





DOCTORAL THESIS

mgr Marta Wyszkowska

Targeted inhibition of human mitochondrial nuclease EXOG, as a means of mitochondrial DNA repair process modulation

Ukierunkowane hamowanie aktywności ludzkiej mitochondrialnej nukleazy EXOG, jako sposób modulacji procesu naprawy mitochondrialnego DNA

A Ph. D. dissertation presented to the Biotechnology Discipline Council of the University of Gdańsk for the award of the doctoral degree in the field of Natural Sciences, in the scientific discipline of Biotechnology

Supervisor: dr hab. Michał R. Szymański, prof. UG
Laboratory of Structural Biology

GDAŃSK 2026

Dokument zanonimizowany za pomocą UG DLP Anonimizator, 2026-10-08 11:49:52

This work was supported by the First Team Grant
(POIR.04.04.00-00-3E44/17-00)
from the Foundation for Polish Science



To my family, whose constant support, patience, and faith in me have been my greatest source of strength.

TABLE OF CONTENTS

1. ABSTRACT **7**

2. STRESZCZENIE **9**

3. ABBREVIATIONS **11**

4. INTRODUCTION **15**

4.1. Mitochondria 15

4.2. Mitochondrial genome 17

4.3. Mitochondrial DNA repair 18

4.3.1. EXOG 22

4.4. Mitochondrial DNA and cancer 25

4.5. DNA damaging agents in cancer chemotherapy 27

4.6. Small-molecule compounds 28

4.7. Strategies for identifying small-molecule EXOG inhibitors 29

4.7.1. Experimental high-throughput screening (HTS) 29

4.7.2. Structure-based virtual screening 30

4.7.3. Compound prioritization and drug-like properties 30

4.8. Post-screening validation and characterization of inhibitor candidates 31

4.8.1. Biochemical validation of candidate inhibitors 31

4.8.2. Selectivity and off-target assessment 31

4.8.3. Mechanistic and structural characterization of inhibitor binding 32

5. AIM OF THE PROJECT **33**

6. MATERIALS **34**

6.1. Strains 34

6.2. Plasmids 34

6.3. Broths and media 34

6.4. Antibiotics 35

7. METHODS **38**

7.1. Preparation of Escherichia coli competent cells 38

7.2. Plasmid DNA isolation 38

7.3. Transformation of E. coli competent cells 38

7.4. Protein purification 38

7.4.1. Human mitochondrial 5'-exonuclease (EXOG) wild-type and variants 38

7.4.2. APE1 39

7.4.3. RNase H1 40

7.4.4. Pol γA 41

7.4.5. Pol γB 42

7.4.6. FEN1 43

7.4.7. MGME1 44

7.4.8. Ligase III 45

7.5. High-throughput small-molecule screening (HTS) of LOPAC®1280 library 46

7.6. EXOG-ATA co-crystallization and data collection 47

7.6.1. Crystallization trials 47

7.6.2. Data collection, structure solution and refinement 48

7.7. Endonuclease activity assay 50
7.8. Inhibition of endonuclease activity 50
7.9. Titration of the ΦX174 DNA with aurintricarboxylic acid and mitoxantrone 51
7.10. DNA and RNA substrates preparation 51
7.11. Exonuclease activity assay 51
7.12. Inhibition of exonuclease activity 52
7.13. ATA inhibition of off-target proteins 53
7.13.1. FEN1 53
7.13.2. RNase H1 53
7.13.3. Pol γAB 54
7.13.4. MGME1 54
7.13.5. Ligase III 54
7.13.6. APE1 54
7.14. Testing of the ATA-related compounds on EXOG exonuclease activity 54
7.15. Biolayer interferometry (BLI) 55
7.15.1. EXOG – ATA binding 55
7.15.2. EXOG-DNA binding 55
7.16. Protein thermal shift assay 56
7.17. Long amplicon PCR assay 56
7.18. MTT assay 57
7.19. LDH assay 57
7.20. Virtual screening of compound libraries 57
7.21. Optimization of the high throughput small-molecule screening (HTS) with the ssDNA substrate 58
7.22. Biochemical testing of virtual-screening derived candidate EXOG inhibitors 59
8. RESULTS AND DISCUSSION 61
8.1. Increased mtDNA integrity coincides with EXOG overexpression in lung adenocarcinoma 61
8.2. Identification of hEXOG inhibitors from the Library of Pharmacologically Active Compounds (LOPAC® 1280) 63
8.2.1. Secondary screen of compounds 68
8.2.2. Inhibition of exonucleolytic activity of EXOG on different RNA/DNA substrates 70
8.2.3. Co-crystallization and soaking of EXOG with ATA 74
8.2.4. In silico evaluation of ATA-EXOG binding 75
8.2.5. Site-directed mutagenesis and testing of changes in inhibition 77
8.2.6. Carboxyl groups of ATA contribute to inhibition of EXOG activity 80
8.2.7. ATA directly binds and stabilizes EXOG protein 82
8.2.8. Preincubation with ATA decreases EXOG-DNA binding affinity and yields stronger inhibition of exonuclease activity 84
8.2.9. ATA inhibits different proteins involved in DNA processing 86
8.2.10. Camptothecin increases the cellular effects of ATA 91
8.3. Virtual screening of compound databases 94
8.3.1. BioAscent database 94
9. SUMMARY 101
10. FUTURE PERSPECTIVES 103
11. SCIENTIFIC ACCOMPLISHMENTS 104

11.1. List of publications in peer-reviewed journals 104
11.2. Manuscripts in preparation 104
11.3. Conferences and workshops 105
11.4. Additional scientific, teaching and outreach activity 106
12. REFERENCES 107

1. ABSTRACT

Mitochondrial DNA (mtDNA) integrity is shaped by several processes including replication, turnover and mitochondrial quality control. Among the repair pathways operating in mitochondria, base excision repair (BER) is the best characterized. In some cancers, including lung adenocarcinoma, elevated mtDNA integrity is associated with cancer-cell survival. mtDNA BER relies in part on enzymes shared with the nuclear pathway, however the 5'-exonuclease EXOG is mitochondria-specific. EXOG contributes to mtDNA BER and has been implicated in the processing of specific 5' repair intermediates. The loss of its activity was reported to result in the accumulation of unrepaired single-strand breaks and, ultimately, apoptosis. These features make EXOG a candidate anti-cancer target: selective inhibition of a mitochondria-specific repair activity could sensitize cancer cells to DNA-damaging chemotherapeutics. To date, however, no selective EXOG inhibitor has been identified.

The aim of this work was to identify a small-molecule modulator of EXOG activity that could serve as a starting point for such an agent. Two complementary strategies were used: high-throughput screening (HTS) of the LOPAC®1280 library and structure-based virtual screening of the BioAscent Compound Collection.

HTS identified aurintricarboxylic acid (ATA) as an inhibitor of both the endonuclease and the 5′-exonuclease activities of EXOG. Its effect on a range of DNA and RNA substrates was characterized, the apparent K_d of the interaction was determined, and a mechanism of inhibition was proposed. ATA bound the free enzyme with high apparent affinity, markedly increased its thermal stability, and, on preincubation, lowered the affinity of EXOG for DNA - observations consistent with a model in which ATA interferes with EXOG-DNA complex formation, possibly at or near the active site. Co-crystallization and soaking did not yield interpretable electron density for the ligand, so a direct structural view of the binding site was not obtained. Residues indicated by *in silico* docking were tested by site-directed mutagenesis. However, none of the variants showed reduced sensitivity to ATA, and their contribution to binding could not be confirmed.

Critically, ATA proved highly non-specific, also inhibiting several other DNA- and RNA-processing enzymes and showing no measurable selectivity for EXOG. Nevertheless, combining ATA with camptothecin (CPT) in a lung adenocarcinoma cell

line (A549) decreased mtDNA integrity, metabolic activity and cell viability - data consistent with, but not directly demonstrating, impairment of mtDNA repair.

Virtual screening yielded three candidate inhibitors, but on further testing these also inhibited unrelated enzymes and were identified as aggregators, consequently, none was validated as a selective inhibitor.

Overall, this work did not identify a selective EXOG inhibitor. It established ATA as a potent but non-specific inhibitor of EXOG and characterized its likely mode of action. ATA–camptothecin combination reduced metabolic activity and mtDNA integrity in A549 cells, which supports the broader strategy of sensitizing cancer cells by reducing mtDNA integrity.

2. STRESZCZENIE

Integralność mitochondrialnego DNA (mtDNA) jest kształtowana przez kilka procesów, w tym replikację, transkrypcję, naprawę i degradację. Spośród szlaków naprawy działających w mitochondriach najlepiej scharakteryzowana jest naprawa przez wycinanie zasad (ang. base excision repair, BER). W niektórych nowotworach, w tym w gruczolakoraku płuca, podwyższona integralność mtDNA wiąże się ze zwiększoną przeżywalnością komórek nowotworowych. Proces BER w mitochondriach opiera się częściowo na enzymach wspólnych ze szlakiem jądrowym, jednak 5′-egzonukleaza EXOG jest specyficzna dla mitochondriów. EXOG uczestniczy w naprawie BER mtDNA, a jego udział wiązany jest z usuwaniem określonych produktów pośrednich naprawy zawierających koniec 5'. Doniesiono, że utrata aktywności białka EXOG prowadzi do gromadzenia się nienaprawionych pęknięć jednoniciowych i ostatecznie do apoptozy. Cechy te czynią EXOG kandydatem celowanej terapii przeciwnowotworowej: wybiórcze zahamowanie specyficznej dla mitochondriów aktywności naprawczej mogłoby uwrażliwiać komórki nowotworowe na chemioterapeutyki uszkadzające DNA. Do tej pory nie zidentyfikowano jednak żadnego selektywnego inhibitora EXOG.

Celem pracy było znalezienie niskocząsteczkowego związku modulującego aktywność EXOG, który mógłby stanowić punkt wyjścia do opracowania takiego inhibitora. Zastosowano dwie komplementarne strategie: wysokoprzepustowe badanie przesiewowe (ang. high-throughput screening, HTS) biblioteki LOPAC®1280 oraz przeszukiwanie wirtualne oparte na strukturze białka, z wykorzystaniem kolekcji związków z komercyjnej biblioteki BioAscent.

W wyniku HTS zidentyfikowano kwas aurintrikarboksylowy (ATA), jako inhibitor zarówno aktywności endonukleolitycznej, jak i 5′-egzonukleolitycznej EXOG. Scharakteryzowano jego wpływ na różne substraty DNA i RNA, wyznaczono stałą dysocjacji (K_d) oraz zaproponowano mechanizm hamowania. ATA wiązał się z wolnym białkiem EXOG z wysokim powinowactwem, wyraźnie zwiększał jego stabilność termiczną, a po preinkubacji obniżał powinowactwo EXOG do DNA, co jest zgodne z modelem, w którym ATA zaburza tworzenie kompleksu EXOG-DNA być może w obrębie miejsca aktywnego lub w jego pobliżu.

Kokrystalizacja i nasączanie kryształów związkiem (ang. soaking) nie dostarczyły interpretowalnej gęstości elektronowej odpowiadającej ligandowi, więc nie uzyskano bezpośredniego obrazu strukturalnego miejsca wiązania. Reszty aminokwasowe zaproponowane w dokowaniu *in silico* zbadano metodą mutagenezy ukierunkowanej, jednak, ponieważ żaden z wariantów nie wykazał obniżonej wrażliwości na ATA, ich dokładny udział w wiązaniu pozostaje do potwierdzenia.

Co istotne, ATA okazał się wysoce nieselektywny. Związek ten hamował również szereg innych enzymów uczestniczących w metabolizmie DNA i RNA i nie wykazywał selektywności względem EXOG. Mimo to, połączenie ATA z kamptotecyną (CPT) w linii komórkowej gruczolakoraka płuca (A549) obniżało integralność mtDNA, aktywność metaboliczną i przeżywalność komórek. Dane te są zgodne z upośledzeniem naprawy mtDNA, choć nie stanowią jego bezpośredniego dowodu.

Przeszukiwanie wirtualne wyłoniło trzy potencjalne inhibitory, jednak w dalszych testach hamowały one także niespokrewnione enzymy i zostały zidentyfikowane jako związki agregujące, przez co żaden z nich nie spełniał kryteriów selektywnego inhibitora.

Podsumowując, w pracy nie zidentyfikowano selektywnego inhibitora EXOG. Ustalono, że ATA jest silnym, lecz nieselektywnym inhibitorem EXOG, i scharakteryzowano jego prawdopodobny mechanizm działania.

Połączenie ATA z kamptotecyną obniżało aktywność metaboliczną i integralność mtDNA w komórkach A549, co wspiera szerszą strategię uwrażliwiania komórek nowotworowych poprzez obniżanie integralności mtDNA.

3. ABBREVIATIONS

Å	Ångstrom
A549	human lung adenocarcinoma cell line
ADP	adenosine diphosphate
APE1	apurinic/apyrimidinic endonuclease 1
AP sites	apurinic/apyrimidinic sites
ATA	aurintricarboxylic acid
ATP	adenosine triphosphate
B	biotin
BER	base excision repair
BHQ-1	Black Hole Quencher 1
BLI	biolayer interferometry
bp	base pair
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
CLogP	calculated logarithm of the partition coefficient
CODASS	Combined Docking And Similarity Searching
COX	cytochrome c oxidase
CPT	camptothecin
Cy3	Cyanine3
Cyt b	cytochrome b
Da	dalton
DDR	DNA-damage response
dL	2-deoxyribonolactone
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNA2	DNA replication helicase/nuclease 2
dNTP	deoxyribonucleoside triphosphate
dRP	5'-deoxyribophosphate
DSB	double-strand break
dsDNA	double-stranded DNA
DTT	dithiothreitol
EC_{50}	half-maximal effective concentration

EDTA	ethylenediaminetetraacetic acid
EndoG	endonuclease G
EtBr	ethidium bromide
EXOG	endo/exonuclease G-like protein
FDA	US Food and Drug Administration
FEN1	flap endonuclease 1
Flsn	fluorescein
G	guanine
GWOVina	Grey Wolf Optimization AutoDock Vina
HeLa	human cervical adenocarcinoma cell line
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HSP	Heavy strand promoter
H-strand	Heavy strand of mitochondrial DNA
HTS	high-throughput screening
IC_{50}	half maximal inhibitory concentration
IPTG	isopropyl β-D-1-thiogalactopyranoside
kbp	kilobase pair
Kd	dissociation constant
kDa	kilodalton
LA-qPCR	long-amplicon quantitative PCR
LB	lysogeny broth
LDH	lactate dehydrogenase
Lig III	DNA ligase III
LLG	log-likelihood gain
LOPAC	Library of Pharmacologically Active Compounds
L-strand	light strand of mitochondrial DNA
LSP	light strand promoter
MGME1	mitochondrial genome maintenance exonuclease 1
MLS	mitochondrial localization sequence
mtBER	mitochondrial base excision repair
mtDNA	mitochondrial DNA
mtSSB	mitochondrial single-stranded DNA-binding protein

MST	microscale thermophoresis
MTT	3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide
MTX	mitoxantrone
MW	molecular weight
MWCO	molecular weight cut-off
NAD^+	nicotinamide adenine dinucleotide
NADH	reduced nicotinamide adenine dinucleotide
nanoDSF	nano differential scanning fluorimetry
NCR	non-coding region
NMR	nuclear magnetic resonance
NNScore	Neural-Network Score
NP-40	Nonidet P 40
nt	nucleotide
NTA	nitrilotriacetic acid
OD	optical density
OriH	heavy strand origin of replication
OriL	light strand origin of replication
OXPHOS	oxidative phosphorylation
P	phosphorylated end
P_i	inorganic phosphate
PAGE	polyacrylamide gel electrophoresis
PARP	poly (ADP-ribose) polymerase
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEG	polyethylene glycol
PEI	poly(ethyleneimine)
Pol γ	DNA polymerase gamma
RMSD	root-mean-square deviation
RNA	ribonucleic acid
ROS	reactive oxygen species
rRNA	ribosomal RNA

RT	room temperature
SAX	streptavidin-coated
SD	standard deviation
SDS	sodium dodecyl sulfate
SEC	size exclusion chromatography
Sf9	*Spodoptera frugiperda* 9 insect cell line
SFM	serum free medium
SSB	single-stranded break
ssDNA	single-stranded DNA
ssRNA	single-stranded RNA
SUMO	small ubiquitin-like modifier
TAE	Tris-acetate-EDTA
TBE	Tris-borate-EDTA
TCEP	Tris(2-carboxyethyl)phosphine hydrochloride
TFAM	mitochondrial transcription factor A
TFZ	translation function Z-score
THF	tetrahydrofuran
T_i	inflection temperature
TKI	tyrosine kinase inhibitor
TOP1	topoisomerase I
tRNA	transfer RNA
V_{max}	maximum reaction velocity

4. INTRODUCTION

4.1. Mitochondria

Mitochondria are membrane-bound organelles which can be found in almost all eukaryotic cells. Electron microscopy studies performed in the 1950s revealed the internal organization of mitochondria that could not be resolved by light microscopy establishing the two-membrane framework and the concept of cristae on which the current understanding of mitochondrial architecture rests (Palade, 1952, 1953; Sjostrand, 1953). Mitochondria consist of an outer membrane, an inner membrane, an intermembrane space and a matrix (Fig. 1). Mitochondria are thought to originate from endosymbiotic ancestors - α-proteobacteria (Gray et al., 1999; Roger et al., 2017). This is supported by the fact that they possess their own genetic material in the form of mitochondrial DNA (mtDNA). This genome is a remnant of much larger genome of the endosymbiont, which was extensively reduced during its transition to mitochondria, when most ancestral genes were lost or relocated to the nucleus through endosymbiotic transfer (Gray et al., 1999; Roger et al., 2017). As a result, mitochondria possess a compact mtDNA molecule and most mitochondrial proteins are encoded by nuclear genes (Mercer et al., 2011; Barshad et al., 2018). The genes in mtDNA encode protein subunits necessary for respiratory chain function together with the ribosomal and transfer RNAs necessary for their translation within the organelle (Anderson et al., 1981). Several additional features of mitochondria are in line with bacterial ancestry. Like bacteria, they are bounded by two membranes and contain compact, usually circular genome, generally lacking introns (Anderson et al., 1981; Young, 2017). They also possess a translation system like that of bacteria – mitochondrial protein synthesis is initiated with N-formylmethionine and mitochondrial ribosomes are sensitive to antibiotics such as chloramphenicol, while at the same time they remain resistant to inhibitors of cytosolic translation (Young, 2017). At the sequence level, early comparative genomics identified the α-proteobacterium *Rickettsia prowazekii* as one of the closest sequenced bacterial relatives of mitochondria (Andersson et al., 1998). Although the alphaproteobacterial ancestry of mitochondria is strongly supported by phylogenetic analyses, the precise identity of their closest bacterial relatives remains unresolved (Martijn et al., 2018).

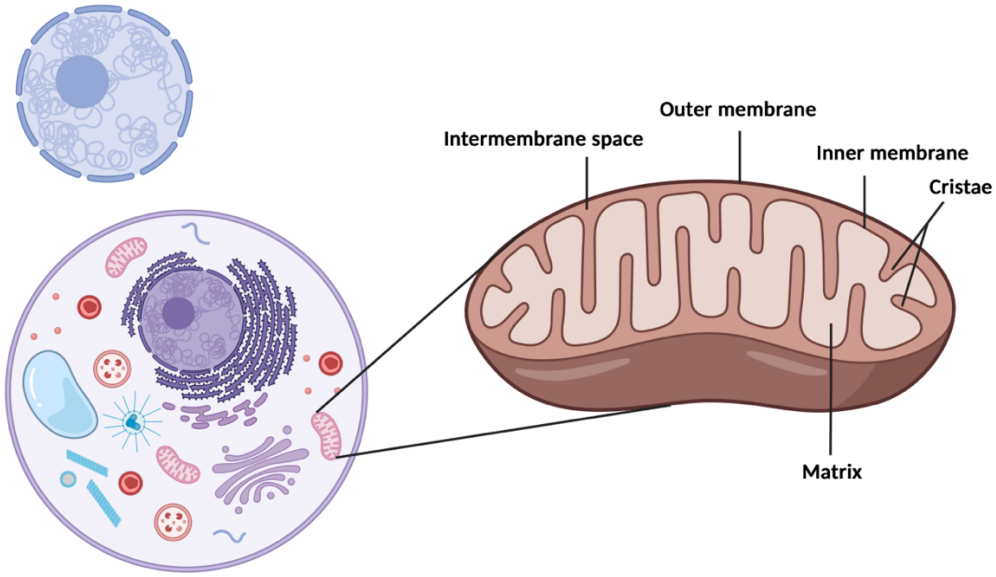


Figure 1. Schematic illustration of mitochondrial structure. *The inner compartment of mitochondria, the matrix, is surrounded by the outer and inner membranes. The inner membrane is folded into cristae where OXPHOS takes place (created with BioRender).*

Coordination of nuclear and mitochondrial encoded gene expression is crucial to ensure the assembly and function of proteins involved in the mitochondrial respiratory chain. Over 1100 nuclear-encoded proteins – around 99% of the mitochondrial proteome – are translated in the cytosol and imported into mitochondria with only 13 proteins encoded by mtDNA (Rath et al., 2021) and interestingly, all factors necessary for replication, transcription and maintenance of mtDNA are encoded by the nuclear genome (Falkenberg and Gustafsson, 2020).

Mitochondria are responsible for energy production through the process of oxidative phosphorylation (OXPHOS) (Osellame et al., 2012). Energy coming from this process is stored in the form of adenosine triphosphate (ATP), which acts as a cellular transporter for chemical energy, that can be released through hydrolysis of high energy bond, resulting in adenosine diphosphate (ADP) and inorganic phosphate (Pi) (Alberts, 2015). The demand for energy varies in different cell types and this is correlated with the number of mitochondria present in cells (D'Erchia et al., 2015). This number can vary from few hundreds of mitochondria that can be found in keratinocytes and circa 5000 of them in high energy-demanding cardiomyocytes (Alberts, 2015). Mitochondria are dynamic

structures which can change their morphology through the balance between fusion and fission (division) processes (Horbay and Bilyy, 2016).

4.2. Mitochondrial genome

Mammalian mtDNA is a 16.6 kb circular, double-stranded DNA molecule containing genes encoding 13 components of the OXPHOS system, 22 tRNAs and 2 rRNAs required for their translation (Fig. 2) (Anderson et al., 1981). Depending on the type of cell, they contain several hundred to hundreds of thousands of mtDNA molecules per cell (Bogenhagen and Clayton, 1974; King and Attardi, 1989; D'Erchia et al., 2015). MtDNA is associated with proteins and organized in nucleoids that are in the mitochondrial matrix, near the inner mitochondrial membrane (Alam, 2003; Bogenhagen et al., 2008; Farge and Falkenberg, 2019).

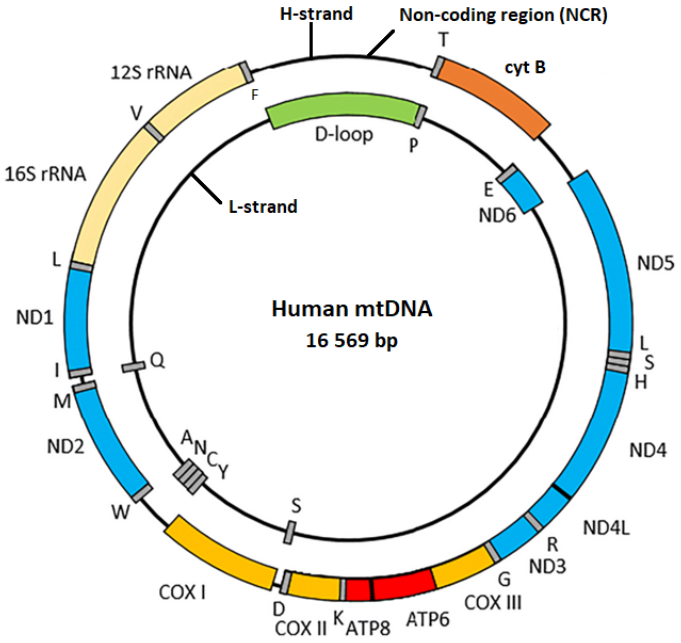


Figure 2. The mitochondrial genome map. *The two strands are called heavy (H-strand) and light strand (L-strand). The long non-coding region (NCR) contains the origin of replication for the H-strand (OriH) and the transcription promoters for both strands. The genes encoded by mtDNA are complex III cytochrome b (Cyt b)—orange; complex I NADH dehydrogenase (ND) genes—blue; complex IV cytochrome c oxidase (COX) genes—yellow; complex V ATP synthase (ATP) genes—red; ribosomal RNA (rRNA)—light yellow; transfer RNA genes—grey boxes.*

The mitochondrial genome is highly compact, lacks introns and contains very limited non-coding regions. The two strands creating this molecule can be distinguished into heavy (H) and light (L) strand based on their difference in guanine and cytosine content (Berk and Clayton, 1974). L-strand encodes one protein and 8 tRNAs whilst H-strand encodes about 12 proteins, 2 rRNAs and 14 tRNAs. mtDNA possesses only two non-coding regions: a smaller region which contains the origin of replication for the L-strand (OriL) (Tapper and Clayton, 1981; Shadel and Clayton, 1997; Uhler and Falkenberg, 2015) and a larger one known as the control region of the non-coding region (NCR) (Anderson et al., 1981). The NCR is one kilobase pair (kbp) in length and contains origin of replication of the H-strand (OriH) and transcription origin for both strands: the H-strand promoter (HSP) and the L-strand promoter (LSP) (Crews et al., 1979; Montoya et al., 1982; Chang and Clayton, 1984). mtDNA is maternally inherited (Giles et al., 1980) and mtDNA copy number is heavily downregulated during spermatogenesis and actively degraded after fertilization (Kaneda et al., 1995; Larsson et al., 1996; Larsson et al., 1997; Sutovsky et al., 1999). Unlike genomic DNA which is packaged by histones and nucleosomes (Kornberg, 1974), mtDNA lacks canonical histone packaging and instead is organized into nucleoids, in which it is bound and compacted by the mitochondrial transcription factor A (TFAM) (Kaufman et al., 2007; Kukat et al., 2015).

4.3. Mitochondrial DNA repair

As mitochondrially encoded proteins are required for cellular energy production, the maintenance and integrity of mtDNA are critical. Three major potential reasons have been linked with the increased sensitivity of mtDNA to damage compared with nuclear DNA. Firstly, mtDNA is located close to the inner membrane OXPHOS complexes, a major site of reactive oxygen species (ROS) production and is therefore directly exposed to ROS (Alexeyev, 2009). Secondly, under the strand-replacement model of mtDNA replication, synthesis proceeds asymmetrically, resulting in the heavy strand exposure as a single-stranded structure for extended periods of time, which was proposed to result in the spontaneous deamination of nucleotides (Tanaka and Ozawa, 1994; Reyes et al., 1998). The extent of such exposure depends on the replication mode, however alternative models – the strand-coupled model and the RITOLS model – propose that the displaced strand is not exposed as naked single-stranded DNA (Falkenberg and Gustafsson, 2020). Thirdly, and as a separate replication-based

mechanism, the asymmetric distribution of dNTPs in mitochondria, i.e. higher levels of dGTP, has been shown to reduce the replication fidelity of DNA polymerase γ *in vitro* and may contribute to the elevated rate of mutagenesis (Song et al., 2005). Together, these factors are thought to increase the mutational burden of mtDNA through distinct routes: oxidative damage, spontaneous deamination, and reduced replication fidelity, respectively. With respect to oxidative damage specifically, mtDNA has long been reported to accumulate more of it than nuclear DNA (Richter et al., 1988), although the magnitude of this difference is debated, with reported estimates ranging from about three-fold to over ten-fold, and some studies finding no clear difference (Ozawa, 1997; Yakes and Van Houten, 1997; Beckman and Ames, 1999). Damage-induced mutations are a result of DNA lesions escaping repair process and becoming fixed as permanent sequence changes during replication. In mitochondria, a major endogenous source of such lesions is oxidative damage by ROS (Cooke et al., 2003). It can generate a wide array of DNA lesions including base modifications, abasic sites (apurinic/apyrimidinic (AP) sites), DNA strand breaks and crosslinks (Cooke et al., 2003; De Bont, 2004). When mtDNA is left unrepaired, the damaged mitochondria can drive increased cellular apoptosis due to its impact on cellular homeostasis (Tann et al., 2011). Mitochondria rely on nuclear-encoded DNA-repair proteins imported from the cytosol. The best characterized and thought to be the most predominant pathway of mtDNA repair is Base Excision Repair (BER) (Fortini et al., 2003). The existence and significance of other pathways, such as mismatch repair and double-strand break repair remains debated (Kazak et al., 2012). BER pathway eliminates DNA lesions such as oxidized and deaminated bases or alkylation damage (De Souza et al., 2008; Krokan and Bjørås, 2013). BER is a multi-enzyme process that can be distinguished into short-patch (insertion of 1 nucleotide) or long-patch (insertion of 2-10 nucleotides) mechanisms. In long-patch BER, strand-displacement synthesis produces a 5′ single-stranded flap that must be removed before ligation. In human mitochondria this is carried out by several nucleases, including the endonucleases FEN1 (Liu et al., 2008) and DNA2 (Copeland and Longley, 2008), the exonuclease MGME1 (Uhler et al., 2016), and, as shown more recently – EXOG (Karlowicz et al., 2025). The mitochondrial BER follows the same general scheme as nuclear BER but has mitochondria-specific features. The steps of this pathway include (I) recognition and excision of the damaged base, (II) removal of the abasic site, (III) 5'-end processing, (IV) gap filling, and (V) DNA ligation (Fig. 3) (Prakash and Doublie, 2015). DNA glycosylases are the first

enzymes initiating the BER process and can be divided into monofunctional and bifunctional glycosylases. Monofunctional glycosylases recognize specific DNA lesions and remove the damaged base leaving the DNA backbone intact (Sampath et al., 2012; Sampath, 2014). Subsequently apurinic/apyrimidinic endonuclease (APE1) cleaves the phosphodiester backbone 5' to the AP site, generating a single-strand break with a 3'-hydroxyl terminus on the upstream side and non-ligatable 5'-deoxyribophosphate (dRP) on the downstream side (Mol et al., 2000). The 5′-dRP moiety is then excised by the dRP lyase activity of DNA Polymerase γ, which subsequently fills the single-nucleotide gap, producing a ligatable substrate (Longley et al., 1998). On the other hand, bifunctional glycosylases possess the intrinsic lyase activity and can cleave the abasic site (Kazak et al., 2012; Stein and Sia, 2017). While dRP is a substrate of the Pol γ 5'-dRP lyase activity, this activity is weak (Pinz and Bogenhagen, 2000), and the oxidized abasic residue 2-deoxyribonolactone (dL) is resistant to it (Liu et al., 2008), thus efficient removal of such 5′-blocking intermediates requires a 5′-exonuclease activity. EXOG has been shown to hydrolyze both dRP and dL moieties (Tann et al., 2011; Karlowicz et al., 2025). EXOG is a nuclease that has both endonuclease and 5'-exonuclease activity and is localized on the inner membrane of mitochondria (Cymerman et al., 2008). Cellular depletion of EXOG results in persistent single-strand breaks in mtDNA, mitochondrial dysfunction and programmed cell death (Tann et al., 2011).

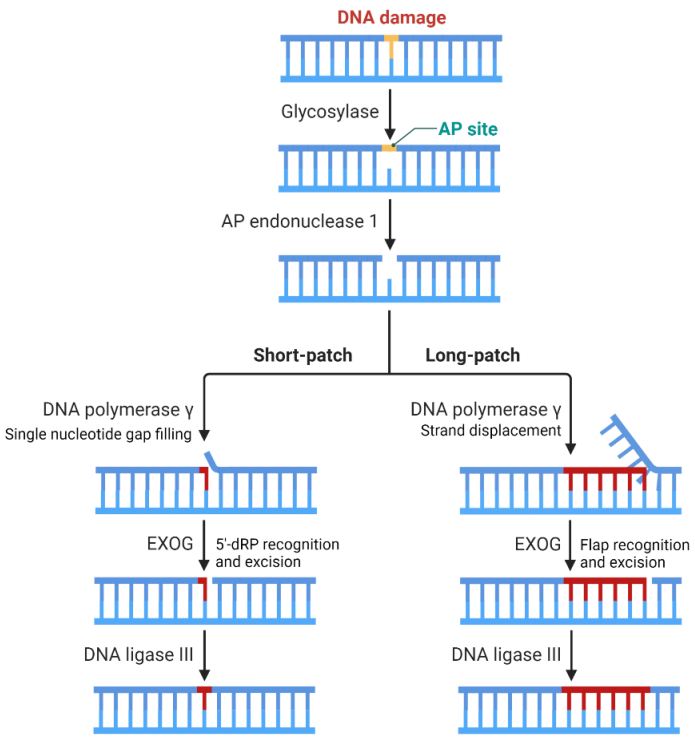


Figure 3. Base excision repair in mitochondria. *BER is initiated by the action of DNA glycosylases, creating abasic site which is further processed by APE1. EXOG is responsible for 5'-end processing creating a repair intermediate for DNA gap filling by Polymerase γ and ligation by Lig III (created with BioRender)*

What is more, EXOG is found in a complex with enzymes involved in mtDNA repair such as: APE1, Pol γ and Ligase III (Lig III) (Tann et al., 2011; Szczesny et al., 2014). Structural and biochemical studies of EXOG revealed a 'tape-measure' mechanism that allows it to precisely excise a dinucleotide from the 5'-end of DNA (Szymanski et al., 2017). This activity is well suited to 5'-end processing in mtBER. Removing two nucleotides eliminates the 5'-blocking residue and leaves a gap of at least two nucleotides, which serves as an optimal substrate for subsequent gap-filling synthesis by DNA Pol γ (He et al., 2013). Several studies have reported that Pol β, a canonical nuclear BER enzyme, also localizes to mammalian mitochondria and may contribute to mtDNA repair alongside Pol γ (Sykora et al., 2017; Prasad et al., 2017), although its presence and role in mitochondria remain debated (Kaufman and Van Houten, 2017). The last step of the BER process is the nick ligation by Lig III (Simsek et al., 2011). The coordinated activity of these proteins is required to maintain mtDNA integrity.

4.3.1. EXOG

Human EXOG is a mitochondrial nuclease that belongs to the ββα-Me family of DNA/RNA non-specific ('sugar-non-specific') nucleases – an evolutionarily conserved family found from bacteria to higher eukaryotes. Within this family, EXOG and its paralog EndoG (Fig. 4) are the mitochondrial members in higher eukaryotes (Cymerman et al., 2008).

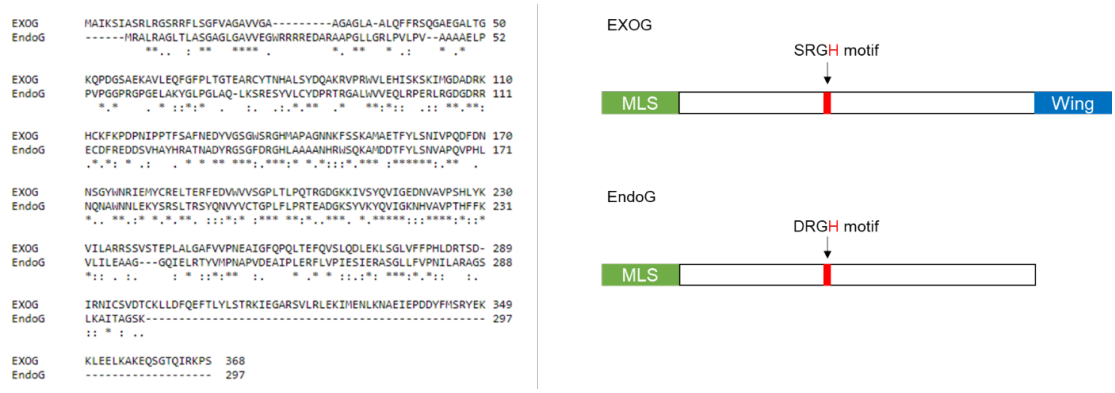


Figure 4. EXOG is highly related to mammalian mitochondrial endonuclease EndoG. *Sequence alignment obtained using Clustal Omega program (left panel). Domain organization of hEXOG and hEndoG; both proteins contain mitochondrial localization signal at N-terminus (marked in green), catalytic domain (white) with conserved DRGH motif in EndoG and SRGH motif in EXOG marked with arrow and additional wing domain (marked in blue) present in EXOG (right panel).*

The presence of EXOG protein most likely resulted from the duplication of an ancestral endo/exonuclease gene in higher eukaryotes. EXOG protein differs from other sugar non-specific nucleases, including EndoG, as in the core domain instead of highly conserved DRGH motif at positions 137 to 140 it has a SRGH motif (Fig. 4, right panel) (Cymerman et al., 2008). What is more, EXOG has an additional wing domain, which differentiates this enzyme in function from EndoG, as it possesses not only endonuclease but also 5'-exonuclease activity (Cymerman et al. 2008). Based on the primary sequence of the protein it has been shown that EXOG possesses a hydrophobic N-terminal transmembrane part (residues 16-38), which corresponds to a transmembrane helix that might anchor the protein to the inner mitochondrial membrane (Cymerman et al., 2008), yet the structure of the full-length enzyme remains unknown.

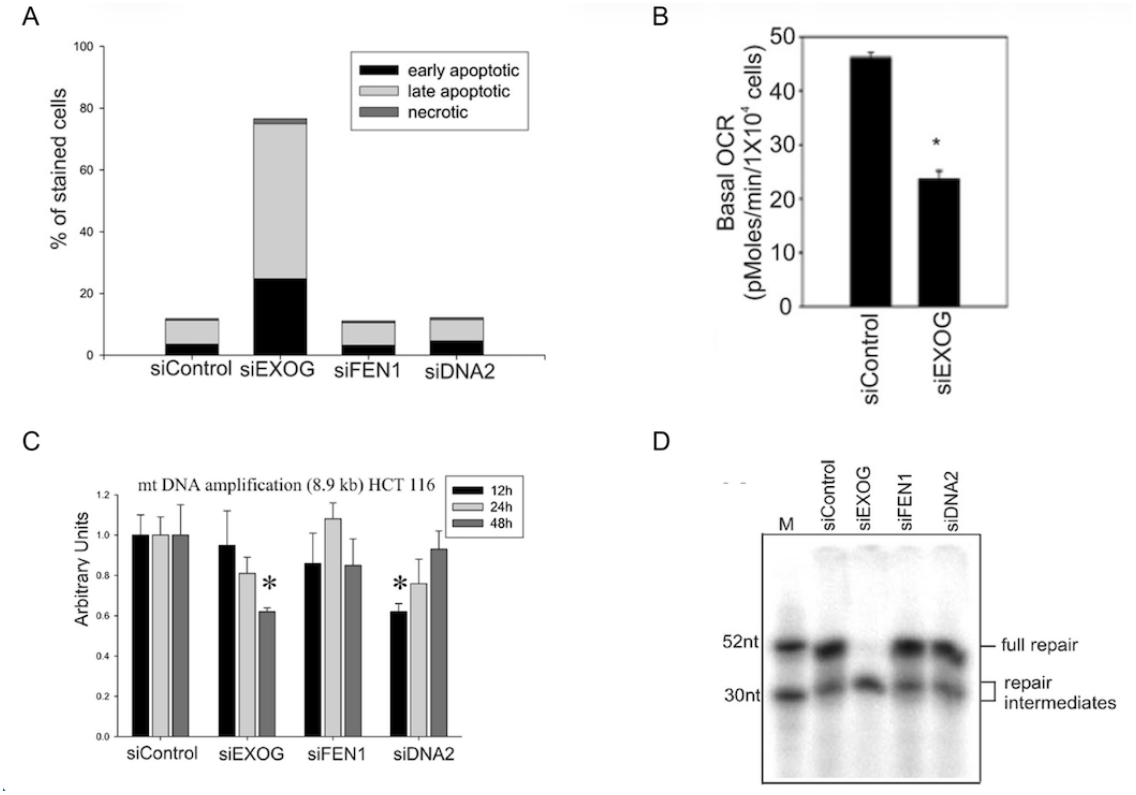


Figure 5. Effect of mitochondrial 5'-exo/endonucleases depletion on cell viability and mtDNA integrity and repair. *(A) EXOG but not FEN1 or DNA2 depletion triggers cell death in HeLa cells. (B) Loss of EXOG causes a decline in OXPHOS capacity. (C) EXOG silencing causes accumulation of mtDNA damage. (D) EXOG provides critical 5'-exonuclease activity for mtBER. (Adapted from Tann et al., 2011).*

All biochemical and structural studies were done on a protein variant lacking part of the N-terminus (in this work ΔN58). Conducted *in vitro* and *in vivo* studies highlight the critical role of this nuclease in integrity of mtDNA, functioning of mitochondria and cell viability (Fig. 5). EXOG's depletion was linked with persistent single-stranded DNA breaks (SSBs) in mtDNA, enhanced ROS levels, which caused severe mitochondrial dysfunction leading to cell death. Ectopic expression of EXOG had a positive impact on cell's resistance to oxidative stress (Tann et al., 2011).

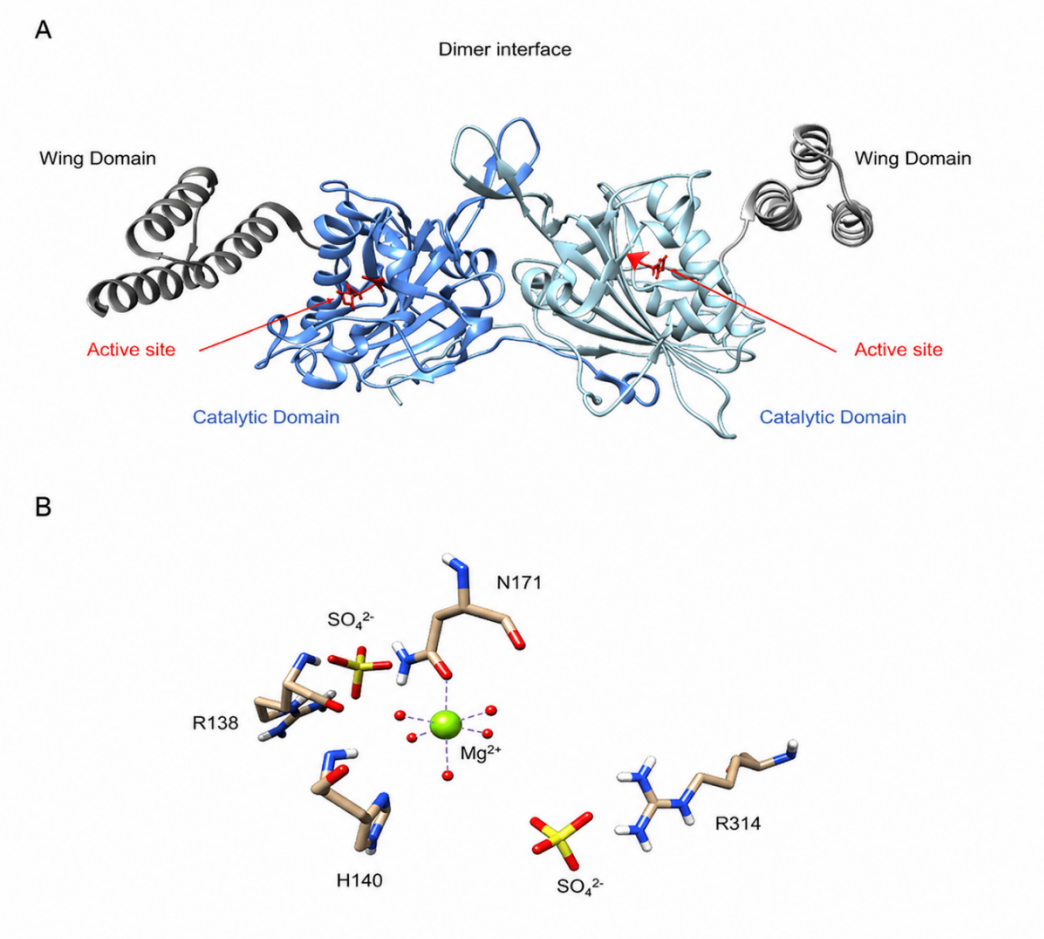


Figure 6. Structure of apo hEXOG (PDB:5T40). *(A) Enzyme consists of two monomers creating a homodimer. Each monomer has a Core (light and dark blue) and Wing domain (light and dark grey). Histidine at position 140 corresponding to protein's active site marked in red. (B) The active site of EXOG showing magnesium ion (light green sphere) with N171 residue and water molecules (red spheres) and R138 and R314 interactions with* SO_4^{2-} *ions.*

Structural and biochemical studies performed on the ΔN58 variant shed light on protein's specific substrate recognition – 'tape measure' mechanism (Fig. 6 A) (Szymanski et al., 2017). EXOG can recognize the 5'-end of dsDNA and precisely excise a dinucleotide. Such specific substrate recognition and cleavage is provided by the wing domain and differentiates this protein from other ββα-Me enzymes. In the apo form ΔN58 EXOG is a symmetrical homodimer, of mass 72.6 kDa. Upon DNA binding, the otherwise symmetrical EXOG homodimer becomes asymmetric: DNA bound at one active site allosterically regulates the activity of the other (Szymanski et al., 2017). In EXOG, only one of the two active sites of the homodimer is catalytically active at a time (half-site reactivity), which has been linked to an ~50-fold difference in DNA-substrate affinity between the monomers (Szymanski et al., 2017).

The EXOG active site contains two residues critical for catalytic activity. One is histidine 140, which is required for substrate cleavage; the other is asparagine 171, which together with five water molecules forms an octahedral coordination of Mg^{2+} ion necessary for substrate catalysis (Fig. 6 B). Active site of EXOG forms a V-shaped cleft which is lined with positively charged residues, relevant for the substrate binding. EXOG was shown to process broad range of substrates including ssDNA, ssRNA, RNA/DNA hybrids as well as RNA, RNA/DNA and DNA flaps (Wu et al., 2019; Karlowicz et al., 2022; Karlowicz et al., 2025). Additionally, research showing the broad-spectrum activity of EXOG on variety of abasic sites, confirms protein indispensable role in 5'-end processing, creating a substrate for Pol γ gap filling synthesis followed by ligation (Szymanski et al., 2022). More recently, EXOG was shown to cleave both free and mtSSB-protected 5′ DNA flaps. The Wing domain provides an additional binding site for the displaced single-strand, and flap removal by EXOG generates ligatable ends that, together with Pol γ, mtSSB and Lig III, are converted into fully processed dsDNA, placing EXOG among the mitochondrial flap nucleases acting in long-patch BER (Karlowicz et al., 2025).

4.4. Mitochondrial DNA and cancer

In normal cells, DNA damage is detected by the DNA-damage response (DDR), which activates cell-cycle checkpoints and induces cell-cycle arrest; this prevents the transmission of damaged DNA to daughter cells. Depending on the context, however, the same response can instead lead to repair, senescence or apoptosis (Jackson and Bartek, 2009). On the other hand, in cancer cells, nuclear DNA damage, mainly increased rate of mutations in proto-oncogenes, tumor suppressor genes or DNA repair genes causes cancer cells to evade organism's antitumor mechanisms, additionally allowing them to achieve high proliferation rate (Patel, et al., 2020). Rapidly dividing cells have high energy demands, and Otto Warburg observed that cancer cells adopt a metabolic phenotype distinct from that of normal cells, characterized by an altered mode of energy production (Warburg, 1956). Malignant cells showed increased glycolysis and reduced oxygen consumption which was later called 'Warburg effect' (Warburg, 1956). Based on the observation that cancer cells have upregulated glycolysis compared with normal cells, the idea emerged that OXPHOS is downregulated in all cancers; this view, however, was later shown to be a misinterpretation, as many cancers retain functional mitochondrial respiration (Koppenol et al., 2011). It has been shown that in some types

of cancer including leukemias, lymphomas, pancreatic ductal adenocarcinoma, melanoma and endometrial carcinoma mitochondrial metabolism is not impaired at all (Moreno-Sanchez et al., 2007; Weinberg and Chandel, 2015). Szczesny et al. (2016) reported, that mtDNA integrity assessed by a long-amplicon PCR assay (LA-qPCR), in which fewer strand breaks in the template yield greater amplification of an 8.9 kb mtDNA fragment normalized to mtDNA copy number, was significantly higher in tumors of lung adenocarcinoma patients (n=20), when compared to the healthy lung adjacent tissue. On the other hand, such phenomenon was not found for the nuclear DNA. What is more, recent studies show that mtDNA integrity is vital for the survival of cancer cells (Wallace, 2012). Tan and his colleagues (Tan et al., 2015) shown that intact mtDNA is critical for tumorigenic potential. Cells lacking mtDNA could not metastasize and only after acquisition of the mtDNA from the host cells, they were able to form tumors. Impairment of mtDNA integrity activates cell death pathways, disturbs regulation of mitochondrial protein expression, diminishes cellular bioenergetics and suppresses cell viability and survival (Van Houten et al., 2016). If faithful repair of the mtDNA is crucial for continuously dividing cancer cells, then inhibition of its repair can be expected to have profound effect on cancer progression.

The dependence of cancer cells on efficient DNA repair has itself become a therapeutic opportunity. The clearest example is the use of poly(ADP-ribose) polymerase (PARP) inhibitors, which exploit synthetic lethality: in tumors deficient in homologous recombination, such as BRCA1/2-mutant cancers, inhibition of PARP-dependent repair leads to an unsustainable accumulation of DNA damage and selective death of cancer cells, while normal cells survive (Petropoulos et al., 2024). A related rationale can be applied to mitochondrial DNA repair, with an additional layer of selectivity. The mtBER repair machinery is not a duplicate of the nuclear one. Although most repair proteins are shared between the two compartments, only two enzymes EXOG and Pol γ are mitochondria specific (Szczesny et al., 2016). Because mtDNA integrity is elevated in certain cancers including lung adenocarcinoma and is maintained by a defined set of enzymes (Szczesny et al., 2016), pharmacological inhibition of mitochondrial DNA repair has been explored as an anti-cancer strategy targeting mitochondrial DNA-processing enzymes (Somuncu et al., 2022). EXOG is a particularly relevant target in this respect, as a mitochondria-specific 5'-exonuclease required to complete mtDNA BER. Its loss of activity results in the accumulation of unrepaired single-strand breaks and ultimately apoptosis (Tann et al., 2011). Selectively disabling this mitochondria-

specific repair activity on which cancer cells particularly rely, provides the rationale for the present work.

4.5. DNA damaging agents in cancer chemotherapy

According to the Centers for Disease Control and Prevention (CDC, 2024), common types of cancer treatment include surgery, hormonal therapy, chemotherapy, radiation therapy, immunotherapy and stem cell (bone marrow) transplant (CDC, 2024) (Fig. 7). Within pharmacological cancer treatment, cytotoxic chemotherapy was for a long time the dominant drug-based approach, later complemented by targeted and immunological therapies (DeVita and Chu, 2008). Chemotherapy is based on differences in the rates of division between cancer and normal cells. The first compounds that were widely used as cancer drugs were discovered during and after the World Wars, after studying soldiers who were victims of sulphur mustard gas use (Chabner and Roberts, 2005). It was found that they had depleted bone marrow and reduced lymph nodes (Krumbhaar and Krumbhaar, 1919). This observation prompted the testing of more stable nitrogen mustards and later these agents together with folate antagonists were recognized to kill cancer cells through their DNA-damaging properties.

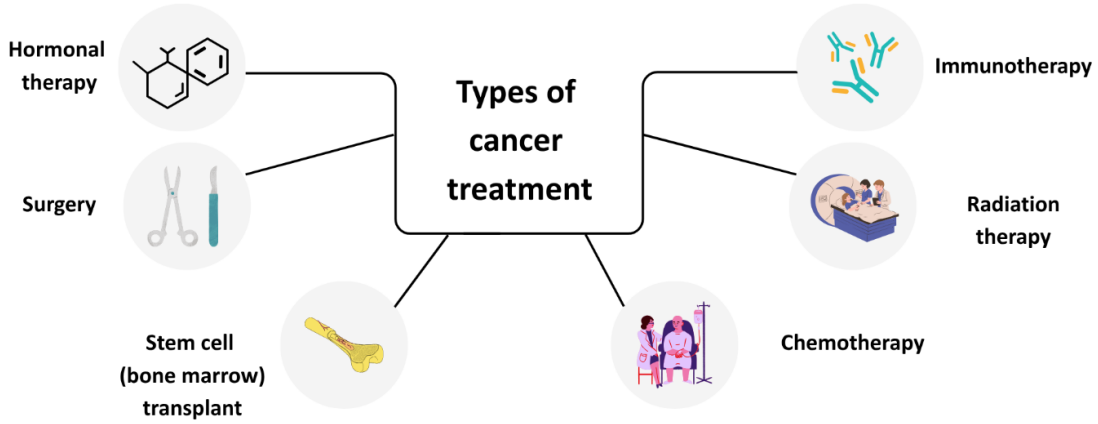


Figure 7. The most common approaches of treating cancer according to Centers for Disease Control and Prevention (CDC). *First line of fighting cancer is surgery, chemotherapy, and radiation, but new types are being developed to increase the effectiveness of treatment (created in Canva).*

This finding inspired the development of several DNA-targeting compounds such as cisplatin, doxorubicin, 5-fluorouracil, etoposide, and gemcitabine (Cheung-Ong et al., 2013). Another group of DNA-targeting drugs inhibits the enzymes that control DNA

topology, such as the topoisomerase I (TOP1) inhibitor camptothecin (CPT), which traps TOP1 on DNA and generates DNA breaks that are cytotoxic to dividing cells (Wall and Wani, 1995; Hertzberg et al., 1989). By creating lesions in DNA, these compounds aim to induce apoptosis through excessive DNA damage (Huschtscha et al., 1996). However, cancer cells may develop resistance to DNA-damaging chemotherapeutic agents through diverse molecular mechanisms. For example, repair enzymes can remove DNA lesions caused by cancer treatment, which will lead to cell's survival and has been described as a mechanism of drug resistance (Chen et al., 2005). As a result such cells have abnormally high DNA repair capacity and thus play important roles in modulating the response of human cancer to chemotherapy (Zheng et al., 2012). This fact makes targeting repair enzymes as onco-targets an interesting topic of research. One way to exploit this, is to combine a conventional DNA-damaging agent with an inhibitor of a DNA repair enzyme: blocking repair sensitizes cancer cells to the induced lesions and can allow lower, less toxic doses of the chemotherapeutic. This work follows that rationale.

4.6. Small-molecule compounds

The main disadvantage of conventional cytotoxic chemotherapy is the fact that it cannot distinguish between cancer cells and healthy cells, which results in high toxicity and side effects. For this reason, the course of treatment shifts its focus from broad-spectrum cytotoxic drugs to targeted therapy. In 2001, the US Food and Drug Administration (FDA) approved the first small-molecule tyrosine kinase inhibitor (TKI) imatinib, which opened the door for the targeted drugs in the cancer treatment. These can be classified into two categories: small compounds and macromolecules (e.g. monoclonal antibodies, antibody-drug conjugates). Small molecule compounds possess some advantages as they are cheaper, easier in storage and transport, have better pharmacokinetic properties. Another distinctive difference between small molecule compounds and, for example antibodies, is their route of administration. Antibodies must be injected intravenously or subcutaneously, due to their high molecular weight, while small molecule compounds can be administered orally and are able to bind their intra- and extracellular targets (Moroz et al., 2016; Wu et al., 2016; Roskoski, 2020).

4.7. Strategies for identifying small-molecule EXOG inhibitors

4.7.1. Experimental high-throughput screening (HTS)

High-throughput screening (HTS) is a widely used strategy for identifying compounds that modulate a biochemical activity of interest (Macarron et al., 2011). In this approach, large collections of compounds, often from thousands to hundreds of thousands, are tested in parallel in a miniaturized, automated assay, which allows many potential modulators to be examined in a single campaign (Hughes et al., 2011; Macarron et al., 2011). The immediate output of such a screen is a set of hit compounds: molecules that show the desired activity above a defined threshold. A hit is, however, only a starting point. Before it can be regarded as a validated lead, its activity must be confirmed in independent assays and further optimized, and even a validated lead is separated from a clinical drug by extensive additional development (Hughes et al., 2011). Experimental HTS therefore serves to identify active compounds within an existing library rather than to deliver drugs directly. A characteristic feature of experimental HTS is that compounds from a physical, commercially available library are tested directly against the target, which requires little prior knowledge of its structure. This makes the approach applicable to targets for which no established inhibitors are yet available. Because a large number of samples must be measured, the readout of such a screen must be fast, sensitive and easy to automate. For this reason, most HTS assays rely on fluorescently labelled reporters, which offer rapid detection, high sensitivity, wide reagent availability and relatively straightforward optimization (Inglese et al., 2007). For the assay to be informative, it must be optimized to give reproducible results and must include positive and negative controls (Hughes et al., 2011). These controls define the dynamic range of the signal and provide a proof of concept. They demonstrate that the measured signal reflects genuine modulation of the target rather than an assay artefact. The separation between the positive and negative controls is commonly expressed by the Z′-factor, which is used to judge whether an assay is sufficiently robust for screening (Zhang et al., 1999).

4.7.2. Structure-based virtual screening

The second strategy used to identify EXOG inhibitors was structure-based virtual screening. In contrast to experimental HTS, this approach is computational. It uses the three-dimensional structure of the target to search compound libraries *in silico* and therefore requires that an experimental structure or a reliable model of protein is available. In such screening compounds are docked into defined binding site on the target to predict their orientation and conformation. Next, scoring function is used to estimate the predicted affinity of each compound for the site (Kitchen et al., 2004; Shoichet, 2004). Because it is performed prior to any biochemical testing, virtual screening allows very large libraries to be examined which reduces number of compounds that need to be tested experimentally (Shoichet, 2004). In principle it can prioritize a small set of candidates predicted to bind the target with high affinity and selectivity, which can then be validated and optimized experimentally. The reliability of predictions is however limited by the scoring functions used. Docking typically returns several poses with similar scores, and incorrect poses give unreliable affinity predictions (Kitchen et al., 2004). For this reason, virtual-screening pipelines commonly combine several docking and scoring methods and apply additional filtering steps to increase confidence in the retained compounds before they are purchased and tested.

4.7.3. Compound prioritization and drug-like properties

Once active compounds have been identified by either experimental or virtual screening, they must be prioritized before further work. A key task at this stage is to distinguish genuine, target-specific hits from artefacts, which is done by re-testing the primary hits in an orthogonal assay. False positives can arise in several ways, including interference with the assay signal (for example, fluorescence interference), colloidal aggregation, promiscuous (pan-assay) inhibition, chemical reactivity of the compound, and assay-specific artefacts (McGovern et al., 2002; Baell and Holloway, 2010).

Aggregating compounds are a particularly common source of false positives: at micromolar concentrations they can form colloidal particles that non-specifically inhibit many unrelated enzymes by adsorbing the protein to their surface (McGovern et al., 2002). This is why orthogonal confirmation and selectivity testing are an essential part of hit prioritization. Beyond specificity, prioritization also considers whether a compound has physicochemical properties compatible with further

development. For orally administered drugs, a widely used guideline is the 'rule of five', formulated by Lipinski and colleagues (Lipinski et al., 2001). It states that poor absorption or permeation is more likely when a compound has a molecular weight above 500 Da, a calculated lipophilicity (CLogP) above 5, more than five hydrogen-bond donors (mainly OH and NH groups), or more than ten hydrogen-bond acceptors (the sum of nitrogen and oxygen atoms). The high molecular-weight limit relates to the poorer intestinal and blood-brain-barrier permeability of larger molecules (Cohen and Bangham, 1972; Pardridge, 1995). Poor absorption is considered likely when at least two of these limits are exceeded. Importantly, the rule of five is a guideline for oral absorption and permeability rather than a strict requirement or a predictor of all pharmacokinetic properties, and Lipinski and colleagues noted that some drugs fall outside these limits (Lipinski et al., 2001). Tool compounds and early inhibitors intended as starting points for target validation need not satisfy the rule, since they are not necessarily developed as oral drugs.

4.8. Post-screening validation and characterization of inhibitor candidates

Hits identified by experimental or virtual screening require biochemical confirmation, selectivity testing and, where possible, characterization of their mechanism of inhibition. These post-screening steps form a major part of the present work.

4.8.1. Biochemical validation of candidate inhibitors

A compound identified in a primary screen is only a candidate until its activity has been confirmed. Validation therefore begins by re-testing the primary hits in an independent, target-specific assay, ideally one with a different readout from that used in the screen, and by establishing a concentration-dependent inhibition of the enzyme from which an IC_{50} can be determined (Hughes et al., 2011). At this stage it is also necessary to exclude the assay artefacts discussed above (Section 4.7.3) and to confirm that the compound acts on the enzyme itself rather than on the substrate or the assay signal.

4.8.2. Selectivity and off-target assessment

Confirming that a compound inhibits its intended target does not establish that it does so selectively, and this distinction is central to the present project. EXOG belongs to a group of functionally related nucleic-acid-processing enzymes, so a compound that inhibits it

may also inhibit its relatives. Several of these enzymes, unlike EXOG, are not confined to mitochondria but also act in the nucleus; a non-selective compound would therefore be expected to reduce DNA-repair capacity in both compartments rather than only in mitochondria, which bears on both its specificity as a tool and its potential toxicity. Selectivity is therefore assessed by testing candidate inhibitors against a panel of related enzymes and known off targets under comparable conditions. For ATA in particular, inhibition of other DNA-processing enzymes such as APE1 and FEN1 had already been reported (Simeonov et al., 2009; Dorjsuren et al., 2011), which made an explicit selectivity assessment especially important.

4.8.3. Mechanistic and structural characterization of inhibitor binding

For the most promising candidates, the final assessment is how inhibition occurs, whether the compound competes with the substrate or acts elsewhere, and where and how it binds the enzyme. Direct binding to the purified protein can be measured by biophysical methods. As an example, a shift in the protein's thermal stability upon addition of the compound is consistent with binding and can be used to estimate its affinity (Niesen et al., 2007), although such a shift does not by itself prove a direct interaction and is therefore usually complemented by a method that measures binding directly, such as biolayer interferometry. The binding interface can be defined by co-crystallography of the enzyme-inhibitor complex. In cases where an interpretable complex cannot be obtained, computational docking combined with site-directed mutagenesis can be used to propose and test candidate contacts, keeping in mind that the absence of interpretable ligand electron density is not in itself evidence that the compound does not bind.

5. AIM OF THE PROJECT

Although the role of EXOG in mtDNA repair is well established, no specific inhibitor of this enzyme has been identified to date. The aim of this project was to find a small-molecule compound that modulates the activity of EXOG – a solely mitochondrial nuclease vital for mitochondrial Base Excision Repair. Such a compound could be used as a potential starting point for an anti-cancer agent with the aim of increasing sensitivity of cancer cells to conventional chemotherapeutics.

To achieve this goal, two complementary and independent strategies were pursued to identify EXOG inhibitors: one based on high-throughput screening of a physical compound library, and one based on structure-based virtual screening.

Part I - HTS-derived compounds:

- high-throughput screening of a commercially available compound library, to identify EXOG inhibitors (performed in collaboration with Bartosz Szczesny, PhD, UTMB);
- characterization of the interaction between EXOG and the identified inhibitor, including its effect on the enzyme's activity towards different DNA and RNA substrates;
- structural studies aimed at identifying the interaction interface;
- assessment of off-target inhibition against other DNA- and RNA-processing enzymes;
- assessment of the effect of the inhibitor on cell viability and mtDNA integrity (cell-based experiments performed by Bartosz Szczesny, PhD, UTMB); these data are presented and discussed in this thesis.

Part II - Virtual-screening-derived compounds

- structure-based virtual screening of compound libraries (performed in collaboration with Douglas Houston, PhD, University of Edinburgh);
- experimental testing and validation of the selected compounds against EXOG.

6. MATERIALS

6.1. Strains

***Escherichia coli* BL21 (DE3) pLysS** F–, ompT, hsdSB (rB–, mB–), dcm, gal, λ(DE3), pLysS, Cam^R

***Escherichia coli* Rosetta 2 (DE3) pLysS** F- ompT hsdSB (rB- mB-) gal dcm (DE3) pLysSRARE2, Cam^R

***Escherichia coli* Rosetta 2 (DE3)** F-ompT hsdSB (rB- mB-) gal dcm (DE3) pRARE2, Cam^R

***Escherichia coli* BL21 Star (DE3)** F-ompT hsdSB (rB-, mB-) galdcmrne131 (DE3)

***Spodoptera frugiperda* (Sf9)** cells derived from the immature ovaries of fall armyworm moth pupae

6.2. Plasmids

pET 22b (+) (69744, Novagen) vector used for overproduction of EXOG and its variants, Pol γB, Amp^R, T7 lac promoter, IPTG-induced

pET 15b vector used for overproduction of APE1 and RNase H1, Amp^R, T7 lac promoter,
IPTG induced (Barnes et al., 2009; Nowotny et al., 2007)

pBacPAK9 transfer vector (Clontech) shuttle vector used for transfer into the baculovirus genome, used for overproduction of Pol γA

pE-SUMO vector used for overproduction of FEN1, Kan^R, T7 lac promoter, IPTG induced
(a gift from Prof. Jane A. Grasby from University of Sheffield, UK)

pET 28a-SUMO vector used for overproduction of Ligase III and MGME1, Kan^R, T7 lac promoter, IPTG induced (McNally and O'Brien, 2017, Szczesny et al. 2013)

6.3. Broths and media

LA 0.5% yeast extract, 1% peptone, 1% NaCl, 1.5% agar

LB 0.5% yeast extract, 1% peptone, 1% NaCl

2xYT 1% yeast extract, 0.6% peptone, 0.5% NaCl

Sf-900™ II SFM (Gibco) - serum-free, protein-free insect cell culture medium optimized for the growth and maintenance of *Spodoptera frugiperda* (Sf9 and Sf21) cells and for large-scale production of recombinant proteins expressed using the baculovirus expression vector system

6.4. Antibiotics

Ampicillin (Sigma, cat. no. A9393) (100 μg ml^{-1})

Chloramphenicol (BioShop Canada, cat. no. CLR201.25) (34 μg ml^{-1})

Kanamycin (BioShop Canada, cat. no. KAN201.25) (25 μg ml^{-1})

Table 1. Oligonucleotides

No.	Sequence (5'-3')	Length (nt)
1	CGCACGTAATACAGTCTAGT-Flsn	20
2	cgcacguaauacagucuagu-Flsn	20
3	B-GCTTTTGCTCCCGGTCACGGTATGTGCGTGCATTATGTCAGATCA-B	45
4	B-CGAAAACGAGGGCCAGTGCCATAC	24
5	B-ACTAGACTGTATTACGTGCG	20
6	$(T)_{15}$-CGCACGTAATACAGTCTAGT-Flsn	35
7	B-ACTAGACTGTATTACGTGCGGGTATGGCACTGGCCCTCGTTTTCG-B	45
8	B-CGAAAACGAGGGCCAGTGCCATAC(**THF**)CGCACGTAATACAGTCTAGT-Flsn	45
9	cgcacguaauACAGTCTAGT-Flsn	20
10	B-GCTTTTGCTCCCGGTCACGGTATGGCGTGCATTATGTCAGATCA-B	44
11	Flsn-CGAAAACGAGGGCCAGTGCCATAC	24
12	ACGAGCCTGCCTGACGTGCGAGTATGGCACTGGCCCTCGTTTTCG	45
13	P-CGCACGTAATACAGTCTAGT-Flsn	20
14	CGCACGTCAGGCAGGCTCGT-B	20
15	CGCACGTCAGGCAGGCTCGT-Cy3	20
16	AGTGAAGGCATGCACCCTTCT-Flsn	21
17	BHQ-1-AGAAGGGTGCATGCCTTC	18

Flsn – fluorescein; B – biotin; P – phosphorylated end; ATGC – DNA; augc – RNA; THF – Tetrahydrofuran; Cy3 – Cyanine3; BHQ-1 – black hole quencher 1

Table 2. Substrates

Name	Substrate	Oligonucleotides (Table 1)
S1	ssDNA	1
S2	ssRNA	2
S3	ssDNA for BioAscent screening	15
S4	dsDNA	1, 5
S5	1-nt gapped dsDNA	1, 3, 4
S6	Nicked 10RNA-10DNA	4, 9, 10
S7	Primer extension substrate	11, 12
S8	Nicked phosphorylated dsDNA	4, 10, 13
S9	Nicked 15-nt $(dT)_{15}$ flap dsDNA	4, 6, 10
S10	dsDNA with AP-site	7, 8
S11	Biotinylated ssDNA for BLI studies	14
S12	dsDNA with fluorescein and BHQ-1 for LOPAC HTS	16, 17

Table 3. Compounds used for inhibition studies

Name of the compound	Catalogue number	Vendor	CAS registry number	Purity
Aurintricarboxylic acid	A1895	Sigma-Aldrich	4431-00-9	≥85%
Aurintricarboxylic acid ammonium salt	A0885	Sigma-Aldrich	569-58-4	-
5,5'-methylenebis(2-hydroxybenzoic acid)	PH011549	Merck	122-25-8	-
p-Rosolic acid	861324	Sigma-Aldrich	603-45-2	85%
Pararosaniline chloride	227880250	Thermo Fisher Scientific	569-61-9	≥85%

7. METHODS

7.1. Preparation of *Escherichia coli* competent cells

Preparation of *Escherichia coli* competent cells was performed using Mix & Go *E. coli* Transformation Kit & Buffer Set (Zymo Research, cat. no. T3001). Briefly, 100 ml of LB medium was inoculated with 1 ml of overnight *E. coli* culture and grown in a shaking water bath at 37 °C until $OD_{600} \approx 0.5$. The culture was then chilled on ice for 10 min and harvested by centrifugation (3000 RCF, 10 min, 4 °C). Next steps were done according to the manual. Finally, cells were aliquoted to 100 µl prior freezing at -80 °C.

7.2. Plasmid DNA isolation

Isolation of plasmid DNA from *E. coli* night cultures was performed using Plasmid Mini isolation kit (A&A Biotechnology, cat. no. 020-50) according to the instructions.

7.3. Transformation of *E. coli* competent cells

Competent cells were thawed on ice and plasmid DNA (50-100 ng per 100 µl of cells) was added. Cells with plasmid DNA were incubated on ice for 5 min and later plated on freshly prepared LA medium plates supplemented with antibiotics adequate for strain or plasmid resistance. In case of kanamycin resistance, after 5 min of incubation on ice, cells were supplemented with 1 ml of LB medium and incubated with shaking for 1 h at 37 °C to improve the efficiency of transformation.

7.4. Protein purification

7.4.1. Human mitochondrial 5'-exonuclease (EXOG) wild-type and variants

Plasmids encoding recombinant human EXOG, lacking the mitochondrial targeting sequence (ΔN58), carrying a C-terminal His_6 tag and its single-point mutants F114A, R138A, F158A, N171A, R183A and R314A were purchased from Epoch Life Sciences. Rosetta 2 (DE3) pLysS *E. coli* cells were transformed with pET 22b (+) plasmid encoding EXOG (ΔN58) wild-type or its variant. Cell culture was grown in fresh LB at 37 °C (orbital shaker) until $OD_{600} \approx 0.6$, then cells were induced with 1 mM IPTG. After induction, temperature was lowered to 20 °C and culture was grown for 18 h. Cells were

harvested by centrifugation (4950 RCF, 10 min, 4 °C). Next, the pellet was resuspended in Buffer UA (20 mM HEPES, pH 7.5, 100 mM NaCl, 0.1 mM EDTA, pH 8, 0.1% CHAPS, 10 mM DTT, 10% sucrose, 10 mM imidazole) with addition of protease inhibitors (E64 (Sigma-Aldrich, cat. no. E3132), pepstatin A (Sigma-Aldrich, cat. no. P5318), and leupeptin (Sigma-Aldrich, cat. no. 62070)) prior to lysis with sonication (50% amplitude, 30 s bursts followed by 1.5 min pauses on ice, total sonication time 5 min). The lysate was centrifuged (31 000 RCF, 30 min, 4 °C) and the pellet fraction with EXOG was collected. The pellet was suspended in Buffer UB (10 mM Tris- HCl pH 8.2, 6 M urea, 10 mM β-mercaptoethanol, 200 mM NaCl, 10 mM imidazole) to solubilize EXOG. Then sample was centrifuged (31 000 RCF, 30 min, 4 °C) and supernatant was loaded onto Protino® NiNTA agarose resin (Macherey-Nagel, cat. no. 745400.100) equilibrated with Buffer UB. Resin was extensively washed with Buffer UB and Buffer UC (10 mM Tris- HCl pH 8.2, 3.5 M urea, 10 mM β-mercaptoethanol, 200 mM NaCl, 10 mM imidazole) and eluted with Buffer UD (10 mM Tris- HCl pH 8.2, 3.5 M urea, 10 mM β-mercaptoethanol, 200 mM NaCl, 200 mM imidazole). Samples containing EXOG were combined and diluted 10 times with Buffer UD for final volume ~550 ml. Diluted sample was dialyzed against refolding buffer FF (50 mM HEPES, pH 7.5, 1.5 M sorbitol, 1 mM TCEP (Tris(2-carboxyethyl)phosphine hydrochloride), 24 mM NaCl, 1 mM KCl) using a dialysis membrane with a 30 kDa cut-off (MWCO) for 5 h at RT. After this time buffer was exchanged to UN300 BF (20 mM HEPES pH 8, 300 mM NaCl, 1 mM TCEP, 5% glycerol) and left overnight at 4 °C. After dialysis, the sample was concentrated using Centricon® Plus-70 MWCO 30 kDa centrifugal filter devices (Merck, cat. no. [redacted]) and as the last step of the purification protein was applied to size exclusion chromatography column Hi-Load™ 16/600 Superdex™ 200 pg (GE Healthcare cat. no. GE28-9893-35) equilibrated in UN300 BF, as a quality control step to compare oligomeric state of EXOG variants. Fractions containing dimeric EXOG were collected, concentration was determined spectrophotometrically using extinction coefficient $\varepsilon_{280(dimer)}$= 77 030 M^{-1} cm^{-1} and protein was stored at -80 °C in UN300 BF.

7.4.2. APE1

BL21 (DE3) pLysS *E. coli* cells were transformed with pET 15b plasmid harboring N-terminal His_6 tag APE1 with thrombin cleavage site. Cells were grown in fresh LB at 37 °C (orbital shaker) until $OD_{600} \approx 0.6$, then cells were induced with 0.5 mM IPTG.

After induction the temperature was lowered to 18 °C and culture was grown for 16 h. Cells were harvested by centrifugation (4950 RCF, 10 min, 4°C). Cells were resuspended in Lysis Buffer (20 mM Tris-HCl pH 8, 0.5 M NaCl, 0.1 mM EDTA) with addition of protease inhibitors (E64 (Sigma-Aldrich, cat. no. E3132), pepstatin A (Sigma-Aldrich, cat. no. P5318), and leupeptin (Sigma-Aldrich, cat. no. 62070)) prior to lysis with sonication (50% amplitude, 30 s bursts followed by 1.5 min pauses on ice, total sonication time 5 min). The lysate was centrifuged (31 000 RCF, 30 min, 4 °C) and supernatant was loaded onto Protino® NiNTA agarose resin (Macherey-Nagel, cat. no. 745400.100) equilibrated with Wash Buffer (20 mM Tris-HCl pH 8, 0.5 M NaCl, 40 mM imidazole, 0.1 mM EDTA) and then extensively washed with the same buffer. Protein was eluted with Elution Buffer (20 mM Tris-HCl pH 8, 0.5 M NaCl, 200 mM imidazole, 0.1 mM EDTA) and fractions containing APE1 were combined. Sample was dialyzed against Thrombin Cleavage Buffer (20 mM Tris-HCl pH 8, 100 mM NaCl, 0.1 mM DTT, 1 mM EDTA, 10% glycerol) for 2 h at 4 °C using a dialysis membrane with a 10 kDa cut-off (MWCO). Then buffer was exchanged for another 2 liters and Thrombin from human plasma (Sigma-Aldrich, cat. no. T6884) was added for overnight cleavage of the His_6 tag. Thrombin was used at a ratio of 10 U per 1 mg of tagged protein. After overnight cleavage protein was applied onto Protino® NiNTA agarose resin (Macherey-Nagel, cat. no. 745400.100) equilibrated with Thrombin Cleavage Buffer to collect protein without His_6 tag. Protein was concentrated using Vivaspin® Turbo 15 MWCO 10 kDa (Sartorius, cat. no. VS15T02) and applied to size exclusion chromatography column Hi-Load™ 16/600 Superdex™ 200 pg (GE Healthcare, cat. no. GE28-9893-35) equilibrated with Thrombin Cleavage Buffer. Fractions containing APE1 were collected, concentration was determined spectrophotometrically using extinction coefficient ε_{280}= 56 567 M^{-1} cm^{-1} and protein was stored at -80 °C in the Thrombin Cleavage Buffer.

7.4.3. RNase H1

pET 15b plasmid encoding recombinant human RNase H1 (27-286 aa) with N-terminal thrombin-cleavable His_6 tag was transformed into BL21 (DE3) pLysS *E. coli* cells. Cells were grown in fresh LB at 37 °C (orbital shaker) until $OD_{600} \approx 0.9$, then cells were induced with 0.4 mM IPTG. After induction the temperature was lowered to 16 °C and culture was grown for 16 h. Cells were harvested by centrifugation (4950 RCF, 10 min, 4 °C) and later resuspended in Lysis Buffer (40 mM HEPES pH 7, 1 M NaCl,

5% glycerol, 2.8 mM β-mercaptoethanol and 10 mM imidazole) with the addition of protease inhibitors (E64 (Sigma-Aldrich, cat. no. E3132), pepstatin A (Sigma-Aldrich, cat. no. P5318), and leupeptin (Sigma-Aldrich, cat. no. 62070)) prior to lysis with sonication (50% amplitude, 30 s bursts followed by 1.5 min pauses on ice, total sonication time 5 min). Lysate was centrifuged (31 000 RCF, 30 min, 4 °C) and soluble fraction was applied to Protino® NiNTA agarose resin (Macherey-Nagel, cat. no. 745400.100) equilibrated with Lysis Buffer containing 60 mM imidazole. Protein was eluted with Lysis Buffer containing 300 mM imidazole. Collected fractions were dialyzed using a dialysis membrane with a 10 kDa MWCO against Buffer B (40 mM HEPES, 100 mM NaCl, 5% glycerol, 2 mM DTT, and 0.5 mM EDTA) for thrombin cleavage that was done overnight at 4 °C. Thrombin (Sigma-Aldrich, cat. no. T6884) was used at a ratio of 10 U per 1 mg of tagged protein. Cleaved protein was loaded onto Protino® NiNTA agarose resin (Macherey-Nagel, cat. no. 745400.100) and flow-through containing cleaved protein was collected. RNase H1 was concentrated using Vivaspin® Turbo 15 (Sartorius) and loaded onto a Hi-Load™ 16/600 Superdex™ 200 gel filtration column (GE Healthcare cat. no. GE28-9893-35). Protein fractions were concentrated using Vivaspin® Turbo 4 MWCO 10 kDa (Sartorius, cat. no. VS04T02) in the Storage Buffer (20 mM HEPES pH 7, 200 mM NaCl, 5% glycerol, 0.5 mM EDTA, 2 mM DTT). Concentration was determined spectrophotometrically using extinction coefficient ε_{280}= 47 565 M^{-1} cm^{-1}. Final protein sample was stored at -80 °C in the Storage Buffer (20 mM HEPES pH 7.0, 200 mM NaCl, 30% glycerol, 2 mM DTT, and 0.5 mM EDTA).

7.4.4. Pol γA

The Pol γA coding sequence (protein lacks MLS and 10 of the 13 sequential glutamines (residues 43–52)) was transferred to the baculovirus genome using pBacPak9 shuttle vector. C-terminally His-tagged Pol γA was expressed in Sf9 insect cells. Cells were infected at a density of 1-2 x 10^6 cells/ml and grown in Sf900-II SFM growth medium at 27 °C. After 72 h cells were centrifuged (500 RCF, 15 min., 4 °C), washed with Sf900-II SFM growth medium and frozen in liquid nitrogen. In the next step, cells were resuspended in Lysis Buffer (15 mM HEPES pH 7.6, 75 mM KCl, 2.25 mM $CaCl_2$, 1.5 mM Magnesium acetate, 250 mM sucrose, 1.25% glycerol, 0.375% NP-40, 0.025% Triton X-100) supplemented with protease inhibitors (E64 (Sigma-Aldrich, cat. no. E3132), pepstatin A (Sigma-Aldrich, cat. no. P5318) and leupeptin

(Sigma-Aldrich, cat. no. 62070). Lysate was centrifuged (1500 RCF, 10 min, 4 °C) and supernatant fraction was mixed stepwise with Buffer K (20 mM HEPES pH 8, 2 M KCl, 5% glycerol) to final 0.5 M KCl concentration. After 30 min the sample was centrifuged (31 000 RCF, 30 min, 4 °C). Supernatant was applied onto Talon Superflow resin (GE Healthcare), resin was washed using Wash Buffer (20 mM HEPES pH 7.2, 0.2 M KCl, 5 mM imidazole, 5% glycerol) and protein was eluted with Elution Buffer (20 mM HEPES pH 7.2, 0.2 M KCl, 200 mM imidazole, 5% glycerol). Protein fractions were combined, concentrated using Vivaspin® Turbo 15 MWCO 50 kDa (Sartorius, cat. no. VS15T32) with simultaneous exchange to SEC Buffer (20 mM HEPES pH 7.5, 140 mM KCl, 10 mM β-mercaptoethanol, 5% glycerol) and applied onto a Hi-Load™ 16/600 Superdex™ 200 gel filtration column (GE Healthcare cat. no. GE28-9893-35). Concentration was determined spectrophotometrically using extinction coefficient ε_{280}= 242 295 M^{-1} cm^{-1} and protein was stored at -80 °C in the SEC Buffer.

7.4.5. Pol γB

pET 22b (+) plasmid encoding Pol γB without MLS was transformed into Rosetta 2 pLysS *E. coli* cells. Cells were grown in fresh LB at 37 °C (orbital shaker) until $OD_{600} \approx 0.6$, then cells were induced with 0.4 mM IPTG. After induction the temperature was lowered to 18 °C and culture was grown for 16 h. Cells were harvested by centrifugation (4950 RCF, 10 min, 4 °C) and later resuspended in Lysis Buffer (20 mM HEPES pH 8, 0.3 M KCl, 1 mM β-mercaptoethanol, 5% glycerol, 0.1% Triton X-100) with addition of protease inhibitors (E64 (Sigma-Aldrich, cat. no. E3132), pepstatin A (Sigma-Aldrich, cat. no. P5318) and leupeptin (Sigma-Aldrich, cat. no. 62070). Cells were lysed through sonication (50% amplitude, 30 s bursts followed by 1.5 min pauses on ice, total sonication time 5 min) and centrifuged (31 000 RCF, 30 min, 4 °C). The soluble fraction was collected. To remove nucleic acid cationic polyelectrolyte poly(ethyleneimine) (PEI) was added to final concentration of 0.1%. Sample was again centrifuged (31 000 RCF, 30 min, 4 °C) to remove the precipitation. Supernatant was applied to Protino® NiNTA agarose resin (Macherey-Nagel, cat. no. 745400.100) equilibrated with Lysis Buffer and subsequently resin was washed extensively with Wash Buffer (20 mM HEPES pH 7.5, 0.2 M KCl, 5 mM imidazole, 5% glycerol). Protein was eluted using Elution Buffer (20 mM HEPES pH 7.5, 0.1 M KCl, 200 mM imidazole, 5% glycerol). Collected fractions were combined and diluted with Dilution Buffer (20 mM HEPES pH 7.5, 5% glycerol) to adjust sample

conductivity before application onto the HiTrap SP HP cation exchange chromatography column (Cytiva, cat. no. 17115101). The column was equilibrated with the Low Salt Buffer (20 mM HEPES pH 7.5, 60 mM KCl, 5% glycerol) and eluted with linear gradient of KCl (60 mM-700 mM). Pol γB-containing fractions were pooled, concentrated using Vivaspin® Turbo 15 MWCO 30 kDa (Sartorius, cat. no. VS15T22) and loaded onto gel filtration column Hi-Load™ 16/600 Superdex™ 200 pg (GE Healthcare, cat. no. GE28-9893-35) equilibrated with (20 mM HEPES pH 7.5, 140 mM KCl, 5 mM β-mercaptoethanol, 5% glycerol). Concentration was determined spectrophotometrically using extinction coefficient ε_{280}= 143 280 M^{-1} cm^{-1} and protein was stored at -80 °C in the SEC buffer.

7.4.6. FEN1

The full-length human FEN1 (amino acids: 2−380) with N-terminal His_6 tag followed by SUMO protein was expressed from the pE expression vector in *E. coli* Rosetta 2 (DE3) competent cells. Overnight culture was diluted into fresh 2xYT media and grown at 30 °C to an OD_{600} ≈ 0.5, at which point the temperature was reduced to 16 °C. The culture was induced with 0.4 mM IPTG and grown for additional 16 h. Cells were harvested by centrifugation (4950 RCF, 10 min, 4 °C) and later lysed by sonication (50% amplitude, 30 s bursts followed by 1.5 min pauses on ice, total sonication time 5 min) in Lysis Buffer (20 mM Tris pH 7, 1 M NaCl, 5 mM imidazole, 0.1% lysozyme). Lysis Buffer was supplemented with protease inhibitors (E64 (Sigma-Aldrich, cat. no. E3132), pepstatin A (Sigma-Aldrich, cat. no. P5318) and leupeptin (Sigma-Aldrich, cat. no. 62070). The lysate was centrifuged (12 000 RCF, 30 min, 4 °C) and loaded onto Talon Superflow (GE Healthcare, cat. no. GE28-9574-99) resin pre-equilibrated with Talon A Buffer (20 mM Tris-HCl pH 7, 1 M NaCl, 5 mM imidazole). The resin was washed with Talon B Buffer (20 mM Tris-HCl pH 7, 500 mM NaCl, 40 mM imidazole, 0.1% Tween 20), after which protein was eluted with Talon Elution Buffer (20 mM Tris-HCl pH 7, 1 M NaCl, 250 mM imidazole). Fractions containing FEN1 were pooled and incubated with SUMO protease (Ulp1) at a Ulp1:FEN1 ratio of 1:250 (w/w) to cleave off the His-SUMO tag. Cleavage was carried out during dialysis against Ulp1 Cleavage Buffer (20 mM Tris-HCl pH 7, 500 mM NaCl, 5 mM imidazole) using dialysis membrane with a 30 kDa MWCO, overnight at 4 °C. FEN1 protein was loaded onto Talon Superflow resin to remove the cleaved His-SUMO tag and SUMO protease. A flow-through fraction with cleaved FEN1 protein was collected.

FEN1 protein was diluted to a final salt concentration of 250 mM NaCl with Anion A1 Buffer (20 mM Tris-HCl pH 8, 1 mM EDTA) and $CaCl_2$ was added to the final concentration of 2 mM. FEN1 was loaded on HiTrap Heparin HP column (GE Healthcare, cat. no. 17-0406-01) that was equilibrated with Heparin A Buffer (25 mM Tris-HCl pH 7.5, 1 mM $CaCl_2$). The column was washed with Heparin A buffer, and protein was eluted with a linear gradient against Heparin B Buffer (25 mM Tris-HCl pH 7.5, 1 M NaCl, 1 mM $CaCl_2$). Fractions containing FEN1 protein were collected and loaded on the Sephacryl S-300 HR column (GE Healthcare, cat. no. GE17-0599-10) equilibrated with SEC Buffer (50 mM HEPES pH 7.5, 1 mM $CaCl_2$, 100 mM KCl, 5 mM DTT, 10% glycerol). Fractions containing FEN1 were dialyzed against Storage Buffer (50 mM HEPES pH 7.5, 1 mM $CaCl_2$, 100 mM KCl, 5 mM DTT, 40% glycerol). Collected fractions were combined and concentration was determined spectrophotometrically using extinction coefficient ε_{280}= 23 107 M^{-1} cm^{-1} after which protein was flash frozen and stored at -80 °C in the Storage Buffer.

7.4.7. MGME1

The full-length human MGME1 lacking MLS sequence (1-21 aa) with N-terminal His_6 tag followed by SUMO protein was expressed from the pET 28 His-SUMO plasmid in *E. coli* BL21 Star™ (DE3) competent cells. Overnight culture was diluted into fresh LB media and grown at 37 °C to an $OD_{600} \approx 0.6$, at which point the temperature was reduced to 18 °C. Culture was then induced with 0.2 mM IPTG, and grown for additional 18 h. Cells were harvested by centrifugation (4950 RCF, 10 min, 4 °C) and disrupted by sonication (50% amplitude, 30 s bursts followed by 1.5 min pauses on ice, total sonication time 5 min) in Lysis Buffer (20 mM Tris-HCl pH 8, 500 mM NaCl, 25 mM imidazole) with addition of protease inhibitors (E64 (Sigma-Aldrich, cat. no. E3132), pepstatin A (Sigma-Aldrich, cat. no. P5318) and leupeptin (Sigma-Aldrich, cat. no. 62070). The lysate was centrifuged (31 000 RCF, 30 min, 4 °C) and loaded onto HisTrap FF crude column (Cytiva, cat. no. 11000458) equilibrated with Lysis Buffer. Resin was washed extensively with Lysis Buffer, and protein was eluted with linear gradient against Elution Buffer (20 mM Tris-HCl 8, 500 mM NaCl, 500 mM imidazole). Fractions containing MGME1 were combined and dialysed, using a dialysis membrane with a 10 kDa MWCO, for 2 hours against Lysis Buffer at 4 °C. The dialysis buffer was then replaced with a fresh portion of Lysis Buffer, SUMO protease (Ulp1) was added at a Ulp1:MGME1 ratio of 1:250 (w/w) and cleavage of the

His-SUMO tag was carried out during dialysis overnight at 4 °C. MGME1 protein was loaded onto HisTrap FF crude column (Cytiva, cat. no. 11000458) to remove the cleaved His-SUMO tag and SUMO protease. Flow-through was concentrated using Vivaspin® Turbo 4 MWCO 10 kDa (Sartorius, cat. no. VS04T02) and loaded onto gel filtration Superdex 30 Increase 10/300 GL column (Cytiva, cat. no. 29219757) equilibrated with SEC Buffer (20 mM Tris-HCl pH 8, 200 mM NaCl, 1 mM DTT). Fractions containing MGME1 were collected and dialyzed against SEC Buffer with 40% glycerol overnight. Concentration was determined spectrophotometrically using extinction coefficient ε_{280}= 47 057 M^{-1} cm^{-1} after which protein was flash frozen and stored at -80 °C.

7.4.8. Ligase III

The full-length human Lig III lacking MLS sequence was expressed from the pET 28 His-SUMO plasmid in *E. coli* Rosetta 2 (DE3) competent cells. Overnight culture was diluted into fresh LB media and grown at 30 °C to an $OD_{600} \approx 0.6$, at which point the temperature was reduced to 16 °C. After that, culture was induced with 0.3 mM IPTG and grown for additional 24 h. Cells were harvested by centrifugation (4950 RCF, 10 min, 4 °C) and later resuspended in Lysis Buffer (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 3 mM EDTA, 10 mM β-mercaptoethanol, 10 % glycerol) with addition of protease inhibitors (E64 (Sigma-Aldrich, cat. no. E3132), pepstatin A (Sigma-Aldrich, cat. no. P5318) and leupeptin (Sigma-Aldrich, cat. no. 62070)). Cells were disrupted by sonication (50% amplitude, 30 s bursts followed by 1.5 min pauses on ice, total sonication time 5 min). The lysate was centrifuged (39 000 RCF, 1 h, 4 °C) and filtrated through 0.45 μm nitrocellulose filter. Next, protein was applied to Cellulose phosphate resin equilibrated with (50 mM Tris-HCl pH 7.5, 250 mM NaCl, 3 mM EDTA, 2 mM β-mercaptoethanol, 10 % glycerol) and then eluted with linear gradient against Elution Buffer (50 mM Tris-HCl pH 7.5, 2 M NaCl, 3 mM EDTA, 2 mM β-mercaptoethanol, 10 % glycerol). Fractions containing Lig III were pooled, combined and applied onto HisTrap Fast Flow (GE Healthcare, cat. no. GE17-5319-01) equilibrated with LigA2 Buffer (50 mM Tris-HCl pH 7.5, 1 M NaCl, 20 mM imidazole, 2 mM β-mercaptoethanol, 10 % glycerol). Resin was washed extensively with LigA2 Wash Buffer (50 mM Tris-HCl pH 7.5, 200 mM NaCl, 20 mM imidazole, 2 mM β-mercaptoethanol, 10 % glycerol) and then protein was eluted with linear gradient against LigA2 Elution Buffer (50 mM Tris-HCl pH 7.5, 200 mM NaCl, 600 mM imidazole, 2 mM β-mercaptoethanol, 10 % glycerol). Samples containing Lig III were

collected, combined and dialyzed using a dialysis membrane with a 30 kDa MWCO, against LigA3 Buffer (50 mM Tris-HCl pH 7.5, 200 mM NaCl, 2 mM β-mercaptoethanol, 10 % glycerol) for 2 h at 4 °C. The dialysis buffer was then replaced with the fresh portion of LigA3 buffer, Ulp1 was added at a Ulp1: Lig III ratio of 1:250 (w/w), and cleavage of the His-SUMO tag was carried out during dialysis overnight at 4 °C. Protein after cleavage was then loaded onto Blue Sepharose 6 Fast Flow resin (GE Healthcare, cat. no. GE17-0948-01) equilibrated with LigA3 buffer and washed extensively with the same buffer. In the next step protein was eluted with linear gradient against LigB3 Elution Buffer (50 mM Tris-HCl pH 7.5, 2 M NaCl, 2 mM β-mercaptoethanol, 10 % glycerol). Fractions containing Lig III were collected and concentrated using Vivaspin® Turbo 4 MWCO 30 kDa (Sartorius, cat. no. VS04T22). Next protein was incubated with 20 mM $MgCl_2$ (20 µl of 1 M per each 1 ml of sample) and 1 mM ATP (10 µl of 100 mM per each 1 ml of sample) for 30 minutes on ice. For the final step protein was loaded onto gel filtration column Hi-Load™ 16/600 Superdex™ 200 pg (GE Healthcare, cat. no. GE28-9893-35) equilibrated with Lig4 SEC Storage Buffer (50 mM NaMOPS pH 7.5, 300 mM NaCl, 1 mM TCEP, 10% glycerol). Fractions containing Lig III were collected and concentration was determined spectrophotometrically using extinction coefficient ε_{280}= 79 988 M^{-1} cm^{-1} after which protein was flash frozen and stored at -80 °C.

7.5. High-throughput small-molecule screening (HTS) of LOPAC®1280 library

To identify candidate EXOG inhibitors, a dsDNA substrate (S12; Table 2) carrying a fluorescein at the 3’ end of the top strand and Black Hole Quencher 1 (BHQ-1) at the 5’ end of the bottom strand was designed to report dinucleotide cleavage by the 5’-exonuclease activity of EXOG. Excision of the 5′ dinucleotide from the top strand releases the quencher at the bottom strand, resulting in an increase in fluorescence. Reactions were carried out in 96-well plates in HK140 buffer (20 mM HEPES, pH 7.4, 140 mM KCl, 10 mM TCEP, 5% glycerol). Each well contained buffer, 100 nM EXOG and the test compound at a final concentration of 12 µM (final dimethyl sulfoxide (DMSO) concentration was 0.12%). After 10 min pre-incubation of protein and compound at 25 °C, 200 nM DNA substrate and 20 mM $MgCl_2$ were added to start the reaction. Fluorescence was measured at excitation/emission wavelengths of 490/525 nm

at 0, 5, 10, 20 and 30 min. Inhibition of EXOG activity was calculated relative to the no-inhibitor control. The LOPAC®1280 library (Library of Pharmacologically Active Compounds; Sigma-Aldrich, cat. no. LO4200) was screened, comprising 1280 compounds supplied in 96-well format (one compound per well) at 10 mM in dimethyl sulfoxide (DMSO). Compounds identified as modulating EXOG activity were purchased from commercial vendors, dissolved in DMSO and stored at -20 °C as 10 mM stock solutions.

7.6. EXOG-ATA co-crystallization and data collection

7.6.1. Crystallization trials

Co-crystallization was carried out to study the interaction sites between EXOG and its inhibitor aurintricarboxylic acid (ATA) (Table 3). Compound was dissolved in ≥99.9% DMSO and stored at −20 °C as 100 mM stock solutions. The EXOG-ATA complex was prepared by mixing EXOG (stock concentration: 7 mg/ml) and ATA at final concentration of 1 mM in the HK140 buffer (20 mM HEPES pH 7.5, 140 mM KCl, 1 mM TCEP, 5% glycerol). Prior to protein-compound complex assembly, EXOG was clarified by centrifugation (10 min, maximum speed, 4 °C). The complex was incubated at room temperature for 20 min before crystallization setup. Crystallization trials were performed using sitting-drop vapor diffusion in 96-well Intelli-Plates (cat. No. 102-0001-20), set up at room temperature using a Mosquito® liquid-handling robot (TTP Labtech). Drops were dispensed in two protein-to-reservoir ratios: 1:1 (400 nl + 400 nl) and 2:1 (800 nl + 400 nl). Nine commercial sparse matrix screens: Crystal Screen, Index, MembFac, Crystal Screen Lite (Hampton Research), Morpheus I, Morpheus II, Morpheus III, Helix and JCSG Plus (Molecular Dimensions) were used for this experiment. Plates were incubated at 18 °C and monitored regularly by light microscopy. Crystallization optimization of the EXOG-ATA complex was prepared by mixing EXOG (stock concentration: 7 mg/ml) and ATA at different concentrations in the HK140 buffer. As described above, prior to protein-compound complex assembly, EXOG was clarified by centrifugation (10 min, maximum speed, 4 °C). The complex was incubated at room temperature for 20 minutes before crystallization setup. Crystallization optimization was performed using sitting-drop vapor diffusion in 24-well Cryschem Plates (Hampton Research, cat. no. HR3-159). Drops were dispensed in protein-to-reservoir ratio: 1:1 (2 µl + 2 µl). The reservoir volume was 500 µl. A custom matrix was prepared based on

condition number 4 (0.1 M Tris-HCl pH 8.5, 2.0 M ammonium sulfate) from Crystal Screen (Hampton Research, cat. no. HR2-130) with the addition of 0.12 M $MgCl_2$ for protein stabilization. Ammonium sulfate varied across columns at six concentrations (1.8 M, 1.925 M, 2 M, 2.1 M, 2.2 M, 2.25 M), and 0.1 M Tris-HCl was varied across rows at four different pH values (8.0, 8.2, 8.5, 8.8). ATA concentration varied across the plates at four concentrations (0.1 mM, 0.2 mM, 0.5 mM, 1 mM). All stock solutions were freshly prepared, and ATA aliquot was freshly thawed before each set up. Plates were incubated at 18 °C and monitored regularly by light microscopy. The best results were obtained for the crystallization condition: 1.925 M ammonium sulphate, 0.1 M Tris pH 8.2, 0.12 M $MgCl_2$ supplemented with 0.25 mM ATA.

7.6.2. Data collection, structure solution and refinement

X-ray diffraction data were collected at the XRD2 beamline of Elettra Sincrotrone Trieste (Trieste, Italy). Data were recorded at a wavelength of 1.0001 Å with the crystals held at 100 K on a Pilatus 6M detector. For each dataset, 720 images were recorded over a 360° rotation (0.5° per image). Data were collected and processed for a total of 100 crystals using autoPROC (Vonrhein et al., 2011) and anisotropy was accounted for with STARANISO. The dataset used for structure solution belonged to space group $P2_1$ (unit-cell parameters a = 73.55 Å, b = 83.84 Å, c = 76.38 Å, β = 115.42°) and was 99.7 % complete to 1.90 Å. Data quality was assessed with phenix.xtriage; the intensity statistics indicated that the data were not twinned, although translational non-crystallographic symmetry was detected. Phases were obtained by molecular replacement with Phaser using the previously published EXOG structure (PDB 5T40) as the search model. A single solution was obtained (LLG = 18420, TFZ = 9.9). An initial model was built automatically with AutoBuild and refined with phenix.refine to a resolution of 1.9 Å, yielding a final R_{work} of 0.165 and R_{free} of 0.189 (data-collection and refinement statistics are given in Table 4). Finally, the refined ligand-free model was used to search for electron density corresponding to ATA using LigandFit. All steps of structure solution and refinement were performed within the Phenix suite (Liebschner et al., 2019).

Table 4. Data collection and refinement statistics.

Data collection	
Beamline	XDR2, Eletra Sincrotrone Tireste
Wavelength (Å)	1.0001
Space group	$P2_1$
Cell dimensions	
a, *b*, *c* (Å)	73.55, 83.84, 76.38
α, β, γ (°)	90, 115.42, 90
Resolution (Å)	66.43 -1.9 (1.93 - 1.90)
Total reflections	396331 (18179)
R_{sym} or R_{merge}	0.069 (0.682)
$I/\sigma I$ $CC_{1/2}$	15.0 (2.8) 0.999 (0.617)
Completeness (%)	99.7 (99.3)
Multiplicity	6.0 (5.6)
Refinement	
Resolution (Å)	66.43 - 1.90
Unique reflections	63 988
R_{work}/R_{free}	0.165/0.189
No. atoms	
Protein	4808
Ligand/ion	12
Water	786
***B*-factors**	
Protein	27.82
Ligand/ion	32.64
Water	40.28
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.829
Ramachandran statistics	
Ramachandran favored (%)	98.32

Ramachandran outliers (%)	0.17
Rotamer outliers (%)	0.38
Clashscore	3.67
C-beta outliers	0
Overall score	1.16

7.7. Endonuclease activity assay

Endonuclease activity was tested on ΦX174 DNA (New England Biolabs, cat. no. N3023S) which is a single-stranded, circular, virion DNA. All the reactions were performed in HK140 Buffer. Each reaction contained 30 nM DNA and increasing concentrations of EXOG (0.4, 0.8, 1.5, 3, 6, 12.5 and 25 nM), with 5 mM $MgCl_2$, in a final volume of 20 µl. After 1 h incubation at 25 °C, reactions were stopped by addition of 5 µl of Stop Buffer 1 (250 mM EDTA, 0.1% SDS), followed by incubation at 65 °C for 5 min. 5 µl of the DNA Gel Loading Dye (6x) (Thermo Fisher Scientific, cat. no. R0611) was added to the sample. GeneRuler™ 1kb DNA Ladder (Thermo Fisher Scientific, cat. no. SM0311) was used as a size marker. Cleavage products were resolved and analyzed on 1% agarose gel in TAE buffer (0.4 M Tris acetate, 0.01 M EDTA) stained with ethidium bromide (EtBr) (0.6 µg/ml). Gel was resolved for 1 h, 100 V. Cleavage products were visualized using Gel Doc (Bio-Rad). Substrate depletion was quantified using ImageJ (http://rsb.info.nih.gov/ij/) by measuring the intact substrate band intensity, relative to a control containing 30 nM DNA and buffer only, in the presence of magnesium ions. The assay was performed in three independent experiments.

7.8. Inhibition of endonuclease activity

Inhibition of endonuclease activity was performed at 25 °C in HK140 Buffer. ΦX174 DNA (New England Biolabs) premix was prepared with 5 mM $MgCl_2$ (final). Reactions (final volume: 20 µl) contained 30 nM DNA, ATA at final concentrations of 62.5 nM, 125 nM, 200 nM, 250 nM, 375 nM, 500 nM, 1 µM, 2 µM, 4 µM and 8 µM (serial dilutions; Table 3) and 20 nM EXOG. After 1 h of incubation reactions were stopped and processed for agarose gel analysis as described in Section 7.7. (Stop Buffer 1; 6× DNA Gel Loading Dye; 1% agarose/TAE with ethidium bromide [0.6 µg/ml]; GeneRuler™ 1 kb DNA Ladder as size marker; 1 h at 100 V; Gel Doc, Bio-Rad). Substrate depletion was quantified in ImageJ by measuring the intact-substrate band

intensity, relative to a no-enzyme control containing 30 nM DNA and the highest ATA concentration (with buffer in place of EXOG), in the presence of magnesium ions. The assay was performed in three independent experiments.

7.9. Titration of the ΦX174 DNA with aurintricarboxylic acid and mitoxantrone

To establish whether ATA or mitoxantrone (MTX) interact directly with the DNA substrate rather than with EXOG, a fixed concentration of ΦX174 DNA (30 nM final) was titrated with increasing concentrations of ATA (15 nM, 30 nM, 60 nM, 125 nM, 200 nM, 250 nM, 375 nM, 500 nM, 1 μM and 2 μM final) or MTX (0.3, 0.7, 1.5, 3, 6.25, 12.5, 25, 50, 100 and 200 μM final) in HK140 Buffer containing 5 mM $MgCl_2$ (final volume 20 μl). Samples were incubated for 1 h at 25 °C and then processed for agarose gel analysis as described in Section 7.7. (Stop Buffer 1; 6× DNA Gel Loading Dye; 1% agarose/TAE with ethidium bromide 0.6 μg/ml; 1 h at 100 V; Gel Doc, Bio-Rad). The amount of DNA was quantified in ImageJ by measuring the intact-substrate band intensity, relative to a control containing 30 nM DNA and buffer only, in the presence of magnesium ions. The assay was performed in three independent experiments.

7.10. DNA and RNA substrates preparation

All the oligonucleotides were purchased from Metabion International AG and are listed in Table 1. To prepare the duplex substrates (Table 2), the fluorescein-labelled strand (4 μM final) was mixed with a slight excess of the complementary unlabeled strand (4.5 μM final) to ensure complete incorporation of the labelled strand into the duplex. Strands were annealed in T2100 Buffer (50 mM Tris–HCl pH 8, 100 mM NaCl, 0.1 mM EDTA, 10% glycerol) at 95 °C for 5 min and allowed to cool slowly to room temperature. Substrates were labelled with fluorescein or Cy3 by the manufacturer as indicated in Table 1.

7.11. Exonuclease activity assay

EXOG activity was tested on ssDNA, ssRNA, dsDNA, 1-nt gapped dsDNA and 15-nt DNA flap dsDNA (S1, S2, S4, S5 and S9 substrates; Table 2). Fluorescently labelled DNA or RNA substrates were prepared in the HK140 BF containing $MgCl_2$ (10 mM final concentration). Reactions (final volume 20 μl) contained 200 nM of the substrate and

increasing EXOG concentrations (1, 5, 10, 25, 50, 100, 200 and 400 nM final). All the reactions were performed at 25 °C and initiated by the addition of protein. After 5 min of incubation reactions were stopped by the addition of 40 µl of the Stop Buffer 2 (78% formamide, 50 mM EDTA, 0.1% SDS), which was then followed by incubation at 95 °C for 4 min. Cleavage products were resolved on 23% denaturing polyacrylamide gels (23% acrylamide, 7 M urea, 0.5 x TBE) run in TBE buffer (130 mM Tris, 45 mM boric acid, 2.5 mM EDTA) using Owl Dual-Gel Vertical Electrophoresis System (Thermo Fisher Scientific, cat. no. P10DS-1). To reach the temperature necessary for denaturing conditions (~55 °C), gel was pre-run for 1 h and 30 min at 40 W. Samples were then resolved for 50 min at 20 W and visualized using Amersham™ Typhoon™ Biomolecular Imager (GE Healthcare). Substrate depletion was quantified using ImageJ by measuring the intact-substrate band intensity, relative to a no-enzyme control containing the same substrate and buffer, in the presence of magnesium ions. The assay was performed in three independent experiments.

7.12. Inhibition of exonuclease activity

Inhibition of the 5'-exonuclease activity of EXOG was performed on the same substrates as mentioned in chapter 7.11. Reactions (final volume 20 µl) contained 200 nM of DNA substrate in HK140 Buffer with 10 mM $MgCl_2$ and ATA at concentrations ranging from 0.1 to 50 µM (depending on the substrate; marked in the figure legends in the Results section). Substrate was pre-mixed with ATA, and the reaction was initiated by adding EXOG to a final concentration of 100 nM. To test the influence of inhibitor preincubation, EXOG was first incubated with ATA (0.5-12.5 µM) for 5 min prior mixing with the substrate. After 5 min, reactions were stopped by the addition of 40 µl of the Stop Buffer 2, followed by incubation at 95 °C for 4 min. Cleavage products were resolved by 23% denaturing polyacrylamide gel electrophoresis as described in 7.11. and visualized using Amersham™ Typhoon™ Biomolecular Imager (GE Healthcare). Substrate depletion was quantified using ImageJ (http://rsb.info.nih.gov/ij/) by measuring the intact-substrate band intensity, relative to a no-enzyme control containing the same substrate and the highest ATA concentration (buffer in place of EXOG), in the presence of magnesium ions. IC_{50} values were determined by fitting a four-parameter logistic function in SigmaPlot. The assay was performed in three independent experiments.

7.13. ATA inhibition of off-target proteins

All activity and inhibition studies for off-target proteins were performed in HK140 BF with 10 mM $MgCl_2$ at 25 °C. Substrates used in these experiments are listed in Table 2. First, the basal activity of each protein was assessed; inhibition by ATA was then determined. For activity assays, 200 nM enzyme-specific DNA/RNA substrate was titrated with the increasing concentrations of the protein (final volume of reaction 20 µl). For inhibition studies 200 nM substrate was first pre-mixed with the ATA and then reaction was initiated by adding protein at a fixed concentration (final volume 20 µl). For each enzyme reactions were stopped after 5 min by the addition of 40 µl of Stop Buffer 2, followed by incubation at 95 °C for 4 min. Reaction products were separated by 23% denaturing PAGE as described in Section 7.11 and visualized using an Amersham™ Typhoon™ Biomolecular Imager (GE Healthcare). Depending on the enzyme substrate depletion (nucleases) or product formation (extension by Pol γ or ligation by Lig III) was quantified using ImageJ (http://rsb.info.nih.gov/ij/) relative to a no-enzyme control containing the same substrate and the highest ATA concentration (buffer in place of protein), in the presence of magnesium ions. The assay was performed in three independent experiments.

7.13.1. FEN1

In the basal activity assay fluorescently labelled nicked 15-nt $(dT)_{15}$ flap dsDNA substrate (S9; Table 2) was mixed with 5-200 nM of FEN1 protein to initiate the reaction. In case of ATA inhibition assay 200 nM of the same DNA substrate was pre-mixed with ATA (0.5-8 µM) before addition of 100 nM of protein.

7.13.2. RNase H1

In the basal activity assay fluorescently labelled nicked 10RNA-10DNA substrate (S6; Table 2) was mixed with 2.5-200 nM of RNase H1 to initiate the reaction. In case of ATA inhibition assay 200 nM of RNA-DNA substrate was pre-mixed with ATA (0.8-12.5 µM) before addition of 100 nM protein.

7.13.3. Pol γAB

Fluorescently labelled primer extension substrate (S7; Table 2) was supplemented with 0.1 mM dNTPs and mixed with 1-200 nM of Pol γAB holoenzyme to initiate the reaction. In case of ATA inhibition assay, 200 nM of primer extension substrate was pre-mixed with ATA (0.3-50 μM) before 100 nM protein addition.

7.13.4. MGME1

200 nM of fluorescently labelled ssDNA substrate (S1; Table 2) was mixed with 2.5-200 nM of MGME1 to initiate the reaction. In case of ATA inhibition assay 200 nM of ssDNA substrate was pre-mixed with ATA (0.3-50 μM) before 100 nM protein addition.

7.13.5. Ligase III

200 nM of nicked phosphorylated dsDNA (S8; Table 2) was supplemented with 1 mM ATP and mixed with 5-200 nM of Lig III to initiate the reaction. In case of ATA inhibition assay, 200 nM of nicked phosphorylated dsDNA substrate was pre-mixed with ATA (0.3-50 μM) before 100 nM protein addition.

7.13.6. APE1

200 nM of fluorescently labelled dsDNA substrate with THF mimicking AP-site (S10; Table 2) was mixed with 1.25-200 nM of APE1 to initiate the reaction. In case of ATA inhibition assay 200 nM of THF substrate was pre-mixed with ATA (0.01-1.5 μM) before 10 nM protein addition.

7.14. Testing of the ATA-related compounds on EXOG exonuclease activity

Inhibition of exonuclease activity of EXOG was performed in HK140 BF at 25 °C. A 3'-fluorescein-labelled ssDNA substrate (S1; Table 2) was prepared in the HK140 BF containing 10 mM $MgCl_2$. Reactions (final volume 20 μl) contained 200 nM DNA substrate and the tested compound (Table 3) at concentrations of 0.5-12.5 μM). The substrate was pre-mixed with the compound, and the reaction was initiated by adding EXOG to a final concentration of 100 nM. After 5 min reactions were stopped by the addition of 40 μl of Stop Buffer 2, followed by incubation at 95 °C for 4 min. Cleavage products were resolved by denaturing polyacrylamide gel electrophoresis (as described

in 7.11.) and visualized using Typhoon fluorescence imager. Substrate depletion was quantified using ImageJ (http://rsb.info.nih.gov/ij/) by measuring the intact-substrate band intensity, relative to a no-enzyme control containing the same substrate and the highest compound concentration (buffer in place of EXOG), in the presence of magnesium ions. IC_{50} values were determined by fitting a four-parameter logistic function in SigmaPlot; for compounds that did not reach full inhibition within the concentration range tested, IC_{50} could not be determined.

7.15. Biolayer interferometry (BLI)

7.15.1. EXOG – ATA binding

BLI experiments were performed on Octet K2 instrument (ForteBio) using Octet® NTA Biosensors (Sartorius, cat. no. 18-5103) at 25 °C with shaking at 1000 RPM. Each assay comprised the following steps:

1. An initial baseline in HK140 BF with addition of 0.05% Triton X-100 (60 s),
2. Loading of C-terminally His-tagged EXOG (5 μg/ml) onto the sensor (300 s),
3. A wash in the same buffer (60 s),
4. An association phase, in which the immobilized protein was incubated with inhibitor at concentrations of 1 nM-10 μM (900 s),
5. A dissociation phase in buffer alone (900 s).

Binding responses at equilibrium (end of the association phase) were plotted against compound concentration, and apparent K_d values were determined by fitting a one-site specific-binding model in GraphPad Prism. Binding curves represent the mean of three independent experiments, error bars represent SD.

7.15.2. EXOG-DNA binding

BLI assays were performed with the use of BLItz instrument (Sartorius) at 25 °C, with shaking at 1500 RPM using Octet® SAX (streptavidin-coated) Biosensors (Sartorius, cat. no. 18-5117) for capture of the biotinylated substrate.
Each assay comprised of the following steps:

1. An initial baseline/wash in HK140 BF with addition of 0.05% Triton X-100 (30 s),
2. Immobilization of 3’-biotinylated 20-nt ssDNA (S11; Table 2) on the sensor (500 s),

3. Wash in the same buffer (300 s),
4. An association phase in which immobilized DNA was titrated with increasing concentrations of EXOG (2-1000 nM; 500 s),
5. A dissociation phase in buffer alone (900 s).

To estimate the influence of ATA on EXOG-DNA binding, the same protocol was used except for the association phase. Here, EXOG (8-2000 nM) was pre-incubated with 1 μM ATA for 5 min and this mixture was then applied during the association phase. The raw data was fitted with GraphPad Prism software using one site specific binding model for EXOG-DNA binding and one site specific binding with Hill slope for EXOG-ATA/DNA binding. Each point on the graph represents an average of three data sets.

7.16. Protein thermal shift assay

To assess whether ATA affects the thermal unfolding of EXOG, nanoDSF measurements were performed using a Tycho NT.6 instrument (NanoTemper) in HK140 BF. 20 μl reactions containing EXOG at 10 μM concentration and increasing concentrations of ATA (10-200 μM), together with an ATA-free control, were incubated at 25 °C for 5 min. Then samples were centrifuged at 15 000 x g for 10 min at 25 °C and applied to Tycho NT.6 capillaries (NanoTemper, cat. no. TY-C001). Intrinsic protein fluorescence at 330 and 350 nm was recorded over a temperature gradient from 35 to 95 °C, and the inflection temperature (T_i) was determined from the 350/330 nm ratio. Data represent the mean of three independent experiments.

7.17. Long amplicon PCR assay

To measure DNA integrity gene-specific semi-quantitative PCR assays using LongAmp Taq DNA Polymerase (New England Biolabs, Ipswich, MA) was used. Total nuclear or mitochondrial DNA was isolated using DNeasy Blood and Tissue Kit (QIAGEN). Nuclear DNA damage was assessed by quantification of the PCR amplification of the 9 kb nuclear-specific DNA fragment using PicoGreen fluorescent dye to detect double-stranded DNA (Quant-iT™ PicoGreen, Molecular Probes). Mitochondrial DNA damage was estimated by quantification of the PCR amplification of the 10 kb mitochondrial-specific DNA fragment using PicoGreen staining. Obtained data was normalized by the PCR amplification of 117 bp mitochondrial genome-specific fragment for correction of

the multiple copies of the mitochondrial genome. Preliminary assays were carried out to ensure the linearity of PCR amplification with respect to the number of cycles and DNA concentration. This experiment was performed by Bartosz Szczesny, PhD (UTMB).

7.18. MTT assay

The MTT method was performed using 3-(4,5-dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT). The compound was added to the cells at a final concentration of 0.5 mg/ml and cells were cultured at 37 °C for 1 h. The cells were washed with PBS, and the formazan dye was dissolved in DMSO. The amount of converted formazan dye was measured at 570 nm with a background measurement at 690 nm on a Powerwave reader (Biotek). The method detects Complex II-dependent mitochondrial activity and is often used to estimate mitochondrial function and/or cell viability. This experiment was performed by Bartosz Szczesny, PhD (UTMB).

7.19. LDH assay

Lactate dehydrogenase (LDH) release was determined as a cytotoxicity assay, a secondary measurement for determination of cell death. 30 µl of supernatant was saved before addition of MTT and mixed with 100 µl freshly prepared LDH assay reagent containing 85 mM lactic acid, 1 mM nicotinamide adenine dinucleotide (NAD^+), 0.27 mM N-methylphenazonium methyl sulfate (PMS), 0.528 mM 2-(4-Iodophenyl)-3-(4-nitrophenyl)-5-phenyl-2H-tetrazolium chloride (INT), and 200 mM Tris (pH 8.2). The changes in absorbance were read kinetically at 492 nm for 15 min (kinetic LDH assay) on a monochromator-based reader (Powerwave HT, Biotek) at 37 °C. LDH activity values are shown as V_{max} for kinetic assays in mOD/min. This experiment was performed by Bartosz Szczesny, PhD (UTMB).

7.20. Virtual screening of compound libraries

Structure-based virtual screening of the BioAscent Compound Collection against EXOG was performed by Dr Douglas Houston (University of Edinburgh) using the CODASS (Combined Docking And Similarity Searching) pipeline. The screened library comprised 124 732 commercially available compounds (BioAscent Compound Cloud Collection) and was docked without prior physicochemical filtering. Docking was carried out against the crystal structure of human EXOG (PDB: 5T40), with the conformational search restricted to a defined box around the binding site. Compounds were first docked with

AutoDock Vina (Trott and Olson, 2010), and the 10 000 top-scoring compounds were retained. These were re-docked with AutoDock 4 (Forli et al., 2016), from which the 1000 top-scoring compounds were retained. To improve the reliability of the predicted binding modes, consensus docking was applied (PoseMatch; Houston and Walkinshaw, 2013): the poses generated by Vina and AutoDock were compared and compounds whose two poses agreed within a root-mean-square deviation (RMSD) of less than 2.0 Å were retained, a criterion reported to identify correctly docked poses with approximately 82% reliability. The retained compounds were rescored with several independent scoring functions – DSX (Neudert and Klebe, 2011), X-Score v1.3 (Li et al., 2014), NNScore 1 (Durrant and McCammon, 2010), NNScore 2 (Durrant and McCammon, 2011) and RFScore-VS v2 (Wójcikowski et al., 2017). A consensus score was calculated from these individual scores (Wang and Wang, 2001) and compounds were then ranked by the consensus score. Where applied, an optional pharmacophore filter was used to discard compounds lacking functional groups considered necessary for binding. Finally, the predicted binding modes of the top-ranked compounds were inspected visually in PyMOL, and compounds with plausible binding modes were selected for purchase from BioAscent and subsequent experimental testing (Section 8.3.1).

7.21. Optimization of the high throughput small-molecule screening (HTS) with the ssDNA substrate

HTS, with ssDNA substrate, was optimized to test experimentally compounds selected *in silico* from the BioAscent database for inhibition of the 5'-exonuclease activity of EXOG. The assay utilized a ssDNA substrate with Cy3 dye at 3'-end (S3; Table 2). The cleavage of the substrate by EXOG resulted in a release of a fluorophore from DNA molecule. Cy3 is highly sensitive to its immediate environment, when tethered to a DNA, its rotational freedom is reduced, which prevents energy-wasting photoisomerization and increases brightness (Spiriti et al., 2011). Cleavage removes this stabilizing environment, returning the dye to a more flexible state that emits less light. In this setup cleavage by EXOG results in a decrease of the DNA fluorescence signal. Measurements were performed using Spark® multimode microplate reader (Tecan) in 384-well OptiPlate™ black plates (Greiner, cat. no. M7935) at 25 °C with the excitation at 528 nm and emission at 573 nm in HK140 BF. Screening was performed in HK140 Buffer. All reactions had a final volume of 50 μl and contained 250 nM substrate. To assess the

EXOG concentration giving a suitable decrease of the fluorescence signal for inhibitor testing, a protein titration was performed. Control measurements included substrate alone, EXOG alone and substrate with EXOG.

Experiment was divided into three steps:

1. measurement of DNA signal (300 s),
2. addition of EXOG (25-200 nM; 300 s),
3. initiation of the reaction by adding $MgCl_2$ to a final concentration of 20 mM (1800 s).

For the inhibition experiments the assay comprised of four steps:

1. measurement of the substrate signal (300 s),
2. addition of ATA (0.01-100 μM final concentrations, 300 s),
3. addition of EXOG (25 nM final concentration, 300 s),
4. and initiation of the reaction by addition of $MgCl_2$ (20 mM final concentration (1800 s)).

The experiment was done in three independent experiments.

7.22. Biochemical testing of virtual-screening derived candidate EXOG inhibitors

Candidate EXOG inhibitors selected by virtual screening (Chapter 7.20.) were tested experimentally in a fluorescence-based EXOG activity assay. In total, 286 compounds were obtained from BioAscent as a 2 mM stocks in DMSO, distributed one compound per well, on 96-well plates. Screening was performed using EnVision® Multilabel Reader (PerkinElmer) in 384-well OptiPlate™ black plates (Greiner) with the excitation at 528 nm and emission at 573 nm. Screening was performed in HK140 Buffer at 25 °C and final reaction volume was 50 μl. Control measurements included substrate alone, EXOG alone, substrate with inhibitor, and substrate with EXOG (no inhibitor). The experiment was divided into four steps:

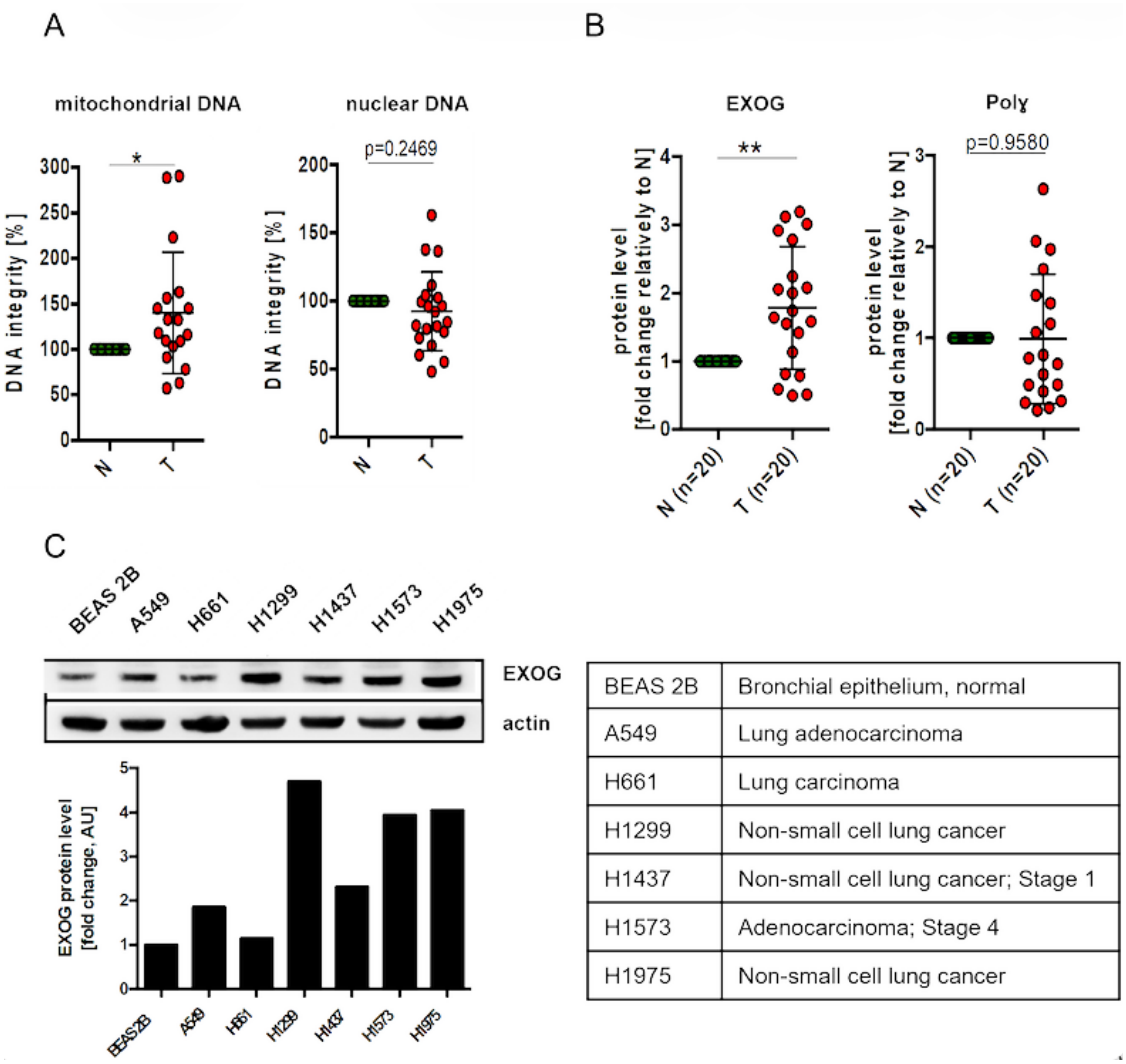
1. recording of substrate signal (250 nM; 300 s),
2. addition of the tested compound (100 μM, 300 s),
3. addition of EXOG (25 nM final, 300 s),
4. and initiation of reaction by adding $MgCl_2$ to 20 mM final (1800 s).

EXOG activity was calculated relative to the no-inhibitor control (substrate + EXOG + Mg^{2+}, defined as 0% inhibition) as: % inhibition = 100 × [1 − (Fluorescence_sample − Fluorescence_background) / (Fluorescence_no-inhibitor − Fluorescence_background)]. Data shown are the mean of three independent experiments.

8. RESULTS AND DISCUSSION

8.1. Increased mtDNA integrity coincides with EXOG overexpression in lung adenocarcinoma

mtDNA has been described as essential for cancer progression (Tan et al., 2015). The differences in integrity level between mitochondrial and nuclear DNA were first compared. For this purpose, a long amplicon PCR assay (LA-qPCR) was used on the DNA isolated from 20 pairs of lung adenocarcinoma tumors and normal adjacent lung tissue. The level of mtDNA integrity was significantly elevated in tumors when compared to the healthy tissue. The same phenomenon was not observed for nuclear DNA, where no significant increase in integrity was detected (Fig. 8 A). As EXOG was shown to be essential for mtDNA repair and lack of its activity leads to apoptosis, its level of expression was next examined in the same tumor and healthy tissue. The results showed almost 2 times higher level of EXOG in human lung adenocarcinoma, when compared to the healthy adjacent tissue (Fig. 8 B). As a control, the expression level of another enzyme involved in mtDNA BER, Pol γ, was also tested. Here, no significant differences were noticed between the two types of tissue (Fig. 8 B). It was additionally tested whether the higher EXOG level is strictly correlated with lung adenocarcinoma or is it a common occurrence in other types of lung cancer. To this end, one cell line derived from healthy bronchial epithelium and six different cancer derived cell lines were tested: A549 (lung adenocarcinoma), H661 (lung carcinoma), H1299 (non-small cell lung cancer), H1437 (non-small cell lung cancer, stage 1), H1573 (adenocarcinoma, stage 4) and H1975 (non-small cell lung cancer). EXOG overexpression was observed in all the tested cancer derived cell lines. The largest change was observed in H1299 cell line where EXOG level was 5 times higher than in normal cells. In H1573 and H1975 this difference was 4-fold. The smallest difference was observed in H661 cell line (Fig. 8 C). All these findings show that EXOG level is elevated both in tissue collected from patients, as well as, in the cell culture experiments and that this elevation co-occurs with the higher mtDNA integrity seen in the same tumors. These data are correlative, so on their own they do not prove that EXOG is responsible for the increased mtDNA integrity. However, EXOG has been previously shown to be essential for mtDNA repair, as its depletion causes persistent single-strand breaks and apoptosis (Tann et al., 2011).



BEAS 2B	Bronchial epithelium, normal
A549	Lung adenocarcinoma
H661	Lung carcinoma
H1299	Non-small cell lung cancer
H1437	Non-small cell lung cancer; Stage 1
H1573	Adenocarcinoma; Stage 4
H1975	Non-small cell lung cancer

Figure 8. High EXOG level is linked with increased mtDNA integrity level in lung adenocarcinoma tumors and several lung cancer cell lines. *(A) Comparison of the mtDNA and nuclear DNA integrity in human lung adenocarcinoma tumors (T, n = 20) and normal adjacent lung tissue (N, set as 100%; n = 20). (B) Protein levels of two enzymes involved in mtDNA BER process – EXOG and Pol γ in human lung adenocarcinoma tumors (T, n = 20) and normal adjacent lung tissue (N, set as 100%; n = 20). (C) EXOG protein level in cancer derived cell lines shown as a fold change in comparison to healthy cells based on the Western Blot analysis.*

Given this established role, the overexpression of EXOG in the same cancers that show elevated mtDNA integrity is consistent with EXOG contributing to this phenotype and raises the possibility that its inhibition could affect cancer cell survival. Because EXOG is localized to mitochondria, a selective EXOG inhibitor could, in principle, modulate mtDNA repair without directly inhibiting nuclear DNA-repair enzymes. All the experiments were performed at The University of Texas Medical Branch (UTMB) in Galveston in collaboration with Bartosz Szczesny, PhD.

8.2. Identification of hEXOG inhibitors from the Library of Pharmacologically Active Compounds (LOPAC® 1280)

The HTS was carried out at The University of Texas Medical Branch in Galveston in collaboration with Bartosz Szczesny, PhD. To perform an efficient screening of the LOPAC® 1280 library, a dsDNA substrate labelled with fluorescein at the 3' end and the Black Hole Quencher 1 (BHQ-1) at the 5' end was designed. In the intact substrate, BHQ-1 quenches the fluorescein signal. Excision of the 5′ dinucleotide by EXOG separates the quencher from the fluorophore and results in an increase in fluorescence, which was assessed in the time-course experiment (Fig. 9 B). The fluorescence signal was measured at five time points to check whether it increased linearly. As controls, substrate (blue) and protein (red) alone were measured. A significant increase in fluorescence was observed only when all the components (dsDNA substrate, EXOG, Mg^{2+}) were present (green). Once the set-up had been established to be suitable for HTS and to give a stable, proportional signal, the screening was performed. In total, 1280 compounds from the LOPAC® library were tested at a fixed concentration of 12 μM. A threshold of 50% inhibition of EXOG was set for a compound to be considered a candidate inhibitor. Six compounds met this criterion (Fig. 9 C). All of them were compared based on the efficiency of inhibition, the compound structure and Lipinski's rule of five (Table 5). The two most potent inhibitors, aurintricarboxylic acid (ATA) and mitoxantrone (MTX), were selected for further study; additionally, both compounds satisfied Lipinski's rule of five.

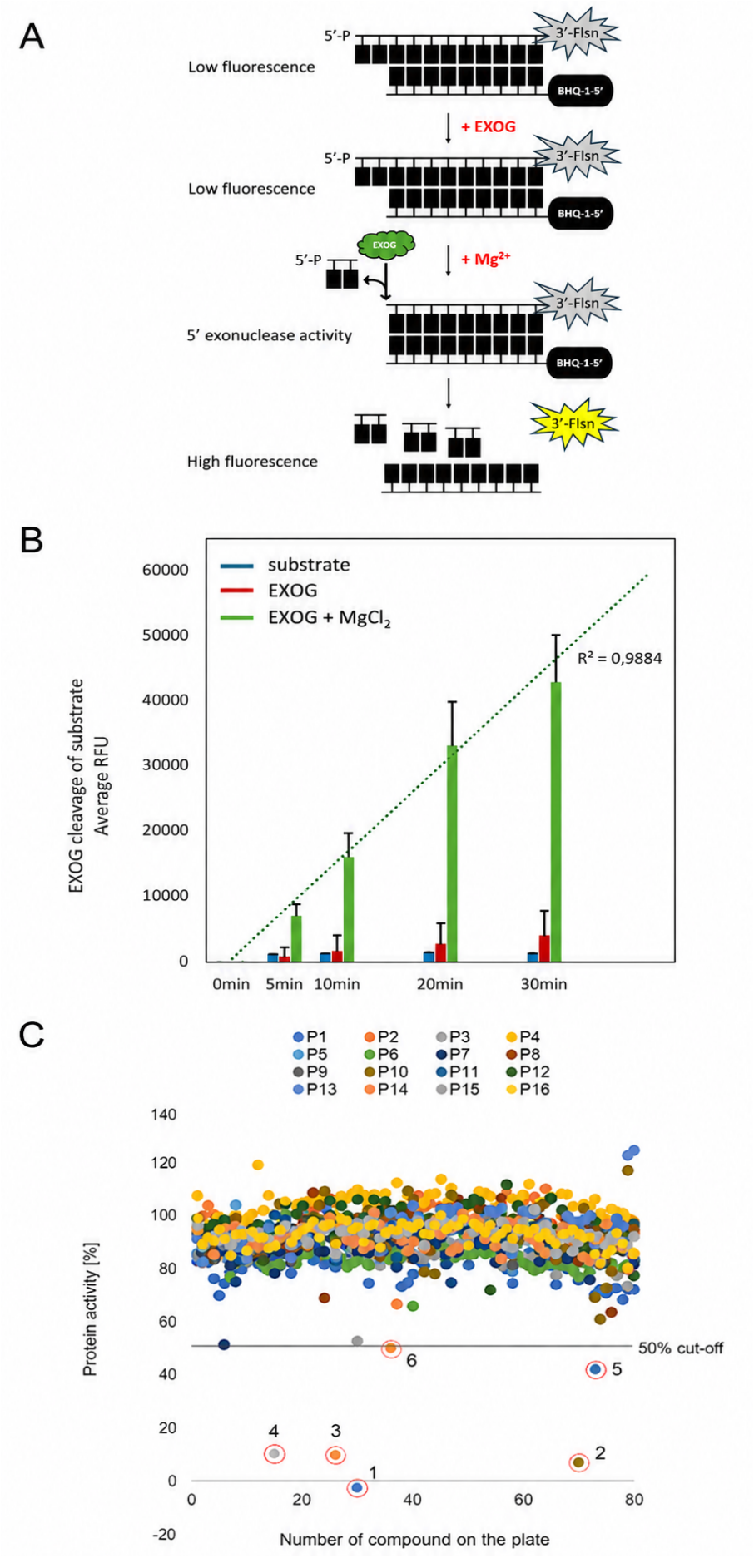
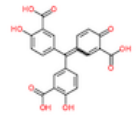
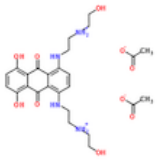
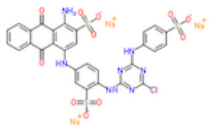
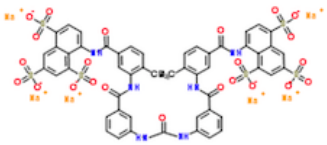
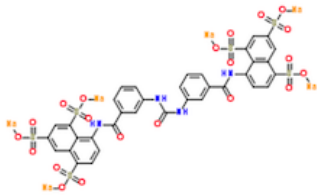

A
5'-P
3'-FIsn
Low fluorescence
BHQ-1-5'
+ EXOG
5'-P
3'-FIsn
Low fluorescence
BHQ-1-5'
EXOG
5'-P
+ Mg2+
3'-FIsn
5' exonuclease activity
BHQ-1-5'
3'-FIsn
High fluorescence
B
substrate
EXOG
EXOG + MgCl2
R² = 0,9884
EXOG cleavage of substrate
Average RFU
60000
50000
40000
30000
20000
10000
0
0min
5min
10min
20min
30min
C
P1 P2 P3 P4
P5 P6 P7 P8
P9 P10 P11 P12
P13 P14 P15 P16
Protein activity [%]
140
120
100
80
60
40
20
0
-20
50% cut-off
1
2
3
4
5
6
0
20
40
60
80
Number of compound on the plate

Figure 9. High-throughput screening of LOPAC® 1280 compound library. *(A) dsDNA substrate containing fluorescein at the 3' end and the Black Hole Quencher 1 (BHQ-1) at the 5'-end was utilized in HTS. Cleavage of dinucleotides by EXOG results in the increase of fluorescence signal in the presence of cofactor (Mg^{2+}). (B) Basal EXOG activity on the fluorescein/BHQ-1 substrate. (C) HTS of LOPAC® 1280 compound library results. Assay was divided into 16 plates marked as P1-P16 (legend). 80 compounds were tested per plate (marked in x-axis). Compounds inhibiting EXOG by at least 50% were marked with red circle and number based on their efficacy.*

Table 5. List of compounds selected based on HTS of LOPAC® 1280 compound library.

LOPAC (Plate, well)	Name	MW [Da]	% inhibition at 12µM	Structure
P01-F5	Aurintricarboxylic acid Compound **1**	422.35	100	
P10-F10	Mitoxantrone hydrochloride Compound **2**	517.41	94	
P14-B5	Reactive Blue 2 Compound **3**	840.11	90	
P15-G3	Suramin sodium salt Compound **4**	1429.18	90	
P11-A11	NF 023 Compound **5**	1162.88	58	
P14-D6	Sodium nitroprusside dehydrate Compound **6**	261.92	50	

ATA (Fig. 10 A) has been described as polyaromatic carboxylic acid derivative (Hunt and Wagner, 1975). It has been shown to inhibit proteins involved in nucleic acid processing, such as APE1 (Simeonov et al., 2009), FEN1 (Dorjsuren et al., 2011), DNase I, RNase A or Taq polymerase (Ghosh et al., 2009). Nuclear magnetic resonance spectroscopy (NMR) studies on bovine pancreatic ribonuclease A performed in 1980 by Gonzalez and colleagues, suggest that ATA competes with the DNA for binding in the active site of the protein and may interfere with electrostatic interactions which contribute to protein-nucleic acid binding (Gonzalez et al., 1980).

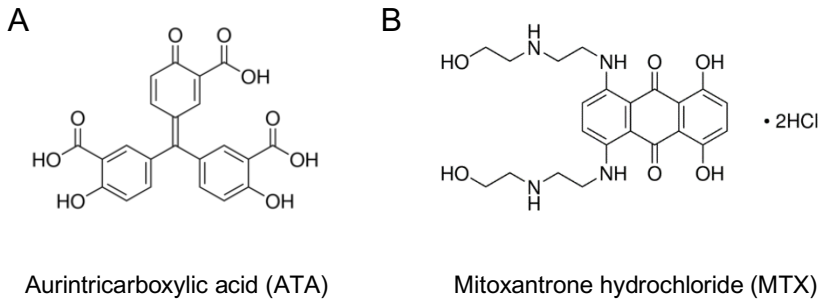


Figure 10. Compounds selected based on the high-throughput screening assay*. (A) Aurintricarboxylic acid (ATA), molecular weight = 422.3 g/mol, compound ID in PubChem: 2259. (B) Mitoxantrone hydrochloride (MTX), molecular weight = 517.41 g/mol, compound ID in PubChem: 51082.*

MTX (Fig. 10 B) is widely used as a potent antitumor drug in the treatment of various cancers such as breast cancer, lymphoma and leukemia (Shenkenberg and Von Hoff, 1986). MTX is an anthracenedione, similar in its structure to anthracyclines but lacking their amino sugar moiety. Compared with anthracyclines it exhibits lower risk of cardiotoxic effects (Shenkenberg and Von Hoff, 1986). Studies have shown that nuclear DNA is the major target for this drug as it can intercalate between DNA base pairs and bind to negatively charged phosphate groups of DNA, thus causing DNA condensation (Li et al., 2005). In addition to intercalating into DNA, MTX stabilizes the transient topoisomerase II-DNA cleavage complex and block the religation step. This leads to persistent, cytotoxic double-strand breaks (Shenkenberg and Von Hoff, 1986).

8.2.1. Secondary screen of compounds

ATA and MTX were taken to further steps of assessment, to check whether they act on EXOG or on the DNA substrate. Direct interaction with DNA is a common mechanism for some anti-cancer compounds, including MTX (Shenkenberg and Von Hoff, 1986). As HTS setup was based on measuring changes in fluorescence signal, an orthogonal method, with label-free substrate was used. Single-stranded, circular ΦX174 virion DNA was chosen as a substrate suitable for nucleolytic cleavage. First, the basal activity of EXOG on this substrate was assessed to identify the protein concentration at which 100% of substrate conversion would be observed (Fig. 11). On this basis 20 nM EXOG concentration was selected as it was the concentration at which the enzyme fully converted the substrate while remaining below the DNA concentration. EXOG concentration below that of the substrate was chosen so that substrate cleavage was limited by the amount of active enzyme rather than by substrate availability. Under these conditions, inhibition of EXOG translates into a measurable decrease in substrate conversion, whereas an excess of enzyme could mask partial inhibition.

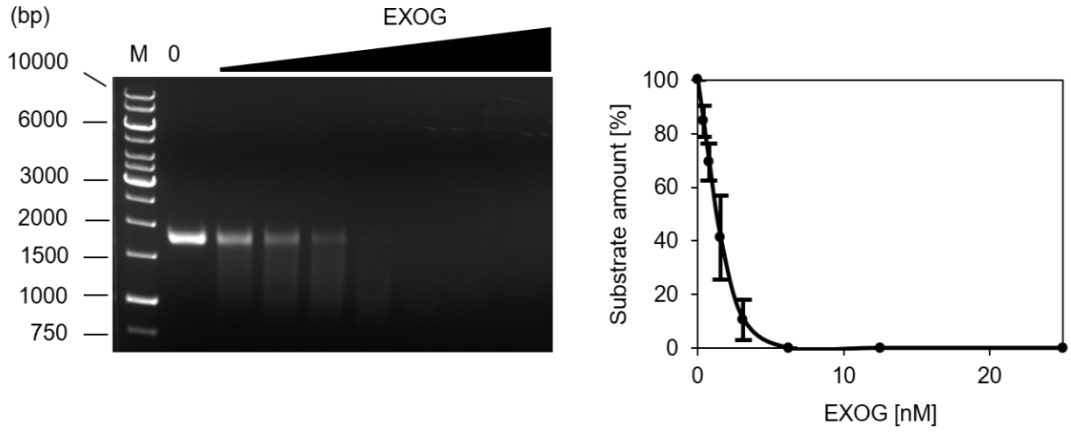


Figure 11. Endonucleolytic cleavage of EXOG on ΦX174 DNA substrate. *30 nM of virion DNA (single-stranded, circular) was incubated with the increasing concentration of EXOG (0.4-25 nM). After 1 h incubation at 25 °C, reactions were stopped and analyzed on 1% agarose gel. Negative control marked as a 0 (left panel). Graph shows the densitometric analysis of the decrease of tested substrate (right panel). Y error bars represent standard deviation (SD).*

The next step was titration of fixed amount of EXOG and substrate (20 nM and 30 nM, respectively) with increasing concentrations of ATA and MTX, to assess IC_{50} (half maximal inhibitory concentration) values (Fig. 12, left panel). For ATA, IC_{50} of0.71 ± 0.03 μM was obtained (Fig. 12, right panel). In the similar range of concentrations as used for ATA (0.06-8 μM), no effect of MTX on the EXOG activity was observed. The MTX concentration was therefore increased up to 200 μM, which resulted in the change of the substrate signal and shift in migration of the DNA. Mitoxantrone interaction with single-stranded substrate most likely accounts for the altered migration, which therefore cannot be interpreted as inhibition of EXOG.

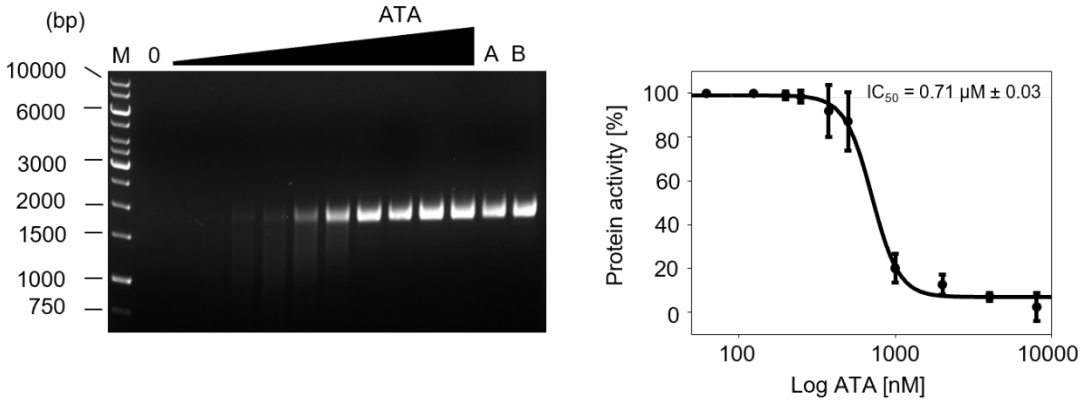


Figure 12. Inhibition of endonucleolytic activity of EXOG by ATA. *ΦX174 DNA was incubated with EXOG and increasing concentrations of ATA (62.5 nM-8 μM). After 1 h of incubation at 25 °C, reactions were stopped and analyzed on a 1% agarose gel. DNA incubated with ATA (8 μM) is marked with A and DNA control is marked with B (left panel). Densitometric analysis of the protein activity inhibition on tested substrate in (right panel). Y error bars represent SD. IC_{50} value is presented as the mean ± SD of three independent experiments.*

To determine whether MTX binds ssDNA, thus altering its staining and migration, and whether ATA shows a similar effect, a fixed amount of DNA was titrated with increasing concentrations of either compound (Fig. 13). ATA did not influence DNA signal or migration (Fig. 13, left panel) but MTX significantly changed bands intensity and at highest concentrations caused a shift in migration. This suggests an interaction with the DNA or interference with its visualization (Fig. 13, right panel).

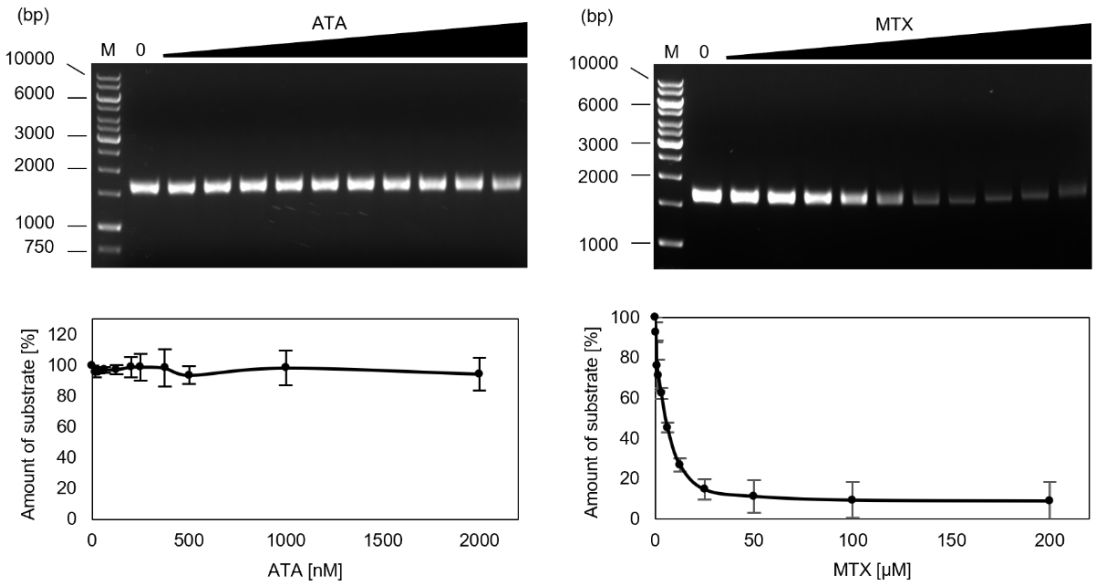


Figure 13. MTX influences DNA staining and migration in agarose gel*. Titration of ΦX174 DNA with increasing concentrations of ATA (15 nM-2 µM) and MTX (0.3-200 µM). Y error bars represent SD.*

This part of the results shows the significance of using orthogonal methods to confirm the inhibition by selected compound. Because the HTS was based on observing a loss of the fluorescence signal as inhibition of enzyme activity, compounds that bind DNA and quench the fluorescence signal could be misclassified as inhibitors. MTX altered the migration and staining of the single-stranded substrate, consistent with an interaction that had previously been described mainly for double-stranded DNA. One possible explanation is the formation of local secondary structure within the single-stranded substrate, providing duplex regions that could accommodate intercalation. MTX was therefore excluded from further analysis, since the aim was to identify an inhibitor specific to EXOG rather than a compound acting on DNA itself.

8.2.2. Inhibition of exonucleolytic activity of EXOG on different RNA/DNA substrates

After validating ATA-mediated inhibition of EXOG activity, its effect on 5'-exonuclease activity was examined. EXOG has been previously shown to efficiently cleave ssDNA, ssRNA and duplex structures (Szymanski et al., 2017, Wu et al., 2019, Karlowicz et al., 2022) all these substrates were therefore included. Additionally, a 15-nt DNA flap substrate was tested, as it mimics the product of strand displacement by Pol γ during long-patch base excision repair pathway.

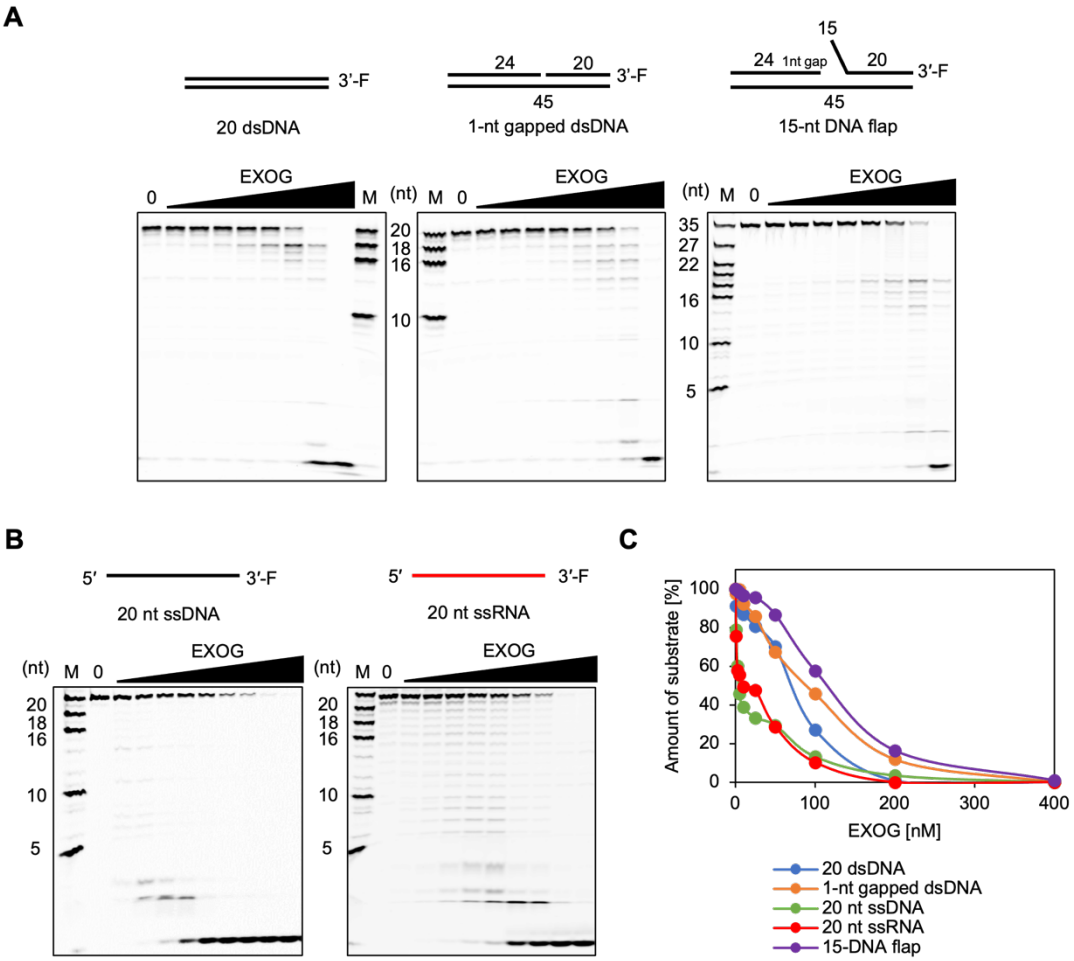


Figure 14. Basal activities of EXOG on different DNA and RNA substrates. *(A) EXOG cleavage of dsDNA, 1-nt gapped dsDNA and 15-nt DNA flap. 200 nM of dsDNA, 1-nt gapped dsDNA or 15-nt DNA flap was incubated with increasing concentrations of EXOG (1-400 nM) (B) EXOG cleavage on ssDNA and ssRNA. 200 nM of ssDNA or ssRNA was incubated with increasing concentrations of EXOG (1-400 nM). For all the reactions after 5 min of incubation at 25 °C, reactions were stopped and analyzed on a 23% denaturing polyacrylamide gel. Corresponding molecular weight standards are marked with M. (C) Densitometric analysis of the decrease of tested substrates in (A) and (B).*

Such an intermediate is further recognized by EXOG for 5'-end processing, generating a ligatable end for Ligase III (Karlowicz et al., 2025). All substrates used in these assays were labelled at 3' end. To ensure that cleavage was measured only from the intended 5'-end, all the unlabeled blunt ends in linear substrates used in this study were blocked with biotin. Basal EXOG activity was first determined on all the tested substrates to choose the enzyme concentration necessary for total substrate cleavage (Fig. 14). For duplex and flap substrates, 200 nM EXOG was required (Fig. 14 A), whereas 100 nM was sufficient for single-stranded substrates (Fig. 14 B), against 200 nM substrate. A fixed concentration of EXOG and substrate was then titrated with increasing concentrations of ATA to assess inhibition (Fig. 15 A and B). Final concentrations and the number of replicates is given in the Methods (Section 7.12.). Inhibition of exonuclease activity was observed for all of the tested substrates and the IC_{50} values were similar ranging from 1.48 ± 0.08 μM for 15-nt DNA flap, 1.58 ± 0.04 μM for ssDNA, 2.58 ± 0.04 μM for 1-nt gapped dsDNA, 3.45 ± 0.06 μM for ssRNA and the highest value 3.81 ± 0.1 μM for dsDNA (Fig. 15 C).

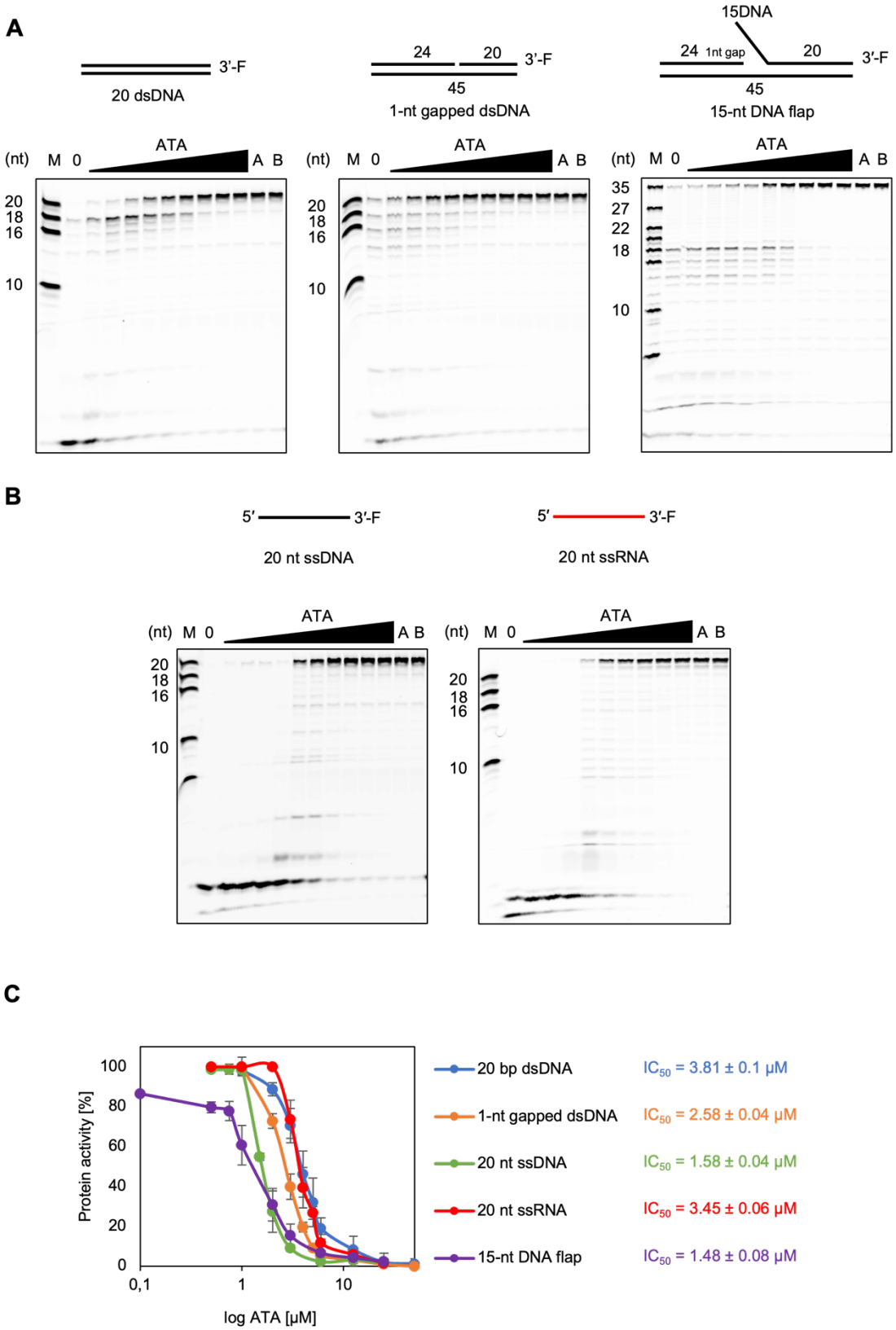

A
3'-F
20 dsDNA
24
20
3'-F
45
1-nt gapped dsDNA
15DNA
24
1nt gap
20
3'-F
45
15-nt DNA flap
(nt) M 0
ATA
A B
20
18
16
10
35
27
22
18
10
B
5'
3'-F
20 nt ssDNA
20 nt ssRNA
C
Protein activity [%]
100
80
60
40
20
0
0,1
1
10
log ATA [μM]
20 bp dsDNA
IC50 = 3.81 ± 0.1 μM
1-nt gapped dsDNA
IC50 = 2.58 ± 0.04 μM
20 nt ssDNA
IC50 = 1.58 ± 0.04 μM
20 nt ssRNA
IC50 = 3.45 ± 0.06 μM
15-nt DNA flap
IC50 = 1.48 ± 0.08 μM

Figure 15. EXOG cleavage of different RNA/DNA substrates is inhibited by increasing concentrations of ATA. *(A) 200 nM of dsDNA (left panel), 1-nt gapped dsDNA (middle panel) or 15-nt DNA flap substrate (right panel) was incubated with 200 nM of EXOG and increasing concentrations of ATA. (B) 200 nM of ssDNA (left panel) or ssRNA (right panel) was incubated with 100 nM of EXOG and increasing concentrations of ATA. After 5 min of incubation at 25 °C, reactions were stopped and analyzed on a 23% denaturing polyacrylamide gel. Corresponding molecular weight standards are marked with M. DNA control is marked with A and DNA incubated with ATA is marked with B. (C) Densitometric analysis of the substrate depletion on the tested substrate in (A) and (B) together with IC_{50} calculations. Y error bars represent SD. IC_{50} value is presented as the mean ± SD of three independent experiments.*

Overall, the IC_{50} values were within the same micromolar range. This lack of substrate dependence is informative on its own. As the IC_{50} values are similar regardless of the substrate, ATA most likely acts at a step common to all these reactions, that is on the enzyme rather than on a particular nucleic acid structure. This already points towards the protein, and not the DNA, as the target of ATA, which is addressed directly in the following sections. Inhibition of EXOG activity on the 15-nt DNA flap provided a rationale for testing whether ATA affects mtDNA integrity in cells.

8.2.3. Co-crystallization and soaking of EXOG with ATA

To identify the ATA-EXOG binding site co-crystallization of EXOG with the compound was attempted. A range of commercial sparse-matrix screens (n=9) was set up using the sitting-drop method (details in Methods 7.6.1). Besides already published crystallization conditions 2.0 M $(NH_4)_2SO_4$, 0.1 M $MgCl_2$ and 0.1 M Tris-HCl pH 8.5, another crystallization condition was found, which was 0.1 M $(NH_4)_2SO_4$, 0.1 M NaCl, 1.0 M HCOONa, 0.05 M Bis-Tris pH 7.0, 20 % w/v PEG 8000. Unfortunately, crystals grown in the second crystallization condition decreased markedly in size upon ATA addition and diffracted more poorly than those from ammonium sulphate condition. Co-crystallization was therefore carried out in the condition with ammonium sulphate, where ATA was used in a range of concentrations from 0.1-1 mM. The largest, best diffracting crystals grew in the middle range of ATA concentration, which was between 0.25-0.5 mM (Fig. 16, right panel). Crystals grown in 2.0 M $(NH_4)_2SO_4$, 0.1 M $MgCl_2$ and 0.1 M Tris-HCl pH 8.5 were also soaked in mother liquor containing ATA for 1-24 h before freezing (Fig. 16, left panel). Approximately 100 datasets with the resolution between 1.77-2.29 Å for both co-crystallization and soaking experiments were

collected. Attempts to place ATA into the difference density with LigandFit yielded a best real-space correlation coefficient of 0.42, below the 0.7 threshold, and the placement was flagged by the program as unlikely to be correct. No interpretable electron density corresponding to the ligand was therefore observed under the conditions applied. This may reflect low or heterogeneous ligand occupancy, ligand disorder, competition with components of the crystallization solution, or crystal-packing effects. The high ammonium sulphate concentration used may also have interfered with ATA binding both for the co-crystallization and soaking experiments. As a result, a direct structural view of the ATA-binding site was not obtained and the interface proposed in this work therefore relies on docking on published structure (PDB: 5T40) (done in collaboration with Douglas Houston from University of Edinburgh) and site-directed mutagenesis rather than on an experimental EXOG-ATA structure.

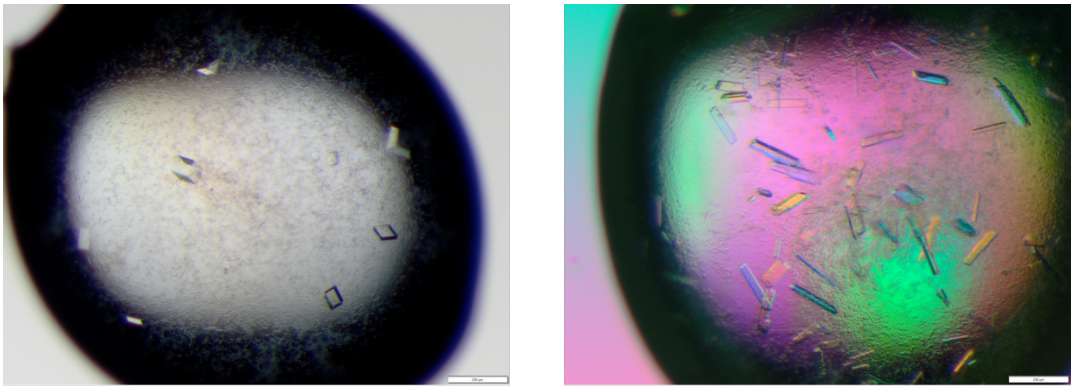


Figure 16. EXOG crystals grown in 2.0 M $(NH_4)_2SO_4$, 0.1 M $MgCl_2$ and 0.1 M Tris-HCl pH 8.5 condition. *Crystals of apo EXOG (left panel) which were used for soaking and as another method to obtain EXOG-inhibitor structure, EXOG was co-crystallized with ATA (right panel).*

8.2.4. *In silico* evaluation of ATA-EXOG binding

To computationally predict possible ATA-binding orientations, *in silico* docking was performed with GWOVina - Fast Protein-Ligand Docking Tool. GWOVina is a variant of Vina that uses the grey wolf optimizer (GWO), random walk for the global search and the Dunbrack rotamer library for side-chain sampling (Wong et al., 2021). Multiple predicted poses of ATA in complex with EXOG were generated, using published structure (PDB: 5T40).

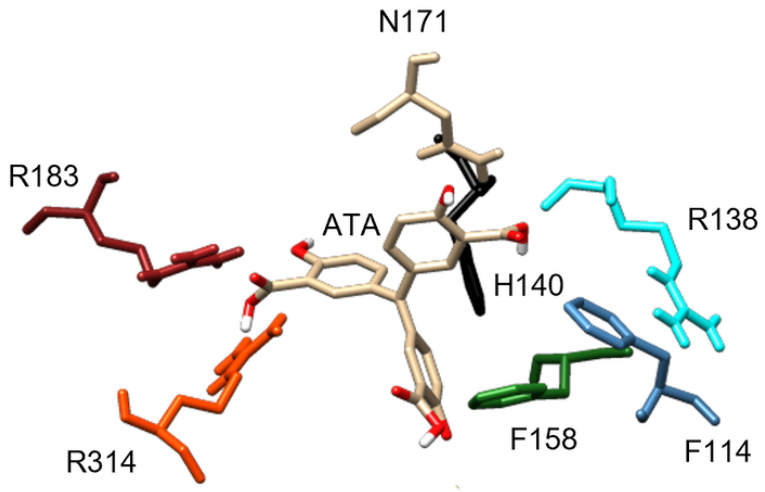


Figure 17. Docking of ATA to EXOG active site using GWOVina resulted in seven possible residues involved in binding: F114 (blue), R138 (cyan), H140 (black), F158 (dark green), R183 (dark red), N171 (beige), R314 (orange).

This analysis identified seven residues near the active site as candidates that may be involved in ATA binding (F114, R138, H140, F158, N171, R183, R314; Fig. 17). These residues were identified by docking and do not represent contacts observed in an experimentally determined complex. Because the co-crystallization and soaking attempts did not yield ATA density (Section 8.2.3), the docking was performed on the apo structure, so the proposed binding site is inferred rather than directly observed. For this reason, the seven residues were treated as candidates to be tested experimentally, rather than as a confirmed interface. Of these, histidine at position 140 lies in the active site and was not pursued for mutagenesis, as its substitution would abolish catalysis and prevent an activity-based readout of ATA binding. Asparagine at position 171, which contributes to coordination of the catalytic magnesium ion was selected for further analysis, however because this residue participates in catalysis, any change in ATA inhibition of the corresponding variant may reflect an indirect effect on enzyme activity, rather than direct disruption of an ATA-binding contact. This limitation was considered when interpreting the mutagenesis data.

8.2.5. Site-directed mutagenesis and testing of changes in inhibition

To probe which of the residues indicated by docking contribute to ATA-mediated inhibition, alanine screening mutagenesis was performed at the six positions selected after excluding the catalytic histidine H140: F114, R138, F158, N171, R183 and R314. All protein variants were purified with size exclusion chromatography (SEC) as the final step (Fig. 18A). No change in the ratio of monomer to dimer in the purified mutants or any significant aggregation were observed, indicating that the substitutions did not alter the oligomeric state of the protein. SEC does not, however, exclude more local effects on folding, stability, substrate binding or catalysis. Final protein samples were resolved on the SDS-PAGE gel to assess the purity. All the obtained protein samples were 95% pure (Fig. 18 B).

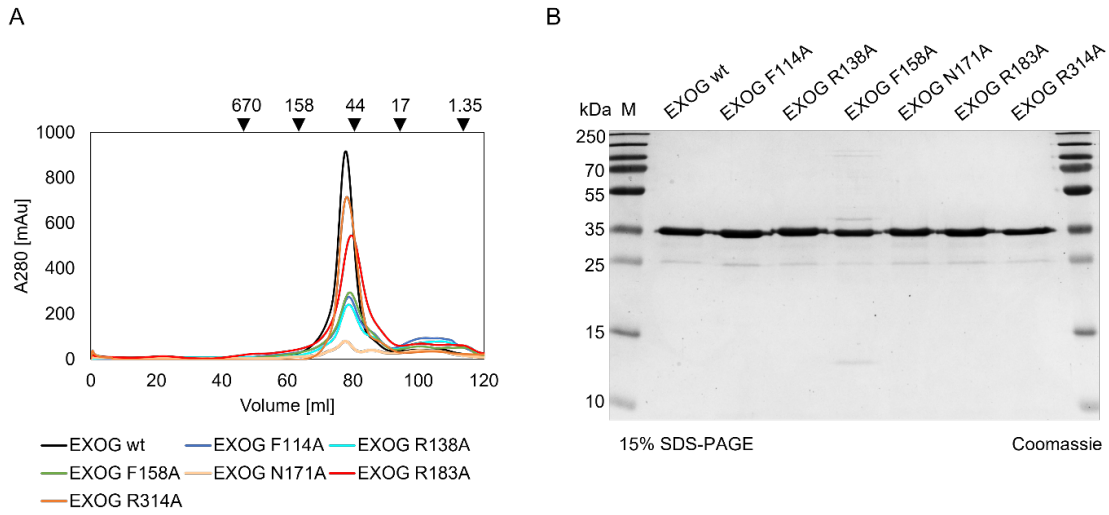


Figure 18. Purification of EXOG wild-type and its variants. *(A) SEC profiles of EXOG and its point mutants F114A, R138A, F158A, N171A, R183A and R314A. (B) 1 µg of EXOG protein and its variants resolved on the SDS-PAGE gel. Protein preparations were of 95% purity.*

Then, a substrate was sought on which all variants would exhibit comparable basal activity at a single enzyme concentration. This is necessary because inhibition is measured as the decrease in activity relative to the uninhibited reaction. If variants differ in their intrinsic cleavage efficiency, differences in apparent inhibition could reflect altered catalytic performance rather than altered ATA binding. Of the three substrates tested (dsDNA, ssDNA and ssRNA), ssRNA gave the most uniform cleavage across all proteins and was therefore used for the inhibition measurements (Fig. 19).

Comparable basal activity on this substrate does not, however, exclude that individual substitutions affect substrate binding or protein stability more generally. Next, fixed concentrations of EXOG, either wild-type or variants, and substrate were titrated with increasing concentrations of ATA, to determine whether the substitutions altered apparent ATA sensitivity (Fig. 20). Four out of six variants, F114A, R138A, F158A and R183A, were more effectively inhibited by ATA than the wild-type protein with IC_{50} values of 1.8 ± 0.1, 0.8 ± 0.1, 1.6 ± 0.2 and 1.3 ± 0.2 µM, respectively, compared with 3.4 ± 0.2 µM for the wild-type. The N171A variant of EXOG lost its activity almost totally after introduction of the lowest ATA concentration (0.5 µM), so that an IC_{50} could not be determined. Only variant R314A was inhibited with the efficiency comparable to the wild-type protein (IC_{50} 3.7 ± 0.1 µM). These results must be interpreted with caution, as the IC_{50} in this assay is measured at a single, fixed substrate concentration and therefore reflects not only ATA binding, but also how well a given variant binds the substrate.

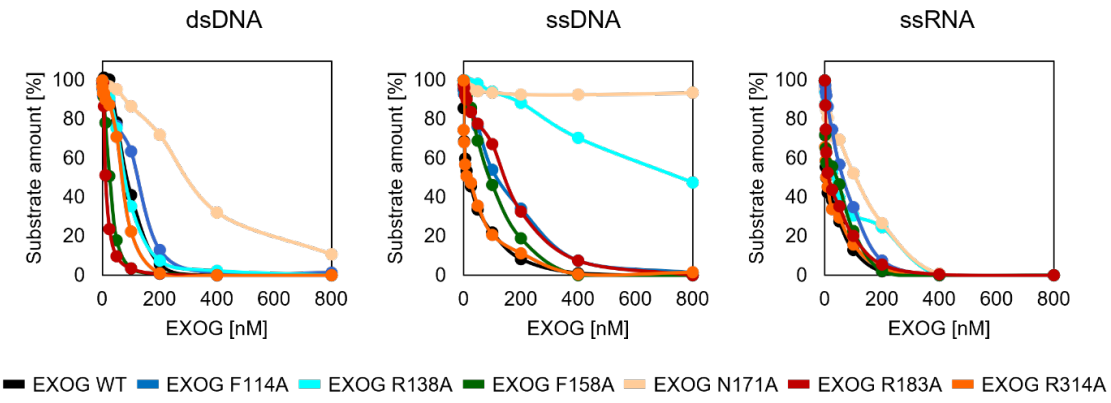


Figure 19. Basal cleavage of EXOG wild-type and its variants on different substrates: dsDNA (left panel), ssDNA (middle panel) and ssRNA (right panel).

ATA is known to interfere with protein-nucleic acid interactions (Gonzalez et al., 1980) and its preincubation with EXOG decreases protein affinity to DNA (Section 8.2.8), consistent with ATA competing with the substrate at or near the active site. A change that weakens substrate binding would lower the IC_{50} even if ATA binding itself is unchanged. Alternatively, replacing a bulky side chain with alanine may increase the accessibility or hydrophobic character of the pocket, thereby promote compound binding and lower the apparent IC_{50} (a plausible contribution for F114A and F158A in particular). In addition, the basal activity of the variants was assessed under conditions of complete

substrate conversion, which may mask differences in catalytic efficiency. For N171A, which contributes to coordination of the catalytic magnesium ion, the pronounced sensitivity to ATA may therefore reflect a reduced catalytic margin rather than enhanced ATA binding. The thermal stability of the variants was not assessed, which further limits interpretation.

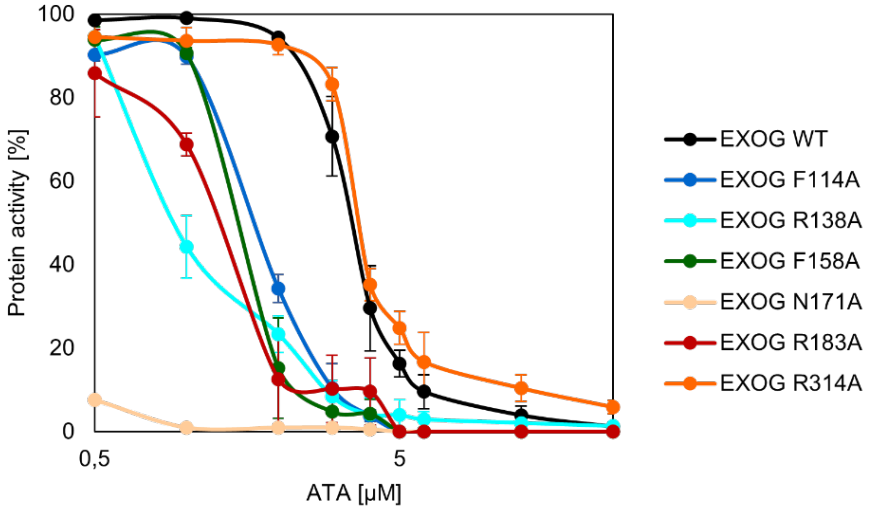


Figure 20. Concentration-dependent inhibition of EXOG wild-type and alanine variants by ATA on an ssRNA substrate. *Y error bars represent SD.*

This is important in the case of R138 and R183, as they are positively charged residues that line the V-shaped active site cleft and contribute to substrate binding (Section 4.3.1). What is more, R138 is seen interacting with a sulfate ion in the apo structure (Fig. 6 B). Their stronger inhibition might therefore result from impaired substrate binding rather than tighter ATA binding. The two remaining variants, N171A and R314A, point in the same direction. The almost total loss of N171A activity upon ATA addition is more easily explained by the catalytic role of N171, which coordinates the active site Mg^{2+} ion (Section 4.3.1), than by its direct involvement in ATA binding. R314A, although indicated by docking, behaved like the wild-type protein and thus does not support the involvement of R314.These data show that substitutions at F114, R138, F158, R183 and N171 alter the apparent sensitivity of EXOG to ATA. Because the assay does not separate effects on ATA binding from effects on substrate binding or catalysis, the results do not establish these residues as direct ATA contacts. To discriminate between these possibilities, a direct measurement of ATA binding to each variant would be needed, for example using biolayer interferometry or protein thermal shift, as performed here for the

wild-type protein (Section 8.2.7), which would show whether the changes in IC_{50} are accompanied by changes in ATA binding.

8.2.6. Carboxyl groups of ATA contribute to inhibition of EXOG activity

Following the mutagenesis screen, the effect of modifying the inhibitor itself was examined (Fig. 21). Three compounds related to ATA (green) in their structure were selected: pararosaniline chloride (red), p-rosolic acid (orange) and 5,5'-methylenebis (2-hydroxybenzoic acid) (purple). These analogues differ from ATA not only in the substituents on the aromatic ring but also in their charge and physicochemical properties (Fig. 21 A). P-rosolic acid shares the triarylmethane core but carries hydroxyl instead of carboxyl groups, pararosaniline chloride is a cationic dye bearing amino groups, while 5,5'-methylenebis(2-hydroxybenzoic acid) has two rather than three carboxyl groups and a methylene bridge in place of the conjugated central carbon. The ATA ammonium salt (blue) was also tested, as it is more soluble in aqueous solvents. A fixed concentration of ssDNA and EXOG was titrated with increasing concentrations of ATA and its related compounds (Fig. 21 B). The change in solubility of ATA did not substantially affect the IC_{50} values as it changed from 1.58 ± 0.04 μM for ATA to 1.83 ± 0.05 μM for ATA ammonium salt. Both pararosaniline chloride and p-rosolic acid did not inhibit 5'-exonuclease activity of EXOG under tested conditions while 5,5'-methylenebis (2-hydroxybenzoic acid) showed weak inhibition within the range tested (at concentrations up to 10 μM). The clearest comparison is p-rosolic acid which retains the triarylmethane core but lacks the carboxyl groups. Its loss of activity suggests that the carboxyl groups and possibly their negative charge contribute to inhibition. For pararosaniline chloride this interpretation is confounded by the reversal of charge, and for 5,5′-methylenebis(2-hydroxybenzoic acid) several features differ simultaneously, so the reduced potency cannot be attributed to the loss of a single ring or carboxyl group alone. Together with the observation that ATA competes with DNA for binding (Section 8.2.8), this dependence on carboxyl groups is consistent with ATA acting as a polyanionic mimic of the nucleic-acid backbone – a mechanism that would also be expected to limit its selectivity among nucleic-acid-binding enzymes.

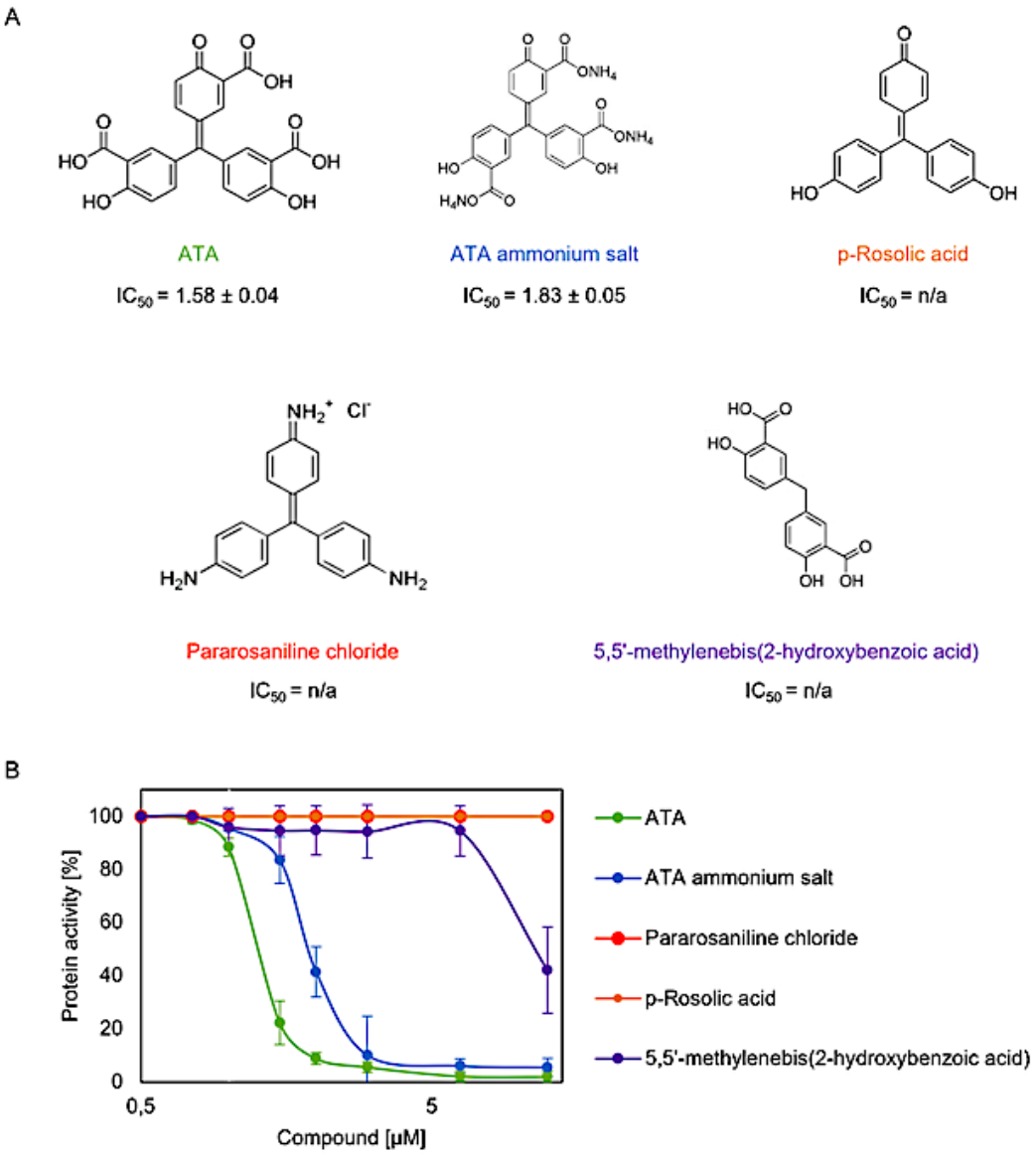


Figure 21. Comparison of exonuclease activity inhibition by ATA and its related compounds. *(A) 2D structures of ATA and its related compounds with calculated IC_{50} values based on protein activity decrease upon titration with given compound shown in (B). IC_{50} values are presented as the mean ± SD of three independent experiments. (B) Plot based on densitometric analysis of substrate decrease in exonuclease activity assay. Y error bars represent SD.*

8.2.7. ATA directly binds and stabilizes EXOG protein

To test the influence of the ATA binding on EXOG stability, nanoDSF was used to measure the intrinsic tryptophan and tyrosine fluorescence of the protein during thermal unfolding in the presence of increasing concentrations of ATA. The results showed a significant change in the inflection temperature of EXOG from 50.6 °C to 67.5 °C in the presence of ATA, suggesting that ATA markedly stabilizes EXOG protein (Fig. 22). Such a shift is consistent with binding, but a change in the thermal unfolding profile does not by itself demonstrate a direct interaction. Direct binding was therefore assessed by biolayer interferometry (BLI).

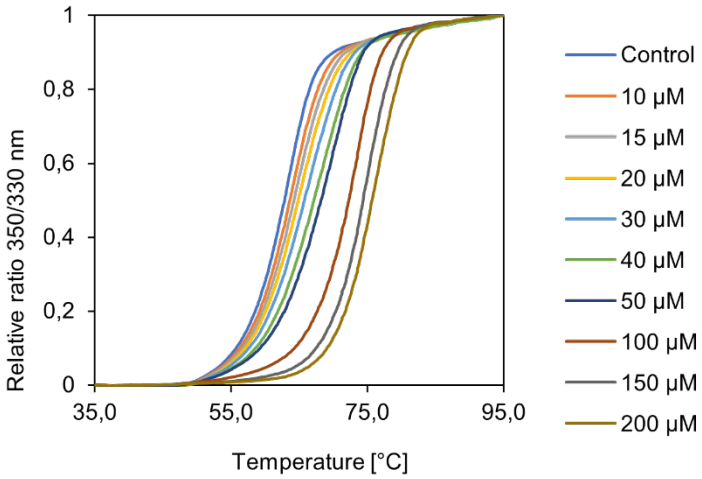


Figure 22. Inflection temperature of EXOG shifts due to presence of increasing concentrations of ATA. *Unfolding profile of EXOG shown as a change in the 350 nm/330 nm ratio signal in temperature gradient. Graph shows an average of three independent datasets.*

BLI allows the interaction between an immobilized ligand and an analyte to be followed in real time (Fig. 23 A). C-terminally His-tagged EXOG protein was immobilized on Octet® NTA biosensor (Fig. 23 B) and titrated with increasing concentrations of ATA to observe real-time ligand binding. The increase in thickness of biolayer allowed to assess apparent dissociation constant (K_d) of 70 ± 9 nM (Fig. 23 C). The obtained K_d of 70 nM is lower than the IC_{50} values measured in the activity assays, which were in the micromolar range (Section 8.2.2).

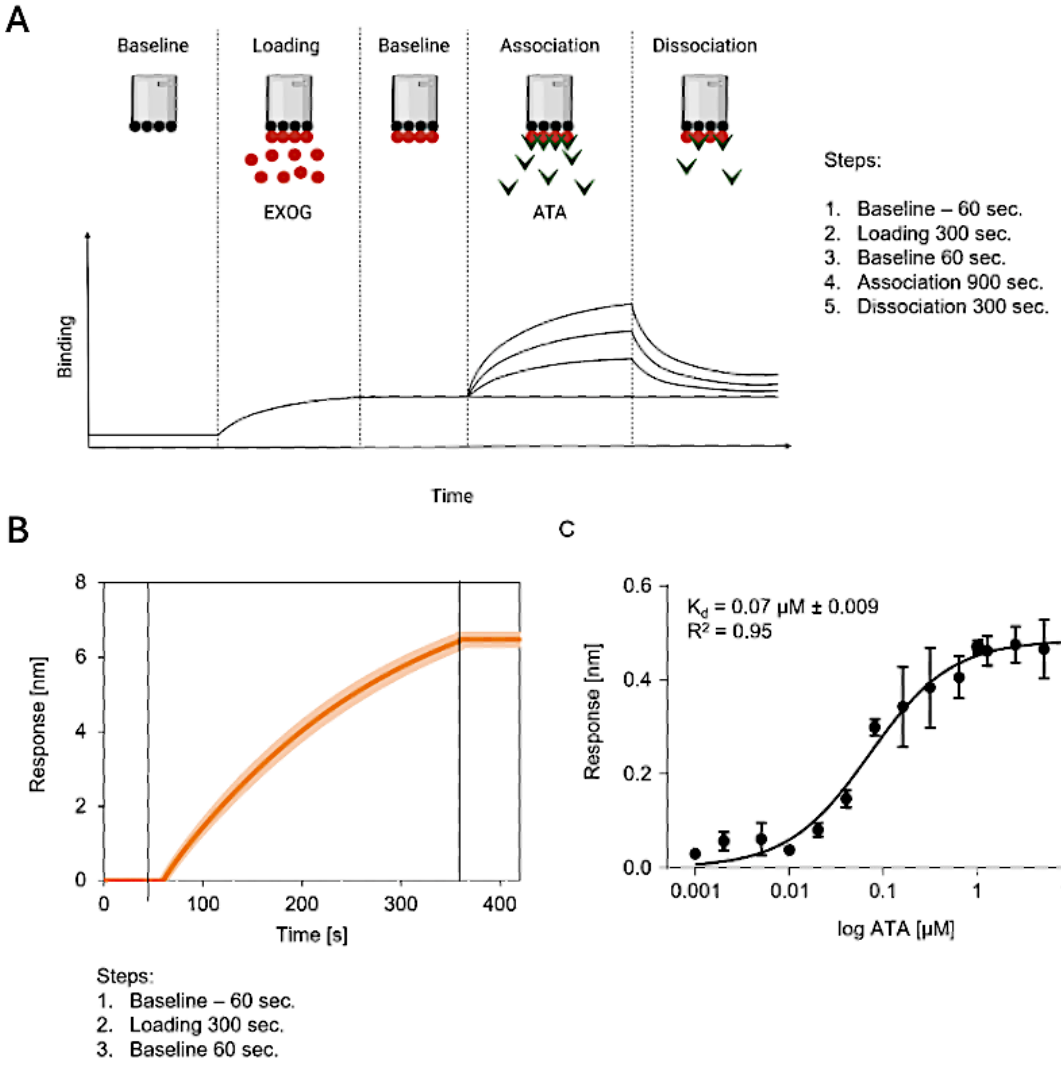


Figure 23. ATA binding to EXOG measured by BLI. *(A) Schematic representation of BLI binding assay. Assay was divided into 5 steps: (1) baseline for sensor washing, (2) loading meaning protein immobilization on the sensor, (3) baseline to wash out unspecifically bound EXOG, (4) association where ATA was titrated and (5) dissociation to wash unbound ATA. (B) C-terminally His-tagged EXOG at concentration of 5 µg/ml was used for binding to NiNTA sensor as indicated in the legend. After step 3 EXOG was titrated with increasing concentrations of used inhibitor. Graph shows an average of three independent experiments. Y error bars (marked in light orange) represent SD. (C) Plot of the increase in thickness of the biolayer upon ATA binding. EXOG immobilized on Octet NTA biosensors was titrated with ATA (1 nM-10 µM). The data was acquired at 25 °C. The raw data was fitted in Prism using one site specific binding model. K_d value is presented as the mean ± SD of three independent experiments. Y error bars on the graph represent SD.*

These values are not directly comparable, because the apparent K_d reflects binding to the immobilized free enzyme, whereas the IC_{50} is a functional value measured in the presence of substrate and Mg^{2+}. However, the strong thermal stabilization and the nanomolar K_d support the conclusion that ATA binds EXOG tightly. Combined with the observation that ATA alone does not alter the DNA substrate (Fig. 13), this supports the interpretation that inhibition results from binding to the protein rather than from an effect on the DNA. The marked stabilization is consistent with strong binding of ATA to EXOG.

8.2.8. Preincubation with ATA decreases EXOG-DNA binding affinity and yields stronger inhibition of exonuclease activity

Both the endonuclease and the 5'-exonuclease activities of EXOG are inhibited by ATA (Section 8.2.1. and 8.2.2.). To establish whether ATA interferes with formation of the EXOG-DNA complex or acts after the substrate is bound, BLI was used. Active protein was used and the buffer was supplemented with EDTA to prevent substrate cleavage. As EXOG binds DNA from the 5'-end (Szymanski et al., 2017), a 3'-end biotinylated 20-nt long ssDNA was immobilized on Octet® SAX Biosensors, leaving the 5'-end accessible, and titrated with increasing concentrations of EXOG to observe real-time protein-DNA binding (Fig. 24 A, left panel). The apparent dissociation constant (K_d) for this interaction was 38.9 ± 5.6 nM. The experiment was then repeated with EXOG pre-incubated for 5 min with 1 μM ATA before addition to the immobilized substrate (Fig. 24 A, right panel). Under these conditions, the apparent half-effective concentration (EC_{50}) increased 4.2-fold to 166.1 ± 5.5 nM. The binding curve was better described by a Hill than by a one-site model. A Hill fit does not however demonstrate a cooperative binding. Here a fixed concentration of ATA (1 μM) was titrated against increasing concentrations of EXOG (8 nM-2 μM), so that the upper end of the titration of the protein is present in molar excess over the inhibitor. Even if every ATA molecule were bound, at least half of the active sites would remain inhibitor-free. Given the apparent K_d of ATA for EXOG (70 ± 9 nM), binding is tight relative to the concentrations used, and free ATA is therefore progressively depleted as the protein concentration increases. The concentration of free inhibitor is thus not constant across the titration but falls as the titrated variable rises. This coupling alone can produce an apparently sigmoidal dependence without any change in the underlying binding mechanism. The data are

consistent with a model in which at low protein concentrations, with the inhibitor no longer saturating, dimers with a single occupied site remain able to engage the immobilized DNA through the free site.

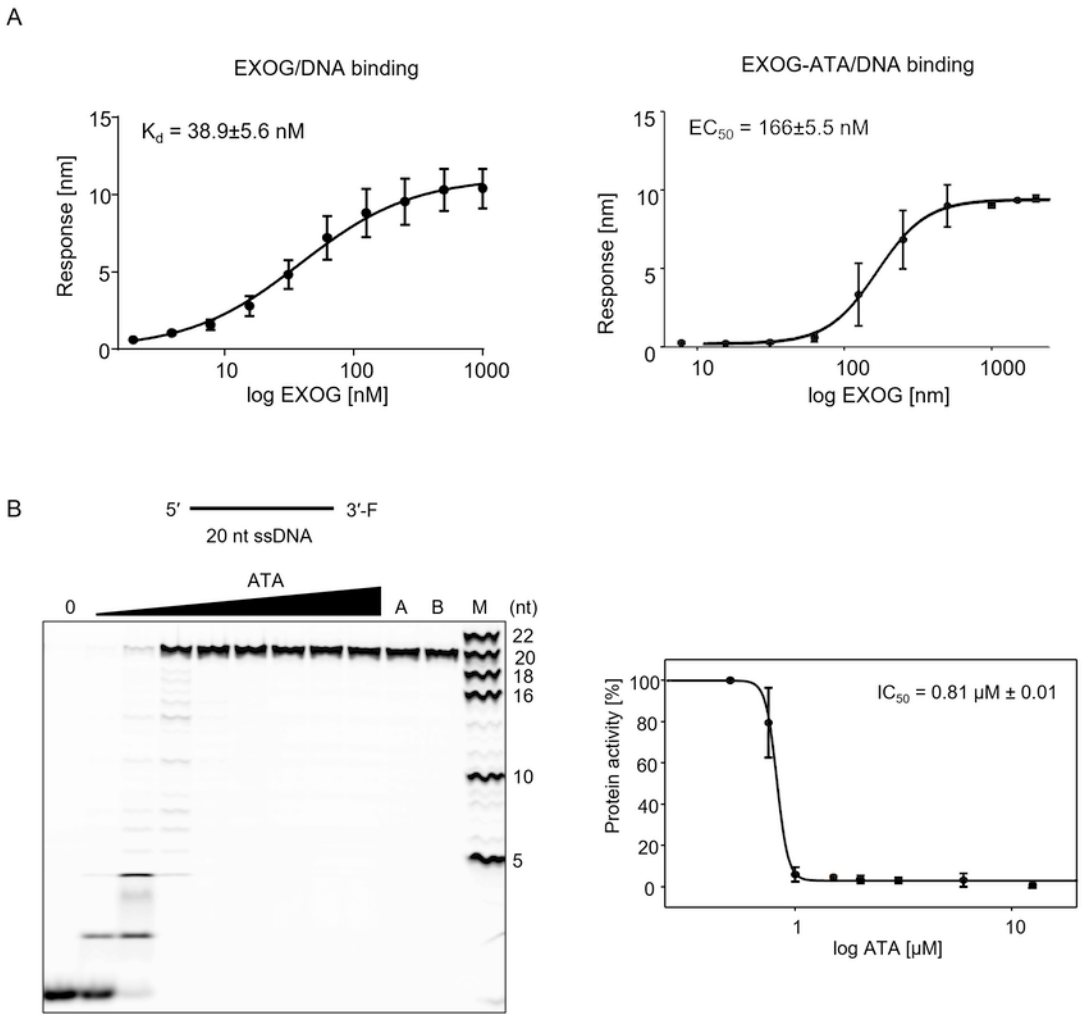


Figure 24. Preincubation of EXOG with ATA increases the efficiency of inhibition and decreases protein affinity to DNA. *(A) ssDNA with 3'-biotin was immobilized on SAX Biosensor and titrated with increasing concentrations of EXOG (left panel). Different EXOG concentrations were preincubated for 5min. with 1 µM ATA upon DNA binding (right panel). The data was acquired at 25 °C. Each point on the graph represents an average of three independent experiments ± SD. Y error bars represent SD. (B) 100 nM EXOG was incubated with different ATA concentrations (0.5-12.5 µM) for 5 min. at 25 °C after which the reaction was started with the addition of 200 nM ssDNA in the presence of magnesium. After 5 min of incubation at 25 °C, reactions were stopped and analyzed on a 23% denaturing polyacrylamide gel. Corresponding molecular weight standards are marked with M. DNA control is marked with A and DNA incubated with ATA is marked with B (left panel). Densitometric analysis of the inhibition of protein activity in the left panel (right panel). Y error bars represent SD. IC_{50} value is presented as the mean ± SD of three independent experiments.*

This model is however itself a description of inhibitor depletion and the present data cannot distinguish it from genuine cooperativity, which is not impossible for EXOG, given the allosteric coupling between its two active sites (Szymanski et al., 2017). Distinguishing the two possibilities would require fitting with an explicit ligand-depletion model or repeating the titration at several ATA concentrations.

Consistent with this, the maximal biolayer response was similar whether EXOG was applied alone or after pre-incubation with ATA, suggesting that the inhibitor shifts the protein concentration required for DNA binding rather than abolishing binding altogether. Taken together, these data support the possibility that ATA interferes with formation of the EXOG-DNA complex, possibly at or near the active site. They do not define the binding site, nor do they exclude other DNA-related effects. The conclusion that inhibition arises from binding to the protein rather than to the DNA rests on the separate observation that ATA alone did not alter the DNA substrate (Fig. 13). Finally, the effect of the order of addition on inhibition efficiency was tested. EXOG was mixed with ATA and allowed to bind for 5 minutes at room temperature after which the reaction was started by addition of ssDNA and Mg^{2+} (Fig. 24 B, left panel). Under these conditions, the apparent IC_{50} dropped from 1.58 ± 0.04 μM to 0.81 ± 0.01 μM (Fig. 24 B, right panel). This result further supports the model that ATA binding prevents substrate binding and its processing by EXOG.

8.2.9. ATA inhibits different proteins involved in DNA processing

ATA has already been reported to inhibit proteins involved in DNA processing such as APE1 (Simeonov et al., 2009) and FEN1 (Dorjsuren et al., 2011). Whether ATA also affects the activity of other proteins of mtDNA repair machinery was therefore tested. For this analysis Polymerase γ, MGME1, RNase H1 and Ligase III were selected, and APE1 and FEN1 were included under the same experimental conditions. The basal activity of each enzyme was first established on several selected substrates typical for their functions (Fig. 25). Then ATA-mediated inhibition was assessed. Surprisingly, ATA inhibited the activity of all proteins (Fig. 26 A, B and C). The most potent inhibition was in case of APE1 (IC_{50} = 5.67 ± 1.82 nM), while the rest of the enzymes were also inhibited quite strongly as IC_{50} values were in micromolar range (Fig. 26 D). ATA therefore shows no measurable selectivity for EXOG over the other enzymes tested. All assays were performed in the same buffer to allow direct comparison. This led to

a hypothesis that inhibition by ATA may be connected to possible similarities in charge of the active site of tested enzymes, as they all bind negatively charged DNA molecule, which also stands in line with the results showing the importance of negatively charged carboxyl groups of ATA. This is also consistent with the published data described earlier, where ATA was shown to compete with the DNA for binding in the active site and interfere with electrostatic interactions contributing to protein-nucleic acid binding (Gonzalez et al., 1980).

This broad activity has direct consequences for the aim of the project. Since ATA inhibits not only EXOG but also APE1, FEN1, Pol γ, MGME1, RNase H1 and Ligase III, it cannot be considered a specific EXOG inhibitor, which was the original goal. Moreover, because APE1, FEN1, RNase H1 and ligase III are not mitochondria-specific but are shared with, or act primarily in, the nucleus, ATA would be expected to reduce DNA repair capacity in both compartments rather than selectively in mitochondria. This limits its use as a selective tool, but it does not exclude its potential as an agent that lowers mtDNA repair capacity in general. This is the reasoning behind testing ATA in combination with a chemotherapeutic, which is described in the next section.

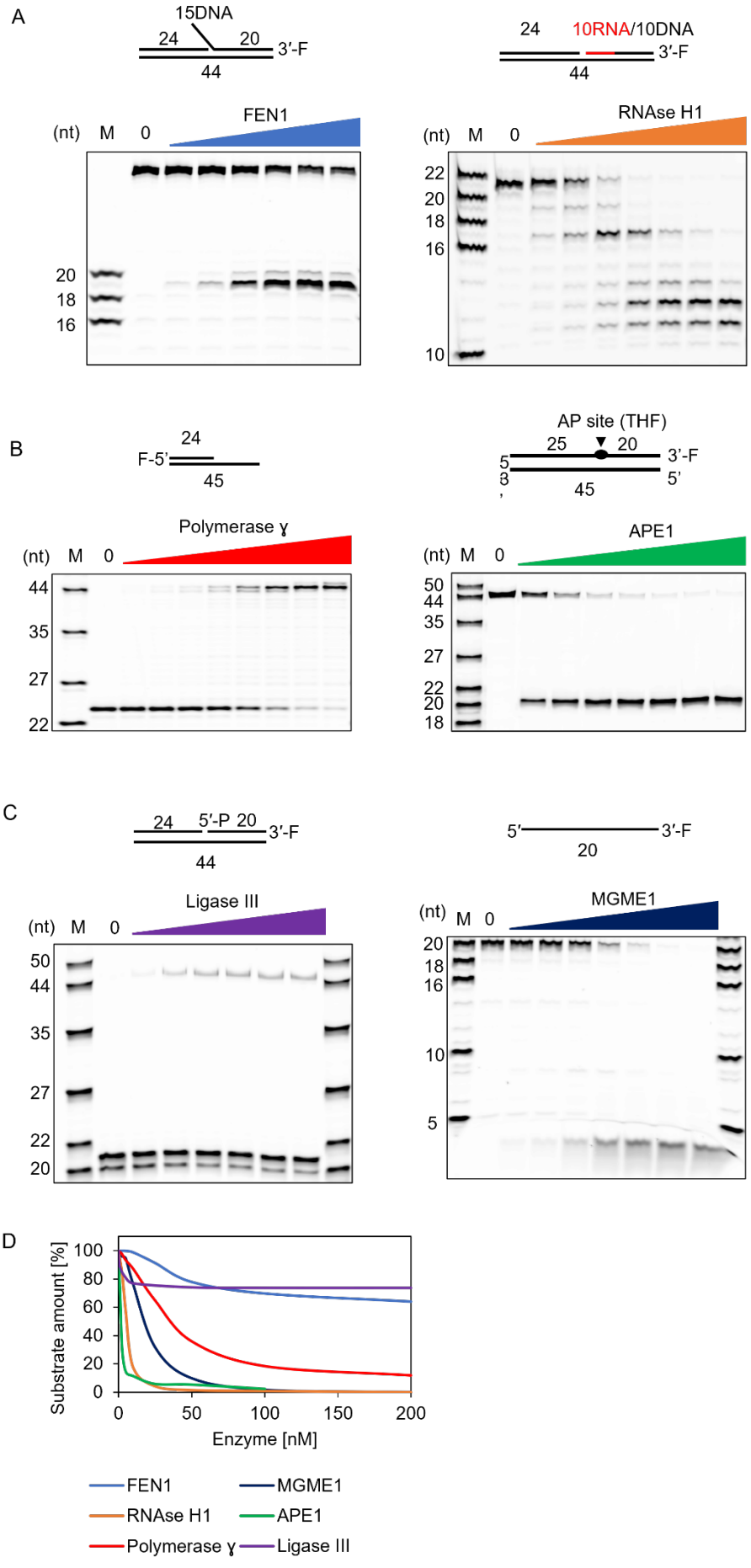
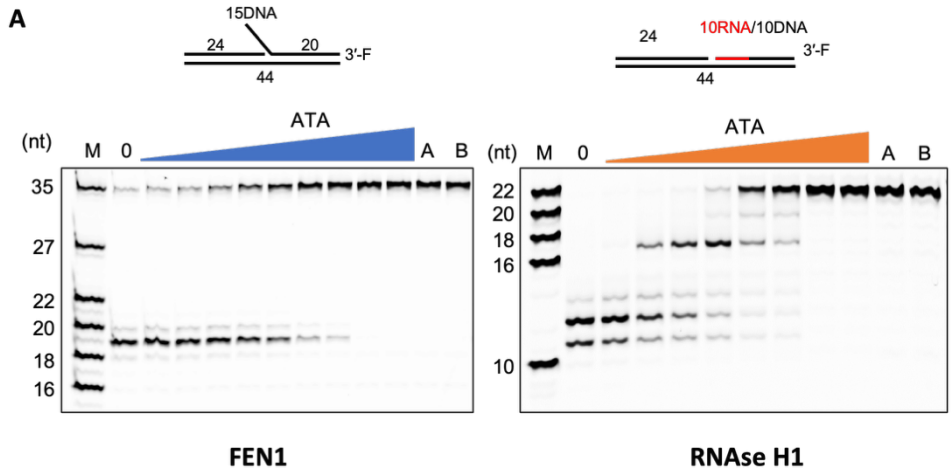
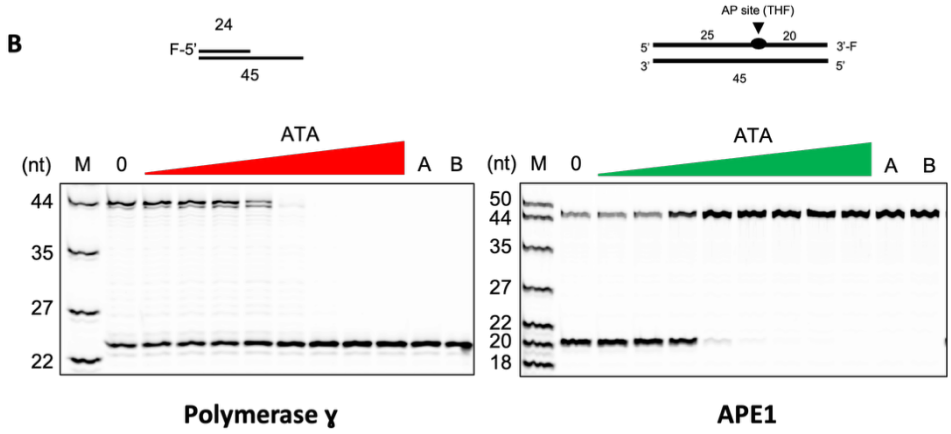
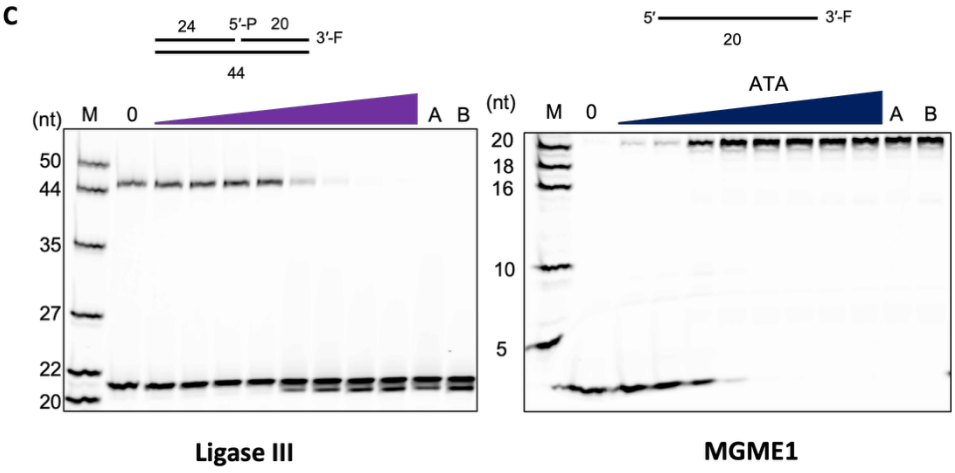

A
15DNA
24
20
3′-F
44
24
10RNA/10DNA
3′-F
44
FEN1
(nt) M 0
20
18
16
RNAse H1
(nt) M 0
22
20
18
16
10
B
24
F-5’
45
AP site (THF)
25
20
5
3’
3’-F
5’
45
Polymerase ɣ
(nt) M 0
44
35
27
22
APE1
(nt) M 0
50
44
35
27
22
20
18
C
24
5′-P 20
3′-F
44
5′
3′-F
20
Ligase III
(nt) M 0
50
44
35
27
22
20
MGME1
(nt) M 0
20
18
16
10
5
D
Substrate amount [%]
100
80
60
40
20
0
0
50
100
150
200
Enzyme [nM]
FEN1
MGME1
RNAse H1
APE1
Polymerase ɣ
Ligase III

Figure 25. (A) Cleavage of various substrates by DNA repair enzymes. *(A) 200 nM of substrate was incubated with increasing concentrations of FEN1 (left) and RNase H1 (right). (B) 200 nM of substrate was incubated with increasing concentrations of Pol γ (left) and APE1 (right). (C) 200 nM of substrate was incubated with increasing concentrations of Ligase III (left) and MGME1 (right). For FEN1, RNase H1, Pol γ, APE1 and MGME1 after 5 min of incubation at 25 °C, reactions were stopped and analyzed on a 23% denaturing polyacrylamide gel. Corresponding molecular weight standards are marked with M. (D) Densitometric analysis of the decrease of tested substrates in (A), (B), (C).*


A
15DNA
24
20
3′-F
44
10RNA/10DNA
ATA
(nt)
M
0
A
B
FEN1
RNAse H1


B
F-5′
AP site (THF)
25
5′
3′
45
Polymerase γ
APE1


C
5′-P
Ligase III
MGME1

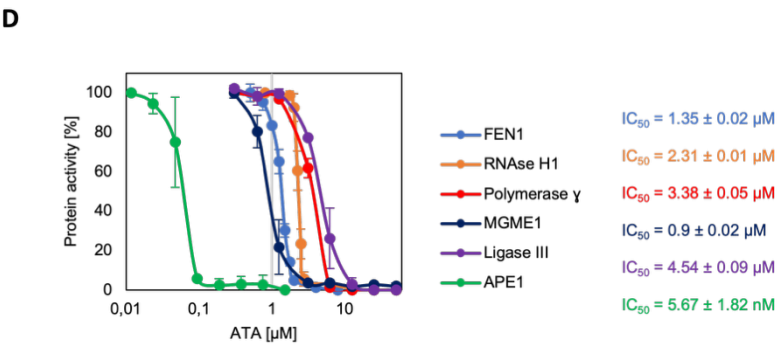

D
Protein activity [%]
ATA [μM]
FEN1
RNAse H1
Polymerase γ
MGME1
Ligase III
APE1
IC50 = 1.35 ± 0.02 μM
IC50 = 2.31 ± 0.01 μM
IC50 = 3.38 ± 0.05 μM
IC50 = 0.9 ± 0.02 μM
IC50 = 4.54 ± 0.09 μM
IC50 = 5.67 ± 1.82 nM

Figure 26. ATA inhibits activities of various components of DNA repair machinery. *(A) 200 nM offlap-DNA was incubated with 100 nM of FEN1 and increasing concentrations of ATA (left panel) 200 nM of RNA/DNA hybrid was incubated with 100 nM RNase H1 and increasing concentrations of ATA. (B) 200 nM of substrate was incubated with 100 nM of Pol γ and increasing concentrations of ATA (left panel). 200 nM of dsDNA 45-mer with AP site was incubated with 10 nM of APE1 and increasing concentrations of ATA (right panel). (C) 200 nM of nicked dsDNA (44-mer) was incubated with 100 nM of DNA Ligase III and increasing concentrations of compound ATA (left panel). 200 nM of 20 nt ssDNA substrate was incubated with 100 nM of MGME1 and increasing concentrations of ATA (right panel). (D) Plot based on densitometric analysis of substrate decrease in A, B and C with calculated IC_{50} values. Y error bars represent SD. IC_{50} value is presented as the mean ± SD of three independent experiments.*

8.2.10. Camptothecin increases the cellular effects of ATA

Because ATA broadly reduces the activity of DNA repair enzymes rather than acting selectively on EXOG, its potential as a general inhibitor of DNA repair was explored in combination with a chemotherapeutic agent. Impairment of mtDNA integrity can, depending on the cell type, severity and context activate cell death pathways reduce cellular bioenergetics and lower cell viability and survival (Van Houten et al., 2016). Combination of chemotherapeutic, which increases DNA damage, with a small molecule compound, that lowers DNA repair capacity, could be potentially used to enhance cancer cell killing, which was tested in the following experiments. CPT is used in treatment of ovarian, colon and small-cell lung cancer (Bertino, 1997; Mondal et al., 2019). From the pharmacological point of view, CPT is a unique compound as its principal cellular target is Topoisomerase I (TOP1) enzyme (Eng et al., 1988; Nitiss and Wang, 1988). TOP1 enzymes are crucial in higher eukaryotes (Morham et al., 1996; Zhang et al., 2000). Their role is to relax supercoiled DNA during transcription or replication (Koster et al., 2005). The mode of action of CPT on TOP1 relies on the fact, that compound traps the enzyme on the DNA during its cleavage intermediate step, causes double strand breaks (DSBs), thus interfering with DNA replication and transcription processes and causing the cytotoxic effect (Hertzberg et al., 1989). The clinical use of CPT alone is limited, however, because high doses cause profound side effects (Creaven et al., 1972). Lowering the effective CPT dose by combination with a small-molecule compound is a promising strategy.

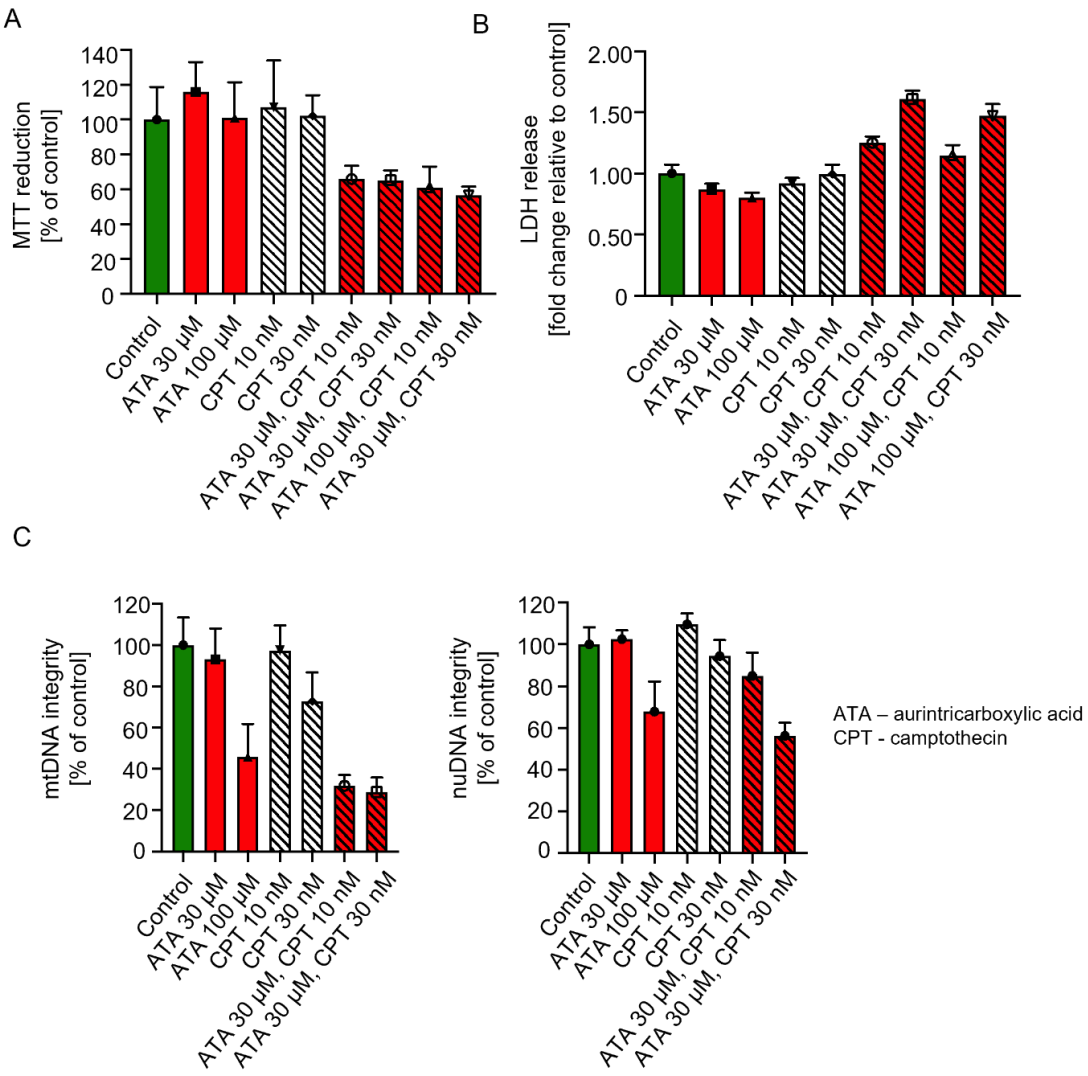


Figure 27. ATA in combination with camptothecin (CPT) affects the bioenergetic status of the cell, decreases mitochondrial and nuclear DNA stability and increases LDH release in lung adenocarcinoma cells. *(A) Metabolic cell activity of A549 cells treated with ATA, CPT or combination of both. (B) Necrotic cell death determined by LDH release of A549 cells treated with ATA, CPT and ATA/CPT. (C) mtDNA and nuclear DNA integrity of A549 cells treated with ATA, CPT or combination of both.*

To test this in cellular model, the A549 human lung adenocarcinoma cell line was used, as lung adenocarcinoma tumors presented an elevated mtDNA integrity level when compared to normal adjacent lung tissue (Fig. 8 A). The following cell-based experiments were performed by Bartosz Szczesny, PhD (UTMB). The influence of ATA

on cellular metabolic activity was first assessed by MTT reduction assay (Fig. 27 A). ATA or CPT alone did not influence cell metabolism but combination of both compounds in all the tested concentrations resulted in drop of the MTT reduction by 40% when compared to control. Cytotoxicity was tested by LDH release (a marker of necrotic cell death, Fig. 27 B). ATA treatment alone did not result in decreased viability of A549 cells. The combination of 30 μM ATA and 30 nM CPT resulted in a 1.6-fold increase in LDH release. Finally, the influence of these compounds on mitochondrial and nuclear DNA integrity was examined (Fig. 27 C). ATA alone at a high concentration (100 μM) significantly decreased mtDNA integrity, whilst combination with CPT (30 μM of ATA and 10 or 30 nM of CPT) allowed using lower concentration of ATA and obtaining a greater effect with a 70% decrease in mtDNA integrity. In case of nuclear DNA integrity, ATA and CPT alone, had similar effect on stability as in case of mtDNA, however combination of these two compounds at 30 μM ATA and 10 nM CPT did not result in significant change of the nuclear DNA integrity. Together, these data indicate that use of combination of ATA and CPT reduces metabolic activity and mtDNA integrity in A549 cells and is consistent with increased sensitivity to CPT. They are consistent with, but do not directly demonstrate inhibition of mtDNA repair.

Taken together, the results presented above provide a consistent picture of how ATA acts on EXOG. ATA inhibits both the endonuclease and the 5'-exonuclease activity of EXOG, with similar potency across different DNA and RNA substrates, which points to the enzyme rather than the substrate as its target. This is further supported by the observation that ATA alone does not alter the DNA substrate. ATA binds the free protein with high apparent affinity and strongly increases its thermal stability. Its preincubation with EXOG lowers the affinity of the enzyme for DNA, consistent with ATA competing with the substrate at or near the active site. Site-directed mutagenesis, guided by docking to the apo structure, showed that several residues in this region alter the apparent sensitivity of EXOG to ATA, however no variant showed reduced sensitivity and these effects could not be separated from changes in substrate binding or catalysis so their exact role in ATA binding remains to be confirmed. The negatively charged carboxyl groups of ATA appear to be important for inhibition, which is consistent with an electrostatic, active-site mode of action shared with the other DNA-processing enzymes that ATA also inhibits. This lack of specificity prevents ATA from serving as a selective EXOG inhibitor, but its ability to lower mtDNA integrity, especially in combination with CPT, supports the

broader idea that reducing mtDNA integrity – potentially by impairing its repair, can sensitize cancer cells with high mtDNA integrity.

8.3. Virtual screening of compound databases

As a second complementary strategy for identifying EXOG candidate inhibitors, *in silico* virtual screening was used. In contrast to the experimental HTS of a physical library, virtual screening allows very large compound databases, up to several millions of compounds, to be searched computationally to identify possible protein inhibitor leads, which can be further optimized. Such approach can result in finding a small number of molecules which have a high affinity and selectivity for their target. In this type of screening, compounds are first docked into a defined binding site on a model of the protein (or across its surface) to predict their location, orientation and conformation. Binding score is then used to estimate their predicted affinity.

This screening was carried out in collaboration with Dr Douglas Houston (University of Edinburgh), who performed the docking pipeline for virtual compounds database: BioAscent. The candidate compounds were then tested experimentally in this work (Chapter 8.3.1.). The screening of two other databases eMolecules and Enamine is still ongoing and is not part of this thesis and aims to incorporate additional filtering to reduce potential off target activity and yield more selective hits (described in Future perspectives).

8.3.1. BioAscent database

To identify the possible EXOG inhibitors from BioAscent database, a cascade of different computational analyses was utilized (Fig. 28). As a first step Vina which is a computational docking tool based on a simple scoring function and rapid conformational search was used. Vina program predicted the binding of 10 000 out of 124 732 compounds present in the library to the active site of EXOG. These compounds were again tested by different docking tool – AutoDock, which allowed to estimate bound conformations and free energies of binding for ligands to EXOG and to compare it with previously obtained poses. Combination of the results from the two programs allowed to improve confidence of the obtained results. Docking often yields several poses with similar scores, and incorrect poses give unreliable affinity predictions. To address this, PoseMatch was used to retain only ligands for which the top Vina and AutoDock poses agreed to within a root-mean-square deviation (RMSD) ≤ 2.0 Å and

only these ligands were considered to be docked correctly. The retained poses were then rescored. The scoring function estimates binding energy from electrostatic and van der Waals forces, a weighted sum of hydrogen-bond and hydrophobic contacts combined with a knowledge-based term reflecting atom-type pairs that occur frequently at a close range and are therefore energetically favorable. This step is performed to evaluate how well each pose for each ligand fits into the binding site and the quality of the interactions it forms with the target. The last step before visual analysis of the bound ligands is pharmacophore filtering. This is a processing of docking results to eliminate molecules which are not fully chemically compatible with the binding site (for example, poses that do not completely fill the site, or that leave unpaired buried hydrogen bond donors or acceptors).

Together these steps yielded 286 compounds, which were purchased from the vendor (BioAscent) and tested experimentally in biochemical assays. The virtual screening was performed by Dr Houston; the experimental testing described below was carried out as part of this work. To assay the compounds in a plate-reader format (384-well plates), different fluorophores attached to ssDNA substrate were first checked for resistance to photobleaching. Based on the obtained preliminary results, ssDNA labelled with Cy3 dye was selected. The substrate was titrated with increasing concentrations of EXOG in the presence of magnesium ions, to observe changes in fluorescence upon cleavage (Fig. 29).

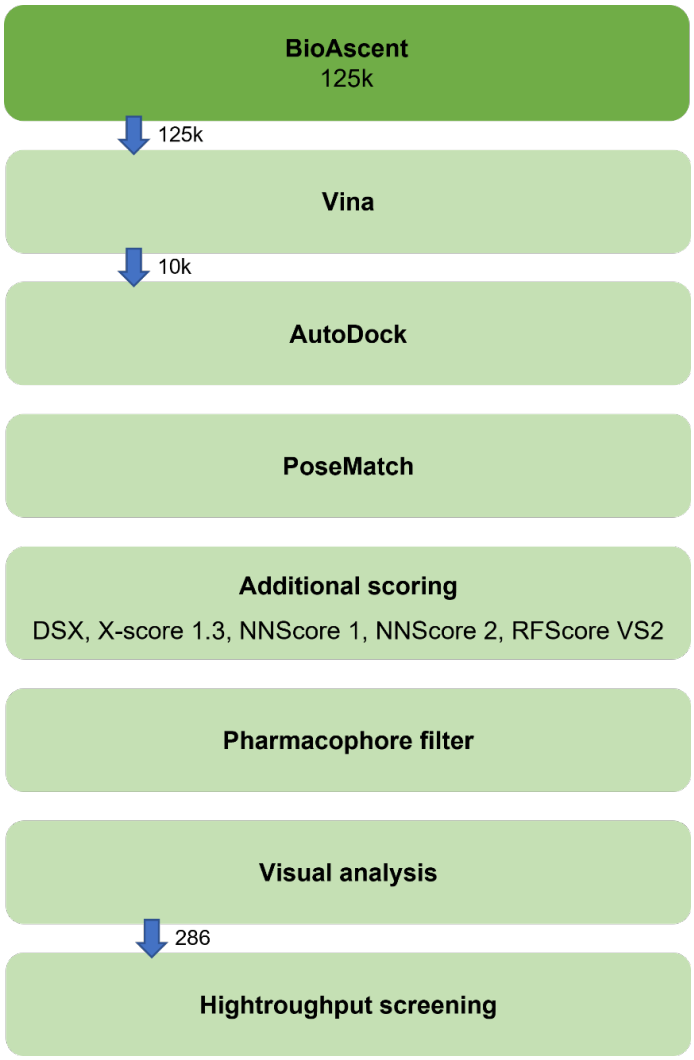


Figure 28. Pipeline of virtual screening used to distinguish potential leads from the BioAscent compound database.

The fluorescence signal decreased with increasing EXOG concentrations (Fig. 29). As a proof of concept, that addition of the inhibitor would reverse this phenomenon, a fixed EXOG concentration of 25 nM, sufficient in terms of signal decrease upon cleavage, was titrated with increasing concentrations of ATA (Fig. 30).

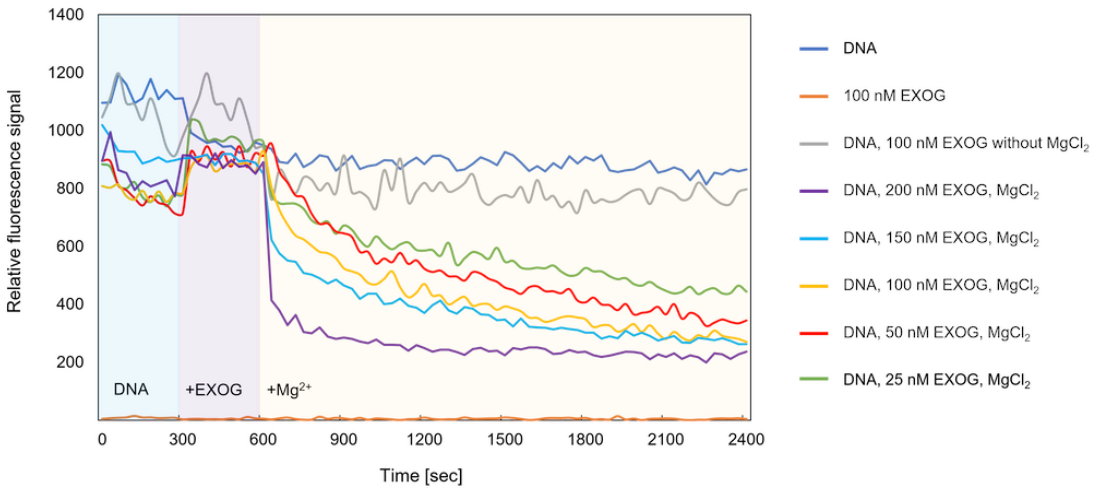


Figure 29. EXOG titration on ssDNA substrate labelled with Cy3 dye. *Activity of EXOG on a substrate was illustrated as a decrease in fluorescence signal proportional to increasing concentrations of the enzyme. Assay was divided into three steps where first fluorescence of DNA alone was measured for 300 s, then EXOG or buffer was added and measured for another 300 s and the last step was addition of $MgCl_2$ or buffer, as marked in legend. As a negative controls signal from DNA (blue), EXOG (orange) and DNA mixed with EXOG without the cofactor (Mg^{2+}) (grey) were measured.*

ATA was added to the substrate before the enzyme and magnesium so that any direct effect of the compound on the fluorescence signal (quenching or enhancement) could be identified. ATA did not significantly influence the measured signal and increasing concentrations of ATA progressively prevented the EXOG-dependent decrease in fluorescence, consistent with inhibition of cleavage. The proposed set up of experiment could be further used as a mean of HTS method and was therefore applied to screen the 286 compounds selected by virtual screening of the BioAscent database. To better visualize obtained results the heatmap showing the change of EXOG activity upon addition of different compounds was utilized (Fig. 31). Only the compounds that inhibited EXOG activity by at least 50% were considered for further assessment using secondary screen. Nine compounds that met this criterion were in positions K1, L1, D2, L3, E4, G5, J13, I18 and L19 marked on the heatmap (Fig. 31). In the secondary screen, inhibition of ssDNA cleavage was analyzed by urea-PAGE, using the same oligonucleotide as in the ATA inhibition studies. Fixed concentrations of EXOG and substrate were titrated with increasing concentrations of each compound.

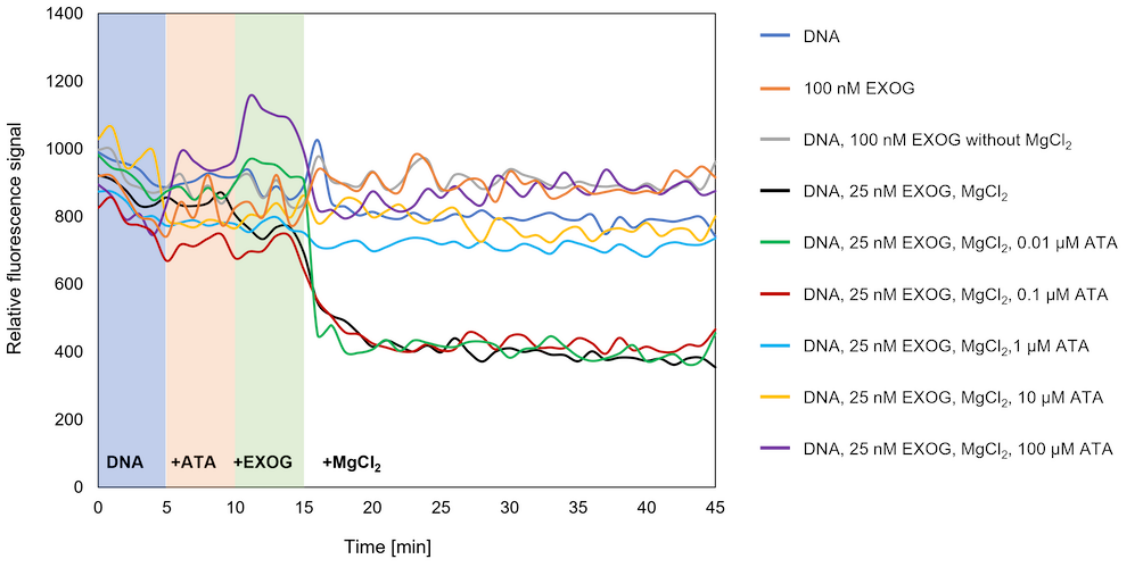


Figure 30. ATA titration of EXOG measured as a change in the substrate fluorescence. *Assay was divided into four steps where first fluorescence of DNA alone was measured for 5 min., then ATA or buffer was added and measured for another 5 min., next step was EXOG or buffer addition and measurement of another 5 min. and the last step was addition of $MgCl_2$ or buffer, as marked in legend, which was measured for 30 min. As a negative controls signal from DNA (blue), EXOG (orange) and DNA mixed with EXOG without the cofactor (Mg^{2+}) (grey) were measured, as a positive control DNA mixed with EXOG and the cofactor (Mg^{2+}) (black) was checked. ATA was added in concentrations ranging from 0.01 µM to 100 µM as marked in the legend.*

Based on the obtained results, the inhibition by three compounds named 0667, 1168 and 4548 (naming based on their vendor numbers) was confirmed (Fig. 32 A). The IC_{50} values were established for all of them, and they were 21.9, 20.2 and 42.4 µM, respectively (Fig. 32 B and C). These values were much higher than the one obtained for the ATA, nevertheless it was hoped that these compounds would inhibit EXOG more selectively than ATA. For this reason, it was checked if they would inhibit some of the previously tested DNA processing enzymes. Unfortunately, they were also inhibiting other proteins. After a closer look, it was recognized that they might be so-called aggregators. Such compounds often give false positive results, and it is quite common in high-throughput screening (Alturki et al., 2018; Viviani et al., 2018; Schencking et al., 2021).

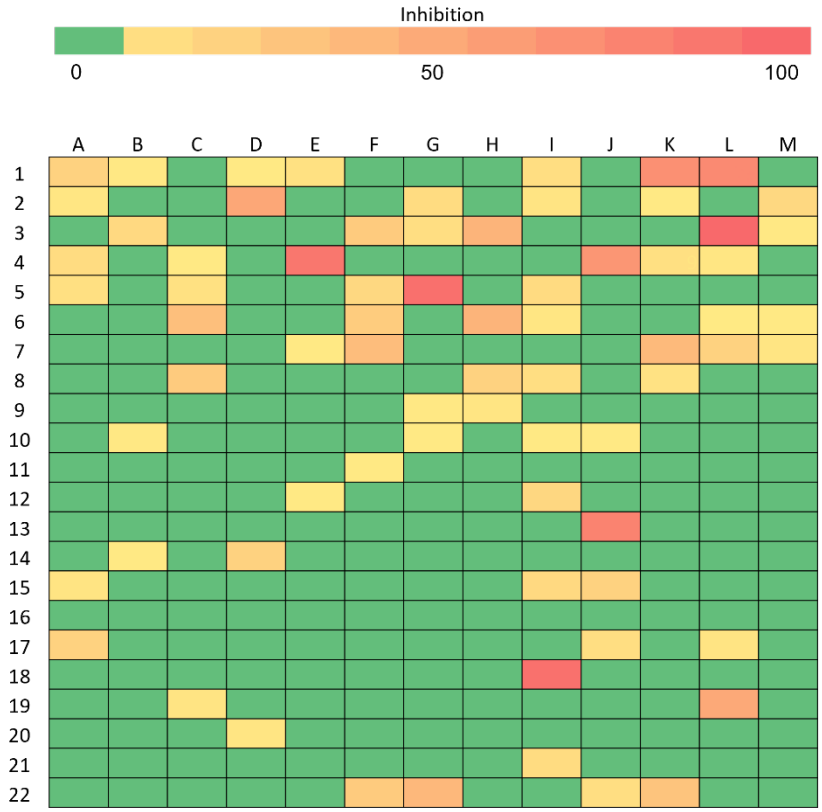


Figure 31. Result of screening of compounds selected from BioAscent database*. Inhibition shown in the color range, where no inhibition is marked as zero (green) and full inhibition is marked as one hundred (red).*

Compounds first recognized as inhibitors do not truly act against their specific biological targets but rather form colloidal aggregates that sequester the enzyme. Such compounds in solution in micromolar or nanomolar concentrations form aggregates and non-specifically inhibit proteins *in vitro* by adsorption to their surface (McGovern et al., 2002). It is often that in buffers containing Mg^{2+}, compound which is an aggregator presents a poor solubility, which was noticed in case of inhibitors 0667, 1168 and 4548, but only in higher concentrations of the compounds. The assumption was that this relates to their chemical characteristics, as the problem was not occurring at lower working concentrations. On the other hand, no problems with solubility in aqueous solutions were observed for ATA, which also inhibits many proteins involved in DNA processing. Unfortunately, even though there are computational methods (Seidler et al., 2003; Irwin et al., 2015), that predict potential aggregators, still experimental approach is the best way to detect aggregator compounds.

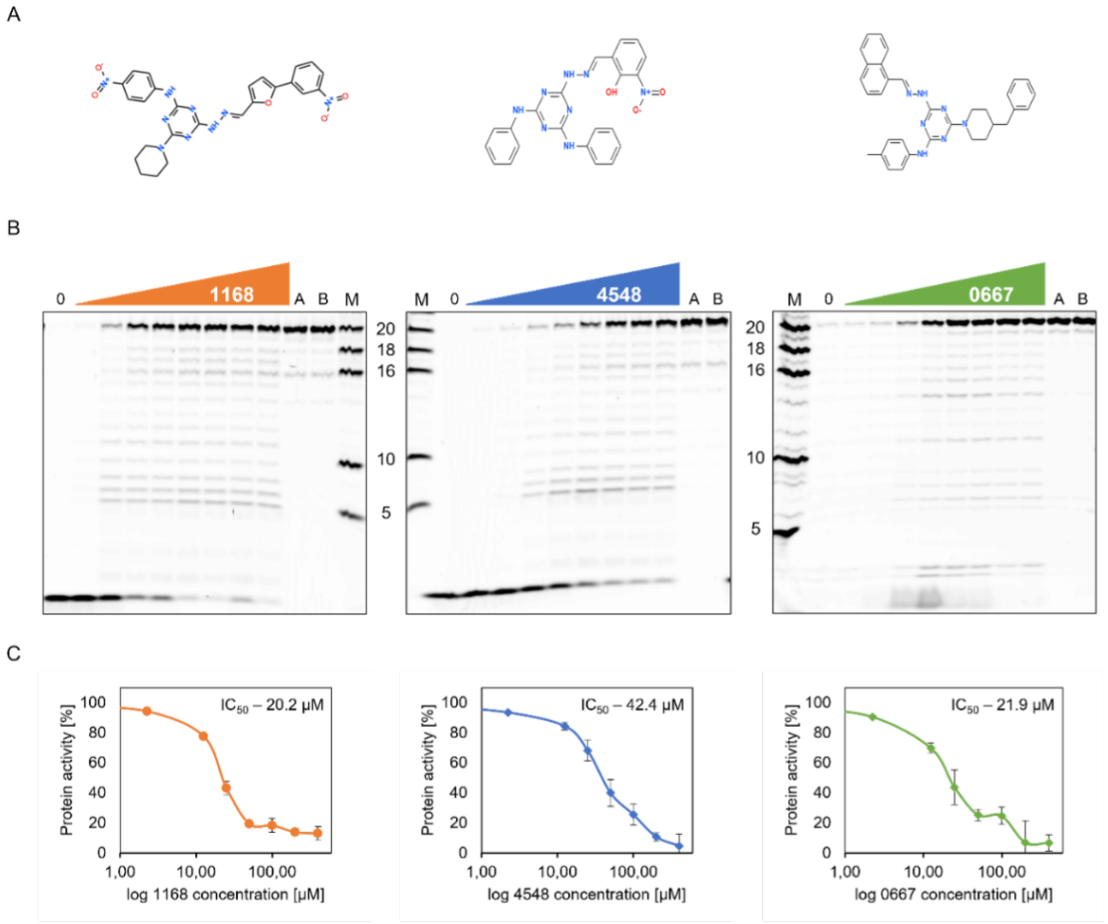


Figure 32. EXOG cleavage of ssDNA substrate is inhibited by increasing concentrations of inhibitors 1168, 4548 and 0667. *(A) Structures of the compounds selected based on the secondary screening: 1168 (left panel), 4548 (middle panel), 0667 (right panel). (B) 200 nM of ssDNA was incubated with 100 nM of EXOG and increasing concentrations of inhibitors 1168 (left panel), 4548 (middle panel) and 0667 (right panel). After 5 min of incubation at 25 °C, reactions were stopped and analyzed on a 23% denaturing polyacrylamide gel. Corresponding molecular weight standards are marked with M. DNA control is marked with A and DNA incubated with highest concentration of the inhibitor used is marked with B. (C) Densitometric analysis of the protein activity on tested substrate in (B) together with IC_{50} calculations. Y error bars represent SD.*

9. SUMMARY

Although the role of EXOG in mitochondrial DNA (mtDNA) repair is well established, no specific inhibitor of this enzyme has been identified to date. The aim of this project was to find a small-molecule compound that modulates EXOG activity and could serve as a starting point for an anti-cancer agent increasing the sensitivity of cancer cells to conventional chemotherapeutics. To this end, two complementary strategies were used: high-throughput screening (HTS) of a commercial compound library and structure-based virtual screening.

The first strategy was HTS of the LOPAC®1280 library, which allowed for the identification of ATA as an EXOG inhibitor. Its influence on EXOG activity towards different DNA and RNA substrates was analyzed, the apparent K_d of the interaction was established, and a mechanism of inhibition was proposed. ATA inhibited both the endonuclease and the 5′-exonuclease activity of EXOG, bound the free protein with high apparent affinity and markedly increased its thermal stability, and its preincubation with EXOG lowered the affinity of the enzyme for DNA. These observations are consistent with ATA competing with the substrate at or near the active site. To identify the residues involved in ATA binding, co-crystallization and soaking studies were performed; however, no interpretable electron density corresponding to ATA was observed under the conditions applied, so a direct structural view of the binding site was not obtained. To address this, *in silico* docking studies were performed, which proposed several candidate residues. Site-directed mutagenesis was then used to obtain point-mutation variants, whose activity was tested in the presence of ATA. Variants F114A, R138A, F158A and R183A were more effectively inhibited by ATA than the wild-type protein, and N171A lost its activity almost totally at the lowest ATA concentration. These substitutions alter the apparent sensitivity of EXOG to ATA and map to the region near the active site; however, because no variant showed reduced sensitivity and the effects could not be separated from changes in substrate binding or catalysis, their exact contribution to ATA binding remains to be confirmed. Testing of ATA-related compounds suggested that the carboxyl groups and the negative charge of ATA contribute to inhibition.

ATA tested *in vitro*, however, turned out to be highly unspecific, as it also inhibited several other proteins involved in DNA and RNA processing, showing no measurable selectivity for EXOG. Nevertheless, the combination of ATA with camptothecin (CPT)

in a lung adenocarcinoma cell line gave promising results, decreasing mtDNA integrity, cellular metabolic activity and cell viability; these data are consistent with, but do not directly demonstrate, impairment of mtDNA repair.

The second strategy was virtual screening, which allowed the selection of three possible EXOG inhibitors; after further analysis, these compounds also inhibited other proteins and turned out to be aggregators and therefore did not provide a validated selective inhibitor.

Taken together, this project did not ultimately identify a selective inhibitor of EXOG. It identified ATA as a potent inhibitor of EXOG activity and characterized its likely mode of action, but also showed that ATA is not specific, as it inhibits several other DNA- and RNA-processing enzymes. The compounds obtained through virtual screening likewise did not provide a validated selective inhibitor. The ATA-CPT combination provided the most encouraging cellular result and supports further investigation of whether reducing mtDNA integrity can sensitize cancer cells to DNA-damaging agents.

10. FUTURE PERSPECTIVES

The search for a specific EXOG inhibitor is still ongoing. Because ATA was found to inhibit several other proteins involved in DNA and RNA processing, further work will focus on identifying a compound that inhibits EXOG selectively, and several complementary directions are being pursued.

The virtual screening approach is being extended to two additional compound databases, eMolecules and Enamine, in collaboration with Douglas Houston, PhD (University of Edinburgh) and Vincent Blay, PhD (University of Valencia). The screening pipeline is being refined to reduce the number of nonspecific hits: the aim is to establish a machine-learning-based scoring function that evaluates multiple ligand poses according to their RMSD from the native pose and eliminates compounds predicted to have off-target activity. Because several of the hits obtained so far, as well as ATA itself, proved to be nonspecific or aggregating, filters intended to remove such compounds at the computational stage would be valuable; as shown in this work, however, experimental testing remains necessary to confirm that a candidate is neither an aggregator nor a broad-spectrum inhibitor.

The candidate compounds prioritized *in silico* will then be validated experimentally, using the biochemical assays established in this work, to confirm inhibition of EXOG and to determine their potency. To assess selectivity, promising candidates will be profiled against the panel of DNA- and RNA-processing enzymes used here: APE1, FEN1, Pol γ, MGME1, RNase H1 and Ligase III – so that compounds inhibiting EXOG but sparing the other enzymes can be prioritized.

Finally, in addition to the active site, another pocket identified in the EXOG core domain is being explored as a possible allosteric site. Because allosteric sites are often less conserved than catalytic ones, they may offer an alternative route to selective inhibition, and a search for allosteric inhibitors is already under way.

11. SCIENTIFIC ACCOMPLISHMENTS

11.1. List of publications in peer-reviewed journals

- Karlowicz, A., Dubiel, A. B., **Wyszkowska, M.**, Hossain, K. A., Czub, J., & Szymanski, M. R. (2025). Mitochondrial exonuclease EXOG supports DNA integrity by the removal of single-stranded DNA flaps. *Nucleic Acids Research*, *53*(5), gkaf099. https://doi.org/10.1093/nar/gkaf099
- Nekulová, M., **Wyszkowska, M.***, Friedlová, N., Uhrík, L., Zavadil Kokáš, F., Hrabal, V., Hernychová, L., Vojtěšek, B., Hupp, T. R., & Szymański, M. R. (2024).

  Biochemical evidence for conformational variants in the anti-viral and pro-metastatic protein IFITM1. *Biological Chemistry*, *405*(5), 311–324. https://doi.org/10.1515/hsz-2023-0327 (done in collaboration with T. Hupp group)
- Karlowicz, A., Dubiel, A. B., Czerwinska, J., Bledea, A., Purzycki, P., **Grzelewska, M.****, McAuley, R. J., Szczesny, R. J., Brzuska, G., Krol, E., Szczesny, B., & Szymanski, M. R. (2022). *In vitro* reconstitution reveals a key role of human mitochondrial EXOG in RNA primer processing. *Nucleic Acids Research*, *50*(14), 7991–8007. https://doi.org/10.1093/nar/gkac581

* Equal contribution of the first two authors.

** Published under the author's maiden name (Grzelewska).

11.2. Manuscripts in preparation

- **Marta Wyszkowska**, Andrzej B. Dubiel, Piotr Purzycki, Douglas Houston, Bartosz Szczesny, Michal R. Szymanski. Targeted inhibition of human mitochondrial nuclease EXOG, as a means of mitochondrial DNA repair process modulation.
- **Marta Wyszkowska**, Beata Walter, Douglas Houston, Michal R. Szymanski. Virtual screening to discover inhibitors of the human mitochondrial nuclease EXOG.

11.3. Conferences and workshops

- Anna Karlowicz, Andrzej B Dubiel, Adela Bledea, Piotr Purzycki, **Marta Grzelewska**, Ryan J McAuley, Bartosz Szczesny, Jolanta Czerwinska, Roman J Szczesny, Michal R Szymanski. *In vitro reconstitution reveals a key role of human mitochondrial EXOG in RNA primer processing.* EMBO Workshop: Molecular Biology of Mitochondrial Gene Maintenance and Expression. 2022, Bro, Sweden.
- Anna Karlowicz, Andrzej B Dubiel, Adela Bledea, Piotr Purzycki, **Marta Grzelewska**, Ryan J McAuley, Bartosz Szczesny, Jolanta Czerwinska, Roman J Szczesny, Michal R Szymanski. *In vitro reconstitution reveals a key role of human mitochondrial EXOG in RNA primer processing*. iNEXT-Discovery 2nd Annual Scientific Meeting. 2022, Warsaw, Poland.
- **Marta Grzelewska**, Adela Bledea, Andrzej B. Dubiel, Michal R. Szymanski. *Solubilization and purification of human mitochondrial inner membrane nuclease EXOG for structural and functional studies.* EMBO Workshop: Tools for Structural Biology of Membrane Proteins. 2019, Hamburg, Germany.
- **Marta Grzelewska**, Adela Bledea, Andrzej B. Dubiel, Michal R. Szymanski. *Purification and structural characterization of EXOG, a human mitochondrial inner membrane nuclease.* The 44th FEBS Congress: From molecules to the living systems. 2019, Kraków.

11.4. Additional scientific, teaching and outreach activity

- **Supervision.** Supervision of Nela Friedlová, a visiting trainee from Masaryk University (Brno, Czech Republic). The internship was carried out within the continuing collaboration with Prof. Ted R. Hupp on the IFITM1 protein. Under my supervision, the trainee was introduced to *in vitro* methods for the analysis of interactions between the IFITM1 and IFITM3 proteins, including biolayer interferometry (BLI) and microscale thermophoresis (MST), October 2022.
- **Scientific outreach.** Co-organisation of the "Klimat dla nauki" Science Picnic and the Open Days of the University of Gdańsk, 22 April 2022.
- **Conference activity.** Session chair at the 12th Biotechnology Students' Conference at IFB UG and MUG, 9-10 June 2022.

12. REFERENCES

Alam, T. I. (2003). Human mitochondrial DNA is packaged with TFAM. *Nucleic Acids Research*, *31*(6), 1640–1645. https://doi.org/10.1093/nar/gkg251

Alberts, B., Johnson, A., Lewis, J., Morgan, D., Raff, M., Roberts, K., & Walter, P. (2015). *Molecular biology of the cell* (6th ed.). Garland Science.

Alexeyev, M. F. (2009). Is there more to aging than mitochondrial DNA and reactive oxygen species? *The FEBS Journal*, *276*(20), 5768–5787. https://doi.org/10.1111/j.1742-4658.2009.07269.x

Alturki, M. S., Fuanta, N. R., Jarrard, M. A., Hobrath, J. V., Goodwin, D. C., Rants'o, T. A., & Calderón, A. I. (2018). A multifaceted approach to identify non-specific enzyme inhibition: Application to Mycobacterium tuberculosis shikimate kinase. *Bioorganic & Medicinal Chemistry Letters*, *28*(4), 802–808. https://doi.org/10.1016/j.bmcl.2017.12.002

Anderson, S., Bankier, A. T., Barrell, B. G., De Bruijn, M. H. L., Coulson, A. R., Drouin, J., Eperon, I. C., Nierlich, D. P., Roe, B. A., Sanger, F., Schreier, P. H., Smith, A. J. H., Staden, R., & Young, I. G. (1981). Sequence and organization of the human mitochondrial genome. *Nature*, *290*(5806), 457–465. https://doi.org/10.1038/290457a0

Andersson, S. G., Zomorodipour, A., Andersson, J. O., Sicheritz-Pontén, T., Alsmark, U. C., Podowski, R. M., Näslund, A. K., Eriksson, A. S., Winkler, H. H., & Kurland, C. G. (1998). The genome sequence of Rickettsia prowazekii and the origin of mitochondria. *Nature*, *396*(6707), 133–140. https://doi.org/10.1038/24094

Baell, J. B., & Holloway, G. A. (2010). New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays. *Journal of Medicinal Chemistry*, *53*(7), 2719–2740. https://doi.org/10.1021/jm901137j

Barnes, T., Kim, W.-C., Mantha, A. K., Kim, S.-E., Izumi, T., Mitra, S., & Lee, C. H. (2009). Identification of Apurinic/apyrimidinic endonuclease 1 (APE1) as the endoribonuclease that cleaves c-myc mRNA. *Nucleic Acids Research*, *37*(12), 3946–3958. https://doi.org/10.1093/nar/gkp275

Barshad, G., Blumberg, A., Cohen, T., & Mishmar, D. (2018). Human primitive brain displays negative mitochondrial-nuclear expression correlation of respiratory genes. *Genome Research*, *28*(7), 952–967. https://doi.org/10.1101/gr.226324.117

Beckman, K. B., & Ames, B. N. (1999). Endogenous oxidative damage of mtDNA. *Mutation Research*, *424*(1–2), 51–58. https://doi.org/10.1016/s0027-5107(99)00007-x

Berk, A. J., & Clayton, D. A. (1974). Mechanism of mitochondrial DNA replication in mouse L-cells: Asynchronous replication of strands, segregation of circular daughter molecules, aspects of topology and turnover of an initiation sequence. *Journal of Molecular Biology*, *86*(4), 801–824. https://doi.org/10.1016/0022-2836(74)90355-6

Bertino, J. R. (1997). Chemotherapy of colorectal cancer: History and new themes. Seminars in Oncology, 24(5 Suppl 18), S18-3–S18-7.

Bogenhagen, D., & Clayton, D. A. (1974). The number of mitochondrial deoxyribonucleic acid genomes in mouse L and human HeLa cells. Quantitative isolation of mitochondrial deoxyribonucleic acid. *The Journal of Biological Chemistry*, *249*(24), 7991–7995.

Bogenhagen, D. F., Rousseau, D., & Burke, S. (2008). The Layered Structure of Human Mitochondrial DNA Nucleoids. *Journal of Biological Chemistry*, *283*(6), 3665–3675. https://doi.org/10.1074/jbc.M708444200

Centers for Disease Control and Prevention. (2024, August 26). *Cancer treatments*. https://www.cdc.gov/cancer-survivors/patients/treatments.html

Chabner, B. A., & Roberts, T. G. (2005a). Chemotherapy and the war on cancer. *Nature Reviews Cancer*, *5*(1), 65–72. https://doi.org/10.1038/nrc1529

Chang, D. D., & Clayton, D. A. (1984). Precise identification of individual promoters for transcription of each strand of human mitochondrial DNA. *Cell*, *36*(3), 635–643. https://doi.org/10.1016/0092-8674(84)90343-X

Chen, Q., Van Der Sluis, P. C., Boulware, D., Hazlehurst, L. A., & Dalton, W. S. (2005). The FA/BRCA pathway is involved in melphalan-induced DNA interstrand cross-link repair and accounts for melphalan resistance in multiple myeloma cells. *Blood*, *106*(2), 698–705. https://doi.org/10.1182/blood-2004-11-4286

Cheung-Ong, K., Giaever, G., & Nislow, C. (2013). DNA-Damaging Agents in Cancer Chemotherapy: Serendipity and Chemical Biology. *Chemistry & Biology*, *20*(5), 648–659. https://doi.org/10.1016/j.chembiol.2013.04.007

Cohen, B. E., & Bangham, A. D. (1972). Diffusion of Small Non-Electrolytes across Liposome Membranes. *Nature*, *236*(5343), 173–174. https://doi.org/10.1038/236173a0

Cooke, M. S., Evans, M. D., Dizdaroglu, M., & Lunec, J. (2003). Oxidative DNA damage: Mechanisms, mutation, and disease. *The FASEB Journal*, *17*(10), 1195–1214. https://doi.org/10.1096/fj.02-0752rev

Copeland, W. C., & Longley, M. J. (2008). DNA2 resolves expanding flap in mitochondrial base excision repair. *Molecular Cell*, *32*(4), 457–458. https://doi.org/10.1016/j.molcel.2008.11.007

Creaven, P. J., Allen, L. M., & Muggia, F. M. (1972). Plasma camptothecin (NSC-100880) levels during a 5-day course of treatment: Relation to dose and toxicity. *Cancer Chemotherapy Reports*, *56*(5), 573–578.

Crews, S., Ojala, D., Posakony, J., Nishiguchi, J., & Attardi, G. (1979). Nucleotide sequence of a region of human mitochondrial DNA containing the precisely identified origin of replication. *Nature*, *277*(5693), 192–198. https://doi.org/10.1038/277192a0

Cymerman, I. A., Chung, I., Beckmann, B. M., Bujnicki, J. M., & Meiss, G. (2008). EXOG, a novel paralog of Endonuclease G in higher eukaryotes. *Nucleic Acids Research*, *36*(4), 1369–1379. https://doi.org/10.1093/nar/gkm1169

De Bont, R. (2004). Endogenous DNA damage in humans: A review of quantitative data. *Mutagenesis*, *19*(3), 169–185. https://doi.org/10.1093/mutage/geh025

De Souza-Pinto, N. C., Wilson, D. M., Stevnsner, T. V., & Bohr, V. A. (2008). Mitochondrial DNA, base excision repair and neurodegeneration. *DNA Repair*, *7*(7), 1098–1109. https://doi.org/10.1016/j.dnarep.2008.03.011

D'Erchia, A. M., Atlante, A., Gadaleta, G., Pavesi, G., Chiara, M., De Virgilio, C., Manzari, C., Mastropasqua, F., Prazzoli, G. M., Picardi, E., Gissi, C., Horner, D., Reyes, A., Sbisà, E., Tullo, A., & Pesole, G. (2015). Tissue-specific mtDNA abundance from exome data and its correlation with mitochondrial transcription, mass and respiratory activity. *Mitochondrion*, *20*, 13–21. https://doi.org/10.1016/j.mito.2014.10.005

DeVita, V. T., & Chu, E. (2008). A History of Cancer Chemotherapy. *Cancer Research*, *68*(21), 8643–8653. https://doi.org/10.1158/0008-5472.CAN-07-6611

Dorjsuren, D., Kim, D., Maloney, D. J., Wilson, D. M., & Simeonov, A. (2011). Complementary non-radioactive assays for investigation of human flap endonuclease 1 activity. *Nucleic Acids Research*, *39*(2), e11–e11. https://doi.org/10.1093/nar/gkq1082

Durrant, J. D., & McCammon, J. A. (2010). NNScore: A Neural-Network-Based Scoring Function for the Characterization of Protein−Ligand Complexes. *Journal of Chemical Information and Modeling*, *50*(10), 1865–1871. https://doi.org/10.1021/ci100244v

Durrant, J. D., & McCammon, J. A. (2011). NNScore 2.0: A Neural-Network Receptor–Ligand Scoring Function. *Journal of Chemical Information and Modeling*, *51*(11), 2897–2903. https://doi.org/10.1021/ci2003889

Eng, W. K., Faucette, L., Johnson, R. K., & Sternglanz, R. (1988). Evidence that DNA topoisomerase I is necessary for the cytotoxic effects of camptothecin. *Molecular Pharmacology*, *34*(6), 755–760.

Falkenberg, M., & Gustafsson, C. M. (2020). Mammalian mitochondrial DNA replication and mechanisms of deletion formation. *Critical Reviews in Biochemistry and Molecular Biology*, *55*(6), 509–524. https://doi.org/10.1080/10409238.2020.1818684

Farge, G., & Falkenberg, M. (2019). Organization of DNA in Mammalian Mitochondria. *International Journal of Molecular Sciences*, *20*(11), 2770. https://doi.org/10.3390/ijms20112770

Forli, S., Huey, R., Pique, M. E., Sanner, M. F., Goodsell, D. S., & Olson, A. J. (2016). Computational protein–ligand docking and virtual drug screening with the AutoDock suite. *Nature Protocols*, *11*(5), 905–919. https://doi.org/10.1038/nprot.2016.051

Fortini, P. (2003). The base excision repair: Mechanisms and its relevance for cancer susceptibility. *Biochimie*, *85*(11), 1053–1071. https://doi.org/10.1016/j.biochi.2003.11.003

Ghosh, U., Giri, K., & Bhattacharyya, N. P. (2009). Interaction of aurintricarboxylic acid (ATA) with four nucleic acid binding proteins DNase I, RNase A, reverse transcriptase and Taq polymerase. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *74*(5), 1145–1151. https://doi.org/10.1016/j.saa.2009.09.024

Giles, R. E., Blanc, H., Cann, H. M., & Wallace, D. C. (1980). Maternal inheritance of human mitochondrial DNA. *Proceedings of the National Academy of Sciences*, *77*(11), 6715–6719. https://doi.org/10.1073/pnas.77.11.6715

Gonzalez, R. G., Haxo, R. S., & Schleich, T. (1980). Mechanism of action of polymeric aurintricarboxylic acid, a potent inhibitor of protein-nucleic acid interactions. *Biochemistry*, *19*(18), 4299–4303. https://doi.org/10.1021/bi00559a023

Gray, M. W., Burger, G., & Lang, B. F. (1999). Mitochondrial Evolution. *Science*, *283*(5407), 1476–1481. https://doi.org/10.1126/science.283.5407.1476

He, Q., Shumate, C. K., White, M. A., Molineux, I. J., & Yin, Y. W. (2013). Exonuclease of human DNA polymerase gamma disengages its strand displacement function. *Mitochondrion*, *13*(6), 592–601. https://doi.org/10.1016/j.mito.2013.08.003

Hertzberg, R. P., Caranfa, M. J., & Hecht, S. M. (1989). On the mechanism of topoisomerase I inhibition by camptothecin: Evidence for binding to an enzyme-DNA complex. *Biochemistry*, *28*(11), 4629–4638. https://doi.org/10.1021/bi00437a018

Horbay, R., & Bilyy, R. (2016). Mitochondrial dynamics during cell cycling. *Apoptosis*, *21*(12), 1327–1335. https://doi.org/10.1007/s10495-016-1295-5

Houston, D. R., & Walkinshaw, M. D. (2013). Consensus Docking: Improving the Reliability of Docking in a Virtual Screening Context. *Journal of Chemical Information and Modeling*, *53*(2), 384–390. https://doi.org/10.1021/ci300399w

Hughes, J., Rees, S., Kalindjian, S., & Philpott, K. (2011). Principles of early drug discovery. *British Journal of Pharmacology*, *162*(6), 1239–1249. https://doi.org/10.1111/j.1476-5381.2010.01127.x

Hunt, D. M., & Wagner, R. R. (1975). Inhibition by aurintricarboxylic acid and polyethylene sulfonate of RNA transcription of vesicular stomatitis virus. *Journal of Virology*, *16*(5), 1146–1153. https://doi.org/10.1128/jvi.16.5.1146-1153.1975

Huschtscha, L., Bartier, W., Ross, C., & Tattersall, M. (1996). Characteristics of cancer cell death after exposure to cytotoxic drugs in vitro. *British Journal of Cancer*, *73*(1), 54–60. https://doi.org/10.1038/bjc.1996.10

Inglese, J., Johnson, R. L., Simeonov, A., Xia, M., Zheng, W., Austin, C. P., & Auld, D. S. (2007). High-throughput screening assays for the identification of chemical probes. *Nature Chemical Biology*, *3*(8), 466–479. https://doi.org/10.1038/nchembio.2007.17

Irwin, J. J., Duan, D., Torosyan, H., Doak, A. K., Ziebart, K. T., Sterling, T., Tumanian, G., & Shoichet, B. K. (2015). An Aggregation Advisor for Ligand Discovery. *Journal of Medicinal Chemistry*, *58*(17), 7076–7087. https://doi.org/10.1021/acs.jmedchem.5b01105

Jackson, S. P., & Bartek, J. (2009). The DNA-damage response in human biology and disease. *Nature*, *461*(7267), 1071–1078. https://doi.org/10.1038/nature08467

Kaneda, H., Hayashi, J., Takahama, S., Taya, C., Lindahl, K. F., & Yonekawa, H. (1995). Elimination of paternal mitochondrial DNA in intraspecific crosses during early mouse embryogenesis. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(10), 4542–4546. https://doi.org/10.1073/pnas.92.10.4542

Karlowicz, A., Dubiel, A. B., Czerwinska, J., Bledea, A., Purzycki, P., Grzelewska, M., McAuley, R. J., Szczesny, R. J., Brzuska, G., Krol, E., Szczesny, B., & Szymanski, M. R. (2022). *In vitro* reconstitution reveals a key role of human mitochondrial EXOG in RNA primer processing. *Nucleic Acids Research*, *50*(14), 7991–8007. https://doi.org/10.1093/nar/gkac581

Karlowicz, A., Dubiel, A. B., Wyszkowska, M., Hossain, K. A., Czub, J., & Szymanski, M. R. (2025). Mitochondrial exonuclease EXOG supports DNA integrity by the removal of single-stranded DNA flaps. *Nucleic Acids Research*, *53*(5), gkaf099. https://doi.org/10.1093/nar/gkaf099

Kaufman, B. A., Durisic, N., Mativetsky, J. M., Costantino, S., Hancock, M. A., Grutter, P., & Shoubridge, E. A. (2007). The Mitochondrial Transcription Factor TFAM Coordinates the Assembly of Multiple DNA Molecules into Nucleoid-like Structures. *Molecular Biology of the Cell*, *18*(9), 3225–3236. https://doi.org/10.1091/mbc.e07-05-0404

Kaufman, B. A., & Van Houten, B. (2017). POLB: A new role of DNA polymerase beta in mitochondrial base excision repair. *DNA Repair*, *60*, A1–A5. https://doi.org/10.1016/j.dnarep.2017.11.002

Kazak, L., Reyes, A., & Holt, I. J. (2012). Minimizing the damage: Repair pathways keep mitochondrial DNA intact. *Nature Reviews Molecular Cell Biology*, *13*(10), 659–671. https://doi.org/10.1038/nrm3439

King, M. P., & Attardi, G. (1989). Human Cells Lacking mtDNA: Repopulation with Exogenous Mitochondria by Complementation. *Science*, *246*(4929), 500–503. https://doi.org/10.1126/science.2814477

Kitchen, D. B., Decornez, H., Furr, J. R., & Bajorath, J. (2004). Docking and scoring in virtual screening for drug discovery: Methods and applications. *Nature Reviews Drug Discovery*, *3*(11), 935–949. https://doi.org/10.1038/nrd1549

Koppenol, W. H., Bounds, P. L., & Dang, C. V. (2011). Otto Warburg's contributions to current concepts of cancer metabolism. *Nature Reviews Cancer*, *11*(5), 325–337. https://doi.org/10.1038/nrc3038

Kornberg, R. D. (1974). Chromatin Structure: A Repeating Unit of Histones and DNA: Chromatin structure is based on a repeating unit of eight histone molecules and about 200 DNA base pairs. *Science*, *184*(4139), 868–871. https://doi.org/10.1126/science.184.4139.868

Koster, D. A., Croquette, V., Dekker, C., Shuman, S., & Dekker, N. H. (2005). Friction and torque govern the relaxation of DNA supercoils by eukaryotic topoisomerase IB. *Nature*, *434*(7033), 671–674. https://doi.org/10.1038/nature03395

Krokan, H. E., & Bjoras, M. (2013). Base Excision Repair. *Cold Spring Harbor Perspectives in Biology*, *5*(4), a012583. https://doi.org/10.1101/cshperspect.a012583

Krumbhaar, E. B., & Krumbhaar, H. D. (1919). The Blood and Bone Marrow in Yelloe Cross Gas (Mustard Gas) Poisoning: Changes produced in the Bone Marrow of Fatal Cases. *The Journal of Medical Research*, *40*(3), 497-508.3.

Kukat, C., Davies, K. M., Wurm, C. A., Spåhr, H., Bonekamp, N. A., Kühl, I., Joos, F., Polosa, P. L., Park, C. B., Posse, V., Falkenberg, M., Jakobs, S., Kühlbrandt, W., & Larsson, N.-G. (2015). Cross-strand binding of TFAM to a single mtDNA molecule forms the mitochondrial nucleoid. *Proceedings of the National Academy of Sciences*, *112*(36), 11288–11293. https://doi.org/10.1073/pnas.1512131112

Larsson, N.-G., Garman, J. D., Oldfors, A., Barsh, G. S., & Clayton, D. A. (1996). A single mouse gene encodes the mitochondrial transcription factor A and a testis–specific nuclear HMG-box protein. *Nature Genetics*, *13*(3), 296–302. https://doi.org/10.1038/ng0796-296

Larsson, N.-G., Oldfors, A., David Garman, J., Barsh, G. S., & Clayton, D. A. (1997). Down-Regulation of Mitochondrial Transcription Factor a During Spermatogenesis in Humans. *Human Molecular Genetics*, *6*(2), 185–191. https://doi.org/10.1093/hmg/6.2.185

Li, N., Ma, Y., Yang, C., Guo, L., & Yang, X. (2005). Interaction of anticancer drug mitoxantrone with DNA analyzed by electrochemical and spectroscopic methods. *Biophysical Chemistry*, *116*(3), 199–205. https://doi.org/10.1016/j.bpc.2005.04.009

Li, Y., Han, L., Liu, Z., & Wang, R. (2014). Comparative Assessment of Scoring Functions on an Updated Benchmark: 2. Evaluation Methods and General Results. *Journal of Chemical Information and Modeling*, *54*(6), 1717–1736. https://doi.org/10.1021/ci500081m

Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L.-W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., … Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in *Phenix*. *Acta Crystallographica Section D Structural Biology*, *75*(10), 861–877. https://doi.org/10.1107/S2059798319011471

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings 1PII of original article: S0169-409X(96)00423-1. The article was originally published in Advanced Drug Delivery Reviews 23 (1997) 3–25. 1. *Advanced Drug Delivery Reviews*, *46*(1–3), 3–26. https://doi.org/10.1016/S0169-409X(00)00129-0

Liu, P., Qian, L., Sung, J.-S., De Souza-Pinto, N. C., Zheng, L., Bogenhagen, D. F., Bohr, V. A., Wilson, D. M., Shen, B., & Demple, B. (2008). Removal of Oxidative DNA Damage via FEN1-Dependent Long-Patch Base Excision Repair in Human Cell Mitochondria. *Molecular and Cellular Biology*, *28*(16), 4975–4987. https://doi.org/10.1128/MCB.00457-08

Longley, M. J., Prasad, R., Srivastava, D. K., Wilson, S. H., & Copeland, W. C. (1998). Identification of 5′-deoxyribose phosphate lyase activity in human DNA polymerase γ and its role in mitochondrial base excision repair in vitro. *Proceedings of the National Academy of Sciences*, *95*(20), 12244–12248. https://doi.org/10.1073/pnas.95.21.1224

Macarron, R., Banks, M. N., Bojanic, D., Burns, D. J., Cirovic, D. A., Garyantes, T., Green, D. V. S., Hertzberg, R. P., Janzen, W. P., Paslay, J. W., Schopfer, U., & Sittampalam, G. S. (2011). Impact of high-throughput screening in biomedical research. *Nature Reviews Drug Discovery*, *10*(3), 188–195. https://doi.org/10.1038/nrd3368

Martijn, J., Vosseberg, J., Guy, L., Offre, P., & Ettema, T. J. G. (2018). Deep mitochondrial origin outside the sampled alphaproteobacteria. *Nature*, *557*(7703), 101–105. https://doi.org/10.1038/s41586-018-0059-5

McGovern, S. L., Caselli, E., Grigorieff, N., & Shoichet, B. K. (2002). A Common Mechanism Underlying Promiscuous Inhibitors from Virtual and High-Throughput Screening. *Journal of Medicinal Chemistry*, *45*(8), 1712–1722. https://doi.org/10.1021/jm010533y

McNally, J. R., & O'Brien, P. J. (2017). Kinetic analyses of single-stranded break repair by human DNA ligase III isoforms reveal biochemical differences from DNA ligase I. *Journal of Biological Chemistry*, *292*(38), 15870–15879. https://doi.org/10.1074/jbc.M117.804625

Mercer, T. R., Neph, S., Dinger, M. E., Crawford, J., Smith, M. A., Shearwood, A.-M. J., Haugen, E., Bracken, C. P., Rackham, O., Stamatoyannopoulos, J. A., Filipovska, A., & Mattick, J. S. (2011). The Human Mitochondrial Transcriptome. *Cell*, *146*(4), 645–658. https://doi.org/10.1016/j.cell.2011.06.051

Mol, C. D., Izumi, T., Mitra, S., & Tainer, J. A. (2000). DNA-bound structures and mutants reveal abasic DNA binding by APE1 DNA repair and coordination. *Nature*, *403*(6768), 451–456. https://doi.org/10.1038/35000249

Molecular biology of the cell (6th ed). (2015). Garland science, Taylor and Francis group.

Mondal, A., Gandhi, A., Fimognari, C., Atanasov, A. G., & Bishayee, A. (2019). Alkaloids for cancer prevention and therapy: Current progress and future perspectives. *European Journal of Pharmacology*, *858*, 172472. https://doi.org/10.1016/j.ejphar.2019.172472

Montoya, J., Christianson, T., Levens, D., Rabinowitz, M., & Attardi, G. (1982). Identification of initiation sites for heavy-strand and light-strand transcription in human mitochondrial DNA. *Proceedings of the National Academy of Sciences*, *79*(23), 7195–7199. https://doi.org/10.1073/pnas.79.23.7195

Moreno-Sánchez, R., Rodríguez-Enríquez, S., Marín-Hernández, A., & Saavedra, E. (2007). Energy metabolism in tumor cells. *The FEBS Journal*, *274*(6), 1393–1418. https://doi.org/10.1111/j.1742-4658.2007.05686.x

Morham, S. G., Kluckman, K. D., Voulomanos, N., & Smithies, O. (1996). Targeted Disruption of the Mouse Topoisomerase I Gene by Camptothecin Selection. *Molecular and Cellular Biology*, *16*(12), 6804–6809. https://doi.org/10.1128/MCB.16.12.6804

Moroz, E., Matoori, S., & Leroux, J.-C. (2016). Oral delivery of macromolecular drugs: Where we are after almost 100 years of attempts. *Advanced Drug Delivery Reviews*, *101*, 108–121. https://doi.org/10.1016/j.addr.2016.01.010

Neudert, G., & Klebe, G. (2011). *DSX*: A Knowledge-Based Scoring Function for the Assessment of Protein–Ligand Complexes. *Journal of Chemical Information and Modeling*, *51*(10), 2731–2745. https://doi.org/10.1021/ci200274q

Niesen, F. H., Berglund, H., & Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nature Protocols*, *2*(9), 2212–2221. https://doi.org/10.1038/nprot.2007.321

Nitiss, J., & Wang, J. C. (1988). DNA topoisomerase-targeting antitumor drugs can be studied in yeast. *Proceedings of the National Academy of Sciences*, *85*(20), 7501–7505. https://doi.org/10.1073/pnas.85.20.7501

Nowotny, M., Gaidamakov, S. A., Ghirlando, R., Cerritelli, S. M., Crouch, R. J., & Yang, W. (2007). Structure of Human RNase H1 Complexed with an RNA/DNA Hybrid: Insight into HIV Reverse Transcription. *Molecular Cell*, *28*(2), 264–276. https://doi.org/10.1016/j.molcel.2007.08.015

Osellame, L. D., Blacker, T. S., & Duchen, M. R. (2012). Cellular and molecular mechanisms of mitochondrial function. *Best Practice & Research Clinical Endocrinology & Metabolism*, *26*(6), 711–723. https://doi.org/10.1016/j.beem.2012.05.003

Ozawa, T. (1997). Genetic and functional changes in mitochondria associated with aging. *Physiological Reviews*, *77*(2), 425–464. https://doi.org/10.1152/physrev.1997.77.2.425

Palade, G. E. (1952). The fine structure of mitochondria. *The Anatomical Record*, *114*(3), 427–451. https://doi.org/10.1002/ar.1091140304

Palade, G. E. (1953). AN ELECTRON MICROSCOPE STUDY OF THE MITOCHONDRIAL STRUCTURE. *Journal of Histochemistry & Cytochemistry*, *1*(4), 188–211. https://doi.org/10.1177/1.4.188

Pardridge, W. M. (1995). Transport of small molecules through the blood-brain barrier: Biology and methodology. *Advanced Drug Delivery Reviews*, *15*(1–3), 5–36.

Patel, J., Baptiste, B. A., Kim, E., Hussain, M., Croteau, D. L., & Bohr, V. A. (2020). DNA damage and mitochondria in cancer and aging. *Carcinogenesis*, *41*(12), 1625–1634. https://doi.org/10.1093/carcin/bgaa114

Petropoulos, M., Karamichali, A., Rossetti, G. G., Freudenmann, A., Iacovino, L. G., Dionellis, V. S., Sotiriou, S. K., & Halazonetis, T. D. (2024). Transcription–replication conflicts underlie sensitivity to PARP inhibitors. *Nature*, *628*(8007), 433–441. https://doi.org/10.1038/s41586-024-07217-2

Pinz, K. G., & Bogenhagen, D. F. (2000). Characterization of a Catalytically Slow AP Lyase Activity in DNA Polymerase γ and Other Family A DNA Polymerases. *Journal of Biological Chemistry*, *275*(17), 12509–12514. https://doi.org/10.1074/jbc.275.17.12509

Prakash, A., & Doublié, S. (2015). Base Excision Repair in the Mitochondria. *Journal of Cellular Biochemistry*, *116*(8), 1490–1499. https://doi.org/10.1002/jcb.25103

Prasad, R., Çağlayan, M., Dai, D.-P., Nadalutti, C. A., Zhao, M.-L., Gassman, N. R., Janoshazi, A. K., Stefanick, D. F., Horton, J. K., Krasich, R., Longley, M. J., Copeland, W. C., Griffith, J. D., & Wilson, S. H. (2017). DNA polymerase β: A missing link of the base excision repair machinery in mammalian mitochondria. *DNA Repair*, *60*, 77–88. https://doi.org/10.1016/j.dnarep.2017.10.011

Rath, S., Sharma, R., Gupta, R., Ast, T., Chan, C., Durham, T. J., Goodman, R. P., Grabarek, Z., Haas, M. E., Hung, W. H. W., Joshi, P. R., Jourdain, A. A., Kim, S. H., Kotrys, A. V., Lam, S. S., McCoy, J. G., Meisel, J. D., Miranda, M., Panda, A., … Mootha, V. K. (2021). MitoCarta3.0: An updated mitochondrial proteome now with sub-organelle localization and pathway annotations. *Nucleic Acids Research*, *49*(D1), D1541–D1547. https://doi.org/10.1093/nar/gkaa1011

Reyes, A., Gissi, C., Pesole, G., & Saccone, C. (1998). Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. *Molecular Biology and Evolution*, *15*(8), 957–966. https://doi.org/10.1093/oxfordjournals.molbev.a026011

Richter, C., Park, J. W., & Ames, B. N. (1988). Normal oxidative damage to mitochondrial and nuclear DNA is extensive. *Proceedings of the National Academy of Sciences*, *85*(17), 6465–6467. https://doi.org/10.1073/pnas.85.17.6465

Roger, A. J., Muñoz-Gómez, S. A., & Kamikawa, R. (2017). The Origin and Diversification of Mitochondria. *Current Biology*, *27*(21), R1177–R1192. https://doi.org/10.1016/j.cub.2017.09.015

Roskoski, R. (2020). Properties of FDA-approved small molecule protein kinase inhibitors: A 2020 update. *Pharmacological Research*, *152*, 104609. https://doi.org/10.1016/j.phrs.2019.104609

Sampath, H. (2014). Oxidative DNA damage in disease—Insights gained from base excision repair glycosylase-deficient mouse models. *Environmental and Molecular Mutagenesis*, *55*(9), 689–703. https://doi.org/10.1002/em.21886

Sampath, H., McCullough, A. K., & Lloyd, R. S. (2012). Regulation of DNA glycosylases and their role in limiting disease. *Free Radical Research*, *46*(4), 460–478. https://doi.org/10.3109/10715762.2012.655730

Schencking, I., Schäfer, E. M., Scanlan, J. H. W., Wenzel, B. M., Emmerich, R. E., Steinmetzer, T., Diederich, W. E., Schlitzer, M., & Hartmann, R. K. (2021). RNase P Inhibitors Identified as Aggregators. *Antimicrobial Agents and Chemotherapy*, *65*(8), e00300-21. https://doi.org/10.1128/AAC.00300-21

Seidler, J., McGovern, S. L., Doman, T. N., & Shoichet, B. K. (2003). Identification and Prediction of Promiscuous Aggregating Inhibitors among Known Drugs. *Journal of Medicinal Chemistry*, *46*(21), 4477–4486. https://doi.org/10.1021/jm030191r

Shadel, G. S., & Clayton, D. A. (1997). MITOCHONDRIAL DNA MAINTENANCE IN VERTEBRATES. *Annual Review of Biochemistry*, *66*(1), 409–435. https://doi.org/10.1146/annurev.biochem.66.1.409

Shenkenberg, T. D., & Von Hoff, D. D. (1986). Mitoxantrone: A New Anticancer Drug with Significant Clinical Activity. *Annals of Internal Medicine*, *105*(1), 67–81. https://doi.org/10.7326/0003-4819-105-1-67

Shoichet, B. K. (2004). Virtual screening of chemical libraries. *Nature*, *432*(7019), 862–865. https://doi.org/10.1038/nature03197

Simeonov, A., Kulkarni, A., Dorjsuren, D., Jadhav, A., Shen, M., McNeill, D. R., Austin, C. P., & Wilson, D. M. (2009). Identification and Characterization of Inhibitors of Human Apurinic/apyrimidinic Endonuclease APE1. *PLoS ONE*, *4*(6), e5740. https://doi.org/10.1371/journal.pone.0005740

Simsek, D., Furda, A., Gao, Y., Artus, J., Brunet, E., Hadjantonakis, A.-K., Van Houten, B., Shuman, S., McKinnon, P. J., & Jasin, M. (2011). Crucial role for DNA ligase III in mitochondria but not in Xrcc1-dependent repair. *Nature*, *471*(7337), 245–248. https://doi.org/10.1038/nature09794

Sjöstrand, F. S. (1953). Electron Microscopy of Mitochondria and Cytoplasmic Double Membranes: Ultra-Structure of Rod-shaped Mitochondria. *Nature*, *171*(4340), 30–31. https://doi.org/10.1038/171030a0

Somuncu, B., Ekmekcioglu, A., Antmen, F. M., Ertuzun, T., Deniz, E., Keskin, N., Park, J., Yazici, I. E., Simsek, B., Erman, B., Yin, W., Erman, B., & Muftuoglu, M. (2022). Targeting mitochondrial DNA polymerase gamma for selective inhibition of MLH1 deficient colon cancer growth. *PLOS ONE*, *17*(6), e0268391. https://doi.org/10.1371/journal.pone.0268391

Song, S., Pursell, Z. F., Copeland, W. C., Longley, M. J., Kunkel, T. A., & Mathews, C. K. (2005). DNA precursor asymmetries in mammalian tissue mitochondria and possible contribution to mutagenesis through reduced replication fidelity. *Proceedings of the National Academy of Sciences*, *102*(14), 4990–4995. https://doi.org/10.1073/pnas.0500253102

Spiriti, J., Binder, J. K., Levitus, M., & Van Der Vaart, A. (2011). Cy3-DNA Stacking Interactions Strongly Depend on the Identity of the Terminal Basepair. *Biophysical Journal*, *100*(4), 1049–1057. https://doi.org/10.1016/j.bpj.2011.01.014

Stein, A. & Sia, E. A. (2017). Mitochondrial DNA repair and damage tolerance. *Frontiers in Bioscience*, *22*(5), 920–943. https://doi.org/10.2741/4525

Sutovsky, P., Moreno, R. D., Ramalho-Santos, J., Dominko, T., Simerly, C., & Schatten, G. (1999). Ubiquitin tag for sperm mitochondria. *Nature*, *402*(6760), 371–372. https://doi.org/10.1038/46466

Sykora, P., Kanno, S., Akbari, M., Kulikowicz, T., Baptiste, B. A., Leandro, G. S., Lu, H., Tian, J., May, A., Becker, K. A., Croteau, D. L., Wilson, D. M., Sobol, R. W., Yasui, A., & Bohr, V. A. (2017). DNA Polymerase Beta Participates in Mitochondrial DNA Repair. *Molecular and Cellular Biology*, *37*(16), e00237-17. https://doi.org/10.1128/MCB.00237-17

Szczesny, B., Brunyanszki, A., Olah, G., Mitra, S., & Szabo, C. (2014). Opposing roles of mitochondrial and nuclear PARP1 in the regulation of mitochondrial and nuclear DNA integrity: Implications for the regulation of mitochondrial function. *Nucleic Acids Research*, *42*(21), 13161–13173. https://doi.org/10.1093/nar/gku1089

Szczesny, B., Marcatti, M., Zatarain, J. R., Druzhyna, N., Wiktorowicz, J. E., Nagy, P., Hellmich, M. R., & Szabo, C. (2016). Inhibition of hydrogen sulfide biosynthesis sensitizes lung adenocarcinoma to chemotherapeutic drugs by inhibiting mitochondrial DNA repair and suppressing cellular bioenergetics. *Scientific Reports*, *6*(1), 36125. https://doi.org/10.1038/srep36125

Szczesny, R. J., Hejnowicz, M. S., Steczkiewicz, K., Muszewska, A., Borowski, L. S., Ginalski, K., & Dziembowski, A. (2013). Identification of a novel human mitochondrial endo-/exonuclease Ddk1/c20orf72 necessary for maintenance of proper 7S DNA levels. *Nucleic Acids Research*, *41*(5), 3144–3161. https://doi.org/10.1093/nar/gkt029

Szymanski, M. R., Karlowicz, A., Herrmann, G. K., Cen, Y., & Yin, Y. W. (2022). Human EXOG Possesses Strong AP Hydrolysis Activity: Implication on Mitochondrial DNA Base Excision Repair. *Journal of the American Chemical Society*, *144*(51), 23543–23550. https://doi.org/10.1021/jacs.2c10558

Szymanski, M. R., Yu, W., Gmyrek, A. M., White, M. A., Molineux, I. J., Lee, J. C., & Yin, Y. W. (2017). A domain in human EXOG converts apoptotic endonuclease to DNA-repair exonuclease. *Nature Communications*, *8*(1), 14959. https://doi.org/10.1038/ncomms14959

Tan, A. S., Baty, J. W., Dong, L.-F., Bezawork-Geleta, A., Endaya, B., Goodwin, J., Bajzikova, M., Kovarova, J., Peterka, M., Yan, B., Pesdar, E. A., Sobol, M., Filimonenko, A., Stuart, S., Vondrusova, M., Kluckova, K., Sachaphibulkij, K., Rohlena, J., Hozak, P., … Berridge, M. V. (2015). Mitochondrial Genome Acquisition Restores Respiratory Function and Tumorigenic Potential of Cancer Cells without Mitochondrial DNA. *Cell Metabolism*, *21*(1), 81–94. https://doi.org/10.1016/j.cmet.2014.12.003

Tanaka, M., & Ozawa, T. (1994). Strand Asymmetry in Human Mitochondrial DNA Mutations. *Genomics*, *22*(2), 327–335. https://doi.org/10.1006/geno.1994.1391

Tann, A. W., Boldogh, I., Meiss, G., Qian, W., Van Houten, B., Mitra, S., & Szczesny, B. (2011). Apoptosis induced by persistent single-strand breaks in mitochondrial genome: Critical role of EXOG (5'-EXO/endonuclease) in their repair. *The Journal of Biological Chemistry*, *286*(37), 31975–31983. https://doi.org/10.1074/jbc.M110.215715

Tapper, D. P., & Clayton, D. A. (1981). Mechanism of replication of human mitochondrial DNA. Localization of the 5' ends of nascent daughter strands. *The Journal of Biological Chemistry*, *256*(10), 5109–5115.

Trott, O., & Olson, A. J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, *31*(2), 455–461. https://doi.org/10.1002/jcc.21334

Uhler, J. P., & Falkenberg, M. (2015). Primer removal during mammalian mitochondrial DNA replication. *DNA Repair*, *34*, 28–38. https://doi.org/10.1016/j.dnarep.2015.07.003

Uhler, J. P., Thörn, C., Nicholls, T. J., Matic, S., Milenkovic, D., Gustafsson, C. M., & Falkenberg, M. (2016). MGME1 processes flaps into ligatable nicks in concert with DNA polymerase γ during mtDNA replication. *Nucleic Acids Research*, *44*(12), 5861–5871. https://doi.org/10.1093/nar/gkw468

Van Houten, B., Hunter, S. E., & Meyer, J. N. (2016). Mitochondrial DNA damage induced autophagy, cell death, and disease. *Frontiers in Bioscience*, *21*(1), 42–54. https://doi.org/10.2741/4375

Viviani, L. G., Piccirillo, E., Cheffer, A., De Rezende, L., Ulrich, H., Carmona-Ribeiro, A. M., & Amaral, A. T. -do. (2018). Be Aware of Aggregators in the Search for Potential Human ecto-5′-Nucleotidase Inhibitors. *Molecules*, *23*(8), 1876. https://doi.org/10.3390/molecules23081876

Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T., & Bricogne, G. (2011). Data processing and analysis with the *autoPROC* toolbox. *Acta Crystallographica Section D Biological Crystallography*, *67*(4), 293–302. https://doi.org/10.1107/S0907444911007773

Wall, M. E., & Wani, M. C. (1995). Camptothecin and taxol: Discovery to clinic—thirteenth Bruce F. Cain Memorial Award Lecture. *Cancer Research*, *55*(4), 753–760.

Wallace, D. C. (2012). Mitochondria and cancer. *Nature Reviews Cancer*, *12*(10), 685–698. https://doi.org/10.1038/nrc3365

Wang, R., & Wang, S. (2001). How Does Consensus Scoring Work for Virtual Library Screening? An Idealized Computer Experiment. *Journal of Chemical Information and Computer Sciences*, *41*(5), 1422–1426. https://doi.org/10.1021/ci010025x

Warburg, O. (1956). On respiratory impairment in cancer cells. *Science*, *124*(3215), 269–270.

Weinberg, S. E., & Chandel, N. S. (2015). Targeting mitochondria metabolism for cancer therapy. *Nature Chemical Biology*, *11*(1), 9–15. https://doi.org/10.1038/nchembio.1712

Wong, K. M., Tai, H. K., & Siu, S. W. I. (2021). GWOVina: A grey wolf optimization approach to rigid and flexible receptor docking. *Chemical Biology & Drug Design*, *97*(1), 97–110. https://doi.org/10.1111/cbdd.13764

Wójcikowski, M., Ballester, P. J., & Siedlecki, P. (2017). Performance of machine-learning scoring functions in structure-based virtual screening. *Scientific Reports*, *7*(1), 46710. https://doi.org/10.1038/srep46710

Wu, C.-C., Lin, J. L. J., Yang-Yen, H.-F., & Yuan, H. S. (2019). A unique exonuclease ExoG cleaves between RNA and DNA in mitochondrial DNA replication. *Nucleic Acids Research*, *47*(10), 5405–5419. https://doi.org/10.1093/nar/gkz241

Wu, P., Nielsen, T. E., & Clausen, M. H. (2016). Small-molecule kinase inhibitors: An analysis of FDA-approved drugs. *Drug Discovery Today*, *21*(1), 5–10. https://doi.org/10.1016/j.drudis.2015.07.008

Yakes, F. M., & Van Houten, B. (1997). Mitochondrial DNA damage is more extensive and persists longer than nuclear DNA damage in human cells following oxidative stress. *Proceedings of the National Academy of Sciences*, *94*(2), 514–519. https://doi.org/10.1073/pnas.94.2.514

Young, M. J. (2017). Off-Target Effects of Drugs that Disrupt Human Mitochondrial DNA Maintenance. *Frontiers in Molecular Biosciences*, *4*, 74. https://doi.org/10.3389/fmolb.2017.00074

Zhang, C. X., Chen, A. D., Gettel, N. J., & Hsieh, T. (2000). Essential functions of DNA topoisomerase I in Drosophila melanogaster. *Developmental Biology*, *222*(1), 27–40. https://doi.org/10.1006/dbio.2000.9704

Zhang, J.-H., Chung, T. D. Y., & Oldenburg, K. R. (1999). A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *Journal of Biomolecular Screening*, *4*(2), 67–73.

Zheng, L., Dai, H., Zhou, M., Li, X., Liu, C., Guo, Z., Wu, X., Wu, J., Wang, C., Zhong, J., Huang, Q., Garcia-Aguilar, J., Pfeifer, G. P., & Shen, B. (2012). Polyploid cells rewire DNA damage response networks to overcome replication stress-induced barriers for tumour progression. *Nature Communications*, *3*(1), 815. https://doi.org/10.1038/ncomms1825