ACADEMIA ROMÂNĂ
Revue Roumaine de Chimie
http://web.icf.ro/rrch/

Rev. Roum. Chim.,
2021, *66*(2), 135-148
DOI: 10.33224/rrch.2021.66.2.04

Dedicated on the occasion of the 90th anniversary
*of **Professor Alexandru T. Balaban***
with admiration for his remarkable contributions to
the fields of Organic, Physical Organic, and Mathematical Chemistry

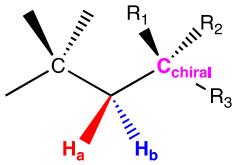
^{1}H NMR AND CONFORMATIONAL ANALYSIS OF DIASTEREOTOPIC METHYLENE PROTONS IN ACHIRAL FLEXIBLE MOLECULES

Mircea D. GHEORGHIU,[a,*] Carlos A. VALDEZ[b] and Ana RACOVEANU[b]

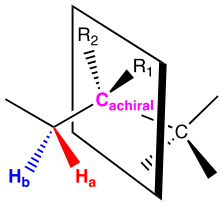
[a] Department of Chemistry, One Washington Square, San José State University, San José, CA 95192-0101, USA
[b] Lawrence Livermore National Laboratory, 7000 East Avenue, Livermore, CA 94550, USA

Received October 16, 2020

Unlike diastereotopic -CH_2- protons from chiral compounds, the diastereotopic protons from achiral compounds may not be easily predictable. For molecules with flexible chain this task is even more problematic. Additionally, the assignment from the experimental ^{1}H NMR spectrum of chemical shifts and coupling constants is very frequently uncertain. Nevertheless, computationally, for both chiral and achiral molecules the -CH_2- protons diastereotopic nature is evident as well as NMR parameters assignments. In the present paper we examine the NMR parameters calculated as a sum of Boltzmann weighted average of the lowest conformations that fits a window less than 10 kJ/mole. A Spartan20 six steps protocol was chosen to select the lower energy conformations. The NMR conformationally averaged parameters obtained for 2-ethyl-2-hyroxybutanoic acid (**EHB)** and citric acid (**CA),** both achiral molecules, and the chiral malic acid (**MA)** fit acceptable well with the experimental data, although the calculation is done for gas phase (conformations dominated by intermolecular hydrogen bonding) and the experiment was carried out in D_2O (conformations controlled by intermolecular hydrogen bonding of highly decorated molecules with OH and COOH that can interact with the solvent). In addition, in the present paper we have examined **EHB**, **CA** and **MA** the lowest energy conformer geometry, Atom in Molecules (AIM) and Natural Bond Orbital protocols to identify the intramolecular hydrogen bonds.



if CIP C_{chiral} > C:
H_a is pro-(S)
H_b is pro-(R)



if CIP $C_{achiral}$ > C:
H_a is pro-(S)
H_b is pro-(R)

INTRODUCTION

Anisochronous[1] (nonequivalent) NMR signals of diastereotopic ligands (protons, fluorine, methyl groups, etc.) have been recognized over the last 50 years to be a significant attraction both in chemical and biochemical research and academia. In achiral NMR solvents, molecules containing constitutionally heterotopic ligands[2-4] and stereochemically diastereotopic ^{1}H and ^{13}C nuclei often give rise to anisochronous resonances with distinct chemical shifts. Ever since and immediately after the discovery in 1957 by W. D. Philips[5] and independently by J. D. Roberts[6] of the non-equivalent diastereotopic property of the *gem*-difluoro (-CF_2-) and methylene protons (-CH_2-)

* Corresponding author: mircea.gheorghiu@sjsu.edu

chemical shifts, a surge of papers have been published with new examples and attempts to rationalize the phenomenon. For about three decades, diastereotopicity was accounted for based on qualitative and speculative conformer distribution, generating abundant concepts and vocabulary that later were tacitly abandoned. Immediately, after computational methods convincingly gained in accuracy and modeling capability, they became the preferential analytical tools to study conformer distribution statistics (Boltzmann distribution) and ultimately the calculation of the weighted average of chemical shifts and coupling constants for flexible molecules.[7-14] The concept of diastereotopicity is crucial in the realm of enzyme biochemistry as highlighted by reactions that specifically involve the recognition of diastereotopic ligands, in these cases protons. For example, fumarase hydratase (EC 4.2.1.2) promotes the water elimination from sodium malate by removing the (pro)-*R* proton[14] while during the Krebs cycle, a pro-*R* hydrogen from citric acid is abstracted and eliminated by aconitase (EC 4.2.1.3).[15]

The stereochemical concept of diastereotopicity denotes the relationship between two groups that if they are replaced with a higher-ranking group will produce diastereoisomers.

Diastereotopicity can be categorized in two groups: the first group comprises molecules that contain diastereotopic ligands and one chiral center and a second group involving diastereotopic ligand surfaces in molecules that are achiral.

In the first group shown in Figure 1(a), the molecule has a prochiral center in the vicinity or over several bonds from the chiral center. With the stereocenter present in the vicinity, the methylene protons become diastereotopic in nature and therefore experimentally they are observed as anisochronous peaks in the ^{1}H NMR spectrum.

A second situation, with very limited number of examples, occurs when the CH_2 protons are nonequivalent despite the absence of stereocenters. Such molecules possess a C_s symmetry element that gives rise to the methylene protons' non-equivalency. In this case, the existence of a stereocenter is not a required structural precondition for protons (ligands) to become diastereotopic. For the second situation, there is some difficulty in delineating diastereotopic nuclei when they are connected to a carbon on a freely rotating σ bond such as those shown in Figure 2.

To this end, we would like to report our in-depth analysis, using a combination of molecular modeling and NMR spectroscopy, of diastereotopic nuclei present in three compounds, two achiral in nature (2-ethyl-2-hydroxybutanoic acid (**EHB**) and citric acid (**CA**)) and one possessing a chiral center (malic acid (**MA**).

EXPERIMENTAL

Citric acid monohydrate (CAS: 5949-29-1; Lot#0833C072) was purchased from Amresco (Fountain Parkway Solon, OH.). DL-Malic acid (CAS: 6915-15-7; Lot# MKBB1053) and 2-ethyl-2-hydroxybutyric acid (CAS: 3639-21-2; Lot# 138436A) were purchased from Aldrich (St. Louis, MO.). ^{1}H NMR (600 MHz) spectra were obtained using a Bruker Avance III 600 MHz instrument equipped with a Bruker TCI 5 mm cryoprobe (Bruker Biospin, Billerica, MA) at 30.0 ± 0.1 °C and were recorded in D_2O and in acetone-d_6. For the spectra recorded in D_2O, the ^{1}H NMR chemical shifts are calibrated with respect to residual partially deuterated water centered at δ = 4.79 ppm, while for those recorded in acetone-d_6 the chemical shifts are calibrated with respect to the residual acetone-d_5 peak centered at δ = 2.05 ppm.[17] For the NMR analyses, all the acid analytes were separately prepared as 0.08 M solutions in D_2O and all solutions had a final volume of 500 μL.

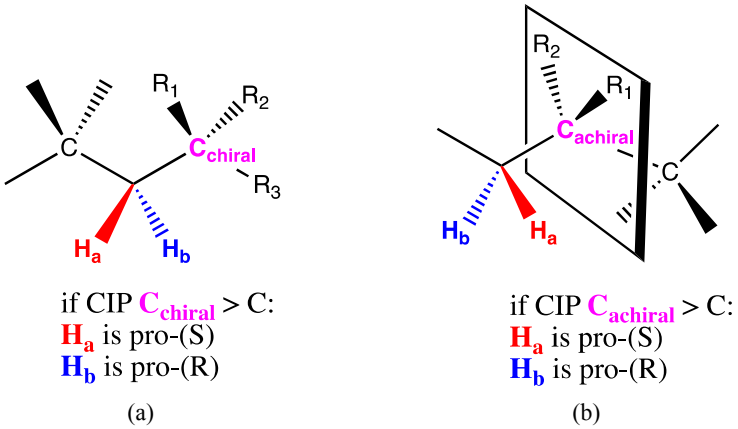
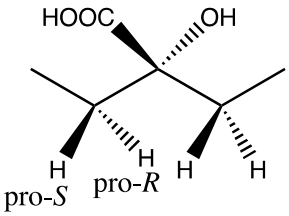
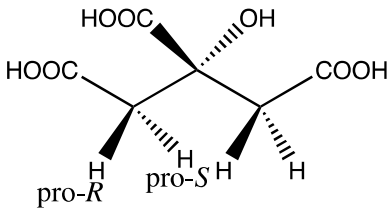


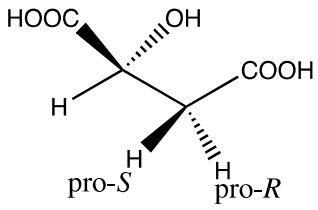
Fig. 1 – (a) Diastereotopic protons in molecules with a chiral center in vicinity and (b) without chiral center in vicinity. CIP stands for Cahn-Ingold-Prelog priorities.[16]



2-ethyl-2-hydroxybutanoic acid
EHB



2-Hydroxy-1,2,3-propanetricarboxylic acid
(citric acid, **CA**)



Hydroxy-butanedioic acid
(malic acid, **MA**)

Fig. 2 – The three hydroxy acids discussed in this paper.

COMPUTATIONAL DETAILS

Calculations have been carried out with the following packages: Spartan 20,[18] Q-Chem 5.3.0 software,[19] AIMAII version 19.10.12,[20] and NBO-7.[21] Geometry optimizations and frequency calculation were done with functionals as the Range-Separated Hybrid Generalized Gradient Approximation wB97X-D[22-23] and wB97X-V[24] (improved over B3LYP[25]), meta-Generalized Gradient Approximation B97M-V[26-27] (include correction for long-range non-bonded interactions). To produce accurate results, for example in NMR calculations, the dual basis set

6-311+G(2DF,2P)[6-311G*] was employed to estimate the Boltzmann factors. It is known that more accurate Boltzmann results are obtained with a large base, however these are time consuming calculations. A more economical time could be obtained using dual basis correction for high accuracy Kohn-Sham density functional theory (DFT) calculations.[28-30] The dual-basis approximation uses *two* atomic orbital basis sets instead of just one. Typically, iterations are performed in the more economic smaller 6-311G* basis set. In the second step an energy correction is calculated in the larger basis set. A single SCF calculation is computed in the larger basis 6-311+G(2DF,2P) in order to build the Fock matrix using the density matrix from

6-311G* basis set calculation. This way, the energy in the large basis is perturbatively approximated at vastly less cost (that could amount to 10 times)[31] than a full SCF calculation carried only in the large basis set:

$$E_{total} = E_{small\ basis} + Tr[(\Delta P) * F]_{large\ basis}$$

in which:

DP = (**P'-P**), is the Kohn-Sham density matrix; **P'** is the projection of the 6-311G* density matrix in the larger space; **P** density matrix defined in the 6-311+G(2DF,2P) basis set.

F the Fock matrix build from **P** (density matrix defined in the 6-311+G(2DF,2P) basis set).

In order to facilitate the understanding the significance of **AIMMII** results, a brief introduction concerning the concepts of Atoms in Molecules is provided next. See more information in reference: https://www.ias.ac.in/article/fulltext/jcsc/128/10/1527-1536.[32] Electron density (r(**r**)), a quantum observable and the energy densities are the vehicle to find an answer for covalent and non-covalent (such as hydrogen bonds) interactions in a molecule. The definition of a ''critical point'' (*CP)* in the electron density is a point in space at which the first derivatives of the electron density vanish:

$$\nabla\rho = \mathbf{i}\frac{d\rho}{dx} + j\frac{d\rho}{dy} + k\frac{d\rho}{dz} = \vec{0}$$

where the zero vector signifies that each individual derivative in the gradient operator ∇ is zero.

A *bond critical point*s (*BCP's*) that appears between two nuclei are points where $\nabla r(r_{bcp})$ vanishes. At critical point two gradient paths are starting and terminate at nuclei producing the *atomic interaction line*. When a structure is fully optimized, all forces on all nuclei vanish and the atomic interaction line becomes a *bond path* (a line of locally maximum density linking two nuclei). The *bond path* is the universal indicator of chemical bonding of all kinds: weak, strong, closed-shell, and open-shell interactions.[33] The collection of bond paths linking the nuclei of bonded atoms in an equilibrium geometry, with the associated critical points, is known as the molecular graph. The second derivatives of electron density are collected in the so-called Hessian matrix when evaluated at $\mathbf{r}_{cp}$. Because the Hessian matrix is real and symmetric by diagonalization becomes:

$$\begin{pmatrix} \frac{\partial^2\rho}{\partial x^2} & \cdots & 0 \\ \vdots & \frac{\partial^2\rho}{\partial y^2} & \vdots \\ 0 & \cdots & \frac{\partial^2\rho}{\partial z^2} \end{pmatrix} - \begin{pmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \lambda_2 & \vdots \\ 0 & \cdots & \lambda_3 \end{pmatrix}$$

The trace of the matrix, which is invariant to rotations of the coordinate system, is known as the Laplacian of the density:

$$\nabla^2\rho(r) = \frac{\partial^2\rho}{\partial x^2} + \frac{\partial^2\rho}{\partial y^2} + \frac{\partial^2\rho}{\partial z^2} = \lambda_1 + \lambda_2 + \lambda_3$$

in which l_1, l_2, l_3 are the curvatures with respect to the three principal axis x', y', z' resulted from the rotation of the original principal axis x, y, z carried out during the diagonalization process.

Critical points are designated as (ω,σ), where ω means the rank of *CP* and σ is its signature. The rank is the number of nonzero eigenvalues of the Hessian of electron density at the critical point defining the principal axes of the curvature of the saddle point; there are no *CP*s with rank less than 3 for the equilibrium structure. The signature is the algebraic sum of the signs of eigenvalues of the Hessian of r. For example, the *BCP* of the hydrogen bond C=O-----H-O is (3,-1) because the rank of an equilibrium geometry (stable structure) is 3 and the signature is -1 since the calculated curvature characteristics of the Hessian matrix one is positive and two have negative eigenvalues. There are four types of stable critical points having three non-zero eigen values:

a. (3, -3) is identified as *nuclear critical point* (*NCP*). All curvatures are negative: r is a local maximum at nuclei.
b. (3, -1) is identified as *bond critical point* (*BCP*). This is a saddle point of r. Two negative curvatures ($-l_1$,$-l_2$) (both perpendicular to the bond path): is a maximum in the plane defined by the corresponding eigenvectors but is a minimum along the third axis along the bond path (l_3, which is perpendicular to this plane).
c. (3, +1) is identified as *ring critical point* (*RCP*). Two positive curvatures: r is a minimum in the plane defined by the corresponding eigenvectors and a maximum along the third axis which is perpendicular to this plane.
d. (3, +3) is identified as *cage critical point* (*CCP*). Three curvatures are positive: r is a local minimum.

Topological parameters such as r, electron density at the *BCP*, the Laplacian at the *BCP*, $\nabla^2 r$ and the energetic properties of *BCP* such as the energy densities G (Lagrangian Form of Kinetic Energy Density), V (potential electron energy density, Virial Field = Potential Energy Density = Trace of Stress Tensor) and H (the total electron energy density) are employed to describe the nature of bonding.

$$Hc = Gc+Vc$$

and the virial theorem (Lagrangian density $L= -¼ \nabla^2 r$)

$$¼ \nabla^2 r = 2Gc+Vc$$

$$L= -¼ \nabla^2 r = K-G$$

RESULTS AND DISCUSSION

1. NMR results

The ^{1}H NMR parameters (chemical shifts and coupling constants) have been calculated for compounds **EHB, MA**, and **CA** as a weighted average of all conformers comprised in a relative window of less or equal with 10 kJ/mol. Then, the calculated NMR parameters are compared with the experimental NMR results obtained at 600 MHz in D_2O (and/or acetone-d_6). It is well-established that the experimental NMR chemical shifts and the coupling constants measured for a molecular system in fast interconversion among its stable conformers are given by the weighted averages of the molar fractions (Boltzmann factors) n_i and the calculated chemical shifts and the coupling constants for the individual conformers in their equilibrium according to the equation (1) and (2).[34-37]

Experimentally, assuming that at room temperature, the rotation around single bonds is fast on the NMR time scale, all the measurements collect only the collective average for all conformers. Therefore, the NMR parameters such as chemical shift (d_i) and the coupling constants (J_i) are averaged over all conformations.[38]

$$\delta = \sum_i n_i \times \delta_i \qquad \text{(equation 1)}$$

and

$$J = \sum_i n_i \times J_i \qquad \text{(equation 2)}$$

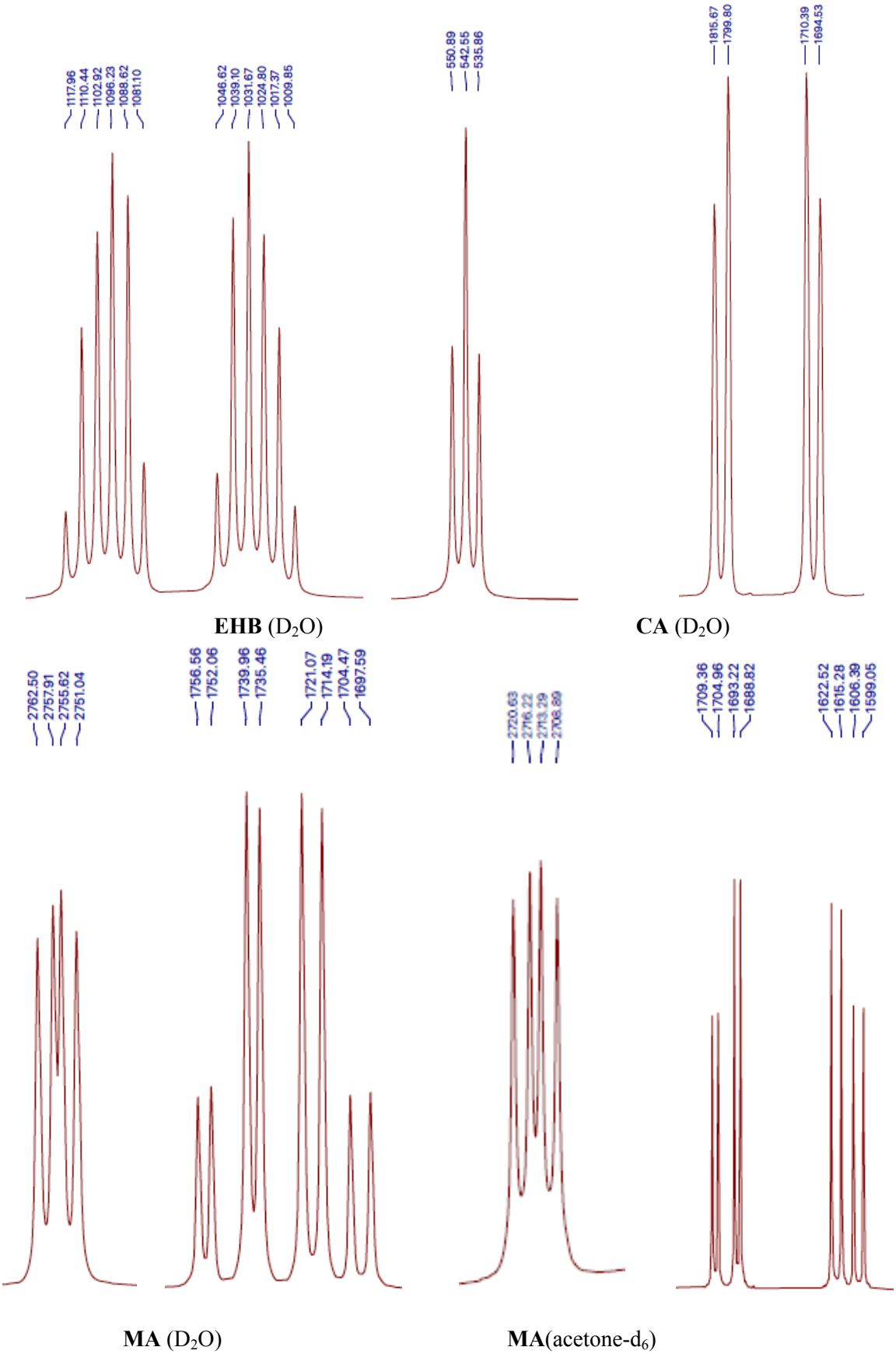


Fig. 3 – Experimental ^{1}H NMRs for 2-ethyl-hydroxybutyric acid (**EHB**), citric acid (**CA**), and malic acid (**MA**).

In equation (1), d is the calculated chemical shift; d_i is the calculated chemical shift of the i^{th} conformer; J is the calculated chemical; coupling, J_i is the calculated coupling constant of the it^{h} conformer, and n_i is the calculated Boltzmann distribution coefficient.

1.1. Calculated ^{1}H NMR parameters for 2-Ethyl-2-hydroxybutyric acid (EHB)

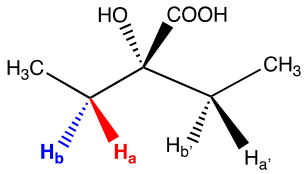
The NMR spectra for the three lowest energy conformers have been calculated based on the geometry optimization resulted from B97X-D/6-31G(d). Boltzmann weights have been calculated as a single point at RWB97XRV/ 6-311+G(2DF,2P) [6-311G*]. Chemical shifts have been calculated as single point with wB97X-D/6-31G* level of theory, and the coupling constants have been calculated with the functional B3LYP using PCJ-0 basis set (which is recommended for NMR calculations), and the option NMR_JISSC=Full (Fermi contact + spin dipolar contribution + paramagnetic spin-orbit contribution + diamagnetic spin-orbit contribution). The lowest energy conformation has a Boltzmann factor of 81.9%. The geometry is presented Figure 8a. The next two conformation have Boltzmann factors 14.2 % and 3.8 %, respectively. The geometries are presented Figure 8a, 8b and 8c. Note that all have the five membered ring intramolecular hydrogen bonded.

The results are presented in Table 1. Comparison of calculated values with the experimental data is gratifying (except the geminal coupling constant), even though the experimental NMR was recorded in D_2O, while calculation is assumed to provide gas-phase NMR data.

The chemical shifts and coupling constants are Boltzmann averaged of the three conformations **8a** (81.9 %), **8b** (14.2%) and **8c** (3.8%). An analysis of conformer distribution provides quite similar results. The calculation was carried out with Spartan20 protocol using method wB97X-V, with the dual basis set 6-311+G(2DF,2P)[6-311G*] and the geometries were obtained with the model wB97XD/6-31G*.

Table 1

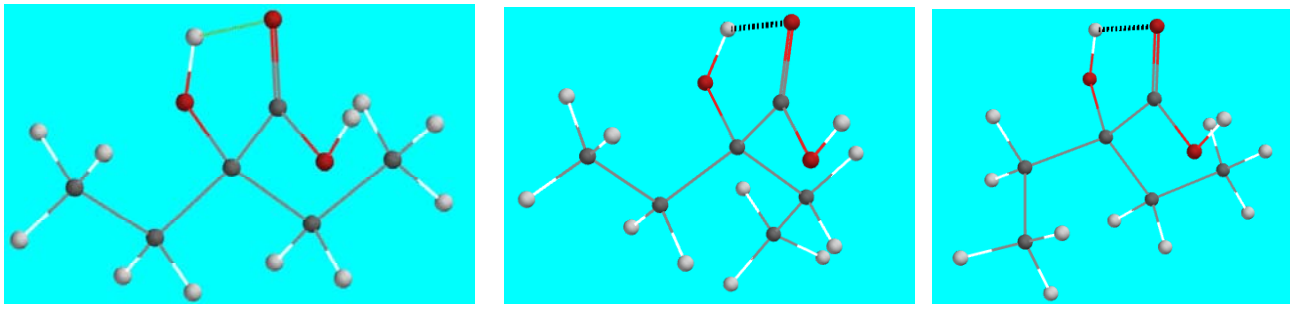
Calculated chemical shifts (600 MHz) and coupling constants for **EHB**.



H_a, $H_{a'}$ are pro-(S)
H_b, $H_{b'}$ are pro-(R)

	H_a ($H_{a'}$) d ppm	H_b ($H_{b'}$) d ppm	CH_3 d ppm	$J_{Ha,Hb}$ (Hz)	$J_{Ha,CH3}$ (Hz)	$J_{Hb,CH3}$ (Hz)
Experimental*	1.83 or 1.71	1.71 or 1.83	0.90	15.02	7.52	7.52
Calculated** (gas phase)	1.87	1.78	0.97	-13.40	7.22	7.46

*In D_2O, 600 MHz
** Calculated (600 MHz) with Spartan'20 release 1.0.0 package.



(a) (b) (c)

Fig. 4 – Room temperature calculated conformations (see also Table 2) of compound **EHB.**

The lowest energy **EHB** conformer has a "W" shape of the four-carbon chain (see Figure 4 (a)). The next two, with relative higher energy conformers, each has the terminal ethyl group rotated with ca. 60°, as it is shown in 4 (b) and 4 (c).

The calculated results (i) confirm the diastereotopicity of methylene protons H_a and H_b, (ii) arguably, assign the chemical shift, fact that is not obvious from experimental results and (iii) suggest a different value for the coupling constant of the diastereotopic methylene hydrogens with the vicinal methyl protons, unfortunately not seen at the resolution provided with a 600 MHz instrument.

1.2. Calculated ^{1}H NMR parameters for Malic acid (MA)

The diastereotopic protons of **MA** are computationally "recognized" the gas phase calculated chemical shifts for protons H_a and H_b are 0.4 ppm apart (see Table 3). In d_6-acetone the same interval is two-times larger than in D_2O. In the gas phase the vicinal coupling constants are smaller than in solution. Perhaps, this fact is due to the change from an intramolecular hydrogen bond-imposed conformation, to intermolecular conformers-imposed conformations as result of the strong interactions of the -OH and -COOH groups with the solvent, yielding a modified conformers population responsible for the average NMR parameters.

1.3. Calculated ^{1}H NMR parameters for Citric acid (CA)

CA has more complicated conformers space than **EHB** or **MA-5** and **MA-6**. The systematic conformer search started at the level of molecular mechanics (MMFF94x) when 725 conformers have been retained (generated by 6 rotatable bonds). The procedure continued through calculation at levels HF/21G followed by wB97X-d/6-31G* and ending with dual base calculation 6-311+G(2df,2p)[6-311G*] using the wB97X-V functional. After successive pruning in the last step only 30 only conformers were considered. Computationally, the CH_2 protons are diastereotopic. Their chemical shift splitting, d_{Hb}-d_{Ha}, is of 0.012 ppm, however, it is smaller than that observed experimentally in D_2O (0.06 ppm).

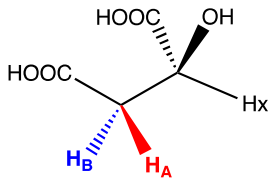
Table 2

EHB conformers relative energetics and statistics calculate with the Spartan 20 protocol described in section **4.1.4**

Conformer	Relative energy (kJ/mol)	Boltzmann factor
4(a)	0.0	0.819
4(b)	4.34	0.142
4(c)	7.58	0.038

Table 3

Calculated Chemical shifts (600 MHz) and Coupling constants for **MA**.



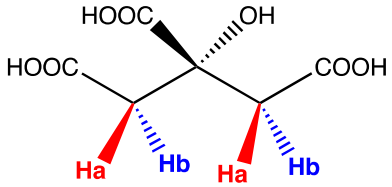
H_A, is pro-(R)
H_B, is pro-(S)

	H_A d ppm	H_B d ppm	H_X d ppm	J_{HAHB} (Hz)	$J_{HA,HX}$ (Hz)	$J_{HB,HX}$ (Hz)
Exp. (D_2O)*	2.91 or 2.85	2.85 or 2.91	4.59	16.6	6.88	4.5
Exp. (acetone- d_6)*	2.83 or 2.68	2.68 or 2.83	4.52	16.17	7.33	4.41
Calculated** (gas phase)	2.75	2.96	4.32	-18.00	4.78	3.18

*600 MHz
** Calculated (600 MHz) Spartan'20 release 1.0.0 package.

Table 4

Calculated Chemical shifts (600 MHz) and Coupling constants for **CA**



	H_a d ppm	H_b d ppm	J_{HaHb} (Hz)
Exp. (D_2O)*	2.91 or 2.85	2.85 or 2.91	16.6
Calculated** (gas phase)	3.06	3.07	-17.7

*600 MHz
** Calculated (600 MHz) Spartan'18 release 1.42 package.

1.4. NMR parameters calculation

Spartan18 and Spartan 20 calculation of NMR parameters is based on a protocol that requires 7 steps:

- Conformer search, geometry optimization with molecular mechanics MMFF94x. A pruning is done to keep molecule with calculated energy less than 40.00 kJ/mole (at most 400)
- All the selected molecules in previous step are subjected to a quick optimization at HF/3-21G(d) level, with TOLG (tolerance of gradient) = 0.0008 and TOLE (tolerance in energy change) =0.00023, pruning window keeps molecules with E<40 kJ/mole (at most 200).
- A single point calculation at level WB97X-D/6-31G(d). Pruning window is <15 kJ/mol (at most 100 molecules).
- All the selected molecules in previous step are subjected to a quick optimization at HF/3-21G(d) level, with TOLG = 0.0008 and TOLE = 0.000076, pruning window keeps molecules with E<10 kJ/mole (at most 50).
- Optimization at level WB97X-D/6-31G(d). Pruning window is <10 kJ/mol (at most 50 molecules), with TOLG = 0.0007 and TOLE = 0.000020
- Using the geometries calculated in the previous step, a single point calculation method WB97X-6 and as basis set 6-311+G(2DF,2P),[6-311G(d)]. Pruning window is <10 kJ/mol (at most 30 molecules). Boltzmann weights are calculated.
- Using the geometries from the optimization step, NMR parameters are calculated JISS=Full\B3LYP\PCJ-0.

2. Equilibrium geometry and frequency calculation

2.1. 2-Ethyl-2-hydroxybutyric acid (EHB)

EHB geometry and frequencies were computed by means of Spartan18 software, using Range Separated Hybrid Generalized Gradient Approximation (RSH-GGA) RWB97X-D/6-311+G(2DF,2P) level of theory. The equilibrium structure (geometry) which is also equilibrium conformer (lowest-energy conformer) is shown in the picture below. The **EHB-5 (**with a hydrogen bond forming a five-membered ring) molecule has a W-shape of the 5-carbons chain (Figure 4). The -OH and -COOH substituents are placed in an approximate orthogonal plane to the 5-carbons chain. The two groups are linked by a weak hydrogen bond. The -C=O-----H-O, the distance is 2.014 Å. Mulliken bond order of hydrogen bond is 0.058. The structure below is a true minimum because has no imaginary frequency (zero order saddle point). The structure without hydrogen bond (which also has no imaginary frequency) is less stable with 19.6 kJ/mol.

The broken lines refer to the hydrogen bond -C=O-----H-O (O2----H7) is also depicted. The $C_1C_2C_3C_4$ dihedral angle 178.9°; $C_5C_4C_3C_2$ -178.9°.

2.2. Malic acid (MA)

Malic acid lowest energy conformation, labelled as **MA-5**, has a five-membered ring hydrogen bond among the geminal OH and -COOH (Figure 5). Malic acid that has a six-membered hydrogen bond, labelled as **MA-6**, has hydrogen bond among the vicinal OH and -COOH. **MA-6** is with ca. 6 kJ/mol higher in energy than **MA-5** [(BM97-M/6-311+G(2dp,2p)]. It is, therefore expected, that **MA-6** is included in the

list of conformers that meets the requirement to be energetically within the 10 kJ/mol window relative to the lowest energy **MA-5**.

2.3. Citric acid (CA)

Citric acid (**CA**) prefers a five membered ring hydrogen bond conformation versus a six membered ring (Figure 6). Thus, starting from a **CA-6,** during the geometry optimization (B97B-V/6-311+G(2df,2p) the output is a **CA-5** conformation. The carbon backbone has a "W" shape, capped with the two COOH groups such as to extend the "W" shape.

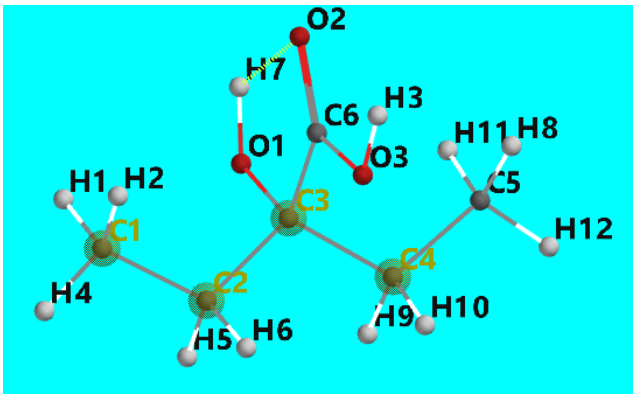
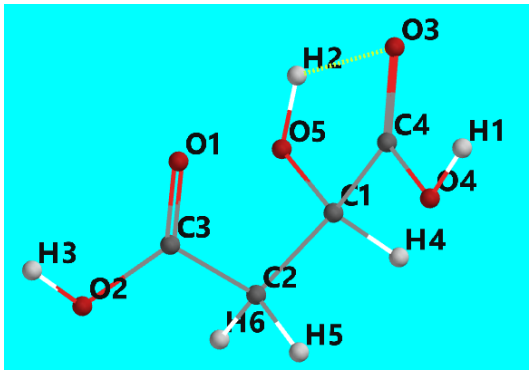
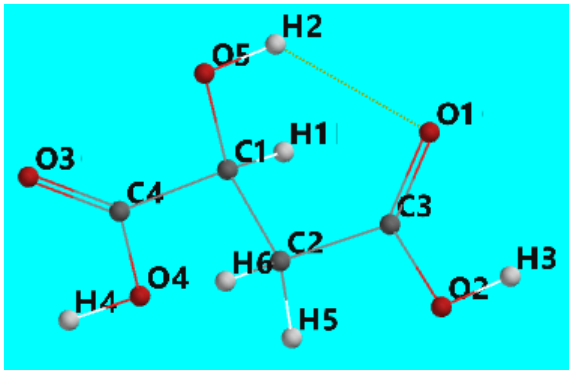


Fig. 4 – The lowest energy conformation of **EHB**.



MA-5
(hydrogen bond among H2---O3)



MA-6
(hydrogen bond among H2---O5)

Fig. 5 – Molecular graphs of MA-**5** and **MA-6**.

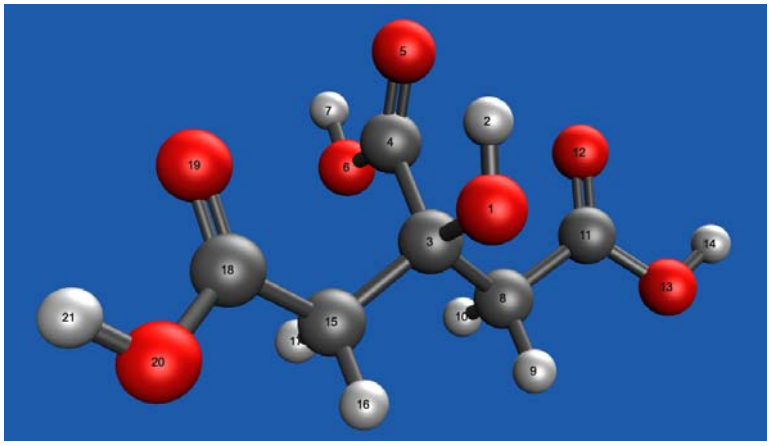
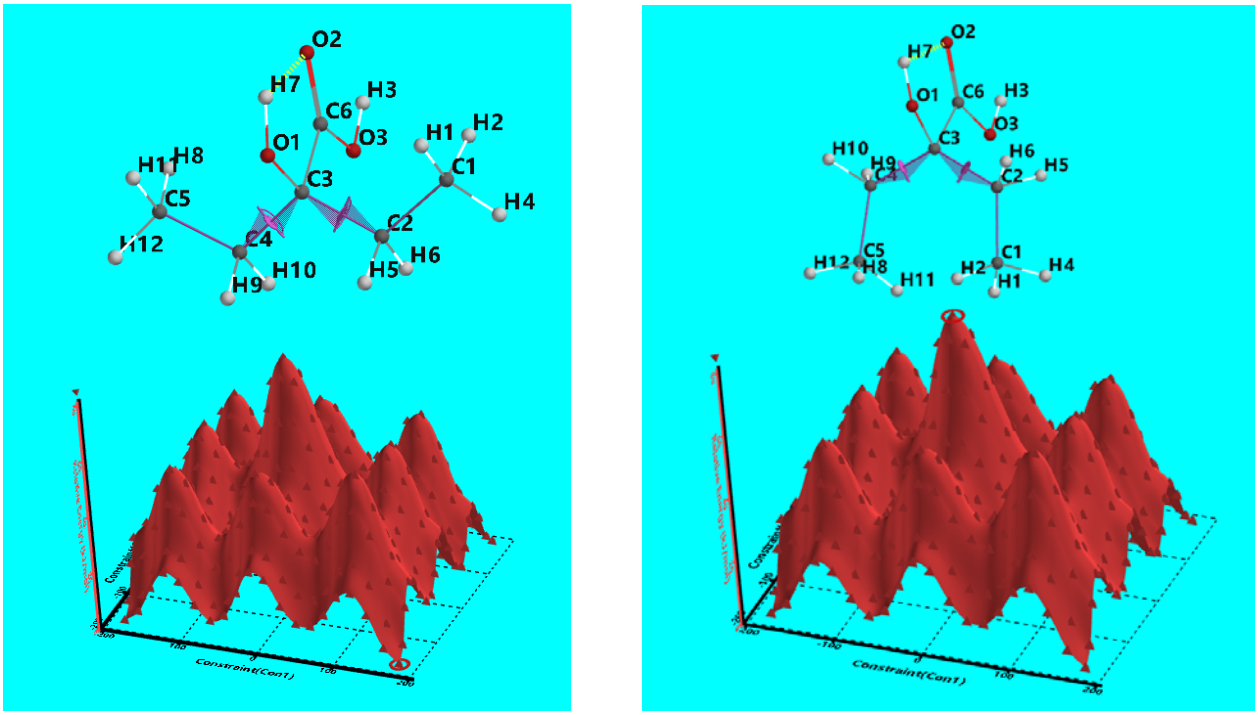


Fig. 6 – A view of the molecular graph of citric acid. C3C8C18C16 = 179.8 C11C8C3C16 = 179.7; C6H2 = 2.035 Å.



(a) (b)

Fig. 6 – The energy (z axis kJ/mol) versus C1C2-C3C4 and C5C4-C3C2 dihedral angles (in °) shown in the structure above with the dihedral constrain markers. In (a) the minimum energy structure is shown by a circle in the right bottom; in (b) the maximum energy structure is shown by a circle at the top.

For all optimized geometry concerning molecules **EHB**, **MA-5, MA-6**, and **CA-5** the frequency calculation did not evidence imaginary frequencies.

3. Conformers distribution

All three compounds chosen to be examined in this paper **EHB**, **MA**, and **CA**, possess flexible structure. Therefore, because of the timescale of NMR, cannot be described for NMR parameters calculation purposes by a single 3D structure. More distinct conformers resulting from rotation around single bonds are likely to exist. Unfortunately, calculations have only rarely been used by practicing chemists to assist in interpretation of spectra. The NMR spectrum for flexible molecules is a weighted average of energy weighted spectra of the individual conformer. Identifying the lowest-energy conformers (say, within a window of 10 kJ/mol) for a molecule with multiple degrees of conformational freedom, is a major and a serious challenge. Additionally, providing accurate Boltzmann weights must result from a systematic search of conformer distribution space. Boltzmann coefficients are utilized to obtain the average chemical shifts and coupling constants. To secure, for example, that the **EHB** conformer shown in Figure 6 is indeed the equilibrium structure (geometry) which is also equilibrium conformer (lowest-energy conformer), we examined the energies (wB97X-D/6-31G*) calculated for 400 conformations resulted by rotating around the dihedral angles C1C2-C3C4 and C5C4-C3C2 (see the embedded structures). Indeed, the ground state the conformation is identical with the structure and energy of conformer shown in Figure 6. The 3-D surface relative energy versus dihedrals C1C2-C3C4 and C5C4-C3C2 are presented below.

For flexible molecules, as are the ones examined in this paper, finding the most reliable lowest energy conformer is a challenging enterprise. The conformers distribution space has been scrutinized with ***searchmethod=systematic*** keyword. The following steps, representing one of the most reliable method of conformations search, have the following protocol for optimization of the conformer distribution. An extensive search starts with the molecular mechanics MMFF94,[39] retaining all the conformers with energy less than 40 kJ/mol than the lowest-energy conformer that found. A pool of 73 molecules has resulted this

way for **EHB** molecule. The next steps with successfully better models are gradually reducing the number of conformers that are passed to the next level. Thus, in the second step for each of 73 molecules, the energy and a new Boltzmann coefficient are calculated at HF/3-21G*level of theory. At the end of this step, 35 conformations were retained. The third step is a single point re-calculation of all 35 conformation at the level of B97X-D/6-31G*, followed by a pruning to retain 18 molecules (out of 35) having the relative energy within the 15 kJ/mol window. In the next step the 18 molecules geometries have been optimized with B97M-V/6-311+G** level and pruning with the window of < 10 kcal/mol, only 3 conformers are retained. In the fifth step each of the 3 conformers are subjected to the geometry optimization at level B97M-V/6-311+G** followed by a pruning with a window < 10 kJ/mole and all the 3 conformation have been retained. In the last step the best Boltzmann coefficients are calculated by a single point calculation carried out using the functional B97M-V and the dual basis set 6-311+G(2DF,2P)[6-311G*]. The results are summarized in Table 2.

4. Hydrogen bond(s). AIMAII and NBO analysis

Intramolecular hydrogen bond (IHB) forms chelated substructures within the examined a-hydroxyacids and is expected that this quite rigid structure to control the conformers distribution and ultimately the NMR parameters of diastereotopic protons. There are two possibilities for forming -C=O….H-O- IHB: either with formation of a five membered ring among the geminal HO- and -C=O (COOH) like in **EHB, MA-5 and CA-5**, or a six membered ring when the -OH form an IHB with the vicinal -C=O (-COOH) as shown for **MA-6**, and **CA-6** only.

A strong support that indeed there is a hydrogen bond among -C=O---H-O is provided by Bader's Quantum Theory of Atoms in Molecules (QTAIM)[32, 40-42] analysis performed with AIMAII (version 19.10.12) [20] and NBO-7 (version 7.0.5)[43] computations. The hydrogen bond path with the C=O either with the a-COOH or the b-COOH are illustrated in Figure 7 and details regarding the parameters are given in Table 4.

For example, if the intramolecular hydrogen bond yields a five-membered ring, the ring critical point (*RCP)* is designed as (3, +1) critical point. It reveals two positive curvatures in the ring surface and a negative curvature along an axis perpendicular to the ring surface. All the calculated parameters with AIMII point strongly for -C=O-----H-O hydrogen bond:

a. Weak hydrogen bonding (< 12.0 kcal/mol) shows Laplacian of electron density $\nabla^2 r(r_{BCP})$ and the local energy density $H(r_{BCP})$ ($-K(r_{BCP})$ both positive.[44-45] For example, for **EHB**, $\nabla^2 r_{BCP}$ = 1.101E-01 and ; $H(r_{BCP})$ = 3.402E-03
b. The calculated electron density and the Laplacian at BCP are within accepted range: for example for **EHB**, $r(r_{BCP})$ = 2.382E-02 au is within the range 2.000 E-03 $< r(r_{BCP}) <$ 4.000E-02, and $\nabla^2 r(r_{BCP})$ = 1.101E-01, is within the range 2.400E-02 $< \nabla^2 r(r_{BCP}) <$ 13.900E-02 au.[46]
c. The energetic topological parameter resulted by dividing the negative of the local kinetic energy density to local potential energy $-G(r_{BCP}) / V(r_{BCP})$ = -(2.413)/(-2.07E-02) = 1.16E+00 being greater than 1 indicates that the nature of interaction is purely non-covalent (for example, hydrogen bond).[45-46]

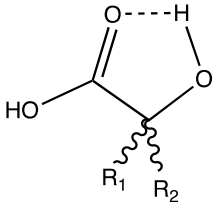
Finally, the last step that confirms a successful AIM analysis is checked with Poincare-Hopf[47] relationship:

$$number_{ncp} - number_{bcp} + number_{rcp} - number_{ccp} = 1$$

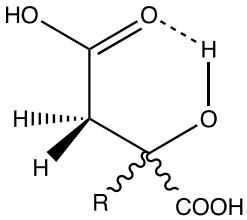
The molecular graph of **EHB** is associated to the following characteristic set: the $number_{ncp}$ is 21, the $number_{bcp}$ is 21, $number_{rcp}$ is 1, $number_{ccp}$ is zero. The number and type of the critical points have to follow the strict topological relationship known as the violation of Poincare-Hopf. This implies an inconsistent characteristic set, that a critical point has been missed, and that a further search for the missing critical point(s) is necessary. The fact that the total topology for **EBH** molecular graph is consistent is the fact that Poincare-Hopf is indeed 1 (21-21+1-0 = 1).

Hydrogen Bonds Critical Points (*BCP*) and Natural Bonds Critical Points (*NBCP*) have also been calculated with the NBO-7 package. Results, analogous to the AIMII calculations, are summarized in Table 6.

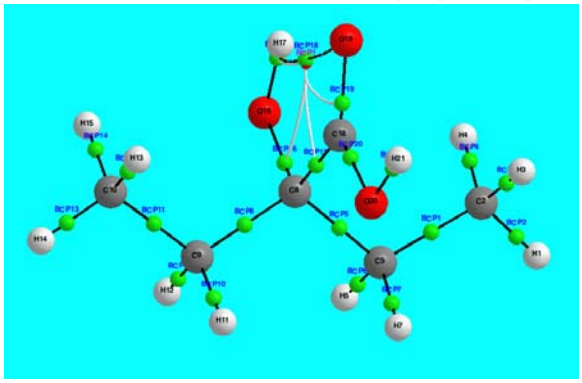
The major NBO-based contributions to hydrogen bonds is similar for **EHB**, **MA-5** and **MA-6**. The bond density r of the hydrogen bond at BCP, for example in **MA-5**, is produced by a 28.2 % contribution of oxygen (O7) lone pair and 31.2 % contribution of O(10)-H(11) [Figure 8 (b)].



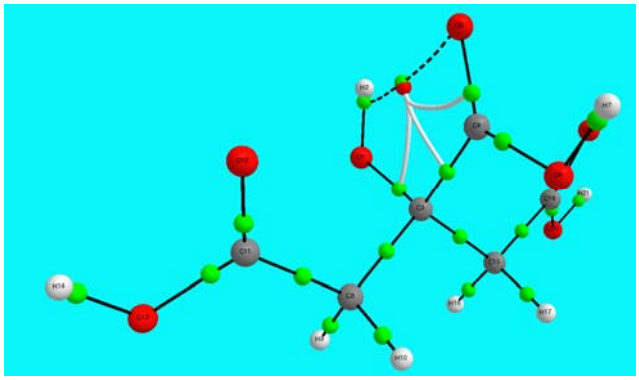
EHB-5: $R_1, R_2 = CH_3$
MA-5: R_1=H, $R_2 = CH_2COOH$
CA-5: $R_1, R_2 = -CH_2COOH$



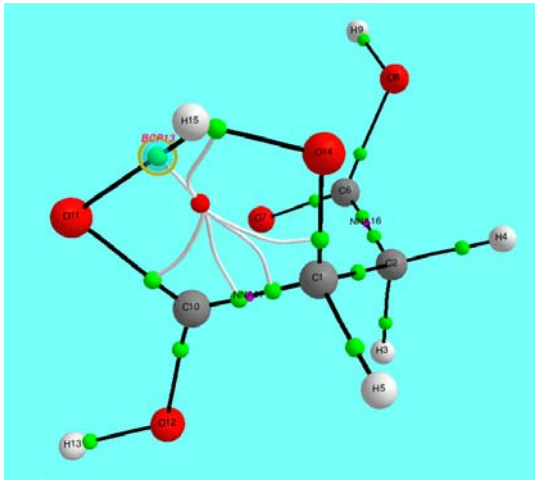
MA-6: R = H
CA-6: R= $-CH_2COOH$



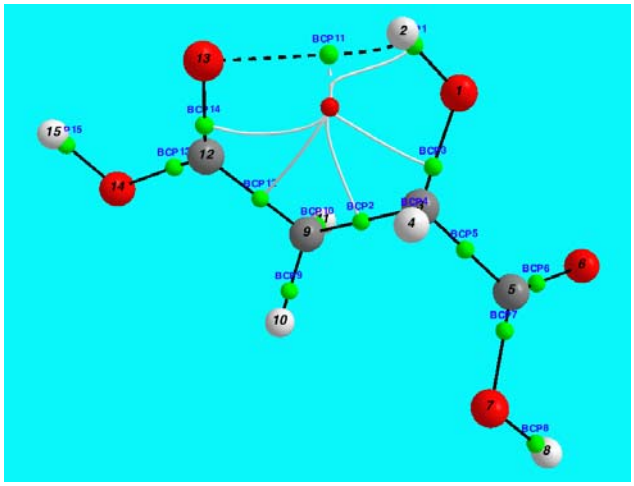
EHB-5 (wB97X-V/6-311++G**)



CA-5 (B97M-V/6-311+G(2df,2p)



MA-five membered ring (**MA-5**)
[B97M-V/6-311+G(2df,2p)]



MA-six membered ring (**MA-6**)
[wB97X-D/6-311+G(2df,2p)]

Fig. 7 – A perspective view of the molecular graph of (a) **EHB**, (b) **CA-5**, and (c) **MA-5** and **MA-6**. Small green spheres represent the *BCP*, and the small red sphere represents the ring critical point (*RCP*). *RCPs* path to *BCPs* are shown in white colors.

Table 5

Calculated **AIMII** parameters for hydrogen bonds*

	$r(r_{BCP})$	$\nabla^2 r(r_{BCP})$	K (-H)	G	V	L
EHB-5	0.02381	+0.11013	-0.00340	+0.02413	-0.02073	-0.02753
CA-5	0.02482	+0.10924	-0.00315	+0.02416	-0.02101	-0.02731
MA-5	0.13021	+0.25103	+0.06614	+0.12890	-0.19505	-0.06276
MA-6	0.01806	+0.06430	-0.001417	+0.014659	-0.01324	-0.01608

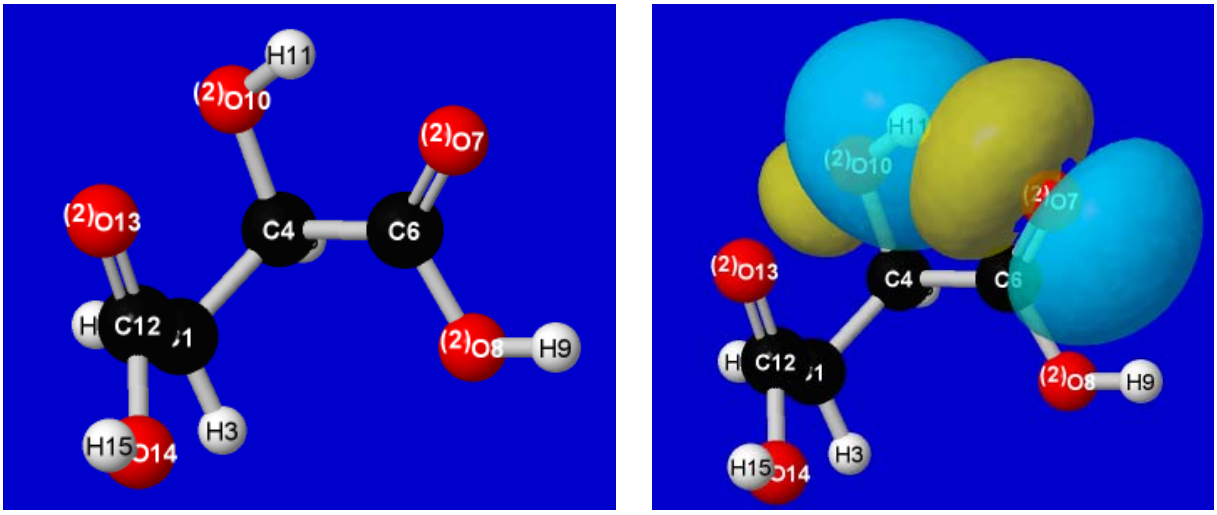
*See the text for meaning of the Table headings and **3. Computational details**.

Table 6

Topological (3,-1) critical points

	r_{CP}	$\nabla^2(r_{CP})$	Laplacian Curvature (l_1, l_2, l_3)		
EHB					
BCP	0.0235	0.1038	0.1377	-0.0076	-0.0262
NBCP	0.0190	0.0646	0.1235	-0.0274	-0.0315
MA-5					
BCP	0.0296	0.1280	0.1878	-0.0220	-0.0378
NBCP	0.0258	0.0936	0.1812	-0.0408	-0.0468
MA-6					
BCP	0.0332	0.1272	0.2207	-0.0450	-0.0485
NBCP	0.0341	0.1159	0.2339	-0.0568	-0.0613

BCP (total density); NBCP (Natural Atomic Orbital atomic densities)



(a) (b)

Fig. 8 – (a) The **MA-5** skeleton and (b) the major NBO contributions of O7 lone pair and O10-H11 bond.

CONCLUSIONS

We have examined computationally and confirmed the CH_2 diastereotopic protons from two Cs-symmetrical compounds (**EHB** and **CA**), compounds that are achiral, and the CH_2 diastereotopic protons from the chiral **MA**. Unlike diastereotopic protons from chiral compounds, at first glance, the diastereotopic protons from achiral compounds may not be easily identifiable, however computationally, their diastereotopic nature is evident. A good matching of calculated versus experimental chemical shifts resulted for **EHB**. The calculated proton NMR parameters for **MA** and **CA** are acceptable, even though they are highly decorated with OH and COOH that can interact with the NMR solvent. Perhaps, the experimental conformations of the **MA** and **CA** in D_2O are determined by the stronger intermolecular hydrogen bonding with the solvent, while in gas phase calculation the dominant conformers are the result of intramolecular hydrogen bonding. In addition to this, the NMR analyses provided the opportunity to examine the lowest energy conformer geometry, AIM and NBO hydrogen bonds.

Acknowledgements. We are grateful to Dr. Todd A. Keith for many clarification of issues conserning **AIMII** package, to Don Garka for assistance of Q-Chem installation, and Yvonne L. Johnson (Lawrence Livermore National Laboratory) for prompt and efficient library assistance.

REFERENCES

1. B. Gerhard and G. R. Franzen, *J. Am. Chem. Soc.*, **1969,** *91*, 3999.
2. K. Mislow and M. Raban, "Stereoisomeric Relationships of Groups in Molecules", In "In Topics in Stereochemisty", N. L. Allinger and E. L. Eliel (Eds.), John Wiley & Sons, Inc, 1967; Vol. 1, 1–38.
3. H. Hirschmann and K. R. Hanson, *Eur. J. Biochem.,* **1971,** *22*, 301.
4. E. L. Eliel, *J. Chem. Educ.,* **1980,** *57*, 52.

5. J. J. Drysdale and W. D. Philips, *J. Am. Chem. Soc.,* **1957,** *79,* 52.
6. P. M. Nair and J. D. Roberts, *J. Am. Chem. Soc.,* **1957,** *79*, 4565.
7. C. B. Braga and R. Rittner, *J. Phys. Chem. A,* **2017,** *121*, 729.
8. K. M. S. Gonçalves, D. R. Garcia, T. C. Ramalho, J. D. Figueroa-Villar and M. P. Freitas, *J. Phys. Chem. A,* **2013,** *117*, 10980.
9. P. Borowski, *J. Magn.Reson.,* **2012,** *214*, 1.
10. F Freire, J. M. Seco, E. Quinoa and R. Riguera, *Org. Lett.,* **2010,** *12*, 208.
11. E. Helms, N. Arpaia and M. Widener, *J. Chem. Educ.,* **2007,** *84*, 1328.
12. T. H. Sefzik, D. Turco, R. J. Iuliucci and J. C. Facelli, *J. Phys. Chem. A,* **2005,** *109*, 1180.
13. C. F. Tormena, M. P. Freitas, R. Rittner and R. J. Abraham, *Magn. Reson. Chem.,* **2002,** *40*, 279.
14. G. J. Kasperek and R. F. Pratt, *J. Chem. Educ.,* **1977,** *54*, 515.
15. E. L. Eliel and S. H. Wilen, "Stereochemistry of Carbon Compounds" John Wiley & Sons: New York, 1994.
16. R. S. Cahn and C. Ingold, *Angew. Chem. Int. Ed. Engl.,* **1966,** *5*, 385.
17. H. E. K. Gottlieb and A. Nudelman, *J. Org. Chem.,* **1997,** *21*, 7512.
18. *Spartan' 18 (version 1.4.4)*, Wavefunction, Inc.: Irvine, CA, 2019.
19. Y. Shao, Z. Gan, E. Epifanovsky, A. T. B. Gilbert, M. Wormit, J. Kussmann, A. W. Lange, A. Behn, J. Deng and X. Feng, *Mol. Phys.,* **2015,** *113*, 184.
20. T. A. Keith, *AIMAll (Version 19.10.12)*, TK Gristmill Software, Overland Park KS, USA, 2019.
21. E. D. Glendening, C. R. Landis and F. Weinhold, *J. Comput. Chem.,* **2013,** *34*, 1429.
22. J.-D. Chai and M. Head-Gordon, *J. Chem. Phys,* **2008,** *128*, 0841106.
23. J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.,* **2008,** *10*, 6615.
24. N. Mardirossian and M. Head-Gordon, *Phys. Chem. Chem. Phys.,* **2014,** *16*, 9904.
25. A. D. Becke, *J. Chem. Phys.,* **1993,** *98*, 5648.
26. N. Mardirossian and M. Head-Gordon, *J. Chem. Phys.,* **2015,** *142*, 074111.
27. N. Mardirossian and M. Head-Gordon, *J. Chem. Phys.,* **2016,** *144*, 214110.
28. R. P. Steele, R. A. Jr. DiStasio, Y. Shao, J. Kong and M. Head-Gordon, *J. Chem. Phys.,* **2006,** *125*, 074108/1.
29. R. P. Steele, Y. Shao, R. A. Jr. DiStasio and M. Head-Gordon, *J. Phys. Chem. A,* **2006,** *110*, 13915.
30. R. P. Steele, R. A. Jr. Di Stasio and M. Head-Gordon, *J. Chem. Theory Comput.,* **2009,** *5*, 1560.
31. R. P. Steele and M. Head-Gordon, *Mol. Phys.,* **2007,** *105*, 2455.
32. P. S. V. Kumar, V. Raghavendra and V. Subramanian, *J. Chem. Sci.,* **2016,** *128*, 1527.
33. R. F. W. Bader, *J. Phys. Chem. A,* **1998,** *102*, 7314.
34. J. Jazwinki, "Theoretical aspects of indirect spin-spin couplings", in "Nuclear Magnetic Resonance", K. Kamieńska-Trela (Ed.), The Royal Society of Chemistry: Cambridge, U.K., Vol. 42, 2013, p. 152−180.
35. T. Helgaker, M. Jaszunski and K. Ruud, *Chem. Rev.,* **1999,** *99*, 293.
36. W. J. E. Parr and T. Schaefer, *Acc. Chem. Res.,* **1980,** *13*, 400.
37. R. J. Abraham and E. Bretschneider, "Medium effects on rotational and conformational equilibria", in "Internal Rotation in Molecule", W. J. Orville-Thomas (Ed.), Academic Press: London, 1974.
38. E. L. Eliel, *Chem. Ind. (London),* **1959**, 568.
39. T. A. Halgren, *J. Comput. Chem.,* **1999,** *20*, 730.
40. R. F. W. Bader, "Atoms in Molecules: A Quantum Theory", Oxford University Pres: Oxford, 1990.
41. R. F. W. Bader, *Chem. Rev.,* **1991,** *91*, 893.
42. R. F. W. Bader, *Acc. Chem. Res.,* **1985,** *18*, 9.
43. E. D. Glendening, J. K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, P. Karafiloglou, C. R. Landis and F. Weinhold, Theoretical Chemistry Institute, University of Wisconsin: Madison, WI 2019.
44. U. Koch and P. L. A. Popelier, *J. Phys. Chem. A,* **1995,** *99*, 9747.
45. I. Rozas, I. Alkorta and J. Elguero, *J. Am. Chem. Soc.,* **2000,** *122*, 11154.
46. M. Ziolkowski, *J. Phys. Chem. A,* **2006,** *110*, 6514.
47. C. F. Matta and R. J. Boyd, "An Introduction to the Quantum Theory of Atoms in Molecules", in "The Quantum Theory of Atoms in Molecules", C. F. Matta and R. J. Boyd (Eds.), WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim, 2007, 1–34.