Explorations of Peptide and Oligonucleotide Binding Sites of Tyrosyl-DNA Phosphodiesterase Using Vanadate Complexes

Douglas R. Davies,[†] Heidrun Interthal,[‡] James J. Champoux,[‡] and Wim G. J. Hol*[,†,§]

Department of Biochemistry, P.O. Box 357742, School of Medicine, University of Washington, Seattle, Washington 98195-7242, Department of Microbiology, P.O. Box 357242, School of Medicine, University of Washington, Seattle, Washington 98195-7242, and Biomolecular Structure Center and Howard Hughes Medical Institute, Seattle, Washington 98195

Received October 1, 2003

Tyrosyl-DNA phosphodiesterase (Tdp1) catalyzes the hydrolysis of a phosphodiester bond between a tyrosine residue and a DNA 3′ phosphate and functions as a DNA repair enzyme that cleaves stalled topoisomerase I−DNA complexes. We previously determined a procedure to crystallize a quaternary complex containing Tdp1, vanadate, a DNA oligonucleotide, and a tyrosine-containing peptide that mimics the transition state for hydrolysis of the Tdp1 substrate. Here, the ability of vanadate to accept a variety of different ligands is exploited to produce several different quaternary complexes with a variety of oligonucleotides, and peptides or a tyrosine analogue, in efforts to explore the binding properties of the Tdp1 DNA and peptide binding clefts. Eight crystal structures of Tdp1 with vanadate, oligonucleotides, and peptides or peptide analogues were determined. These structures demonstrated that Tdp1 is able to bind substituents with limited sequence variation in the polypeptide moiety and also bind oligonucleotides with sequence variation at the 3′ end. Additionally, the tyrosine analogue octopamine can replace topoisomerase I derived peptides as the apical ligand to vanadate. The versatility of this system suggests that the formation of quaternary complexes around vanadate could be adapted to become a useful method for structure-based inhibitor design and has the potential to be generally applicable to other enzymes that perform chemistry on phosphate esters.

Introduction

Tyrosyl-DNA phosphodiesterase (Tdp1), a member of the phospholipase D superfamily, is a recently discovered enzyme that catalyzes the hydrolysis of a phosphodiester bond between a tyrosine residue and a DNA 3′ phosphate.[1,2] Tdp1 achieves this hydrolysis via a two-step mechanism in which a conserved histidine residue first performs a nucleophilic attack on the phosphate moiety of the substrate, releasing tyrosine and forming a covalent enzyme−DNA complex.[3] In the second step, the intermediate phosphohistidine complex is hydrolyzed by nucleophilic attack by a water molecule likely activated by a second conserved histidine residue, yielding 3′ phosphate DNA and the native enzyme.[4] Because the 3′ phosphotyrosine substrate is only known to exist in eukaryotic cells as part of the transient covalent enzyme−DNA intermediate formed when a type IB DNA topoisomerase nicks double-stranded DNA, it was suggested that Tdp1 functions as a DNA repair enzyme to remove such complexes when they become stalled on the DNA in the cell.[1] Such stalled topoisomerase I−DNA complexes can be induced by various forms of DNA damage such as nicks, gaps, or abasic sites that prevent religation of DNA by the enzyme. Treatment with topoisomerase I poisons such as the anticancer drug camptothecin also leads to topoisomerase I−DNA complexes with prolonged lifetimes (reviewed in refs 5−7). Genetic studies of *Saccharomyces cerevisiae* have confirmed that Tdp1 does play a direct role in the repair of topoisomerase I−DNA lesions.[8,9]

A stalled topoisomerase I−DNA complex is a problematic substrate for Tdp1, since the scissile phosphodiester bond between the catalytic tyrosine residue of topoisomerase I and the DNA 3′ phosphate is partially buried within the complex.[10] A proposal that proteolysis or some other form of degradation of the topoisomerase I−DNA complex must precede Tdp1 activity[1] is supported by biochemical[3] and structural experiments,[11] although the exact extent of, and the agents responsible for, the degradation that occurs in vivo are unknown. In any case, the natural substrate of Tdp1 is large and complex, consisting of tyrosine or possibly a tyrosine-containing peptide moiety linked to a single strand of DNA via a 3′ phosphodiester bond. In vitro experiments with human Tdp1 have demonstrated that the peptide moiety of suitable substrates can be as small as a single tyrosine residue or as large as the 6.3 kDa C-terminal domain of topoisomerase I.[12] The peptide moiety of the substrate need not even be a peptide, as demonstrated by the discovery that both human and *S. cerevisiae* Tdp1 are capable of removing glycolate from a 3′ phosphoglycolate DNA oligonucleotide.[13] Similarly, studies with human[14] and *S. cerevisiae* Tdp1[7] show that the enzyme is capable of cleaving substrates in vitro with substantial length diversity in the DNA moiety of the substrate.

The suitability of human Tdp1 as a drug target for the treatment of cancer has been the subject of much

* To whom correspondence should be addressed. Phone: 206-685-7044. Fax: 206-685-7002. E-mail: hol@gouda.bmsc.washington.edu.
[†] Department of Biochemistry.
[‡] Department of Microbiology.
[§] Biomolecular Structure Center and Howard Hughes Medical Institute.

10.1021/jm030487x CCC: $27.50
Published on Web 01/13/2004

speculation.[2,15,16] Early conjecture that Tdp1 inhibitors would be effective in combined drug therapy with camptothecin for treatment of cancers was based on the presumption that the activity of Tdp1 would counteract the effects of camptothecin and its derivatives. Indeed, *S. cerevisiae* Tdp1 deletion mutants are slightly hypersensitive to camptothecin and hypersensitive to a "toxic topoisomerase I" that is defective in religation.[9] However, genetic studies in yeast have revealed alternative pathways for repair of topoisomerase I−DNA lesions that are independent of Tdp1. Specifically the RAD1−RAD10 and Mus81/Mms4 endonucleases have been identified as repair alternatives to Tdp1 in yeast.[9] Additional light was shed on the biological role of Tdp1 by the recent discovery of a heritable condition in humans that leads to spinocerebellar ataxia with axonal neuropathy (SCAN1).[17] The specific mutation in SCAN1 is the replacement of His493 in Tdp1 with arginine, a change that is predicted to interfere with the second step of catalysis.[3,4] The effect of this form of ataxia is most pronounced in neurons that rarely divide, suggesting that an essential role of Tdp1 in humans may be the repair of topoisomerase I induced DNA damage that interferes with transcription. These results indicate that blocking the activity of Tdp1 may have significantly different effects in different organisms or different situations. Although *S. cerevisiae* Tdp1 deletion mutants are not vastly hypersensitive to camptothecin, the effect of topoisomerase I poisons on mammalian or human cells where the Tdp1 repair pathway is blocked has yet to be reported. Further research is needed to determine the potential benefits of human Tdp1 inhibitors in combination with camptothecin for the treatment of cancer. Similarly, selective inhibitors targeted toward Tdp1 of eukaryotic pathogens may also hold potential for the treatment of parasitic infections by combination drug therapy, since it is known that *Plasmodium falciparum* is sensitive to camptothecin.[18]

Irrespective of the ultimate suitability of human Tdp1 as a drug target, the enzyme is very well suited to demonstrate a new method for the rapid crystallographic screening of substrates, substrate analogues, and potential inhibitors through binding of these molecules to vanadate in the active site of the enzyme. Recently, a crystal structure of human Tdp1 containing vanadate in a transition-state mimic of the first step of the Tdp1 reaction was solved.[11] Through cocrystallization, Tdp1, a hexanucleotide and an octapeptide were made to self-assemble around a central vanadate moiety to form a transition-state mimic for the hydrolysis of the phospohodiester bond of the peptide−DNA substrate. In the resultant quaternary complex, the DNA oligonucleotide was located in a narrow, positively charged groove that extends approximately 20 Å from the active site of the enzyme. The peptide moiety of the substrate mimic occupied a portion of a large bowl-shaped binding cleft that is located on the side of the active site opposite the DNA-binding groove. The central vanadate moiety exhibited trigonal bipyramidal geometry consistent with an S_N2 nucleophilic attack on phosphate, with His263 of Tdp1 comprising one apical ligand (representing the attacking nucleophile), the tyrosine residue of the peptide comprising the other apical ligand (representing the leaving group), and the DNA 3′ hydroxyl comprising one equatorial ligand.[11] This method of forming quaternary complexes around vanadate is predicted to be modular and versatile; any compound that binds to Tdp1 in such a way to present a suitable ligand to vanadate should be able to participate in the formation of such a complex.

We have tested this concept by investigating a number of different compounds for their ability to form cocrystal complexes with Tdp1 and vanadate. The previously described quaternary complex around vanadate utilized a peptide with a sequence derived from the topoisomerase I substrate. The resulting structure provided clues to the interaction of Tdp1 with its natural substrate. Altering the sequence of the peptide moiety of the substrate analogue, however, provides an opportunity to search for tighter and more extensive interactions with the Tdp1 peptide-binding pocket as a first step toward the design of specific inhibitors of the enzyme. The use of a non-peptide analogue of tyrosine, octopamine, was also employed to demonstrate the ability of non-peptide molecules to occupy the peptide-binding site of Tdp1. We also varied both the sequence and the length of the DNA moiety of the substrate analogue in order to determine if Tdp1 has sequence or length specificity. Here, we report eight structures of Tdp1−vanadate complexes that include peptides of varying length and sequence, non-peptide analogues of tyrosine, and oligonucleotides of varying length and sequence. Taken together, these structures represent not only an exploration of the DNA and peptide binding sites of Tdp1 but also a demonstration of the potential for utilizing the vanadate-complex method for the rapid crystallographic evaluation of potential inhibitors in support of structure-based drug design efforts.

Results

Explorations of the Peptide Binding Site. In the Tdp1−vanadate complex previously solved with an eight-residue topoisomerase I derived peptide (NH_3−KLNYLDPR−COOH, called 8WT), only the first five residues of the synthetic peptide, corresponding to topoisomerase I residues Lys720−Leu724, were ordered in the electron density map. These residues made only three direct hydrogen bonds with the enzyme and buried only 575 Å^2 of solvent-accessible surface area of Tdp1.[11] The C-terminal ordered residue of the topoisomerase I derived peptide, Leu724, made no hydrogen bonds with Tdp1 but was close enough to Trp590 to be in hydrophobic contact with the edge of the indole side chain. Trp590 is in turn located near Trp605, forming a hydrophobic pocket on one side of the peptide-binding cleft. Modeling indicated that interactions between the hydrophobic pocket in the peptide-binding cleft of Tdp1 formed by Trp590 and Trp605 and the topoisomerase I derived peptide could potentially be increased by substitutions of the residue corresponding to Leu724 of topoisomerase I. Synthetic peptides were made that incorporated the following changes: a pentapeptide incorporating a leucine to lysine substitution (5L724K, NH_3−KLNY**K**−COOH) and an octapeptide incorporating a leucine to tyrosine change (8L724Y, NH_3−KLNY**Y**DPR−COOH). The lysine substitution was chosen for its potential to make a hydrogen-bond contact with the carbonyl oxygen of Trp590, which is located at the bottom of the hydrophobic binding pocket. The

Table 1. Data Collection and Refinement Statistics for Tdp1−Vanadate Complexes

	Tdp1−vanadate−AGTT−8L724Y	Tdp1−vanadate−AGTC−5L724K	Tdp1−vanadate−AGTA−octopamine	Tdp1−vanadate−AGTC−octopamine	Tdp1−vanadate−AGTG−octopamine	Tdp1−vanadate−AGTT−octopamine	Tdp1−vanadate−GTT−octopamine	Tdp1−vanadate−TT−octopamine
λ (Å)	1.0000	0.9791	0.9791	0.9791	1.0781	0.9791	0.9791	0.9791
a (Å)	50.076	49.820	49.811	49.791	49.716	49.800	49.692	49.862
b (Å)	104.904	104.832	104.688	104.894	104.557	104.711	104.750	104.578
c (Å)	193.832	193.539	193.310	194.124	193.180	193.751	194.033	193.377
resoln (Å)	50.0−1.7	50.0 −2.2	50.0−2.1	50.0−2.0	50.0−2.22	50.0−2.1	50.0−2.3	50.0−1.9
mosaicity (deg)	0.317	0.453	0.350	0.319	0.548	0.979	0.653	0.386
unique reflns	103 225	53 356	57 880	68 032	47 285	56 460	46 407	68 069
completeness	90.6 (64.6)	99.9 (99.8)	95.6 (85.6)	97.4 (91.2)	92.7 (80.3)	94.2 (91.7)	99.1 (50.5)	84.1 (77.4)
av I/σ	21.04(2.00)	12.99(2.62)	13.45(2.74)	15.29(3.38)	21.25(3.41)	10.50(1.80)	13.72(2.53)	14.54(1.76)
redundancy	4.52	5.26	4.93	4.15	5.84	3.19	4.90	1.94
R_{sym}	7.1 (48.2)	12.5 (52.7)	11.4 (48.0)	8.5 (36.8)	8.1 (39.0)	10.8 (60.1)	11.8 (50.5)	6.0 (43.1)
Refinement								
resoln range (Å)	50.0−1.7	50.0−2.2	50.0−2.1	50.0−2.0	50.0−2.22	50.0−2.1	50.0−2.3	50.0−1.9
reflns (free)	102 833 (5161)	52 449 (2674)	57 538 (2917)	67 981 (3424)	47 189 (2398)	56 410 (2852)	45 642 (2312)	67 910 (3387)
R_{cryst} (R_{free})	19.3 (21.3)	18.8 (22.7)	19.5 (23.3)	19.6 (22.8)	18.7 (23.1)	19.9 (23.9)	19.9 (23.9)	20.3 (23.4)
rms deviations								
bond lengths (Å)	0.012	0.010	0.012	0.010	0.010	0.011	0.011	0.011
bond angles (deg)	1.335	1.167	1.343	1.161	1.219	1.245	1.241	1.224
chirality	0.102	0.085	0.084	0.086	0.087	0.091	0.088	0.091
av *B* factors								
main chain	22.933	24.222	23.178	22.692	29.041	28.123	26.845	28.956
side chain	25.440	25.860	24.623	24.705	30.457	29.713	28.158	30.304
ligands	30.748	38.035	29.081	27.252	34.785	31.562	36.418	
waters	29.637	28.646	26.496	25.323	31.061	29.573	29.145	29.496

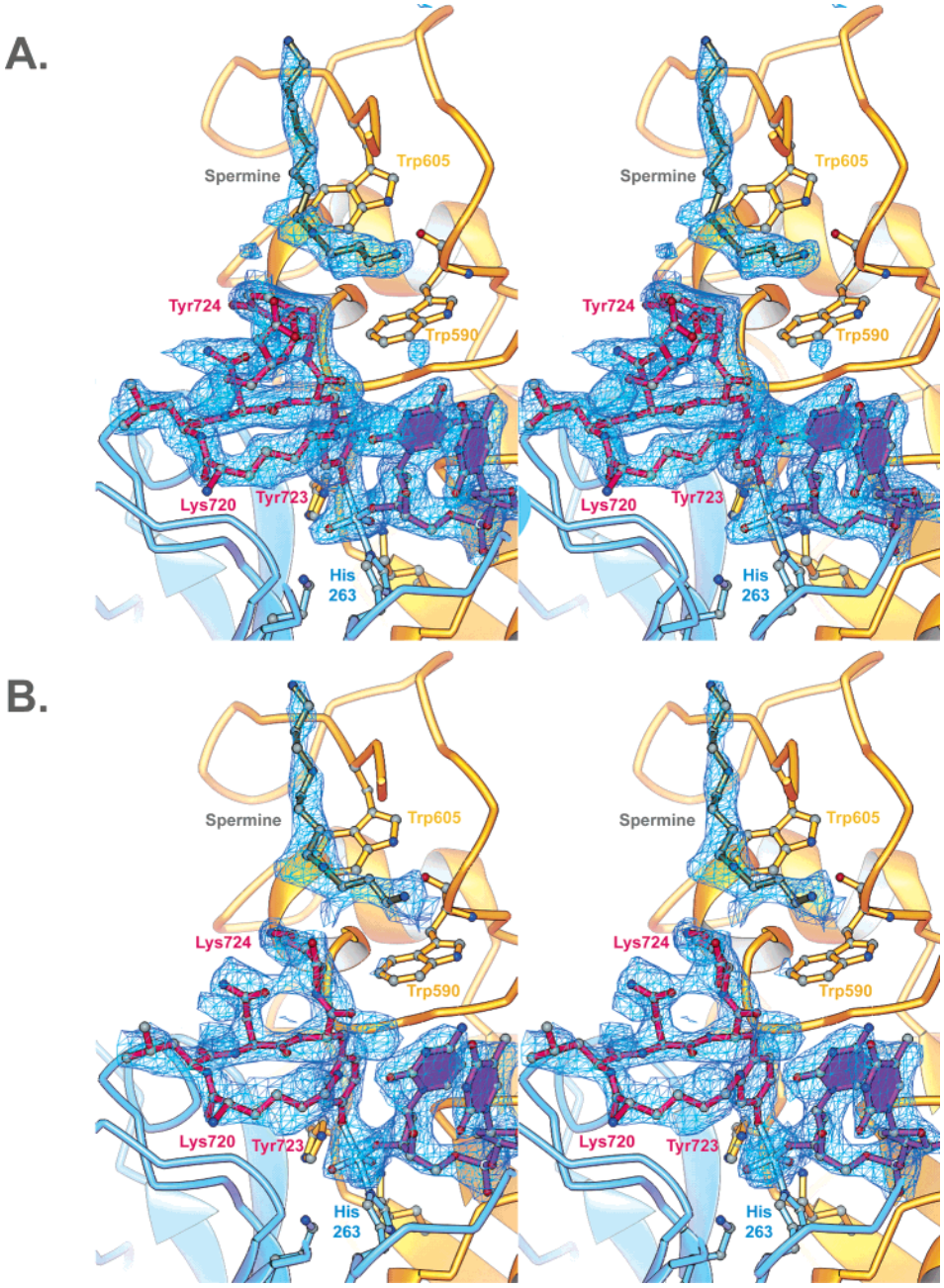
tyrosine substitution was designed to form more general hydrophobic and stacking interactions with the side chain of Trp590 and also to test the specificity of Tyr723 for binding to vanadate in the active site of Tdp1.

All of the quaternary complexes around vanadate described in this work were formed using the method described previously,[11] substituting various components for the peptide and oligonucleotide moieties. All complexes crystallized in the same crystal form, with similar unit cell dimensions, and were solved by molecular replacement. Data collection and refinement statistics for all eight structures are listed in Table 1.

Cocrystallization of the 8L724Y peptide, vanadate, and a DNA oligonucleotide with the sequence 5′ AGTT 3′ resulted in a structure of the quaternary complex that was solved by molecular replacement at a resolution of 1.7 Å (Figure 1A). Similarly, a structure of the vanadate complex with 5L724K and the oligonucleotide 5′ AGTC 3′ was solved at a resolution of 2.2 Å (Figure 1B). In both structures, the conformation of residues Lys720−Tyr723 of the topoisomerase I derived peptide was very similar to the conformation observed in the previously solved structure with the wild-type topoisomerase I peptide. Interestingly, there was no indication in the electron density maps for the 8L724Y structure that the "non-native" tyrosine at position 724 served as an alternative ligand for vanadate in the quaternary complex, even as a minor product. An attempt was also made to replace the topoisomerase I derived peptide with a different tyrosine-containing octapeptide, [Asn[1],Val[5]] angiotensin II (NH_3−NRVYVHPF−COOH). However, the resulting structure did not contain evidence of a quaternary complex (data not shown).

Unexpectedly, neither Tyr724 in 8L724Y nor Lys724 in 5L724K was found in these structures to interact with the hydrophobic pocket formed by Trp590 and Trp605. The side chains of both Tyr724 (Figure 1A) and Lys724 (Figure 1B) were rotated from the targeted hydrophobic binding pocket and oriented toward the center of the peptide-binding cleft directly opposite the DNA-binding groove. The change in orientation of these altered residues was largely due to a rotation of the χ_1 angle relative to the conformation of Leu724 in the previously solved peptide−DNA−vanadate complex of approximately +40° for lysine in 5L724K and +90° for tyrosine in 8L724Y. Inspection of the relatively high-resolution structure of the quaternary complex with 8L724Y revealed that Tyr724 was prevented from binding in the targeted hydrophobic pocket by the presence of a spermine molecule (Figure 1A,C). Spermine was present as an additive necessary for the crystallization of Tdp1 in the crystal form used for the formation of quaternary complexes.

The spermine molecule was bound at a crystal-packing interface between subunit A of Tdp1 from one asymmetric unit and subunit B of Tdp1 from a neighboring asymmetric unit. One hydrogen bond links the terminal amino group of spermine to the backbone oxygen of Ser449 on the symmetry-related B subunit, and a water-mediated hydrogen bond connects this amino group to the backbone nitrogen of Met604 on the proximal A subunit (Figure 1C). A secondary amino group of spermine (N5) is also within hydrogen-bonding distance of the backbone nitrogen of residue Val606 of the proximal A subunit. Additionally, the backbone carbonyl oxygens of Ser450 and Ser600 of the symmetry-related B subunit are located approximately 3.6 Å from amino groups of spermine. These distances are very long for hydrogen-bonding interactions but may represent hydrogen-bonding opportunities for minor alternative conformations of spermine not present in our model. While approximately one-half of the spermine molecule is involved in this crystal-packing interface, the remaining half of the molecule rests on the faces of the indole side chains of Trp590 and Trp605, blocking the hydrophobic binding pocket targeted by the 8L724Y and 5L724K peptides (Figure 1). In the B subunit of Tdp1, the crystal-packing interaction that formed part of the



spermine-binding site is absent. Consequently, the electron density for the spermine molecule was weaker and appeared to be partially disordered. Nevertheless, the conformations of the topoisomerase I derived peptides are very similar in both the A and B subunits. Spermine has subsequently been located at the same crystal-packing interface in all structures where it was used as a crystallization additive. Re-examination of the moderate-resolution quaternary complex described previously[11] reveals that spermine is likely located in this structure as well.

The tyrosine analogue octopamine was employed in place of topoisomerase I derived peptides to evaluate the ability of non-peptide substituents to occupy the peptide-binding site of Tdp1 in vanadate complexes. This analysis provides a proof of principle test for the use of quaternary vanadate complexes for inhibitor design, since it exhibits the versatility of the approach and obviates any reliance on polypeptides, which are poor drug molecules. Octopamine has a secondary alcohol functional group rather than a carboxylate group but retains the phenol moiety of the tyrosine side chain. Octopamine is also more soluble than tyrosine, making it more suitable for the formation of quaternary complexes with vanadate, Tdp1, and DNA.

The structures of six Tdp1−vanadate−DNA−octopamine complexes have been solved with resolutions between 2.0 and 2.3 Å (Table 1, Figures 2 and 3). In all cases, the octopamine molecule binds to vanadate through the phenolic oxygen in the same manner as tyrosine. Octopamine, like tyrosine, contains a single chiral center. Although octopamine used in the formation of quaternary complexes was a racemic mixture, electron density maps for the complexes indicate that the (*S*)-enantiomer is preferred when the 3′ nucleotide is thymidine or guanosine, but both the (*S*)- and (*R*)-enantiomers are present when the 3′ nucleotide is adenosine or cytidine (Figures 2−4).

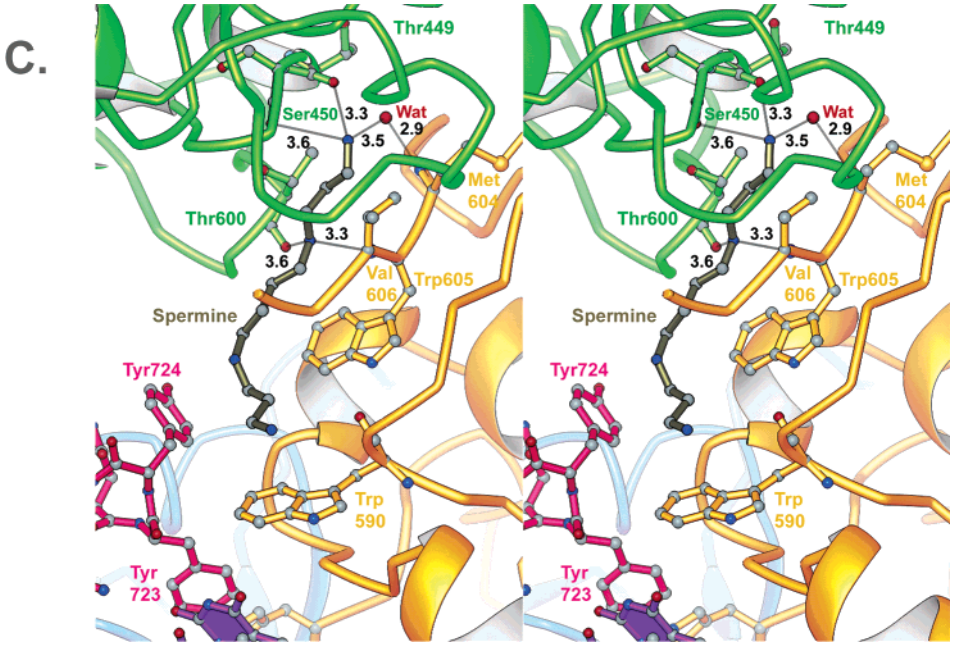


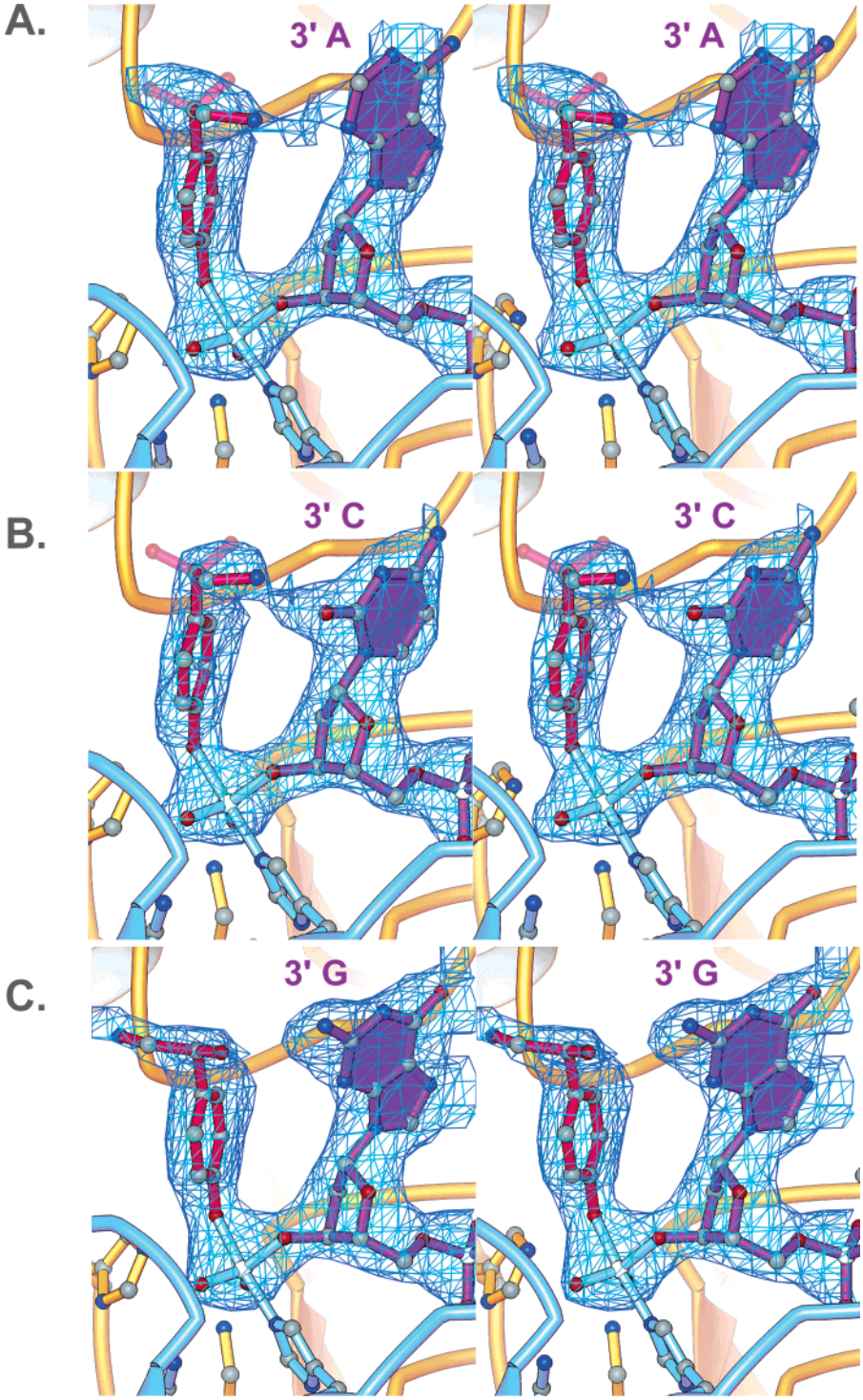
Figure 1. Stereoviews of Tdp1−vanadate−DNA−modified peptide complexes. Tdp1 is displayed as a ribbon structure with the N-terminal domain (residues 162−350) colored blue and the C-terminal domain (residues 351−608) colored yellow. Histidine and lysine residues of the active site and tryptophan residues of the hydrophobic binding pocket in the peptide-binding cleft are displayed as ball-and-stick structures with bonds colored to match the corresponding domain of Tdp1. Ball-and-stick structures also represent the vanadate (blue bonds), DNA (purple), peptide (magenta), and spermine (dark-gray). (A) Tdp1 quaternary complex with peptide 8L724Y. Blue contours represent the 2.0σ level of the $F_o - F_c$ electron density map from the final model with the peptide, vanadate, DNA, and spermine removed. (B) Tdp1 quaternary complex with peptide 5L724K. Blue contours represent the 2.0σ level of the $F_o - F_c$ electron density map from the final model with the peptide, vanadate, DNA, and spermine removed. (C) Close-up view of spermine in the 8L724Y structure. Potential hydrogen bonds are displayed as thin black lines with distances in angstroms labeled on the figure. All figures were generated with the program RIBBONS.[27]

Explorations of the DNA Binding Groove. To explore the DNA-binding groove, quaternary vanadate complexes were formed using single-stranded oligonucleotides varying in length and sequence. An oligonucleotide with the sequence 5′ AGAGTT 3′ was used previously to form the quaternary complex with the topoisomerase I derived peptide.[11] In this structure, the backbone of only the three residues at the 3′ end and only two full nucleotide residues at the 3′ end were visible in the electron density map. Furthermore, of the two monomers of Tdp1 in the asymmetric unit of the crystal, only the A subunit contained well resolved DNA or peptide. Crystal packing interactions were likely responsible for exclusion of DNA from the B subunit, since the three visible nucleotide residues extend nearly to the end of the DNA-binding groove and the area adjacent to this groove is more constricted by symmetry-related molecules for the B subunit.

Four shorter oligonucleotides, AGTT, AGTA, AGTG, and AGTC, were used to examine the effect of sequence differences at the 3′ end, and the oligonucleotides GTT and TT were employed to test the effect of oligonucleotide length on binding to Tdp1. These structures were all solved by molecular replacement at resolutions ranging from 2.0 to 2.3 Å (Table 1). In all cases, quaternary complexes of oligonucleotides, vanadate, and Tdp1 were formed with octopamine as the component of the quaternary complex bound in the peptide-binding cleft. As was the case with the hexanucleotide-containing complex, three nucleotide residues at the 3′ end of the tetranucleotides AGTT, AGTA, AGTG, and AGTC were visible in the electron density maps, with the guanine base of the third residue from the 3′ end disordered. In each case, the identity of the 3′ base was clearly identifiable as thymine, adenine, guanine, and cytosine, respectively (Figure 2). Unlike the case with the hexanucleotide, structures solved with tetranucleotides exhibited well-defined electron density for the DNA moiety in both subunits of Tdp1 in the asymmetric unit of the crystals. Although four-residue oligonucleotides yielded better electron density than six-residue oligonucleotides, DNA shorter than four residues resulted in inferior structures. The three-residue oligonucleotide GTT produced a crystal structure in which the 5′ G residue was completely disordered (Figure 3B). The dinucleotide TT resulted in a structure in which neither the DNA nor the octopamine moieties were well defined as ligands for vanadate (Figure 3C).

Discussion

Specificity of the Peptide Binding Site and Implications for Inhibitor Design. Since the function of Tdp1 is the removal of stalled topoisomerase I−DNA complexes, the peptide-binding cleft of Tdp1 has most likely evolved to specifically bind polypeptides corresponding to the region around the catalytic tyrosine residue of topoisomerase I. Interestingly, the conformations of the 8WT, 8L724Y, and 5L724K peptides in vanadate complexes with Tdp1 are significantly different from the conformation of the corresponding residues in the crystal structures of topoisomerase I bound to DNA.[10,11,19] However, in the topoisomerase I structure these residues are adjacent to a flexible linker that connects the core of the enzyme to the C-terminal domain that contains the catalytic tyrosine residue and thus may be free to adopt alternative conformations if



the topoisomerase I−DNA complex is partially degraded. Although the topoisomerase I derived peptide makes only three direct hydrogen bonds with Tdp1 and buries less than 600 Å^2 of solvent-accessible surface area, the interaction appears to confer a degree of specificity. The 8L724Y peptide contained a second tyrosine residue in addition to Tyr723, either of which in principle would be suitable ligands for vanadate. However, there was no evidence for any binding of Tyr724 to vanadate, indicating that the context of a tyrosine residue in a peptide is an important factor in the formation of quaternary complexes with vanadate. Consistent with this observation is the inability of an angiotensin II peptide to participate in the formation of a quaternary vanadate complex with Tdp1, despite the fact that like the topoisomerase I derived peptide, angiotensin is an octapeptide with a tyrosine residue at the fourth position from the N terminus.

Although peptide sequence is important for binding to Tdp1−vanadate complexes, optimization appears to be possible. The altered peptide 8L724Y did not bind to a targeted hydrophobic pocket in the expected manner, but the structure of the complex with 8L724Y did yield the most complete picture of any octapeptide bound to Tdp1−vanadate to date (Figure 1A). The binding of the crystallization additive spermine is an intriguing discovery that suggests a potential avenue for inhibitor design. If a spermine-like moiety could be linked to a

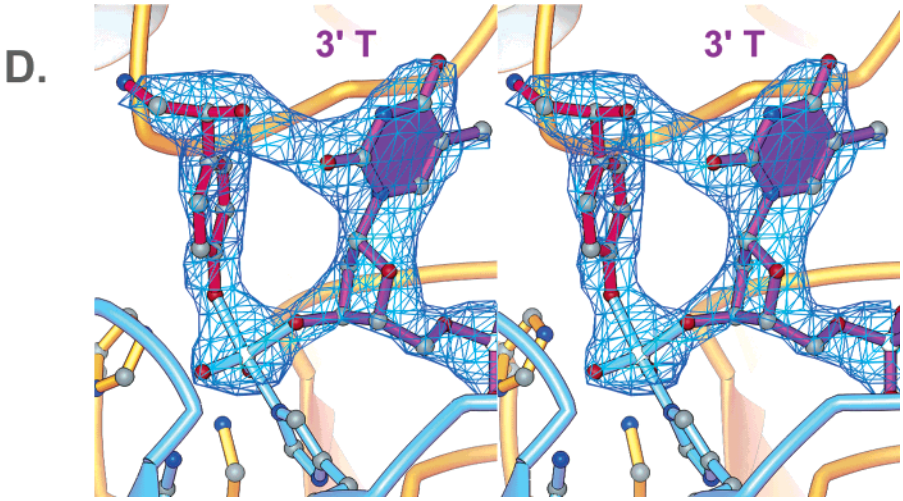


Figure 2. Explorations of the human Tdp1 DNA sequence specificity, with stereoviews of structures of quaternary complexes of Tdp1, vanadate, octopamine, and oligonucleotides. Tdp1 is colored as in Figure 1. Blue contours represent electron density for $F_o - F_c$ omit maps for the final refined structures with vanadate, octopamine, and DNA removed from the model. All electron density maps are displayed at the 2.5σ contour level except where otherwise noted. (A) AGTA and (B) AGTC are contoured at 3.0σ. The presence of both (*R*)- and (*S*)-isomers of octopamine in the AGTA and AGTC structures was best modeled by splitting the secondary alcohol oxygen into both possible positions with occupancy of 0.5 at each position. Semitransparent bonds and atoms for the secondary alcohol in the figures display these alternative positions. In both the (C) AGTG and (D) AGTT structures, only the (*S*)-isomer of octopamine was modeled in the peptide-binding site.

peptide or peptidomimetic moiety, the resulting molecule could potentially serve as a highly specific binder in the peptide-binding cleft of Tdp1 and as a crystallization additive. Linkage of a spermine moiety to a compound capable of binding to vanadate would be, in principle, straightforward because of the presence of reactive amino groups at the termini of the spermine molecule. While omission of spermine from the crystallization solution abrogates crystal growth, crystals can still be obtained when spermine is replaced by spermidine in crystal drops. However, even this shorter-chain molecule partially blocks the targeted hydrophobic binding pocket[20] (data not shown).

The peptide-binding cleft of human Tdp1 appears to be well suited as a target for potential inhibitors of Tdp1. Features such as the hydrophobic pocket formed by Trp590 and Trp605 provide potential binding sites for inhibitors. Furthermore, the peptide-binding site may provide the best opportunity for the targeting of specific inhibitors to the Tdp1 of a specific pathogen. The DNA-binding groove contains many residues that are conserved among Tdp1 orthologues from different species, whereas the peptide-binding pocket shares much less sequence conservation between different species.[11] The potential for such species selectivity may be exploited in using Tdp1 as a drug target for the treatment of eukaryotic pathogenic parasitic infections.

DNA Binding Site. The binding of different oligonucleotides as shown in the quaternary vanadate complexes with AGTT, AGTA, AGTG, and AGTC in the DNA-binding site is a clear demonstration that Tdp1 has relaxed specificity for DNA sequence at the 3′ end. This relaxed specificity is not surprising in light of the DNA sequence specificity of topoisomerase I that forms part of the Tdp1 substrate. Although topoisomerase I has a preference for cleaving DNA at the 3′ end of a thymidine residue, topoisomerase I will also cleave DNA at other sites.[21] The relaxed 3′ sequence specificity of Tdp1 is also consistent with the proposal that the ability of the enzyme to remove glycolate from a 3′ phosphoglycolate−DNA substrate in vitro is indicative of a secondary role for Tdp1 in DNA repair.[13] Phosphoglycolate lesions can result from fragmentation of DNA ribose sugars due to ionizing radiation. Since such damage could theoretically occur anywhere in the DNA, any corresponding repair enzyme should be capable of removing glycolate from the 3′ end of any DNA sequence. All crystal structures of Tdp1 solved today show a DNA-binding site that can accommodate only a single strand of DNA, whereas Tdp1 can act on both single-stranded and double-stranded DNA in vitro. As disussed in more detail earlier,[12] the enzyme may undergo a conformational change in order to accommodate double-stranded DNA or, alternatively, it may melt duplex DNA prior to binding of the substrate in the active site. The exact nature of the in vivo substrates for Tdp1 remains elusive.

In light of the new crystal structures, the length requirements of oligonucleotides for crystallographic studies of Tdp1 can be well understood. A hexanucleotide was used in the first quaternary vanadate complex but was only visible in one of the two molecules of Tdp1 in the asymmetric unit of the crystals.[11] Tetranucleotides were observed in both subunits of Tdp1 in crystals but only for the three residues at the 3′ end (Figure 3A). A trinucleotide was observed but with the entire 5′ residue invisible in the electron density map (Figure 3B), and a quaternary complex formed with a dinucleotide resulted in no interpretable electron density for the DNA moiety at all (Figure 3C). Oligonucleotides longer than four residues extend beyond the end of the DNA-binding groove and have the potential to clash with symmetry-related Tdp1 molecules. The poor binding of oligonucleotides shorter than four residues may be understood in view of the number of phosphate-binding sites in the DNA-binding groove of Tdp1. Excluding the active site, there are three phosphate-binding sites in the DNA-binding groove: the one nearest the active site

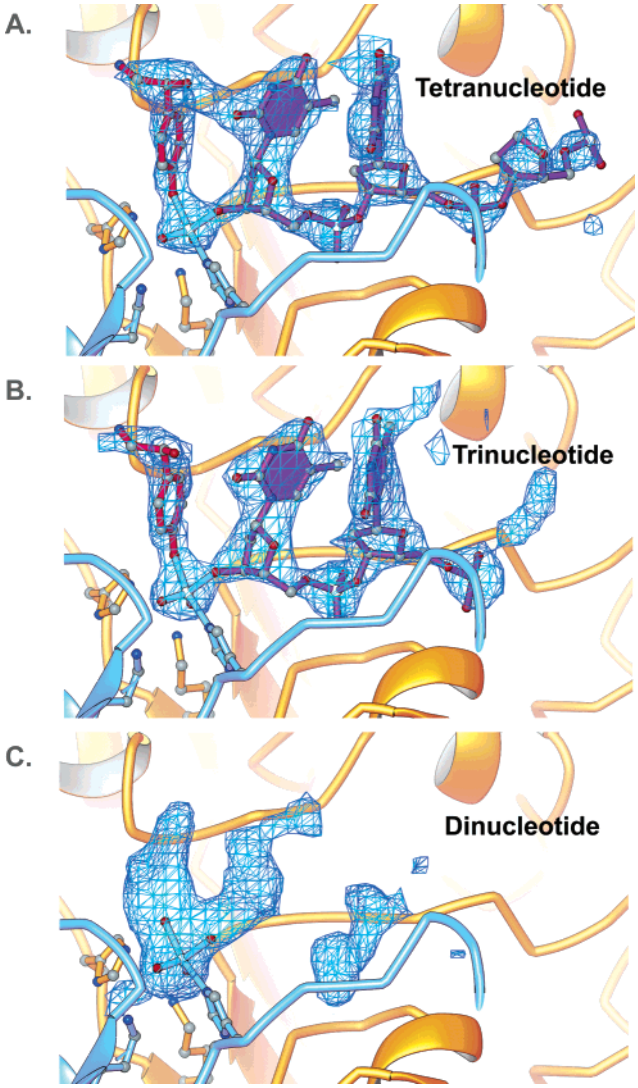


Figure 3. Exploration of DNA length specificity in Tdp1 crystals. Coloration is the same as in Figures 1 and 2, and electron density contours represent the 2.5σ level of $F_o - F_c$ omit maps calculated from final refined structures with vanadate, DNA, and octopamine removed, if present. (A) Quaternary complex with tetranucleotide AGTT. The electron density map was sufficient to allow the building of the backbone of the three 3′ residues and the whole nucleotide of the two 3′ residues of the oligonucleotide. (B) Quaternary complex with trinucleotide GTT. The quality of the electron density allowed only the building of the two 3′ residues of the oligonucleotide. (C) Electron density obtained from cocrystallization of octopamine, vanadate, and dinucleotide TT. In this structure, only the vanadate moiety of the substrate analogue formed by octopamine, vanadate, and DNA was built into the model because of the poor quality of the resulting electron density map.

formed by the side chains of Ser400 and Ser518, the one formed by the backbone nitrogens of Ser403 and Ala520, and the third formed by the side chains of Ser403 and Lys469. Since synthesized oligonucleotides used for the formation of quaternary complexes with vanadate did not include 5′ phosphates, only the tetranucleotides possessed enough phosphates to bind all three phosphate-binding sites in the DNA-binding groove.

The nature of the nucleotide residue at the 3′ end of the oligonucleotides exerts an influence on the binding of octopamine in the peptide-binding site. For the

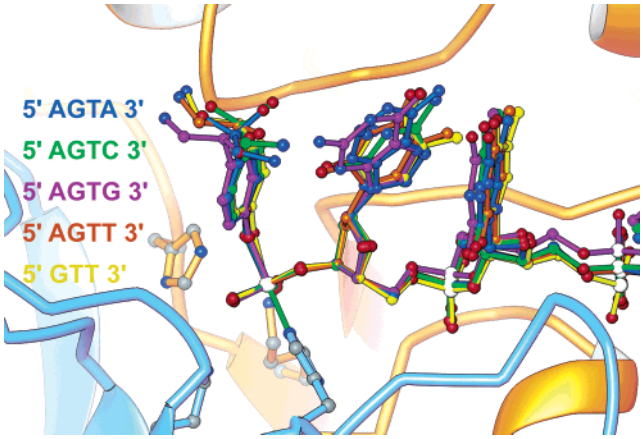


Figure 4. Superposition of Tdp1−vanadate−octopamine−DNA complexes. Tdp1 model is from the 2.0 Å structure of Tdp1−vanadate−octopamine−AGTC. The bonds for the vanadate−octopamine−DNA portions of the quaternary structures are colored as follows: AGTA, blue; AGTC, green; AGTG, purple; AGTT, orange; GTT, yellow.

structures with AGTA and AGTC, initial difference maps and residual difference maps following the inclusion of octopamine in the structure indicated that both enantiomers of octopamine were bound to vanadate (parts A and B of Figure 2). However, in each case, the methylamino moiety of octopamine produced the best fit when oriented toward the 3′ base. In contrast, structures solved with thymidine or guanosine at the 3′ position (AGTT, GTT, and AGTG) appeared to select for binding of only the (*S*)-isomer (Figures 2C,D and 3B). In the cases of the AGTT and GTT structures, the reason for this selectivity appears to be due to a subtle shift of the thymine base in the DNA-binding groove relative to the positions of adenine and cytosine (Figure 4). Perhaps because of steric constraints imposed by the C5 methyl group, the nucleoside portion of thymidine is rotated slightly about the C3′−C4′ bond in the ribose sugar to position the O2 atom of the base approximately 0.5 Å to one side of the active site of Tdp1 compared to the position of O2 in cytosine or the position of N3 in adenosine. This shift would make the distance between the amino group of octopamine and the hydrogen bond accepting O2 of thymidine too short. Instead, the hydroxyl group of the secondary alcohol moiety of octopamine is within hydrogen-bonding distance of O2. In the case of the AGTG structure, the protrusion of the N2 amino group in guanine places similar constraints on the conformation of the octopamine, specifically selecting for the (*S*)-isomer.

Significance of Vanadate Complexes for Functional Studies and Inhibitor Screening. The experiments described in this work illustrate the versatility of the formation of vanadate complexes in the active site of Tdp1. Quaternary vanadate complexes appear to be insensitive to the identity of the nucleotide base at the 3′ end of the oligonucleotide. For crystallographic reasons, tetramer oligonucleotides appear to be optimal for forming quaternary complexes. The variety of functional groups that have been shown to be suitable ligands for vanadate in quaternary complexes with Tdp1 suggests the possibility that large libraries of compounds may be suitable for the formation of quaternary complexes with a view toward inhibitor screening.

Materials and Methods

Crystallization, Freezing Conditions, and Data Collection. Human Tdp1 Δ1−148 was expressed and purified as previously described.[3] All Tdp1−vanadate complexes were prepared by cocrystallization. Concentrated Tdp1 (approximately 6.5 mg/mL in 250 mM NaCl, 15 mM Tris (pH 8.2), and 1 mM EDTA) was combined with activated sodium orthovanadate[22] to a final vanadate concentration of 2 mM. After incubation of the Tdp1−vanadate mixture on ice for 1−5 min, additional components of the complex (e.g., 5 mM peptides or 50 mM octopamine, and 2−4 mM DNA oligonucleotides) were added. Complexes were crystallized by vapor diffusion in sitting drops by combining equal volumes of protein solution and a crystallant typically containing 19−21% PEG 3000, 200 mM NaCl, 10 mM spermine, and 100 mM HEPES or MOPS (pH 7.8). To prepare crystals for flash-cooling, they were transferred directly to a cryoprotectant containing 5 mM sodium orthovanadate, 15% PEG 3000, 100 mM HEPES (pH 7.8), 375 mM NaCl, 35% PEG DME 250, and peptide or octopamine and DNA at the same concentrations used in crystallization. Within 30−90 s of transference to cryoprotectant, crystals were mounted in loops and flash-cooled by plunging directly into liquid nitrogen. Crystals were orthorhombic ($P2_12_12_1$; approximate cell dimensions $a = 50$ Å, $b = 105$ Å, $c = 194$ Å). Usually, dozens of crystals appeared in each drop within 1−2 days and usually reached full size in 1 week. Crystals had a platelike habit, typically reached sizes of 0.2 mm × 0.1 mm × 0.05 mm, and diffracted to approximately 2.0 Å with synchrotron radiation.

All data were collected at the Advanced Light Source beamline 8.2.2 on an ADSC Quantum 315 detector. The data were integrated, scanned, and reduced with the HKL 2000 software.[23]

Phase Determination, Model Building, and Structural Refinement. The structures of Tdp1 were solved by molecular replacement with AMORE[24] using residues 160−608 of apo-Tdp1 (PDB accession number 1jy1 [12]) as a search model. Additional model building and addition of waters were carried out using XtalView.[25] Further refinement was carried out using Refmac 5,[24,26] and alternating cycles of manual rebuilding were carried out in XtalView. Noncrystallographic symmetry restraints were not applied because significant differences in loop structure between the two subunits were observed.

Refinement of Octopamine in Quaternary Complexes. A solution of racemic octopamine was a component in the formation of the vanadate complexes for oligonucleotide specificity and length experiments. Fourier difference maps calculated from the initial molecular replacement solutions of the AGTA and AGTC complexes had features resembling both (*R*)- and (*S*)-isomers of octopamine. Each enantiomer of octopamine was individually modeled into the peptide-binding pocket and refined with Refmac5. In these cases, Fourier difference maps suggested the presence of the alternative enantiomer, regardless of the model used. For the final round of refinement, the secondary alcohol oxygen atom was split into two positions with an occupancy of 0.5 and the octopamine residue was refined as both enantiomers simultaneously. This approach gave the flattest difference maps for the AGTA and AGTC structures.

Coordinates. The atomic coordinates for the structures described in this work have been deposited in the RCSB Protein Data Bank (accession codes 1RFF, 1RFI, 1RG1, 1RG2, 1RGT, 1RGU, and 1RHO).

Acknowledgment. The authors thank Daniel Mitchell, Claire O'Neal, Stewart Turley, and Dr. Jan Abendroth for assistance with data collection, and Francis Athappilly for computing support. We also gratefully acknowledge the staff of beamline 8.2.2 of the Advanced Light Source. This work was supported by NIH Grant GM49156 to J.J.C. and NIH Grant CA65656 to W.J.G.H.

References

(1) Yang, S. W.; Burgin, A. B., Jr.; Huizenga, B. N.; Robertson, C. A.; Yao, K. C.; et al. A eukaryotic enzyme that can disjoin dead-end covalent complexes between DNA and type I topoisomerases. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 11534−11539.
(2) Pouliot, J. J.; Yao, K. C.; Robertson, C. A.; Nash, H. A. Yeast gene for a Tyr-DNA phosphodiesterase that repairs topoisomerase I complexes. *Science* **1999**, *286*, 552−555.
(3) Interthal, H.; Pouliot, J. J.; Champoux, J. J. The tyrosyl-DNA phosphodiesterase Tdp1 is a member of the phospholipase D superfamily. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12009−12014.
(4) Davies, D. R.; Interthal, H.; Champoux, J. J.; Hol, W. G. J. Insights into substrate binding and catalytic mechanism of human tyrosyl-DNA phosphodiesterase (Tdp1) from vanadate and tungstate-inhibited structures. *J. Mol. Biol.* **2002**, *324*, 917−932.
(5) Strumberg, D.; Pilon, A. A.; Smith, M.; Hickey, R.; Malkas, L.; et al. Conversion of topoisomerase I cleavage complexes on the leading strand of ribosomal DNA into 5′-phosphorylated DNA double-strand breaks by replication runoff. *Mol. Cell. Biol.* **2000**, *20*, 3977−3987.
(6) Pourquier, P.; Pommier, Y. Topoisomerase I-mediated DNA damage. *Adv. Cancer Res.* **2001**, *80*, 189−216.
(7) Pommier, Y. Diversity of DNA topoisomerases I and inhibitors. *Biochimie* **1998**, *80*, 255−270.
(8) Vance, J. R.; Wilson, T. E. Repair of DNA strand breaks by the overlapping functions of lesion-specific and non-lesion specific DNA 3′-phosphatases. *Mol. Cell. Biol.* **2001**, *21*, 7191−7198.
(9) Liu, C.; Pouliot, J. J.; Nash, H. A. Repair of topoisomerase I covalent complexes in the absence of the tyrosyl-DNA phosphodiesterase Tdp1. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 14970−14975.
(10) Redinbo, M. R.; Stewart, L.; Kuhn, P.; Champoux, J. J.; Hol, W. G. J. Crystal structures of human topoisomerase I in covalent and noncovalent complexes with DNA. *Science* **1998**, *279*, 1504−1513.
(11) Davies, D. R.; Interthal, H. I.; Champoux, J. J.; Hol, W. G. J. Crystal structure of a transition state mimic for Tdp1 assembled from vanadate, DNA, and a topoisomerase I-derived peptide. *Chem. Biol.* **2003**, *10*, 139−147.
(12) Davies, D. R.; Interthal, H.; Champoux, J. J.; Hol, W. G. J. The Crystal Structure of Human Tyrosyl-DNA Phosphodiesterase, Tdp1. *Structure* **2002**, *10*, 237−248.
(13) Inamdar, K. V.; Pouliot, J. J.; Zhou, T.; Lees-Miller, S. P.; Rasouli-Nia, A.; et al. Conversion of phosphoglycolate to phosphate termini on 3′ overhangs of DNA double-strand breaks by the human tyrosyl-DNA phosphodiesterase. *J. Biol. Chem.* **2002**, *277*, 27162−27168.
(14) Interthal, H.; Champoux, J. Unpublished results.
(15) Pouliot, J. J.; Robertson, C. A.; Nash, H. A. Pathways for repair of topoisomerase I covalent complexes in *Saccharomyces cerevisiae*. *Genes Cells* **2001**, *6*, 677−687.
(16) Vance, J. R.; Wilson, T. E. Yeast Tdp1 and Rad1−Rad10 function as redundant pathways for repairing Top1 replicative damage. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 13669−13674.
(17) Takashima, H.; Boerkoel, C. F.; John, J.; Saifi, G. M.; Salih, M. A. M.; et al. Mutation of TDP1, encoding a topoisomerase I-dependent DNA damage repair enzyme, in spinocerebellar ataxia with axonal neuropathy. *Nat. Genet.* **2002**, *32*, 267−272.
(18) Bodley, A.; JN, C.; Shapiro, T. Effects of camptothecin, a topoisomerase I inhibitor, on *Plasmodium falciparum*. *Biochem. Pharmacol.* **1998**, *55*, 709−711.
(19) Stewart, L.; Redinbo, M. R.; Qiu, X.; Hol, W. G.; Champoux, J. J. A model for the mechanism of human topoisomerase I. *Science* **1998**, *279*, 1534−1541.
(20) Davies, D. R. Unpublished results.
(21) Been, M. D.; Burgess, R. R.; Champoux, J. J. Nucleotide sequence preference at rat liver and wheat germ type I DNA topoisomerase breakage sites in duplex SV40 DNA. *Nucleic Acids Res.* **1984**, *12*, 3097−3114.
(22) Gordon, J. A. Use of vanadate as protein−phosphotyrosine phosphatase inhibitor. *Methods Enzymol.* **1991**, *201*, 477−482.
(23) Otwinowski, Z.; Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **1997**, *276*, 307−326.
(24) Collaborative Computational Project Number 4. The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr.* **1994**, *D50*, 760−763.
(25) McRee, D. E. XtalView/Xfit−a versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **1999**, *125*, 156−165.
(26) Murshudov, G. N.; Vagin, A. A.; Lebedev, A.; Wilson, K. S.; Dodson, E. J. Efficient anisotropic refinement of macromolecular structures using FFT. *Acta Crystallogr.* **1999**, *D55*, 247−255.
(27) Carson, M. Ribbons. *Methods Enzymol.* **1997**, *277*, 493−505.

JM030487X