Preprint

This is the submitted version of a paper published in *Journal of Physical Organic Chemistry*.

The Importance of Orientation of Exo-cyclic Groups in a Naphthoxyloside: A Specific Rotation Calculation Study

Paula N. Goulart,[a] Clarissa O. da Silva,[a,*] and GöranWidmalm[b,*]

a. Departamento de Química – ICE/UFRRJ Rodovia BR465, km 7, Seropédica, RJ, Brazil 23897-000. E-mail: clarissa-dq@ufrrj.br

b. Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, SE-106 91, Stockholm, Sweden. E-mail: goran.widmalm@su.se

Abstract: 2-naphthyl β-D-xylopyranoside (XylNap) inhibitsβ-1,4-galactosyltransferase 7 (β4GalT7)and thereby growth of tumor cells both *in vitro* and *in vivo*. The binding pocket of β4GalT7 has a defined orientation of hydrogen bond acceptors and hydrophobic moiety. Knowing the orientation of the hydroxyl and naphthyl groups of this molecule would help in the development of more efficient inhibitors. In this work we have tried, for the first time, to determine the exo-cyclic hydroxyl and aglycon groups orientation of XylNap, using *ab initio* descriptions, and calculation of the specific rotation values, in methanol solutions, using two different solvent descriptions: a dielectric continuum approach (PCM) and a microsolvated+continuum approach (MS+PCM). In the PCM approach $[\alpha]_D$ = -59 deg/(dm(g/cm^3)) whereas for the MS+PCM approach $[\alpha]_D$ = -29 deg/(dm(g/cm^3)). The latter is in excellent agreement with the experimentally determined value in methanol solution, viz., $[\alpha]_D$= -30 deg/(dm(g/cm^3)). This agreement allow us to say that the hydroxyl groups have similar orientations in xylose and XylNap, and the naphthyl group has a very well defined ψ dihedral angle value in the most abundant conformations.

Keywords: xylopyranose; naphthyl xylose; specific rotation

.

Introduction

Proteoglycans are glycoconjugates consisting of a protein core with covalently linked glycan chains. Nearly all mammalian cells produce these proteoglycans that, *inter alia*, are secreted into the extracellular matrix or inserted into the plasma membrane. [1] The carbohydrate portion consists of linear polysaccharides known as glycosaminoglycans (GAGs) having disaccharide building blocks of hexosamine and uronic acids or galactose. In these GAGs two classes of polysaccharides, namely, chondroitin sulfate/dermatan sulfate and heparan sulfate/heparin, are joined via a tetrasaccharide linker to a serine residue of the peptide chain. The first sugar residue of the core region is the five-carbon sugar D-xylose, [2,3] a rare sugar in humans but omnipresent in wood. The D-Xyl*p*-L-Ser structure acts as an acceptor for the β4GalT7 galactosyl transferase, [4] which adds the second sugar of the tetrasaccharide core region.

Priming of GAG chains is possible using different 2-naphthyl β-D-xylopyranoside (XylNap) derivatives [5] or triazole ring-containing xylosides.[6] Interestingly, a modified XylNap, where position 6 of the naphthyl group has been hydroxylated, can inhibit growth of tumor cells both *in vitro* and *in vivo*. Due to its unique structure, D-xylose and Nap-derived glycosides thereof can fit snugly in the binding pocket of the β4GalT7 enzyme, with key hydrogen bonds between sugar hydroxyl groups and protein residues as well as stacking interactions to the hydrophobic patches of the sugar and, when present, also to the naphthyl aglycone group.[7] Importantly, modifications of the xylose moiety of these glycosides can render inhibitors to the galactosylation process and may be valuable tools in the development of anti-tumor agents.

The intrinsic flexibility of xylose and derivatives thereof, [8] the hydrogen bonding abilities, and the favoured orientation of the naphthyl group are all of vital importance in understanding its conformational behavior. To gain further insight into its molecular properties we herein investigate the XylNap system by first principles quantum mechanics methods and validate the predicted conformational preferences by comparing the experimentally determined specific rotation of the molecule in methanol solution to that computed by *ab initio* methods.

Specific Rotation ($[\alpha]_D$)

Recently the possibility of using a theoretical description to reproduce specific rotation (SR or $[\alpha]_{589}$) values of monosaccharides in aqueous solution has been suggested in the literature to validate monosaccharide conformational sampling [9,10,11]. This procedure has been proven to be effective for glucose [9], xylose [12], levoglucosan[13], and other carbohydrate systems.[11]

Determining the theoretical SR value for such molecules is a challenging task because it consists of computation of the relative abundance of each conformation and its individual contribution to the final SR value. The latter is described as an average in a solution, weighted by the Boltzmann population values of the corresponding conformer. Even for monosaccharides (structurally the most simple carbohydrates), the polyhydroxylated compounds exhibit a huge conformational versatility for the hydroxyl groups, and many conformations are possible as the energy differences among them are very small at room temperature (with such differences often being smaller than the error of the computational method used to describe them). For rigid molecules at their equilibrium geometries, Stephens and coworkers [14] reported a RMS deviation of approx.

28.9 for calculated and experimental SR values for molecules with small SR values in a study of 65 compounds, using a TDDFT/GIAO B3LYP/aug-cc-pVDZ//B3LYP/6-31G(d) description. In these cases, the optical activity was related to the unique chiral center present in the molecules studied.

Despite of these facts, in two recent studies we have demonstrated for monosaccharides that: 1) the difference in SR values among different conformations is larger than the difference due to changes in the description level, and 2) that the hydroxyl group orientation of each chiral center influences the individual SR value of the particular conformer [15]. Both characteristics are prerequisites for a property that is intended to be used in a validation procedure of any conformational set assumed as able to physically represent the corresponding monosaccharide's aqueous solution behavior. Therefore, in this work we have chosen to test the ability of the SR property to unravel the XylNap conformational preferences.

Methods

Due to very small energy differences among monosaccharide conformations (their hydroxyl groups may take several different orientations), quantum mechanical calculations were chosen to describe the conformations in this work. With this methodological choice, the protocol adopted for the conformational search is a delicate task, since the topography of the potential energy surface is not automatically covered as in molecular dynamics methods.

The density functional B3LYP was employed together with the 6-31+G(d,p) basis set functions in all geometry optimization and frequency calculations [16,17,18]. For the isolated system, the calculations were performed in the Jaguar [19] program and those mimicking the methanol solution utilized the Gaussian03 program.[20]

The 1-3-6 nomenclature, here presented for the first time, was introduced to identify all conformations investigated. The circle in Figure 1 being the Newman projection for each C-O bond, is a transversal cross section of the axis of the Ox-Cx bond, where x=1, 2, 3 or 4. It is divided into three sectors, each identified by a number corresponding to the dashed line; these numbers define a range of values inside the sector that may be accepted for the corresponding dihedral angle, which defines the staggered orientation of the corresponding group (hydroxyl or naphthyl).

For instance, the conformation presented in Figure 2 is named 3636 because the dihedral angle values that define the groups orientation are ϕ = C_β-O1-C1-O5= 275.28°, θ_2 = H2-O2-C2-C1= 70.95°, θ_3 = H3-O3-C3-C2 = 297.71° and θ_4 = H4-O4-C4-C3 = 56.90°. There is another important dihedral angle denoted ψ_{Nap} = C1-O1-C_β-C_α= 201.14°, for which there is no preferred orientation, *a priori*, since it is not sterically hindered. Its value is explicitly indicated in the conformer identification (for instance 3636_125.00°).

The initial conformations for the isolated system

The three most abundant conformations found for β-D-xylose in aqueous solution were utilized (3636, 6636 and 1111) [12] and the naphthyl group was linked to the O1 atom, replacing the hydrogen atom of this hydroxyl group, in each case. The geometry of these structures was optimized (Figure 2).

The all-antiperiplanar 1111 conformation was not obtained, even considering different values for ψ_{Nap}, since when ϕ dihedral values are limited to sector 1, steric hindrance and a 1,3-diaxial interaction between the naphthyl and the hydroxyl groups at C2 in the non-*exo*-anomeric conformation will occur. The torsion angle ψ_{Nap} was not restricted during the geometry optimization calculations, unless otherwise mentioned.

The search for stable conformations

In this step, 81 starting conformations were investigated (3 for each θ_n, n=2-4, and 3 for ϕ) and their initial geometries were obtained by assuming that the staggered orientation of the hydroxyl and naphthyl groups are preferred to any eclipsed one. It should be emphasized that the ψ_{Nap} dihedral angle was always optimized without any restraints. Geometry optimization calculations were performed for the system in the gas-phase, and 39 stable conformations were found in a range of 9.63 kcal/mol from the global energy minimum.

Scanning the ψ_{Nap} degree of freedom

The ψ_{Nap} dihedral angle was fully scanned from 0 to 360°, in 18 steps of 20° each, for conformers 3636 (counter clockwise) [21] and 6636, both of which have a counter clockwise orientation around the pyranose ring, because they were the most stable ones for xylose. The energy range covered by variation of the ψ_{Nap} dihedral angles was found to be 3.5 kcal/mol, at the description level adopted, by rotating the O1-C_β bond. Minimum energy values were found for six different ψ_{Nap} values, viz., 15°, 100°, 150°, 200°, 250° and 330°. They were combined with all stable conformations found in the range of up to 7 kcal/mol (from the global energy minimum), chosen since it is twice the energy gap found for O1-C_β bond rotation. Twelve conformations were chosen for ϕ and θ_n, n=2-4, that were found within this energy gap: 3111, 3113, 3161, 3163, 3166, 3336, 3611, 3613, 3636, 6636, 6111, 6336, and a new set of 72 (12x6) geometry optimization calculations were performed. The starting geometry of each conformation investigated was obtained from the combination of ψ_{Nap}, ϕ and θ_n (n=2-4) values. Forty-three different stable conformations were obtained by this procedure, within a range of 8.55 kcal/mol from the global energy minimum.
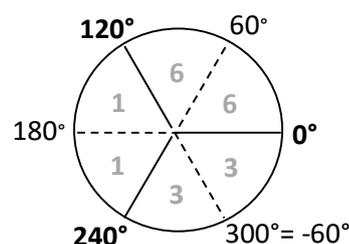
Figure 1: 1-3-6 nomenclature, based on the orientation of the group defined by the dihedral angle values in the sequence ϕ, θ_2, θ_3 and θ_4: C_β-O1-C1-O5 and Hn-On-Cn-C(n-1), n=2-4.
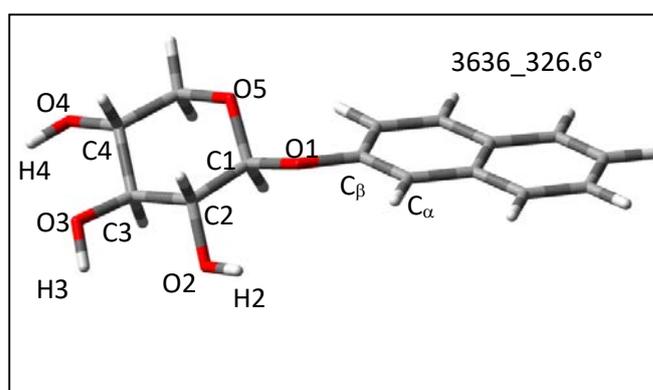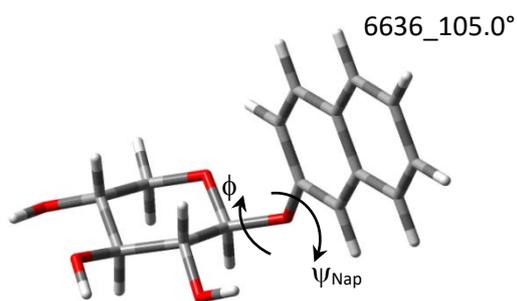
6636_105.0°

Figure 2: Optimized geometry (B3LYP/6-31+G(d,p)) of XylNap for the 3636_326.6° and 6636_105.0° conformations.

Solvation using a polarizable continuum model for methanol and specific rotation calculations

From the 43 stable conformations found, those comprised in a range of 5 kcal/mol from the global energy minimum were selected to **be 'solvated in methanol'**. The solvation process was considered using the Integral Equation Formalism of the Polarizable Continuum Model (PCM) [22], in which the solute is introduced into a cavity with a molecular shape opened in a continuum dielectric medium, formed by interlocking spheres centered on atoms or groups of atoms. The radii of the spheres are 2.40 Å for CH or CH_2 groups, 2.28 Å for C and CH if C has an sp^2 hybridization, 1.80 Å for an O atom and 1.44 Å for an H atom bonded to the oxygen atom of hydroxyl groups [23]. In the PCM model we computed the solvation energy (G_{solv}) as a summation of the electrostatic (G_{elect}) and non-electrostatic ($G_{non-elect}$) components (the latter component is obtained from the computation of the dispersion, repulsion and cavitation terms [24]), corrected by thermal (translational, rotational and vibrational corrections) and entropic contributions for the solute at 298.15K.

In the evaluation of the conformations found we have used the specific rotation value of XylNap in methanol solution. The specific rotation ($[\alpha]_D$) values were calculated using only B3LYP/6-311++G(2d,2p), as suggested in the literature for rigid systems [25]. The individual $[\alpha]_D$ values were obtained for each conformation found as representative in PCM, following the TD-DFT/GIAO approach, which describes the optical rotation at a frequency ν of a chiral molecule in an achiral solvent (dilute solution) as

$$\phi(\nu) = \frac{16\pi^3 N_m \nu^2}{c^2} \widetilde{\beta}_m \quad (1)$$

where $\phi(\nu)$ is the rotation in radians per centimeter, N_m is the number of solute molecules/cm³, and $\widetilde{\beta}_m$ is the effective electric dipole-magnetic dipole polarizability of the solute molecule in the solvent. All details of the mathematical treatment may be found in the original paper and references therein. [26] The specific rotation (SR) can be obtained from the optical rotation by dividing the optical rotation by the concentration of the sample, in grams/cm³. The results are then reported as specific rotation (SR or $[\alpha]_D$) values

calculated at the sodium D line wavelength. The specific rotation $[\alpha]_D$ is the observed rotation corrected for concentration c (g/mL) and a defined path length l (dm), specifically, $[\alpha]_D = \alpha_{obs}/(c \times l)$. In a rigorous treatment, the experimental SR value itself should also be investigated with respect to its dependence on the concentration [27]. In this work, we assume that each conformer contributes to the SR final value by an amount defined by its relative abundance. The final SR solution value is calculated as a weighted average of these individual values. The weights are obtained from the Boltzmann distribution calculated for the system. We model a methanolic XylNap solution when it reaches the thermodynamic equilibrium at room temperature conditions, in an infinite dilution situation, that is assumed to have an extensive entropy [28].

Microsolvation in the gas-phase and in PCM

An explicit methanol molecule from the bulk was also considered, to recover, to some extent, the effects of hydrogen bonds that are established between the solvent and the hydroxyl groups of the solute, which sometimes may affect the $[\alpha]_D$ individual values. [13] To do so, it is necessary to properly identify the preferential site for this solvent molecule for the three different hydroxyl groups of the molecule. Considering that the nature of this solute-solvent interaction is predominantly electrostatic [29], the charge values located on the apparent surface (obtained from PCM calculations) were compared for the hydroxyl groups, and the methanol molecule was subsequently oriented at the preferential site, chosen as the one that has the largest charge value. If such a preferred site on the solute is an oxygen atom, the hydrogen bond was tested with both lone pairs of the solute, generating two different clusters to be investigated, while if it is a hydrogen atom, the two possibilities refer to the lone pairs of the oxygen atom of the explicit solvent molecule (methanol). If two atoms have very similar values of their charges located on the apparent surface, both centers were investigated. New geometry optimization calculations were then performed; first in the gas-phase and then using the PCM. Subsequently, frequency calculations were performed again, Boltzmann population values were obtained and specific rotation values recalculated.

Results and Discussion

From the geometry optimization calculations of the 81 conformations, 39 stable conformers were found (see Electronic Supplementary Information – ESI for energetic ordering), with very different values for the ψ_{Nap} torsion angle. There is a larger flexibility for this degree of freedom when the ϕ is approximately around 300°, as can be seen from Figure 3. Although we considered all staggered orientations possible for ϕ and θ_n torsion angles, n=2-4, ψ_{Nap} was not fully scanned at this point.

In order to take into account at least some part of the conformational versatility that the naphthyl group brings to the molecular system, in Figure 4 we report the scan performed for ψ_{Nap}, for both the 3636 and 6636 conformations. Thus, six different ψ_{Nap} values are identified defining local minima to the system: 15°, 100°, 150°, 200°, 250° and 330°. This indicates that there is an interaction between the naphthyl

substituent and the xylose hydroxyl groups that, at least to some extent, affects the global conformation.

Then, the combination of these six ψ_{Nap} dihedral angle values with the 12 conformations found previously (energy range < 7 kcal/mol) generated 72 conformers that furnished, after new geometry optimization calculations, the 43 stable conformations identified in Table 1. During the geometry optimizations, the orientation of the hydroxyl groups of these conformers was not changed (for instance, conformer 3613 was not converted into 3111), even with the proximity of the naphthyl group.

Again, the conformations with ϕ around 60° present less conformational versatilities for ψ_{Nap} than those with ϕ around 300°, and are also less abundant (vide infra). All 43 conformations cover a range of 8.55 kcal/mol, and those that are 5 kcal/mol above the global energy minimum, were selected to be solvated using the PCM. Subsequently, 17 conformations were selected and information on these is compiled in Table 2. After geometry optimization calculations using the PCM, no appreciable changes were observed, except for some conformers where ψ_{Nap} values changed slightly.

Comparing the results obtained for both isolated and solvated systems, we note that the global energy minimum has the same conformation, and that two conformations (3636 and 3111) are responsible for more than 50% of the XylNap in MeOH solution (PCM), even though there is a larger conformational diversity in solution than for the isolated system. However, the computed theoretical value found for $[\alpha]_D$ from such conformational set suggests that there is room for improvements in our description.

For these conformations, the quantities ΔE (the electronic energy for the isolated system) and ΔG_{solv} (standard Gibbs free energy for the solvated system) values are shown in Figure 5, and in Table 2 we have highlighted the stable conformations found in MeOH solution with relative abundance larger than 1%; the eleven different conformations of XylNap are presented in Figure 6.

For this conformational set obtained from the polarizable continuum medium description mimicking a MeOH solution, the ψ_{Nap} values are around 10° and 200°, while in the gas-phase (isolated system) other naphthyl orientations are observed (see Table 2). In both naphthyl orientations found as stable in methanol solutions, the π–electrons can be considered as 'aligned', to some extent, with the oxygen lone pairs of the O1 atom.

Table 1: Geometrical parameters for the 43 stable conformations found for XylNap, as an isolated system.

| ϕ and θ_n, n=2-4	ψ_{Nap}/°
3111	3.29, 20.11, 131.70, 206.98
3113	19.91, 100.00, 148.69, 205.06, 328.10
3161	22.10, 122.11, 210.10, 330.00
3163	19.81, 107.10, 147.12, 202.09, 328.10
3166	19.82, 103.02, 145.89, 200.48, 215.05, 250.10
3336	19.78, 104.00, 207.36
3611	14.58, 22.68, 100.56, 201.16
3613	13.39, 106.10, 201.66
3636	15.86, 109.76, 203.31
6111	100.03, 255.00
6336	104.63, 323.01
6636	134.79, 323.34

Table 2: ψ_{Nap} values for the system in the gas phase and methanol solution using the PCM; energy differences (ΔE is the relative electronic energy in the gas phase and ΔG_{solv} is the relative standard Gibbs free energy in the PCM (kcal/mol)). Population and specific rotation ($[\alpha]_D$ in deg/(dm(g/cm³))) values in MeOH solution.

| Conf.	ψ_{Nap} (gas)/°	ΔE	ψ_{Nap} (PCM)/°	ΔG_{solv} (PCM)	Pop. (PCM) %	$[\alpha]_D$
3636	15.86	0.00	12.12	0.00	41.47	-31.05
3636	203.31	0.69	201.14	0.68	13.13	-125.61
3636	109.76	1.50	124.46	2.43	0.68	138.14
3111	20.11	3.13	10.58	0.57	15.79	-66.07
3336	207.36	3.64	200.32	2.02	1.37	-136.45
3336	19.78	3.65	10.84	1.42	3.79	-40.34
6636	323.34	3.80	323.10	6.55	0.00	-139.70
3111	206.98	3.90	197.49	1.32	4.46	-138.35
6636	134.79	3.95	130.74	5.70	0.00	296.51
3611	14.58	3.99	12.05	0.96	8.15	-58.13
3113	19.91	4.25	10.72	1.29	4.71	-29.83
3166	19.82	4.38	10.89	1.68	2.40	-32.00
3111	131.70	4.59	131.34	2.74	0.40	114.80
3611	201.16	4.71	199.24	3.45	0.12	-152.57
3613	13.39	4.91	11.77	1.58	2.86	-25.64
3113	205.06	5.00	198.41	1.83	1.87	-120.61
3166	200.48	5.00	199.63	2.23	0.96	-
Calc.						-59.27
Expt.						-30.04[30]

In Table 2, the results reported for all conformations in the gas-phase refer to electronic energy, ONLY, without ZPE or any other physical correction. On the other hand, all results presented for the conformations in PCM with or without micro-solvation, contain ZPE, thermal and entropic corrections, that lead the reported quantities to the status of $G^0_{298.15K}$. The description in the gas phase is performed to furnish an improved electronic density to describe the intramolecular interactions, since the solute-solvent interaction is treated in PCM as a perturbation of the Hamiltonian of the system in the gas-phase.
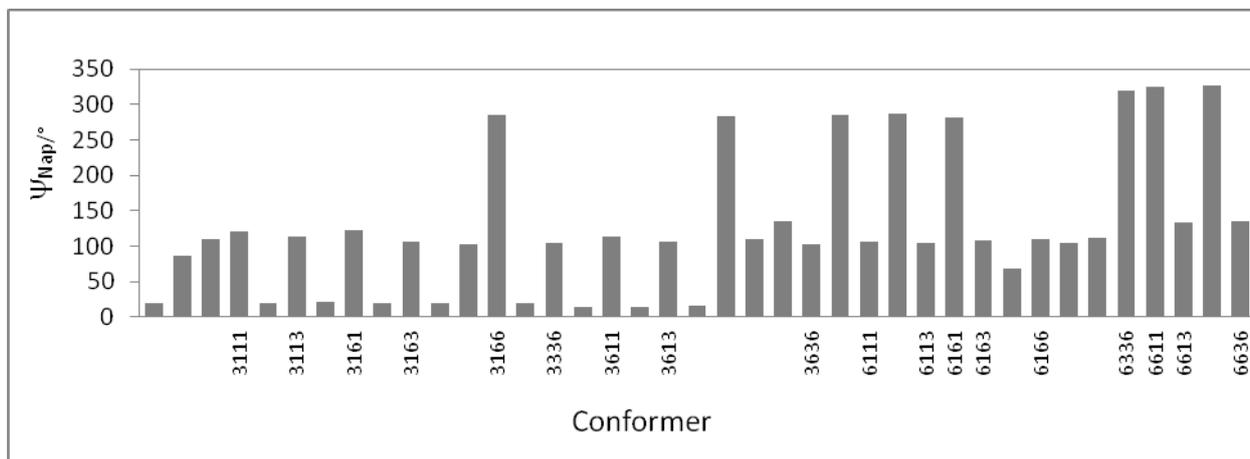
Figure 3: Naphthyl dihedral angle identification of the stable conformers found, considering all 81 staggered orientations for ϕ and θ_n (n=2-4) torsion angles.
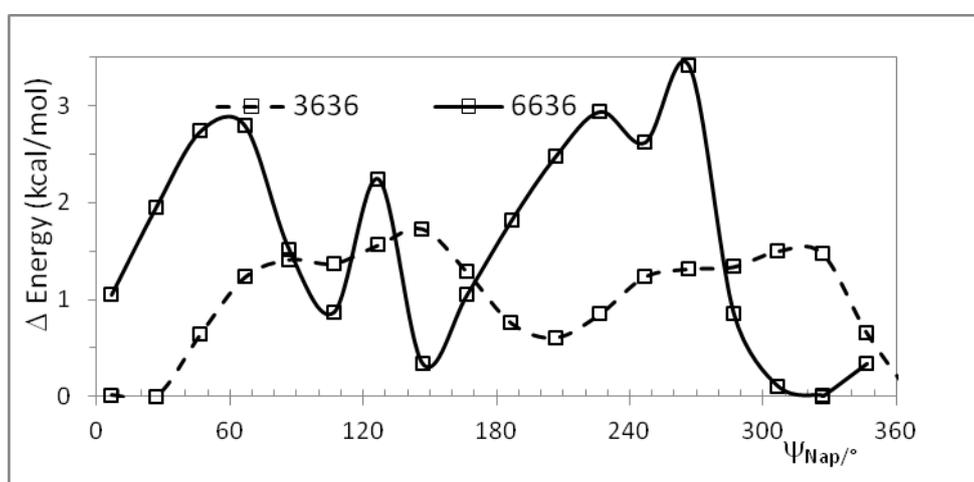
**Figure 4:** Scan on the ψ_{Nap} dihedral angle for 3636 and 6636 conformations.
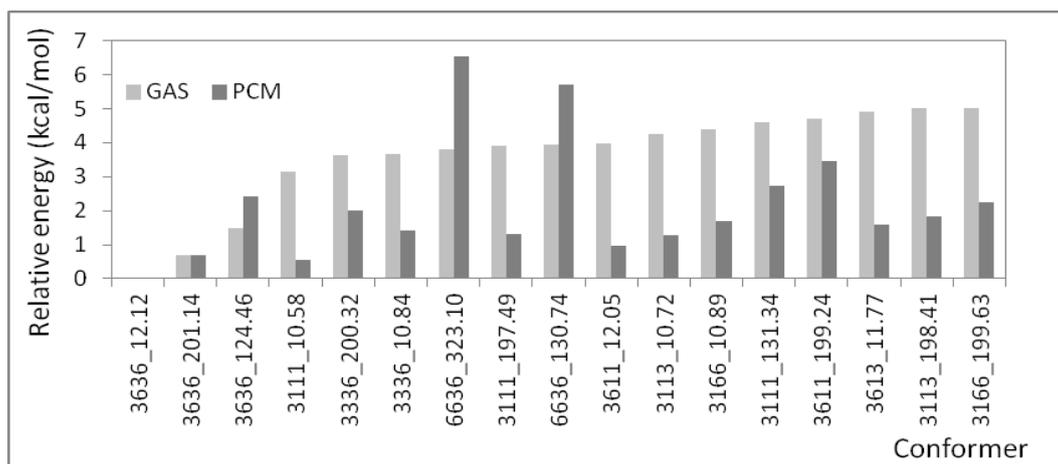
Figure 5: ΔE and ΔG$_{solv}$ (kcal/mol) for XylNap representative conformations in the gas-phase and MeOH solution, corresponding to room temperature.

5

Even though very large deviations can be observed for calculated and experimental specific rotations values of rigid molecules [26], the validation of the stable conformations found is questionable if a disagreement as that reported in Table 2 is found. Therefore, a microsolvation step (MS) was considered, to evaluate how the introduction of specific solute-solvent interactions affects the individual values calculated for $[\alpha]_D$. This choice is based on the electrostatic nature [29] of the hydrogen bond interaction, which is established between the hydroxyl groups of XylNap and the hydroxyl group of methanol, and the effects that such interaction may have on the SR property [31].

However, for a polyhydroxylated compound, the identification of the solvation preferential site (even qualitatively) is not immediate. Thus, a quantitative criterion was adopted, considering the electrostatic nature of the hydrogen bond between solute and solvent. The charge values located on the apparent surface of charges of the excluding solvent surface in the PCM were collected for each atom of the different hydroxyl groups of each stable conformation (see Figure 2 for numbering of reference atoms and Graph 1 in the ESI). The site chosen for microsolvation was the one with the highest charge value found. As previously mentioned in the Methods section, two different hydrogen bonds for each conformation were tested, considering both oxygen lone pairs of an O atom, if it was the preferential site. For conformation 3111_197.49° O2 and H4 presented very similar charge values on their cavity surfaces, and then, both sites were investigated. We propose that these different solvation clusters need to be considered, because they may be related to different solvation patterns, which are known to have different physico-chemical properties as presented and discussed elsewhere. [32,33,34]
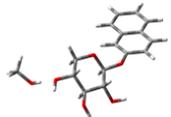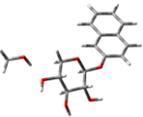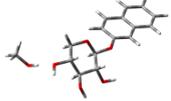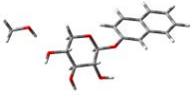
It should be mentioned that the explicit consideration of solvent molecules during the geometry optimization may change the conformation, and the final identification of the conformation is needed.

In Figure 7, eleven conformations of micro-solvated XylNap using the PCM are reported, with their corresponding Boltzmann population and specific rotation values. It might be argued whether it is reasonable to consider all these clusters in the average weighted by the Boltzmann population values, because in doing so, two different solvation sites are considered for the same conformation. It might also be considered as an unbalanced description, because some conformations have only one solvation site (3113_12.24° and 3613_14.00°), while others have four (3111_197.49°). However, the impact of such an unbalanced weighted average is minimized due to the small contribution to the Boltzmann population final v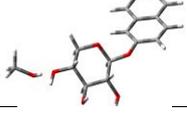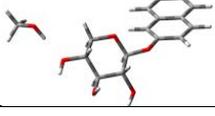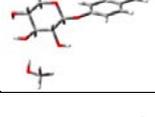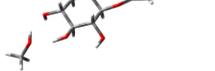alue obtained in the mentioned cases. They were reported with the purpose to evaluate how the micro-solvation affects the $[\alpha]_D$ individual values. With this purpose, several aspects of data reported in Table 3 may be analyzed. In Figure 8 we compare the effects of considering a solvent molecule explicitly for each conformer, on the specific rotation individual value, on the geometry and finally on the energy value of each conformation, respectively, which was found as abundant using the PCM.
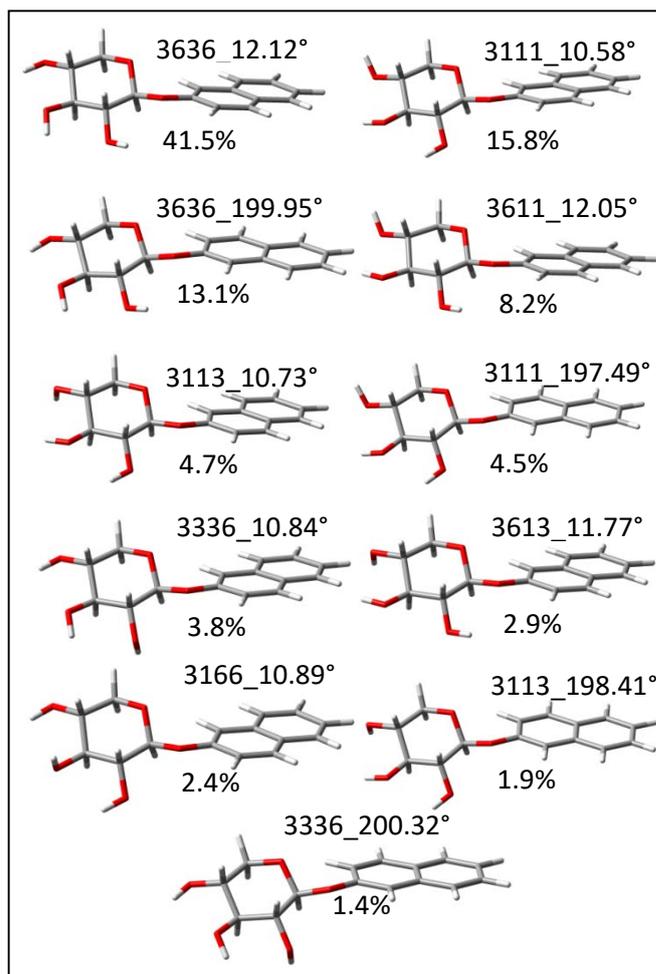
Figure 6: The eleven most abundant conformations found using the PCM for XylNap in MeOH solution at room temperature conditions. Boltzmann population values are indicated in %.

In Figure 8 (a), we circled the most abundant (in PCM and MS+PCM) conformations. The individual $[\alpha]_D$ values have changed significantly in both descriptions for 3111 with torsion angle values of 10.58° and 197.49° for ψ_{Nap}. Particularly, in the latter case when the micro-solvation site is the H4 atom. Then, it can be concluded that it is not necessarily a geometrical change in the carbohydrate moiety induced by the presence of a methanol molecule that brings about a change in the $[\alpha]_D$ values [35], since there is no appreciable change in the ψ_{Nap} values (or even in other dihedral angles of the molecule) as can be seen from Figure 8 (b).

Conformer	Cluster 1	Cluster 2
3636_12.12°		
3636_201.14°		
3111_10.58°		
3336_200.32°		
3336_10.84°		
3111_197.49°		
3611_12.05°*		
3113_10.72°		
3166_10.89°		
3613_11.77°		
3113_198.41°		
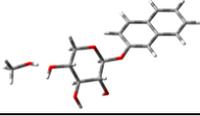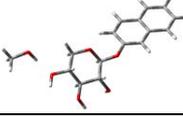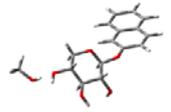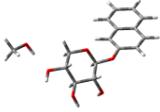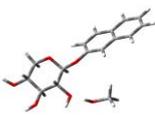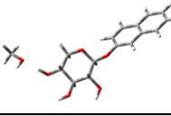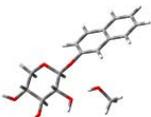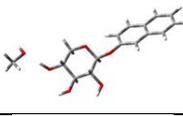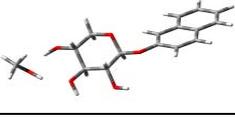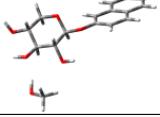
Figure 7: Conformers of micro-solvated XylNap using the PCM. The corresponding energy and property values are reported in Table 3. Full geometric information can be found in the ESI file. *Cluster 1 was transformed during geometry optimization to 3636_11.46°.

Table 3: Torsion angle ψ_{Nap} values of XylNap for the system in methanol solution as described by PCM and micro-solvation+PCM for Clusters 1 and 2, respectively, being the solvation site(s) reported, population (%) and specific rotation ($[\alpha]_D$ in deg/(dm(g/cm³))) values.

	In PCM			MS+PCM			
Conf.	$\psi_{Nap}/°$	Pop. (%)	$[\alpha]_D$	$\psi_{Nap}/°$	Solv.site	Pop. (%)	$[\alpha]_D$
3636	12.12	41.47	-31.05	13.05	O4	0.90	-41.75
				13.92	O4	1.59	-31.39
3636	201.14	13.13	-125.61	199.10	O4	0.90	-122.10
				198.99	O4	0.50	-145.15
3111	10.58	15.79	-66.07	13.09	O2	0.60	-13.98
				9.85	O2	0.13	-74.69
3336	200.32	1.37	-136.45	199.96	O4	0.04	-138.36
				192.56	O4	0.01	-66.91
3336	10.84	3.79	-40.34	11.64	O4	0.09	-50.32
				12.80	O4	0.08	-90.93
3111	197.49	4.46	-138.35	197.65	O2	0.02	-142.42
				204.22	O2	0.27	-198.73
				195.92	H4	3.41	-73.68
				198.69	H4	2.61	-179.24
3611	12.05	8.15	-58.13	11.46 [a]	O3	83.96	-21.21
				14.59	O3	3.12	-18.20
3113	10.72	4.71	-29.83	12.24 [b]	O2	0.00	11.08
3166	10.89	2.40	-32.00	12.55	O4	0.05	-28.96
				11.87	O4	0.08	-54.08
3613	11.77	2.86	-25.64	14.00 [b]	O3	1.02	16.56
3113	198.41	1.87	-120.61	196.44	O4	0.05	-153.12
				196.93	O4	0.06	-92.98
Calc.			-59.27				-29.68
Expt.			-30.04[29]				-30.04[30]

a) Conformation 3611 (Cluster 1) changed to 3636 during the geometry optimization calculations;
b) only one geometry optimization of the two possible ones converged for the solvation site.

Another important change refers to the 3611 conformation, which in Cluster 1 (one of two hydrogen bond cases investigated - see Figure 7) was converted to a 3636 conformation. Then its $[\alpha]_D$ value should be compared properly, and in doing so, it can be verified that it has not changed appreciably, i.e., from -31.39° in the 3636_13.92° to -21.21° in the optimized and transformed conformation. The main difference among them is the microsolvation site, O4 in the first and O3 in the latter case.

In fact, from Figure 8(c), it is evident that this conformation has become the dominant one for this system, and it could be, at a first glance, attributed to two hydrogen bonds that are established between O3···HO$_{methanol}$···HO4. However, one of the conformations found for 3113_198.41° (Cluster 2) presents a similar hydrogen bond-like interaction between O4···HO$_{methanol}$···HO3, but its abundance is not significant. It should be emphasized that this kind of hydrogen bond interaction is not introduced in any starting geometry. On the contrary, the starting geometries were built maximizing the interaction with each lone pair considered. These resulting structures are obtained during geometry optimization calculations and correspond to minima on the potential energy surface, properly characterized by real vibrational frequency values found.

From the above discussion, the improvement of the theoretically computed final $[\alpha]_D$ value obtained when a solvent molecule is explicitly considered, may be assigned to the large stabilization obtained for the 3636 conformation, together with the decrease in magnitude of its $[\alpha]_D$ value.

If only one solvation cluster for each conformer is considered in a weighted average of the theoretically calculated $[\alpha]_D$ value, namely the most stable situation (system with lowest standard Gibbs free energy), the final $[\alpha]_D$ value found is -24.08 deg/(dm(g/cm^3)), still close to the experimental value, if no other ring conformation was considered besides $^4C_1$. In fact, this ring conformation was found to correspond to 93.7% of the occurrence of XylNap in MeOH solution, while 6.3% are in the ^2S_O ring conformation, based on NMR $^3J_{HH}$ coupling constants. [5]
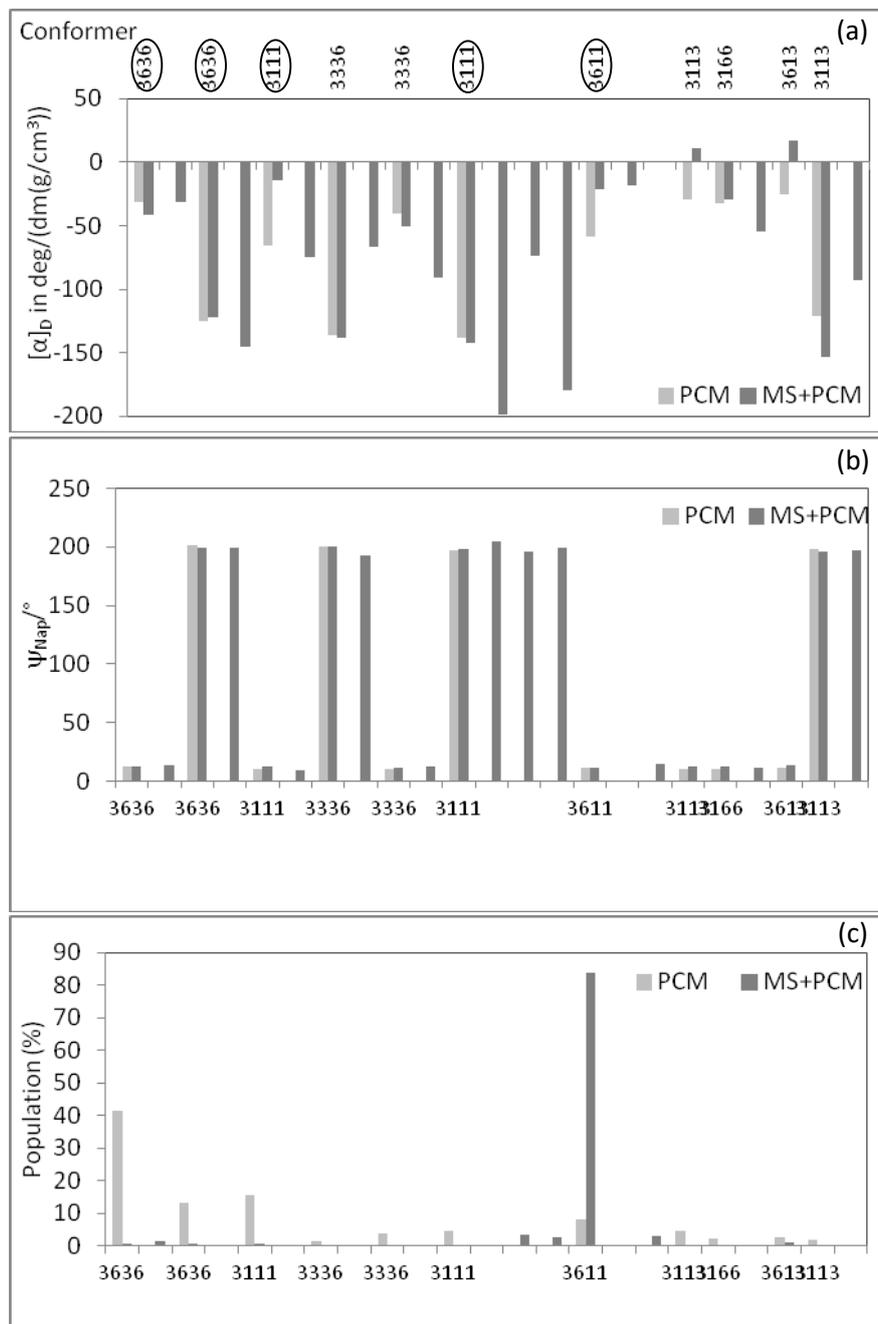
Figure 8: (a) $[\alpha]_D$, (b) ψ_{Nap} and (c) Boltzmann population values for conformations of XylNap listed in Table 3 and Figure 7, following the same sequence of conformations. In (a) the most abundant conformations in both descriptions are circled.

In fact, another full conformational search may perhaps not be justified for the ^2S_O ring conformation. However, we have performed additional calculations with only one conformation, which qualitatively was supposed to be abundant among several others, due to the cooperative effect [36]. The molecular geometry was optimized in the gas-phase and in PCM (Figure 9). Its occurrence in MeOH solution in PCM was found to be 0.39%, and its theoretical individual contribution is $[\alpha]_D$= -339.10 deg/(dm(g/cm³)), when the calculation level and basis set employed are the same as described

above. When the $[\alpha]_D$ weighted average in PCM, comprises the ^2S_O conformation it changes from -59.27 to -60.35 deg/(dm(g/cm³)). The microsolvation step was not considered. It should be mentioned that in all previous systems already studied, no individual conformation has presented such a negative $[\alpha]_D$ value. We are prompted to conclude that the contributions to $[\alpha]_D$ from ^2S_O conformations are always negative, considering the variation of this property already observed for different monosaccharides. [9,10,13,15]
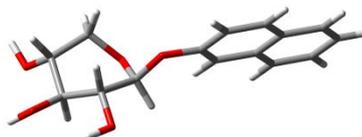
9

Figure 9: The ^2S_O 3111_10.66° conformation of XylNap, with geometrical coordinates optimized for MeOH solution (in PCM and at the B3LYP/6-31+G(d,p) level of theory). The pseudo-dihedral angle defined by the C1-C3-C4-C5 atoms is 2.93°, i.e., the four atoms are essentially in the same plane, which is required for the canonical ^2S_O conformation.

Conclusions

Considering the previous discussion, it can be concluded that xylose [10,] and its derivative XylNap in aqueous and MeOH solution populate in the conformations 3636, 3111 and 6636, independently of the solvent description adopted (dielectric continuum or micro-solvation + continuum).

The naphthyl group in the XylNap derivative brings about a larger conformational versatility to the molecule when the molecule is described as an isolated system than when it is solvated in MeOH (see the values in the second and fourth columns of Table 2). In this latter situation, two orientations for the ψ_{Nap} dihedral angle are preferred: approximately 15° and 200°. It can be stated that in an environment with a low dielectric constant, rotation of the naphthyl group experiences a low barrier of approximately 3.5 kcal/mol.

The experimental specific rotation value is excellently reproduced by the set of conformers identified as the most abundant ones, but only after an explicit solvent molecule is added from the bulk. Such a molecule does not have a geometrical effect on the orientation of the naphthyl group, and only a small effect on the hydroxyl groups of XylNap, as can be inferred from the SR values when they are compared for 3636_12.12° in PCM and MS+PCM (ψ_{Nap}= 13.05° and ψ_{Nap}= 13.92°, respectively) and 3611_12.05° (that was converted during geometry optimizations into 3636_11.46°), which are respectively: -31.05, -41.75, -31.39 and -21.21, all in deg/(dm(g/cm³)). The largest effect was observed on the relative abundance of these conformations that were found respectively as 41.47%, 0.90%, 1.59% and 83.96%, depending on how the solvent interactions are considered. This kind of result reinforces the idea that a suitable solvent/environment description is mandatory to describe the behavior of carbohydrates, and also that the SR analysis allows differentiation of contributing states in the search for representative conformations and their relative abundance.

The XylNap compound is here proposed to be represented in MeOH solution by a combination of the following conformations: 3636_12.12°, 3636_201.14°, 3111_10.58° and 3111_197.49°. The first conformation corresponds to the global minimum and the others to local minima located above it by 0.68 kcal/mol, 0.57 kcal/mol and 1.32 kcal/mol, respectively, in $\Delta G^0_{298.15K}$ values in PCM. These free energy differences become 2.69 kcal/mol, 2.93 kcal/mol and 1.90 kcal/mol, respectively, in the MS+PCM description.

Electronic Supplementary Information (ESI) available: electronic energy values and geometrical data for the most stable conformations are reported.

References

[1] J. D. Esko, K. Kimata, U. Lindahl in *Essentials of Glycobiology*. 2nd ed., A. Varki, R. D. Cummings, J.D. Esko *et al.*, editors. Chapter 16,Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; **2009**.

[2] K. Thorsheim, A. Siegbahn, R. E. Johnsson, H. Stålbrand, S. Manner, G. Widmalm, U. Ellervik., *Carbohydr. Res.* **2015**, *418*, 65-88.

[3] C. Brusa, M. Muzard, C. Rémond, R. Plantier-Royon. *RSC Adv.*, **2015**, *5*, 91026-91055.

[4] A. Siegbahn, S. Manner, A. Persson, E. Tykesson, K. Holmqvist, A. Ochocinska, J. Ronnols, A. Sundin,K. Mani, G. Westergren-Thorsson,G. Widmalm, U. Ellervik. *Chem. Sci.,* **2014**, *5*, 3501-3508.

[5] A.Siegbahn, U. Aili, A. Ochocinska, M. Olofsson, J. Rönnols, K. Mani, G. Widmalm, U. Ellervik. *Bioorg. Med. Chem.* **2011**, *19*, 4114-4126.

[6] A. Chatron-Colliet, C. Brusa, I. Bertin-Jung, S. Gulberti, N. Ramalanjaona, S. Fournel-Gigleux, S. Brézillon, M. Muzard, R. Plantier-Royon, C. Rémond, Y. Wegrowski. *Chemical Biology & Drug Design,* **2016**, DOI: 10.1111/cbdd.12865.

[7] A. Siegbahn, K. Thorsheim, J. Ståhle, S. Manner, C. Hamark, A. Persson, E. Tykesson, K. Mani, G. Westergren-Thorsson, G. Widmalm, U. Ellervik. *Org. Biomol. Chem.,* **2015**, *13*, 3351-3362.

[8] J.Rönnols, S. Manner, A. Siegbahn, U. Ellervik, G. Widmalm. *Org. Biomol. Chem.,* **2013**, *11*, 5465-5472.

[9] C. O.da Silva, B. Mennucci, T. Vreven. *J. Org. Chem.,* **2004**, *69*, 8161-8164.

[10] R. R. Andrade, C. O. da Silva. *Carbohyd. Res.,* **2012**, *350*, 62-70.

[11] J. Kaminský, I. Raich, K. Tomčáková, P. Bouř. *J. Comp. Chem.,* **2010**, *31*, 2213-2224.

[12] R. R. Andrade, C. O. da Silva. *Mini-Rev. Org. Chem.,* **2011**, *8*, 239–248.

[13] A. V. Orlova, R. R. Andrade, C. O. da Silva, A. I. Zinin, L. O. Kononov. *ChemPhysChem.,* **2014**, *15*, 195-207.

[14] P. J. Stephens, D. M. McCann, J. R.Cheeseman, M. J. Frisch. *Chirality,* **2005**, *17*, S52-S64.

[15] B. A. França, C. O. da Silva. *Phys. Chem. Chem. Phys.,* **2014**, *16*, 13096-13102.

[16] G. I. Csonka. *J. Mol. Struct. (THEOCHEM),* **2002**, *584*, 1–4.

[17] G. I. Csonka, A. D. French, G. P. Johnson, C. A. Stortz. *J. Chem. Theory Comp.,* **2009**, *5*, 679–692.

[18] M. Marianski, A. Supady, T. Ingran, M. Schneider, C. Baldauf. *J. Chem. Theory Comp.*, **2016**, *12*, 6157-6168.

[19] Jaguar, version 7.9, Schrödinger, LLC, New York, NY, 2012.

[20] Gaussian 03, Revision C.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C.Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, Gaussian, Inc., Wallingford CT, 2004.

[21] C. Höög, G. Widmalm. *J. Phys. Chem. B,* **2001**, *105*, 6375-6379.

[22] B. Mennucci, E. Cancès, J.Tomasi. *J. Chem. Phys.,* **1997**, *107*, 3032-3037.

[23] A. Bondi. *J. Phys. Chem.,* **1964**, *68*, 441-451.

[24] B. Mennucci, R. Cammi. Eds. Continuum Solvation Models in Chemical Physics: from theory to applications. Wiley & Sons, **2007**.

[25] P. J. Stephens, F. J. Devlin, J. R. Cheeseman, M. J. Frisch. *J. Phys. Chem. A,* **2001**, *105*, 5356-5371.

[26] B. Mennucci, J. Tomasi, R. Cammi, J. R. Cheeseman, M. J. Frisch, F. J. Devlin, S. Gabriel, P. J. Stephens. *J. Phys. Chem. A,* **2002**, *106*, 6102-6113.

[27] P. L. Polavarapu, A.Petrovic, F. Wang. *Chirality,* **2003**, *15*, S143-S149.

[28] C. Tsallis. Introduction to non-extensive statistical mechanics: approaching a complex world. Springer-Verlag, **2009**.

[29] S. J. Grabowski. *Chem. Rev.,* **2011**, *111*, 2597-2625.

[30] J. Rönnols, S. Manner, U. Ellervik, G. Widmalm. *Org. Biomol. Chem.,* **2014**, *12*, 8031-8035.

[31] F. Lipparini, F. Egidi, C. Cappelli, V.Barone. *J. Chem. Theo. Comp.,* **2013**, *9*, 1880-1884.

[32] L. O. Kononov, N. N. Malysheva, A. V. Orlova. *Eur. J. Org. Chem.,* **2009**, 611-616.

[33] a) M. Sedlák.*J. Phys. Chem. B*2006, **110**, 4329–4338; b) M. Sedlák.*J. Phys. Chem. B*2006, **110**, 4339–4345; c) M. Sedlák.*J. Phys. Chem. B*2006, **110**, 13976–13984.

[34] G. Cainelli, P. Galletti, D.Giacomini. *Chem. Soc. Rev.,* **2009**, *38*, 990–1001, and references cited therein.

[35] P. Mukhopadhyay, P. Wipf, D. N. Beratan. *Acc. Chem. Res.,* **2009**, *42*, 809-819.

[36] R. A. Klein. *J. Am. Chem. Soc.,* **2001**, *124*, 13931-13937.

[37] P. E. Mason, G. W. Neilson, J. E. Enderby, M. L. Saboungi, J. W. Brady. *J. Am. Chem. Soc.*, **2005**, *127*, 10991-10998.