Theor Chem Acc (2002) 107:229–240
DOI 10.1007/s00214-002-0324-3

Theoretical
Chemistry Accounts

Regular article

A theoretical investigation of the mechanism for the reaction between a quebrachitol derivative and N_3^-

Hélio F. Dos Santos[1,2], **Mauro V. De Almeida**[2], **Wagner B. De Almeida**[1,3]

[1]Núcleo de Estudos em Química Computacional
[2]Departamento de Química, ICE, Universidade Federal de Juiz de Fora, Campus Universitário, Martelos, 36036-330 Juiz de Fora, MG, Brazil,
[3]Laboratório de Química Computacional e Modelagem Molecular, Departamento de Química, ICEx, Universidade Federal de Minas Gerais, Campus Universitário, Pampulha, 31270-901 Belo Horizonte, MG, Brazil
e-mail: wagner@netuno.qui.ufmg.br; Fax: +55-31-34995700

Received: 10 July 2001 / Accepted: 20 December 2001 / Published online: 8 April 2002

Abstract. The reaction between a mesylated compound and sodium azide was previously studied experimentally at a temperature of 140 °C using dimethylformamide as a solvent. The product was assigned on the basis of the analysis of the NMR spectra. In this work semiempirical (AM1 and PM3), ab initio (Hartree–Fock and MP2) and density functional theory (BLYP functional) quantum mechanical calculations, using continuum models for describing the solvent effect, were carried out for this process to better understand the reaction mechanism. Three distinct mechanisms involving a carbocation and epoxide intermediates, and a transition-state structure for direct attack of the N_3^- species to the reactant were investigated. The theoretically calculated preferred reaction pathway passing through an epoxide intermediate agrees nicely with the experimental proposal, providing a good example of where theoretical calculations can be of great help to definitively elucidate the reaction mechanism.

Key words: Quebrachitol – Cyclitol – Solvent effect – Theoretical calculation

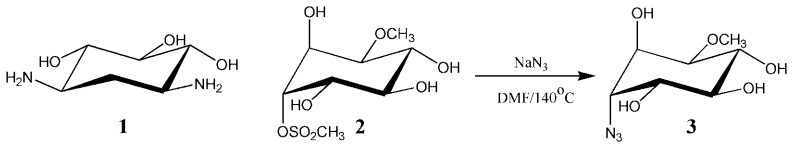
1 Introduction

The aminocyclitol moiety (deoxystreptamine, **1**) is present in the structure of distinct chemically useful aminoglycoside antibiotics, such as neomycin, kanamycin and gentamycin [1, 2]. These compounds are also components of several substances of biological interest, such as pancrastitatin [3]. Furthermore, Kozikowski and coworkers [4, 5, 6] have found that some azido inositols are selective inhibitors of the growth of v-sis-transformed NIH 3T3 cells.

In connection with our interest in cyclitol synthesis, we have developed an approach to the synthesis of the mesylated compound **2** (Scheme 1) [7]. Surprisingly, the reaction of **2** with sodium azide in dimethylformamide (DMF) at 140 °C for 16 h gave exclusively the azido compound **3** in 70% yield, with total retention of the configuration. Proof of the structures of compounds **2** and **3** was provided by NMR analysis [7].

Seeking a better understanding of the reaction mechanism, in the present work quantum mechanical calculations were carried out for this reaction. Semiempirical approaches were used in a first stage and in a second step ab initio (Hartree–Fock, HF, and Møller–Plesset second-order perturbation theory, MP2) and density functional theory (DFT) (using the BLYP functional) calculations were performed aiming to assess the influence of basis set and electronic correlation effects on the reaction process. The solvent effects were included using continuum solvation models [8]. It would be ideal to solve the theoretical problem of the chemical reaction mechanism in solution using a uniform approach based on first principles, employing, for example, the state-of-the-art correlated ab initio CCSD(T) approach with close-to-completeness basis sets, for structure and energy calculations. However, such a dreamed of theoretical procedure is beyond our present computational capabilities for molecular systems involving more than a dozen of atoms. Therefore, combined approximated treatments are needed. The results reported in this article provide strong support for the use of a viable computational procedure for studying reaction mechanisms taking into account solvent effects, which can lead to very satisfactory predictions. Such a procedure can be applied in the study of other chemical reactions in solution.

Correspondence to: W.B. De Almeida



Scheme 1.

2 Calculation

The geometry of the stationary points located on the potential-energy surface (PES) for the reaction of compound **2** with sodium azide producing compound **3** (Scheme 1) were initially fully optimized without any symmetry constraint using the quantum mechanical semiempirical molecular orbital methods AM1 [9] and PM3 [10] as implemented in the MOPAC 93 package [11]. The PRECISE and GNORM = 0.1 options were used in all the calculations for tightening the convergence criteria for all optimizations. In a second stage, full geometry optimizations were performed at the ab initio HF level of theory employing the 3-21+G* and 6-31++G** basis sets [12]. In addition, single-point energy calculations were done using MP2 [13] and DFT [14], with the generalized-gradient BLYP functional [15] approaches implemented in the Gaussian 98 package [16]. Three distinct mechanisms (Scheme 2) involving a carbocation and epoxide intermediates and a transition-state structure (TS-1) for direct attack of the N_3^- species on the reactant were investigated (see also Scheme 3). The epoxide formation was analyzed considering two routes. First, the internal displacement of the mesylate by the OH group which makes a charged intermediate (interm-1) and second the loss of hydroxyl proton yielding the alkoxide (interm-2) (see Scheme 3). The structures of the reactants and the products obtained for the three processes were characterized as true minimum-energy or TS structures on the PES through harmonic frequency calculations (all real frequencies correspond to a minimum-energy structure).

The solvent (DMF, $\varepsilon = 36.7$) effect was assessed using continuum models: conductor-like screening model (COSMO) [17] as implemented in the semiempirical MOPAC 93 package, and self-consistent reaction field (SCRF) [18], isodensity polarized continuum model (IPCM) [19] and Polarizable Continuum Model (PCM) [20] as implemented in the ab initio Gaussian 98 package. The solvation Gibbs free energies were evaluated using the well-known thermodynamic cycle [21], where the reaction Gibbs free energy in solution, ΔG^{sol}, is calculated for each process as the sum of two contributions: a gas-phase reaction free energy, ΔG^{gas}, and a solvation reaction free energy term calculated with the continuum approach, ΔG^{solv}.

$$\Delta G^{sol} = \Delta G^{gas} + \Delta G^{solv} \quad (1)$$

The gas-phase reaction free energy is the sum of two parts: electronic plus nuclear repulsion energy (ΔE_{ele}) and thermal contribution including zero-point energy ($\Delta G_T^{gas} = \Delta H - T\Delta S$), as given in Eq. (2). The last term in Eq. (2), named the thermal correction, is evaluated using the calculated quantum mechanical vibrational frequencies.

Process I:

$$\text{compound 2} \rightarrow \text{carbocation} + CH_3SO_3^- \quad \text{(Ia)}$$

$$\text{carbocation} + N_3^{(-)} + CH_3SO_3^- \rightarrow \text{P1-ax} + \text{P1-eq} + CH_3SO_3^- \quad \text{(Ib)}$$

Process II:

$$\text{compound 2} + N_3^{(-)} \rightarrow \text{TS-1} \quad \text{(IIa)}$$

$$\text{TS-1} \rightarrow \text{P1-eq} + CH_3SO_3^- \quad \text{(IIb)}$$

Process III:

$$\text{compound 2} \longrightarrow \text{interm-1} + CH_3SO_3^- \quad \text{(IIIa)}$$

$$\text{compound 2} + N_3^{(-)} \longrightarrow \text{interm-2} + HN_3 \quad \text{(IIIb)}$$

$$\text{compound 2} \longrightarrow \text{epoxide} + CH_3SO_3H \quad \text{(IIIc)}$$

$$\text{epoxide} + N_3^{(-)} + CH_3SO_3H \longrightarrow \text{TS-2} \longrightarrow \text{P1-ax} + CH_3SO_3^- \quad \text{(IIId)}$$

$$\text{epoxide} + N_3^{(-)} + CH_3SO_3H \longrightarrow \text{TS-3} \longrightarrow \text{P2-eq} + CH_3SO_3^- \quad \text{(IIIe)}$$

Scheme 2.

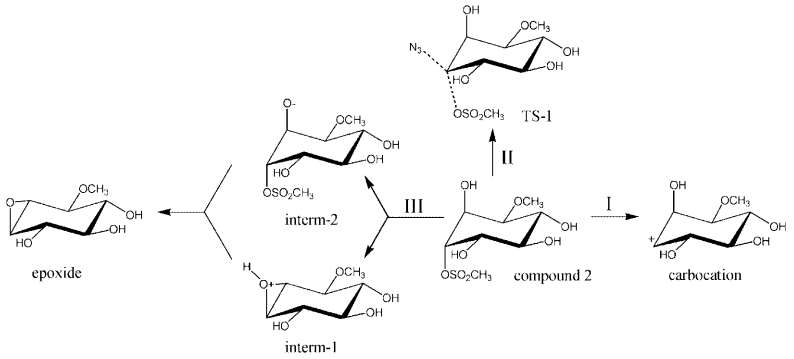
$$\Delta G^{gas} = \Delta E_{ele} + \Delta G_T^{gas} \quad (2)$$

The fully optimized gas-phase geometries were used for the solvation energy calculation. It was assumed that the geometry relaxation effects, within the continuum model, are not significant even for charged species [22]. Only the electrostatic contribution was included in the solvation free energy [20].

It is important to mention that for the evaluation of ΔG_T^{gas} harmonic frequency calculations have to be performed and a proper analysis of the low-frequency vibrational modes is required. Some of these modes are not true vibrations and must be treated separately as discussed in detail in Refs. [23, 24]. We have shown recently [25] that ΔG_T^{gas} can be evaluated satisfactorily using harmonic frequencies calculated at the HF level employing a moderate basis set. Then, a higher level of calculation can be used to generate the ΔE_{ele} term and also an adequate model for describing solvent effects can be utilized for the evaluation of the second term in Eq. (1). Therefore, a very satisfactory and computationally viable procedure for studying chemical reactions in solution can be proposed.

3 Results and discussion

The analysis of the main dihedral angles relevant for the specification of the conformation of compound **2**, carbocation, interm-1, interm-2, epoxide and TS-1 structure, evaluated using distinct theoretical methods, showed that the HF/3-21+G* level of calculation is sufficient for predicting the geometry of the stationary points present on the PES for the reacting system. The set of optimized structural parameters are available on request to the authors. The single-point correlated DFT



Scheme 3.

and MP2 energy calculations, using the fully optimized HF/3-21+G* equilibrium geometries, can be performed for an improved description of the energetics of the chemical reaction. In Tables 1, 2, 3 and 4 we show that this procedure works fairly well and then proceed in a computationally feasible way to study mechanisms of chemical reactions involving larger molecular species.

Tables 1 and 2 report reaction energies, ΔE_{ele}, Gibbs free energies, ΔG, and Gibbs free energy barriers, $\Delta G^{\#}$, for processes Ia, IIa and IIIa–IIIc corresponding to the first stage of the reaction mechanism. The solvent effect (DMF) was evaluated using distinct continuous models and thermal corrections (ΔG_T^{gas}) were calculated for 140 °C through HF harmonic frequencies, as described in Sect. 2. The values not corrected for internal rotational low-frequency modes are given in parentheses (HF/3-21+G*). It can be seen that this correction amounts to about 7 kcal mol^{-1} (TS-1 structure), 6 kcal mol^{-1} (carbocation), 4 kcal mol^{-1} (interm-1), 2 kcal mol^{-1} (interm-2) and 2 kcal mol^{-1} (epoxide), so

Table 1. Reaction energy, ΔE, Gibbs free energy, ΔG, and Gibbs free-energy barrier, $\Delta G^{\#}$, for processes Ia and IIa (kcal mol^{-1} at 140 °C) calculated in the gas phase (ΔE_{ele} and ΔG^{gas}) and solution (ΔG^{sol} in dimethylformamide, *DMF*, $\varepsilon = 36.7$). ΔG_T^{gas} is the thermal correction including zero-point energy, ΔE^0. The solvation energy contribution, ΔG^{solv}, is given in *parentheses*

Process Ia: carbocation	ΔE_{ele}	ΔG_T^{gas}	ΔG^{gas}	ΔG^{sol}			
				COSMO	IPCM[a]	PCM[b]	SCRF[c]
AM1	171	−28.0	143	18.3 (−125)	–	–	–
PM3	171	−27.0	144	20.5 (−124)	–	–	–
HF/3-21+G*	170	−29.8 (−24)[e]	140	–	26.2 (−114)	36.2 (−104)	136 (−4.0)
HF/6-31++G**	157	−29.6[f]	127	–	–	28.0 (−99)	–
HF/6-31+G*//3-21+G*	160	−29.8[d]	130	–	17.2 (−113)	27.1 (−103)	126 (−4.0)
BLYP/6-31+G*// HF/3-21+G*	145	−29.8[d]	115	–	12.6 (−103)	16.1 (−99.1)	112 (−3.2)
BLYP/6-31++G**// HF/6-31++G**	142	−29.8[d]	112	–	11.3 (−101)	15.5 (−96.7)	–
MP2/6-31+G*// HF/3-21+G*	174	−29.8[d]	144	–	–	43.2 (−101)	–
MP2/6-31++G**// HF/6-31++G**		−29.6[f]		–	–	45.1	–
Process-IIa: TS-1[g]	$\Delta E_{ele}^{\#}$	$\Delta G_T^{\#gas}$	$\Delta G^{\#gas}$	$\Delta G^{\#sol}$			
				COSMO	IPCM[a]	PCM[b]	SCRF[c]
AM1	29.3	14.8	44.1	74.4 (30.3)	–	–	–
PM3	36.5	15.1	51.6	89.2 (37.6)	–	–	–
HF/3-21+G*	11.3	18.1 (11)[e]	29.4	–	36.5 (7.1)	43.4 (14.0)	26.9 (−2.5)
HF/6-31+G*//3-21+G*	20.6	18.1[d]	38.7	–	45.1 (6.4)	52.9 (14.2)	35.7 (−3.0)
BLYP/6-31+G*// HF/3-21+G*	8.9	18.1[d]	27.0	–	39.4 (12.4)	44.3 (17.3)	26.0 (−1.0)
MP2/6-31+G*// HF/3-21+G*	11.8	18.1[d]	29.9	–	–	44.7 (14.8)	–

[a] Isodensity set to 0.0004
[b] PCM (UAHF). Initial number of tesserae = 100. Only the electrostatic contribution to the solvation energy was included
[c] Cavity radius (Å): 5.06 (**2**), 4.40 (carbocation), 3.55 ($CH_3SO_3^-$), 3.08 (N_3^-), 5.48 (TS-1$^-$)
[d] Thermal correction (ΔG_T^{gas}) from HF/3-21+G* level
[e] The HF/3-21+G* thermal free-energy corrections not corrected for internal rotational low-frequency modes
[f] Thermal free-energy correction calculated using the HF/6-31++G** fully optimized geometry for harmonic frequency calculation
[g] The expected convergence criterion for geometry optimization could not be attained with the 6-31++G** basis set, therefore the respective harmonic frequency calculations were not performed

Table 2. Reaction energy, ΔE, Gibbs free energy, ΔG, and Gibbs free-energy barrier, $\Delta G^{\#}$, for processes IIIa, IIIb and IIIc leading to the epoxide formation (kcal mol^{-1} at 140 °C) calculated in the gas phase (ΔE_{ele} and ΔG^{gas}) and solution (ΔG^{sol} in DMF, ε = 36.7). ΔG_T^{gas} is the thermal correction including zero-point energy, ΔE^0. The solvation energy contribution, ΔG^{solv}, is given in *parentheses*

	ΔE_{ele}	ΔG_T^{gas}	ΔG^{gas}	ΔG^{sol}			
				COSMO	IPCM[a]	PCM[b]	SCRF[c]
Process IIIa: interm-1							
AM1	185	−25.2	160	29.1 (−131)	–	–	–
PM3	183	−25.2	158	27.3 (−130)	–	–	–
HF/3-21+G*	165	−25.9 (−22)[e]	139	–	28.0 (−111)	0.0 (−139)	135 (−4.2)
HF/6-31+G*//3-21+G*	156	−25.9[d]	130	–	22.0 (−108)	−6.0 (−136)	126 (−3.9)
BLYP/6-31+G*// HF/3-21+G*	149	−25.9[d]	123	–	20.0 (−103)	−10.0 (−133)	120 (−3.4)
MP2/6-31+G*// HF/3-21+G*	165	−25.9[d]	139	–	–	4.0 (−135)	–
Process IIIb: interm-2							
AM1	15.2	−3.13	12.1	25.5 (13.5)	–	–	–
PM3	29.5	−3.24	26.3	36.4 (10.2)	–	–	–
HF/3-21+G*	19.3	−5.68 (−4.4)[e]	13.6	–	23.4 (9.8)	1.4 (−12.2)	5.8 (−7.8)
HF/6-31+G*//3-21+G*	21.0	−5.68[d]	15.3	–	26.0 (10.7)	8.9 (−6.4)	8.3 (−7.0)
BLYP/6-31+G*// HF/3-21+G*	16.1	−5.68[d]	10.4	–	23.0 (12.6)	11.0 (0.59)	4.7 (−5.7)
MP2/6-31+G*// HF/3-21+G*	28.8	−5.68[d]	23.1	–	–	20.6 (−2.5)	–
Process IIIc: epoxide							
AM1	37.1	−25.9	11.2	−2.3 (−13.5)	–	–	–
PM3	23.8	−26.3	−2.5	−14.4 (−12)	–	–	–
HF/3-21+G*	63.4	−25.3 (−23)[e]	38.1	–	13.9 (−24.2)	24.3 (−14)	34.1 (−4.0)
HF/6-31++G**	32.0	−26.1[f]	5.9	–	–	−0.42 (−6.3)	–
HF/6-31+G*//3-21+G*	37.5	−25.3[d]	12.2	–	−7.5 (−19.7)	1.3 (−10.9)	8.8 (−3.4)
BLYP/6-31+G*// HF/3-21+G*	26.5	−25.3[d]	1.2	–	−12.4 (−14)	−7.5 (−8.7)	−1.3 (−2.5)
BLYP/6-31++G**// HF/6-31++G**	24.6	−25.3[d]	−0.7	–	−11.9 (−11)	−5.4 (−4.7)	–
MP2/6-31+G*// HF/3-21+G*	40.3	−25.3[d]	15.0	–	–	5.4 (−9.6)	–
MP2/6-31++G*// HF/6-31++G**		−26.1[f]		–	–	7.3	–

[a] Isodensity set to 0.0004
[b] PCM (UAHF). Initial number of tesserae = 100. Only the electrostatic contribution to the solvation free energy (ΔG^{solv}) was included
[c] Cavity radius (Å): 5.06 (**2**), 4.53 (interm-1), 3.55 ($CH_3SO_3^-$), 3.08 (N_3^-), 4.83 (interm-2), 3.09 (HN_3)
[d] Thermal correction (ΔG_T^{gas}) from HF/3-21+G* level
[e] The HF/3-21+G* thermal free-energy corrections not corrected for internal rotational low-frequency modes
[f] Thermal free-energy correction calculated using the HF/6-31++G** fully optimized geometry for harmonic frequency calculation

care is needed with the low-frequency modes. If the energy differences are small this correction must be taken into account [25], otherwise wrong predictions can be made. It can also be seen from the second column of Tables 1 and 2 (processes Ia and IIIc) that thermal corrections evaluated at the HF/3-21+G* level differ from those evaluated at the HF/6-31++G** level by less than 1 kcal mol^{-1}. Therefore, for the range of relative energy values involved in the processes described here the HF/3-21+G* level of calculation is quite satisfactory for the evaluation of thermal energy corrections. On performing HF full-geometry optimizations for the transition-state structures, TS-1, TS-2 and TS-3, the expected convergence criterion was not achieved with the 6-31++G** basis set; therefore, harmonic frequency calculations and further correlated single-point energy calculations using the HF/6-31++G** geometry could not be carried out.

The solvation energy contribution to the total Gibbs free-energy value, i.e. ΔG^{solv}, is given in parentheses (starting in the fourth column of Tables 1, 2, 3, 4, respectively). From the last column of Tables 1 and 2 it can be seen that the SCRF approach is insensitive both to the level of theory and to the molecular structure, predicting values of ΔG^{solv} in the range from −8 to −1 kcal mol^{-1}. Attention should be paid to the charged and neutral species, where the SCRF method gives essentially the same solvation energy. So, it can be concluded that the SCRF approach is inappropriate for describing the solvent effect in the chemical reaction studied here. The same behavior reported in Tables 1 and 2 and discussed in this and preceding paragraphs is found in Tables 3 and 4 for the TS-2, TS-3 and the P1-ax, P1-eq and P2-eq structures.

The description provided by the semiempirical AM1 and PM3 approaches using the COSMO model for solvent effects follow closely the HF/6-31+G*//HF/3-21+G* IPCM calculation, in the sense that the same qualitative trend is predicted. A disagreement was observed only for the epoxide formation, where the AM1–

Table 3. Free energy barrier, $\Delta G^{\#}$, for the two distinct processes involving the epoxide intermediates IIId and IIIe (kcal mol^{-1} at 140 °C) calculated in the gas phase and solution using DMF (ε = 36.7) as a solvent. ΔG_T^{gas} is the thermal correction including ΔE^0. The solvation energy contribution, ΔG^{solv}, is given in *parentheses*

	$\Delta E_{ele}^{\#}$	$\Delta G_T^{\#gas}$	$\Delta G^{\#gas}$	$\Delta G^{\#sol}$			
				COSMO	IPCM[a]	PCM[b]	SCRF[c]
Process IIId: TS-2[f]							
AM1	15.0	15.7	30.7	62.1 (31.4)	–	–	–
PM3	18.8	16.2	35.0	72.5 (37.5)	–	–	–
HF/3-21+G*	−17.5	18.4 (17)[e]	0.9	–	34.2 (33.3)	28.2 (27.3)	8.8 (7.9)
HF/6-31+G*//3-21+G*	4.9	18.4[d]	23.3	–	53.8 (30.5)	48.7 (25.4)	29.3 (6.0)
BLYP/6-31+G*// HF/3-21+G*	−9.8	18.4[d]	8.6	–	34.5 (25.9)	33.1 (24.5)	14.2 (5.6)
MP2/6-31+G*// HF/3-21+G*	−12.2	18.4[d]	6.2	–	–	31.2 (25.0)	–
Process IIIe: TS-3[f]							
AM1	17.0	15.9	32.9	62.5 (29.6)	–	–	–
PM3	21.6	16.9	38.5	83.2 (44.7)	–	–	–
HF/3-21+G*	−10.7	19.4 (16)[e]	8.7	–	42.8 (34.1)	34.3 (25.6)	16.4 (7.7)
HF/6-31+G*//3-21+G*	11.9	19.4[d]	31.3	–	62.4 (31.1)	55.3 (24.0)	37.1 (5.8)
BLYP/6-31+G*// HF/3-21+G*	−4.7	19.4[d]	14.7	–	41.6 (26.9)	38.2 (23.5)	–
MP2/6-31+G*// HF/3-21+G*	−5.3	19.4[d]	14.1	–	–	38.0 (23.9)	–

[a] Isodensity set to 0.0004
[b] PCM (UAHF). Initial number of tesserae = 100. Only the electrostatic contribution to the solvation free energy (ΔG^{solv}) was included
[c] Cavity radius (Å): 6.72 (TS-2), 6.57 (TS-3)
[d] Thermal correction (ΔG_T^{gas}) from HF/3-21+G* level
[e] The HF/3-21+G* thermal free-energy corrections not corrected for internal rotational low-frequency modes
[f] The expected convergence criterion for geometry optimization could not be attained with the 6-31++G** basis set, therefore the respective harmonic frequency calculations were not performed

COSMO calculation predicted that the route through interm-2 would be favored slightly relative the one involving interm-1 (see Table 2). So, the AM1/PM3–COSMO calculations can provide qualitatively the correct description of the reaction mechanism. Nevertheless, for an assessment of the size of the energy barriers, such as those calculated for process IIa (Table 1) an ab initio correlated level of theory is required.

The results reported in Tables 1, 2, 3 and 4 show explicitly the three components of the total Gibbs free energy in solution, ΔG^{sol}, i.e. the gas-phase electronic plus nuclear repulsion term, ΔE_{ele}, the thermal energy correction, ΔG_T^{gas}, and the solvation energy contribution, ΔG^{solv}, in addition to the Gibbs free energy for the gas-phase species ($\Delta G^{gas} = \Delta E_{ele} + \Delta G_T^{gas}$). By comparing the energy decomposition for the distinct theoretical level and distinct solvent models we can assess unambiguously the behavior of each method used. Among the various computational procedures utilized we can take the MP2/6-31+G*//HF/3-21+G* level of calculation as the best and most viable computational approach and use it as a reference for comparisons. In view of the size of the molecular systems treated here this is a justifiable assumption. The first comment that can be made is the confirmation that the HF level is good for structural predictions; however, it is not consistently satisfactory for energy-difference predictions. Then, we can use HF geometries in single-point energy post-HF calculations. The behavior of the BLYP/6-31+G* functional for the calculation of ΔE_{ele} is to consistently underestimate it compared to the MP2/6-31+G* values, the degree of disagreement varying depending on the molecular species involved, with the largest deviation occurring for the carbocation species (about 30 kcal mol^{-1}, Table 1).

The MP2/6-31+G* gas-phase relative Gibbs free-energy predictions, ΔG^{gas}, at 140 °C yielded the following values for the carbocation, TS-1 and interm-2, respectively, 144, 29.9 and 23.1 kcal mol^{-1}, showing that process IIIb passing through an epoxide intermediate is favored in relation to the direct attack of the N_3^- species. Therefore, in the gas phase the formation of the epoxide intermediate would be the most probable process, passing through the alkoxide interm-2. It can be seen that this prediction is also obtained at all the levels of calculation reported in Tables 1 and 2, except for the HF/3-21+G* result, where the epoxide intermediate is higher in energy than the TS-1 structure by about 9 kcal mol^{-1}, confirming that the small 3-21+G* basis set is not adequate for relative-energy predictions. It should be mentioned that if thermal corrections are neglected and only the relative energies of isolated species in a perfect vacuum (ΔE_{ele}) are considered, process IIa, via the transition-state TS-1 structure, would be the preferred one relative to process IIIb (Tables 1, 2).

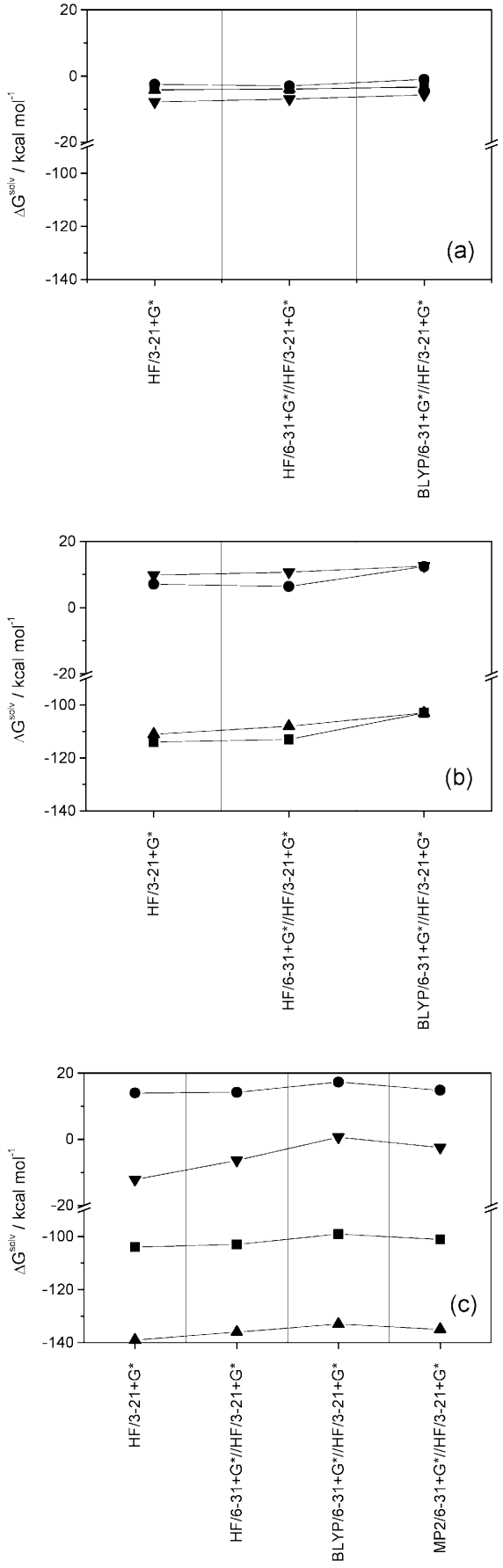
Let us now discuss the solvation energy term (IPCM and PCM), which certainly deserves our attention. The effect of the basis set is not very pronounced, with the maximum deviation between the HF/6-31+G*//HF/3-21+G* and HF/3-21+G* results being about 4 kcal mol^{-1} (IPCM) and 6 kcal mol^{-1} (PCM), respectively, for processes IIIc and IIIb. The effect of the

Table 4. Thermodynamic analysis (kcal mol^{-1}) for the possible products resulting from the reaction compound $\mathbf{2} + N_3^- \rightarrow$ product + $CH_3SO_3^-$ in DMF at 140 °C, 1 atm. ΔG_T^{gas} is the thermal correction including ΔE^0. The solvation energy contribution, ΔG^{solv}, is given in *parentheses*

	ΔE_{ele}	ΔG_T^{gas}	ΔG^{gas}	ΔG^{sol}			
				COSMO	IPCM[a]	PCM[b]	SCRF[c]
P1-ax							
AM1	−24.0	−5.1	−29.1	−22.9 (6.2)	–	–	–
PM3	−8.4	−5.5	−13.9	−2.8 (11.1)	–	–	–
HF/3-21+G*	−35.7	−4.9 (2.4)[e]	−40.6	–	−37.2 (3.4)	−39.0 (1.6)	−38.4 (2.2)
HF/6-31++G**	−28.6	−4.9[d]	−33.5	–	–	−30.4 (3.1)	–
HF/6-31+G*//3-21+G*[d]	−31.2	−4.9[d]	−36.1	–	−33.4 (2.7)	−33.5 (2.6)	−34.8 (1.3)
BLYP/6-31+G*// HF/3-21+G*	−23.0	−4.9[d]	−27.9	–	–	−24.7 (3.2)	−25.6 (2.4)
BLYP/6-31++G**// HF/6-31++G**	−22.5	−4.9[d]	−27.4	–	−25.0 (2.4)	−23.5 (3.9)	–
MP2/6-31+G*// HF/3-21+G*	−19.0	−4.9[d]	−23.9	–	–	−21.1 (2.8)	–
MP2/6-31++G**// HF/6-31++G**	–	−4.9[d]	–	–	–	−20.2	–
P1-eq							
AM1	−28.2	−3.7	−31.9	−25.5 (6.4)	–	–	–
PM3	−12.3	−5.2	−17.5	−1.1 (16.4)	–	–	–
HF/3-21+G*	−26.1	−5.2 (1.4)[e]	−31.3	–	−44.0 (−13)	−34.3 (−3.0)	−29.9 (1.4)
HF/6-31++G**	−23.7	−5.2[d]	−28.9	–	–	−28.8 (0.1)	–
HF/6-31+G*//3-21+G*[d]	−25.1	−5.2[d]	−30.3	–	−41.9 (−12)	−31.3 (−1.0)	−29.6 (0.7)
BLYP/6-31+G*// HF/3-21+G*	−18.5	−5.2[d]	−23.7	–	–	−23.4 (0.3)	−22.1 (1.6)
BLYP/6-31++G**// HF/6-31++G**	−19.1	−5.2[d]	−24.3	–	−30.1 (−5.8)	−22.8 (1.5)	–
MP2/6-31+G*// HF/3-21+G*	−13.2	−5.2[d]	−18.4	–	–	−18.8 (−0.4)	–
MP2/6-31++G**// HF/6-31++G**	–	−5.2[d]	–	–	–	−18.7	–
P2-eq							
AM1	−24.7	−4.5	−29.2	−23.8 (5.4)	–	–	–
PM3	−10.6	−5.3	−15.9	−3.2 (12.7)	–	–	–
HF/3-21+G*	−24.5	−5.0 (0.7)[e]	−29.5	–	−50.5 (−21)	−33.3 (−3.8)	−27.0 (2.5)
HF/6-31++G**	−26.8	−5.0[d]	−31.8	–	–	−29.9 (1.9)	–
HF/6-31+G*//3-21+G*[d]	−25.4	−5.0[d]	−30.4	–	−50.3 (−20)	−32.1 (−1.7)	−28.8 (1.6)
BLYP/6-31+G*// HF/3-21+G*	−17.7	−5.0[d]	−22.7	–	–	−23.5 (−0.8)	−20.6 (2.1)
BLYP/6-31++G**// HF/6-31++G**	−21.0	−5.0[d]	−26.0	–	−33.2 (−7.2)	−23.5 (2.5)	–
MP2/6-31+G*// HF/3-21+G*	−11.4	−5.0[d]	−16.4	–	–	−17.9 (−1.5)	–
MP2/6-31++G**// HF/6-31++G**	–	−5.0[d]	–	–	–	−19.0	–

[a] Isodensity set to 0.0004
[b] PCM (UAHF). Initial number of tesserae = 100. Only the electrostatic contribution to the solvation free energy (ΔG^{solv}) was included
[c] Cavity radius (Å): 4.64 (P1-ax), 4.79 (P1-eq), 4.64 (P2-eq)
[d] Thermal correction (ΔG_T^{gas}) from HF/3-21+G* level
[e] The HF/3-21+G* thermal free-energy corrections not corrected for internal rotational low-frequency modes

electronic correlation can be assessed through the BLYP and MP2 calculations compared with the HF/6-31+G* ones. The correlation effect is substantial for the IPCM calculation, amounting to about 10 kcal mol^{-1} (BLYP/6-31+G*//HF/3-21+G*) in the carbocation case, while the corresponding PCM variation is 4 kcal mol^{-1}. The greatest correlation effects on the solvation energy using the PCM were found for process IIIb: 7 kcal mol^{-1} (BLYP/6-31+G*//HF/3-21+G*) and 4 kcal mol^{-1} (MP2/6-31+G*//HF/3-21+G*). So it can be seen that the IPCM is much more sensitive to electron correlation effects than the PCM. The way that electron correlation is handled in DFT calculations is different from the standard post-HF treatment, like MP2. So, it is interesting to compare the BLYP and MP2 results. We discuss the PCM results, where a large set of calculations was performed. First, regarding the basis set effects it can be seen from Tables 1 (process Ia) and 2 (process IIIc) that the increase of the basis set to 6-31++G** does not affect appreciably the MP2 results (maximum change of about 2 kcal mol^{-1}), considering that the energy differences involved here are of the order of more than about



10 kcal mol^{-1}, therefore not affecting our predictions. The corresponding effect on the BLYP values is slightly larger. It should be said that the effects of basis set and electron correlation on the equilibrium structures of the products reported in Table 4, regarding the PCM calculation, are much smaller. Some of the species addressed in Tables 1, 2 and 3 are charged species (and also TS structures), which may be expected to be more difficult to treat. Finally, by comparing the IPCM and PCM results through Tables 1, 2, 3 and 4 the same overall trend is observed when the basis set and the level of theory are varied. However, when the solvation energies found for the distinct processes are compared, the behavior of the continuum models used is quite different. The analysis depicted in Fig. 1 is very interesting, showing that the SCRF method is not very sensitive to the reaction processes, with the solvation energy for the reactions ranging from -1 to -8 kcal mol^{-1}, as previously discussed. The IPCM results are grouped in two sets of values depending on the process. For the reactions involving anions as products (IIa and IIIb) the Gibbs solvation free energy is small and is very similar. Processes Ia and IIIa, which involve cations as products, presented similar and very negative solvation energies. The values from the PCM are sensitive to the process and to the structure involved in the reaction mechanism. We then make use of the PCM results in the discussion of the reaction mechanism. It is also opportune to say by analyzing the results in Fig. 1 that the HF/3-21 + G* level of calculation is sufficient for the evaluation of ΔG^{solv} (IPCM and PCM) to a precision of about 3 kcal mol^{-1}, compared to the MP2/6-31 + G* values. The only exception was process IIIb (interm-2), where a variation of about 10 kcal mol^{-1} between the results from HF/3-21 + G* and MP2/6-31 + G*//HF/3-21 + G* was found. So, if the energy differences involved are much larger than 3 kcal mol^{-1}, the HF/3-21 + G* level can safely be used to generate the solvation energy contribution. Then, in this case, the thermal correction (ΔG_T^{gas}) and the solvation part (ΔG^{solv}) can be evaluated using the HF level and the electronic plus nuclear repulsion contribution (ΔE_{ele}) is calculated using a correlated level of theory (e.g. MP2).

It is important to mention that by analyzing the results reported in Tables 1 and 2, if the solvent effects are considered neglecting the thermal corrections, the TS-1 structure would be preferred over the interm-1 one by about 4 kcal mol^{-1}. Therefore, it can be seen that the thermal correction plays a crucial role for the correct description of the mechanism of reaction between a quebrachitol derivative and sodium azide. This result stresses the necessity of a computationally feasible method for harmonic frequency calculation (such as HF/3-21 + G*) in order to obtain satisfactory thermal energy correction values.

◀

Fig. 1. Solvation-energy dependence on the basis set and the level of theory: **a** self-consistent reaction field, **b** isodensity polarized continuum model and **c** polarized continuum model–united atom model for Hartree–Fock (*HF*). Ia, carbocation (*squares*); IIa, TS-1 (*circles*); IIIa, interm-1 (*up triangles*); IIIb, interm-2 (*down triangles*)

It is relevant from the computational point of view to assess the ability of the BLYP functional to describe the energetics of the processes described in Sect. 2 by comparing the BLYP/6-31+G* and MP2/6-31+G* results, using the PCM for solvent effects. For the negatively charged transition-state structures TS-1, TS-2 and TS-3, we found good agreement, the maximum deviation being about 2 kcal mol^{-1}. For the intermediate carbocation and epoxide structures the agreement is not consistent, with a deviation of 27 (carbocation), 14 (interm-1) and 9.6 kcal mol^{-1} (interm-2), with the BLYP energies being underestimated. For the minimum-energy structures of the products (P1-ax, P1-eq and P2-eq) the agreement is reasonable with a maximum deviation of about 5 kcal mol^{-1}; again the BLYP functional underestimates the stabilization energies (the energies are more negative). Despite this rather large variation in the relative energies, the BLYP/6-31+G* prediction follows closely the MP2/6-31+G* trend, in the sense that the predicted Gibbs free-energy differences would lead to the same conclusions regarding the reaction mechanism. However, the MP2 energy values are used in the following discussion of the reaction mechanism, since MP2 is our best level of calculation.

Now that a comparison between the distinct approaches has been made we turn to the analysis of the reaction mechanism. The first reaction pathway investigated passes through a carbocation intermediate (Scheme 3), first step, and then the attack of the N_3^- species occurs, leading to two possible products: axial (with retention of configuration), named P1-ax, or equatorial, named P1-eq, as shown by Scheme 2 (process I). The formation of the carbocation is the key step for the reaction to take place via this mechanism. The relevant HF/3-21+G* fully optimized three-dimensional structures located on the PES corresponding to process I are depicted in Fig. 2. Values of the total energy and the Gibbs free-energy difference both in the gas phase and in solution are reported in Table 1. It can be seen that the carbocation is enormously stabilized by solvent effects, the Gibbs free energy in DMF at 140 °C being 43.2 kcal mol^{-1} (MP2/6-31+G*//HF/3-21+G*).

The second process investigated was the direct attack of N_3^- on compound **2**, passing through a transition state, named TS-1 (see Scheme 3), followed by elimination of the $CH_3SO_3^-$ species, forming an equatorial product (P1-eq) with inversion of configuration (Scheme 2, process II). The HF/3-21+G* fully opti-

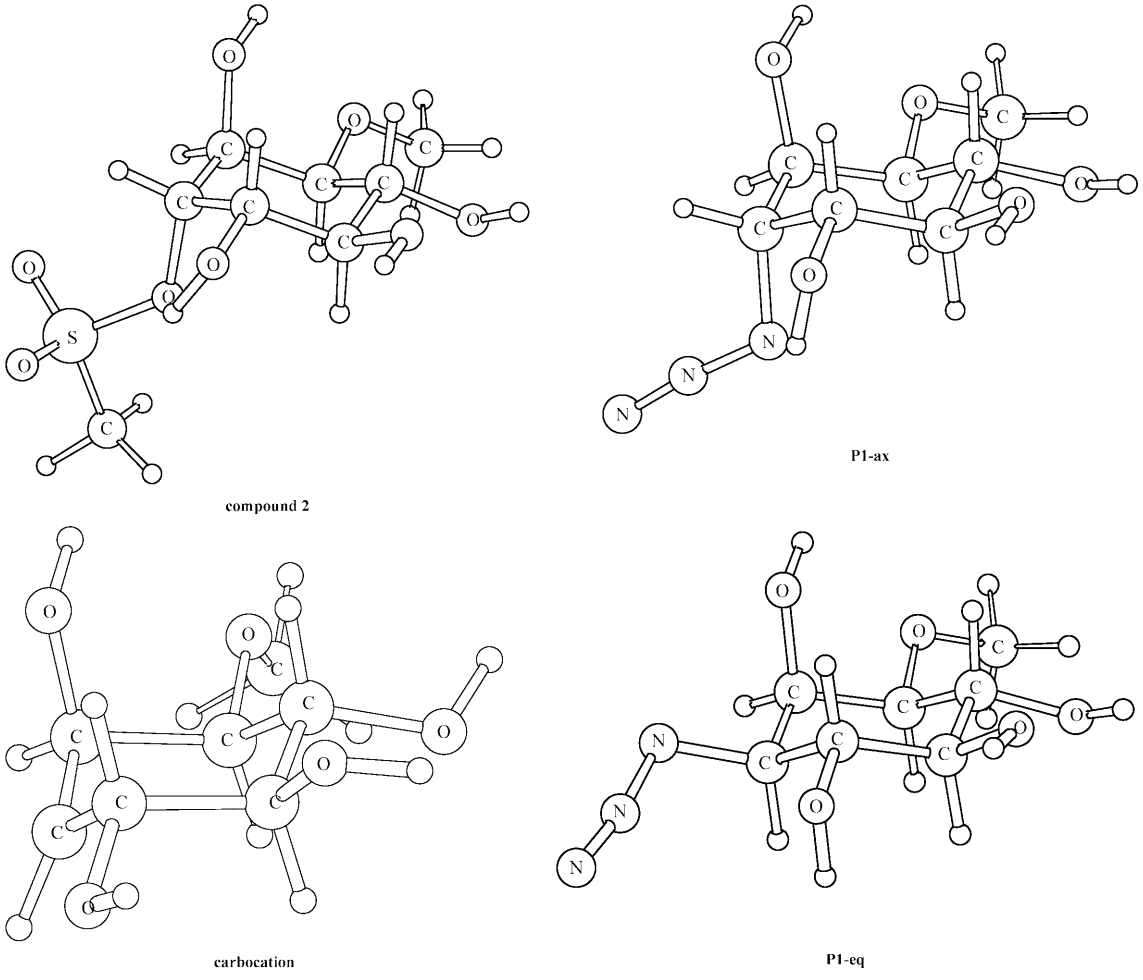


Fig. 2. HF/3-21+G* fully optimized structures for the mesylated compound **2**, the carbocation corresponding to process I and the possible reaction products: axial (*P1-ax*) and equatorial (*P1-aq*)

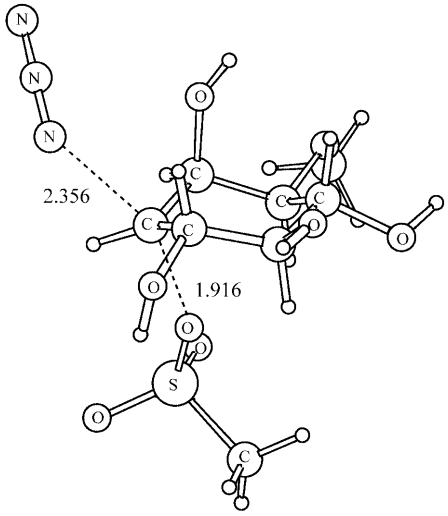


Fig. 3. HF/3-21+G* fully optimized structure for transition state TS-1 corresponding to process II (values in angstroms)

mized transition-state structure TS-1 located on the PES for process IIa is depicted in Fig. 3 and Table 1 reports the values of the total energy and the Gibbs free-energy difference both in the gas phase and in solution. Differently from the carbocation case, the TS-1 structure is not stabilized by the solvent, with the energy barrier being higher when the solvent effect is included. The energy barrier in solution for process IIa is 44.7 kcal mol^{-1}, which makes this process comparable to the formation of the carbocation.

Finally, we studied the reaction mechanism through an epoxide intermediate (see Scheme 3), where the determining step for this process is the formation of the epoxide species as shown in Scheme 2, process III. Two distinct routes were analyzed: IIIa, passing through a positively charged intermediate (interm-1), and IIIb, through the alkoxide intermediate (interm-2). In the latter mechanism the azide acts as a basis for ionizing the hydroxyl proton. The energy results are given in Table 2 and the HF/3-21+G* epoxide minimum-energy structures are shown in Fig. 4. The epoxide formed can undergo two possible transformations leading to two distinct products, depending on the position of the attack of the N_3^- species, passing through two different TS structures, named TS-2 and TS-3, as shown in Scheme 2 (processes IIId and IIIe) and Fig. 5. The calculated relative energy values for processes IIId and IIIe are given in Table 3.

From Tables 1 and 2 it can be seen that the gas-phase and solution Gibbs free-energy barrier for the epoxide formation is notably smaller than the respective values for the carbocation formation and the TS-1 transition-state structure. The Gibbs free-energy difference in DMF favors the epoxide intermediate by more than 30 kcal mol^{-1} (MP2). The internal displacement of the mesylate by the OH group (process IIIa) is favored relative to process IIIb by about 16 kcal mol^{-1} (PCM-MP2/6-31+G*//HF/3-21+G*). So the epoxide formation should pass through interm-1. It is interesting to note that the epoxide is less stable than interm-1 in

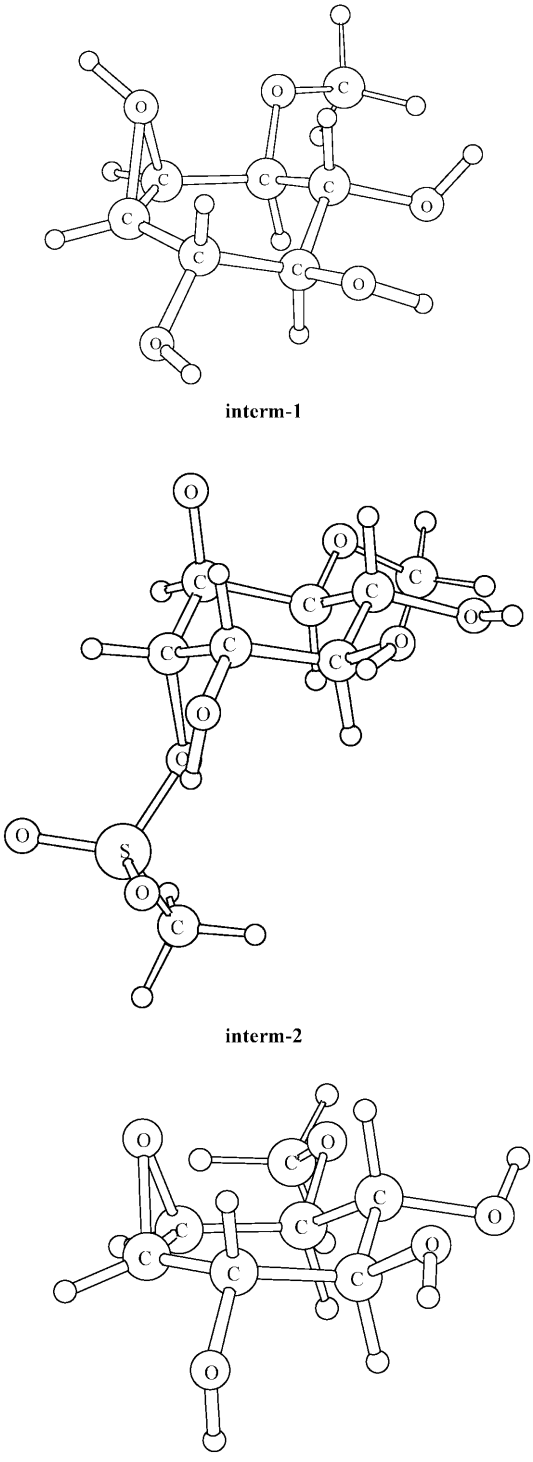


Fig. 4. HF/3-21+G* fully optimized structure for the epoxide intermediates corresponding to process III

solution by only 1.4 kcal mol^{-1} (PCM-MP2). If the uncorrected thermal corrections to the Gibbs free energy are used, ΔG^{sol} at PCM-MP2/6-31+G*//HF/3-21+G* is 8.0 (interm-1) and 7.7 kcal mol^{-1} (epoxide). So, it can be said that the epoxide formation should occur with a very low energy barrier. This important result confirms unambiguously that the reaction mechanism is definitively via the epoxide formation. The fate

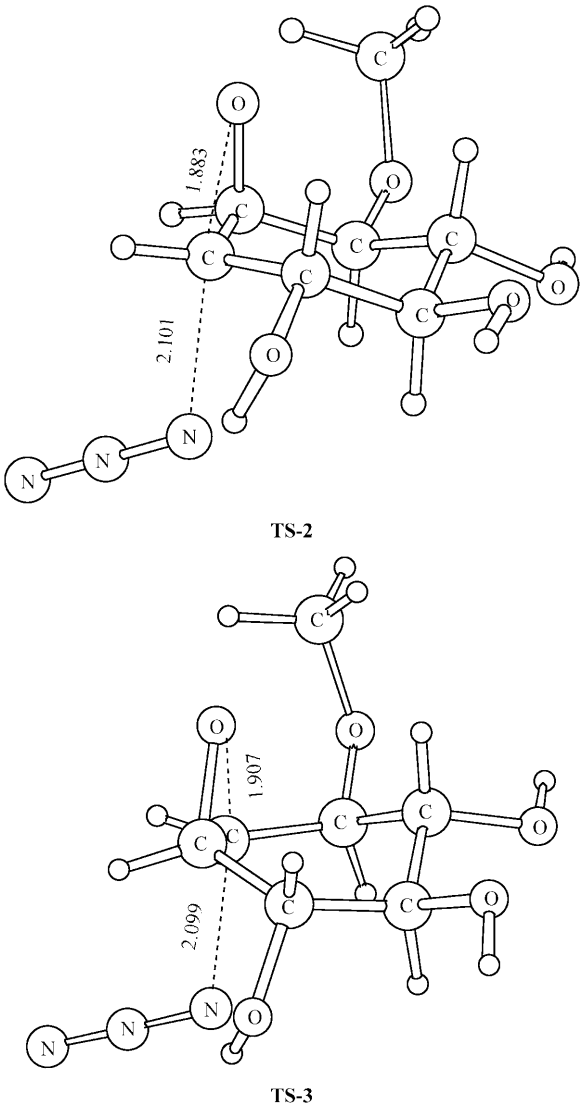


Fig. 5. HF/3-21+G* fully optimized structures for transition states *TS-2* and *TS-3*, corresponding to processes IIIb and IIIc, respectively (values in angstroms)

of the epoxide is crucial to decide which product is formed. It can be seen from Table 3 that process IIId leading to the axial P1-ax product is more favorable than process IIIe by a reasonable amount of about 7 kcal mol^{-1}; therefore, the main product of the reaction between compound **2** and N_3^- should be the axial structure with retention of configuration (P1-ax) given in Fig. 2, as proposed experimentally. The structure of product P2-eq formed through the pathway from process IIIe is shown in Fig. 6, and is rather different from the P1-eq spatial arrangement with the nitrogen atom being attached to a different carbon atom, adjacent to the methoxyl group.

Table 4 reports the stabilization energies of the possible products resulting from the reaction between compound **2** and N_3^- in DMF at 140 °C. From Table 4 it can be seen that the three possible products are thermodynamically favorable, with the Gibbs free energy of solvation differing within about 3 kcal mol^{-1}. The order of

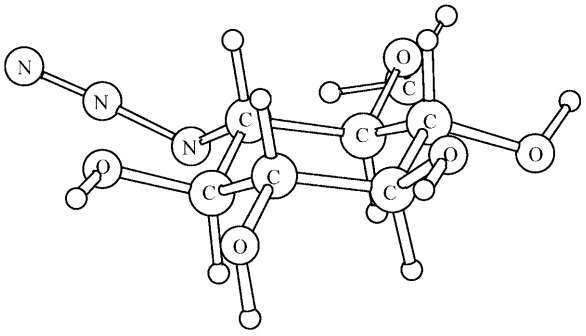


Fig. 6. HF/3-21+G* fully optimized structure of product P2-eq corresponding to process IIIc

the relative stabilization in solution is P1-ax > P1-eq > P2-eq. The P1-ax compound is only about 2 kcal mol^{-1} more stable than P1-eq and favored by about 3 kcal mol^{-1} in relation to P2-eq. Nevertheless the mechanisms that lead to the P1-eq and P2-eq products in solution are unfavorable in relation to the formation of the axial product by about 12 and about 8 kcal mol^{-1}, respectively.

The MP2/6-31+G*//HF/3-21+G* energy-profile diagram for the reaction between a quebrachitol derivative and sodium azide is depicted in Fig. 7, where the energy heights for the distinct mechanisms can be easily visualized. From Fig. 7 some final and concluding observations can be made. From the thermodynamic point of view, all three products could be formed since the relative energy differences involved are within the variation found among the theoretical methods used for gas-phase and solution calculations. Despite this, P1-ax is still the thermodynamical product. However, by analyzing the reaction mechanism (Fig. 7) we can undoubtedly affirm that P1-eq has a negligible chance to be formed. The same is true for the P2-eq product. Therefore, the energy diagram given in Fig. 7 provides very strong support for the epoxide mechanism leading to the P1-ax product, on a sound basis, in agreement with the experimental finding. As the energy difference favoring the epoxide mechanism is not small at all (more than 30 kcal mol^{-1}) we can be quite confident that this is the preferred mechanism in solution at 140 °C, a conclusion that could be reached using any of the theoretical approaches described in Sect. 2, beside MP2.

4 Conclusions

In this work we investigated theoretically the mechanism for the chemical reaction between a quebrachitol derivative (**2**) and the N_3^- species in solution (DMF) at 140 °C. This reaction was previously studied experimentally using the same conditions, with the main product being assigned on the basis of the analysis of the NMR spectra. We performed semiempirical (AM1 and PM3), ab initio (HF and MP2) and DFT (BLYP functional) calculations, using the 3-21+G*, 6-31+G* and 6-31++G** basis sets, for the stationary points present on the PES for this reaction in order to provide an explanation for the surprisingly observed axial product,

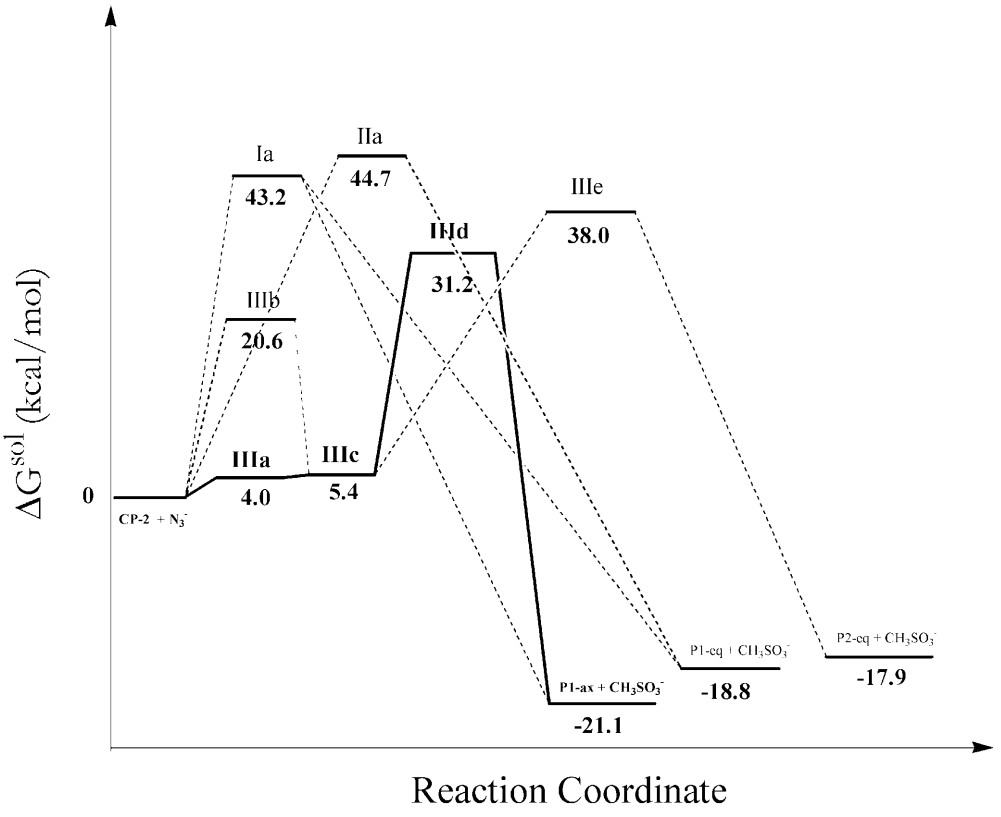


Fig. 7. Energy-profile diagram for the reaction between compound **2** (*CP-2*) and N_3^-. The Gibbs free-energy values (ΔG^{sol}) were obtained at the MP2/6-31+G*//HF/3-21+G* level using the polarized continuum solvation model (dimethylformamide, $\varepsilon = 36.7$). The most favorable mechanism is represented in *bold*

with retention of configuration, in the experiment. The solvent effect was evaluated using continuum models at the semiempirical level (COSMO) and ab initio and DFT levels (SCRF, IPCM, PCM). We investigated three distinct reaction pathways involving a carbocation, an epoxide and a TS structure for a direct attack of the N_3^- species on compound **2**. The preferred mechanism in solution is through an epoxide intermediate followed by attack of the azide species leading, according to the Gibbs free-energy barrier in solution, to the axial product P1-ax, with retention of configuration, as the main reaction product. The MP2/6-31+G*//HF/3-21+G* calculation favors the key step of process III, i.e. the epoxide formation, by more than 30 kcal mol^{-1} in relation to the formation of the carbocation intermediate or the TS-1 transition-state structure (processes I and II, respectively). This excellent agreement between the theoretical predictions and experimental findings, confirming the assignments of the NMR spectra, strongly stimulates combined theoretical/experimental projects related to the elucidation of mechanisms of chemical reactions in solution.

5 Supplementary material

The optimized coordinates of all the structures considered in the present work and the main structural parameters are available on request to the authors.

Acknowledgements. This research was supported by the Brazilian agencies Conselho Nacional de Desenvolvimento Científico e Tecnológico and Fundação de Amparo a Pesquisa do Estado de Minas Gerais. W.B.DeA. would like to thank the members of the Departamento de Química, Universidade Federal de Juiz de Fora for the very pleasant stay during the time that this work was carried out and the Pró-Reitoria de Pós-Graduação e Pesquisa for a visiting professor fellowship. The authors also thank the CENAPAD-MG/CO-NAR-UFJF for providing computational facilities.

References

1. Da Silva ET, Le Hyaric M, Machado AS, De Almeida MV (1998) Tetrahedron Lett 39: 6659
2. (a) Ogawa S, Isaka A (1991) Carbohydr Res 210: 105; (b) Arjona O, De Dios A, Plumet J, Saez B (1995) Tetrahedron Lett 8: 1319
3. (a) Aceña JL, Arjona O, Iradier F, Plumet J (1996) Tetrahedron Lett 37: 105; (b) Pettit GR, Gaddamidi V, Cragg GM, Herald DL, Sagawa Y (1984) J Chem Soc Chem Commun 1693
4. Kozikowski AP, Fauq AH, Powis G, Kurian P, Crews FT (1992) J Chem Soc Chem Commun 362
5. Kozikowski AP, Fauq AH, Powis G, Melder DC (1990) J Am Chem Soc 112: 4528
6. Brunn G, Fauq AH, Chow S, Kozikowski AP, Gallegos A, Powis G (1994) Cancer Chemother Pharmacol 35: 71
7. De Almeida MV, Figueiredo RM, Dos Santos HF, Da Silva AD, De Almeida WB (2001) Tetrahedron Lett 42: 2767
8. Truong TN (1998) Int Rev Phys Chem 17: 525
9. Dewar MJS, Zoebish EG, Healy EF, Stewart JJP (1985) J Am Chem Soc 107: 902
10. Stewart JJP (1989) J Comput Chem 209: 221
11. Stewart JJP (1993) MOPAC 93 manual. Fujitsu, Tokyo
12. (a) Binkley JS, Pople JS, Hehre WJ (1980) J Am Chem Soc 102: 939; (b) Clark T, Chandrasekhar J, Spitznagel GW, Schleyer PvR (1983) J Comput Chem 4: 294
13. Møller C, Plesset MS (1934) Phys Rev 46: 618

14. Parr RG, Yang W (1989) Density functional theory of atoms and molecules. Oxford University Press, New York
15. (a) Becke AD (1988) Phys Rev A 38: 3098; (b) Lee C, Yang W, Parr RG (1993) Phys Rev B 37: 785
16. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA Jr, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma K, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Head-Gordon M, Replogle ES, Pople JA (1998) Gaussian 98, revision A.6. Gaussian, Pittsburgh, Pa
17. Klamt A, Schürmann G (1993) J Chem Soc Perkin Trans 2: 799
18. Onsager L (1936) J Am Chem Soc 58: 1486
19. Foresman JB, Keith TA, Wiberg KB, Snoonian J, Frish MJ (1996) J Phys Chem 100: 16098
20. (a) Cossi M, Barone V, Camimi R, Tomasi J (1996) Chem Phys Lett 255: 327; (b) Barone V, Cossi M, Tomasi J (1997) J Chem Phys 107: 3210
21. (a) De Almeida WB, Dos Santos HF, Zerner MC (1998) J Pharm Sci 87: 1101; (b) Dos Santos HF, De Oliveira LFC, Dantas SO, Santos PS, De Almeida WB (2000) Int J Quantum Chem 80: 1076
22. (a) Dos Santos HF, De Almeida WB, Zerner MC (1998) J Pharm Sci 87: 190; (b) Dos Santos HF (2001) Quim Nova (in press)
23. Ayala PY, Schlegel HB (1998) J Chem Phys 108: 2315
24. Ochterski JW (2000) Thermochemistry in Gaussian. http://www.gausssian.com/thermo.htm
25. Dos Santos HF, Rocha WR, De Almeida WB (2001) (submitted)