Article:
Michel, M., Visnes, T., Homan, E.J. et al. (12 more authors) (2019) Computational and experimental druggability assessment of human DNA glycosylases. ACS Omega, 4 (7). pp. 11642-11656. ISSN 2470-1343

https://doi.org/10.1021/acsomega.9b00162

ACS OMEGA

Article

Cite This: ACS Omega 2019, 4, 11642−11656

http://pubs.acs.org/journal/acsodf

Computational and Experimental Druggability Assessment of Human DNA Glycosylases

Maurice Michel,*[,†,▽] Torkild Visnes,[†,‡,▽] Evert J. Homan,[†] Brinton Seashore-Ludlow,[§] Mattias Hedenström,[∥] Elisée Wiita,[†] Karl Vallin,[†] Cynthia B. J. Paulin,[†] Jiaxi Zhang,[⊥] Olov Wallner,[†] Martin Scobie,[†] Andreas Schmidt,[⊥] Annika Jenmalm-Jensen,[§] Ulrika Warpman Berglund,[†] and Thomas Helleday*[,†,#]

[†]Science for Life Laboratory, Department of Oncology-Pathology, Karolinska Institutet, S-171 76 Stockholm, Sweden

[‡]Department of Biotechnology and Nanomedicine, SINTEF Industry, N-7465 Trondheim, Norway

[§]Chemical Biology Consortium Sweden, Science for Life Laboratory, Division of Translational Medicine and Chemical Biology, Department of Medical Biochemistry and Biophysics, Karolinska Institutet, S-171 21 Stockholm, Sweden

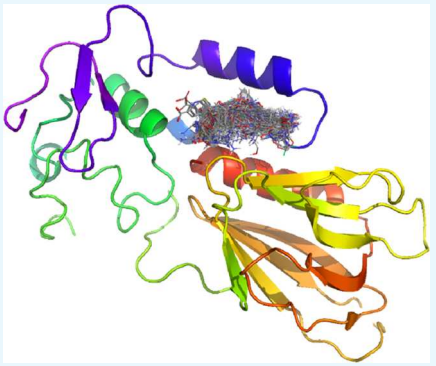
[∥]Department of Chemistry, Umeå University, S-901 87 Umeå, Sweden

[⊥]Institute of Organic Chemistry, Clausthal University of Technology, Leibnizstrasse 6, D-38678 Clausthal-Zellerfeld, Germany

[#]Sheffield Cancer Centre, Department of Oncology and Metabolism, University of Sheffield, S10 2RX Sheffield, U.K.

Supporting Information

ABSTRACT: Due to a polar or even charged binding interface, DNA-binding proteins are considered extraordinarily difficult targets for development of small-molecule ligands and only a handful of proteins have been targeted successfully to date. Recently, however, it has been shown that development of selective and efficient inhibitors of 8-oxoguanine DNA glycosylase is possible. Here, we describe the initial druggability assessment of DNA glycosylases in a computational setting and experimentally investigate several methods to target endonuclease VIII-like 1 (NEIL1) with small-molecule inhibitors. We find that DNA glycosylases exhibit good predicted druggability in both DNA-bound and -unbound states. Furthermore, we find catalytic sites to be highly flexible, allowing for a range of interactions and binding partners. One flexible catalytic site was rationalized for NEIL1 and further investigated experimentally using both a biochemical assay in the presence of DNA and a thermal shift assay in the absence of DNA.



INTRODUCTION

The concept of inflicting damage to DNA in tumor cells to treat cancer, e.g., by radiation therapy or chemotherapeutics such as antimetabolites and DNA intercalators, has formed the foundation of modern clinical oncology.[1,2] The success of these first-line cancer treatments have prompted increased attention toward enzymes that repair damaged bases and to the development of corresponding small-molecule inhibitors for cancer therapy. We and others have shown that Mut T homolog 1 (MTH1) inhibition leads to incorporation of oxidized bases such as 7,8-dihydro-8-oxoguanine (8-oxoG) into DNA and selectively kills cancer cells.[3−7] Despite the unclear underlying biology of MTH1 inhibition,[8] it is evident that cancer cells depend on protective repair pathways to tolerate increased oxidative stress. Therefore, we argue that further inhibition of the main DNA repair pathways for repair of oxidized nucleobases, in particular the base excision repair proteins 8-oxoguanine DNA glycosylase (OGG1), Mut Y homolog (MUTYH), or NEIL1, could lead to powerful combination therapies.[9−15] OGG1, the DNA glycosylase responsible for repairing the bulk of 8-oxoG in mammals,[16,17] has recently been validated preclinically as a drug target by us and others, proving druggable with selective small molecules.[18,19] The importance of managing 8-oxoG levels is also facilitated by MUTYH, which removes adenine misincorporated opposite to 8-oxoG.[20] This initiates recycling of the damaged DNA strand back to OGG1, which otherwise fails to recognize 8-oxoG unless it is base-paired with cytosine. NEIL1, on the other hand, has a unique substrate range, removing all products formed from further oxidation and fragmentation of 8-oxoG, but also thymine glycol (Tg), oxidized cytosine and uracil.[12,21−24] Mice devoid of these DNA glycosylases are viable and grow old, suggesting that potential inhibitors would show little on-target toxicity in a rodent model.[25,26]

DNA glycosylases exist in DNA-bound and -unbound conformations.[19,27−32] It is of interest whether a DNA-bound or -unbound state facilitates or restricts the binding of small molecules. Thus, one major challenge is to be able to target one

Received: January 17, 2019
Accepted: May 21, 2019
Published: July 5, 2019

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

population of a DNA glycosylase with a small molecule, given that this is a requirement for conveying a certain phenotype.[19] Computational binding-site prediction, for example, is a suitable method to investigate chemotype preference of DNA glycosylases using available crystal structures of single isolated protein species. However, literature concerning druggability of any DNA glycosylase is nonexistent and reported findings are only applicable in the broadest sense by transferring knowledge from glycosylases and RNA-, DNA-, nucleotide-, and carbohydrate-binding proteins.[3,33−36] Additionally, these previous studies based on crystal structures have considered the relevant proteins to be rigid and not flexible, a scenario that is not applicable to DNA glycosylases.

Druggability is defined as the ability of a protein to specifically bind rule-of-five-compliant small molecules with high affinity.[37−39] A high druggability score and the induction of a therapeutic effect by small-molecule binding in a living system are characteristics of a good drug target. Several computational and empirical methods to assess protein druggability have been reported over the past years.[40−42] Computational druggability predictions are less time-consuming and relatively cheap compared to experimental methods. Given the availability of structural information, i.e., high-resolution crystallographic data, they allow for the rapid evaluation of target suitability for a drug discovery campaign. A number of computational methods predicting protein-binding sites and their druggability are available,[33,43−46] spanning the entire spectrum from geometric to energy-based and from rigid proteins to systems allowing for high flexibility.

High-throughput screening (HTS) of large druglike compound libraries has yielded a number of hits for NEIL1 and OGG1 with micromolar (μM) potency.[18,47,48] However, target screening using rule-of-three-compliant fragment libraries may be more productive, since it can assess the target's druggability. Furthermore, fragment screening also covers a larger chemical space and typically yields hits with higher ligand efficiencies, which are often more amenable for further lead generation than μM druglike hits.[35,49−51] Techniques commonly used for fragment screening are biophysical methods such as surface plasmon resonance (SPR), nuclear magnetic resonance (NMR), differential scanning fluorimetry (DSF), and X-ray crystallography.[52]

Here, we performed a computational binding-site prediction and druggability assessment of available high-resolution structures of human DNA glycosylases using different algorithms and experimentally assessed the computationally derived druggability scores for NEIL1. We show structures identified in the biochemical HTS, assess their target engagement in DSF, and importantly, suggest a broadly applicable adaptation of DSF for DNA-binding proteins.

RESULTS AND DISCUSSION

Sequence Alignment Reveals Low Sequence Conservation of DNA Glycosylases. Initially, we compiled a list of all available high-resolution crystal structures of human DNA glycosylases (Table 1) and assessed consensus protein sequence by alignment of FASTA at Clustal Omega at EMBL.[3,27−30,53−64] Sequence similarity was extraordinary low, with an average of 15.5% (Figure 1B).[65] This is in accordance with earlier reports that assort the low sequence similarity of especially the uracil-binding glycosylases to the major role of uracil flipping.[66] This seems to be due to a conformational strain caused by the enzyme on uracil-containing DNA, which is much bigger compared to the catalytic action of individual polar residues. Thus, it was observed that despite their low sequence similarity, bacterial uracil−DNA glycosylase and mismatch-specific uracil−DNA glycosylase exhibit an almost identical global folding. Even so, two different groups can be distinguished by the phylogenetic tree (Figure 1A,B): NEIL1, NEIL2, and NEIL3 with a similarity between 19.6 and 26.1%, and MPG, UNG, SMUG1, MUTYH, NTH1, and OGG1 with a similarity between 11.9 and 24.1%. Furthermore, the root-mean-square deviation (RMSD) of atomic (Cα) positions for the pairs available for NEIL1 (0.84 Å), UNG (1.05 Å), MBD4 (0.88 Å), and OGG1 (0.85 Å) are small and suggest that possible protein flexibility is either caused by conformational changes of the amino acid side chains or by existence of only a few higher-valued amino acids.

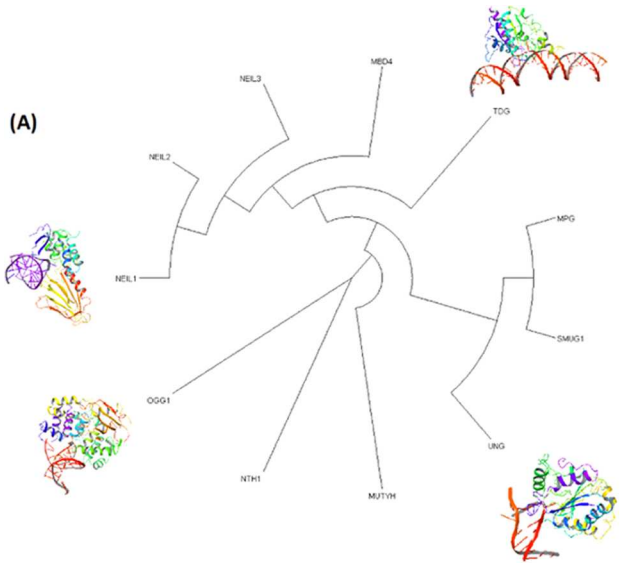
Table 1. List of Analyzed Crystal Structures of Human DNA Glycosylases[a]

protein name	PDB code	w/DNA	w/o DNA	apo	reference
MBD4	4E9E			X	Moréra et al.[56]
MBD4	4OFA	X			Ouzon-Shubeita et al.[57]
MBD4	4OFA		X		Ouzon-Shubeita et al.[57]
MPG	1EWN	X			Lau et al.[64]
MPG	1EWN		X		Lau et al.[64]
MUTYH	3N5N			X	Luncsford et al.[62]
NEIL1	1TDH			X	Doublié et al.[30]
NEIL1	5ITY	X			Zhu et al.[29]
NEIL1	5ITY		X		Zhu et al.[29]
NEIL3	3W0F			X	Liu et al.[61]
OGG1	5AN4			X	Zander et al.[27]
OGG1	2XHI	X			Dalhus et al.[28]
OGG1	2XHI		X		Dalhus et al.[28]
TDG	5HF7	X			Coey et al.[60]
TDG	5HF7		X		Coey et al.[60]
UNG	2HXM			X*	Krosky et al.[59]
UNG	1EMH	X			Parikh et al.[58]
UNG	1EMH		X		Parikh et al.[58]
MTH1	3Q93			X	Tresauges et al.[69]
PTP1B	2HNP			X	Barford et al.[76]

[a]MTH1 and PTP1B were included as reference proteins; * ligand removed.

Computational Binding-Site Analysis by DogSite Predicts Good Druggability. We then predicted binding pockets and assessed druggability of the high-resolution crystal structures. Generally, hydrophilic sites are considered less druggable than more hydrophobic sites. This effect should in theory yield low scores for DNA-binding regions of the proteins. We identified between 1 and 11 potential binding sites for each DNA glycosylase. Except for DNA-bound MBD4 (4OFA),[57] at least two sites were found to be druggable by the definition of DogSite (drug score > 0.5). For all structures except apo MBD4 (4E9E),[56] DNA-bound NEIL1 (5ITY),[29] apo OGG1 (5AN4),[27] and DNA-bound TDG (5HF7),[60] the active site with catalytic amino acid residues was identified as the top-ranking binding pocket. Apo NEIL1 (1TDH)[30] was the only structure where two distinct binding pockets were identified close to amino acid residues involved in the catalytically active site. Finally, closely situated binding sites in apo, DNA-bound, and DNA-stripped structures scored differently and, in addition, were found to have different size and altered amino acid composition.

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656



(B)

	MBD4	MPG	UNG	SMUG1	MutYH	NTH1	OGG1	TDG	NEIL3	NEIL1	NEIL2
MBD4	100.0										
MPG	14.4	100.0									
UNG	12.4	20.2	100.0								
SMUG1	12.5	21.1	20.3	100.0							
MutYH	15.6	13.5	16.2	20.8	100.0						
NTH1	15.1	13.6	20.6	22.8	21.2	100.0					
OGG1	14.2	19.7	20.2	11.9	23.5	24.1	100.0				
TDG	14.4	8.6	7.9	6.6	10.7	8.5	14.4	100.0			
NEIL3	15.0	14.6	10.4	8.7	17.1	11.4	11.5	15.7	100.0		
NEIL1	14.0	15.0	10.0	14.2	12.9	13.9	16.1	14.2	19.6	100.0	
NEIL2	13.1	15.4	14.6	10.9	15.2	16.4	18.4	20.1	24.2	26.1	100.0

Figure 1. Phylogenetic tree (A) and sequence alignment (B) of human DNA glycosylases. Despite recognizing damaged DNA by flipping out the corresponding base, human DNA glycosylases have a surprisingly low sequence similarity even among those classified in the same protein family such as the endonuclease VIII-like members.

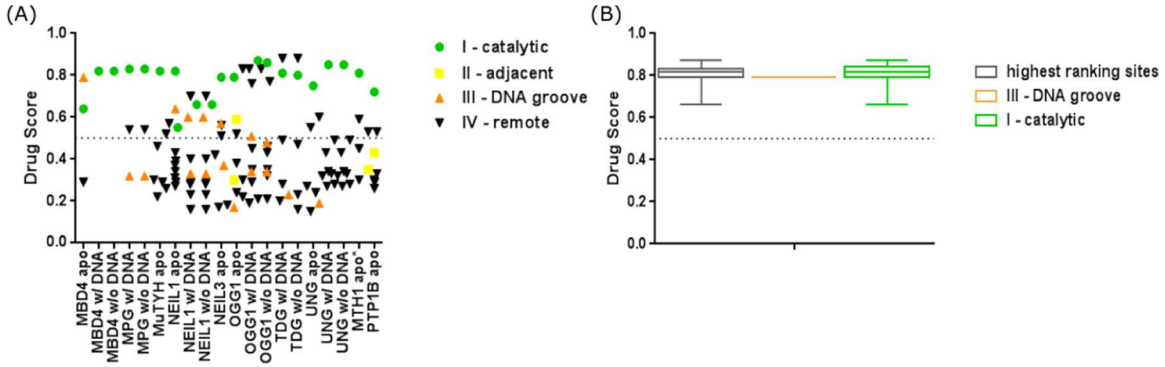


Figure 2. (A) DogSite was used to predict binding sites in available high-resolution crystal structures of human DNA glycosylases. Each predicted binding site was further assessed for its DrugScore, i.e., computational derived druggability score. Values between 0.5 and 1.0 define a pocket as druggable, of which values closer to 1.0 are considered the best. Identified binding sites of DNA glycosylases can be sorted into four categories: I catalytic site (green); II catalytic adjacent site (yellow); III part of DNA groove (orange); IV other remote sites (black). Except for apo MBD4, DNA-bound NEIL1, apo OGG1, and DNA-bound TDG, the algorithm identified the catalytic sites as the highest ranking pocket. Surprisingly, most high scoring sites reached a DrugScore as high as the binding site of MTH1, a well-characterized and druggable protein. (B) To induce a therapeutic effect, perturbation of DNA binding is a sufficient mode of action for a prospective small-molecule drug. Therefore, for each protein, only the single highest ranking DrugScore of category I, II, or III was used to calculate a mean druggability score for human DNA glycosylase. With an average DrugScore of 0.80, DNA glycosylases identify as favorably druggable targets (MTH1 0.81, PTP1B 0.72).

Although therapeutically targeting DNA glycosylases with small molecules requires inhibition of DNA−protein complex formation, this may be achieved without direct ligand binding to the active site. Despite possible involvement in protein regulation or allosteric sites, inhibition of remote sites is not generally considered a requisite for eliciting a therapeutic effect.[67] We therefore assigned the predicted binding sites to one of four categories: Active site with catalytic amino acid residue I, active-site adjacent pocket II, DNA-binding groove III, and remote binding sites IV (Figure 2A). For each protein, the single, top-ranking pocket for the categories I, II, and III was used to calculate a mean druggability score. With a druggability score of 0.80, DNA glycosylases were identified as favorably druggable. The druggable reference protein MTH1 (3Q93)[3,7,68,69] was found to score slightly higher with 0.81. The protein tyrosine phosphatase 1B (PTP1B) was included as a negative control. PTP1B is widely accepted as a potential therapeutic target for type 2 diabetes and obesity,[70−72] and for this reason, numerous potent and structurally diverse PTP1B inhibitors have been discovered over the last two decades.[73−75] Most of the inhibitors incorporate phosphotyrosine (pTyr) mimetics and bind to the enzyme's catalytic site. However, the highly conserved and polar nature of the PTP1B catalytic site in combination with a catalytically active cysteine renders development of PTP1B inhibitors with good selectivity and permeability a challenging task. So far, only three small-molecule PTP1B inhibitors, including allosteric drugs, have progressed into clinical trials.[75] The catalytic site of PTP1B (2HNP)[76] scored 0.72 by combination of two small sites, which separately scored 0.41 and 0.56. Interestingly, OGG1 (2XHI),[28] TDG (5HF7),[60] and UNG (1EMH)[58] show especially high drug scores of >0.85 for the catalytic site, while 0.66 for apo MBD4 (4E9E)[56] was the only one indicating a challenging drug discovery campaign (Figure 2B). In contrast to DNA-, RNA-,

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

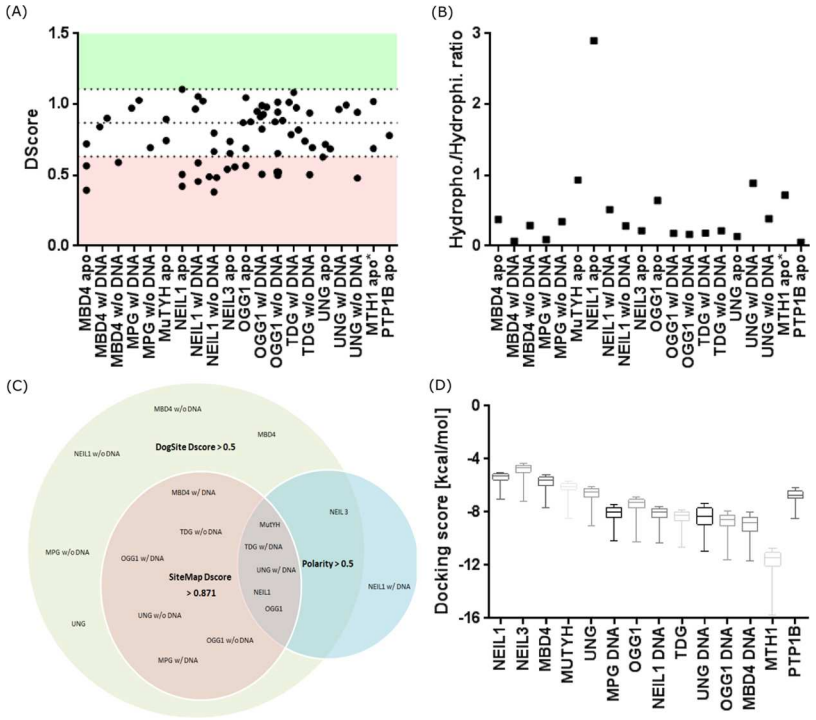


Figure 3. (A) As a second algorithm, SiteMap of Maestro from Schrödinger was used to predict binding sites and calculate the corresponding Dscores. Values below 0.871 reflect a challenging nature of the target, while proteins reaching values larger than 1.108 are considered excellent drug targets. With the highest ranking pocket of category I, II, or III, DNA glycosylases yield an average DScore of 0.89 (MTH1 1.020, PTP1B 0.770) and are predicted to be druggable targets. (B) Polarity of active sites are a main concern when targeting DNA-binding proteins. Indeed, DNA glycosylases bound to DNA tend to exhibit higher hydrophilicity than apo structures. These values, however, do not reflect a general trend that excludes druggability of DNA glycosylases judging by pocket polarity. (C) Venn diagram illustrating that all proteins considered druggable by SiteMap are also considered druggable by DogSite. A large percentage of these show sufficient polarity or hydrophobic/hydrophilic ratio to be targeted by classic rule-of-five-compliant small molecules. (D) Docking scores generated during in silico screening of a subset of the ZINC library to the highest ranking site are an additional measure to estimate the druggability of a target. Here, the top-1000 ranked fragments against every DNA glycosylase cover a broad spectrum with docking scores as low as those reached against MTH1 and higher scores for the hydrophobic site of apo NEIL1 and beyond PTP1B.

and carbohydrate-binding proteins and based on remarkably high DogSite generated druggability scores, this indicates that DNA glycosylases are indeed druggable.[33−36]

Binding-Site Assessment with SiteMap Identifies Dependency of Pocket Polarity on the Presence of DNA. Schrödingers SiteMap is commonly used to assess protein surfaces for binding pockets and their druggability, i.e., Dscore. The Dscore favors proteins with more hydrophobic binding pockets, and proteins reaching Dscores higher than 1.108 are considered excellent drug targets, while proteins with values up to 0.871 are considered difficult.[77,78] For comparison, the largest identified site for the undruggable PTP1B (2HNP)[76] showed a Dscore of 0.77, while MTH1 (3Q93)[69] scored 1.02. Between one and eight sites were found for DNA glycosylases by SiteMap, with Dscores ranging from nondruggable to druggable (Figure 3A). The single highest scoring sites of categories I, II, and III reached a mean Dscore of 0.89, scoring below the Dscore of MTH1 (3Q93)[69] of 1.02, but clearly above PTP1B (2HNP).[76] The range of Dscores between 0.59 and 1.11 on the other hand reflects a rather distributed picture of the different target proteins. Although slightly lower than in DogSite, these results confirm the finding of good druggability of human DNA glycosylases.

One important characteristic of a druggable binding pocket is a favorable hydrophobic/hydrophilic balance. This is the main point of concern regarding druggability of DNA-binding proteins, since high-affinity DNA binding requires a network of polar amino acid residues. Thus, it was of interest to determine whether the polarity of the highest ranking sites allows for the binding of druglike molecules, which are usually more hydrophobic than DNA. When investigated for these parameters using the incorporated function in SiteMap, the single highest scoring sites of human DNA glycosylases score on average 0.49, respectively, 0.69 for MTH1 (3Q93)[69] and 0.05 for PTP1B (2HNP),[76] with a range of 0.07−2.90 (Figure 3B). Apo structures tend to be higher ranking, as only UNG with bound DNA (1EMH)[58] reaches hydrophobicity levels higher than MTH1 (3Q93).[69] As an extreme example, apo NEIL1 (1TDH)[30] scored a hydrophobic/hydrophilic ratio of 2.90. In contrast, the corresponding DNA-bound proteins score substantially lower (0.16−0.38), suggesting major conformational changes upon DNA binding. In total, only MUTYH (3N5N),[62] NEIL1 (1TDH),[30] NEIL1 w/DNA (5ITY),[29] OGG1 (5AN4),[27] TDG w/DNA (5HF7),[60] and UNG w/DNA (1EMH)[58] are feasible drug targets judging by a pocket polarity over 0.5 (Figure 3B).

In Silico Docking of ZINC Fragments to Human DNA Glycosylases. As an additional method to investigate the obtained druggability scores and pocket polarities, we performed virtual screens of a subset of the open-access ZINC database against the highest scoring binding site of each DNA glycosylase, as identified by SiteMap.[79,80] An obvious advantage of this screening campaign is that top-ranked fragments could provide potential starting points for hit expansion and lead

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

generation. We hypothesized that the docking score profiles of the top-ranked 1000 fragments would reflect the druggability of the respective sites, i.e., protein pockets with good druggability scores should yield lower average fragments scores. By applying a cascade of docking steps with increasing scoring precision, the top-ranked 1000 fragments against each DNA glycosylase were retained and the median Glide XP docking scores for each of the top-ranked 1000 fragment sets were calculated. Figure 3C summarizes the docking score profiles and shows the scores of the 1000 top-ranked fragments against human DNA glycosylases. The fragment docking profile of MTH1 (3Q93)[69] clearly stands out with the top-1000 fragments scoring between −10.7 and −15.7 kcal/mol, and median docking score of −11.4 kcal/mol, supporting the in silico druggability assessments and experimental data of MTH1 being highly druggable. In support of this are results obtained from our in-house fragment screen against purified MTH1 using DSF, which yielded a hit rate of 10% (unpublished results) and the high DSF fragment screen hit rate of 23% obtained by Rahm et al.[81] Furthermore, the enzymatic HTS resulting in the TH588 series of MTH1 inhibitors showed a hit rate of 4%.[3] In contrast, PTB1B (2HNP),[76] known to be a challenging target for small-molecule intervention at the catalytic site, displayed a median docking score of −6.7 kcal/mol, with scores for the top-1000 fragments ranging from −6.2 to −8.5 kcal/mol. For the DNA glycosylases, median fragment docking scores are found between −4.8 and −9.0 kcal/mol, and with an average median docking score of −7.1 kcal/mol. Importantly, a clear difference in docking score profiles for apo and DNA-bound structures of the same protein, e.g., NEIL1, MBD4, OGG1, and UNG, can be observed. This effect can mostly be attributed to the slightly different location of the binding sites identified by SiteMap and used for docking in this experiment. Interestingly, DNA-bound structures of DNA glycosylases tend to perform better than apo structures, as the five poorest scoring proteins are apo structures. Apo NEIL1 (1TDH)[30] showed a poor docking score profile with a median docking score of −5.3 kcal/mol. This stands in sharp contrast to the earlier observation of its high druggability scores obtained in all algorithms. The high Dscore (1.11) of the largest pocket, adjacent to the catalytic site in the DNA-bound structure, is primarily caused by the highly favorable hydrophobic/hydrophilic balance of 2.90. The ZINC fragment subset, however, was tailored toward fragments with good H-bond formation possibilities. This may explain the suboptimal complementarity of a large proportion of fragments with this hydrophobic site.

Arg242 Flexibility Alters NEIL1 Binding-Site Polarity. In an attempt to structurally rationalize the findings from above for NEIL1, we inspected both the highest ranking binding sites in apo (1TDH)[30] and DNA-bound (5ITY)[29] forms. A major difference can be observed in close proximity, as inward movement of a flexible loop brings the terminal guanidine group of Arg242 in position to coordinate the tautomerized lactim moiety of the former lactam thymine glycol substrate in the DNA-bound structure, allowing for energetically favored stabilization (Figure 4).[29] This conformation is absent in the apo structure where Arg242 points outward, rendering this particular region a hydrophobic site. This observation is further supported by the presence of three glycine residues in the direct vicinity of Arg242 (motif GGRGY), as glycine in higher numbers are known to enhance flexibility of protein active sites.[82,83] Taken together, the microenvironment of Arg242 suggests possible application in inhibitor design, where prevention of Arg242 inward movement allows for the required

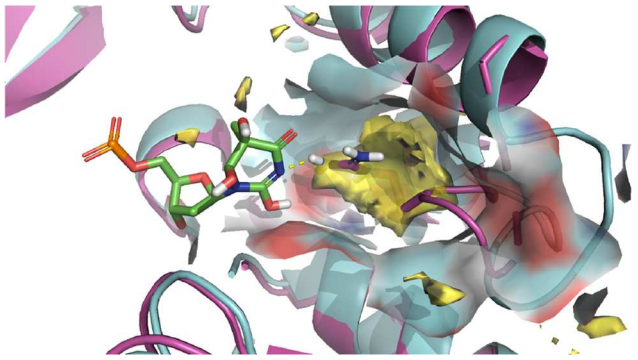


Figure 4. Superimposed structures of apo NEIL1 (turquoise, 1THD) and DNA-bound NEIL1 (purple, 5ITY) close to the catalytic pocket show little conformational flexibility of the peptide backbone (RMSD 0.84 Å). However, during DNA-binding events, the hydrophobic patch (yellow) in the apo state is occupied by the flexible (motif GGRGY) and polar guanidine side chain of Arg242. This inward directed movement enables coordination of the lactim group of thymine glycol and highlights DNA binding by NEIL1 as a process under allosteric control in a conformationally active pocket.

hydrophobicity to target NEIL1 with small molecules. Considering this importance of Arg242 and the large substrate specificity of NEIL1, we decided to perform a binding site prediction allowing for amino acid flexibility in DNA glycosylases.

Catalytic Sites Are Conformationally Flexible. The process of predicting binding sites and their druggability with DogSite and SiteMap is solely based on the proteins' crystal structure and thus cannot compensate for the conformational flexibility, which DNA glycosylases in general and NEIL1 in particular seem to require upon DNA binding. CryptoSite is a web-based server, which can detect holo sites in apo structures of proteins by interrogating flexibility of single amino acid residues.[44] A cryptic value above 0.10 identifies a position of a potential cryptic network, i.e., cryptic site. The structures of human DNA glycosylases yield an average of 0.30 with values from 0.12 to 0.46 for their respective catalytically active sites (Figure 5A, left). Furthermore, an extended network of amino acids around these pockets shows the same tendency of elevated cryptic values (Figure 5A, right). On average, 39 residues are involved in an anticipated movement, ranging from 15 for MPG devoid of DNA (1EWN)[64] to 69 residues for MUTYH (3N5N).[62] This confirms the good observed druggability of the space surrounding the catalytic sites in DNA glycosylases from above and indicates a concerted movement of a three-dimensional amino acid network during DNA binding. These findings, taken together with the results obtained with DogSite and SiteMap indicate a rather good druggability of human DNA glycosylases and, in addition, suggest DNA binding to occur in conformationally active sites (Figure 5B). Based on our DogSite and SiteMap results, these often highest scoring pockets may therefore be targeted with small molecules that initiate a number of conformational changes.[84,85] Earlier investigations of bacterial formamidopyrimidine DNA glycosylase (Fpg) suggest that syn- and anti-orientation of 8-oxo-guanine nucleobase and furanose are not discriminated during base recognition in a living system. This observation and our recent investigation of inhibitor binding to mOGG1 (6G3Y)[19] confirm the obtained inducible flexibility in the present study and stand in sharp contrast to earlier reports based on rigid protein conformation.[19,33−36] The contrast in druggability scores provided by the pair apo NEIL1

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

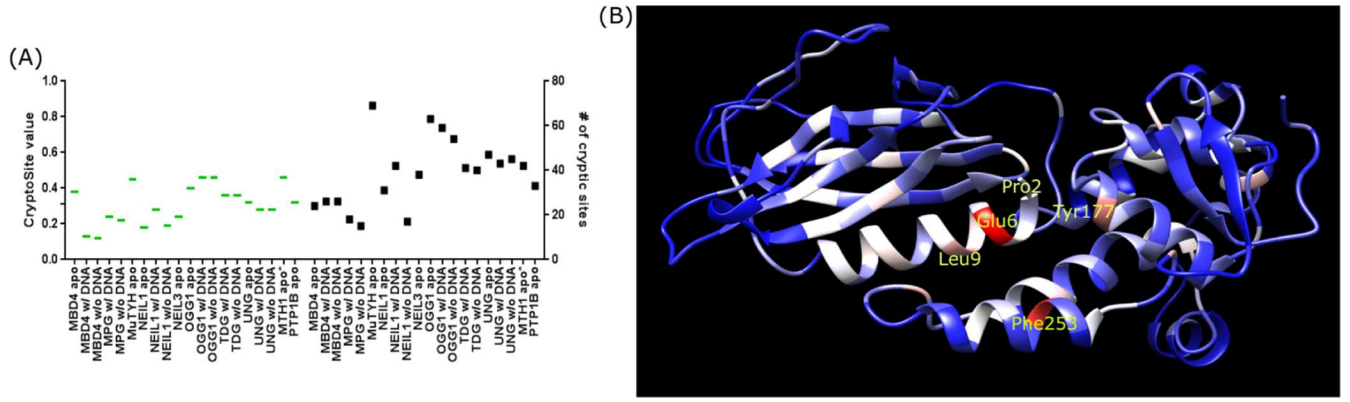


Figure 5. (A) Druggability assessment by CryptoSite allows for interrogation of single amino acid flexibility that in the form of larger clusters in close proximity causes formation of allosteric sites. Each amino acid is considered to be a cryptic site when reaching CryptoSite values above 0.10. Most identified cryptic sites in human DNA glycosylases are positioned close to the catalytic site (values in green) or pocket and form a comprehensive three-dimensional network with one another (number of amino acids participating in black), supporting the hypothesis of allosteric sites controlling DNA binding. (B) Illustration of identified highest ranking cryptic sites in apo NEIL1 (1THD; increasing CryptoSite value from blue to red) confirms the sites to be part of the catalytically active pocket.

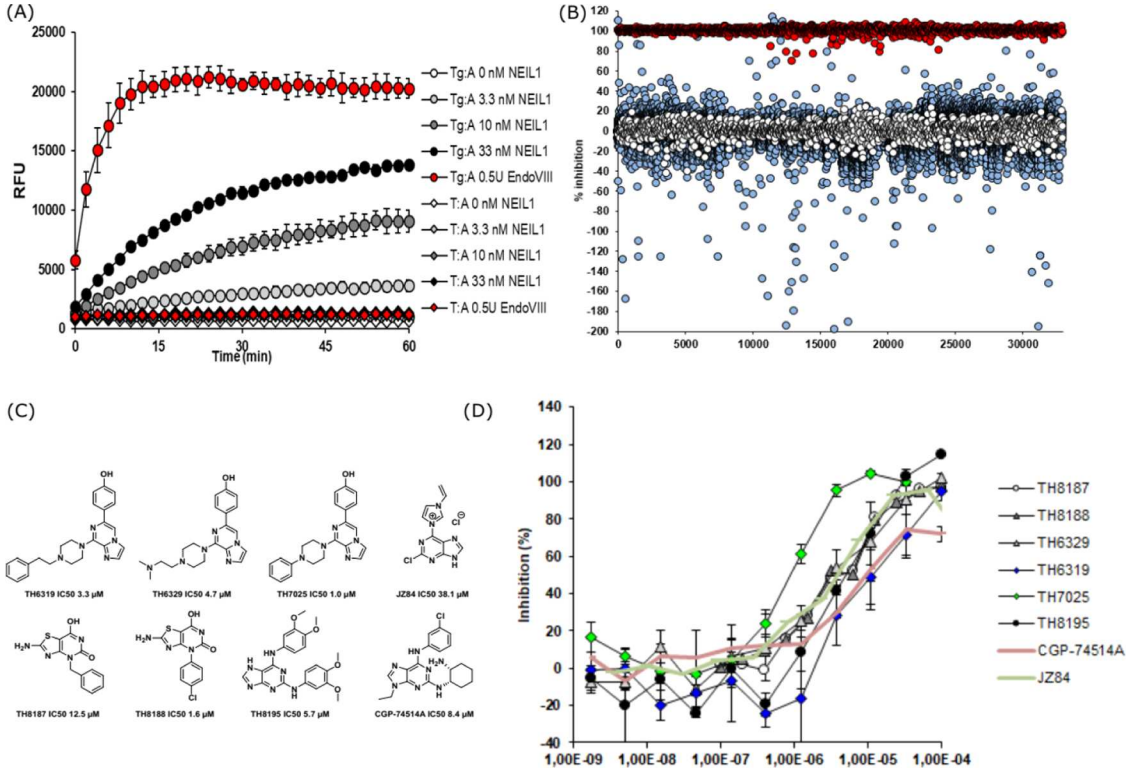


Figure 6. (A) Fluorigenic biochemical assay. An oligonucleotide (10 nM) containing thymine glycol or canonical thymine opposite adenine (Tg/A and T/A, respectively) was incubated with the indicated concentations of NEIL1, or 0.5 units of the *E.coli* ortholog endonuclease VIII (EndoVIII). Enzymatic treatment caused a concentration-dependent increase in fluorescence for Tg/A, but not T/A substrate. (B) High-throughput screen for NEIL1 inhibitors; 20 nM Tg/A substrate was incubated with 10 nM NEIL1, in the presence of 10 μM test compounds (light blue circles), 0.1% dimethyl sulfoxide (DMSO) (white circles), or 100 μM positive control inhibitor CGP-74514A[48] (red circles). A total of 64 compounds inhibited NEIL1 > 36%; 49 compounds are excluded from the diagram due to excessive autofluorescence (<−200% apparent NEIL1 inhibition). (C) Selected chemotypes found in the high-throughput screen. (D) Representative IC_{50} curves of control compound CGP-74514A and confirmed hits.

(1TDH)[30] and DNA-bound NEIL1 (5ITY)[29] most likely reflects the broad substrate specificity of NEIL1 and possible syn- and anti-orientation of each substrate.[86,87] Considering DogSite, SiteMap, and polarity, NEIL1 is among the most promising proteins in the family of DNA glycosylases (Figure 3). For these reasons, we decided to investigate NEIL1 for its experimental druggability using several experimental screening methods.

NEIL1 Activity Assay Development and HTS. NEIL1 is a bifunctional DNA glycosylase with comparatively broad substrate specificity, excising oxidized pyrimidines followed by an associated β- and δ-lyase activity acting on the apurinic (AP) site.[22] Thus, to develop a biochemical activity assay, a duplex oligonucleotide containing an internally quenched fluorophore was used. A NEIL1 substrate nucleobase in the form of thymine glycol (Tg) was positioned six base pairs from the 5′-terminal

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

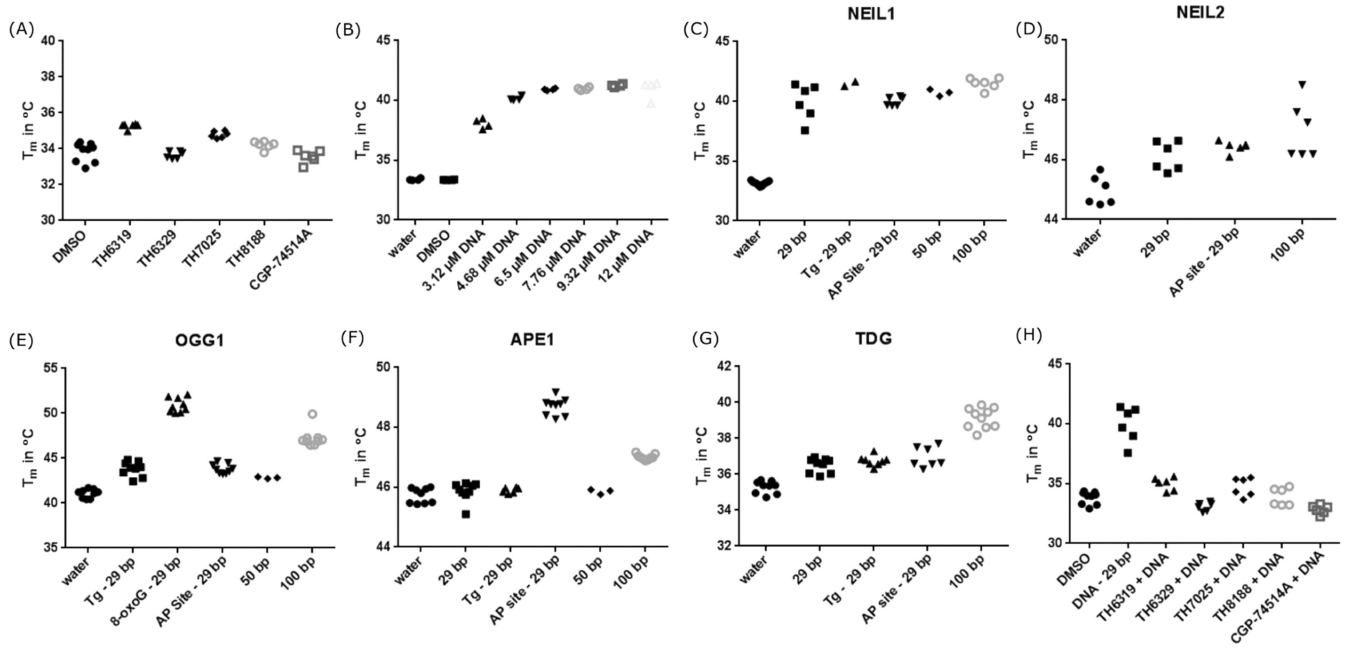


Figure 7. In vitro target engagement confirmation using differential scanning fluorimetry. (A) Hits from the HTS were incubated with NEIL1Δ56 at 200 μM concentration in the absence of DNA, and thermal stabilization was assessed. However, NEIL1Δ56 was only marginally stabilized by hit compounds. (B) Thermal stabilization of NEIL1 Δ56 by undamaged double-stranded DNA is dose-dependent and peaks above physiological temperature at >40 °C. (C) The same effect is observed for longer DNA oligos (50 bp, 100 bp) and for damaged DNA modifications known to be recognized by NEIL1 (Tg, AP). (D−G) NEIL2, OGG1, APE1, and TDG are thermally stabilized by DNA in the same substrate- and/or length-dependent manner. (H) DNA does not stabilize NEIL1Δ56 in the presence of hit compounds. An adapted version of the target engagement assay with DNA being added after incubating NEIL1Δ56 with the HTS hits confirms target engagement.

fluorophore. Excision of thymidine glycol, followed by incision at the AP site by the associated lyase activity of NEIL1 led to a localized melting of the DNA helix, causing the fluorophore to be released from the quencher.[88] Both NEIL1 and *E. coli* endonuclease VIII (Nei) enzymes were able to cause release of fluorescence from the oligo duplex containing Tg. No such increase was observed for the control oligonucleotide containing a normal thymine instead of thymine glycol. Importantly, a NEIL1 concentration-dependent increase in signal with a signal-to-noise ratio at around ∼20 was observed (Figure 6A). The assay was used to perform a high-throughput screen, as an orthogonal method for small-molecule hit confirmation and for determination of IC_{50} values.

A total of 30 114 compounds were screened to identify potential inhibitors. As positive control for inhibition, the previously identified NEIL1 inhibitor analogue CGP-74514A was employed at 100 μM concentration.[48] Using this setup, an average Z' factor of 0.69 (n = 86 plates) was achieved, indicating that the assay was sufficiently robust to identify inhibitors. Indeed, 64 compounds (an initial hit rate of 0.21%) were found to inhibit the enzymatic reaction by 36% or more (Figure 6B); 12 of these were excluded from further analysis due to structural frequent hitter behavior and pan-assay interference alerts. The remaining 52 compounds were then tested in three-dose confirmation assays and counterscreened for DNA intercalation. Here, 14 compounds failed to produce a dose-dependent inhibition and 17 compounds were found to intercalate DNA to some extent. The remaining 21 compounds were subjected to dose−response assays, where the chemotypes of TH8187 and TH6319, as well as adenine analogues of CGP-74514A were confirmed to be low μM inhibitors of NEIL1 (Figure 6C).[48,89]

DSF Fails to Validate Target Engagement of Biochemical Activity Screen Hits. Differential scanning fluorimetry[90,91] is a fast and readily available method to experimentally assess protein druggability and target engagement in vitro. In addition, the advantages of fragment-based drug design render DSF the ideal tool to add complementary chemical starting points to a medicinal chemistry campaign. Initially, the hits from HTS and their analogues were used to perform validation of target engagement. The melting temperature T_m of NEIL1 was found to be below physiological temperatures. Furthermore, earlier reported ligands for NEIL1[48] and chemotypes obtained in the screen above failed to significantly stabilize NEIL1Δ56 (Figure 7A). Only TH8187 and TH6319 and analogues thereof were able to stabilize about 1.3 °C at 200 μM in a 96-well plate setup.

DNA-Binding Proteins Are Thermally Stabilized by DNA. NEIL1 interacts with DNA in a mechanism where a nucleobase is "flipped out" and continuously interrogates the respective DNA strand for damaged nucleobases.[92] This process should stabilize NEIL1 through the formation of several interaction patterns. Possible phosphate−amino group salt bridges and multiple hydrogen bonds are readily observed when inspecting the respective thymine glycol-bound crystal structure (5ITY). Indeed, when added to the DSF assay, DNA stabilized NEIL1 to higher melting temperatures (Figure 7B). Furthermore, this effect was enhanced, when thymine glycol or a THF analogue of an abasic site was incorporated into DNA oligonucleotides (Figure 7C). We then investigated whether this effect was applicable to other DNA-binding proteins. OGG1, APE1, NEIL2, and TDG were stabilized in a length-dependent manner by DNA. Additionally, OGG1 showed increased T_m in the presence of 8-OxoG and the abasic-site analogue, while NEIL2 and APE1 were stabilized by the abasic-site analogue (Figure 7C−F).

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

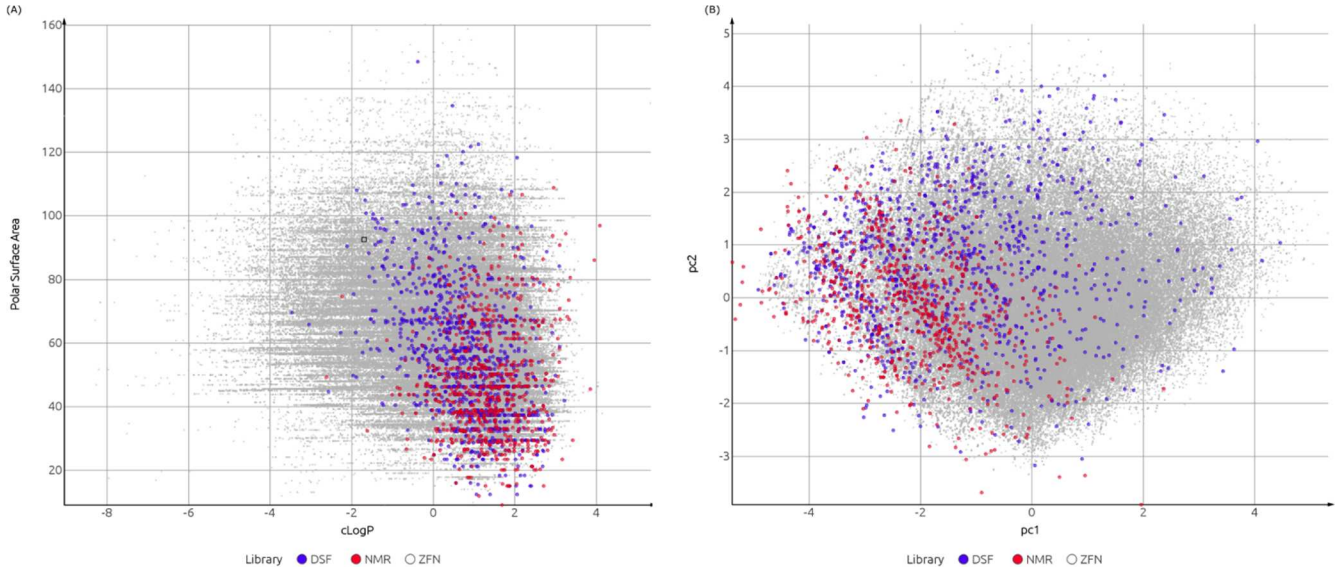


Figure 8. Comparison of the 200 K member virtual fragment library selected from ZINC Fragments Now (ZFN, gray), the laboratory for chemical biology at Karolinska Institutet (LCBKI) DSF fragment library (DSF, blue), and the Maybridge Ro3 NMR library (NMR, red) with respect to: (A) polarity and hydrophobicity, expressed as calculated polar surface area versus cLogP; (B) physicochemical property coverage and diversity, expressed as the first two principal components (pc1 and pc2) obtained from a principal component analysis (PCA) on six Lipinski-type properties.

Modified DNA−DSF Assay Validates Target Engagement by NEIL1 HTS Hits. Based on the observation from above, a compound that inhibits thymine glycol excision in a biochemical activity assay should also result in a different T_m for NEIL1 in the presence of DNA. Consequently, the HTS hits were incubated with NEIL1 and, in a second step, DNA was added. The HTS hits induced a T_m shift between 5.0 and 7.1 K and on average 6.1 K relative to the T_m of 40.0 °C measured for the NEIL1Δ56-DNA control (Figure 7H). These results indicate a reduced formation of a DNA−protein complex in the presence of inhibitor and confirm target engagement in a category I, II, or III pocket. In summary, we suggest DSF in the presence of DNA as an additional alternative method for assessment of in vitro target engagement of DNA-binding proteins.

DSF Fragment Screen against NEIL1 and OGG1. In contrast to the biochemical activity screen, a fragment screen in the absence of DNA relies on an unbiased protein−binder interaction and should in principle pick up hits that bind all available binding sites on the protein. In fragment screens, observed hit rates are usually much higher compared to conventional HTS and typically lie between 3 and 10%. We performed a DSF-based fragment screen of NEIL1 in the absence of DNA using a library that consists of 1000 fragments covering diverse chemical space (Figure 8), and 200 hand-picked nucleobase-like fragments. In this particular screen, we defined a hit as stabilizing more than the positive control TH8187 or 3 times the standard deviation of all DMSO controls. Using these definitions, a screen of the 1200 fragment library (average T_m 36.4 °C with a standard deviation of 0.94 °C) gave an initial hit rate of 4.3%, or 52 fragments. After revalidation in DSF, purity confirmation by liquid chromatography-mass spectrometry, and, if required, identity confirmation by NMR, a final set of 25 confirmed hits remained, corresponding to a hit rate of 2.1%. Next, validation of these fragments was performed in the orthogonal biochemical activity assay using a concentration series; 10 compounds or 40% of identified fragment hits showed inhibition of NEIL1-mediated Tg repair in the low μM range and thus confirmed interaction with pocket categories I, II, or III.

This hit rate is at the lower end of commonly observed fragment screening hit rates and stands in contrast to the computational assessment of NEIL1. To investigate whether this observation is reproduced for other DNA glycosylases, we expanded our fragment screening campaign to OGG1. We and others have recently shown that OGG1 can be specifically inhibited by druglike small molecules.[18,19] T_m of OGG1 in DSF was 43.2 °C, a hit was considered to stabilize at least 1 °C, and TH5487[19] was used as a positive control. The corresponding fragment screen (average T_m of 43.1 °C and a standard deviation of 0.82 °C) initially yielded 62 fragment hits or a hit rate of 5.2%. After revalidation in DSF, and purity and identity confirmation, 35 molecules or 2.9% of the initial library remained. These compounds, occupying different chemical space as previously published OGG1 inhibitors,[18,19] were then tested in the biochemical OGG1 activity assay, where 16 fragments (46%) showed inhibition of 8-oxoG excision. Despite the slightly higher hit rate compared to NEIL1, these results indicate that DNA glycosylases show lower hit rates than expected by computational assessment.

NEIL1 NMR Fragment Screen. Fragment screens performed in different assays, such as DSF, SPR, and saturation transfer difference (STD)-NMR, may yield a different number of binder and not necessarily confirm small-molecule binding of a different method. To assess whether the experimental druggability scores identified for NEIL1 and OGG1 in DSF could be confirmed by a different biophysical screening method, an STD-NMR screen was performed against NEIL1 in the absence of DNA. In STD-NMR, small-molecule binders to the surface of the target protein are identified after differencing the pure protein NMR spectra from a spectrum with small molecules added to the same protein. Due to little spectral overlapping of small molecules, typical samples contain between 5 and 10 individual compounds. This approach allows for the rapid screen of larger-fragment libraries. In the present case, using an 800-member subset of the commercially available

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

Maybridge Ro3 fragment library, 75 fragment hits or an initial hit rate of 9.4% were found to bind NEIL1 with mostly weak signal responses. Hit validation was performed with the 10 compounds giving the strongest signal. Eight molecules (hit rate, 1.0%) were confirmed to engage NEIL1 and subsequently tested in the biochemical assay. Here, four compounds or 50% showed inhibition of NEIL1-mediated Tg repair. This confirms the findings of the DSF screens, as 40−50% of binders are inhibitors of DNA-binding events. Questioning whether unsuitable library chemotype composition could account for the values at the lower end of typically observed hitrates, the screening libraries from LCBKI for DSF and Maybridge Ro3-1000 for NMR were compared to the exhaustive ZINC fragment library subset (Figure 8). The observed overlap in chemical space suggests no chemotype bias for the performed screens. In summary, experimental hit rates are lower than expected by our initial computational assessment. Nonetheless, a large percentage of indentified binders are inhibitors of catalytic activity and thus are suitable chemical starting points for an ongoing drug discovery campaign against NEIL1.

SUMMARY

Here, we described a comprehensive computational druggability assessment of human DNA glycosylases. We then assessed these findings in an experimental setting using NEIL1 and attempted to rationalize the difficulties in developing selective and potent druglike inhibitors for DNA-binding proteins. Recently, DNA glycosylases have been drawing considerable attention, with the report of a first-in-class OGG1 inhibitor,[18] followed subsequently by a comprehensive study for the application of other OGG1 inhibitors in inflammation.[19] Despite these results, and to the best of our knowledge, no articles explicitly reporting on the druggability assessment of DNA binding molecules are available. Hajduk et al. consider proteins, which bind DNA or RNA, and carbohydrate-binding glycosylases as representatives of highly challenging target families.[33,35] The findings reported in the present article suggest that this conclusion is not valid for DNA glycosylases. In contrast, when interrogated by computational measures, DNA glycosylases in general appear to be feasable drug targets and OGG1, MUTYH, NEIL1, UNG, and TDG, in particular, were found to be favorable drug-binding proteins (Figures 2, 3, and 5). Our study further confirms that DNA-binding proteins are difficult to assess with regard to which protein state is investigated. Multiple examples were found to be contradictory when DNA was either removed from or retained in the same high-resolution structure of a protein. Finally, our computational assessment suggests that catalytic sites of DNA glycosylases are conformationally active, which is yet unsupported by a number of high-resolution crystal structures of the same protein (Figure 5). This observation allows for flexible and dynamic DNA binding, enabling multiple protein conformations that also allow small-molecule binding. These mechanisms, which were rationalized for NEIL1 (Figures 4 and 5), allow for differences in binding of certain chemotypes in the different stages, i.e., the presence or absence of DNA.

By employing several screening methods, we carried out an experimental investigation of our NEIL1 druggability assessment to determine if it was verifiable in vitro. Initially, a biochemical assay using a NEIL1 substrate was applied in a HTS to screen for small-molecule inhibitors (Figure 6). Here, identified inhibitors failed to stabilize NEIL1 in the absence of DNA when evaluated using DSF. In the presence of DNA, thermal stabilization was observed for several DNA binding proteins. Furthermore, compared to the T_m observed in these scenarios, biochemically active compounds inhibited thermal stabilization by DNA and thus were confirmed as binders of NEIL1 (Figure 7). To identify binders of NEIL1, we performed a DSF screen using 1200 rule-of-three-compliant fragments. We identified 25 compounds (2.1%), of which 40% were confirmed to inhibit NEIL1 enzyme activity in the biochemical assay. To investigate whether this low hit rate was typical for DNA glycosylases, we performed a DSF screen against OGG1. Here, 35 compounds (2.9%) were confirmed as binders, of which 46% were inhibitors of 8-oxoG excision. Furthermore, an NMR fragment screen was performed against NEIL1; 8 fragments were identified to bind NEIL1, of which 50% inhibited NEIL1-mediated Tg removal. Overall, the hit rates at the lower end of the range typically observed for fragment screening reflect the challenges with screening DNA glycosylases for starting points of a prospective medicinal chemistry campaign. This stands in contrast to the computational druggability assessments in this article. Nevertheless, DNA glycosylases emerge as novel drug targets,[18,19,93] as they are involved in increased DNA repair necessary in a number of diseases and general inflammation. The present study and its findings lay the foundations as a basis for how to address this challenging family of enzymes and to exploit them as potential drug targets.

MATERIALS AND METHODS

Computational Protein Preparation. Available crystal structures of human DNA glycosylases with the highest resolution were imported into Maestro (Schrödinger Suite 2018-1, Schrödinger, LLC, New York, NY, 2018). NEIL1, OGG1, MBD4, and TDG were accessible in DNA-bound and DNA-unbound states. MPG and UNG were only available as co-crystals with DNA. Crystal structures of SMUG1, NTHL1, and NEIL2 have not yet been disclosed by the scientific community. Human MUTYH and mouse NEIL3 are available as deletion mutants in DNA-unbound state only. MTH1 (NUDT1), a well-characterized nucleotide-binding protein, was included as positive control in its apo state. PTP1B was added as a negative control. The structures were then prepared using the Protein Preparation Wizard[94−96] as implemented in Schrödinger Suite 2018-1. Briefly, protein data bank (pdb) structures were processed by automatically assigning bond orders, adding hydrogens, creating zero-order bonds to metals, creating possible disulfide bridges, adding missing side chains, deleting waters beyond 5.0 Å of hetero groups (if present), and generating hetero states at pH 7.0. Protonation and metal charge states were then generated for the hetero groups and visually inspected, and the most likely states based on hydrogen-bonding pattern and state penalty were selected. The hydrogen-bonding networks were optimized automatically, by sampling water orientations and optimization of hydroxyls, Asn, Gln, and His states using ProtAssign. Subsequently, all waters with less than four H bonds to nonwaters were removed, and finally, the structures were submitted to a restrained minimization in the optimized potentials for liquid simulations (OPLSs) 3 force field,[97−100] until the heavy-atom positions had converged to a root-mean-square deviation (RMSD) of atomic positions of 0.30 Å. If present, DNA was removed in a copy of the structure and both protein states were kept. Depending on their online availability, protein sets comprising apo, DNA-bound, and DNA-stripped structures were submitted to SiteMap or exported as protein data bank (pdb) files and imported to DogSite[45] and CryptoSite[44] for subsequent computational

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

binding-site prediction and druggability assessment. Sequence alignment was performed at Clustal Omega at the EMBL and phylogenetic trees were generated using Dendroscope.[101]

In Silico Fragment Screening.

(1) Fragment subset selection: a subset of the ZINC Frags Now set[79] was created by applying a number of filters implemented in a Knime workflow (Knime 3.5.2[102,103]). Foremost, only fragments available from a list of 19 preferred suppliers, composed of a team of experienced medicinal chemists, were considered. These were then filtered using a cascade of structural filters, including REOS,[104] PAINS,[105] and a set of in-house filters (ScrapFilter). Lipinski-type descriptors (SlogP, TPSA, AMW, NumLipinskiHBA, NumLipinskiHBD, NumRotatableBonds, NumHeavyAtoms, NumRings, NumAromaticRings) were then calculated using the RDKit Descriptor Calculation node. An additional descriptor HetRatio was then calculated as the ratio of NumLipinskiHBA and NumHeavyAtoms, and fragments with HetRatio <0.2 or >0.5 were filtered out. Finally, remaining outliers were removed by applying Gaussian *Z*-score normalization on the descriptor space, filtering fragments with descriptor values deviating more than 3 units from the mean. The entire filtering cascade reduced the original input file of 704 041 structures as downloaded from ZINC to 205 891 fragments.

(2) Ligand preparation: the selected fragment subset was then prepared for docking using LigPrep (Schrodinger): the OPLS3 force field was used for minimizations; possible ionization states at pH 7.0 ± 2.0 were generated using Epik,[106,107] metal-binding states were added, and tautomers were generated; specified chiralities were retained and at most four stereoisomers were generated per structure. This yielded 345 044 structures for docking.

(3) Docking grid generation: Glide docking grids[108−110] were generated for each target protein by focusing the grid box on the center of the site with the highest Dscore as determined by SiteMap.[77,78]

(4) Virtual screening: The virtual screening workflow as implemented in Schrödinger Suite was used for docking, scoring, and ranking of the top-1000 fragments against the sites with the highest Dscore as determined by SiteMap. The workflow comprised a cascade of docking steps with increased accuracy (Glide HTVS → SP → XP), where the top-10% ranked ligands were passed on to the next step. After Glide XP docking, the top-1000 ranked fragments were retained for druggability assessment based on docking scores.

Fragment Library Comparison. The 205 K fragment subset selected from ZINC Fragments Now, DSF fragment library, and the NMR fragment library were compared with respect to physicochemical properties coverage. To this effect, the structures were imported into DataWarrior 4.7.2[111] and six Lipinski-type properties were calculated: molecular weight, cLogP, H-bond acceptor count, H-bond donor count, polar surface area, and rotatable bond count. Principal component analysis (PCA) was then applied to the descriptor matrix to be able to visualize this six-dimensional space and compare the libraries in two dimensions.

Recombinant Proteins. APE1 was expressed from pET14b, and OGG1 from pNIC28a with N-terminal histidine tags. TDG, NEIL1, and NEIL2 were expressed from pET-28a. NEIL1 and NEIL2 contained C-terminal histidine tags, whereas TDG contained an N-terminal histidine tag. The pET-28a-NEIL1Δ56 expression construct was generated using FastCloning.[112] APE1 and OGG1 were expressed and purified as described here.[19] Full-length NEIL1 and NEIL2, NEIL1Δ56, and TDG were purified by the Protein Science Facility at Karolinska Institutet. *Escherichia coli* endonuclease VIII were purchased from New England Biolabs (M0299).

HTS Compound Library. Screening for inhibitors of NEIL1 was performed at the Laboratories of Chemical Biology Karolinska Institutet (LCBKI), part of Chemical Biology Consortium Sweden, using a compound concentration of 10 μm in 384-well microtiter plates (black OptiPlates, PerkinElmer). Rule-of-five-compliant, quality-controlled compounds originating from Enamine were screened, along with an LCBKI in-house compound library (donated by Biovitrum AB). All assay plates contained test compounds in columns 1−22, whereas negative and positive controls in columns 23 and 24 were used for normalization. In this layout, the negative controls represent the uninhibited enzyme, corresponding to empty wells or wells with the equivalent amount of DMSO as the compound containing wells, whereas the positive controls represent the completely inhibited enzyme, owing to the presence of 100 μM of the previously discovered NEIL1 inhibitor CGP-74514A[48]

HTS-Compatible DNA Glycosylase Activity Assays. The NEIL1 activity assay was performed at room temperature in black 384-well plates (black OptiPlates, PerkinElmer) using final concentrations of 25 mM Tris-HCl pH 8.0, 20 mM KCl, 1 mM EDTA, 0.5 mM DTT, 0.1% bovine serum albumin, 0.0025% Tween-20, 10 nM NEIL1 enzyme, and 20 nM Tg/A substrate in a final volume of 25 μL. The Tg/A substrate was a duplex oligonucleotide where 5′-(dT-FAM)CTG CCA YCA CTG CGT CGA CCT G-3′ was annealed to a 25% surplus of 5′-CAG GTC GAC GCA GTG CTG GCA G(dT-DABCYL)-3′. "Y" signifies thymidine glycol, and "dT-FAM" and "dT-DABCYL" signify fluorescein and dabcyl, respectively, both coupled to dT (TriLink BioTechnologies). Briefly, compounds dissolved in DMSO were dispensed using an Echo 550 (Labcyte), followed by transfer of enzyme and substrate solutions by a MultiDrop (PerkinElmer), followed by centrifugation at 1000 g to remove bubbles. The reactions were read after 20 min in an Envision plate reader using a 485 nm filter with a bandwidth of 14 nm for excitation and a 535 nm filter with a bandwidth of 25 nm for emission. OGG1 activity was measured as described here.[19]

DNA Intercalation Counterscreen. DNA intercalation was measured by incubating 20 nM unlabeled, undamaged oligonucleotides in NEIL1 reaction buffer (excluding bovine serum albumin) in the presence of 100 nM Thiazole Orange (Sigma-Aldrich 390 062) as described in ref 113.

DSF Fragment Library. The laboratory for chemical biology at Karolinska Institutet (LCBKI)-based fragment library consists of 1200 rule-of-three-compliant substances, which have been filtered for undesirable functionalities and PAINS. Along with a versatile set of chemotypes, the library also contains a population of small-molecule nucleotide analogues for specifically targeting DNA-, nucleotide-, and nucleoside-binding proteins. All substances were obtained commercially from different vendors (Enamine, Activate Scientific, Combi-Blocks, Sigma-Aldrich, Fisher Scientific, AK Scientific, abcr Chemie, Fluorochem, Spirochem, KeyOrganics, 1Click Chemistry, VWR, TCI, ChemBridge, Carbosynth) or from the NCI Developmental Therapeutics Program and are stored at 4 °C as 50 mM DMSO stock solutions.

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

DSF Fragment Screen. White BioRad 384-well plates with duplicates of each compound, DMSO, and positive control were prepared. Next, 9.8 μL of protein buffer, containing 25 mM Tris-acetate pH 7.5, 150 mM NaCl, 10% glycerol, 1 mM DTT, 5X SYPRO Orange, and 4 μM NEIL1, lacking 56 amino acids dispensable for activity at the disordered C-Terminus, was added to each well, yielding a final protein concentration of 3.92 μM and a final compound concentration of 1 mM. Similarly, 4 μM OGG1 was screened in 25 mM Tris-acetate pH 7.5, 50 mM $CaCl_2$, 10% glycerol, 1 mM DTT, 5X SYPRO Orange. The plates were sealed with BioRad MicroSeal 'B' and subjected to a temperature gradient from 20 to 95 °C in a Roche Light Cycler 480 II. The resulting fluorescence at 465−580 nm with an excitation wavelength of 465 nm was measured. Visual quality control of each graph was performed, and the resulting data were processed and imported to a GraphPad Prism template provided by Niesen et al.[90] For optimization and DNA experiments not concerning the fragment screens, black BioRad Hard-Shell PCR 96-well thin-wall plates were used in a BioRad C1000 Touch Thermal Cycler with 510 nm excitation wavelength and measurement at 465 nm. NEIL2, APE1, and TDG were prepared in a protein buffer containing 4 μm protein, 25 mM Tris-acetate pH 7.5, 150 mM NaCl, 10% glycerol, 1 mM DTT, and 5X SYPRO Orange.

DNA Oligo Preparation and DSF Use. DNA oligos of different sequences were obtained from IDT and TriLink: 50 base pairs: (5′-TAC GCT AGT ATG CGT TCT TCC TTC CAG GGG TAT GTG GCT GCG TGG TCA AA-3′ and 5′-TTT GAC CAC GCA GCC ACA TAC CCC TGG AAG GAA GAA CGC ATA CTA GCG TA-3′), 100 base pairs: (5′-CCT GAC GCT AGT GCA AAA CAC CCA AGC GAC CCT GAC AGT GCG AAT TGG CGA GCC TTA AGC TCT TTC GTT TGC TGA CGA GCG TTG CTG CGA GAC TGG CTT G-3′ and 5′-CAA GCC AGT CTC GCA GCA ACG CTC GTC AGC AAA CGA AAG AGC TTA AGG CTC GCC AAT TCG CAC TGT CAG GGT CGC TTG GGT GTT TTG CAC TAG CGT CAG G-3′), APE abasic-site analogue, 29 base pairs: (5′-TGT ATC GAT ACC XTC AAC CTC GAG GAA TT and AAT TCC TCG AGG TTG AGG TAT CGA TAC A-3′, where X signifies tetrahydrofuran), 100 base pairs: (5′-CTT CAA CCT CGA GGA ATG TAT CGA TAC CTT CAA CCT CGA GGA ATG TAT CGA TAC CTT CAA CCT CGA GGA ATG TAT CGA TAC CTT CAA CCT CGA GGA ATT-3′ and 5′-AAT TCC TCG AGG TTG AAG GTA TCG ATA CAT TCC TCG AGG TTG AAG GTA TCG ATA CAT TCC TCG AGG TTG AAG GTA TCG ATA CAT TCC TCG AGG TTG AAG-3′), thymine glycole control, 29 base pairs: (5′-TGT ATC GAT ACC TTC AAC CTC GAG GAA TT and 5′-AAT TCC TCG AGG ATG ACG GTA TCG ATA CA-3′), 8-oxo G, 29 base pairs: (5′-TGT ATC GAT ACC YTC AAC CTC GAG GAA TT and 5′-AAT TCC TCG AGG TTG ACG GTA TCG ATA CA-3′, where Y signifies 8-oxoG), thymine glycol, 29 base pairs: (5′-TGT ATC GAT ACC ZTC AAC CTC GAG GAA TT-3′ and 5′-AAT TCC TCG AGG TTG AAG GTA TCG ATA CA-3′, where Z signifies thymine glycol); 200 μM DNA oligo stock solutions were prepared in a buffer system containing 25 mM Tris-acetate pH 7.5, 150 mM NaCl, and 10% glycerol and stored at −20 °C. The required amounts of the complementary strands were combined in a 1:1 ratio and heated to 90 °C for 1 min. Storing at room temperature for 1 h and then cooling on ice allows for DNA double-strand generation in a 200 μM stock solution. After optimization of DNA concentration, 8 μM DNA was used for a standard experiment.

Statistical Analysis DSF Fragment Screen. Statistical analysis of all graphs was performed with the DSF template provided by Niesen et al.[90] The nonlinear fit performed by GraphPad Prism considers half the difference between a pretransitional minimum and a post-transitional maximum to be the melting temperature of the protein.

NMR Fragment Screen. NEIL1 in 20 mM HEPES, 300 mM NaCl, 10% glycerol, 2 mM TCEP, pH 7.5 was dialyzed against 1× PBS, pH 7.4 overnight, and added to fragment mixtures containing 10 compounds, each leading to a final concentration of 10 μm NEIL1 and 100 μM of compound. The fragment mixtures were prepared from 100 mM DMSO-d_6 stock solutions, diluted in 1× PBS pH 7.4 containing 10% D_2O and 50 μm trimethylsilylpropanoic acid (TMSP-d_4) as a chemical shift reference. A Bruker SamplePro liquid handling system with cooling was used to transfer samples to 3 mm NMR tubes. A total of 800 compounds, selected from the Maybridge Ro3 1000 fragment library (www.maybridge.com), were used in the screening. Both T_2-filtered spectra[114] and saturation transfer difference (STD) spectra[115] were recorded on all samples. A Carr−Purcell−Meiboom−Gill (CPMG) spin-lock of 200 ms was used in the T_2-filtered experiment, and in the STD experiment, on- and off-resonance saturations at 0.85 and −40 ppm were achieved with 40 and 50 ms Gaussian-shaped pulses, respectively. A total of 64 scans were recorded for both experiments. Reference spectra on fragment mixtures without NEIL1 were recorded separately. The NMR screen was performed within the NMR for Life (www.nmrforlife.se) initiative at the Swedish NMR Centre, Umeå, Sweden. All spectra were recorded on a Bruker 600 MHz Avance III HD spectrometer equipped with a 5 mm TCI cryoprobe and a SampleJet sample changer with cooling option. Spectral processing and analysis was performed in Topspin 4.0.1 (www.bruker.com). Hit validation was performed by repeating the experiments described above on individual compounds instead of mixtures.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.9b00162.

Raw data of protein assessment by DogSite, SiteMap, CryptoSite, as well as HTS data (PDF)

SiteMap and DogSite results; CryptoSite; IC 50; top 1000 zinc scores; DSF (XLSX)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: maurice.grube@scilifelab.se (M.M.).

*E-mail: t.helleday@sheffield.ac.uk (T.H.).

ORCID

Maurice Michel: 0000-0003-3261-2493

Torkild Visnes: 0000-0003-1047-988X

Author Contributions

$^{\nabla}$M.M. and T.V. contributed equally to the work.

Notes

The authors declare the following competing financial interest(s): T.V., O.W., and T.H. are listed as inventors on a provisional patent application No. 62/636983, covering OGG1 inhibitors. The patent is fully owned by a non-profit public foundation, the Helleday Foundation, and T.H. and U.W.B. are

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

member of the foundation board developing OGG1 inhibitors towards the clinic. An inventor reward scheme is under discussion. The remaining authors declare no competing financial interests.

■ ACKNOWLEDGMENTS

The authors gratefully acknowledge LCBKI for supplying the DSF fragment library and help during development of the glycosylase assay and the HTS for NEIL1, NMR for Life for the NMR screen using the Maybridge Ro3 fragment library (provided by Umeå Centre for Microbial Research (UCMR) and Laboratories for Chemical Biology Umeå (LCBU)), the Protein Science Facility of Karolinska Institutet for expressing and purifying NEIL1Δ56 and NEIL1. They thank the operators of CryptoSite at the University of California, San Francisco; DogSite at the Zentrum für Bioinformatik, University of Hamburg; and Clustal Omega at the EMBL in Heidelberg for using their infrastructure and computing setup. Olga Loseva is acknowledged for expressing and purifying OGG1. Dana Michel is acknowledged for her critical discussion and proofreading of the manuscript. Funding was obtained from the Swedish Foundation for Strategic Research (T.H.), Swedish Cancer Society (T.H.), the Swedish Children's Cancer Foundation (T.H.), ERC Tarox-695376 (T.H.) and Swedish Pain Relief Foundation (T.H.).

■ REFERENCES

(1) Curtin, N. J. DNA Repair Dysregulation from Cancer Driver to Therapeutic Target. *Nat. Rev. Cancer* **2012**, *12*, 801−817.
(2) Helleday, T.; Petermann, E.; Lundin, C.; Hodgson, B.; Sharma, R. A. DNA Repair Pathways as Targets for Cancer Therapy. *Nat. Rev. Cancer* **2008**, *8*, 193−204.
(3) Gad, H.; Koolmeister, T.; Jemth, A.-S.; Eshtad, S.; Jacques, S. A.; Ström, C. E.; Svensson, L. M.; Schultz, N.; Lundbäck, T.; Einarsdottir, B. O.; et al. MTH1 Inhibition Eradicates Cancer by Preventing Sanitation of the DNTP Pool. *Nature* **2014**, *508*, 215−221.
(4) Narwal, M.; Jemth, A.-S.; Gustafsson, R.; Almlöf, I.; Warpman Berglund, U.; Helleday, T.; Stenmark, P. Crystal Structures and Inhibitor Interactions of Mouse and Dog MTH1 Reveal Species-Specific Differences in Affinity. *Biochemistry* **2018**, *57*, 593−603.
(5) Pudelko, L.; Rouhi, P.; Sanjiv, K.; Gad, H.; Kalderén, C.; Höglund, A.; Squatrito, M.; Schuhmacher, A. J.; Edwards, S.; Hägerstrand, D.; et al. Glioblastoma and Glioblastoma Stem Cells Are Dependent on Functional MTH1. *OncoTargets* **2017**, *8*, 84671−84684.
(6) Warpman Berglund, U.; Sanjiv, K.; Gad, H.; Kalderén, C.; Koolmeister, T.; Pham, T.; Gokturk, C.; Jafari, R.; Maddalo, G.; Seashore-Ludlow, B.; et al. Validation and Development of MTH1 Inhibitors for Treatment of Cancer. *Ann. Oncol.* **2016**, *27*, 2275−2283.
(7) Huber, K. V. M.; Salah, E.; Radic, B.; Gridling, M.; Elkins, J. M.; Stukalov, A.; Jemth, A.-S.; Göktürk, C.; Sanjiv, K.; Strömberg, K.; et al. Stereospecific Targeting of MTH1 by (S)-Crizotinib as an Anticancer Strategy. *Nature* **2014**, *508*, 222−227.
(8) Samaranayake, G. J.; Huynh, M.; Rai, P. MTH1 as a Chemotherapeutic Target: The Elephant in the Room. *Cancers* **2017**, *9*, No. 47.
(9) Oka, S.; Ohno, M.; Tsuchimoto, D.; Sakumi, K.; Furuichi, M.; Nakabeppu, Y. Two Distinct Pathways of Cell Death Triggered by Oxidative Damage to Nuclear and Mitochondrial DNAs. *EMBO J.* **2008**, *27*, 421−432.
(10) Ohno, M.; Sakumi, K.; Fukumura, R.; Furuichi, M.; Iwasaki, Y.; Hokama, M.; Ikemura, T.; Tsuzuki, T.; Gondo, Y.; Nakabeppu, Y. 8-Oxoguanine Causes Spontaneous *de Novo* Germline Mutations in Mice. *Sci. Rep.* **2015**, *4*, No. 4689.
(11) Eshtad, S.; Mavajian, Z.; Rudd, S. G.; Visnes, T.; Boström, J.; Altun, M.; Helleday, T. HMYH and HMTH1 Cooperate for Survival in Mismatch Repair Defective T-Cell Acute Lymphoblastic Leukemia. *Oncogenesis* **2016**, *5*, No. e275.
(12) Dou, H.; Mitra, S.; Hazra, T. K. Repair of Oxidized Bases in DNA Bubble Structures by Human DNA Glycosylases NEIL1 and NEIL2. *J. Biol. Chem.* **2003**, *278*, 49679−49684.
(13) Krokeide, S. Z.; Laerdahl, J. K.; Salah, M.; Luna, L.; Cederkvist, F. H.; Fleming, A. M.; Burrows, C. J.; Dalhus, B.; Bjørås, M. Human NEIL3 Is Mainly a Monofunctional DNA Glycosylase Removing Spiroimindiohydantoin and Guanidinohydantoin. *DNA Repair* **2013**, *12*, 1159−1164.
(14) Liu, M.; Bandaru, V.; Bond, J. P.; Jaruga, P.; Zhao, X.; Christov, P. P.; Burrows, C. J.; Rizzo, C. J.; Dizdaroglu, M.; Wallace, S. S. The Mouse Ortholog of NEIL3 Is a Functional DNA Glycosylase in Vitro and in Vivo. *Proc. Natl. Acad. Sci.* **2010**, *107*, 4925−4930.
(15) Hu, J.; de Souza-Pinto, N. C.; Haraguchi, K.; Hogue, B. A.; Jaruga, P.; Greenberg, M. M.; Dizdaroglu, M.; Bohr, V. A. Repair of Formamidopyrimidines in DNA Involves Different Glycosylases ROLE OF THE OGG1, NTH1, AND NEIL1 ENZYMES. *J. Biol. Chem.* **2005**, *280*, 40544−40551.
(16) Klungland, A.; Rosewell, I.; Hollenbach, S.; Larsen, E.; Daly, G.; Epe, B.; Seeberg, E.; Lindahl, T.; Barnes, D. E. Accumulation of Premutagenic DNA Lesions in Mice Defective in Removal of Oxidative Base Damage. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13300−13305.
(17) Souza-Pinto, N. C. de; Eide, L.; Hogue, B. A.; Thybo, T.; Stevnsner, T.; Seeberg, E.; Klungland, A.; Bohr, V. A. Repair of 8-Oxodeoxyguanosine Lesions in Mitochondrial DNA Depends on the Oxoguanine DNA Glycosylase (OGG1) Gene and 8-Oxoguanine Accumulates in the Mitochondrial DNA of OGG1-Defective Mice. *Cancer Res.* **2001**, *61*, 5378−5381.
(18) Tahara, Y.; Auld, D.; Ji, D.; Beharry, A. A.; Kietrys, A. M.; Wilson, D. L.; Jimenez, M.; King, D.; Nguyen, Z.; Kool, E. T. Potent and Selective Inhibitors of 8-Oxoguanine DNA Glycosylase. *J. Am. Chem. Soc.* **2018**, *140*, 2105−2114.
(19) Visnes, T.; Cázares-Körner, A.; Hao, W.; Wallner, O.; Masuyer, G.; Loseva, O.; Mortusewicz, O.; Wiita, E.; Sarno, A.; Manoilov, A.; et al. Small-Molecule Inhibitor of OGG1 Suppresses Proinflammatory Gene Expression and Inflammation. *Science* **2018**, *362*, 834−839.
(20) Ohtsubo, T.; Nishioka, K.; Imaiso, Y.; Iwai, S.; Shimokawa, H.; Oda, H.; Fujiwara, T.; Nakabeppu, Y. Identification of Human MutY Homolog (HMYH) as a Repair Enzyme for 2-Hydroxyadenine in DNA and Detection of Multiple Forms of HMYH Located in Nuclei and Mitochondria. *Nucleic Acids Res.* **2000**, *28*, 1355−1364.
(21) Hazra, T. K.; Izumi, T.; Boldogh, I.; Imhoff, B.; Kow, Y. W.; Jaruga, P.; Dizdaroglu, M.; Mitra, S. Identification and Characterization of a Human DNA Glycosylase for Repair of Modified Bases in Oxidatively Damaged DNA. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 3523−3528.
(22) Bandaru, V.; Sunkara, S.; Wallace, S. S.; Bond, J. P. A Novel Human DNA Glycosylase That Removes Oxidative DNA Damage and Is Homologous to Escherichia Coli Endonuclease VIII. *DNA Repair* **2002**, *1*, 517−529.
(23) Takao, M.; Kanno, S.-I.; Kobayashi, K.; Zhang, Q.-M.; Yonei, S.; van der Horst, G. T. J.; Yasui, A. A Back-up Glycosylase in Nth1 Knock-out Mice Is a Functional Nei (Endonuclease VIII) Homologue. *J. Biol. Chem.* **2002**, *277*, 42205−42213.
(24) Fleming, A. M.; Burrows, C. J. Formation and Processing of DNA Damage Substrates for the HNEIL Enzymes. *Free Radicals Biol. Med.* **2017**, *107*, 35−52.
(25) Rolseth, V.; Luna, L.; Olsen, A. K.; Suganthan, R.; Scheffler, K.; Neurauter, C. G.; Esbensen, Y.; Kuśnierczyk, A.; Hildrestrand, G. A.; Graupner, A.; et al. No Cancer Predisposition or Increased Spontaneous Mutation Frequencies in NEIL DNA Glycosylases-Deficient Mice. *Sci. Rep.* **2017**, *7*, No. 4384.
(26) Sampath, H. Oxidative DNA Damage in Disease—Insights Gained from Base Excision Repair Glycosylase-deficient Mouse Models. *Environ. Mol. Mutagen.* **2014**, *55*, 689−703.
(27) Zander, U.; Hoffmann, G.; Cornaciu, I.; Marquette, J. P.; Papp, G.; Landret, C.; Seroul, G.; Sinoir, J.; Röwer, M.; Felisaz, F.; et al. Automated Harvesting and Processing of Protein Crystals through

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

Laser Photoablation. *Acta Crystallogr., Sect. D: Struct. Biol.* **2016**, *72*, 454−466.
(28) Dalhus, B.; Forsbring, M.; Helle, I. H.; Vik, E. S.; Forstrøm, R. J.; Backe, P. H.; Alseth, I.; Bjørås, M. Separation-of-Function Mutants Unravel the Dual-Reaction Mode of Human 8-Oxoguanine DNA Glycosylase. *Structure* **2011**, *19*, 117−127.
(29) Zhu, C.; Lu, L.; Zhang, J.; Yue, Z.; Song, J.; Zong, S.; Liu, M.; Stovicek, O.; Gao, Y. Q.; Yi, C. Tautomerization-Dependent Recognition and Excision of Oxidation Damage in Base-Excision DNA Repair. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 7792−7797.
(30) Doublié, S.; Bandaru, V.; Bond, J. P.; Wallace, S. S. The Crystal Structure of Human Endonuclease VIII-like 1 (NEIL1) Reveals a Zincless Finger Motif Required for Glycosylase Activity. *Proc. Natl. Acad. Sci.* **2004**, *101*, 10284−10289.
(31) Bruner, S. D.; Norman, D. P.; Verdine, G. L. Structural Basis for Recognition and Repair of the Endogenous Mutagen 8-Oxoguanine in DNA. *Nature* **2000**, *403*, 859−866.
(32) Sharma, N.; Chakravarthy, S.; Longley, M. J.; Copeland, W. C.; Prakash, A. The C-Terminal Tail of the NEIL1 DNA Glycosylase Interacts with the Human Mitochondrial Single-Stranded DNA Binding Protein. *DNA Repair* **2018**, *65*, 11−19.
(33) Hajduk, P. J.; Huth, J. R.; Tse, C. Predicting Protein Druggability. *Drug Discovery Today* **2005**, *10*, 1675−1682.
(34) Hajduk, P. J.; Huth, J. R.; Fesik, S. W. Druggability Indices for Protein Targets Derived from NMR-Based Screening Data. *J. Med. Chem.* **2005**, *48*, 2518−2525.
(35) Aretz, J.; Wamhoff, E.-C.; Hanske, J.; Heymann, D.; Rademacher, C. Computational and Experimental Prediction of Human C-Type Lectin Receptor Druggability. *Front. Immunol.* **2014**, *5*, No. 323.
(36) Rampazzo, C.; Tozzi, M. G.; Dumontet, C.; Jordheim, L. P. The Druggability of Intracellular Nucleotide-Degrading Enzymes. *Cancer Chemother. Pharmacol.* **2016**, *77*, 883−893.
(37) Hopkins, A. L.; Groom, C. R. The Druggable Genome. *Nat. Rev. Drug Discovery* **2002**, *1*, 727−730.
(38) Cheng, A. C.; Coleman, R. G.; Smyth, K. T.; Cao, Q.; Soulard, P.; Caffrey, D. R.; Salzberg, A. C.; Huang, E. S. Structure-Based Maximal Affinity Model Predicts Small-Molecule Druggability. *Nat. Biotechnol.* **2007**, *25*, 71−75.
(39) Chen, I.-J.; Hubbard, R. E. Lessons for Fragment Library Design: Analysis of Output from Multiple Screening Campaigns. *J. Comput.-Aided Mol. Des.* **2009**, *23*, 603−620.
(40) Barril, X. Druggability Predictions: Methods, Limitations, and Applications. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2013**, *3*, 327−338.
(41) Vidler, L. R.; Brown, N.; Knapp, S.; Hoelder, S. Druggability Analysis and Structural Classification of Bromodomain Acetyl-Lysine Binding Sites. *J. Med. Chem.* **2012**, *55*, 7346−7359.
(42) Kozakov, D.; Hall, D. R.; Napoleon, R. L.; Yueh, C.; Whitty, A.; Vajda, S. New Frontiers in Druggability. *J. Med. Chem.* **2015**, *58*, 9063−9088.
(43) Fauman, E. B.; Rai, B. K.; Huang, E. S. Structure-Based Druggability Assessment—Identifying Suitable Targets for Small Molecule Therapeutics. *Curr. Opin. Chem. Biol.* **2011**, *15*, 463−468.
(44) Cimermancic, P.; Weinkam, P.; Rettenmaier, T. J.; Bichmann, L.; Keedy, D. A.; Woldeyes, R. A.; Schneidman-Duhovny, D.; Demerdash, O. N.; Mitchell, J. C.; Wells, J. A.; et al. CryptoSite: Expanding the Druggable Proteome by Characterization and Prediction of Cryptic Binding Sites. *J. Mol. Biol.* **2016**, *428*, 709−719.
(45) Volkamer, A.; Kuhn, D.; Rippmann, F.; Rarey, M. DoGSiteScorer: A Web Server for Automatic Binding Site Prediction, Analysis and Druggability Assessment. *Bioinformatics* **2012**, *28*, 2074−2075.
(46) Kozakov, D.; Grove, L. E.; Hall, D. R.; Bohnuud, T.; Mottarella, S. E.; Luo, L.; Xia, B.; Beglov, D.; Vajda, S. The FTMap Family of Web Servers for Determining and Characterizing Ligand-Binding Hot Spots of Proteins. *Nat. Protoc.* **2015**, *10*, 733−755.
(47) Donley, N.; Jaruga, P.; Coskun, E.; Dizdaroglu, M.; McCullough, A. K.; Lloyd, R. S. Small Molecule Inhibitors of 8-Oxoguanine DNA Glycosylase-1 (OGG1). *ACS Chem. Biol.* **2015**, *10*, 2334−2343.
(48) Jacobs, A. C.; Calkins, M. J.; Jadhav, A.; Dorjsuren, D.; Maloney, D.; Simeonov, A.; Jaruga, P.; Dizdaroglu, M.; McCullough, A. K.; Lloyd, R. S. Inhibition of DNA Glycosylases via Small Molecule Purine Analogs. *PLoS One* **2013**, *8*, No. e81667.
(49) Aretz, J.; Kondoh, Y.; Honda, K.; Anumala, U. R.; Nazaré, M.; Watanabe, N.; Osada, H.; Rademacher, C. Chemical Fragment Arrays for Rapid Druggability Assessment. *Chem. Commun.* **2016**, *52*, 9067−9070.
(50) Aretz, J.; Baukmann, H.; Shanina, E.; Hanske, J.; Wawrzinek, R.; Zapol'skii, V. A.; Seeberger, P. H.; Kaufmann, D. E.; Rademacher, C. Identification of Multiple Druggable Secondary Sites by Fragment Screening against DC-SIGN. *Angew. Chem., Int. Ed.* **2017**, *56*, 7292−7296.
(51) Edfeldt, F. N. B.; Folmer, R. H. A.; Breeze, A. L. Fragment Screening to Predict Druggability (Ligandability) and Lead Discovery Success. *Drug Discovery Today* **2011**, *16*, 284−287.
(52) Erlanson, D. A.; Fesik, S. W.; Hubbard, R. E.; Jahnke, W.; Jhoti, H. Twenty Years on: The Impact of Fragments on Drug Discovery. *Nat. Rev. Drug Discovery* **2016**, *15*, 605−619.
(53) Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R. A New Bioinformatics Analysis Tools Framework at EMBL−EBI. *Nucleic Acids Res.* **2010**, *38*, W695−W699.
(54) Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; et al. Clustal W and Clustal X Version 2.0. *Bioinformatics* **2007**, *23*, 2947−2948.
(55) Xu, C.; Tempel, W.; Wernimont, A. K.; Bountra, C.; Arrowsmith, C. H.; Edwards, A. M.; Min, J. Crystal Structure MBD4 MBD Domain in Complex with Methylated CpG DNA 2013, DOI: 10.2210/pdb4lg7/pdb.
(56) Moréra, S.; Grin, I.; Vigouroux, A.; Couvé, S.; Henriot, V.; Saparbaev, M.; Ishchenko, A. A. Biochemical and Structural Characterization of the Glycosylase Domain of MBD4 Bound to Thymine and 5-Hydroxymethyuracil-Containing DNA. *Nucleic Acids Res.* **2012**, *40*, 9917−9926.
(57) Ouzon-Shubeita, H.; Lin, Y.-L.; Lee, S. Structure of MBD4 Bound to G:T Mispair DNA. DOI: 10.2210/pdb4ofa/pdb.
(58) Parikh, S. S.; Walcher, G.; Jones, G. D.; Slupphaug, G.; Krokan, H. E.; Blackburn, G. M.; Tainer, J. A. Uracil-DNA Glycosylase−DNA Substrate and Product Structures: Conformational Strain Promotes Catalytic Efficiency by Coupled Stereoelectronic Effects. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 5083−5088.
(59) Krosky, D. J.; Bianchet, M. A.; Seiple, L.; Chung, S.; Amzel, L. M.; Stivers, J. T. Mimicking Damaged DNA with a Small Molecule Inhibitor of Human UNG2. *Nucleic Acids Res.* **2006**, *34*, 5872−5879.
(60) Coey, C. T.; Malik, S. S.; Pidugu, L. S.; Varney, K. M.; Pozharski, E.; Drohat, A. C. Structural Basis of Damage Recognition by Thymine DNA Glycosylase: Key Roles for N-Terminal Residues. *Nucleic Acids Res.* **2016**, *44*, 10248−10258.
(61) Liu, M.; Imamura, K.; Averill, A. M.; Wallace, S. S.; Doublié, S. Structural Characterization of a Mouse Ortholog of Human NEIL3 with a Marked Preference for Single-Stranded DNA. *Structure* **2013**, *21*, 247−256.
(62) Luncsford, P. J.; Chang, D.-Y.; Shi, G.; Bernstein, J.; Madabushi, A.; Patterson, D. N.; Lu, A.-L.; Toth, E. A. A Structural Hinge in Eukaryotic MutY Homologues Mediates Catalytic Activity and Rad9−Rad1−Hus1 Checkpoint Complex Interactions. *J. Mol. Biol.* **2010**, *403*, 351−370.
(63) Svensson, L. M.; Jemth, A.-S.; Desroses, M.; Loseva, O.; Helleday, T.; Högbom, M.; Stenmark, P. Crystal Structure of Human MTH1 and the 8-oxo-dGMP Product Complex. *FEBS Lett.* **2011**, *585*, 2617−2621.
(64) Lau, A. Y.; Wyatt, M. D.; Glassner, B. J.; Samson, L. D.; Ellenberger, T. Molecular Basis for Discriminating between Normal and Damaged Bases by the Human Alkyladenine Glycosylase, AAG. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 13573−13578.
(65) Aravind, L.; Koonin, E. V. The Alpha/Beta Fold Uracil DNA Glycosylases: A Common Origin with Diverse Fates. *Genome Biol.* **2000**, *1*, No. RESEARCH0007.1.

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

(66) Barrett, T. E.; Savva, R.; Panayotou, G.; Barlow, T.; Brown, T.; Jiricny, J.; Pearl, L. H. Crystal Structure of a G:T/U Mismatch-Specific DNA Glycosylase: Mismatch Recognition by Complementary-Strand Interactions. *Cell* **1998**, *92*, 117−129.
(67) Baptiste, B. A.; Katchur, S. R.; Fivenson, E. M.; Croteau, D. L.; Rumsey, W. L.; Bohr, V. A. Enhanced Mitochondrial DNA Repair of the Common Disease-Associated Variant, Ser326Cys, of HOGG1 through Small Molecule Intervention. *Free Radical Biol. Med.* **2018**, *124*, 149−162.
(68) Ellermann, M.; Eheim, A.; Rahm, F.; Viklund, J.; Guenther, J.; Andersson, M.; Ericsson, U.; Forsblom, R.; Ginman, T.; Lindström, J.; et al. Novel Class of Potent and Cellularly Active Inhibitors Devalidates MTH1 as Broad-Spectrum Cancer Target. *ACS Chem. Biol.* **2017**, *12*, 1986−1992.
(69) Tresaugues, L.; Siponen, M. I.; Arrowsmith, C. H.; Berglund, H.; Bountra, C.; Collins, R.; Edwards, A. M.; Ekblad, T.; Flodin, S.; Flores, A. et al. Crystal Structure of Human 8-Oxo-DGTPase (MTH1). 2011, DOI: 10.2210/pdb3q93/pdb.
(70) Cho, H. Protein Tyrosine Phosphatase 1B (PTP1B) and Obesity. *Vitam. Horm.* **2013**, *91*, 405−424.
(71) Tamrakar, A. K.; Maurya, C. K.; Rai, A. K. PTP1B Inhibitors for Type 2 Diabetes Treatment: A Patent Review (2011 - 2014). *Expert Opin. Ther. Pat.* **2014**, *24*, 1101−1115.
(72) Vieira, M. N. N.; Lyra e Silva, N. M.; Ferreira, S. T.; De Felice, F. G. Protein Tyrosine Phosphatase 1B (PTP1B): A Potential Target for Alzheimer's Therapy? *Front. Aging Neurosci.* **2017**, *9*, 7.
(73) Krishnan, N.; Konidaris, K. F.; Gasser, G.; Tonks, N. K. A Potent, Selective, and Orally Bioavailable Inhibitor of the Protein-Tyrosine Phosphatase PTP1B Improves Insulin and Leptin Signaling in Animal Models. *J. Biol. Chem.* **2018**, *293*, 1517−1525.
(74) Zhang, S.; Zhang, Z.-Y. PTP1B as a Drug Target: Recent Developments in PTP1B Inhibitor Discovery. *Drug Discovery Today* **2007**, *12*, 373−381.
(75) Mullard, A. Phosphatases Start Shedding Their Stigma of Undruggability. *Nat. Rev. Drug Discovery* **2018**, *17*, 847−849.
(76) Barford, D.; Flint, A. J.; Tonks, N. K. Crystal Structure of Human Protein Tyrosine Phosphatase 1B. *Science* **1994**, *263*, 1397−1404.
(77) Halgren, T. New Method for Fast and Accurate Binding-Site Identification and Analysis. *Chem. Biol. Drug Des.* **2007**, *69*, 146−148.
(78) Halgren, T. A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377−389.
(79) Irwin, J. J.; Sterling, T.; Mysinger, M. M.; Bolstad, E. S.; Coleman, R. G. ZINC: A Free Tool to Discover Chemistry for Biology. *J. Chem. Inf. Model.* **2012**, *52*, 1757−1768.
(80) Irwin, J. J.; Shoichet, B. K. ZINC − A Free Database of Commercially Available Compounds for Virtual Screening. *J. Chem. Inf. Model.* **2005**, *45*, 177−182.
(81) Rahm, F.; Viklund, J.; Trésaugues, L.; Ellermann, M.; Giese, A.; Ericsson, U.; Forsblom, R.; Ginman, T.; Günther, J.; Hallberg, K.; et al. Creation of a Novel Class of Potent and Selective MutT Homologue 1 (MTH1) Inhibitors Using Fragment-Based Screening and Structure-Based Drug Design. *J. Med. Chem.* **2018**, *61*, 2533−2551.
(82) Yan, B. X.; Sun, Y. Q. Glycine Residues Provide Flexibility for Enzyme Active Sites. *J. Biol. Chem.* **1997**, *272*, 3190−3194.
(83) Saavedra, H. G.; Wrabl, J. O.; Anderson, J. A.; Li, J.; Hilser, V. J. Dynamic Allostery Can Drive Cold Adaptation in Enzymes. *Nature* **2018**, *558*, 324−328.
(84) Blank, I. D.; Sadeghian, K.; Ochsenfeld, C. A Base-Independent Repair Mechanism for DNA Glycosylase—No Discrimination Within the Active Site. *Sci. Rep.* **2015**, *5*, No. 10369.
(85) Kreppel, A.; Blank, I. D.; Ochsenfeld, C. Base-Independent DNA Base-Excision Repair of 8-Oxoguanine. *J. Am. Chem. Soc.* **2018**, *140*, 4522−4526.
(86) Banerjee, A.; Yang, W.; Karplus, M.; Verdine, G. L. Structure of a Repair Enzyme Interrogating Undamaged DNA Elucidates Recognition of Damaged DNA. *Nature* **2005**, *434*, 612−618.
(87) Sung, R.-J.; Zhang, M.; Qi, Y.; Verdine, G. L. Sequence-Dependent Structural Variation in DNA Undergoing Intrahelical Inspection by the DNA Glycosylase MutM. *J. Biol. Chem.* **2012**, *287*, 18044−18054.
(88) Madhusudan, S.; Smart, F.; Shrimpton, P.; Parsons, J. L.; Gardiner, L.; Houlbrook, S.; Talbot, D. C.; Hammonds, T.; Freemont, P. A.; Sternberg, M. J. E.; et al. Isolation of a Small Molecule Inhibitor of DNA Base Excision Repair. *Nucleic Acids Res.* **2005**, *33*, 4711−4724.
(89) Zhang, J.; Hübner, E. G.; Namyslo, J. C.; Nieger, M.; Schmidt, A. Purine-Substituted Imidazolium Mesomeric Betaines and Their Tautomeric N-Heterocyclic Carbenes. Formation of a Cyclic Borane Adduct. *Org. Biomol. Chem.* **2018**, *16*, 6801−6808.
(90) Niesen, F. H.; Berglund, H.; Vedadi, M. The Use of Differential Scanning Fluorimetry to Detect Ligand Interactions That Promote Protein Stability. *Nat. Protoc.* **2007**, *2*, 2212−2221.
(91) Chilton, M.; Clennell, B.; Edfeldt, F.; Geschwindner, S. Hot-Spotting with Thermal Scanning: A Ligand- and Structure-Independent Assessment of Target Ligandability. *J. Med. Chem.* **2017**, *60*, 4923−4931.
(92) Lee, A. J.; Wallace, S. S. Hide and Seek: How Do DNA Glycosylases Locate Oxidatively Damaged DNA Bases amidst a Sea of Undamaged Bases? *Free Radicals Biol. Med.* **2017**, *107*, 170−178.
(93) Visnes, T.; Grube, M.; Fekry Hanna, B. M.; Benitez-Buelga, C.; Cázares-Körner, A.; Helleday, T. Targeting BER Enzymes in Cancer Therapy. *DNA Repair* **2018**, *71*, 118.
(94) Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and Influence on Virtual Screening Enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221−234.
(95) Jacobson, M. P.; Friesner, R. A.; Xiang, Z.; Honig, B. On the Role of the Crystal Environment in Determining Protein Side-Chain Conformations. *J. Mol. Biol.* **2002**, *320*, 597−608.
(96) Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J. F.; Honig, B.; Shaw, D. E.; Friesner, R. A. A Hierarchical Approach to All-Atom Protein Loop Prediction. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 351−367.
(97) Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; et al. OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *J. Chem. Theory Comput.* **2016**, *12*, 281−296.
(98) Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W. Prediction of Absolute Solvation Free Energies Using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *J. Chem. Theory Comput.* **2010**, *6*, 1509−1519.
(99) Jorgensen, W. L.; Tirado-Rives, J. The OPLS [Optimized Potentials for Liquid Simulations] Potential Functions for Proteins, Energy Minimizations for Crystals of Cyclic Peptides and Crambin. *J. Am. Chem. Soc.* **1988**, *110*, 1657−1666.
(100) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *J. Am. Chem. Soc.* **1996**, *118*, 11225−11236.
(101) Huson, D. H.; Richter, D. C.; Rausch, C.; Dezulian, T.; Franz, M.; Rupp, R. Dendroscope: An Interactive Viewer for Large Phylogenetic Trees. *BMC Bioinf.* **2007**, *8*, No. 460.
(102) Berthold, M. R.; Cebron, N.; Dill, F.; Gabriel, T. R.; Kötter, T.; Meinl, T.; Ohl, P.; Sieb, C.; Thiel, K.; Wiswedel, B. KNIME: The Konstanz Information Miner. In *Data Analysis, Machine Learning and Applications*; Studies in Classification, Data Analysis, and Knowledge Organization; Springer: Berlin, Heidelberg, 2008; pp 319−326.
(103) Berthold, M. R.; Cebron, N.; Dill, F.; Gabriel, T. R.; Kötter, T.; Meinl, T.; Ohl, P.; Sieb, C.; Thiel, K.; Wiswedel, B.. *KNIME: The Konstanz Information Miner*. 2007.
(104) Walters, W. P.; Murcko, M. A. Prediction of "Drug-Likeness". *Adv. Drug Delivery Rev.* **2002**, *54*, 255−271.
(105) Baell, J. B.; Holloway, G. A. New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *J. Med. Chem.* **2010**, *53*, 2719−2740.
(106) Shelley, J. C.; Cholleti, A.; Frye, L. L.; Greenwood, J. R.; Timlin, M. R.; Uchimaya, M. Epik: A Software Program for PK<Subscrip-

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656

t><Emphasis Type = "Italic">a</Emphasis></Subscript> Prediction and Protonation State Generation for Drug-like Molecules. *J. Comput.-Aided Mol. Des.* **2007**, *21*, 681−691.
(107) Greenwood, J. R.; Calkins, D.; Sullivan, A. P.; Shelley, J. C. Towards the Comprehensive, Rapid, and Accurate Prediction of the Favorable Tautomeric States of Drug-like Molecules in Aqueous Solution. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 591−604.
(108) Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra Precision Glide: Docking and Scoring Incorporating a Model of Hydrophobic Enclosure for Protein−Ligand Complexes. *J. Med. Chem.* **2006**, *49*, 6177−6196.
(109) Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; et al. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J. Med. Chem.* **2004**, *47*, 1739−1749.
(110) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 2. Enrichment Factors in Database Screening. *J. Med. Chem.* **2004**, *47*, 1750−1759.
(111) Sander, T.; Freyss, J.; von Korff, M.; Rufener, C. DataWarrior: An Open-Source Program For Chemistry Aware Data Visualization And Analysis. *J. Chem. Inf. Model.* **2015**, *55*, 460−473.
(112) Li, C.; Wen, A.; Shen, B.; Lu, J.; Huang, Y.; Chang, Y. FastCloning: A Highly Simplified, Purification-Free, Sequence- and Ligation-Independent PCR Cloning Method. *BMC Biotechnol.* **2011**, *11*, 92.
(113) Tse, W. C.; Boger, D. L. A Fluorescent Intercalator Displacement Assay for Establishing DNA Binding Selectivity and Affinity. *Acc. Chem. Res.* **2004**, *37*, 61−69.
(114) Hajduk, P. J.; Olejniczak, E. T.; Fesik, S. W. One-Dimensional Relaxation- and Diffusion-Edited NMR Methods for Screening Compounds That Bind to Macromolecules. *J. Am. Chem. Soc.* **1997**, *119*, 12257−12261.
(115) Mayer, M.; Meyer, B. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chem., Int. Ed. 38* (), 1784−1788. DOI: 10.1002/(SICI)1521-3773(19990614)38:12<1784::AID-ANIE1784>3.0.CO;2-Q.

DOI: 10.1021/acsomega.9b00162
ACS Omega 2019, 4, 11642−11656