10010–10019 *Nucleic Acids Research, 2013, Vol. 41, No. 22* *Published online 9 September 2013*
doi:10.1093/nar/gkt791

Sequence selectivity of the cleavage sites induced by topoisomerase I inhibitors: a molecular dynamics study

Fung-Ming Siu[1,2,*] and Yves Pommier[3,*]

[1]Center for High Performance Computing, Institute of Advanced Computing and Digital Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 1068 Xueyuan Boulevard, University Town of Shenzhen, Xili Nanshan, Shenzhen 518055, China, [2]Department of Chemistry and Institutes of Molecular Technology for Drug Discovery and Synthesis, State Key Laboratory of Synthetic Chemistry, The University of Hong Kong, Pokfulam Road, Hong Kong and [3]Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, Bethesda, MD, USA

Received July 2, 2013; Revised August 1, 2013; Accepted August 9, 2013

ABSTRACT

Topoisomerase IB (Top1) inhibitors, such as camptothecin (CPT), stabilize the Top1-DNA cleavage complex in a DNA sequence-dependent manner. The sequence selectivity of Top1 inhibitors is important for targeting specific genomic sequences of therapeutic value. However, the molecular mechanisms underlying this selectivity remain largely unknown. We performed molecular dynamics simulations to delineate structural, dynamic and energetic features that contribute to the differential sequence selectivity of the Top1 inhibitors. We found the sequence selectivity of CPT to be highly correlated with the drug binding energies, dynamic and structural properties of the linker domain. Chemical insights, gained by per-residue binding energy analysis revealed that the non-polar interaction between CPT and nucleotide at the +1 position of the cleavage site was the major (favorable) contributor to the total binding energy. Mechanistic insights gained by a potential of mean force analysis implicated that the drug dissociation step was associated with the sequence selectivity. Pharmaceutical insights gained by our molecular dynamics analyses explained why LMP-776, an indenoisoquinoline derivative under clinical development at the National Institutes of Health, displays different sequence selectivity when compared with camptothecin and its clinical derivatives.

INTRODUCTION

Topoisomerase IB (Top1) is an essential enzyme required for DNA replication and RNA transcription. Top1 is composed of 765 amino acids and has four distinct domains: the NH_2-terminal domain (1-214), the core domain (215-635), the linker (636-712) and the COOH-terminal domain (713-765). The Top1 enzyme catalyzes the relaxation of DNA superhelicity by reversibly nicking one strand of DNA to generate a Top1-DNA cleavage complex that allows controlled DNA swiveling. Top1 inhibitors act as interfacial inhibitors (1). They intercalate into the DNA base pairs at the cleavage site and form specific hydrogen bonds with Top1 to stabilize the Top1-DNA-inhibitor ternary complex (Figure 1A). The stabilized ternary complex produces collisions with replication and transcription forks leading to DNA double strand breaks and cell death (Supplementary Scheme S1) (2). The discovery and development of Top1 inhibitors as anticancer drugs is one of the significant achievements in cancer chemotherapy (3,4). Topotecan and irinotecan, analogs of the natural alkaloid camptothecin (CPTs, Supplementary Chart S1), have been approved by the US Food and Drug Administration for the treatment of ovarian and colon cancers, respectively. The CPT class of anticancer drugs targets only Top1 and the poisoning of Top1-DNA cleavage complexes occurs only at specific cleavage sites along the DNA chain. This sequence selectivity is a key step in preferentially directing the action of drugs onto specific genomic sequences of therapeutic interest (5–8). The sequence selectivity exhibited by Top1 inhibitors is 'intrinsically' dependent on the base

*To whom correspondence should be addressed. Tel: +86 755 86392336; Fax: +86 755 86392299; Email: xiao.fm@siat.ac.cn
Correspondence may also be addressed to Yves Pommier. Tel: +1 301 496 5944; Fax: +1 301 402 0752; Email: pommier@nih.gov

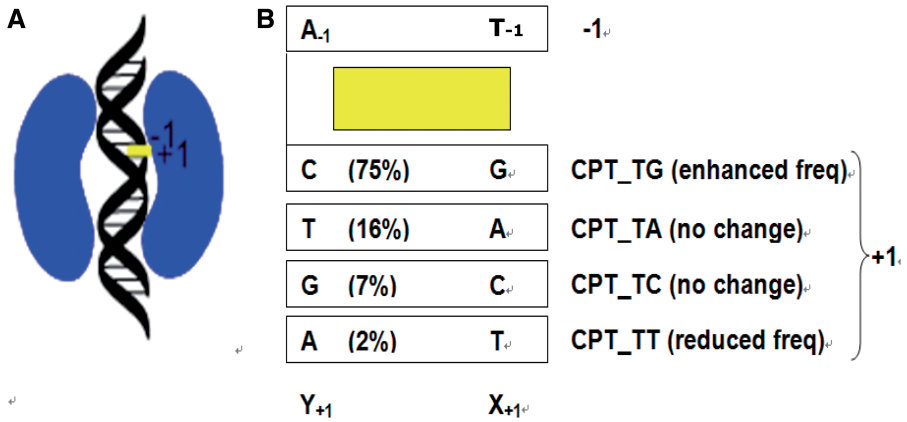

Figure 1. (**A**) Schematic diagram of a Top1-DNA-inhibitor ternary complex. Top1, DNA and CPT are shown in kidney-, helix- and rectangle-shape, respectively. (**B**) Schematic diagram of the CPT at the cleavage site, with its flanking base pairs. The frequencies of base sequences observed in CPT-induced DNA cleavage in SV40 DNA are given in parenthesis. Experimental base frequencies revealed that base frequencies at DNA cleavage sites was significantly (99.9% confidence interval) enhanced at the TG cleavage site in the presence of CPT, but no significant preference was seen for the TA and TC cleavage sites. The action of CPT was almost abolished at the TT cleavage site. Its base frequency was significantly (99.9% confidence interval) reduced in the presence of CPT.

sequences immediately preceding (−1) and following (+1) the cleavage site. CPT binding has a strict requirement for T at the −1 site and G at the +1 site (9). Analysis of sequence selectivity around Top1 cleavage sites in SV40 DNA and 30mer duplex oligonucleotides indicated that CPT action is strongly dependent on the base pair at the +1 position (9–12). Recent reports showed that CPT interfered with transcription regulation in a sequence-selective manner, leading to alterations in gene expression that may be relevant for cancer therapy (13,14). However, the biochemical and biophysical determinants underlying the sequence selectivity remains unclear.

In the present study, we investigated the impact of various local DNA sequences on the energetic, dynamic and structural properties of Top1-DNA-drug ternary complexes. This is the first molecular dynamics (MD) study on the sequence selectivity of Top1 inhibitors. We first examined four ternary complexes namely, CPT_TG, CPT_TA, CPT_TC and CPT_TT (Figure 1B). Experimental base frequencies (9–12) revealed that base frequencies at DNA cleavage sites was significantly (99.9% confidence interval) enhanced at the TG cleavage site in the presence of CPT, but no significant preference was seen for the TA and TC cleavage sites. The action of CPT was almost abolished at the TT cleavage site; its base frequency was significantly (99.9% confidence interval) reduced in the presence of CPT. In other words, sequence selectivity of CPT is in the following order of preference (Figure 1B): CPT_TG (frequency enhanced, compared with the frequency of the cleavage site in the absence of CPT) > CPT_TA (no change) ~ CPT_TC (no change) > CPT_TT (frequency reduced). Potential of mean force (PMF) is defined as the free-energy changes along an arbitrarily defined path (15). In the present study, PMF study is performed to mimic the drug dissociation process. This is the first PMF studies on Top1-DNA-drug ternary complexes. We further demonstrate that MD approach could be used to study the sequence selectivity of an indenoisoquinoline derivative NSC725776 (LMP-776, Supplementary Chart S1). LMP-776 produces Top1 cleavage at unique genomic position compared with those resulting from CPT treatment, causes cell cycle arrest in both S and G2/M, exerts part of their anti-tumor effect through antiangiogenesis. In contrast to CPT, overexpression of ABCG2 and ABCB1 pumps do not induce resistance for LMP-776. These favorable pharmaceutical properties render LMP-776 a promising anticancer drug candidate (2,3). LMP-776 is currently under clinical development at the National Institutes of Health (NIH).

MATERIALS AND METHODS

MD simulations

The initial geometries of the Top1-DNA-drug ternary complexes were modeled from the crystallographic structures of ternary complexes (PDB 1K4T, 1SC7) (16,17). The CPT-bound and LMP-776-bound ternary complexes were modeled using atomic positions from the 1K4T and 1SC7 structures, respectively (18–22). The topotecan and MJ-III-65 were changed to CPT and LMP-776, respectively; the 5′-sulfhydryl group was replaced by a 5′-hydroxyl group at the cleavage site in the ternary complexes using the Accelrys DS Visualizer. The lactone form of CPT was assumed (23–25). All MD simulations were carried out with the Particle Mesh Ewald molecular dynamics (PMEMD) module in AMBER12 using the ff10 variant and general AMBER force field (GAFF) force fields (26,27). Although MD simulations are known to be limited by force field approximations, AMBER-based force fields have been successfully used to reflect key features of the structural dynamics of DNA and protein molecules (28–30). Counter ions (20 Na^+) were added to keep the whole simulated system neutral. The whole system was immersed in the rectangular box of TIP3P water molecules. The water box extended 8 Å away from any solute atom. This resulted in ~28 270 water molecules used for solvation. Electrostatic interactions were taken

into account using the Particle Mesh Ewald method with a cutoff distance of 10 Å (31). We used SHAKE to constrain all bonds involving hydrogen atoms (32). Energy-minimization was executed by the steepest descent method for the first 1000 steps and by the conjugated gradient method for the subsequent 2500 steps. Relaxation of solvent molecules and Na^+ ions was initially performed keeping solute atoms restrained to their initial positions for 200ps. The temperature was controlled by Langevin dynamics and was kept constant at 300 K. Each model was simulated for 100 ns. Over 100 000 frames were uniformly extracted over the 100 ns of simulations (33), unless otherwise specified.

Analysis of trajectories

The trajectories were analyzed using the AMBER12 packages (26), the VMD program (34) and the GROMACS MD package v4.5.5 (35).

Root-mean-square deviations

The cpptraj module in AMBER12 (26) was used to calculate the root-mean-square deviations (RMSD). RMSD is defined as the root-mean-square-average distance between protein C_α atoms of two optimally superimposed macromolecules, i.e. superimposed macromolecules at each frame (S_i) and the starting geometry (S_0). After a mass-weighted least square fitting to the starting geometry, the RMSD were calculated using Equation (1):

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{i=N}\delta_i^2} \quad (1)$$

where N is the number of equivalent protein C_α atoms in each macromolecules, δ is the distance between N pairs of equivalent atoms (i.e. protein C_α atoms). The RMSD values reach an average value of ~3 Å (Supplementary Figure S1A). Excluding the linker domain from the analysis, the RMSD values reduce to ~2 Å (Supplementary Figure S1B). Stable plateaus were reached in all the systems well before 30 ns of the simulation time.

Cosine content (c_i)

Based on the production MD trajectories (the last 70 ns), non-weighted covariance matrix of Cα atoms of the models were calculated to obtain cosine content (c_i) using the GROMACS MD package v4.5.5 (35). The cosine content (c_i) of a principal component is used as an negative indicator for bad sampling (36). It ranges between 0 (no cosine) and 1 (perfect cosine), when the first of three principal components is similar to a cosine with half a period; the sampling is far from converged. The cosine contents of the first three principal components are far from a perfect cosine ($c_i < 0.1$), indicating that our systems are not bad sampling. Thus, all the analyses listed later in the text were calculated over the production MD trajectories (i.e. the last 70 ns).

Dynamic cross-correlation coefficients ($C_{i,j}$)

The dynamic cross-correlation coefficients were done with cpptraj module in the AMBER package (26), taking into account only the coordinates of the protein Cα, nucleotide N1, N1@CPT, N1@LMP-776 atom (Supplementary Chart S1), as they contain enough information to describe the largest system motions (37). The dynamic cross-correlation coefficients were computed based on Equation (2):

$$C_{i,j} = \frac{(\Delta r_i \bullet \Delta r_j)}{\left(\sqrt{\langle \Delta r_i^2 \rangle} \bullet \sqrt{\langle \Delta r_j^2 \rangle}\right)} \quad (2)$$

where Δr_i is the displacement from the mean position of the ith atom and <>represents the time average over the whole trajectory. Positive $C_{i,j}$ values represent a correlated motion between residues i and j (i.e. the residues move in the same direction). Negative values of $C_{i,j}$ represent an anti-correlated motion between residues i and j (i.e. they move in opposite directions).

Root-mean-square fluctuations

The per-residue root-mean-square fluctuations (RMSF) is defined as a root-mean-square-average distance between a protein C_α atom (x_i) and its time-averaged position of the equivalent atom ($\tilde{x}_i$) in a set of S structures that were extracted from the frames of the production MD trajectories. RMSF were computed using Equation (3):

$$RMSF = \sqrt{\frac{1}{S}\sum_{i=1}^{S}(x_i - \tilde{x}_i)^2} \quad (3)$$

Clustering

The clustering was calculated using the cpptraj module in AMBER12 (26). Structures from the frames of the production trajectories were used to perform average-linkage hierarchical clustering, based on their mass-weighted RMSD distance matrix.

DNA-linker angle

The DNA-linker angle is defined as the angle between N1@A_{+12}−N1@X_{+1}−Cα@ASP677 (Supplementary Scheme S2). X_{+1} and A_{+12} are the nucleotides at the +1 and +12 positions of the intact DNA strand.

Hydrogen bond

Hydrogen bonds were defined as the default setting at the cpptraj program. Hydrogen bonds are determined via the distance between the heavy atoms using a cutoff of 3.5 Å and the angle between the acceptor, hydrogen, and donor atoms using a cutoff of 135 degrees, unless otherwise specified.

Molecular mechanics-generalized born surface area

The drug binding energies were calculated using the mmGBSA.py module (igb = 2) in AMBER12 (26). The

per-residue decomposition energies reflect the relative strength of each specific residue, i.e. its importance to the binding. The post-processing molecular mechanics-generalized born surface area (MM-GBSA) method, using the modified generalized Born model (38), is known to be good at qualitative accuracy (39–43). We herein focused our discussions on qualitative comparison. The total binding energy (ΔG_{tot}) was decomposed into contributions from electrostatic energies (ΔE_{elec}), van der Waals interaction energies (ΔE_{vdw}), electrostatic solvation (ΔG_{GB}) and non-polar solvation ($\Delta G_{non\text{-}polar}$) free energies as shown in Equation (4):

$$\Delta G_{tot} = \Delta E_{elec} + \Delta E_{vdw} + \Delta G_{GB} + \Delta G_{non-polar} \quad (4)$$

The entropy was not considered because of its high computational demand and relatively low accuracy of prediction (44). The per-residue decomposition with 1–4 VDW and 1–4 EEL terms were added to internal potential terms.

PMF

PMF was calculated to study the free-energy profiles for drug dissociation. The PMF is a function of the distance between the CPT and the ternary complex (χ), which can be used to monitor the drug dissociation process (Supplementary Scheme S3). The distance (χ) is defined as the distance between P@A_{-1} and N1@CPT (Supplementary Chart S1).

Umbrella sampling was performed by calculating a series of dynamic simulations with biased potential in a format of $K(\chi - \chi_i)^2$, where K = 5 or 10 or 20 kcal mol^{-1} and χ_i ranged from 0 to 50 Å. Umbrella sampling windows were separated by 1 Å. Simulations with K = 5 were carried out for at least 0.2 ns; the rest of the simulations were carried out with K = 10 and 20 for at least 0.4 ns each, until convergence is reached. All the simulations were combined to construct the final PMF. Multiple starting structures were extracted from the all-atom simulations. Each simulation was started from the end point of the previous one. Adjacent histograms were found to show significant overlaps, indicating that all values of a distance within the desired range had been appropriately sampled. The unbiased free energy was obtained using the weight histogram analysis method (http://membrane.urmc.rochester.edu/content/wham) with convergence criterion of 10^{-4} and a bin width of 0.2 Å (45). The convergence of final PMF was tested by comparing the free energy change calculated using the full sampling to that calculated using shorter sampling time. For each PMF models, at least 50 ns of simulations were carried out until the PMF profile converged.

RESULTS AND DISCUSSION

Binding energies of CPT to the Top1-DNA complex

The total binding energies of CPT to the Top1-DNA complex ranged from −54.9 kcal mol^{-1} for CPT_TG to −49.9 kcal mol^{-1} for CPT_TT (Table 1), which were considered to be reasonable for two reasons. First, in the previous studies by Cushman *et al.* (46) (−32 kcal mol^{-1}) and Siu *et al.* (47) (−26 to −31 kcal mol^{-1}), the reported binding energies of a Top1 inhibitor to its flanking base pairs were in the same order of magnitude. The more favorable binding energies obtained in the present study could be attributed to contributions from protein residues and other DNA bases, which were not considered in the two previous studies (46,47). Second, the experimental frequencies of cleavage sites correlate strongly with the total binding energies ($R^2 = 0.8$, Table 1). To probe properties that underpin the sequence selectivity, we identified their correlations with various per-residue decomposition energies (Table 1). A strong correlation between the experimental frequencies and the binding energy between CPT and the nucleotide at the +1 position of the scissile strand (X_{+1}) was observed ($R^2 = 0.9$, Table 1). The non-polar interaction ($\Delta E_{vdw} + \Delta G_{non\text{-}polar}$, Table 2) was the major favorable contributor to the total binding energy. Taken together, our results indicate that the non-polar interaction between CPT and the X_{+1} is a predominant, if not exclusive, factor governing the sequence selectivity of CPT. We would like to emphasize that given there are only four data points for the correlation analyses, no quantitative analysis were discussed.

Table 1. The per-residue decomposition energies (kcal mol^{-1})[a]

Decomp. En.[a]	CPT_TG	CPT_TA	CPT_TC	CPT_TT
T_{-1}	−4.8	−5.0	−5.0	−5.0
X_{+1}	−9.1	−6.6	−5.8	−6.4
Y_{+1}	−2.4	−3.2	−5.0	−3.3
A_{-1}	−5.4	−4.5	−5.6	−5.2
Arg364	−2.5	−2.5	−2.2	−2.4
Lys532	−1.8	−1.7	−1.1	−0.9
CPT	−27.2	−26.0	−25.4	−24.5
Others	−1.7	−2.0	−2.2	−2.2
Total	**−54.9**	**−51.5**	**−52.3**	**−49.9**

[a]The total binding energies are highlighted in bold. The per-residue decomposition energies reflect the relative importance of each specific residue, including flanking base pairs at the cleavage site, Arg364, Lys532, CPT and other residues, to the total drug binding energies.

Table 2. The MM-GBSA per residue decomposition energies (kcal mol^{-1})[a] for the interaction between CPT and the X_{+1} base at the cleavage site

Decomp. En.	CPT_TG	CPT_TA	CPT_TC	CPT_TT
ΔE_{elec}	−2.9	−1.5	−0.9	−1.8
ΔE_{vdw}	−7.8	−6.9	−5.6	−5.7
ΔG_{GB}	2.2	2.4	1.2	1.5
$\Delta G_{non\text{-}polar}$	−0.6	−0.6	−0.5	−0.4
Electrost[b]	−0.7	0.9	0.3	−0.3
Non-polar[c]	**−8.4**	**−7.5**	**−6.1**	**−6.1**
ΔG_{tot}	−9.1	−6.6	−5.8	−6.4

[a]The major favorable contributor is highlighted in bold.
[b]Electrost = $\Delta E_{elec} + \Delta G_{GB}$.
[c]Non-polar = $\Delta E_{vdw} + \Delta G_{non\text{-}polar}$.

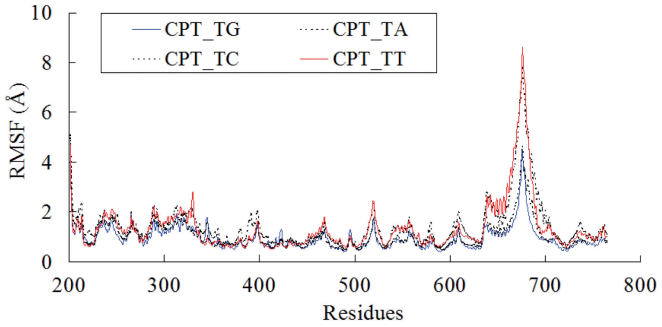

Figure 2. RMSF of CPT-bound ternary complexes. The solid lines represent the ternary complex with highest observed frequency (CPT_TG) and lowest observed frequency (CPT_TT).

Flexibility of the linker domain and its orientation with respect to the DNA

The per-residue RMSF calculated from the production trajectories showed that the four complexes underwent similar fluctuations (Figure 2). As observed in previous studies (24,48), the linker domain (residues 636–712) was the most flexible region. Moreover, the flexibility of the linker domain is qualitatively correlated with the experimental frequencies of the sequence selectivity. First, for the linker domain, the RMSF maximum value for the preferred CPT_TG complex (2.1 Å) is lower than those of the unfavorable CPT_TA/TC/TT complexes (2.4–4.7 Å, Figure 2). The linker and core domains are connected by residues 633–643, whose overall mobility increases as the number of orientations of the linker domain increases (24). Consistently, our analysis revealed that the RMSF maximum value of residues 633–643 is the lowest for the preferred CPT_TG complex (1.2 Å) when compared with those of the unfavorable CPT_TA/TC/TT complexes (1.5–2.1 Å, Figure 2). Cluster analysis further confirmed that the conformational space visited by the preferred CPT_TG complex (nine clusters) is less than that of the unfavorable complexes (CPT_TA/TC/TT complexes, 11–27 clusters). Collectively, our study revealed that the sequence selectivity increases as the flexibility of the linker domain decreases.

Comparing the average structures of the CPT-bound ternary complexes (Figure 3), a notable difference was observed in the linker domain. The angles formed between the DNA and the linker domain (θ) in these complexes were determined (Supplementary Scheme S2). The linker domain of the preferred CPT_TG complex adopts a position that is closer to the DNA when compared with that of the unfavorable CPT_TA/TC/TT complexes (Figure 3). The experimental frequencies of cleavage sites correlate strongly with the DNA-linker angle ($R^2 = 0.9$). We postulate that interaction(s) between the linker domain and the DNA duplex is needed to stabilize the ternary complexes, which leads to the higher observed frequencies.

Long-range interactions between CPT and protein residues

To examine why the orientation of the linker domain changed dramatically, we studied the hydrogen bonds (Supplementary Figure S2) and the correlative motion of the CPT-bound ternary complexes. The correlative motion of CPT with protein residues (Figure 4) were investigated by means of the dynamic cross-correlation coefficients ($C_{CPT,j}$), with positive and negative $C_{CPT,j}$ values signify correlated and anti-correlated motions, respectively (37).

The $C_{CPT,j}$ (Figure 4) values are highly positive for the Arg364 and Lys532 residues. Single amino acid mutations at these residues are known to confer CPT resistance (49). The highly positive $C_{CPT,Arg364}$ values for the four CPT-bound ternary complexes revealed that CPT moves along the same direction with the Arg364 residue. A hydrogen bond is observed between N1 of CPT and the side-chain of Arg364 (Supplementary Figure S2). This hydrogen bond was also observed in the X-ray crystallographic structure of a Top1-DNA-CPT complex (17). The highly positive $C_{CPT,Lys532}$ values also revealed that motion of CPT correlate with that of the Lys532 residue (Figure 4). The side-chain of Lys532 forms hydrogen bond with the lactone ring of CPT (O21@CPT, Supplementary Chart S1) for >80% of the simulation time (Supplementary Figure S2). Lys532 is a known catalytic residue making a hydrogen bond with carboxylate O21 of topotecan in the toptoecan-Top1-DNA ternary X-ray crystallographic structure (16).

In contrast, CPT moves in an anti-correlated way with the linker domain (residues 636–712). The $C_{CPT,linker}$ values (Figure 4) indicates that the strength of the anti-correlated motion is in the order of CPT_TT (most negative) > CPT_TA ~ CPT_TC > CPT_TG (least negative). As there is no direct hydrogen bond between CPT and the linker domain, the anti-correlated motion is expected to be mediated through DNA-linker interactions. Hydrogen bonds are observed between the backbone nitrogen of the Arg634 and Gly717 residues (linker domain) and the phosphodiester groups of G_{+2} and A_{+3} bases at the scissile strand, respectively. In addition, the free 5′-OH group of the +1 base at the scissile strand (X_{+1}) forms intermolecular hydrogen bond with the lateral chain of the Asn722 residue. This hydrogen bond presents for >80% of the simulation time for all the complexes. The Arg634 forms a loop (residues 634–640) connecting the linker to the core domain. This loop is known to have a direct impact on the flexibility of the linker domain (24), as the loop and the linker domain are also absent in the X-ray crystallographic structure of the Top1-DNA binary complex (PDB: 1A31), but observable in the drug-bound ternary complexes (PDB: 1K4T) (Supplementary Figure S3) (50). Gly717Val mutated Top1 was isolated in a CPT-resistant cell line (51). In Top1, the Asn722 amino acid residue is immediately NH_2-terminal to the catalytic Tyr723. The mutation Asn722Ser in Top1 leads to CPT resistance (52,53). Collectively, these findings indicate that there are long-range interactions between the cleavage site and the linker domain.

The linker flexibility has been shown to affect enzyme catalytic activity, Top1 sensitivity to CPT and long-range effects on the flexibility of the linker domain. In the absence of CPT, Lys681Ala and Thr718Ala mutations

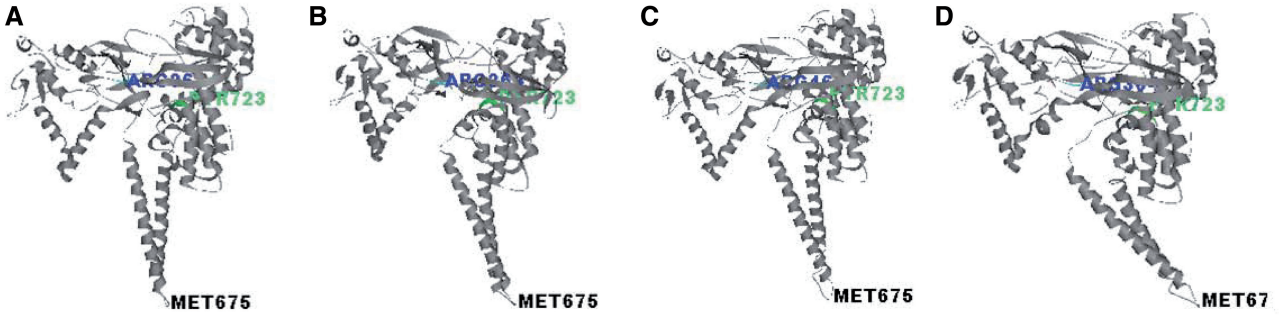

Figure 3. Average structures of (**A**) CPT_TG, (**B**) CPT_TA, (**C**) CPT_TC and (**D**) CPT_TT complexes. The DNA and CPT are not shown, and the ARG364, PTR723 and MET675 residues are labeled.

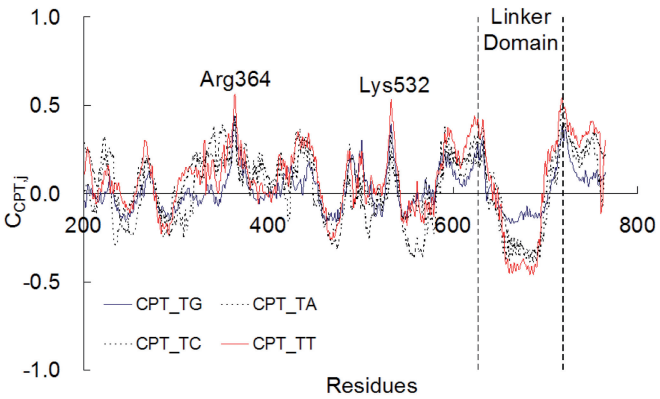

Figure 4. Dynamic cross-correlation coefficients ($C_{CPT,j}$) between CPT and protein residues in Top1.

render the flexibility of the linker domain reduces. The Lys681Ala mutant has a CPT-like behavior: the stability of the cleavage complex is enhanced and the distance between Lys532 and the 5′OH of the +1 base phosphates is increased (54). The Thr718Ala mutant shows a decreased efficiency in DNA relaxation (55). We note with interest that although the Ala653Pro mutation in the linker domain confers CPT resistance and renders the linker domain more flexible (56), the mutations in Asp677Gly, Val703Ile and Ile728Thr renders the mutant more sensitive to CPT and higher binding affinity to DNA (57). It is also well-known that the reconstituted top1 lacking the linker domain has a reduced affinity for DNA (58). X-ray crystallographic (50) and MD studies (24) revealed that the linker domain is highly flexible in the drug-free binary Top1-DNA complexes, when compared with that of the CPTs-bound ternary complexes (Supplementary Figure S3). Collectively, the CPT-induced or mutation-mediated inhibitory effects on Top1 increase as the mobility of the linker domain decreases. Accordingly, we found the sequence selectivity of CPT is correlated with the flexibility of the linker domains. We conceive that both the inhibitory and sequence selectivity effects are associated with the stability of the cleavage complexes, which has long-range effects on the flexibility of the linker domain.

Drug development prospective

Having established the driving force for the site selectivity of CPT, we extended our study to non-CPTs, with an aim to evaluate the potential application of MD approach for studying the site selectivity of non-CPT Top1 poisons (59). CPTs have several undesirable properties, including E-ring inactivation by lactone ring opening, affinity for cellular efflux transporters and relatively high toxicity, which limits their therapeutic index. Non-CPT Top1 inhibitors could have distinct pharmacological activities because they target distinct genomic sites due to their different sequence selectivity; they also overcome the chemical instability of CPT and lead to distinct pharmacological activity (59–62). An indenoisoquinoline NSC725776 (LMP-776, Supplementary Chart S1) is currently in clinical trial at the National Institutes of Health (USA). The DNA cleavage pattern induced by LMP-776 is different from that of CPT. Notably, LMP-776 prefers a cleavage site at the CG site [site 44 in the original article (3,63)], rather than the TG site (the preferred site of CPT).

In the present study, we compared the ternary complexes of LMP-776 bound with Top1-DNA complexes with CG versus TG as the cleavage sites (referred to as LMP_CG and LMP_TG, Figure 5). In agreement with the trend observed in CPT, the drug binding energy of the LMP_CG complex (−60.6 kcal mol^{-1}) was more favorable than the binding energy of the LMP_TG complex (−54.3 kcal mol^{-1}). The flexibility of the linker domain of LMP_CG (2.7 Å) was also lower than that of LMP_TG (2.9 Å, Figure 5B). In addition, the DNA-linker angle (θ) of the LMP_CG (14°) complex was significantly smaller (*t*-test: $P < 1E\text{-}5$) than that of the LMP_TG complex (29°, Figure 5C). Collectively, LMP-776 is consistent with CPT that sequence selectivity was governed by drug binding energies, flexibility of the linker domain and the DNA-linker angle. To target specific genomic sequences of therapeutic value, developing top1 inhibitors with different preference on the cleavage site is of high importance. Determining the base frequencies of various top1 inhibitors experimentally is time-consuming and tremendous. We have illustrated that the MD approach reported in the present study can be used to study the sequence selectivity of CPT and non-CPTs.

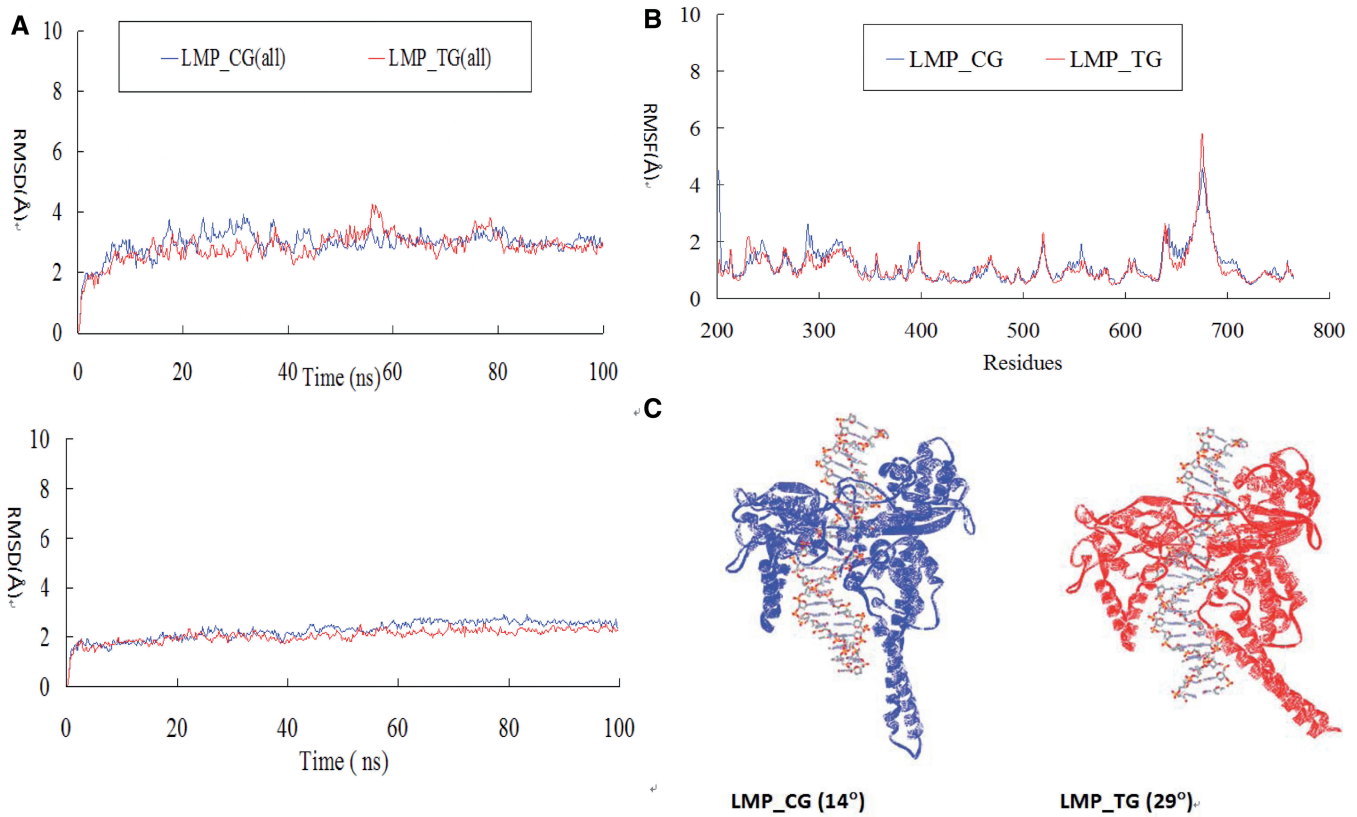

Figure 5. Data for LMP-776 (NSC725776). (**A**) RMSD with and without the linker domain; (**B**) RMSF; (**C**) Average Structures for the LMP-bound ternary complex.

We envision MD approach can be used to *in silico* screen compounds that target specific genomic sequences.

Mechanistic perspective

The perturbed mobility of the linker domain has been reported to be associated with alteration in religation rate and drug resistance (56,64–67). Drug dissociation is the rate determining step when compared with DNA religation. To investigate the underlying mechanism for the sequence selectivity of the drugs, we modeled the drug dissociation process using the PMF approach. PMF is defined as the free energy changes along the drug dissociation distance (χ) (Supplementary Scheme S3). Using multiple starting geometries, we calculated the PMFs and found that CPT dissociates towards to the NH_2-terminal domain (Figure 6A). The residues 190–210 at the NH_2-terminal domain are known to be important for Top1 activity (68–72): The physical contact of Trp203 with His346 and of Trp205, Trp206 with the hinge region of the core domain controls the DNA rotation process (73). Site mutations at these tryptophan residues are known to confer CPT resistance (68). Using MD approach, Chillemi *et al.* (71) found that the residues 203–207 at the NH_2-terminal domain and the linker domain participate in a network of correlated movements, involved in the Top1 catalytic cycle and associated with the DNA relaxation activity. Our results shed light on the potential role of the NH_2-terminal domain in the drug dissociation process that may warrant further study.

A closer inspection of the PMFs of the CPT_TG and CPT_TT complexes revealed that the DNA-linker angle in the CPT_TG complex is smaller than that of the CPT_TT complex (Figure 6B), and the drug dissociation energy barrier for CPT_TG is significantly higher than that of the CPT_TT complex (Figure 6C). The smaller DNA-linker angle in the CPT_TG complexes, when compared with that of CPT_TT, implied that the linker domain and DNA remain in close contact during the drug dissociation process in the former complex. The linker domain is known to play important roles in stabilizing the cleavage complex. The stabilized CPT_TG complex renders drug dissociation more difficult, as represented by a higher drug dissociation barrier. We therefore conclude that the DNA-linker interaction also play a role in impeding drug dissociation process. In the absence of CPT, the DNA sequence does not affect the dissociation rate constants or the rate of association between DNA and Top1 (74), as the four sequences were similarly cleaved (9). DNA sequence does not appear to play a dominating role in the rate of drug-mediated DNA uncoiling (75). Top1-mediated DNA breakage is enhanced the greatest at those sites where closure of the break is the slowest and that the CPT-mediated religation rate is sequence dependent (76–78). Collectively, previous experiments together with our findings imply that the drug dissociation step plays

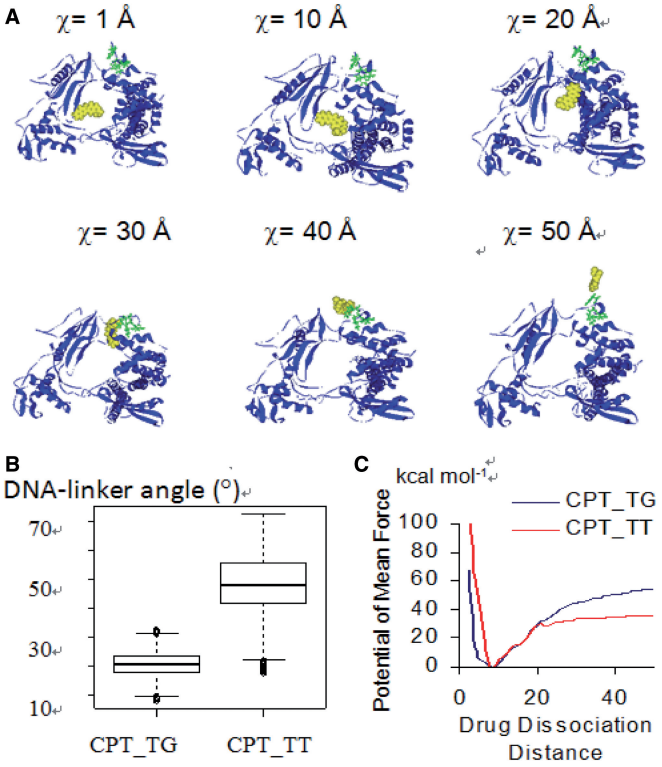

Figure 6. (**A**) Snapshots of the ternary complex as the drug dissociate from the cleavage complex. The CPT and the three Trp203/205/206 residues in the NH2-terminal of Top1 are shown in CPK and ball and stick styles, respectively; the DNA strands are not shown for the ease of visualization. (**B**) Angles formed between DNA and the linker domain of the complexes during the drug dissociation process. (**C**) The PMF of CPT_TG complex is higher than that of the CPT_TT complex.

a critical role in the sequence selectivity of Top1 inhibitors.

Final remarks

Our study provides novel insights into the mechanistic determinants for the sequence selectivity of Top1 inhibitors. The sequence selectivity of CPT was revealed to be correlated with energetic (drug binding energy), dynamic (flexibility of the linker domain) and structural (angle formed between the linker domain and DNA) properties of the drug-bound ternary complexes. Our structural analysis and PMF study further reveal that the hydrogen bonding network between DNA and the linker domain and the drug dissociation step plays a role in the sequence selectivity induced by CPT. We also demonstrated the use of MD approaches to study the sequence selectivity of non-CPTs Top1 inhibitor. The same approach can be envisioned to aid the development of additional Top1 inhibitors.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The Visiting Scientist position granted to F-M.S. by the National Institutes of Health (US Government) and professional leave granted to F-M.S. by Professor CM Che and the University of Hong Kong are gratefully acknowledged. This study used the high-performance computational capabilities of the Biowulf Linux cluster at the National Institutes of Health, Bethesda, Maryland (http://biowulf.nih.gov) and the high-performance computational capabilities of the Gridpoint Linux cluster at the University of Hong Kong, and the National Supercomputing Center at Shenzhen.

FUNDING

Center of Cancer Research, National Cancer Institute Intramural Program at the National Institutes of Health (US Government); National Science Foundation of China (NSFC), [21203237] and Shenzhen Basic Research Grant; [JCYJ20130401170306875] seed funding from the Committee on Research and Conference Grants (CRCR) of the University of Hong Kong; the Areas of Excellence Scheme administrated by the University Grants Committee of the Hong Kong Special Administrative Region, China [AoE/P-10/01], Innovation and Technology Fund [ITF-Tier 2, ITS/134/09FP]; National Key Basic Research Program of China [2013CB834802]; Special Equipment Grant [SEG HKU09]; University Development Fund and Strategic Research Fund. Funding for open access charge: Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences; NSFC [21203237]; and Shenzhen Basic Research Grant [JCYJ20130401170306875].

Conflict of interest statement. None declared.

REFERENCES

1. Pommier,Y. and Marchand,C. (2012) Interfacial inhibitors: targeting macromolecular complexes. *Nat. Rev. Drug. Discov.*, **11**, 25–36.
2. Pommier,Y., Leo,E., Zhang,H. and Marchand,C. (2010) DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chem. Biol.*, **17**, 421–433.
3. Antony,S., Agama,K.K., Miao,Z.H., Takagi,K., Wright,M.H., Robles,A.I., Varticovski,L., Nagarajan,M., Morrell,A., Cushman,M. *et al.* (2007) Novel indenoisoquinolines NSC 725776 and NSC 724998 produce persistent topoisomerase I cleavage complexes and overcome multidrug resistance. *Cancer Res.*, **67**, 10397–10405.
4. Wang,X., Zhou,X. and Hecht,S.M. (1999) Role of the 20-hydroxyl group in camptothecin binding by the topoisomerase I-DNA binary complex. *Biochemistry*, **38**, 4374–4381.
5. Oussedik,K., Francois,J.C., Halby,L., Senamaud-Beaufort,C., Toutirais,G., Dallavalle,S., Pommier,Y., Pisano,C. and Arimondo,P.B. (2010) Sequence-specific targeting of IGF-I and IGF-IR genes by camptothecins. *FASEB J.*, **24**, 2235–2244.
6. Arimondo,P.B., Laco,G.S., Thomas,C.J., Halby,L., Pez,D., Schmitt,P., Boutorine,A., Garestier,T., Pommier,Y., Hecht,S.M. *et al.* (2005) Activation of camptothecin derivatives by conjugation to triple helix-forming oligonucleotides. *Biochemistry*, **44**, 4171–4180.
7. Baranello,L., Bertozzi,D., Fogli,M.V., Pommier,Y. and Capranico,G. (2010) DNA topoisomerase I inhibition by camptothecin induces escape of RNA polymerase II from

promoter-proximal pause site, antisense transcription and histone acetylation at the human HIF-1α gene locus. *Nucleic Acids Res.*, **38**, 159–171.
8. Vekhoff,P., Duca,M., Guianvarc'h,D., Benhida,R. and Arimondo,P.B. (2012) Sequence-specific base pair mimics are efficient topoisomerase IB inhibitors. *Biochemistry*, **51**, 43–51.
9. Jaxel,C., Capranico,G., Kerrigan,D., Kohn,K.W. and Pommier,Y. (1991) Effect of local DNA sequence on topoisomerase I cleavage in the presence or absence of camptothecin. *J. Bio. Chem.*, **266**, 20418–20423.
10. Jaxel,C., Kohn,K.W. and Pommier,Y. (1988) Topoisomerase I interaction with SV40 DNA in the presence and absence of camptothecin. *Nucleic Acids Res.*, **16**, 11157–11170.
11. Kjeldsen,E., Mollerup,S., Thomsen,B., Bonven,B.J., Bolund,L. and Westergaard,O. (1988) Sequence-dependent effect of camptothecin on human topoisomerase I DNA cleavage. *J. Mol. Biol.*, **202**, 333–342.
12. Porter,S.E. and Champoux,J.J. (1989) The basis for camptothecin enhancement of DNA breakage by eukaryotic topoisomerase I. *Nucleic Acids Res.*, **17**, 8521–8532.
13. Solier,S., Barb,J., Zeeberg,B.R., Varma,S., Ryan,M.C., Kohn,K.W., Weinstein,J.N., Munson,P.J. and Pommier,Y. (2010) Genome-wide analysis of novel splice variants induced by topoisomerase I poisoning shows preferential occurrence in genes encoding splicing factors. *Cancer Res.*, **70**, 8055–8065.
14. Capranico,G., Marinello,J. and Baranello,L. (2010) Dissecting the transcriptional functions of human DNA topoisomerase I by selective inhibitors: implications for physiological and therapeutic modulation of enzyme activity. *Biochim. Biophys. Acta*, **1806**, 240–250.
15. Roux,B. (1995) The calculation of the potential of mean force using computer simulations. *Comp. Phys. Commun.*, **91**, 275–282.
16. Staker,B.L., Hjerrild,K., Feese,M.D., Behnke,C.A., Burgin,A.B. Jr and Stewart,L. (2002) The mechanism of topoisomerase I poisoning by a camptothecin analog. *Proc. Natl Acad. Sci. USA*, **99**, 15387–15392.
17. Staker,B.L., Feese,M.D., Cushman,M., Pommier,Y., Zembower,D., Stewart,L. and Burgin,A.B. (2005) Structures of three classes of anticancer agents bound to the human topoisomerase I-DNA covalent complex. *J. Med. Chem.*, **48**, 2336–2345.
18. Ivanova,B. and Spiteller,M. (2012) Structure and properties of camptothecin derivatives, their protonated forms, and model interaction with the topoisomerase I-DNA complex. *Biopolymers*, **97**, 134–144.
19. Laco,G.S. (2011) Evaluation of two models for human topoisomerase I interaction with dsDNA and camptothecin derivatives. *PLoS One*, **6**, e24314.
20. Drwal,M.N., Agama,K., Wakelin,L.P., Pommier,Y. and Griffith,R. (2011) Exploring DNA topoisomerase I ligand space in search of novel anticancer agents. *PLoS One*, **6**, e25150.
21. Lauro,G., Romano,A., Riccio,R. and Bifulco,G. (2011) Inverse virtual screening of antitumor targets: pilot study on a small database of natural bioactive compounds. *J. Nat. Prod.*, **74**, 1401–1407.
22. Lauria,A., Ippolito,M. and Almerico,A.M. (2007) Molecular docking approach on the Topoisomerase I inhibitors series included in the NCI anti-cancer agents mechanism database. *J. Mol. Model.*, **13**, 393–400.
23. Chillemi,G., Coletta,A., Mancini,G., Sanna,N. and Desideri,A. (2010) An amber compatible molecular mechanics force field for the anticancer drug topotecan. *Theor. Chem. Acc.*, **127**, 293–302.
24. Mancini,G., D'Annessa,I., Coletta,A., Sanna,N., Chillemi,G. and Desideri,A. (2010) Structural and dynamical effects induced by the anticancer drug topotecan on the human topoisomerase I - DNA complex. *PLoS One*, **5**, e10934.
25. Sanna,N., Chillemi,G., Grandi,A., Castelli,S., Desideri,A. and Barone,V. (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. *J. Am. Chem. Soc.*, **127**, 15429–15436.
26. Case,D.A., Cheatham,T.E. III, Darden,T., Gohlke,H., Luo,R., Merz,K.M. Jr, Onufriev,A., Simmerling,C., Wang,B. and Woods,R.J. (2005) The Amber biomolecular simulation programs. *J. Comput. Chem.*, **26**, 1668–1688.
27. Perez,A., Marchan,I., Svozil,D., Sponer,J., Cheatham,T.E. III, Laughton,C.A. and Orozco,M. (2007) Refinement of the AMBER force field for nucleic acids: improving the description of α/γ conformers. *Biophys. J.*, **92**, 3817–3829.
28. Perez,A., Lankas,F., Luque,F.J. and Orozco,M. (2008) Towards a molecular dynamics consensus view of B-DNA flexibility. *Nucleic Acids Res.*, **36**, 2379–2394.
29. Cao,Z.W., Xue,Y., Han,L.Y., Xie,B., Zhou,H., Zheng,C.J., Lin,H.H. and Chen,Y.Z. (2004) MoViES: molecular vibrations evaluation server for analysis of fluctuational dynamics of proteins and nucleic acids. *Nucleic Acids Res.*, **32**, W679–W685.
30. Maffeo,C., Luan,B. and Aksimentiev,A. (2012) End-to-end attraction of duplex DNA. *Nucleic Acids Res.*, **40**, 3812–3821.
31. Darden,T., York,D. and Pedersen,L.G. (1993) Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.*, **98**, 10089–10092.
32. Ryckaert,J.P., Ciccotti,G. and Berendsen,H.J.C. (1977) Numerical-integration of cartesian equations of motion of a system with constraints: molecular dynamics of N-alkanes. *J. Comp. Phys.*, **23**, 327–341.
33. Chu,W.T., Zhang,J.L., Zheng,Q.C., Chen,L. and Zhang,H.X. (2013) Insights into the folding and unfolding processes of wild-type and mutated SH3 domain by molecular dynamics and replica exchange molecular dynamics simulations. *PLoS One*, **8**, e64886.
34. Humphrey,W., Dalke,A. and Schulten,K. (1996) VMD: visual molecular dynamics. *J. Mol. Graph.*, **14**, 33–38, 27–38.
35. Hess,B., Kutzner,C., van der Spoel,D. and Lindahl,E. (2008) GROMACS 4: algorithms for high efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.*, **4**, 435–447.
36. Hess,B. (2002) Convergence of sampling in protein simulations. *Phys. Rev. E*, **65**, 031910.
37. Chillemi,G., D'Annessa,I., Fiorani,P., Losasso,C., Benedetti,P. and Desideri,A. (2008) Thr729 in human topoisomerase I modulates anti-cancer drug resistance by altering protein domain communications as suggested by molecular dynamics simulations. *Nucleic Acids Res.*, **36**, 5645–5651.
38. Onufriev,A., Bashford,D. and Case,D.A. (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins*, **55**, 383–394.
39. Gohlke,H., Kiel,C. and Case,D.A. (2003) Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J. Mol. Biol.*, **330**, 891–913.
40. Hou,T., Zhang,W., Case,D.A. and Wang,W. (2008) Characterization of domain-peptide interaction interface: a case study on the amphiphysin-1 SH3 domain. *J. Mol. Biol.*, **376**, 1201–1214.
41. Balius,T.E. and Rizzo,R.C. (2009) Quantitative prediction of fold resistance for inhibitors of EGFR. *Biochemistry*, **48**, 8435–8448.
42. McGillick,B.E., Balius,T.E., Mukherjee,S. and Rizzo,R.C. (2010) Origins of resistance to the HIVgp41 viral entry inhibitor T20. *Biochemistry*, **49**, 3575–3592.
43. Zoete,V. and Michielin,O. (2007) Comparison between computational alanine scanning and per-residue binding free energy decomposition for protein-protein association using MM-GBSA: application to the TCR-p-MHC complex. *Proteins*, **67**, 1026–1047.
44. Lee,V.S., Tue-ngeun,P., Nangola,S., Kitidee,K., Jitonnom,J., Nimmanpipug,P., Jiranusornkul,S. and Tayapiwatana,C. (2010) Pairwise decomposition of residue interaction energies of single chain Fv with HIV-1 p17 epitope variants. *Mol. Immunol.*, **47**, 982–990.
45. Grossfield,A. *WHAM: The Weighted Histogram Analysis Method, version 2.0.7*.
46. Xiao,X. and Cushman,M. (2005) An ab initio quantum mechanics calculation that correlates with ligand orientation and DNA cleavage site selectivity in camptothecin-DNA-topoisomerase I ternary cleavage complexes. *J. Am. Chem. Soc.*, **127**, 9960–9961.

47. Siu,F.M. and Che,C.M. (2008) Persistence of camptothecin analog-topoisomerase I-DNA ternary complexes: a molecular dynamics study. *J. Am. Chem. Soc.*, **130**, 17928–17937.
48. Redinbo,M.R., Stewart,L., Champoux,J.J. and Hol,W.G. (1999) Structural flexibility in human topoisomerase I revealed in multiple non-isomorphous crystal structures. *J. Mol. Biol.*, **292**, 685–696.
49. Pommier,Y., Pourquier,P., Urasaki,Y., Wu,J. and Laco,G.S. (1999) Topoisomerase I inhibitors: selectivity and cellular resistance. *Drug Resist. Updat.*, **2**, 307–318.
50. Redinbo,M.R., Stewart,L., Kuhn,P., Champoux,J.J. and Hol,W.G. (1998) Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA. *Science*, **279**, 1504–1513.
51. Wang,L.F., Ting,C.Y., Lo,C.K., Su,J.S., Mickley,L.A., Fojo,A.T., Whang-Peng,J. and Hwang,J. (1997) Identification of mutations at DNA topoisomerase I responsible for camptothecin resistance. *Cancer Res.*, **57**, 1516–1522.
52. Fujimori,A., Harker,W.G., Kohlhagen,G., Hoki,Y. and Pommier,Y. (1995) Mutation at the catalytic site of topoisomerase I in CEM/C2, a human leukemia cell line resistant to camptothecin. *Cancer Res.*, **55**, 1339–1346.
53. Sirikantaramas,S., Yamazaki,M. and Saito,K. (2008) Mutations in topoisomerase I as a self-resistance mechanism coevolved with the production of the anticancer alkaloid camptothecin in plants. *Proc. Natl Acad. Sci. USA*, **105**, 6782–6786.
54. Fiorani,P., Tesauro,C., Mancini,G., Chillemi,G., D'Annessa,I., Graziani,G., Tentori,L., Muzi,A. and Desideri,A. (2009) Evidence of the crucial role of the linker domain on the catalytic activity of human topoisomerase I by experimental and simulative characterization of the Lys681Ala mutant. *Nucleic Acids Res.*, **37**, 6849–6858.
55. Chillemi,G., Fiorani,P., Castelli,S., Bruselles,A., Benedetti,P. and Desideri,A. (2005) Effect on DNA relaxation of the single Thr718Ala mutation in human topoisomerase I: a functional and molecular dynamics study. *Nucleic Acids Res.*, **33**, 3339–3350.
56. Fiorani,P., Bruselles,A., Falconi,M., Chillemi,G., Desideri,A. and Benedetti,P. (2003) Single mutation in the linker domain confers protein flexibility and camptothecin resistance to human topoisomerase I. *J. Bio. Chem.*, **278**, 43268–43275.
57. Scaldaferro,S., Tinelli,S., Borgnetto,M.E., Azzini,A. and Capranico,G. (2001) Directed evolution to increase camptothecin sensitivity of human DNA topoisomerase I. *Chem. Biol.*, **8**, 871–881.
58. Stewart,L., Ireton,G.C. and Champoux,J.J. (1997) Reconstitution of human topoisomerase I by fragment complementation. *J. Mol. Biol.*, **269**, 355–372.
59. Pommier,Y. and Cushman,M. (2009) The indenoisoquinoline noncamptothecin topoisomerase I inhibitors: update and perspectives. *Mol. Cancer Ther.*, **8**, 1008–1014.
60. Fukasawa,K., Komatani,H., Hara,Y., Suda,H., Okura,A., Nishimura,S. and Yoshinari,T. (1998) Sequence-selective DNA cleavage by a topoisomerase I poison, NB-506. *Int. J. Cancer*, **75**, 145–150.
61. Teicher,B.A. (2008) Next generation topoisomerase I inhibitors: Rationale and biomarker strategies. *Biochem. Pharm.*, **75**, 1262–1271.
62. Marco,E., Laine,W., Tardy,C., Lansiaux,A., Iwao,M., Ishibashi,F., Bailly,C. and Gago,F. (2005) Molecular determinants of topoisomerase I poisoning by lamellarins: comparison with camptothecin and structure-activity relationships. *J. Med. Chem.*, **48**, 3796–3807.
63. Antony,S., Jayaraman,M., Laco,G., Kohlhagen,G., Kohn,K.W., Cushman,M. and Pommier,Y. (2003) Differential induction of topoisomerase I-DNA cleavage complexes by the indenoisoquinoline MJ-III-65 (NSC 706744) and camptothecin: base sequence analysis and activity against camptothecin-resistant topoisomerases I. *Cancer Res.*, **63**, 7428–7435.
64. Stewart,L., Ireton,G.C. and Champoux,J.J. (1999) A functional linker in human topoisomerase I is required for maximum sensitivity to camptothecin in a DNA relaxation assay. *J. Bio. Chem.*, **274**, 32950–32960.
65. Losasso,C., Cretaio,E., Palle,K., Pattarello,L., Bjornsti,M.A. and Benedetti,P. (2007) Alterations in linker flexibility suppress DNA topoisomerase I mutant-induced cell lethality. *J. Bio. Chem.*, **282**, 9855–9864.
66. Losasso,C., Cretaio,E., Fiorani,P., D'Annessa,I., Chillemi,G. and Benedetti,P. (2008) A single mutation in the 729 residue modulates human DNA topoisomerase IB DNA binding and drug resistance. *Nucleic Acids Res.*, **36**, 5635–5644.
67. Gongora,C., Vezzio-Vie,N., Tuduri,S., Denis,V., Causse,A., Auzanneau,C., Collod-Beroud,G., Coquelle,A., Pasero,P., Pourquier,P. *et al.* (2011) New Topoisomerase I mutations are associated with resistance to camptothecin. *Mol. Cancer*, **10**, 64.
68. Laco,G.S. and Pommier,Y. (2008) Role of a tryptophan anchor in human topoisomerase I structure, function and inhibition. *Biochem. J.*, **411**, 523–530.
69. Frohlich,R.F., Andersen,F.F., Westergaard,O., Andersen,A.H. and Knudsen,B.R. (2004) Regions within the N-terminal domain of human topoisomerase I exert important functions during strand rotation and DNA binding. *J. Mol. Biol.*, **336**, 93–103.
70. Vassallo,O., Castelli,S., D'Annessa,I., della Rocca,B.M., Stella,L., Knudsen,B.R. and Desideri,A. (2011) Evidences of a natively unfolded state for the human topoisomerase IB N-terminal domain. *Amino Acids*, **41**, 945–953.
71. Chillemi,G., Redinbo,M., Bruselles,A. and Desideri,A. (2004) Role of the linker domain and the 203-214 N-terminal residues in the human topoisomerase I DNA complex dynamics. *Biophys J.*, **87**, 4087–4097.
72. Christensen,M.O., Barthelmes,H.U., Boege,F. and Mielke,C. (2003) Residues 190-210 of human topoisomerase I are required for enzyme activity in vivo but not in vitro. *Nucleic Acids Res.*, **31**, 7255–7263.
73. Redinbo,M.R., Champoux,J.J. and Hol,W.G. (2000) Novel insights into catalytic mechanism from a crystal structure of human topoisomerase I in complex with DNA. *Biochemistry*, **39**, 6832–6840.
74. Pond,C.D., Holden,J.A., Schnabel,P.C. and Barrows,L.R. (1997) Surface plasmon resonance analysis of topoisomerase I-DNA binding: effect of Mg^{2+} and DNA sequence. *Anticancer Drugs*, **8**, 336–344.
75. Koster,D.A., Czerwinski,F., Halby,L., Crut,A., Vekhoff,P., Palle,K., Arimondo,P.B. and Dekker,N.H. (2008) Single-molecule observations of topotecan-mediated TopIB activity at a unique DNA sequence. *Nucleic Acids Res.*, **36**, 2301–2310.
76. Tanizawa,A., Kohn,K.W., Kohlhagen,G., Leteurtre,F. and Pommier,Y. (1995) Differential stabilization of eukaryotic DNA topoisomerase I cleavable complexes by camptothecin derivatives. *Biochemistry*, **34**, 7200–7206.
77. Kjeldsen,E., Svejstrup,J.Q., Gromova,I.I., Alsner,J. and Westergaard,O. (1992) Camptothecin inhibits both the cleavage and religation reactions of eukaryotic DNA topoisomerase I.J. *Mol. Biol.*, **228**, 1025–1030.
78. Tanizawa,A., Fujimori,A., Fujimori,Y. and Pommier,Y. (1994) Comparison of topoisomerase I inhibition, DNA damage, and cytotoxicity of camptothecin derivatives presently in clinical trials. *J. Natl Cancer Inst.*, **86**, 836–842.