

# Raman optical activity of a flavone C-diglycoside: Aqueous solution conformations and absolute configuration



Ítalo H. Calisto<sup>a</sup>, Maysa Furlan<sup>b</sup>, Ewan W. Blanch<sup>c</sup>, João M. Batista Jr.<sup>a,\*</sup>

<sup>a</sup> Department of Chemistry, Federal University of São Carlos—UFSCar, Rod. Washington Luís km 235, São Carlos, SP 13565-905, Brazil

<sup>b</sup> Institute of Chemistry, São Paulo State University—UNESP, Rua Prof. Francisco Degni, 55, Araraquara, SP 14800-060, Brazil

<sup>c</sup> School of Applied Sciences, RMIT University, GPO Box 2476, Melbourne, VIC 3001, Australia

## ARTICLE INFO

### Article history:

Received 29 April 2016

Received in revised form 11 July 2016

Accepted 8 August 2016

Available online 8 August 2016

### Keywords:

DFT

QM/MM

ROA

Disaccharides

Molecular dynamics

## ABSTRACT

The combination of experimental and theoretical Raman optical activity data for the flavone C-diglycoside isoswertisin-4'-methyl-ether-2'' $\alpha$ -L-rhamnoside provided detailed information about its conformational preferences in aqueous solution as well as the absolute configuration of the glycosidic linkage. This work also demonstrated the importance of explicit solvation for accurate predictions of the conformational ensemble and vibrational optical activity properties of natural products in water, as opposed to gas-phase or polarizable continuum model calculations.

© 2016 Elsevier B.V. All rights reserved.

## 1. Introduction

Flavone C-glycosides are among the most numerous groups of glycosylated flavonoids found in nature, and investigations of these compounds remain one of the most active areas of flavonoid research. Despite their importance and wide occurrence, the determination of the anomeric configuration of flavonoid glycosides based exclusively on NMR coupling constants may be questionable in many cases [1]. A recent report from our research group [2] revealed that the conformational preferences of the saccharide moiety of the ellagitannin corilagin, in different solvents, were erroneously interpreted as the presence of its  $\alpha$ -configured anomer, isocorilagin. In this latter case, the correct configuration of the molecule in methanol and DMSO solutions was determined by using vibrational and electronic circular dichroism spectroscopies. Vibrational optical activity (VOA) methods have been frequently used for configurational assessment of different classes of natural products [3] however, no reports so far included flavone diglycosides. Electronic circular dichroism (ECD) is often used in flavonoid chemistry [4], but the lack of UV chromophores within the glycosidic portion of the molecules limits its applicability. On the other hand, vibrational circular dichroism (VCD) and Raman optical activity (ROA) have been

demonstrated to be very sensitive to carbohydrate structure and stereochemistry [5,6]. ROA is especially useful for water-soluble molecules and some previous spectra-structure relationships have been established for the stereochemistry of disaccharides [7]. In this work, we describe the application of experimental and calculated Raman/ROA data for the stereochemical characterization of the flavone C-diglycoside isoswertisin-4'-methyl-ether-2'' $\alpha$ -L-rhamnoside (**1**), isolated from *Peperomia obtusifolia* (Piperaceae) [8]. The configuration for the  $\alpha$ -L-Rha-(1 $\rightarrow$ 2)- $\beta$ -D-Glc linkage in **1** was first determined based solely on the NMR coupling constants measured in DMSO- $d_6$ . However, both  $\alpha$  and  $\beta$ -configured molecules would generate coupling constants of similar magnitudes due to equatorial-equatorial or axial-equatorial arrangements of H-1''' and H-2'''. Therefore, in order to tackle this problem, we decided to apply a computational approach for both anomers including molecular dynamics (MD) simulations with explicit water molecules, followed by quantum mechanics/molecular mechanics (QM/MM) geometry optimizations and Raman/ROA calculations to best reproduce experimental data. This procedure was compared to regular molecular mechanics (MM) conformational searches, and both gas-phase and polarizable continuum model (PCM) vibrational optical activity calculations.

## 2. Results and discussion

The doublet at  $\delta$  4.95 ( $J$  = 1.5 Hz, DMSO- $d_6$ ) observed by Mota et al. [8] in the  $^1\text{H}$  NMR spectrum of **1** was originally assigned to

\* Corresponding author.

E-mail address: [joaombj@hotmail.com](mailto:joaombj@hotmail.com) (J.M. Batista).

H-1''' and interpreted as the  $\alpha$  linkage between D-glucose and L-rhamnose. As carbohydrates in aqueous solution are favorable samples for ROA studies providing rich and informative band structure over a wide range of the vibrational spectrum, the Raman/ROA spectra of **1** was obtained (Fig. 1) in order to confirm the NMR assignment. The backscattered scattered circular polarization (SCP) ROA spectrum of **1** in water revealed the presence of characteristic carbohydrate bands, including the +,– (low to high wavenumber) couplet centered at  $885\text{ cm}^{-1}$ , as well as the band pattern +,– (low to high wavenumber) at around  $400\text{ cm}^{-1}$ , which are considered spectral markers for the  $\alpha$  (1  $\rightarrow$  4) glycosidic link in D-maltose [7]. However, many bands in the region of aromatic skeleton vibrations as well as C=C stretches were also observed, even though the flavone moiety contained no additional chiral centers. In order to interpret the complex ROA patterns observed for **1** in water, and also to confirm the origin of the couplet centered at  $885\text{ cm}^{-1}$ , theoretical calculations were performed.

The first approach involved a regular conformational search at the MM level of theory using the Monte Carlo algorithm for both anomers of **1**. Then, the lowest-energy conformers identified were geometry optimized at the B3LYP/6-31G(d) level in water using PCM followed by Raman/ROA calculations ("one-step" [9] B3LYP/PCM(water)/TZVP). Fig. 2 shows the comparison between experimental Raman/ROA for **1** and the Boltzmann-averaged calculated spectra for the  $\alpha$  and  $\beta$  epimers. Interestingly, no calculation was capable of reproducing the large –,+ (low to high wavenumber) couplet-like signal centered at  $1600\text{ cm}^{-1}$  in the experimental spectrum (bands k and l). Additionally, only the  $\beta$ -anomer gave rise to a couplet at around  $885\text{ cm}^{-1}$  with the same sign pattern

observed experimentally (bands e and f). At this point, it was not clear whether **1** had a  $\beta$  link between the two sugar molecules or if the +,– couplet around  $885\text{ cm}^{-1}$  was a result of the presence of an L-configured sugar. Therefore, to answer these questions, a new computational approach based on hybrid QM/MM calculations was used. Recent literature reports show that the use of the two-layer ONIOM [10] method with electronic embedding scheme improves significantly the agreement between experimental and calculated ROA spectra for monosaccharides [10,11]. This is particularly useful when the conformational search is carried out using MD with explicit water molecules.

For the conformational search of **1**, MD simulations under periodic box conditions with explicit water molecules were carried out for both anomers. Snapshots were taken at every 10 ps, and the geometry optimization and harmonic frequency calculations for different conformers were performed using the two-layer ONIOM method. The high layer/QM region included compound **1** and was treated at the B3LYP/6-31G(d) level of theory. The low layer/MM region was comprised of all water molecules and was treated with AMBER force field. All optimized conformers for both glycosidic linkage configurations were then selected for "two-step" [9] Raman/ROA calculations at the B3LYP/rDPS level of theory also employing the ONIOM method. As a large variance in calculated energies is expected in systems containing multiple solvent molecules, which in turn results in inaccurate Boltzmann weightings, in this stage we decided to analyze the Raman and ROA features of each individual conformer in order to find the best agreement with experimental data.

For the  $\alpha$ -anomer, three conformers (a–c) were found to populate the sample in water solution (Fig. 3 and Table S1). Out of

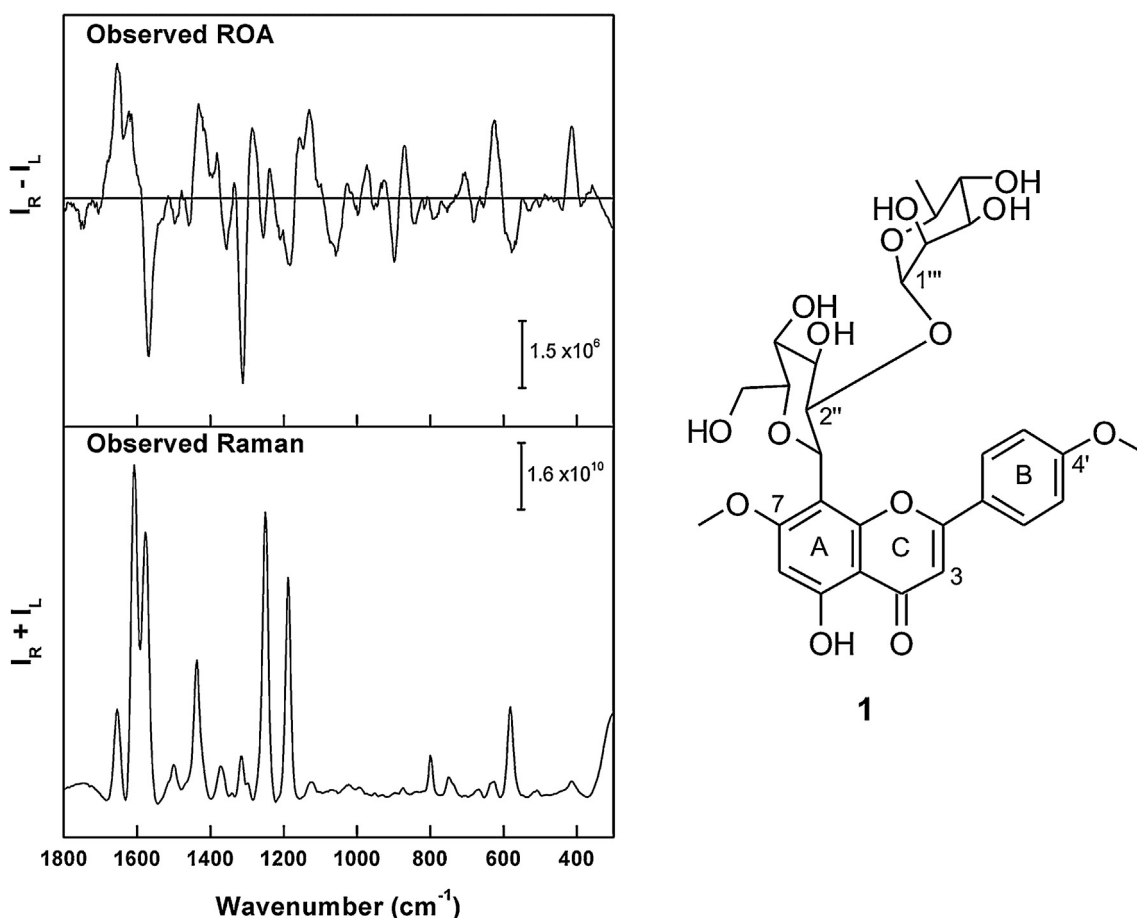
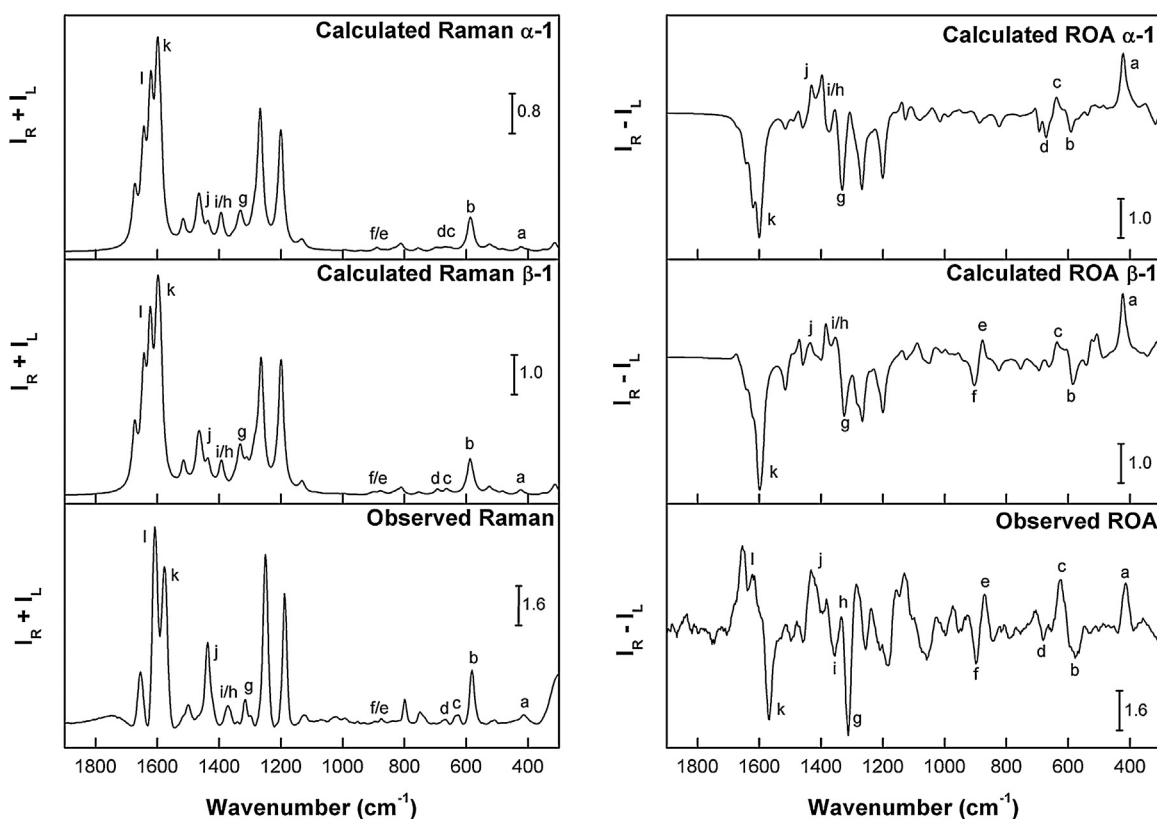
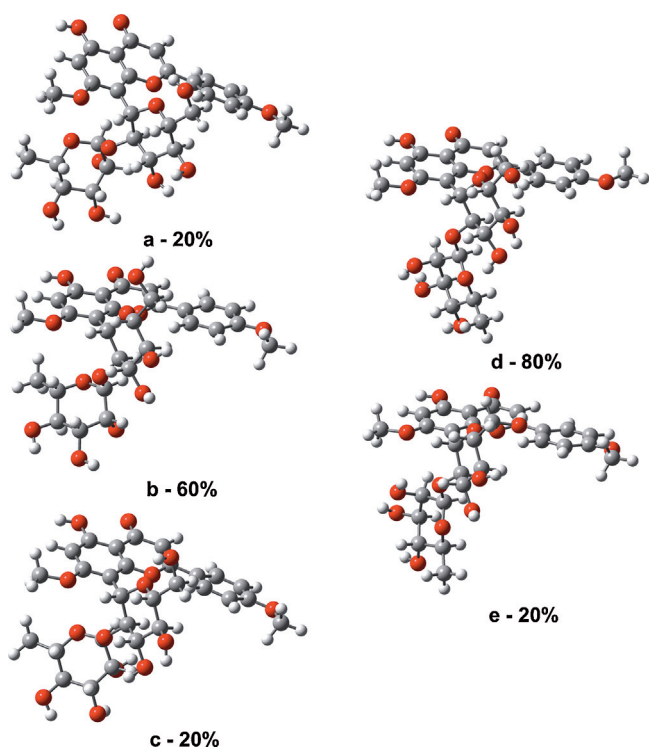


Fig. 1. (Left) Observed Raman and ROA spectra in water. (Right) Chemical structure of isoswertisin-4'-methyl-ether-2'' $\alpha$ -L-rhamnaside (**1**).



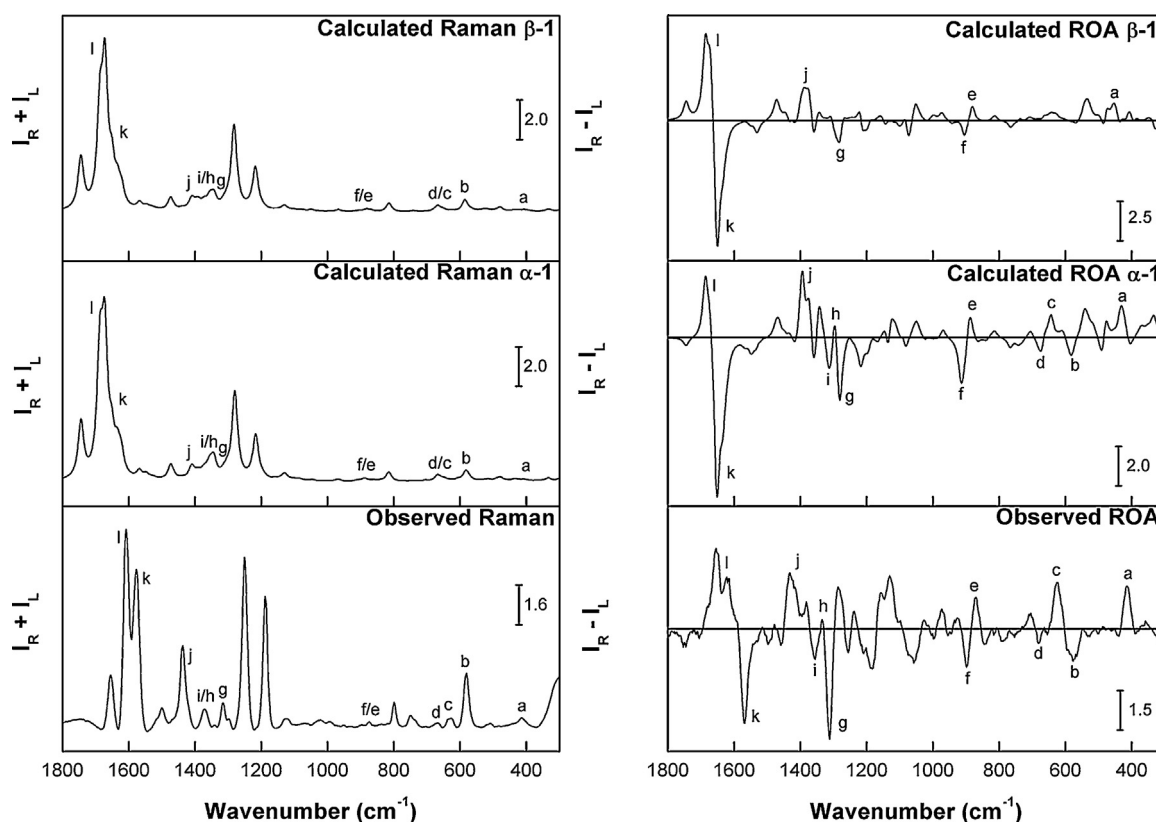
**Fig. 2.** Observed and calculated (B3LYP/TZVP//B3LYP/TZVP in water using PCM) Raman and ROA spectra for the Boltzmann average of  $\alpha$ - and  $\beta$ -configured **1**. Please refer to Supplementary material (Figs. S1 and S2) for lowest-energy conformers. Lowercase letters represent band assignments. Scale bars indicate relative intensities in arbitrary units.



**Fig. 3.** (Left) Optimized structures of conformers **a**, **b**, and **c** identified for  $\alpha$ -**1** using the two-layer ONIOM method at the B3LYP/6-31G(d) level. (Right) Optimized structures of conformers **d** and **e** identified for  $\beta$ -**1** using the two-layer ONIOM approach at the B3LYP/6-31G(d) level. Percentages represent contributions of individual conformers to the final spectra of  $\alpha$ - and  $\beta$ -**1** (Please see Fig. 4). Explicit water molecules were omitted for the sake of clarity. Please refer to Supplementary material (Figs. S3 and S4) for solvated molecules.

these, conformer **b** was predicted to make the largest contribution to the final spectrum. Conformers **a** and **c** differed mainly with respect to the O—H rotamers of the sugar molecules, while conformer **b** presented an additional important difference. This conformer was the only one to present a nearly planar arrangement of rings B and C of the flavone moiety ( $\omega = -3.64^\circ$  C3, C2, C1', C2'). Such an arrangement allowed for the coupling of the aromatic ring skeleton vibrations of rings A and B with the C=C stretching mode of C2 and C3, resulting in the couplet observed at around  $1600\text{ cm}^{-1}$  (bands k and l). Based on the analysis of all other conformers identified after MD simulations, departures of roughly  $\pm 25^\circ$  from the above-mentioned planar arrangement still lead to coupled vibrations, however, the ROA intensities do not split in a  $-$ ,  $+$  couplet. Instead, either a positive or negative band is observed at around  $1600\text{ cm}^{-1}$  with its sign directly related to the sign of the dihedral angle C3, C2, C1', C2'. A very similar trend was observed for the two conformers identified for the  $\beta$ -anomer (**d** and **e**) (Fig. 3 and Table S1). Interestingly, the nearly planar arrangement of rings B and C was found only after MD simulations followed by QM/MM geometry optimization using the two-layer ONIOM approach with explicit solvation. All other conformational searches and geometry optimization steps resulted in rotated B and C rings and, consequently, in poor reproduction of experimental data at higher wavenumbers. The same rotated geometry was also observed after the geometry optimization of conformer **b** containing the first solvation shell (see Supplementary material). This latter finding evidences the role of the whole water hydrogen bond network topology in stabilizing the planar conformation of the flavone moiety.

Once the importance of explicit solvation in the conformational preferences of **1** was demonstrated, the next step was to evaluate the possible influence of explicit water molecules in the ROA spectrum of **1**. To that end, conformers **b** ( $\alpha$ -anomer) and **d**



**Fig. 4.** Observed and calculated Raman and ROA spectra for  $\alpha$ - (conformers **a**, **b** and **c**) and  $\beta$ -configured **1** (conformers **d** and **e**) using the two-layer ONIOM method at the B3LYP/rDPS level. Lowercase letters represent band assignments. Scale bars indicate relative intensities in arbitrary units.

( $\beta$ -anomer) had their Raman and ROA properties calculated without the presence of water molecules at the B3LYP/rDPS level. The same calculations were also performed with implicit solvation by using PCM. The calculated Raman/ROA spectra for both anomers, with and without explicit water molecules, were significantly different (see Supplementary material). It indicates the importance of incorporating MM partial charges in the QM Hamiltonian. Therefore, for water-soluble natural product molecules, MD simulations and QM/MM are the methods of choice for chiroptical property calculations, as opposed to classic single molecule gas-phase or PCM approaches.

Following the identification of the most abundant conformers for  $\alpha$ - and  $\beta$ -**1**, thorough analyses of their experimental and calculated Raman and ROA spectra were performed in order to determine the stereochemistry of the glycosidic linkage in **1** (Fig. 4). The analysis was first focused on the region around  $885\text{ cm}^{-1}$  (bands **e** and **f**). As described above, a +,– (low to high wavenumber) couplet in this part of the spectrum is considered as a marker for the  $\alpha$  ( $1 \rightarrow 4$ ) glycosidic link in D-maltose [7]. The band around  $400\text{ cm}^{-1}$  (band **a**) was not further investigated since it contained large contributions from ring torsion modes of ring B. As can be seen in the calculated spectra for  $\alpha$ - and  $\beta$ -**1**, the +,– couplet at  $885\text{ cm}^{-1}$  is common to both anomers, rendering it not useful for this assignment. The couplet in question was shown to arise from breathing vibrations of the aromatic rings coupled with  $\text{CH}_2$  rocking motion and small deformations of the glucose ring. The same pattern was observed for both diastereoisomers since they share similar conformations for the flavone moiety and the same configuration for the C-linked D-glucose molecule. Curiously, however, this region contains small-amplitude ROA intensities with +,– signs (low to high wavenumber) for the  $\alpha$ -anomer and –, + for the  $\beta$ -configured isomer, which arise from C–C stretches and  $\text{CH}_3$  deformation modes of L-rhamnose. Therefore, this region of

the spectrum would possibly be useful for the stereochemical assignment of the glycosidic link between D-glucose and L-rhamnose in the absence of the flavone molecule. The –, +,– (low to high wavenumber) band pattern centered at around  $630\text{ cm}^{-1}$  (bands **b**, **c**, and **d**), on the other hand, contains large contributions from C–C–O and C–O–C bendings involving the glycosidic linkage. These bands are observed for the  $\alpha$ -configured isomer only.

Finally, the focus was directed towards the fingerprint region at around  $1330\text{ cm}^{-1}$ , where significant differences between the ROA spectra of  $\alpha$ - and  $\beta$ -**1** could be identified. The +,– (low to high wavenumber) couplet-like ROA band (bands **h** and **i**), which was also observed only for the  $\alpha$ -configured molecule, originated from C–H bendings of the glucose moiety as well as C–H bendings and C–C stretches of L-rhamnose. Even though, some of these vibrational modes involved C–H bendings from the flavone rings, the sugar bending and stretching modes described above included the anomeric carbon of L-rhamnose (C-1''') as well as C-2'' of D-glucose, which are the atoms that form the glycosidic bond in question. Therefore, the configuration for the L-Rha-( $1 \rightarrow 2$ )- $\beta$ -D-Glc linkage in **1** was securely confirmed to be  $\alpha$  (1'''-S), which is in accordance with the NMR analysis in  $\text{DMSO}-d_6$ .

### 3. Conclusions

The combination of Raman/ROA spectroscopy with state-of-the-art theoretical calculations for a water-soluble flavone C-diglycoside confirms the exquisite sensitivity of vibrational optical activity to molecular chirality. However, the enhanced sensitivity of ROA requires very accurate DFT calculations in order to interpret experimental data. These calculations are, in turn, highly dependent on the correct description of the conformational ensemble in solution. In this work, we demonstrated the advantage of using MD



simulations with explicit solvation followed by a hybrid QM/MM method in order to describe the correct conformations and the absolute configuration of the glycosidic link of isoswertisin-4'-methyl-ether-2'' $\alpha$ -l-rhamnoside in water solution. The presence of a nearly planar conformation for the flavone moiety, which accounts for important ROA observables, was detected only by applying the computational protocol presented above. Classic MM conformational searches and geometry optimizations without explicit water molecules failed to reproduce experimental findings. On top of that, the results presented herein evidence the risks of interpreting ROA spectra of (di)saccharides based solely on empirical analysis, especially for sugar molecules containing chiral or achiral large substituents.

## 4. Materials and methods

### 4.1. Experimental

Isoswertisin-4'-methyl-ether-2'' $\alpha$ -l-rhamnoside (**1**) was isolated and purified from the aerial parts of *Peperomia obtusifolia* (Piperaceae) following chromatographic procedures previously described [12]. Its structural identification was based on NMR and MS analyses, which were in accordance with literature values [8]. Fluorescence background was removed by leaving the sample in the laser beam for 2 h (photobleaching) under the same conditions used for Raman/ROA measurements. The Raman and ROA spectra of **1** were recorded with a BioTools ChiralRaman spectrometer operated via Critical Link LLC software. The instrument was set up with a scattered circular polarization strategy in backscattering geometry using a SpectraPhysics Millenium Pro Nd:VO<sub>4</sub> laser with an excitation wavelength of 532 nm, laser power of 600 mW at the sample, spectral resolution of 7 cm<sup>-1</sup>, sample concentration at 35 mg/mL in water, and spectral acquisition time of 24 h. Raw ROA data was used without any preprocessing; however, the Raman data was baseline-corrected by applying an asymmetric least squares algorithm using MatLab software v. R2012a.

### 4.2. Calculations

All DFT calculations were carried out at 298 K in gas phase or in water by using PCM in its integral equation formalism version (IEFPCM) incorporated in Gaussian 09 software [13]. In a second approach, DFT calculations were performed using the two-layer ONIOM method. The high layer/QM region was treated at the B3LYP/6-31G(d) level of theory for geometry optimizations and harmonic frequency calculations, and at the B3LYP/rDPS for the calculations of ROA tensors. The low layer/MM region was treated with AMBER force field with electronic embedding scheme. The first set of calculations were performed for the arbitrarily chosen (1'''S)- and (1'''R)-**1** ( $\alpha$ - and  $\beta$ -anomers, respectively). The conformational searches were carried out at the molecular mechanics level of theory with the Monte Carlo algorithm employing the MM+ force field incorporated in HyperChem 8.0.10 software package. Initially, 29 ( $\alpha$ -anomer) and 19 ( $\beta$ -anomer) conformers with relative energy (rel E.) within 6 kcal mol<sup>-1</sup> of the lowest-energy conformer were selected and further geometry optimized at the B3LYP/6-31G(d) level. The 3 ( $\alpha$ -anomer) and 4 ( $\beta$ -anomer) conformers with rel E. <1.6 kcal mol<sup>-1</sup>, which corresponded to more than 90% of the total Boltzmann distributions, were selected for Raman and ROA spectral calculations. Raman and ROA calculations were carried out using the so-called “one-step” (2n + 1) approach at the B3LYP/PCM/TZVP level in water. The incident light frequency was 532 nm. Each spectrum was plotted as a sum of Lorentzian bands with half-widths at half-maximum of 10 cm<sup>-1</sup>. The calculated wavenumbers

were multiplied with a scaling factor of 0.978 and the Boltzmann-averaged composite spectra were plotted using Origin software. Molecular dynamics simulations were performed for the  $\alpha$ - and  $\beta$ -anomers of **1** in periodic box conditions with 298 explicit water molecules for  $\alpha$ -**1** and 300 water molecules for  $\beta$ -**1** using the AMBER force field as implemented in HyperChem 8.0.10 software package. A single molecule of  $\alpha$ - or  $\beta$ -**1** was placed at the center of the box. The lowest-energy conformer of each anomer found during the first MM conformational search was selected as input for MD simulations using the following parameters: run time of 1 ns, time step of 0.001 ps, simulation temperature of 300 K. The different conformers obtained from individual snapshots taken at every 10 ps were geometry optimized and had harmonic frequencies and ROA tensors calculated using the two-layer ONIOM approach described above. Final Raman and ROA spectra were generated using the “two-step” (n + 1) procedure. The incident light frequency was 532 nm. Each spectrum was plotted as a sum of Lorentzian bands with half-widths at half-maximum of 10 cm<sup>-1</sup>. The calculated wavenumbers were multiplied with a scaling factor of 0.978 and the spectra were plotted using Origin software.

## Acknowledgments

The authors thank São Paulo Research Foundation (FAPESP, Grants# 2014/25222-9 and 2015/07089-2) for funding. This research was also supported by resources supplied by the Center for Scientific Computing (NCC/GridUNESP) of the São Paulo State University (UNESP). The authors are also grateful to Dr. Shaun T. Mutter and Dr. James R. Cheeseman for the helpful discussions about QM/MM and ROA calculations.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.vibspec.2016.08.002>.

## References

- [1] N.C. Veitch, R.J. Grayer, Nat. Prod. Rep. 28 (2011) 1626–1695.
- [2] R.F. Sprenger, S.S. Thomasi, A.G. Ferreira, Q.B. Cass, J.M. Batista Jr., Org. Biomol. Chem. 14 (2016) 3369–3375.
- [3] J.M. Batista Jr., E.W. Blanch, V.S. Bolzani, Nat. Prod. Rep. 32 (2015) 1280–1302.
- [4] X.-C. Li, D. Ferreira, Y. Ding, Curr. Org. Chem. 14 (2010) 1678–1697.
- [5] K. Monde, T. Taniguchi, N. Miura, S.-I. Nishimura, J. Am. Chem. Soc. 126 (2004) 9496–9497.
- [6] S.T. Mutter, F. Zielinski, C. Johannessen, P.L.A. Popelier, E.W. Blanch, J. Phys. Chem. A 120 (2016) 1908–1916.
- [7] F. Zhu, N.W. Isaacs, L. Hecht, G.E. Tranter, L.D. Barron, Chirality 18 (2006) 103–105.
- [8] J.S. Mota, A.C. Leite, M.J. Kato, M.C.M. Young, V.S. Bolzani, M. Furlan, Nat. Prod. Res. 25 (2011) 1–7.
- [9] J.R. Cheeseman, M.J. Frisch, J. Chem. Theory Comput. 7 (2011) 3323–3334.
- [10] S. Dapprich, I. Komáromi, K.S. Byun, K. Morokuma, M.J. Frisch, J. Mol. Struct. 462 (1999) 1–12.
- [11] S.T. Mutter, F. Zielinski, J.R. Cheeseman, C. Johannessen, P.L.A. Popelier, E.W. Blanch, Phys. Chem. Chem. Phys. 17 (2015) 6016–6027.
- [12] J.S. Mota, A.C. Leite, J.M. Batista Jr., S.N. López, D.L. Ambrósio, G.D. Passerini, M.J. Kato, M.C.M. Young, V.S. Bolzani, R.M.B. Cicarelli, M. Furlan, Planta Med. 75 (2009) 620–623.
- [13] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H.P. Hratchian, A.F. Izmaylov, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J.M. Millam, M. Klene, J.E. Knox, 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, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P. Salvador, J.J. Dannenberg, S. Dapprich, A.D. Daniels, O. Farkas, J.B. Foresman, J.V. Ortiz, J. Cioslowski, D.J. Fox, GAUSSIAN 09 (Revision A.02), Gaussian, Inc, Wallingford, CT, 2009.