

Conformational study of arbutin by quantum chemical calculations and multivariate analysis

Cuahtémoc Araujo-Andrade^a, Susy Lopes^b, Rui Fausto^b, Andrea Gómez-Zavaglia^{b,c,*}

^aUnidad Académica de Física de la Universidad Autónoma de Zacatecas, Zacatecas, Mexico

^bDepartment of Chemistry, University of Coimbra, 3004-535 Coimbra, Portugal

^cCentro de Investigación y Desarrollo en Criotecología de Alimentos (Conicet La Plata, UNLP), 1900 La Plata, Argentina

ARTICLE INFO

Article history:

Received 23 February 2010

Received in revised form 31 March 2010

Accepted 1 April 2010

Available online 7 April 2010

Keywords:

Arbutin

Conformational analysis

Multivariate analysis

Intramolecular interactions

ABSTRACT

A conformational study of the molecule of arbutin (4-hydroxyphenyl- β -D-glucopyranoside) has been undertaken. The molecule is composed by a glucopyranoside moiety bound to a phenol ring. It has eight conformationally relevant dihedral angles, five of them related with the orientation of the hydroxyl groups and the remaining three taking part in the skeleton of the molecule. A systematic search on the conformational space of arbutin was performed using molecular orbital methods, followed by the identification of structural similarities between the different conformers, using multivariate analyses. This approach allowed the grouping of conformers according to their structural affinity and the establishment of correlations between their structures and several properties. Intramolecular interactions involving OH groups were also investigated and correlations between spectroscopic, structural and thermodynamic properties established. The developed strategy might be useful to investigate the structure and structure/properties correlations in other conformationally flexible molecules.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Arbutin (IUPAC name: (2R,3S,4S,5R,6S)-2-hydroxymethyl-6-(4-hydroxyphenoxy)oxane-3,4,5-triol, also known as hydroquinone- β -D-glucopyranoside) is an abundant solute in the leaves of many freezing- or desiccation-tolerant plants. It has been used pharmaceutically in humans for centuries, either as plant extracts or, in more recent decades, in the purified form. Arbutin acts as an anti-septic or antibacterial agent on the urinary mucous membranes while converting into hydroquinone in the kidney [1]. It is also used as a depigmenting agent (skin whitening agent) as it inhibits melanin synthesis by inhibition of tyrosinase activity [2].

From a chemical point of view, the arbutin molecule has eight conformationally relevant dihedral angles, five of them related with the orientation of the hydroxyl groups and the remaining three taking part in the skeleton of the molecule. The number of hydroxyl groups and multitude of possible conformations and intermolecular interactions for which they are relevant are responsible for most of the arbutin relevant physico-chemical properties and biological activity [3–6], but are also a source of complexity for molecular structural studies.

In this study, a systematic search on the conformational space of arbutin was performed, conformers were grouped according to

structural similarities using multivariate analyses and correlations between their structures and several properties established. The strategy here used to analyze such a conformationally complex system as arbutin (with 35 atoms and eight conformationally relevant dihedral angles) might be useful to investigate the structure and structure/properties correlations in other conformationally flexible molecules, in particular carbohydrates and other biologically relevant substances.

Intramolecular interactions involving OH groups were also investigated and correlations between spectroscopic, structural and thermodynamic properties established.

2. Computational details

The semi-empirical PM3 method [7] was used to perform a systematic preliminary conformational search on the arbutin potential energy surface (PES). It provided a quick assessment of the main features of the conformational space of the molecule, which were later on taken into account in the subsequent, more reliable analysis performed at higher level of theory. This preliminary conformational search was carried out using the HyperChem Conformational Search module (CyberChem, Inc.© 2004) [8]. Taking into account the high flexibility of the arbutin molecule, a random search appeared as the most appropriate way to perform the conformational search [9–11]. In this approach, the generation of new starting conformations for the energy minimization uses a random

* Corresponding author.

E-mail address: angoza@qui.uc.pt (A. Gómez-Zavaglia).

variation of the dihedral angles of previously found conformers [9,10]. The method searches on until no new minima are generated.

The eight dihedral angles defining the conformational isomers of arbutin (Fig. 1) were considered in the random search: $C_2C_1O_{23}C_{24}$, $C_1O_{23}C_{24}C_{25}$, $O_6C_5C_7O_{11}$, $C_5C_7O_{11}H_{15}$, $C_3C_4O_{10}H_{14}$, $C_2C_3O_9H_{13}$, $C_1C_2O_8H_{12}$ and $C_{26}C_{27}O_{34}H_{35}$. Conformations with energies lower than 50 kJ mol^{-1} were stored while higher-energy conformations or duplicate structures were discarded. The structures obtained from this conformational search were used as start points for the construction of the input files later used in the higher level quantum chemical calculations. These latter were performed with Gaussian 03 [12] at the DFT level of theory, using the 6-311++G(d,p) basis set [13] and the three-parameter density hybrid

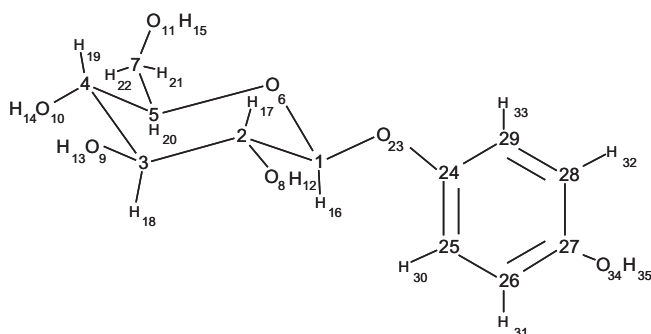


Fig. 1. Arbutin molecule, with atom numbering scheme.

functional abbreviated as B3LYP, which includes Becke's gradient exchange correction [14] and the Lee, Yang and Parr [15] and Vosko, Wilk and Nusair correlation functionals [16]. Conformations were optimized using the Geometry Direct Inversion of the Invariant Subspace (GDIIIS) method [17]. The optimized structures of all conformers were confirmed to correspond to true minimum energy conformations on the PES by inspection of the corresponding Hessian matrix. Vibrational frequencies were calculated at the same level of theory.

The multivariate studies were performed using The Unscrambler software (v9.8) [18].

3. Results and discussion

3.1. Geometries and energies

Arbutin has eight conformationally relevant rotational axes. After re-optimizing at the DFT(B3LYP)/6-311++G(d,p) level of theory the structures obtained from the preliminary semi-empirical random conformational search, 130 conformers were found, all of them belonging to the C_1 symmetry point group.

Table 1 presents the calculated relative energies and dipole moments of all the conformers calculated at the DFT(B3LYP)/6-311G++(d,p) level of theory. The values corresponding to the dihedral angles defining the conformers are provided in Table S1.

To group the conformers according to their main structural similarities, the following strategy was followed, which intends to be a fast and general procedure to perform such operation in medium size (30–60 atoms) conformationally flexible molecules:

Table 1
DFT(B3LYP)/6-311++G(d,p) relative energies, including zero point vibrational contributions (ΔE), and dipole moments (μ) for the 130 most stable conformers of arbutin.^a

Form ^b	ΔE	μ	Form ^b	ΔE	μ	Form ^b	ΔE	μ	Form ^b	ΔE	μ
29 (B)	0 ^c	1.90	110 (B)	14.88	4.00	221 (A)	20.64	4.03	113 (C)	27.40	3.14
60 (B)	0.54	4.29	115 (B)	14.98	3.70	1 (B)	21.09	2.48	133 (B)	28.34	2.69
57 (C)	0.93	2.80	141 (C)	15.30	3.43	4 (C)	21.20	3.97	107 (C)	28.52	2.95
88 (B)	1.18	4.12	26 (B)	15.58	2.71	204 (A)	21.40	3.06	123 (B)	28.57	3.10
53 (C)	1.82	2.77	36 (B)	15.62	2.22	16 (B)	22.17	3.52	104 (C)	28.98	2.94
47 (C)	2.18	3.91	76 (C)	15.63	3.39	100 (A)	22.17	4.18	253 (A)	29.15	2.67
52 (C)	2.68	2.96	211 (A)	16.15	3.38	130 (A)	22.27	1.65	237 (A)	29.17	2.64
56 (B)	2.88	2.04	143 (A)	16.67	5.00	153 (A)	22.32	3.51	157 (B)	29.27	3.72
200 (B)	3.29	4.48	22 (C)	16.69	1.71	39 (B)	22.42	2.01	99 (A)	29.32	4.99
112 (B)	3.88	2.36	51 (B)	16.87	2.36	44 (C)	22.51	2.78	222 (A)	29.75	2.56
274 (B)	4.43	4.15	83 (C)	16.89	2.95	168 (C)	22.71	3.37	137 (C)	29.82	4.21
134 (C)	8.97	1.70	280 (C)	16.90	4.23	164 (A)	22.73	3.96	289 (A)	29.83	3.80
33 (C)	9.64	0.93	24 (C)	16.95	1.31	262 (A)	22.94	3.43	284 (A)	30.00	2.32
136 (B)	9.75	2.88	59 (B)	17.01	3.84	161 (A)	23.03	1.91	145 (C)	30.01	4.73
75 (B)	10.04	3.46	40 (C)	17.37	1.47	119 (B)	23.78	2.62	261 (C)	30.11	5.34
120 (C)	10.73	1.14	303 (B)	17.50	4.32	175 (A)	23.92	1.65	270 (A)	30.27	3.57
185 (C)	10.84	2.48	9 (B)	17.53	2.35	139 (A)	24.02	4.28	69 (C)	30.43	3.24
173 (B)	11.10	3.07	202 (A)	17.56	2.08	250 (A)	24.66	1.75	298 (A)	30.60	3.69
103 (B)	11.28	2.32	7 (B)	17.61	3.72	263 (A)	24.68	2.75	246 (A)	31.08	3.42
178 (C)	11.32	2.92	58 (B)	17.62	2.60	180 (C)	24.96	3.61	306 (B)	31.16	4.58
150 (B)	11.36	3.66	45 (B)	17.94	3.91	95 (A)	24.97	3.00	114 (A)	31.42	2.61
41 (C)	11.53	2.72	209 (A)	17.96	3.30	73 (C)	24.97	0.68	232 (B)	32.71	5.65
288 (C)	12.40	3.05	10 (B)	18.07	1.36	142 (A)	24.97	3.00	215 (A)	32.72	1.30
109 (B)	12.59	2.07	98 (B)	18.12	4.12	195 (C)	25.13	2.32	307 (A)	34.23	3.47
154 (C)	13.25	4.68	6 (C)	18.21	3.00	151 (A)	25.20	2.32	282 (A)	35.46	4.52
63 (B)	13.36	3.15	304 (C)	18.81	5.25	79 (B)	25.30	2.78	297 (A)	35.75	4.05
74 (B)	13.77	4.23	158 (A)	18.98	2.06	177 (C)	25.72	5.28	301 (A)	36.94	3.67
85 (C)	13.82	2.21	3 (C)	19.22	2.52	186 (B)	25.86	3.48	207 (A)	37.31	4.34
111 (B)	13.98	3.47	11 (C)	19.34	2.91	43 (A)	26.03	4.82	293 (A)	39.42	3.63
62 (C)	13.99	2.30	166 (A)	19.37	2.97	160 (A)	26.19	3.93	264 (A)	41.46	3.01
70 (B)	14.07	3.42	311 (A)	19.77	5.33	198 (A)	26.27	1.55	313 (A)	45.68	6.40
121 (A)	14.16	3.45	225 (A)	20.25	1.93	105 (A)	26.37	2.98			
156 (A)	14.50	4.96	71 (C)	20.29	1.94	116 (B)	27.01	3.49			

^a Energies in kJ mol^{-1} , Dipole moments in Debyes ($1 \text{ D} = 3.33564 \times 10^{-30} \text{ C m}$).

^b The numbers designating the conformers correspond to the number of the structure submitted to optimization in the preliminary PM3 calculations; the group to whom conformers belong is given in parenthesis.

^c The calculated total energy (with zero point vibrational energy contribution) for the lowest energy conformer of arbutin (conformer 29) is $-2608343.813 \text{ kJ mol}^{-1}$.

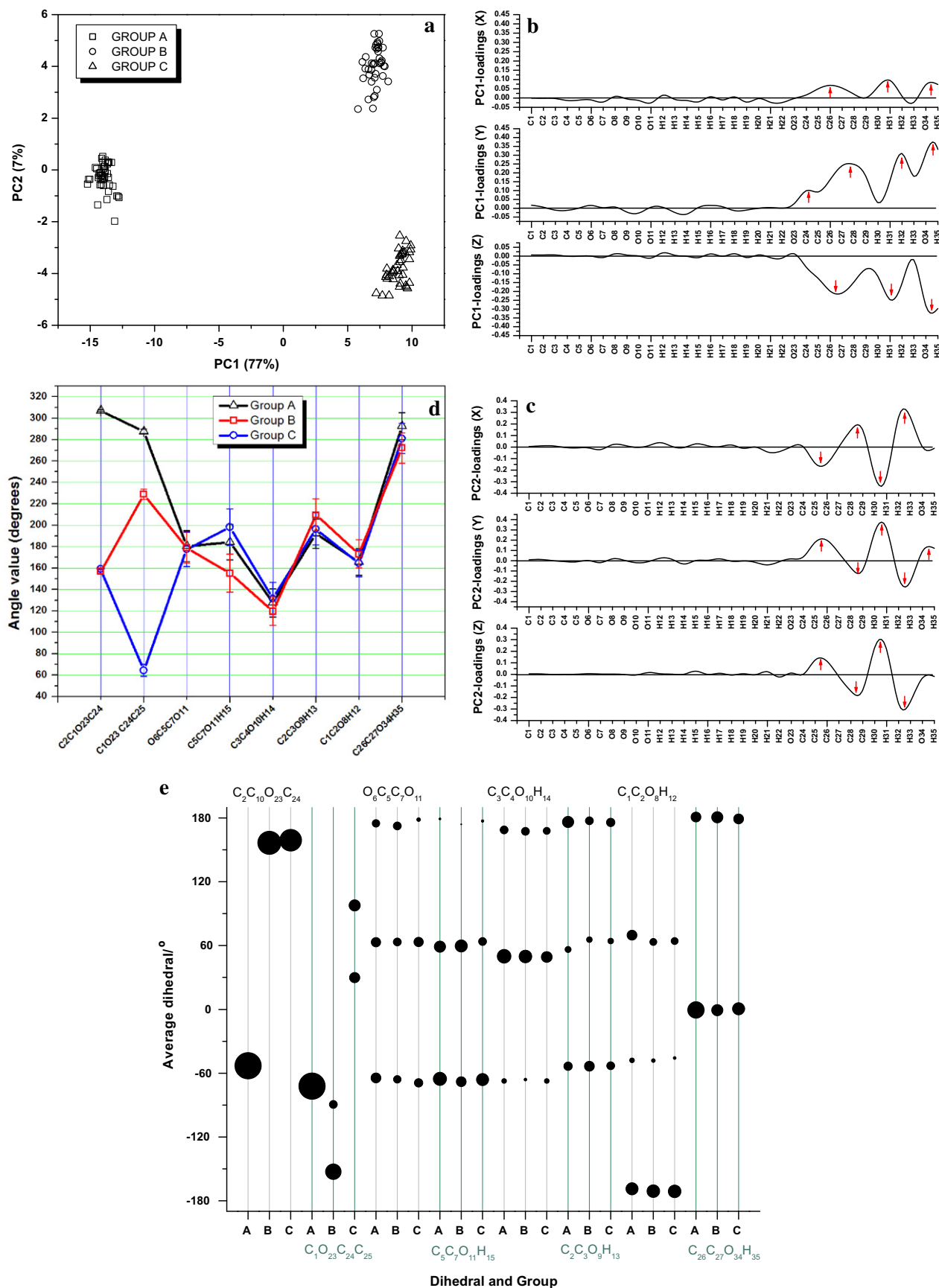


Fig. 2. PCA-scores (PC2 vs. PC1; **plot a**) and the corresponding loadings (**plots b, c**), grouping arbutin conformers in terms of structural similarity (see text for detailed description). **Plot d:** total average values and standard deviations of the eight conformationally relevant dihedral angles of arbutin in the three groups of conformers resulting from the PCA analysis. **Plot e** was built as follows: values of each dihedral angle were grouped according to their similarity and the average value for each set marked in the graph by means of a circle whose diameter is proportional to the cardinal of the set.

- (a) All conformers were oriented equally in the Cartesian referential, in such a way that the main structurally rigid fragment of the molecular skeleton (in this case the glucopyranoside ring) is placed as close as possible to the axes origin: atoms C₁, C₂ and C₃ were located in the origin, positive y axis and xy plane (first quadrant), respectively.
- (b) All Cartesian coordinates of the 130 low energy conformers of arbutin were then used to perform a principal component analysis (PCA). The table of data (data matrix) was built as follows: each line corresponds to a conformer and the columns to the Cartesian coordinates, the first 35, to the x coordinates of the 35 atoms of arbutin, the second 35 to the y coordinates and the last 35 to the z coordinates. In practice, this method is very fast to implement, once the desired orientation of the molecule has been used as input geometry in the Gaussian 03 DFT calculations; then, the optimized geom-

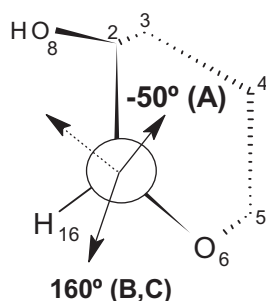


Fig. 3. Newman projection along the C₁–O₂₃ bond of arbutin. The arrows represent the orientation of the O₂₃–C₂₄ bond in conformers of Group A (ca. -50°) and B,C (ca. 160°). The orientation of the O₂₃–C₂₄ bond according to the remaining staggered position around the C₁–O₂₃ bond (indicated by the dashed arrow) was not found in any of the 130 lowest energy conformers of the molecule due to steric interactions between the phenol ring and the O₈–H₁₂ group.

etries resulting from these calculations are already correctly oriented in the Cartesian frame and in the proper format to be provided as input to the PCA calculations.

Fig. 2 depicts the PCA-scores (PC2 vs. PC1; frame a) obtained from this analysis and the corresponding loadings (frames b and c). A good separation of arbutin conformers in three groups can be observed. Group A appears well separated from Groups B and C along the first principal component (PC1), which explained 77% of total variance. Group B is separated from Group C along the PC2. The loading values of PC1 and PC2 evidentiate the atoms' coordinates contributing the most to structurally distinguish the conformers of arbutin in the chosen reference system. According to these values, the spatial orientation of the phenol ring (atoms 25–35) relatively to the reference glucopyranoside fragment is the main factor allowing for the differentiation among conformers. The relative spatial orientation of the phenol ring is determined by the dihedrals interconnecting the glucopyranoside and phenol rings, C₂C₁O₂₃C₂₄ and C₁O₂₃C₂₄C₂₅, which are then shown to be of first importance in structural terms. This can be also noticed very clearly in the additional graphs shown in Fig. 2, where the eight conformationally relevant dihedral angles in the three groups of conformers of arbutin are plotted in two ways:

In frame e, for each Group (A, B and C) the values of each dihedral angle were grouped according to their similarity and the average value for each subset marked in the graph by means of a circle whose diameter is proportional to the cardinal of the subset. It can be noticed that only the C₂C₁O₂₃C₂₄ and C₁O₂₃C₂₄C₂₅ dihedrals allow for a discrimination among groups: in Group A, C₂C₁O₂₃C₂₄ is ca. -50° and C₁O₂₃C₂₄C₂₅ ca. -70° , all conformers (#48) obeying this general pattern; on the other hand, in both Groups B and C the C₂C₁O₂₃C₂₄ dihedral is around 160° , but the C₁O₂₃C₂₄C₂₅ one may exhibit two different ranges of values in each group: ca. -150° (#28) and -90° (#14) in Group B and ca. 30° (#19) and 100° (#21) in Group C. All the other dihedrals show similar distributions of values in all groups.

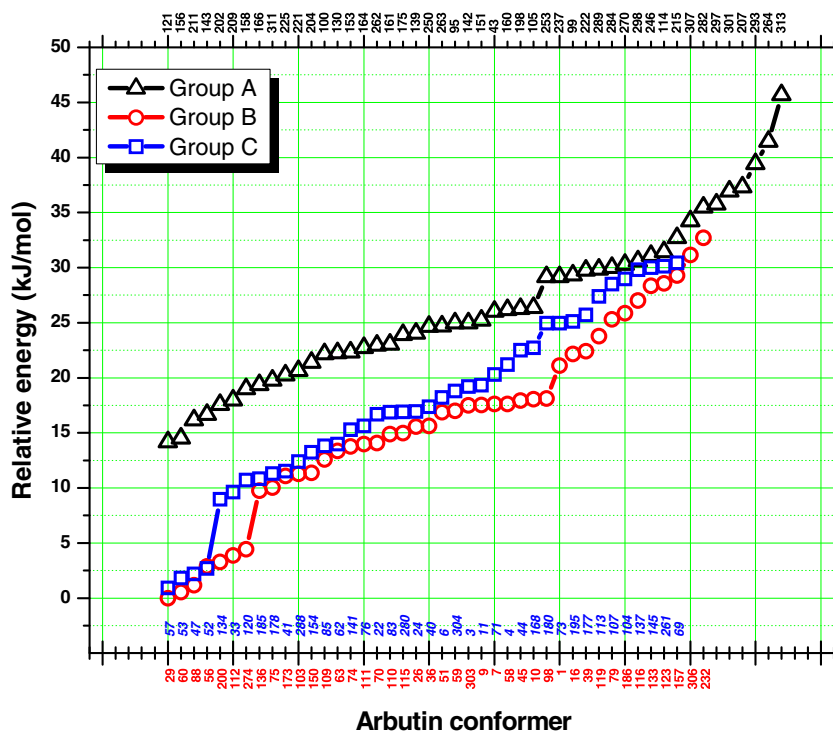


Fig. 4. Relative energies of the 130 lowest energy conformers of arbutin (the energy of the conformational ground state was taken as reference). The numbers designating the conformers correspond to the numbers of the structures submitted to optimization in the preliminary PM3 calculations.

In the alternative representation (frame *d*), the means and standard errors of the means (standard deviation of the sample divided by the square root of the sample size) of the eight conformationally relevant dihedral angles of arbutin in the three groups of conformers are plotted (the angles were first converted to the 0–360° range). In this representation, the individual ranges corresponding to the characteristic values for each dihedral in the different groups cannot in general be seen (this is possible only in cases where the values gravitate around a single value, as for the $C_2C_1O_{23}C_{24}$ dihedral), but the plotted global average values are also good descriptors of the distribution of values assumed for a given dihedral angle in the different groups. From this graph, it is then also easily noticed that, except for $C_2C_1O_{23}C_{24}$ and $C_1O_{23}C_{24}C_{25}$ dihedral angles, the plots exhibit similar profiles for the three groups.

According to the values assumed by the $C_2C_1O_{23}C_{24}$ and $C_1O_{23}C_{24}C_{25}$ dihedral angles, in Group A the phenol ring is placed above and nearly perpendicular to the glucopyranoside ring, whereas in both Groups B and C the phenol ring is pointing out of the glucopyranoside moiety, oriented to the side of the oxygen atom from the glucopyranoside ring (Fig. 3); in each group, two different orientations of the phenol ring may exist (see Fig. 2e). Noteworthy, there are no low energy conformers exhibiting the third

possible staggered orientation of the $C_2C_1O_{23}C_{24}$ dihedral (see Fig. 3), which is a result of the destabilizing interactions between the phenol ring and the O_8H_{12} group (see Fig. 1).

The relative energies of the conformers in the three groups (taken as reference the energy of the conformational ground state) are plotted in Fig. 4. From this point of view, Groups B and C are equivalent. However, Group A is clearly different from the other two groups, exhibiting an absence of conformers with relative energies below 15 kJ mol^{-1} and integrating the highest energy forms among the conformers investigated. This trend of the conformers belonging to Group A to have higher energies than those belonging to Groups B and C can be correlated with the different orientations of the phenol ring relatively to the glucopyranoside ring in Group A and both Groups B and C (see Figs. 2 and 3): as already mentioned, in Group A the phenol ring is placed above the glucopyranoside ring, thus in a considerably more crowded position compared with that it occupies in the conformers belonging to Groups B and C.

Once the influence of the relative position of the two rings in arbutin on the relative energy of the conformers has been evaluated, the preferred conformations assumed by the substituents of the glucopyranoside ring and their influence on energies were

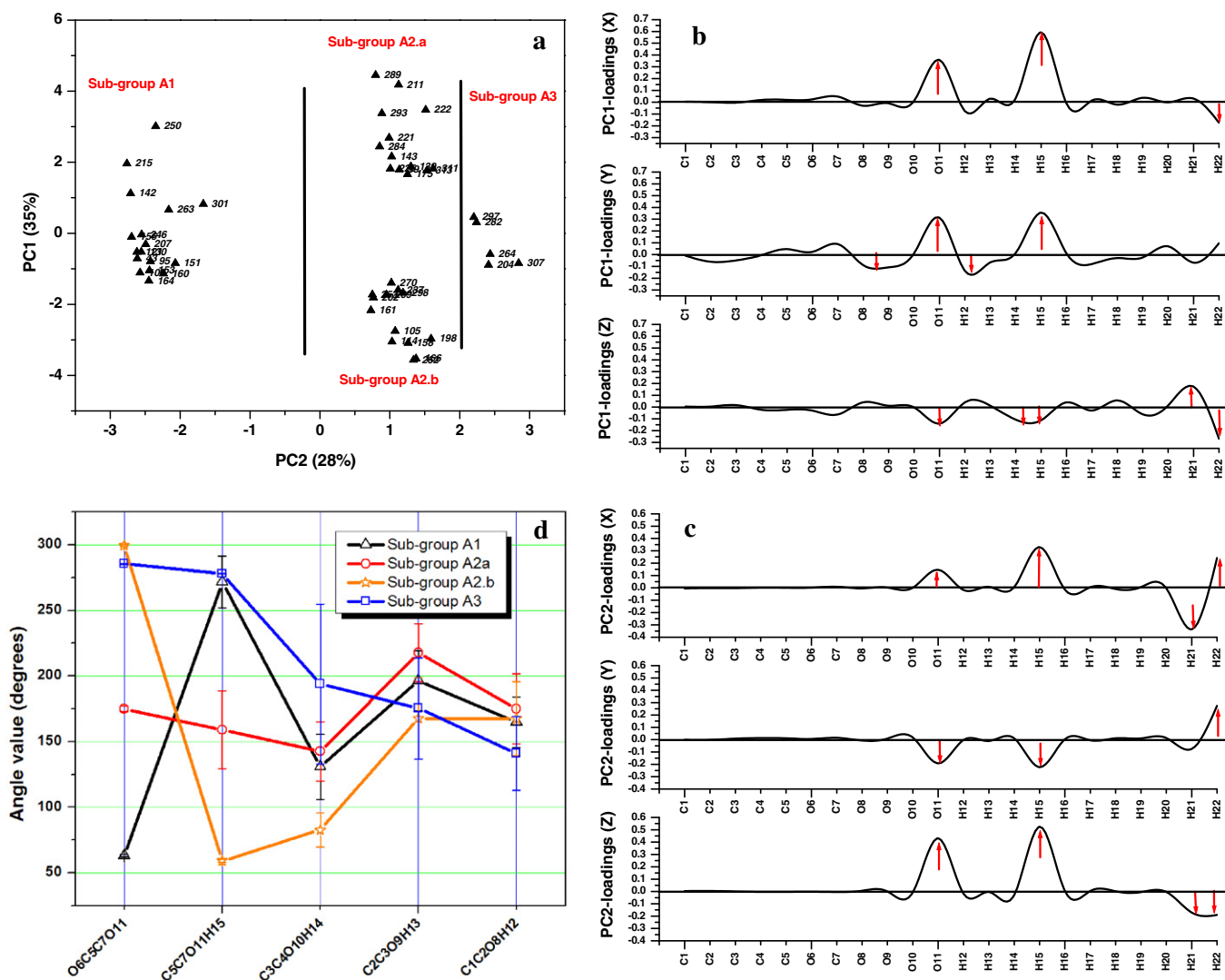


Fig. 5. PCA-scores (PC2 vs. PC1; plot *a*) and the corresponding loadings (plots *b*, *c*), grouping arbutin conformers belonging to Group A in terms of structural similarity in the conformations of the substituents of the glucopyranoside ring (see text for detailed description). Plot *d*: total average values and standard errors of the means of the five conformationally relevant dihedral angles of the glucopyranoside ring arbutin in the four subgroups of conformers resulting from the PCA analysis.

investigated in deeper detail. To this aim, we performed an additional PCA analysis on each of the previously determined groups of conformers (A, B, and C), using only the coordinates of the first 22 atoms of the molecule (*i.e.*, excluding the phenol ring; see Fig. 1). This strategy allowed for the elimination of information that is not relevant for a conformational analysis within the glucopyranoside ring. The results of this analysis are shown in Figs. 5–8.

When PC1 and PC2 are plotted together, the PCA-scores plot for Group A (Fig. 5; frame a) shows four clear groupings, denominated here as subgroups A1, A2.a, A2.b and A3. If all elements of these four subgroups are projected over the PC1 axis, three groups can be distinguished, with one of them constituted by subgroups A1 and A3, other formed by subgroup A2.a and the third one by subgroup A2.b. On the other hand, projecting the elements over the PC2 axis allows also to distinguish three groups, but this time corresponding to A1, (A2.a, A2.b) and A3. The observation of the loadings plots for PC1 and PC2 (Fig. 5; frames b and c) allows to conclude that the positions of atoms C₇, H₂₁, H₂₂, O₁₁ and H₁₅ are highly related with the formation of the four subgroups observed in the PCA-scores plot, *i.e.*, the conformation exhibited by the CH₂OH substituent at C₅ is the main discriminating factor among subgroups. In consonance with this observation, when the mean

values of the five conformationally relevant dihedral angles associated with the substituted glucopyranoside ring in each subgroup are plotted (Fig. 5; frame d), it is possible to observe that the dihedral angles associated with the CH₂OH substituent (O₆–C₅–C₇–O₁₁ and C₅–C₇–O₁₁–H₁₅) are those allowing the discrimination among the four subgroups. The remaining three dihedral angles show similar patterns of variation regarding means and standard errors of the means. In the case of the C₅–C₇–O₁₁–H₁₅ dihedral, one can promptly correlate the three groups corresponding to the projection of the PCA subgroups over the PC1 axis with the three dihedral mean values shown in the plot: *ca.* 60°, 160° and 275° (–85°), respectively for A2.b, A2.a and (A1, A3). These dihedral angle values correspond roughly to the three expected stable staggered positions of the O₁₁H₁₅ group (*gauche+*, *anti* and *gauche-*).

The PCA-scores plot for the conformers belonging to the Group B (Fig. 6; frame a) shows only two clear groupings of conformers, where PC2 is the component separating these two groups the best. Both PC1 (not plotted) and PC3 have little influence in the clear definition of the two clusters. The loadings plot of PC2 (Fig. 6; frame b) shows that the clusters are, like in the case of Group A, also determined by the positions of atoms C₇, H₂₁, H₂₂, O₁₁ and

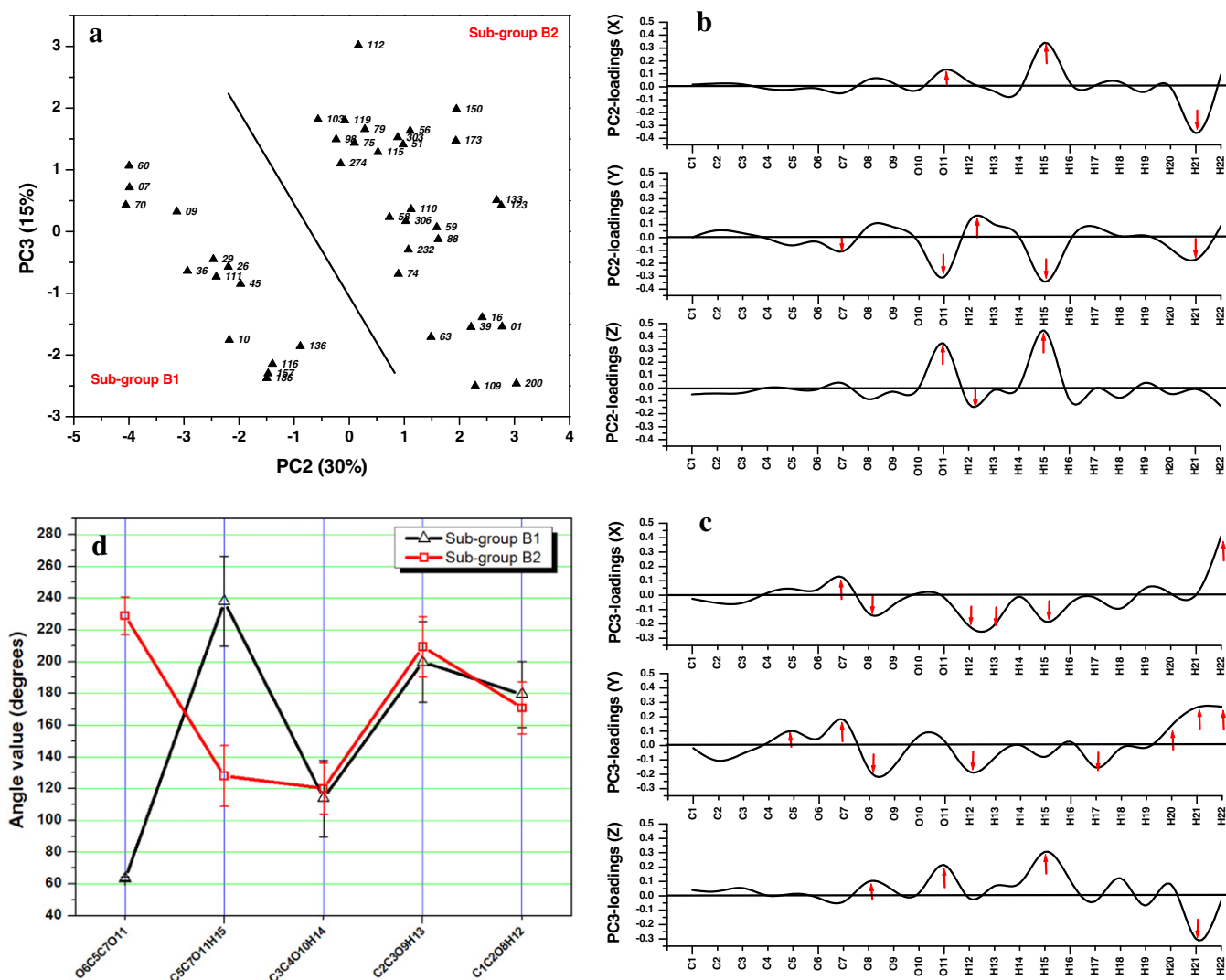


Fig. 6. PCA-scores (PC3 vs. PC2; plot a) and the corresponding loadings (plots b, c), grouping arbutin conformers belonging to Group B in terms of structural similarity in the conformations of the substituents of the glucopyranoside ring (see text for detailed description). **Plot d:** total average values and standard errors of the means of the five conformationally relevant dihedral angles of the glucopyranoside ring arbutin in the two subgroups of conformers resulting from the PCA analysis.

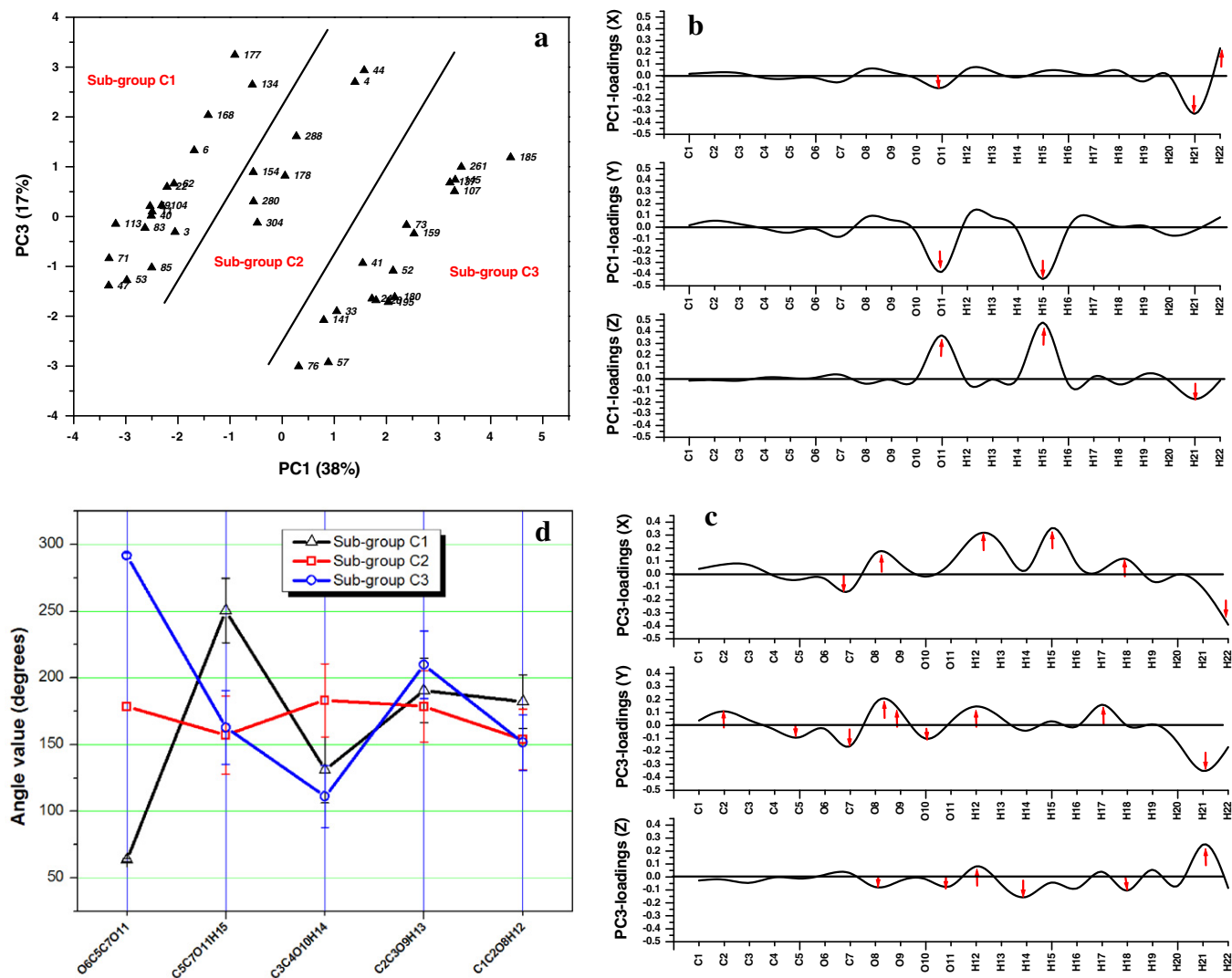


Fig. 7. PCA-scores (PC3 vs. PC1; **plot a**) and the corresponding loadings (**plots b, c**), grouping arbutin conformers belonging to Group A in terms of structural similarity in the conformations of the substituents of the glucopyranoside ring (see text for detailed description). **Plot d:** total average values and standard errors of the means of the five conformationally relevant dihedral angles of the glucopyranoside ring arbutin in the three subgroups of conformers resulting from the PCA analysis.

H₁₅, i.e., by the conformation of the CH₂OH fragment. As expected, these observations are in agreement with the dihedral angles' mean values plot (Fig. 6; frame d), which clearly reveals that, the O₆–C₅–C₇–O₁₁ and C₅–C₇–O₁₁–H₁₅ dihedral angles are the internal coordinates that mainly discriminate the conformers belonging to subgroups B1 or B2.

A similar analysis made for conformers belonging to Group C allows concluding that three subgroups (C1, C2 and C3) can be defined (Fig. 7; frame a), once again resulting mainly from different conformations assumed by the CH₂OH substituent (Fig. 7; frames b–d).

Regarding the energies of the conformers, subgroups are not strongly discriminative. However, subgroups A3, C2 and, in less extent B1, include conformers that show a trend to be less stable than the remaining subgroups of each main group (Fig. 8). In typical members of these subgroups, the specific destabilizing interactions are respectively H₁₅···H₁₉ (A3), H₂₁···H₁₉ (C2) and H₁₅···H₂₀ (B1).

In summary, PCA analyses based on atomic Cartesian coordinates of the properly oriented in the Cartesian system conformers of arbutin allowed the grouping of these conformers by structural analogies, which could be related with the conformationally relevant dihedral angles. Among them, the dihedrals interconnecting

the glucopyranoside and phenolic rings and those associated with the CH₂OH fragment were found to be the most relevant ones. In spite of the fact that there is not a strong general correlation between the different groups of conformers and the relative conformational energies, trends were found in energies that could be correlated with structural details of the conformers characteristic of a given group (or subgroup).

3.2. Intramolecular interactions involving the OH groups

The hydroxyl groups in arbutin can participate in different types of intramolecular interactions either with other OH groups, the glucopyranoside ring oxygen atom or the ring-bridging O₂₃ atom. The exception is the phenolic O₃₄H₃₅ group, which for the isolated molecule does not interact with any other group in the molecule. Table 2 shows the different intramolecular interactions in which each OH group can participate.

Most of the interactions involving the OH groups in arbutin are related with vicinal OH groups. The nature of the intramolecular interactions involving vicinal hydroxyl groups has been the matter of several discussions. In fact, referring them as hydrogen bonds is controversial [19–22]. Indeed, though some studies described vic-

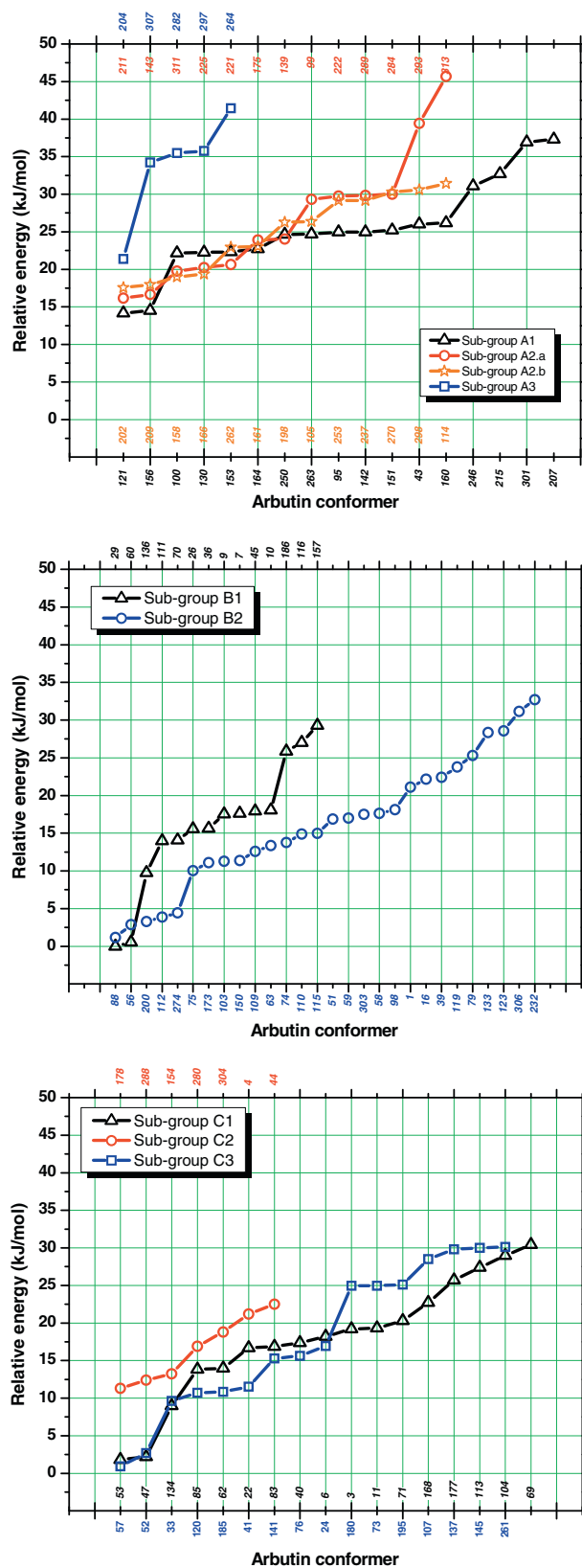


Fig. 8. Relative energies of the 130 lowest energy conformers of arbutin (the energy of the conformational ground state was taken as reference) according to its distribution by the different subgroups resulting from the PCA analyses (see text for details). The numbers designating the conformers correspond to the numbers of the structures submitted to optimization in the preliminary PM3 calculations.

Table 2

Role of the hydroxyl groups in intramolecular interactions in arbutin.^a

Conformer	O ₁₁ H ₁₅	O ₁₀ H ₁₄	O ₉ H ₁₃	O ₈ H ₁₂	O ₃₄ H ₃₅	O ₆	O ₂₃
O ₁₁ H ₁₅		A, D	-	-	-	A	-
O ₁₀ H ₁₄	A, D		A, D	-	-	-	-
O ₉ H ₁₃	-	A, D		A, D	-	-	-
O ₈ H ₁₂	-	-	A, D		-	-	A
O ₃₄ H ₃₅	-	-	-	-		-	-

^a D, electron donor; A, electron acceptor; -, no interaction.

inal intramolecular O–H···OH attractive interactions as corresponding to weak hydrogen bonds on the basis of geometrical parameters and spectroscopy data [20–22], others [19] consider that the topological analysis of the electron density shows no evidence for the existence of a hydrogen bond in such cases. Taking into account these considerations, in this work we will refer to the vicinal O–H···OH attractive interactions just as ‘O–H···OH intramolecular interactions’, avoiding their classification as hydrogen bonds. Stabilizing interactions between non-vicinal OH groups are no controversial regarding its nature and will be here treated as representing H-bonds. As it will be shown below, in all cases, however, the existing attractive interactions between OH groups correlate with longer OH distances and lower OH stretching frequencies of the electron acceptor group.

Table 3 shows the average OH bond lengths and OH stretching frequencies associated to each OH group in arbutin. Data for the groups and subgroups resulting from the PCA structural analysis were also given in this table.

As it could be expected, the O₃₄H₃₅ bond length is the shortest OH bond length in the molecule (96.26 pm), since this group does not interact with any other group in the molecule. Accordingly, the O₃₄H₃₅ stretching frequency is predicted to occur at the highest value among all OH stretching modes (3838 cm⁻¹). The similar chemical environment in all conformers is expressed in the fact that the calculated frequency for $\nu_{O_34H_{35}}$ is practically constant in all forms.

In the opposite extreme is the O₁₁H₁₅ group. The values corresponding to the O₁₁H₁₅ bond lengths and $\nu_{O_{11}H_{15}}$ vibrational mode are within the 96.14–96.68 pm and 3765–3838 cm⁻¹ ranges, respectively. These larger ranges of possible values demonstrate that the O₁₁H₁₅ group can establish either relatively strong hydrogen bonds with O₁₀H₁₄, interact weakly with O₆ or be free of any kind of interaction with other groups. The longest bond lengths and lowest frequency values occur in conformers where the O₁₁H₁₅ group is H-bonded to O₁₀H₁₄, whereas the shortest bond lengths and highest frequency values are characteristic of those conformers where the O₁₁H₁₅ group is free of any H-bond like interaction. Note that, compared to the remaining OH groups in the molecule, the O₁₁H₁₅ group is particularly flexible conformationally, since rotation around two bonds (C₅C₇ and C₇O₁₁), instead of around a sole bond like in all the other groups is allowed. This fact lets the O₁₁H₁₅ group easily adopt more adequate orientations to minimize the energy of the molecule, in particular by establishing more favorable H-bonds with O₁₀H₁₄.

On the other hand, even though the O₁₁H₁₅···O₁₀H₁₄ H-bond existing in several conformers is certainly among the strongest intramolecular interactions involving OH groups in arbutin, when all conformers are taken into account the calculated average value for $\nu_{O_{11}H_{15}}$ (3817 cm⁻¹) is the second highest (just preceded by $\nu_{O_{34}H_{35}}$; see Table 3), exactly because in several conformers the group O₁₁H₁₅ is not interacting. Very interestingly, when the PCA structural groups were considered, the highest average values for

$\nu_{O_{11}H_{15}}$ occur for subgroups A3, B1 and C2 (see Table 3), i.e., those subgroups that were found to be formed by conformers gradually less stable than the remaining subgroups of each main group (see Fig. 8). On the other hand, the lowest $\nu_{O_{11}H_{15}}$ average values correspond to subgroups A2.a, B2 and C1, which integrate most of the conformers where the $O_{11}H_{15} \cdots O_{10}H_{14}$ hydrogen bond is present.

Among all the global average OH bond length and ν_{OH} stretching frequency values (Table 3), those associated with the $O_{10}H_{14}$ group are maximum (96.47 pm) and minimum (3804 cm^{-1}), respectively, i.e., they indicate that this group is in average taking part in stronger interactions. In fact, the large majority of the conformers possess this group oriented to the $O_{11}H_{15}$ group, i.e., involved in a relatively strong H-bond forming a six-membered ring, whereas most of the remaining forms have this group pointing to the vicinal O_9H_{13} group. In turn, the O_9H_{13} group follows $O_{11}H_{15}$ in the series, because in all the studied arbutin conformers this group is interacting either with $O_{10}H_{14}$ or O_8H_{12} . Both these interactions are however, of the vicinal type and thus, not as strong as the $O_{11}H_{15} \cdots O_{10}H_{14}$ hydrogen bond above discussed. Finally, O_8H_{12} was found to be the hydroxyl group of arbutin exhibiting OH bond lengths and stretching frequencies compatible with an involvement in weaker interactions. Indeed, among the possible interactions of this group, the vicinal $O_8H_{12} \cdots O_9H_{13}$ can be expected to be the strongest, whereas the alternative one, $O_8H_{12} \cdots O_{23}$, is certainly a considerably weak interaction.

Hydrogen bond or hydrogen bond like interaction strengths can be estimated from OH bond lengths or OH stretching frequencies. Ilogensen, Rozenberg and coworkers have developed a quite general Eq. (1) where the redshift observed in the ν_{OH} vibrational mode in the presence of the H-bond type interaction can be used to estimate the interaction enthalpy [23–28]:

$$\Delta H^2 = 1.92 * [\Delta(\nu_{OH}) - 40] \quad (1)$$

ΔH is the interaction enthalpy (in kJ mol^{-1}) and $\Delta(\nu_{OH}) = \nu_o - \nu_{H}$ is given in cm^{-1} ; ν_o and ν_{H} correspond to values of ν_{OH} (in cm^{-1}) for free and interacting OH groups, respectively.

For the estimation of ΔH , in the present study a value of 3855 cm^{-1} (the highest calculated OH stretching frequency value in the studied arbutin conformers) was considered for the frequency of a free OH group (ν_o). In addition, since the DFT calculated frequencies are known to be overestimated by ca. 4% relatively to the experimental values [29,30], the $\nu_o - \nu_{H}$ values directly obtained from the calculations were scaled down by 0.96, in order to obtain redshift values correlatable to experiment. The ΔH average values for the different types of OH \cdots O interactions resulting from application of Eq. (1) are presented in Table 4.

The results shown in this Table fully support the conclusions made above on the basis of the structural (OH bond lengths) and spectroscopic (OH stretching frequencies) data. In particular: (i) the $O_{11}H_{15} \cdots O_{10}H_{14}$ and $O_{10}H_{14} \cdots O_{11}H_{15}$ H-bond type interactions

Table 3
Average OH distances (pm) and OH stretching frequencies (cm^{-1}) in arbutin conformers.^a

	All forms	A	B	C	A1	A2.a	A2.b	A3	B1	B2	C1	C2	C3
<i>Bond length</i>													
$O_{11}H_{15}$	96.39	96.39	96.42	96.36	96.41	96.42	96.41	96.26	96.38	96.44	96.39	96.28	96.36
$O_{10}H_{14}$	96.47	96.48	96.47	96.46	96.42	96.65	96.38	96.52	96.39	96.51	96.41	96.70	96.42
O_9H_{13}	96.47	96.47	96.47	96.47	96.48	96.47	96.45	96.49	96.48	96.47	96.48	96.47	96.46
O_8H_{12}	96.43	96.42	96.44	96.44	96.43	96.41	96.44	96.40	96.44	96.44	96.45	96.43	96.43
$O_{34}H_{35}$	96.26	96.27	96.26	96.26	96.27	96.27	96.27	96.26	96.26	96.26	96.26	96.26	96.26
<i>Frequency</i>													
$\nu_{O_{11}H_{15}}$	3817	3815	3815	3822	3813	3810	3815	3830	3818	3807	3816	3835	3821
$\nu_{O_{10}H_{14}}$	3804	3803	3805	3804	3814	3772	3821	3797	3819	3798	3815	3759	3815
$\nu_{O_9H_{13}}$	3809	3809	3808	3809	3807	3809	3811	3805	3808	3808	3806	3810	3810
$\nu_{O_8H_{12}}$	3813	3814	3813	3813	3814	3816	3810	3820	3814	3813	3811	3815	3814
$\nu_{O_{34}H_{35}}$	3838	3838	3838	3838	3838	3838	3838	3838	3838	3838	3838	3838	3838

^a See Fig. 1 for atom numbering.

Table 4
Average energies ($-\Delta H/\text{kJ mol}^{-1}$) for intramolecular interactions involving OH groups in arbutin as electron acceptor group, calculated from the $\Delta(\nu_{OH})$ frequencies according to Eq. (1) (see text).^a

Electron acceptor group	Electron donor group	All forms	A1	A2.a	A2.b	A3	B1	B2	C1	C2	C3
$O_{11}H_{15}$	(global)	4.13									
	O_6	2.56	3.18	^b	1.84	^b	3.26	2.11	^b	2.06	2.71
	$O_{10}H_{14}$	8.58	^b	8.65	^b	^b	^b	8.80	7.38	^b	^b
$O_{10}H_{14}$	(global)	5.53									
	$O_{11}H_{15}$	11.00	^b	11.08	^b	^b	^b	11.02	10.89	^b	^b
	O_9H_{13}	3.61	3.43	3.79	3.43	4.82	3.11	3.98	4.80	4.04	2.98
O_9H_{13}	(global)	4.39									
	$O_{10}H_{14}$	2.88	3.60	1.87	^b	3.15	2.88	2.96	2.07	2.29	3.53
	O_8H_{12}	4.28	3.53	4.29	3.81	^b	4.41	4.62	^b	4.06	4.56
O_8H_{12}	(global)	3.70									
	O_9H_{13}	3.60	4.06	3.87	5.15	1.89	3.16	2.98	2.67	3.89	4.02
	O_{23}	2.95	^b	^b	3.12	^b	^b	^b	^b	2.62	^b
$O_{34}H_{35}$		^b	^b	^b	^b	^b	^b	^b	^b	^b	^b

^a See Fig. 1 for atom numbering.

^b No interaction where this OH substituent acts as electron acceptor has been observed among the conformers of this group.

(in particular the latter) are the strongest stabilizing interactions in the molecule, with enthalpies amounting to *ca.* -8.5 and -11.0 kJ mol^{-1} , respectively; (ii) these interactions occur for subgroups A2.a, B2 and C1, which are composed by conformers bearing these two types of interactions; (iii) the vicinal $\text{OH}\cdots\text{OH}$ interactions have associated enthalpies of -3.6 ± 1.9 kJ mol^{-1} , whereas those involving O_6 and O_{23} as electron donors have the lowest enthalpies (average: -2.8 ± 0.3 kJ mol^{-1}); (iv) globally, the interactions involving the O_8H_{12} group are the weakest among those involving OH groups as electron acceptor groups.

The enthalpy values obtained for the strongest OH intramolecular interactions in arbutin (H-bond interactions) are of comparable strength of those found for other compounds bearing $\text{OH}\cdots\text{O}$ intramolecular H-bonds, such as, for example, butanediol isomers ($-\Delta H$: 1,3-butanediol, $12\text{--}14$ kJ mol^{-1} [31]; 1,4-butanediol, $19\text{--}13$ kJ mol^{-1} [32]), glycolic (CH_2OHCOOH), oxalic ($\text{HOOC}=\text{COOH}$) and pyruvic ($\text{CH}_3\text{COCOCH}_3$) acids ($-\Delta H$: 8.8, 7.6 and 12 kJ mol^{-1} [33–35], respectively).

4. Conclusion

Quantum chemical calculations in large flexible molecules still represent a challenge. The high number of combinations of conformationally relevant dihedral angles makes it an extremely complicated task, unless a systematic approach is carried out. In this work, a systematic conformational procedure has been developed to be applied to the arbutin molecule. Such procedure is simple, requires relatively modest computational facilities and may be used also to investigate the structure and establish structure/properties correlations in other conformationally flexible molecules.

According to calculations, 130 minima were found on the potential energy surface of arbutin. These conformers were grouped according to their structural similarities and correlations between their structures and several properties, including relative energies and spectroscopic properties, established. The grouping of the conformers was made using PCA analyses based on atomic Cartesian coordinates of the properly oriented in the Cartesian system conformers of arbutin. The characteristic structural features of each group of conformers could later on be related with the conformationally relevant dihedral angles of the molecule. Among them, the dihedrals interconnecting the glucopyranoside and phenol rings and those associated with the CH_2OH fragment were found to be the most relevant ones.

Intramolecular interactions involving OH groups were also investigated and correlations between spectroscopic, structural and thermodynamic properties established, in particular between νOH vibrational frequencies and OH bond lengths and strengths of the $\text{OH}\cdots\text{OH}$ (or $\text{OH}\cdots\text{O}$) interactions.

Finally, this work represents a new simple/fast approach for the structural analysis of complex molecules and its aim was also to show the potential applications of the developed methodology.

Acknowledgements

This work was supported by ANPCyT (Project PICT(2006)/00068), CYTED (Network 108RT0362) and CONACYT-CONICET (Mexico, Argentina) (bilateral Project res. No. 962/07–05–2009). AGZ is member of the Research Career Conicet, Argentina. SL acknowledges FCT for the Grant SFRH/BD/29698/2006.

Appendix A. Supplementary data

Table S1, providing the calculated values of the conformationally relevant dihedral angles of the 130 conformers of arbutin optimized at the B3LYP/6–311++G(d,p) level of theory. Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.molstruc.2010.04.002.

References

- [1] J. Wittig, S. Wittemer, M. Veit, *J. Chromatogr. B* 761 (2001) 125.
- [2] J.L. O'Donoghue, *J. Cosmet. Dermatol.* 5 (2006) 196.
- [3] D. Hinch, A.E. Oliver, J.H. Crowe, *Biophys. J.* 77 (1999) 2024.
- [4] A.E. Oliver, L.M. Crowe, P.S. de Araujo, E. Fisk, J.H. Crowe, *Biochim. Biophys. Acta* 1302 (1996) 69.
- [5] A.E. Oliver, L.M. Crowe, *J.H. Crowe, Seed Sci. Res.* 8 (1998) 211.
- [6] S. Turner, T. Senaratna, D. Touchell, E. Bunn, K. Dixon, B. Tan, *Plant Sci.* 160 (2001) 489.
- [7] J.P. Stewart, *J. Comput. Chem.* 10 (1989) 209.
- [8] HyperChem Conformational Search module, Tools for Molecular Modeling. Hypercube, Inc., 1115 NW 4th St., Gainesville, FL 32608 USA, 2002.
- [9] M. Saunders, *J. Am. Chem. Soc.* 109 (1987) 3150.
- [10] M. Saunders, K.N. Houk, Y.-D. Wu, W.C. Still, J.M. Lipton, G. Chang, W.C. Guida, *J. Am. Chem. Soc.* 112 (1990) 1419.
- [11] A.E. Howard, P.A. Kollman, *J. Med. Chem.* 31 (1988) 1669.
- [12] Gaussian 03, Revision C.02, 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, J.A. Pople, Gaussian, Inc., Wallingford CT, 2004. Copyright © 1994–2003, Gaussian, Inc.
- [13] M. Frisch, M. Head-Gordon, J. Pople, *Chem. Phys. Lett.* 166 (1990) 281.
- [14] A. Becke, *Phys. Rev. A* 38 (1988) 3098.
- [15] C. Lee, W. Yang, R. Parr, *Phys. Rev. B* 37 (1988) 785.
- [16] S. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* 58 (1980) 1200.
- [17] P. Csaszar, P. Pulay, *J. Mol. Struct. (Theochem.)* 114 (1984) 31.
- [18] The Unscrambler software (v9.8) CAMO Software AS, Nedre Vollgate 8, 0158 Oslo (Norway), 2009.
- [19] R.A. Klein, *J. Am. Chem. Soc.* 124 (2002) 13931.
- [20] L.P. Kuhn, *J. Am. Chem. Soc.* 74 (1952) 2492.
- [21] A.J. Lopes Jesus, M.T.S. Rosado, I.D. Reva, R. Fausto, M.E. Eusébio, J.S. Redinha, *J. Phys. Chem. A* 110 (2006) 4169.
- [22] E. Fishman, T.L. Chen, *Spectrochim. Acta A* 25 (1969) 1231.
- [23] M. Rozenberg, A. Lowenschuss, Y. Marcus, *Carbohydr. Res.* 328 (2000) 307.
- [24] M. Rozenberg, A. Lowenschuss, Y. Marcus, *Carbohydr. Res.* 394 (1997) 183.
- [25] M. Rozenberg, G. Shoham, I. Reva, R. Fausto, *Phys. Chem. Chem. Phys.* 7 (2005) 2376.
- [26] M. Rozenberg, *Spectrochim. Acta A* 52 (1996) 1559.
- [27] M. Rozenberg, G. Shoham, I.D. Reva, R. Fausto, *Spectrochim. Acta A* 60 (2004) 463.
- [28] A.V. Iogansen, *Spectrochim. Acta A* 55 (1999) 1585.
- [29] L.A. Curtiss, K. Raghavachari, P.C. Redfern, J.A. Pople, *Chem. Phys. Lett.* 270 (1997) 419.
- [30] J.A. Pople, H.B. Schlegel, R. Krishnan, D.J. Defrees, J.S. Binkley, M.J. Frisch, R.A. Whiteside, R.F. Hout, W.J. Hehre, *Int. J. Quant. Chem. Quant. Chem. Symp.* 15 (1981) 269.
- [31] M.T.S. Rosado, A.J. Lopes Jesus, I.D. Reva, R. Fausto, J.S. Redinha, *J. Phys. Chem. A* 113 (2009) 7499.
- [32] A.J. Lopes Jesus, M.T.S. Rosado, I.D. Reva, R. Fausto, M.E.S. Eusébio, J.S. Redinha, *J. Phys. Chem. A* 112 (2008) 4669.
- [33] S. Jarmelo, T.M.R. Maria, M.L.P. Leitão, R. Fausto, *Phys. Chem. Chem. Phys.* 3 (2001) 387.
- [34] R. Fausto, E.M.S. Maçôas, *J. Mol. Struct.* 563/564 (2001) 29.
- [35] I.D. Reva, S. Stepanian, L. Adamowicz, R. Fausto, *J. Phys. Chem.* 105 (2001) 4773.