6**, **B7) → F3**, (**B9**, **B10) → F4** and **B11 → F5**. In this substitution, **B3** with the secondary carbamate group has no corresponding isomer in the lactic acid isomers **(F)**. From the substitution relation, lactic acid resembles alanine in terms of isomer stability (**B1 → F1**, **B5 → F2**). β-Hydroxycarbonate **F3** can be found at an energy slightly higher than α-hydroxycarbonate **F2**. This characteristic is observed also in other alanine compounds (**B7**, **C3**, **D11**, **E2**).

3.7 Lactonitrile isomers ($C_3H_5NO$, **G**)

To complete our study, we sought for the isomers of hydroxy nitril, $HOCH(CH_3)CN$. In this case, we found that two isomers are more stable than 2-hydroxypropanenitrile (**G3**) (Figure 10). These isomers are isocyanatoethane (**G1**) and 3-hydroxypropanenitrile (**G2**). The isocyanato compound **G1** is significantly more stable than the hydroxy nitrile compounds **G2** and **G3** ($\Delta E$(**G2**)=15.5 kcal $mol^{-1}$ and $\Delta E$(**G3**)=16.5 kcal $mol^{-1}$) (Table 7). Hydroxypropanenitriles (**G2, G3**) can assume three conformations as a result of the rotation of the hydroxy group. β-Hydroxynitrile **G2** is slightly lower in energy than α-hydroxynitrile **G3**. The five-membered ring isomers of the dihydro-oxazoles (**G4**, **G5)** are characterized by higher energies with respect to **G3** ($\Delta E$(**G4**)=21.3 kcal $mol^{-1}$, $\Delta E$(**G5**)=27.9 kcal $mol^{-1}$). Ethyl cyanate (**G6**) is further located at a higher energy ($\Delta E$(**G6**)=28.6 kcal $mol^{-1}$). Having assessed the high stability of **G1**, a correspondence of hydroxy nitrils (**G**) with amid nitrils (**E**) can be established by introducing a substitution of an amid ($NH_2$) to the hydroxy (OH): **E1 → G3**, **E2 → G2**, **E4 → G6, E5 → G4**. As the order of these substitutions are different, the stabilities in the isomers between the amid nitrils (**E**) and hydroxynitriles (**G**) differ consistently with this reordering.

3.8 Origin of homochirality

The present results indicated that the thermodynamical stability of amino nitrile is remarkably advantageous, and that amino nitrile is a key chiral molecule in comparison with other precursors and amino acids in the interstellar and circumstellar medium. Recently, the first chiral molecule, propylene oxide, was detected in Sgr B2(N) (McGuire *et al.*, 2016). Theoretical investigations carried out by

Hori and coworkers suggest that the photolysis of propylene oxide or a cation precursor can induce the homochirality (Hori *et al.*, 2022). If the enantiomeric excess of chiral isomer generates from photoreactions under circular polarized light (CPL) (Fukushima *et al.*, 2020), the most stable and longest-living species, aminonitriles, become the best candidates responsible for the enantiomeric excess of amino acids. Therefore, we expect that the 2-aminopropanenitrile (3-aminopropanenitrile) might be a crucial molecule contributing to the origin of homochirality for all the chiral α-amino acids (β-amino acids). Although 2-aminopropanenitrile was not yet detected in SgrB2(N) (Møllendal *et al.*, 2012), observations and analyses for aminonitriles are the worth considering targets.

On the other hand, aminoacetonitrile, the precursor of glycine, has been detected in Sagittarius (Sgr) B2 (Belloche *et al.*, 2008). Amino acids are much less stable upon UV irradiation and cosmic rays in interstellar regions than nitriles (Bernstein *et al.*, 2004). Previous theoretical studies have shown that the formation of aminoacetonitrile requires the overcoming of a relatively low energy barrier ($\Delta U$ = 8-9 kcal mol$^{-1}$), though the hydrolysis of aminoacetonitrile to glycine is characterized by a higher electronic energy barrier ($\Delta U$ = 27-34 kcal mol$^{-1}$), which cannot proceed at cryogenic temperatures (Woon 2008; Rimola *et al.*, 2010). Assuming a typical temperature of $T$ < 100 K in the molecular cloud, the energy differences of $\Delta E$ = 3.2 and 5.7 kcal mol$^{-1}$ correspond to the population differences of less than $1.0\times10^{-7}$ and $3.5\times10^{-13}$ according to the Boltzmann distribution. It should be noted that abundance in the ISM is not determined only by the thermal stability but environmental effects such as adsorption and reactivity with atomic hydrogen can also contribute the abundance (Fourré *et al.*, 2020). Nonetheless, it becomes clear that a small relative energy difference of even $\Delta E$ = 3.2 kcal mol$^{-1}$ contributes to the relative probability in the ISM. These reactions are deeply different from ordinary chemical reactions occurring in laboratory at $T$ = 300K, which can overcome energy barrier up to the values of about $\Delta E$ = 20 kcal mol$^{-1}$ by thermal activation (Shoji *et al.*, 2022).

### 3.9 Limitation of the present theoretical approach

In the present theoretical investigation, amino acids and their chiral precursors are assumed to be in the gas phase. In the ISM, the dominant molecular formation occurs on the surface of dust and ice, and the processes of adsorption and diffusion of molecules on the surfaces can play a role in the chemical evolution of interstellar molecules (Wakelam *et al*., 2017; Vidali, 2013). Indeed, isomers of isocyanic acid (HNCO) have been shown to be characterized by different adsorption energies with an upper limit of 19.4 kcal mol$^{-1}$ (Lattelais M. et al, 2015). Nonetheless, even for the large perturbation

on relative energies due to surfaces adsorption, HNCO and HOCN isomers remain the most and second-most stable isomers, respectively. Thus, this energetic order is not jeopardized by the vicinity of an adsorbing surface. This is consistent with our results and a confirmation that the relative energies obtained in the gas phase (MEP) are reliable and represent a reliable predictive quantity for the identification of molecules. Further theoretical investigations on the relative stabilities on icy surfaces will be desirable and will represent the target of forthcoming studies.

## 4. Conclusion

To investigate the stabilities of alanine and the related chiral molecules among their isomers, six chiral molecules formed during the prebiotic processes and achiral glycine were targeted in the present theoretical study. The molecular mechanism of the formation of homochirality of amino acids is intimately related to the origin of life on Earth. The molecules of amino acids and their precursors are large-size compared to the molecules detected in ISM, and their conformational isomers as well as their structural isomers should be searched explicitly for their relative stabilities.

The approach originally proposed by Fourré and coworkers (Fourré *et al.,* 2016; Fourré *et al.,* 2020), has been demonstrated to be efficient for searching ISM molecules provided that the number of conformers is not too large. Nonetheless, it can fail in finding exhaustively all the relevant conformers and unexpected bonding, especially for larger and complex molecules. To overcome this difficulty, we developed a new isomer search algorithm, referred to as RC method by introducing random selections for the degrees of freedoms of bond connection and conformation. The successful application of the RC method to glycine allowed us to move to the actual target of this study. More precisely, two isomers, *N*-methylcarbamic acid (**A1**) and methyl carbamate (**A2**), could be obtained and their more stable four and one conformers were obtained completely for **A1** and **A2**, respectively. Glycine(**A3**) remains higher in energy than the most stable isomer by $\Delta E$(**A3**)=10.0 kcal mol$^{-1}$. All the energetics obtained at the B3LYP were refined within the CCSD(T) approach. Relative energies computed at the CCSD(T) level, shown in parentheses in Tables 1-7, turned out to be qualitatively similar to those obtained at the B3LYP//6-311++G(d,p) level. Hence, our analysis was mainly based on the B3LYP/6-311++G(d,p) results. Moreover, CCSD(T) calculations become clearly computationally demanding for larger-size molecules.

The RC method was applied to chiral molecules including alanine and alanine precursors. Based on the relationship of the molecular frameworks in different compositions, we found that there are

structural similarities between the stable isomers. Among the stable isomers of alanine and alanine precursors (**B**-**E**), all the most stable ones **B1**, **C1**, **D1** assume a linear chain framework like *N*-ethylcarbamic acid (**B1**) with the ethyl group, whereas alanine and the precursors (**B5**, **C2**, **D9**), which have the α-amino acid framework like alanine (**B5**), are all higher energy isomers by more than 5.7 kcal $mol^{-1}$. The higher energies of α-amino acid and the precursors can be easily explained by the group substitutions, the main chain part of the amino acid, $NH_2$-CH-COOH, is higher in energy with respect to the methylcarbamic acid part (-$CH_2$-NH-COOH).

Analogously, the hydroxy acids and the precursors (**F** and **G**) have the same relationship as alanine, however, only the α-hydroxy acid (**F2**) is energetically low as 3.2 kcal $mol^{-1}$. This result suggests that stability among the isomers during the formation of α-hydroxy acid is highest in lactic acid (**F2**). Another remarkable feature is the fact that the corresponding isomers of the β-amino/hydroxy acid are similarly stable over the precursors (**B7, C3**, **D11**, **E2**, **F3**, **G2**) comparable to ones of the α-amino/hydroxy acid (**B5, C2**, **D9**, **E1**, **F2**, **G3**). This feature can explain the wide variety of amino acids detected in meteorites.

Despite their structural similarities to **B5**, 2-aminopropanenitrile (**E1**) is the most stable among the isomers. This unique property of **E1** can be ascribed to two different features: on one hand, the amino and nitrile groups are not stably connected, and, on the other hand, the nitrile group acts as a terminal moiety in the molecular structure.

The intrinsic stability of 2-aminopropanenitrile is a convincing piece of evidence on which the origin of homochirality of amino acids can be rationalized. To generate the enantiomeric excesses of L-amino acids, the aminonitrile state is a fundamental element, since it is the easiest to form and the one characterized by the longest lifetime. As such, it plays a major role in the enantiomeric excess issue. The aminonitrile intermediates is a promising candidate to induce the homochirality of amino acids in prebiotic syntheses. In this respect, this works can stimulate forthcoming theoretical and experimental investigations to consolidate or disprove the outcome of our modeling.

Acknowledgements

This research was supported by JST, PRESTO Grant Number JPMJPR19G6, Japan and JSPS KAKENHI ground numbers 20H05453 and 20H05088. This work used computational resources of (1) Cygnus system through Multidisciplinary Cooperative Research Program in CCS, University of Tsukuba, and (2) supercomputer Fugaku provided by the RIKEN Center for Computational Science

(project ID: hp210115). M.B. thanks the HPC Center at the University of Strasbourg funded by the Equipex Equip@Meso project and the CPER Alsacalcul/Big Data, and the Grand Equipement National de Calcul Intensif (GENCI) under allocation DARI-A0100906092.

## Author Disclosure Statement

No competing financial interests exist.

## Supplementary Material

Supplementary Tables about (S1) advantage of the RC method, (S2) UFF parameters and (S3-S7) molecular structures for all the stable conformational and structural isomers, and XYZ coordinates of all the stable isomers.

References

Altwegg, K., Halsiger, H., Bar-Nun, A. et al. (2016) Prebiotic chemicals-amino acid and phosphorus-in the coma of comet 67P/Churyumov-Gerasimenko. *Sci Adv* 2:e1600285.

Belloche, A., Menten, K. M., Comitto, C. *et al.* (2008) Detection of amino acetonitrile in Sgr B2(N). *Astro Astro* 482:179.

Belloche, A., Menten, K. M., Comito, C. *et al.* (2008) Detection of amino acetonitrile in Sgr B2(N) *Astro Astro* 492:769-773.

Bernstein, M.P., Ashbourn, F.M., Sandford, S.A., and Allamandola, L.J. (2004) The Lifetimes of Nitriles (CN) and Acids (COOH) during Ultraviolet Photolysis and Their Survival in Space. *Astrophys J* 601:365-370.

Elsila, J.E., Aponte, J.C., Blackmond, D.G. *et al.* (2016) Meteoritic Amino Acids: Diversity in Compositions Reflects Parent Body Histories. *ACS Cent Sci* 2(6):370-379.

Fourré, I., Matz, O., Ellinger, Y., and Guillenmin, J.-C. (2020) Relative thermodynamic stability of the [C,N,O] linkages as an indication of the most abundant structures in the ISM. *Astro Astro* 639:A16.

Fourré, I., Rosset, L., Chevreau, H., and Ellinger, Y. (2016) About the detection of urea in the interstellar medium: the energetic aspect. *Astro Astro* 589:A18.

Frisch, M.J., Trucks, G.W., Schlegel, H.B. *et al.* (2016) Gaussian, Inc., Wallingford CT.

Garrod, R.T. (2013) A three-phase chemical model of hot cores:the formation of glycine. *Astrophys J*, 765:60.

Fukushima H, Yajima H, Umemura M. (2020) High circular polarization of near-infrared light induced by micron-sized dust grains. *Mon. Not. R. Astron. Soc.* 496:2762–2767.

Glavin, D.P., Burton, A.S., Elsila, J.E. *et al.* (2020) The Search for Chiral Asymmetry as a Potential Biosignature in our Solar System. *Chem Rev* 120:4660-4689.

Grady, M. M., Wright, I. P., Engrand, C., Siljestrom, S. (2018) The Rosetta Mission and the Chemistry of Organic Species in Comet 67P/Churyumov-Gerasimenko. *Elements* 14:95-100.

Hori, Y., Nakamura, H., Sakawa, T., *et al.* (2022) Theoretical Investigation of the Possibility of the Formation of Propylene Oxide Homochirality in Space. will be published somewhere.

Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics. *J Mol Graph* 14:33-38.

Kayanuma, M., Kidachi, K., Shoji, M., Komatsu, Y., Sato, A., Shigeta, Y., Aikawa, and Y., Umemura, M. (2017) A theoretical study of the formation of glycine via hydantoin intermediate in outer space

environment. *Chem Phys Lett* 687:178-183.

Koga, T. and Naraoka, H. (2017) A new family of extraterrestrial amino acids in the Murchison meteorite. *Sci Rep* 7:636.

Lattelais, M., Pauzat, F., Ellinger, Y., and Ceccarelli, C. (2009) Interstellar complex organic molecules and the minimum energy principle. *The Astrophys J* 696:L133-L136.

Lattelais, M., Pauzat, F., Ellinger, Y. and Ceccarelli, C. (2010) A new weapon for the interstellar complex organic molecule hunt: the minimum energy principle. *Astro Astro* 519:A30.

Lattelais, M., Pauzat, F., Ellinger, Y., Ceccarelli, C. (2015) Differential adsorption of CHON isomers at interstellar grain surfaces. Astro Astro 578:A62.

McGuire, B.A., Carroll, P.B., Loomis, R.A., *et al.* (2016) Discovery of the interstellar chiral molecule propylene oxide ($CH_3CHCH_2O$). *Science* 352:1449–1452.

Møllendal, H., Margulès, L., Belloche, A. *et al.* (2012) Rotational spectrum of a chiral amino acid precursor, 2-aminopropionitrile, and searches for it in Sagittarius B2(N). *Astro Astro* 538:A51.

Ohno, K., Kishimoto, N., Iwamoto, T., Satoh, H., and Watanabe, H. (2021) High performance global exploration of isomers and isomerization channels on quantum chemical potential energy surface of $H_5C_2NO_2$. *J Comput Chem* 42:192-204.

Rappe, A.K., Casewit, C.J., Colwell, K.S., Goddard III, W.A., and Skiff, W.M. (1992) UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J Am Chem Soc* 114(25):10024-10035.

Rimola, A., Sodupe, M., and Ugliengo, P. (2010) Deep-space glycine formationvia Strecker-type reactions activated by ice water dust mantles. A computational approach. *Phys Chem Chem Phys* 12:5285-5294.

Sato, A., Kitazawa, Y., Ochi, T., Shoji, M., Komatsu, Y., Kayanuma, M., Aikawa, Y., Umemura, M., and Shigeta, Y. (2018) First-principles study of the formation of glycine-producing radicals from common interstellar species. *Mol Astrophys* 10:11-19.

Shoji, M., *et al.* (2022) Peculiar Conformational Change and Electron Transfer realized by the Quinone Cofactor during the Catalytic Cycle of Bacterial Copper Amine Oxidase. will be published somewhere.

Singh, A., Shivani, Misra, A., and Tandon, P. (2013) Quantum chemical analysis for the formation of glycine in the interstellar medium. *Res Astro Astro* 13(8):912-920.

Woon, D.E. (2008) Pathways to Glycine and Other Amino Acids in Ultraviolet-irradiated Astrophysical Ices Determined via Quantum Chemical Modeling. *The Astrophys J* 571:L177-L180.

Table 1. Relative energies ($\Delta E$) and number of stable conformations ($N_{conf.}$) in isomers in glycine composition ($C_2H_5NO_2$) as obtained by the RC method.

| Isomer | $\Delta E$ / kcal mol$^{-1}$ [a] | $N_{conf.}$ [b] | Molecule |
|---|---|---|---|
| **A1** | 0(0), 1.3(1.4), 7.7(6.7), 8.8 (8.0) | 4(4) | $CH_3NHCOOH$, *N*-methylcarbamic acid |
| **A2** | 4.6(4.7),12.8(12.0) | 1(1) | $CH_3OCONH_2$, methyl carbamate |
| **A3** | <u>10.0(8.4)</u>, 10.7(9.4), 11.5(10.1) | 0(0) | $H_2NCH_2COOH$, glycine |
| **A4** | 10.3(10.5), 10.9(11.7) | 0(0) | $H_2NCOCH_2OH$, glycolamide |
| **A5** | 18.0(16.9) | 0(0) | $NH_2CH_2OCHO$, aminomethyl formate |

[a] ZPE corrected relative energies ($\Delta E$) with respect to the most stable isomer at the theoretical level B3LYP//6-311++G(d,p). Relative energies computed at the CCSD(T)/cc-pVTZ level are shown in parentheses.

[b] Number of stable conformations with respect to the most stable conformation of glycine (**A3**), whose $\Delta E$ is underlined. Values in parentheses are the results obtained at the CCSD(T)/cc-pVTZ level.

Table 2. Relative energies ($\Delta E$) and number of stable conformations ($N_{conf.}$) in isomers in alanine composition ($C_3H_7NO_2$) as obtained by the RC method.

| Isomer [a)] | $\Delta E$ / kcal mol$^{-1}$ [b)] | $N_{conf.}$ [c)] | Molecule |
|---|---|---|---|
| **B1** | 0(0), 1.4(1.4), 7.5(6.5), 9.1(8.1) | 4(4) | $CH_3CH_2NHCOOH$, *N*-ethylcarbamic acid |
| **B2** | 3.5(3.7), 10.9(10.6), 12.4(11.8) | 1(1) | $NH_2COOCH_2CH_3$, ethylcarbamate |
| **B3** | 7.5(7.1), 15.9(14.4) | 1(1) | $(CH_3)_2NCOOH$, *N*-dimethylcarbamic acid |
| **B4*** | 8.9(9.3), 9.1(8.7), 9.4(9.5), 9.5(9.5) | 4(4) | $NH_2COCH(OH)CH_3$, 2-hydroxypropanamide |
| **B5*** | <u>10.3(7.9)</u>, 10.6(8.3), 11.4(9.2), 11.5(9.0), 12.2(9.6) | 0(0) | $NH_2CH(CH_3)COOH$, alanine |
| **B6** | 10.4(10.6) | 0(0) | $NH_2COCH_2CH_2OH$, 3-hydroxypropanamide |
| **B7** | 10.9( 9.1), 11.1( 9.3), 11.4( 9.6), 11.9(10.5 ), 12.9(11.4) | 0(0) | $NH_2CH_2CH_2COOH$, 3-aminopropanoic acid (β-alanine) |
| **B8** | 11.1(11.9), 12.5(13.2) | 0(0) | $CH_3NHCOOCH_3$, methyl methyl carbamate |
| **B9** | 11.4(11.9), 11.4(11.9) | 0(0) | $CH_3CONHCH_2OH$, *N*-(hydroxymethyl)acetamide |
| **B10** | 11.9(11.4), 12.7(12.8) | 0(0) | $NH_2CH_2OCOCH_3$, aminomethyl formate |
| **B11*** | 13.6(13.2) | 0(0) | $CHONHCH(OH)CH_3$, *N*-(1-hydroxyethyl)formate |

[a] Isomers contain a chiral center are marked by asterisk.

[b] ZPE corrected relative energies ($\Delta E$) with respect to the most stable isomer at the theoretical level B3LYP//6-311++G(d,p). Relative energies computed at the CCSD(T)/cc-pVTZ level are shown in parentheses.

[c] Number of stable conformations with respect to the most stable conformation of alanine (**B5**), whose $\Delta E$ is underlined. Values in parentheses are the results obtained at the CCSD(T)/cc-pVTZ level.

Table 3. Relative energies ($\Delta E$) and number of stable conformations ($N_{conf.}$) for isomers in alanine amide composition ($C_3H_8N_2O$) as obtained by the RC method.

| Isomer [a] | $\Delta E$ / kcal mol$^{-1}$ [b] | $N_{conf.}$ [c] | Molecule |
|---|---|---|---|
| **C1** | 0(0), 0.2(0.5),1.2(1.6), 1.6(1.7) | 4(4) | $NH_2CONHCH_2CH_3$, *N*-ethylurea |
| **C2*** | <u>5.7(4.4)</u>, 5.8(4.7), 7.5(6.2), 8.4(6.9), 9.0(7.6) | 0(0) | $NH_2CHCH_3CONH_2$, alanine amide |
| **C3** | 7.1(6.6), 8.2(7.4), 8.8(8.0), 9.5(9.3) | 0(0) | $NH_2CH_2CH_2CONH_2$, 3-amino propan amide |
| **C4** | 7.2(7.8), 8.1(9.0), 12.8(13.2) | 0(0) | $CH_3NHCONHCH_3$, *N*,*N*'-dimethylurea |
| **C5** | 8.3(7.8) | 0(0) | $(CH_3)_2NCONH2$, *N*,*N*-dimethylurea |
| **C6** | 8.9(8.8), 10.0(10.3), 12.2(12.1) | 0(0) | $NH_2CH_2NHCOCH_3$, *N*-(aminomethyl) acetamide |
| **C7** | 11.7(12.1) | 0(0) | $NH_2CH_2CONHCH_3$, 2-amino-*N*-methylacetamide |
| **C8*** | 12.2(11.1) | 0(0) | $NH_2CH(CH_3)NHCHO$, *N*-(1-aminoethyl)formamide |

[a] Isomers contain a chiral center are marked by asterisk.

[b] ZPE corrected relative energies ($\Delta E$) with respect to the most stable isomer at the theoretical level B3LYP//6-311++G(d,p). Relative energies computed at the CCSD(T)/cc-pVTZ level are shown in parentheses.

[c] Number of stable isomers with respect to the most stable conformation of alanine amide (**C2**), whose $\Delta E$ is underlined. Values in parentheses are the results obtained at the CCSD(T)/cc-pVTZ level.

Table 4. Relative energies ($\Delta E$) and number of stable conformations ($N_{conf.}$) for isomers in 2-aminopropanal composition ($C_3H_7NO$) as obtained by the RC method.

| Isomer [a)] | $\Delta E$ / kcal mol$^{-1}$ [b)] | $N_{conf.}$ [c)] | Molecule |
|---|---|---|---|
| **D1** | 0(0), 1.0(1.2) | 2(2) | $NH_2COCH_2CH_3$, propanamide |
| **D2** | 4.2(5.1), 6.8(7.5), 6.8(7.5) | 3(3) | $CH_3NHCOCH_3$, *N*-methylacetamide |
| **D3** | 8.8(9.3), 9.6(10.4) | 2(2) | $CH_3CH_2NHCHO$, *N*-ethylformamide |
| **D4** | 13.2(10.5), 16.6(13.6), 16.8(13.9), 17.4(14.6) | 4(4) | $HN=C(OH)CH_2CH_3$, propanimidic acid |
| **D5** | 14.3(14.5) | 1(1) | $(CH_3)_2NCHO$, *N*,*N*-dimethylformamide |
| **D6** | 17.9(17.0), 18.7(18.0), 19.9(19.1), 20.8(19.9) | 4(4) | $NH_2CH_2COCH_3$, 1-aminopropan-2-one |
| **D7** | 19.0(17.6), 21.5(19.6) | 2(2) | $CH_3NC(CH_3)OH$, *N*-methylethanimidic acid |
| **D8** | 23.7(22.0), 25.9(23.5), 26.0(24.4), 27.0(25.2), 27.8(25.9) | 1(0) | $CH_3CH_2N=CHOH$, ethylmethanimidic acid |
| **D9*** | <u>23.9(21.7)</u>, 23.9(21.7), 24.6(22.9), 24.9(22.6), 24.9(22.9), 25.1(23.0) | 0(0) | $NH_2CH(CH_3)CHO$, 2-aminopropanal |
| **D10** | 24.7(22.6) | 0(0) | $CH_3C=NHOCH_3$, methyl ethanimidate |
| **D11** | 25.2(23.4), 25.9(24.1), 26.1(24.7), 26.3(24.8), 26.8(25.4), 26.9(25.5), 27.0(25.8), 27.2(25.8), 27.2(26.1) | 0(0) | $NH_2CH_2CH_2CHO$, 3-aminopropanal |
| **D12** | 32.7(31.7) | 0(0) | $CH_3N=CHOCH_3$, methyl methylmethanimidate |

[a] Isomers contain a chiral center are marked by asterisk.

[b] ZPE corrected relative energies ($\Delta E$) with respect to the most stable isomer at the theoretical level B3LYP//6-311++G(d,p). Relative energies computed at the CCSD(T)/cc-pVTZ level are shown in parentheses.

[c] Number of stable isomers with respect to the most stable conformation of 2-aminopropanal(**D9**), whose $\Delta E$ is underlined. Values in parentheses are the results obtained at the CCSD(T)/cc-pVTZ level.

Table 5. Relative energies ($\Delta E$) and number of stable conformations ($N_{conf.}$) for isomers in 2-aminopropanenitrile composition ($C_3H_7NO$) as obtained by the RC method.

| Isomer [a)] | $\Delta E$ / kcal mol$^{-1}$ [b)] | $N_{conf.}$ [c)] | Molecule |
|---|---|---|---|
| **E1**[*] | <u>0(0)</u>, 1.8(1.6), 1.9(1.8) | 0(0) | $NH_2CH(CH_3)CN$, 2-aminopropanenitrile |
| **E2** | 0.1(0.9), 0.2(0.9), 0.5(1.4),0.6(1.7), 1.5(2.4) | 0(0) | $NH_2CH_2CH_2CN$, 3-aminopropanenitrile |
| **E3** | 3.5(7.8), 4.2(8.6), 4.5(8.8), 4.7(11.9) | 0(0) | $NH_2(CH_2CH)C=NH$, propanimidamide |
| **E4** | 4.1(6.7), 4.1(6.8), 4.3(6.8) | 0(0) | $NCNCH_2CH_3$, ethylcyanamide |
| **E5** | 5.3(6.8) | 0(0) | $NCH_2NHCH_2CH_2$, imidazoline |
| **E6** | 6.3(12.8) | 0(0) | $HN=CHNHCH=CH_2$, *N*-ethylmethanimidamide |
| **E7** | 7.3(9.7) | 0(0) | $HN=CH(CH_3)C=NH$, propane-1,2-diimine |
| **E8** | 7.9(8.9), 9.5(10.6) | 0(0) | $CH_3NHCH_2CN$, (methylamino)acetonitrile |

[a] Isomers contain a chiral center are marked by asterisk.

[b] ZPE corrected relative energies ($\Delta E$) with respect to the most stable isomer at the theoretical level B3LYP//6-311++G(d,p). Relative energies computed at the CCSD(T)/cc-pVTZ level are shown in parentheses.

[c] Number of stable conformations with respect to the most stable conformation of 2-aminopropanenitrile (**E1**), whose $\Delta E$ is underlined. Values in parentheses are the results obtained at the CCSD(T)/cc-pVTZ level.

Table 6. Relative energies ($\Delta E$) and number of stable conformations ($N_{conf.}$) for isomers in lactic acid composition ($C_3H_6O_3$) as obtained by the RC method.

| Isomer [a)] | $\Delta E$ / kcal mol$^{-1}$ [b)] | $N_{conf.}$ [c)] | Molecule |
|---|---|---|---|
| **F1** | 0(0), 0.7(1.2), 1.0(1.2), 3.2(3.6), 11.1(10.8) | 4(3) | $CH_3CH_2OCOOH$, etabonic acid |
| **F2*** | <u>3.2(3.2)</u>, 5.1(5.6), 5.2(5.7) | 0(0) | $CH_3CH(OH)COOH$, lactic acid |
| **F3** | 4.0(4.5), 5.8(6.8) | 0(0) | $CH_2(OH)\ CH_2COOH$, 3-hydroxypropanoic acid |
| **F4** | 6.4(7.2) | 0(0) | $CH_3COOCH_2OH$, hydroxymethyl acetate |
| **F5*** | 8.6(8.7), 11.3(11.0) | 0(0) | $CH_3CH(OH)OCHO$, 1-hydroxyethyl formate |

[a] Isomers contain a chiral center are marked by asterisk.

[b] ZPE corrected relative energies ($\Delta E$) with respect to the most stable isomer at the theoretical level B3LYP//6-311++G(d,p). Relative energies computed at the CCSD(T)/cc-pVTZ level are shown in parentheses.

[c] Number of stable conformations with respect to the most stable conformation of lactic acid (**F2**), whose $\Delta E$ is underlined. Values in parentheses are the results obtained at the CCSD(T)/cc-pVTZ level.

Table 7. Relative energies ($\Delta E$) and number of stable conformations ($N_{conf.}$) for isomers in lactonitrile composition ($C_3H_5NO$) as obtained by the RC method.

| Isomer [a)] | $\Delta E$ / kcal mol$^{-1}$ [b)] | $N_{conf.}$ [c)] | Molecule |
|---|---|---|---|
| **G1** | 0(0) | 1(1) | $CH_3CH_2$-N=C=O, isocyanatoethane |
| **G2** | 15.5(9.7), 15.8(10.4), 16.3(10.8), 16.6(11.0), 17.3(11.6) | 3(1) | $NCCH_2CH_2OH$, 3-hydroxypropanenitrile |
| **G3*** | <u>16.5(10.3)</u>, 16.8(10.5), 18.3(11.8) | 0(0) | $CH_3CH(OH)CN$, 2-hydroxypropanenitrile |
| **G4** | 21.3(15.5) | 0(0) | c-$CH_2CH_2NCHO$-, 4,5-dihydro-1,3-oxazole |
| **G5** | 27.9(23.6), 28.2(23.7) | 0(0) | $CH_3CH_2OCN$, ethyl cyanate |
| **G6** | 28.6(22.9) | 0(0) | c-$CH_2CH$=$NCH_2O$-, 2,5-dihydro-1,3-oxazole |

[a] Isomers contain a chiral center are marked by asterisk.

[b] ZPE corrected relative energies ($\Delta E$) with respect to the most stable isomer at the theoretical level B3LYP//6-311++G(d,p). Relative energies computed at the CCSD(T)/cc-pVTZ level are shown in parentheses.

[c] Number of stable conformations with respect to the most stable conformation of lactonitrile (**G3**), whose $\Delta E$ is underlined. Values in parentheses are the results obtained at the CCSD(T)/cc-pVTZ level.

Figures

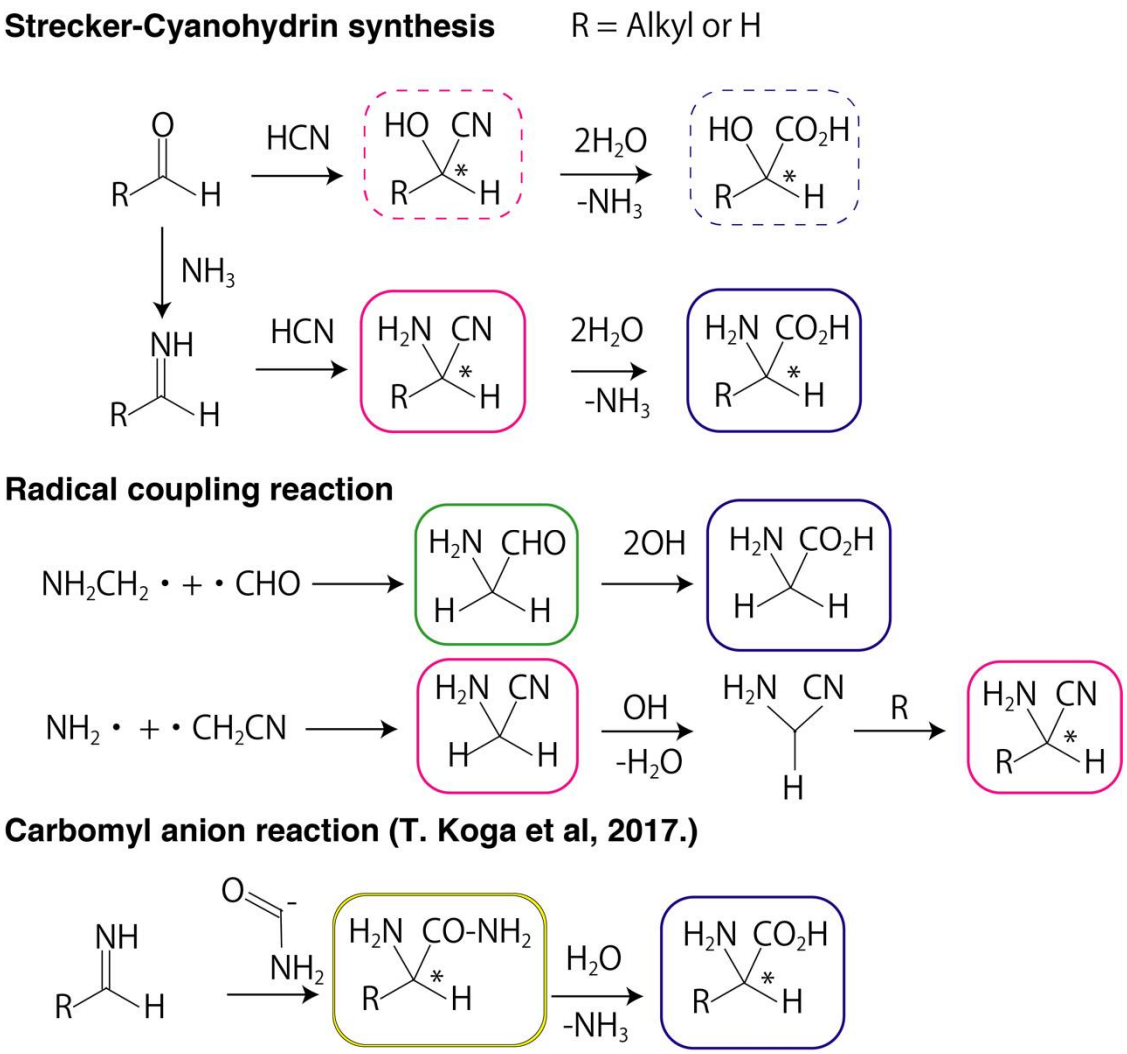


Figure 1. Proposed formation pathways for α-amino acids and α-hydroxy acids. Intermediate compounds of nitrile (pink), aldehyde (green), carbonylic acid (blue) and amide (yellow) are framed in different colors for the sake of clarity. Hydroxyl compounds are framed by the dashed lines in the corresponding colors. The chiral centers are indicated by an asterisk (*).

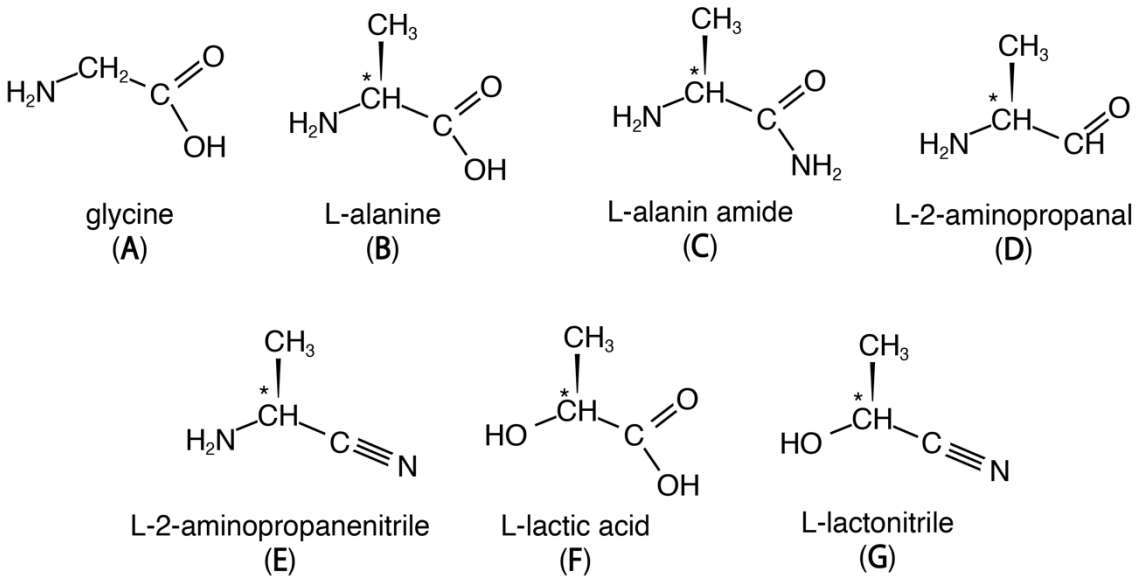


Figure 2. Molecular structures investigated in the present study. The chiral centers are indicated by an asterisk (*).

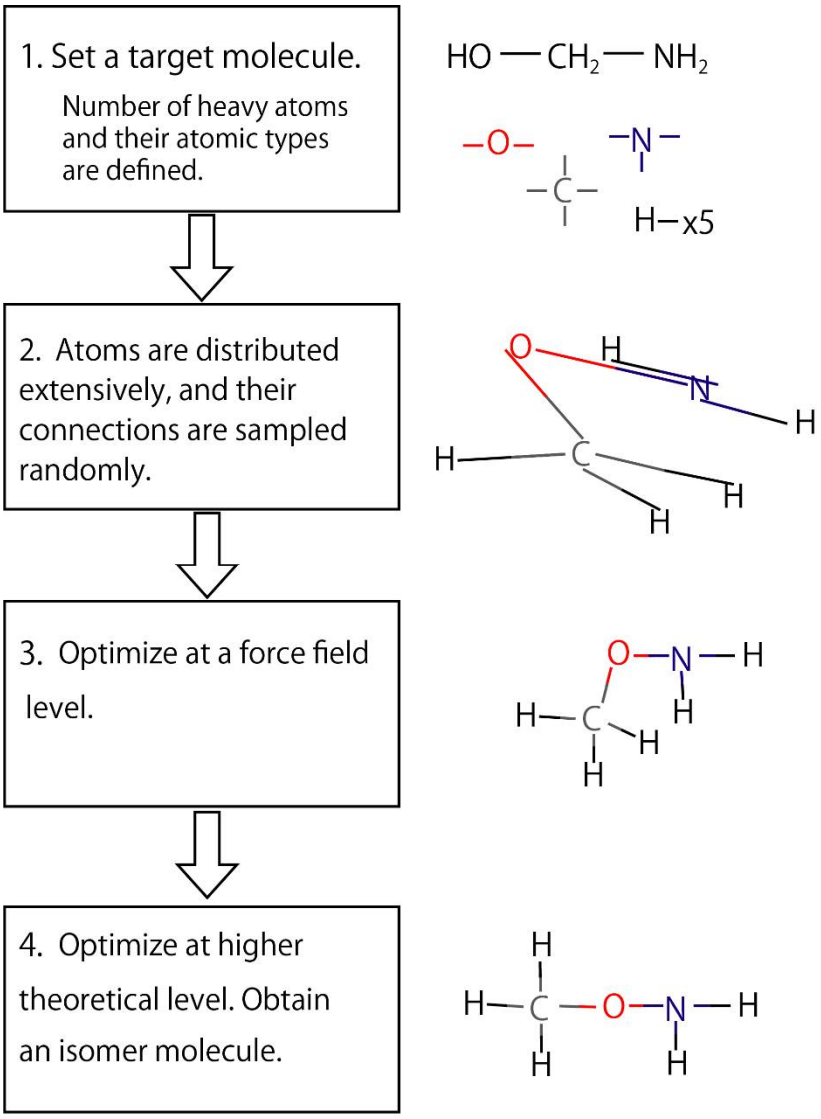


Figure 3. Schematic flowchart of the random connection (RC) method proposed and implemented in the present work. By repeating the procedures 2-4, a number of isomers in different conformations are generated.

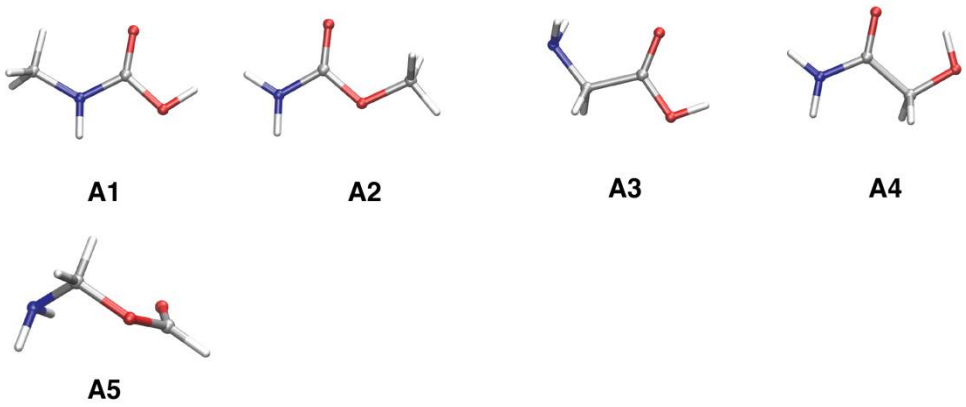


Figure 4. Stable isomers of the glycine composition ($C_2H_5NO_2$). Most stable conformers for each isomer are shown. Glycine is **A3**. Here and in the following figures, a standard color code is used: gray for C, blue for N, red for O and white for H.

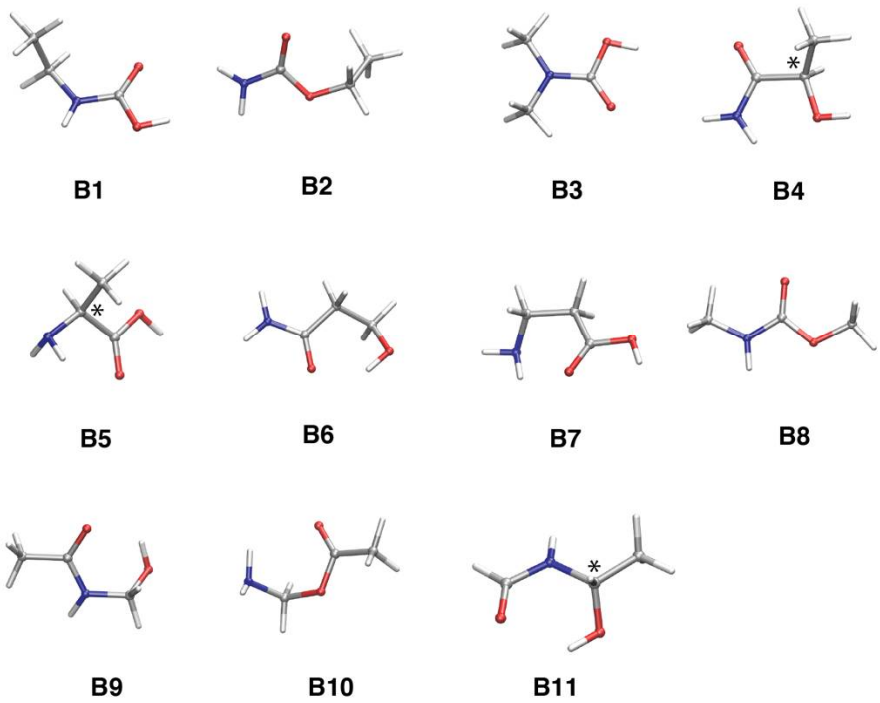


Figure 5. Stable isomers of the alanine composition ($C_3H_7NO_2$). Most stable conformers for each isomer are shown. Alanine is labeled as **B5**. All the chiral centers are indicated by an asterisk (*).

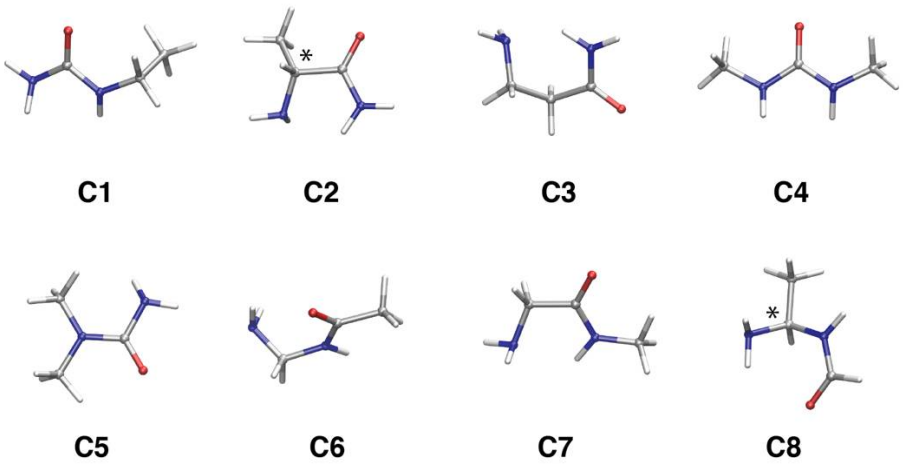


Figure 6. Stable isomers of the alanine amide composition ($C_3H_8N_2O$). Most stable conformers for each isomer are shown. Alanineamide is labeled as **C2**. All the chiral centers are indicated by an asterisk (*).

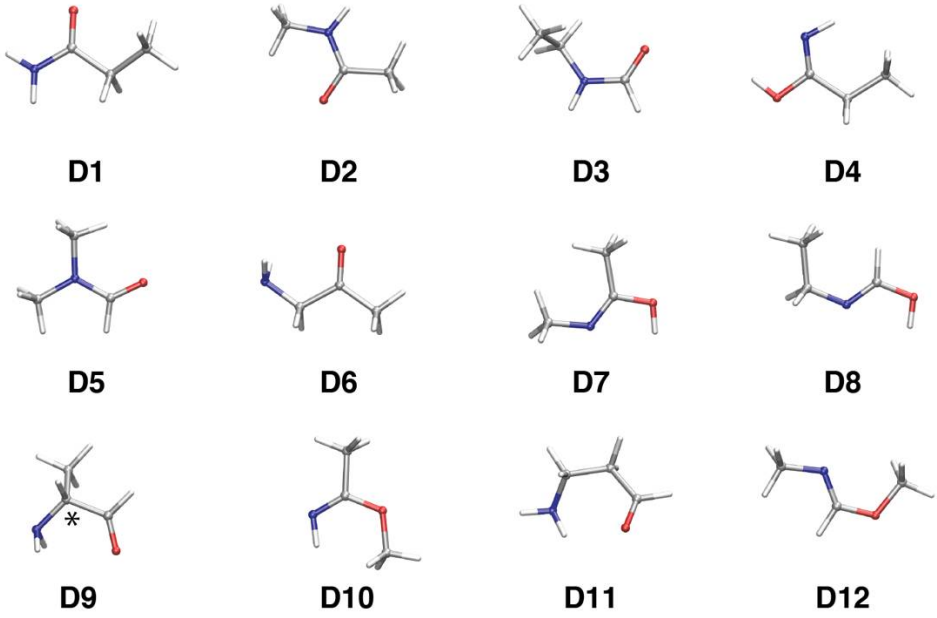


Figure 7. Stable isomers of the 2-aminopropanal composition ($C_3H_7NO$). Most stable conformers for each isomer are shown. 2-aminopropanal is labeled as **D9**. All the chiral centers are indicated by an asterisk (*).

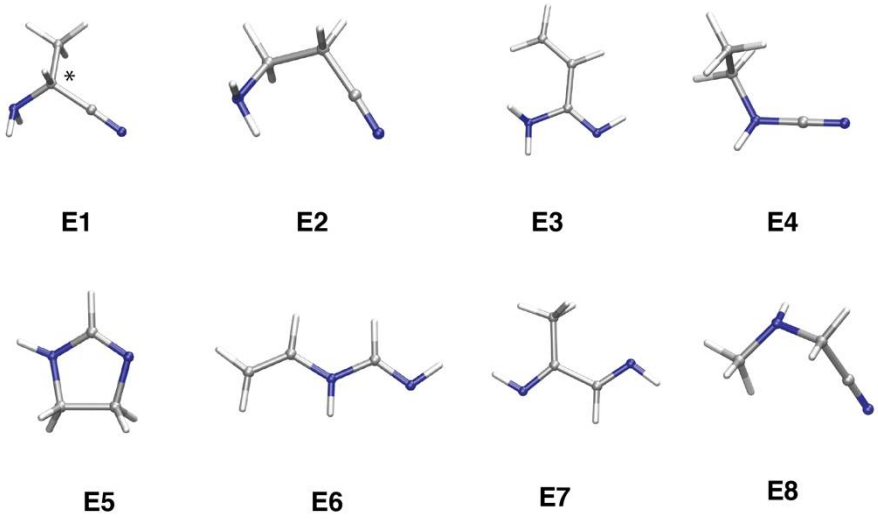


Figure 8. Stable isomers of the 2-aminopropanenitrile composition ($C_3H_6N_2$). Most stable conformers for each isomer are shown. 2-aminopropanenitrile is labeled as **E1**. All the chiral centers are indicated by an asterisk (*).

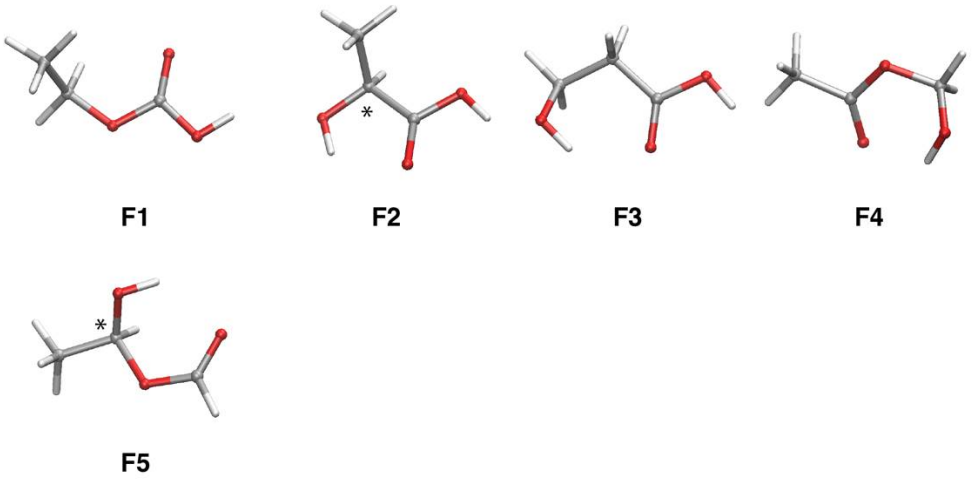


Figure 9. Stable isomers of the lactic acid composition ($C_3H_6O_3$). Most stable conformers for each isomer are shown. lactic acid is labeled as **F2**. All the chiral centers are indicated by an asterisk (*).

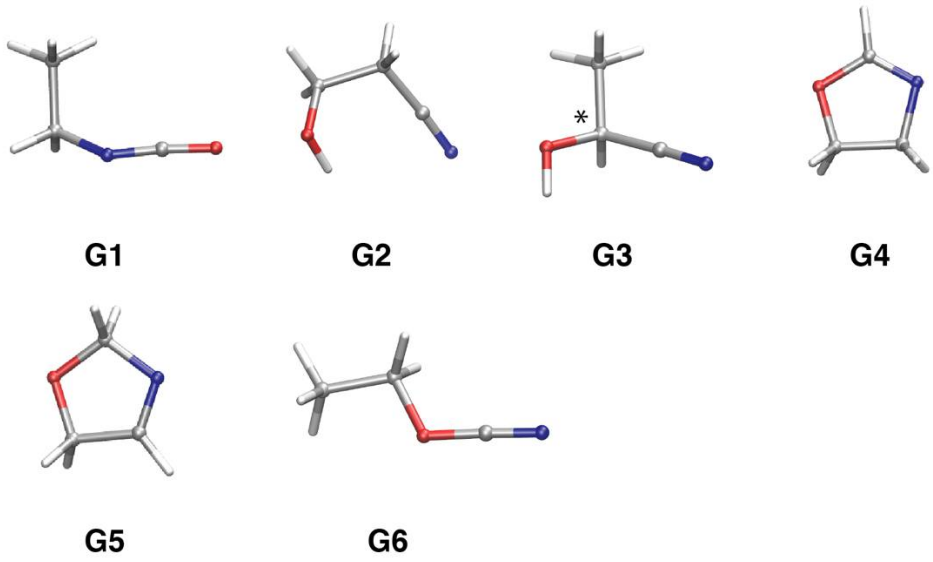


Figure 10. Stable isomers of lactonitrile composition ($C_3H_5NO$). Most stable conformers for each isomer are shown. Lactonitrile is labeled as **G3**. All the chiral centers are indicated by an asterisk (*).