CHEMISTRY
A EUROPEAN JOURNAL

DOI: 10.1002/chem.200700732

Binding of Topotecan to a Nicked DNA Oligomer in Solution

W. Bocian,*[a] R. Kawęcki,[b] E. Bednarek,[a] J. Sitkowski,[a, b] M. P. Williamson,[c] P. E. Hansen,[d] and L. Kozerski*[a, b]

Abstract: Topotecan (TPT) is in clinical use as an antitumor agent. It acts by binding to the covalent complex formed between nicked DNA and topoisomerase I, and inserts itself into the single-strand nick, thereby inhibiting the religation of the nick and acting as a poison. A crystal structure analysis of the ternary complex has shown how the drug binds (B. L. Staker, K. Hjerrild, M. D. Feese, C. A. Behnke, A. B. Burgin, L. Stewart, *Proc. Natl. Acad. Sci. U.S.A.*, **2002**, *99,* 15387–15392), but has left a number of unanswered questions. Herein, we use NMR spectroscopy and molecular modeling to show that the solution structure of a complex of TPT with nicked natural DNA is similar, but not identical to the crystal conformation, and that other geometries are of very low population. We also show that the lactone form of TPT binds approximately 40 times more strongly than the ring-opened carboxylate.

Keywords: alkaloids · antitumor agents · DNA binding · molecular dynamics · NMR spectroscopy

Introduction

Topoisomerases are essential enzymes that relieve the torsional stress produced in DNA during replication by cutting the DNA on one strand and thus allowing the broken strand to rotate around the uncut strand.[1,2] In human topoisomerase I (topI) the reaction is mediated by Tyr723, which attacks the phosphate backbone to create a 3′-phosphotyrosine.[3] After rotation of the cut strand, the DNA is re-ligated by attack of the liberated 5′-OH on the phosphotyrosine.[2] This re-ligation step is normally much faster than the initial cleavage, which implies that the concentration of nicked DNA is very low. A variety of small aromatic molecules, such as the natural product camptothecin (CPT) or its derivative topotecan (TPT), can intercalate into the nick site generated by topI and stabilize the DNA–topI complex.[4–7] This leads to cell death, which explains why TPT and its analogues are typically described as poisons as they convert an essential enzyme into something that damages DNA. The molecular mechanism for cell damage is thought to arise from the generation of double-strand breaks at the intercalation site in rapidly dividing cells as a result of stalling the replication complex.[8–11] Such DNA poisons show great promise as drugs against solid tumors, of which TPT (Hycamptin) and irinotecan (Camptosaur) are currently in clinical use.[12–16] This has led to considerable effort in developing other analogues as improved drugs.[17–22] The aim of this paper is to provide a greater understanding of the bind-

[a] Dr. W. Bocian, Dr. E. Bednarek, J. Sitkowski, Prof. Dr. L. Kozerski
National Medicines Institute
00-725 Warszawa, Chełmska 30/34 (Poland)
Fax: (+48) 22-8410652
E-mail: lkoz@icho.edu.pl

[b] Dr. R. Kawęcki, J. Sitkowski, Prof. Dr. L. Kozerski
Institute of Organic Chemistry, Polish Academy of Sciences
01-224 Warszawa, Kasprzaka 44 (Poland)
Fax: (+48) 22-6326681

[c] Prof. Dr. M. P. Williamson
Department of Molecular Biology and Biotechnology
University of Sheffield, Sheffield S10 2TN (UK)
Fax: (+44) 114-222-2800
E-mail: M.Williamson@sheffield.ac.uk

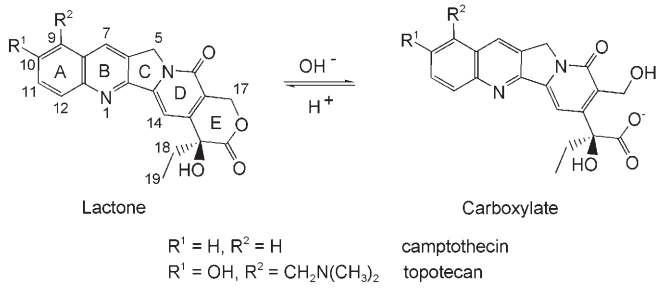
[d] Prof. Dr. P. E. Hansen
Department of Life Sciences and Chemistry, Roskilde University
DK-4000 Roskilde (Denmark)
Fax: (+45)-46743011
E-mail: poulerik@ruc.dk

Supporting information for this article is available on the WWW under http://www.chemeurj.org or from the author. It contains details of ^{1}H NMR chemical shifts of nicked decamer–TPT and of a nicked decamer tethered to PEG_6, NOE analyses, free-energy analysis for DNA–TPT complexes, comparison of structural parameters derived from NMR spectroscopy and crystal structures, starting structures for MD calculations, chemical shift changes for nicked DNA on interaction with TPT, experimental NOESY spectrum, aromatic expansion of NOESY spectra of neat decamer and decamer titrated with TPT, minor-groove views of the eight MD-derived structures of nicked decamer–TPT complexes.

ing of TPT to nicked DNA to assist in designing improved drugs.

Recent crystal structures of the ternary top1–DNA–CPT and top1–DNA–TPT complexes[6,8] have thrown light on the way these molecules bind to nicked DNA. In the crystal structures, both CPT and TPT intercalate into the nick site as if they were normal base pairs; this widens the interbase distance by approximately 3.6 Å and leaves the DNA helix essentially linear. The major interaction is hydrophobic ring-stacking. There are no hydrogen bonds to the DNA in the crystal structures, but there are a small number of nonconserved hydrogen bonds to topI. Specifically, one hydrogen bond between the 20(*S*)-hydroxy group of TPT ring E and Asp533 (in both the lactone and carboxylate forms, see below), and probably the same hydrogen bond in the CPT complex as well as between N1 of CPT and Arg364. In an indenoisoquinoline complex there is a probable bifurcated hydrogen bond to Arg364 from a carbonyl group for which no equivalent functional group exists in CPT. Thus, there appears to be no hydrogen-bonding interactions either with DNA or with topI that are key to determining specificity. This observation highlights the need to investigate the complexes in more detail.

TPT and CPT exist in an equilibrium between the lactone and the ring-opened carboxylate form depending on the pH (Scheme 1), at pH 7.4 $>80\%$ is the carboxylate.[23–26] Most



Scheme 1.

biological assays indicate that the carboxylate form of TPT is not active as a topI inhibitor. However, both crystal structures reveal the presence of both forms in the crystal, apparently in approximately the same ratio as seen in solution. It was not possible to determine the relative affinities of the two forms, although NMR studies suggest little if any binding of the carboxylate form.[27] Modeling studies of DNA interactions with the CPT family have been published, but throw little light on this problem.[28–32] Therefore, there is a question as to whether the bound carboxylate form seen in the crystal is due to the carboxylate binding to the ternary complex with the same affinity as the lactone or is due to hydrolysis of the lactone, possibly when it is already bound into the DNA nick site. This is an important question for the design of compounds with improved efficacy.

The crystal structures have greatly advanced our understanding of the binding of the camptothecin family to nicked DNA. However, there are still some important details that need to be addressed. The first is whether CPT and TPT can bind in other places and/or orientations to those seen in the crystal structures. Only one bound conformation is seen in the crystal structures, but this does not preclude binding in other modes because it may be that the form seen is the only one that crystallizes. As an example of this problem, many ligand complexes with dihydrofolate reductase have been shown to adopt multiple ligand conformations in solution, although in crystal structures generally only a single form is observed, which is not always the physiologically important form.[33] The second question is the relative affinity of the carboxylate and lactone to nicked DNA, which is an important question for the design of effective analogues for use as drugs. Herein, we have addressed these questions by using NMR spectroscopy. We have used a stable synthetic nicked decamer GCGTT↓GTCGC,[34] which we have previously shown to adopt a standard B-DNA conformation. We show, by using a method developed previously by us,[35] that there is only one significant binding mode and that it is similar to that observed in the crystal structures, although the conformation is not identical. Importantly, the method does not assume or require a single conformation and thus offers a relatively unbiased search of conformational space. We also show that the carboxylate form binds to nicked DNA approximately 40 times more weakly than the lactone, and hence, is probably not biologically significant, even when bound to the ternary complex.

Results and Discussion

The crystal structures of complexes of topI–DNA with CPT and its analogues all show the drugs intercalated and binding in the same orientation. However, there are four possible coplanar stacking orientations, and the fact that only one orientation was seen in the crystal structures does not necessarily mean that only one orientation is possible or biologically relevant. Moreover, there have been experimental studies demonstrating minor groove binding. It is therefore important to check explicitly for minor populations of alternative conformations. The standard restrained molecular dynamics/simulated annealing method for NMR structure calculation is a minimization towards a single optimum and is inherently poor at identifying minor alternative conformers. We have therefore taken a different approach here; we have used modeling to calculate the structures of the four possible conformations and compared these with experimental measurements to allow the identification of minor conformers, if present.

The other question addressed herein is the binding affinity of the lactone and carboxylate forms of TPT, which was investigated by measuring the diffusion rate of TPT in the absence and presence of DNA. We discuss this question first.

Nicked DNA decamer binding by TPT lactone and carboxylate: In the presence of nicked DNA, TPT is in equilibrium between free and DNA-bound forms. Its diffusion coefficient is therefore a weighted average of free and bound forms, which permits calculation of the binding affinity.[36] This analysis is complicated by the self-association of TPT in solution, which also affects the diffusion coefficient[35] and implies that it is first necessary to measure the self-association; this can also be carried out by using diffusion coefficients (Table 1). Because the lactone and carboxylate forms have different NMR signals (e.g., for the C-20 methyl group), we have been able to determine the binding affinities of both forms from the same solution by using PFG (pulsed field gradient) experiments performed at pH 7.05 at which both forms are significantly populated. By measuring the binding affinities of both forms in the same solution, we are effectively using the best possible internal control on the accuracy of the results. The relative diffusion coefficients for the methyl groups of the two forms are shown in Figure 1. This experiment yielded values for K_a of 3.78 and 0.1 mM^{-1} for the lactone and carboxylate forms, respectively (Table 1). This demonstrates that binding of the carboxylate is very weak, in line with its lack of biological activity, but in contrast to conclusions from the ternary complex crystal structure, which showed that the carboxylate form can be equally well-accommodated in an intercalation pocket in an enzymically prepared nick structure.

We have previously measured the affinity of TPT for the DNA octamer $d(GCGATCGC)_2$; TPT binds to the ends of the octamer with a K_a value of 1.5 mM^{-1}.[35] Thus, TPT binds more strongly to nicked DNA than to terminal base pairs, but only by a factor of 2.5. Our nicked decamer is held together by poly(ethylene glycol) linkers attached as a loop to both ends of the stem that holds the ends together. It therefore does not contain readily available terminal base pairs although it is possible to envisage binding of TPT onto the ends of the stem. However, in practice no such binding is observed, as demonstrated below.

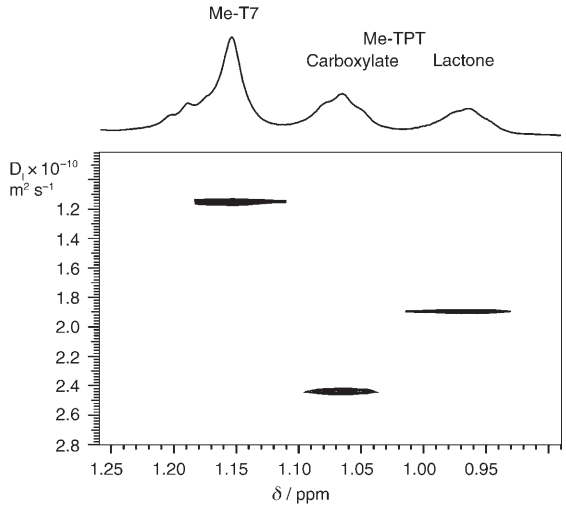


Figure 1. A DOSY spectrum showing diffusion coefficient versus methyl chemical shift.[37] A weaker interaction of the TPT-carboxylate form with the nicked DNA decamer (sample **2a**) is revealed compared with the TPT-lactone form.

Binding of TPT to DNA affects the kinetics and energetics of the ring-opening reaction. In free aqueous solution, the lactone/carboxylate mixture reached equilibrium within 2 h at pH 7 with a ratio of 20:80 lactone/carboxylate. However, in the presence of DNA at pH 7, equilibrium was only reached after 24 h and the lactone/carboxylate ratio was 1:1. Therefore, DNA binding slows down the interconversion. This result is not surprising because DNA binding reduces the exposure of the lactone ring to the solvent. The change in the lactone/carboxylate ratio demonstrates that binding to DNA favors the lactone form. This is consistent with the lactone being the preferred bound form of TPT.[38]

Figures S1 and S2 of the Supporting Information show regiospecific changes to the chemical shifts induced by TPT on interaction with DNA and some intermolecular NOE cross-peaks for the complex, respectively. Table S3 in the Supporting Information comprises the complete experimental NOE data along with back-calculated values for eight computed conformations of the complex. Tables S4 and S5 in the Supporting Information summarize the results of the NOE analysis, averaging over 800 structures within a 9 ns molecular dynamics (MD) run. These data demonstrate that TPT binds at the nick site rather than in the minor groove,[39] the major groove, or at the ends of the DNA stem. However, if we assume that TPT may adopt multiple con-

Table 1. PGSE data for the binding of TPT to nicked decamer DNA duplex.[a]

Sample	pH	Concentration [mM]	D_i (TPT) (free) [10^{-10} m^2s^{-1}]	D_i (TPT) (complexed) [10^{-10} m^2s^{-1}]	D_i(DNA) [10^{-10} m^2s^{-1}]	K_a [mM^{-1}]	Signal observed
2[b]	5.95	C_{DNA}, 0.96	–	–	1.13	–	T-7 CH_3
		C_{TPT}, 0.96	3.05 ± 0.04	1.68 ± 0.01	–	3.78 ± 0.1	C-20 CH_3
2a[c]	7.01	C_{DNA}, 0.96	–	–	1.18	–	T-7 CH_3
		C_{TPT}, 0.96	3.05 ± 0.04	2.39 ± 0.01	–	0.1 ± 0.01[d]	C-20 CH_3
TPT	5.0	C_{TPT}, 0.10	4.07 ± 0.13	–	–	–	C-20 CH_3
TPT	6.0	C_{TPT}, 0.64	3.43 ± 0.13	–	–	–	C-20 CH_3
TPT	5.9	C_{TPT}, 0.96	3.05 ± 0.04	–	–	–	C-20 CH_3

[a] D_i denotes the diffusion coefficient, K_a is the affinity constant. [b] TPT present in solution only in the lactone form. [c] TPT present in solution in both carboxylate and lactone forms. [d] Carboxylate form in the presence of the lactone.

formations in fast exchange, the data do not unambiguously show which sets of conformations are allowed. We therefore used molecular modeling to address this question.

Molecular modeling: Having established that TPT binds exclusively in the nick site, we used molecular modeling to generate a set of low-energy conformers and then compared these with the experimental data to test for the presence of alternative conformers. Because MD is good at producing low-energy conformations, but poor at comparing their absolute energies, we do not have great confidence in the energies generated from the MD analysis. Therefore, our methodology, which builds on our earlier work on the analysis of multiple conformations,[35,40,41] takes all of the conformations within a wide energy range of the lowest-energy conformer and searches for any conformers or sets of conformers that are consistent with the experimental data. Eight conformational families were investigated that consist of four pairs of structures (**A** and **B**) that differ only in the orientation of the CH_2NMe_2 group. Each conformation was modeled over a long simulation time in explicit water after an extended equilibration period, and the averaged conformation was calculated. It turned out that during a long simulation time the TPT in complexes **St2A**, **St2B**, **St3A**, and **St3B** moved into the minor groove, whereas the other structures remained intercalative. The free energies and entropies of these structures are listed in Table S6 in the Supporting Information, from which it can be seen that structures **St1** can probably be rejected because of their high energy, but the other three pairs are energetically similar. It is thus clear that molecular modeling alone cannot distinguish between correct and incorrect conformers, in agreement with earlier results.[30,31] We have therefore used experimental NOEs and chemical shifts to eliminate incorrect conformers. Tables S4 and 5S in the Supporting Information summarize the results of NOE analysis, averaging over 800 structures within a 9 ns MD run, and clearly show that only the conformational family **St4** displays acceptable agreement between calculated and experimental NOEs, with better agreement for the structure **St4A**. We calculated a semi-quantitative NOE *R* factor for each of the structures **St1A**, **St1B**, **St2A**, **St2B**, **St3A**, **St3B**, **St4A**, and **St4B** against the experimental intermolecular NOEs, which are 55, 53, 47, 45, 47, 47, 3, and 16%, respectively. This clearly demonstrates that **St4A** gives the best fit to the experimental NOE intensities, **St4B** is acceptable, and the others deviate too widely to be acceptable. Significantly, inclusion of other structures into an ensemble with **St4** did not improve the fit. In other words, the NOE data demonstrate that conformations **St1**, **St2**, and **St3** are not populated in solution to any significant extent. Based on experience with other systems[40–42] we can be confident that no more than 10% of other structures are present and probably much less, which implies an energy penalty of at least 1.3 kcal mol^{-1} against other structures being present in the complex. The absence of the minor-groove conformations **St2** and **St3** was also demonstrated by the chemical shifts. The averaged structures of **St4A** and **St4B** are shown in Figure 2, and all eight structures are shown in Figure S4 in the Supporting Information.

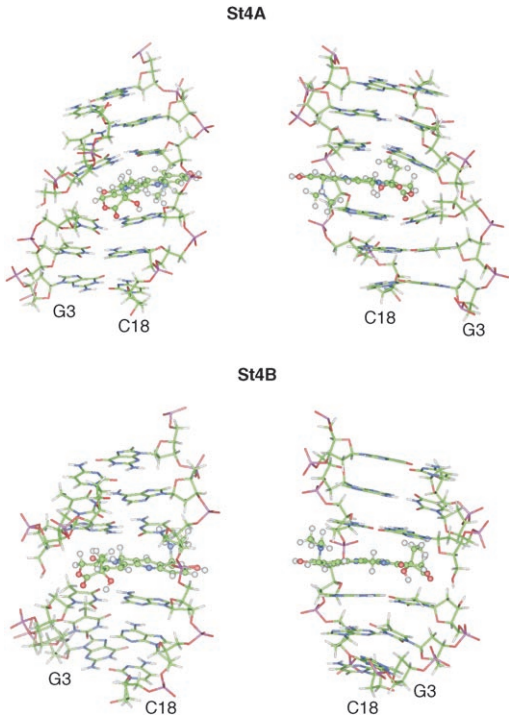


Figure 2. Views of the minor grooves of the MD-derived structures of the nicked decamer–TPT complexes that fit best the experimental NOEs. Only six base-pair units flanking the nick are shown for clarity. Each structure represents the minimized average of structures from the last 1 ns out of a 10 ns MD run.

The NOEs and chemical shifts therefore demonstrate that the intercalated structure **St4** is the only plausible conformer in solution. NOEs and chemical shifts are thus useful in restricting the conformational space explored by molecular modeling.

Comparison with the X-ray structure: The calculations reported above demonstrate that the experimental parameters are consistent with **St4A** being the only significantly populated conformer in solution. The geometry of **St4A** is also similar to that seen in the X-ray structures of ternary complexes.[6,8] This is confirmed by the good agreement between NOEs back-calculated from the X-ray structure and the experimental NOEs (Table S4). However, **St4A** is not identical to the crystal structure (Table S7). The NOE *R* factor for the crystal structure is 22%, which demonstrates reasonable agreement with the experimental data, but implies that the solution structure of the binary complex is better matched by **St4A** than by the crystal structure. Our analysis of the NOEs also suggests that structure **St4** is a better model for the binary complex than the crystal structure. This may imply some structural rearrangement caused by the presence of topI.

An additional finding of this study concerns the binding of the lactone and carboxylate. The PFG result demonstrates that the carboxylate binds approximately 40 times more weakly than the lactone, in agreement with the biological inactivity of the former. On the other hand, the X-ray ternary structures contain both forms in approximately their solution ratios, although we have been careful not to read too much into this result. The result reinforces the conclusion of the X-ray studies that the major interaction is ring-stacking with few if any structure-specific hydrogen bonds. Independent support for this hypothesis has been furnished by a study on luotonin A, which is a close analogue of campthothecin and has an unsubstituted aromatic E ring that stabilizes the human DNA topI–DNA covalent binary complex and mediates topI-dependent cytotoxicity in intact cells.[19] The authors concluded that even in the absence of the 20(*S*)-hydroxy group, the analogue is able to intercalate into nicked DNA in a similar manner to CPT.

Biological significance: The nicked DNA decamer used in this work is a useful model of the topI–DNA complex. It contains a native-like hydroxy group at the 5′ nick site rather than a thiol. The nick was introduced into the TT↓GT motif, which is commonly considered to be the preferred site of cleavage. Thus, the environment of TPT is similar to that found in vivo, although it lacks the presence of topI. The crystal structures, however, show two direct hydrogen bonds at most between TPT and the protein, and as discussed above, the binding of TPT appears to be dominated by hydrophobic stacking interactions with nicked DNA. Moreover, the structure of DNA in the ternary complex is similar to that found herein, as shown above. It is therefore likely that the binary complex studied herein is similar in geometry to the in vivo ternary complex. Thus, our finding that only structure **St4** is present to any significant extent implies that only this binding mode need be considered in any drug design program as the additional binding energy required to stabilize any other geometry is too large.

Nevertheless, the in vivo complex is different to that studied herein in that it also contains topI. The presence of the enzyme is likely to stabilize the more elongated DNA structure required for the intercalation of TPT and contributes at least one additional hydrogen bond. The energy contributed by a hydrogen bond is generally quoted as being in the range of 2 to 6 kJ mol^{-1} (0.5–1.5 kcal mol^{-1}),[43] although in solvent-exposed situations it is often even lower than this.[44] Thus, the extra stabilization of the TPT complex as a result of the enzyme is likely to be of the order of 0.5 to 1 kcal mol^{-1}, which is the same magnitude as kT (0.6 kcal mol^{-1}). We therefore anticipate that the presence of the enzyme will not have a large effect on the conclusions reached herein.

Our observation of the stabilization by DNA of the lactone form is consistent with earlier observations and with the greater affinity of DNA for the lactone. The slowing down of the lactone/carboxylate exchange kinetics may indicate involvement of the lactone carbonyl in a specific interaction that hinders nucleophilic attack of the carbonyl carbon atom by the bulk solvent. The crystal structure of the TPT ternary complex[6] reveals no direct hydrogen bonds that involve the lactone or carboxy group, although there are two water-mediated interactions, which generally have very little effect on the energy.[44] We therefore suggest that the lactone is likely to bind more tightly in the ternary complex and in the binary complex, which implies that the approximately equal amounts of lactone and carboxylate seen in the crystal structure may arise from slow hydrolysis after binding. The implication for drug design is that the lactone is strongly preferred.

Both X-ray and NMR spectroscopy studies have now demonstrated that TPT intercalates into the nick, essentially lengthening the DNA by one base pair. In the X-ray structure of the ternary complex, there is only one hydrogen bond between the enzyme and TPT, between 20(*S*)-OH and Arg533. Almost all of the interactions with TPT are therefore hydrophobic stacking interactions. The limited number of interactions explains the low binding affinity of TPT for DNA measured herein.

Our complementary studies herein and previously, in which we showed two conformational families in equilibrium for TPT-binding to the uncleaved DNA octamer d-(GCGATCGC)$_2$,[35] indicate that unstrained wild-type DNA is not a target for TPT. Only nicked DNA is a target for TPT. This is confirmed by the fact that TPT binds only to the edge GC base pair in the octamer, which mimics one half of a nick, and not to the stem of the duplex. It is important to point out that both studies show the same geometry for TPT-stacking against a GC base pair. This implies that binding selectivity is a cooperative, complementary property based on the contributions from both partners; that is, a properly prepared nick composed of the most suitable bases for an attractive interaction and relevant pharmacophores in the ligand capable of supplying the complementary functions for binding interactions with DNA.

This paper also demonstrates how ab initio calculations and the measurements of NOEs and chemical shifts can be used to determine the structure of weakly bound ligands to DNA and to distinguish between different binding modes.

Experimental Section

Synthesis and purification of DNA oligomers: The nicked duplex decamer oligonucleotide was synthesized as described earlier[34] on a 1-μmole scale with an ABI 394 DNA synthesizer by using phosphoramidite chemistry starting on T5. A, T, G, and C phosphoramidites were purchased from ABI/Perkin–Elmer, Warrington (UK). The PEG-6 spacer (18-*O*-dimethoxytritylhexaethyleneglycol,1-[(2-cyanoethyl)-(N,N-diisopropyl)]phosphoramidite; (Glen Research, Sterling, Virginia (USA)) was coupled with the oligonucleotide during the normal synthesis cycle between G1–C20 and G11–C10. The oligonucleotide was purified by ion-exchange chromatography on a HiTrap-Q column (Pharmacia Biotech) by using gradient elution with ammonium bicarbonate solution (0.1–0.8 M) and desalted on Sephadex G-10.

Sample preparation: Samples for NMR spectroscopy measurements were prepared by dissolving the nicked decamer in K_3PO_4 or Na_2HPO_4 buffer (38 mM) that contained sodium chloride (38 mM) plus [D$_4$]TSP (TSP = tri-

methylsilylpropionate). The pH was adjusted to 7.05 (sample **1**). Samples in H_2O contained 10% (v/v) D_2O. Samples in both H_2O and D_2O were 0.96 mM in oligonucleotide, as measured by UV spectrophotometry. Titration of the DNA with TPT was performed by using a stock solution of TPT (5 mM) in H_2O (≈pH 4), thus producing a 1:1 TPT/DNA ratio by adding 100 μL of the TPT solution. The titration was performed by adding 10 μL aliquots to allow changes in the DNA signals to be monitored and to permit their assignment. At the start of the titration, TPT was completely in its lactone form and equilibrium with the carboxylate was reached only after 24 h. At the end of the titration, the pH was adjusted to 5.6 (sample **1a**). TPT is more soluble at this pH and the equilibrium was shifted to the lactone form. A sample in D_2O was prepared independently, in the same buffer at pH 5.95 (sample **2**). After acquiring the necessary NMR data the pH was changed to 7.01 (sample **2a**) to enhance the population of the carboxylate form. TPT reference solutions for PFGSE experiments and DNMR studies were prepared in the same buffer and concentrations were checked by UV spectrophotometry. The diffusion experiments were run in D_2O solution (sample **2a**) by using a 170 ms sine-shaped gradient pulse incremented from 1 to 50 G cm^{-1}.

2D NMR spectra: Spectra of samples **1**, **1a**, **2**, and **2a** were measured by using a 500 MHz Varian INOVA spectrometer and sample **2** was measured by using a Varian 800 MHz spectrometer at the Carlsberg National NMR Center for Biomolecular Research, Copenhagen. NOESY spectra[45] in H_2O (samples **1** and **1a**) at 30 °C were recorded by using a Varian INOVA 500 MHz spectrometer employing the TPPI method[46] with a 250 ms mixing time. The spectra of **2** and **2a** in D_2O were obtained by using a Varian INOVA 500 MHz spectrometer under similar conditions. The NOESY spectra of sample **2** (pH 5.95, 1:1 TPT/DNA molar ratio) were acquired by using a Varian INOVA 800 MHz spectrometer with a mixing time of 200 ms. TOCSY spectra[47] were acquired in phase-sensitive mode by using a Varian INOVA 500 MHz spectrometer with isotropic mixing times from 0.01 to 0.12 s.

Computational methods

General: All calculations were carried out by using the AMBER 6.0 program[48] with the PARM99[49] parameter set. The nucleic acid molecules were neutralized by Na^+ cations. The molecules were surrounded by a periodic box of water described by the TIP3P potential[50] extended to a distance of 10 Å from any solute atom. The number of explicit water molecules included in the simulations varied from 2477 to 2545. The force-field parameters for TPT were selected by analogy to existing parameters in the force field. Charges were derived by using the RESP[51] multiconformational charge-fitting procedure. The ab initio electrostatic potential for RESP was calculated by using GAUSSIAN 98[52] at the HF/6-31G* level of theory, and two low-energy conformers of topotecan were used. As suggested recently,[26] the geometry of the tautomeric TPT forms and a protonated form of NMe_2 were calculated at the B3LYP/6-31G* level of theory. Solvent effects were included by means of the CPCM variant of the polarizable continuum model (PCM).

Starting structures: Our NMR spectroscopy decamer structure[34] was used (PDB code: 1G1N) as the starting structure for simulation. Topotecan was docked into the nick of the decamer parallel to the G^6–C^{15} and T^5–A^{14} base pairs. There are only four possible stacking orientations (**St1**–**St4**) for TPT in a nick, based on the two axes of a molecular plane (Scheme S1). It is assumed that all local changes in the stacking geometry of TPT in a nick (e.g., small-scale movement within the nick) will be sufficiently probed in a dynamic run. During the initial calculations it became apparent that the dimethylamino group very rarely reorients on the timescale of our MD simulations. Therefore, two conformations of the CH_2NMe_2 group in TPT were considered, designated A and B and placed on one or the other face of the molecular plane, to give eight starting structures in all. In conformation A the dimethylamino and ethyl groups are on opposite sides of the molecular plane, whereas in conformation B they are placed on the same side of the TPT plane.

Molecular dynamics: The particle mesh Ewald (PME) method[53] was used to treat long-range electrostatic interactions with a cubic B-spline interpolation and a 10^{-5} tolerance for the direct space sum cut-off. A 9 Å cut-off was applied to the nonbonded Lennard–Jones interactions. The SHAKE algorithm was applied to constrain all bonds that involved hydrogen atoms with a tolerance of 10^{-5} Å^2 and a 1 fs time step was used in the dynamics simulation. All systems used the same minimization and equilibration protocols. First, the water molecules and counterions were minimized for 1000 steps of steepest descent and 4000 of conjugate gradient method with the DNA and topotecan restrained by 10 kcal mol^{-1} Å^{-2} to the initial positions, followed by a second unrestrained minimization. The next steps of the equilibration protocol were 15 ps constant volume MD with 5 kcal mol^{-1} Å^{-2} restraints on the DNA and topotecan with the system gradually heated from 10 to 300 K by using the Berendsen coupling algorithm[54] with a coupling parameter of 1 ps. Then, by using 50 ps constant pressure MD with a 1 ps pressure relaxation time, the density of the system was adjusted close to 1 g cm^{-3}. During a subsequent 35 ps of constant volume and temperature dynamics the restraints on the TPT and DNA base pairs were gradually reduced to 0.1 kcal mol^{-1} Å^{-2}.

Docking procedure: A coarse intercalative complex was constructed manually by using our NMR decamer structure and TPT. Then a single distance restraint (TPT-H17b/G6-H4′, based on an experimental NOE) was introduced to hold the two molecules together and the energy of this restraint was gradually increased from zero up to 10 kcal mol^{-1} Å^2, with a 4 Å upper threshold, during 200 ps MD, followed by gradual relaxation (600 ps) and a final 2 ns unrestrained MD equilibration. This procedure led to the structure **St1A**. The other starting structures were generated by using this template by rotation of the TPT molecule along one of its axes (**St2A**, **St3A**, and **St4A**) or by rotating the dimethylamino group (**St1B**, **St2B**, **St3B**, and **St4B**), followed by 1 ns unrestrained equilibration. This simplified docking procedure was used because it made it possible to omit the initial generation of a nick-widening in DNA, which is the most difficult part of the docking procedure. The docking procedures were followed by a 9 ns unrestrained MD production run for each of the eight structures **St1A**–**St4B**, and additionally for one reference structure **St0** for a pure nicked DNA.

Calculations at the B3LYP/6-31G* CPCM level of theory suggest that there is a double potential well for the location of the proton in the $Me_2N\cdots HO{-}C10$ fragment, so that two tautomers are possible.[26] Independent 10 ns MD runs were therefore conducted for both tautomers of neutral TPT (**St4A-OH** and **St4A-NH**, with the proton on oxygen and nitrogen, respectively) and for TPT protonated on the NMe_2 group (**St4A-prot**), with essentially identical results (Table S5).

Back-calculations of NOE effects:[55] NOE effects were calculated by making the approximation that correlation times in the DNA–TPT complex are the same for all nuclei and neglecting any exchange effects. The NOEs were calculated for each snapshot structure taken from the MD simulations, using the cytosine H5–H6 cross-peak volumes for reference, and reported as an average over the trajectory [Eq. (1), in which $\bar{\eta}$ is the percentage of the calculated average NOE relative to the H5–H6 cytosine NOE, r_i is the distance between the observed protons in the i snapshot structure for a total number of N structures, and r_0 is the reference distance of 2.47 Å between the H5–H6 cytosine protons]. When protons from a rapidly rotating methyl group are involved in an interaction, the expression r^{-6} is replaced by $<r^{-3}>^2$. The NOE R factor was calculated by using Equation (2), in which A_{obs} and A_{calcd} are the observed and calculated NOEs, respectively.[56,57] The NOEs used are the observed intermolecular NOEs reported in Table S3.

$$\bar{\eta} = \frac{\sum_i \frac{r_i^{-6}}{r_0^{-6}}}{N} \times 100\,\% \quad (1)$$

$$R = \frac{\sum |A_{obs} - A_{calcd}|}{\sum |A_{obs}|} \times 100\,\% \quad (2)$$

Acknowledgements

The authors thank Glaxo SmithKline for a generous gift of TPT. Help from Dr. W. Koźmiński, Warsaw University, is gratefully acknowledged. Thanks are also expressed to Agnieszka Ulkowska for help in preparing the NMR samples. We also wish to thank the Danish Instrument Center

for NMR Spectroscopy of Biological Molecules at Carlsberg Laboratory for the use of the Varian 800 MHz spectrometer.

[1] J. C. Wang, *Annu. Rev. Biochem.* **1996**, *65*, 635–692.
[2] L. Stewart, M. R. Redinbo, X. Y. Qiu, W. G. J. Hol, J. J. Champoux, *Science* **1998**, *279*, 1534–1541.
[3] M. R. Redinbo, L. Stewart, P. Kuhn, J. J. Champoux, W. G. J. Hol, *Science* **1998**, *279*, 1504–1513.
[4] L. P. Rivory, J. Robert, *Bull. Cancer* **1995**, *82*, 265–285.
[5] W. J. Slichenmyer, E. K. Rowinsky, R. C. Donehower, S. H. Kaufmann, *J. Natl Cancer Inst.* **1993**, *85*, 271–291.
[6] B. L. Staker, K. Hjerrild, M. D. Feese, C. A. Behnke, A. B. Burgin, L. Stewart, *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15387–15392.
[7] Y. Pommier, G. Kohlhagen, K. W. Kohn, F. Leteurtre, M. C. Wani, M. E. Wall, *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 8861–8865.
[8] B. L. Staker, M. D. Feese, M. Cushman, Y. Pommier, D. Zembower, L. Stewart, A. B. Burgin, *J. Med. Chem.* **2005**, *48*, 2336–2345.
[9] T. K. Li, L. F. Liu, *Annu. Rev. Pharmacol. Toxicol.* **2001**, *41*, 53–77.
[10] P. Pourquier, Y. Pommier, *Adv. Cancer Res.* **2001**, *80*, 189–216.
[11] L. F. Liu, P. Duann, C. T. Lin, P. Darpa, J. X. Wu, *Ann. N.Y. Acad. Sci.* **1996**, *803*, 44–49.
[12] F. Zunino, G. Pratesi, *Expert Opin. Invest. Drugs* **2004**, *13*, 269–284.
[13] C. H. Takimoto, J. Wright, S. G. Arbuck, *Biochim. Biophys. Acta Gene Struct. Expression* **1998**, *1400*, 107–119.
[14] M. Palumbo, B. Gatto, C. Sissi in *DNA Topoisomerase-targeted Drugs* (Eds.: M. Demeunynck, C. Bailly, W. D. Wilson), Wiley, **2002**, pp. 503–537.
[15] J. F. Pizzolato, L. B. Saltz, *Lancet* **2003**, *361*, 2235–2242.
[16] H. K. Wang, S. L. Morris-Natschke, K. H. Lee, *Med. Res. Rev.* **1997**, *17*, 367–425.
[17] K. J. Cheng, N. J. Rahier, B. M. Eisenhauer, R. Gao, S. J. Thomas, S. M. Hecht, *J. Am. Chem. Soc.* **2005**, *127*, 838–839.
[18] N. J. Rahier, B. M. Eisenhauer, R. Gao, S. J. Thomas, S. M. Hecht, *Bioorg. Med. Chem.* **2005**, *13*, 1381–1386.
[19] A. Cagir, S. H. Jones, R. Gao, B. M. Eisenhauer, S. M. Hecht, *J. Am. Chem. Soc.* **2003**, *125*, 13628–13629.
[20] G. Capranico, M. Palumbo, S. Tinelli, M. Mabilia, A. Pozzan, F. Zunino, *J. Mol. Biol.* **1994**, *235*, 1218–1230.
[21] A. Tanizawa, K. W. Kohn, G. Kohlhagen, F. Leteurtre, Y. Pommier, *Biochemistry* **1995**, *34*, 7200–7206.
[22] R. T. Crow, D. M. Crothers, *J. Med. Chem.* **1992**, *35*, 4160–4164.
[23] M. C. Wani, A. W. Nicholas, M. E. Wall, *J. Med. Chem.* **1987**, *30*, 2317–2319.
[24] C. Jaxel, K. W. Kohn, M. C. Wani, M. E. Wall, Y. Pommier, *Cancer Res.* **1989**, *49*, 1465–1469.
[25] R. P. Hertzberg, M. J. Caranfa, S. M. Hecht, *Biochemistry* **1989**, *28*, 4629–4638.
[26] N. Sanna, G. Chillemi, A. Grandi, S. Castelli, A. Desideri, V. Barone, *J. Am. Chem. Soc.* **2005**, *127*, 15429–15436.
[27] S. Yao, D. Murali, P. Seetharamulu, K. Haridas, P. N. V. Petluru, D. G. Reddy, F. H. Hausheer, *Cancer Res.* **1998**, *58*, 3782–3786.
[28] Y. Fan, J. N. Weinstein, K. W. Kohn, L. M. Shi, Y. Pommier, *J. Med. Chem.* **1998**, *41*, 2216–2226.
[29] X. Y. Wang, X. A. Zhou, S. M. Hecht, *Biochemistry* **1999**, *38*, 4374–4381.
[30] J. E. Kerrigan, D. S. Pilch, *Biochemistry* **2001**, *40*, 9792–9798.
[31] G. S. Laco, J. R. Collins, B. T. Luke, H. Kroth, J. M. Sayer, D. M. Jerina, Y. Pommier, *Biochemistry* **2002**, *41*, 1428–1435.
[32] S. Mazzini, M. C. Bellucci, S. Dallavalle, F. Fraternali, R. Mondelli, *Org. Biomol. Chem.* **2004**, *2*, 505–513.
[33] J. Feeney, *Angew. Chem.* **2000**, *112*, 298–321; *Angew. Chem. Int. Ed.* **2000**, *39*, 290–312.
[34] L. Kozerski, A. P. Mazurek, R. Kawecki, W. Bocian, P. Krajewski, E. Bednarek, J. Sitkowski, M. P. Williamson, A. J. G. Moir, P. E. Hansen, *Nucleic Acids Res.* **2001**, *29*, 1132–1143.
[35] W. Bocian, R. Kawecki, E. Bednarek, J. Sitkowski, A. Pietrzyk, M. P. Williamson, P. E. Hansen, L. Kozerski, *Chem. Eur. J.* **2004**, *10*, 5776–5787.
[36] J. S. Gounarides, A. D. Chen, M. J. Shapiro, *J. Chromatogr. B Anal. Technol. Biomed. Life Sci.* **1999**, *725*, 79–90.
[37] H. Barjat, G. A. Morris, S. Smart, A. G. Swanson, S. C. R. Williams, *J. Magn. Reson., Ser. B* **1995**, *108*, 170–172.
[38] D. Z. Yang, J. T. Strode, H. P. Spielmann, A. H. J. Wang, T. G. Burke, *J. Am. Chem. Soc.* **1998**, *120*, 2979–2980.
[39] J. Bunkenborg, C. Behrens, J. P. Jacobsen, *Bioconjugate Chem.* **2002**, *13*, 927–936.
[40] L. Kozerski, R. Kawecki, P. Krajewski, P. Gluzinski, K. Pupek, P. E. Hansen, M. P. Williamson, *J. Org. Chem.* **1995**, *60*, 3533–3538.
[41] L. Kozerski, P. Krajewski, K. Pupek, P. G. Blackwell, M. P. Williamson, *J. Chem. Soc., Perkin Trans. 2* **1997**, 1811–1818.
[42] R. Lavery, H. Sklenar, *J. Biomol. Struct. Dyn.* **1988**, *6*, 63–91.
[43] A. R. Fersht, *Structure and Mechanism in Protein Science*, W. H. Freeman, New York, **1999**.
[44] H. F. Xie, D. N. Bolam, T. Nagy, L. Szabo, A. Cooper, P. J. Simpson, J. H. Lakey, M. P. Williamson, H. J. Gilbert, *Biochemistry* **2001**, *40*, 5700–5707.
[45] J. Jeener, B. H. Meier, P. Bachmann, R. R. Ernst, *J. Chem. Phys.* **1979**, *71*, 4546–4553.
[46] G. Bodenhausen, H. Kogler, R. R. Ernst, *J. Magn. Reson.* **1984**, *58*, 370–388.
[47] C. Griesinger, G. Otting, K. Wüthrich, R. R. Ernst, *J. Am. Chem. Soc.* **1988**, *110*, 7870–7872.
[48] AMBER 6.0, D. A. Pearlman, D. A. Case, J. C. Caldwell, G. L. Seibel, U. C. Singh, P. Weiner, P. A. Kollman, UCSF, San Francisco, CA (USA), **1991**.
[49] J. M. Wang, P. Cieplak, P. A. Kollman, *J. Comput. Chem.* **2000**, *21*, 1049–1074.
[50] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
[51] C. I. Bayly, P. Cieplak, W. D. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269–10280.
[52] Gaussian 98 (Revision A7), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. J. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, J. A. Pople, Pittsburgh, PA, **1998**.
[53] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577–8593.
[54] H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684–3690.
[55] D. Neuhaus, M. P. Williamson, *The Nuclear Overhauser Effect in Structural and Conformational Analysis*, Wiley-VCH, New York, **2000**.
[56] S. P. Edmondson, *J. Magn. Reson.* **1992**, *98*, 283–298.
[57] C. Gonzalez, J. A. C. Rullmann, A. M. J. J. Bonvin, R. Boelens, R. Kaptein, *J. Magn. Reson.* **1991**, *91*, 659–664.

Received: May 14, 2007
Published online: January 23, 2008

