Published in final edited form as:
J Med Chem. 2008 August 14; 51(15): 4553–4562. doi:10.1021/jm8001668.

Identification and Validation of Human DNA Ligase Inhibitors Using Computer-Aided Drug Design

Shijun Zhong[†], **Xi Chen**[‡], **Xiao Zhu**[†], **Barbara Dziegielewska**[‡], **Kurtis E. Bachman**[§,#], **Tom Ellenberger**[‖], **Jeff D. Ballin**[⊥], **Gerald M. Wilson**[⊥], **Alan E. Tomkinson**[*,‡,§], and **Alexander D. MacKerell Jr.**[*,†]

Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland 21201, Radiation Oncology Research Laboratory, Department of Radiation Oncology, School of Medicine, University of Maryland, Baltimore, Maryland 21201, University of Maryland Marlene and Stewart Greenebaum Cancer Center, School of Medicine, University of Maryland, Baltimore, Maryland 21201, Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110, and Department of Biochemistry and Molecular Biology, School of Medicine, University of Maryland, Baltimore, Maryland 21201

[†]School of Pharmacy, University of Maryland. [‡]Department of Radiation Oncology, School of Medicine, University of Maryland. [§]University of Maryland Marlene and Stewart Greenebaum Cancer Center, School of Medicine, University of Maryland. [‖]Washington University School of Medicine. [⊥]Department of Biochemistry and Molecular Biology, School of Medicine, University of Maryland.

Abstract

Linking together of DNA strands by DNA ligases is essential for DNA replication and repair. Since many therapies used to treat cancer act by causing DNA damage, there is growing interest in the development of DNA repair inhibitors. Accordingly, virtual database screening and experimental evaluation were applied to identify inhibitors of human DNA ligase I (hLigI). When a DNA binding site within the DNA binding domain (DBD) of hLigI was targeted, more than 1 million compounds were screened from which 192 were chosen for experimental evaluation. In DNA joining assays, 10 compounds specifically inhibited hLigI, 5 of which also inhibited the proliferation of cultured human cell lines. Analysis of the 10 active compounds revealed the utility of including multiple protein conformations and chemical clustering in the virtual screening procedure. The identified ligase inhibitors are structurally diverse and have druglike physical and molecular characteristics making them ideal for further drug development studies.

Introduction

Target-based virtual database screening has become a useful tool for the identification of inhibitors for protein—ligand and protein—protein interactions.[1–4] One million or more compounds may be screened to identify those with a high probability of binding to a site on a target macromolecule. The selected compounds are then subjected to experimental assay; hit

*To whom correspondence should be addressed. For A.E.T.: phone, 410-706-2365; fax, 410-706-6666; atomkinson@som.umaryland.edu. For A.D.M.: phone, 410-706-7442; fax, 410-706-5017; alex@outerbanks.umaryland.edu..
[#]Present address: Discovery & Translational Medicine, GlaxoSmithKline, 709 Swedeland Road, King of Prussia, PA 19406.

rates of 5% or more are often reported.[5] In the present work, virtual database screening[5] in combination with experimental assays has been utilized to identify low molecular weight inhibitors of human DNA ligase I (hLigI[a]).[6]

DNA ligases catalyze the joining of interruptions in the phosphodiester backbone of double-stranded DNA, making them essential enzymes for DNA repair and replication. In addition, they are an indispensable reagent in molecular biology research for generating recombinant DNA. DNA ligases are members of the larger nucleotidyl transferase family that also includes RNA ligases and mRNA capping enzymes. In the first step of the ligation reaction, DNA ligases react with a nucleotide cofactor, either NAD^+ or ATP, to form the covalent enzyme—AMP intermediate. Next the AMP moiety is transferred to the 5′-phosphate termini in duplex DNA, forming the DNA adenylate intermediate. Finally, the nonadenylated enzyme catalyzes phosphodiester bond formation between the 3′-hydroxyl and 5′-phosphate termini.

There are three human *LIG* genes, *LIG1*, *LIG3*, and *LIG4* that encode ATP-dependent DNA ligases.[7] The *LIG1* gene product, hLigI, joins Okazaki fragments during lagging strand DNA replication and also participates in DNA excision repair.[8] Several distinct DNA ligase polypeptides that function in nuclear DNA repair, mitochondrial DNA metabolism, and germ cell development are encoded by the *LIG3* gene.[7] The *LIG4* gene product, hLigIV, completes the repair of DNA double strand breaks by nonhomologous end joining and V(D)J recombination events that generate diversity in immunoglobulin and T-cell receptor loci during immune system development.[7]

Because of their involvement in DNA replication and DNA repair, DNA ligase inhibitors are likely to be antiproliferative and to potentiate the cytotoxicity of DNA damaging agents, properties that may have clinical utility in the treatment of cancer, in particular malignancies with an altered DNA damage response. Attempts to identify human DNA ligase inhibitors by screening of chemical and natural product libraries have met with limited success.[9,10] The recent determination[6] of an atomic resolution structure of hLigI bound to nicked DNA by X-ray crystallography allowed us to utilize a rational, structure-based approach to identify DNA ligase inhibitors.

In the complex formed by hLigI on DNA with a nonligatable nick, three hLigI domains encircle and interact with the nicked DNA duplex.[6] Two of these domains, an adenylation domain (AdD) and an OB-fold domain (OBD), are present in other DNA ligases and nucleotidyl transferases. In contrast, the DNA binding domain (DBD, residues Asp262 to Ser535) is restricted to eukaryotic ATP-dependent DNA ligases.[7] Notably, the DBD is the predominant DNA binding activity within hLigI and stimulates joining in trans by a hLigI fragment containing the adenylation and OB-fold domains.[6] On the basis of these properties, we chose to focus on identifying compounds that bind to the DBD and inhibit hLig1 activity by interfering with its interaction with nicked DNA.

Methods

CADD Screening

The in silico identification of compounds with a high probability of binding to and inhibiting DNA ligase involved the following steps, i.e., identification of a putative ligand binding site on the interface between the DBD and bound DNA (Figure 1), molecular dynamics (MD) simulations for the generation of multiple protein conformations to address the flexibility of the binding site in the screening process (Table 1), preliminary screening of over a million compounds, secondary docking of 50 000 compounds from the preliminary screen against the crystal structure and the MD generated structures, and final selection of compounds for experimental assay.

Protein Structure Preparation

The crystal structure of hLig1, obtained from the Protein Data Bank[11] (PDB, http://www.rcsb.org/) (PDB identifier 1×9n),[6] was truncated, keeping only the DBD. Hydrogen atoms were then added followed by local energy minimization with the program CHARMM.[12] The minimization involved 100 conjugate gradient (CONJ) steps with the positions of atoms identified in the crystallographic structure fixed at their experimental values. Calculations were performed using the CHARMM all-atom protein force field including the CMAP modification[13,14] with default cutoffs for the nonbond interactions. The resulting DBD structure was used in the preliminary docking (see below).

To obtain multiple conformations of the protein for secondary docking, an MD simulation was performed for 5 ns on the DBD using stochastic boundary conditions.[15] First, the structure was minimized for 200 steepest descent[16] (SD) steps in vacuum. The binding region was then solvated by overlaying the protein with a 35 Å water sphere centered on the geometric center of the three residues His337, Arg449, and Gly453 defining the binding pocket (see below). Water molecules within 2.8 Å to any protein non-hydrogen atom were removed. A test equilibration MD simulation showed a tendency for the water to move toward the protein and away from the surface of the sphere associated with the deletion of water overlapping the protein. Therefore, the water ball was assigned a boundary potential of 30 Å for all remaining calculations. All atoms were divided into three radial shells, i.e., the central region, an edge region from 26 to 30 Å, and an outer region beyond 30 Å, which comprised only protein atoms. Atoms in the outer region were constrained to their energy-minimized positions, atoms in the edge region were harmonically restrained with a force constant of 5 (kcal/mol)/Å, and the central region was not subject to any type of restraints. The density of the water sphere was maintained using a quartic potential via the miscellaneous mean field potential (MMFP) module[17,18] in CHARMM. Parameters defining the potential were force 0.25, droff 28.5, and p1 2.25, which yields a local well of −0.31 kcal/mol at the edge of the sphere. Nonbond interaction lists were heuristically updated out to 14 Å with the electrostatic and Lennard Jones (LJ) interactions truncated at the range of 10–12 Å using force switching.[19] Following a 500-step SD minimization, the protein was subjected to a 5 ns MD simulation at 300 K using the velocity verlet (VVER)[20] integrator, an integration time step of 2 fs, and SHAKE to constrain all covalent bonds involving hydrogen atoms.[21] Coordinates were saved every 5 ps, yielding a total of 1000 conformations from which additional structures were selected for the secondary docking. Selection of conformations for docking was performed via clustering based on pairwise root-mean squared differences of the position of residues defining the binding site, i.e., the residues Glu300—Arg305 on the loop between helixes 3 and 4 according to the helix order in 1×9n.pdb, Ser334-His337 at the end of helix 5, Pro341-Asp351 on the loop following the short helix 6, and residues Gly448—Glu456 on the loop between helixes 12 and 13. Clustering was performed with NMRCLUST[22] with representative structures from the four biggest clusters chosen and used in the secondary docking.

Identification of Putative Binding Site

A putative DNA binding site within the DBD was identified using the sphere sets used in the program DOCK[23] in combination with residues implicated in DNA binding by X-ray crystallography. In the present study, we focused on three residues, His337, Arg449, and Gly453, that are located in the central region of the DBD and make direct contacts with the DNA substrate, as shown in Figure 1. Other residues comprising the binding site include Gly448, Arg451, and Ala455. Generation of the sphere set used the minimized crystallographic structure, with the hydrogen atoms deleted, to compute the Connolly solvent accessible surface[24,25] via the subroutine DMS which is implemented in the program MIDAS[26] (recently updated to Chimera[27]). The solvent accessible surface was computed via DMS using the surface points of the probe sphere as required when hydrogen atoms are not present, with the

density of points set to 0.5 as suggested for proteins. Second, spheres with radii ranging from 1.2 to 4 Å, complementary to the protein surface, were generated by the subroutine SPHGEN implemented in the package DOCK.[23] Each sphere contacts two protein surface points and lies on the normal of one of the two points. This procedure generates a very large number of spheres, which are filtered by selecting only the largest sphere associated with each surface atom. Next, spheres within 8 Å of all three reference residues, His337, Arg449, and Gly453, were selected and selected spheres on the periphery of the putative binding pocket manually deleted. This yielded the final sphere set shown in Figure 1, which was used to direct the in silico database screening. This protocol was used to generate sphere sets for the multiple conformations of the protein generated from the MD simulation.

In Silico Compound Databases

An in-house database of more than 4.3 million low molecular weight compounds has been developed at the University of Maryland Computer-Aided Drug Design Center. This database comprises three types of files, i.e., 2D SD format files originally from the commercial vendors, 3D MOL2 format files for docking, and binary MDB format files for use in the program MOE (Chemical Computing Group Inc., Montreal, Canada). Compound preparation included removal of the smaller components in entries containing salts (e.g., counterions), addition of hydrogen atoms, assignment of the protonation state, geometry optimization using the MMFF94[28,29] force field level with either SYBYL (Tripos Associates, St. Louis, MO) or MOE (Chemical Computing Group, Canada), and assignment of atomic partial charges based on CM2 charge model computed at the semiempirical quantum chemical AM1 level using AMSOL.[30,31] For the present study preliminary screening used approximately 1 500 000 compounds from vendors chosen on the basis of their reliability with respect to availability of compounds. Vendors include Chembridge (371 000), Chemdiv (750 000), Maybridge (60 000), MDD (33 000), Nanosyn (47 000), Specs (232 000), Timtec (165 000) and Tripos (80 000), where the values in parentheses represent the approximate number of compounds associated with each company. Recently, the compounds in the collections from these companies have been shown to typically have druglike characteristics.[32]

Docking and Final Compound Selection

Docking computations were performed using DOCK4.0[33] with parameters previously used in our laboratory.[34–36] Kollman partial atomic charges for the DBD were assigned using the program SYBYL. Database searching in DOCK is performed via a fragment-based build up procedure.[37] In this approach one or more anchor fragments (e.g., rigid units, such as rings with five or more atoms) are overlaid on the spheres in 200 orientations. The remainder of the ligand is then built layer by layer, with a rotation about each added bond in 10° increments to identify the most favorable orientation based on the total ligand—protein interaction energy. Thus, the docking procedure accounts for ligand flexibility while the protein is treated as rigid. From the preliminary docking using only a single conformation of the protein, the top 50 000 compounds were selected on the basis of the normalized van der Waals attractive (Va) energy, as described below. These compounds are then subjected to a second round of docking where the crystallographic conformation and four additional conformations of the protein from the MD simulation (Table 1) were targeted to account for protein flexibility. The ligands were separately docked into each protein conformation, with the most favorable score from all five conformations assigned to rank that ligand. The score used in the second docking run is the total interaction energy including electrostatic and van der Waals interactions. In addition, the ligand was subjected to additional optimization by increasing the maximum anchor fragment orientations from 200 to 500, performing minimization of the anchor at each cycle of ligand buildup and minimizing the five inner layers upon addition of each layer of the ligand.

Compound clustering was performed using the Tanimoto similarity index[38,39] based on BIT_MACCS fingerprints[40] which is implemented in the MOE software package. The BIT_MACCS fingerprints are used to compute the pairwise Tanimoto similarity matrix **S** which contains the similarity metric between the molecular fingerprints of compounds *i* and *j*. The matrix element $S(i,j)$, i.e., the Tanimoto coefficient (Tc) is defined as $Tc = c(i,j)/u(i,j)$, where $c(i,j)$ is the number of common features in the fingerprints of molecules *i* and *j* and where $u(i,j)$ is the number of all features in the union of the fingerprints of molecules *i* and *j*.[41] Two compounds are regarded as similar if $S(i,j)$ is equal to or greater than a predefined similarity threshold. Then, from matrix **S**, another binary matrix **O** is created where each matrix element $O(i,j)$ has the value 1 if $S(i,j)$ is equal to or greater than the predefined similarity threshold, or 0 otherwise. Two molecules *i* and *j* are then grouped into a cluster if the overlap between the two row vectors $\mathbf{O}(i)$ and $\mathbf{O}(j)$ is greater than or equal to a predefined overlap threshold. In the present work a similarity threshold of 70% and an overlap threshold of 40% were used.

Compounds and Proteins for Experiments

Compounds identified by CADD screening were purchased from Chembridge, Chemdiv, Maybridge, MDD, Nanosyn, Specs, Timtec, and Tripos. The 10 mM stocks were prepared in DMSO and stored at −20 °C. The molecular mass of the active compounds was confirmed by mass spectrometry at the University of Maryland School of Pharmacy facility. Recombinant human DNA ligase I was purified as described previously.[42,43] T4 DNA ligase was purchased from NEB.

DNA Joining Assays

Candidate ligase inhibitors identified by CADD were assayed for their ability to inhibit hLigI and T4 DNA ligase using a high throughput, fluorescence energy transfer-based DNA joining assay.[42] Duplicate mixtures (30 μL) containing 10 pmol of nicked DNA substrate and either 0.25 pmol of hLigI or 10 units of T4 DNA ligase were incubated in the presence or absence of 100 μM of the putative inhibitor. DNA binding by the candidate DNA ligase inhibitors was measured by displacement of ethidium bromide from DNA as previously described.[44]

A radioactive-gel-based DNA ligation assay was performed as previously described.[45] A 25-mer (5′-CGC CAG GGT TTT CCC AGT CAC GAC C-3′) and a 5′ [^{32}P] end-labeled 18-mer (5′-GTA AAA CGA CGG CCA GTG-3′) were annealed to a complementary 44-mer oligonucleotide, generating a linear duplex with a central nick. DNA joining mixtures (30 μL) containing 0.5 pmol of labeled DNA substrate, and hLigI (0.02 pmol) in ligation buffer were incubated in the absence or presence of ligase inhibitors at 25 °C for 30 min.

Electrophoretic Mobility Shift Assay

A labeled linear duplex with a nonligatable nick was incubated with hLig1 in ligation buffer (30 μL total volume) with or without ligase inhibitors for 120 min at 25 °C. After the addition of an equal volume of native gel buffer (160 mM Tris-HCl, pH 6.8, 20% glycerol, 1.4 M 2-mercaptoethanol, 0.05% bromophenol blue), samples were separated by electrophoresis through a 12% native polyacrylamide gel and detected in the dried gel by phosphor-Imager analysis.

Kinetic Analysis of Ligase Inhibitors

To measure the initial rates of ligation, hLigI (0.05 pmol) was incubated with 0.5–100 pmol of the fluorescent, nicked DNA substrate, and various concentrations of the ligase inhibitors. K_i values were obtained from Lineweaver—Burk double-reciprocal plots and curve fitting using PRISM, version 3.03 (GraphPad).

Cell Proliferation Assays

The ability of the compounds identified by CADD to inhibit proliferation of normal mammary epithelial MCF10A cells and colon carcinoma HCT116 cells was determined using a Biomek FX laboratory automation workstation (Beckman Coulter, Inc., Fullerton, CA). On day 0, 20 μL of complete medium containing the appropriate number of cells (150–300) was plated per well of a 384-well tissue plate (Fisher Scientific, Hampton, NH) and incubated overnight at 37 °C with 5% CO_2 and 90% humidity. Next day, day 1, compounds were prepared by serial dilution with complete medium to yield the concentration 100 μM, and 20 μL was added to each well containing 20 μL of medium and cells yielding the final concentration 50 μM in 40 μL volume. Plates were incubated for an additional 3 days (days 2–5) until control cells (0.5% DMSO) reached ~70–80% confluency. On day 6, 40 μL of lysis/detection solution containing 1.2% Igepal CA-630 (Sigma) and a 1:1000 dilution of SYBR Green I nucleic acid stain (Molecular Probes, Eugene, OR) were added to each well. Following an overnight incubation at 37 °C, total fluorescence was measured using a Fluorostar Galaxy plate reader with a 485 nm excitation filter and 520 nm emission filter set (BMG Labtech, Inc., Durham, NC). Data were exported to a custom program that determined growth inhibition by dividing each individual fluorescence value by the average of fluorescence values obtained with cells treated with DMSO alone. Compounds that showed at least 40% growth inhibition compared with the DMSO-only controls inhibition of one or both of the cell lines were scored as “hits”.

The activity of hits from the initial screen was further validated using the MTT assay. Briefly, MCF10A and HCT116 cells were seeded in 96-well plates at 300 and 1200 cells per well, respectively, and allowed to adhere overnight. The following day, serial dilutions of compounds in media were added to the cells in a final volume 200 μL. After incubation for 5 days, MTT reagent (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltatrazolium) was added and incubation was continued for 4 h. Formazan crystals generated by reduction of the MTT reagent in mitochondria were solubilized with isopropanol prior to the measurement of absorbance at 560 nm wavelength in a plate reader.

Results

In Silico Database Screening

A putative DNA binding pocket within the DBD of hLigI was chosen as the target for a multitiered in silico database screening procedure, based on regions of the DBD in direct contact with the DNA in the X-ray structure of hLigI complexed with nicked DNA (Figure 1). In the first step of the screen to identify compounds with a high probability of binding to the DBD of hLig1, ligand posing of 1.5 million compounds was based on the total interaction energy between the ligands and the protein, with ligand ranking performed using the normalized van der Waals attractive (Va) energy. Use of the Va energy parameter selects for compounds with significant steric overlap with the binding pocket and avoids compounds with highly favorable electrostatic interactions that do not fit well into the pocket. In addition, normalization procedures correct for the tendency of compound selection based on interaction energies to bias toward high molecular weight (MW) compounds.[46] Distributions of MW using different normalization procedures and the distributions of normalized scores are shown in parts a and b of Figure 2, respectively. On the basis of $N^{2/3}$ normalization, a total of 50 000 compounds with a molecular weight distribution centered around 300 Da were selected for further analysis.

Secondary screening of the 50 000 compounds applied additional energy minimization during docking and partially addressed protein flexibility[47,48] via the inclusion of four additional, structurally diverse conformations obtained from an MD simulation. Overall, the five conformations of the DBD are similar (Figure 3A), indicating that significant structural

changes in the protein did not occur during the MD simulation. However, a detailed comparison of the orientation of the residues lining the binding region shows that there is significant diversity across the five conformations. In Figure 3B, the orientations of three residues His337, Arg449, and Gly453, located in the central site of the binding region, are shown. Table 1 gives the root-mean-squared deviations (rmsd) of the residues in the binding region, including residues Glu300—Arg305, Ser334—His337, Pro341—Asp351, and Gly448—Glu456. Differences in the rmsd values between the crystal structure and the four conformations from MD simulation range from 1.4 to 2.5 Å, indicating significant conformational variation in the binding pocket.

Compounds were ranked based on the most favorable normalized total interaction energy of each ligand against the five protein conformations. At this stage the total interaction energy includes electrostatic interactions as well as steric considerations in the selection process. The MW and energy distributions for different powers of N normalization of the 1000 compounds with the most favorable normalized total interaction energies are shown in parts a and b of Figure 4, respectively. These compounds were selected on the basis of $E/(N^{2/5})$ normalization, yielding a MW distribution consistent with compounds known to be pharmacologically active or used in lead compound optimization studies.[49,50]

Final selection of compounds for in vitro biochemical assays involved maximizing chemical diversity and was based on their druglike or leadlike compound properties.[49,50] Diversity was maximized by clustering the compounds based on chemical fingerprints using the Tanimoto similarity index. This yielded approximately 200 clusters of chemically similar compounds, with one or two compounds from each cluster selected on the basis of druglike or leadlike compound properties as defined by Lipinski's rule of five.[51] These rules include molecular weight (MW < 500), adequate solubility expressed by the octanol/water partition coefficient ($-5 < \log P(\text{o/w}) < 5$), number of hydrogen bond acceptors (Ha < 10), number of hydrogen bond donors (Hd < 5), number of rotatable single bonds (Rot < 10), and number of rings (Ring < 5). The final compounds typically also satisfy the slightly stricter rules of Oprea.[49] However, for clusters in which the criteria were not met, compounds were still selected for experimental assay. From this process, 233 compounds were selected for experimental testing. Distributions of the physical and molecular properties of the 233 compounds are presented in Figure 5, showing them to indeed fulfill Lipinski's rule of five.

Experimental Assays

Of the 233 compounds identified by in silico screening, 192 were available from commercial vendors. These compounds were assayed for inhibitory activity in DNA joining assays with purified enzymes. In parallel, they were assayed for their ability to inhibit proliferation of cultured human cells.

By use of a fluorescence-based high throughput ligation assay, the 192 compounds were assayed for their ability to inhibit hLigI and T4 DNA ligase. Although the bacteriophage enzyme is ATP-dependent like hLigI and has similar adenylation and OB-fold domains, it lacks a DBD domain and so should be insensitive to molecules that bind to the DBD of hLigI.[7] The activity of the 16 compounds that inhibited hLigI by >50% (Table 2) was confirmed in a radioactive gel-based ligation assay.[45] Of the 16 compounds that inhibited hLigI by >50% (Table 2), 6 were also active against T4 DNA ligase by >50%. One mechanism by which a compound may nonspecifically inhibit human DNA ligase is by binding to DNA rather than the ligase, thereby interfering with the enzyme—substrate interaction. In accord with this idea, **192**, which inhibits both hLigI and T4 DNA ligase (Table 2), reduced DNA binding of the DNA intercalating agent ethidium bromide (Figure 6B) whereas four other compounds **25**, **67**, **82** and **189**, which inhibit hLigI but not T4 DNA ligase (Table 2), had no effect on DNA binding by ethidium bromide (Figure 6A), indicating that **25**, **67**, **82**, and **189** do not bind DNA.

The aggregation of small molecules into large aggregates that sequester and inhibit enzymes is another mechanism that generates nonspecific false positives in high throughput screening.[52,53] To eliminate this possibility, we examined the activity of the seven hLigI inhibitors that have not been extensively characterized, **1**, **25**, **64**, **113**, **190**, **197**, and **200** in the presence of 0.1% Triton X-100. For six of these compounds, the presence of the nonionic detergent did not reduce the inhibitory activity of the small molecules by more than 5% (data not shown). In the case of **190**, inhibition was reduced from 57% to 39% by the inclusion of Triton X-100. Thus, the in silico screen yielded at least nine compounds that specifically inhibit hLigI (Table 1), a hit rate of 4–5%, significantly higher than the hit rate of 0.1% reported for an experimental screen of a random chemical library.[9]

In a previous study, we determined IC_{50} values of 8, 12, and 4 μM for the hLigI inhibitors **67**, **82**, and **189**, respectively.[45] Using the same radioactive gel-based ligation assay, we have determined the IC_{50} values for the other seven hLigI inhibitors (Figure 7). The IC_{50} values for the 10 hLigI specific inhibitors identified by the in silico screen range from 0.6 to 25 μM. Since the screen was targeted against a DNA binding pocket within the DBD of hLigI, the resultant inhibitors should interfere with interaction of hLigI with nicked DNA. In accord with this prediction, we have shown that two inhibitors, **67** and **189**, act as simple competitive inhibitors with respect to nicked DNA whereas the other inhibitor **82** acts uncompetitively and enhances formation of a hLigI—nicked DNA complex.[45] Although the uncompetitive inhibitor may bind to any region of hLigI, it is conceivable that it contacts both the hLigI DBD and nicked DNA in a manner analogous to the mechanism by which camptothecin inhibits topoisomerase I.[54,55] An attractive feature of the uncompetitive inhibitor is that it may be possible to directly visualize its binding site by determining the structures of cocrystals of the inhibitor and hLigI—nicked DNA. To determine the mechanisms by which the other seven hLigI inhibitors act, we used the same electrophoretic mobility shift assay with a linear DNA substrate containing a single nonligatable nick. Of the seven inhibitors, four reduced formation of the hLigI—nicked DNA complex by more than 50%, indicating that they competitively inhibit DNA binding (Figure 8A). As expected, the nonspecific inhibitor **192**, which binds to DNA, also blocks formation of the hLigI—DNA complex (Figure 8A). The three inhibitors, **113**, **190** and **200**, that had weak or no detectable inhibitory activity in the electrophoretic mobility shift assay also had the highest IC_{50} values (Figure 7), and so it is possible that competitive inhibition by these compounds would be evident at higher concentrations. To confirm the validity of the electrophoretic mobility shift assay for identifying competitive inhibitors, we measured the kinetics of DNA joining in the absence or presence of compound **25**. The resulting Lineweaver—Burk plots indicate that **25** is a simple competitive inhibitor with respect to the DNA substrate, with a K_i value of 4 ± 2 μM.

All 192 compounds were also assayed for their ability to inhibit proliferation of two human cell lines, MCF10A cells from normal breast epithelial and HCT116 cells from a colon cancer, at 50 μM. Sixteen compounds inhibited proliferation of one or both cell lines by more than 40%. There was considerable overlap with the 10 compounds identified as inhibitors of hLigI. Notably, the five hLigI specific inhibitors **64**, **67**, **82**, **189** and **190** also inhibited cell proliferation.

CADD Analysis of Active Compounds

Structures of the 10 hLigI specific inhibitors are shown in Figure 9. As may be seen, the structures are chemically diverse, as verified by the calculation of pairwise Tanimoto similarity indices between the compounds (Table 3). The largest Tc value between two compounds is 69%, and the majority of the values are less than 50%, indicating a low degree of similarity between compounds. Previous studies have indicated that a value of 85% or more is associated with compounds that will have similar biological activities.[56] The inclusion of chemical

diversity in compound selection has the desirable effect of identify structurally dissimilar compounds for drug optimization, thereby increasing the probability of identifying active compounds. This may be seen by analyzing the energy scores of the selected compounds. Presented in Figure 10 are the distributions of the total interaction energy scores for the 1000 compounds from the secondary screen, for the 233 compounds selected from the top 1000 based on diversity and physical properties, and for the 10 active compounds (Figure 10). Consideration of diversity and physical properties led to the selection of more compounds with less favorable interaction energies. Notably, many of the active compounds would not have been selected if the top 233 scoring compounds were selected on the basis of interaction energies and so would not have been identified.

The importance of the inclusion of multiple conformations of the putative binding site from the MD simulation in the in silico screen may be determined by simply identifying the conformation from which the 10 active compounds were selected. Of the 10 hLigI specific inhibitors, one (**67**) was based on the crystal conformation, none was based on MD conformation C2 (time point 2.015 ns of the simulation), two (**82**, **113**) were based on MD conformation C3 (time point 2.335 ns of the simulation), four (**1**, **25**, **190**, **197**) were based on MD conformation C4 (time point 2.950 ns of the simulation), and three (**64**, **189**, **200**) on conformation C5 (time point 3.795 ns of the simulation). Thus, the inclusion of multiple conformations is leading to the identification of additional active compounds, emphasizing the utility of this component of the screening procedure.

The docked orientations of the above 10 active compounds are shown together in Figure 11. The protein conformation is that from the crystallographic study with the orientation of the compounds extracted from the individual conformations following alignment of the protein conformations as shown in Figure 3A. All the inhibitors occupy the targeted site, consistent with the docking methodology. However, they do sample different regions of the binding site. Such difference may contribute to differential selectivities of activity of the inhibitors for different ligases.

Three of the active compounds, **67**, **82**, and **189**, have been subjected to more extensive biological characterization.[45] Although all three compounds are predicted to bind in the putative binding site, they do exhibit some level of variability in the binding orientations, as shown in Figure 11. Interestingly, while all three compounds inhibit hLigI but not T4 DNA ligase (Table 2), their activity versus the other human DNA ligases differs significantly. While **82** inhibits only hLigI, **67** inhibits both hLigI and hLigIII, and **189** inhibits all three human DNA ligases. Presumably, differences in the specificities of the inhibitors for the three human DNA ligases reflect a combination of differences in the binding modes of the structurally diverse inhibitors and differences in the molecular architecture of the targeted DNA binding pocket between the three human DNA ligases. Importantly, inhibitors with defined specificities for the different human DNA ligases will be invaluable reagents for elucidating the physiological roles of human DNA ligases.

Consistent with the inclusion of physical properties in the selection process, all the hLigI specific inhibitors fall into the druglike range according to Lipinski's rule of five (Table 2)[51] while still spanning a range of physical properties. On the basis of the structural diversity of the active compounds, their druglike but wide range of physical properties, and their activity in biochemical and cell culture assays, this collection of novel inhibitors of hLigI may be of utility as the starting point for the development of cancer therapeutic agents.

Summary

A combined process involving MD simulation, virtual database screening, and experimental validation has been used to identify small molecules that interfere with the interaction between

human DNA ligases and nicked DNA. Out of 192 compounds that were identified as potentially binding to the target site in hLigI by in silico screening, 10 compounds specifically inhibited hLigI in DNA joining assays, a hit rate of approximately 5%. The active compounds were structurally diverse while having druglike physical and molecular properties. Notably, several of the active compounds also inhibited hLigIII and/or hLigIV, making them a valuable set of reagents for research purposes. Furthermore, several of the active compounds inhibited cell proliferation and specifically enhanced the killing of cancer cells by DNA damaging agents used to treat cancer.[45] Together the properties of these novel DNA ligase inhibitors make them promising lead compounds for development into anticancer drugs to be used either alone or in combination with genotoxic chemotherapy agents and radiation therapy.

Acknowledgments

This work was supported in part by the U.S. National Institutes of Heath (Grants GM47251, ES12512, and GM57479 to A.E.T., Grant CA102428 to G.M.W., Grant GM52504 to T.E., a Structural Cell Biology of DNA Repair Program Grant CA92584 to A.E.T. and T.E.). Financial and computational support from the University of Maryland Computer-Aided Drug Design Center is acknowledged.

a Abbreviations:

AdD, adenylation domain
AMP, adenosine monophosphate
AMSOL, a semiempirical quantum chemical program
ATP, adenosine 5′-triphosphate
BIT_MACCS, a code system to express molecular fingerprints
CADD, computer-aided drug design
CHARMM, Chemistry at Harvard Macromolecular Mechanics, a package of computational programs for molecular modeling
Chimera, software for the visualization of molecular structure and some simple molecular modeling tasks
CM2, a charge model used in electronic structure theoretical calculations
CMAP, grid-based correction for the φ, ψ dihedral dependence of the energy used in CHARMM force field
compd, compound
CONJ, conjugate gradient method for the minimization of molecular geometry
CPK, ball-and-stick molecular model proposed by Corey, Pauling, and Koltun
DBD, DNA binding domain
DMS, program to generate the molecular surface
DOCK, software for docking ligands against a putative binding site
Ha, the number of hydrogen bond acceptors
Hd, the number of hydrogen bond donors
hLigI, human DNA ligase I
hLigIV, human DNA ligase IV
LIG, human DNA ligase gene that encodes ATP dependent DNA ligase
LIG1, human DNA ligase gene 1 that encodes ATP dependent DNA ligase I
LIG3, human DNA ligase gene 3 that encodes ATP dependent DNA ligase III
LIG4, human DNA ligase gene 4 that encodes ATP dependent DNA ligase IV
LJ, Lennard-Jones potential
log *P*, the logarithm of the ratio of the solubility of a compound in octanol to its solubility in water
MD, molecular dynamics
MIDAS, software to view molecular structure and perform some simple calculations
MMFF94, 1994 release of the Merck molecular force field
MMFP, miscellaneous mean field potential

MOE, software provided by Chemical Computing Group for handling compound database
MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium
MW, molecular weight
NMRCLUST, software for the clustering of protein conformations
OBD, OB-fold domain
PDB, Protein Data Bank
Ring, the number of rings in a molecular structure
Rot, the number of rotatable bonds in a molecule
PRISM, a software for scientific graphing, curve fitting, and statistics
rmsd, root-mean-squared deviation
SD, steepest descent method for the minimization of molecular geometry
SHAKE, method to constrain covalent bonds during a molecular modeling calculation
SPHGEN, software to generate spheres complementary to molecular surface
SYBYL, software provided by Tripos for the treatment of compound database
Va, van der Waals attractive energy
V(D)J, a gene containing a variable (V) gene, a diverse (D) gene and a functional joining (J) gene, which be used in genetic recombination
VVER, velocity verlet integrator used in molecular dynamics simulation

References

(1). Schneidman-Duhovny D, Nussinov R, Wolfson HJ. Predicting molecular interactions in silico: II. Protein—protein and protein—drug docking. Cur. Med. Chem 2004;11:91–107.

(2). Méndez R, Leplae R, Lensink MF, Wodak SJ. Assessment of CAPRI predictions in rounds 3–5 shows progress in docking procedures. Proteins 2005;60:150–169. [PubMed: 15981261]

(3). Mohan V, Gibbs AC, Cummings MD, Jaeger EP, DesJarlais RL. Docking: successes and challenges. Curr. Pharm. Des 2005;11:323–333. [PubMed: 15723628]

(4). Ghosh S, Nie A, An J, Huang Z. Structure-based virtual screening of chemical libraries for drug discovery. Curr. Opin. Chem. Biol 2006;10:194–202. [PubMed: 16675286]

(5). Zhong S, Macias AT, MacKerell AD Jr. Computational identification of inhibitors of protein—protein interactions. Curr. Top. Med. Chem 2007;7:63–82. [PubMed: 17266596]

(6). Pascal JM, O'Brien PJ, Tomkinson AE, Ellenberger T. Human DNA ligase I completely encircles and partially unwinds nicked DNA. Nature 2004;432:473–478. [PubMed: 15565146]

(7). Tomkinson AE, Vijayakumar S, Pascal JM, Ellenberger T. DNA ligases: structure, reaction mechanism, and function. Chem. Rev 2006;106:687–699. [PubMed: 16464020]

(8). Levin DS, McKenna AE, Motycka TA, Matsumoto Y, Tomkinson AE. Interaction between PCNA and DNA ligase I is critical for joining of Okazaki fragments and long-patch base-excision repair. Curr. Biol 2000;10:919–922. [PubMed: 10959839]

(9). Sun D, Urrabaz R. Development of non-electrophoretic assay method for DNA ligases and its application to screening of chemical inhibitors of DNA ligase I. J. Biochem. Biophys. Methods 2004;59:49–59. [PubMed: 15134906]

(10). Tan GT, Lee S, Lee IS, Chen J, Leitner P, Besterman JM, Kinghorn AD, Pezzuto JM. Natural-product inhibitors of human DNA ligase I. Biochem. J 1996;314(Part 3):993–1000. [PubMed: 8615799]

(11). Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res 2000;28:235–242. [PubMed: 10592235]

(12). Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J. Comput. Chem 1983;4:187–217.

(13). MacKerell AD Jr. Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J,

Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B 1998;102:3586–3616.

(14). Mackerell AD Jr. Empirical force fields for biological macromolecules: overview and issues. J. Comput. Chem 2004;25:1584–1604. [PubMed: 15264253]

(15). Brooks CL III, Karplus M. Deformable stochastic boundaries in molecular dynamics. J. Chem. Phys 1983;79:6312–6325.

(16). Snyman, JA. Practical Mathematical Optimization: An Introduction to Basic Optimization Theory and Classical and New Gradient-Based Algorithms. Springer-Verlag; New York: 2005. p. 257

(17). Beglov D, Roux B. Dominant salvation effects from the primary shell of hydration. Approximation for molecular-dynamics simulations. Biopolymers 1995;35:171–178.

(18). Beglov D, Roux B. Finite representation of an infinite bulk system. Solvent boundary potential for computer-simulations. J. Chem. Phys 1994;100:9050–9063.

(19). Steinbach PJ, Brooks BR. New spherical-cutoff methods of long-range forces in macromolecular simulations. J. Comput. Chem 1994;15:667–683.

(20). Swope WC, Andersen HC, Berens PH, Wilson KR. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. J. Chem. Phys 1982;76:637–649.

(21). Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. J. Comput. Phys 1977;23:327–341.

(22). Kelley LA, Gardner SP, Sutcliffe MJ. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. Protein Eng 1996;9:1063–1065. [PubMed: 8961360]

(23). Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A geometric approach to macromolecule-ligand interactions. J. Mol. Biol 1982;161:269–288. [PubMed: 7154081]

(24). Connolly M. Analytical molecular surface calculation. J. Appl. Crystallogr 1983;16:548–558.

(25). Connolly ML. Solvent-accessible surfaces of proteins and nucleic acids. Science 1983;221:709–713. [PubMed: 6879170]

(26). Ferrin TE, Huang CC, Jarvis LE, Langridge R. The MIDAS display system. J. Mol. Graphics 1988;6:13–27.

(27). Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera, a visualization system for exploratory research and analysis. J. Comput. Chem 2004;25:1605–1612. [PubMed: 15264254]

(28). Halgren TA. MMFF VI MMFF94s option for energy minimization studies. J. Comput. Chem 1999;20:720–729.

(29). Halgren TA. MMFF VII characterization of MMFF94, MMFF94s, and other widely available force fields for conformational energies and for intermolecular-interaction energies and geometries. J. Comput. Chem 1999;20:730–748.

(30). Chambers CC, Hawkins GD, Cramer CJ, Truhlar DG. Model for aqueous solvation based on class IV atomic charges and first solvation shell effects. J. Phys. Chem 1996;100:16385–16398.

(31). Li J, Zhu T, Cramer CJ, Truhlar DG. New class IV charges model for extracting accurate partial charges from wave functions. J. Phys. Chem. A 1998;102:1820–1831.

(32). Sirois S, Hatzakis G, Wei D, Du Q, Chou K-C. Assessment of chemical libraries for their druggability. Comput. Biol. Chem 2005;29:55–67. [PubMed: 15680586]

(33). Ewing TJ, Makino S, Skillman AG, Kuntz ID. DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases. J. Comput.-Aided Mol. Des 2001;15:411–428. [PubMed: 11394736]

(34). Markowitz J, Chen J, Gitti R, Baldisseri DM, Pan YP, Udan R, Carrier F, MacKerell AD Jr. Weber DJ. Identification and characterization of small molecule inhibitors of the calcium-dependent S100B-p53 tumor suppressor interaction. J. Med. Chem 2004;47:5085–5093. [PubMed: 15456252]

(35). Hancock CN, Macias A, Lee EK, Yu SY, MacKerell AD Jr. Shapiro P. Identification of novel extracellular signal-regulated kinase docking domain inhibitors. J. Med. Chem 2005;48:4586–4595. [PubMed: 15999996]

(36). Huang N, Nagarsekar A, Xia G, Hayashi J, MacKerell AD Jr. Identification of non-phosphate-containing small molecular weight inhibitors of the tyrosine kinase p56 Lck SH2 domain via in silico screening against the pY + 3 binding site. J. Med. Chem 2004;47:3502–3511. [PubMed: 15214778]

(37). Leach AR, Kuntz ID. Conformational analysis of flexible ligands in macromolecular receptor sites. J. Comput. Chem 1992;13:730–748.

(38). Tanimoto, T. IBM Internal Report. Nov. 1957

(39). Godden JW, Xue L, Bajorath J. Combinatorial preferences affect molecular similarity/diversity calculations using binary fingerprints and Tanimoto coefficients. J. Chem. Inf. Comput. Sci 2000;40:163–166. [PubMed: 10661563]

(40). Durant JL, Leland BA, Henry DR, Nourse JG. Reoptimization of MDL keys for use in drug discovery. J. Chem. Inf. Comput. Sci 2002;42:1273–1280. [PubMed: 12444722]

(41). Willett P, Barnard JM, Downs GM. Chemical similarity searching. J. Chem. Inf. Comput. Sci 1998;38:983–996.

(42). Chen X, Pascal J, Vijayakumar S, Wilson GM, Ellenberger T, Tomkinson AE. Human DNA ligases I, III, and IV. Purification and new specific assays for these enzymes. Methods Enzymol 2006;409:39–52. [PubMed: 16793394]

(43). Mackey ZB, Niedergang C, Murcia JM, Leppard J, Au K, Chen J, de Murcia G, Tomkinson AE. DNA ligase III is recruited to DNA strand breaks by a zinc finger motif homologous to that of poly (ADP-ribose) polymerase. Identification of two functionally distinct DNA binding regions within DNA ligase III. J. Biol. Chem 1999;274:21679–21687. [PubMed: 10419478]

(44). Srivastava SK, Dube D, Tewari N, Dwivedi N, Tripathi RP, Ramachandran R. *Mycobacterium tuberculosis* NAD^+-dependent DNA ligase is selectively inhibited by glycosylamines compared with human DNA ligase I. Nucleic Acids Res 2005;33:7090–7101. [PubMed: 16361267]

(45). Chen X, Zhong S, Zhu X, Dziegielewska B, Ellenberger T, Wilson GM, MacKerell AD Jr. Tomkinson AE. Rational design of human DNA ligase inhibitors that target cellular DNA replication and repair. Cancer Res 2008;68:3169–3177. [PubMed: 18451142]

(46). Pan Y, Huang N, Cho S, MacKerell AD Jr. Consideration of molecular weight during compound selection in virtual target-based database screening. J. Chem. Inf. Comput. Sci 2003;43:267–272. [PubMed: 12546562]

(47). Carlson HA. Protein flexibility and drug design: how to hit a moving target. Curr. Opin. Chem. Biol 2002;6:447–452. [PubMed: 12133719]

(48). Carlson HA, McCammon JA. Accommodating protein flexibility in computational drug design. Mol. Pharmacol 2000;57:213–218. [PubMed: 10648630]

(49). Oprea TI, Davis AM, Teague SJ, Leeson PD. Is there a difference between leads and drugs? A historical perspective. J. Chem. Inf. Comput. Sci 2001;41:1308–1315. [PubMed: 11604031]

(50). Lipinski CA. Drug-like properties and the causes of poor solubility and poor permeability. J. Pharmacol. Toxicol. Methods 2000;44:235–249. [PubMed: 11274893]

(51). Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv. Drug Delivery Rev 2001;46:3–26.

(52). Feng BY, Shelat A, Doman TN, Guy RK, Shoichet BK. High-throughput assays for promiscuous inhibitors. Nat. Chem. Biol 2005;1:146–148. [PubMed: 16408018]

(53). McGovern SL, Helfand BT, Feng B, Shoichet BK. A specific mechanism of nonspecific inhibition. J. Med. Chem 2003;46:4265–4272. [PubMed: 13678405]

(54). Hertzberg RP, Caranfa MJ, Hecht SM. On the mechanism of topoisomerase I inhibition by camptothecin: evidence for binding to an enzyme—DNA complex. Biochemistry 1989;28:4629–4638. [PubMed: 2548584]

(55). Staker BL, Hjerrild K, Feese MD, Behnke CA, Burgin AB Jr. Stewart L. The mechanism of topoisomerase I poisoning by a camptothecin analog. Proc. Natl. Acad. Sci. U.S.A 2002;99:15387–15392. [PubMed: 12426403]

(56). Brown RD, Martin YC. An evaluation of structural descriptors and clustering methods for use in diversity selection. SAR QSAR Environ. Res 1998;8:23–39. [PubMed: 9517009]

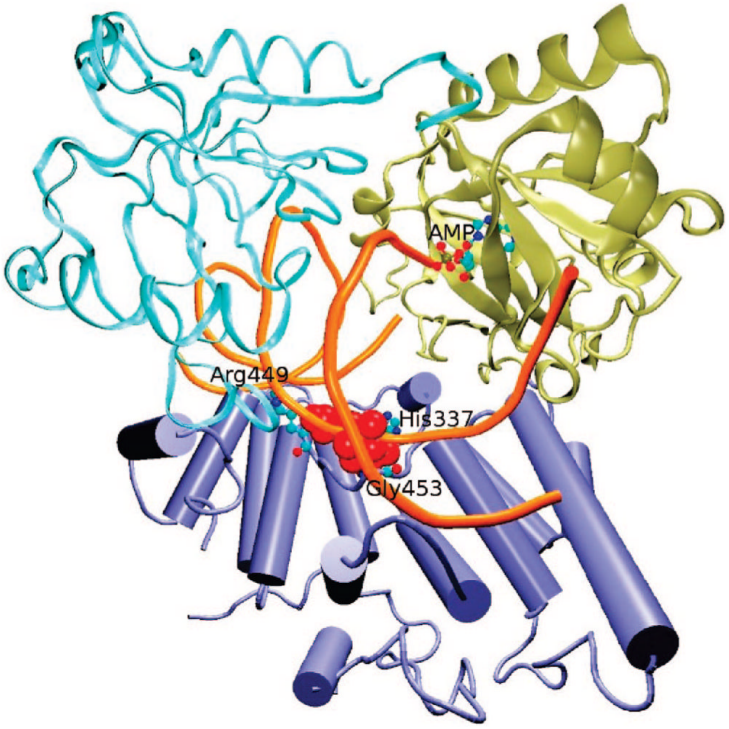


Figure 1.
The DNA substrate (orange tube) is encircled by three domains of human DNA ligase I, i.e., the DNA binding domain (DBD) containing residues Asp262—Ser535 (ice-blue carton), the adenylation domain (AdD) Pro536—Asp748 (wide tan ribbon), and the OB-fold domain (OBD) Tyr749—Ser901 (narrow cyan ribbon). The AMP cofactor (in CPK representation) is located in AdD. The putative binding site on DBD is represented by red spheres, and the three residues defining the binding pocket, His337, Arg449 and Gly453, are shown in CPK representation.

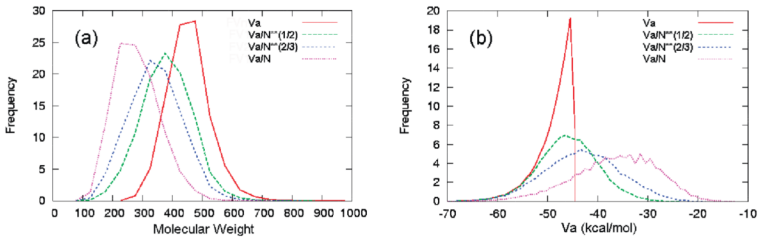


Figure 2.

Distributions of (a) the molecular weight and (b) the van der Waals attractive energy Va of the 50 000 compounds selected, via different normalizations by *N*, the number of heavy atoms of the compound, from the preliminary screening. The normalization $Va/N^{2/3}$, as shown by the blue dotted line, is used to select compounds.

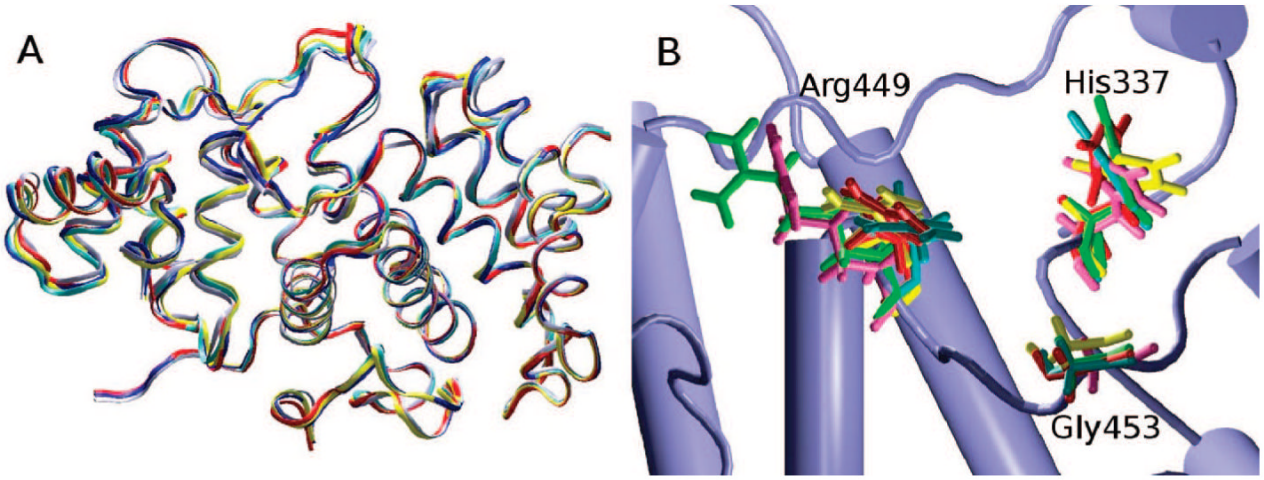


Figure 3.
(A) Alignment of the crystal structure and four conformations selected from the 5 ns MD simulation. (B) Orientations of three of the residues lining the binding site.

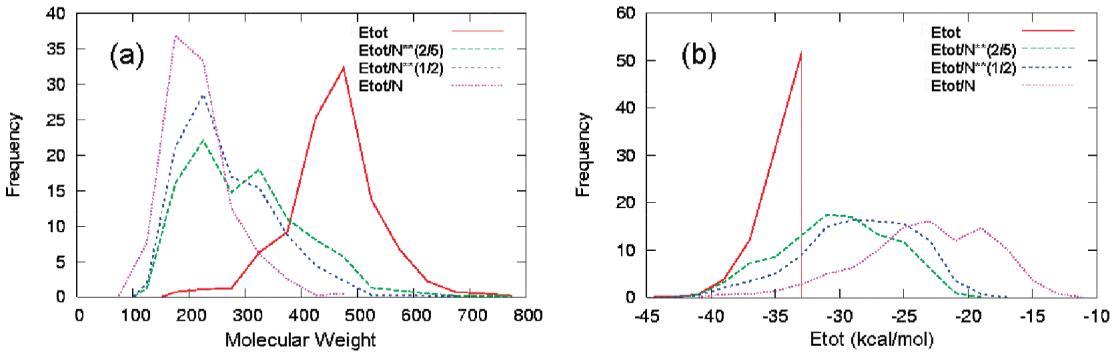


Figure 4.

Distributions of (a) the molecular weight and (b) the total interaction energy, Etot, of the 1000 compounds selected via different normalization schemes from the secondary docking.

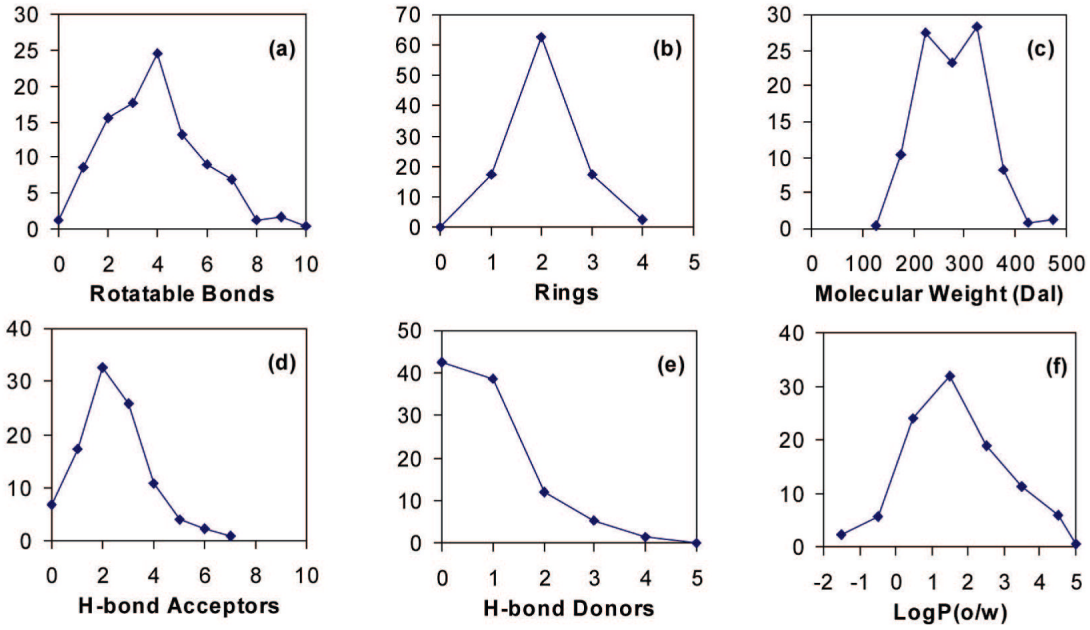


Figure 5.
Distributions of physical and molecular properties of the 233 selected compounds. The *Y* axes present the % frequency.

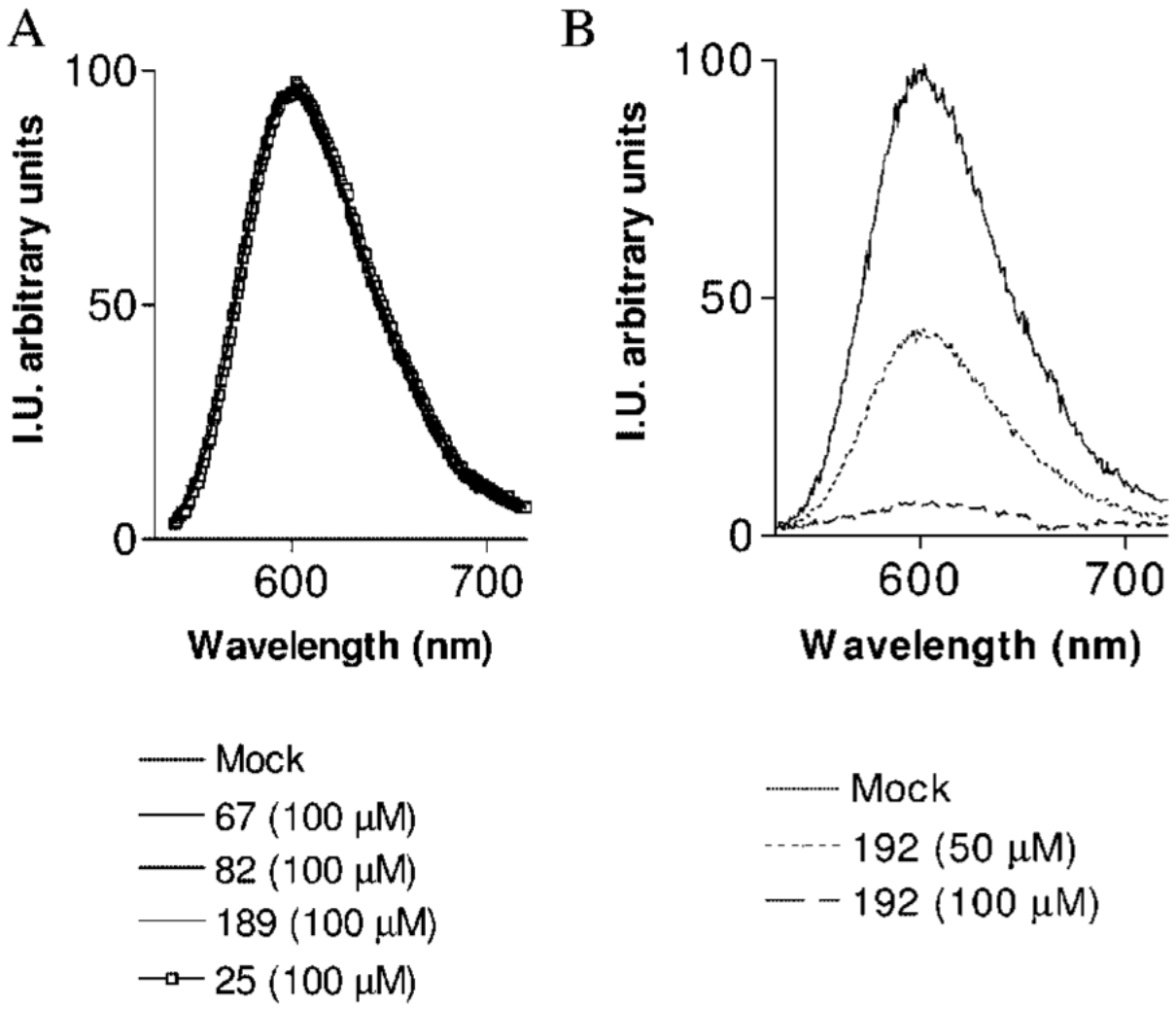
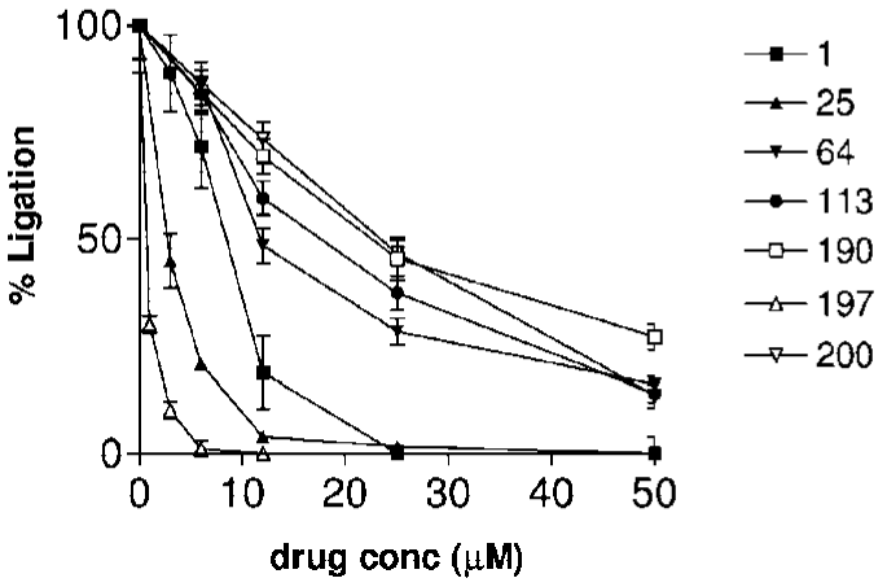


Figure 6.

Assessment of interaction of inhibitors with DNA by EtBr replacement assay. There is no binding of **25**, **82**, **67**, and **189** at 100 μM with DNA (A). In contrast, EtBr is displaced from DNA by **192** in a concentration (0, 50, 100 μM) dependent manner (B).



Drug No.	IC50
1	10 ± 5
25	4 ± 2
64	13 ± 5
113	18 ± 5
190	25 ± 6
197	0.6 ± 0.5
200	25 ± 6

Figure 7.

Concentration dependent inhibition of hLigI by small molecule inhibitors of hLigI identified by CADD. Left: Radioactively labeled nicked DNA substrate (0.5 pmol) and hLig1 (0.02 pmol) were incubated with varying concentrations of the indicated compounds. After agarose gel electrophoresis, ligation was detected and quantitated by phosphorimaging. The results of three independent experiments are shown graphically. Right: The IC_{50} values of the hLigI inhibitors are shown.

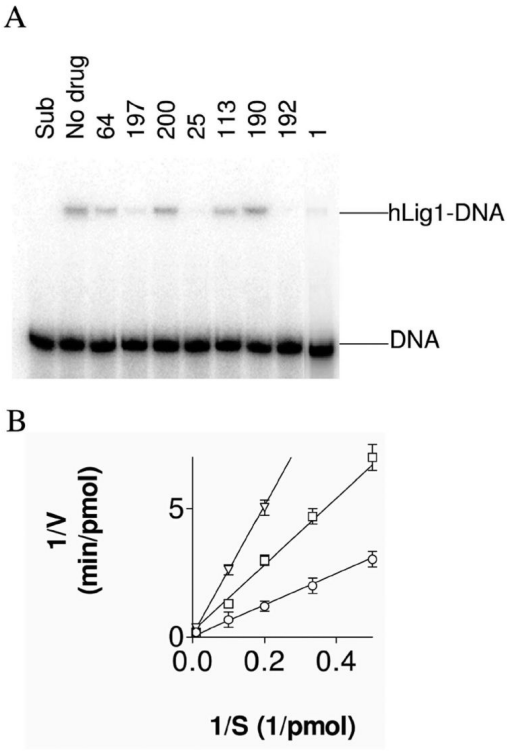


Figure 8.

Effect of inhibitors on formation of hLigI—DNA complex. A radioactively labeled DNA substrate containing a single nonligatable nick (1 pmol) and hLigI (1 pmol) were incubated without or with the indicated compounds at 100 μM. After separation by native gel electrophoresis, labeled DNA and DNA—protein complexes were detected by phosphorImager analysis (A). Compound **25** is a simple competitive inhibitor with respect to nicked DNA. hLig1 (0.05 pmol) was incubated in the absence and presence of L25 at 12 μM (□) and 25 μM (▽) with increasing amounts of a linear nicked DNA substrate. Lineweaver—Burk double-reciprocal plots of initial reaction velocity versus substrate concentration are shown (B).

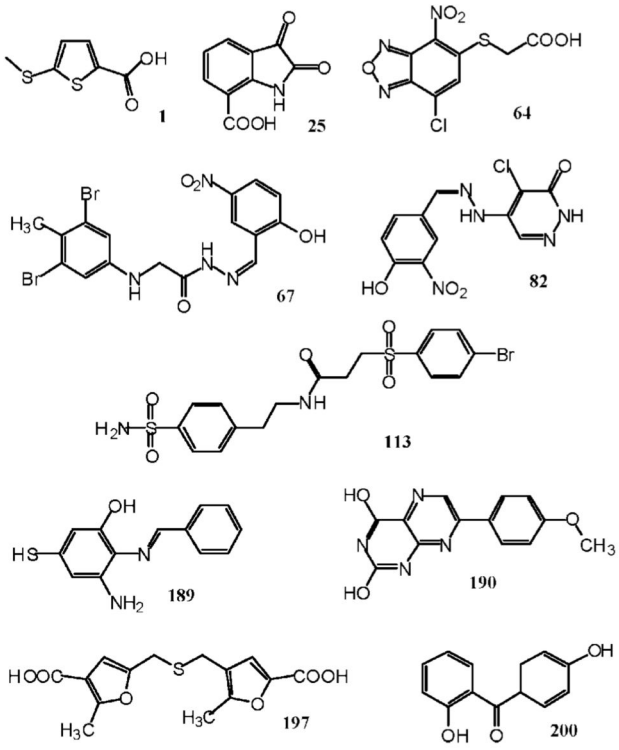


Figure 9.
Structures of the 10 active compounds.

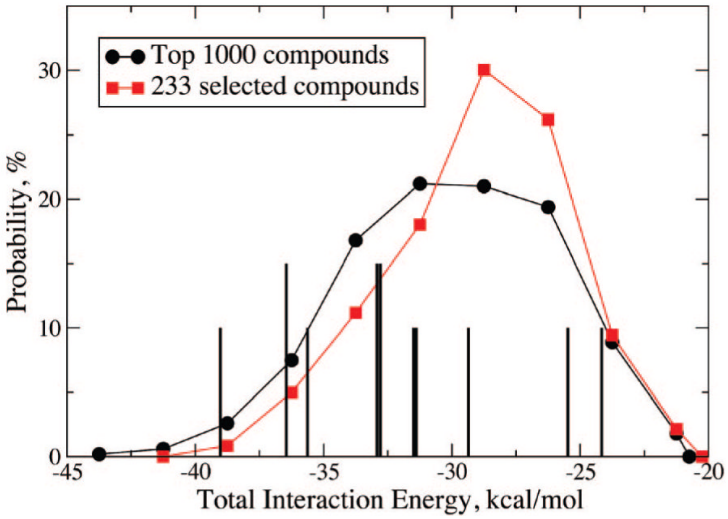


Figure 10.

Distribution of the total interaction energy scores for the top 1000 compounds (black circles), the selected 233 compounds (red squares), and the 10 compounds that specifically inhibit hLigI activity by >50% (spikes). The scores for the three characterized compounds **67**, **82**, and **189** are shown as longer spikes; the wider spike is associated with two compounds having similar scores.

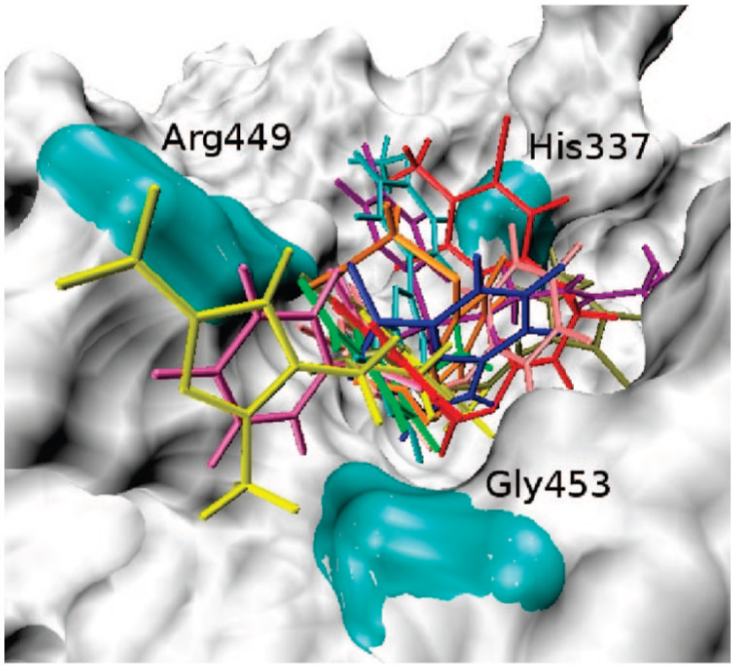


Figure 11.
Perspective view of the docked orientations of the 10 characterized compounds in the putative binding site surrounded by residues His337, Arg449, and Gly453 colored in cyan on the surface of the hLig1 DBD.

Table 1

The rmsd Values in Å between Each Pair of the Five Conformations Used for Database Screening, Including the Crystal Structure (1×9n) and the Four MD Generated Conformations (C2—C5)[a]

	1×9n	C2	C3	C4	C5
1×9n	0				
C2	2.18	0			
C3	2.29	1.82	0		
C4	2.23	1.65	1.42	0	
C5	2.45	2.15	1.74	1.43	0

[a]Only residues related to the binding region are used in the calculation (see text).

Table 2

Molecular Properties of the 16 Active Compounds, Their Experimental Inhibition of hLig1, and Their Growth Inhibition of Cultured Human Cells[a]

compd	Rot	Ring	MW	Ha	Hd	log *P*	hLig1_Inh% (100 μM)	Growth_Inh% (50 μM)	
								MCF10A	HCT116
1	2	1	174.24	2	1	1.92	60.5 ± 5.0		
25	1	2	190.13	2	1	0.24	59.3 ± 7.5		
32	6	3	385.38	6	1	2.10	83.4 ± 3.1		
64	4	2	288.65	2	0	2.15	60.6 ± 0.54	40	
67	6	2	485.11	2	2	4.91	78.3 ± 8.5	70	70
82	4	2	308.66	3	2	4.84	59.8 ± 1.7	70	40
113	8	2	475.38	5	2	1.42	53.4 ± 7.6		
123	4	1	228.32	4	0	1.51	66.9 ± 3.5		
175	1	2	173.15	2	0	1.12	54.2 ± 16.5		
180	2	3	301.21	2	1	2.44	58.3 ± 6.0	70	70
189	2	2	244.28	3	1	1.96	68.8 ± 4.9	70	70
190	2	3	268.23	5	0	1.78	52.5 ± 6.6	70	40
192	2	3	274.28	2	1	1.56	68.2 ± 1.5		70
197	6	2	308.31	0	0	1.87	91.0 ± 6.9		
200	2	2	212.2	1	0	3.04	50.7 ± 10.9		
202	4	4	390.42	1	0	5.08	95.0 ± 6.0	50	

[a]Nonboldfaced compounds are those that also inhibit T4 DNA ligase by >50%. compd is the compound label, Rot the number of rotatable bonds, Ring the number of rings, MW the molecular weight, Ha the number of H-bond acceptors, Hd the number of H-bond donors, and log *P* the octanol/water partition coefficient. DNA joining and cell culture assays were performed as described in Methods. Inhibition of joining and proliferation is expressed as a percentage of values obtained with DMSO alone.

Table 3

Similarity between the 10 Active Compounds (Bold Numbers) Based on the Tanimoto Index (Tc%)

	1	**25**	**64**	**67**	**82**	**113**	**189**	**190**	**197**	**200**
1	100									
25	26	100								
64	24	39	100							
67	9	38	53	100						
82	11	42	53	69	100					
113	14	22	34	45	30	100				
189	16	40	40	37	43	24	100			
190	15	43	28	32	39	18	52	100		
197	35	40	35	24	21	18	18	22	100	
200	19	40	19	25	27	13	25	35	35	100