Mode of action of DNA-competitive small molecule inhibitors of tyrosyl DNA phosphodiesterase 2

Article (Accepted Version)


This version is available from Sussex Research Online: http://sro.sussex.ac.uk/60620/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
Mode of action of DNA-competitive small molecule inhibitors of tyrosyl DNA phosphodiesterase 2

Peter Hornyak\textsuperscript{1,2}, Trevor Askwith\textsuperscript{3}, Sarah Walker\textsuperscript{3}, Emilia Komulainen\textsuperscript{1}, Michael Paradowski\textsuperscript{3}, Lewis E. Pennicott\textsuperscript{3}, Edward J. Bartlett\textsuperscript{1,5}, Nigel C. Brissett\textsuperscript{1}, Ali Raoof\textsuperscript{4}, Mandy Watson\textsuperscript{4}, Allan M. Jordan\textsuperscript{4}, Donald J. Ogilvie\textsuperscript{4}, Simon E. Ward\textsuperscript{3}, John R. Atack\textsuperscript{3}, Laurence H. Pearl\textsuperscript{2*}, Keith W. Caldecott\textsuperscript{1*}, Antony W. Oliver\textsuperscript{2*}

*corresponding authors

1. Genome Damage and Stability Centre, School of Life Sciences, University of Sussex, Falmer, BN1 9RQ, UK
2. Cancer Research UK DNA Repair Enzymes Group, Genome Damage and Stability Centre, School of Life Sciences, University of Sussex, Falmer, BN1 9RQ, UK
3. Sussex Drug Discovery Centre, School of Life Sciences, University of Sussex, Falmer, BN1 9QJ, UK
4. Drug Discovery Unit, Cancer Research UK Manchester Institute, University of Manchester, Manchester, M20 4BX, UK
5. Current Address: Health Research Innovation Centre, The University of Calgary, T2N 4Z6, Canada

Tel: +44 1273 876544 Fax: +44 1273 877856; E-mail: laurence.pearl@sussex.ac.uk
Tel: +44 1273 875519; Fax: +44 1273 678121; E-mail: k.w.caldecott@sussex.ac.uk
Tel: +44 1273 678349; Fax: +44 1273 678121; E-mail: antony.oliver@sussex.ac.uk
ABSTRACT
TDP2 is a 5'-tyrosyl DNA phosphodiesterase important for the repair of DNA adducts generated by non-productive (abortive) activity of topoisomerase II. TDP2 facilitates therapeutic resistance to topoisomerase poisons, which are widely used in the treatment of a range of cancer types. Consequently, TDP2 is an interesting target for the development of small molecule inhibitors that could restore sensitivity to topoisomerase-directed therapies. Previous studies identified a class of deazaflavin-based molecules that showed inhibitory activity against TDP2 at therapeutically useful concentrations, but their mode of action was uncertain. We have confirmed that the deazaflavin series inhibits TDP2 enzyme activity in a fluorescence-based assay, suitable for HTS-screening. We have gone on to determine crystal structures of these compounds bound to a ‘humanised’ form of murine TDP2. The structures reveal their novel mode of action as competitive ligands for the binding site of an incoming DNA substrate, and point the way to generating novel and potent inhibitors of TDP2.

(155 words)
INTRODUCTION

TDP2 (also known as TTRAP [1] or EAPII [2]) is a 5'-tyrosyl DNA phosphodiesterase that can remove the covalently-attached abortive complexes generated by DNA Topoisomerase II (TOP2) when it is unable to complete its normal catalytic cycle [3, 4]. TDP2 has also been co-opted by picornaviruses as the 'unlinkase' that releases nascent RNA from its VPg protein primer during viral replication [5].

TOP2 generates transient double-strand breaks (DSBs) in genomic DNA that relieve torsional stress introduced during transcription and replication [6]. The mechanism of DNA strand cleavage by TOP2 involves formation of a covalent reaction intermediate between the side-chain of a catalytic tyrosine in the enzyme, and the 5'-phosphate at the site of the DNA-strand break. TOP2 functions as a dimer, and produces a staggered DSB, with the nicks in the two strands separated by four base pairs. Normally, following passage of a distal segment of duplex DNA through the break, TOP2 facilitates attack on the phosphotyrosyl bonds by the free 3'-OH at the nicks, resealing the break and releasing the enzyme. In some circumstances TOP2 is unable to complete the reaction cycle, leaving it covalently attached to the DSB 5'- termini in a highly toxic, abortive, 'cleavage complex' that can severely disrupt replication and transcription. Fortunately, this type of DNA lesion is efficiently repaired, in part though the action of TDP2, which hydrolyses the phosphotyrosyl bond in a magnesium cofactor-dependent reaction, leaving ligatable 5'-phosphate moieties.

TOP2 'poisons' such as doxorubicin (adriamycin) are known to intercalate into DNA (PDB: 1D12; [7]); amsacrine (mAMSA) and etoposide (VP-1) have also been demonstrated to bind selectively at the nick sites of a TOP2-DNA complex (PDB: 4G0U, 3QX3; [8]) and prevent resealing of the strands. These interactions promote the cumulative formation of abortive TOP2-DNA complexes, leading ultimately to cell death in rapidly proliferating cells. Consequently, TOP2 poisons are potent anticancer agents [9], and are currently used in the clinic either as monotherapies, or as components of combination therapies for a range of tumour types [10-12]. However, the abortive complexes promoted by TOP2-poisons can be rescued by the action of TDP2, so that variations in cellular TDP2 activity (natural or tumour-specific) can influence the clinical efficacy of these drugs, and are therefore an important determinant of response to chemotherapy in the individual patient and tumour.
Consequently, TDP2 has attracted interest as a target for small molecule inhibitors that might have utility in overcoming innate or acquired resistance to TOP2 poisons.

A series of potent TDP2 inhibitors based on a deazaflavin scaffold have previously been described, but their molecular mode of action could not be determined and the rational exploitation of structure-activity relationships was consequently limited [13]. Using a ‘humanised’ mouse TDP2 expression construct, we have now determined X-ray crystal structures of the catalytic domain of TDP2 in complex with deazaflavin inhibitors, revealing the unusual molecular interactions underpinning their mode of action. These results provide a platform for structure-based drug design that will greatly assist the future development of small molecule inhibitors of human TDP2.

RESULTS
Structure of the human TDP2 catalytic domain
We obtained crystals of the catalytic domain of human TDP2 (FIGURE 1A; hTDP2\textsuperscript{CAT} – residues 113-362), in its unliganded form, using protein expressed in \textit{Escherichia coli} and purified using standard chromatographic procedures (see MATERIALS AND METHODS). The expression construct contains a single-point mutant – Cys273Ser – which removes a potentially reactive surface cysteine, and that in our hands significantly improved the expression levels and solubility of the recombinant protein.

The structure of hTDP2\textsuperscript{CAT} was determined by molecular replacement, with the catalytic domain of mouse TDP2 as a search model (PDB: 4GYZ). However, crystals of hTDP2\textsuperscript{CAT} typically diffracted both poorly and anisotropically. Through iterative screening of many crystals, we were eventually able to collect a dataset that could be truncated isotropically to a maximum resolution of 3.1 Å (see TABLE 1). The fold of the human protein is highly similar to that of mouse [14], worm [15], and zebrafish [15] (FIGURE 1B,C) but with some important differences – see expanded description below.

Deazaflavin inhibitors of TDP2
A series of small molecule deazaflavin inhibitors of human TDP2 have previously been described [13], with examples showing \textit{in vitro} biochemical activity against the isolated
enzyme in the nanomolar concentration range (FIGURE 2A). However, the precise mode of action of these compounds was unclear, and in the absence of an experimentally-determined binding site, the structure activity relationships (SAR) were difficult to rationalise.

As the original HTS screens were performed with full-length human TDP2, we first sought to confirm binding of these compounds to hTDP2CAT by thermal denaturation [16]. Using the most potent reported deazaflavin – compound 163 [13] we observed a shift in the melting temperature (Tm) of hTDP2CAT from 48.7°C in the unliganded state, to 51.5°C in the presence of compound 163 at a concentration of 300 μM. These results are indicative of moderately tight binding of 163 to hTDP2CAT (FIGURE 2B).

We also confirmed inhibition of catalytic activity by three deazaflavin compounds, by using a modified form of a fluorescence-based assay [17]. A phosphotyrosyl moiety conjugated to Fluorescein isothiocyanate was readily removed from the 5’ end of a single-stranded oligonucleotide substrate by hTDP2CAT in a dose-dependant manner (SUPPLEMENTARY FIGURE S1 A,B). We could subsequently determine Km for the oligonucleotide substrate (208 ± 23 nM), Kcat for the reaction under the experimental conditions tested (1.58 ± 0.06 sec⁻¹; FIGURE 2C), as well as IC₅₀ values for each of the deazaflavins; 163: 19, 148: 103, 149: 489 nM (FIGURE 2D). The IC₅₀ rank order is in agreement with the previously reported EC₅₀ values, which used full-length human TDP2 and BIOMOL Green as a detection/quantification agent [13].

**Low resolution co-crystal structure of hTDP2CAT bound to 163**

As with the apo-hTDP2CAT crystals, co-crystals with 163 presented severe problems with anisotropy during data collection. Again, multiple crystals were screened before a dataset that could be truncated isotropically, to a maximum resolution of 3.4Å, was obtained (TABLE 1). Despite the low resolution, which prevents detailed characterisation of specific protein-drug interactions, clear difference electron density maps identified the binding site and general pose of this class of inhibitor bound to TDP2 (SUPPLEMENTARY FIGURE S2).
The experimentally determined binding-mode of 163 is surprisingly different from that predicted by molecular docking studies [13], in which the pyrimidodione ring of the tricyclic core, was hypothesised to bind to the magnesium cofactor – required for catalytic activity of TDP2 – and made few (if any) hydrogen bonds directly to the protein.

Instead, our data indicate that the tricyclic deazaflavin core of 163 unexpectedly packs against an extended hydrophobic patch, formed by the side chains of Trp297, Thr299, Leu305, Ile 307, Cys311, Leu313 and Phe315, which lies at the mouth of the channel leading to the active site, more than 9Å from the catalytic Mg$^{2+}$ ion bound by the carboxyl side chains of Asp122 and Glu152 (FIGURE 2E).

Due to the persistently poor crystallisation behaviour of the human protein, and the need to characterise fully this unusual and unexpected binding mode of the deazaflavin inhibitor, we switched species and used the murine form of the enzyme (mTDP2$^{\text{CAT}}$), which yielded well-behaved crystals diffracting to high resolution; as in previous studies [14]. As a preliminary experiment we tested the ability of 163 to inhibit mTDP2$^{\text{CAT}}$ in a gel-based assay. Unexpectedly, whilst 163 readily inhibited the catalytic activity of hTDP2$^{\text{CAT}}$ – preventing conversion of 5'-phospho-tyrosyl termini to 5'-phosphates on a DNA duplex substrate – it failed to inhibit the mouse protein (FIGURE 2F).

Comparison of the mouse and human TDP2 amino acid sequences revealed substantial differences in a segment of polypeptide, denoted the ‘M7-motif’ [14], delimited by highly conserved ‘TWDT’ (amino acids 296-299 in humans) and ‘RFDR’ (aa 314-317) motifs. A multiple amino acid sequence alignment of this region (FIGURE 3A) shows it to be conserved in length (14 residues) amongst different organisms, but poorly conserved in amino acid sequence and somewhat variable in conformation (FIGURE 3B). Significantly, the M7-motif maps to the binding site we observe for 163 in hTDP2$^{\text{CAT}}$, where human TDP2 residues Cys311 and Leu313, along with Ala309, coalesce to form a ‘hydrophobic-shelf’ on which the deazaflavin core of 163 sits. Replacement of Cys311 and Leu313 in hTDP2 with the equivalent residues Tyr321 and His323 in mTDP2 disrupts this hydrophobic-shelf and explains the resistance of the mouse enzyme to the deazaflavin-based inhibitors (FIGURE 3C).
Development of a protein surrogate: m2hTDP2\textsuperscript{CAT}

Although hTDP2\textsuperscript{CAT} is effectively inhibited by 163 and related deazaflavins, its poor crystallographic behaviour greatly limits the degree to which structure-led approaches can be used in order to optimise these compounds further for clinical development. Conversely, mTDP2\textsuperscript{CAT}, while crystallographically well-behaved and able to yield high resolution structures, does not bind deazaflavin inhibitors due to sequence differences in the M7-motif.

To overcome these problems, we designed a series of mutations in the M7-motif and associated structure, designed to ‘humanise’ the inhibitor-binding site of mTDP2 while retaining the favourable murine surface that facilitates well-ordered crystals. Rather than exchange the entire mouse M7-motif for that of the human protein, we decided as an initial approach only to ‘humanise’ the amino acids expected to be in the proximity of the bound compound, and not those which would simply point towards solvent. We therefore created the expression construct m2hTDP2\textsuperscript{CAT} (m2h = mouse-‘to’-human) by mutating the following mouse residues to their human equivalent: Glu242Gly, Gln278Arg, Tyr321Cys and His323Leu.

The m2hTDP2\textsuperscript{CAT} protein retained full phosphodiesterase activity, and consistent with our prediction that the resistance of the mouse protein to deazaflavins was due to the variant M7-motif, this activity in m2hTDP2\textsuperscript{CAT} was inhibited by 163 (FIGURE 3D). As no further optimisation of the ‘humanised’ construct was therefore required, m2hTDP2\textsuperscript{CAT} was used for all subsequent structural analyses.

High-resolution co-crystal structure of m2hTDP2\textsuperscript{CAT} bound to 163

We were able to validate our approach, by determining the X-ray crystal structure of m2hTDP\textsuperscript{CAT} bound to 163, at a resolution of 1.8 Angstrom (Table 1). Superposition of the low resolution hTDP2\textsuperscript{CAT} and high-resolution m2hTDP2\textsuperscript{CAT} structures in complex with 163, confirms that the binding site is conserved across both proteins (SUPPLEMENTARY FIGURE S3). However, the higher resolution data obtained with the m2hTDP2\textsuperscript{CAT} surrogate, allows a far more detailed characterisation of the interaction of the inhibitor with the TDP2 catalytic domain (FIGURE 4A, SUPPLEMENTARY FIGURE S2).
The guanidinium head group of Arg276† is involved in an extensive hydrogen-bond network, including both carbonyl and nitrogen groups of the 163 pyrimidodione ring, as well as the backbone carbonyl and side chain (via a water molecule) of Asp 277. As in the fully human protein, amino acids of the M7-motif coalesce to form the ‘hydrophobic shelf’, which permits extensive van der Waals interactions with the tricyclic deazaflavin core; these residues include Trp307, Ile317, Ala319, Cys321‡, Leu323 and Phe325 (FIGURE 3A). The tetrazole R1 substituent of 163, is directed towards solvent, but packs against the carbon-atoms of the Arg241 sidechain, providing additional van der Waals interactions.

**High-resolution co-crystal structure of m2hTDP2CAT bound to 148**

Raoof et al. also described and characterised two hydroxyphenyl substituents at the R1 position of the deazaflavin core, in combination with the cyano group at the R3 position: compound 147 (3-hydroxyphenol) and 148 (4-hydroxyphenol) with reported EC50 values of 50 nm and 90 nM respectively [13] (FIGURE 2A). We were successful in obtaining diffracting crystals of 148 in complex with m2hTDP2CAT, determining the structure at a resolution of 1.7 Ångstrom (TABLE 1). In this instance, we also supplemented the purification buffers with manganese chloride (iso-structural with magnesium, but readily discernible in electron density maps) in order to more fully define any interaction of the inhibitors with the metal-dependent catalytic centre of the enzyme.

In the 148 complex the side chains of Arg276 and Asp277 make the same set of interactions with the carbonyl and nitrogen groups of the pyrimidodione ring, as with 163. However, the side chain of Arg241 occupies a different position, breaking its hydrogen bond interaction with Asp277, and instead packing in a π-π interaction against the face of the tricyclic core of the deazaflavin (FIGURE 4B, SUPPLEMENTARY FIGURE S2). Interestingly, the para-hydroxyl group of the R1 substituent in 148, points directly towards the catalytic core of the protein and makes a series of water-mediated contacts to the side chains of Asp272 and Asn274. Electron density maps also clearly showed the position of a tartrate molecule (picked up from the crystallisation mother liquor), which is hydrogen-bonded to the side chain of Ser239, as well as an extensive series of water-mediated interactions with several amino acid residues surrounding the catalytic centre. The tartrate

---

† all numbering is for the mouse protein, the equivalent in human is calculated by subtraction of 10.
‡ text in italics represent a ‘humanised’ amino acid residue.
molecule also completes the octahedral coordination of the Mn\(^{2+}\) ion (marking the position of the catalytic Mg\(^{2+}\) cofactor) (FIGURE 4C). Moreover, its position marks a potentially useful 'molecular interaction path' that could be utilised in future iterations of TDP2 inhibitor compounds.

**Mode of action of deazaflavin inhibitors of TDP2**

Several structures for the catalytic domain of TDP2 in complex with nucleic acid have been reported [14, 15]. However, there is disagreement between these structures regarding the path that the bound DNA follows across the protein surface; potentially as a result of the different substrates used (double vs single-stranded) or possibly due to different crystal packing interactions.

However, taking PDB entries 4GZ1 and 4F1H as representative examples of the two different modes of TDP2 – product DNA interaction, we note that although the DNA trajectories are different (FIGURE 5A) the position of the first (Nuc1) and second nucleotides (Nuc2) of the previously modified strand are highly similar in both structures (FIGURE 5B).

Regardless of which of these TDP2 — product DNA structures actually represents a physiologically relevant binding mode, comparison with the structures reported here shows that the deazaflavins occupy a position equivalent to that of Nuc2, with the pendant group sitting at the position of the 5'-phosphate, and the deazaflavin core at the position of the ribose and base moieties.

**Discussion**

The proteins that constitute the DNA Damage Response (DDR) are increasingly being recognised as attractive targets for the development of new drugs, which exploit the inherent genomic instability found in many cancers [18, 19]. This field has been galvanised by the approval of olaparib as a monotherapy for the treatment of specific subsets of patients with advanced ovarian cancer. Olaparib is an inhibitor of the DNA damage sensor PARP1, which exploits cancer-associated genetic defects in one DNA repair system (homologous recombination-mediated DSB repair) by inhibiting a second system (single-strand break repair) to achieve potent and selective 'synthetic lethality' against tumour cells [20].
DDR enzymes that utilise cofactors – such as PARPs or protein kinases for example – fall within the traditional canon of drug development targets; i.e. where a specific cofactor-competitive inhibitor can be readily found. However, many biologically interesting DNA repair ‘targets’ for cancer drug discovery lack such cofactors and instead act directly on DNA. These present the challenge of discovering compounds capable of out-competing an often substantial protein-DNA interaction potentially dominated by highly polar interactions with the negatively charged sugar-phosphate backbone.

The discovery of deazafavins as TDP2 inhibitors [13] highlighted the possibility that uncharged and ‘drug-like’ competitive inhibitors of DNA binding to a DNA repair enzyme, could be found and developed. However, the mode of action of these was unclear, and limited further optimisation of these compounds towards clinical development. The structural analysis we describe here reveals a very unusual binding mode in which the deazaflavin core and pendant groups act as nucleoside mimetics, binding to the hydrophobic ‘shelf’ that can support the exposed face of the last base-paired nucleotide at the transition from double-stranded to single-stranded DNA in the mouse TDP2$^{\text{CAT}}$ — product DNA structure (PDB: 4GZ1) Although this deazaflavin series is specific for TDP2, many enzymes involved in DNA repair and manipulation engage with a dsDNA-ssDNA transition, and structural studies of their complexes with DNA reveal the presence of functionally equivalent ‘shelf’ structures that could also provide binding sites for DNA-competitive small ligands [21-23] that could have utility as selective inhibitors of the DNA repair pathways such enzymes mediate.

Interestingly, TDP2 has recently also been identified as a host factor required for the replication of picornaviruses in mammalian cells [5]; encoding the essential ‘unlinkase’, required to cleave a unique covalent interaction made between the virally-encoded protein VPg and RNAs produced by the virion. It is therefore feasible that small molecule inhibitors of TDP2 may also find utility as novel anti-viral agents, for the treatment of picornoviral infections; which includes rhinovirus and enterovirus pathogens, both of which produce high morbidity rates in infants and children [5, 24].

It is worth noting here (as previously reported [13]) that deazaflavin-based compounds are, in themselves, highly unlikely to be suitable as anti-cancer agents; in most part due to the
general inability of the chemical scaffold to penetrate the cell membrane of mammalian cells. However, with the data gleaned from the high-resolution crystal structures presented here, we now have a detailed understanding of their mode of action, which will in turn, inform the discovery and development of alternative scaffolds, and may eventually lead to the conversion of this highly useful ‘tool compound’ into a potential *bona fide* drug candidate for the treatment of disease.
MATERIALS AND METHODS

Expression Constructs

hTDP2\textsuperscript{CAT}

DNA encoding the catalytic domain of human TDP2 (amino acids 113-362) was cloned into pET-28a (Novagen, Merck Chemicals, Nottingham, UK). Using standard site-directed mutagenesis techniques, a single point mutant – Cys273Ser – was subsequently introduced into this expression construct; improving the behaviour of the recombinant protein, qualitatively with respect to both solubility and aggregation.

mTDP2\textsuperscript{CAT} / m2hTD2\textsuperscript{CAT}

Synthetic genes codon-optimised for \textit{E.coli} expression were sub-cloned into a in-house modified form of pET-15b, which encodes a N-terminal His-SUMO affinity/solubility tag that is cleavable by SENP1 protease (Genscript, Piscataway NJ, USA). The mTDP2\textsuperscript{CAT} expression construct encodes amino acids Leu118-Leu370 of the full-length gene. The m2hTDP2\textsuperscript{CAT} construct encodes the same amino acid range, but with the following series of mutations: Glu242Gly, Gln278Arg, Tyr321Cys and His323Leu.

Protein Expression and Purification

hTDP2\textsuperscript{CAT}

Transformed colonies of \textit{E.coli} strain B834(DE3) were used to inoculate a 250-ml flask containing 50 ml of Turbo broth (Molecular Dimensions, Newmarket, UK) supplemented with kanamycin (50 µg/ml). The inoculated culture was grown at 37°C, at 220rpm, in an orbital shaking incubator until the optical density at 600 nm reached ~1.5. The culture was then stored overnight at 4°C.

The following day, 20 ml of the overnight culture was used to inoculate a 2-l flask containing 1 l of Turbo-broth supplemented with antibiotic as before. Cultures were grown at 37°C, at 220rpm, in an orbital shaking incubator until the optical density at 600 nm again reached ~1.5. Recombinant protein expression was then induced by the addition of 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG), and the culture incubated for a further 3.5 hours at a reduced temperature of 30°C. Cells were then harvested by centrifugation, and the resulting cell pellet stored at -80°C until required.
The cell pellet arising from 4 l of culture was resuspended in buffer A: 50 mM HEPES·NaOH pH 7.5, 250 mM NaCl, 10 mM imidazole, 0.5 mM TCEP, supplemented with protease inhibitors (Roche, Burgess Hill, UK), then disrupted by sonication, and the resulting lysate clarified by high-speed centrifugation at 40 000 x g, for 60 minutes at 4°C.

The supernatant was applied to a batch/gravity flow column containing 10 ml of Talon resin (TaKaRa Bio, Saint-Germain-en-Laye, France) pre-equilibrated in buffer A. After incubation (with rolling) for a period of 1h at 4°C, any unbound material was removed by sequential washes with buffer A. Retained protein was then eluted by application of the same buffer supplemented with 300 mM imidazole pH 7.5.

The eluate was diluted, in order to reduce the salt concentration to < 150 mM NaCl, and then applied to a 5ml HiTrap Heparin HP column (GE Healthcare, Little Chalfont, UK) equilibrated in buffer B: 50 mM HEPES·NaOH pH7.5, 150 mM NaCl, 1 mM TCEP. Unbound material was first removed by application of 5 column volumes of the same buffer. Any retained protein was eluted, through application of a linear salt gradient (0.15 to 1 M NaCl) to the column.

A HiLoad 16/600 Superdex 75pg size exclusion column (GE Healthcare) was used to purify hTDP2-CAT to homogeneity in 20 mM HEPES·NaOH pH7.5, 300 mM NaCl, 0.5 mM TCEP.

mTDP2CAT / m2hTDP2CAT
A similar protocol was also used to purify both mTDP2CAT and m2hTDP2CAT, with the following alterations: E.coli strain: Rosetta 2(DE3)pLysS (Novagen); antibiotic selection: 50 µg/ml ampicillin / 34 µg/ml chloramphenicol. The His-SUMO affinity/solubility tag was cleaved overnight at 4°C by the addition of SENP1 protease, after the initial Talon capture step.

Thermal Denaturation
For thermal denaturation, samples containing protein at 1.7 µM and 5 x SYPRO Orange (diluted from a 5000 x stock supplied in DMSO; Sigma-Aldrich, Gillingham, UK) were prepared in sample buffer: 50mM HEPES·NaOH pH 7.5, 300mM NaCl, 0.5mM TCEP. Either DMSO alone, or compound (dissolved in 100% v/v DMSO at its maximum solubility) was then added; producing a final concentration of 3% v/v DMSO. Denaturation curves
were monitored in 96-well PCR plates using a Roche LightCycler 480 II, with 465 and 580 nm filters for excitation and emission wavelengths, respectively. The program was as follows: 1 min at 20°C, followed by a continuous increment of 0.03°C/s to a final temperature of 85°C. Temperature midpoints (Tm) for each folded to unfolded transition were determined by non-linear regression fitting of a modified Boltzmann model to normalized data in Prism 6.0 (version 6.0h, GraphPad Software, La Jolla, CA USA).

\[
Y = \frac{(a_n X + b_n) + (a_d X + b_d)}{1 + e^{-m(X - T_m)}}
\]

Where: \(a_n\) and \(a_d\) are the slopes, and \(b_n\) and \(b_d\) the y-intercepts, of the native and denatured baselines respectively. \(T_m\) is the temperature midpoint of the transition from native to denatured states, and \(m\) represents a generic slope factor.

**Fluorescence-based Enzyme Activity Assay**

Substrate oligonucleotide containing a 5’ phosphotyrosine (5’-Y-GATCTAAAAGACT-3’) conjugated to Fluorescein isothiocyanate (FITC) was purchased from Midland Certified Reagents (Midland, TX, USA).

This assay is a modified version of that reported by [17]. Briefly, enzyme assays were run in black 384 well plates, in 50 mM Tris.HCl pH 8.0, 10 mM MgCl₂, 80 mM KCl, 0.05% (v/v) Tween-20, 1 mM DTT. 50 pmole of hTDP2CAT was combined with 25 nM substrate oligonucleotide, in a final volume of 15 µl, for a period of 10 minutes, before addition of quench reagents (Gyrasol Technologies, Kansas, USA; sensor diluted 1:15 with enhancer buffer). Fluorescence was subsequently measured in a PHERAstar multimode plate reader (BMG Labtech GmbH, Ortenberg, Germany) with an excitation wavelength of 490nm, and an emission wavelength of 520nm. The built-in analysis tools of Prism 6.0 was used to examine all experimental data.

**Gel-based Enzyme Activity Assay**

This assay is a modified version of that reported by [3, 4]. Briefly recombinant TDP2CAT proteins were diluted with reaction buffer: 50 mM Tris-HCl pH 7.5, 50 mM KCl, 1 mM MgCl₂,
1 mM DTT to produce final concentrations of 1nM (for hTDP2CAT and m2hTDP2CAT) or 3nM (for mTDP2CAT) in a volume of 16.2 µl – either in the presence or absence of compound (50 nM for hTDP2CAT and m2hTDP2CAT, 150 nM for mTDP2CAT; 1% v/v final DMSO concentration). To start the reaction 1.8 µl of radio-labelled 5' phosphotyrosyl DNA substrate was added to the mixture, then incubated at 37 °C. At 2, 5, and 10 minute time-points, 6 µl of the reaction mix was withdrawn, and stopped by the addition of formamide loading buffer. Samples were then analysed by denaturing PAGE, visualised by phosphorimager, with intensities of substrate and product bands measured by GelAnalyzer2010 software (http://www.gelanalyzer.com).

**Crystallisation and data collection**

**hTDP2CAT**

Crystals of hTDP2CAT were grown at 20 °C in 24-well hanging-drop vapour-diffusion plates, mixing 1 µl of protein at 7.5 mg/ml with 1 µl of 100 mM Bis-Tris propane pH 7.0, 0.5 M NaCl, 0.05 M magnesium acetate, 1.5% v/v Trimethylamine N-oxide, equilibrated against 500 µl of the same solution. Crystals were swiped successively through buffers containing increasing concentrations of cryo-protectant, before being plunged into liquid nitrogen; 30% v/v glycerol was sufficient to prevent ice-formation.

Diffraction data to 3.1 Å resolution were collected at the Diamond Light Source (DLS, Didcot, UK) on beamline I04. Crystals grew in space group P3, 2 1, with two molecules of hTDP2CAT comprising the asymmetric unit.

**hTDP2CAT/163**

Crystals of hTDP2CAT/163 were grown at 20 °C in 24-well hanging-drop vapour-diffusion plates, mixing 1 µl of complex (1 : 2.7 molar ratio protein : compound) at ~7.5 mg/ml, with 1 µl of 1.2M D/L-Malic acid pH7.0, 0.1M Bis-Tris propane pH 7.0, 3% v/v DMSO, equilibrated against 500 µl of the same solution. Crystals were swiped successively through buffers containing increasing concentrations of cryo-protectant, before being plunged into liquid nitrogen; 40% w/v sucrose was sufficient to prevent ice-formation.

Diffraction data to 3.4 Å resolution were collected at the Diamond Light Source (DLS, Didcot, UK) on beamline I04. Crystals grew in space group P6₅ 2 2, with a single molecule of hTDP2CAT in complex with 163 comprising the asymmetric unit.
**m2hTDP2\textsuperscript{CAT}/163**

Crystals of m2hTDP2\textsuperscript{CAT}/163 were grown at 20 °C in 24-well hanging-drop vapour-diffusion plates, mixing 1 µl of complex (1 : 2.7 molar ratio protein : compound) at ~7.5 mg/ml, with 1 µl of 0.1M Bis-Tris propane pH 7.5, 0.2 M sodium citrate, 20% w/v PEG3350, 0.3% v/v DMSO, equilibrated against 500 µl of the same solution. Crystals were swiped successively through buffers containing increasing concentrations of cryo-protectant, before being plunged into liquid nitrogen; 8% v/v glycerol + 5% w/v glucose was sufficient to prevent ice-formation.

Diffraction data to 1.8 Å resolution were collected at the Diamond Light Source (DLS, Didcot, UK) on beamline I04. Crystals grew in space group P2\textsubscript{1}, with two molecules of m2hTDP2\textsuperscript{CAT}, both in complex with 163, comprising the asymmetric unit.

**m2hTDP2\textsuperscript{CAT}/148**

Crystals of m2hTDP2\textsuperscript{CAT}/148 were grown at 20 °C in 24-well hanging-drop vapour-diffusion plates, mixing 1 µl of complex (1 : 2.7 molar ratio protein : compound) at ~7.5 mg/ml, with 1 µl of 0.1M Bis-Tris propane pH 6.5, 0.2 M sodium tartrate, 20% w/v PEG3350, 0.3% v/v DMSO, equilibrated against 500 µl of the same solution. Crystals were swiped successively through buffers containing increasing concentrations of cryo-protectant, before being plunged into liquid nitrogen; 8% v/v glycerol + 5% w/v glucose was sufficient to prevent ice formation.

Diffraction data to 1.7 Å resolution were collected at the Diamond Light Source (DLS, Didcot, UK) on beamline I04. Crystals grew in space group P2\textsubscript{1}, with two molecules of m2hTDP2\textsuperscript{CAT}, both in complex with 148, comprising the asymmetric unit.

**Phasing, model building and refinement**

All diffraction data were collected at 100K. Data were integrated using the software package XDS [25], and then processed using the Pointless / Aimless / Ctruncate pipeline of the CCP4 software suite [26-28] [29]. Phases were obtained by molecular replacement, using PHASER with 4GYZ (mTDP2-CAT) as a search model [30]. An iterative combination of manual building in Coot [31] and refinement with either phenix.refine [32] or BUSTER [33] produced the final models.
Compounds
Compounds were provided either by the Cancer Research UK Manchester Institute Drug Discovery Unit or were re-synthesised in-house by LP or MP, following the published protocols [13].

Accession Numbers
5J3P, 5J3S, 5J3Z and 5J42 for apo-hTDP2CAT, hTDP2CAT+163, m2hTDP2CAT+163 and m2hTDP2CAT+148 respectively.

Funding
Cancer Research UK Programme Grants C302/A14532 (LHP and AWO), C6563/A16771 (KWC) and C480/A11411 (DJO).

Acknowledgements
We thank Mark Roe for assistance with X-ray diffraction data collection. We also thank Diamond Light Source, Didcot, for access to synchrotron radiation.
**Figure Legends**

**Figure 1. Structure of the catalytic domain of human TDP2**
(A) Schematic representation of human TDP2, highlighting the relative positions of its component domains; where UBA-L = Ubiquitin associated-like domain. The amino acid boundaries of the hTDP2\textsuperscript{CAT} expression construct are indicated by the dotted box. The position of the Cys273Ser mutation introduced to improve protein solubility is also indicated.
(B) Superposition of TDP2 catalytic domains. Structures for human, mouse, worm and zebrafish TDP2\textsuperscript{CAT} are each shown in ribbon representation. RMSD values, over the indicated number of equivalent atom positions, were calculated using PyMOL (www.pymol.org).
(C) Stereo-view of the X-ray crystal structure of hTDP2\textsuperscript{CAT}. Molecular cartoons are coloured from blue \(\rightarrow\) red, from the visible N- to C-terminus. The position of the M7-loop or motif is additionally highlighted.

**Figure 2. Small molecule Inhibitors of human TDP2**
(A) Chemical drawings of four deazaflavin-based inhibitors of human TDP2, as described by [13]. Compound numbering and EC\textsubscript{50} values are as previously reported.
(B) Thermal denaturation profiles for hTDP2\textsuperscript{CAT} in the absence or presence of 300 \(\mu\text{M}\) 163. The experimental data (data points coloured in black) were fitted with a modified Boltzmann model to obtain the indicated temperature midpoints (T\textsubscript{m}) for the apo- (red curve) or liganded forms (blue curve) of the protein.
(C) Michaelis-Menten Curve for hTDP2\textsuperscript{CAT} using the Fluorescence-based enzyme assay. 50 pmole of hTDP2\textsuperscript{CAT} was incubated with substrate oligonucleotide for a period of 10 minutes, before the addition of Gyrasol reagents. Experimental data were fitted with a standard Michaelis-Menten equation (Prism 6, Graphpad software) in order to determine \(K_m\) and \(V_{\text{max}}\) parameters. Data points are the mean of 3 independent experiments, with error bars representing 1 standard deviation.
(D) Dose response curves for inhibition of hTDP2\textsuperscript{CAT} by increasing concentrations of the indicated deazaflavin. See associated key for details, and IC\textsubscript{50} values determined. Data points are the mean of 6 replicates, with error bars representing 1 standard deviation.
(E) Molecular cartoon highlighting the binding position of 163, relative to the active site of hTDP2\textsuperscript{CAT}. The compound packs against an extended hydrophobic shelf (light grey molecular surface + stick representation) formed by the side-chains of Trp297, Thr299,
Leu305, Ile307, Cys311, Leu313 and Phe315. The compound binds at a distance (> 9Å) from the active site of the enzyme, as indicated by the sidechains of Asp122 and Glu152, and the magnesium cofactor (purple sphere).

(F) Top; Representative image for the gel-based enzyme activity assay. hTDP2\textsuperscript{CAT} efficiently catalyses the conversion of a 5’ phosphotyrosylated (5’-Y) DNA duplex substrate to a 5’ phosphorylated product (5’-P) with time. Whilst 1\textsuperscript{63} clearly inhibits the human enzyme, it is ineffective against the mouse equivalent. Bottom; quantification of gel-based assay. Data represent the mean of three independent experiments, with errors bars representing one standard deviation.

**Figure 3. The M7-motif of TDP2**

(A) Multiple amino acid sequence alignment for the M7-motif of TDP2. This region is flanked by the two highly conserved ‘TWDT’ (coloured orange) and ‘RFDR’ (coloured cyan) motifs. Amino acid boundaries and UniProt accession codes are as indicated. Partially conserved amino acids are additionally identified by the bold, underlined text.

(B) Superposition of the M7-motif taken from X-ray crystal structures of human (this study), mouse (PDB:4GYZ), zebrafish (4F1H) and worm TDP2 (4FVA). The spatial arrangement of the ‘TWDT’ and ‘RFDR’ motifs is highly conserved. The loop connecting these two elements is conserved in length, but varies somewhat in conformation. For clarity, numbering is only shown for the human protein.

(C) Molecular cartoon of 1\textsuperscript{63} in complex with hTDP2\textsuperscript{CAT} highlighting the position of two amino acid residues (Leu313 and Cys311; black text on white background) that are not conserved in the M7-motif of mouse TDP2 (equivalent residues His323 and Tyr231; white text on grey background), but are essential for compound binding.

(D) Top; Representative image for the gel-based enzyme activity assay using m2hTDP2\textsuperscript{CAT}. The introduction of four amino acid changes into mTDP2\textsuperscript{CAT} – to generate m2hTDP2\textsuperscript{CAT} – restores the inhibitory activity of 1\textsuperscript{63}. Labelling as per figure 2D. Bottom; quantification of gel-based assay. Data represent the mean of three independent experiments, with errors bars representing one standard deviation.

**Figure 4. Molecular interactions of 163 and 148 with m2hTDP2\textsuperscript{CAT}**

(A) Molecular interactions made by 1\textsuperscript{63}.

(B) Molecular interactions made by 1\textsuperscript{48}.
(C) The para-hydroxyl of the 148 pendant group makes a series of water-mediated interactions with the side chain of Asp 272 and Asn274, in close proximity to the active site of the enzyme.

Molecular cartoons are shown throughout, with selected side chains and compounds shown in stick representation. Amino acids labelled in bold italic text, represent those that have been mutated from their mouse counterpart, in order to generate m2hTDP2\textsuperscript{CAT}. Please see associated key for details of the colouring scheme.

**Figure 5. Deazaflavin binding relative to bound DNA**

(A) The deazaflavin-scaffold interacts with the ‘hydrophobic shelf’ formed by residues of the M7-motif. Molecular surface representations for: (Left) m2hTDP2\textsuperscript{CAT} in complex with 163. (Middle) mouse TDP2\textsuperscript{CAT} in complex with a double-stranded product DNA (PDB: 4GZ1). (Right) zebrafish TDP2\textsuperscript{CAT} in complex with a single-stranded product DNA (PDB: 4F1H).

(B) Superposition of structures reveals that the first two nucleotides of either product DNA (Nuc1, Nuc2) occupy highly similar positions when bound to TDP2. The deazaflavins occupy a position equivalent to Nuc2, with the pendant group sitting at the position of the 5’-phosphate, and the deazaflavin core at those of the ribose and base moieties.

In both panels, the carbons of the previously modified 5’ nucleotide are coloured yellow, and those of the deazaflavin 163 in green.
References


<table>
<thead>
<tr>
<th></th>
<th>hTDP2-CAT</th>
<th>hTDP2-CAT + 163</th>
<th>m2hTDP2-CAT + 163</th>
<th>m2hTDP2-CAT + 148</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>P 3 2 1</td>
<td>P 6 2 2</td>
<td>P 2 1</td>
<td>P 2 1</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>92.54, 92.54, 119.07</td>
<td>68.52, 68.52, 209.51</td>
<td>90.00, 42.80, 108.75</td>
<td>61.05, 42.87, 109.05</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90, 90, 120</td>
<td>90, 90, 120</td>
<td>90, 90, 120</td>
<td>90, 90, 120</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>47.79-3.10 (3.31-3.10)</td>
<td>45.22-3.40 (3.67-3.40)</td>
<td>42.8-1.80 (1.84-1.80)</td>
<td>42.07-1.70 (1.73-1.70)</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.133 (0.731)</td>
<td>0.256 (1.876)</td>
<td>0.051 (0.166)</td>
<td>0.043 (0.810)</td>
</tr>
<tr>
<td>Rfree</td>
<td>0.103 (0.571)</td>
<td>0.168 (1.111)</td>
<td>0.050 (0.164)</td>
<td>0.042 (0.791)</td>
</tr>
<tr>
<td>Mn I/σI</td>
<td>9.6 (1.7)</td>
<td>6.4 (1.1)</td>
<td>9.7 (2.0)</td>
<td>16.3 (1.0)</td>
</tr>
<tr>
<td>CC 1/2</td>
<td>1.00 (0.94)</td>
<td>0.99 (0.54)</td>
<td>0.99 (0.87)</td>
<td>1.00 (0.80)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.8 (99.7)</td>
<td>99.9 (100)</td>
<td>98.0 (79.7)</td>
<td>99.2 (98.9)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>4.8 (4.9)</td>
<td>6.3 (6.8)</td>
<td>2.9 (2.1)</td>
<td>3.1 (3.2)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>47.79-3.10</td>
<td>45.22-3.40</td>
<td>42.8-1.80</td>
<td>42.07-1.70</td>
</tr>
<tr>
<td>No. unique reflections</td>
<td>11128</td>
<td>7545</td>
<td>51379</td>
<td>112934</td>
</tr>
<tr>
<td>Rwork / Rfree</td>
<td>0.26 / 0.30</td>
<td>0.25 / 0.27</td>
<td>0.17/ 0.20</td>
<td>0.16 / 0.20</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>3553</td>
<td>1732</td>
<td>3992</td>
<td>3968</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>8</td>
<td>29</td>
<td>168</td>
<td>142</td>
</tr>
<tr>
<td>Water</td>
<td>7</td>
<td>---</td>
<td>604</td>
<td>693</td>
</tr>
<tr>
<td>B-factors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>98.2</td>
<td>119.9</td>
<td>23.5</td>
<td>26.6</td>
</tr>
<tr>
<td>Ligand/ion</td>
<td>84.3</td>
<td>78.7</td>
<td>42.1</td>
<td>30.9</td>
</tr>
<tr>
<td>Water</td>
<td>43.4</td>
<td>--</td>
<td>40.7</td>
<td>44.9</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond lengths (Å)</td>
<td>0.002</td>
<td>0.001</td>
<td>0.008</td>
<td>0.020</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>0.551</td>
<td>0.435</td>
<td>0.98</td>
<td>1.654</td>
</tr>
<tr>
<td>Reference model§</td>
<td>4GYZ</td>
<td>4GYZ</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Ramachandran</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Outliers (%)</td>
<td>5</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Favoured (%)</td>
<td>92.4</td>
<td>92.6</td>
<td>97.9</td>
<td>97.9</td>
</tr>
<tr>
<td>MolProbity score</td>
<td>1.8</td>
<td>1.9</td>
<td>1.28</td>
<td>1.50</td>
</tr>
</tbody>
</table>

*Values in parentheses are for highest-resolution shell

†Data isotropic to this resolution

§PDB entry used as ‘reference model’ during refinement
Figure 1

A

hTDP2_CAT

CATALYTIC DOMAIN

UBA-L

1

26 63

113

C273S

362

B

RMSD(#atoms)

human

mouse

zebrafish

worm

0.493 (1406)

0.520 (1320)

0.723 (1120)

C

M7-motif

C (359)

N (112)

D

M7-motif

C (359)

N (112)
**A**

![Chemical structures](image)

- **163** EC\(_{50}\): 40 nM
- **147** EC\(_{50}\): 50 nM (3-OH)
- **148** EC\(_{50}\): 90 nM (4-OH)
- **149** EC\(_{50}\): 250 nM

**B**

![Graph showing normalised fluorescence vs temperature](image)

- apo T\(_{m}\): 48.7 °C
- +163 T\(_{m}\): 51.5 °C

**C**

![Graph showing velocity (RFU min\(^{-1}\)) vs substrate concentration](image)

- \(K_m = 208 \pm 22.8 \text{ nM}\)
- \(V_{max} = 4763 \pm 183 \text{ RFU min}^{-1}\)
- \(k_{cat} = 1.58 \pm 0.06 \text{ s}^{-1}\)

**D**

![Graph showing % inhibition vs log [compound] nM](image)

- IC\(_{50}\)
  - +163: 19 ± 3.2 nM
  - +148: 103 ± 4.5 nM
  - +149: 479 ± 1.9 nM

**E**

![Extended hydrophobic shelf of the protein](image)

- Ile307
- Leu305
- Trp297
- Glu152
- Asp122
- Phe315
- 9.1 Å

**F**

![Bar graph showing % product (5'-P) vs time (minutes)](image)

- hTDP2\(^\text{CAT}\)
- mTDP2\(^\text{CAT}\)
- 5'-Y
- 5'-P

- 2 5 10 2 5 10 2 5 10

- hTDP2\(^\text{CAT}\) + 163
- mTDP2\(^\text{CAT}\) + 163
- hTDP2\(^\text{CAT}\) mTDP2\(^\text{CAT}\) + 163
<table>
<thead>
<tr>
<th>Species</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>TYDP2_HUMAN</td>
<td>TRWDSNLGITAACKLRFDR</td>
</tr>
<tr>
<td>TYDP2_BOVIN</td>
<td>TRWDSNLGIAANCCLRDFR</td>
</tr>
<tr>
<td>TYDP2_CHICK</td>
<td>TRWDSNLGIESKCLRDFR</td>
</tr>
<tr>
<td>TYDP2_RAT</td>
<td>TRWDSNLGIAANCCLRDFR</td>
</tr>
<tr>
<td>TYDP2_MOUSE</td>
<td>TRWDSNLGIAANCCLRDFR</td>
</tr>
<tr>
<td>TYDP2_DANRE</td>
<td>TRWDSNLGIAANCCLRDFR</td>
</tr>
<tr>
<td>TYDP2_XENLA</td>
<td>TRWDSNLGIAANCCLRDFR</td>
</tr>
<tr>
<td>TYDP2_XENTR</td>
<td>TRWDSNLGIAANCCLRDFR</td>
</tr>
<tr>
<td>TYDP2_CAEBR</td>
<td>TRWDSNLGIAANCCLRDFR</td>
</tr>
<tr>
<td>TYDP2_CAEEL</td>
<td>TRWDSNLGIAANCCLRDFR</td>
</tr>
</tbody>
</table>

**A**

M7-motif

**B**

Human

Mouse

Zebrafish

**C**

Divergent region

Arg266

Arg276

Ala309

Ala319

Tyr321

His323

Cys311

Leu313

**D**

% Product (5'-P)

0 20 40 60 80

Time (minutes)

2 5 10 2 5 10

m2hTDP2CAT

hTDP2CAT + 163
Hornyak et al. Figure 5

A

Arg241 | 163

'Divergent region' | 'TWDT'

RFDR

'Hydrophobic shelf'

Divergent region | 'TWDT'

mouse + dsDNA

zebrafish + ssDNA

B

TOP2

OH

163

Pendant group

Deazaflavin core

mouse + dsDNA

zebrafish + ssDNA

TOP2

OH

163

Nuc1

5'-nucleotide

Nuc2

Nuc3

Nuc2

Nuc3

5'-nucleotide

5'-nucleotide
Figure S1. Gyrasol reagent-based enzyme assay

(A) Chemical drawing of the Fluorescein isothiocyanate (FITC) - 5’ phosphotyrosyl linkage.

(B) Progression curve for hTDP2<sup>CAT</sup>. Increasing concentrations of enzyme (as indicated) were incubated with a fixed concentration of substrate, and the change in fluorescence monitored over time. Data points are the mean of 9 measurements, with error bars representing 1 standard deviation.
Figure S2. Representative omit electron density maps

(A) hTDP2\textsuperscript{CAT} + 163. 2mF\textsubscript{o}-DF\textsubscript{c} map contoured at 1.2\(\sigma\) (grey mesh), mF\textsubscript{o}-DF\textsubscript{c} contoured at 3\(\sigma\) (green mesh).

(B) m2hTDP2\textsuperscript{CAT} + 163. 2mF\textsubscript{o}-DF\textsubscript{c} map contoured at 1.2\(\sigma\) (grey mesh), mF\textsubscript{o}-DF\textsubscript{c} contoured at 3\(\sigma\) (green mesh).

(C) m2hTDP2\textsuperscript{CAT} + 148. 2mF\textsubscript{o}-DF\textsubscript{c} map contoured at 1.2\(\sigma\) (grey mesh), mF\textsubscript{o}-DF\textsubscript{c} contoured at 3.5\(\sigma\) (green mesh).

In each case, in order to generate the maps, the ligand has been excluded from refinement.

Density for ligand in each 2mF\textsubscript{o}-DF\textsubscript{c} map has been suppressed, in order to improve the figure clarity.

Amino acid side chains are show in stick representation, with carbon atoms coloured in cyan.

Compounds are also shown in stick representation, with carbon atoms coloured in green.
Figure S3. The binding mode of 163 is identical in hTDP2$^{\text{CAT}}$ and m2hTDP2$^{\text{CAT}}$.
A superposition of the two structures is shown in molecular cartoon representation. See associated key for colour scheme.