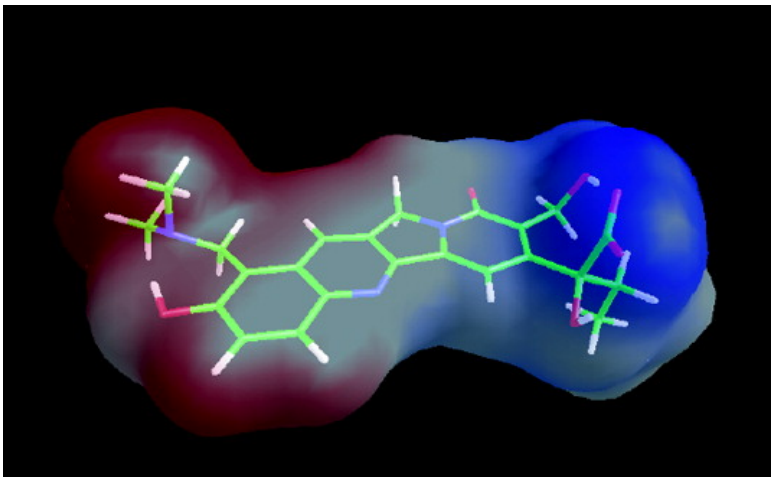
Article

New Hints on the pH-Driven Tautomeric Equilibria of the Topotecan Anticancer Drug in Aqueous Solutions from an Integrated Spectroscopic and Quantum-Mechanical Approach

Nico Sanna, Giovanni Chillemi, Andrea Grandi, Silvia Castelli, Alessandro Desideri, and Vincenzo Barone

J. Am. Chem. Soc., **2005**, 127 (44), 15429-15436• DOI: 10.1021/ja052637u • Publication Date (Web): 14 October 2005



J|A|C|S
ARTICLES
Published on Web 10/14/2005

New Hints on the pH-Driven Tautomeric Equilibria of the Topotecan Anticancer Drug in Aqueous Solutions from an Integrated Spectroscopic and Quantum-Mechanical Approach

Nico Sanna,[§] Giovanni Chillemi,[§] Andrea Grandi,[§] Silvia Castelli,[†] Alessandro Desideri,*[,†] and Vincenzo Barone*[,‡]

Contribution from CASPUR, Consortium for Supercomputing in Research, Via dei Tizii 6/b, 00185 Roma, Italy, INFM and Dipartimento di Biologia, Università di Roma "Tor Vergata", Via della Ricerca Scientifica, 00133 Roma, Italy, and Dipartimento di Chimica, Università di Napoli "Federico II", Complesso Universitario di Monte S. Angelo, Via Cintia, 80126 Napoli, Italy

Received April 22, 2005; E-mail: desideri@uniroma2.it; baronev@unina.it

Abstract: The equilibria between the different forms of the topotecan anticancer drug have been studied at moderately acidic and physiological pH by an integrated computational tool rooted in the density functional theory and its time-dependent extension together with the polarizable continuum model. The results allow an unbiased selection between the different possible tautomeric forms and provide invaluable complements to experimental data. The ultraviolet−visible topotecan spectrum, recorded at moderately acidic pH, is accurately reproduced only by TD-DFT computations including solvent effects. Comparison of the experimental and calculated bands of the UV−vis spectrum at physiological pH indicates the presence of an equilibrium among different forms that is tuned by the microenvironment embedding the drug. The quantitative agreement between TD-DFT/PCM computations and experiments allows the identification of unequivocal spectroscopic signatures for different forms of topotecan.

1. Introduction

The camptothecins belong to a class of antitumor agents whose lead compound, camptothecin (CPT), is a naturally occurring alkaloid extracted from the Chinese tree *Camptotheca acuminata*.[1] CPT and its derivatives have been shown to be effective in the treatment of a broad spectrum of human cancers. Topotecan (Hycamtin; GlaxoSmithKline), for example, is the most commonly used agent for the treatment of ovarian carcinoma (the fifth leading cause of cancer deaths in women) and is the only single agent currently approved in the United States for the treatment of small-cell lung cancer (SCLC) recurrent disease, which is among the most lethal malignancies.[2−4] Topotecan, when compared to its parental compound CPT, has increased solubility and stability, decreased toxicity, and a shorter half-life with a consequent lack of drug accumulation when more daily doses are given.[5−7] The molecular target of the CPT family is topoisomerase I, which belongs to a ubiquitous class of enzymes essential for the control of DNA supercoiling generated by replication, transcription, and recombination.[8] Eukaryotic topoisomerase I (Top1) relaxes DNA superhelical tension by introducing a transient single-strand break in one strand of duplex DNA and forming a covalent phosphotyrosyl bond with the 3′-end of the broken DNA strand. CPT drugs stabilize the covalent Top1−DNA complex, thus producing DNA damage after collisions with replication forks or transcription complexes.[9] Selectivity for cancer cells is due to their high proliferation rate, which implies a high level of topoisomerase I activity. The drug displays a high degree of toxicity also against normal tissues that show enhanced proliferative rates, such as the bone marrow, gastrointestinal tract, and hair follicles. Therefore, a full understanding of the factors contributing to tumor sensitivity and resistance is of considerable interest.[10] In particular, the definition of the structural factors of the CPT compound family determining its interaction with the Top1−DNA complex may lead to the engineering of new topoisomerase I poisons with increased efficiency. CPT and its derivatives coexist in two forms at physiological pH: the lactone form (see Figures 1a,b, 2a,b), believed to be the active one, where the E ring is in a closed conformation, and the hydrolyzed open carboxylate form[2] (see Figure 3a,b).

[§] CASPUR.
[†] Università di Roma "Tor Vergata".
[‡] Università di Napoli "Federico II".

(1) Wall, M. E.; Wani, M. C.; Cook, C. E.; Palmer, K. H.; McPhail, A. T.; Sim, G. A. *J. Am. Chem. Soc.* **1966**, *88*, 3888−3890.
(2) Bailly, C. *Crit. Rev. Oncol. Hematol.* **2003**, *45*, 91−108.
(3) Nagourney, R. A.; Sommers, B. L.; Harper, S. M.; Radecki, S.; Evans, S. S. *Br. J. Cancer* **2003**, *89*, 1789−1795.
(4) Garcia-Carbonero, R.; Supka, J. G. *Clin. Cancer Res.* **2002**, *8*, 641−661.
(5) Burke, T. G.; Mi, Z. *J. Med. Chem.* **1994**, *37*, 40−46.
(6) Verweij, J.; Lund, B.; Beijnen, J.; Planting, A.; de Boer-Dennert, M.; Koier, I.; Rosing, H.; Hansen, H. *Ann. Oncol.* **1993**, *4*, 673−678.
(7) Rothenberg, M. L. *Ann. Oncol.* **1997**, *8*, 837−855.
(8) Pommier, Y.; Pourquier, P.; Fan, Y.; Strumberg, D. *Biochim. Biophys. Acta* **1998**, *1400*, 83−106.
(9) Pommier, Y.; Redon, C.; Rao, V. A.; Seiler, J. A.; Sordet, O.; Takemura, H.; Antony, S.; Meng, L. H.; Liao, Z. Y.; Kohlhagen, G.; Zhang, H. L.; Kohn, K. W. *Mutat. Res.* **2003**, *532*, 173−203.
(10) Pommier, Y.; Pourquier, P.; Urasaki, Y.; Wu, J.; Laco, G. S. *Drug Resist. Updates* **1999**, *2*, 307−318.

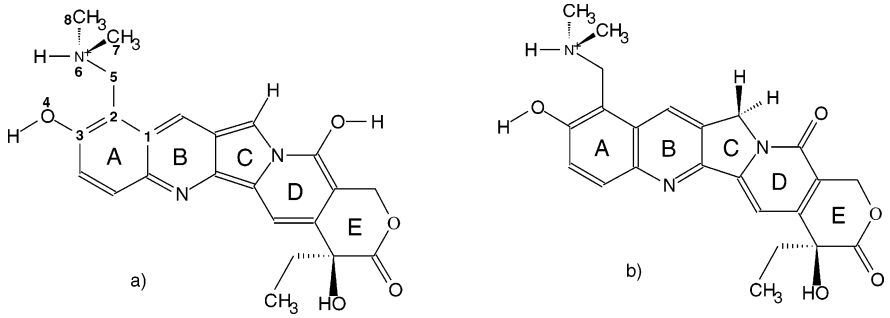


Figure 1. Molecular structures of topotecan type 1: (a) eno form; (b) keto form.

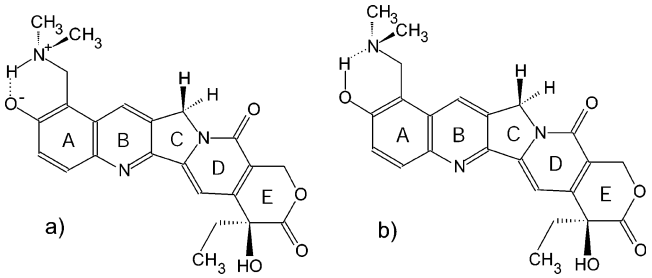


Figure 2. Molecular structures of topotecan type 2. Keto tautomers: (a) *NH* form; (b) *OH* form.

The equilibrium is tuned by the chemical environment, and in particular, it is shifted toward the lactone form upon binding to different DNA sequences.[11] The molecular mechanism that leads to the formation of a topotecan−Top1−DNA complex is not fully understood, although the X-ray diffraction structure of the ternary complex has been recently solved.[12] The 3D structure shows that the human topoisomerase I is covalently joined to a double-stranded DNA, bound to topotecan in equilibrium between the lactone and carboxylate forms, suggesting that also the open form can contribute to the topoisomerase inhibition.[12] In the structure, topotecan intercalates between upstream (−1) and downstream (+1) base pairs, displacing the downstream DNA and thus preventing the religation of the cleaved strand. The coexistence in the crystal of lactone and carboxylate forms[8] suggests that the equilibrium between the two forms can be modulated by subtle enzyme or DNA conformational changes.

A powerful approach to identify the chemical factors influencing the electronic properties of the drug molecule is provided by computational methods. Both classical and quantum-mechanical (QM) studies provide useful information concerning the drug−enzyme interaction. From the classical side, we have carried out molecular dynamics (MD) simulations of the protein−DNA complex that have allowed us to describe the main dynamical properties of the enzyme and their role in DNA recognition.[13−16] More recently, we applied a combined computational and biochemical approach to fully characterize the CPT mimic T718A Top1 mutant, which allowed us to propose a novel interpretation for the mutant lethal phenotype.[17] From the quantum-mechanical point of view, the prediction and interpretation of the structural and spectroscopic behavior of the topotecan molecular system cast several problems since topotecan can exist in different structures due to the flexibility of the fragments attached to the leading CPT polycyclic system and because of the chemical ability to establish keto/enolic tautomeric equilibria on the D ring oxygen (Figure 1). The study of the relative stability of the different forms is computationally quite demanding since it is strongly dependent on the solvent surrounding the molecule. For these reasons the computational studies on the topotecan and the CPT family molecules have been, up to now, limited to in vacuo basic semiempirical (parametric) quantum-mechanical calculations.[18] As a matter of fact, the prediction and interpretation of spectroscopic parameters is a long-standing goal of computational chemistry, but, until very recently, it faced the dilemma of choosing between accurate methods applicable only to very small model systems[19,20] and fast methods based on extensive parametrization, which should be recalibrated for each new class of systems or properties.[21] In the last years, however, the density functional theory (DFT) has produced a third class of reliable linear scaling methods, which includes few parameters, besides not depending on a particular property and/or molecular structure. After great success for ground-state properties, implementation of the powerful time-dependent (TD) formalism[22,23] extends their field of application to excited electronic states. In particular, the so-called hybrid functionals, including some fraction of Hartree−Fock exchange,[24] provide remarkably accurate results for structural, thermodynamic, and spectroscopic parameters. However, physicochemical properties are often strongly tuned by environmental effects, so that solute−solvent interactions cannot be neglected for a meaningful comparison with experimental data. Thanks to significant improvements both in the physical model and in the computational implementation, continuum solvent models (in particular the so-called polarizable continuum model, PCM) offer to-date the best compromise between accuracy and computational efficiency[25] for ground and excited

(11) Yang, D. Z.; Strode, J. T.; Spielmann, H. P.; Wang, A. H. J.; Burke, T. G. *J. Am. Chem. Soc.* **1998**, *120*, 2979−2980.
(12) Staker, B. L.; Hjerrild, K.; Feese, M. D.; Behnke, C. A.; Burgin, A. B.; Stewart, L. *Proc. Natl. Acad. Sci.* **2002**, *99*, 15387−15392.
(13) Chillemi, G.; Castrignanó, T.; Desideri, A. *Biophys. J.* **2001**, *81*, 490−500.
(14) Chillemi, G.; Fiorani, P.; Benedetti, P.; Desideri, A. *Nucleic Acids Res.* **2003**, *31*, 1525−1535.
(15) Fiorani, P.; Bruselles, A.; Falconi, M.; Chillemi, G.; Desideri, A. *I. J. Biol. Chem.* **2003**, *278*, 43268−43275.
(16) Chillemi, G.; Redinbo, M.; Bruselles, A.; Desideri, A. *Biophys. J.* **2004**, *87*, 4087−4097.
(17) Chillemi, G.; Fiorani, P.; Castelli, S.; Bruselles, A.; Benedetti, P.; Desideri, A. *Nucleic Acids Res.* **2005**, *33*, 3339−3350.
(18) Strel'tsov, S. A.; Grokhovskii, S. L.; Kudelina, I. A.; Oleinikov, V. A.; Zhuze, A. L. *Mol. Biol.* **2001**, *35*, 432−441.
(19) Andersson, K.; Malmqvist, P. A.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218−1226.
(20) Noojen, M.; Bartlett, R. J. *J. Chem. Phys.* **1997**, *106*, 6441−6448.
(21) Zerner, M. C. In *Reviews in Computational Chemistry*, Vol. 2; Lipkowitz, K. B., Boyd, D. B., Eds.; VCH Publishers: New York, 1991.
(22) Casida, M. K. In *Recent Advances in Density Functional Methods, Part I*; World Scientific: Singapore, 1995.
(23) Adamo, C.; Scuseria, G. E.; Barone, V. *J. Chem. Phys.* **1999**, *111*, 2889−2899.
(24) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

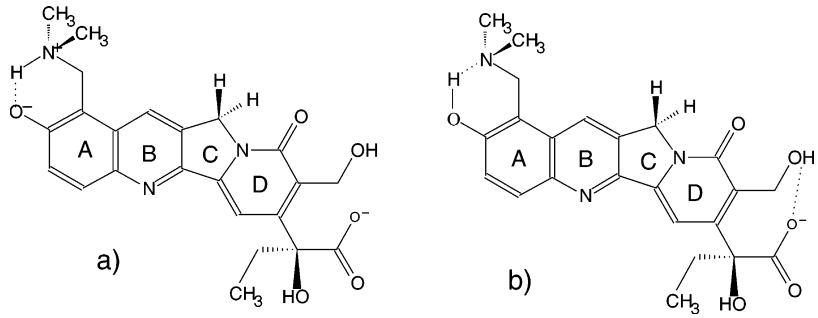


Figure 3. Molecular structures of topotecan type 3. Keto tautomers: (a) *NH* form; (b) *OH* form.

electronic states.[26] Some works along these lines have shown that the TD-DFT/PCM approach is very promising for the study of solvent shifts of electronic spectra.[27,28]

In this study we have applied an integrated computational and experimental (UV−vis) strategy to analyze the electronic and structural properties of the topotecan molecule in water solution at different pH. This approach has allowed us to select the most stable tautomers of the topotecan drug at moderately acidic pH and to accurately reproduce the corresponding experimental UV−vis absorption bands. The application of this methodology to the study of the topotecan drug at physiological pH has shown the coexistence in solution of several topotecan ketonic isomers with different hydrogen bond patterns. For the first time a direct structure−spectroscopic correlation for the topotecan drug in solution has been established, providing a remarkable agreement between theoretical results and measured UV−vis absorption bands at almost neutral pH. The procedure to develop a classical drug force field, starting from QM structural and electronic properties, is also sketched.

2. Materials and Methods

2.1. Experimental Details. Topotecan was a kind gift of Dr. C. Pisano from Sigma Tau. The molar extinction coefficient $\epsilon_{380} = 22\,000$ M^{-1} cm^{-1} at pH 6.8 has been used to determine the topotecan concentration. The samples used for the absorption spectra are composed by 5 μM topotecan in sodium acetate buffer 10 mM (pH 5.3) or Tris-HCl 10 mM (pH 6.8 and 9.5). The UV−vis spectra have been recorded with a Cary 5 spectrophotometer.

2.2. Computational Details. All the computations have been performed with the Gaussian03 package[29] using an increasing level of complexity for the description of the molecular electronic state. Starting from X-ray crystallographic data[12] of the topotecan ternary complex with the DNA−Top1 system, the minimum energy structures have been fully optimized at the HF/3-21G level both in vacuo and in aqueous solution, including in the latter case solvent effects by means of the CPCM variant of the polarizable continuum model (PCM).[26] Correlation effects were next taken into account via the density functional theory (DFT) using the B3LYP[24] hybrid functional and the 6-31G* basis set. Improved energies have been finally obtained by single-point B3LYP/6-311+G** and MP2/6-311+G**[30,31] computations at B3LYP/6-31G* geometries.

Since UV−vis spectra of large molecules are often computed by the semiempirical ZINDO method with remarkable success,[32] we performed single-point computations at this level on the structures previously optimized. Unfortunately, the results were quite disappointing, so that we resorted to the TD-DFT approach[22] at the B3LYP/6-31G* level coupled to a PCM treatment taking into proper account nonequilibrium solvent effects.[33,34] The results have been compared with analogous calculations employing the augmented 6-31+G* basis set, to evaluate the effect of the diffuse functions on the UV absorption spectra. Although in principle the TD-DFT approach has inherent problems with extended π systems, the results reported in the following suggest that valence excitations of topotecan are well represented by a TD-DFT approach employing hybrid functionals, as already shown on related systems.[23]

3. Results

3.1. General Strategy. Topotecan can undergo several equilibria depending on the pH of the solution. In particular, the following four pK values have been experimentally determined: $pK_{a1} < 0.8$, $pK_{a2} \simeq 3.6$, $pK_{a3} = 6.5$, $pK_{a4} = 10.7$ (see ref 18 and references therein). At concentrations low enough to avoid dimer formation ($<10^{-5}$ M) a lactone form with charge +1 (Figure 1) dominates between $4.0 < pH < 6.0$, whereas around neutral pH two different forms should be present, namely, a neutral lactone form (Figure 2) and a carboxylate form with charge −1 (Figure 3). To verify the accuracy of the QM procedure in reproducing the electronic distribution of the topotecan, we have first carried out calculations on the moderately acidic pH region for both eno and keto lactonic forms (Figure 1a,b, respectively) in order to determine the most stable tautomer. After this validation, an analogous procedure has been applied to the study of the different equilibria present around neutral pH. An independent check of the computational results is provided by a comparison between UV−vis spectra obtained from TD-DFT computations and purposely recorded at different pH values.

3.2. Analysis of Topotecan Moderately Acidic pH Forms. At moderately acidic pH the topotecan has a charge +1 (see ref 18), and it exists as a lactone form represented by two tautomeric structures: an eno form characterized by an aromatic C ring and protonated on the oxygen of the D ring (Figure 1a); a keto form protonated on the C ring (Figure 1b). In vacuo geometry optimizations at different QM levels agree in indicating that only the eno species corresponds to a true energy minimum, whereas this is not the case for the corresponding keto tautomer (see Table 1).

(25) Amovilli, C.; Barone, V.; Cammi, R.; Cances, E.; Cossi, M.; Mennucci, B.; Pomelli, C. S.; Tomasi, J. *Adv. Quantum Chem.* **1999**, *32*, 227−237.
(26) Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Chem. Phys.* **2002**, *117*, 43−54.
(27) Adamo, C.; Barone, V. *Chem. Phys. Lett.* **2000**, *330*, 152−160.
(28) Barone, V.; de Biani, F. F.; Ruiz, E.; Sieklucka, B. *J. Am. Chem. Soc.* **2001**, *123*, 10742−10743.
(29) Frisch, M. J.; et al. *Gaussian 03 Rev. B4*; Gaussian, Inc.: Wallingford, CT, 2004.
(30) Binning, R. C.; Curtiss, L. A. *J. Comput. Chem.* **1990**, *11*, 1206−1216.
(31) Saebo, S.; Almlof, J. *Chem. Phys. Lett.* **1989**, *154*, 83−89.
(32) Thompson, M. A.; Zerner, M. C. *J. Am. Chem. Soc.* **1991**, *113*, 8210−8215.
(33) Cossi, M.; Barone, V. *J. Chem. Phys.* **2000**, *112*, 2427−2435.
(34) Cossi, M.; Barone, V. *J. Chem. Phys.* **2001**, *115*, 4708−4717.

Table 1. Absolute and Relative Energies (au and kJ mol^{-1}, respectively) of the Topotecan Type 1 Tautomer Structures[a]

method	type 1a (eno) [au]	type 1b (keto) [kJ mol^{-1}]
Geometry Optimization		
HF/3-21G	−1414.481611	not found
HF/3-21G+PCM	−1414.578513	−158.6
B3LYP/6-31G*+PCM	−1431.134223	−123.8
Single-Point Energy		
(B3LYP/6-311+G**//HF/3-21G)+PCM	−1431.521328	−116.5
(B3LYP/6-311+G**//B3LYP/6-31G*)+PCM	−1431.531470	−112.0
(MP2/6-311+G**//B3LYP/6-31G*)+PCM	−1427.587208	−85.9

[a] Data refer to the molecular geometries of Figure 1a (eno) and Figure 1b (keto) at different levels of computational methods.

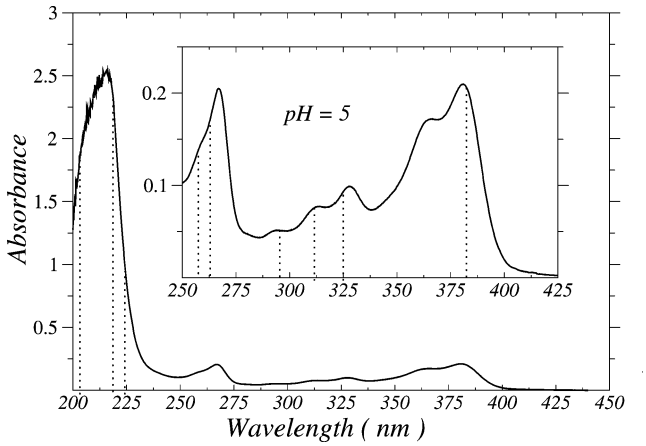


Figure 4. UV−vis experimental spectrum at pH 5.0. In the inset the 250−425 nm region is depicted. Dotted bars lines refer to the calculated TD-B3LYP/6-31G*+PCM absorption bands of the keto tautomer (type 1b). Concentration of the drug was 5.5 × 10^{-6} M.

Only inclusion of solvent effects through PCM permits finding an energy minimum for the keto tautomer that is even more stable than the eno one. In fact, at the PCM(HF/3-21G) level the keto tautomer is the absolute energy minimum, 158.6 kJ/mol more stable than the enolic form (Table 1). This stability trend is confirmed both by PCM(B3LYP/6-31G*) geometry optimizations (ΔE = 123.9 kJ/mol) and by single-point computations at these geometries with the more extended 6-311+G** basis set (ΔE = 112.0 and 85.9 kJ/mol at the B3LYP and MP2 level, respectively). Differently from what is observed for the eno tautomer, PCM calculations do not show any significant effect on the structural parameters of the keto tautomer. The optimized structures obtained at the HF/3-21G level with and without PCM show a root-mean-square deviation as low as 0.62 Å between the positions of the 31 heavy atoms. It is also noteworthy that use of HF/3-21G or B3LYP/6-31G* geometries leads to very similar results in single-point B3LYP/6-311+G** energy evaluations. This can be profitably used for geometry optimization of larger systems (vide infra).

To validate the energy results, the UV−vis absorption spectrum of the drug has been recorded in buffered solution at pH 5 (Figure 4). Next a set of UV−vis absorption spectra have been calculated at the TD-B3LYP/6-31G* (in vacuo and with PCM) and TD-B3LYP/6-31+G* (PCM) levels of computation, starting from the PCM(B3LYP/6-31G*)-optimized molecular structures for the type 1a and 1b topotecan forms. Since the dimensions of the system prevent us from reporting all the values of the absorption bands (more than 40), we show in Table 2

Table 2. Experimental Band Maxima and Calculated Vertical Transitions (in nm) of the Eno(a)/Keto(b) Stable Structures of Topotecan at pH 5 (Type 1)[a]

B3LYP/6-31G*		B3LYP/6-31G*+PCM		B3LYP/6-31+G*+PCM		
(a)	(b)	(a)	(b)	(a)	(b)	exptl
602*(0.12)*	482*(0.15)*	569*(0.15)*		573*(0.15)*		
			382*(0.44)*		379*(0.50)*	384$^{(s)}$
	362*(0.19)*					368$^{(sh)}$
		340*(0.81)*	325*(0.12)*	347*(0.71)*		330$^{(w)}$
317*(0.88)*			312*(0.11)*	314*(0.14)*	313*(0.14)*	316$^{(w)}$
290*(0.15)*		294*(0.44)*	296*(0.16)*	300*(0.35)*	297*(0.15)*	296$^{(w)}$
280*(0.15)*		273*(0.18)*	262*(0.22)*	277*(0.33)*	263*(0.40)*	270$^{(s)}$
260*(0.17)*		257*(0.30)*	256*(0.25)*	261*(0.21)*		
250*(0.17)*	247*(0.24)*	254*(0.47)*		259*(0.27)*		
246*(0.22)*	244*(0.36)*				1tht	
			225*(0.20)*	225*(0.15)*	230*(0.52)*	
			224*(0.28)*			
238*(0.23)*	225*(0.15)*	220*(0.16)*	218*(0.20)*			220$^{(vs)}$
		212*(0.21)*	202*(0.17)*	207*(0.10)*		

[a] Only the bands with oscillator strengths larger than 0.1 (in parentheses) are reported. Legend: (vs) very strong; (s) strong; (w) weak; (sh) shoulder.

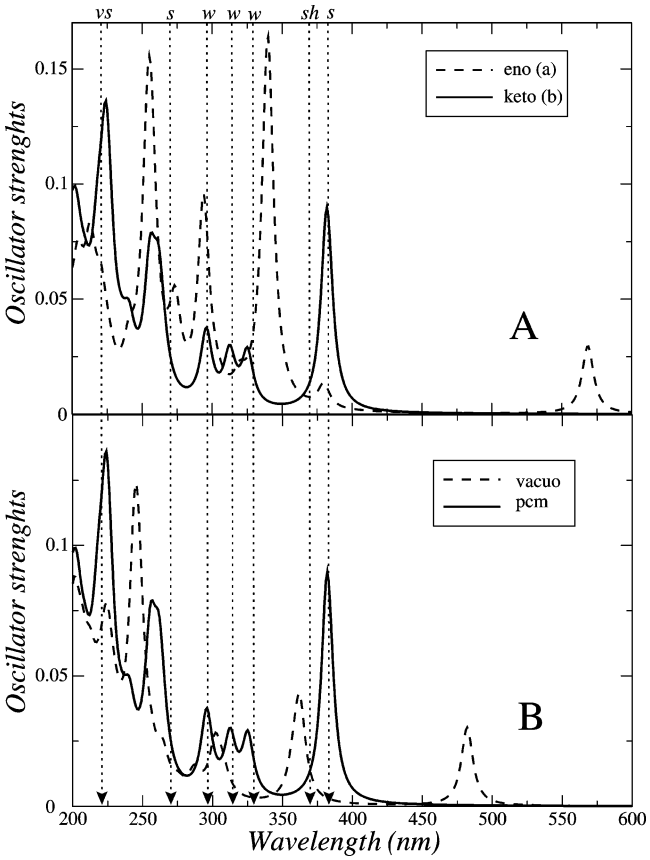


Figure 5. UV−vis theoretical spectra calculated at the TD-B3LYP/6-31G* level of computation. Panel A: eno (a) and keto (b) tautomers (PCM calculations), represented with dashed and solid lines, respectively. Panel B: in vacuo and PCM calculations for the keto (b) tautomer, represented with dashed and full lines, respectively. Dotted lines refer to the absorption band maxima of the experimental spectrum at pH 5 (Legend: (vs) very strong; (s) strong; (w) weak; (sh) shoulder).

only the transitions with oscillator strengths larger than 0.1, in comparison with the experimental band maxima. Moreover, in Figure 5 we report the spectra obtained by TD-B3LYP/6-31G* calculations in which each transition is represented by a Gaussian with half-bandwidth of 10 nm. Panel A of Figure 5 shows the results for the eno (a) and keto (b) tautomers with the inclusion of the solvent effect via PCM. The data clearly

Table 3. Absolute and Relative Energies (au and kJ mol^{-1}, respectively) of the Two Stable Keto Conformers of Type 2 Topotecan Structures at Different Levels of Computational Methods[a]

method	type 2a [kJ mol^{-1}] N protonated	type 2b [au] O protonated
Geometry Optimization		
HF/3-21G	not found	−1414.138162
HF/3-21G+PCM	19.4	−1414.175534
B3LYP/6-31G*+PCM	16.9	−1430.734125
HF/3-21G+PCM$^{(3wat)}$	−26.3	−1641.016082
Single-Point Energy		
(B3LYP/6-311+G**//HF/3-21G)+PCM	10.6	−1431.122878
(B3LYP/6-311+G**//B3LYP/6-31G*)+PCM	11.6	−1431.130965
(MP2/6-311+G**//B3LYP/6-31G*)+PCM	12.8	−1427.176589
(B3LYP/6-311+G**//HF/3-21G)+PCM$^{(3wat)}$	−15.7	−1660.531511

[a] The (3wat) symbol represents the calculations on the complexes in the presence of three water molecules.

demonstrate that the experimental position of the peaks (dotted lines in Figure 5) are well reproduced only by the type 1b keto tautomer, particularly in the region 280−600 nm, with the exception of the shoulder at 368 nm. The essential role of the PCM can be appreciated by comparing the theoretical absorption spectra for the keto tautomer either with or without the PCM. These data, reported in Figure 5B, show that the spectrum, calculated without solvent effect, does not reproduce at all the experimental results. Inclusion of diffuse functions in the calculation of the UV absorption spectra at the TD-B3LYP/6-31+G*+PCM level leads to some differences in the calculated bands in the 200−250 nm region, while the 250−600 nm region is unaffected (Table 2). For this reason in the following we will refer to TD-B3LYP/6-31+G*+PCM calculations, while reporting also the less computationally demanding TD-B3LYP/6-31G*+PCM results.

In general terms, our results indicate that proper QM calculations including solvent effects are able to identify the preferred molecular structure of topotecan in solution, which at pH 5 is the keto form (Figure 1, type 1b). TD-DFT computations performed at the same level reproduce the UV−vis experimental data, confirming that the PCM is able to appropriately model at the same time chemical structures and spectroscopic properties.

3.3. Analysis of Topotecan Neutral pH Forms. HF/3-21G geometry optimization of topotecan at neutral pH for both keto and eno forms confirms that the keto structures correspond to the lowest energy minima (data not shown for the eno forms). HF and DFT geometry optimizations of topotecan keto forms have then been carried out at neutral pH, either in vacuo or including the solvent effect through PCM, considering that topotecan at this pH can exist either as a lactone form with a closed E ring, hereafter called type 2 (Figure 2), or as a carboxylate form with an open E ring, hereafter called type 3 (Figure 3). In vacuo calculations of the lactone form, starting from the structure depicted in Figure 2a (type 2a), result in an intramolecular hydrogen migration from O4$^-$···HN6$^+$ (type a) to O4H···N6 (type b) moieties. Only by inclusion of solvent effects via PCM could we obtain two energy minima for both the type 2a and 2b structures. Molecular geometry optimization and single-point energy calculations, carried out at the same levels of computational complexity used for moderately acidic pH calculations, give rise to the energetic behavior reported in Table 3. In all cases the type 2b structure is the most stable one, but the quite small energy difference with structure 2a (11.6 kJ/mol with B3LYP and 12.8 kJ/mol with MP2, employing the 6-311+G** basis set) suggests that both forms of topotecan can coexist in aqueous solution at pH 6.8.

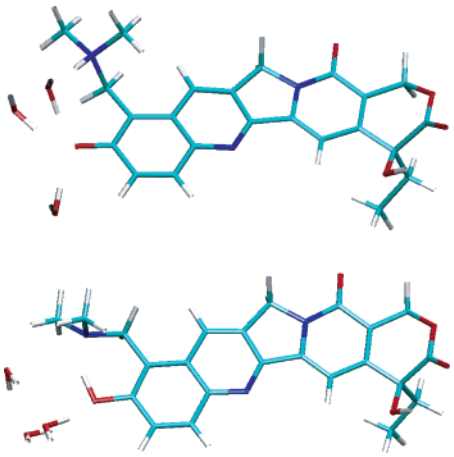


Figure 6. Topotecan type 2 neutral lactone isomers with three water molecule complexes. The 3D views of type 2a and 2b lactone forms are reported in the upper and bottom panels, respectively.

Once again single-point energy evaluations by a more extended basis set at HF/3-21G and B3LYP/6-31G* geometries give very close results. This finding allows the use of the cheaper HF/3-21G approach for investigating the role of specific water molecules. As a matter of fact, a PCM(HF/3-21G) geometry optimization, carried out including three explicit water molecules located in the proximity of the two protonation sites (Figure 6), reverts the energy stability, indicating the type 2a form as the most stable. This result is further confirmed by PCM(B3LYP/6-311+G**//HF/3-21G) single-point energy calculation (Table 3, bottom part). In both cases the energy difference is again just a few kJ/mol, thus suggesting that the microenvironment plays a major role in selecting the most stable form. This result must be taken into consideration when the drug is inserted in its natural binding site, represented by the Top1−DNA binary complex.

In vacuo calculations of the carboxylate form, starting from the type 3a structure depicted in Figure 3, show again a spontaneous transformation of the O4$^-$···HN6$^+$ moiety into its O4H···N6 counterpart (type 3b structure). Once more, only the PCM inclusion in the calculations has permitted us to obtain an energy minimum also for the type 3a (N protonated) structure. The most stable structure in solution, at any level of calculation, is type 3b, similarly to what is observed for the lactone form. In this case, the energy differences between the type 3a and 3b structures are relatively large, that is, about 30 kJ/mol at our best computation level (B3LYP or MP2/6-311+G**//B3LYP/6-31G* +PCM; see Table 4). Explicit inclusion of three water molecules located in the proximity of the two protonation sites again reverts the relative stability; that is, the type 3a structure becomes the most stable one. However, once again the energy difference between the type 3a/b structures is quite small, being in the range 8−18 kJ/mol (see Table 4). The microenvironment (polarity, hydrophobicity, ionic strength, dielectric constant) will then play a key role in selecting the preferred structure at pH 6.8.

The experimental topotecan UV−vis spectra in the 250−450 nm range recorded at pH 5 and pH 6.8 are compared in Figure 7. Absorption spectra have been computed for all four forms

Table 4. Absolute and Relative Energies (au and kJ mol^{-1}, respectively) of the Two Stable Keto Conformers of Type 3 Topotecan Structures at Different Levels of Computational Methods[a]

method	type 3a [kJ mol^{-1}] N protonated	type 3b [au] O protonated
Geometry Optimization		
HF/3-21G	not found	−1489.190203
HF/3-21G + PCM	57.3	−1489.309107
B3LYP/6-31G*+PCM	59.8	−1506.694279
HF/3-21G+PCM(3wat)	−18.4	−1716.151001
Single-Point Energy		
(B3LYP/6-311+G**//HF/3-21G)+PCM	27.8	−1507.133155
(B3LYP/6-311+G**//B3LYP/6-31G*)+PCM	31.3	−1507.143369
(MP2/6-311+G**//B3LYP/6-31G*)+PCM	29.4	−1503.008713
(B3LYP/6-311+G**//HF/3-21G)+PCM(3wat)	−8.5	−1736.543807

[a] The (3wat) symbol represents the calculations on the complexes in the presence of three water molecules.

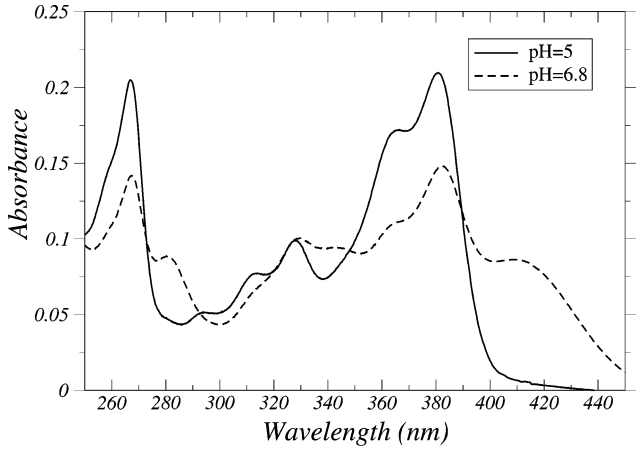


Figure 7. Combined UV−vis experimental spectrum recorded at two pH values in the range 250−450 nm. Full line: pH 5. Dashed line: pH 6.8. Concentration of the drug was 5.5×10^{-6} M.

depicted in Figures 2 and 3 starting from the PCM(B3LYP/6-31G*)-optimized structures and carrying out TD calculations at the B3LYP/6-31G*+PCM and B3LYP/6-31+G*+PCM levels. The results reported in Table 5 indicate that it is not possible to identify a unique form that best fits the experimental UV−vis spectrum, suggesting that at this pH several topotecan structures are in equilibrium. However, comparison of experimental and calculated UV−vis band positions permits us to identify bands that can be confidently attributed to a specific form.

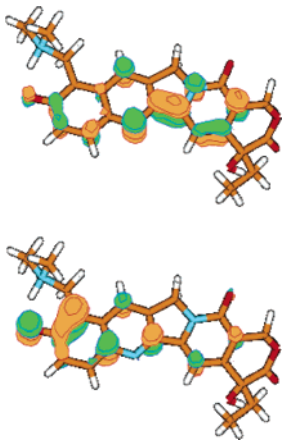


Figure 8. HOMO (bottom) and LUMO (top) of the type 2a neutral lactone form, calculated at the B3LYP/6-31G*+PCM level.

In fact, the experimental band at 416 nm can be reproduced only by the type a forms (Table 5), indicating that this band can be taken as a diagnostic for a deprotonated O4 atom. Moreover the experimental band at 384 nm can be reproduced only by the type b forms and can then be taken as a diagnostic for a protonated O4 atom. As a matter of fact, this band is the most intense one at pH 5, when O4 is surely protonated (Figure 7). Moreover, at pH 9.5 (spectrum not shown) the band at 420 nm is the most intense one, while the band at 384 nm completely disappears, confirming our interpretation.

TD-DFT computations indicate that both bands are dominated by HOMO−LUMO electronic transitions, and in all cases a significant contribution to frontier orbitals is provided by the C3O4 moiety, which is involved in the intramolecular H-bond with N6 (see Figures 8 and 9). This explains why passage from $O4^{-}\cdots HN6^{+}$ (type a) to O4H···N6) (type b) forms displaces the HOMO−LUMO band by as much as 40 nm (Table 5). As already seen for type 1, the TD/DFT calculations do not reproduce the experimental shoulder at 364 nm, even though both the type 2a and 3a forms show the presence of an

Table 5. Experimental Band Maxima and Calculated Vertical Transitions (in nm) of the Most Stable Structures of Topotecan at pH 6.8 (types 2/3)[a]

structure type 2				structure type 3				
(a)		(b)		(a)		(b)		
M1	M2	M1	M2	M1	M2	M1	M2	exptl
426(*0.54*)	426(*0.56*)			421(*0.51*)	424(*0.55*)			416(w)
		381(*0.55*)	382(*0.56*)			388(*0.50*)	387(*0.55*)	384(s)
357(*0.05*)	358(*0.05*)			355(*0.02*)	356(*0.05*)			364(sh)
340(*0.28*)	341(*0.29*)			347(*0.30*)	346(*0.31*)			345(w)
		320(*0.13*)	322(*0.16*)			323(*0.17*)	325(*0.17*)	330(w)
287(*0.24*)	292(*0.27*)	297(*0.14*)	298(*0.11*)	297(*0.11*)			301(*0.11*)	282(w)
			271(*0.26*)		278(*0.32*)	274(*0.21*)	262(*0.36*)	268(s)
251(*0.17*)	255(*0.11*)					255(*0.14*)		
	253(*0.12*)				251(*0.18*)	255(*0.11*)		
244(*0.30*)	249(*0.17*)		246(*0.24*)					
224(*0.15*)	229(*0.13*)	229(*0.20*)	236(*0.15*)	235(*0.23*)	232(*0.14*)	227(*0.55*)	232(*0.48*)	220(vs)
219(*0.12*)	224(*0.14*)		226(*0.24*)		225(*0.16*)	225(*0.10*)		
217(*0.16*)	223(*0.32*)							
217(*0.18*)			214(*0.23*)		217(*0.19*)			

[a] Calculations are carried out at the TD-B3LYP/6-31G*+PCM (M1) and TD-B3LYP/6-31+G*+PCM (M2) levels of computation. Only the bands with oscillator strengths greater than 0.1 (in parentheses) are reported, with the exception of the bands close to 364 nm. Legend: (vs) very strong; (s) strong; (w) weak; (sh) shoulder.

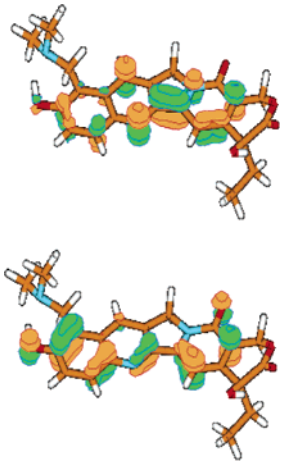


Figure 9. HOMO (bottom) and LUMO (top) of the type 2b neutral lactone form calculated at the B3LYP/6-31G*+PCM level.

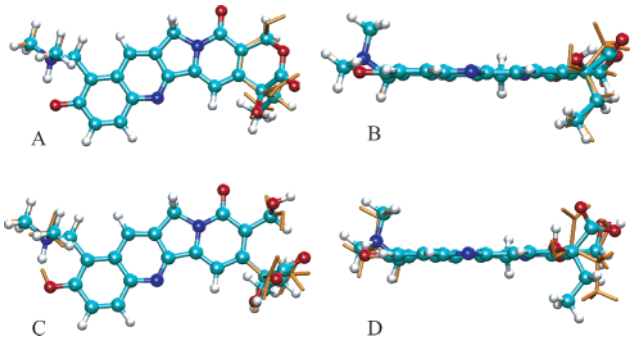


Figure 10. Topotecan type 2a−3a−3b molecular structure comparisons. Upper A/B panels: 3D top/in-plane view of type 2a−3a lactone and carboxylate forms superimposed (type 3a in orange stick). Bottom C/D panels: 3D top/in-plane view of type 3a−3b carboxylate forms superimposed (type 3b in orange stick)

absorption band with low oscillator strength (see Table 5). Furthermore, there is no evidence of any specific feature discriminating between the type 2 and type 3 forms. In conclusion, the calculated UV−vis electronic transitions (Table 5) and the experimental absorption bands (Figure 7) recorded at physiological pH strongly suggest that the spectrum is composed by contributions of all four type 2/3 drug isomers in chemical equilibrium at this pH.

While a comprehensive analysis of the equilibrium between different topotecan forms is outside the scope of the present paper, it is worth investigating the role of the solvent in tuning the reaction *topotecan type 2b* + OH^- → *type 3b*. Calculations, performed at the B3LYP/6-311+G** level starting from the B3LYP/6-31G* geometries indicate that the exothermicity of the reaction is reduced from 78.8 kJ/mol in vacuo to 29.9 kJ/mol in aqueous solution. This result confirms that the biological environment plays a fundamental role in selecting the preferred drug structure to interact with.

3.4. Conformational and Electrostatic Analysis of Neutral pH Topotecan Forms. The study of the interaction of topotecan with DNA or with the DNA−Top1 binary complex requires an accurate description of the conformational and electrostatic properties of the drug.

The conformational differences between structures type 2a and 3a can be appreciated overlapping the two molecules on their A−D rings (Figure 10 panels A, B). The figure shows that the largest geometrical deformations involve atoms belonging to the E ring for structure 2a and to the corresponding open carboxyl/alcoholic moieties in structure 3a. As a matter of fact, the presence or absence of a ring in this region does not affect at all the structure of the remaining part of the molecule. In particular, the atoms not belonging to the rings (in Figure 1a labeled as O4, C5, N6, C7, and C8) retain the same orientation in the 2a and 3a configurations, as can be better appreciated from the parameters reported in Table 6. The data shown in the table indicate that bond lengths and dihedral angles are very close for the 2a and 3a configurations, as they are for the 2b and 3b configurations.

Table 6. Leading Structural Parameters of the Topotecan Keto Conformers at the B3LYP/6-31G* Level of Computation[a]

	distance (Å)		dihedral angle (deg)			
type	N_6H	O_4H	$C_1C_2C_5N_6$	$C_2C_5N_6C_8$	$C_2C_5N_6C_7$	$HO_4C_3C_2$
2a	1.09	1.54	−141.5	−160.5	71.7	−20.2
2b	1.69	1.01	−146.1	−163.1	72.1	−13.2
3a	1.09	1.56	−141.6	−160.6	71.8	−20.0
3b	1.70	1.01	−146.1	−163.1	72.1	−13.2

[a] Angles in degrees, internuclear distances in Å. Atom numbering as in Figure 1a.

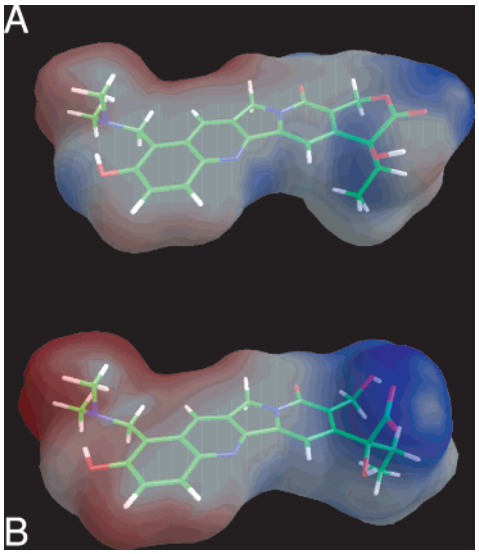


Figure 11. Molecular structure (sticks) and electrostatic potential surface (EPS) of type 2b (panel A) and 3b (panel B).

Therefore, geometrical differences in this part of the molecule occur only because of the hydrogen migration from O4 (b form) to N6 (a form), while the opening of the E ring does not produce any significant geometric perturbation. On the other hand, migration of the hydrogen from O4 to N6 induces structural changes around the carboxylate moiety, as can be appreciated from Figure 10 (panels C, D), where structures type 3a and 3b are overlapped on their A−D rings. Such differences do not occur in the 2a and 2b forms, indicating that the type 3a/3b forms are characterized by a higher degree of flexibility, when compared to the type 2a/2b forms.

The QM calculations also permit us to evaluate the electrostatic distribution of the molecule in the different configurations. The electrostatic potential surface (EPS) is quite different between the four topotecan forms, present at neutral pH. As an example, the type 2b and 3b EPS, represented in Figure 11, indicates that the potential distribution is dependent on the drug structure. The structural and electronic topotecan properties

obtained in this work then give crucial information for the investigation of the drug interactions with its biological target.

4. Conclusions

The present paper is devoted to the identification of the factors playing a role in determining the overall structural and spectroscopic properties of the topotecan anticancer drug in aqueous solution at acid and neutral pH.

DFT/PCM computations provide a coherent picture of the topotecan behavior, confirmed by the close agreement between computed and experimental UV−vis spectra at moderately acidic pH.

TD-DFT/PCM results are also in remarkable agreement with the UV spectrum recorded around neutral pH when assuming the contemporary presence of different tautomeric forms, characterized by different and well-defined spectroscopic signatures. This is confirmed by the very close stability of these forms issuing from a refined discrete/continuum solvent model.

Our results provide for the first time an accurate structural and electronic description of topotecan to be used in the selection of the preferred form bound to the topotecan−DNA−Top1 ternary complex and in the identification of the major protein−DNA anchor sites.

Acknowledgment. The authors wish to thank the CASPUR consortium for the computing facilities, the Italian Minister of Health, and Italian Minister of University and Research (MIUR) for financial support.

Supporting Information Available: Complete Ref 29. This material is available free of charge via the Internet at http://pubs.acs.org.

JA052637U