Published in final edited form as:

Biochem J. 2008 May 1; 411(3): 523–530. doi:10.1042/BJ20071436.

Role of tryptophan anchor in human topoisomerase I structure, function, and inhibition

Gary S. LACO¹ and Yves POMMIER²

¹Laboratory of Biochemistry, Lake Erie College of Osteopathic Medicine, Bradenton, FL, 34211

²Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD, 20892, USA

Abstract

Human topoisomerase I (Top1) relaxes supercoiled DNA during cell division and transcription. Top1 is composed of 765 amino acids (aa) and contains an unstructured N-terminal domain of 200 aa, and a structured functional domain of 565 aa that binds and relaxes supercoiled DNA. Here we examined the region spanning the junction of the N-terminal domain and functional domain (junction region). Analysis of several published Top1 structures [Redinbo, M.R., Stewart, L., Champoux, J. J., Hol, W. G., 1999, *J. Mol. Biol.* 292, 685-96 and Staker, B. L., Hjerrild, K., Feese, M. D., Behnke, C. A., Burgin, A. B., Jr., Stewart, L., 2002, *Proc. Natl. Acad. Sci. USA*, 99, 15387-92] revealed that three Trp formed a network of aromatic stacking interactions and electrostatic interactions that anchored the N-terminus of the functional domain to sub-domains containing the nose cone and active site. Mutation of the three Trp (203/205/206) to Ala either individually, or together, *in silico* revealed that the individual Trp's contribution to the Trp "anchor" were additive. When the three Trp were mutated to Ala *in vitro*, the resulting mutant Top1 differed from wild-type Top1 in that it lacked processivity, exhibited resistance to camptothecin, and was inactivated by urea. The results indicate that the Trp anchor stabilizes the N-terminus of the functional domain and prevents the loss of Top1 structure and function.

Keywords

camptothecin; interaction energy score; distributive relaxation; processive relaxation; topoisomerase I; Trp anchor

INTRODUCTION

Human topoisomerase I (Top1) is a type 1B enzyme that relaxes supercoiled DNA during transcription and cell division [1-4]. Top1 binds supercoiled DNA and attacks a backbone phosphate with the active-site tyrosine (Tyr723), resulting in a tyrosyl-phosphate bond to the 3' end of the -1 scissile-strand cleavage-site deoxynucleoside, leaving a free 5'OH on the +1 scissile-strand cleavage-site deoxynucleoside (+1 deoxynucleoside) [5-7]. After Top1 makes the above covalent complex with DNA, the +1 deoxynucleoside rotates out of the helix where it is trapped in a network of electrostatic interactions with Top1 residues [8-10]. After the DNA is fully relaxed, rotation of the +1 deoxynucleoside back into the helix is favored so that its 5'-

Address correspondence to: Gary S. Laco, Lake Erie College of Osteopathic Medicine-Bradenton, 5000 Lakewood Ranch Blvd., Bradenton, Florida 34211. Phone: 941-782-0690; Fax: 941-782-5729; E-mail: glaco@lecom.edu..

The non-standard abbreviations are: +1 scissile-strand cleavage-site deoxynucleoside, +1 deoxynucleoside; CPT, camptothecin; NLS, nuclear localization signal; Top1, human topoisomerase I.

OH can attack the tyrosyl-phosphate bond, religate the DNA backbone, and allow Top1 to disassociate from the fully relaxed DNA [9]. As a result, Top1 is a processive enzyme since it releases DNA only after it is fully relaxed.

Top1 is composed of 765 aa [11], and can be divided into two general domains (Fig. 1): 1) the N-terminal domain, which has not been shown to have tertiary structure and so is referred to here as being “unstructured”, contains 4 basic nuclear localization signals (NLS) [12] as well as an acidic NLS [13]; and 2) the functional domain, nearly all of which is resolved in several X-ray crystal structures of Top1 in complex with DNA (Fig. 1) [14,15]. The functional domain contains three previously described sub-domains: 1) the core that contains the majority of amino acids that interact with DNA and defined here to include aa 201-635; 2) the flexible linker that joins the core and C-terminus (aa 636-697); and 3) the C-terminus that contains the catalytic Tyr723 (aa 698-765) [6].

Mo *et al.* analyzed the NLS's in the Top1 N-terminal domain for their role in Top1 nuclear and nucleolar localization [13]. They determined that NLS 1 (Fig. 1) is not sufficient for Top1 nuclear localization in HeLa cells, and that while NLS 3 and 4 are not important for nuclear localization NLS 4 is important for nucleolar localization. Interestingly, they found that NLS 2, a typical basic NLS, and a newly discovered acidic NLS were each independently sufficient for nuclear/nucleolar localization of Top1 [13]. They speculated that these two NLS's may be used by two different nuclear receptors [13].

The region of Top1 that spans the C-terminus of the N-terminal domain and the N-terminus of the functional domain is designated here as the junction region (aa 186-215; Fig. 1, inset). Several groups have analyzed residues within this junction region for their role in Top1 activity and came to different conclusions. Lisby *et al.* concluded that aa 191-206 (including NLS 4) were important for camptothecin (CPT) sensitivity because their p67 construct (aa 207-765 plus 10 non-native N-terminal residues) was resistant to CPT during relaxation of negatively supercoiled DNA, although it was not resistant when a dsDNA fragment was used [16]. However, their results did not exclude the deletion of aa 201-206, which lacks a NLS, from being solely responsible for the resistance of their p67 construct [16]. In contrast, Christensen *et al.* found that aa 190-210 were not required *in vitro* but were required *in vivo* [17]. However, Christensen *et al.* deleted aa 190-210 from full length Top1 resulting in aa 169-189 having the opportunity to make substitute Van der Waals and electrostatic interactions in place of the deleted residues, and those substitute interactions could be more stable *in vitro* than *in vivo* [17]. Alsner *et al.* reported that deletion of aa 141-210 from full length Top1 resulted in a Top1 that had full activity *in vitro*, however, as a result of the deletion Top1 aa 71-140 had the opportunity to make substitute interactions in place of the deleted residues [18]. The above recombinant Top1s contained either non-native N-terminal residues, or displaced native residues due to internal deletions. In both cases, residues had the opportunity to make potentially stabilizing interactions with Top1 functional domain residues and this could complicate the interpretation of the above results. Interestingly, while NLS 4 has been part of all Top1s in published Top1/DNA X-ray crystal structures, it has always been disordered [7, 14,15,19].

Here we examined the minimal structural requirements for Top1 relaxation of positively supercoiled DNA *in vitro*. In the process of doing that, we found a cluster of three Trp within the junction region that were found to play a key role in Top1 structure, function, and inhibition.

EXPERIMENTAL

Constructs

The following plasmid constructs were used in this study: pTop65 has an initiator Met codon followed by Top1 codons 214-765; pTop68 has an initiator Met codon followed by 6 His codons and 4 codons encoding the Factor Xa protease cleavage site, and then Top1 codons 201-765; pTop68W/A is the same as pTop68 except that Trp codons 203/205/206 were mutated to Ala codons; pTop69NLS has an initiator Met codon followed by 6 His codons and 4 codons encoding the Factor Xa protease cleavage site, and then Top1 codons 189-765. All Top1 constructs contained 5' *Bam*H I and 3' *Eco*R I restriction sites that were used to clone them into pBacgus-1 (Novagen, Inc., Madison, WI). The pBacgus-1 vector expresses β -glucuronidase and it is required for cytoplasmic accumulation of Top1 due to its ability to protect Top1 from proteases in the insect cell cytoplasm (data not shown).

Insect cells and baculovirus

Sf9 insect cells were used for generating and amplifying all of the Top1 recombinant baculoviruses, and HiFive insect cells (Invitrogen, Carlsbad, CA) were used for recombinant baculovirus directed Top1 expression as previously described [20].

Recombinant Top1 expression and purification

Recombinant Top1 baculovirus infected insect cells in media were centrifuged at 2,000 g and the pellet was resuspended in ice cold lysis buffer (25 mM Tris pH 7.5, 100 mM NaCl, 5 mM EDTA, 0.5 % NP40, 4 mM β -mercaptoethanol, and 10 mM PMSF) and then incubated on ice for 10 min. The sample was then centrifuged at 6,000 g for 20 min and the supernatant was then dialyzed against 25 mM Tris pH 7.5, and 4 mM β -mercaptoethanol to remove the EDTA. Top65, which does not contain a 6-His tag, was purified from the insect cell lysate on a HiTrap Q HP column (Amersham Biosciences, Uppsala, Sweden) in buffer A (25 mM Tris pH 7.5, 5 mM β -mercaptoethanol, 5 mM EDTA) and eluted with buffer B (buffer A with 1 M KCl). The Top65 peak fraction was dialyzed against buffer A, made 50% glycerol, and stored at -20°C. Top69NLS, Top68W/A, and Top68 containing insect cell lysates were batch bound first to Q Fast Flow media, and then SP Fast Flow media, in 25 mM Tris, pH 7.5 (Amersham Biosciences, Uppsala, Sweden). The Top1 containing supernates were then made 2 M KCl and 5 mM PMSF, and bound to HiTrap chelating HP media (Amersham Biosciences, Uppsala, Sweden). The respective Top1s were eluted with an imidazole gradient and dialyzed against buffer A. Top69NLS and Top68 were loaded onto a HiTrap SP HP column in buffer A and eluted with buffer B, while Top68W/A was purified on a HiTrap Q HP column as described for Top65. The peak protein fractions were dialyzed against buffer A containing 5 mM DTT, made 50% glycerol and stored at -20°C. All Top1s were 90-95% pure as judged by Coomassie Blue stained SDS/PAGE gels (data not shown).

Cleavage of the 6-His tag from Top1 constructs

Top69NLS, Top68, and Top68W/A contained an N-terminal 6-His tag followed by a Factor Xa protease cleavage site. The 6-His tag was removed by Factor Xa protease (Amersham Biosciences, Uppsala, Sweden and New England BioLabs, Beverly, MA), a serine protease, during 5 hr incubation at 21°C in 5 mM CaCl₂, 100 mM NaCl, and 50 mM Tris, pH 8.0. The Factor Xa protease/Top1 molar ratio was 1:10. Reactions were then made 1 mM AEBSF to irreversibly inactivate the Factor Xa protease. Cleavage of the 6-His tag was monitored by running aliquots of the proteins, before and after cleavage, in SDS/PAGE. The SDS/PAGE gels were then stained with GelCode® 6xHis Protein Tag Stain (Pierce, Rockford, IL), followed by illumination with UV to visualize His-tagged proteins. On average, ~95% of the 6-His tags were removed from Top69NLS, Top68, and Top68W/A (data not shown). The above

SDS/PAGE gels were subsequently stained with Coomassie Blue, and it was found that there was no significant non-specific cleavage of Top69NLS and Top68 by Factor Xa protease. However, ~25% of Top68W/A was cleaved non-specifically, most likely at a cryptic factor Xa cleavage site [21]. Purified proteins were quantitated using Coomassie Plus reagent at 595 nm using BSA as a standard (Pierce, Rockford, IL). Note: after factor Xa protease cleavage the Top69NLS, Top68, and Top68W/A proteins contained only native N-terminal residues (Experimental/Constructs). Top1s without His-tags were stored in 50% glycerol at -20°C. Note: after cleavage to remove the His-tag from Top68W/A it was used within five days as it gradually lost activity during storage at -20°C.

Top1 supercoiled DNA relaxation assays

Reaction mixtures (10 µl) contained 0.12 µg positively supercoiled phiX174 RF DNA (Invitrogen, Carlsbad, CA) and one of the following enzymes at the indicated concentration: Top69NLS, 1.5 nM; Top68, 1.5 nM; and Top68W/A, 6 nM. When indicated either CPT, or urea, was added to give the indicated concentration in the reaction mix prior to the addition of enzyme. Note: Top68W/A and Top65 50% glycerol stocks were diluted in reaction buffer on ice and then used immediately to prevent loss of activity, all other Top1s were stable when diluted. Reactions were terminated by making reactions 1% SDS, and then 4 µl of 4X loading buffer (20% glycerol, 0.15% bromophenol blue) was added, with reactions then loaded into 0.8% agarose TBE gels. After electrophoresis, the gels were stained with 1X buffer solution containing 10 µg/ml of ethidium bromide, destained using 1 mM MgSO₄, and visualized by transillumination with UV light (300 nm). Assays were done at least three times and representative gels are shown.

End-labeling of oligonucleotides

Oligonucleotides were based on a high affinity tetrahymena rDNA Top1 cleavage site [22]. The oligonucleotides were obtained from Oligosetc (PAGE purified; Wilsonville, OR). The single-stranded scissile-strand oligonucleotide was 3'-end-labeled with ³²P-dideoxyadenosine (Amersham Biosciences, Piscataway, NJ) to give a 33-mer oligonucleotide 5'AAAAAGACT/
GAAACGTTTTACAACAATTAAAA*3' with / indicating the cleavage site, t indicating the -1 thymine to which the Top1 active site Tyr723 makes a covalent bond, and A* indicating the ³²P-dideoxyadenosine label. Oligonucleotides were annealed with a 5 molar excess of unlabeled non-scissile strand oligonucleotide to form dsDNA in annealing buffer (Tris 10 mM pH 7.4, 100 mM NaCl) by heating for 4 min at 95°C and then cooling at a linear rate from 90°C to 24°C in 60 min.

Top1 dsDNA cleavage assays

The end-labeled 33-mer dsDNA substrate (~10 nM) was incubated with 10 µM CPT and one of the following enzymes at the indicated concentration: Top69NLS, 7.5 nM; Top68, 7.5 nM; and Top68W/A, 30 nM. Reactions were incubated for 15 min at 23°C in assay buffer (Tris 10 mM pH 7.4, 100 mM NaCl, 5 mM MgCl₂, 0.1 mM DTT). Note: the Top68W/A 50% glycerol stock was diluted in reaction buffer on ice and then used immediately to minimize loss of activity, the other Top1s were stable when diluted. Reactions were stopped by addition of SDS to 1%. To prepare the sample for electrophoresis, three volumes of loading buffer (98% formamide, 1 mg/ml xylene cyanol and 1 mg/ml bromophenol blue) were added to reaction mixtures. Samples were run in twenty percent denaturing polyacrylamide gels (7 M urea) at 60 W for ~2 h. Imaging and quantification were performed using a PhosphorImager and ImageQuant software, respectively (Molecular Dynamics, Sunnyvale, CA). The experiment was replicated and a representative gel is shown.

Computational chemistry

The X-ray crystal structure by Staker *et al.* of a Top1/DNA/inhibitor complex was modified using Discovery Studio 2.0 (Accelrys, San Diego, CA) [23]: 1) the crystallographic PEG, Hg, oxidized topotecan, and all waters were removed; 2) amino acids side chains with missing atoms were built out; 3) hydrogens were added and then the molecule was typed with the CFF force field; 4) indicated Trp were mutated to Ala. The wild-type and Trp/Ala mutant Top1 structures were then minimized with no constraints using steepest descent followed by conjugate gradient with an implicit distant dependent dielectric solvent model (dielectric constant of 3) until a final convergence of 0.5 was reached. The interaction energy between Top1 residues 201-214 for the wild-type and W/A mutants, and Top1 aa residues 215-765 and bound DNA were then calculated. The Trp anchor was analyzed for intramolecular H-bonds using the default H-bond parameters.

RESULTS AND DISCUSSION

A structure-based analysis and mutagenesis of Top1 was carried out in order to examine how the unstructured N-terminal domain and the highly structured functional domain coexisted in a solvated environment. In the original X-ray crystal structure of Top1 in covalent complex with DNA by Redinbo *et al.* their Top1 started with aa 175, however, the first residue to be resolved was Ile215 [7]. In order to determine if the N-terminus of the functional domain was Ile215, Top1 was first truncated to aa 214 to give Top65 (Fig. 2). The native Gly214 was included in Top65 to allow the initiator methionine to be removed by an endogenous protease during protein expression in insect cells [24]. The resulting Top65 was purified via anion exchange chromatography from the cytoplasm of insect cells (Experimental). Interestingly, Top65 did not bind to cation exchange media. This was unexpected since the Top1 used by Redinbo *et al.* bound to cation media and not anion media [7]. After dialysis to remove the elution salt, Top65 was found to be 1000-fold lower in activity (Fig. 3A) than a Top1 containing a NLS (Top69NLS; Fig. 2 and 3B). In addition, Top65 was distributive in relaxing supercoiled DNA in that it only removed one to several supercoils before religating the DNA, and so left a broad ladder of partially relaxed supercoiled DNA at lower concentrations of enzyme (Fig. 3A). Only at the highest concentration of Top65 (3 μ M) was the supercoiled DNA fully relaxed in 20 min resulting in the accumulation of 4 dominant fully relaxed DNA isomers that differed in helicity by one turn (+/-) of the helix (Fig. 3A). These results indicated that one or more Top1 residues in between aa 188-214 were important for Top1 structure and function. Lisby *et al.* reported the importance of NLS 4 (aa 191-197) for full *in vitro* function of Top1 [16], while Frohlich *et al.* reported that a Top1 containing aa 203-765 retained sensitivity to CPT and processivity and that this indicated that the major part of the N-terminal domain was dispensable for controlled strand rotation [25].

An *in silico* analysis of Top1 was first carried out in order to clarify the role of residues in Top1 structure that were not resolved by X-ray crystallography (Experimental). The first residue resolved in the Top1/DNA/inhibitor structure by Staker *et al.* was Gln201, however, of the side chain atoms only the C-beta was resolved [23]. Here we built out the Gln201 side chain and then extended the N-terminus to include Glu198-200 to determine at what point these residues stopped interacting with the functional domain residues. After minimization it was found that the Gln201 side chain OE1 and His346 ring ND1 were within range to make an electrostatic interaction at 3.5 Å (Fig. 5A). The interaction energy score between Gln201 and His346 (core sub-domain/nose cone) was calculated and found to be -4.08 kcal/mol (-0.99 Van der Waals, -3.08 electrostatic). At the same time, Glu198-200 were found to be solvent exposed and make no significant interactions with functional domain residues (data not shown). These findings are consistent with residues 175-200 always being disordered in Top1 structures [7, 14, 19, 23]. As a result, the N-terminal domain was defined as residues 1-200 and the functional domain was defined as residues 201-765.

Due to the *in vitro* results with Top65 as well as the above *in silico* experiments we decided to test the role of residues 189-200 (including NLS 4) versus residues 201-214 on Top1 activity. The following Top1s were engineered: Top69NLS, aa 189-765; and Top68, aa 201-765 (Fig. 2). The recombinant Top1s were purified, all non-native amino acids were removed from their N-termini (Experimental), and they were then tested for activity in relaxing supercoiled DNA (Fig. 3B). Top69NLS and Top68 relaxed similar amounts of supercoiled DNA in the 20 min time course and both were highly processive resulting in the accumulation of the 4 dominant fully relaxed DNA isomers at all time points (Fig. 3B). This result indicated that the loss of one or more residues in between aa 200-214 were responsible for Top65's 1000-fold lower activity and distributive relaxation (Top65, aa 214-765; Fig. 2).

Inspection of aa 201-213 in a Top1/DNA/inhibitor X-ray crystal structure by Staker *et al.* revealed a cluster of three Trp, 203/205/206 (Fig. 2), although the N-terminal aa 175-200 of that Top1 were still disordered [23]. As shown in Fig. 5 and Fig. 6, Trp 203/205/206 make an intertwined network of aromatic stacking interactions (Van der Waals) and H-bonds (electrostatics) with functional domain residues including: 1) Trp203 makes a face-to-face stacking interaction with His346 (core sub-domain/nose cone), a backbone oxygen mediated electrostatic interaction with the Trp205 backbone nitrogen, and a ring nitrogen mediated H-bond with the Glu208 OD2 (core sub-domain); 2) Trp205, in addition to the H-bond with Trp203, makes a backbone oxygen mediated H-bond with the Glu208 backbone nitrogen (core sub-domain), and a face-to-edge stacking interaction with Trp206; 3) Trp206, in addition to the above stacking interaction with Trp205, makes a ring nitrogen mediated H-bond to the Asp757 OD2 (C-terminus sub-domain/active site). We reasoned that the two aromatic stacking interactions and four electrostatic interactions made by the three Trp could act as an anchor to stabilize the N-terminus of the functional domain. The lack of the Trp anchor in Top65 could be responsible for the 1000-fold loss in activity, distributive relaxation, as well as the change in the solution conformation of the enzyme as indicated by its chromatographic behavior (Experimental).

Here we first analyzed Trp 203/205/206 *in silico* to determine what role they had in stabilizing the N-terminus of the functional domain. Trp 203/205/206 were either mutated individually, or together, to Ala and then minimized until a final convergence of 0.5 was reached (Experimental). Next, an interaction energy score was calculated between Top1 aa 201-214, for the wild-type Top1 and the W/A mutants, and Top1 aa 215-765 with bound DNA (Table 1). The mutation of the Trp 203/205/206 to Ala (Top68W/A) resulted in a 54% reduction in the Van der Waals energy, and this is consistent with the smaller volume of Ala ($90 \text{ \AA}^3$) versus Trp ($228 \text{ \AA}^3$) [26]. In addition, the electrostatic energy was also reduced due to the loss of Trp mediated H-bonds/electrostatic interactions (Table 1 and Fig. 6). The loss in Van der Waals energy for the three individual Top1 W/A mutants equaled the loss in Van der Waals for the Top68W/A triple mutant indicating that the roles of Trp 203/205/206 in anchoring the N-terminus of the functional domain were additive. The interaction energy score total for Top68W/A aa 210-214 was 38% less than the interaction energy score for wild-type Top1 aa 201-214. The loss of Van der Waals and electrostatic interactions in Top68W/A, due to the three W/A mutations, have the potential to destabilize the solvent exposed N-terminus of the functional domain (Fig. 5B).

As a result of the *in silico* data, Trp 203/205/206 were mutated *in vitro* to Ala to give Top68W/A (Fig. 2). Similar to Top65, Top68W/A also bound to anion exchange media (Experimental). However, Top68W/A eluted from the column at a lower salt concentration than Top65 (150 mM KCl for Top68W/A versus 260 mM KCl for Top65). This result indicated that the solution conformation of Top68W/A was also different from that of Top69NLS and Top68. In addition, after removal of the His-tag Top68W/A was found to gradually lose activity over a period of several weeks when stored in 50% glycerol at -20°C , presumably due to irreversible unfolding

of the protein since no degradation was detected by SDS/PAGE (data not shown). This indicated that the non-native His-tag stabilized the N-terminus of Top68W/A, likely through Van der Waals and electrostatic interactions with functional domain residues. When Top68W/A, without the His-tag, was tested for activity in a supercoiled DNA relaxation assay it was found to give a remarkably broad distribution of partially relaxed supercoiled DNA at early time points. However, many of the partially relaxed DNA isomers were not clearly visible due to their relatively low individual concentrations (Fig. 3B, compare the 5 min lanes for Top68 and Top68W/A for supercoiled DNA and partially to fully relaxed DNA isomers). At the later time points for Top68W/A the fully relaxed DNA isomers accumulated (Fig. 3B). The concentration of Top68W/A used in the assay was 4-fold greater than Top69NLS and Top68, so that Top68W/A completely relaxed the supercoiled DNA within the 20 min time course. These results indicated that Top68W/A relaxed supercoiled DNA distributively, like Top65, in that it only removed from one to several supercoils from the DNA before dissociating (Fig. 3B). In contrast, Top69NLS and Top68 relaxed supercoiled DNA processively in that they dissociated from the DNA only after it was fully relaxed (Fig. 3B). The supercoiled DNA relaxation assay results for Top68W/A suggested that the loss of the Van der Waals and electrostatic interactions made by Trp 203/205/206, due to their mutation to Ala, destabilized the N-terminus of the functional domain (Fig. 5). Interestingly, Frohlich *et al.* mutated Trp 205 to Gly and that resulted in a Top1 that had similar activity as wild-type Top1 on positively supercoiled DNA [27]. This lack of effect by the single Trp205Gly mutation supports our mutation of Trp 203/205/206 to Ala.

The sensitivity of the Top1 constructs to inhibition by CPT in a supercoiled DNA relaxation assay were next tested. Top69NLS and Top68 were inhibited nearly 100% by 10 μ M CPT. In contrast, Top68W/A was only partially inhibited by CPT and so a significant fraction of the supercoiled DNA was relaxed after 20 min (Fig. 3C). The resistance of Top68W/A to CPT was quantitated using a 33-mer dsDNA containing a high affinity Top1 cleavage site (Fig. 4A) [22]. Top69NLS, Top68, and Top68W/A were assayed on the dsDNA in the presence of 10 μ M CPT (Fig. 4B). When the amount of the Top69NLS 23-mer oligonucleotide cleavage product was set to 100%, Top68W/A was found on a per-molecule-basis to be inhibited 9-fold less by CPT. This result takes into account that a 4-fold molar excess of Top68W/A, relative to Top69NLS and Top68, was used in the dsDNA assay. The 4-fold greater Top68W/A concentration was used to be consistent with the concentration used in the supercoiled DNA assays. The resistance of Top68W/A to CPT is likely due to it making fewer stable Top68W/ADNA covalent complexes for CPT to trap in either a supercoiled DNA relaxation assay, or dsDNA cleavage assay (Fig. 3C and Fig. 4). When Frohlich *et al.* tested their Top1 Trp205Gly mutant for resistance to CPT on positively supercoiled DNA it was found to be as sensitive as wild-type Top1, however, the mutant had resistance to CPT when assayed on negatively supercoiled DNA [27]. The Frohlich *et al.* Trp205Gly mutation would weaken the Trp anchor (Table 1) and the Trp205 ring nitrogen is 3.8 Å from the Ser432 backbone oxygen. Ser432 is part of the hinge region loop (aa 429-436) and *in silico* studies have indicated that Top1 interacts differently with positively versus negatively supercoiled DNA, including that the hinge region loop extends during relaxation of negatively supercoiled DNA [28].

We thought that if the N-terminus of Top68W/A was destabilized it would be sensitive to a non-ionic denaturing agent like urea. Urea allows the unfolding of hydrophobic regions of proteins including β -sheet and extended conformations, but not α -helical regions [29]. Urea has been utilized *in vitro* to test the stability of other enzymes in which the N-termini were modified [30]. As a result, the enzymatic stability of Top68W/A was tested in a supercoiled DNA relaxation assay in the presence of increasing concentrations of urea (Experimental). And while 0.8 molar urea had no effect on Top69NLS and Top68, it dramatically reduced the activity of Top68W/A (Fig. 3D). The results of the urea assay indicates that mutation of Trp 203/205/206 to Ala destabilized the N-terminus of the functional domain, and confirms the

importance of the Van der Waals and electrostatic interactions made by Trp 203/205/206 in anchoring the N-termini of Top69NLS and Top68.

The Trp anchor is contained within a WKWW motif (Fig. 1 and 2), and we were interested in seeing if this motif was found in other topoisomerases. When all published topoisomerase I amino acid sequences were analyzed, a WXWW motif (X = any amino acid) was found between the respective N-terminal domain and functional domain of nuclear Top1s from slime mold to humans (Fig. 7), with the exception of plant and bacterial Top1s (see below). Interestingly, *Saccharomyces cerevisiae* (*S. cerevisiae*) replaces human Top1 Trp203 with Tyr (aa 131 in *S. cerevisiae*) and human Top1 Asp208 with Lys (aa 136 in *S. cerevisiae*) [31]. In human Top1, Trp203 stacks with His346 and makes a ring nitrogen mediated H-bond with Asp208 (Fig. 6). *In silico* mutation of human Top1 Trp203 to Tyr, followed by minimization, revealed that Tyr at position 203 could still stack with His346, but could not H-bond with Asp208. However, *in silico* incorporation of both the Trp203Tyr and Asp208Lys mutations into human Top1, followed by molecular dynamics simulations of Lys208 and then minimization revealed that Tyr203 could stack with His346 while making a ring oxygen mediated H-bond with the Lys208 NZ (data not shown). This *in silico* data suggests that *S. cerevisiae* Top1 Tyr131 may make similar stacking and side chain H-bond interactions as does human Top1 Trp203, although the 3-dimensional structure of *S. cerevisiae* Top1 aa 131-136 has not been reported.

The WXWW motif was not found in plant topoisomerases including *Arabidopsis thaliana* and this is consistent with the corresponding N-terminal domain and junction region sequences having low sequence identity with human Top1 [32]. However, starting with *Arabidopsis thaliana* Top1 aa 370 and human Top1 aa 216, the Top1s do have regions of high sequence identity within the functional domain [32]. In the case of bacterial Top1, the structures of two bacterial Top1s have been resolved by X-ray crystallography and include the complete N-termini [33,34]. Since these bacterial Top1s do not have an extended/unstructured N-terminal domain, as does human Top1, they may not need a Trp anchor and this is supported by the fact that they lack a WXWW motif [33,35].

The human Top1 nose cone α -helices contain four Lys (310/317/324/328) and Arg316 that do not H-bond with fully relaxed DNA (Fig. 5A & B) [14]. However, they have the potential to interact with supercoiled DNA [19]. Interaction of the four nose cone Lys and Arg316 with supercoiled DNA may prevent Top1 from dissociating from the DNA until it is fully relaxed, at which point we propose those H-bond interactions would be lost and so favor Top1's dissociation from the fully relaxed DNA. At the same time the Top1 active site is involved in recognizing when DNA is fully relaxed via the rotated +1 deoxynucleoside [8-10]. Since Top65 and Top68W/A lack the Trp anchor, the N-termini of the respective functional domains may be destabilized with their N-termini becoming solvated and so exposing the backbone carbonyl oxygens. The backbone carbonyl oxygens, each with a partial charge of -0.27, are the dominate backbone charge and their exposure could be responsible for the fact that Top65 and Top68W/A bound to anion exchange media and not cation exchange media (Experimental). In contrast, Top69NLS and Top68 both contain the Trp anchor and bound to cation exchange media (Experimental). The end result for Top65 and Top68W/A may be that the nose cone, as well as the active site, lose the ability to sense when DNA is fully relaxed resulting in Top65 and Top68W/A religating "partially relaxed" supercoiled DNA (Fig. 3A and B). The results support the conclusion that the Trp anchor stabilizes the N-terminus of the functional domain and prevents loss of Top1 structure and function *in vitro*.

A similar experimental approach, using structure directed mutagenesis and *in vitro* urea denaturation assays, was used to study the human extracellular matrix proteins fibronectin and tenascin. Fibronectin and tenascin both contain multiple fibronectin type III domains (fnIII) each of which contains several "anchors" that are held together by networks of Van der Waals

and electrostatic interactions (reviewed by Vogel [36]). The fnIII domain's anchors break sequentially, under strain, to allow stepwise lengthening of the protein [37,38]. When Cota *et al.* mutated Tyr36 to Ala in the third fnIII domain of tenascin the anchor was destabilized and resulted in a 2.8-fold increased sensitivity to urea denaturation [39]. However, in contrast to the fnIII anchors, the Top1 Trp anchor is required to stay assembled to prevent loss of the functional domain structure and function.

In vivo the Top1 N-terminal domain needs to be unstructured to allow access to nuclear import receptors and nucleolin [40-42]. In addition, the N-terminal domain anchors Top1 to nucleoli fibrillar centers, and nucleolar organization regions, of mitotic chromosomes [43]. Interestingly, Top1 functional domain sub-domains have been shown *in silico* to undergo dynamic motions of up to 80-90 Å in the absence of DNA [44]. When Top1 is anchored to a chromosome via the N-terminal domain, and the functional domain is undergoing dynamic motion either in the absence of DNA, or in the process of binding and relaxing supercoiled DNA, the potential exists for the N-terminus of the functional domain to be strained (Fig. 5). When Christensen *et al.* deleted residues 190-210 of Top1, including the Trp anchor, and analyzed the construct *in vivo* it was found to be insensitive to CPT treatment [17]. This result is consistent with our *in vitro* data presented here which demonstrated that mutation of Trp 203/205/206 to Ala resulted in Top68W/A that was 9-fold less sensitive to CPT inhibition (Fig. 4). Together, these results support the proposal that the Top1 Trp anchor plays a key role in stabilizing the N-terminus of the functional domain as well as the nose cone and active site *in vitro* and *in vivo*, resulting in a highly processive enzyme which as a consequence of the processivity is sensitive to inhibition by camptothecin.

Acknowledgments

We thank Dr. Jane Sayer for advice and careful review of the manuscript, as well as Drs. Ross Longley, Mark Nicklaus, and Megan Peach for thoughtful discussions.

GSL was supported by a National Cancer Institute postdoctoral fellowship and a Lake Erie College of Osteopathic Medicine intramural grant.

REFERENCES

1. Champoux JJ, Dulbecco R. An activity from mammalian cells that untwists superhelical DNA--a possible swivel for DNA replication (polyoma-ethidium bromide-mouse-embryo cells-dye binding assay). *Proc. Natl. Acad. Sci. USA* 1972;69:143-146. [PubMed: 4333036]
2. Champoux JJ. DNA topoisomerases: structure, function, and mechanism. *Annu. Rev. Biochem* 2001;70:369-413. [PubMed: 11395412]
3. Pommier Y, Pourquier P, Fan Y, Strumberg D. Mechanism of action of eukaryotic DNA topoisomerase I and drugs targeted to the enzyme. *Biochim. Biophys. Acta* 1998;1400:83-105. [PubMed: 9748515]
4. Chen AY, Liu LF. DNA Topoisomerases: essential enzymes and lethal targets. *Annu. Rev. Pharmacol. Toxicol* 1994;94:194-218.
5. Wang JC. DNA topoisomerases. *Annu. Rev. Biochem* 1996;65:635-692. [PubMed: 8811192]
6. Champoux JJ. Domains of human topoisomerase I and associated functions. *Prog. Nucleic. Acids Res. Mol. Biol* 1998;60:111-132.
7. Redinbo MR, Stewart L, Kuhn P, Champoux JJ, Hol WG. Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA. *Science* 1998;279:1504-1513. [PubMed: 9488644]
8. Pommier Y, Laco GS, Kohlhaagen G, Sayer JM, Kroth H, Jerina DM. Position-specific trapping of topoisomerase I-DNA cleavage complexes by intercalated benzo[a]-pyrene diol epoxide adducts at the 6-amino group of adenine. *Proc. Natl. Acad. Sci. USA* 2000;97:10739-10744. [PubMed: 10995470]

9. Laco GS, Collins JR, Luke BT, Kroth K, Sayer JM, Jerina DM, Pommier Y. Human topoisomerase I inhibition: docking camptothecin and derivatives into a structure-based active site model. *Biochemistry* 2002;41:1428–1435. [PubMed: 11814334]
10. Laco GS, Wu D, Kohlhagen G, Sayer JM, Jerina DM, Burke TG, Curran DP, Pommier Y. Analysis of human topoisomerase I inhibition and interaction with the cleavage site +1 deoxyguanosine, via in vitro experiments and molecular modeling studies. *Bioorg. Med. Chem* 2004;12:5225–5235. [PubMed: 15351405]
11. D'Arpa P, Machlin PS, Ratrie H 3rd, Rothfield NF, Cleveland DW, Earnshaw WC. cDNA cloning of human DNA topoisomerase I: catalytic activity of a 67.7-kDa carboxyl-terminal fragment. *Proc. Natl. Acad. Sci. USA* 1988;85:2543–2547. [PubMed: 2833744]
12. Stewart L, Ireton GC, Parker LH, Madden KR, Champoux JJ. Biochemical and biophysical analyses of recombinant forms of human topoisomerase I. *J. Biol. Chem* 1996;271:7593–7601. [PubMed: 8631793]
13. Mo YY, Wang C, Beck WT. A novel nuclear localization signal in human DNA topoisomerase I. *J. Biol. Chem* 2000;275:41107–41113. [PubMed: 11016921]
14. Redinbo MR, Stewart L, Champoux JJ, Hol WG. Structural flexibility in human topoisomerase I revealed in multiple non-isomorphous crystal structures. *J. Mol. Biol* 1999;292:685–696. [PubMed: 10497031]
15. Leshner DT, Pommier Y, Stewart L, Redinbo MR. 8-Oxoguanine rearranges the active site of human topoisomerase I. *Proc. Natl. Acad. Sci. USA* 2002;99:12102–12107. [PubMed: 12209008]
16. Lisby M, Olesen JR, Skouboe C, Krogh BO, Straub T, Boege F, Velmurugan S, Martensen PM, Andersen AH, Jayaram M, Westergaard O, Knudsen BR. Residues within the N-terminal domain of human topoisomerase I play a direct role in relaxation. *J. Biol. Chem* 2001;276:20220–20227. [PubMed: 11283003]
17. Christensen MO, Barthelmes HU, Boege F, Mielke C. Residues 190–210 of human topoisomerase I are required for enzyme activity in vivo but not in vitro. *Nucleic Acids Res* 2003;31:7255–7263. [PubMed: 14654701]
18. Alsner J, Svejstrup JQ, Kjeldsen E, Sorensen BS, Westergaard O. Identification of an N-terminal domain of eukaryotic DNA topoisomerase I dispensable for catalytic activity but essential for in vivo function. *J. Biol. Chem* 1992;267:12408–12411. [PubMed: 1319995]
19. Stewart L, Redinbo MR, Qiu X, Hol WG, Champoux JJ. A model for the mechanism of human topoisomerase I. *Science* 1998;279:1534–1541. [PubMed: 9488652]
20. Prasad GS, Stura EA, McRee DE, Laco GS, Hasselkus-Light C, Elder JH, Stout CD. Crystal structure of dUTP pyrophosphatase from feline immunodeficiency virus. *Protein Sci* 1996;5:2429–2437. [PubMed: 8976551]
21. Eaton D, Rodriguez H, Vehar GA. Proteolytic Processing of Human Factor-Viii - Correlation of Specific Cleavages by Thrombin, Factor Xa, and Activated Protein-C with Activation and Inactivation of Factor-Viii Coagulant Activity. *Biochemistry* 1986;25:505–512. [PubMed: 3082357]
22. Bonven BJ, Gocke E, Westergaard O. A high affinity topoisomerase I binding sequence is clustered at DNAase I hypersensitive sites in Tetrahymena R-chromatin. *Cell* 1985;41:541–551. [PubMed: 2985282]
23. Staker BL, Hjerrild K, Feese MD, Behnke CA, Burgin AB Jr, Stewart L. The mechanism of topoisomerase I poisoning by a camptothecin analog. *Proc. Natl. Acad. Sci. USA* 2002;99:15387–15392. [PubMed: 12426403]
24. Laco GS, Kent SB, Beachy RN. Analysis of the proteolytic processing and activation of the rice tungro bacilliform virus reverse transcriptase. *Virology* 1995;208:207–214. [PubMed: 11831702]
25. Frohlich RF, Andersen FF, Westergaard O, Andersen AH, Knudsen BR. Regions within the N-terminal domain of human topoisomerase I exert important functions during strand rotation and DNA binding. *J. Mol. Biol* 2004;336:93–103. [PubMed: 14741206]
26. Tsai J, Taylor R, Chothis C, Gerstein M. The packing density in proteins: standard radii and volumes. *J. Mol. Biol* 1999;290:253–266. [PubMed: 10388571]
27. Frohlich RF, Veigaard C, Andersen FF, McClendon AK, Gentry AC, Andersen AH, Osheroff N, Stevnsner T, Knudsen BR. Tryptophane-205 of human topoisomerase I is essential for camptothecin

- inhibition of negative but not positive supercoil removal. *Nucleic Acids Res* 2007;35:6170–6180. [PubMed: 17827209]
28. Sari L, Andricioaei I. Rotation of DNA around intact strand in human topoisomerase I implies distinct mechanisms for positive and negative supercoil relaxation. *Nucleic Acids Res* 2005;33:6621–6634. [PubMed: 16314322]
 29. Bennion BJ, Daggett V. The molecular basis for the chemical denaturation of proteins by urea. *Proc. Natl. Acad. Sci. USA* 2003;100:5142–5147. [PubMed: 12702764]
 30. Laco GS, Fitzgerald MC, Morris GM, Olson AJ, Kent SB, Elder JH. Molecular analysis of the feline immunodeficiency virus protease: generation of a novel form of the protease by autoproteolysis and construction of cleavage-resistant proteases. *J. Virol* 1997;71:5505–5511. [PubMed: 9188624]
 31. Thrash C, Bankier AT, Barrell BG, Sternglanz R. Cloning, characterization, and sequence of the yeast DNA topoisomerase I gene. *Proc. Natl. Acad. Sci. USA* 1985;82:4374–4378. [PubMed: 2989818]
 32. Kieber JJ, Tissier AF, Signer ER. Cloning and Characterization of an Arabidopsis thaliana Topoisomerase I Gene. *Plant Physiol* 1992;99:1493–1501. [PubMed: 16669064]
 33. Lima CD, Wang JC, Mondragon A. Three-dimensional structure of the 67K N-terminal fragment of E. coli DNA topoisomerase I. *Nature* 1994;367:138–146. [PubMed: 8114910]
 34. Patel A, Shuman S, Mondragon A. Crystal structure of a bacterial type IB DNA topoisomerase reveals a preassembled active site in the absence of DNA. *J. Biol. Chem* 2006;281:6030–6037. [PubMed: 16368685]
 35. Krogh BO, Shuman S. A poxvirus-like type IB topoisomerase family in bacteria. *Proc. Natl. Acad. Sci. USA* 2002;99:1853–1858. [PubMed: 11830640]
 36. Vogel V. Mechanotransduction involving multimodular proteins: converting force into biochemical signals. *Annu. Rev. Biophys. Biomol. Struct* 2006;35:459–488. [PubMed: 16689645]
 37. Craig D, Krammer A, Schulten K, Vogel V. Comparison of the early stages of forced unfolding for fibronectin type III modules. *Proc. Natl. Acad. Sci. USA* 2001;98:5590–5595. [PubMed: 11331785]
 38. Krammer A, Lu H, Isralewitz B, Schulten K, Vogel V. Forced unfolding of the fibronectin type III module reveals a tensile molecular recognition switch. *Proc. Natl. Acad. Sci. USA* 1999;96:1351–1356. [PubMed: 9990027]
 39. Cota E, Hamill SJ, Fowler SB, Clarke J. Two proteins with the same structure respond very differently to mutation: the role of plasticity in protein stability. *J. Mol. Biol* 2000;302:713–725. [PubMed: 10986129]
 40. Mo YY, Wang P, Beck WT. Functional expression of human DNA topoisomerase I and its subcellular localization in HeLa cells. *Exp. Cell Res* 2000;256:480–490. [PubMed: 10772820]
 41. Mao Y, Mehl IR, Muller MT. Subnuclear distribution of topoisomerase I is linked to ongoing transcription and p53 status. *Proc. Natl. Acad. Sci. USA* 2002;99:1235–1240. [PubMed: 11805286]
 42. Edwards TK, Saleem A, Shaman JA, Dennis T, Gerigk C, Oliveros E, Gartenberg MR, Rubin EH. Role for nucleolin/Nsr1 in the cellular localization of topoisomerase I. *J. Biol. Chem* 2000;275:36181–36188. [PubMed: 10967121]
 43. Christensen MO, Barthelmes HU, Boege F, Mielke C. The N-terminal domain anchors human topoisomerase I at fibrillar centers of nucleoli and nucleolar organizer regions of mitotic chromosomes. *J. Biol. Chem* 2002;277:35932–35938. [PubMed: 12119295]
 44. Chillemi G, Bruselles A, Fiorani P, Bueno S, Desideri A. The open state of human topoisomerase I as probed by molecular dynamics simulation. *Nucleic Acids Res* 2007;35:3032–3038. [PubMed: 17439970]

Figure 1. Schematic diagram of human topoisomerase I

N-terminal domain and functional domain are indicated, as are the basic nuclear localization signals (NLS) 1-4 and an acidic NLS (A). The shaded N-terminus of the functional domain (aa 201-240) corresponds to the red backbone in Fig. 5. Inset; junction region residues including the N-terminal domain NLS 4 (underlined) and the N-terminal residues of the functional domain.

Figure 2. Human topoisomerase I constructs

Amino acids spanning the N-terminal domain and functional domain, and the respective N-termini of the Top65, Top68, Top68W/A, and Top69NLS constructs. *Trp 203/205/206 that were mutated to Ala in Top68W/A.

Figure 3. Supercoiled DNA relaxation assays

Nicked circular DNA (N) and supercoiled DNA (SC) indicated on left. A) Top65 (μM) supercoiled DNA relaxation assay, all reactions were stopped at 20 min, the 4 dominant fully relaxed DNA isomers are visible in the 3 μM lane; C, no added enzyme; B) Time course supercoiled DNA relaxation assay, Top69NLS and Top68 at 1.5 nM, Top68W/A at 6 nM, time points of 0 min, 5 min, 10 min, 20 min indicated; C) Time course with 10 μM CPT supercoiled DNA relaxation assay; same time points and enzyme concentrations as in B; D) Urea concentration (M) supercoiled DNA relaxation assay; same enzyme concentrations as in B, all reactions stopped at 20 min. 4X indicates that 4-fold more Top68W/A was used compared to Top69NLS and Top68.

Figure 4. *In vitro* dsDNA cleavage assay

A) Schematic diagram of the 3'-labeled 33-mer dsDNA substrate (^{32}P -dideoxyadenosine label is indicated by *), with Top1 (black circle) making a covalent bond to the DNA backbone resulting in the 23-mer oligonucleotide cleavage product. B) Top1 dsDNA cleavage assay with $10\mu\text{M}$ CPT and 10 nM dsDNA. Reaction products were separated in denaturing PAGE. Lane C, control reaction with no added Top1; lane 1, Top69NLS (7.5 nM); lane 2, Top68 (7.5 nM); lane 3, Top68W/A (30 nM). The 33-mer oligonucleotide is indicated on the left, as is the 23-mer oligonucleotide cleavage product.

Figure 5. Top1/DNA and Trp anchor

A) Top1 in complex with DNA (aqua), both shown as a backbone ribbons. The functional domain backbone is red (aa 201-240) with remainder in blue. Trp 203/205/206 are rendered as CPK, with Trp203 top (gold), Trp205 middle (red), and Trp206 bottom (gold). The N-terminus of the functional domain is to the right of Trp203 and ends with Gln201 (red). His346 (black) is above Trp203, Asp757 (black) is to the lower right of Trp206, and Glu208 (black) is to the immediate left of Trp203, Ile215 (red) is to the far left of Trp205. The nose cone (upper right) is indicated and the 4 Lys and 1 Arg shown left to right: Arg316, red; Lys310 yellow; Lys328 yellow; Lys324 red; Lys317 yellow. The active site Tyr723 side chain and covalent bond to the DNA is red (lower center). The linker sub-domain is indicated.

B) Top1/DNA in 5A was rotated 90° to the left in the horizontal plane, nose cone residues from left to right: Lys328, yellow; Lys324 red; Lys317 yellow; Arg316, red; Lys310 yellow. C)

Close up of Trp anchor in 5B (stereo view) with backbone as ribbon, side chain carbons, gold; oxygens, red; nitrogens, blue. Trp 203/205/206 shown with residues with which they make electrostatic and stacking interactions: His346, core sub-domain/nose cone; Asp757, C-terminal sub-domain/active site. The Trp203 and Trp205 backbone oxygen and nitrogen, and Asp208 backbone nitrogen, are also shown.

Figure 6. Trp anchor electrostatic and stacking interactions

Flattened view of electrostatic and stacking interactions between Trp 203/205/206 and His346 (core sub-domain/nose cone) and Asp757 (C-terminal sub-domain/active site). Carbon, gold; oxygen, red; nitrogen, blue.

Human	KKPKKEE---EQK <u>WXWW</u> EEERYPEGI
Norway Rat	KKVKKEEDI-EDIWEWWKEEKKPAGV
Chicken	KKPKKEE---EQKWKWWEEERYPEGI
African Clawed Frog	KKVKKEE---EEKWKWWEEERHRDGI
Fruit fly	KKAKKEE---EQKWKWWEEERYPEGI
Nematode	KKPKKEE---EQKWKWWEEERYPEGI
Slime Mold	KKPKKEE---EQKWKWWEEERYPEGI
Yeast*	EEEDKKAKEEEEE Y KWWE K ENEDDTI

Figure 7. Alignment of Top1 junction region amino acid sequences

Indicated nuclear Top1s with junction region aa shown in single letter code. The conserved WXWW motif that forms the Trp anchor in human Top1 is in bold and underlined. *Yeast (*Saccharomyces cerevisiae*) contains a Tyr in place of human Top1 Trp203, and a Lys in place of human Top1 Glu208, both are in bold. Plant and bacterial Top1s lack a WXWW motif and so are not shown (see Results and Discussion).

Table I

Wild-type and mutant Trp anchor interaction energy scores

Energy:	Total	Van der Waals	Electrostatic
Top1 wild-type	-94.05 [*]	-72.10	-21.95
Top1 W203A	-85.00	-62.70	-22.30
Top1 W205A	-79.73	-58.65	-21.08
Top1 W206A	-81.82	-62.43	-19.39
Top168W/A ^{**}	-58.16	-39.59	-18.57

Interaction energy scores between aa 201-214 of wild-type and W/A mutant Top1s and Top1 aa 215-765 and bound DNA.

^{*} All values are in Kcal/mol.^{**} Top68W/A, 203/205/206 W/A triple mutant.