Thiopurine Derivative-Induced Fpg/Nei DNA Glycosylase Inhibition: Structural, Dynamic and Functional Insights

Charlotte Rieux, Stéphane Goffinont, Franck Coste, Zahira Tber, Julien Cros, Vincent Roy, Martine Guerin, Virginie Gaudon, Stéphane Bourg, Artur Biela, et al.

International Journal of
Molecular Sciences

MDPI

Article

Thiopurine Derivative-Induced Fpg/Nei DNA Glycosylase Inhibition: Structural, Dynamic and Functional Insights

Charlotte Rieux [1,†], Stéphane Goffinont [1,†], Franck Coste [1], Zahira Tber [2], Julien Cros [1], Vincent Roy [2,3,*], Martine Guérin [1,3], Virginie Gaudon [1], Stéphane Bourg [2], Artur Biela [1,‡], Vincent Aucagne [1], Luigi Agrofoglio [2,3], Norbert Garnier [1,3,*] and Bertrand Castaing [1,*]

1 Centre de Biophysique Moléculaire, UPR4301 CNRS, rue Charles Sadron, CEDEX 2, F-45071 Orléans, France; charlotte.rieux44@yahoo.fr (C.R.); stephane.goffinont@cnrs-orleans.fr (S.G.); franck.coste@cnrs-orleans.fr (F.C.); julien.cros@cnrs-orleans.fr (J.C.); martine.guerin@univ-orleans.fr (M.G.); virginie.gaudon@cnrs-orleans.fr (V.G.); biela.artur@gmail.com (A.B.); vincent.aucagne@cnrs-orleans.fr (V.A.)

2 Institut de Chimie Organique et Analytique, UMR7311 CNRS-Orleans University, Université d'Orléans, Pôle de Chimie, rue de Chartres, F-45100 Orléans, France; tber.zahira@yahoo.com (Z.T.); stephane.bourg@cnrs-orleans.fr (S.B.); luigi.agrofoglio@univ-orleans.fr (L.A.)

3 Université d'Orléans, UFR Sciences et Techniques, rue de Chartres, 45100 Orléans, France

* Correspondence: vincent.roy@univ-orleans.fr (V.R.); norbert.garnier@cnrs-orleans.fr (N.G.); bertrand.castaing@cnrs-orleans.fr (B.C.)

† These authors contributed equally to this work.

‡ Present address: Department of Cell Biology and Imaging, Institute of Zoology and Biomedical Research, Jagiellonian University, 9 Gronostajowa Street, PL30387 Kraków, Poland.

Received: 29 February 2020; Accepted: 14 March 2020; Published: 17 March 2020

Abstract: DNA glycosylases are emerging as relevant pharmacological targets in inflammation, cancer and neurodegenerative diseases. Consequently, the search for inhibitors of these enzymes has become a very active research field. As a continuation of previous work that showed that 2-thioxanthine (2TX) is an irreversible inhibitor of zinc finger (ZnF)-containing Fpg/Nei DNA glycosylases, we designed and synthesized a mini-library of 2TX-derivatives (TXn) and evaluated their ability to inhibit Fpg/Nei enzymes. Among forty compounds, four TXn were better inhibitors than 2TX for Fpg. Unexpectedly, but very interestingly, two dithiolated derivatives more selectively and efficiently inhibit the zincless finger (ZnLF)-containing enzymes (human and mimivirus Neil1 DNA glycosylases hNeil1 and MvNei1, respectively). By combining chemistry, biochemistry, mass spectrometry, blind and flexible docking and X-ray structure analysis, we localized new TXn binding sites on Fpg/Nei enzymes. This endeavor allowed us to decipher at the atomic level the mode of action for the best TXn inhibitors on the ZnF-containing enzymes. We discovered an original inhibition mechanism for the ZnLF-containing Fpg/Nei DNA glycosylases by disulfide cyclic trimeric forms of dithiopurines. This work paves the way for the design and synthesis of a new structural class of inhibitors for selective pharmacological targeting of hNeil1 in cancer and neurodegenerative diseases.

Keywords: BER; DNA glycosylase; Fpg/Nei; hNeil1; disulfide; cyclophane; DNA repair inhibitors; zinc finger oxidation

1. Introduction

DNA constitutive elements are continually subjected to the deleterious effects of physical and chemical agents from endogenous and environmental sources. Induced DNA structural/chemical changes interfere with DNA transactions, such as replication, transcription and recombination [1].

Reactive oxygen species ($OH\cdot$, $O_2^{-}\cdot$, $^1O_2$, H_2O_2, etc.), resulting from cell respiratory metabolism and inflammatory processes or from water radiolysis by ionizing radiation, Fenton reaction or photo activation processes, are responsible for the formation of numerous oxidation/degradation products of nucleobases such as the miscoding abasic site and 8-oxoguanine [2–4]. Oxidized bases can mislead or block replication and transcription machinery and result in mutations or cell death. These structural DNA changes initiate inflammation, carcinogenesis and age-related neurodegenerative processes [5–8]. To counteract these adverse effects, organisms, bacteriophages and viruses have evolved numerous DNA repair strategies in which the basic principles have been conserved during evolution [9]. Among these strategies, the major way to repair oxidized bases is the base excision repair (BER) pathway [10]. BER is initiated by DNA glycosylases that recognize and remove base lesions. The resulting abasic (AP) site can be excised by the combined action of AP endonucleases, AP lyases and dRp lyases. These reactions lead to one or more nucleotide gap. Finally, a DNA polymerase and a DNA ligase cooperate to fill in the gap and restore the DNA integrity.

In addition to being involved in the maintenance of genetic material (DNA repair), DNA glycosylases are also key enzymes involved in a myriad of other physiological processes [11]. They participate in the maturation of immunoglobulin antigenicity (somatic hyper-mutagenesis [SM] and class-switch recombination [CSR] via uracil-DNA glycosylase), the maintenance of telomeres (via the DNA glycosylases Neil3) and in enzymatic mechanisms of active DNA demethylation (via the DNA glycosylases SMUG1, TDG and MBD4). Human DNA glycosylases, such as hOgg1 and hNeil1, are emerging as new pharmacological targets for small-molecule modulators, given their role in a wide range of physiological and/or possible pathological processes. In Huntington disease (HD), Ogg1 (and/or Neil1)-initiated repair of 8-oxoguanine (8-oxoG, or oxidized pyrimidines) in CAG triplets is proposed to trigger iterative oxidation–excision cycles that contribute to the somatic instability of the *huntingtin* gene, through a CAG repeat expansion [12–15]. Strikingly, somatic CAG repeat instability in HD is highest in the striatum, the tissue preferentially affected by the disease, and unbalanced BER enzyme activities seems to be responsible for the tissue-selectivity of the disease [13]. Thus, selective Ogg1/Neil1 inhibitors directed in the striatum might prevent CAG repeat expansion. In another example, a small interfering RNA (siRNA)-screening approach highlighted synthetic lethal interactions between the thymidylate synthase (TS) pathway and several human DNA glycosylases (hOgg1, hNeil1) in osteosarcoma cells [16]. In a more recent study, a new mechanism has been proposed to sustain proliferation in RAS transformed cells through increased BER capability [17]. In such a mechanism, RAS-transformed cells use hOgg1 stimulation to overcome the anti-proliferative effects of excessive oxidative DNA damage. All these observations may provide new therapeutic windows in cancer therapy that might be exploited with selective drugs that specifically target Ogg1 and Neil1.

While recent studies have demonstrated the relevance of the research to design innovative anticancer strategies, only a few reported the search for hOgg1 and hNeil1 inhibitors [18–21]. In previous work, we initiated this study on DNA glycosylases from the structural Fpg/Nei superfamily [18,22,23]. These enzymes recognize and excise oxidized bases in DNA by catalyzing the cleavage of the *N*-glycosydic bond between the damaged base and its associated sugar. These enzymes are bifunctional DNA glycosylase associated with an AP lyase activity that involves the successive cleavages of phosphodiester bonds at 3′ and 5′ sides of the resulting AP site by a βδ-elimination mechanism. Fpg/Nei DNA glycosylases are composed of two globular domains, an N-terminal domain rich in beta-structures and a C-terminal domain rich in alpha-structures. All enzymes display a helix-two turns-helix (H2TH) motif that is involved in DNA binding. Most of them contain an original zinc finger (ZnF) motif (defining a new class of ZnF) comprising a β-hairpin structure with generally four cysteine residues coordinating the Zn^{2+} ion (type $–C\text{-}X_{2(3)}\text{-}C_{(H)}\text{-}X_{16\text{-}18}\text{-}C\text{-}X_2\text{-}C\text{-}$) [24–26]. A small number of enzymes, such as hNeil1 and mvNei1, lack ZnF and, instead, have a very similar β-hairpin motif free of zinc (called zincless finger, [ZnLF]) [27,28]. Site directed mutagenesis and crystal structures of enzymes bound to damaged DNA show that ZnF and ZnLF play exactly the same role in DNA binding. Although structurally related and displaying a slight overlap in substrate specificity, the enzymes from

the Fpg subfamily remove oxidized purines, whereas those from the Nei subfamily are more specific for oxidized pyrimidines. Only the bacterial Fpg enzymes can excise 8-oxoG, the major oxidation product of purines. Thus, they are considered to be the functional homologs of the structurally unrelated eukaryote 8-oxoguanine-DNA glycosylase 1 (Ogg1) [29].

In an effort to exploit the extrahelical base lesion recognition mechanism used by these enzymes for removing oxidized bases, 2-thioxanthine (2TX) was identified among other purine and pyrimidine derivatives as an inhibitor of the bacterial Fpg and Nei enzymes and human hNeil2 [18,30]. Surprisingly, enzyme kinetic experiments with the *Escherichia coli* Fpg protein proposed an uncompetitive inhibition mode. In other words, the effective inhibitor target is probably not the active site of the enzyme. According to the uncompetitive inhibition mode, 2TX only binds the enzyme/substrate complex. This interaction is favored by prior binding of the enzyme to its DNA substrate. In fact, we demonstrated that both free and bound enzymes are targets for 2TX, with a slight preference for the bound enzyme (compatible with mixed inhibition rather than an uncompetitive or non-competitive inhibition). Studies in solution coupled with crystal structure analysis revealed that two ZnF cysteine residues are possible targets for 2TX. This effect results in the loss of zinc (observed both in solution and in crystal structures), the covalent attachment of 2TX to cysteine by a disulfide bond and, thus, the irreversible inhibition of the enzyme. Other 2TX enzyme target sites, however, are not excluded, but the irreversible character of the inhibition at a high 2TX concentration compromises the correct interpretation of enzymatic kinetics data. Although the ZnF oxidation mechanism mediated by 2TX remains unclear, it does explain why hNei1, which lacks a ZnF, is resistant to 2TX and why a strong disulfide reducer, such as tris(2-carboxyethyl)phosphine hydrochloride (TCEP), protects the ZnF-containing enzymes from the 2TX inhibitory effect [18].

In this work, we synthetized a small library of 2TX derivatives and evaluated their effects on bacterial LlFpg (from *Lactococcus lactis*) and eukaryotic Neil1 models (the human hNeil1 and mimivirus mvNei1 DNA glycosylases). We identified inhibitors that are more efficient than 2TX for the ZnF-containing Fpg/Nei DNA glycosylases. Surprisingly, some of them inhibited ZnLF-containing Nei-like proteins, mainly hNeil1 and mvNei1. Based on chemical, biochemical, mass spectrometry, computational and structural approaches, we decipher at the atomic level the mechanism of action of best inhibitors of the ZnF-containing enzymes (LlFpg) and propose a possible inhibitory mechanism for the enzymes of the Fpg/Nei superfamily that do not contain ZnF.

2. Results and Discussion

2.1. Inhibition of the Bacterial Fpg Activity by 2TX Derivatives

The compound 2TX was identified in a DNA base analog screening as an uncompetitive inhibitor of the *E. coli* formamidopyrimidine-DNA glycosylase (EcFpg) [30]. We confirmed the inhibitory effect of 2TX on ZnF-containing enzymes from the Fpg/Nei DNA glycosylase structural superfamily (including LlFpg, EcNei and hNeil2) [18]. Although the precise mode of action of 2TX remains to be clarified, we established in solution and by X-ray analysis that—unexpectedly—2TX chemically and selectively targets the two most exposed cysteine residues of the ZnF in these enzymes. Consequently, 2TX covalently attaches to cysteine through a disulfide bond, and the zinc ion is released [18]. In order to find more selective and efficient inhibitors, and to clarify the inactivation mode through the thiol/thione group, we prepared a mini-library of 2TX-derivatives (TXn) (see Supplementary Information for their synthesis and Figure S1 for their structures).

TXn were screened for their ability to inactivate the 8-oxoG-DNA glycosylase/AP lyase activity of LlFpg (our Fpg model for X-ray structure investigations). Some of these compounds are thiol-free and the others are monothiolated or dithiolated compounds (Figure S1). As expected, the compounds without the thiol/thione group were unable to efficiently inhibit the excision of 8-oxoG-containing DNA by LlFpg (Figure 1a). However, the presence of a thiol/thione group on the tested compound seemed insufficient to inhibit the enzyme. Indeed, 40 μM of some thio-compounds, such as TX15,

TX20, TX21 and TX22, had a very low or unmeasurable effect on the Fpg 8-oxoG-DNA glycosylase/AP lyase activity (Figure 1a). The apparent half-maximal inhibitory concentration ($IC50_{app}{}^{A}$) for the best compounds were obtained from dose-response curves (Figure 1b and Table 1).

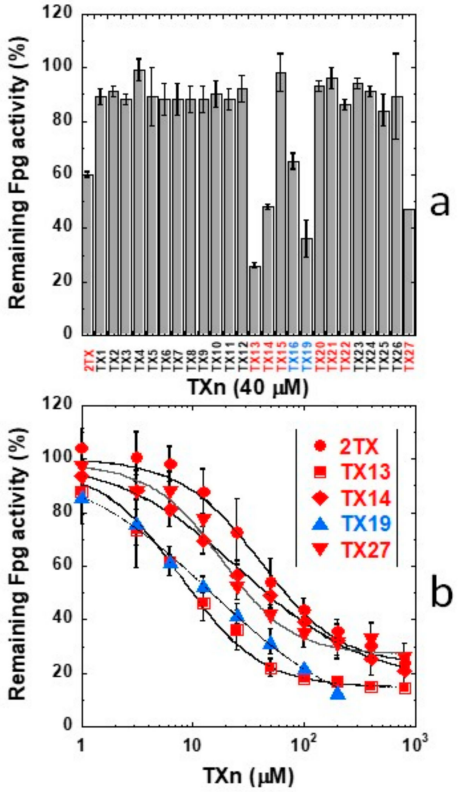
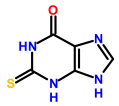
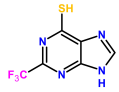
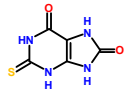
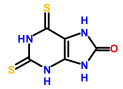
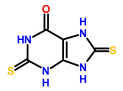
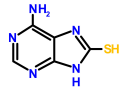


Figure 1. Effect of 2-thioxanthine (2TX)-derivatives on Fpg activity. A total of 20 nM of 5′-[^{32}P]-8-oxoG-containing 24-bp DNA duplex (labeled on the damaged strand) was incubated with 2 nM of LlFpg alone (dimethyl sulfoxide (DMSO) control) or with the indicated 2TX and 2TX-derivatives (TXn) concentrations, as described in the "Materials and Methods" section. (**a**) Screening of TXn for Fpg inhibition. Abbreviations and structures of TXn are presented in Figure S1. Averages and standard deviations were obtained from at least three independent experiments. Mono- and di-thiolated compounds are highlighted in red and blue, respectively. (**b**) Dose-response curves for best inhibitors. Each data point represents the mean value obtained from at least three independent measurements. Half-maximal inhibitory inhibition (IC50) values ($IC50_{app}{}^{A}$) of the more efficient TXn are indicated in Table 1 and were extracted from these data using Origin software.

The two tested dithio-compounds TX16 and TX19 inhibited Fpg in the same range of magnitude as efficient monothio-compounds. This finding indicates a non-cumulative effect of thiol functions for Fpg inhibition (Table 1). The best inhibitor, TX13 (2-trifluoromethyl-6-thiopurine), is a monothio-compound; it was seven times more effective than 2TX, with an $IC50_{app}{}^{A}$ of 7.5 µM (Table 1). The 2- trifluoromethyl group of TX13 might be responsible for a better interaction of the compound with Fpg because its replacement by a methyl group in TX4 drastically decreased the inhibitory effect despite the presence of the thiol/thione group (Figure 1a). The position of the thio-group (thione/thiol) does not appear essential for inhibition and can occupy positions 2, 6, 8, 2 and 6, or, 2 and 8 on the purine aromatic ring of TXn effective inhibitors. As already shown for 2TX [18], TCEP (a strong reducer of disulfide bonds) counteracted the inhibitory effect of TXn (Figure S2).

Table 1. Effect of TXn on Fpg/Nei catalytic and DNA binding activities.

Compound		$IC50_{app}$ [A] (μM) (Fpg/Nei enzyme catalysis)			$IC50_{app}$ [B] (μM) (Fpg DNA binding) *	
		LlFpg	*hNEIL1*	*mvNEI1*	*C1*	*C2*
2TX *		48 ± 4	>500	>500	113.3 ± 17 (453)	NA (9.7)
TX13 *		7.5 ± 1.0	>250	>500	6.1 ± *0.5* (98)	NA (1.3)
TX14 *		28 ± 7	>500	>500	25.1±2.4 (100)	NA (2.8)
TX16 **		41 ± 7	21 ± 1	14 ± 1	ND	NA
TX19 **		15 ± 1	19 ± 1	124 ± 4	17.1 ± 2.4 (68)	NA (2)
TX27 *		20 ± 2	160 ± 16	277 ± 18	44.19 ± 8.0 (177)	NA (4)

$IC50_{app}$ corresponds to the half-maximal inhibitory concentration of a compound with ([A]) and ([B]) being relative to the catalytic activity at 37 °C and the DNA binding activity at 4 °C, respectively. ND and NA indicate not-determined and not-applicable, respectively. * Determined with 0.1 nM of 14-mer [THF:C] DNA duplex and 0.25 and 10 nM of Fpg for C1 and C2, respectively. C1 corresponds to the specific lesion recognition complex (lesion recognition complex (LRC), Fpg/DNA = 1/1), and C2 corresponds to the non-specific binding of a second Fpg molecule to the preformed C1 (Fpg/DNA = 2/1; see text below). The numbers in parentheses indicate the number of TXn equivalents to be added to one protein equivalent to observe 50% inhibition of DNA binding activity.

The purine moiety of efficient TXn compounds seems to be crucial because the thio-compound TX21 did not inhibit Fpg (Figure 1a and Figure S1). Considering that Fpg repairs oxidized purines, this last observation suggests that TXn might interact with the enzyme active site. Kinetic experiments with LlFpg in the presence of 2TX partially support this last hypothesis (Figure S3). Although LlFpg was inhibited by 2TX, as was its *E. coli* homolog (EcFpg), the raw kinetic data from all the experiments appeared to be very complicated and difficult to interpret. This outcome suggests that several inhibitory mechanisms concomitantly contribute to the final inhibitory process. In previous work, we identified an irreversible inhibitory mechanism that resulted in the oxidation of the ZnF in the protein. In this case, no inhibitory constant (Ki) can be determined. This mechanism (also supported by the inhibition suppression by TCEP, Figure S2) is likely superimposed on other inhibitory mechanisms, a factor that makes it difficult to analyze kinetic data. To clarify, we separately analyzed kinetic data obtained for low and high 2TX concentrations (Figure S3). As expected, we observed a linear dependence of the reciprocal velocity (1/v) as a function of the reciprocal substrate concentrations (1/S) with and without the inhibitor. The kinetic constant values for the excision of 8-oxoG by LlFpg were K_m = 5.60 nM, V_{max} = 0.34 fmol.min^{-1} and k_{cat} = 0.09 min^{-1}. These values compare well with those determined previously for the excision of FapyG and 8-oxoG by EcFpg [31–33]. At low 2TX concentrations, the inhibitory process appears partially competitive, with an apparent K_i of 2.19 ± 0.56 μM (K_{ic}). This value is consistent with an interaction of 2TX in the active site of the enzyme. This inhibitory mechanism has not been previously described for the EcFpg enzyme [30]. At high 2TX concentrations, the inhibition displays a clear uncompetitive mechanism (decrease of V_M and K_M, their ratio remaining relatively invariant) with an apparent K_i of 24.88 ± 6.56 μM (K_{iu}) as already

observed with the EcFpg [30]. This mode of inhibition is in agreement with the possibility that 2TX can bind outside the active site on the free and bound enzyme. We will later discuss that the difficulty in interpreting the enzymatic kinetics data in the presence of 2TX (and TXn) can also be explained by the different redox equilibrium states of these compounds in dimethyl sulfoxide (DMSO), which cannot be easily controlled and adds additional complexity. The loss of the inhibition in the presence of TCEP also suggests that the reduced forms of 2TX and TXn are weakly or not effective in inhibiting the enzyme (see below).

2.2. Inhibition of the Bacterial Fpg DNA Binding Activity by 2TX and TXn

To complement this study, we examined the effect of TXn on the binding of Fpg to a short DNA duplex containing a tetrahydrofuran (THF), a non-cleavable AP site analog. The enzyme specifically recognizes THF and forms a stable abortive complex with DNA [18,34–36]. Qualitative and quantitative binding experiments were performed by electrophoretic mobility shift assay (EMSA), and DNA binding conditions were optimized for 8% DMSO (required for TXn dissolution but did not negatively affect enzyme stability and activity) and 0.6% bovine serum albumin (BSA; required to avoid an enzyme dilution negative effect and to obtain reproducible results; Figure S4). A representative titration experiment (with and without TCEP) of the 14-mer [THF:C] DNA duplex by LlFpg is shown in Figure 2a. Fpg formed two stable complexes with this short DNA probe that were easily observed by EMSA: a high-affinity complex C1 corresponding to the lesion recognition complex (LRC), with nanomolar affinity [34,35,37], and a complex C2 only observed at higher Fpg concentrations. C2 results in C1 titration by a second Fpg molecule. This finding illustrates that the crystal structure of Fpg bound to 8-oxoG-carbanucleotide-containing DNA [38] (Figure S5). The stability of each preformed complex was then examined by EMSA in the presence of 2TX or TXn. In these experiments, the choice to pre-form the complex before adding 2TX or TXn in the reaction mixture was based on the observation that pre-incubation of Fpg for 20 min at 4 °C with 2TX or TXn inhibitors in a concentration range varying from micromolar to millimolar completely abolished the enzyme's ability to bind to DNA. This data unambiguously indicates that 2TX and TXn effective inhibitors can interact with the enzyme even in the absence of DNA. This result is consistent with the non-competitive inhibition of Fpg by 2TX. This mechanism was also partially suggested by enzymatic kinetic studies at a high concentration of inhibitors and by previous work (Figure S3) [18]. Indeed, this mechanism suggests that 2TX can inactivate the enzyme by indifferently interacting with the enzyme alone or with the enzyme/substrate complex outside the enzyme active site. The enzyme inactivation mediated by 2TX through ZnF oxidation that directly affects the DNA binding properties of the enzyme is also in agreement with this last observation [18]. Dose-response curves of preformed C1 in the presence of increasing thio-compound concentrations, which we showed effectively inhibit Fpg activity, induced the dissociation of the Fpg/DNA lesion recognition complex C1 (Figure 2b).

The inhibitor concentrations that induced the loss of the half-maximal binding ($IC50_{app}{}^{B}$) were determined from these curves and are reported in Table 1. $IC50_{app}{}^{B}$ compared well with $IC50_{app}{}^{A}$. As observed above for the enzyme activity, strong reducers, such as TCEP, abolished the negative effects of 2TX and TXn on Fpg/DNA complex stability and formation. These findings again indicate that inhibition by thio-compounds essentially relies on the oxidation of the Fpg ZnF (for an example, see experiment with TX19, Figure S4). Surprisingly—and for high concentrations of TXn (above 0.1 mM)—the effect of some TXn on the stability of the complex C2 (preformed complex before treatment) in the presence of TCEP was highly variable. For a [1/1] molar ratio between TCEP and TX13 or TX27, C2 was strongly destabilized, but C1 remained stable (Figure S6). Under the same conditions, 2TX, TX14, TX15 and TX19 had a moderate effect on C2. This troubling and unexpected observation led us to evaluate this effect with the TXn small library compounds. To prevent a side effect due to a possible partial reduction by TCEP, we performed EMSA in the presence of a 6.7-fold molar excess of TCEP with respect to 2TX and TXn after complex formation (Figure 3).

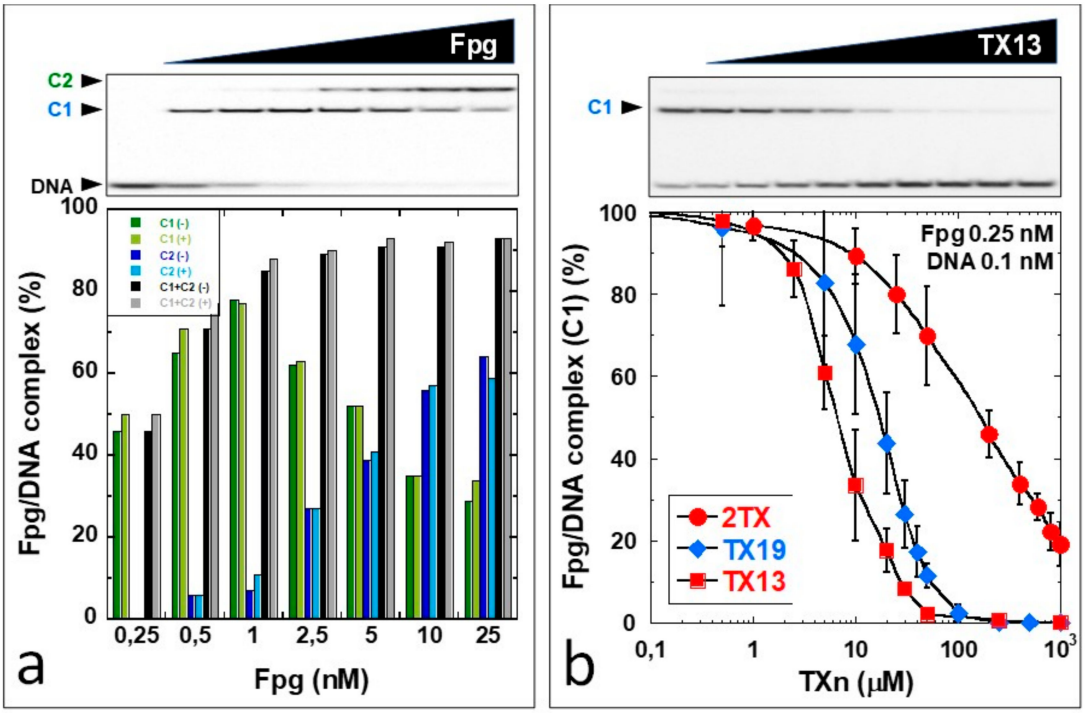


Figure 2. The effect of 2TX and TXn on Fpg DNA binding to abasic site-containing DNA. A 14-mer DNA duplex containing a tetrahydrofurane (THF) site as an abasic site analog was incubated with a limiting Fpg concentration (0.25 nM) with and without increasing concentrations of TXn. Incubation mixtures were then analyzed by electrophoretic mobility shift assay (EMSA), as described in the 'Materials and Methods'. (**a**) Titration experiments with (+) and without (-) 2 mM tris(2-carboxyethyl)phosphine hydrochloride *(TCEP)* C1 and C2 are for the [1/1] and [2/1] Fpg/DNA complexes, respectively (see detailed in the text). (**b**) Fpg DNA binding concentration-dependence for 2TX and the more efficient TXn inhibitors. A representative EMSA autoradiography with 0.25 nM of Fpg alone (lane no) and 0.5, 2.5, 5, 10, 20, 40, 50, 250, 1000 μM is shown in the upper panel. $IC50app^B$ corresponds to the TXn amount needed to halve the C1 complex formation and was extracted from the dose-response curves (lower panel; Table 1).

As expected in the presence of TCEP, thiol/thione function in TXn is not an absolute criterion because among the best C2 destabilizers, we found compounds that lack thiol group, including TX10 and TX11. On the contrary, the presence of an alkyl group (trifluoromethyl and methyl groups) at position two of the purine derivatives (TX6, TX7, TX10, TX11 and TX12, Figure S1) was favorable for destabilizing C2. These data suggest that some TXn compounds interact with Fpg (here LlFpg) with an unknown enzyme binding site. This finding results in incompetency of a second Fpg molecule to form the C2 complex with the 14-mer [THF:C] DNA duplex. Obviously, the binding of these compounds to this exosite had a limited effect on the stability of the Fpg molecule bound to the lesion (LRC, C1; Figure S5a). The crystal structure of LlFpg bound to 14-mer [8-oxoG:C] DNA duplex (which differs from that used in this study only by the damaged nature present in the duplex: an 8-oxoG lesion in place of THF) revealed that the second Fpg molecule can bind to the overhanging base positioned at the 5′ end of the damaged strand (Figure S5b) [38]. In the structure, the first Fpg molecule (Mol-1 in Figure S5a) already positioned on the base lesion, does not physically interact with the second one (Mol-2 in Figure S5a). Combining these biochemical and structural observations, we propose that some TXn can selectively bind Fpg outside the active site, at an exosite. This action impairs its ability to bind to the overhanging base at the DNA duplex end. According to this hypothesis, some TXn can mimic the overhanging base of the DNA duplex and displace it, a phenomenon that destabilizes C2. When a high affinity TXn occupies the exosite, Fpg appears unable to form the C2 complex. Molecular docking experiments and crystal structures presented below will confirm this hypothesis.

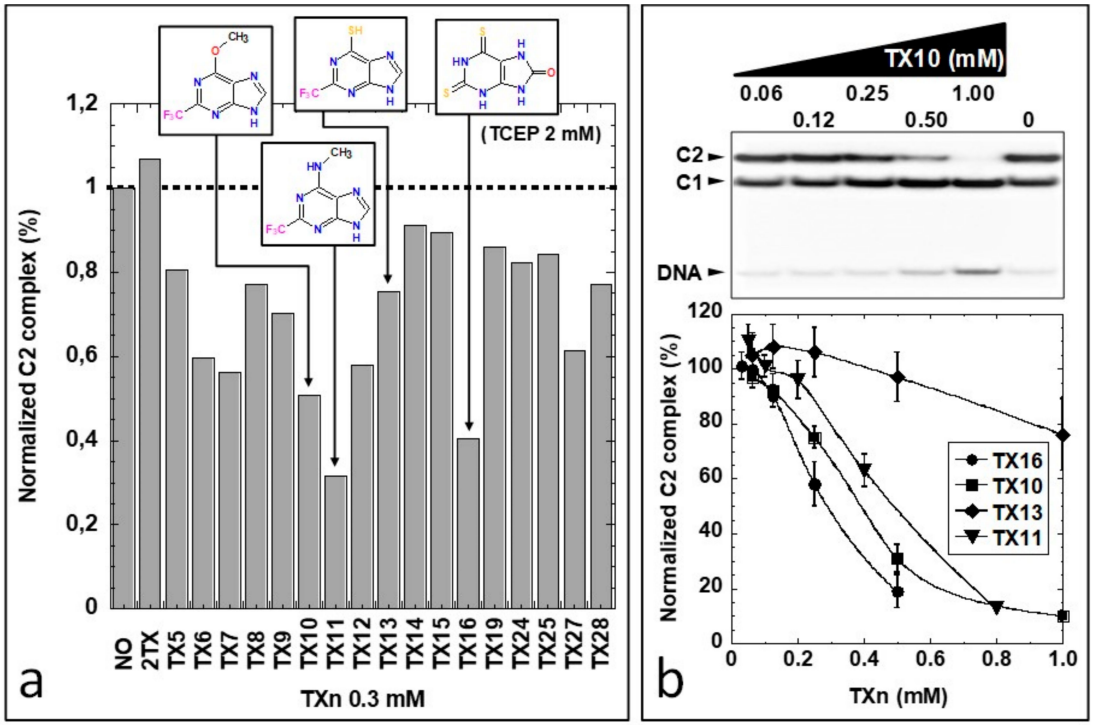


Figure 3. Impact of TXn on the non-specific Fpg/DNA complex C2 stability in highly reducing conditions. EMSA was performed as described in the 'Material & Methods', except TCEP was added at a 2 mM final concentration in each binding assay. (**a**) Representative experiment at 0.3 mM 2TX and more efficient TXn; (**b**) Dose-response with the best TXn. Upper panel: Representative EMSA experiment with TX10. Lower panel: Dose-response curves with the indicated TXn.

2.3. Selective Inhibition of ZnLF-Containing Fpg/Nei DNA Glycosylases by Dithio-TXn

In previous work, we showed that 2TX inhibits ZnF-containing Fpg/Nei DNA glycosylases (EcFpg, EcNei, LlFpg and hNeil2) [18,30]. At high concentrations, this inhibition essentially results in ZnF oxidation associated with zinc release because the zincless (ZnLF)-containing enzyme appears resistant to 2TX. To complement this present study, we evaluated the TXn effect on the catalytic and DNA binding activities of two ZnLF enzymes: hNeil1 and mvNei1. Regarding the capacity to oxidize cysteine by the thiolated TXn, the two proteins are good complementary models: hNeil1 displays eight free cysteine residues whereas mvNei1 lacks cysteine (note that LlFpg only contains four cysteine residues, all involved in Zn^{2+} ion coordination in the ZnF motif). As anticipated, 2TX and all tested monothiolated TXn compounds did not significantly affect hNeil1 and mvNei1, as opposed to the effect on LlFpg (Figure 4).

Interestingly and unexpectedly, the two dithiolated TXn (TX16 and TX19) evaluated in the study displayed apparent inhibitory effects on both the catalytic and DNA binding activities of hNeil1 and mvNei1. One of the two thiol/thione groups substituted position two of the purine ring and the second substituted position six or eight of the purine analogs (TX16 and TX19, respectively). Although the monothiolated TXn were ineffective on these enzymes, hNeil1 inhibition might be attributed to the potential chemical reactivity of free cysteine residues but not for mvNei1, which is completely devoid of cysteine. The best inhibitor was TX16, with an $IC50app^{A}$ of 21 and 14 μM for hNeil1 and mvNei1, respectively (Table 1).

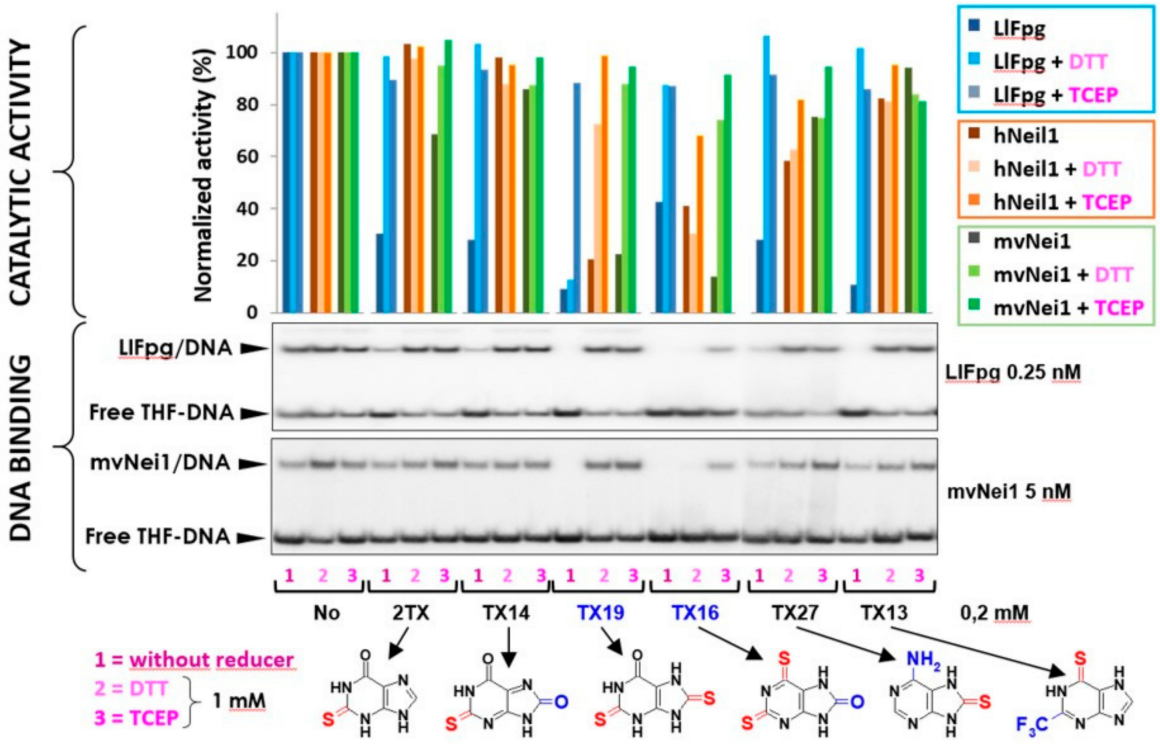


Figure 4. Comparative analysis of the effect of 2TX and TXn on Fpg/Nei DNA glycosylases. Enzyme activities were measured as described in the Figure 1 legend using 0.1 nM of a 5′-[^{32}P]-24-mer DNA duplex containing either an 8-oxoG opposite C for LlFpg or a dihydroxyuracil (DHU) opposite G for hNeil1 and mvNei1. The DNA probe was incubated with the indicated protein concentration and 0.2 mM TXn compounds in the presence or absence of 1 mM dithiothreitol (DTT) and TCEP (a representative experiment is illustrated by the upper histograms). The DNA probe for EMSA was 0.1 nM of 5′-[^{32}P]-24-mer DNA duplex containing THF with LlFpg and mvNei1 (bottom panels: gel autoradiography) with the indicated conditions and as described in 'Materials and Methods'.

Based on our knowledge, TX16 is among the best inhibitors of ZnLF-containing Fpg/Nei enzymes discovered to date. In previous work, using a high-throughput screen of a small molecule library, the Lloyd group identified purine analogs with an IC50 of 10–30 μM for hNeil1 (determined by a radioassay similar to our assay) [20]. Although it is difficult to conclude the TX16 and TX19 mechanism of inhibition of hNeil1 inhibition, mvNei1 inhibition cannot be attributed to an attack of a cysteine. The simplest hypothesis to explain this result is that TX16 (TX19) interacts with mvNei1 (hNeil1) either at the active site and/or another site(s) to impair the ability of the enzyme to bind and/or metabolize its DNA substrate. This mode of inhibition might also exist with LlFpg, a hypothesis that would explain why it was difficult to extract from enzymatic kinetics a precise mode of inhibition because several mechanisms of inhibition contribute to the final inhibition observed (Figure S3).

The most astonishing finding in this study lies in the observation that strong reducers, such as TCEP and to a lesser extent dithiothreitol (DTT), suppressed the inhibition of the hNeil1 and mvNei1 DNA binding and catalytic activities mediated by TX16 and TX19 (Figure 4). If we consider that TCEP is the more powerful reducing agent to break disulfide bond (as compared to DTT) and that the cysteine-free enzyme mvNei1 was inhibited by TX16 and TX19, the simplest hypothesis to explain this last observation is to propose that the reducer has a direct action on enzymes (protein oxidation protection) as well as on TX16 and TX19. According to this proposition, we suggest that disulfide forms (and not other possible oxidized forms) of TX16 and TX19 are potentially responsible for the inhibition of hNeil1 and mvNei1. Due to the presence of two thiol/thione groups in TX16 and TX19, we cannot exclude the possibility of having many oligomeric forms of these compounds in our stock solutions.

2.4. Molecular Mechanisms for Fpg/Nei Inhibition by Disulfide forms of 2TX and TXn

The simplest way to evaluate the molecular homogeneity of the 2TX and TXn stock DMSO solutions used in this study is content analysis by mass spectrometry (MS). Given that direct analysis of samples by matrix-assisted laser desorption/ionization-time of flight (MALDI/TOF) or electrospray ionization (ESI) MS is impossible due to the presence of DMSO, we used high-resolution ESI-MS coupled with liquid chromatography separation (LC-HRMS). LC-HRMS analysis was performed with stock solutions of the monothio 2TX and TX14 and the dithio TX19 (Figure S7). The 2TX and TX14 stock solutions unambiguously contained two major compounds: one having the expected mass of their reduced forms (Figure 5a, $2TX_{Red}$, $TX14_{Red}$) and a second one with a mass consistent with their corresponding disulfide dimeric forms (Figure 5a, $2TX_{Ox}$ and $TX14_{Ox}$). The LC-HRMS spectrum of TX19 displays a more complex LC-profile, a finding that indicates a greater heterogeneity of this compound in DMSO. The molecular species eluted in peaks one and two may correspond to a reduced form of TX19 that lost one sulfur atom (Δmass= −17 u.m.). The compound eluted in peak three corresponds to the reduced form of TX19 (Figure 5a, $TX19_{Red}$). Two other compounds were later eluted separately later and display the same molecular mass expected for the two cyclic trimeric disulfide isomers of TX19 (Figure 5a, $TX19_{Ox1}$ and $TX19_{Ox2}$).

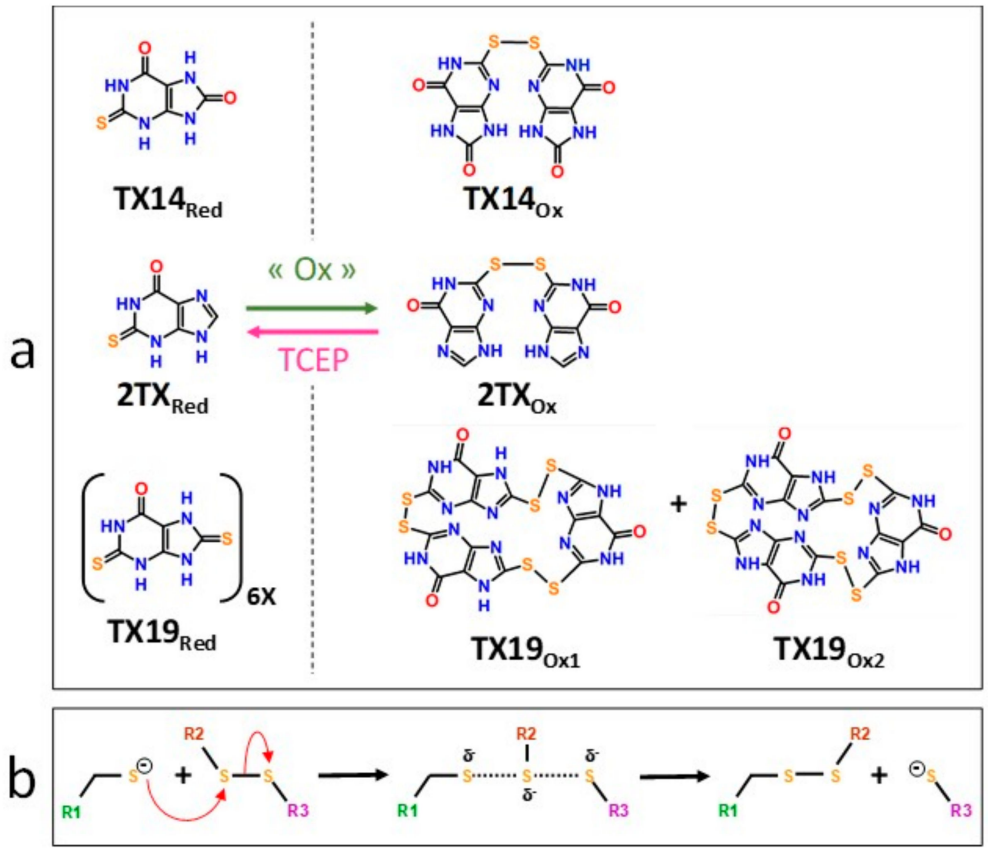


Figure 5. Disulfide forms of 2TX, TX14 and TX19 are responsible for the inhibition of Fpg/Nei DNA glycosylases. (**a**) Disulfide forms of 2TX, TX14 and TX19 identified by LC-HRMS (see the mass spectra in Figure S7). (**b**) Proposed thiol/disulfide exchange mechanism for the inhibition of zinc finger-containing Fpg/Nei enzymes by disulfide forms of 2TX, TX14 or TX19. R1 indicates cysteine thiolate-containing zinc finger Fpg/Nei enzyme, R2 and R3 belong to disulfide form of thio-compounds and can be identical or different.

The diversity of TX19 oxidized species present in the DMSO stock solution appears to be restricted to the two possible cyclic isomers of the trimeric disulfide species (no disulfide linear species, such as dimers, trimers and other oligomers; no other cyclic oligomeric disulfide oxidized species involving less

or more than three monomers). The possibility of having multiple redox states in 2TX and TXn stock solutions raises the question of which form(s) of the molecule is(are) responsible(s) for the Fpg/Nei DNA glycosylase inhibition we observed. Previous experiments strongly suggest the involvement of disulfide forms of 2TX and TXn in the inhibition of Fpg/Nei DNA glycosylases because TCEP completely abolished the inhibition of all the enzymes tested in this work (Figure 4).

To confirm this hypothesis, we established experimental procedures to prepare the homogeneous disulfide forms of the monothio-compounds 2TX and TX14 by using sodium periodate ($NaIO_4$), a more appropriate oxidant that we used to produce disulfide species rather than other oxidized species (Figures S8 and S9). To avoid DMSO-mediated oxidation, the dry powders of the homogeneous reduced and oxidized species were dissolved in 50 mM sodium phosphate, pH8.5. The effect of each homogeneous redox form of these compounds was then evaluated separately on the catalytic activity and DNA binding properties of LlFpg (Figure 6). Only the disulfide forms $2TX_{Ox}$ and $TX14_{Ox}$ inhibited Fpg catalytic activity, with an IC50 of 5.5 and 2.7 µM, respectively (right panel of Figure 6a,b). By contrast, 400 µM of $2TX_{Red}$ and $TX14_{Red}$ did not inhibit the enzyme (left panel, Figure 6a). The $IC50^{A}$ values for activity were approximatively 10 times greater than the $IC50_{app}{}^{A}$ determined with the redox mixtures in DMSO (Table 1). We drew similar conclusions with regard to the effect of disulfide of 2TX and TX14 forms on the DNA binding properties of LlFpg ($IC50^{B}$, Figure 6c). The $IC50^{B}$ obtained for homogeneous species was always 6 to 10 times greater than the $IC50^{A}$, data that suggest the catalytic activity and DNA binding experimental conditions were too different to be compared (incubation temperature, DNA probe and protein concentrations, etc.). Nevertheless, the two measured activities varied in the same direction. As observed above (Table 1), $TX14_{Ox}$ was twice as effective as $2TX_{Ox}$ for inhibiting LlFpg, data that again show the selectivity of TXn. As expected, given the results from experiments that use redox mixtures in DMSO, neither $2TX_{Ox}$ nor $TX14_{Ox}$ inhibited hNeil1 and mvNei1 (Figure 6d). Here again, there was a selectivity of monothio compounds for proteins with ZnF. The inhibition of LlFpg by 2TX and TX14 (and by other monothiol-compounds TX13 and TX27 and dithio-compounds TX16 and TX19) surely resulted from a thiol/disulfide exchange mechanism between thiolates of ZnF cysteine and the disulfide forms of 2TX and TXn (monothio- or dithio-compounds) (Figure 5b). Even if each cysteine thiolate of ZnF is potentially involved in the thiol/disulfide exchange mechanism, we previously showed that only the more exposed Cys248 and Cys268 of the LlFpg ZnF are susceptible to form a disulfide bond with 2TX [18]. However, we cannot completely exclude that a similar mechanism participates in the inhibition of hNeil1. Notably, we failed to observe TX16 and TX19 covalent adducts by mass spectrometry (MALDI-TOF) following a long incubation of hNeil1 with these compounds. The low reactivity of the free cysteine residues of hNeil1 might be linked to the fact that they are buried in the structure of the protein and, therefore, not accessible to the solvent and thus to TX16 and TX19 [27]. Obviously, a thiol/disulfide exchange mechanism is completely excluded with mvNei1, which has no cysteine. Despite many attempts, we failed to prepare the soluble and homogeneous forms $TX19_{Ox1}$ and $TX19_{Ox2}$. Although we have no doubt that the oxidized trimeric cyclic forms of TX19 are responsible for the inhibition of hNeil1 and mvNei1 proteins, we cannot provide experimental proof of such a process.

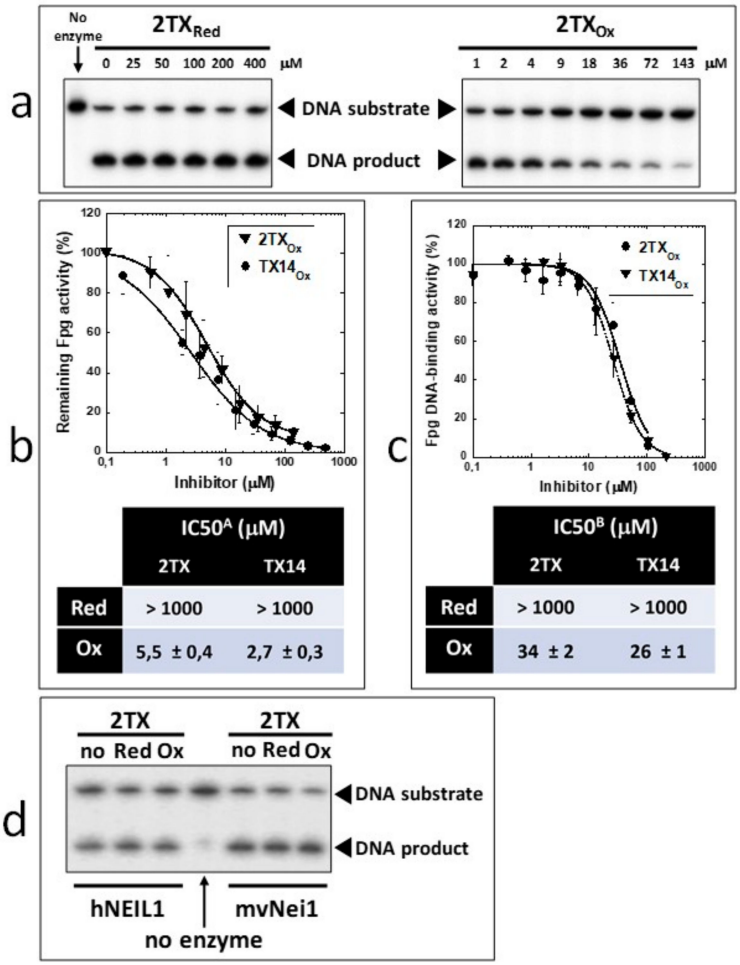


Figure 6. Comparative analysis of the effect of disulfide forms of 2TX and TX14 on the Fpg/Nei DNA glycosylases. Homogeneous reduced and disulfide forms of 2TX and TX14 were freshly prepared at 10 in 50 mM sodium phosphate (pH8.0). Catalytic and DNA binding activities of Fpg/Nei DNA glycosylases were performed as described in the 'Materials and Methods' using the radiolabeled [8-oxoG:C] and [DHU:C] DNA duplexes as substrate for the Fpg and hNeil1 (mvNei1) catalytic activity, respectively, and the 14-mer [THF:C] DNA duplex for DNA binding activity of all enzymes. IC50^{A} and IC50^{B} were determined from three independent experiments. (**a**) Representative autoradiography of 8-oxoG-DNA incubated with Fpg in the presence of the indicated 2TX$_{Red}$ and 2TX$_{Ox}$ concentrations. **(b)** and **(c)** dose-response curves and IC50 determination for Fpg catalysis and DNA binding, (**d**) the effect of 2TX$_{Red}$ and 2TX$_{Ox}$ on the hNeil1 and mvNei1 activity.

2.5. Structural Insights into the Mechanism of Fpg/Nei DNA Glycosylase Inhibition by 2TX and TXn

2.5.1. Proposed Binding Sites from Blind and Flexible Docking of Reduced and Oxidized Forms of 2TX and TX19 on Free and DNA Bound LlFpg and hNeil1

To exhaustively search for potential interaction sites by docking, we considered the intrinsic flexibility of receptors (free and DNA bound proteins) and ligands (conformations and tautomers) as described in Materials and Methods Section. Figure 7 presents an overview of the key steps of the blind and flexible docking procedure we used.

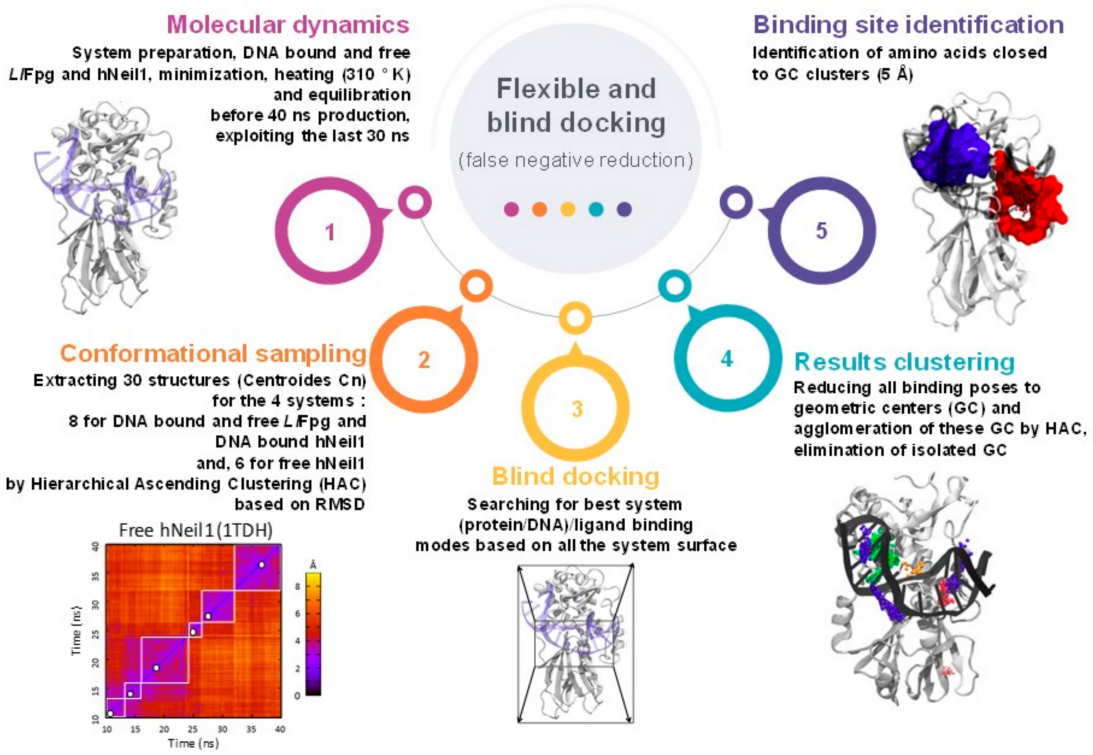


Figure 7. Flexible and blind docking strategy.

We performed docking experiments with $2TX_{Red/Ox}$ and $TX19_{Red/Ox1/Ox2}$ as ligands on free and DNA bound LlFpg and hNeil1 (for example, the docking of $TX19_{Ox}$ conformers is described in pictures, Figure S10). With $2TX_{Red}$ and $TX19_{Red}$, two Fpg/hNei1 binding sites were proposed: site I (including the active site) and site II, which mainly includes the two DNA binding motifs found in all Fpg/Nei DNA glycosylases (the ZnF or ZnLF loop that contains the strictly conserved R260 or R276 residue, respectively, and the H2TH motif, [29]) (Figure 8 and Figure S11). Binding site I is partially composed of amino acids that are directly involved in catalysis (such as the strictly conserved P1 and E2 residues) and partly in the lesion capping loop (LCL, the αF-β9 loop in which residues differ between Fpg and Nei subfamily; Figure S11, Table S1) [22,29,39]. The potential binding of the thio-compounds in the active site is consistent with the competitive inhibition mode suggested by Fpg kinetics experiments analyzed at low 2TX concentrations (see text above and Figure S3). The binding site II area is partly constituted by the strictly conserved LlFpg residues L162 and E163 localized in the helix αD of the H2TH motif and Q165 localized in the H2TH first turn (L165/M162, D167/D163 and Q168/Q164 of hNeil1/mvNei1, respectively; Figure S11). The strictly conserved arginine of the ZnF and ZnLF loop (R260 and R276 for LlFpg and hNeil1, respectively), which neutralizes the DNA phosphate groups located on both sides of the damage nucleoside, is often mobilized to interact with reduced forms of 2TX/TXn. When DNA is present in the structure, it seems that hydrophobic interactions and hydrogen bonds DNA also contribute to the binding of reduced and oxidized 2TX/TXn in site II (Figure S12). This phenomenon appears especially true for nucleotide phosphates localized at the 3′ side of the THF damaged nucleoside (see below). The zinc coordination sphere of the LlFpg ZnF was not identified as a $2TX_{red}/TX19_{red}$ binding site by docking experiments. The binding of $2TX_{red}$ and $TX19_{red}$ to site II (i.e., outside the active site) on free and DNA bound enzyme might be associated with uncompetitive or non-competitive inhibition, as suggested for LlFpg by kinetics experiments analyzed at high 2TX concentrations (Figure S3). However, the biochemical data about the effect of 2TX/TX19 in the presence of TCEP (i.e., reduced species) is not consistent with this hypothesis. This discrepancy suggests that binding of $2TX_{red}/TX19_{red}$ to sites I and/or II would not be associated with Fpg/Nei enzyme inhibition.

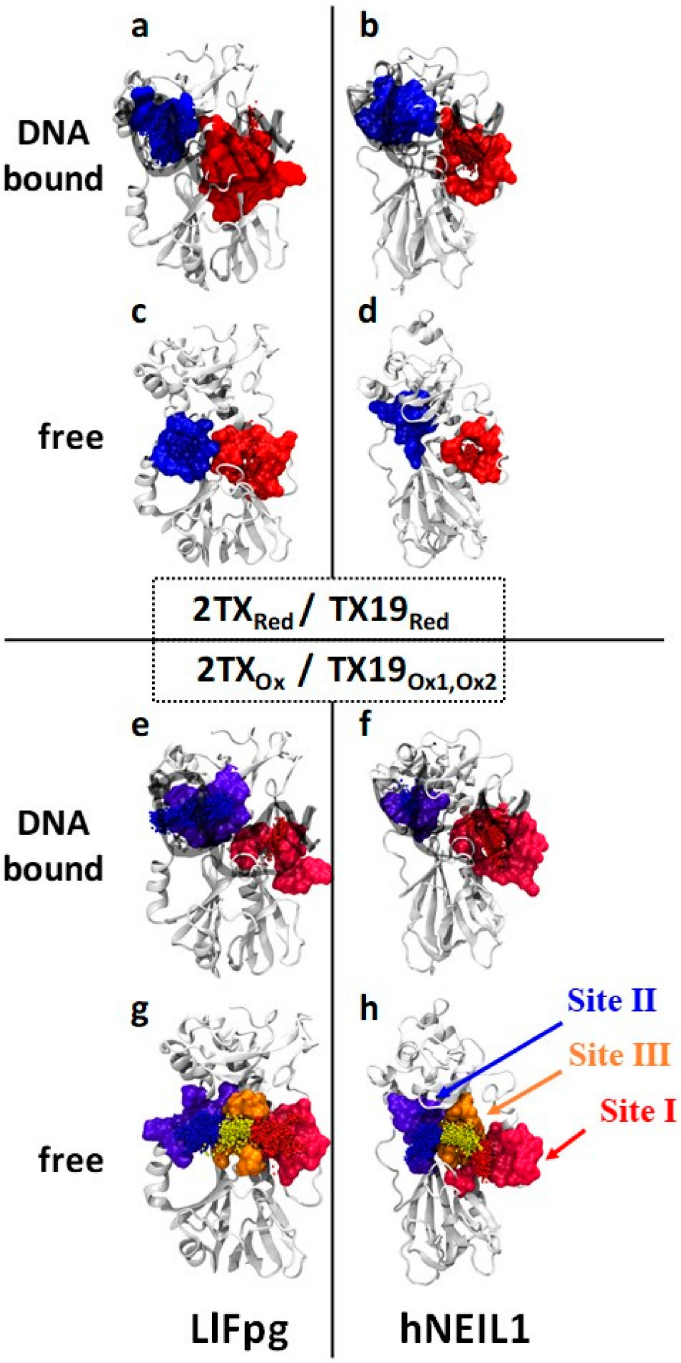


Figure 8. Binding sites of the 2TX and TX19 proposed by blind docking. (**a**–**d**) Docking of 2TX$_{red}$ and TX19$_{red}$ on free and DNA bound LlFpg and hNeil1 as indicated. (**e**–**h**) Docking of 2TX$_{ox}$ and TX19$_{ox}$. Sites in red, blue and orange correspond to site I, II and III, respectively. Proteins are represented by a light grey ribbons and DNA are drawn by translucent black ribbons. Small spheres indicate the geometric centers of blind docking poses of reduced and disulfide forms of 2TX and TX19.

To complement the biochemical data that showed only the disulfide forms of 2TX and TX19 are responsible for enzyme inhibition (Figure 6), we performed blind docking experiments with 2TX$_{ox}$ and TX19$_{ox}$ disulfide forms (Figure 6a, Figures S10 and S12). They revealed two major binding sites on both free and DNA bound enzymes that correspond to binding sites I and II already identified with 2TX$_{red}$ and TX19$_{red}$ (Figure 8e,f). Interestingly, a third TX19$_{ox}$ binding site (named binding site III), which is located between sites I and II, is also proposed by docking experiments, but only on the free enzymes (Figure 8g,h, Figure 9 and Figure S10). Amino acid residues that line the binding site

III are listed in Table S2, and some of them belong to binding sites I or II. Binding site III is hidden when the Fpg/Nei enzymes are bound to DNA (Figure 8e,f). Based on these data, we propose that LlFpg and hNeil1 inhibition by TX19 can result from the binding of its disulfide forms (cyclic trimeric forms, Figure 6a) to site III. The occupation of site III by $TX19_{ox}$ might prevent DNA binding and thus catalysis. The best docking poses in site III of $TX19_{Ox}$ (CTR1R, CTR1S and CTR2) on free enzymes clash with DNA present in the crystal structure of hNeil1 bound to thymine glycol (Tg)-containing DNA (Figure 9). In particular, CTR1S superimposes with flip out Tg base damage in the hNeil1 LRC complex [28]. Thus, docking experiments can provide a reason for the inhibition of hNeil1 and mvNei1 by the dithio-compounds TX16 and TX19 (Figure 4 and Table 1).

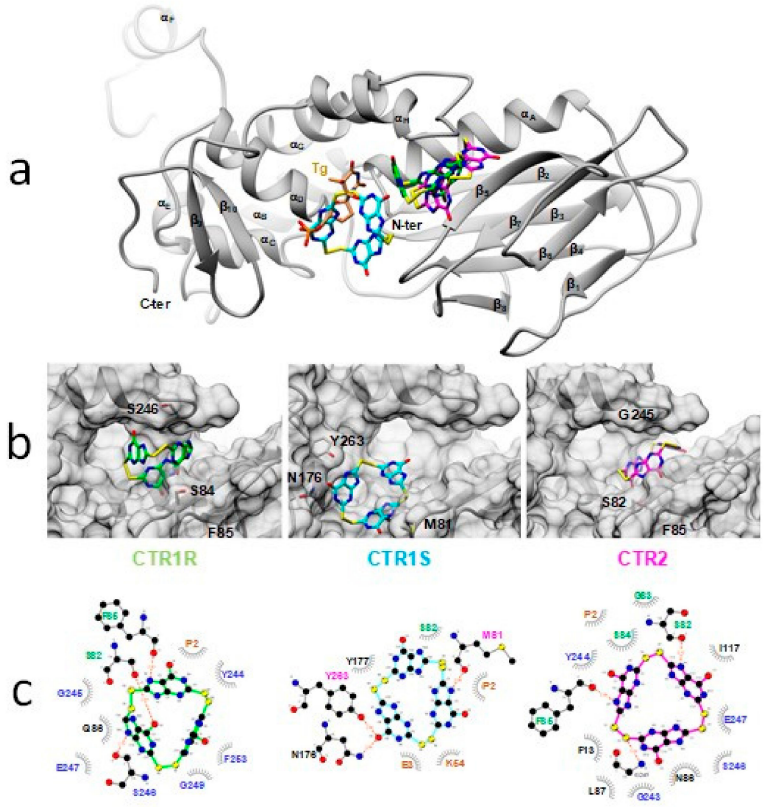


Figure 9. Best docking poses of the three circular trimeric disulfide forms of $TX19_{Ox}$, CTR1R, CTR1S ($TX19_{Ox1}$) and CTR2 ($TX19_{Ox2}$) observed at site III of the free hNeil1 protein (**a**) Superimposition of each docked ligand on the crystal structure of hNeil1 bound to thymine glycol (Tg)-containing DNA (1ITY). For clarity, DNA contained in the crystal model has been deliberately omitted and only the protein in grey ribbon and Tg in brown are shown. This view shows the steric clash between Tg and docked ligands, especially CTR1S. (**b**) Zoom of best docked poses for the three ligands. The solvent accessible surface area in the vicinity of the docking site of hNEIL1 is shown in grey. (**c**) 2D LIGPLOT representations that show protein residues mobilized to stabilize each docked ligand in site III. Hydrogen bonds between protein residues and ligands are shown by red dashed lines and hydrophobic interactions by grey arcs of circle. Catalytic residues of hNeil1 are highlighted in orange. Residues of hNeil1 known to interact directly with the flipped oxidized pyrimidine Tg are colored in magenta. Residues in blue are involved in the lesion capping loop that is essential for stabilizing the flipped damaged nucleoside inside the active site pocket, and residues in green are involved in the intercalation loops of the protein inserted in the DNA minor groove in the hole resulting from the flip out of the damaged nucleoside (Figure S11b). A similar observation can be made for the free LlFpg.

2.5.2. Crystal Structures of Thiopurine Derivatives Bound to Fpg/DNA Complex Partially Confirm Docking Experiments

To further our understanding of the TXn mode of action, we solved several crystal structures of LlFpg bound to 14-mer [THF:C] (LRC) in the presence of thio-compounds. To this end, crystals of the [LlFpg/DNA] complex were soaked in a saturated 2TX/TXn solution of mother liquor, as previously described [18]. Using this method, we successfully resolved the structures of LRC with the moderately potent inhibitor 2TX (new crystal structure), high-efficiency inhibitors TX13, TX19, TX20, TX27, low-efficiency inhibitor TX15 and, TX19 in the presence of TCEP (Table S1). Despite several attempts, we failed to obtain 3D structures with the efficient inhibitors TX14 and TX16, both of which harbor a C8-oxo substitution in the purine ring. In all cases, these crystal structures displayed one molecule of Fpg by the DNA duplex. We did not observe a second molecule binding to the DNA duplex extremity, as we observed with LlFpg bound to 14-mer duplex containing 8-oxoG in place of THF (Figure S5) [38]. In the absence of reducer and as previously observed with 2TX, all crystal structures showed that 2TX/TXn induced oxidation of the ZnF cysteine residues and the concomitant loss of the Zn^{2+} ion [18]. In the present structure of LlFpg/DNA with 2TX, all ZnF cysteine residues are oxidized. The two more exposed LlFpg cysteine residues (Cys248 and Cys268) form an intermolecular disulfide bridge with one 2TX molecule and the two other cysteine residues (Cys245 and Cys265) establish an intramolecular disulfide bond between them (Figure 10a). By showing the 2TX-induced Fpg oxidation product, this crystal structure exquisitely supports the thiol/disulfide exchange mechanism we proposed above for the inhibition of ZnF-containing Fpg/Nei DNA glycosylases. With TXn, there was no clear electron density near the sulfur atoms of Cys248 and Cys268. According to the thiol/disulfide molecular mechanism (Figure 10b), this last observation suggests either the intermolecular disulfide bonds are too labile in the used crystallization conditions (spontaneous hydrolysis after oxidation) or the covalently linked TXn molecules are too dynamic to generate an interpretable electron density.

Similar to the 2TX-containing crystal structure, the intramolecular disulfide bridge Cys245-S-S-Cys265 is also formed in the presence of TXn. Not observed previously, one 2TX molecule that is non-covalently bound to the enzyme is inserted at the protein-DNA interface in the vicinity of the ZnF loop and the H2TH motif (Figure 10a,b), the two DNA binding domains that characterize the Fpg/Nei DNA glycosylase superfamily [29]. This new 2TX binding site is included in site II proposed by the blind docking experiments described above (Figure 8). A similar non-covalent binding at site II is also observed with TXn (Figure 10c–g). Among the best docking poses observed with 2TX and TX19, some of them are very similar to the corresponding crystal structures (Figure S13). This last observation validates the docking procedure we used. The strictly conserved residues R260 of the ZnF loop and K57 of the β2-β3 loop of the N-terminal LlFpg domain are directly involved in the recognition of 2TX/TXn (Figure S11). The guanidinium group of the R260 side chain interacts with the unique sulfur atom of 2TX, TX15, TX20 and TX27, with the sulfur atom at the C8 position of TX19 (with or without TCEP) or with the N1 atom of the purine moiety of TX13. The epsilon amino group of the K57 side chain interacts with the sulfur atom of all 2TX/TXn compounds. R260 of ZnF of LlFpg (R277 of the ZnLF of hNeil1 and mvNei1) and K57 (K54 and K60 in hNeil1 and mvNei1, respectively) are key Fpg/Nei residues that are involved in DNA binding [36]. R260 and K57 are essential to neutralize the electronegativity of DNA phosphates that border the damaged nucleoside. They allow the enzyme to bend DNA. This DNA bending facilitates the extrusion of the damaged nucleoside outside the DNA helix and its stabilization inside the active site pocket that exposes it for catalysis [28,36,40]. In addition to its role in DNA binding, K57 is also proposed as an accessory catalytic residue for DNA glycosylase activity of the EcFpg [38,41,42]. Through binding at site II, the present crystal structures indicate that 2TX/TXn in their reduced forms do not disturb the local structure of the complex (the DNA backbone conformation and the side chain orientation of R260 and K57 are unchanged). In the absence of TCEP, we did not observe the fixation of the 2TX/TXn disulfide forms at sites I and II because they are probably consumed during oxidation of the protein ZnF and transformed into reduced species. This potential is consistent with the thiol/disulfide exchange mechanism we proposed (Figure 5b). For Fpg, the presence of TCEP in the

incubation mixtures prevents ZnF oxidation by 2TX/TXn [18]. The crystal structure of Fpg bound to 14-mer [THF:C] in the presence of TX19 and TCEP revealed that the native structure of ZnF is preserved while one molecule of reduced form of TX19 is still observed at site II (Table S1). As anticipated by docking experiments on the enzyme/DNA complexes, the non-covalent binding site (docking site II) can be seen as a composite area partly formed by the protein and by the DNA backbone phosphates of the two-three nucleotides at the 3′ side of the damaged nucleotide (Figure 10b–g): the phosphate group of Thy(8) interacts with the sulfur group of TXn and the phosphate group of the nucleotide Thy(9) with the N1 atom of 2TX and TX13, the N7 atom of TX13, TX19 and TX27, the N9 atom of TX20, and the NH2 group substituted the C6 atom of the TX27 (Figure 10b). However, the binding of $TX19_{red}$ does not alter the local DNA structure any more than that of the protein. Thus, it is very likely that the binding of the reduced 2TX/TXn species to site II should have little or no impact on Fpg/Nei protein activities and DNA binding. In the presence of TCEP, the disulfide forms of 2TX/TXn are reduced and are, thus, unable to inhibit the Fpg/Nei proteins (Figure 4). In addition, we suggest that the occupation of site II by some mono- or dithio-TXn or TXn without the thiol/thione group is incompatible with the fixation of a second Fpg molecule at the end of the DNA duplex (Figures 2a and 3). Thus, the binding of such small molecules to site II may impair the formation of the C2 complex with or without TCEP observed in EMSA experiments with Fpg and in the crystal structure of LlFpg bound 8-oxoG-containing 14-mer DNA duplex (Figure 3, Figures S5 and S6). Indeed, fine inspection of the crystal structures of C2 [38] with that of the lesion recognition complex C1 resolved in the presence of 2TX, TX13, TX15, TX19, TX20 and TX27 showed that the 5′-overhanging cytosine of the damage-containing strand recognized by the second molecule (Mol-2) of LlFpg in C2 clashes with 2TX/TXn bound at the site II (Figure 11). This structural feature provides a satisfactory explanation to understand why some TXn (with or without thiol/thione group) in the presence of TCEP inhibit the formation of C2 (Figure 3 and Figure S6). At this stage of our investigations, we cannot exclude that the binding of the disulfide forms of monothio 2TX/TXn to site II contributes to the global inhibition of the ZnF-containing Fpg/Nei DNA glycosylases observed. As suggested by docking experiments, only the cyclic disulfide forms of the dithio species TX16 and TX19 might be responsible for the inhibition of the ZnLF-containing Fpg/Nei subclass enzymes (hNeil1 and mvNei1) by interacting more particularly with site III of the free enzyme. This proposal is relevant if site III identified by the docking experiments actually exists in hNeil1 and mvNei1. mvNei1, which is completely devoid of cysteine, is a good model to try to answer this question. Unfortunately, and despite numerous attempts, we did not succeed in crystallizing mvNei1 in the presence of the redox equilibrium mixtures of TX16 and TX19.

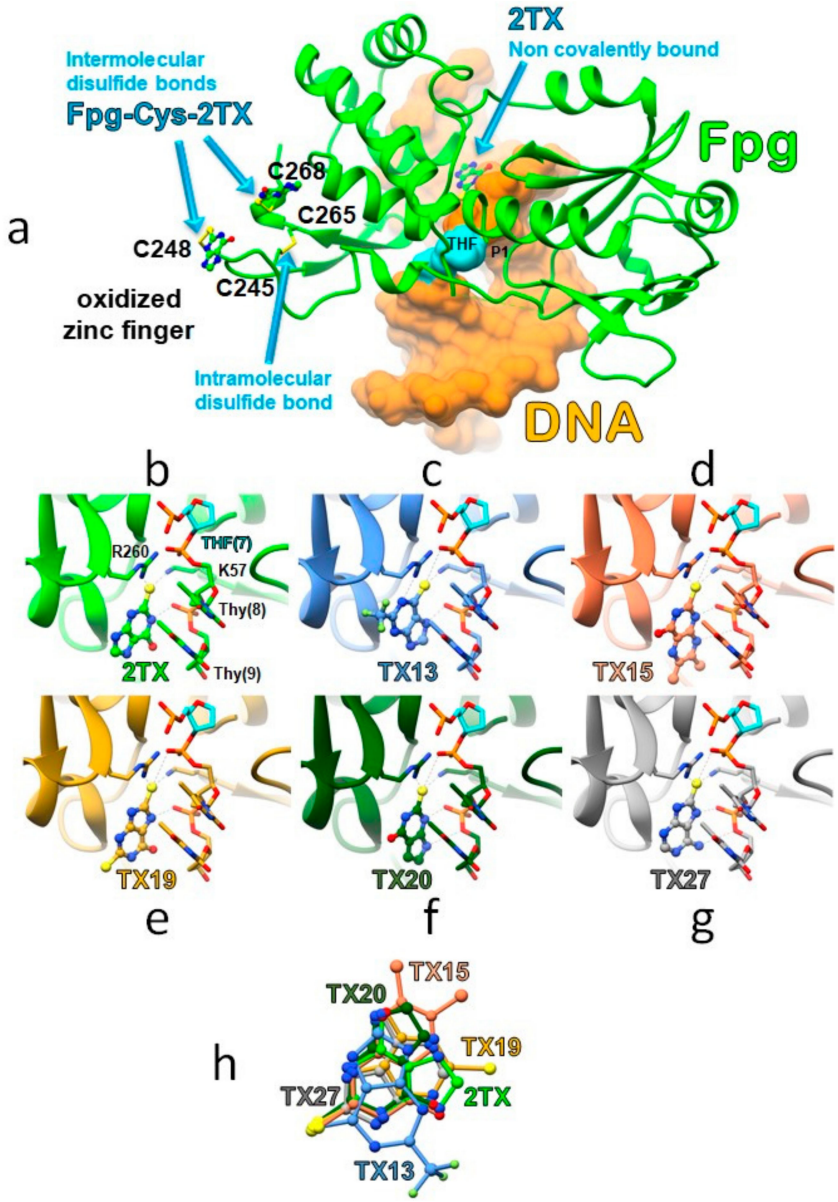


Figure 10. Crystal structures of LlFpg bound to 14-mer [THF:C] DNA in the presence of 2TX/TXn. (**a**) Overview of the structure with 2TX, (**b**) to (**g**) zoom of the non-covalent binding site (i.e., binding site II) with 2TX, TX13, TX15, TX19, TX20 and TX27, respectively, in the absence of TCEP and, (**h**) superimposition of 2TX/TXn inside the LlFpg non-covalent binding site based on a Cα peptide backbone alignment of the protein. The THF abasic site analog is highlighted in magenta. The N-terminal active site proline 1 (P1) is indicated. Ball-and-stick model was used for representing 2TX and TXn. Sulfur atoms are indicated by yellow balls and the intra- and inter-molecular disulfide bonds by yellow sticks. The direct interactions between thiopurines and the side chain of Fpg residues K57 and R260 and between thiopurines and the DNA backbone phosphates at the 3′ side of THF are indicated by grey dashed lines.

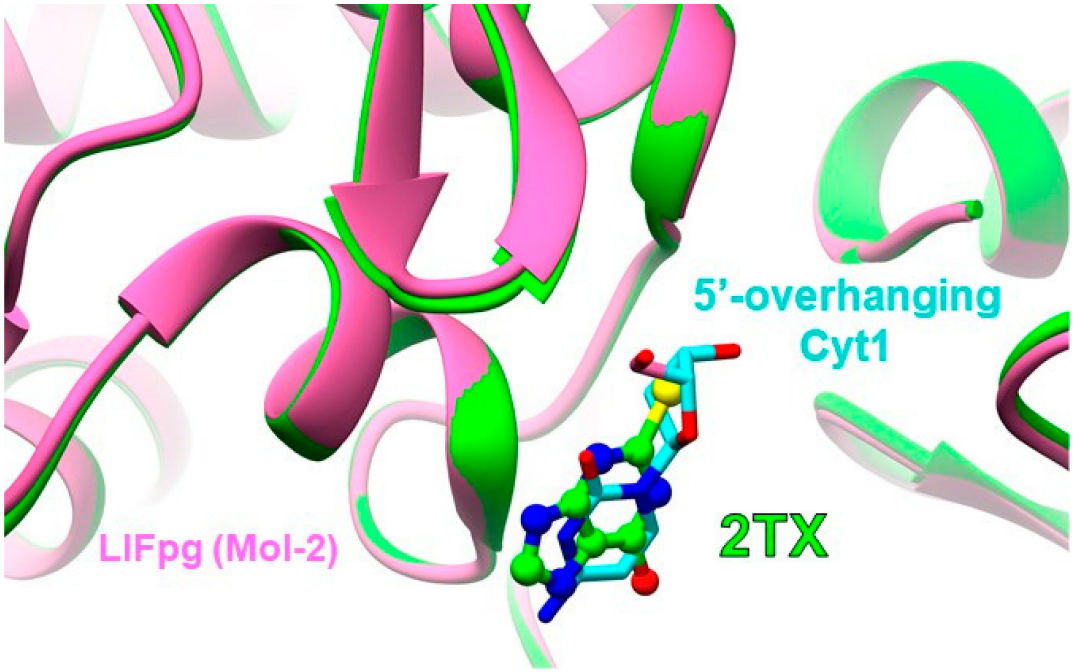


Figure 11. Superimposition of the crystal structure of LlFpg bound to an 8-oxoG-containing 14-mer DNA duplex (in *magenta*) with the crystal structure of LlFpg bound to a THF that contains the same DNA duplex in the presence of 2TX (in green) In this view, only the second LlFpg molecule bound to the extremity of the duplex (Mol-2) is visible (for an overview, see Figure S5, PDBid 4CIS). For clarity, DNA is voluntarily omitted. Only the 5'-overhanging cytosine (Cyt1) of the damaged strand in 4CIS is shown (in cyan).

3. Materials and Methods

3.1. Chemicals and 2TX Derivatives Mini-Library

Sodium periodate (NaIO4) and TCEP were purchased from Sigma-Aldrich. The chemicals 2TX and 2-thio-6,8-dioxopurine (TX14) used to prepare the corresponding homo-disulfide compounds were supplied by MP Biomedicals and ChemCruz, respectively.

Design and synthesis procedures for the TXn used in this study are described in Supplementary Information. All molecule powders were dissolved at 4–10 mM final concentrations in a buffer containing 5 mM HEPES/NaOH (pH7.5), 80% DMSO, 40 mM KCl and 1% glycerol and stored for months at −20 °C.

3.2. Enzymes and DNA Probes

The wild type LlFpg was overproduced in the *fpg* defective *E. coli* strain BH540 and purified, as described previously [22,43]. The human hNeil1 nucleotide sequence was cloned in a pET30a+ expression vector (gift from Dr Susanne Wallace). The mimivirus Neil1 open reading frame (mvNei1) was codon optimized (Genscript) and cloned in pET22b+. The hNeil1 and mvNei1 proteins 6His-tagged at the C-terminus, were overproduced in *E. coli* Rosetta2 (DE3) strain in auto-inducible medium. The 6His-tagged proteins were purified by three liquid chromatography steps using HisTrap™ HP, HiTrap™ SP HP, and Hiload™ Superdex™ 75 pre-packed columns (GE Healthcare, Chicago, Illinois, USA). The hNeil1 and mvNei1 proteins were finally concentrated at 142 and 316 μM, respectively, in a buffer that contained 20 mM HEPES/NaOH (pH7.6), 500 mM NaCl, 1 mM TCEP and 10% glycerol. The truncated form of hNeil1 (ΔhNeil1) was used for EMSA experiments. The ORF coding for ΔhNeil1 fused with a C-terminal 6His-tag was codon optimized (Genscript) and cloned in pET30a+. The recombinant pET30a+ was used to transform the BL21 (DE3) strain. Overproduction was performed

as described above for full length hNeil1. The protein was purified as described above, concentrated at 195 μM in buffer that contained 20 mM HEPES/NaOH (pH 7.5), 150 mM NaCl, 1 mM DTT and 50% Glycerol. After purification, all proteins were stored at −80 °C after flash cooling in liquid nitrogen.

All synthetic oligodeoxyribonucleotides (ODNs) used in this study were purchased from Eurogentec. For catalytic and DNA binding activities (EMSA), the lesion-containing strand was labeled at its 5′ terminus with [γ^{32}P]-ATP using T4 polynucleotide kinase. The 5′-[^{32}P]-labeled 24-mer lesion-containing strand (5′-CTGATCGATGA**X**GCCTGACATGAT where X is for dihydrouracil (DHU) or 8-oxoguanine (8-oxoG)) was annealed with the cold complementary strand (5′-ATCATGTCAGGCYTCATCGATCAG where Y indicates C or G). The resulting DNA duplexes 24-mer 8-oxoG:C (OG:C) and 24-mer DHU:G (DHU:G) were used as substrates for measuring the catalytic activity of Fpg and hNeil1 (mvNei1), respectively. The 5′-[^{32}P]-labeled 14-mer lesion-containing strand 5′-CTCTTTXTTTCTCG, where X is the tetrahydrofuran (THF) AP site analog, was annealed with the cold complementary strand (5′-GCGAGAAACAAAGA) to generate the 14-mer THF:C DNA duplex (THF:C) used for EMSA experiments (see below). The cold 14-mer THF:C DNA duplex was also used to crystallize Fpg/DNA complex with 2TX/TXn, as previously described [44].

3.3. Enzymes Assays for Inhibition and IC50 Determination

Fpg, hNeil1 and mvNei1 are bifunctional DNA glycosylases associated with concerted DNA glycosylase activity (cleavage of the *N*-glycosidic bond between the base lesion and the C1′ of the deoxyribose) and AP lyase activity that comprises the successive cleavage of the phosphodiester bonds at the 3′ and 5′ of the resulting abasic (AP) site, respectively (β, δ-elimination). Thus, AP sites that result from the excision of the damaged base are quantitatively converted into one nucleoside gap by these enzymes. The DNA glycosylase/AP lyase activity of each enzyme was measured by incubating 20 nM of 24-mer DNA duplexes ([OG:C] and [DHU:G] for Fpg and hNeil1/mvNei1, respectively) 5′-labeling on the damaged strand (see above) with 2 nM of Fpg, or mvNei1 for 15 min or, with 20 nM of hNeil1 for 30 min at 37 °C in 10 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1 mM EDTA, 0.1 mg/mL BSA and 8% DMSO. The reaction was stopped by the addition of 1 volume of formamide loading buffer (96% formamide, 20 mM Tris/HCl pH 8.0 and 2 mM Na_2EDTA), heated at 80 °C for 2 min and then analyzed by 20% Urea-PAGE, as previously described [45]. After electrophoresis, the gels were exposed for autoradiography, revealed using Typhoon imager and quantified with ImageQuant software. Under these conditions, only 50% of the substrate is converted to the cleavage product (limiting enzyme concentration conditions). For an initial evaluation of the inhibitory effect, 40 μM of 2TX/TXn was used. The thio-compound was added to the reaction mixture before the final addition of the enzyme (competition condition). For the best identified inhibitors, a dose-response curve was generated and $IC50^A$ determined (Figure 1 and Table 1).

The DNA binding activity was measured using EMSA [37]. Briefly, 0.1 nM of the 5′-labeled 14-mer [THF:C] (see above) was incubated with the enzyme at 4 °C for 30 min in 25 mM HEPES/NaOH (pH 7.5), 150 mM NaCl, 5% glycerol, 0.1 mg/mL BSA and 8% DMSO. The incubation mixture was then loaded onto a native 10% polyacrylamide gel (Acryl/Bis 19/1) and electrophoresed at 25 V/cm for 90 min at 4 °C. After electrophoresis, the gels were dried and exposed overnight for autoradiography using a Typhoon imager and quantified with ImageQuant software. To evaluate the inhibitor effect of 2TX/TXn on the DNA binding activity of each Fpg/Nei enzyme, we first determined the experimental conditions required to bind 50% of the DNA probes without an inhibitor in the presence of 8% DMSO (limiting enzyme concentration). Under these conditions, we obtained the half-maximal DNA binding 0.25, 7 and 5 nM of Fpg, hNeil1 and mvNei1. We used this limited enzyme concentrations to evaluate the inhibitor effect of 2TX/TXn on the enzyme DNA binding activity. This effect was determined by the addition of the thio-compound in the pre-equilibrium mixture of the enzyme and DNA probe (30 min) and incubated for an additional 30 min at 4 °C. For the best inhibitors of Fpg, a dose-response curve was established from at least three independent experiments and the IC50 was determined ($IC50^B$, Table 1).

3.4. Molecular Dynamic Simulation and Docking

Four 3D structures were used as receptors for ligand docking: (i) the 3D model of free Fpg from *L. lactis* (LlFpg) obtained by removing DNA duplex from the X-ray structure of the enzyme bound to THF-containing 14-mer DNA duplex, (ii) the X-ray structure of LlFpg/DNA-THF complex (PDBidcode: 1PM5), (iii) the X-ray structure of free hNeil1 (PDBidcode: 1TDH) for which missing residues were reconstructed using the Prime module from Schrödinger (Schrödinger, LLC, New York, NY, 2017) (https://www.schrodinger.com/) and (iv) the X-ray structure of hNeil1 bound to THF-DNA (PDBidcode: 5ITT).

The four systems were placed in a rectangular box with explicit TIP3P water molecules with a 10 Å buffer. To ensure the neutrality of the systems, counter ions were added. The systems were then energy minimized. Molecular dynamics (MD) simulations of 50 ps were then performed during which the temperature was increased in 50 K steps to reach the target temperature of 310 K. Equilibration were performed in seven MD steps of 50 ps, with positional restraints on the heavy atoms that incrementally decreased from 100 to 0.0 kcal/(mol.Å^2). For each system, a 40 ns MD production run was performed in the NPT ensemble and conformations were stored every 1 ps for analysis. For energy minimization and MD simulations SHAKE algorithm [46] was used with a 2 fs time step for integration of the equation of motion. The particle mesh Ewald summation [47] was used for electrostatics calculation with a 12 Å cutoff. Periodic boundary conditions were applied. For MD simulation, the Amber package (Amber 11) with the ff12SB force field was used (https://ambermd.org/).

For the two systems containing a THF (tetrahydrofuran), the geometry and charge distribution of the THF residue were computed using Gaussian 03 program (https://gaussian.com/) and fitted with the RESP program [48] for compatibility with Cornell et al.'s force field [49] adapted to Parm98 parameters. For each system, a Hierarchical Ascending Clustering (HAC) method by average linkage distance calculation was used to classify system conformations in clusters based on a 2D RMSD map (each configuration of a trajectory was compared to all the others). We extracted one structure, called centroid, for each cluster. A centroid is representative of the system flexibility within a single cluster. The thirty collected centroids were then used for docking studies.

Reduced and disulfide forms of 2TX were sketched within the Sybyl suite and their 3D structures calculated with Concord and their energy minimized. For the disulfide dimer of 2TX, several conformational search techniques (Random search, GA search and Confort with the default parameters set in Sybyl package) were used to have a representative set of conformers. To obtain an exhaustive (as possible) set of cyclic conformers for docking with TX19$_{Ox1,Ox2}$, we used several methods such as Random search, Grid search and simulated annealing, as implemented in the Sybyl package (Certara, NJ, USA), all with default parameters. To prepare the docking with Surflex-Dock 2.706 Surflex Manual: Docking and Similarity Version 2.7 (May 2012) by BioPharmics LLC (Santa Fe, CA, USA), a protomol was generated for each protein structure (centroids) in automatic mode with default parameters from the Sybyl interface. Protomols can be seen as an 'imprinting' an interaction cavity of a protein. The docking studies were performed in GeomX mode; the other options were set as -spindense 9.0, +premin, +remin, -multistart 6, -div_rms 0.5, -ndock_final 20 and +ring. For each pre-calculated conformer, Surflex-Dock will generate 6 subsequent conformers prior to the docking. Twenty poses are proposed at the end.

3.5. Crystallization and X-Ray Structure Determination

LlFpg/DNA complexes (70–160 µM) were prepared by mixing in a 1.5 molar excess the 14-mer [THF:C] DNA duplex with LlFpg and crystallized at room temperature in the presence of 0.5% DMSO and with (for TX19) or absence of TCEP using a previously described method [18,22,23]. Crystals with 2TX, TX13, TX15, TX19, TX19 (+TCEP), TX20 or TX27 were obtained by soaking crystals of LlFpg/DNA complex in 2TX (TXn)-saturated mother liquor. Soaked crystals were then flash cooled in liquid nitrogen. X-ray diffraction data sets were collected at 100 K at beamline PROXIMA-1 and PROXIMA-2 (SOLEIL, Paris) or ID30A-3/MASSIF-3 (ESRF, Grenoble). All data were processed using the XDS

package [50] and the CCP4 program SCALA [51]. The 3D structures were solved by molecular replacement using the crystal structure of the wild type LlFpg protein bound to 14-mer [THF:C] (PDB ID code 1PM5, [35]) as previously described [18]. X-ray data collection and refinement statistics are presented in Table S3. All structure cartoons were generated with CCP4MG [52].

4. Conclusions

DNA glycosylases, which initiate base excision repair (BER), constitute the first line of defense against oxidative stress resulting in nucleobase oxidation. Paradoxically, the very high level of oxidative stress in viral and bacterial infections, in cancer cells and in the cells of patients suffering from inflammatory diseases, such as Huntington's disease, give reason to hope that the DNA glycosylases hOgg1 and hNeil1 are relevant pharmacological targets in precise pathologic situations [13,15,17,53–56]. Although many DNA repair enzymes, such as PARP1, are considered as therapeutic targets [57], the advantage of selectively targeting DNA glycosylases in cancer and neurodegenerative diseases has emerged only recently. The redundancy of DNA glycosylase substrate specificity and the observation that single KO mouse of hNeil1 and hOgg1 are viable and fertile have long hampered the search for inhibitors of these enzymes. Yet this situation is ideal in the case of a synthetic lethality interaction between the pathologic context and DNA glycosylase [58,59]. Over the last few years, the search for inhibitors of DNA glycosylases hNeil1 (hNeil2) and hOgg1 has become a very active field [18–21,56,60,61]. We started this research in 2014 by discovering that 2-thioxanthine is an irreversible inhibitor of the bacterial enzymes Fpg and Nei and the human enzyme hNeil2, three enzymes belonging to the Fpg/Nei DNA glycosylase structural superfamily having a zinc finger (ZnF) [18]. Interestingly, only hNeil1 from this enzyme superfamily is resistant to the inhibitory effect of 2TX. We clearly demonstrated that 2TX is able to irreversibly oxidize the Fpg ZnF. Although this inhibition mechanism remains to be elucidated, this finding provided an explanation permitting to understand why hNeil1 is resistant to 2TX knowing that this enzyme does not have ZnF.

In the present work, we described the discovery and biochemical characterization of new structural Fpg and hNeil1/mvNei1 inhibitors. By screening a small library of 2TX derivatives (TXn), the initial goal of the study was to find inhibitors that were more effective than 2TX and to decipher the mode of action of these thio-compounds on the Fpg/Nei DNA glycosylases. Four TXn (mono- or dithio-thiopurines) are better Fpg inhibitors than 2TX. Especially, 2-trifluoro-6-thiopurine (TX13) is seven times more effective than 2TX. As anticipated by our previous study, most TXn tested are unable to inhibit hNeil1 and mvNei1. What was not expected at the beginning of this study, however, is that TX16 and TX19, the two dithiopurines evaluated, also efficiently inhibit hNeil1 and mvNei1, DNA glycosylases devoid of ZnF. Knowing that mvNei1 is free of cysteine, its inhibition by TX16 and TX19 seemed very mysterious. Even more curious, TCEP, a very powerful reducer of disulfide bridges, annihilates the inhibitor effect of TX16 and TX19 on hNeil1 and mvNei1 glycosylase activity and DNA binding. A similar effect was observed for the action of TCEP in the presence of monothio-compounds. This suggested that TCEP acts directly on 2TX/TXn.

By combining mass spectrometry, chemistry, biochemistry and X-ray structure analysis, we demonstrated that only the disulfide forms of 2TX, TX14 and TX19 (in equilibrium with the reduced form in our DMSO mixtures of 2TX/TXn) are responsible for the inhibition of the Fpg/Nei enzymes. Although reduced 2TX/TXn are able to bind the free and DNA bound enzymes, as proposed by blind docking and confirmed by X-ray crystal structures of ternary complexes Fpg-DNA-2TX/TXn, their binding either in site II or in the active site (site I) observed with or without TCEP is probably not associated with inhibition. Disulfides are known to inhibit several enzymes according to various mechanisms. Unsymmetrical aromatic disulfides obtained from monothiolated compounds inhibit the SARS-CoV M^{pro} (an important target to fight severe acute respiratory syndrome-associated coronavirus, SARS-CoV) probably by interacting with the catalytic cysteine containing the protease M^{pro} [62]. Li et al. reported the synthesis of unsymmetrical aromatic disulfides as novel inhibitors ("herbicides") of the acetohydroxyacid synthase, a plant and microorganism enzyme involved in the

biosynthesis of branched-chain amino acids [63]. Reynolds and co-workers published unsymmetrical aryl-alkyl disulfides that selectively inhibit the methicillin-resistant *Staphylococcus aureus* and *Bacillus anthracis* [64]. The use of disulfides to selectively target the human ZnF- and ZnLF-containing Fpg/Nei DNA glycosylases (hNeil1, hNeil2 and hNeil3) in therapy is, however, conditioned to their activity in plasma (antioxidant power mediated by human serum albumin, HSA; [65]) and in the presence in cells of reduced glutathione (G-SH), which remains to be determined. Taking into account the potential reduction of the disulfide inhibitors by HSA, all enzyme assays presented in this work were performed in the presence of BSA (the bovine functional homolog of HSA). Regarding G-SH, the mixed disulfide Glutathione G-S-S-compounds retain in vivo activity [66]. In pathologic contexts associated with high levels of oxidative stress, the antioxidant power of G-SH and HSA can be greatly reduced and, thus, are incapable of countering the action of disulfides. Some aromatic disulfides were shown to be active in vivo as antitumor agents [67].

In the case of ZnF-containing Fpg/Nei enzymes, we demonstrated that ZnF oxidation results from a thiol/disulfide exchange mechanism between cysteine residues coordinating the Zn^{2+} ion and the 2TX/TXn disulfides. Thus, these disulfides act as Zn ejectors, which results, in our case, in an irreversible inhibition. Such a mechanism has already been proposed for the inhibition of metallo-β-lactamases by symmetrical disulfides, which induced the selective oxidation of the Zn(II)-coordinated cysteine center in the active site [68]. The natural symmetrical disulfide Psammaplin A isolated from sponge, and its analogs were described as antitumor agents by strongly inhibiting DNA methyltransferase (DNMT) and histone deacethylase (HDAC) interfering with their respective Zn center [67]. A thiol/disulfide exchange mechanism resulting in ZnF oxidation and Zn release has also been proposed for the inhibition of the HIV-1 nucleocapside protein by the Zn ejector DIBA (a symmetrical disulfide) [69]. Blind and flexible docking also suggests that inhibition partly results from the binding of the disulfide to the active site (site I) and/or site II.

The most striking result of this study is the discovery that disulfide forms of dithio-TXn (TX16 and TX19) are also capable of inhibiting ZnLF-containing Fpg/Nei DNA glycosylases like hNeil1 and mvNei1. For TX19, we showed that the thio-compound exists in DMSO solution in equilibrium between $TX19_{Red}$ and its two cyclic trimeric disulfide forms $TX19_{Ox1,Ox2}$. Through binding at sites I and II, EMSA experiments and blind and flexible docking suggest that the inhibitory effect of $TX19_{Ox}$ could be effective on both the free and DNA bound hNeil1 and mvNei1. Docking of TX19 disulfides revealed a third binding site (site III) only accessible on the free Fpg and hNeil1. If this site really exists, the binding of Fpg and hNeil1 to DNA could be inhibited when a disulfide of TX19(TX16) occupies site III. The self-assembly property of the aromatic dithio-compounds, such as TX16 and TX19 in cyclic disulfides (disulfide cyclophanes), has been known for a long time by chemists. Thus, symmetrical and unsymmetrical disulfides are useful tools for creating small molecule libraries by dynamic combinatorial chemistry (DCLs) [70]. More recently, this method was adapted to search for enzyme inhibitors using the protein as a molecular trap (named "protein-directed combinatorial chemistry", PDCC) [71]. Combined with a fragment-based screening method, PDCC is an alternative to the expensive high-throughput screening (HTS) of chemical libraries and enables the exploration of a larger chemical space in a more efficient manner. Our work modestly illustrates the interest of this strategy to find new enzyme inhibitors and potential protein target sites that can inspire the design of new classes of inhibitors of hNeil1. The discovery that disulfide cyclophanes inhibit Fpg and hNeil1 (mvNei1) constitutes the starting point in the exploration of new avenues to find new potent inhibitors with original structures using PDCC combined with fragment-base screening. We are in the process of implementing this strategy for hNeil1 and hOgg1.

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/1422-0067/21/6/2058/s1.

Author Contributions: B.C., L.A. and N.G. conceptualized the research and designed all experiments. C.R. and S.B. carried out computational experiments. F.C. and A.B. performed X-ray data collection at synchrotron SOLEIL and solved X-ray crystal structures. Z.T. and V.R. synthetized TXn derivatives. S.G. carried out enzyme assays

(catalysis and DNA binding), produced, and purified disulfide forms of 2TX and TX14 and participated in crystallization experiments. V.A. participated in the preparation of 2TX and TX14 disulfides. J.C. participated in IC50 determination. V.G. and M.G. produced and purified Fpg, hNEIL1 and mvNei1. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Ligue contre le cancer du Grand-Ouest (CSIRGO, AO 2015 to 2019), the Région Centre-Val de Loire (APP IR 2013 and 2017), grant numbers 2013-00082978 and 2017-00117252, the radiation protection board of France Electricity (EDF), grant number RB2015-10 and by the Cancéropôle Grand-Ouest (project CONCERTO), grant number 2018-001240994. All research funding was obtained by BC. CR was funded by a thesis grant from the University of Orleans and JC by a thesis grant from the Ligue nationale contre le cancer.

Acknowledgments: We are very grateful for the help provided by Quoc-Tuan Do and Philippe Bernard from Greenpharma (AS, 3 Allée du Titane, 45000 Orléans) (https://www.greenpharma.com/) for their help in docking experiments with disulfide cyclophanes. We thank Samia Aci-Sèche (Institut de Chimie Organique et analytique, UMR7311, CNRS-Université d'Orléans) for her help in parameterizing THF-containing DNA required for the preparation of the 3D structure of the LlFpg/DNA and hNeil1/DNA complexes before molecular dynamic simulations. We thank the staff of the ESRF (Grenoble, France) and SOLEIL (Saclay, France) synchrotron facilities for their help in X-ray diffraction data collection. We are grateful to Loïc Habert for her excellent experimental assistance in the synthesis of thiopurines. We thank Cyril Colas, Guillaume Gabant and Martine Cadène of the mass spectrometry platform of the 'Centre de Biophysique Moléculaire' for the analysis of our samples by LC-HRMS.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

EMSA	Electrophoretic mobility shift assay
2TX	2-thioxanthine
TXn	2TX derivatives
LRC	Lesion recognition complex
BER	Base excision repair
LCL	Lesion capping loop
DMSO	Dimethylsulfoxide
THF	Tetrahydrofuran
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
DTT	Dithiothreitol
8-oxoG	8-oxoguanine
AP site	Abasic site
ZnF	Zinc Finger
ZnLF	Zincless Finger

References

1. Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **1993**, *362*, 709–715. [CrossRef] [PubMed]
2. Cadet, J.; Davies, K.J.A.; Medeiros, M.H.; Di Mascio, P.; Wagner, J.R. Formation and repair of oxidatively generated damage in cellular DNA. *Free Radic. Biol. Med.* **2017**, *107*, 13–34. [CrossRef] [PubMed]
3. Cheng, K.C.; Cahill, D.S.; Kasai, H.; Nishimura, S.; Loeb, L.A. 8-Hydroxyguanine, an abundant form of oxidative DNA damage, causes G—-T and A—-C substitutions. *J. Biol. Chem.* **1992**, *267*, 166–172. [PubMed]
4. Schaaper, R.M.; Kunkel, T.A.; Loeb, L.A. Infidelity of DNA synthesis associated with bypass of apurinic sites. *Proc. Natl. Acad. Sci. USA* **1983**, *80*, 487–491. [CrossRef]
5. Basu, A.K. DNA Damage, Mutagenesis and Cancer. *Int. J. Mol. Sci.* **2018**, *19*. [CrossRef] [PubMed]
6. Czarny, P.; Wigner, P.; Galecki, P.; Sliwinski, T. The interplay between inflammation, oxidative stress, DNA damage, DNA repair and mitochondrial dysfunction in depression. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **2018**, *80*, 309–321. [CrossRef] [PubMed]
7. Leandro, G.S.; Sykora, P.; Bohr, V.A. The impact of base excision DNA repair in age-related neurodegenerative diseases. *Mutat. Res.* **2015**, *776*, 31–39. [CrossRef]
8. Sahan, A.Z.; Hazra, T.K.; Das, S. The Pivotal Role of DNA Repair in Infection Mediated-Inflammation and Cancer. *Front. Microbiol.* **2018**, *9*, 663. [CrossRef]

9. Scharer, O.D. Chemistry and biology of DNA repair. *Angew. Chem.* **2003**, *42*, 2946–2974. [CrossRef]
10. Krokan, H.E.; Bjoras, M. Base excision repair. *Cold Spring Harb. Perspect. Biol.* **2013**, *5*, a012583. [CrossRef]
11. Wallace, S.S. Base excision repair: A critical player in many games. *DNA Repair* **2014**, *19*, 14–26. [CrossRef] [PubMed]
12. Budworth, H.; Harris, F.R.; Williams, P.; Lee, D.Y.; Holt, A.; Pahnke, J. Suppression of Somatic Expansion Delays the Onset of Pathophysiology in a Mouse Model of Huntington's Disease. *PLoS Genet.* **2015**, *11*, e1005267. [CrossRef] [PubMed]
13. Goula, A.V.; Berquist, B.R.; Wilson, D.M., 3rd; Wheeler, V.C.; Trottier, Y.; Merienne, K. Stoichiometry of base excision repair proteins correlates with increased somatic CAG instability in striatum over cerebellum in Huntington's disease transgenic mice. *PLoS Genet.* **2009**, *5*, e1000749. [CrossRef] [PubMed]
14. Goula, A.V.; Goula, A.V.; Stys, A.; Chan, J.P.; Trottier, Y.; Festenstein, R.; Merienne, K. Transcription elongation and tissue-specific somatic CAG instability. *PLoS Genet.* **2012**, *8*, e1003051. [CrossRef] [PubMed]
15. Kovtun, I.V.; Liu, Y.; Bjoras, M.; Klungland, A.; Wilson, S.H.; McMurray, C.T. OGG1 initiates age-dependent CAG trinucleotide expansion in somatic cells. *Nature* **2007**, *447*, 447–452. [CrossRef]
16. Taricani, L.; Shanahan, F.; Pierce, R.H.; Guzi, T.J.; Parry, D. Phenotypic enhancement of thymidylate synthetase pathway inhibitors following ablation of Neil1 DNA glycosylase/lyase. *Cell Cycle* **2010**, *9*, 4876–4883. [CrossRef]
17. Ramdzan, Z.M.; Ginjala, V.; Pinder, J.B.; Chung, D.; Donovan, C.M.; Kaur, S.; Leduy, L.; Dellaire, G.; Ganesan, S.; Nepveu, A. The DNA repair function of CUX1 contributes toradioresistance. *Oncotarget* **2017**. [CrossRef]
18. Biela, A.; Coste, F.; Culard, F.; Guerin, M.; Goffinont, S.; Gasteiger, K.; Cieśla, J.; Winczura, A.; Kazimierczuk, Z.; Gasparutto, D.; et al. Zinc finger oxidation of Fpg/Nei DNA glycosylases by 2-thioxanthine: Biochemical and X-ray structural characterization. *Nucleic Acids Res.* **2014**, *42*, 10748–10761. [CrossRef] [PubMed]
19. 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**. [CrossRef] [PubMed]
20. 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*, e81667. [CrossRef]
21. Tahara, Y.K.; 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. [CrossRef] [PubMed]
22. Coste, F.; Ober, M.; Le Bihan, Y.V.; Izquierdo, M.A.; Hervouet, N.; Mueller, H.; Carell, T.; Castaing, B. Bacterial base excision repair enzyme Fpg recognizes bulky N7-substituted-FapydG lesion via unproductive binding mode. *Chem. Biol.* **2008**, *15*, 706–717. [CrossRef]
23. Le Bihan, Y.V.; Angeles Izquierdo, M.; Coste, F.; Allerm, P.; Culard, F.; Gehrke, T.H.; Essalhi, K.; Carell, T.; Castaing, B. 5-Hydroxy-5-methylhydantoin DNA lesion, a molecular trap for DNA glycosylases. *Nucleic Acids Res.* **2011**, *39*, 6277–6290. [CrossRef] [PubMed]
24. Castaing, B.; Geiger, A.; Seliger, H.; Nehls, P.; Laval, J.; Zelwer, C.; Boiteux, S. Cleavage and binding of a DNA fragment containing a single 8-oxoguanine by wild type and mutant FPG proteins. *Nucleic Acids Res.* **1993**, *21*, 2899–2905. [CrossRef]
25. O'Connor, T.R.; Graves, R.J.; de Murcia, G.; Castaing, B.; Laval, J. Fpg protein of Escherichia coli is a zinc finger protein whose cysteine residues have a structural and/or functional role. *J. Biol. Chem.* **1993**, *268*, 9063–9070. [PubMed]
26. Tchou, J.; Michaels, M.L.; Miller, J.H.; Grollman, A.P. Function of the zinc finger in Escherichia coli Fpg protein. *J. Biol. Chem.* **1993**, *268*, 26738–26744.
27. Doublie, 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. USA* **2004**, *101*, 10284–10289. [CrossRef]
28. 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. USA* **2016**, *113*, 7792–7797. [CrossRef]

29. Boiteux, S.; Coste, F.; Castaing, B. Repair of 8-oxo-7,8-dihydroguanine in prokaryotic and eukaryotic cells: Properties and biological roles of the Fpg and OGG1 DNA N-glycosylases. *Free Radic. Biol. Med.* **2017**, *107*, 179–201. [CrossRef]
30. Speina, E.; Ciesla, J.M.; Graziewicz, M.A.; Laval, J.; Kazimierczuk, Z.; Tudek, B. Inhibition of DNA repair glycosylases by base analogs and tryptophan pyrolysate, Trp-P-1. *Acta Biochim. Pol.* **2005**, *52*, 167–178. [CrossRef] [PubMed]
31. Boiteux, S.; O'Connor, T.R.; Lederer, F.; Gouyette, A.; Laval, J. Homogeneous Escherichia coli FPG protein. A DNA glycosylase which excises imidazole ring-opened purines and nicks DNA at apurinic/apyrimidinic sites. *J. Biol. Chem.* **1990**, *265*, 3916–3922.
32. Karakaya, A.; Jaruga, P.; Bohr, V.A.; Grollman, A.P.; Dizdaroglu, M. Kinetics of excision of purine lesions from DNA by Escherichia coli Fpg protein. *Nucleic Acids Res.* **1997**, *25*, 474–479. [CrossRef] [PubMed]
33. Tudek, B.; Van Zeeland, A.A.; Kusmierek, J.T.; Laval, J. Activity of Escherichia coli DNA-glycosylases on DNA damaged by methylating and ethylating agents and influence of 3-substituted adenine derivatives. *Mutat. Res.* **1998**, *407*, 169–176. [CrossRef]
34. Castaing, B.; Fourrey, J.L.; Hervouet, N.; Thomas, M.; Boiteux, S.; Zelwer, C. AP site structural determinants for Fpg specific recognition. *Nucleic Acids Res.* **1999**, *27*, 608–615. [CrossRef] [PubMed]
35. Pereira de Jesus, K.; Serre, L.; Zelwerm, C.; Castaing, B. Structural insights into abasic site for Fpg specific binding and catalysis: Comparative high-resolution crystallographic studies of Fpg bound to various models of abasic site analogs-containing DNA. *Nucleic Acids Res.* **2005**, *33*, 5936–5944. [CrossRef]
36. Serre, L.; Pereira de Jesus, K.; Boiteux, S.; Zelwer, C.; Castaing, B. Crystal structure of the Lactococcus lactis formamidopyrimidine-DNA glycosylase bound to an abasic site analog-containing DNA. *EMBO J.* **2002**, *21*, 2854–2865. [CrossRef]
37. Castaing, B.; Boiteux, S.; Zelwer, C. DNA containing a chemically reduced apurinic site is a high affinity ligand for the E. coli formamidopyrimidine-DNA glycosylase. *Nucleic Acids Res.* **1992**, *20*, 389–394. [CrossRef]
38. Sadeghian, K.; Flaig, D.; Blank, I.D.; Schneider, S.; Strasser, R.; Stathis, D.; Winnacker, M.; Carell, T.; Ochsenfeld, C. Ribose-protonated DNA base excision repair: A combined theoretical and experimental study. *Angew. Chem.* **2014**, *53*, 10044–10048. [CrossRef]
39. Duclos, S.; Aller, P.; Jaruga, P.; Dizdaroglu, M.; Wallace, S.S.; Doublie, S. Structural and biochemical studies of a plant formamidopyrimidine-DNA glycosylase reveal why eukaryotic Fpg glycosylases do not excise 8-oxoguanine. *DNA Repair* **2012**, *11*, 714–725. [CrossRef]
40. Imamura, K.; Averill, A.; Wallace, S.S.; Doublie, S. Structural characterization of viral ortholog of human DNA glycosylase NEIL1 bound to thymine glycol or 5-hydroxyuracil-containing DNA. *J. Biol. Chem.* **2012**, *287*, 4288–4298. [CrossRef]
41. Rogacheva, M.; Ishchenko, A.; Saparbaev, M.; Kuznetsova, S.; Ogryzko, V. High resolution characterization of formamidopyrimidine-DNA glycosylase interaction with its substrate by chemical cross-linking and mass spectrometry using substrate analogs. *J. Biol. Chem.* **2006**, *281*, 32353–32365. [CrossRef] [PubMed]
42. Saparbaev, M.; Sidorkina, O.M.; Jurado, J.; Privezentzev, C.V.; Greenberg, M.M.; Laval, J. Repair of oxidized purines and damaged pyrimidines by E. coli Fpg protein: Different roles of proline 2 and lysine 57 residues. *Environ. Mol. Mutagenesis* **2002**, *39*, 10–17. [CrossRef] [PubMed]
43. Duwat, P.; de Oliveira, R.; Ehrlich, S.D.; Boiteux, S. Repair of oxidative DNA damage in gram-positive bacteria: The Lactococcus lactis Fpg protein. *Microbiology* **1995**, *141*, 411–417. [CrossRef] [PubMed]
44. Pereira de Jesus, K.; Serre, L.; Hervouet, N.; Bouckson-Castaing, V.; Zelwer, C.; Castaing, B. Crystallization and preliminary X-ray crystallographic studies of a complex between the Lactococcus lactis Fpg DNA-repair enzyme and an abasic site containing DNA. *Acta Cryst. D Biol. Crystallogr.* **2002**, *58*, 679–682. [CrossRef] [PubMed]
45. Le Meur, R.; Culard, F.; Nadan, V.; Goffinont, S.; Coste, F.; Guerin, M.; Loth, K.; Landon, C.; Castaing, B. The nucleoid-associated protein HU enhances 8-oxoguanine base excision by the formamidopyrimidine-DNA glycosylase. *Biochem. J.* **2015**, *471*, 13–23. [CrossRef] [PubMed]
46. Ryckaert, J.P.; Ciccotti, G.; Berendsen, H.J.C. Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327–341. [CrossRef]
47. Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593. [CrossRef]

48. Wang, J.; Cieplak, P.; Kollman, P.A. How, does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules. *J. Comput. Chem.* **2000**, *21*, 1049–1074. [CrossRef]
49. Cornell, W.D.; Cieplak, P.; Bayly, C.I.; Gould, I.R.; Merz, K.M.; Ferguson, D.M.; Spellmeyer, D.C.; Fox, T.; Caldwell, J.W.; Kollman, P.A. A second Generation Force Field for the Simulation of Proteins. *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197. [CrossRef]
50. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 125–132. [CrossRef]
51. Winn, M.D.; Ballard, C.C.; Cowtan, K.D.; Dodson, E.J.; Emsley, P.; Evans, P.R.; Keegan, R.M.; Krissinel, E.B.; Leslie, A.G.; McCoy, A.; et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 235–242. [CrossRef] [PubMed]
52. McNicholas, S.; Potterton, E.; Wilson, K.S.; Noble, M.E. Presenting your structures: The CCP4mg molecular-graphics software. *Acta Crystallogr. D Biol. Crystallogr.* **2011**, *67*, 386–394. [CrossRef] [PubMed]
53. Goula, A.V.; Merienne, K. Abnormal base excision repair at trinucleotide repeats associated with diseases: A tissue-selective mechanism. *Genes (Basel)* **2013**, *4*, 375–387. [CrossRef] [PubMed]
54. Li, G.; Yuan, K.; Yan, C.; Fox, J., 3rd; Gaid, M.; Breitwieser, W.; Bansal, A.K.; Zeng, H.; Gao, H.; Wu, M. 8-Oxoguanine-DNA glycosylase 1 deficiency modifies allergic airway inflammation by regulating STAT6 and IL-4 in cells and in mice. *Free Radic. Biol. Med.* **2012**, *52*, 392–401. [CrossRef]
55. Touati, E.; Michel, V.; Thiberge, J.M.; Ave, P.; Huerre, M.; Bourgade, F.; Klungland, A.; Labigne, A. Deficiency in OGG1 protects against inflammation and mutagenic effects associated with H. pylori infection in mouse. *Helicobacter* **2006**, *11*, 494–505. [CrossRef]
56. Visnes, T.; Cazares-Korner, 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. [CrossRef] [PubMed]
57. Gavande, N.S.; VanderVere-Carozza, P.S.; Hinshaw, H.D.; Jalal, S.I.; Sears, C.R.; Pawelczak, K.S. DNA repair targeted therapy: The past or future of cancer treatment? *Pharmacol. Ther.* **2016**, *160*, 65–83. [CrossRef]
58. Helleday, T. Cancer phenotypic lethality, exemplified by the non-essential MTH1 enzyme being required for cancer survival. *Ann. Oncol.* **2014**, *25*, 1253–1255. [CrossRef]
59. Visnes, T.; Grube, M.; Hanna, B.M.F.; Benitez-Buelga, C.; Cazares-Korner, A.; Helleday, T. Targeting BER enzymes in cancer therapy. *DNA Repair* **2018**, *71*, 118–126. [CrossRef]
60. Grin, I.R.; Rieger, R.A.; Zharkov, D.O. Inactivation of NEIL2 DNA glycosylase by pyridoxal phosphate reveals a loop important for substrate binding. *Biochem. Biophys. Res. Commun.* **2010**, *394*, 100–105. [CrossRef]
61. Michel, M.; Visnes, T.; Homan, E.J.; Seashore-Ludlow, B.; Hedenstrom, M.; Wiita, E.; Vallin, K.; Paulin, C.B.J.; Zhang, J.; Wallner, O.; et al. Computational and Experimental Druggability Assessment of Human DNA Glycosylases. *ACS Omega* **2019**, *4*, 11642–11656. [CrossRef] [PubMed]
62. Wang, L.; Bao, B.B.; Song, G.Q.; Chen, C.; Zhang, X.M.; Lu, W.; Wang, Z.; Cai, Y.; Li, S.; Fu, S.; et al. Discovery of unsymmetrical aromatic disulfides as novel inhibitors of SARS-CoV main protease: Chemical synthesis, biological evaluation, molecular docking and 3D-QSAR study. *Eur. J. Med. Chem.* **2017**, *137*, 450–461. [CrossRef] [PubMed]
63. Li, Z.S.; Wang, W.M.; Lu, W.; Niu, C.W.; Li, Y.H.; Li, Z.M.; Wang, J.G. Synthesis and biological evaluation of nonsymmetrical aromatic disulfides as novel inhibitors of acetohydroxyacid synthase. *Bioorganic Med. Chem. Lett.* **2013**, *23*, 3723–3727. [CrossRef] [PubMed]
64. Turos, E.; Revell, K.D.; Ramaraju, P.; Gergeres, D.A.; Greenhalgh, K.; Young, A.; Sathyanarayan, N.; Dickey, S.; Lim, D.; Alhamadsheh, M.M.; et al. Unsymmetric aryl-alkyl disulfide growth inhibitors of methicillin-resistant Staphylococcus aureus and Bacillus anthracis. *Bioorg. Med. Chem.* **2008**, *16*, 6501–6508. [CrossRef] [PubMed]
65. Rabbani, G.; Ahn, S.N. Structure, enzymatic activities, glycation and therapeutic potential of human serum albumin: A natural cargo. *Int. J. Biol. Macromol.* **2019**, *123*, 979–990. [CrossRef] [PubMed]
66. Rice, W.G.; Baker, D.C.; Schaeffer, C.A.; Graham, L.; Bu, M.; Terpening, S.; Clanton, D.; Schultz, R.; Bader, J.P.; Buckheit, R.W., Jr.; et al. Inhibition of multiple phases of human immunodeficiency virus type 1 replication by a dithiane compound that attacks the conserved zinc fingers of retroviral nucleocapsid proteins. *Antimicrob. Agents Chemother.* **1997**, *41*, 419–426. [CrossRef]

67. Hong, S.; Shin, Y.; Jung, M.; Ha, M.W.; Park, Y.; Lee, Y.J.; Shin, J.; Oh, K.B.; Lee, S.K.; Park, H.G. Efficient synthesis and biological activity of Psammaplin A and its analogs as antitumor agents. *Eur. J. Med. Chem.* **2015**, *96*, 218–230. [CrossRef]
68. Boerzel, H.; Koeckert, M.; Bu, W.; Spingler, B.; Lippard, S.J. Zinc-bound thiolate-disulfide exchange: A strategy for inhibiting metallo-beta-lactamases. *Inorg. Chem.* **2003**, *42*, 1604–1615. [CrossRef]
69. Garg, D.; Torbett, B.E. Advances in targeting nucleocapsid-nucleic acid interactions in HIV-1 therapy. *Virus Res.* **2014**, *193*, 135–143. [CrossRef]
70. Otto, S. Dynamic molecular networks: From synthetic receptors to self-replicators. *Acc. Chem. Res.* **2012**, *45*, 2200–2210. [CrossRef] [PubMed]
71. Mondal, M.; Hirsch, A.K. Dynamic combinatorial chemistry: A tool to facilitate the identification of inhibitors for protein targets. *Chem. Soc. Rev.* **2015**, *44*, 2455–2488. [CrossRef] [PubMed]

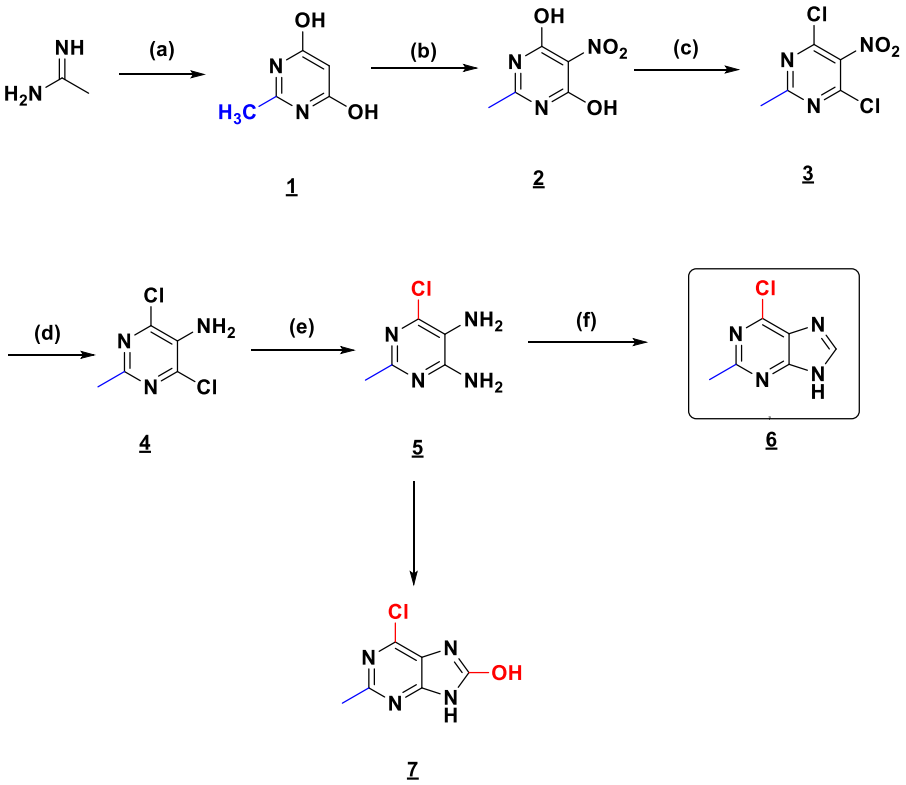
SUPPLEMENTARY INFORMATION

Thiopurine derivative-induced Fpg/Nei DNA glycosylase inhibition: structural, dynamic and functional insights

Charlotte Rieux[1,2], Stéphane Goffinont[1], Franck Coste[1], Zahira Tber[2], Julien Cros[1], Vincent Roy[2,3], Martine Guérin[1,3], Julien Cros[1], Virginie Gaudon[1], Stéphane Bourg[2], Artur Biela[1], Vincent Aucagne[1], Luigi Agrofoglio[2,3], Norbert Garnier[1,3] and Bertrand Castaing[1]

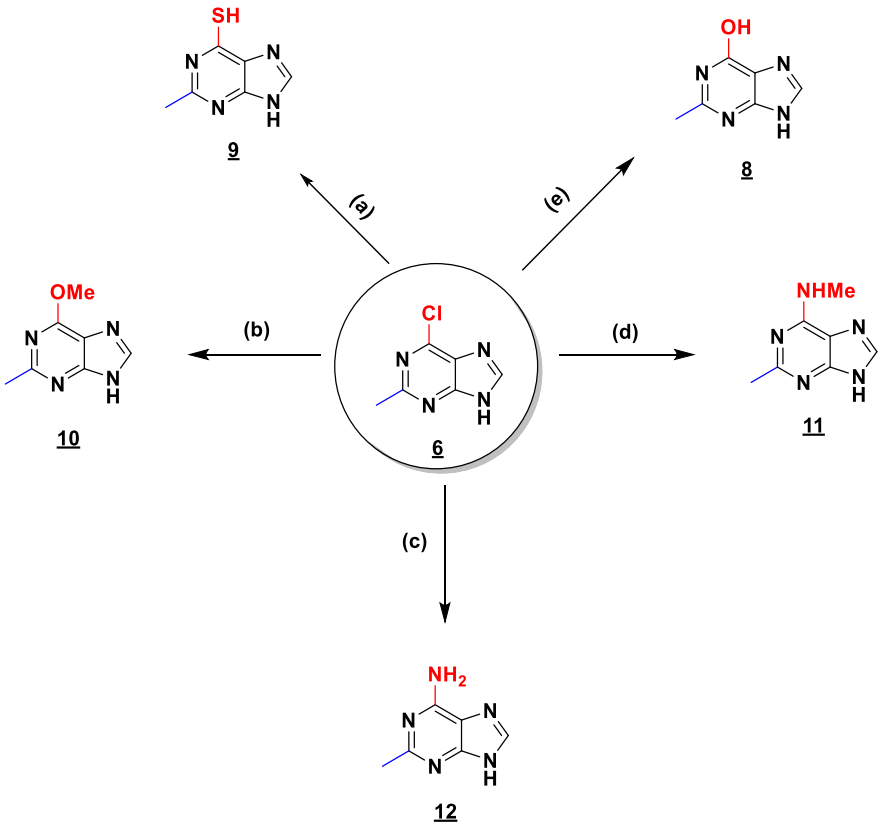
(1) Centre de Biophysique Moléculaire, UPR4301 CNRS, rue Charles Sadron, F-45071 Orléans Cédex 2, France
(2) Institut de Chimie Organique et Analytique, UMR7311 CNRS-Orleans University, Université d'Orléans, Pôle de Chimie, rue de Chartres, F-45100 Orléans, France
(3) Université d'Orléans, UFR Sciences et Techniques, rue de Chartres, 45100, Orléans, France

● Chemical synthesis strategies for the design of the small 2-thioxanthine derivative (TXn) library



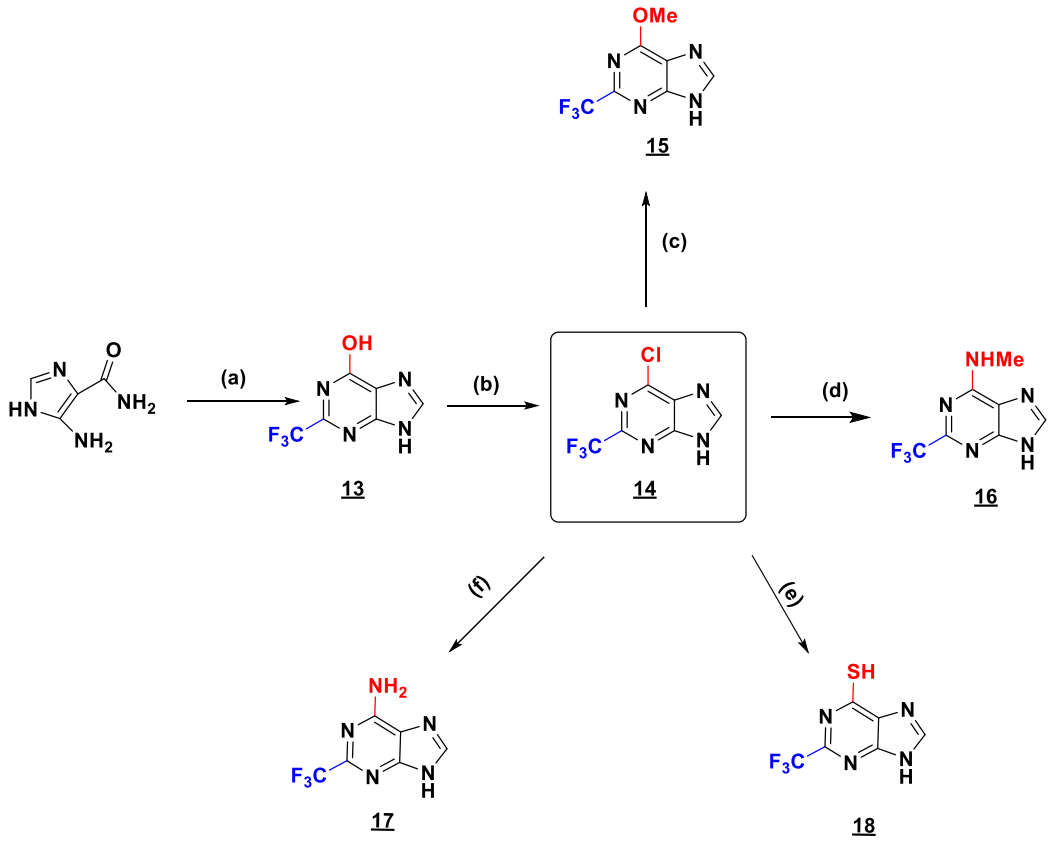
(a) EtONa, diethylmalonate, 85°C, 3 h, 69%; (b) HNO_3 fumant, acetic acid, 0°C, 3 h, 67%; (c) $POCl_3$, *N*,*N* diethylaniline, 115°C , 2h 30, 64%; (d) Fe , AcOH, 65°C, 82%; (e) NH_4OH, 120°C, 25 min, Mw, 70%; (f) Triethylorthoformate, 75 min, 100°C, 84%; (g) Urea, 160°C, 1h, 20%.

Scheme 1



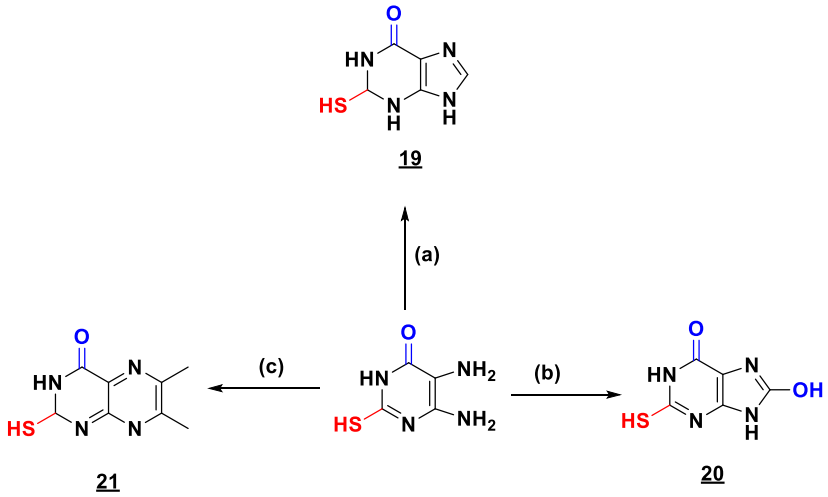
(a) Thiourea, EtOH, reflux, 1h, 35%; (b) NaOH, MeOH, Mw, 100°C, 4min, 81%; (c) 1) NaN_3, $EtOH/H_20$, 80°C , 2 h, 2) $SnCl_2,2H_20$, reflux, 2 h, 74%; (d) $MeNH_2$, EtN_3, BuOH, 140°C, Mw, 15 min, 58%; (e) $HCOOH/H_2O$ (1/1), 40°C , 20 h, 80%.

Scheme 2



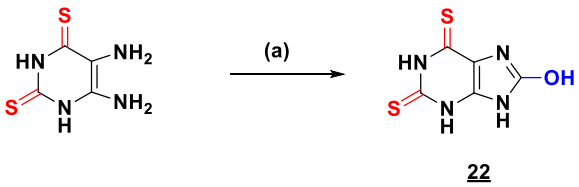
(a) trifluoroacétamide, 160°C, 4 h, 48%; (b) $SOCl_2$, DMF, $CHCl_3$, reflux, 64%; (c) NaOH, MeOH, Mw, 100 °C, 40 min, 79%;(d) $MeNH_2$,BuOH, Et_3N, Mw, 140°C, 15 min, 53%; (e)Thiourea, EtOH, reflux, 4h, 54%; (f) 1) NaN_3 , $EtOH/H_2O$, 2h , 80 °C, 2) $SnCl_2$,$2H_2O$, EtOH, 2 h, reflux, 67%

Scheme 3



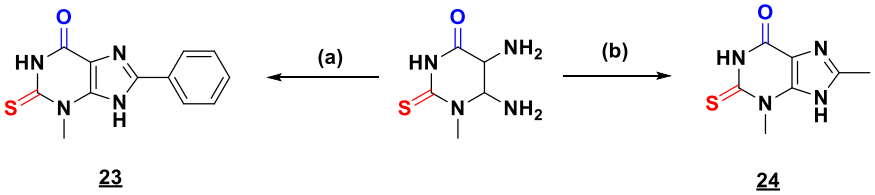
(a) 1) formic acid, 2 h, reflux; cooled, 2) H_2NCHO, 2 h, 175-185°C, 97%; (b) Urea, 170°C, 30 min, 90% ; (c) Diacetyl, DMF, 100°C, 5h, 67%.

Scheme 4



(a) Urea, 170°C, 45 min, 60%

Scheme 5

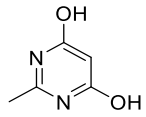


(a) triethoxymethylbenzene, reflux, 15h, 90%; (b) triethylorthoacetate,reflux, 15h, 60%

Scheme 6

The reactions were monitored by thin-layer chromatography (TLC) analysis using silica gel (60 F254) plates. The compounds were visualized by UV irradiation. Flash column chromatography was performed on silica gel 60 (230−400.13 mesh, 0.040 0.063 mm). The infrared spectra of compounds are given in cm^{-1}. Further, 1 H and 13C NMR spectra were recorded at 250 nm (13C, 62.9 MHz) or at 400 nm (13C, 100.62 MHz). Chemical shifts are given in parts per million using tetramethylsilane (TMS) as an internal standard. Coupling constants (J) are reported in Hertz (Hz). All reactions under microwave irradiation were performed using the Biotage Initiator microwave in 2-5 mL or 10-20 mL sealed tubes. High-resolution mass spectra (HRMS) were performed on a quadrupole analyzer.

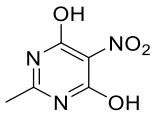
4,6-Dihydroxy-2-methylpyrimidine 1 (CAS : 1194-22-5)



To a stirred solution of sodium ethoxide, prepared from sodium (210 mg, 9.13 mmol) and ethanol (4 mL), acetamidine hydrochloride (290 mg, 3.066 mmol, 1 eq) was added in one portion (rinsed with ethanol, 0.5 mL). After 5 min, diethyl malonate (0.41 mL, 2.7 mmol, 0.9 eq) was added in the above suspension (rinsed with ethanol, 1 mL). The reaction mixture was refluxed (85°C) for 3 h, allowed to cool to room temperature, and diluted with water (4 mL). After all the precipitate was dissolved, concentrated hydrochloric acid (0.6 mL) was added dropwise. The precipitate was collected, washed with water (2 mL), ethanol (2 mL) and ether (2 mL), and then dried under a vacuum.

White solid (69%); **IR:** 3078, 2793, 2584, 1622, 1573, 1450, 1325, 1041, 982, 933; **^{1}H-NMR (DMSO-D_6, 250 MHz):** δ 4.94 (s, 1 H, H-5), 3.33 (s, 2 H), 2.20 (s, 3 H, CH_3); **HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_5H_7N_2O_2$: 127.05020; found: 127.05019.

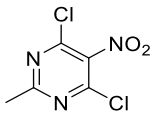
2-Methyl-5-nitro-4,6-dihydroxypyrimidine 2 (CAS : 53925-27-2)



To a mixture of nitric acid (fuming, 0.5 mL, 6 eq) and acetic acid (glacial, 0.25 mL) at 10–15°C, **1** (246 mg, 1.86 mmol, 1 eq) was added in portions over a period of 30 min. The reaction mixture was stirred for 3 h at room temperature, cooled to 0°C, diluted with cold water (0–5°C, 0.4 ml), stirred for 3 min more, and then filtered. The solid was washed with cold water (0–5°C, 0.2 ml), ethanol (1 mL) and ether (2 mL) and dried under a vacuum.

White solid (67%); **IR**: 3537, 3458, 2815, 2711, 1673, 1627, 1361, 1299, 1275, 1050, 785; **^{1} H-NMR (DMSO-D_6, 250 MHz):** δ 12.98 (brs, 2H), 2.31 (s, 3H, CH_3); **HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_5H_6N_3O_4$: 154.02471 found : 154.02476

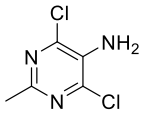
2-Methyl-5-nitro-4,6-dichloropyrimidine **3** (CAS : 13162-43-1)



To a suspension of **2** (241 mg, 1.36 mmol, 1 eq) in phosphorus oxychloride (1 mL) *N*,*N*diethylaniline (0.28 mL, 1.76 mmol) was added dropwise. The reaction mixture was refluxed (115°C bath) for 2.5 h and then cooled to room temperature. Excess phosphorus oxychloride was evaporated under a vacuum. The residue was diluted with ether and poured onto ice. The water layer was extracted with ether (3x5 mL). The combined organic extract was washed with saturated solutions of sodium bicarbonate and sodium chloride, dried over sodium sulfate, and evaporated under a vacuum.

Light brown crystals (64 %); **IR**: 3537, 3464, 2923, 2847, 2701, 1673, 1627, 1589, 1360, 1280, 1048, 784; **^{1}H NMR (DMSO-D$_6$, 250 MHz):** δ 2.7 (s, 3H); **HRMS (ESI) :** not detected.

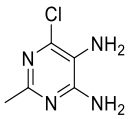
5-Amino-4,6-dichloro-2-methylpyrimidine **4** (CAS : 39906-04-2)



The product **3** (20.2 g, 0.097 mol, 1 eq) was dissolved in a mixture of hydrochloric acid (10 mL) and ethanol (200 mL). One portion of iron powder (16.4 g, 0.293 mol, 3 eq) was added to this solution. The mixture was then refluxed for 8 h, cooled to room temperature, and filtered through a pad of Celite. The filtrate was concentrated in vacuo. The residue was extracted with EtOAc, and the organic extract was washed with 1 N NaOH, water and brine and dried over anhydrous $MgSO_4$. It was then filtered and concentrated in vacuo to a tan solid. Purification by recrystallization from water yielded the pure product.

White crystals (82 %); **IR**: 3449, 3336, 3341, 3293, 1607, 1505, 1430, 1315, 1123, 913, 788; **^{1}H NMR (DMSO-D$_6$, 250 MHz):** δ 5.88 (brs, 2H), 2.40 (s, 3H); **HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_5H_6N_3Cl_2$: 177.99332 found : 177.99323

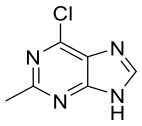
2-Methyl-4,6-dichloro-5-aminopyrimidine **5** (CAS : 933-80-2)



The product **4** (1.00 g) and ammonium hydroxide (5.0 mL, 28.0% - 30.0 %) were placed in a microwave vial. The vial was sealed and heated in a microwave reactor at 120 °C for 25 min. The reaction mixture was cooled to room temperature and concentrated to give 6-chloro-2-methylpyrimidine-4,5-diamine.

Light brown solid (70 %); **IR :** 3347, 3293, 3131, 3034, 2800, 1657, 1545, 1397, 1236, 865; **^{1}H NMR (DMSO-D$_6$, 250 MHz)** δ 6.65 (s, 2H, NH_2), 4.71 (s, 2H, NH_2), 2.18 (s, 3H, CH_3)**; HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_5H_8N_4Cl$: 159.04320 found : 159.04311

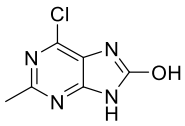
2-Methyl-6-chloropurine 6 (CAS : 100859-35-6)**: TX1**



The compound **5** (0.5 g, 3.155 mmol, 1 eq) was suspended in triethyl orthoformate (5.62 mL, 33.8 mmol, 10.7 eq), placed in a preheated oil bath (100 °C) and stirred for 75 min. The reaction mixture was cooled to room temperature, concentrated, treated with hexanes and filtered. The solid was washed with hexanes and dried to yield 6-chloro-2-methyl-9H-purine.

Light brown solid (84%); **IR:** 3110, 2955, 2923, 2768, 2714, 2673, 1613, 1592, 1386, 1292, 1224, 952, 886, 810; **^{1}H NMR (DMSO-D$_6$, 250 MHz)**: δ 13.7 (brs, 1 H, NH), 8.55 (s, 1H, H-8), 2.65 (s, 3 H, CH_3)**; HRMS (ESI) : m/z (M+H)** $^+$ calculated for $C_6H_6N_4Cl$: 169.02755 found : 169.02747

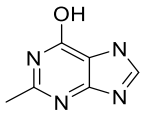
6-chloro-1,7-dihydro-2-methyl- 8*H*-Purin-8-one 7 (CAS: 98138-74-0)**: TX2**



The compound **5** (0.100 g, 0.63 mmol, 1 eq) and urea (0.213g, 3.5 mmol, 5.5 eq) are mixed in a flask with stirring and heated to 160 ° C for 45 min. The agitation became difficult. The residue was purified by silica gel column chromatography.

White solid (20%); **IR**: 3129, 3021, 2980, 2777, 1738, 1593, 1426, 1392, 1222, 1182, 1043, 1005, 901, 711, **^{1}H NMR (MeOH-D$_4$, 250 MHz):** δ 2.79 (s, 3H, CH_3)**; HRMS (ESI): m/z (M+H)** $^+$ calcd for $C_6H_6N_4ClO$: 185.02246 found : 185.02213

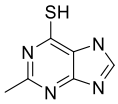
2-Methylhypoxanthine 8 (CAS : 5167-18-0)**: TX3**



A solution of **6** (100 mg, 5.94 mmol, 1 eq) in a 1:1 mixture of water (8.7mL) and formic acid (8.7 mL) was stirred for 20 h at 40 °C. After evaporation of all volatiles, the residue was purified by silica gel column chromatography (EtOAc/MeOH, 9:1)

White solid (80%); **IR** 3092, 3027, 2927, 2797, 2679, 1686, 1364, 1260, 1185, 957, 807; **^{1}H NMR (DMSO-D$_6$, 250 MHz):** δ 13.10 (brs, 1 H, NH), 12.09 (brs, 1 H, OH), 8.12 (s, 1H, H-8), 2.35 (s, 3H, CH_3), **HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_6H_7N_4O$: 151.06143 found : 151.061224.

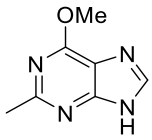
2-Methyl-6-mercaptopurine 9 (CAS: 38917-31-6)**: TX4**



A solution of **6** (200 mg, 1.19 mmol, 1 eq) and thiourea (399 mg, 5.24 mmol, 4.42 eq) in ethanol was stirred at reflux for 1h. After cooling to room temperature, the result was filtered, and the residue was purified by silica gel column chromatography (EtOAc/MeOH, 8:2)

Beige solid (35%), **IR:** 3010, 2946, 25581551, 1418, 1356,1191, 1137, 1007, 794, **^{1}H NMR (DMSO-D$_6$, 250 MHz)**: δ 8.61 (s, 3H), 2.70 (s, 1H), **HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_6H_7N_4S$: 167.038594 , found : 167.038676

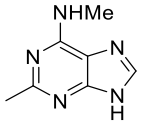
6-methoxy-2-methyl-9*H*-Purine 10 (CAS: 1198-45-4)**: TX5**



A 0.5-2 mL vial was loaded a solution of **6** (0.1 g, 0.593 mmol, 1 eq) and NaOH (71 mg, 1.77 mmol, 3 eq) in methanol (1.2 mL). After microwaving at 100°C for 40 min, the volatiles were evaporated under reduced pressure. The resulting residue was then purified by silica gel chromatography (EtOAc/ MeOH).

White solid (81%) ; **IR :** 3008, 2945, 2752, 2546, 1613, 1591, 1566, 1397, 1382, 1268, 1116, 949, 794, **^{1}H NMR (DMSO-D$_6$, 250 MHz)**: δ 13.16 (brs, 1H, NH), 8.25 (s, 1 H, H-8), 4.05 (s, 3H, CH_3), 2.52 (s, 3H, CH_3), **HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_7H_9N_4O$: 165.077087 found : 165.077014

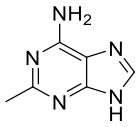
2-Methyl-6-methylaminopurine 11 (CAS : 90375-79-4)**: TX6**



A 2-5 mL vial was charged with **6** (0.150 g, 0.88 mmol, 1 eq), *n*-butanol (1.81 mL), methylamine (0.21mL, 5.33 mmol, 6 eq), and distilled triethylamine (0.37 mL, 2.6 mmol, 3 eq). After microwaving at 140 °C for 15 min, the solvent was removed under pressure, and the residue was purified by silica gel column chromatography (EtOAc/ MeOH).

White solid (58 %); **IR:** 3274, 3208, 3059, 2983, 2942, 1606, 1384, 1330, 1264, 1137, 931, 825, 704; **^{1}H NMR (DMSO-D$_6$, 250 MHz):** δ 12.63 (brs , NH_2) , 7.95 (s, 1H , H-8) , 7.38 (1 H, NH), 2.96 (s, 3H), 2.4 (s, 3H), **HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_7H_{10}N_5$: 164.09307, found : 164.09296.

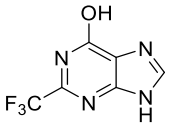
2-Methyladenine 12 (CAS: 1445-08-5)**: TX7**



To a solution of compound **6** (100 mg, 0.593 mmol, 1 eq) in a 5:1 mixture of EtOH/H_2O (6.7 mL/ 1.35 mL) NaN_3 (115 mg, 1.77 mmol, 3 eq) was added. The mixture was stirred for2 h at 80°C and evaporated under reduced pressure. The crude product was dissolved in EtOH (8 mL) and $SnCl_2$, $2H_2O$ (400 mg , 1.77 mmol, 3 eq) was added. The solution was refluxed for 2 h and concentrated *in vacuo*. The product was purified by silica gel column chromatography (EtOAc / MeOH, 8/2).

White solid (74 %); **IR:** 3325, 3151, 2755, 2695, 1664 (NH_2), 1593, 1426, 1393, 1264, 930, 810, **^{1}H NMR (DMSO-D_6, 250 MHz):** δ 12.69 brs (NH), 7.96 (H-8), 6.94 (NH_2), 2.36 (CH_3); **HRMS (ESI): m/z (M+H)$^+$** calculated for $C_6H_8N_5$: 150.07742 found : 150.07725

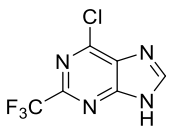
1,7-dihydro-2-(trifluoromethyl)-6*H*-Purin-6-one 13 (CAS: 2268-14-6)**: TX8**



Trifluoroacetamide (174 g, 1.54 mol, 10 eq) was heated in an oil bath at 110°C. After the trifluoroacetamide melted, 50 g of 5-aminoimidazole-4-carboxamide hydrochloride (25.0 g, 0.154 mol) was added. The mixture was warmed to reflux (bath temp 160°C to 165°C) for 4 h, cooled to ambient temperature and the rocky solid was triturated with 1 L of ether and stirred for 2 h. The solid was collected by suction filtration. The solid was dispersed in water, stirred for 1 h and collected by suction filtration. The solid was dried under a vacuum.

Blue solid (48%),**IR:** 3151, 3034, 2939, 2800, 1727, 1687, 1376, 1336, 1190, 1143 (CF_3), 1108, 959, 786, **^{1}H NMR (DMSO-D_6, 250 MHz): δ** 13.75 (br s, 1 H, NH), 8.40 (s, 1 H), **^{19}F NMR (DMSO-D_6, 235 MHz):** δ -67.76 (s, 3F), **HRMS (ESI): m/z (M+H)$^+$** calculated for $C_6H_4N_4OF_3$: 205.03317 found : 205.3346

6-Chloro-2-trifluoromethyl-9*H*-purine 14 (CAS: 1998-63-6)**: TX9**

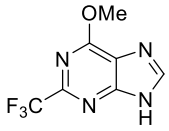


A mixture of 1,9-dihydro-2-(trifluoromethyl)-6H-purin-6-one (5.00 g, 24.5 mmol) and chloroform (100 mL) was refluxed with stirring. A solution prepared by dropwise addition of thionyl chloride (8.9 mL,

122 mmol, 5 eq) to cold dimethylformamide (8.92 g, 122 mmol, 5 eq) was added to the refluxing mixture. After 1.5 h at reflux, the reaction was poured into ice water (400 mL). The layers were separated, and the chloroform phase was washed with water (4x 100 mL). The pH of the aqueous phase was adjusted to 7 with saturated sodium bicarbonate and extracted with ether (3x 400 mL). The combined ether and chloroform extracts were dried with $MgSO_4$ and concentrated to dryness under reduced pressure .

White solid (64%), **IR:** 3 119, 3069, 2961, 2929, 2806, 1358, 1196, 1142 (CF_3), 1000, 872 ; **^{1}H NMR (DMSO-D_6, 250 MHz):** δ 14.45 (brs, 1 H, NH), 8.95 (s, 1 H, H-8); **^{19}F NMR (DMSO-D_6, 235 MHz)** :δ - 67.30 (s, 3F); **HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_6H_3N_4F_3Cl$: 222.99928, found : 222.99912.

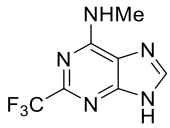
6-methoxy-2-(trifluoromethyl)-1H-Purine 15 (CAS: 658707-77-8): TX10



A 0.5-2 mL vial was loaded a solution of 6-chloro-2-trifluoromethyl-9H-purine (0.150 g, 0.673 mmol, 1 eq) and NaOH (81 mg, 2 mmol, and 3 eq) in methanol (1, 37 ml). After microwaving at 100°C for 40 min, the volatiles were distilled under reduced pressure; the resulting residue was then purified by silica gel chromatography. (EtOAc/ MeOH).

White solid (79 %), **IR:** 3119, 3005, 2951, 2781, 1628, 1391, 1285, 1138, 1117, 921, 804; ^{1}H NMR (DMSO-D_6, 250 MHz): 8.62 (s, 1H), 4.15 (s, 3H) ; **^{19}FNMR (DMSO-$d_{6,\ 235}$)** : -67.68; **HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_7H_6N_4F_3O$: 219.04882 found : 219.04871.

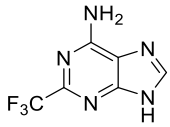
N-methyl-2-(trifluoromethyl)-1*H*-Purin-6-amine 16 (CAS: 18925-07-0): TX11



A 2-5 mL vial was charged with 6-chloro-2-trifluoromethyl-9H-purine (0.100 g, 0.45 mmol, 1 eq), *n*-butanol (0.90 mL), methylamine (0.106 mL, 2.66 mL, 6 eq), and distilled triethylamine (0.188 mL, 1.33 mmol, 3 eq). After microwave at 140 °C for 15 min, the solvent was removed under pressure, and the residue was washed with EtOAc. The product was dried under reduced pressure.

White solid (53%), **IR:** 2974, 2942, 2601, 1628, 1474, 1444, 1396, 1226, 1142, 1035, 936; **^{1}H NMR (DMSO-D_6, 250 MHz):** *δ* 8.25 (s, 1H), 8.15 (s, 1H), 3.05 (s, 3H); **^{19}F NMR (DMSO-D_6, 235 MHz)**: δ -68.06 (s); **HRMS (ESI): m/z (M+H)**$^+$ calculated for $C_7H_6N_5F_3$: 218.06480 found : 218.06477.

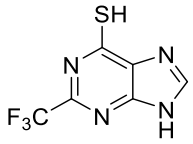
2-(Trifluoromethyl)adenine 17 (CAS: 2993-06-8): TX12



To a solution of 6-chloro-2-trifluoromethyl-9H-purine (0.150 g, 0.673 mmol, 1eq) in a 5:1 mixture of EtOH/H_2O (7.5 mL / 1.5 mL), NaN_3 (131 mg, 2 mmol, 3 eq) was added. The mixture was stirred for 2 h at 80°C and evaporated under reduced pressure. The crude product was dissolved in EtOH (9 mL) and $SnCl_2$, 2H20 (456 mg, 2 mmol, 3 eq) was added. The solution was refluxed for 2 h and concentrated in vacuo. The product was washed with EtOAc and dried under reduced pressure.

White solid (67%), **IR:** 3345, 3154, 3068, 3033, 2850, 1668, 1615, 1436, 1189, 1141, 1129, 1052, 950, 805; **^{1}H NMR (DMSO-D_6, 250 MHz):** δ 13.52 (brs, 1H, NH), 8.34 (s, 1H, H-8), 7.79 (s, 2 H, NH_2); **^{19}F NMR (DMSO-D_6, 235 MHz):** δ -67.9 (s, 3F) ;**HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_6H_5F_3N_5$: 204.04915 found : 204.04890.

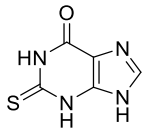
2-(trifluoromethyl)-Purine-6-thiol 18 (CAS: 1996-33-4): TX13



A solution of **14** (200 mg, 1.19 mmol, 1 eq) and thiourea (397 mg, 5.24 mmol, 4.42 eq) in ethanol was stirred at reflux for 4 h. After evaporation, the residue was purified by silica gel column chromatography (EtOAc/MeOH, 8:2).

Beige solid (54%), **IR:** 3348, 2724, 2358, 2196, 1652, 1580, 1551, 1418, 1356, 1191, 1137, 1007, 892, **^{1}H NMR (MeOH-D_6, 250 MHz)**: δ 8.25 (s, 1H). **HRMS (ESI): m/z (M+H)** $^+$ calculated for $C_6H_4F_3N_4S$: 221.01032, found : 221.01035.

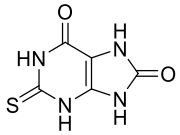
2-THIOXANTINE 19 (CAS: 2487-40-3) (Commercial) : 2TX



The compound 4,5-diamino-6-hydroxy-2-mercaptopyrimidine (MW 158, Aldrich 95%, used as supplied, 500 mg) was refluxed with 90% formic acid (50 mL) for 2 h in a 3-neck flask with mechanical stirring. The mixture initially solidified and an additional 0.5 mL formic acid was added. The mixture was cooled on ice and crude 4-amino-5-formamido-6-hydroxy-2-mercaptopyrirnidine was filtered on a large sinter. It was covered with a rubber dam and dried under vacuum for 30 min giving a light-yellow product. The filtration cake was suspended in formamide (1.12 mL) in a large beaker and heated at 175-185 °C (bath temperature) in a liquid paraffin bath with occasional hand stirring for 2 h. Considerable frothing occurred in the early stages. The mixture was cooled to room temperature and filtered with a dam on a large sinter, then dissolved in approximately 10 mL of 1M NaOH and precipitated with glacial AcOH at 10 °C. It was then filtered, pressed dry and vacuum dried at 95 °C for 2 h.

Light yellow Solid (97%), **IR:** 3585, 3100, 1678, 1565, 1422, 1209, 977, 950, 804, 517, **^{1}H NMR (DMSO-*d_6*, 250 MHz):** δ 13.40 (s, 2H), 12.19 (s, 1H), 8.06 (s, 1H); **HRMS (ESI): m/z (M+H)**$^+$ calculated for $C_5H_5N_4OS$: 169.017858, found :169,01162.

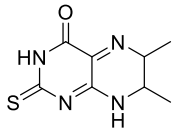
2-Mercapto-6, 8-purinediol 20 (CAS: 15986-31-9)**: TX14**



The compound 5,6-diamino-2-thiouracil (1g, 6.3 mmol, 1 eq) and urea (1.96 g, 32.7 mmol, 5 eq) were mixed in a flask with stirring and heated to 170 °C for 30 min. The agitation became difficult. Sodium hydroxide was added to dissolve all solid material. The hot solution was filtered. The hot filtrate was acidified with concentrated hydrochloric acid. The precipitate was filtered and washed by water and acetone.

Yellow Solid (90%), **IR:** 3675, 2987, 1933, 1589, 1505, 1465, 1250, 1066, 858, 656, 544**.** , **^{1}H NMR (DMSO-D$_6$, 250 MHz):** δ 12.19 (s, 1H), 11.51 (s, 1H), 11.00 – 10.76 (s, 1H). **HRMS (ESI): m/z (M+H)$^+$** calculated for $C_5H_5N_4O_2S$:185.01277, found : 185.01300.

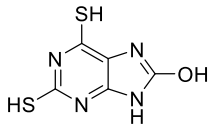
6, 7 –dimethyl-4-hydroxy-2-mercaptopterdine 21 (CAS: 54030-51-2)**: TX15**



The compound 5,6-diamino-2-thiouracil (0.58 g, 3.67 mmol) was dissolved in DMF (20 mL) and diacetyl (0.63 g, 7.34 mmol) was added. The solution was stirred for 5 h at 100 °C, and then the solvent was evaporated under reduced pressure. The resulting orange oil was triturated with ether. The resulting solid was filtered and washed with ether.

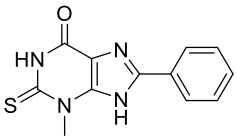
Beige Solid (67%), **IR**: 3675, 3188, 2987, 1694, 1539, 1393, 1378, 1270, 1127, 1065, 502 ; **^{1}H NMR (DMSO-d$_6$, 250 MHz) δ** 13.13 (s, 1H), 12.60 (s, 1H), 2.56 (s, 3h), 2.53 (s, 3h) ;**HRMS (ESI): m/z (M+H)$^+$** calculated for $C_8H_9N_4OS$: 209.04915, found :209.04936.

1,2,3,6,7,9-hexahydro-2,6-dithioxo-8H-Purin-8-one 22(CAS : 6703-93-1)**: TX16**



The compound 4,5-diamino-2,6-dimercaptopyrimidine (1.72 mmol, 1eq) and urea (8.95 mmol, 5.2 eq) were mixed in a flask with stirring and heated to 170 ° C for 45 min. The agitation became difficult. Potassium hydroxide was added to dissolve all solid material. The hot solution was treated with Norit and filtered. The hot filtrate was acidified with concentrated hydrochloric acid. The precipitate was filtered and washed by water.
Yellow solid (60%), **IR:** 3006, 1694, 1584, 1486, 1358, 1302, 1169, 1132, 990, 840, 721; **HRMS (ESI): m/z (M+H) + calcd for** $C_5H_5N_4OS_2$: 200.98992, found: 200.99019.

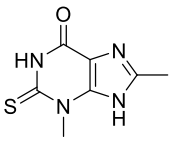
3-Methyl-8-phenyl-2-thiohypoxanthine 23 (CAS: 103258-00-0)**: TX17**



A mixture of 5, 6-diamino-1-methyl-2-thioxo-2,3-dihydro-1H-pyrimidin-4-one (1 eq) and (triethoxymethyl)benzene (10 eq) was heated under reflux for 15h. Cooling the solution yielded the desired product.

Yellow Solid (90 %), **IR**: 3675, 3126, 2987, 1678, 1405, 1217, 1171, 1066, 781, 703 ; **^{1}H NMR (DMSO-d$_6$, 250 MHz)** δ 3.74 (s, 3H), 2.36 (s, 3H) ; **HRMS (ESI): m/z (M+H)$^+$** calculated for $C_{12}H_{11}N_4OS$: 259.06481, found: 259.06479.

3, 8-dimethyl-2-thioxanthine 24 (CAS : 91725-06-3)**: TX18**



A mixture of 5, 6-diamino-1-methyl-2-thioxo-2, 3-dihydro-1H-pyrimidin-4-one (1 eq) and an triethylorthoacetate (10 eq) was heated under reflux for 15h. Cooling the solution yielded the desired product.

Yellow solid (60%), **IR**: 3675, 2987, 1405, 1250, 1065, 891, 510, 501 ; **^{1}H NMR (DMSO-d$_6$, 250 MHz)** δ 3.74 (s, 3H), 2.36 (s, 3H) ; **HRMS (ESI): m/z (M+H)$^+$** calculated for $C_7H_8N_4OS$: 197.049158, found :197.049281.

• Supplementary tables

Table S1: Amino acid residues of free and DNA bound LlFpg and hNeil1 involved in the docking site areas for $2TX_{Red}$ and $TX19_{Red}$

	Site I	Site II
Fpg/DNA	P1 E2 P4 E5 V6 T8 T83 D85 A86 P87 R88 E89 V108 I172 G215 G216 S217 S218 I221 Y222 S223 A224 L225 GG226 S227 T228	K154 P158 Y159 L161 E162 Q163 A258
Free Fpg	P1 E2 P4 E5 V6 V9 G216 S217 S218 I219 A224 L225 G226 S227	R55 G56 K57 Y58 K129 I131 Q163 G168
hNEIL1/DNA	E5 Y176 V235 V236 Q237 L238 G239 G240 R241 Y243 G244 S245 G248 E249 D251 F252 A253 F255	R158 C161 E162 A163 L165 D166 Q167 R168 G174 D272 R273 H274 R276
Free hNEIL1	L6 G239 Y243 G244 S245 E246 F252 A253 F255 R256	K53 E54 L165 D166 Q167 F170 N171 G172 I173 G174 N175

Table S2: Amino acid residues of free and DNA bound Fpg and hNeil1 defined the area of docking sites I, II and III for the disulfide forms $2TX_{Ox}$ and $TX19_{Ox}$ ($TX19_{Ox1}$ and $TX19_{Ox2}$, see Figure 6)

	Site I	Site III	Site II
Fpg/DNA	P1 E2 S217 S218 I219 R220 T221 Y222 S223 A224	-	Y159 L160 L161 E162 Q163 T164 L165 V166 A167 G168 L169 G170 V257 A258 G259 R260
Free Fpg	**P1 E2** M75 E76 G216 S217 S218 I219 G226 Q232	**P1 E2 K57** M75 G170 N171 I172	**K57** K129 K130 I131 G132 P133 G134 P135 E162 Q163 A167 G168 L169
hNEIL1/DNA	P1 E2 E5 V235 V236 Q237 L238 G239 G240 R241 G242 Y243 G244 S245 G248 E249 F252 A253 F255	-	L165 D166 Q167 R168 F169 F170 N171
Free hNEIL1	**P1 E2** L8 S10 S83 F84 G239 G240 G244 S245 E246 G248 E249 A253 F252 F255	**P1 E2** E3 E5 L6 **K53 N175** Y176 L177 P232 K233 V236 Q237	**K53** E54 L55 D166 Q167 R168 G172 I173 G174 **N175**

The site III is observed only with $TX19_{Ox}$ and is not accessible in the enzymes bound to DNA. Amino acids in bold are shared by the site III and the other binding sites.

Table S3: X-ray data collection and refinement statistics

	2TX	TX13	TX15	TX19	TX19 (TCEP)	TX20	TX27
Data collection statistics							
Radiation source	PROXIMA-1	ID30A-3	PROXIMA-1	PROXIMA-1	PROXIMA-2	PROXIMA-1	ID30A-3
Wavelength (Å)	0.97857	0.96770	0.97857	0.97857	0.98010	0.97857	0.96770
$P4_12_12$ spacegroup with cell dimensions: $a = b$, c (Å)	91.36, 141.03	91.81, 141.62	91.67, 141.04	92.18, 139.73	91.63, 141.86	91.42, 140.28	91.96, 142.18
Resolution range (Å)	47.63-1.82 (1.92-1.82)	47.21-2.00 (2.04-2.00)	47.72-2.00 (2.11-2.00)	47.66-2.25 (2.37-2.25)	47.84-1.95 (2.00-1.95)	47.53-2.25 (2.37-2.25)	59.13-1.76 (1.79-1.76)
Total observations	352566 (50454)	292071 (15391)	239478 (34643)	106520 (15359)	213 417 (14586)	135979 (19588)	440572 (22689)
Unique reflections	53367 (7546)	41111 (2048)	41266 (5925)	29044 (4146)	44101 (3023)	28957 (4144)	61222 (3017)
Completeness (%)	98.8 (97.7)	99.1 (99.9)	99.8 (99.9)	99.4 (99.2)	99.0 (97.3)	99.9 (99.9)	99.9 (99.3)
Multiplicity	6.6 (6.7)	7.1 (7.5)	5.8 (5.8)	3.7 (3.7)	4.8 (4.8)	4.7 (4.7)	7.2 (7.5)
R_{merge}[a] (%)	5.2 (88.3)	5.6 (95.0)	6.4 (78.5)	6.0 (70.6)	6.4 (88.0)	4.8 (68.1)	5.6 (95.7)
Average $I/\sigma(I)$	18.4 (2.1)	18.7 (2.1)	13.5 (2.1)	14.3 (1.9)	14.7 (1.5)	17.3 (2.1)	18.3 (2.2)
$CC_{1/2}$ (%)	99.9 (73.0)	99.9 (66.6)	99.8 (72.9)	99.8 (99.2)	99.9 (64.8)	99.9 (73.3)	99.9 (75.6)
Refinement and model statistics							
Resolution range (Å)	47.63-1.82	45.91-2.00	45.83-2.00	41.57-2.25	47.84-1.95	47.53-2.25	59.13-1.76
Number of reflections used	53339	41035	41221	29004	44066	28910	61222
R_{work}[b] / R_{free}[c] (%)	15.58/17.23	17.39/19.50	16.58/18.55	16.35/20.19	16.75/19.56	16.00/19.79	16.92/18.75
Average B values (Å²)							
All atoms	43.3	50.1	52.6	53.3	37.4	62.1	38.5
Protein atoms	37.7	46.7	48.6	48.9	33.7	57.3	33.3
DNA atoms	60.0	63.5	67.0	72.7	47.6	81.6	50.2
Inhibitor atoms	74.1	59.9	51.0	39.3	38.6	57.0	30.0
Glycerol atoms	50.5	55.5	52.8	61.0	55.6	54.3	56.7
Water atoms	47.6	45.8	53.7	50.8	43.1	54.4	49.1
Root mean square deviation from ideality							
Bond lengths (Å)	0.007	0.008	0.008	0.008	0.016	0.011	0.010
Bond angles (°)	1.09	0.98	0.95	1.18	1.48	1.26	1.00
Ramachandran analysis (% of residues)							
Favoured regions / Allowed regions / Outliers	97.4/2.6/0.0	97.0/3.0/0.0	97.3/2.7/0.0	97.0/3.0/0.0	96.7/3.3/0.0	97.7/2.3/0.0	97.7/2.3/0.0
Number of atoms							
Protein	2117	2137	2105	2091	2168	2108	2211
DNA	557	557	557	696	557	557	557
Inhibitor	33	14	14	12	12	11	11
Glycerol atoms	6	6	6	6	12	6	12
Water atoms	281	112	177	126	312	75	449
PDBid	6RO2	6RNR	6RP7	6RP0	6ROK	6RNO	6RNM

[a] $R_{merge} = \sum_h \sum_i |I_{h,i} - \langle I \rangle_h| / \sum_h \sum_i I_{h,i}$ where $\langle I \rangle_h$ is the mean intensity of the symmetry-equivalent reflections.
[b] $R_{work} = \sum_h ||F_o| - |F_c|| / \sum_h |F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes, respectively, for reflection h.
[c] R_{free} is the R-value for a subset of 5% of the reflection data, which were not included in the crystallographic refinement.

• Supplementary figures (Figures S)

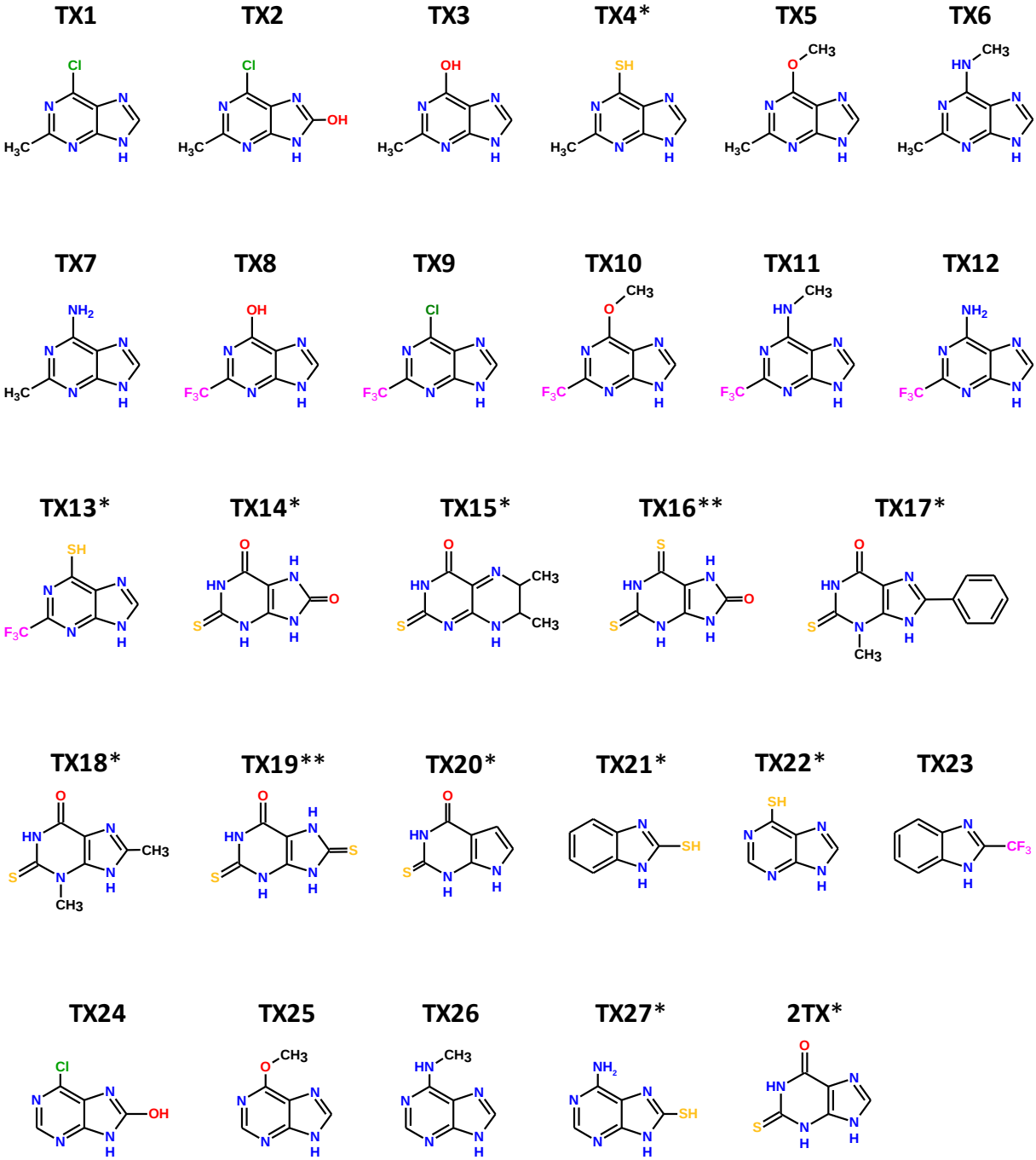


Figure S1. **Small library of thioxanthine derivatives** The design and synthesis are described in the supplementary information. * and ** indicate monothio- and dithio-compounds, respectively.

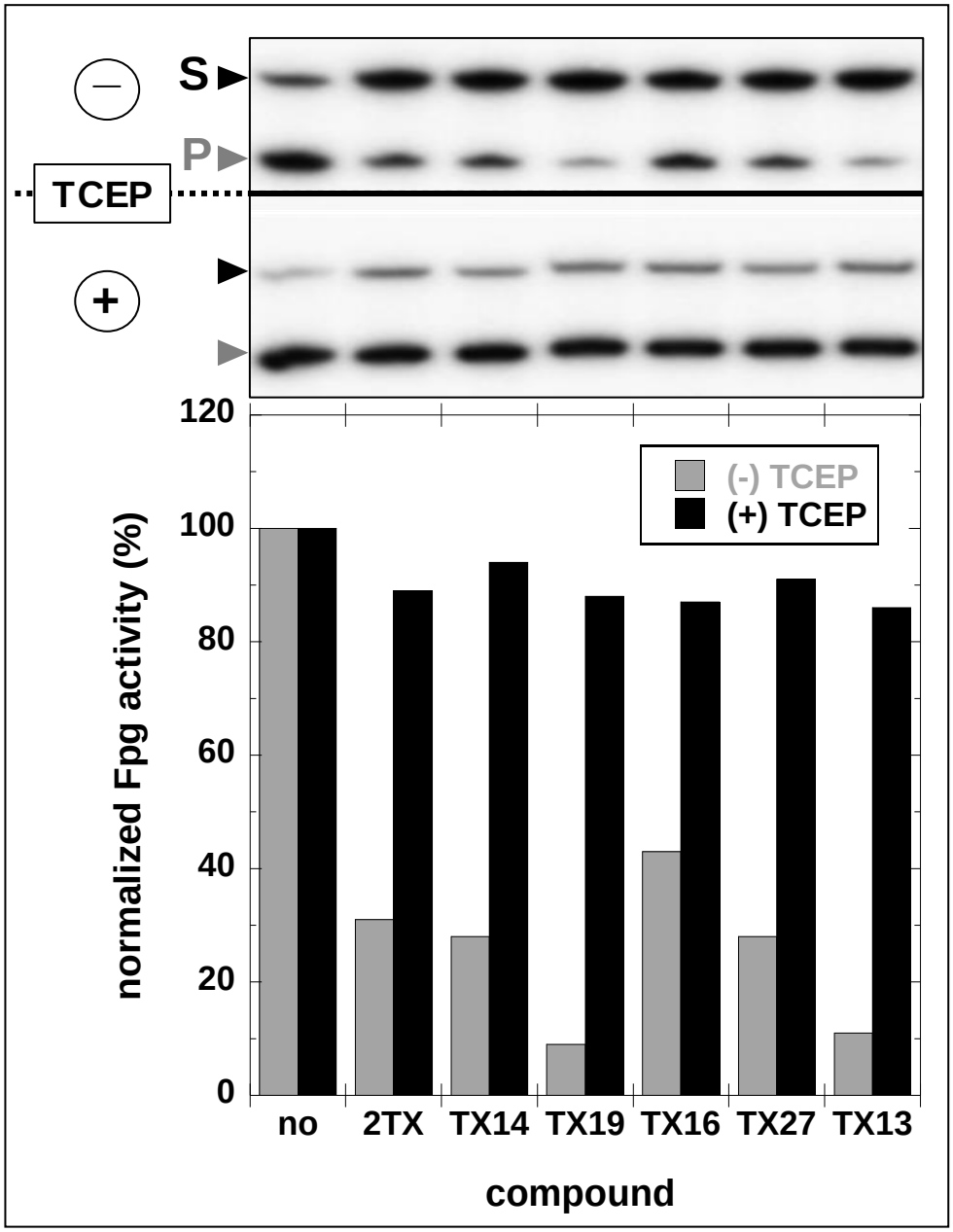


Figure S2. Effect of TCEP on 2TX and TXn-induced Fpg inhibition activity Fpg and 24-bp 8-oxoG-containing DNA duplex (S) were incubated as described in the Figure 1 caption, alone (lane 'no') or with 0.2 mM of 2TX and TXn inhibitors as indicated, in the presence (+) or absence (-) of 2 mM of TCEP. (P) indicates the Fpg DNA cleavage product. The assays were then analyzed by Urea-PAGE as described in the 'Materials and Methods'. Representative Urea-PAGE autoradiography and corresponding quantification are shown.

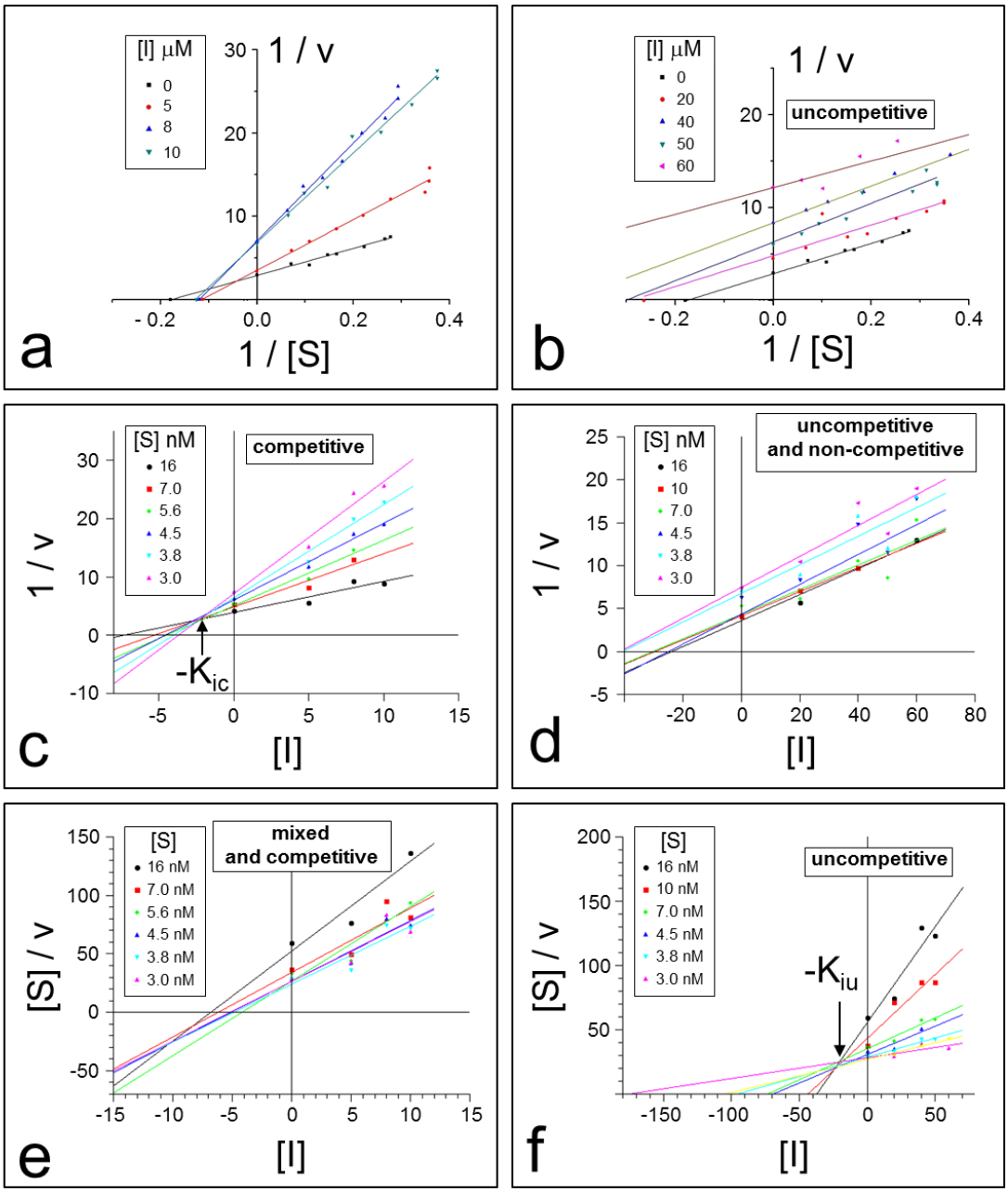


Figure S3. Double-reciprocal plots for the effects of 2TX on the velocity of *Ll*Fpg catalyzed the excision of 8-oxoG I (inhibitor) indicates 2TX and S represents the enzyme substrate, a 24-mer DNA duplex containing one 8-oxoG opposite cytosine (see the 'Materials and Methods'). K_{iu} and K_{ic} correspond to the inhibition constants. **(a)** and **(b)** Lineweaver-Burck plots, **(c)** and **(d)** Dixon plots and **(e)** and **(f)** Cornish-Bowden plots for low and high 2TX concentrations, respectively. The lines drawn are obtained by applying the following equations:

- for (a) and (b): $1/v = 1/V_{max} + (1/[S].K_m/V_{max}).(1+[I]/K_i)$;
- for (c) and (d): $1/v = [(1+K_m/[S])/(V_{max}.K_i)].[I] + 1/V_{max}.(1+K_m/[S])$;
- for (e) and (f): $[S]/v = K_m/V_{max}.(1+[I]/K_i) + [S]/V_{max}.(1+[I]/K_{iu})$

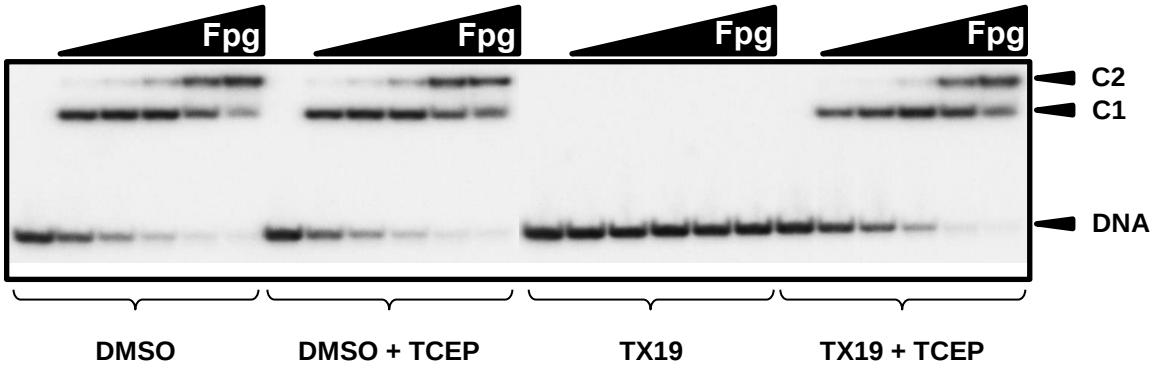


Figure S4. Effect of TCEP on the Fpg DNA binding properties in the presence of TX19

0.1 nM of 5'[^{32}P]-THF-containing 14-mer DNA duplex was incubated with 0, 0.25, 0.5, 1, 5 or 25 nM Fpg alone or after 30 min Fpg preincubation at 4°C with 0.3 mM TX19 and in the presence or absence of 1 mM TCEP. At equilibrium, incubation mixtures were analyzed by EMSA as described in the 'Materials and Methods'.

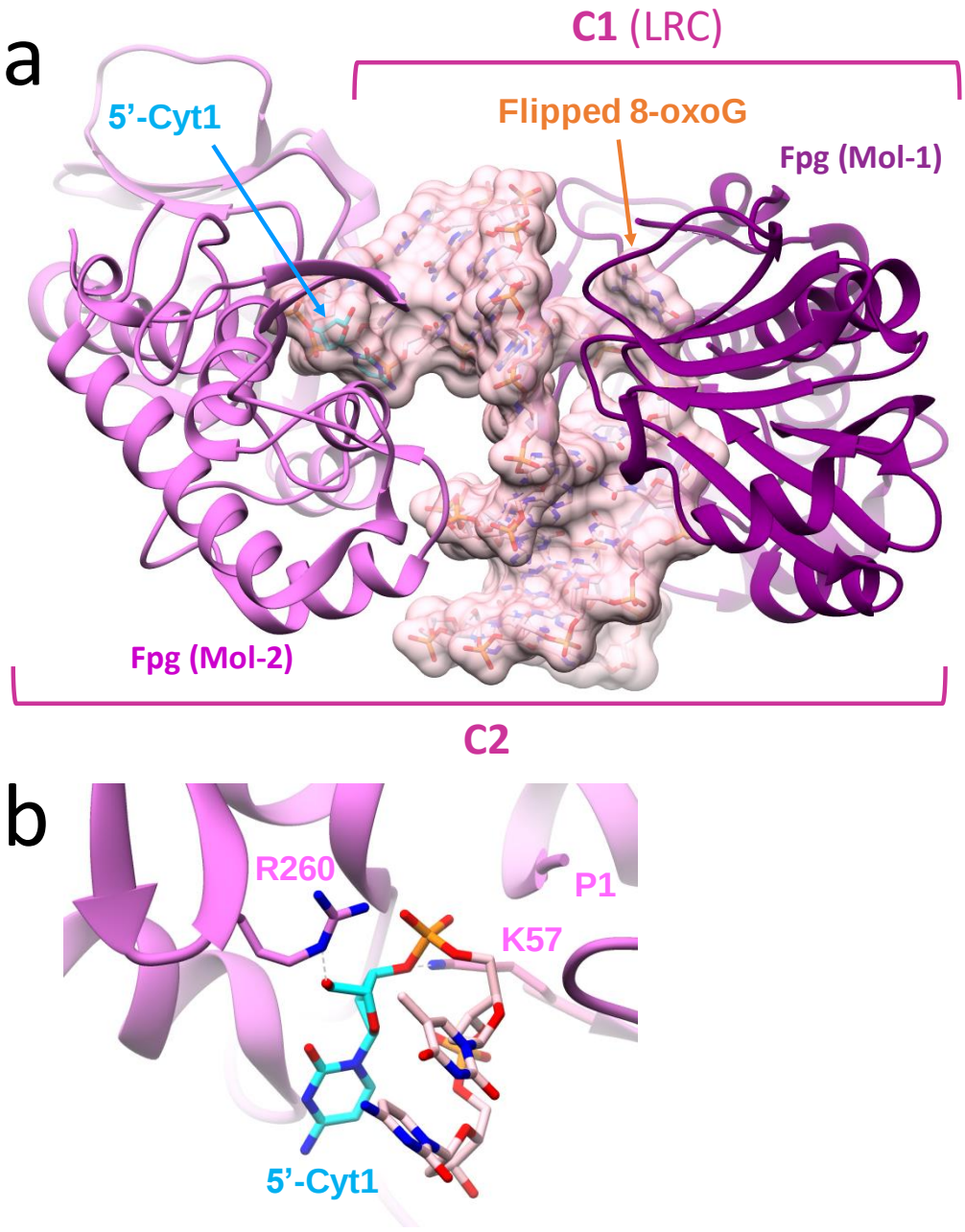


Figure S5. Crystal structure of LlFpg bound to 14-mer c8-oxoG:C DNA duplex (PDB id code 4CIS) **(a)** An overview. **(b)** Zoom of the interaction of a second Fpg molecule (Mol-2). Mol-2 binds to the overhanging 5'-cytosine-containing damaged strand (5'-Cyt(1)) following the binding of a first Fpg molecule (Mol-1) on the base damage 8-oxoG (C1, LRC, lesion recognition complex). C2 might correspond to the C2 complex observed in EMSA experiments with LlFpg (see text). c8-oxoG indicates the carbanucleoside analog of 8-oxoG.

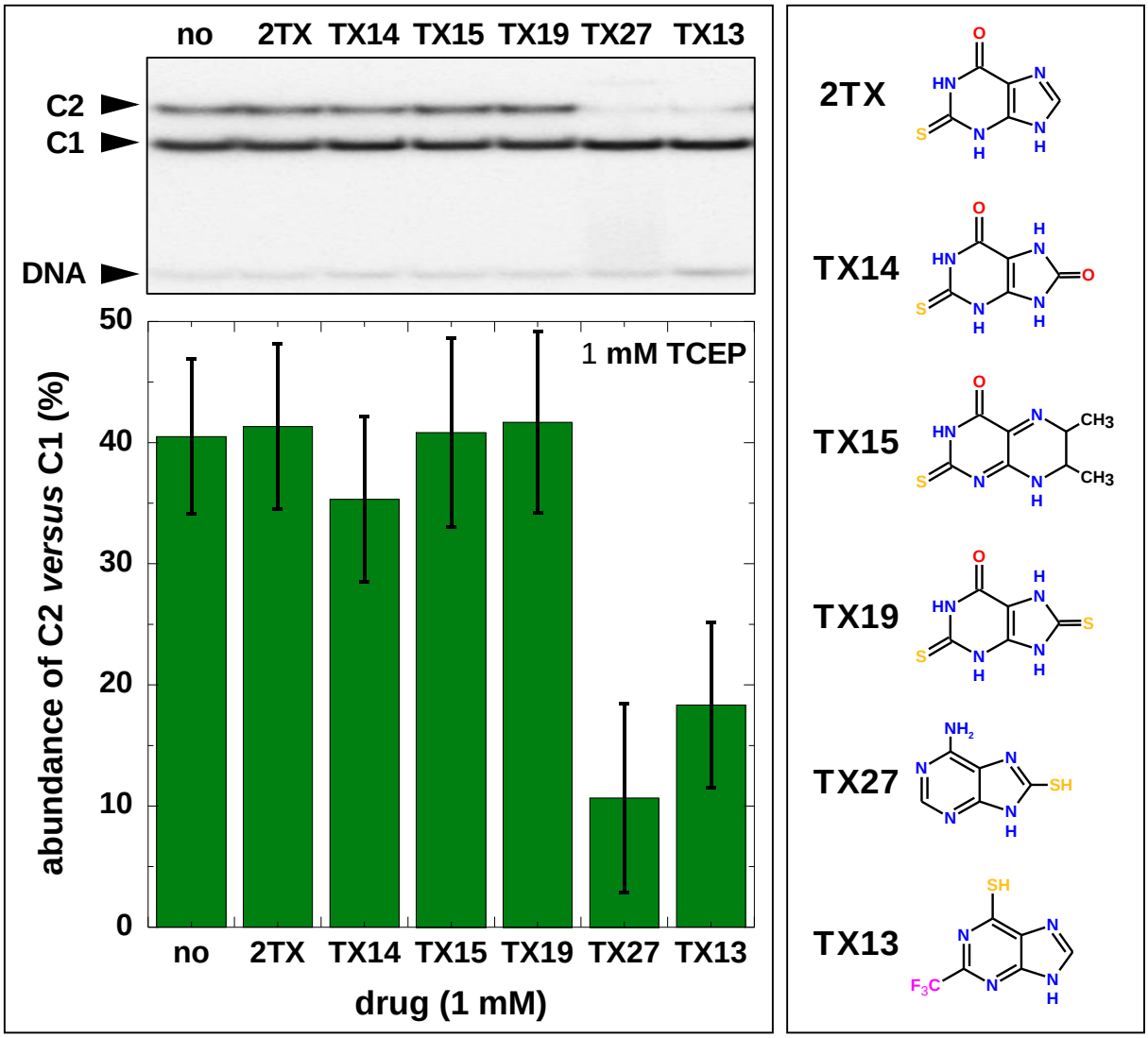


Figure S6. Destabilization of the C2 Fpg/DNA complex by TX13 and TX27 under moderate reducing conditions *Left panel*: at the top is a representative autoradiography of an EMSA experiment that shows the effect of the indicated thiocompounds (1 mM) on the C2 complex. Binding experiments were performed with a limited Fpg concentration, namely the amount required to form approximately 40% of C2; at the bottom: EMSA quantification from at least three independent experiments. *Right panel*: chemical structures of 2TX/TXn used in this experiment.

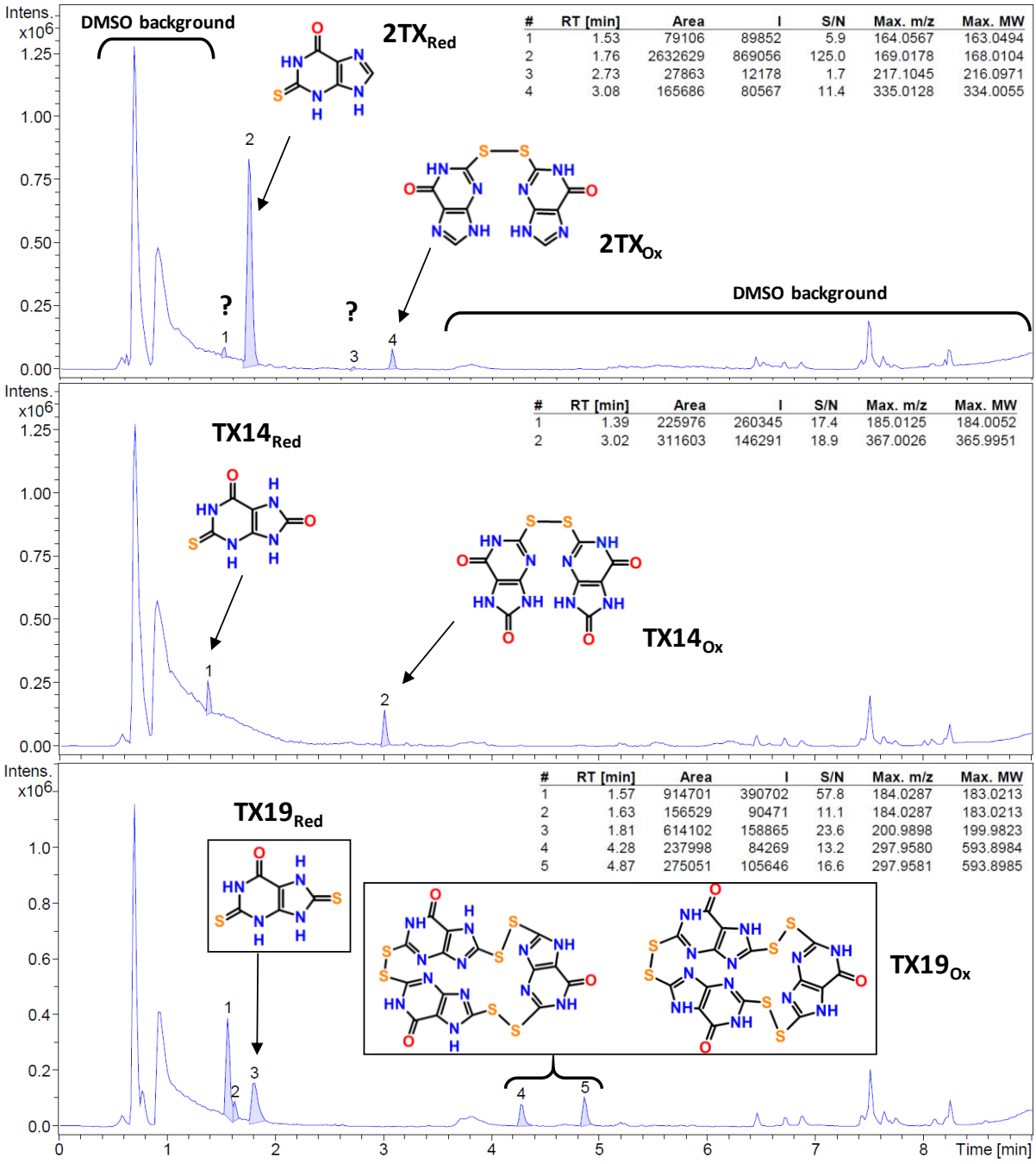


Figure S7. LC-HRMS spectra of 2TX, TX14 and TX19 Stock solutions in 80% DMSO were stored at -80°C. An aliquot of each sample at 1.2 mM was freshly thawed and diluted 1:28 with water and loaded onto a Zorbax Eclipse Plus C18 column (2.1x150 mm, 1.8 mm; Agilent). LC conditions are described in the '*Materials and Methods*'. The peaks highlighted by a *blue* background correspond to molecules present in stock solutions of 2TX, TX14 and TX19. The other peaks are related to a background noise of DMSO ions partly eluted in the dead volume of the column (proximal peaks) or resolved by the column (distal peaks).

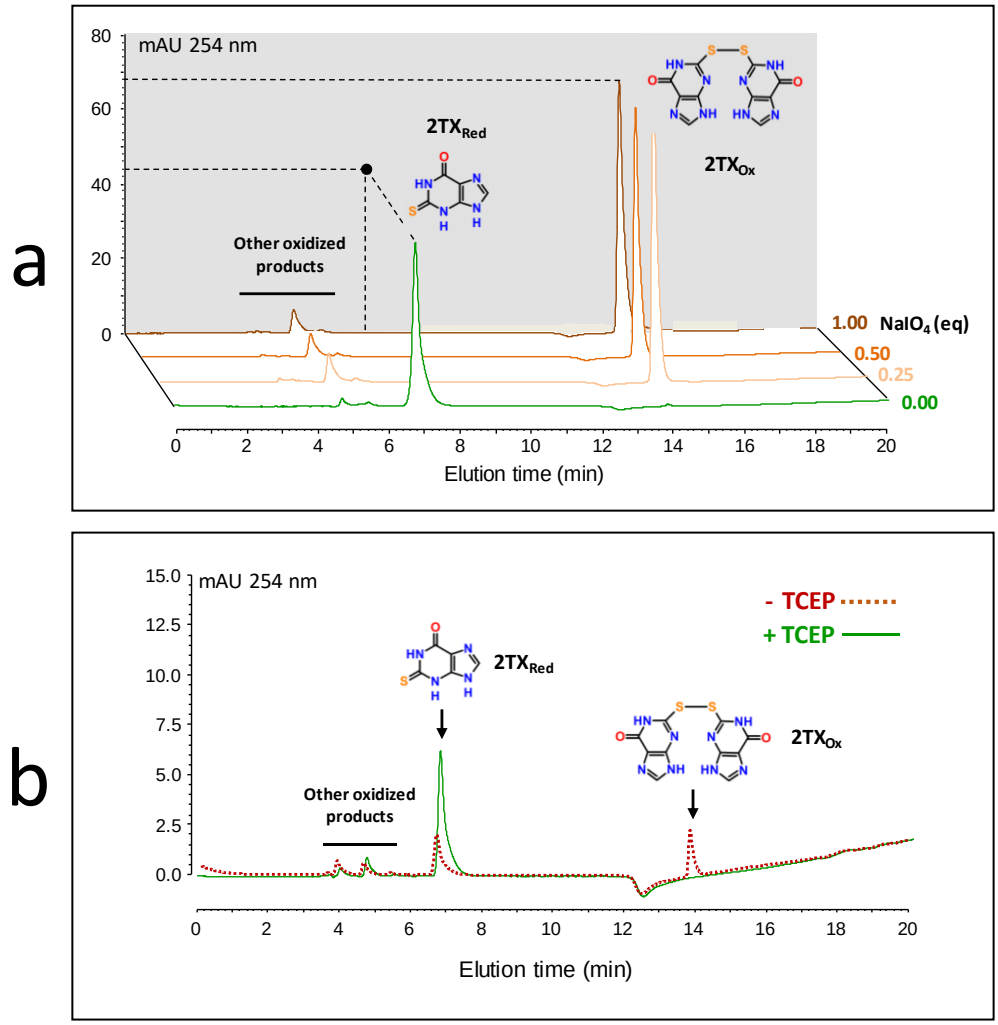


Figure S8. Preparation of 2TX$_{Ox}$ and TX14$_{Ox}$ Reduced forms of 2TX (TX14) were incubated with $NaIO_4$ and the resulting incubation mixture were then analyzed by HPLC on an RP-C4 column. The isolated oxidized species were analyzed by LC-HRMS as described in the '*Materials & Methods*' . **(a)** Oxidation reaction of one equivalent of 2TX$_{Red}$ followed by HPLC as function of $NaIO_4$ equivalent added. **(b)** Effect of TCEP on the chromatogram profile showing the interconversion of 2TX$_{Ox}$ in 2TX$_{Red}$.

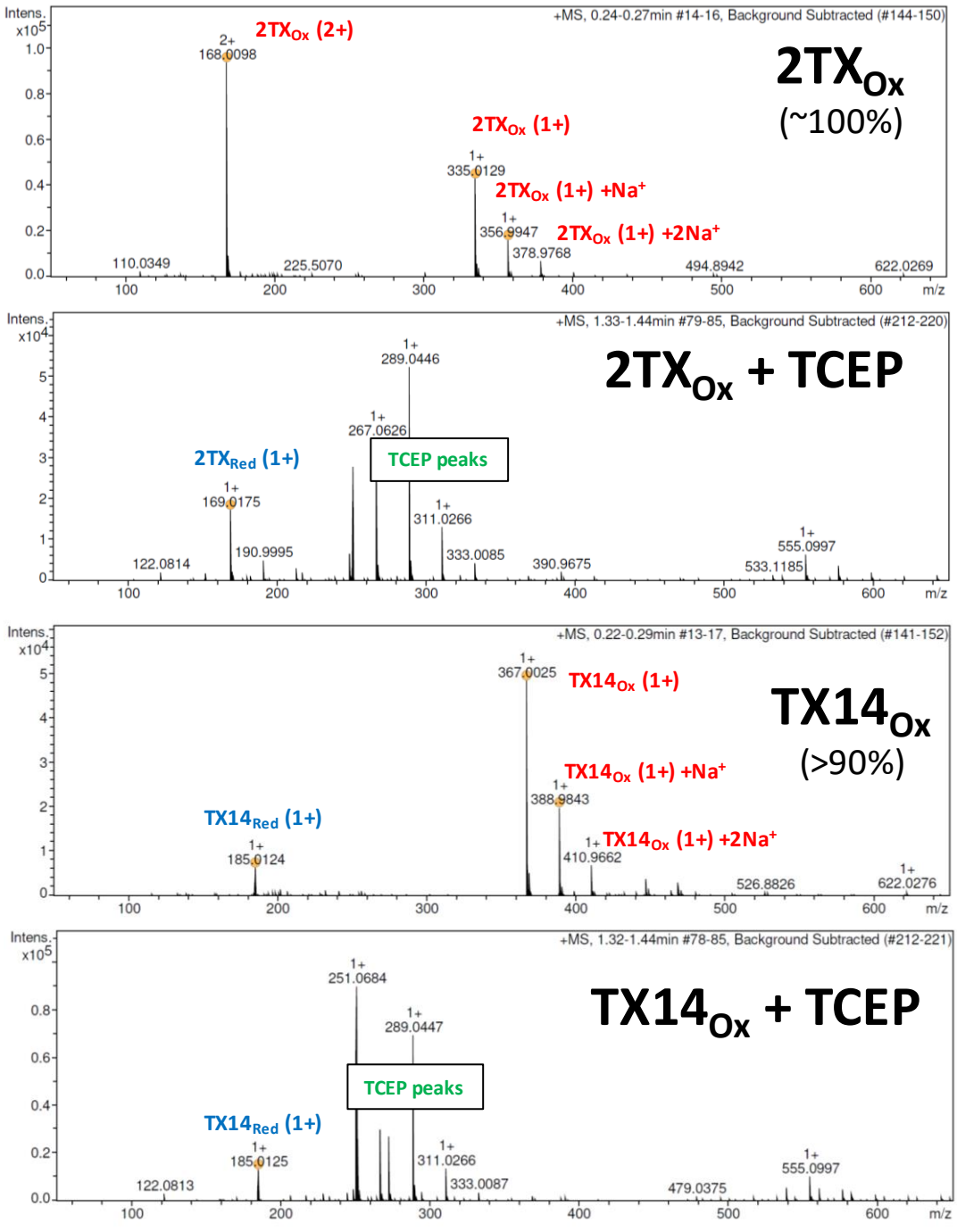


Figure S9. **High-resolution mass spectra of purified $2TX_{Ox}$ and $TX14_{Ox}$ +/- TCEP**

Figure S10. Blind and flexible docking of $TX19_{Ox}$ in pictures

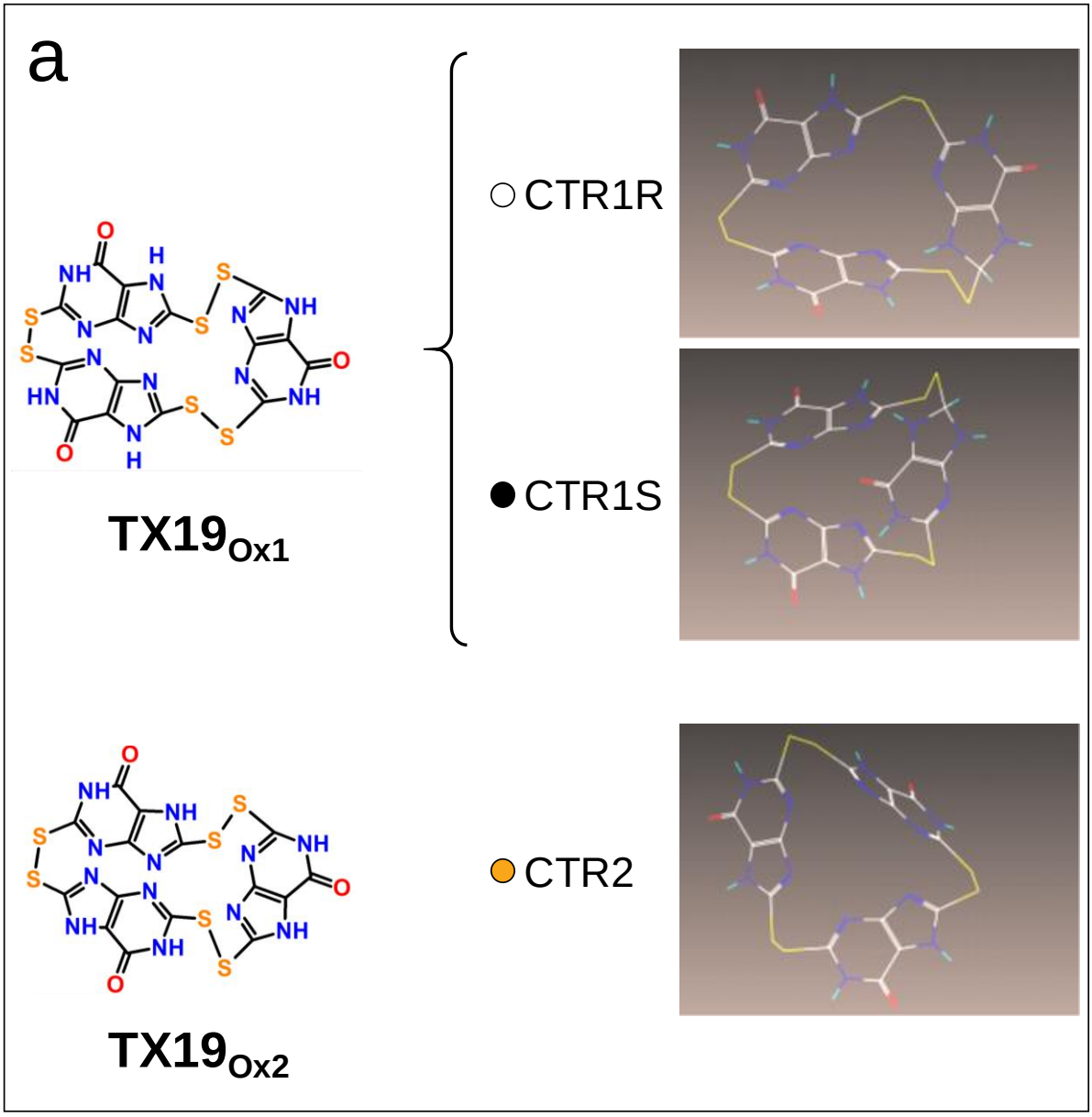


Figure S10a: *Structures of $TX19_{Ox}$* There are two possible combinations for the disulfide trimeric cyclic forms of TX19 ($TX19_{Ox1}$ and $TX19_{Ox2}$) that correspond to the two species identified by LC-HRMS (**Figure S7**). The presence of an asymmetric centre in $TX19_{Ox1}$ generates the possibility of two possible configurations: CTR1 'R' and 'S'. For $TX19_{Ox2}$, there is only one possible configuration, namely CTR2. Several representative conformers (considering the intrinsic dynamic and major tautomers at physiological pH) of each TX19 disulfide form were generated and used for blind docking on each conformation extracted from MD simulations (centroids Cn, **Figure 7**) of free and DNA bound LlFpg and hNeil1.

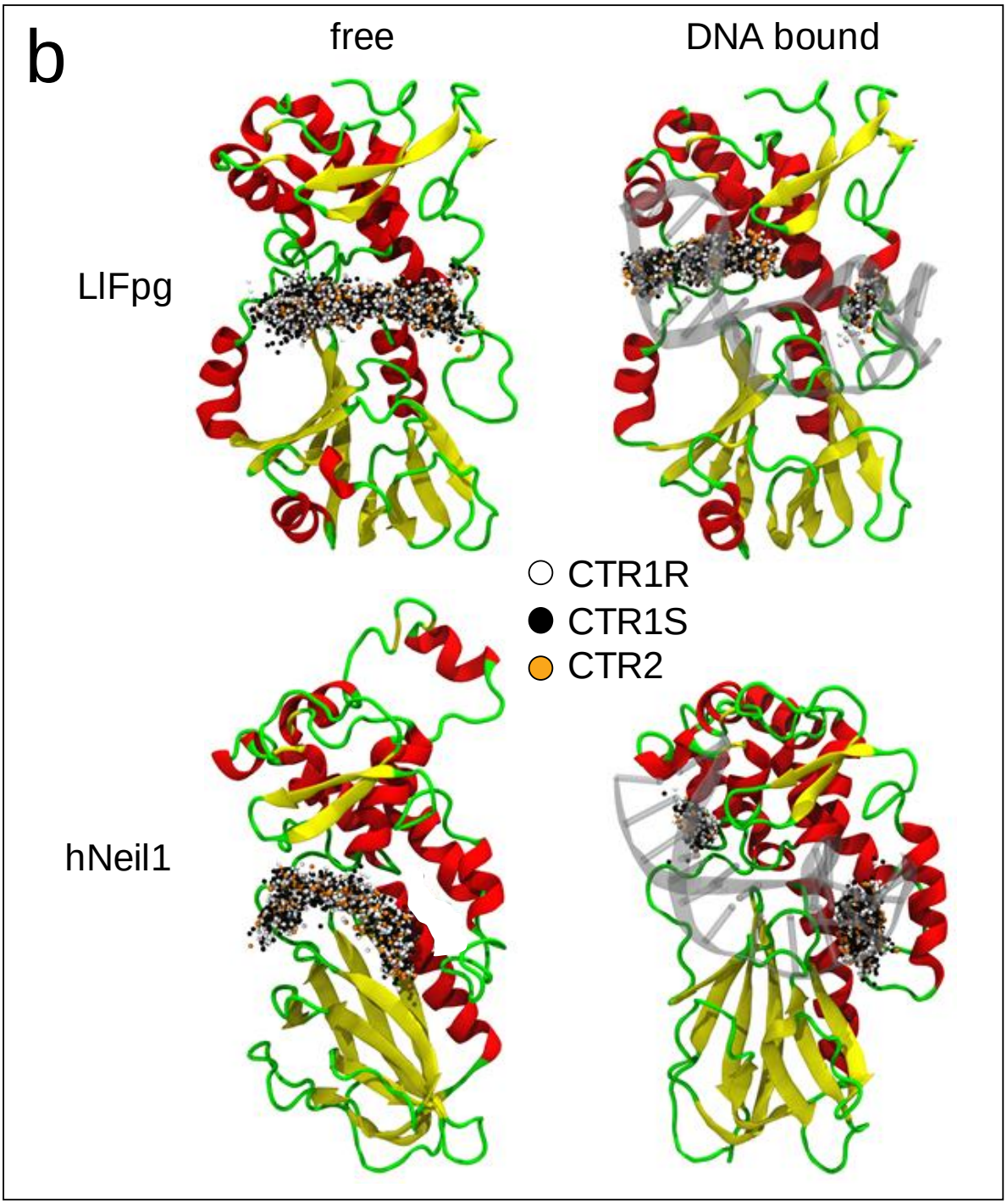


Figure S10b: ***Overview of the docking of disulfide forms of TX19 on free and DNA bound enzymes*** Ribbon presentations correspond to the crystal structures of the free and DNA bound Fpg and hNeil1 as indicated. Alpha-helices and beta-strands are coloured in *red* and *yellow*, respectively. Each docking poses obtained with the conformers of the cyclic trimeric disulfide forms of TX19 (CTR1R, CTR1S and CTR2, as indicated) are shown by small spheres that correspond to the geometric centre of the considered docked ligand.

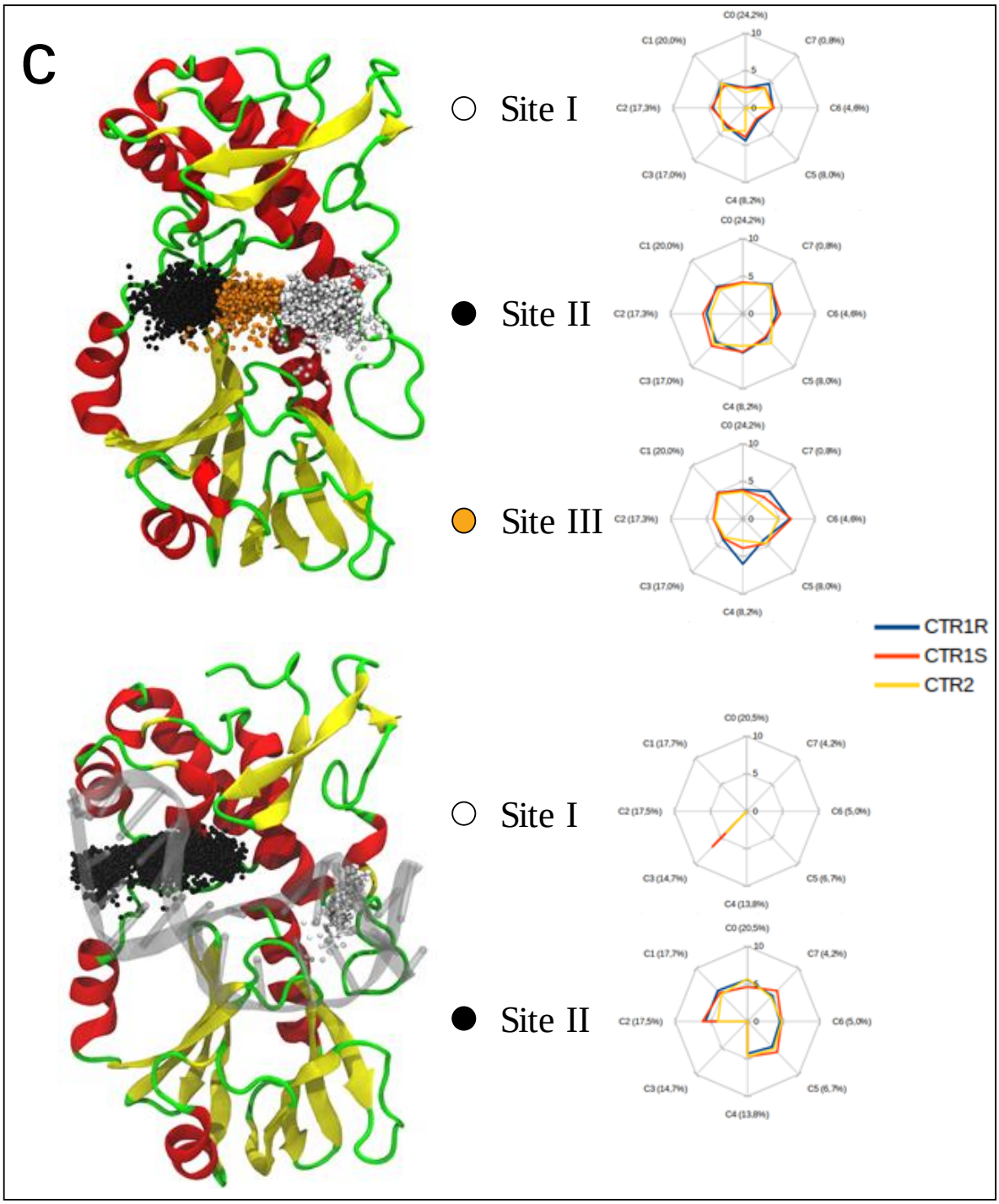


Figure S10c: ***Docking of disulfide forms of TX19 on free and DNA bound Fpg visualized on the preferential docking sites I, II and III*** Ribbon presentations correspond to the crystal structure of free and DNA bound Fpg (1PM5 without DNA and 1PM5, respectively). DNA is colored in *grey*. Docking poses obtained with the conformers of the cyclic trimeric disulfide forms of TX19 (CTR1R, CTR1S and CTR2, as indicated) are shown by small spheres that correspond to the geometric centre of the ligand. The docking pose scores obtained for sites I, II and III are plotted as function of the centroids Cn extracted from MD simulations using Kiryat (Radar) charts presented on the right of the figure.

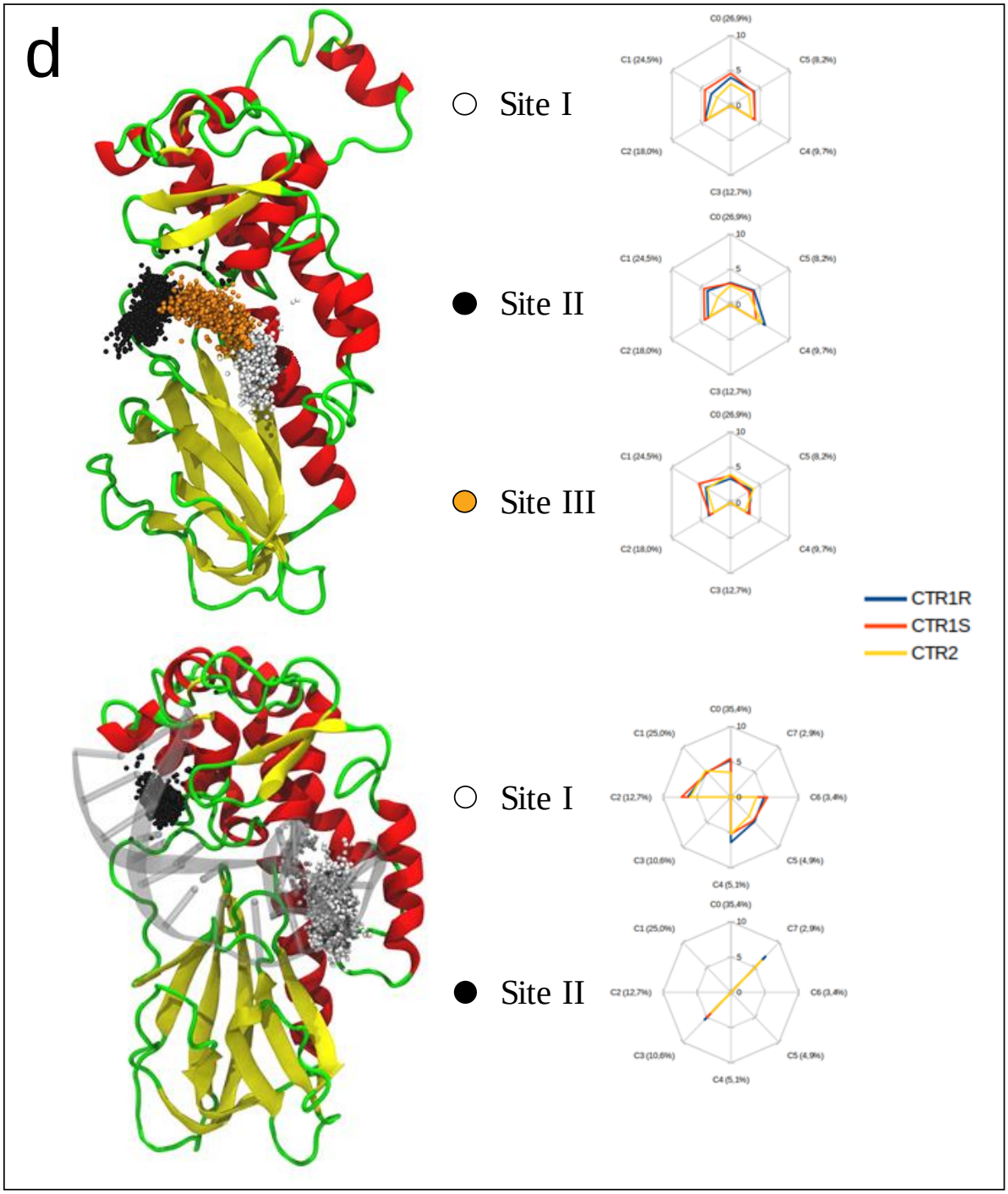


Figure S10d: ***Docking of disulfide forms of TX19 on free and DNA bound hNeil1 visualized on the preferential docking sites I, II and III*** For details, see the Figure S10c legend.

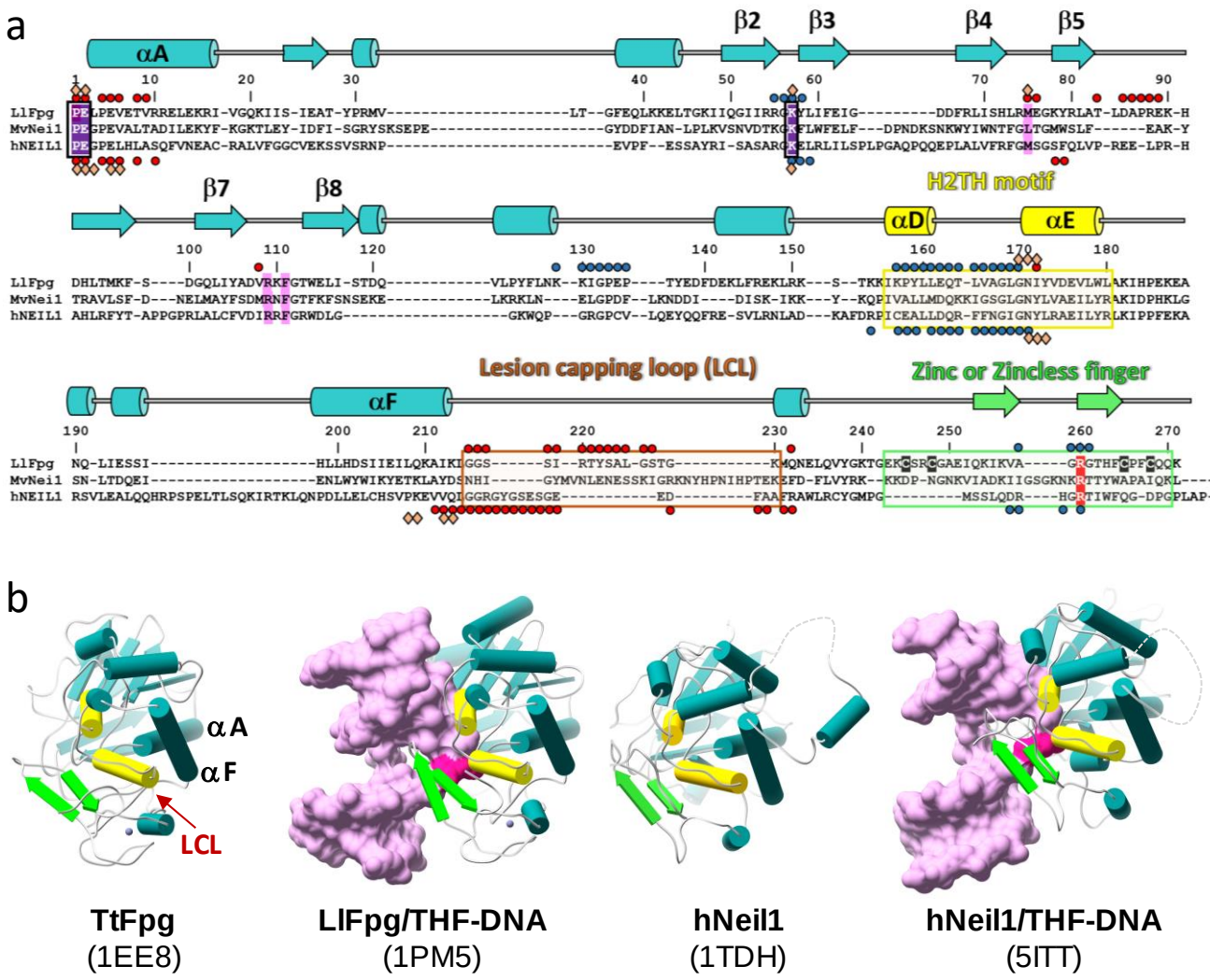
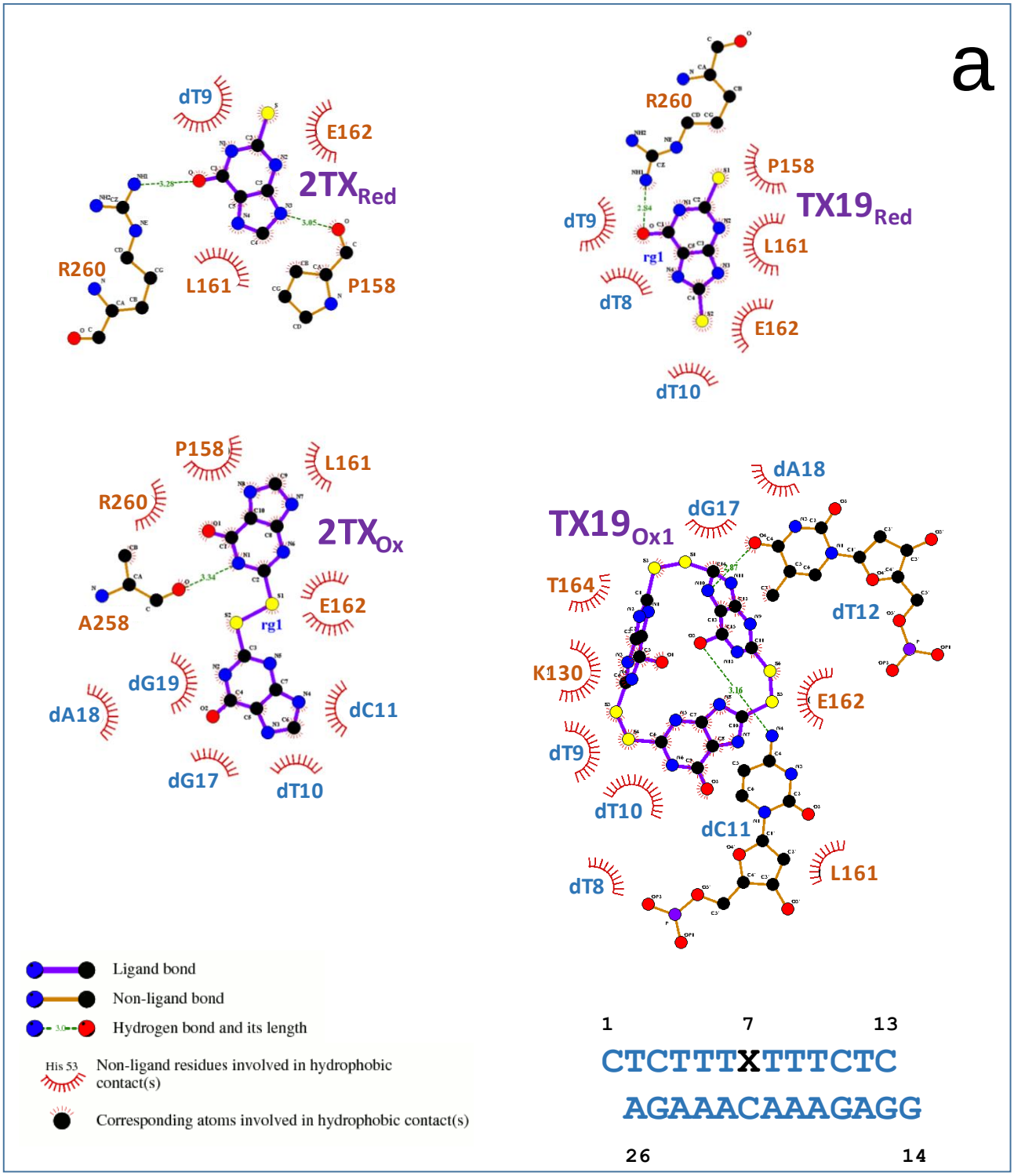
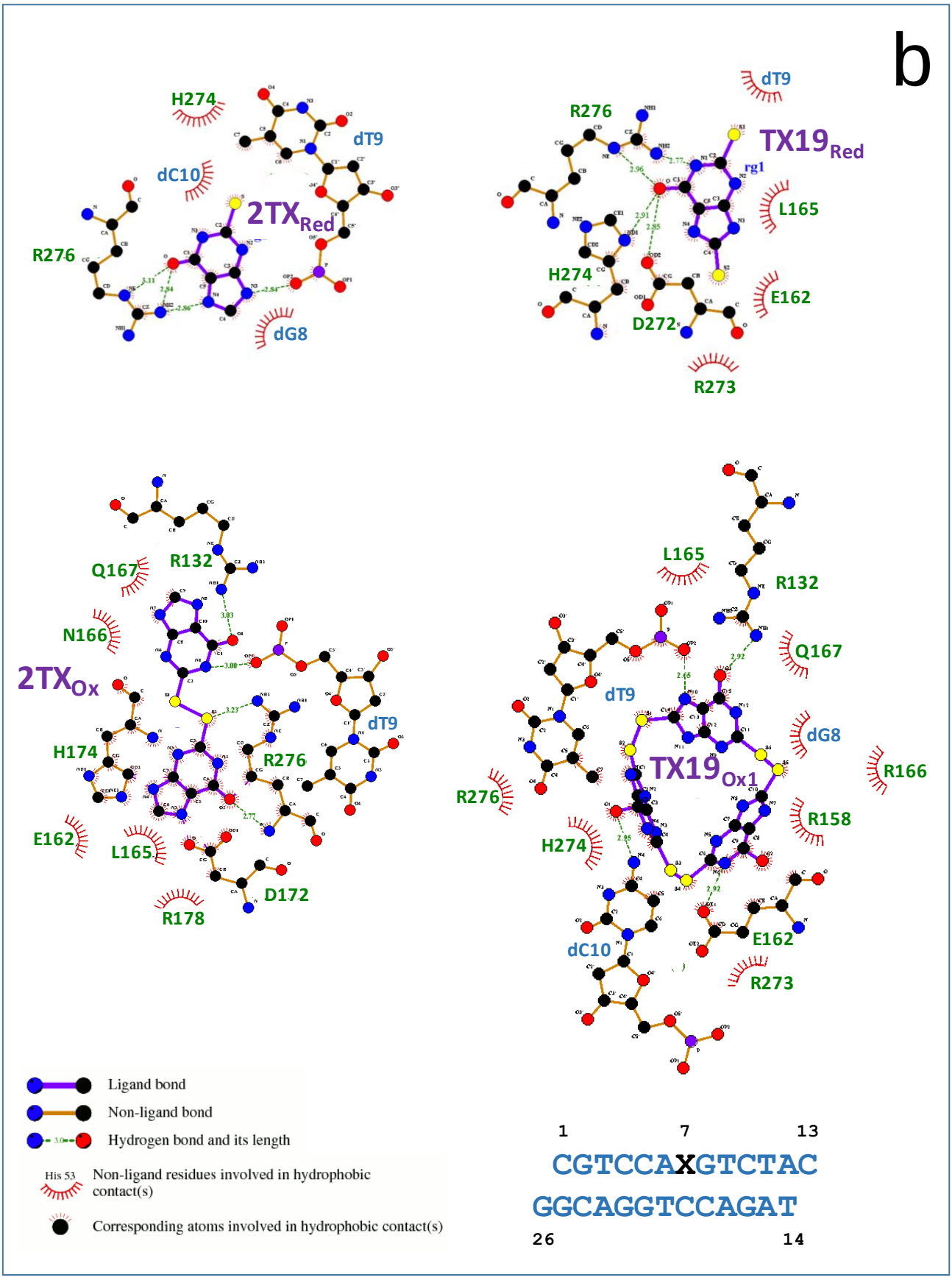


Figure S11. Structures of LlFpg, MvNei1 and hNeil1 ***(a)*** *Primary structure alignment based on 3D-structures* The indicated α-helices (cylinders) and β-strands (arrows) are from LlFpg (observed in 1PM5). P1, E2 and K57 catalytic residues are boxed in *black*. Residues of the intercalation triad are in *black* letters on *light purple* background. The strictly conserved arginine (R) in the *beta* hairpin loop of the zincless and zinc finger is indicated by a *white* letter on a *red* background. The indicated residue numbers correspond to LlFpg. Predicted unstructured, the C-terminal of MvNei1 and hNeil1 proteins has been willingly omitted. The upper (LlFpg) and lower (hNeil1) mall cycles and lozenges highlight residues found by the flexible and blind docking involved in the binding area of the reduced and disulfide forms of 2TX and TX19 (see Table S1 and S2 for more details). *Blue* and *red* cycles are for residues involved in the area of binding sites I and II. *Orange* lozenges indicate residues involved in binding site III. ***(b)*** *Overviews of crystal structures of free or DNA bound Fpg/hNeil1 enzymes used in this study* The Zn^{2+} ion of Fpg zinc finger is indicated by a small *blue* sphere. The H2TH motif and the zinc (or zincless for hNeil1 and MvNei1) finger are highlighted in *yellow* and *green*, respectively. DNA is in *light purple* and THF in *purple*. Ll, h, Tt and Mv are for *Lactococcus lactis*, human, *Thermus thermophilus* and Mimivirus, respectively. The figure is adapted from Boiteux *et al.* (2017) *Free Radical Biology & Medecine* **107**, 179-201.

Figure S12. 2D-LigPlot representations of the best docking poses at site II of reduced and disulfide forms of 2TX and TX19 in DNA bound proteins (a) and **(b)** are for LlFpg/DNA and hNeil1/DNA complexes, respectively. LlFpg and hNeil1 amino acid residues involved in the interaction with ligands are coloured in *orange* and *green*, respectively. Nucleotides involved in interactions are indicated in *blue* and numbered in accordance with the indicated sequence of DNA duplexes.




b
H274
dT9
dC10
2TX$_{Red}$
R276
dG8
dT9
R276
TX19$_{Red}$
rg1
L165
H274
E162
D272
R273
Q167
R132
N166
2TX$_{Ox}$
dT9
H174
R276
E162
L165
D172
R178
L165
R132
Q167
dT9
dG8
TX19$_{Ox1}$
R166
R276
R158
H274
dC10
E162
R273
Ligand bond
Non-ligand bond
Hydrogen bond and its length
His 53
Non-ligand residues involved in hydrophobic contact(s)
Corresponding atoms involved in hydrophobic contact(s)
1
7
13
CGTCCAXGTCTAC
GGCAGGTCCAGAT
26
14

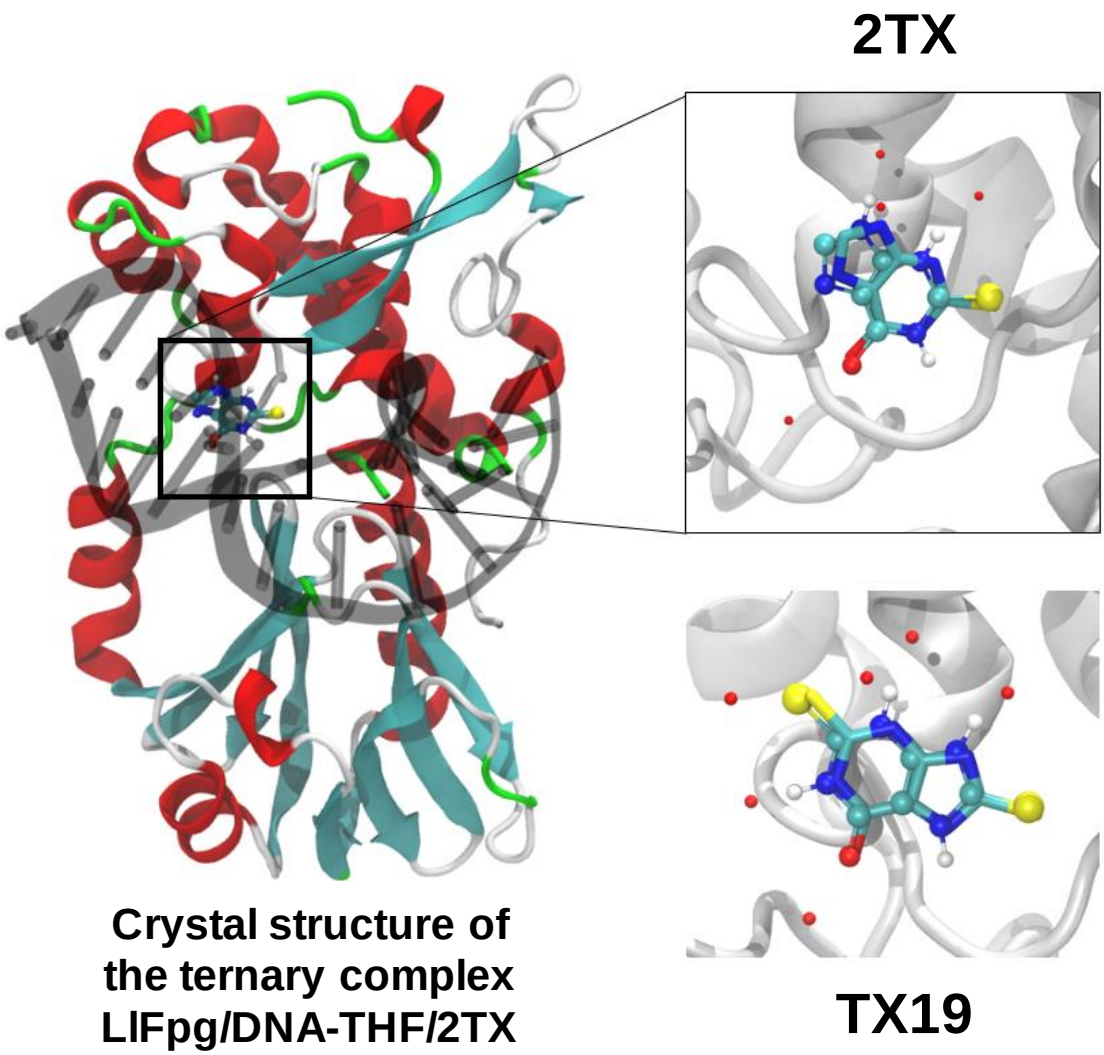


Figure S13. Comparison between crystal structures and docking experiments
Among the poses of 2TX and TX19 having the best docking scores, some of them show a mode of binding at site II that is very similar to that observed in the corresponding crystallographic structures. The selected docking poses of 2TX and TX19 are represented by ball-and-thin sticks and are superimposed on the crystal structures of Fpg/DNA/2TX(TX19) ternary complexes. 2TX and TX19 in the crystal structures are represented by thick sticks. For clarity, the docked Fpg/DNA structure is deliberately omitted. Only docked 2TX and TX19 are represented.