Bioorganic & Medicinal Chemistry 18 (2010) 2347–2355

Contents lists available at ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Erratum

Virtual screening using ligand-based pharmacophores for inhibitors of human tyrosyl-DNA phospodiesterase (hTdp1)

Iwona E. Weidlich [a,†], Thomas S. Dexheimer [b], Christophe Marchand [b], Smitha Antony [b], Yves Pommier [b], Marc C. Nicklaus [a,*]

[a] Chemical Biology Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, DHHS, Frederick, MD 21702, USA
[b] Laboratory of Molecular Pharmacology, Center for Cancer Research, National Cancer Institute, National Institutes of Health, DHHS, Bethesda, MD 20892, USA

ARTICLE INFO

Keywords:
Human tyrosyl-DNA phosphodiesterase
Tdp1
Inhibitors
Small molecule docking
Drug design
Virtual screening

ABSTRACT

Human tyrosyl-DNA phosphodiesterase (hTdp1) inhibitors have become a major area of drug research and structure-based design since they have been shown to work synergistically with camptothecin (CPT) and selectively in cancer cells. The pharmacophore features of 14 hTdp1 inhibitors were used as a filter to screen the ChemNavigator *iResearch* Library of about 27 million purchasable samples. Docking of the inhibitors and hits obtained from virtual screening was performed into a structural model of hTdp1 based on a high resolution X-ray crystal structure of human Tdp1 in complex with vanadate, DNA and a human topoisomerase I (Top1)-derived peptide (PDB code: 1NOP). We present and discuss in some detail 46 compounds matching the three-dimensional arrangement of the pharmacophoric features. The presented novel chemotypes may provide new scaffolds for developing inhibitors of Tdp1.

1. Introduction

Human Tdp1 as a drug target for the treatment of cancer has been the subject of much speculation[1–4] but the exploration of all potential benefits of human Tdp1 inhibitors for the treatment of cancer is far from complete. Previous work has suggested that inhibitors of Tdp1 could act synergistically with camptothecin in a combined anticancer therapeutic regimen and also proposed that therapeutic selectivity may be achieved by combining inhibitors of both topoisomerase I and Tdp1 since a significant number of tumors have defective DNA repair and checkpoint pathways.[4,5]

Human tyrosyl-DNA phosphodiesterase is a member of the phospholipase D (PLD) superfamily. hTdp1 repairs DNA topoisomerase I (Top1) covalent complexes by catalyzing the hydrolysis of the phosphodiester bond between a tyrosine and a DNA 3′-phosphate, and appears to be responsible for repairing the unique protein–DNA linkage that occurs in the cell.[6,7] Tdp1 is a monomer composed of two similar domains that are related by a pseudo-twofold axis of symmetry. Each domain contributes conserved histidine, lysine and asparagine residues to form a single active site (Fig. 1).[4,6–10] Mutagenesis studies with the human enzyme in which the invariant histidines and lysines of the HKD motifs are changed, confirm that these highly conserved residues are essential for hTdp1 activity.[8]

The availability of crystal structures usually helps with the understanding of an enzyme's reaction mechanism. Eight crystal structures of Tdp1 with vanadate, oligonucleotides, and/or peptides or peptide analogues have been determined so far.[6,7,9] The use of vanadate as a central moiety in high-order complexes has the potential to be a general method for capturing protein–substrate interactions for phosphoryl transfer enzymes.[9] Although vanadate and tungstate are not attractive pharmacologically, both have been useful in co-crystallization studies of Tdp1.[10] It was established that the active site of the quaternary structure exhibits trigonal bipyramidal geometry for the vanadate ion, which allowed the proposal of a transition state of an S_N2 nucleophilic attack on the phosphate and thus the development of a reaction mechanism for human Tdp1.[4]

In the crystal complex, the DNA oligonucleotide is 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 occupies a portion of a large bowl-shaped binding cleft that is located on the side of the active site opposite the DNA-binding groove (Fig. 1).[6,7,9]

In this paper we report on a series of novel structures identified on the basis of 14 Tdp1 inhibitors serving as a training set for the in silico approaches applied. We analyze the modes of interaction between these molecules and hTdp1 using molecular docking tools.

DOI of original article: 10.1016/j.bmc.2010.02.008
* Corresponding author. Tel.: +1 301 846 5971; fax: +1 301 846 6033.
E-mail addresses: iweidlic@helix.nih.gov, mn1@helix.nih.gov (M.C. Nicklaus).
† On extended leave from the Poznań University of Medical Sciences, Faculty of Pharmacy, Poland.

0968-0896/$ - see front matter © 2010 Published by Elsevier Ltd.
doi:10.1016/j.bmc.2010.02.009

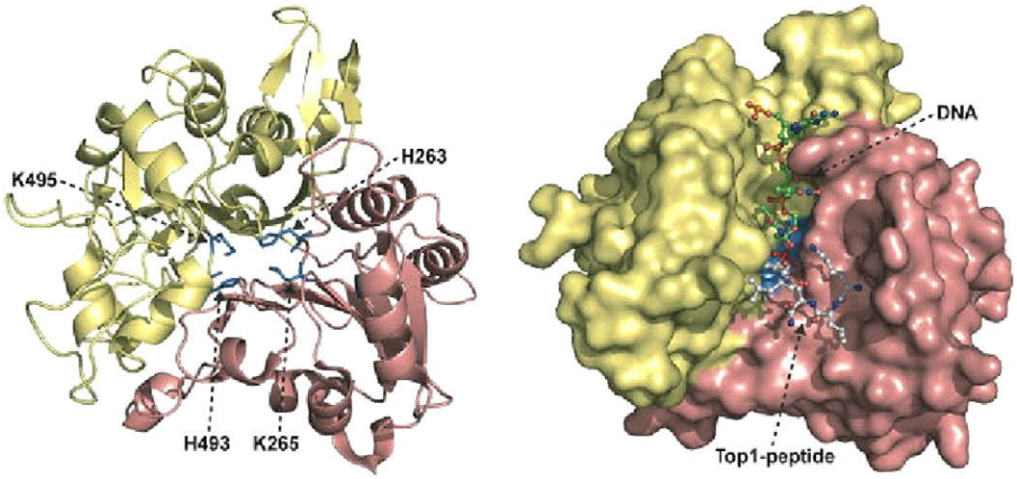

Figure 1. The N-terminal and C-terminal domains consisting of residues 1–350 (brown) and 351–608 (yellow), respectively. Left: Ribbon diagram of Tdp1 (Δ148-truncated human Tdp1). Arrows indicate active site residues H263, K265, H493, and K495, which are shown as stick structures in blue. Right: Structure of the Tdp1–vanadate–peptide–DNA complex (1NOP). Tdp1 is shown as a molecular surface and the vanadate–peptide–DNA substrate mimic is shown as ball-and-stick structures with the DNA colored in green, the vanadate in red, and the Tdp1-derived peptide colored in white (figure from Ref. 4).

2. Results and discussion

2.1. Structural data used

All crystal structures of Tdp1 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. It has been noted[10] that canonical duplex DNA is unlikely to bind to Tdp1 in the absence of conformational change in the nucleotides at the 3′ end. The backbone of a strand of DNA forming Watson–Crick base pairs with nucleotides −1 to −3 would likely clash with loop 229–232 of Tdp1.[10] Oligonucleotides longer than four residues extending beyond the end of the DNA bonding groove have the potential to clash with symmetry-related Tdp1 molecules.[6] It was shown that in order for Tdp1 to remove the topoisomerase I peptide with Tyrosine from the end of duplex DNA, a conformational change in Tdp1 or the DNA must occur.[8] A crystal structure of a reconstituted human topoisomerase I covalently bound to double-stranded DNA shows that the scissile phosphotyrosine bond is buried deep within this complex and would be inaccessible to Tdp1.[10]

These findings and considerations both prompted and allowed us to exclude from our model a fragment of the human Top1-DNA complex consisting of a single-stranded 12-mer oligonucleotide linked to the active site tyrosine[11] via the 3′ end of the DNA.[5] This helps avoid a steric clash of this fragment with hTdp1 and determine how known inhibitors may be binding at the active site of hTdp1.[6] Our intention was to study contacts of active inhibitors with Tdp1 because any compound that binds to Tdp1 in such a way should be able to participate in the formation of such a complex.

We used the Tdp1 crystal structure with the PDB entry code 1NOP and a training set of 14 compounds[12] active in a Tdp1 gel based assay that had been previously developed (see Section 4). Four of these 14 ligands are shown in Figure 2. The physico-chemical properties and IC_{50} values of compounds **1**–**14** are shown in Table 1. The remaining structures will be published in a later publication.

2.2. Ligand-based pharmacophores design

The pharmacophores based on the inhibitors **1**–**14** were designed by using the HipHop method as implemented in the program CATALYST 4.11 (Accelrys).[14] The obtained pharmacophore

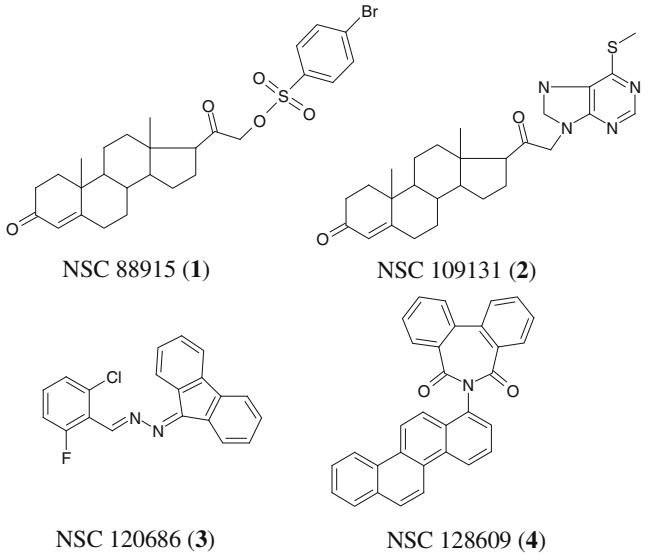

Figure 2. Four of the 14 hTdp1 inhibitors (IC_{50} ~1.8–35 μM) used in the training set for model development. Molecular structures are available from the Enhanced NCI Database Browser: http://cactus.nci.nih.gov/ncidb2/.[13]

Table 1
Physico-chemical properties and IC_{50} values of training set of Tdp1 inhibitors (**1**–**14**)

Compd	MW	NRB	NAR	NHA	NHD	IC_{50} (μM)
1	549.5	5	1	5	0	8.43
2	478.6	4	2	6	0	35
3	334.7	2	3	2	0	14
4	445.5	1	6	1	0	5
5	638.8	12	3	5	3	6
6	387.5	7	1	3	0	6
7	381.4	2	3	3	0	10
8	394.4	3	3	4	0	8
9	401.8	2	3	3	0	12
10	411.5	3	3	4	0	8.0
11	564.6	5	5	6	0	1.8
12	505.6	4	5	4	0	8.0
13	554.0	4	5	4	0	20
14	415.9	2	3	3	0	1.8

NRB—number of rotatable bonds; NAR—number of aromatic rings; NHA—number of hydrogen bond acceptors; NHD–number of hydrogen bond donors.

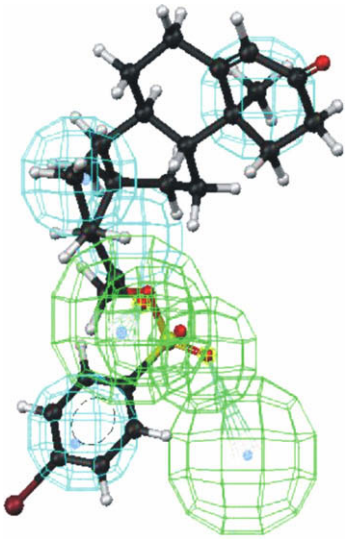

Figure 3. Pharmacophore with six chemical features derived from compound **1**. Green: hydrogen bond acceptors; cyan: hydrophobic centers.

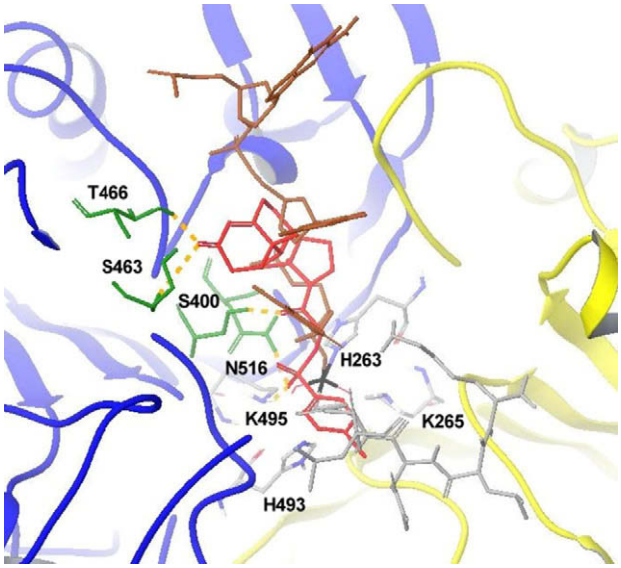

Figure 5. Cartoon representations of the structure of Human Tdp1 with inhibitor (**1**) (red) docked in it. The N-terminal domain of Tdp1 (residues 162–350) is colored yellow, the C–terminal domain (residues 351–608) blue. The active site residues H263, K265, H493, and K495 are depicted as ball and stick structures (located close to the vanadate group). The hydrogen bonds between the best docking pose of inhibitor **1** and Tdp1 (residues N516, S400, K495, S463, and T466, shown in green) are represented by yellow dotted lines. Only hydrogen bonds with a length $\leqslant$3 Å are shown. For clarity only polar hydrogen atoms are shown. The vanadate group is colored in black, the DNA and 3′-phospotyrosyl linkage of the Tdp1 substrate in brown and grey, respectively.

hypotheses for describing the training set compounds were derived using chemical features, then clustered and filtered down to six pharmacophore models, each of which consisted of two H-bond acceptors and four hydrophobic centers. These depict the features the ligand should possess to achieve binding to the important active site residues of the enzyme. Figure 3 shows one of the pharmacophore models derived from inhibitor **1**.

2.3. Database virtual screening

The six pharmacophores were then used as filters to screen the ChemNavigator *iResearch* Library[15] of purchasable screening samples, containing more than 27 million unique structure records. The binding mode of each hit is defined by the interaction pattern of its pharmacophore.

The 248,341 hits obtained in the pharmacophore searches were processed with PIPELINE PILOT 5.0[16] (SciTegic) by applying organic Lipinski 'rule of five' and HTS filters,[17] and removing duplicate molecules (Fig. 4).

Successful drug discovery requires high quality lead structures which may need to be more drug-like than is commonly accepted.[18] Toxicity and poor pharmacokinetics should be eliminated in the early stages of drug discovery. Physico-chemical and ADME/Tox properties were therefore calculated for the filtered set of 102,712 hits using the programs QIKPROP 3.0[19] (Schrödinger) and ADMET PREDICTOR 2.3.0[20] (Simulations Plus, Inc.). Of the 102,712 hits obtained from the PIPELINE PILOT protocol, 88,246 compounds passed the ADME/Tox filter.

2.4. Docking studies

In the crystal structure, vanadate is covalently linked to H263, which means vanadate could be replaced by a variety of ligands as has been suggested in previous work.[7,9,10,21] In order to gain insight into possible binding modes of the known inhibitors **1**–**14**, these compounds were docked into the ligand-binding domain of the Tdp1 active site (H263, K265, H493 and K495) using the program GLIDE.[21] The model preparation procedure is described in the protein preparation section (see Section 4).

The initial docking of the training set of 14 active compounds was intended to help us choose significant interactions for further analysis of the ChemNavigator *iResearch* Library[15] filtered set of 88,246 hit ligands aligned to the designed pharmacophores.

Compound **1** is one of the drug-like active compounds among the 14 training set inhibitors in this study.

We noticed that the vanadate participating in the 3′-phospotyrosyl linkage of the Tdp1 substrate is mimicked by the sulfonyl ester moiety of compound **1**, while the phenyl bromide and steroid moieties of the inhibitor are substituted for the tyrosyl portions of the Tdp1 substrate and DNA, respectively (see Fig. 5).

All 88,246 selected hits were docked into the Tdp1 active site in the same way as the training set of 14 active compounds before. It is well known that the scores calculated by docking programs do not usually permit the exact reproduction of the binding mode of assayed compounds. However, the analysis of all reproduced poses for each of the hits would be time-consuming. We chose the best

Figure 4. PIPELINE PILOT protocol to filter the inhibitors obtained in CATALYST 4.11[14] pharmacophore searches.

scoring pose for each of the 88,246 docked compounds. We then used the properties calculated for the training set of 14 inhibitors, as well as the protocol for GLIDE[22] scoring/ranking of the poses, and the structural characteristics of the binding site to make rational decisions on the ranges of values that are reasonable for new potential inhibitors of Tdp1. The obtained docked poses did not necessarily have to strictly conform with the pharmacophore models we used to screen the ChemNavigator Library.[15] We picked the best output poses based on the analysis of interactions with the binding site which we deemed important for good binding, but we also mapped each high-scored pose with the pharmacophores using the program MOE.[23]

We finally selected 178 compounds based on the best docking score, energy of the model (Emodel), and requirement that a good

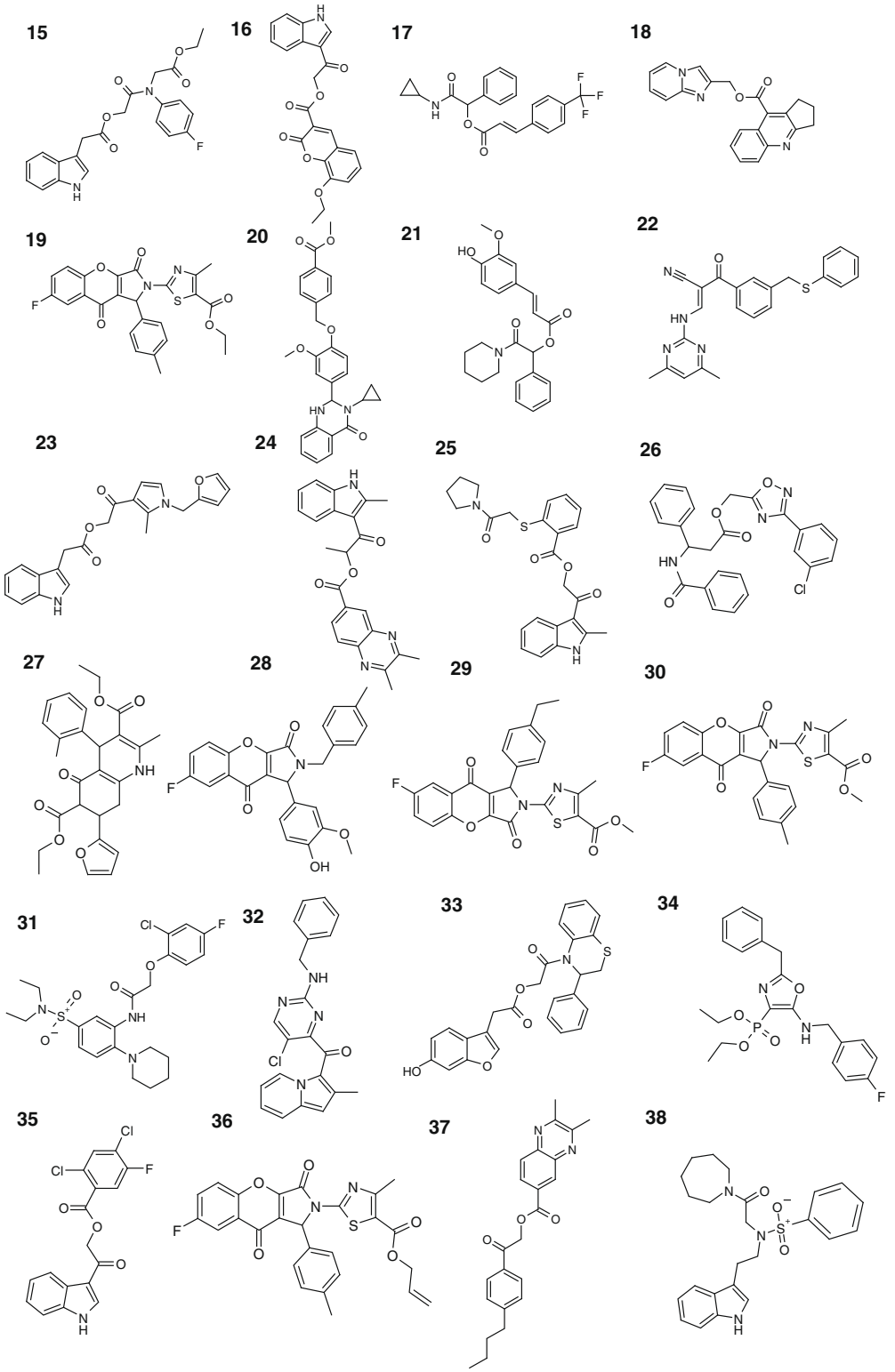

Figure 6. Structures of the compounds **15–60** (see Table 2) identified by virtual screening.

inhibitor should be docked into the active site. Specifically, a good inhibitor should form at least one hydrogen bond with the residues N516, S400, T261 or K495, H263 and have the hydrophobic aromatic ring oriented toward the Tdp1 active site. A hydrophobic interaction with the Tdp1 local hydrophobic pocket seems to be important for the inhibition, too (Fig. 7). We compared the formed hydrogen bonds and hydrophobic interactions with the Tdp1 active site with the proposed binding mode of inhibitors **1**–**14** based on the mentioned crystal structure (Fig. 5).

Based on this evaluation of the docking poses and the previously calculated ADME/Tox properties, 46 compounds were chosen for closer analysis.

The structures of these compounds are shown in Figure 6, their docking scores and physico-chemical properties in Table 2.

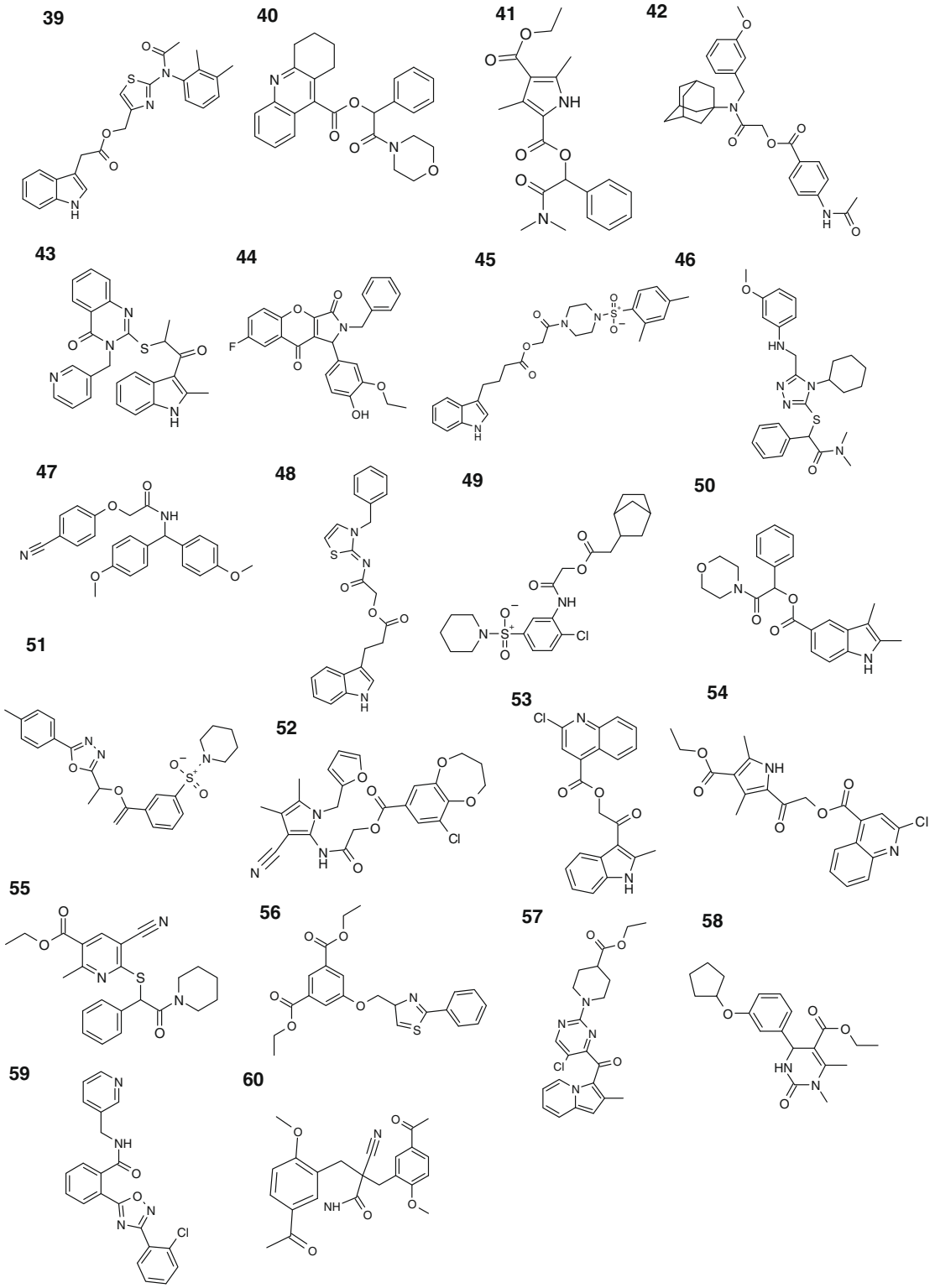

Figure. 6 (*continued*)

2.5. In silico docking versus known in vitro results

The analysis of the optimized complexes shows that the interactions with the residues N516, S400, T261 and K495 are common to the binding complex of, for example, compounds **15**, **16**, and **17**. The hydrophobic aromatic ring present in these compounds is oriented toward the Tdp1 active site, where it forms hydrophobic interactions with residues H263, Y204 and C205 or H493, S400 and P461 (Figs. 7 and 8).

We speculate that this conformation may be important for binding of these hits. One also notices that a dione chemotype is present in a number of these structures. Compounds **15** and **16** are of the 1*H*-indole-3-yl dione chemotype, which might be a chemical feature beneficial for forming a protein–ligand complex. The carbonyl groups of ligands **1**, **15**, **16** and **17** form hydrogen bonds with residues T261 and N516, which also appears to be a feature important for binding. The difference in the inhibitory activity may therefore follow the differences in binding modes.

The ligand-binding region of Tdp1 contains a deep local hydrophobic pocket which is partially occupied by a steroid moiety for compound **1** and hydrophobic aromatic rings for compounds **15**–**18** (Fig. 8). Since compounds **1** and **15**–**17** exhibit the same binding mode, the Tdp1 local hydrophobic pocket may be proposed as a location for future inhibitor modifications.

The hydrogen bonds with residues N516 and T261 and hydrophobic interactions with the local hydrophobic pocket (Fig. 8—left) and residues H263 Y204 and C205 or H493, S400 and P461 (Fig. 7—right) occur for ligands **15**–**17** and seems to be important for binding.

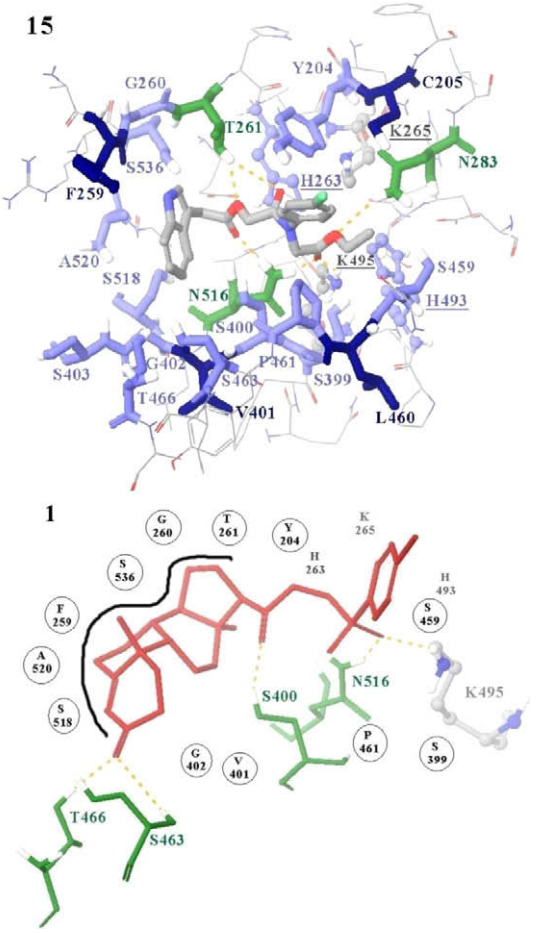

Figure 7. Predicted binding modes of compounds **15** and **1** (red colored molecule), in the hTdp1 active site (H263, K265, H493, K495). Compound **15**—the ligands are drawn in tube representation, where the coloring of atoms for compound **15** is: carbon—grey; nitrogen—blue; oxygen—red; fluorine—green. Human Tdp1 active site and hTdp1 residues are displayed in ball and stick and tube representations, respectively. Amino acid residues common to the binding pocket of both molecules are labeled. For clarity, only polar hydrogens of the Tdp1 enzyme are shown. Significant enzyme/ligand polar interactions and the hydrophobic pocket of the hTdp1 active site are shown as yellow dashed lines and blue colored residues (dark blue: very hydrophobic; light blue: less hydrophobic), respectively. Compound **1**—two-dimensional interaction map of **1** docked into the Tdp1 active site. Inhibitor **1** is drawn in tube representation, where all atoms are colored in red. Amino acid residues capable of polar interactions are labeled in green. Amino acid residues significant for hydrophobic interactions are marked in the circle. The black line indicates the Tdp1 local hydrophobic pocket. Amino acid residues at 4.5 Å of the ligand are shown. Figure generated with the program MAESTRO (MACROMODEL 9.6, Schrodinger Inc.).[24]

2.6. Outlook

The importance of the phenyl sulfonyl ester moiety has become even more evident through very recent SAR studies described in related work,[25] in which we have tried to establish

Table 2
Docking scores and physico-chemical properties for compounds **15**–**60**

Compd	Structure ID[a]	Gscore (kJ/mol)	MW	NRB	NAR	NHA	NHD
15	52,835,583	−9.13	412.4	10	3	5	1
16	52,927,501	−8.98	391.3	7	4	6	1
17	59,619,269	−8.43	389.3	8	2	3	1
18	52,842,479	−8.4	343.3	4	4	4	0
19	40,766,804	−9.06	478.5	5	4	6	0
20	51,227,591	−8.15	458.5	8	3	6	1
21	59,437,235	−8.39	395.4	7	2	5	1
22	60,130,600	−8.19	400.5	7	3	6	1
23	52,835,439	−9.94	390.4	8	4	3	1
24	59,760,832	−9.87	387.4	5	4	5	1
25	37,839,769	−9.39	436.5	8	3	5	1
26	59,580,155	−9.36	461.9	9	4	5	1
27	60,885,592	−9.12	463.5	8	2	6	1
28	47,615,294	−9.1	445.4	4	4	5	1
29	61,213,802	−9.07	478.5	5	4	6	0
30	40,766,803	−9.02	464.4	4	4	6	0
31	59,439,419	−8.9	498.0	9	2	5	1
32	52,858,487	−8.85	376.8	5	4	4	1
33	59,599,445	−8.8	459.5	6	4	5	1
34	61,163,199	−8.79	418.4	6	3	5	1
35	59,748,938	−8.75	366.1	5	3	3	1
36	40,766,805	−8.71	490.5	6	4	6	0
37	59,760,842	−8.67	376.4	8	3	5	0
38	52,979,973	−8.62	439.5	7	3	3	1
39	52,835,344	−8.61	433.5	7	4	4	1
40	59,421,950	−8.61	430.5	5	3	5	0
41	59,501,436	−8.6	372.4	10	2	2	1
42	52,832,237	−8.59	490.6	9	2	5	1
43	59,733,246	−8.58	454.5	6	5	5	1
44	40,766,814	−8.56	445.4	5	4	5	1
45	52,833,421	−8.55	497.6	9	3	5	1
46	59,731,164	−8.54	479.6	9	3	6	1
47	59,427,503	−8.54	402.4	8	3	5	1
48	59,496,773	−8.53	419.5	8	4	6	1
49	28,287,192	−8.51	469.0	5	1	2	1
50	59,681,693	−8.45	392.4	5	3	4	1
51	59,433,631	−8.4	455.5	7	3	6	0
52	59,755,537	−8.31	483.91	7	3	6	1
53	51,229,339	−8.3	378.8	5	4	4	1
54	51,229,223	−8.29	414.8	8	3	6	1
55	59,697,207	−8.27	423.5	7	2	6	0
56	57,531,210	−8.24	411.4	10	3	6	0
57	53,176,932	−8.22	426.9	6	3	6	0
58	61,491,537	−8.14	358.4	6	2	4	1
59	61,525,041	−8.12	390.8	5	4	4	1
60	52,917,020	−8.09	422.4	9	2	6	1

[a] ChemNavigator structure ID;[15] NRB—number of rotatable bonds; NAR—number of aromatic rings; NHA—number of hydrogen bond acceptors; NHD—number of hydrogen bond donors.

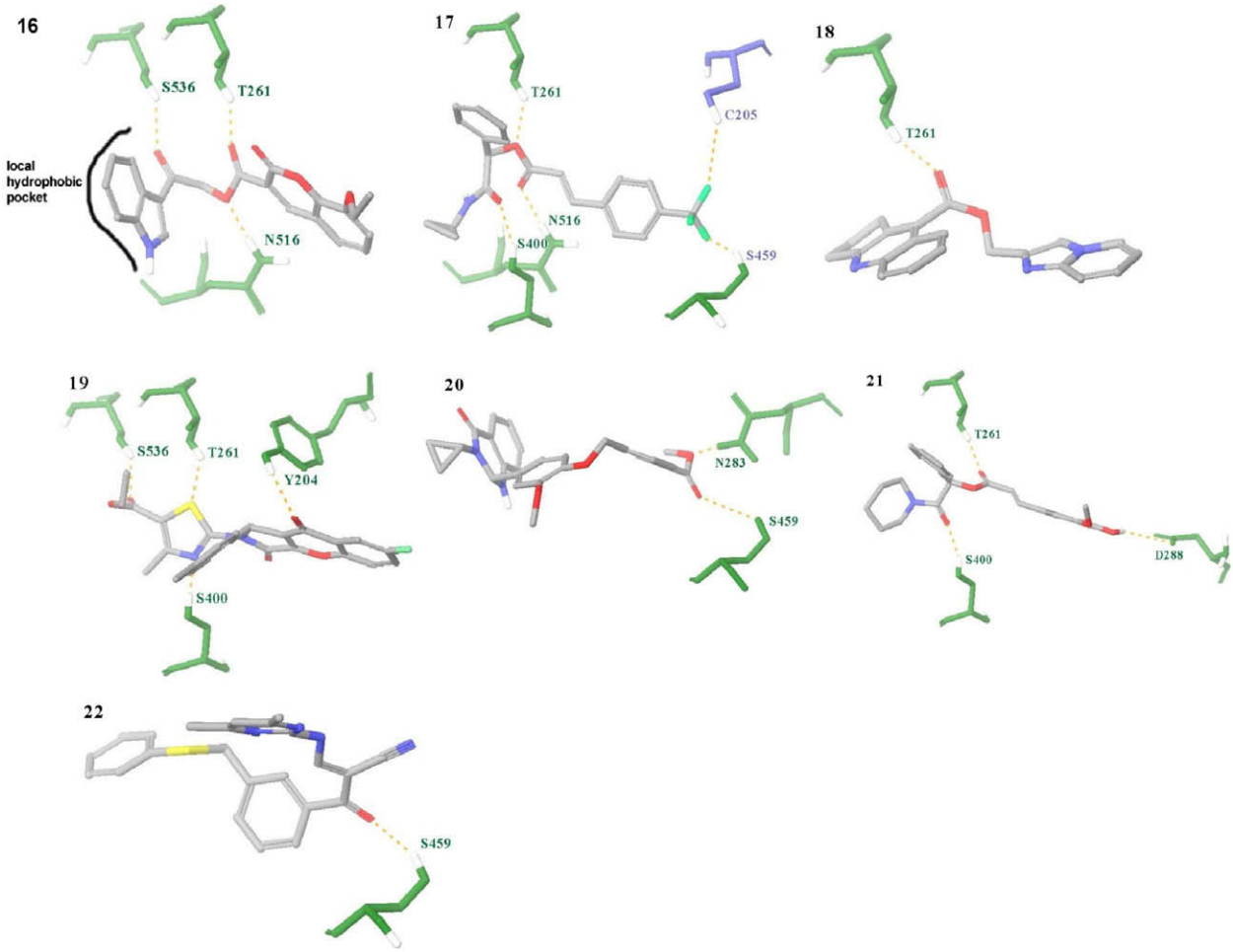

Figure 8. Predicted binding modes of ligands **16**–**22** (tube representation). The orientation and residue colors are identical to those of Figure 7. The black line indicates the Tdp1 local hydrophobic pocket. For clarity only polar hydrogens of amino acid residues are shown.

the relative contribution of specific substructures toward the overall Tdp1 inhibitory activity. Toward this goal, we deconstructed **1** into two fragments that were separately available for assaying. Progesterone and methyl *p*-toluene sulfonate were used in place of the steroid and phenyl sulfonyl ester moieties, respectively. Progesterone was found to be weakly active, whereas methyl *p*-toluene sulfonate was inactive at concentrations of up to 1 mM. Testosterone, also assayed, exhibited activity similar to that of progesterone. One therefore concludes that both a steroid and a phenyl sulfonyl ester moiety need to be present in molecule **1**—and, by extension, in close analogs thereof—for activity against Tdp1.

3. Conclusion

Despite the fact that a crystal structure of Tdp1 complexed with a small-molecule inhibitor has not yet been solved, the in silico docking work reported in this paper allowed us to present a number of possible ligands. These results allow us to put forward suggestions for the future design of Tdp1 inhibitors such as that phenyl sulfonyl ester moiety and (1*H*-indol-3-yl) acetic acid derivatives may lead to active chemotypes of Tdp1 inhibitors.

As soon as a crystal structure of Tdp1–ligand complex will be published, we will use this to check and fine-tune our docking procedure, which in turn will hopefully help improve future lead optimization.

Likewise, an increased number of assay results obtained for much larger screening decks will improve our understanding of structural features beneficial for Tdp1 inhibitory activity. To move toward this goal, a qHTS[26] discovery process using more than 300,000 screening samples has been started for this target at NIH to generate additional small molecule lead compounds and better understanding of this system. This effort has recently been organized as a trans-institute collaboration between the Division of Cancer Treatment and Diagnosis (DCTD) and the Center for Cancer Research (CCR) at the National Cancer Institute (NCI) with the NIH Chemical Genomics Center (NCGC) and other NIH departments in the context of the newly formed Chemical Biology Consortium at the NCI.[27]

4. Materials and methods

4.1. Molecular modeling

4.1.1. Pharmacophore generation

Diverse conformations for the training set compounds were generated using the BEST method as implemented in the program CATALYST 4.11 (Accelrys).[14] We used the recommended maximum number of conformers (255) and the default energy threshold 20 kcal/mol for generating conformers used as input for developing automated hypotheses. We generated the predictive common features hypotheses from a training set of 14 ligands using the

HipHop method, which finds chemical features shared by a set of compounds. The conformations from the training set of molecules were mapped to obtained hypotheses using the best fit method. The collection of obtained feature-based hypotheses was ranked and clustered using the hierarchical average linkage method, from which finally six pharmacophore models were selected.

4.1.2. Generation of the ChemNavigator filtered set

Twenty-seven million unique structure records from the ChemNavigator *iResearch* Library[15] were used to create the 3D multiconformation database in CATALYST. The six pharmacophore models were used as a query to search this database using a shell script. The resulting filtered set contained 248,341 hits, which were subsequently processed with a PIPELINE PILOT 5.0[16] (SciTegic) protocol by applying the organic Lipinski 'rule of five' and HTS filters,[17] and removing duplicate molecules. This filtering stage let pass 102,712 hits, which were in turn subjected to an ADME/Tox filter (again in PIPELINE PILOT 5.0). A total of 88,246 hits passed this filter and were prepared for further docking.

4.1.3. Preparation of protein structure

The crystal structure of hTdp1 (PDB entry code 1NOP)[10] was downloaded from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB-PDB), and read in and processed, including addition of hydrogen atoms, with the program MAESTRO.[24,28]

After the hydrogen atoms were assigned, all water molecules and the cocrystallized DNA–peptide substrate were deleted; so was a vanadate ion O_4V^{3-}, since the OPLS-2005 force field used by GLIDE has no parameters for vanadate.[22]

We know that the function of the residues K265 and K495 in human Tdp1 is substrate binding and transition state stabilization.[8] We also know from the crystal structure that two oxygen atoms of the vanadate are within hydrogen bonding distance of the amino groups of K265 and K495. We observed that most neutral polar and apolar compounds could not fit in the lysine cationic cavity, which mostly binds anionic ligands.[29] For this reason, residues K265 and K495 were neutralized before docking.

Only the N-terminal domain of Tdp1 (residues 162–350) and the C-terminal domain (residues 351–608) from the crystal structure of 1NOP were used as building blocks in the construction of the molecular model for the docking procedure. After these modifications, a series of minimizations using the OPLS-2005 force field were carried out.

4.1.4. Preparation of ligand structures

Samples of compounds NSC 88915, NSC 109131, NSC 120686 and NSC 128609, which formed part of the training set, were obtained from the 1981 diversity set of the Developmental Therapeutics Program (DTP) of the NCI, NIH. Their structures are also publicly available from the Enhanced NCI Database Browser: http://cactus.nci.nih.gov/ncidb2/. Information about the other ten compounds will be presented in a future publication.[12]

Additional molecular construction and modeling of the training set of 14 ligands were performed using the building tools available in MACROMODEL 9.6 (Schrödinger Inc.).[24] The ligands were minimized using the OPLS-2005 force field.

In order to get information of the ranges of values for physicochemical properties of ligands that are observed with molecules that are known to bind to Tdp1 active site, we calculated these properties using different programs: MOE,[23] QIKPROP (Schrödinger),[22] and ADMET PREDICTOR 2.3.0 (Simulations Plus, Inc.).[20]

The physico-chemical properties NRB (number of rotatable bonds); NAR (number of aromatic rings); NHA (number of hydrogen bond acceptors); NHD (number of hydrogen bond donors) were calculated with the program MOE.[23] For filtering and some property calculations, we used the program PIPELINE PILOT 5.0 from SciTegic.[16]

The preparation procedure in GLIDE required the preparation of the structures in the appropriate ionization state. We used LigPrep[30] for the filtered set of 102,712 hit ligands. This 2D to 3D conversion program generates accurate energy minimized molecular structures, expands tautomeric and ionization states, ring conformations, and stereoisomers to produce broad chemical and structural diversity from a single input structure. LigPrep generated ligand libraries (at a rate of approximately one ligand per second) with the desired structural and chemical features for further computational analyses in QIKPROP (Schrödinger)[19] and GLIDE (Schrödinger).[22]

After filtering the structures through the protocol created in PIPELINE PILOT 5.0 (SciTegic)[15] (Fig. 4), we calculated pharmaceutically relevant properties utilizing the program QIKPROP (Schrödinger),[19] specifically using the command 'nosa'.

To quickly eliminate non-promising hits we applied the Schrödinger utility program PROPFILTER. The filtered set of 102,712 hits was treated by Schrödinger filters to eliminate compounds that did not meet the following criteria: molecular weight $\leqslant$500; octanol/water partition coefficient log $P_{o/w} \leqslant 5$; aqueous solubility log $S > -6$; polar surface area FISA $\leqslant$175; 5 or fewer hydrogen bond donors; and 10 or fewer hydrogen bond acceptors.

The program ADMET PREDICTOR 2.3.0 (Simulations Plus, Inc.)[20] was used for predictive modeling of ADME/Tox properties. We applied ADME/Tox properties (BBB quantitative assessment of blood–brain–barrier permeability; S_w [mg/ml] native water solubility; Tox MRTD [mg/kg/day] qualitative assessment of the maximum recommended therapeutic dose administered as an oral dose; Tox ER filter qualitative assessment of estrogen receptor toxicity in rats; Tox BRM Mouse [mg/kg/day] TD50, which is defined as the oral dose of a compound required to induce tumors in 50 percent of a mouse population after exposure over a standard lifetime; P_{eff} [cm/s $\times 10^4$] human jejunal effective permeability; MDCK [cm/s $\times 10^7$] apparent MDCK (Madin–DarbyCanine Kidney) cells-on-sheet (COS) permeability)[20] via a PIPELINE PILOT protocol, which resulted in 88,246 compounds passing the specified ADME/Tox criteria (BBB:low; S_w <0.1; Tox MRTD $\leqslant$1.5; Tox ER filter-nontoxic; Tox BRM Mouse >10; P_{eff} >0.2; MDCK >0).

4.1.5. Docking

The 14 **1**–**14** described inhibitors first, and subsequently the 88,246 hits, were docked into our model of the hTdp1 active site using the program GLIDE (Schrödinger)[22] with the Extra Precision mode. Grid files were generated with residues H263, K265, H493 and K495 at the center of the binding box defining the space through which the center of the docked ligand is allowed to move. The size of the cube box was set to 16 Å edge length in order to explore large region of the protein.

To conduct more precise analysis of docked poses of the 88,246 hits in addition to the scoring/ranking method, we mapped the output poses to the pharmacophores using absolute positioning in MOE.[23] In this way the coordinate frame of reference query was the same as the frame of reference of the docked poses.

4.2. Preparation of human Tdp1 and substrates 1–14 for gel assays

4.2.1. Reagents

HPLC-purified oligonucleotides and tyrosyl nucleotides were purchased from the Midland Certified Reagent Co. (Midland, TX).

4.2.2. Preparation of human Tdp1

Human Tdp1 expressing plasmid pHN1910 (a gift from Dr. Howard Nash, Laboratory of Molecular Biology, National Institute

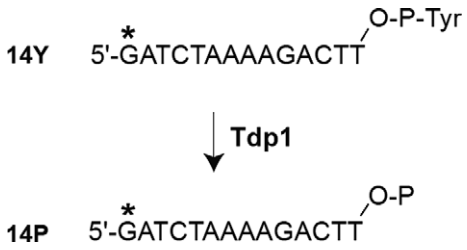

Figure 9. Substrates and products for the hTdp1 assay.

of Mental Health, NIH) was constructed using vector pET-15b (Novagen, Madison, WI) with full-length human Tdp1 and an additional His-tag sequence of MGSSHHHHHHSSGLVPRGSHMLEDP in its N terminus. The His-tagged human Tdp1 was purified from Novagen BL21 cells using a chelating sepharose™ fast flow column (Amersham Biosciences, Piscataway, NJ) according to the company's protocol. The collected fractions were assayed immediately for Tdp1 activity. Fractions that showed Tdp1 activity were pooled and dialyzed in 20% glycerol, 50 mM Tris–HCl, pH 8.0, 100 mM NaCl, 10 mM β-mercaptoethanol, and 2 mM EDTA. Dialyzed samples were divided into aliquots and stored at −80 °C. Tdp1 concentration was determined using the Bradford protein assay (Bio-Rad Laboratories, Hercules, CA). Tdp1 purity was determined as a single ~70 kDa band representing over 95% of the detectable proteins stained by Coomassie after SDS–polyacrylamide gel electrophoresis (SDS–PAGE).[5,26]

4.2.3. Preparation of Tdp1 substrates for gel assays

HPLC-purified oligonucleotides 14Y (see Fig. 9) was labeled at the 5′-end with [γ-^{32}P] ATP (Perkin–Elmer Life and Analytical Sciences, Boston, MA) by incubation with 3′-phosphatase-free T4 polynucleotide kinase (Roche Diagnostics, Indianapolis, IN) according to the manufacturer's protocols. Unincorporated nucleotides were removed by Sephadex G-25 spin-column chromatography (mini Quick Spin Oligo columns; Roche Diagnostics). (10 mM Tris–HCl, pH 7.5, 100 mM NaCl, and 10 mM $MgCl_2$), heated to 96 °C, and allowed to cool down slowly (over 2 h) to room temperature.[5,26]

4.2.4. Tdp1 gel assays

Unless indicated otherwise, Tdp1 assays were performed in 20 μl mixtures in Tdp1 assay buffer (50 mM Tris–HCl, pH 7.5, 80 mM KCl, 2 mM EDTA, 1 mM dithiothreitol and 40 μg/ml bovine serum albumin). Twenty-five nanomolar 5′-^{32}P-labeled substrate (14Y, Fig. 9) was reacted with 1 ng of Tdp1 (0.7 nM) in the absence or presence of inhibitor for 30 min at 25 °C. Reactions were stopped by the addition of 60 μl of gel loading buffer [96% (v/v) formamide, 10 mM EDTA, 1% (w/v) xylene cyanol, and 1% (w/v) bromphenol blue]. Twelve-microliter aliquots were resolved in 20% denaturing polyacrylamide (AccuGel; National Diagnostics, Atlanta, GA) (19:1) gel containing 7 M urea. After drying, gels were exposed overnight to PhosphorImager screens (GE Healthcare Bio-Sciences Corp., Piscataway, NJ). Screens were scanned, and images were obtained with the Molecular Dynamics software. Densitometry analyses were performed using the ImageQuant 5.2 software (GE Healthcare Bio-Sciences Corp., Piscataway, NJ). Tdp1 activity was determined by measuring the fraction of substrate converted into 3′-phosphate DNA product by densitometry analysis of the gel image.[5,26] The representative results were consistently reproduced at least three times.

Acknowledgment

This publication is the result of research supported, in whole or in part, by direct costs funded by NIH.

References and notes

1. Pouliot, J. J.; Yao, K. C. *Science* **1999**, *286*, 552–555.
2. Pouliot, J. J.; Robertson, C. A. *Genes Cells* **2001**, *6*, 677–687.
3. Vance, J. R.; Wilson, T. E. *PNAS* **2002**, *99*, 13669–13674.
4. Dexheimer, T. S.; Antony, S.; Marchand, C.; Pommier, Y. *Anti-Cancer Agents Med. Chem.* **2008**, *8*, 381–389.
5. Antony, S.; Marchand, C.; Stephen, A. G.; Thibaut, L.; Agama, K. K.; Fisher, R. J.; Pommier, Y. *Nucleic Acids Res.* **2007**, *35*, 4474–4484.
6. Davies, D. R.; Interthal, H.; Champoux, J. J.; Hol, W. G. J. *J. Med. Chem.* **2004**, *47*, 829.
7. Davies, D. R.; Interthal, H.; Champoux, J. J.; Hol, W. G. J. *Structure* **2002**, *10*, 237–248.
8. Interthal, H.; Pouliot, J. J.; Champoux, J. J. *PNAS* **2001**, *98*, 12009–12014.
9. Davies, D. R.; Champoux, J. J. *J. Mol. Biol.* **2002**, *324*, 917–932.
10. Davies, D. R.; Interthal, H.; Champoux *Chem. Biol.* **2003**, *10*, 139–147.
11. Coupez, B.; Möbitz, H. *Chem. Modell.: Appl. Theor.* **2006**, *4*, 1–22.
12. Update on the publication status of these compounds can be obtained from Christophe Marchand at marchanc@mail.nih.gov.
13. The Enhanced NCI Database Browser, made available on the web server http://cactus.nci.nih.gov of the CADD Group, Lab. of Medicinal Chemistry, NCI, is an interface to the Open NCI Database, which is a collection of compounds submitted to DTP (Developmental Therapeutics Program), NCI, from a wide variety of academic and industrial labs around the world for screening against cancer and AIDS.
14. CATALYST, Release 4.11, Accelrys Software, San Diego, 2005.
15. The ChemNavigator *iResearch* Library is a compilation of commercially available chemical compounds, in particular screening samples. http://www.chemnavigator.com/cnc/products/iRL.asp.
16. PIPELINE PILOT, version 5.0, SciTegic, San Diego, CA, 2005.
17. This component filters out molecules that are likely to be poor candidates for high-throughput screening, including those containing non-organic atom types, reactive substructures, and those with a molecular weight >150.
18. Proudfoot, J. R. *BMCL* **2002**, *12*, 1647–1650.
19. QIKPROP, version 3.0, Schrödinger, LLC, New York, NY, 2007.
20. ADMET PREDICTOR, Version 2.3.0, Simulations Plus, Lancaster, CA, 2007.
21. Liao, Z.; Tribaut, L.; Pommier, Y. *Mol. Pharmacol.* **2006**, *70*, 366–372.
22. GLIDE, version 5.0, Schrödinger, LLC, New York, NY, 2008.
23. Chemical Computing Group's Molecular Operating Environment (MOE), version 1008.10, 2008.
24. MACROMODEL 9.6, Schrodinger, LLC, New York, NY, 2008.
25. Dexheimer, T. S.; Gediya, L. K.; Stephen, A. G.; Weidlich, I.; Antony, S.; Marchand, C.; Interhal, H.; Nicklaus, M. C.; Fisher, R.; Njar, V. C.; Pommier, Y. *J. Med. Chem.* **2009**, *52*, 7122–7131.
26. Marchand, C.; Lea, W. A.; Jadhav, A.; Dexheimer, T. S.; Austin, C. P.; Inglese, J.; Pommier, Y.; Simeonov, A. *Mol. Cancer Ther.* **2009**, *8*, 240–248.
27. http://dtp.nci.nih.gov/docs/CBC/cbc_index.html.
28. Rishton, G. M. *Drug Discovery Today* **2003**, *8*, 86–96.
29. Graves, A. P. *J. Mol. Biol.* **2008**, *377*, 914–934.
30. LIGPREP, Schrödinger, LLC, New York, NY, 2007.