5872–5879 *Nucleic Acids Research, 2006, Vol. 34, No. 20* *Published online 24 October 2006*
doi:10.1093/nar/gkl747

Mimicking damaged DNA with a small molecule inhibitor of human UNG2

Daniel J. Krosky, Mario A. Bianchet[1], Lauren Seiple, Suhman Chung, L. Mario Amzel[1] and James T. Stivers*

Department of Pharmacology and Molecular Sciences and [1]Department of Biophysics and Biophysical Chemistry of the Johns Hopkins Medical School, 725 North Wolfe Street, Baltimore, MD 21205, USA

Received September 5, 2006; Revised September 22, 2006; Accepted September 26, 2006

ABSTRACT

Human nuclear uracil DNA glycosylase (UNG2) is a cellular DNA repair enzyme that is essential for a number of diverse biological phenomena ranging from antibody diversification to B-cell lymphomas and type-1 human immunodeficiency virus infectivity. During each of these processes, UNG2 recognizes uracilated DNA and excises the uracil base by flipping it into the enzyme active site. We have taken advantage of the extrahelical uracil recognition mechanism to build large small-molecule libraries in which uracil is tethered via flexible alkane linkers to a collection of secondary binding elements. This high-throughput synthesis and screening approach produced two novel uracil-tethered inhibitors of UNG2, the best of which was crystallized with the enzyme. Remarkably, this inhibitor mimics the crucial hydrogen bonding and electrostatic interactions previously observed in UNG2 complexes with damaged uracilated DNA. Thus, the environment of the binding site selects for library ligands that share these DNA features. This is a general approach to rapid discovery of inhibitors of enzymes that recognize extrahelical damaged bases.

INTRODUCTION

The RNA base uracil is one of the most prevalent non-canonical bases found in genomic DNA (1). It arises from spontaneous or intentional enzymatic deamination of cytosine in DNA (2–5), or alternatively, by misincorporation of dUTP in place of TTP during DNA replication (6). Both pathways for uracil incorporation are forms of DNA damage, and accordingly, an elaborate uracil base excision repair (UBER) mechanism is present in all organisms to reverse this damage (Figure 1A) (7). Without repair, U/G mismatches lead to T/A transition mutations and corresponding changes in protein sequence. Although U/A base pairs arising from misincorporation of dUTP are not mutagenic, if large numbers of uracils are inserted on both strands of replicated DNA this can lead to disruptions in gene expression, and even double strand DNA breaks can arise from the base excision repair process (8). Although the accidental appearance of uracil in DNA is well-appreciated, it has become apparent that enzymatic deamination of cytosine to uracil in DNA plays a key role in the processes of somatic hypermutation and class switch recombination in B cells (2,9,10), in certain B cell lymphomas (11), and as an innate host defense mechanism against retroviral infection (12). In addition, the widely used chemotherapeutic agent 5-fluorouracil (5-FU) promotes uracil misincorporation into DNA by increasing the ratio [dUTP]/[TTP] in the cell, suggesting inhibitors of UBER could serve as sensitizers during 5-FU therapy (8,13). In general, these diverse roles for uracil indicate that small molecule UBER inhibitors might be very useful investigational or therapeutic agents.

In humans, the first step in the UBER pathway, cleavage of the glycosidic bond of deoxyuridine in DNA, is catalyzed by the powerful nuclear enzyme uracil DNA glycosylase (UNG2) (14). This extensively studied glycosylase uses an extrahelical recognition mechanism in which the uracil base that is originally embedded in the DNA base stack is ultimately extruded into the enzyme active site (Figure 1B) (15). Thus by 'uracil flipping' the enzyme can recognize the unique structural features of uracil that allows catalysis to proceed (7). Although the uracil is attached to a large duplex DNA substrate, most of the enzyme–substrate interactions involve the base itself. Perhaps not surprisingly, the uracil base alone has been found to be a product inhibitor of the enzyme ($K_i \sim 300$ μM at physiological pH) (16).

We recently published a general strategy for rapid discovery of small molecule inhibitors of UNG2 and other UBER enzymes called 'substrate fragment tethering' (SFT) (17), which is an efficient variation of the combinatorial target-guided ligand assembly method of Ellman *et al*. (18). The basic approach is extremely simple and involves tethering a chemical library of aldehydes to pieces of substrates (such as uracil) that already bind weakly to an enzyme active site. As shown in Figure 2, UNG2 library synthesis involves

*To whom correspondence should be addressed. Tel: +1 410 502 2758; Fax: +1 410 955 3023; Email: jstivers@jhmi.edu

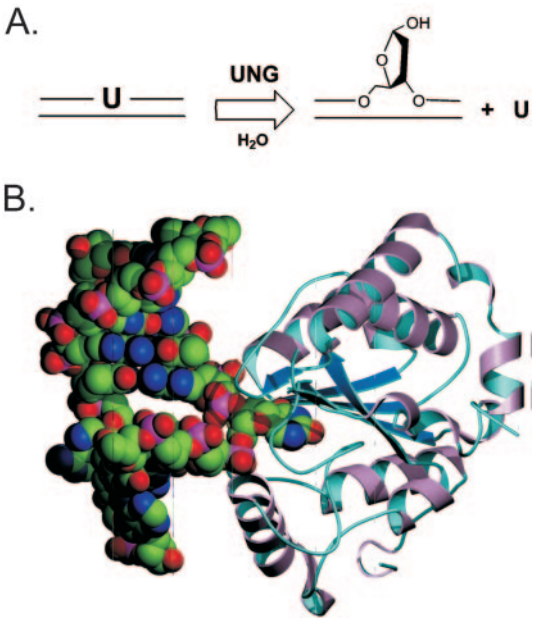

Figure 1. Uracil DNA base excision repair and extrahelical recognition of uracil. (**A**) Uracil in the context of a U/A or U/G base pair is repaired by a series of enzymatic reactions that restore the integrity of the DNA sequence. The first enzyme in the pathway is uracil DNA glycosylase (UNG) that hydrolytically cleaves the *N*-glycosidic bond connecting the uracil to the deoxyribose, leaving an abasic site and free uracil. Three other enzymes complete the repair process in humans: an abasic site endonuclease (APE1), a dual-function repair polymerase (pol β) that inserts the correct nucleotide and eliminates the abasic site via a β-elimination reaction, and finally, DNA ligase. (**B**) Structure of uracilated DNA bound to human UNG2 (PDB code 1EMH).

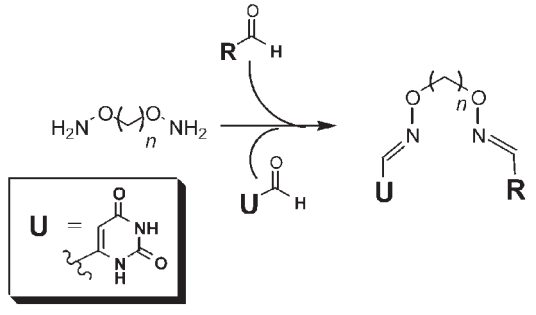

Figure 2. Chemistry of substrate fragment tethering. In this approach, a suitable substrate fragment (such as uracil) is identified and derivatized at a nonperturbing position with an aldehyde functional group. In the case of UNG2, the substrate fragment ii 6-formyluracil. The substrate fragment is tethered to one end of a bifunctional alkyloxyamine linker of variable length ($n = 2$–6), which is then derivatized on the other end with a library of aldehyde binding elements (RCHO). Although statistical mixtures of all possible oximes result (25% each homodimer derived from uracil-CHO or RCHO, and 50% of the heterodimer derived from uracil-CHO and RCHO), this poses no difficulty because the crude mixtures are directly screened for inhibitory activity. Once active mixtures are identified, the specific inhibitory components can be rapidly identified by deconvolution to uncover the linker length that gave rise to the observed inhibition. The structures of the 215 aldehyde library members used in this study are reported in Supplementary Table S1.

highly efficient formation of oxime linkages between a bivalent alkyloxyamine linker, a uracil aldehyde derivative, and each library aldehyde member. Thus, the uracil fragment targets the entire tethered molecule to the active site where the library pieces can then explore adjacent binding pockets. SFT has the following strengths: (i) library synthesis is economical and very rapid and can be performed in microtiter plate format, (ii) the reactions are extremely efficient and no purification of any products is required, (iii) mixtures of flexible linkers are used in each reaction which allows multiple tethering lengths to be probed simultaneously in activity screens, and (iv) the method is easily adaptable to any desired target.

Here we report the results from high-throughput screening (HTS) of an SFT library derived from tethering 6-formyl uracil to a library of 215 different aldehyde-binding elements. Of the two hits identified in this screen, the most potent SFT ligand was co-crystallized with UNG2 to yield a high-resolution structure of the complex. This first portrait of a bound SFT ligand shows how a small molecule can surprisingly mimic the hydrogen bonding and electrostatic interactions of a larger DNA substrate. Thus, the environment of the binding site appears to select for library ligands that share molecular features of DNA. This efficient approach should be easily adaptable to other DNA repair glycosylases that recognize extrahelical bases.

MATERIALS AND METHODS

Reagents and general methods

All chemicals were purchased from commercial sources without further purification unless otherwise stated. The ^{1}H NMR spectra were recorded on a 400 MHz Varian Innova instrument in hexadeuteriodimethylsulfoxide (DMSO-d_6). The chemical shifts of protons are given in p.p.m. with DMSO as an internal standard.

Synthesis of tethered oxime libraries

The 6-formyluracil tethered library was synthesized as described previously (17). Briefly, a set of 215 alkyl and aryl aldehydes (Supplementary Table S1) was selected for library synthesis and coupled to 6-formyluracil using the *O*,*O*′-diaminoalkanediol linkers as follows. To each 0.5 ml well of a Matrix microtiter plate was added a DMSO stock solution of AcOH (20 µl, 150 mM, 3 µmol), 6-formyluracil (20 µl, 150 mM, 3 µmol) and a single alkyl or aryl aldehyde (20 µl, 150 mM, 3 µmol). The plate was carefully agitated to make the solutions homogenous. To each of the mixtures was added a DMSO solution of the *O*,*O*′-diaminoalkanediol linkers containing each of the five linker lengths in equal proportion (22 µl, 150 mM, 3.3 µmol total amine equivalents). The plate was sealed, further agitated and incubated in an oven for 12 h at 37°C.

Deconvolution of inhibitory mixtures

The two active mixtures containing compounds **1** and **2** were deconvoluted with respect to linker length by individually synthesizing each oxime dimer using a *single* diaminoalkanediol linker per reaction. At this stage we did not separate the homodimers from the heterodimers in the mixtures. The linker length dependence of the inhibition is reported in Supplementary Table S2. The corresponding compounds were then synthesized in larger scale and purified for complete analysis of their inhibition properties as described below.

Synthesis of 1

Solutions (0.15 M) of 4-carboxybenzaldehyde (888 µl, 0.165 mmol), 6-formyluracil (888 µl, 0.165 mmol) and acetic acid (888 µl, 0.165 mmol) in DMSO were added to a reaction vessel. The reaction was initiated by the addition of 0.15 M *O*,*O*′-diaminoethanediol (888 µl, 0.165 mmol) in DMSO, and incubated at 37°C for 36 h. The desired heterosubstituted compound was purified by direct injection of the reaction mixture onto a Phenomenex Aqua reversed phase C-18 HPLC column (250 mm, 10 mm, 5 µm) using gradient elution from 0 to 65% CH_3CN in 0.1 M aqueous TEAA over the course of 2 h using UV detection at 320 nm. Fractions containing **1** were combined and concentrated *in vacuo*. The compound was precipitated using ice-cold water, centrifuged, washed twice with ice-cold water and dried *in vacuo*. This yielded **1** as a white powder (9.6 mg, 0.028 mmol) in 34% yield. ^{1}H NMR (400 MHz, DMSO-d_6): δ 11.19 (s, 1H), 10.81 (s, 1H), 8.37 (s, 1H), 7.98 (s, 1H), 7.96 (d, 3H), 7.72 (d, 2H), 5.79 (s, 1H), 4.66 (d, 4H); ^{13}C NMR (400 MHz, DMSO-d_6): δ 167.6, 163.9, 151.0, 148.7, 144.5, 142.5, 129.8, 126.9, 102.0, 73.2, 71.9; (*m*/*z*): $[M+H]^+$ calcd for $C_{15}H_{14}N_4O_6$, 347.0986; found, 347.0991.

Synthesis of 2

The synthesis and purification was identical to that reported above for **1** except that 0.15 M solutions of 3-carboxybenzaldehyde and *O*,*O*′-diaminopropanediol were used. ^{1}H NMR (400 MHz, DMSO-d_6): δ 11.16 (s, 1H), 10.78 (s, 1H), 8.34 (s, 1H), 8.18 (s, 1H), 7.94 (d, 2H), 7.83 (d, 2H), 7.53 (t, 1H), 5.76 (s, 1H), 4.27 (t, 2H), 4.20 (t, 2H), 2.10 (m, 2H); ^{13}C NMR (400 MHz, DMSO-d_6): δ 166.9, 163.9, 151.0, 148.3, 144.6, 142.2, 132.5, 131.5, 130.9, 130.5, 129.2, 127.6, 101.7, 71.8, 70.5, 28.4; (*m*/*z*): $[M+H]^+$ calcd for $C_{15}H_{14}N_4O_6$, 361.1143; found, 361.1153.

High-throughput screening of oxime library

The DNA substrate in this HTS assay was synthesized using standard phosphoramidite DNA solid-phase chemistry using reagents purchased from Glen Research. The DNA was purified using anion exchange chromatography followed by desalting using reversed phase methods. The sequence and size was confirmed using analytical denaturing PAGE and MALDI-MS. The substrate is a single-stranded 28mer DNA hairpin containing nine U·A base pairs and a hexapolyethylene glycol (PEG_6) linker (5′-FAM-GCA CUU AAG AAU UG-PEG_6-CA AUU CUU AAG UGC-DABSYL-3′). The UNG2 HTS assay has been described previously (17).

IC_{50} determinations

To a 96-well plate was added 5 µl compound **1** in DMSO, followed by 75 µl of 66.5 nM PEG-U9 hairpin in reaction buffer (20 mM Tris–HCl, pH 8.0, 50 mM KCl, 0.2 mM $MgCl_2$ and 0.05% Brij-35). Eight different inhibitor concentrations were used in the range of 0.045–100 µM. Reactions were initiated by the addition of 20 µl of 0.5 nM human UNG in reaction buffer. The final concentrations of reagents in the assay are 20 mM Tris–HCl, pH 8.0, 50 mM KCl, 0.2 mM $MgCl_2$, 0.05% Brij-35, 5% DMSO, 0.1 nM human UNG, 50 nM PEG-U9 hairpin DNA and 0–100 µM **1**. Wells containing DMSO vehicle only or no UNG2 were used as negative controls and background, respectively. The plates are incubated at ambient temperature in a fluorescence plate reader for 30 min, and the progress of the reaction was monitored every 5 min (Ex. 485 nm/Em. 520 nm). Percent inhibition versus log concentration of **1** data were fit to a four parameter sigmoidal dose–response equation (Equation 1) using Prism 4.0 (GraphPad Software Inc., San Diego, CA, USA).

$$\%\ \text{Inhibition} = \frac{\text{Min}(\text{Max} - \text{Min})}{1 + 10^{((\log IC_{50} - \log[1]) * n)}}. \qquad \mathbf{1}$$

Mechanism of inhibition

To a 96-well plate was added 5 µl compound in DMSO, followed by 75 µl PEG-U9 hairpin in reaction buffer (20 mM Tris–HCl, pH 8.0, 50 mM KCl, 0.2 mM $MgCl_2$ and 0.05% Brij-35). Eight different DNA concentrations were used in the range of 27.5–1100 nM. Reactions were initiated by the addition of 20 µl of 0.5 nM human UNG in reaction buffer. The final concentrations of reagents in the assay are 20 mM Tris–HCl, pH 8.0, 50 mM KCl, 0.2 mM $MgCl_2$, 0.05% Brij-35, 5% DMSO, 0.1 nM human UNG, 27.5–1100 nM PEG-U9 hairpin DNA and 0–128 µM of **1**. The plates were incubated at ambient temperature in a fluorescence plate reader for 60 min, and the progress of each reaction was monitored every 30 s ($\lambda_{ex} = 485$ nm, $\lambda_{em} = 520$ nm). Afterwards, *Escherichia coli* UNG was added to each well to drive the reactions to completion, and the total change in fluorescence corresponding to complete consumption of the substrate was calculated (ΔFU_{tot}). These values were used to calculate initial molar velocities (i.e. [product]/s = $\Delta\mu M/\Delta FU_{tot} \times$ FU/s). Mechanisms of inhibition and their corresponding inhibitor dissociation constants were determined by Lineweaver–Burk slope and intercept replot analysis.

Cell culture studies

Currently, there exists no simple assay to monitor the efficacy of inhibitors of UNG within living cells. In order to evaluate the ability of compound **1** to inhibit UNG *in vivo*, inhibitor was added to cells in culture, lysates were carefully prepared to minimize dilution of the inhibitor, and UNG activity was assayed as follows (Figure 3). PC-3 human prostate adenocarcinoma cells were grown in Modified Eagle's Medium (Mediatech) supplemented with 10% fetal bovine serum (Hyclone) and penicillin/streptomycin. Cells at 70% confluency were trypsinized, resuspended in growth media and counted using a hemocytometer. The cells were then centrifuged and resuspended to a density of 1×10^6 cells/ml. Aliquots of 1×10^6 cells were pelleted and resuspended in 46.5 µl growth media with or without 3 mM compound **1** to achieve a final concentration of 3 mM inhibitor. The cells were incubated at 37 °C, 5% CO_2 for 30 min before being spun down and resuspended in 7 µl UNG lysate reaction buffer [10 mM Tris–HCl, pH 8.0, 60 mM NaCl, 1 mM DTT, 1 mM EDTA, 0.002% Brij 35, protease inhibitors (Roche Complete, Mini)]. The cells were lysed by freeze-thawing five times, the lysates were centrifuged at 14 000 r.p.m. for 30 min and the supernatant was transferred to a fresh tube, and the volume was now increased by 3.5 µl based on the number of cells pelleted (106) and the average cell volume of 3.5 pl for PC-3 cells (19). Thus, it is calculated that 3.5 µl × 3 mM (1) = 10 nmol of compound **1** was trapped

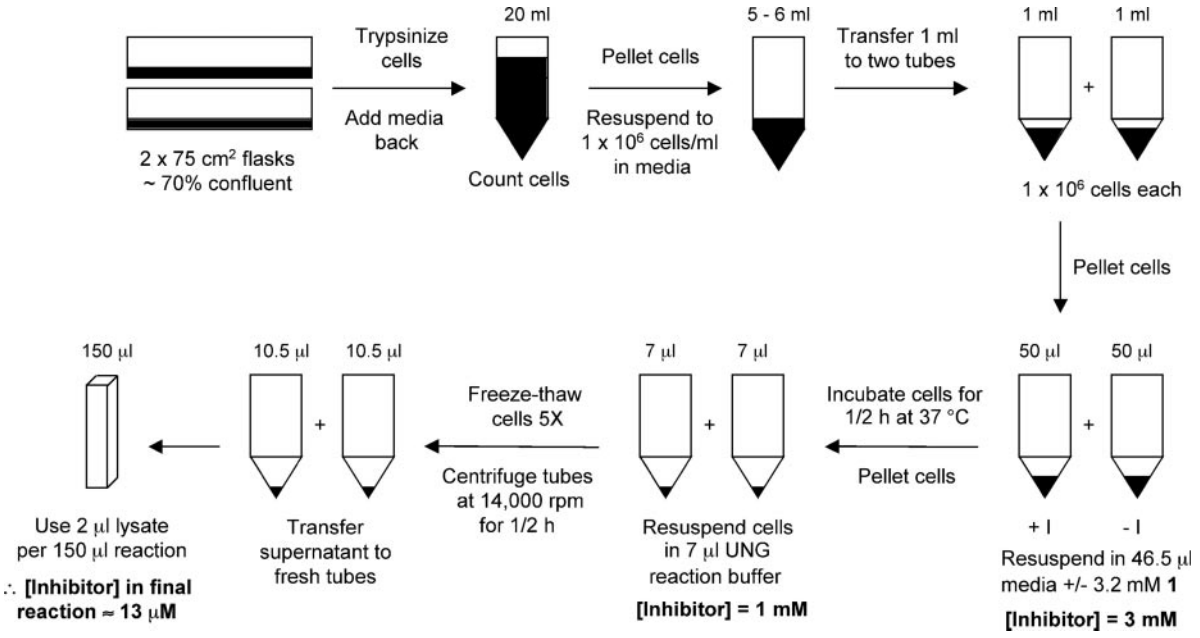

Figure 3. Assay for small molecule inhibition of UNG in cells.

in the pelleted cells. The UNG activity was measured using the PEG-U9 hairpin substrate and 2 µl lysate in a total reaction volume of 150 µl. Assuming equilibration of **1** across the cell membrane, its final concentration in the fluorescence reaction was 13 µM. Initial reaction velocities were calculated using the total change in fluorescence as described (1) and normalized to protein content measured using the BioRad protein assay.

Crystallization of the complex of UNG2 and 3-(2)-A8

Human UNG2 was expressed and purified as described previously (20). A solution of human UNG2 (112.5 µl, 44.2 mg/ml) in a buffer containing 50 mM Tris–OAc, pH 7.0, 150 mM NaCl and 1 mM DTT, was mixed with **1** (12.5 µl, 16.8 mM) in 100 mM Tris–HCl, pH 8.0 and 5% DMSO. The mixture was allowed to incubate at ambient temperature for 30 min, and then centrifuged at 10 000× *g* for 5 min. Co-crystallization conditions were screened using the Nextal PEG Suite library. A total of 300 nl of the complex was mixed with an equal volume of precipitant, and allowed to crystallize at 22°C using the hanging drop method. Crystals were observed within 48 h with 0.2 M potassium thiocyanate, 20% PEG 3350. X-ray diffraction data were collected from a flash frozen crystal in its unmodified mother liquor at the National Synchrotron Light source at Brookhaven National Laboratory (beam line X6A) using a wavelength of 1.1 Å with a ADSC CCD detector Quantum-4. The package HKL2000 (21) was used for data reduction.

The structure was determined by molecular replacement with the program MOLREP using the uncomplexed UNG2 structure (1KHZ) as the searching model. After an initial rigid-body refinement, compound **1** was placed in a difference Fourier electron density. The final model of the UNG2-**1** complex, refined using REFMAC5 (22) with isotropic temperature factors, shows all non-glycine residues in allowed regions of the Ramachandran plot and excellent stereochemistry (Table 1). Riding hydrogens of protein atoms were used in REFMAC5. The structural statistics are reported in Table 1.

RESULTS AND DISCUSSION

We previously used SFT to identify two uracil-tethered inhibitors of UNG2 from a small aldehyde library containing only 14 different aldehyde binding elements (17). Although these SFT compounds had K_i values between 0.3 and 3 µM, the non-uracil binding elements were derived from unstable di- or trihydroxybenzaldehydes. Thus, these initial compounds were prone to air oxidation to the inactive quinone forms, and were not suitable for structural or cell culture studies. We therefore turned to screening a much larger 215 member aldehyde library using our high-throughput molecular beacon fluorescence assay (see Supplementary Data) (17). This screening effort resulted in the identification of two new active mixtures (Figure 4). After deconvolution to identify the linker length that gave rise to inhibition, and purification of the individual inhibitory compounds, the IC_{50} values were determined (Figure 5A).

The library binding elements that gave rise to the observed inhibition shared a common chemical structure The two most potent compounds **1** and **2**, with IC_{50} values of 9 and 11 µM, respectively, both shared formate-substituted benzaldehyde functional groups and short alkyl chain linker lengths of $n = 2$ or 3 (Figure 4). These structure-activity trends suggested the presence of a binding pocket directly adjacent to the uracil binding site that depends on positioning of the negatively charged formate groups of either **1** or **2**. To date, tethering library binding elements to the 6-formyl uracil substrate fragment has brought about increases in binding affinity of up to −3 kcal/mol as compared to the *O*-methyl

Table 1. Crystallographic data collection and refinement statistics

Space group	P212121	
Cell dimensions	$a = 43.2$, $b = 69.1$ and $c = 70.4$ Å	
Resolution range (Å)	49.3–1.3	
R_{sym} (Last shell)[a]	0.06 (0.49)	
Completeness (Last shell)	98.0 (92.4)%	
Multiplicity (Last shell)	4.9 (4.2)	
$I/\sigma(I)$ (Last shell)	10.2 (2.6)	
Number of reflections	51 470	
Refinement		
F Data cutoff in σ (F) units	0	
	Number of atoms	Average B (Å^2)
Protein	1808	11.1
Solvent	315	21.0
Ligand	25	13.8
Total	2148	12.6
R-value		0.18
R_{free} (test set of 5%)		0.21
Stereochemical constraints		
Bond length rms (Å)	0.007	
Bond angles rms (degrees)	1.26	
Improper angles rms (degrees)	0.07	

[a]$R_{sym} = \sum_h \sum_j |I_{hj} - \langle I_h \rangle| / \sum_h \sum_j |I_{hj}|$, where h represents a unique reflection and j means symmetry equivalent indices, I is the observed intensity, and $\langle I \rangle$ is the mean value of I.

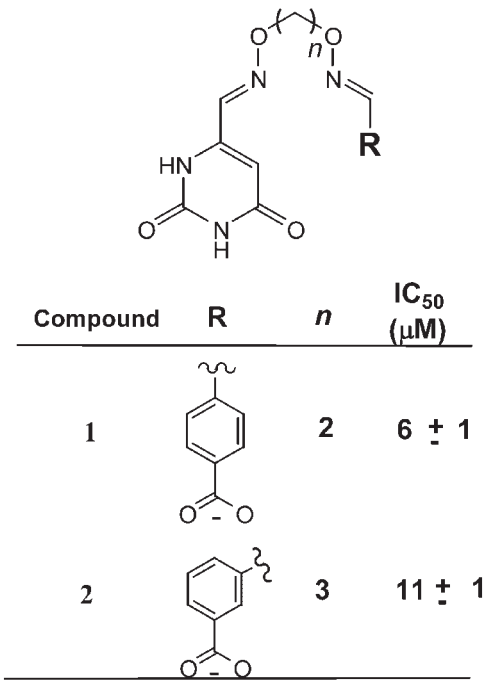

Figure 4. Structure of inhibitory compounds identified from high-throughput screening. IC_{50} values were determined for each purified compound and a full mode-of-inhibition analysis and structural characterization was performed for compound **1**. The IC_{50} value for compound **1** is equivalent to its true K_i.

oxime derived from 6-formyl uracil and methoxyamine ($K_i = 45$ μM) (17).

Since **1** showed the highest activity we investigated its mode of inhibition in greater detail. Interestingly, most of our previously reported SFT inhibitors, as well as the uracil base itself, showed surprisingly complex modes of inhibition with either competitive or partial uncompetitive binding to

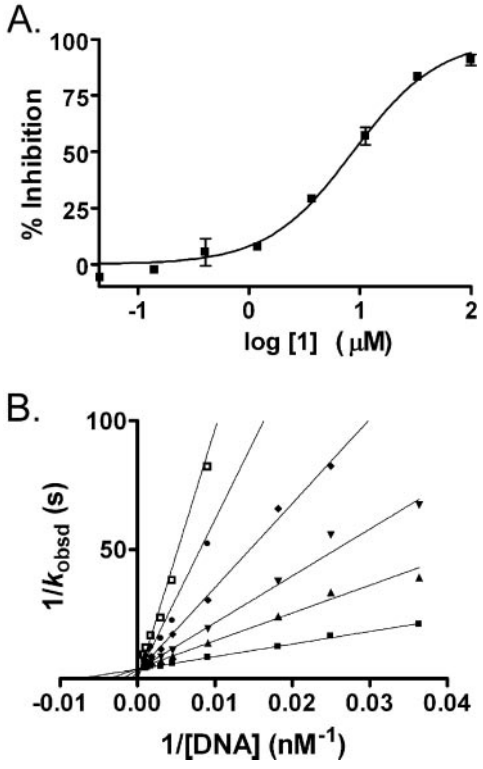

Figure 5. Inhibition by **1**. (A) Concentration dependence of inhibition. The curve is a nonlinear least-squares fit to Equation 1 ($IC_{50} = 9 \pm 1$ μM). (B) Mode-of-inhibition analysis for compound **1**. Linear competitive inhibition was observed: $K_i = 6 \pm 1$ μM.

two distinct uracil binding sites (17). This surprising complexity, which is not entirely understood at a structural level, was attributed to the presence of a second weak uracil binding site that may be occupied transiently during the process of uracil flipping into the active site. In contrast with the previous complexity, the inhibition patterns for **1** indicated simple competitive inhibition (Figure 5B), with linear Lineweaver–Burk slope replots (data not shown). SFT inhibitor **1** also showed an IC_{50} value against full-length UNG2 that was only 25% greater than the catalytic domain (the full length UNG2 was assayed using cell extracts). The full-length UNG2 protein differs from the catalytic domain by a 90 amino acid N-terminal extension that is involved in nuclear localization and other protein interactions (23). Thus, mode of inhibition analysis indicates that **1** competes for binding to the extrahelical uracil binding site observed in the uracilated-DNA complex shown in Figure 1B and that its inhibitory potency is not affected by the N-terminal extension present in nuclear UNG2.

We also investigated the potency of **1** in cell culture. Since there is no simple marker for assessing UNG2 inhibition in cell culture, an *ex vivo* assay was developed to assess whether the inhibitor enters cells and binds to UNG2. In this assay, cells are treated with a single high concentration of inhibitor (3 mM), and then carefully diluted cell extracts are prepared for fluorometric assay of UNG2 activity (17). Assuming full equilibration of the inhibitor across the cell membrane, and taking into account extract dilutions and measured cell

numbers and volumes (19), the concentration of **1** in the final UNG2 assay was calculated to be 13 μM (Figure 3). This concentration consistently gave rise to 20 ± 4% inhibition of UNG2 activity relative to control extracts prepared identically and in parallel, which is <70% inhibition expected from a competitive inhibitor with a $K_i = 6$ μM. This difference may reflect that (i) **1** is poorly membrane permeable, or (ii) that **1** is not metabolically stable in the intracellular environment. In this respect, oximes are known to be reduced by microsomal NADH cytochrome b5 reductase (24,25).

To evaluate the structural basis for inhibition, UNG2 was cocrystallized with **1** and diffraction data were collected to 1.3 Å resolution and refined to an R_{factor} and R_{free} of 0.19 and 0.22, respectively (PDB ID 2HXM, Figure 6A). In contrast to damaged DNA binding, which leads to a contraction of the active site structure (15) (26), binding of **1** led to only minor structural changes as compared to uncomplexed UNG2 (pdb code 1AKZ), with an r.m.s. deviation over 221 C-α atoms of only 0.58 Å. Despite the differences in induced fit binding as compared to damaged DNA, **1** remarkably shares many of the binding interactions observed in the uracilated-DNA complex (Figure 6B and C). The DNA binding site of UNG2 is composed of a uracil recognition pocket flanked by a deep groove which is predominantly involved in accommodating the single strand of DNA that contains the extrahelical uracil. In the DNA complex, specific hydrogen bonding and aromatic stacking interactions with the extrahelical uracil involving Asn204, His268, Gln144 and Phe158 are observed (Figure 6B). In addition, UNG2 also makes important interactions via neutral and charged hydrogen bonds with the 5′ and 3′ phosphodiester groups of the deoxyuridine and the 3′ phosphodiester group of the 3′ adjacent nucleotide using the γ-hydroxyls of Ser169, Ser270 (data not shown) and Ser247, respectively (Figure 6B). The uracil base of **1** shares the uracil interactions seen with the uracilated-DNA complex, with the exception of the catalytically important short hydrogen bond between uracil O2 and His268 (16,27,28). The planar oxime linkage at the uracil side of the tether is observed to extend directly over the space that is occupied by the deoxyribose ring of deoxyuridine in the DNA complex, but then, the alkane linker sharply kinks such that the oxime linkage connecting to the benzylformate moiety nearly perfectly superimposes the path taken by the sugar phosphate backbone of the DNA 3′ to the deoxyuridine nucleotide (Figure 6C). This trajectory of the linker presents the carboxylate substituent of the benzyl ring such that it forms a charged tridentate hydrogen bond with the backbone amide groups of Ser247 and Tyr248 and the γ hydroxyl of Ser247 (Figure 6B). These interactions with the carboxylate group mimic those of the 3′ phosphodiester group of the nucleotide directly adjacent to deoxyuridine in the DNA complex (Figure 6C). In addition, the oxime oxygen on the uracil side of the tether accepts a hydrogen bond from the γ-hydroxyl of Ser169 thereby mimicking the interaction of the 5′-phosphate of dUrd in the DNA complex. Due to differences in induced fit binding between **1** and uracilated-DNA, the catalytic His268 is too far from uracil O2 to form the strong hydrogen bond seen in the DNA complex. Instead, His268 stacks over the benzyl ring of **1** to form a 3.6 Å π–π aromatic interaction (Figure 6B). Overall, **1** shares three of the four hydrogen bond interactions with the uracil base observed in the DNA structure and three of the five DNA backbone hydrogen bonds.

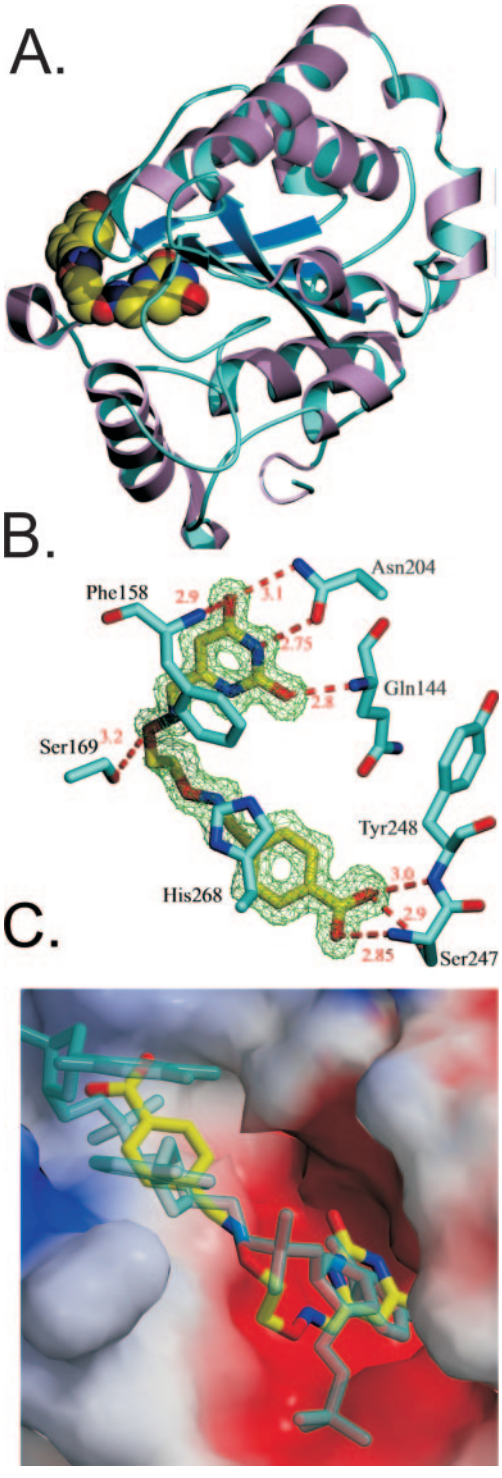

Figure 6. Interactions of **1** and damaged DNA with the active site of UNG2. (**A**) Global structure of inhibitor-UNG2 complex. (**B**) Discrete interactions of **1** with the active site of UNG2. The $2F_o–F_c$ electron density map is shown at a contour level of 1σ. (**C**) Overlay of **1** (gold) with the region of the damaged DNA strand (turquoise) containing uracil and the adjacent two 3′ nt (1EMH).

This structure also provides useful insights into the inhibition provided by compound **2**, as well as our previously

characterized tight-binding SFT compound (**3**, $K_i = 300$ nM) that contains a 2,4,5-trihydroxybenzyl substituent (17).

Manual docking studies suggest that the longer three carbon linker of **2** is used to extend its *m*-formate substituent such that it can serve as a DNA phosphate mimic as observed for the *p*-formate substituent of **1**. Our previously reported tight-binding SFT compound contained a *p*-hydroxyl substituent and also a linker that is one carbon longer than **1** suggesting that the *p*-hydroxyl is positioned to form neutral hydrogen bonding interactions in the same pocket occupied by the *p*-formate group of **1**. Apparently, the 20-fold higher affinity of this previous SFT ligand arises from favorable presentation of all of its hydroxyl substituents. It is interesting to note that the 215 member aldehyde SFT library contains only two carboxylate compounds, and both of these were detected as inhibitors in HTS when the correct linker length was employed (i.e. compounds **1** and **2**). Thus, binding elements possessing molecular features similar to the DNA substrate arise more frequently as inhibitors, suggesting that libraries enriched in such motifs might have higher hit rates.

The SFT approach may find general utility in targeting enzymes that recognize extrahelical bases. The flexible alkane tether appears to be an accommodating scaffold that allows favorable presentation of binding elements that are complementary to the DNA binding surface of the enzyme. More generally, substrate fragments (or weak binding ligands) that target enzyme active sites should make excellent starting places for rapid inhibitor development by this or other tethering approaches (29).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR online.

ACKNOWLEDGEMENTS

We thank M. Allaire and V. Stojanoff for help during the data collection at the National Synchrotron Light Source. This work was supported by NIH grants GM56834-11 (J.T.S) and GM066895 (L.M.A). D.J.K. was supported by the DOD Breast Cancer Research Program (DAMD17-03-1-1251). Funding to pay the Open Access publication charges for this article was provided by NIH/GMS.

Conflict of interest statement. None declared.

REFERENCES

1. Krokan,H.E., Drablos,F. and Slupphaug,G. (2002) Uracil in DNA—occurrence, consequences and repair. *Oncogene*, **21**, 8935–8948.
2. Petersen-Mahrt,S.K., Harris,R.S. and Neuberger,M.S. (2002) AID mutates *E.coli* suggesting a DNA deamination mechanism for antibody diversification. *Nature*, **418**, 99–103.
3. Lindahl,T. (1974) An n-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. *Proc. Natl Acad. Sci. USA*, **71**, 3649–3653.
4. Lindahl,T. and Nyberg,B. (1974) Heat-induced deamination of cytosine residues in deoxyribonucleic acid. *Biochemistry*, **13**, 3405–3410.
5. Nilsen,H., Haushalter,K.A., Robins,P., Barnes,D.E., Verdine,G.L. and Lindahl,T. (2001) Excision of deaminated cytosine from the vertebrate genome: role of the smug1 uracil-DNA glycosylase. *EMBO J.*, **20**, 4278–4286.
6. Nilsen,H., Rosewell,I., Robins,P., Skjelbred,C.F., Andersen,S., Slupphaug,G., Daly,G., Krokan,H.E., Lindahl,T. and Barnes,D.E. (2000) Uracil-DNA glycosylase (ung)-deficient mice reveal a primary role of the enzyme during DNA replication. *Mol. Cell*, **5**, 1059–1065.
7. Stivers,J.T. and Jiang,Y.L. (2003) A mechanistic perspective on the chemistry of DNA repair glycosylases. *Chem. Rev.*, **103**, 2729–2759.
8. Seiple,L., Jaruga,P., Dizdaroglu,M. and Stivers,J.T. (2006) Linking uracil base excision repair and 5-fluorouracil toxicity in yeast. *Nucleic Acids Res.*, **34**, 140–151.
9. Rada,C., Williams,G.T., Nilsen,H., Barnes,D.E., Lindahl,T. and Neuberger,M.S. (2002) Immunoglobulin isotype switching is inhibited and somatic hypermutation perturbed in ung-deficient mice. *Curr. Biol.*, **12**, 1748–1755.
10. Di Noia,J. and Neuberger,M.S. (2002) Altering the pathway of immunoglobulin hypermutation by inhibiting uracil-DNA glycosylase. *Nature*, **419**, 43–48.
11. Nilsen,H., Stamp,G., Andersen,S., Hrivnak,G., Krokan,H.E., Lindahl,T. and Barnes,D.E. (2003) Gene-targeted mice lacking the ung uracil-DNA glycosylase develop B-cell lymphomas. *Oncogene*, **22**, 5381–5386.
12. Priet,S., Sire,J. and Querat,G. (2006) Uracils as a cellular weapon against viruses and mechanisms of viral escape. *Curr. HIV Res.*, **4**, 31–42.
13. Malet-Martino,M. and Martino,R. (2002) Clinical studies of three oral prodrugs of 5-fluorouracil (capecitabine, UFT, S-1): A review. *Oncologist*, **7**, 288–323.
14. Stivers,J.T. and Drohat,A.C. (2001) Uracil DNA glycosylase: Insights from a master catalyst. *Arch. Biochem. Biophys.*, **396**, 1–9.
15. Parikh,S.S., Walcher,G., Jones,G.D., Slupphaug,G., Krokan,H.E., Blackburn,G.M. and Tainer,J.A. (2000) Uracil-DNA glycosylase-DNA substrate and product structures: conformational strain promotes catalytic efficiency by coupled stereoelectronic effects. *Proc. Natl Acad. Sci. USA*, **97**, 5083–5088.
16. Drohat,A.C. and Stivers,J.T. (2000) *Escherichia coli* uracil DNA glycosylase: NMR characterization of the short hydrogen bond from His187 to uracil O2. *Biochemistry*, **39**, 11865–11875.
17. Jiang,Y.L., Krosky,D.J., Seiple,L. and Stivers,J.T. (2005) Uracil-directed ligand tethering: an efficient strategy for uracil DNA glycosylase (ung) inhibitor development. *J. Am. Chem. Soc.*, **127**, 17412–17420.
18. Maly,D.J., Choong,I.C. and Ellman,J.A. (2000) Combinatorial target-guided ligand assembly: identification of potent subtype-selective c-src inhibitors. *Proc. Natl Acad. Sci. USA*, **97**, 2419–2424.
19. Halgunset,J., Lamvik,T. and Espevik,T. (1988) Butyrate effects on growth, morphology, and fibronectin production in PC-3 prostatic carcinoma cells. *Prostate*, **12**, 65–77.
20. Slupphaug,G., Eftedal,I., Kavli,B., Bharati,S., Helle,N.M., Haug,T., Levine,D.W. and Krokan,H.E. (1995) Properties of a recombinant human uracil-DNA glycosylase from the ung gene and evidence that ung encodes the major uracil-DNA glycosylase. *Biochemistry*, **34**, 128–138.
21. Otwinoski,Z. and Minor,W. (1997) Processing of x-ray diffraction data in oscillation mode. *Methods Enzymol.*, **276**, 307–325.
22. Ccp4. (1994) The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr.*, **D50**, 760–763.
23. Kavli,B., Sundheim,O., Akbari,M., Otterlei,M., Nilsen,H., Skorpen,F., Aas,P.A., Hagen,L., Krokan,H.E. and Slupphaug,G. (2002) Hung2 is the major repair enzyme for removal of uracil from U:A matches, U:G mismatches, and U in single-stranded DNA, with hsmug1 as a broad specificity backup. *J. Biol. Chem.*, **277**, 39926–39936.
24. Kurian,J.R., Bajad,S.U., Miller,J.L., Chin,N.A. and Trepanier,L.A. (2004) Nadh cytochrome b5 reductase and cytochrome b5 catalyze the microsomal reduction of xenobiotic hydroxylamines and amidoximes in humans. *J. Pharmacol. Exp. Ther.*, **311**, 1171–1178.
25. Heberling,S., Girreser,U., Wolf,S. and Clement,B. (2006) Oxygen-insensitive enzymatic reduction of oximes to imines. *Biochem. Pharmacol.*, **71**, 354–365.
26. Werner,R.M., Jiang,Y.L., Gordley,R.G., Jagadeesh,G.J., Ladner,J.E., Xiao,G., Tordova,M., Gilliland,G.L. and Stivers,J.T. (2000) Stressing-out DNA? The contribution of serine-phosphodiester interactions in catalysis by uracil DNA glycosylase *Biochemistry*, **39**, 12585–12594.

27. Drohat,A.C., Xiao,G., Tordova,M., Jagadeesh,J., Pankiewicz,K.W., Watanabe,K.A., Gilliland,G.L. and Stivers,J.T. (1999) Hetronuclear NMR and crystallographic studies of wild-type and H187Q *Escherichia coli* uracil DNA glycosylase: electrophilic catalysis of uracil expulsion by a neutral histidine 187. *Biochemistry*, **38**, 11876–11886.
28. Drohat,A.C. and Stivers,J.T. (2000) NMR evidence for an unusually low N1 pK_a for uracil bound to uracil DNA glycosylase: implications for catalysis. *J. Am. Chem. Soc*., 1840–1841.
29. Erlanson,D.A., Wells,J.A. and Braisted,A.C. (2004) Tethering: fragment-based drug discovery. *Annu. Rev. Biophys. Biomol. Struct*., **33**, 199–223.