

Molecular docking approach on the Topoisomerase I inhibitors series included in the NCI anti-cancer agents mechanism database

Antonino Lauria · Mario Ippolito ·
Anna Maria Almerico

Received: 12 September 2006 / Accepted: 27 September 2006 / Published online: 28 October 2006
© Springer-Verlag 2006

Abstract Topoisomerase I (Top1) is an essential enzyme participating to all those processes associated with separation of DNA strands. It manages superhelical tensions through the transient breakage of one strand of duplex DNA, followed by the unwinding of supercoiled DNA. Camptothecins, a class of alkaloids extracted from the wood of a Chinese tree, were found to be potent inhibitors of Topoisomerase I. The National Cancer Institute (NCI) Anti-cancer Agents Mechanism Database contains several camptothecin derivatives, classified as selective Top1 inhibitors. In this work we performed molecular docking studies on 24 camptothecin-like inhibitors present in this database (using Autodock 3.0.5). In order to consider the different orientations of the active site residues, docking was performed using four different structures of a Top1-DNA complex. The results obtained allowed us to analyze some conformations adopted by the inhibitors during active site binding, confirming the role of hydrogen bond and contributed to clarify the loss of activity due to single point mutations.

Keywords Molecular docking · Topoisomerase I inhibitors · Autodock · DNA

Introduction

DNA Topoisomerases are enzymes common to all living organism. They participate to all metabolic processes

associated with separation of DNA strands, such as transcription, repair and replication. Topoisomerases are classified as type I and II on the basis of two different sequences and functions. The types II (Top2) are dimeric enzymes which break both strands of duplex DNA in an ATP-dependent manner, whereas the type I (Top1) enzymes are monomeric entities and require no additional energy cofactor to manage DNA superhelical tensions through the cleavage of one strand of duplex DNA. This reaction is carried out via a nucleophilic attack of an active site tyrosine OH group (Tyr723) which becomes covalently bound to the 3'-phosphate of DNA at the site of nicking and forms the “*cleavage complex*” (Fig. 1). This transient break allows helical supertension to be released by the unwinding of positively supercoiled DNA or the rewinding of the negatively supercoiled DNA [1, 2].

Experiments of protease digestion [3, 4] revealed that human Top1 is comprised of four major domains (Fig. 2). The NH₂-terminal domain (24 kDa) is comprised between Met1 and Lys197, and seems dispensable for in vitro activity. Residues Glu198 to Ile651 form the highly conserved “core domain” (54 kDa), followed by a short unconserved 3 kDa linker (Asp 652 to Glu696). This linker has been found to be highly positive charged and may bind directly to DNA, thus contributing to the processivity of the whole enzyme. It has also been proven that this region is essential for enhancing the sensitivity of some inhibitors [5]. Finally the C-terminal domain, comprised between Gln697 and Phe795, is highly conserved and contains the active site Tyr723.

Camptothecin (CPT, Fig. 3), the major alkaloid extracted from the stem wood of the Chinese tree *Camptotheca acuminata*, was originally discovered and isolated 40 years ago [6]. Since its discovery, it was demonstrated that camptothecin and its analogs cause cell death during the

A. Lauria (✉) · M. Ippolito · A. M. Almerico
Dipartimento Farmacochimico, Tossicologico e Biologico,
Università di Palermo,
Via Archirafi 32,
90123 Palermo, Italy
e-mail: lauria@unipa.it

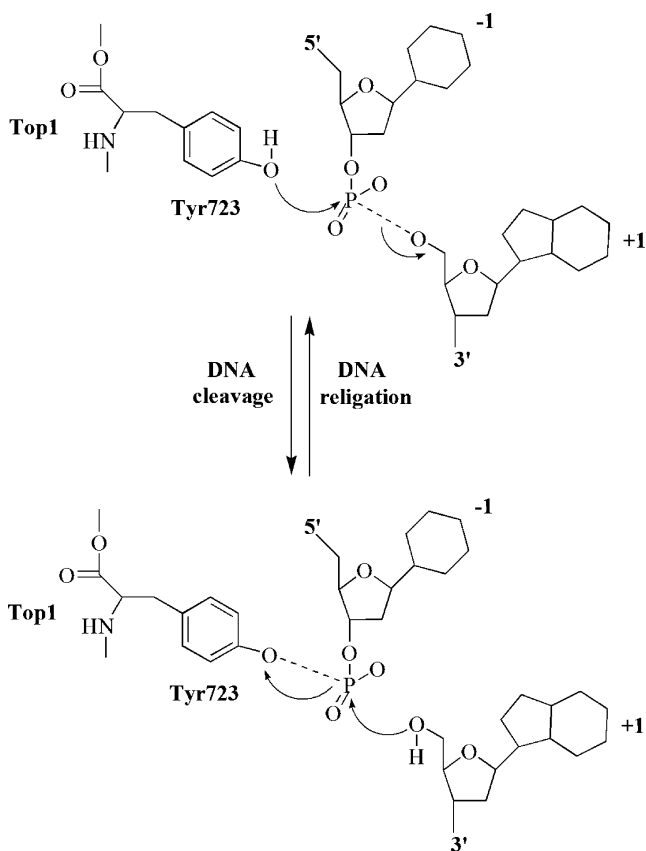


Fig. 1 DNA strand breaking by Top1

S-phase of the cell cycle. This cytotoxicity is thought to be caused by the cessation of DNA synthesis followed by a double-strand breakage when the replication fork encounters the camptothecin-trapped Top1-DNA complex [7] (termed as “*ternary complex*”). The exact mechanism by which CPT stabilizes the ternary complex is not fully cleared. To date there are many models proposed by several research groups [8], involving both intercalation and interaction with the DNA scissile strand and groove binding. Each one of these models is supported by the observation of covalent and H-bonding between the CPT, DNA bases and catalytic residues, such as Asn722, Lys532, Asp533, Arg364, Asn352, and Tyr723. Many structure-activity relationships have also been observed [2]. In general, camptothecins are planar molecules with an arc shape; substitutions on A and B ring do not give any perturbation of activity; positions 7, 9, 10 and sometimes 11 could be altered without loss of activity. Modification of the C and D rings could diminish the activity, probably causing

Fig. 2 Domain organization of human Topoisomerase I

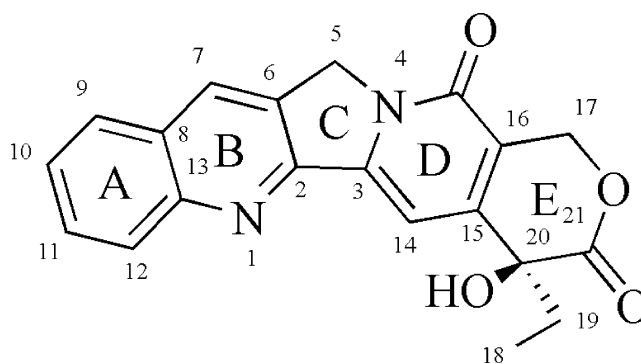
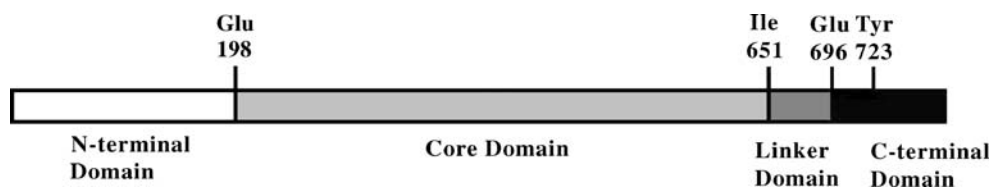


Fig. 3 Structure of Camptothecin

the disruption of planarity. The 20-OH is required for activity enhancement, since it might participate in H-bonding interactions with other moieties [9].

Molecular modelling studies have been facilitated by the recent availability of several X-ray structures of Top1, crystallized both with CPT [10] and with other inhibitors [11].

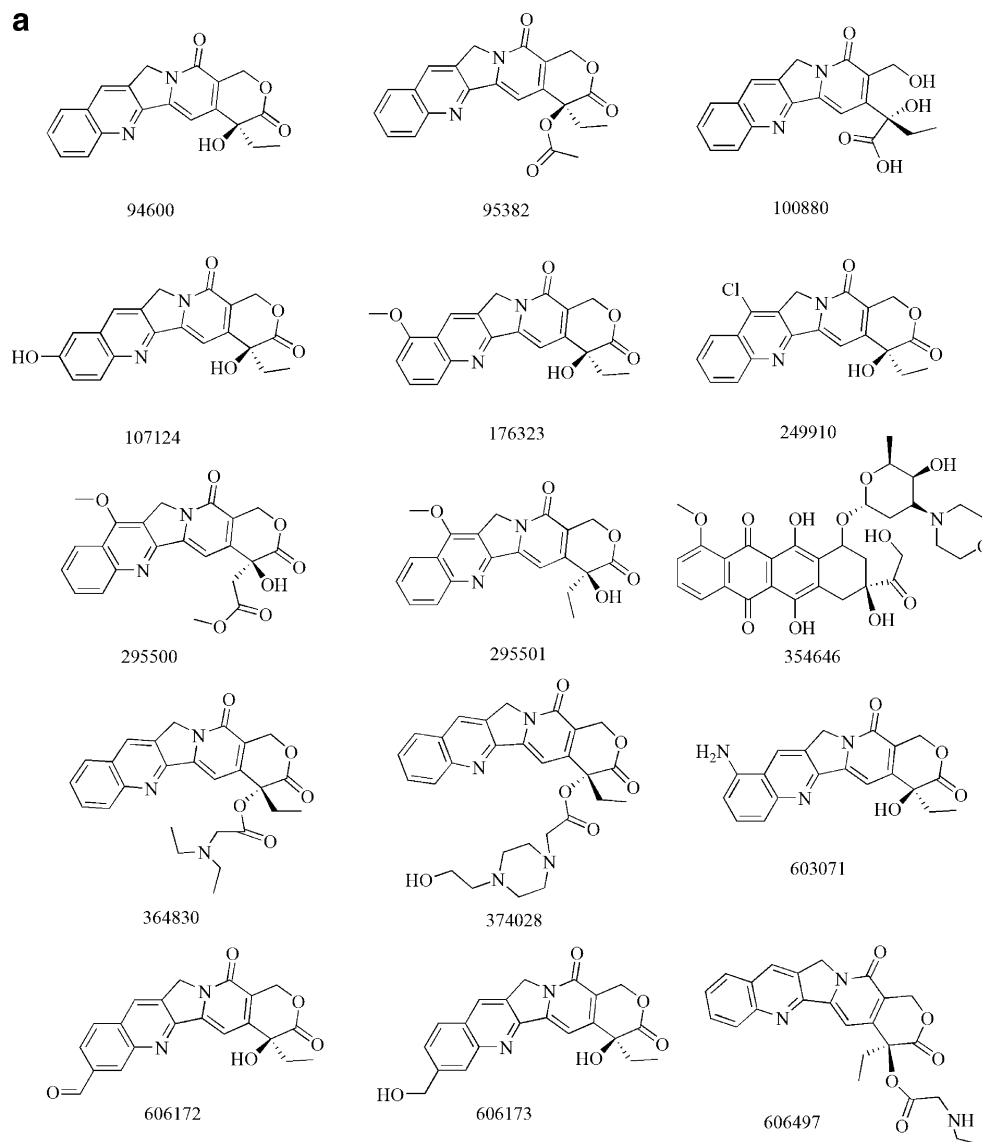
Several camptothecin derivatives are present in the NCI databases, classified as Topoisomerase I inhibitors agents into the Anti-cancer Agent Mechanism Databases. This is a set of 122 compounds with anti-cancer activity and a reasonably well-known mechanism of action, for which are available drug screening data that measure their ability to inhibit growth of a panel of 60 human tumor lines [12].

In the present work we performed molecular docking studies on a series of 24 Topoisomerase I inhibitors present in the NCI Anti-cancer Agents Mechanism Database, using four different Top1 structures deriving from four different Top1-DNA-inhibitor complexes, with the aim of investigate the binding modes of these derivatives.

Materials and methods

The structures of 24 Top1 inhibitors (23 camptothecin derivatives and one non-camptothecin compound) (Fig. 4) present in the Anti-cancer Agents Mechanism Database [http://www.dtp.nci.nih.gov/docs/cancer/searches/standard_mechanism.html] were generated using XDrawChem 1.9.8 and energy minimized with MOPAC (PM3 Method [13], RMS gradient 0.0100). Docking was performed with Autodock 3.0.5 [14, 15], a program combining a rapid energy evaluation through precalculated grids of affinity potentials with a variety of search algorithms to find

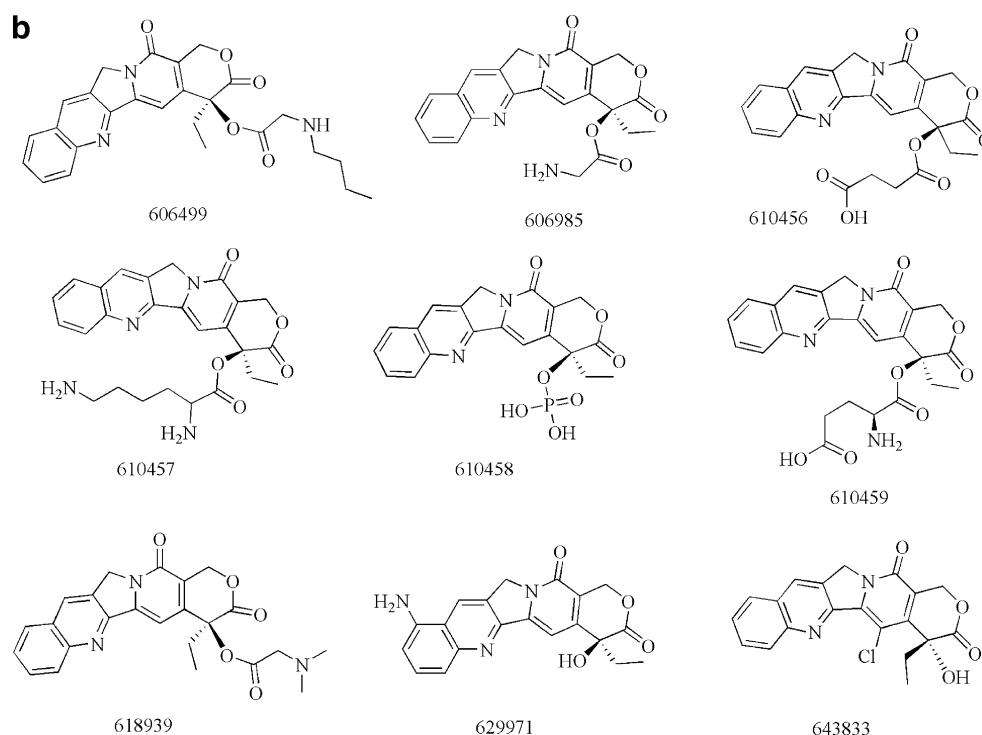
Fig. 4 a, b NCI Topoisomerase I inhibitors involved in this study



suitable binding position for a ligand on a given macromolecule. This program keeps the macromolecule rigid, while allows torsional flexibility for the ligand. Docking to macromolecule was carried out using an empirical free-energy function and Lamarckian Genetic Algorithm, with an initial population of 50 randomly placed individuals, a maximum number of 10^6 energy evaluations, a mutation rate of 0.02, a crossover rate of 0.80, and an elitism value of 1. For the local search, the so-called pseudo-Solis and Wets algorithm was applied using a maximum of 300 iterations per local search. The probability of performing local search on an individual in the population was 0.06 and the maximum number of consecutive failures before doubling or halving the local step size was 4. Fifty independent docking runs were carried out for each ligand. Results differing by less than 1.0 Å in positional root-mean-square deviation (rmsd) were clustered together and represented by the result with the most favorable free energy of binding.

The docking procedure was carried out using four different X-ray structures of a Top1-DNA complex [11] crystallized with four different inhibitors (Fig. 5a,b: Camptothecin 94600, 1, orange; Topotecan, 2, red; Indolocarbazole SA315F, 3, blue; Indenoisoquinoline MJ238, 4, green), selected from the Brookhaven Protein Data Bank (PDB codes: 1T8I, 1K4T, 1SEU, 1SC7). In these structures Top1 is bound to the oligonucleotide sequence 5'-AAAAA GACTTsX-GAAAAATTTT-3', where "s" is the 5'-bridging phosphorothiolate of the cleaved strand and "X" represents any of the four bases A, G, C or T. Before performing docking experiment the structure of the original ligand was removed from each complex, and Gasteiger–Marsili charges [16] were assigned. Solvation parameters were added to the final macromolecule structure using Addsol utility of Autodock. The grid maps representing the protein in the actual docking process were calculated with Autogrid. The grids (one for each atom type in the ligand,

Fig. 4 (continued)



plus one for electrostatic interactions) were chosen to be sufficiently large to include the entire width of the DNA fragment in which the original inhibitor was posed, a portion of minor and major grooves, and the active site-residues of Top1. The dimensions of the grids were thus $60 \times 50 \times 30$, with a spacing of 0.375 \AA between the grid points and the center close to the cavity left by the ligand after its removal.

Results and discussion

The accurate prediction of protein–ligand interaction geometries is essential for the success of virtual screening approaches in drug design. This process requires docking tools that should be able to generate suitable conformations of a ligand within a protein binding site, and reliable energetic evaluation, which can easily indicate the quality of a hypothetical interaction. Since Top1 active site can bind different classes of inhibitors, it can be useful to consider the differences and the analogies between the available structures of Top1–DNA–drug complexes. Therefore we examined the active sites deriving from four X-ray structures of Topoisomerase I crystallized with different inhibitors.

A comparison of the four Top1 structures through a backbone superimposition (Fig. 6) revealed that all macromolecules are quite similar except for the orientation of some active site residues, such as Arg364, Lys532, Asp533,

Asn722, and Tyr723. The presence of different orientations in major groove-directed residues is also supported by previous Molecular Dynamics simulations [17]. These conformational changes could be correlated to the different characteristics (steric, electrostatic, constitutional) of the various inhibitors (Fig. 5a,b).

A comparative analysis of the four ligands originally crystallized within the four Top1–DNA complexes used in this study revealed some common features. An obvious evidence is the presence of a planar heterocyclic ring system, confirming an intercalative-like mechanism of action (Fig. 7). A hydrogen-bond accepting moiety on the minor groove side of the inhibitors is also a conserved feature; the interaction between this H-bond acceptor and Arg364 is observed in all the ternary complexes.

In order to measure the steric hindrance for the four crystallized ligand, calculations of Van der Waals surfaces were carried out using a 1.4 \AA probe (Fig. 8). The values obtained, in the range $516 \div 641 \text{ \AA}^2$, gave a further evidence of the fact that the active site is quite flexible, in order to vary its size to accommodate diverse structures.

These size variations are quite difficult to consider during a docking experiment. Therefore we used all the four Top1–DNA structures, in order to keep account of active-site flexibility even if the docking procedure considers the macromolecule as a rigid entity while the ligand is flexible. The results deriving from the docking of the 24 NCI inhibitors with each of the four macromolecules are summarized in Table 1. All values are expressed in terms of

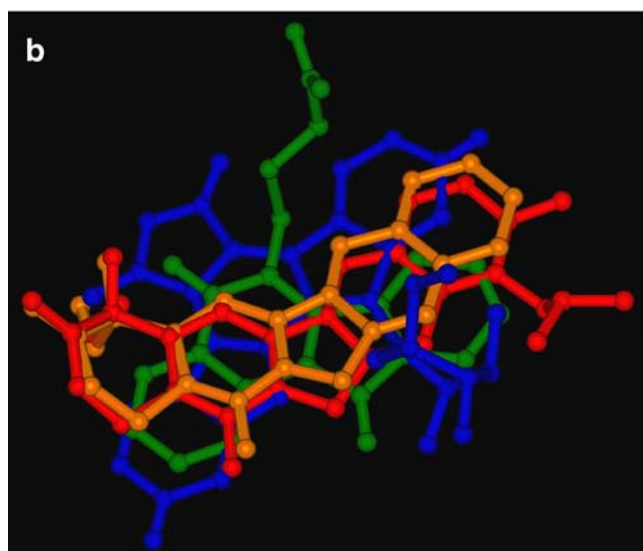
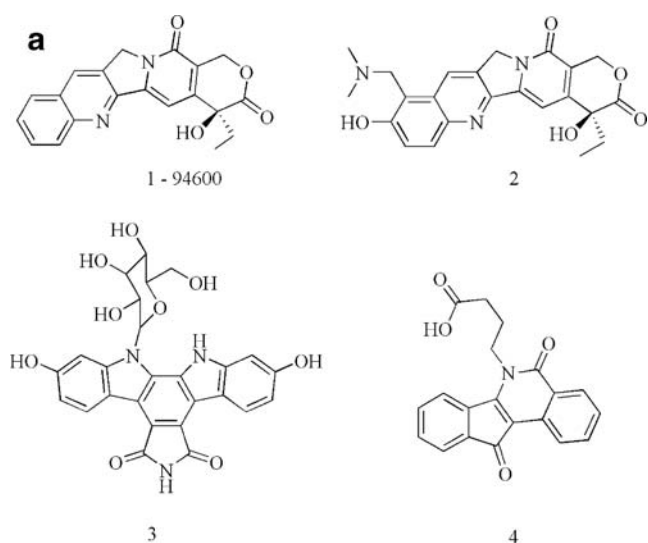


Fig. 5 **a, b** Structures (**a**) and superimposition (**b**) of Top1 inhibitors originally crystallized with K4T (Camptothecin 94600, **1**, orange), 1T8I (Topotecan, **2**, red), 1SEU (SA315F, **3**, blue), 1SC7 (MJ238, **4**, green)

ΔG . A mean of all the docking energies was also calculated, in order to get a trend of the goodness of the results.

As shown in Table 1 the range for the mean of ΔG values goes from -25.31 to -33.63 kcal/mol. From these data it emerges that compounds 354646 (a non-camptothecin derivative with $\Delta G = -33.63$ kcal/mol) and 610458 (a camptothecin derivative providing $\Delta G = -32.70$ kcal/mol) showed the highest mean binding affinity. Compounds 364830 and 618939 (both camptothecin derivatives, with $\Delta G = -26.82$ and $\Delta G = -25.31$ kcal/mol, respectively) had the lowest mean energy values.

An analysis of the best conformations obtained with docking reveals some common features in the general

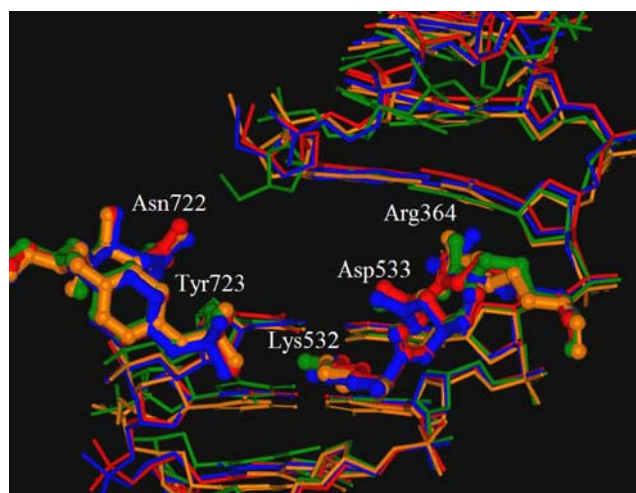


Fig. 6 Superimposition of active site from the four Top1-DNA complexes (RMS between 0.63 and 0.87 Å)

orientation of the derivatives into the active site. All conformations generally follow the “Intercalation Model” proposed by Pommier and co-workers [18], which predicts that the A-ring of the drug points in toward the backbone of DNA non-scissile strand, with the E-ring pointed in direction of Asn722 probably interacting with this residue. This latter observation is important, since it can offer an explanation for the Asn722Ser mutation associated to camptothecin resistance [19].

In order to illustrate these results, the docked conformations of the ligands providing lowest ΔG values are discussed herein. The binding mode for compound 100880, the camptothecin Na salt, is illustrated in Fig. 9a. This derivative is the open-lactone form of camptothecin,

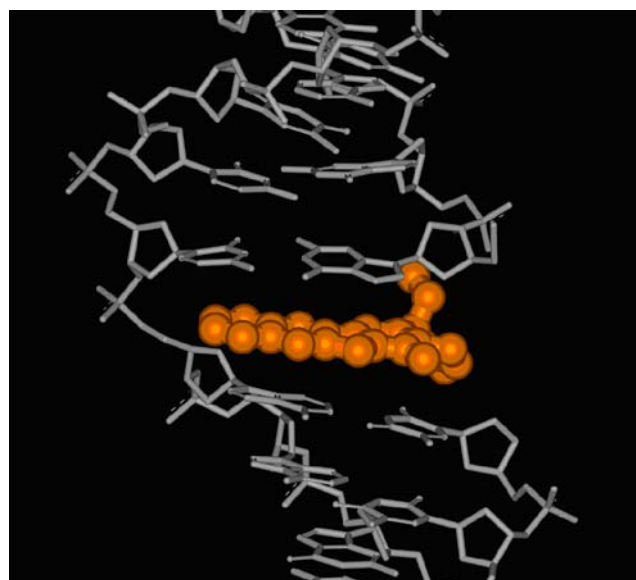
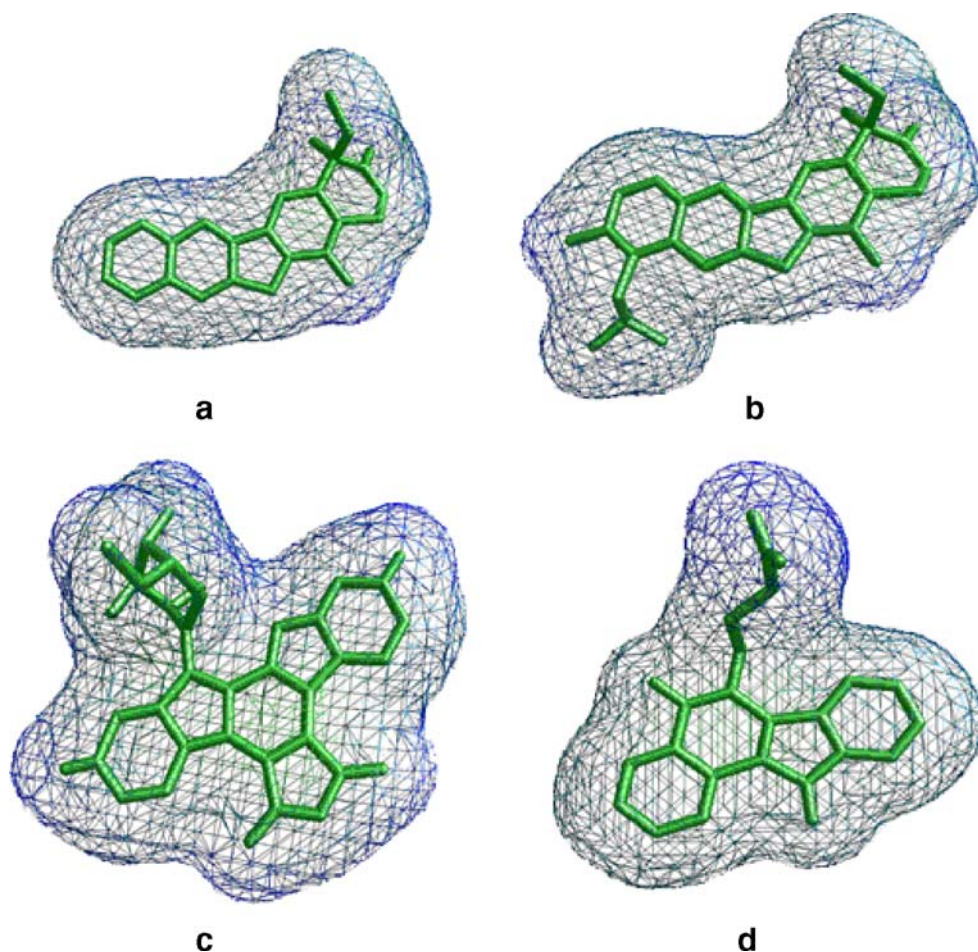


Fig. 7 Camptothecin intercalated into DNA (from 1T8I)

Fig. 8 Van der Waals surfaces: **a** Camptothecin (516 Å²); **b** Topotecan (612 Å²); **c** SA315F (641 Å²); **d** MJ238 (526 Å²)



which is thought to be the real active drug in vitro [20]. The docking results for this ligand are all satisfactory, with values in the range -30.19 – -34.55 kcal/mol.

Compound 295500, a camptothecin derivative with a OCH_3 substituent in position 7 and a $\text{CH}_2\text{COOCH}_3$ in position 20 provided very good binding interactions with 1K4T and 1T8I (ΔG values = -33.25 and -33.90 kcal/mol, respectively), although values obtained in the case of 1SEU and 1SC7 are less satisfactory.

Compound 374028 (Fig. 9b) is another camptothecin derivative, with a bulky substituent on C_{20} . The ΔG mean value obtained for this structure is -30.57 kcal/mol, with the highest value of -32.37 kcal/mol on 1SEU and the lowest of -29.07 kcal/mol on 1K4T. The best conformation obtained generally follows the binding pattern described previously, with the planar moiety intercalated between the base pairs and the E-ring directed toward Tyr723. In this case the substituent on C_{20} is docked in the region comprised between Asp533 (3.68 Å) and Tyr723 (3.17 Å). The ring skeleton is intercalated between the base pairs, thus allowing π -interactions. However, a longer distance (4.08 Å) was measured between N_1 and Arg364, suggesting the absence of interactions. The good score provided by this

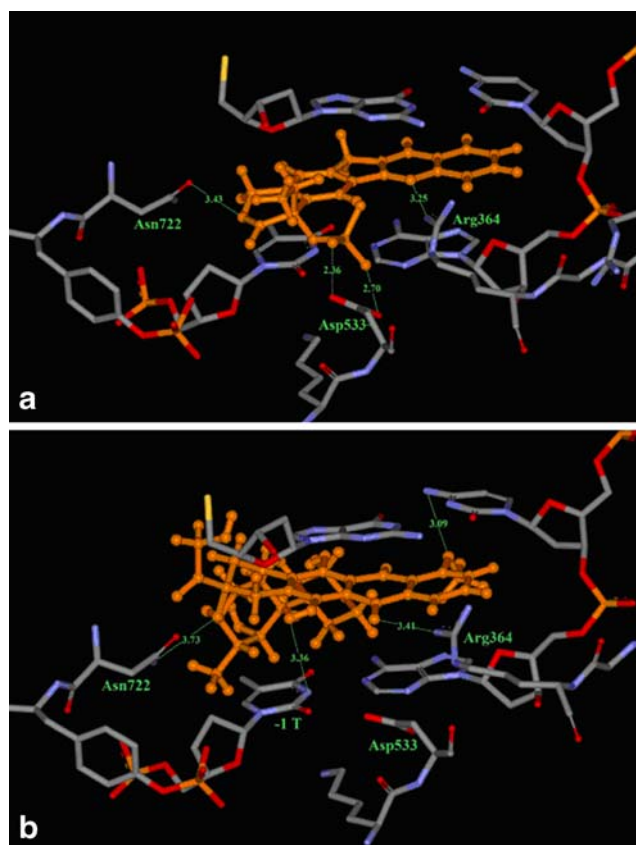
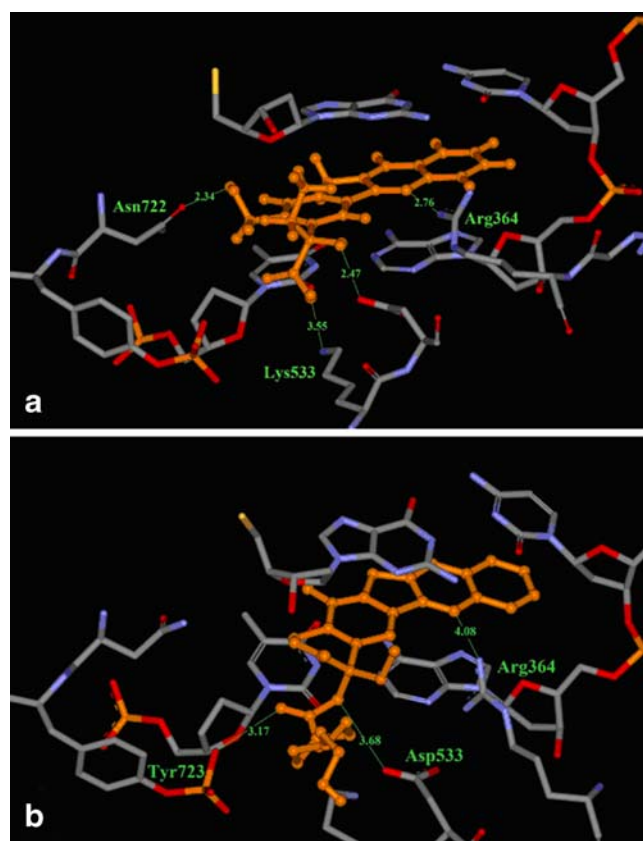
conformation is probably given by the abundance of H-bond accepting groups in the C_{20} substituent. It also seems to permit some degree of flexibility and it could be accommodated in the minor groove region.

Another derivative providing good binding interactions is 606172. The mean energy value obtained for this compound is -30.64 kcal/mol, with low values on 1K4T and 1T8I (-32.66 and -31.97 kcal/mol, respectively). It is interesting to consider that a poorly substituted compound such as this one gives better results on smaller active sites. In contrast, a derivative with a bulky substituent, such as 374028, provides better ΔG values on larger active sites, even if the mean energy values are quite similar in the two cases (-30.57 against -30.64 kcal/mol).

An interesting binding mode is given by compound 610458 (Fig. 10a). This derivative, consisting in a 20-phosphate substituted camptothecin, has been docked into the binding site with very good results ($\Delta G_{\text{mean}} = -32.70$ kcal/mol, $\Delta G_{\text{best}} = -35.15$ kcal/mol on 1K4T). These results are probably due to the presence of the phosphate group, the OH moieties of which interact with carbonyl terminal of Asp533 (2.36 and 2.70 Å, respectively). Other possible interactions can occur between the *O*-lactone

Table 1 Results of dockings using the four Top1-DNA structures. All energies are given in kcal mol⁻¹

NCI ID	1K4T	1T8I	1SEU	1SC7	Mean
94600	-30.46	-31.01	-27.15	-28.31	-29.23
95382	-30.20	-30.87	-25.07	-25.14	-27.82
100880	-33.44	-34.55	-30.19	-30.89	-32.27
107124	-31.65	-31.40	-25.74	-26.85	-28.91
176323	-30.63	-31.12	-27.05	-27.88	-29.17
249910	-30.20	-30.72	-25.99	-26.90	-28.45
295500	-33.25	-33.90	-28.81	-28.83	-31.20
295501	-28.67	-29.06	-27.85	-27.03	-28.15
354646	-35.40	-31.15	-33.81	-34.17	-33.63
364830	-27.49	-26.64	-28.87	-24.26	-26.82
374028	-29.07	-29.30	-32.37	-31.53	-30.57
603071	-31.14	-31.75	-27.17	-28.41	-29.62
606172	-32.66	-31.97	-29.26	-28.67	-30.64
606173	-31.40	-30.13	-28.46	-26.36	-29.09
606497	-27.12	-26.85	-28.66	-28.53	-27.79
606499	-30.67	-30.00	-29.42	-30.01	-30.03
606985	-26.10	-26.32	-28.56	-27.89	-27.22
610456	-28.86	-28.33	-29.70	-30.43	-29.33
610457	-26.01	-25.64	-27.26	-30.96	-27.47
610458	-35.15	-35.13	-32.26	-28.27	-32.70
610459	-29.43	-29.95	-29.42	-32.78	-30.40
618939	-27.96	-23.85	-24.10	-25.31	-25.31
629971	-31.22	-31.81	-27.54	-28.64	-29.80
643833	-28.58	-27.20	-27.14	-25.89	-27.20

**Fig. 10** Docked conformations of compounds 610458 (a) and 354646 (b)**Fig. 9** Docked conformations of compounds 100880 (a) and 374028 (b)

and Asn722 (3.43 Å), and between N₁ and Arg364. A deep intercalation completes the binding pattern for this derivative.

Compound 354646 is the sole non-camptothecin like inhibitor of Top1 present in the NCI Anti-cancer Agents Mechanism Database, with an anthracyclin skeleton. In our experiments it provided low ΔG values with every macromolecule (mean $\Delta G = -33.63$ kcal/mol), giving the best binding affinity ($\Delta G = -35.40$ kcal/mol) with 1K4T. The best binding conformation predicted by our models is illustrated in Fig. 10b. There are many possible interactions between active site residues and the drug, due to the large presence of H-bond interacting groups such as carbonyl and hydroxyl ones, accompanied by an intercalation mode which favors π -interactions between the drug aromatic rings and the stacking base pairs.

The analysis of the conformations obtained with molecular docking can also clarify the role of some Top1 residues in determining the resistance to CPT derivatives. In fact several point mutations have been identified, including Arg364 [21], Asp533 [22], and Asn722 [17], which result in a lack of activity of camptothecins. In our models these residues are the best H-bonding site between the drug and Top1, interacting respectively with N₁ of camptothecin ring system (at a distance ≈ 3.25 Å), with the substituent on C₂₀

(distance in the range 2.47–3.60 Å, depending upon the nature of the substituent) and with the oxygen atoms on E-ring; thus, it is clear how a loss of such interactions can determine a loss of activity and easily leads to drug-resistance phenomena.

Conclusions

In this contribution we used a docking procedure with the aim to investigate the binding conformations adopted by a set of 24 Topoisomerase I inhibitors from the NCI Anti-cancer Agents Mechanism Database. The use of four different structures of Top1-DNA complex crystallized with four inhibitors of various sizes permitted us to give account of the eventual variations in the orientation of active site residues, thus allowing to dock structures with bulky substituents. In general, we observed low ΔG values for all derivatives. Poorly substituted camptothecins generally fit well into protein structure in which a simple camptothecin-like drug was crystallized (1K4T, 1T8I), whereas highly substituted compounds are better docked in those macromolecule crystallized with larger inhibitors.

The analysis of the best docked conformations permitted us to investigate the binding mode of compounds involved in this study, confirming the role of some amino acids present in the active site as previously proposed. These results can also clarify and explain the role of some single point mutation in developing resistance to camptothecins, and are useful in order to understand the structural features required to improve the performance of camptothecin derivatives as Topoisomerase I inhibitors.

References

1. Been MD, Champoux JJ (1984) *J Mol Biol* 180:515
2. Pommier Y (1998) *Biochimie* 80:255
3. Stewart L, Ireton GC, Champoux JJ (1996) *J Biol Chem* 271:7602
4. Stewart L, Ireton GC, Champoux JJ (1997) *J Mol Biol* 269:355
5. Stewart L, Ireton GC, Champoux JJ (1999) *J Biol Chem* 274:32950
6. Wall ME, Wani MC, Cook CE, Palmer KH, McPhail AT, Sim GA (1966) *J Am Chem Soc* 88:3888
7. Hsiang YH, Hertzberg R, Hecht S, Liu LF (1985) *J Biol Chem* 260:14873
8. Thomas CJ, Rahier NJ, Hecht SM (2004) *Bioorg Med Chem* 12:1585
9. Wang X, Zhou X, Hecht SM (1999) *Biochemistry* 38:4374
10. Staker BL, Hjerrild K, Feese MD, Behnike CA, Burgin AB, Stewart L (2002) *Proc Natl Acad Sci USA* 99:15387
11. Staker BL, Feese MD, Cushman M, Pommier Y, Zembower D, Stewart L, Burgin AB (2005) *J Med Chem* 48:2336
12. Monks A, Scudiero D, Skehan P, Shoemaker R, Paull K, Vistica D, Hose C, Langley J, Cronise P, Vaigro-Wolff A, Gray-Goodrich M, Campbell H, Mayo J, Boyd M (1991) *J Natl Cancer Inst* 83:757
13. Stewart JJP (1989) *J Comp Chem* 10:209
14. Goodsell DS, Olson AJ (1990) *Proteins* 8:195
15. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) *J Comput Chem* 19:1639
16. Gasteiger J, Marsili M (1980) *Tetrahedron* 36:3219
17. Chillemi G, Fiorani L, Benedetti P, Desideri A (2003) *Nucleic Acids Res* 31:1525
18. Fan Y, Weinstein JN, Kohn KW, Shi LM, Pommier Y (1998) *J Med Chem* 41:2216
19. Fujimori A, Harker WG, Kohlhagen G, Hoki Y, Pommier Y (1995) *Cancer Res* 55:1339
20. Hertzberg RP, Caranfa MJ, Holden KG, Jakas DR, Gallagher G, Mattern MR, Mong SM, Bartus JO, Johnson RK, Kingsbury WD (1989) *J Med Chem* 32:715
21. Li XG, Haluska P, Hsiang YH, Bearti AK, Kufe DW, Liu LF, Rubin EH (1997) *Biochem Pharmacol* 53:1019
22. Saleem A, Ibrahim N, Patel M, Li XG, Gupta E, Mendoza J, Pantazis P, Rubin EH (1997) *Cancer Res* 57:5100