

DISS. ETH No. 26484

Gene-targeted identification of DNA lesions to unravel mechanisms of mutagenesis

A thesis submitted to attain the degree of

DOCTOR OF SCIENCES of ETH Zurich

(Dr. sc. ETH Zurich)

presented by

CLAUDIA ALOISI

M. Sc. in Chemistry, Università di Pisa, Italy

Born on 24.12.1990

Citizen of Italy

accepted on the recommendation of

Prof. Dr. Shana J. Sturla

Prof. Dr. Orlando D. Schärer

Prof. Dr. Hanspeter Nägeli

Prof. Dr. Martin Jinek

Dr. Hailey Gahlon

2019

Ogni “così fu” è un frammento, un enigma, una casualità orrida
fin quando la volontà che crea non dica anche: “ma così volli che fosse!”
finché la volontà che crea non dica anche “ma io così voglio! così vorrò!”

Also sprach Zarathustra, Friedrich Nietzsche

Table of Contents

Acknowledgements	1
Abstract	3
Sommario	5
Abbreviations	7
Chapter 1 - Introduction: DNA Adduct-Directed Synthetic Nucleosides	11
1.1 Abstract.....	13
1.2 Introduction	14
1.3 Synthetic Nucleosides that Pair specifically with <i>O</i> ⁶ -Alkylguanine adducts.....	15
1.3.1 DNA Adduct Hybridization Nanoprobes	18
1.4 Synthetic nucleosides as probes to Investigate Translesion DNA synthesis	19
1.5 DNA Damage Detection and Amplification with Synthetic Nucleosides.....	251
1.6 Conclusion.....	285
1.7 Overview of thesis work	285
1.8 References	28
Chapter 2: A Gene-Targeted Polymerase Amplification Strategy to Identify <i>O</i>⁶-Methyl-Guanine Damage	33
2.1 Abstract.....	35
2.2 Introduction, results and discussion.....	35
2.3 Supplementary Information (Experimental Section)	42
2.4 References	47
Chapter 3: Sequence-specific Quantitation of Mutagenic DNA damage via Polymerase Amplification with an Artificial Nucleotide	49
3.1 Abstract.....	51
3.2 Introduction	51

3.3 Results and Discussion.....	53
3.3.1 Artificial nucleotide analogue is specifically incorporated opposite <i>O</i> ⁶ -CMG DNA adduct	53
3.3.2 Structural basis for ExBenziTP selective incorporation opposite <i>O</i> ⁶ -CMG by KlenTaq M747K.....	55
3.3.3 ExBenzi is required for efficient amplification of a damaged DNA	57
3.3.4 Quantification of ExBenzi nucleoside by mass spectrometry is diagnostic of <i>O</i> ⁶ -CMG levels in DNA	59
3.4 Conclusion.....	61
3.5 Supplementary Information (Experimental Section)	62
3.6 References	75
Chapter 4: MGMT and NER alleviate <i>O</i>⁶-Carboxymethyl-Guanine DNA Adduct Mutagenicity in Cells	79
4.1 Abstract.....	81
4.2 Introduction.....	81
4.3 Results.....	83
4.3.1 Reduced MGMT function sensitizes cells to carboxymethylating agent azaserine	83
4.3.2 Azaserine is more mutagenic in cells when MGMT is inhibited	85
4.3.3 Sequencing data for <i>hprt</i> gene in MGMT-depleted V79-4 cells	86
4.3.4 <i>O</i> ⁶ -CMG adduct levels induced by azaserine in MGMT-depleted V79-4 cells.....	87
4.3.5 Cells deficient in NER are more sensitive to carboxymethylating agent azaserine.....	88
4.3.6 Cells deficient in NER are not sensitized to carboxymethylation when depleted of MGMT	89
4.3.7 Azaserine is more mutagenic in cells deficient in NER	89
4.3.8 Azaserine mutagenicity is significantly higher when both NER and MGMT are depleted in cells.....	90

4.3.9 Sequencing data for <i>hprt</i> gene in NER CHO cells.....	91
4.3.10 <i>O</i> ⁶ -CMG adduct levels induced by azaserine in MGMT-depleted V79-4 cells.....	92
4.3.11 <i>In vitro</i> NER biochemical assays on <i>O</i> ⁶ -CMG DNA are contradictory regarding the role of NER in <i>O</i> ⁶ -CMG recognition and repair	92
4.4 Discussion and Future studies	95
4.5 Material and Methods.....	97
4.6 Supplementary Figures.....	102
4.7 References	107
Chapter 5: Summary and outlook	113
5.1 Artificial nucleotides as markers of presence, location and amount of mutagenic DNA damage	113
5.2 Future directions of artificial nucleotide-based strategies for the detection of rare DNA damage events.....	114
5.3 References	116
Chapter 6 - Appendix A: Modulation of Cytotoxicity by Transcription-Coupled Nucleotide Excision Repair Is Independent of the Requirement for Bioactivation of Acylfulvene	119
6.1 Abstract.....	121
6.2 Introduction.....	121
6.3 Material and Methods.....	123
6.4 Results.....	126
6.4.1 Cytotoxicity of AF and iso-MO in human fibroblasts with different DNA repair proficiencies.....	126
6.4.2 AF and its activated analog iso-Mo inhibit RNA synthesis.....	127
6.4.3 PTGR1 levels in WT, XP and CSB cells	128
6.4.4 Transient depletion of certain NER factors in cells can enhance AF cytotoxicity.....	129
6.5 Discussion	130

6.6 Conclusion.....	133
6.7 Supplementary Information	133
6.8 References	136
Chapter 7 - Appendix B: Mechanism of RNA polymerase II stalling by DNA alkylation	139
7.1 Abstract.....	141
7.2 Introduction	141
7.3 Results.....	144
7.3.1 DNA Alkylation Impairs DNA Synthesis in Cells	144
7.3.2 Pol II Transcription Can Stall at Minor-Groove Alkylation Adducts	144
7.3.3 Pol II is Tolerant to Smaller Modifications but Errors Arise.....	146
7.3.4 DNA Alkylation Alters Pol II Incorporation and Extension Efficiency	147
7.3.5 Structural Investigation of Pol II Stalling by DNA Alkylation.....	149
7.4 Discussion	151
7.5 Materials and Methods	152
7.6 Supplementary Information	155
7.7 References	158
Curriculum vitae	163

Acknowledgements

I would like to thank my PhD advisor *Prof. Dr. Shana Sturla* for her guidance and support. It has been a great experience to have the opportunity to work with you, professionally and personally. You gave me the freedom and the trust that allowed me to grow as an independent scientist, and you always listened to me whatever I had to say. It has been an amazing learning curve to carry out my doctoral studies in your group, and I have learnt from you so many important aspects of what it takes to be in academia that rarely I could have learnt elsewhere. As we came to share opinions and anticipate each other thoughts, it seems like it is time for me to go soon, but I will do so with a heavy heart. Thanks for all!

I thank the members of my PhD Committee. *Prof. Dr. Hanspeter Nägeli* and *Prof. Dr. Martin Jinek*, thanks for all your critical and insightful inputs. It was a pleasure to meet every year and receive your valuable support.

My gratitude goes to *Prof. Dr. Orlando Schärer*. Your patient help and incredible knowledge have accompanied me since I started in Shana's group. Organizing a course and writing grant proposals was only the beginning of a fruitful, exciting collaboration. I am grateful for having had the priceless opportunity to research in your lab in South Korea: it was one of the most remarkable experience of my PhD. It was great to discuss science with you and be your student for a while. Thanks for your kind hospitality, and for all your support in my next career steps.

A very special thanks goes to *Dr. Hailey Gahlon*. I wish I had met you down the line of my doctoral studies: working with you has been an absolute pleasure, and the most enjoyable period of my PhD. Your immediate, knowledgeable feedback and your discreet supervision made a massive difference in the most gel-challenging days. Your positive attitude has taught me and encouraged me a lot. Thanks a lot for being also so supportive of my next career path.

Thank you, *Prof. Dr. Maureen McKeague*, for being the kindest, best colleague one could ever meet. I wonder how you could do some many smart things and yet never refuse once to help the others, with a smile. For a while, I jokingly regretted that I did not have the luck of being your supervisee and learn all the tricks from you. Little did I know, I would have found instead a caring, supportive and amazing friend in you, and an inspiring person to look up to. I miss you a lot and a wish we could still have our evening teas and strolls from time to time.

A big thanks goes to *Dr. Marianna Stamou*, with whom I started this adventure in Zurich. Thanks for being such a helping, caring person, always ready to brainstorm with me and give me smart advice, even if we worked on different projects. I am grateful that I met you, and that we could share concerns and overcome them healthily. You have been such a support and friend.

Now, how to thank all the amazing members of the group? For all the laughter, the shared frustrations, the scientific and less scientific brainstorming, the organization of group events and retreats. For all the gym sessions, coffee sessions, for the kindergarten-fun and all the pranks. And then again I think of Christmas (and basement) parties, singing and playing cello in the lab, cold swims and late beers at the lake (oh well late for whoever was left). How to thank each one of you for the last four years? The risk is that I could go on for pages. But there is no need to name anyone, you guys know. The laughs and sorrows that some of us shared will always be shared, and in many ways will always bind us. After all, taking a PhD is like going to war!! But it is also a beautiful journey.

To the group, thanks for all the support, for listening, for helping out. I found some good friends in you, no matter what the age, no matter what the background or the individual differences.

Thanks to all the people I directly researched with: I had lots of positive interactions, it was absolutely great doing science together, and at times lots of fun too.

Thanks to the amazing students and supervisee I had the pleasure to work with and share experiences with!! It has been a real honor to have seen you grow into outstanding scientists. Each one of you has so much strength, so much potential, and you are doing the best use of it: keep it up!

Then, I need to keep my word on something I have been joking about: I am grateful to caffeine and to the ASVZ gym association for all the positive chemistry of neurotransmitters that kept me going even during the most difficult times!

Finally, thanks to friends and to my family, for their unreasonable faith in me and unconditional love. Most importantly, to the family that has nothing to do with genes.. It was no dream. It never was :) and I am so embarrassingly lucky!

Abstract

The integrity of our genomic material is continuously threatened by exposure to endogenous and exogenous DNA-damaging chemicals. The resulting chemical adducts to DNA can initiate adverse biological consequences, including cell death and DNA mutations that potentially lead to cancer. Our capacity to relate mutagenesis with specific DNA adducts is limited by the lack of strategies to measure the early event of DNA damage within genomic loci. Furthermore, inadequate knowledge regarding the cellular processes that handle the DNA adducts hinders our understanding of mutagenicity-driving factors in response to DNA damage in cells. The work described in this thesis concerns addressing both the gap in technology and the gap in knowledge with regard to a type of mutagenic, biologically relevant DNA adducts. The main achievements of this work were the development of strategies for the detection of the DNA adducts in a target DNA sequence, and the identification of cellular mechanisms that alleviate the mutagenicity induced by the DNA adducts.

In *Chapter 1*, topics further discussed in this thesis are introduced. DNA adduct-directed artificial nucleotides have been developed as a basis for interrogating damaged DNA in duplex hybridization or polymerase-mediated synthesis contexts. Strategies based on artificial nucleotides are employed for the work of *Chapters 2* and *3*. An *overview* of the studies presented in this thesis is given at the end of the chapter.

In *Chapter 2*, a strategy to detect the mutagenic O^6 -methylguanine DNA adduct in the DNA sequence of a cancer-relevant gene is developed. A DNA primer containing an artificial nucleoside analogue that stabilizes O^6 -methylguanine in the duplex allowed replication of DNA by an engineered DNA polymerase only when the artificial nucleotide was paired with O^6 -methylguanine. The artificial nucleotide-modified primer effectively marked presence and position of O^6 -methylguanine in the sequence. Furthermore, thanks to high specificity of replication, the modified DNA primer could be extended and amplified only when O^6 -methylguanine damaged DNA was present, and signal was obtained that linearly increased with the amount of O^6 -methylguanine in the sample. The presented strategy represents the first instance of polymerase amplification-based O^6 -methylguanine detection to single base resolution. This approach could be used to detect O^6 -methylguanine in other cancer-relevant gene sequences.

In *Chapter 3*, the scope and performance of artificial nucleotide-based detection of DNA adducts reported in the previous chapter is expanded by the introduction of an analytical measurement that allows for the quantification of the DNA adducts. A heterocyclic imidic nucleotide triphosphate analogue is synthesized for the detection and quantification of the mutagenic O^6 -carboxymethylguanine (O^6 -CMG), which has been associated to cancer development linked to meat consumption. The nucleotide analogue was incorporated selectively opposite O^6 -CMG over undamaged G by an engineered DNA polymerase, and was required for replication of DNA past O^6 -

CMG in the template strand. Molecular modelling studies identified base pairing and interaction with the polymerase as driving factors for the selective incorporation. The high template-selectivity meant that a DNA primer complementary to the target DNA containing O^6 -CMG could be amplified. The amplified DNA contained the artificial nucleotide and marked presence and location of O^6 -carboxymethylguanine. By developing an analytical method for the quantification of the incorporated nucleotide, trace amounts of O^6 -CMG were detected, and the amount of detected artificial nucleoside increased linearly with the initial amount of O^6 -carboxymethylguanine. The amplified quantitative signal obtained in this work for mutagenic DNA damage in a DNA sequence lays the foundation to the study of the cause-effect relationship between DNA damage and mutations.

In *Chapter 4*, we investigated the impact of DNA repair on the cellular phenotype induced by O^6 -CMG, to address a pervasive gap of knowledge regarding repair and mutagenicity of O^6 -CMG. Based on few previous studies performed on O^6 -CMG and on structurally similar DNA adducts, two important cellular DNA repair pathways were investigated for their capacity to alleviate toxicity and mutagenicity of a O^6 -CMG-inducing drug. Cells impaired in those repair pathways mutated at a higher rate than repair-proficient cells, likely due to O^6 -CMG persistence and accumulation. Adduct level measurement and characterization of the mutational spectra are underway and expected to provide further insight regarding the impact and contribution of those repair pathways. Despite pending experimentation, the findings in this work answered a longstanding question and demonstrated that O^6 -CMG mutagenicity is significantly increased upon inadequate DNA repair.

In *Chapter 5*, findings of the work of this thesis are summarized. Future opportunities and challenges are critically discussed and recommendations for future directions are presented.

Finally, in *Chapters 6 and 7*, the cellular DNA repair of DNA damage induced by the cancer drug acylfulvene was identified, and repair initiation was characterized using a combination of molecular biology, biochemistry and crystallography studies. Cellular toxicity is on the opposite spectrum of the biological consequences induced by DNA damage discussed so far, i.e. mutagenesis. DNA damage-induced cellular toxicity is exploited in chemotherapy to induce cell death of fast-dividing cancer cells. However, cellular processes such as insufficient bioactivation and DNA repair can decrease susceptibility of cancer cells to the action of DNA-damaging chemotherapeutic drugs. In the last chapters of this thesis, the repair of DNA damage induced by acylfulvene is studied by interrogating the contribution and the mode of action of an important DNA repair pathway. The identification of a potential, pivotal cellular mechanism of drug resistance achieved in this work means that therapeutic strategies to tackle mechanisms of resistance could be devised, such as the combined treatment with specific repair inhibitors.

Sommario

L'integrità del nostro materiale genetico è continuamente minacciata dall'esposizione a sostanze chimiche endogene ed esogene. Gli addotti chimici al DNA che si generano possono innescare condizioni biologiche avverse, inclusa morte cellulare e mutazioni del DNA che potenzialmente portano al cancro. La nostra capacità di correlare la mutagenesi alla presenza di specifici addotti al DNA è limitata dalla carenza di approcci in grado di misurare i danni al DNA in definiti loci genici. Inoltre, un'inadeguata conoscenza dei processi cellulari coinvolti nella gestione dei danni al DNA ostacola la nostra comprensione dei fattori che guidano la mutagenicità nelle cellule in risposta ai danni al DNA. Il lavoro descritto in questa tesi riguarda l'affrontare sia il divario nella tecnologia che il divario nella conoscenza degli addotti del DNA con carattere mutagenico che hanno rilevanza biologica. I principali risultati di questo lavoro includono lo sviluppo di strategie per la rilevazione di addotti mutageni al DNA in una sequenza bersaglio del DNA e l'identificazione di meccanismi cellulari che alleviano la mutagenicità indotta da tali addotti al DNA.

Nel *Capitolo 1* vengono introdotti alcuni argomenti discussi in questa tesi. Nucleotidi artificiali sono stati sviluppati sulla base di addotti al DNA per studiare l'ibridizzazione nella doppia elica di tali danni o per studiare sintesi del DNA mediata dalla polimerasi. Strategie basate su nucleotidi artificiali sono impiegate per il lavoro dei *Capitoli 2* e *3*. Una panoramica degli studi presentati in questa tesi è fornita alla fine del capitolo.

Nel *Capitolo 2* viene sviluppata una strategia per rilevare l'addotto mutagenico O^6 -metilguanina nella sequenza del DNA di un gene rilevante per il cancro. Un iniziatore di DNA contenente un analogo nucleosidico artificiale che stabilizza l' O^6 -metilguanina ha consentito la replicazione del DNA da parte di una DNA polimerasi ingegnerizzata solo quando il nucleotide artificiale era accoppiato con O^6 -metilguanina. L'iniziatore contenente il nucleotide artificiale era in grado di indicare la presenza e la posizione di O^6 -metilguanina nella sequenza generata. Inoltre, grazie all'elevata specificità della replicazione, l'iniziatore di DNA modificato poteva essere amplificato solo in presenza di DNA danneggiato con O^6 -metilguanina ed il segnale misurato aumentava linearmente con la concentrazione di O^6 -metilguanina nel campione. La strategia presentata rappresenta il primo esempio del rilevamento di O^6 -metilguanina basata sull'amplificazione della polimerasi con risoluzione alla singola base e potrebbe essere estesa ad altre sequenze geniche di interesse per il cancro.

Nel *capitolo 3*, l'ambito d'uso e le prestazioni della rilevazione degli addotti al DNA basata su nucleotidi artificiali discussa nel capitolo precedente e altrove è ampliata con l'introduzione di una misurazione analitica che consente la quantificazione dell'addotto di DNA. Un analogo imidico eterociclico di trifosfato nucleotidico è sintetizzato per il rilevamento e la quantificazione della mutagenica O^6 -carbossimetilguanina (O^6 -CMG), che è stata associata allo sviluppo del cancro legato al consumo di carne. L'analogo nucleotide è stato incorporato in una sequenza di DNA da una DNA polimerasi

ingegnerizzata opposto a O^6 -CMG ma non G in maniera selettiva, ed era necessario per la replicazione del DNA dopo la presenza di O^6 -CMG. Studi di modellistica molecolare hanno identificato l'accoppiamento di base e l'interazione con la polimerasi come fattori trainante dell'incorporazione selettiva. L'elevata selettività nei riguardi del DNA stampo ha permesso di amplificare un iniziatore di DNA complementare al DNA target contenente O^6 -CMG. Il DNA amplificato conteneva il nucleotide artificiale e questo indicava la presenza e la posizione di O^6 -CMG. Grazie ad un metodo analitico sviluppato per la quantificazione del nucleotide incorporato, sono state rilevate tracce di O^6 -CMG e la quantità di nucleoside artificiale rilevata aumentava linearmente con la quantità iniziale di O^6 -CMG. Il segnale quantitativo ottenuto per il danno del DNA mutageno in una sequenza di DNA ottenuta con questo lavoro pone le basi per lo studio della relazione causa-effetto tra danno al DNA e mutazioni.

Nel capitolo 4 viene affrontato la profonda mancanza di conoscenze riguardi la riparazione cellulare di O^6 -CMG. Sulla base della letteratura precedente sull' O^6 -CMG e su addotti strutturalmente simili, sono state investigate due importanti vie di riparazione del DNA cellulare per la loro capacità di alleviare la tossicità e la mutagenicità di un farmaco che induce l' O^6 -CMG nelle cellule. Le cellule che erano compromesse nella loro capacità di riparare DNA mutavano ad un ritmo più elevato rispetto alle cellule con capacità di riparazione intatte, probabilmente a causa della persistenza e dell'accumulo di O^6 -CMG. La misurazione del livello di addotto e la caratterizzazione delle variazioni mutazionali sono in corso e dovrebbero fornire ulteriori informazioni sull'impatto e sul contributo di tali meccanismi di riparazione. Nonostante la sperimentazione sia ancora in corso, i risultati di questo lavoro rispondono ad un vecchio quesito sulla riparazione del DNA e dimostrano che la mutagenicità di O^6 -CMG aumenta significativamente in caso di riparazione inadeguata del DNA.

Nel capitolo 5, sono riassunti i risultati del lavoro di questa tesi. Opportunità e sfide future vengono discusse criticamente e vengono presentate raccomandazioni per direzioni future.

Infine, nei *Capitoli 6 e 7*, viene identificata la riparazione dei danni al DNA indotti da un agente chemioterapico, e l'iniziazione della riparazione viene caratterizzata con una combinazione di studi di biologia molecolare, biochimica e cristallografia in uno sforzo collaborativo. La tossicità cellulare è sullo spettro opposto delle conseguenze biologiche indotte dai danni al DNA affrontate nel capitolo precedente. La tossicità cellulare indotta dai danni al DNA viene comunemente sfruttata nella chemioterapia per indurre la morte di cellule tumorali che si dividono rapidamente. Tuttavia, processi cellulari come l'insufficiente bioattivazione e la riparazione del DNA possono ridurre la suscettibilità delle cellule tumorali all'azione di farmaci chemioterapici che danneggiano il DNA. In questi due capitoli, la riparazione dei danni indotti da uno specifico farmaco chemioterapico viene affrontata studiando il contributo e la modalità di azione di un importante meccanismo di riparazione del DNA. L'identificazione di un probabile ma al contempo importante meccanismo cellulare di resistenza ai farmaci raggiunta in questi lavori implica che potrebbero essere elaborate strategie terapeutiche per contrastare tale meccanismo, come il trattamento combinato con specifici inibitori della riparazione.

Abbreviations

3d-Me-A	3-deaza-3-methyl-adenosine
3d-Natpht-A	3-deaza-3-methoxynaphtylethyl-adenosine
3d-Phen-A	3-deaza-3-phenethyl-adenosine
6-TG	6-thioguanine
8-oxoG	8-Oxoguanine
A	Adenine
AA	Aristolochic acid
AAF	Acetylaminofluorene
AF	Acylfulvene
apc	Adenomatous polyposis coli gene
ATL	Alkyltransferase-like
Benzi	2(1 <i>H</i>)-Benzimidazolone
BIM	1 <i>H</i> -Benzimidazole
BOA	2-Benzoxazolinone
C	Cytosine
CHO	Chinese hamster ovary cells
CPDs	Cyclobutane pyrimidine dimers
CRC	Colorectal cancer
CSB	Cockayne Syndrome group B
DDR	DNA damage response
DMSO	Dimethyl sulfoxide
DNA	DeoxyriboNucleic Acid

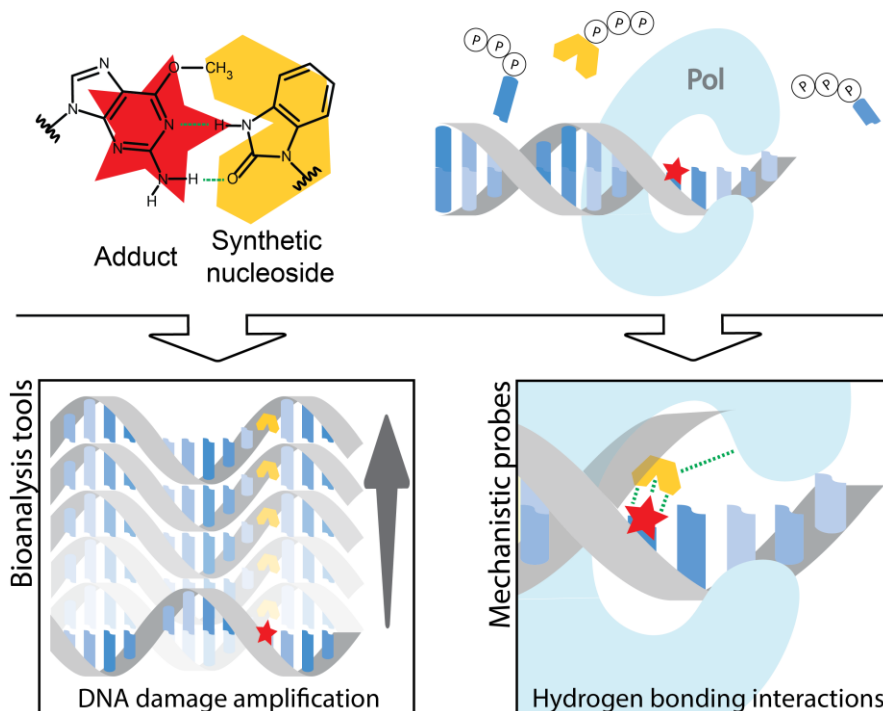
dNTP	2'-Deoxynucleoside-5'-O-triphosphate
Dpo4	DNA polymerase IV of <i>Sulfolobus solfataricus</i>
EC	Elongation complex
Em	Emission
EMS	Ethyl methanesulfonate
EMSA	Electrophoretic mobility shift assay
ERCC1	Excision repair cross-complementation group 1
ESI	Electrospray ionization
Ex	Excitation
ExBenzi	1 <i>H</i> -Naphtho[2,3- <i>d</i>]imidazole-2(3 <i>H</i>)-one
ExBIM	1 <i>H</i> -Naphtho[2,3- <i>d</i>]imidazole
FAM	Fluorescein amidite
G	Guanine
GG-NER	Global Genome-Nucleotide Excision Repair
HCEC	Human Corneal Epithelial Cells
HPLC	High performance liquid chromatography
hprt	Hypoxanthine-guanine phosphoribosyltransferase gene
HPRT	Hypoxanthine-guanine phosphoribosyltransferase
HRMS	High resolution Mass Spectrometry
k_{cat}	Catalytic turnover
k_{cat}/K_M	Catalytic efficiency
KDA	potassium diazoacetate
<i>KlenTaq</i>	Large fragment of <i>Taq</i> DNA polymerase

K_M	Michaelis-Menten constant
<i>KRAS</i>	Kirsten rat sarcoma viral oncogene homologue
LC-MS	Liquid chromatography-mass spectrometry
Mg^{2+}	Magnesium ion
MGMT	<i>O</i> ⁶ -methylguanine-DNA-methyltransferase
MMS	Methyl methanesulfonate
MS	Mass Spectrometry
MS	Molecular sieves
NER	Nucleotide excision repair
NMR	Nuclear magnetic resonance
nt	Nucleotide
<i>O</i> ⁶ -BnG	<i>O</i> ⁶ -Benzylguanine
<i>O</i> ⁶ -CMG	<i>O</i> ⁶ -Carboxymethylguanine
<i>O</i> ⁶ -MeG	<i>O</i> ⁶ -Methylguanine
p53	Tumor protein p53
PCR	Polymerase chain reaction
PE	Plating Efficiency
Per	1 <i>H</i> -Perimidin-2(3 <i>H</i>)-one
Peri	1 <i>H</i> Perimidine
PLS	Postlesion DNA synthesis
Pol	Polymerase
PTGR1	NADPH-dependent prostaglandin reductase 1
qPCR	quantitative PCR

Abbreviations

RNA	Ribonucleic acid
RT-PCR	Reverse transcriptase-PCR
SMRT	Single molecule real-time sequencing
siRNA	Small interfering RNA
SPR	Plasmon Resonance
T	Thymine
<i>Taq</i>	<i>Thermus aquaticus</i>
TC-NER	Transcription Coupled-Nucleotide Excision Repair
TLS	Translesion DNA synthesis
TM	Transcriptional Mutagenesis
TP	Triphosphate
WT	Wild type
XPA	Xeroderma pigmentosum complementation group A
XPC	Xeroderma pigmentosum complementation group C

Chapter 1 – Introduction: DNA Adduct-Directed Synthetic Nucleosides



Reproduced with permission from

Michael H. Rätz, Claudia M. N. Aloisi, Hailey L. Gahlon, and Shana J. Sturla*, DNA Adduct-Directed Synthetic Nucleosides, *Accounts of Chemical Research* **2019**, 52, 5, 1391-1399

Copyright © 2019 American Chemical Society

C.M.N Aloisi wrote paragraphs about DNA damage detection, edited the document, responded and addressed reviewers' comments. M.H Rätz wrote paragraphs about DNA translesion synthesis, edited the document, responded and addressed reviewers' comments. H.L Gahlon wrote about synthetic nucleotides and edited the document. S.J Sturla wrote about nanoparticles and edited the document.

1.1 Abstract

Chemical damage to DNA is a key initiator of adverse biological consequences due to disruption of the faithful reading of the genetic code. For example, O^6 -alkylguanine (O^6 -alkylG) DNA adducts are strongly miscoding during DNA replication when the damaged nucleobase is a template for polymerase-mediated translesion DNA synthesis (TLS). Thus, mutations derived from O^6 -alkylG adducts can have severe adverse effects on protein translation and function and are an early event in the initiation of carcinogenesis. However, the low-abundance of these adducts places significant limitations on our ability to relate their presence and biological influences with resultant mutations or disease risk. As a consequence, there is a critical need for novel tools to detect and study the biological role of alkylation adducts. Incorporating DNA bases with altered structures that are derived synthetically is a strategy that has been used widely to interrogate biological processes involving DNA. Such synthetic nucleosides have contributed to our understanding of DNA structure, DNA polymerase (Pol) and repair enzyme function, and to the expansion of the genetic alphabet. This accounts article describes our efforts towards creating and applying synthetic nucleosides directed at DNA adducts. We have synthesized a variety of nucleosides with altered base structures that complement the altered hydrogen-bonding capacity and hydrophilicity of O^6 -alkylG adducts. The heterocyclic perimidinone-derived nucleoside Per was the first of such adduct-directed synthetic nucleosides; it specifically stabilized O^6 -benzylguanine (O^6 -BnG) in a DNA duplex. Structural variants of Per were used to determine hydrogen-bonding and base-stacking contributions to DNA duplex stability in templates containing O^6 -BnG as well as O^6 -methylguanine (O^6 -MeG) adducts. We created synthetic probes able to stabilize damaged over undamaged templates and established how altered hydrogen bonding or base-stacking properties impact DNA duplex stability as a function of adduct structures. This knowledge was then applied to devise a hybridization-based detection strategy involving gold nanoparticles that distinguish damaged from undamaged DNA by colorimetric changes. Furthermore, synthetic nucleosides were used as mechanistic tools to understand chemical determinants e.g. hydrogen bonding, π -stacking, and size and shape deviations that impact the efficiency and fidelity of DNA adduct bypass by DNA Pols. Finally, we reported the first example of amplifying alkylated DNA, accomplished by combining an engineered polymerase and synthetic triphosphate for which incorporation is templated by a DNA adduct. The presence of the synthetic nucleoside in amplicons could serve as a marker for the presence and location of DNA damage at low levels in DNA strands. Adduct-directed synthetic nucleosides have opened new concepts to interrogate the levels, locations and biological influences of DNA alkylation.

1.2 Introduction

DNA is the molecule of genetic information storage and its accurate replication is essential to all life. DNA is, however, often subject to chemical damage resulting from daily life, such as exposure to chemicals from diet, environment and drugs, leading to chemical alkylation and formation of covalent DNA adducts. DNA adduct structures are diverse, ranging from small base modifications to bulky adducts and crosslinks.¹ Such nucleobase alterations impose serious biochemical consequences, since they can impede replication and, if not effectively repaired, induce cytotoxicity or lead to genomic instability.¹

Translesion DNA synthesis (TLS) is crucial for the prevention of replication fork collapse and replication-associated double strand breaks and relies on DNA polymerases (Pols) that have the capacity to replicate past DNA adducts.² This capacity arises from structural features of TLS Pols, such as open active sites and domains promoting DNA binding interactions. Based on structural classifications, this process is mediated in human cells by the Y-family Pols η , ι , κ , and Rev1, and the B-family Pol ζ . These enzymes have large solvent-exposed active sites, lack proofreading activity, and are inherently error-prone; they are, therefore, responsible for damage-associated mutation events as a basis of carcinogenesis.³

Understanding biochemical interactions and biological consequences of DNA adduct formation is key in elucidating carcinogenesis mechanisms and mitigating cancer risk, as well as improving the efficacy and safety of DNA alkylating drugs used in cancer chemotherapy. Complementary objectives in this context include (1) understanding the biochemical mechanisms governing evasion of DNA repair, (2) the impact of chemical alkylation on replication and altered transcription fidelity, and (3) characterizing the distribution and overall burden of chemical damage to the DNA from exposure to DNA alkylating agents. The impact of DNA adduct formation on biochemical processes and cell fate depends both on the biochemical competences of DNA repair and tolerance, as well as the chemistry of DNA damage. For standard DNA bases, as well as a variety of non-canonical structures, synthetic nucleosides, including Watson-Crick like systems, isosteric shape mimics, and self-pairing hydrophobic analogues have found widespread applications in the expansion of the genetic code,⁴ the investigation of polymerase mechanism,⁵ and the optimization of aptamers.⁶ Furthermore, synthetic nucleosides have been created that pair with isoguanine,⁷ abasic sites,^{8,9} and 7,8-dihydro-8-oxoguanine (8-oxoG).¹⁰ Such probes have great potential as DNA damage sensors and molecular tools to study enzymes, e.g. TLS Pols, but limited strategies are available for the detection of DNA adducts. The use of synthetic nucleosides as specific base pairing partners for alkylation adducts could overcome limitations in amplification-based adduct detection. However, knowledge of physical and chemical properties that drive probe:damage specificity and stabilize probe-adduct interactions in duplex DNA and in enzyme active sites are critical parameters to address. In this accounts article we summarize our advances in

the development of adduct-directed nucleosides and describe their use as chemistry-based tools to address mechanisms of mutation and detect DNA adducts.

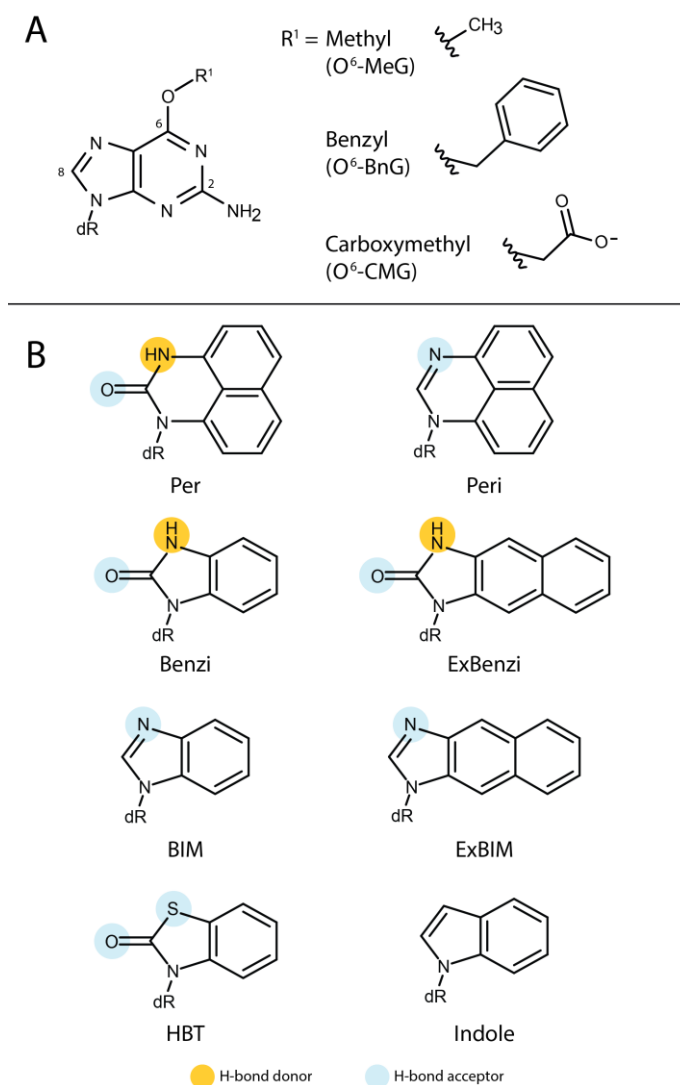


Figure 1. (A) Structures of DNA adducts O^6 -MeG, O^6 -BnG, and O^6 -CMG targeted by adduct-directed nucleosides; dR = 2'-deoxyribose. (B) Synthetic nucleoside probes designed as molecular probes to investigate translesion DNA synthesis or for pairing with O^6 -alkylG adducts.

1.3 Synthetic Nucleosides that Pair specifically with O^6 -Alkylguanine adducts

We have a significant interest in O^6 -alkylG adducts as small but mighty modifiers of the genetic code that are highly relevant to exposures from the environment, diet, and chemotherapeutics (Figure 1 A). O^6 -alkylG adducts arise from S_N1 -like chemical reactions with the exocyclic oxygen at the 6-position of

deoxyguanosine. For example, the formation of the biologically relevant O^6 -alkylG adduct O^6 -MeG derives from exposure to the chemotherapeutic drug temozolomide¹¹ and from various methyl nitrosamines present in tobacco.¹² O^6 -MeG also forms endogenously and is repaired by alkylguanine alkyltransferase, which transfers the methyl group to a cysteine in the active site of the protein.¹³ Another biologically relevant O^6 -alkylG adduct is O^6 -carboxymethylG (O^6 -CMG, Figure 1 A), which arises from the endogenous nitrosation of glycine and subsequent reaction with DNA. The occurrence of O^6 -CMG is strongly associated with diet, in particular with high red meat consumption.^{14, 15} Despite the low abundance of these O^6 -alkylG adducts, forming at rates on the order of 100-1000 O^6 -MeG per human genome (approximately 1-10 O^6 -MeG per 10^8 nts),¹⁶ they are highly mutagenic.

The high mutagenicity of O^6 -alkylG adducts arises from their tendency to misinsert T. Pols often misincorporate T opposite O^6 -alkylG adducts because the O^6 -alkylG:T base pair has a higher geometric similarity to a Watson-Crick base pair as compared to a wobble base pair involving C.¹⁷ The misinsertion of T, and following a subsequent round of DNA replication, leads to GC→AT transition mutations.¹⁸ These types of point mutations are frequently observed in colorectal cancers.¹⁹ Although O^6 -alkylG formation is relevant to disease etiology and therapy, there is limited understanding of the distribution of O^6 -alkylG in the genome and its relationship with the mutation landscape of human cancers. Therefore, our research has focused on tailoring synthetic probes to bind selectively with O^6 -alkylG as a strategy for locating DNA adducts in the human genome.

The design of new DNA adduct:synthetic base pairs focused on exploiting chemical properties of O^6 -alkylG adducts through hydrogen-bonding and π - π -stacking interactions to yield a base pair with increased thermodynamic stability. Based on computational modeling, we hypothesized that a nucleoside derived from the heterocycle perimidinone (Per, Figure 1 B), consisting of a conjugated π -(naphthyl) system, a hydrogen-bond accepting carbonyl group, and an imino-based hydrogen-bond donor, could complement O^6 -alkylG adducts.²⁰ Per was synthesized from diaminonaphthalene and incorporated into DNA duplexes. Duplexes containing the O^6 -BnG:Per pair were more thermally stable than duplexes with G:Per, as indicated by a 4 °C higher T_m value. This example was the first stable DNA base pair with a biologically accessible alkylation adduct.²⁰

To investigate chemical interactions essential for increased duplex stability, we expanded our probe repertoire by preparing a small library of non-natural nucleosides (Figure 1 B) with different size, hydrogen-bonding and π -stacking properties, and evaluated the stability of duplexes in which the probes were paired with DNA adducts. Peri, derived from a perimidine, is similar in size to Per, but lacks the hydrogen-bond-accepting carbonyl group (Figure 1 B). It instead contains an imidazole hydrogen-bond acceptor. Nucleosides derived from benzimidazole (BIM, Figure 1A) and benzimidazolinone (Benzi, Figure 1 B) are comprised of a benzopyrrole-derived bicyclic ring systems and equipped with the same hydrogen-bond-donor-acceptor configuration as Peri and Per, respectively. To test their potential to stabilize duplexes containing O^6 -alkylG adducts, we paired Peri, BIM, or Benzi opposite O^6 -MeG or O^6 -BnG (Figure 1 A) in duplex DNA and characterized the thermodynamic stabilities. DNA duplexes

were generally more stable when the synthetic nucleotides were placed opposite O^6 -MeG or O^6 -BnG rather than canonical nucleobases. For example, in instances where BIM, Benzi and Peri were paired opposite O^6 -MeG, an increase of T_m of 4 °C was observed as compared to these probes paired opposite G.²¹ Interestingly, the larger Per and Peri probes stabilized duplexes containing O^6 -BnG better, whereas BIM and Benzi stabilized duplexes containing O^6 -MeG better (the measured T_m for O^6 -MeG:BIM and O^6 -MeG:Benzi were 2.4 °C higher as compared to BIM and Benzi paired opposite O^6 -BnG). This relationship suggested that BIM and Benzi might stabilize DNA on the basis of hydrogen-bonding interactions whereas the larger Per and Peri probes may utilize π -stacking interactions.²¹

Elucidation of the Per: O^6 -BnG structure in a DNA duplex revealed that Per adopts a *syn* conformation opposite O^6 -BnG and engages in an interstrand intercalation in the DNA duplex.^{22, 23} This orientation allows for a large overlap with the benzyl moiety and suggests contributions of inter-strand π -stacking to stability.²³ There was no evidence for contributions of hydrogen bonding. These data suggested first, while duplexes with the synthetic probes paired with O^6 -MeG and O^6 -BnG adducts were both stabilized to a similar extent, the molecular basis for stability with O^6 -MeG was probably not the same interstrand stacking interaction. Structural data to address the basis of interactions with the methyl adduct remain an active pursuit. Second, we hypothesized that the aromatic rings of Per and Peri could be re-oriented to achieve more effective overlap of the π -surface, leading to the synthesis and evaluation of expanded analogs.

The synthetic nucleosides ExBIM and ExBenzi were created as analogs to BIM and Benzi with expanded (Ex) π -surfaces to test if more overlap with the neighboring bases could improve their function as probes. ExBenzi and ExBIM were paired opposite G and O^6 -MeG in DNA duplexes and their thermodynamic stabilities were characterized. Both nucleosides stabilized the adduct-containing duplex more than BIM and Benzi²¹ and were adduct specific.²⁴ Dangling-end T_m measurements, where the synthetic nucleoside is unpaired and overlaps with the adjacent terminal base pair, confirmed that ExBIM and ExBenzi stabilized the duplex more than Benzi and BIM, and supported the hypothesis for enhanced surface area overlap with the neighboring base on the opposing strand.²⁴ Further, while Benzi, BIM, ExBenzi and ExBIM are fluorescent (quantum yields $\Phi_f = 0.35; 0.82; 0.69; 0.44$, respectively), the fluorescence of ExBIM was the most sensitive to the altered chemical environment imposed by the presence of the adduct. Thus, ExBIM showed the highest adduct selectivity. In particular, ExBIM: O^6 -MeG duplexes had a 2-4-fold higher fluorescence emission at 370 nm (excitation at 250 nm) as compared to ExBIM paired opposite canonical bases or other alkylation-adducts.²⁴ The selectivity for stabilizing alkylated DNA and the enhanced fluorescence properties make these elongated synthetic nucleosides potential candidates for DNA hybridization-based O^6 -alkylG sensors.

1.3.1 DNA Adduct Hybridization Nanoprobes

Aside from fluorescence-based strategies, combining nanotechnology with selective hybridization of oligonucleotide probes has been extremely effective in bioanalysis involving both environment-sensitive changes in fluorescence and in using nanoparticles as biosensors. Hybridization-mediated aggregation of gold nanoparticles (Au-NPs) is effective in bioanalysis due to high absorption coefficients and surface plasmon resonance (SPR) properties, related to shape, size and interparticle distance. Au-NPs were seen as a particularly promising reporter candidate for the presence of DNA adducts in defined sequences due to the sharp transitions that result from the cooperative nature of Au-NPs functionalized with oligonucleotides.²⁵ However, limitations in using standard Au-NP targeting approaches for adduct detection led us to develop novel combinations of functionalized Au-NPs as DNA adduct nanoprobes.

We integrated our DNA adduct-directed synthetic nucleotides with Au-NPs to detect O^6 -alkylG adducts within a specific sequence and as a minority component in a DNA sample.²⁶ The specificity of this Au-NP strategy allowed visual discrimination between modified and unmodified DNA. Thus, Au-NPs were conjugated to two sets of oligonucleotides, one functionalized with a sequence that terminated with Per, and the other functionalized with a sequence complementary to the target DNA when aligned tail-to-tail. A Au-NP suspension is red in color, however when a target sequence containing a DNA adduct was added, the color changed to purple as a result of coupling of localized plasmon fields (Figure 2).

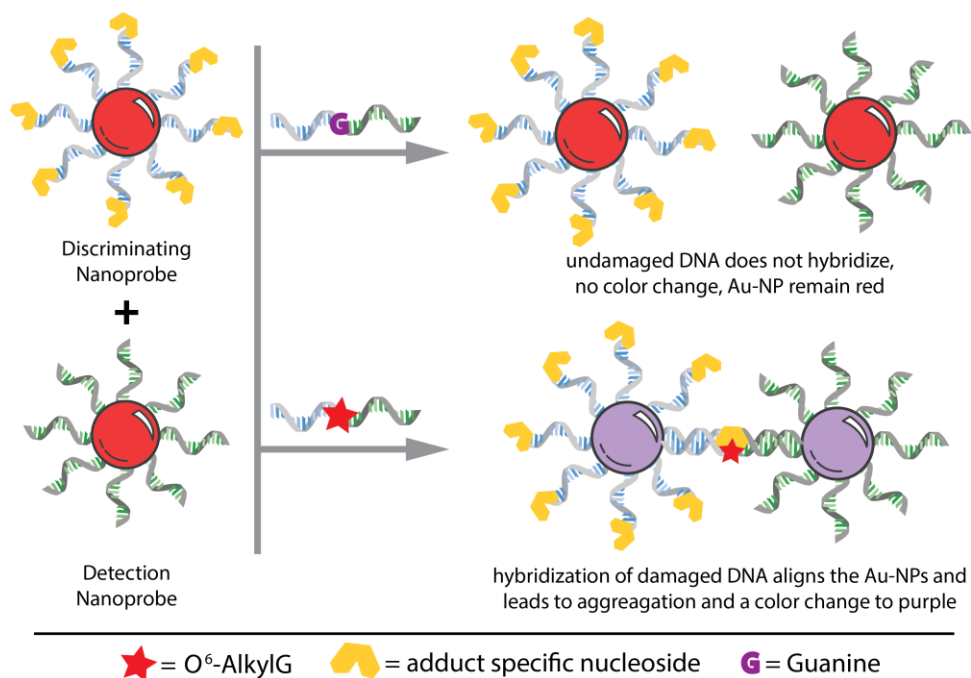


Figure 2. Schematic representation of nanoprobe-based detection of O^6 -alkylG adducts.

Aggregates containing DNA adducts were more thermally stable due to the higher binding affinity of strands with Per to strands with O^6 -BnG vs. canonical bases. Thus, in the absence of DNA adduct targets, samples dispersed and remained red (Figure 2). The discrimination for matched targets allows one to quantitatively relate the increase in thermal stability of aggregates with increasing concentrations of O^6 -BnG, even in the presence of mixed targets with adduct levels below 1% of unmodified DNA.

Next, we focused on increasing probe sensitivity and targeting biologically relevant adduct structures in human cancer-relevant gene sequences.²⁷ For example, ExBenzi and ExBIM were inserted into oligonucleotide-functionalized Au-NP-based nanoprobe targeting O^6 -MeG in the human *KRAS* gene.²⁷ The functionalized Au-NPs exhibited a characteristic SPR band at 530 nm and remained dispersed/red until a target containing an adduct was added, which caused the detection and discriminating probes to come into close proximity to one another and change color (Figure 2). This process, involving the selective detection of one methyl group in at least 13 bases, was found to be effective also in the presence of mixed targets, useful for sequences of varying lengths and sequence contexts, as well as in the presence of genomic DNA, notably without heating and under non-stringent salt conditions, often used in mismatch discrimination for single nucleotide polymorphisms in unamplified genomic samples.

1.4 Synthetic Nucleosides as Probes to Investigate Translesion DNA synthesis

Adduct-directed nucleoside probes have also been exploited to probe the catalytic center of enzymes involved in repair and bypass of O^6 -alkylG adducts e.g. TLS Pols or alkylguanine alkyltransferase.²⁸ Translesion DNA synthesis continues DNA replication in the face of DNA damage, but the low fidelity of TLS Pols is responsible for many mutations observed in cancer.³ Therefore, elucidating the mechanisms of DNA adduct bypass by TLS Pols is crucial to understand DNA damage-induced mutagenesis. Much of the current knowledge for how Pols work stems from studies with synthetic nucleosides designed to probe hydrogen bonding,²⁹⁻³² shape complementarity,³³⁻³⁷ hydrophobic packing,³⁸⁻⁴⁰ desolvation, and base-stacking^{41, 42} interactions. These studies showed that the Pol specificity and efficiency varies between Pol families and is substrate dependent. TLS Pols share structural-similarities with replicative Pols, but they possess unique, solvent-exposed active sites, that allow for the specialized bypass of DNA lesion.⁴³ Still, molecular mechanisms of TLS are not well understood due to multiple interactions between the Pol, the DNA and the incoming dNTP. In addition, the structural diversity of DNA adducts that Pols encounter adds great complexity to our understanding for chemical determinants involved in DNA adduct bypass. Therefore, new chemical strategies are required to decipher molecular interactions which guide DNA Pol-mediated TLS.

TLS is characterized by a nucleotide insertion and subsequent extension step to replicate past DNA damage. To understand chemical factors impacting nucleotide selection for insertion opposite O^6 -MeG,

we tested the ability of the archaeal DNA polymerase Dpo4 to insert nucleoside analogs opposite O^6 -MeG.⁴⁴ We found that synthetic dNTPs with a wide range of size and hydrogen-bonding capacities were generally inserted with a higher frequency opposite G than O^6 -MeG.

To be able to interrogate a wider array of sizes and hydrogen-bonding capacities, and to understand structural features of nucleotides that promote the extension step after insertion of a nucleotide at a DNA adduct, i.e. postlesion DNA synthesis (PLS), we tested the capacity of Dpo4 to extend from O^6 -MeG and O^6 -BnG adducts paired with adduct-directed synthetic probes (Figure 1 B).^{22, 45} Thus, Benzi, BIM, Peri and Per were placed at the primer terminus, directly opposite G, O^6 -MeG and O^6 -BnG (Figure 1 A) and primer-extension analysis as well as steady-state kinetic parameters were determined. We found that the larger naphthyl-containing probes, Peri and Per, were a complete block for Dpo4 to extend from O^6 -MeG or O^6 -BnG, while the smaller probes, BIM and Benzi were extended by Dpo4 in an error-free manner.

Regarding TLS, Dpo4 extended from Benzi paired opposite O^6 -MeG and O^6 -BnG, but did not extend from Benzi paired opposite G. Assuming Benzi adopts a syn conformation opposite adducts, as was observed in the crystal structure of its analog Per, Benzi can form two hydrogen bonds with O^6 -MeG and O^6 -BnG. In the case of Benzi paired opposite G, a clash is likely to occur between the N1 hydrogen on G and the NH hydrogen on Benzi. This clash could inhibit Dpo4 from extending the strand. Overall, these data support a model whereby hydrogen bonding at the DNA terminus promotes O^6 -alkylG bypass by Dpo4. While Dpo4 is an informative model enzyme to interrogate lesion bypass mechanisms, information was still lacking regarding how DNA Pols likely to be involved in extension during replication perform PLS.

Human DNA Pol ζ is a B-family enzyme involved in the extension step during lesion bypass. For example, for abasic sites, insertion is often performed by DNA Pol δ , Rev1 or Pol η , while extension is efficiently carried out by DNA Pol ζ .⁴⁶ To define the role of hydrogen bonding and base stacking for human Pol ζ to perform PLS, we examined its extension from nucleoside probes Benzi, HBT, BIM and Indole as well as C and U (Figure 3) placed opposite G, O^6 -MeG, N^1 -MeG and N^2 -MeG. The hybridization of each primer and template resulted in 24 permutations of oligonucleotides to systematically test the effect of hydrogen bonding at the DNA terminus for DNA Pol ζ PLS. Plotting the k_{cat}/K_m as a function of each base pair at the DNA terminus suggested that Pol ζ displayed the highest catalytic efficiency by (1.3-1.5 $\mu\text{M}^{-1}\text{min}^{-1}$) with DNA containing terminal base pairs with three hydrogen bonds, whereas termini with no capacity for hydrogen bonding e.g. O^6 -MeG:Indole (Figure 3) displayed the lowest catalytic efficiency (0.005-0.12 $\mu\text{M}^{-1}\text{min}^{-1}$).

In some instances, termini anticipated to engage in two hydrogen bonds with the syn oriented probe did not follow the trend, for example O^6 -MeG:Uracil, (two hydrogen bonds, Figure 3) was less efficiently extended by Pol ζ than O^6 -MeG:HBT (one hydrogen bond, Figure 3) with catalytic efficiencies of 0.057 $\mu\text{M}^{-1}\text{min}^{-1}$ and 0.063 $\mu\text{M}^{-1}\text{min}^{-1}$, respectively. This finding suggests that while hydrogen bonding is important, stacking interactions also influence DNA Pol ζ catalysis.⁴⁷

These examples highlight unique chemical insight provided by using nucleoside analogs as a strategy to interrogate mechanistic details of nucleotide insertion and extension by DNA Pols. By systematically changing the structure of synthetic nucleosides, we have demonstrated the importance of hydrogen bonding for Dpo4 and human DNA Pol ζ to insert and extend nucleotides during DNA synthesis template by DNA containing O^6 -alkylG.

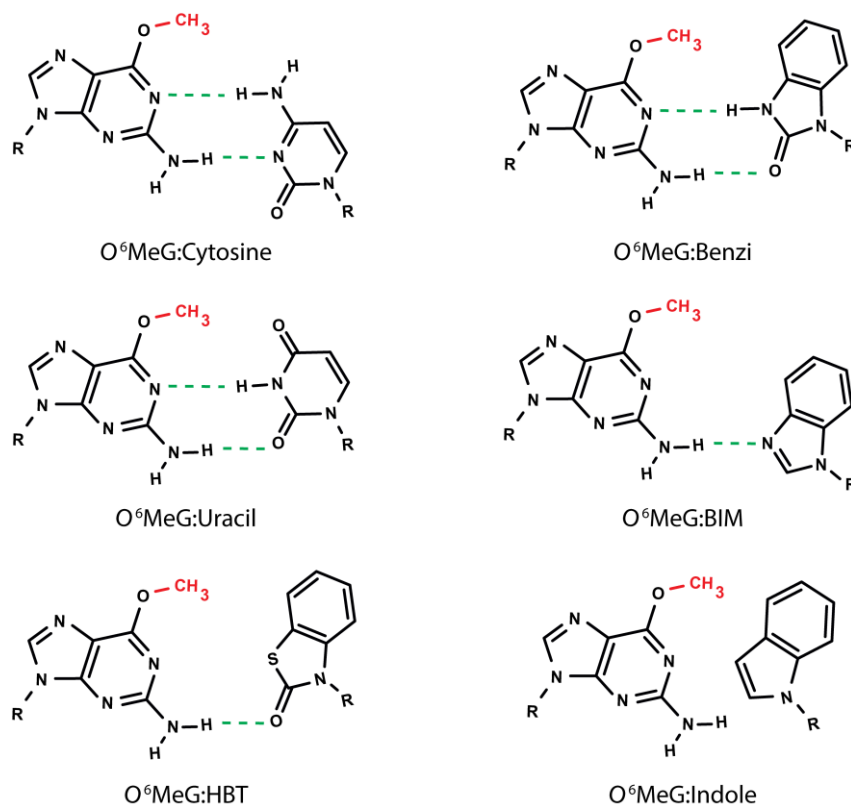


Figure 3. Base pair variations with different hydrogen-bond donor and acceptor functionalities. Proposed base pair structures of O^6 -MeG paired with C, U, Benzi, BIM, HBT, and indole.

1.5 DNA Damage Detection and Amplification with Synthetic Nucleosides

The extremely low abundance and structural diversity of DNA alkylation products makes damage detection difficult, and the current state-of-the-art for quantifying total DNA damage burden is mass spectrometry.⁴⁸⁻⁵⁰ However, mass spectrometric approaches usually provide limited sequence context. In recent years, diverse strategies have emerged to map DNA adducts within genomic locations, such as single-molecule real-time (SMRT) sequencing⁵¹ and nanopore sequencing.^{52, 53} Combinations of immunoaffinity enrichment with polymerase blockage and high throughput sequencing have produced the first maps of bulky adducts like those arising from chemicals and UV damage.⁵⁴ A complementary

approach has been to use a DNA repair enzyme to create gapped DNA sites that can be marked with an amplifiable synthetic base pair⁵³ or a oligonucleotide code sequence.⁵⁵ For O^6 -alkylG adducts, adduct-directed synthetic probes offer a possible chemical basis for locating damaged bases within sequence contexts by polymerase-mediated amplification.

Motivated to decipher the complex relationship between the occurrence of DNA adducts and adverse biological effects, we have developed detection strategies that integrate synthetic nucleotides and engineered Pols to measure O^6 -alkylG adducts in defined DNA sequences. By designing primers to target specific areas of the genome (e.g. mutation hotspots), we were able to locate alkylation-damage in a sequence-specific manner. We anticipate that the amplification of the DNA damage signal using synthetic nucleosides coupled to sensitive detection methods would allow us to overcome the detection of low abundance adducts and form a basis for future DNA-adduct-detection in biological samples.

To assess synthetic nucleosides as damage probes, we evaluated the incorporation of the synthetic triphosphates BIMTP and BenziTP opposite O^6 -BnG in single-nucleotide primer extension studies with engineered Pols. Amongst the Pols tested, the KlenTaq mutant KTqM747K was the most specific enzyme and incorporated BenziMP more frequently (79%) than natural dNTPs opposite O^6 -BnG, and only to a small extent opposite G (12%). Further, BenziMP was 25-fold more efficiently incorporated opposite O^6 -BnG than opposite G, and incorporation of BenziMP opposite O^6 -BnG was more efficient than incorporation of natural dNTPs (17-fold and 5-fold more than dCMP and dTMP, respectively). These data demonstrate the adduct-specific incorporation of BenziTP over canonical nucleotides.

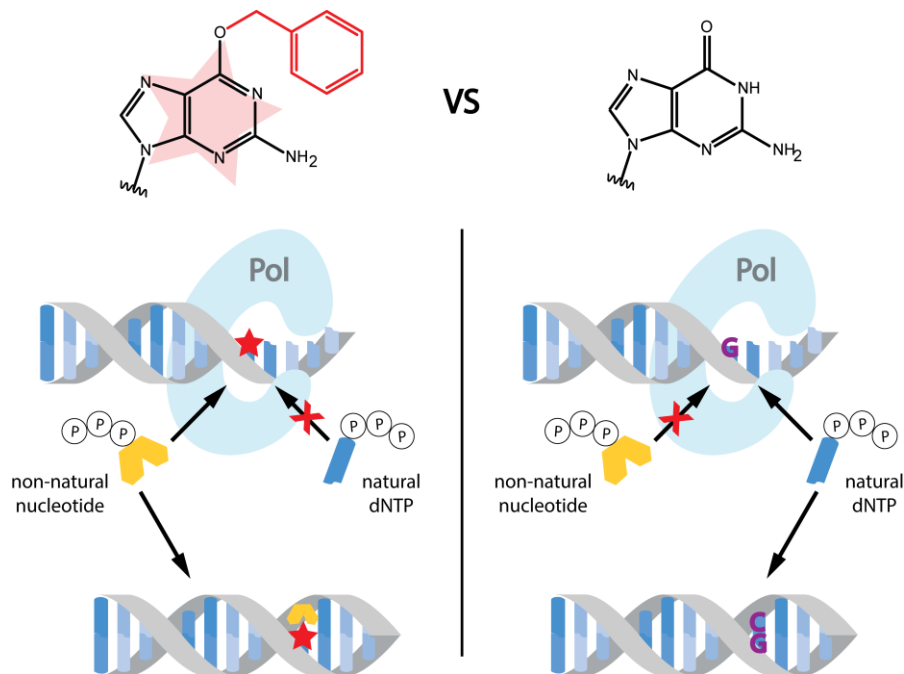


Figure 4. Specific polymerase-mediated incorporation of a non-natural nucleotide (BenziTP) opposite O^6 -BnG in the presence of natural dNTPs.

In primer-extension assays with all four natural dNTPs, KTqM747K was blocked by O^6 -BnG, and bypass and subsequent full-length extension occurred only after adding BenziTP (Figure 4). The high Pol specificity and the requirement of BenziTP for full-length extension allowed us to amplify the signal from BenziTP and hence from O^6 -BnG, in an experimental set-up involving the use of a small amount of the damaged template and excess of the primer, by iteratively repeating primer-extension steps. To our knowledge, this was the first example of linear amplification of an alkylation adduct, establishing an exciting outlook for future PCR-like approaches for analyzing DNA adducts.⁵⁶

To expand the proof-of-principle findings with the model DNA adduct O^6 -BnG to adducts of biological relevance, we tested signal-amplification of O^6 -MeG (Figure 1 A) and O^6 -carboxymethylguanine (O^6 -CMG, Figure 1 A). KTqM747K bypassed both adducts by incorporation of BenziMP and misincorporation of dTMP. Although KTqM747K catalyzes the insertion of BenziTP more efficiently opposite O^6 -MeG than O^6 -CMG, its insertion opposite O^6 -CMG is more specific as much lower incorporation of natural dCMP and dTMP was observed (41% and 92% respectively dCMP and dTMP opposite O^6 -MeG vs. 17% and 32% opposite O^6 -CMG).⁵⁷ Indeed, when performing full-length primer-extension in the absence of BenziTP, O^6 -MeG was bypassed and over 90% of primer was extended, while replication was stalled at the O^6 -CMG site and BenziTP was required for full primer synthesis. Identical to O^6 -BnG, iterative primer-extension reactions allowed for the amplification of O^6 -CMG in the presence of BenziTP. On the basis of steady-state kinetic parameters, nucleotide incorporation opposite the adducts was up to 150-fold higher as compared to G.⁵⁷

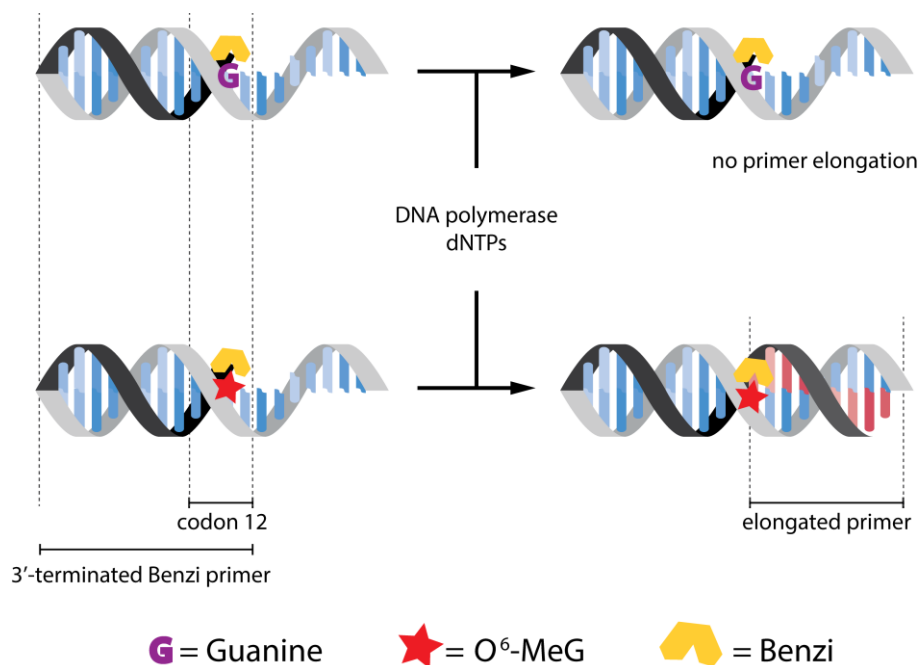


Figure 5. KlenTaq specifically extends past a Benzi: O^6 -MeG but not a Benzi:G base pair. The discrimination in elongation allows for the detection of O^6 -MeG through linear amplification.

Apart from specifically inserting a non-natural nucleotide opposite O^6 -alkylG, we found that some Pols have a specific capacity to extend from non-natural nucleotides paired with O^6 -alkylG. For example, Dpo4 is proficient in extending from Benzi: O^6 -MeG but unable to extend a Benzi:G base pair.^{22, 45} However, this specificity is sequence-dependent. KlenTaq and the KTqM747K mutant are equally specific in their extension capacity from Benzi paired with O^6 -MeG vs. G but remain unaffected by the DNA sequence context.⁵⁸ To exploit this selectivity as a basis for the detection of O^6 -MeG, we prepared primers containing Benzi at the 3'-end. Primers were tailored to be complementary to a cancer hotspot region of the *KRAS* sequence in a way that Benzi was paired opposite the second base of *KRAS* codon 12 (Figure 5),⁵⁸ which is mutated in 30% of cancers.⁵⁹ Repeated cycles of annealing to a template containing G or O^6 -MeG and extension of the Benzi primer allowed us to amplify the damage signal and detect amplicons containing Benzi only when paired opposite O^6 -MeG. In mixtures containing both G and O^6 -MeG templates, the signal intensity uniquely depended on the amount of O^6 -MeG and increased linearly with the amount of O^6 -MeG. The specific positioning of Benzi, together with the extension properties of KlenTaq, allowed us to catalyze linear amplification of the O^6 -MeG signal in a specific DNA sequence context.

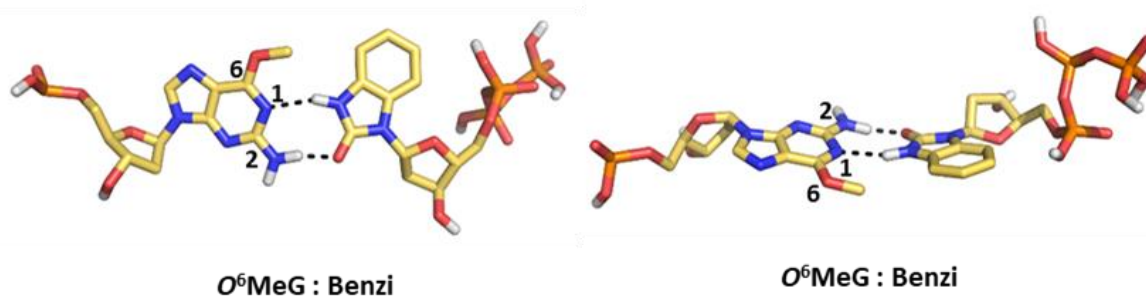


Figure 6. Base-pairing interactions of O^6 -MeG with BenziTP (PDB: 5OXJ). Hydrogen bondings are shown as dashed lines.

To investigate the structural basis for base-pairing selectivity for Benzi with O^6 -alkylG adducts, KTqM747K was crystallized in a ternary complex with BenziTP paired opposite O^6 -MeG.⁶⁰ The KTqM747K $_{O^6\text{-MeG:BenziTP}}$ structure displayed a closed enzyme conformation, which superimposed well with the KTqM747K $_{G:dCTP}$ structure. BenziTP was found to adopt a *syn*-conformation and form two hydrogen bonds with the templating O^6 -MeG via the *N1* on O^6 -MeG and the NH donor on BenziTP (3.2 Å) as well as between NH₂ donor on O^6 -MeG and carbonyl group on BenziTP (2.9 Å) (Figure 6).⁶⁰ The position of the methyl group in relation to BenziTP was not clearly defined, however, the distal position had a higher refined occupancy. The distal orientation positions the methyl group closer to the phenyl ring and misaligns BenziTP, resulting in widening the O^6 -MeG:Benzi base pair and a slightly larger propeller twist as compared to G:C. As a consequence, the enzyme is not as tightly closed as the

ternary KTqM747K_{G:dCTP} complex, resulting in more flexibility of the finger domain and potential reduced catalytic efficiency through enhanced conformational flexibility. These data provide key insight for further design and optimization of enzyme- and synthetic probe combinations for further challenges in achieving DNA damage amplification in native biological samples.

1.6 Conclusion

DNA adduct-directed nucleosides are a class of synthetic nucleosides that bind selectively in DNA with DNA adducts that are relevant in human disease. Herein, we described the path from their design and physical characterization in DNA duplexes to applications in mechanistic biochemical studies and in building sequence-targeted DNA detection strategies. The structures were optimized on the basis of exploring combinations of hydrogen-bonding, hydrophobic and π -stacking interactions for DNA duplex stabilization. For DNA duplex interactions, stabilization of adduct-containing DNA appeared to be promoted most effectively through hydrophobic aromatic ring system interactions and on the basis of these interactions, a notable advance was the creation of gold nanoparticles decorated with synthetic probe-containing DNA as a basis for hybridization detection of trace amounts of DNA in mixtures of unmodified genomic DNA. On the other hand, in the active site of DNA polymerases, the altered hydrogen-bonding properties of the probes were consequential for efficient extension process from alkylated guanine adducts by TLS Pols. Finally, the probes, combined with engineered polymerases, were the basis of achieving a breakthrough in amplifying alkylated DNA for the first time with synthetic triphosphate incorporation as a marker for the presence of DNA damage. Further research in this area is ongoing to couple sensitive detection strategies to report on the presence of markers for DNA adducts and define their genomic distribution. Overarching goals are to anticipate biological consequences of chemically induced DNA damage, such as cytotoxicity and mutagenicity, by understanding how discreet chemical binding interactions in DNA duplexes and with DNA-interacting enzymes drive these outcomes.

1.7 Overview of thesis work

The aims of the work presented in this doctoral thesis were to create gene-targeted tools for detecting mutagenic DNA adducts, and to investigate the biological consequences of such adducts, in particular the impact of DNA repair on adduct-induced mutagenicity. Although the aims of this thesis differ from an experimental point of view, they tackle two important pieces of the same jigsaw: understanding and predicting mutagenesis that arises from DNA damage.

In *Chapter 2* and *3*, artificial nucleotides are employed for the detection of mutagenic O^6 -alkylguanines, exploiting principles of selective base pairing and polymerase-catalyzed replication past DNA adducts, which have been discussed in this introductory review chapter, *Chapter 1*.

In *Chapter 2*, we developed a strategy to detect O^6 -methylguanine (O^6 -MeG) in a DNA sequence to the single base resolution. Heterocyclic imidic nucleoside analogue Benzi is inserted in a DNA primer, which is hybridized opposite DNA containing O^6 -MeG in the target sequence of the *k-ras* gene. The engineered polymerase KlenTaq synthesized DNA past O^6 -MeG:Benzi base pair, but completely stalled when Benzi was opposite undamaged G. The high specificity of bypass when Benzi paired with O^6 -MeG vs. G meant that, by cycling repeating reactions of primer extension, Benzi-modified DNA primer was amplified only in the presence of O^6 -MeG, and that O^6 -MeG-damaged DNA was detected from a mixture of damaged and undamaged DNA. A fluorescent signal associated with DNA amplicons linearly increased with the amount of O^6 -MeG DNA in the sample. Molecular modeling analyses suggested that O^6 -MeG and Benzi paired more favourably than G and Benzi, and therefore adopted a stable planar conformation, which was superimposable to that of natural base pairs. The presented strategy of O^6 -MeG detection to the base resolution could be expanded to other cancer-relevant gene sequences, thereby broadening our current understanding of the relationship between DNA damage and disease initiation and progression.

In *Chapter 3*, the aromatic imidic nucleotide ExBenziTP, equipped with an expanded conjugated π -system compared to the previously reported BenziTP, was synthesized for the detection and quantification of the mutagenic O^6 -carboxymethylguanine (O^6 -CMG) DNA adduct, advancing in scope and performance the qualitative sensing of DNA adducts using artificial nucleotides. ExBenzi displayed improved catalytic efficiency of incorporation opposite O^6 -CMG by the bypass-proficient polymerase KlenTaq M747K. ExBenziTP was required for replication past O^6 -CMG and for the full synthesis of a DNA primer annealed opposite damaged DNA. The incorporation of ExBenzi opposite undamaged G was poor, probably due to the energetically demanding interaction between ExBenziTP and G, as computed by molecular modelling. Instead, ExBenziTP established strong hydrogen bonds with O^6 -CMG, and energetically favorable ionic and Van der Waals interactions with the surrounding enzymatic residues, which might be a basis of ExBenzi selective incorporation over natural dNTPs. The high template-selectivity and selectivity over natural dNTPs measured for ExBenzi meant that DNA past O^6 -CMG DNA adducts could be replicated and amplified, by reiterating reactions of primer extension. Amplified DNA contained ExBenzi and effectively marked the presence and the location of O^6 -CMG. By developing an analytical method for ExBenzi quantification, we demonstrated that ExBenzi detected from amplified DNA linearly increased with the amount of O^6 -CMG in the initial sample, therefore providing a strategy for the quantification of O^6 -CMG in a target DNA sequence. With the work presented in this chapter, we achieved for the first time an amplified quantitative signal of trace levels

of mutagenic DNA damage at a specific location in a DNA sequence, laying the foundation to the study of the relationship between DNA damage and mutations.

In *Chapter 4*, to understand how DNA mutations arise from O^6 -CMG in cells, we investigated the impact of DNA repair on O^6 -CMG cellular phenotype. Based on findings from few previous studies regarding the repair of O^6 -CMG or other O^6 -alkylGs, we focused on investigating the role of nucleotide excision repair (NER) and O^6 -methylguanine-DNA-methyltransferase (MGMT) protein in repairing O^6 -CMG. We employed an O^6 -CMG-inducing drug, azaserine, and measured cellular responses to it upon depletion of MGMT or NER. Azaserine was significantly more toxic and mutagenic in cells where MGMT was chemically inhibited. Similarly, defined deficiencies in the NER machinery sensitized cells to carboxymethylation, and resulted in increased mutagenicity. Cells impaired in both repair pathways produced the highest number of mutants, raising the compelling question on whether MGMT and NER proteins might crosstalk over O^6 -CMG repair. Despite pending experimentation, with the work presented in *Chapter 4*, a longstanding question in DNA repair was unraveled and two major factors affecting O^6 -CMG mutagenicity in cells were identified, namely MGMT and the NER machinery. Ongoing and future planned studies include the measurement of adduct level upon repair impairment, the characterization of azaserine-induced mutational spectra as a function of DNA repair capacities, and the investigation of the mode of action of the MGMT and NER.

1.6. References

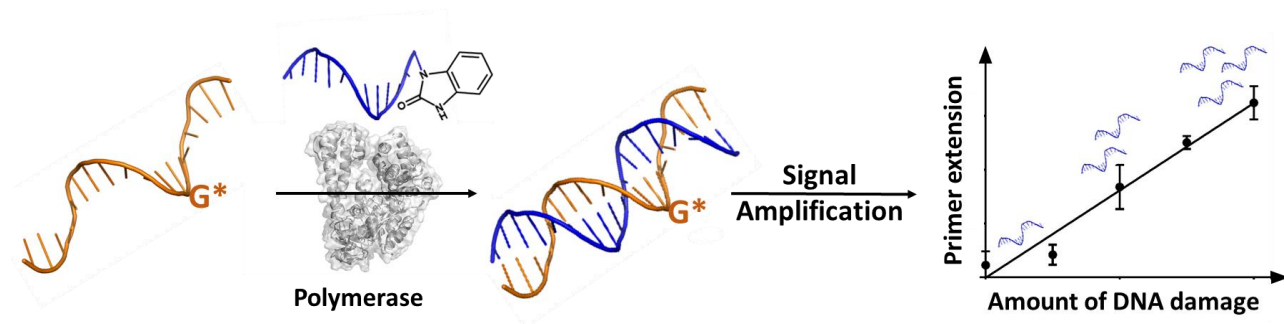
1. Roos, W. P.; Thomas, A. D.; Kaina, B. DNA damage and the balance between survival and death in cancer biology. *Nat. Rev. Cancer* **2016**, *16*, 20-33.
2. Waters, L. S.; Minesinger, B. K.; Wiltrout, M. E.; D'Souza, S.; Woodruff, R. V.; Walker, G. C. Eukaryotic translesion polymerases and their roles and regulation in DNA damage tolerance. *Microbiol. Mol. Biol. Rev.* **2009**, *73*, 134-154.
3. Lange, S. S.; Takata, K.; Wood, R. D. DNA polymerases and cancer. *Nat. Rev. Cancer* **2011**, *11*, 96-110.
4. Feldman, A. W.; Romesberg, F. E. Expansion of the Genetic Alphabet: A Chemist's Approach to Synthetic Biology. *Acc. Chem. Res.* **2018**, *51*, 394-403.
5. Jung, K. H.; Marx, A. Nucleotide analogues as probes for DNA polymerases. *Cell Mol. Life Sci.* **2005**, *62*, 2080-2091.
6. Rothlisberger, P.; Gasse, C.; Hollenstein, M. Nucleic Acid Aptamers: Emerging Applications in Medical Imaging, Nanotechnology, Neurosciences, and Drug Delivery. *Int. J. Mol. Sci.* **2017**, *18*.
7. Switzer, C. Y.; Moroney, S. E.; Benner, S. A. Enzymic recognition of the base pair between isocytidine and isoguanosine. *Biochemistry* **1993**, *32*, 10489-10496.
8. Matray, T. J.; Kool, E. T. A specific partner for abasic damage in DNA. *Nature* **1999**, *399*, 704-708.
9. Zhang, X.; Donnelly, A.; Lee, I.; Berdis, A. J. Rational attempts to optimize non-natural nucleotides for selective incorporation opposite an abasic site. *Biochemistry* **2006**, *45*, 13293-13303.
10. Taniguchi, Y.; Kawaguchi, R.; Sasaki, S. Adenosine-1,3-diazaphenoxazine derivative for selective base pair formation with 8-oxo-2'-deoxyguanosine in DNA. *J. Am. Chem. Soc.* **2011**, *133*, 7272-7275.
11. Roos, W. P.; Batista, L. F.; Naumann, S. C.; Wick, W.; Weller, M.; Menck, C. F.; Kaina, B. Apoptosis in malignant glioma cells triggered by the temozolomide-induced DNA lesion O^6 -methylguanine. *Oncogene* **2007**, *26*, 186-197.
12. Brown, B. G.; Chang, C. J.; Ayres, P. H.; Lee, C. K.; Doolittle, D. J. The effect of cotinine or cigarette smoke co-administration on the formation of O^6 -methylguanine adducts in the lung and liver of A/J mice treated with 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK). *Toxicological sciences : an official journal of the Society of Toxicology* **1999**, *47*, 33-39.
13. Pegg, A. E. Multifaceted Roles of Alkyltransferase and Related Proteins in DNA Repair, DNA Damage, Resistance to Chemotherapy, and Research Tools. *Chem. Res. Toxicol.* **2011**, *24*, 618-639.
14. Vanden Bussche, J.; Hemeryck, L. Y.; Van Hecke, T.; Kuhnle, G. G.; Pasmans, F.; Moore, S. A.; Van de Wiele, T.; De Smet, S.; Vanhaecke, L. O^6 -carboxymethylguanine DNA adduct formation and lipid peroxidation upon in vitro gastrointestinal digestion of haem-rich meat. *Molecular nutrition & food research* **2014**, *58*, 1883-1896.
15. Hemeryck, L. Y.; Rombouts, C.; De Paepe, E.; Vanhaecke, L. DNA adduct profiling of in vitro colonic meat digests to map red vs. white meat genotoxicity. *Food and chemical toxicology : an international journal published for the British Industrial Biological Research Association* **2018**, *115*, 73-87.
16. Foiles, P. G.; Miglietta, L. M.; Akerkar, S. A.; Everson, R. B.; Hecht, S. S. Detection of O^6 -methyldeoxyguanosine in human placental DNA. *Cancer Res.* **1988**, *48*, 4184-4188.

17. Warren, J. J.; Forsberg, L. J.; Beese, L. S. The structural basis for the mutagenicity of *O*⁶-methyl-guanine lesions. *Proc. Natl. Acad. Sci.* **2006**, *103*, 19701-19706.
18. Fu, D.; Calvo, J. A.; Samson, L. D. Balancing repair and tolerance of DNA damage caused by alkylating agents. *Nat. Rev. Cancer.* **2012**, *12*, 104-120.
19. Gottschalg, E.; Scott, G. B.; Burns, P. A.; Shuker, D. E. Potassium diazoacetate-induced p53 mutations in vitro in relation to formation of *O*⁶-carboxymethyl- and *O*⁶-methyl-2'-deoxyguanosine DNA adducts: relevance for gastrointestinal cancer. *Carcinogenesis* **2007**, *28*, 356-362.
20. Gong, J.; Sturla, S. J. A synthetic nucleoside probe that discerns a DNA adduct from unmodified DNA. *J. Am. Chem. Soc.* **2007**, *129*, 4882-4883.
21. Gahlon, H. L.; Sturla, S. J. Hydrogen bonding or stacking interactions in differentiating duplex stability in oligonucleotides containing synthetic nucleoside probes for alkylated DNA. *Chemistry* **2013**, *19*, 11062-11067.
22. Gahlon, H. L.; Schweizer, W. B.; Sturla, S. J. Tolerance of base pair size and shape in postlesion DNA synthesis. *J. Am. Chem. Soc.* **2013**, *135*, 6384-6387.
23. Kowal, E. A.; Lad, R. R.; Pallan, P. S.; Dhummakupt, E.; Wawrzak, Z.; Egli, M.; Sturla, S. J.; Stone, M. P. Recognition of *O*⁶-benzyl-2'-deoxyguanosine by a perimidinone-derived synthetic nucleoside: a DNA interstrand stacking interaction. *Nucleic Acids Res.* **2013**, *41*, 7566-7576.
24. Dahlmann, H. A.; Berger, F. D.; Kung, R. W.; Wyss, L. A.; Gubler, I.; McKeague, M.; Wetmore, S. D.; Sturla, S. J. Fluorescent Nucleobase Analogues with Extended Pi Surfaces Stabilize DNA Duplexes Containing *O*⁶-Alkylguanine Adducts. *Hel. Chim. Acta* **2018**, *101*, e1800066.
25. Trantakis, I. A.; Bolisetty, S.; Mezzenga, R.; Sturla, S. J. Reversible aggregation of DNA-decorated gold nanoparticles controlled by molecular recognition. *Langmuir* **2013**, *29*, 10824-10830.
26. Trantakis, I. A.; Sturla, S. J. Gold nanoprobe for detecting DNA adducts. *Chem. Commun.* **2014**, *50*, 15517-15520.
27. Trantakis, I. A.; Nilforoushan, A.; Dahlmann, H. A.; Stauble, C. K.; Sturla, S. J. In-Genes Quantification of *O*⁶-Methylguanine with Elongated Nucleoside Analogues on Gold Nanoprobes. *J. Am. Chem. Soc.* **2016**, *138*, 8497-8504.
28. McKeague, M.; Otto, C.; R az, M. H.; Angelov, T.; Sturla, S. J. The Base Pairing Partner Modulates Alkylguanine Alkyltransferase. *ACS Chem. Biol.* **2018**, *13*, 2534-2541.
29. Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. Enzymatic Incorporation of a New Base Pair into DNA and RNA Extends the Genetic Alphabet. *Nature* **1990**, *343*, 33-37.
30. Ohtsuki, T.; Kimoto, M.; Ishikawa, M.; Mitsui, T.; Hirao, I.; Yokoyama, S. Unnatural base pairs for specific transcription. *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 4922-4925.
31. Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Yokogawa, T.; Nishikawa, K.; Yokoyama, S. An unnatural base pair for incorporating amino acid analogs into proteins. *Nature biotechnology* **2002**, *20*, 177-182.
32. Hoshika, S.; Leal, N. A.; Kim, M. J.; Kim, M. S.; Karalkar, N. B.; Kim, H. J.; Bates, A. M.; Watkins, N. E., Jr.; SantaLucia, H. A.; Meyer, A. J.; DasGupta, S.; Piccirilli, J. A.; Ellington, A. D.; SantaLucia, J., Jr.; Georgiadis, M. M.; Benner, S. A. Hachimoji DNA and RNA: A genetic system with eight building blocks. *Science (New York, N.Y.)* **2019**, *363*, 884-887.
33. Guckian, K. M.; Krugh, T. R.; Kool, E. T. Solution structure of a DNA duplex containing a replicable difluorotoluene-adenine pair. *Nat. Struct. Biol.* **1998**, *5*, 954-959.

34. Morales, J. C.; Kool, E. T. Efficient replication between non-hydrogen-bonded nucleoside shape analogs. *Nat. Struct. Biol.* **1998**, *5*, 950-954.
35. Delaney, J. C.; Henderson, P. T.; Helquist, S. A.; Morales, J. C.; Essigmann, J. M.; Kool, E. T. High-fidelity in vivo replication of DNA base shape mimics without Watson-Crick hydrogen bonds. *Proc. Natl. Acad. Sci.* **2003**, *100*, 4469-4473.
36. Hirao, I.; Kimoto, M.; Mitsui, T.; Fujiwara, T.; Kawai, R.; Sato, A.; Harada, Y.; Yokoyama, S. An unnatural hydrophobic base pair system: site-specific incorporation of nucleotide analogs into DNA and RNA. *Nat Methods* **2006**, *3*, 729-735.
37. Kimoto, M.; Kawai, R.; Mitsui, T.; Yokoyama, S.; Hirao, I. An unnatural base pair system for efficient PCR amplification and functionalization of DNA molecules. *Nucleic Acids Res* **2009**, *37*, e14.
38. Seo, Y. J.; Hwang, G. T.; Ordoukhanian, P.; Romesberg, F. E. Optimization of an unnatural base pair toward natural-like replication. *J Am Chem Soc* **2009**, *131*, 3246-3252.
39. Betz, K.; Malyshev, D. A.; Lavergne, T.; Welte, W.; Diederichs, K.; Dwyer, T. J.; Ordoukhanian, P.; Romesberg, F. E.; Marx, A. KlenTaq polymerase replicates unnatural base pairs by inducing a Watson-Crick geometry. *Nature chemical biology* **2012**, *8*, 612-614.
40. Malyshev, D. A.; Dhami, K.; Lavergne, T.; Chen, T.; Dai, N.; Foster, J. M.; Correa, I. R., Jr.; Romesberg, F. E. A semi-synthetic organism with an expanded genetic alphabet. *Nature* **2014**, *509*, 385-388.
41. Reineks, E. Z.; Berdis, A. J. Evaluating the contribution of base stacking during translesion DNA replication. *Biochemistry* **2004**, *43*, 393-404.
42. Zhang, X.; Lee, I.; Berdis, A. J. Evaluating the contributions of desolvation and base-stacking during translesion DNA synthesis. *Org. Biomol. Chem.* **2004**, *2*, 1703-1711.
43. Yang, W.; Gao, Y. Translesion and Repair DNA Polymerases: Diverse Structure and Mechanism. *Annu. Rev. Biochem.* **2018**, *87*, 239-261.
44. Stornetta, A.; Angelov, T.; Guengerich, F. P.; Sturla, S. J. Incorporation of nucleoside probes opposite *O*⁶-methylguanine by *Sulfolobus solfataricus* DNA polymerase Dpo4: importance of hydrogen bonding. *Chembiochem* **2013**, *14*, 1634-1639.
45. Gahlon, H. L.; Boby, M. L.; Sturla, S. J. *O*⁶-alkylguanine postlesion DNA synthesis is correct with the right complement of hydrogen bonding. *ACS Chem. Biol.* **2014**, *9*, 2807-2814.
46. Haracska, L.; Unk, I.; Johnson, R. E.; Johansson, E.; Burgers, P. M.; Prakash, S.; Prakash, L. Roles of yeast DNA polymerases delta and zeta and of Rev1 in the bypass of abasic sites. *Genes. Dev.* **2001**, *15*, 945-954.
47. Ráz, M. H.; Sturla, S. J.; Gahlon, H. L. Hydrogen bonding interactions at the DNA terminus promote extension from methylguanine lesions by human extender DNA polymerase ζ . *Biochemistry* **2018**.
48. Randerath, K.; Reddy, M. V.; Disher, R. M. Age- and tissue-related DNA modifications in untreated rats: detection by 32P-postlabeling assay and possible significance for spontaneous tumor induction and aging. *Carcinogenesis* **1986**, *7*, 1615-1617.
49. Turteltaub, K. W.; Felton, J. S.; Gledhill, B. L.; Vogel, J. S.; Southon, J. R.; Caffee, M. W.; Finkel, R. C.; Nelson, D. E.; Proctor, I. D.; Davis, J. C. Accelerator Mass-Spectrometry in Biomedical Dosimetry - Relationship between Low-Level Exposure and Covalent Binding of Heterocyclic Amine Carcinogens to DNA. *Proc. Natl. Acad. Sci.* **1990**, *87*, 5288-5292.
50. Singh, R.; Farmer, P. B. Liquid chromatography-electrospray ionization-mass spectrometry: the future of DNA adduct detection. *Carcinogenesis* **2006**, *27*, 178-196.

51. Clark, T. A.; Spittle, K. E.; Turner, S. W.; Korlach, J. Direct detection and sequencing of damaged DNA bases. *Genome Integr.* **2011**, *2*, 10.
52. Schibel, A. E.; An, N.; Jin, Q.; Fleming, A. M.; Burrows, C. J.; White, H. S. Nanopore detection of 8-oxo-7,8-dihydro-2'-deoxyguanosine in immobilized single-stranded DNA via adduct formation to the DNA damage site. *J. Am. Chem. Soc.* **2010**, *132*, 17992-17995.
53. Riedl, J.; Ding, Y.; Fleming, A. M.; Burrows, C. J. Identification of DNA lesions using a third base pair for amplification and nanopore sequencing. *Nat. Commun.* **2015**, *6*, 8807.
54. Hu, J.; Adebali, O.; Adar, S.; Sancar, A. Dynamic maps of UV damage formation and repair for the human genome. *Proc. Natl. Acad. Sci.* **2017**, *114*, 6758-6763.
55. Wu, J.; McKeague, M.; Sturla, S. J. Nucleotide-Resolution Genome-Wide Mapping of Oxidative DNA Damage by Click-Code-Seq. *J. Am. Chem. Soc.* **2018**, *140*, 9783-9787.
56. Wyss, L. A.; Nilforoushan, A.; Eichenseher, F.; Suter, U.; Blatter, N.; Marx, A.; Sturla, S. J. Specific incorporation of an artificial nucleotide opposite a mutagenic DNA adduct by a DNA polymerase. *J. Am. Chem. Soc.* **2015**, *137*, 30-33.
57. Wyss, L. A.; Nilforoushan, A.; Williams, D. M.; Marx, A.; Sturla, S. J. The use of an artificial nucleotide for polymerase-based recognition of carcinogenic O⁶-alkylguanine DNA adducts. *Nucleic Acids Res.* **2016**, *44*, 6564-6573.
58. Aloisi, C.M.N.; Sturla, S.J.; Gahlon, H.L. A gene-targeted polymerase-mediated strategy to identify O⁶-methylguanine damage. *Chem Commun* **2019**.
59. Imamura, Y.; Morikawa, T.; Liao, X.; Lochhead, P.; Kuchiba, A.; Yamauchi, M.; Qian, Z. R.; Nishihara, R.; Meyerhardt, J. A.; Haigis, K. M.; Fuchs, C. S.; Ogino, S. Specific mutations in KRAS codons 12 and 13, and patient prognosis in 1075 BRAF wild-type colorectal cancers. *ClinCancer Res.* **2012**, *18*, 4753-4763.
60. Betz, K.; Nilforoushan, A.; Wyss, L. A.; Diederichs, K.; Sturla, S. J.; Marx, A. Structural basis for the selective incorporation of an artificial nucleotide opposite a DNA adduct by a DNA polymerase. *Chem. Commun.* **2017**, *53*, 12704-12707.

Chapter 2: A Gene-Targeted Polymerase-Mediated Strategy to Identify O^6 -Methyl-Guanine Damage



Claudia M. N. Aloisi, Shana J. Sturla, and Hailey L. Gahlon*

Manuscript published in Chemical Communications **2019**, 55, 3895

C.M.N. Aloisi designed and performed linear amplification experiments, molecular modeling studies, interpreted data, and wrote the manuscript. S.J. Sturla interpreted data and wrote the manuscript. H.L. Gahlon synthesized oligonucleotides, performed primer extension assay, interpreted data, and wrote the manuscript.

2.1 Abstract

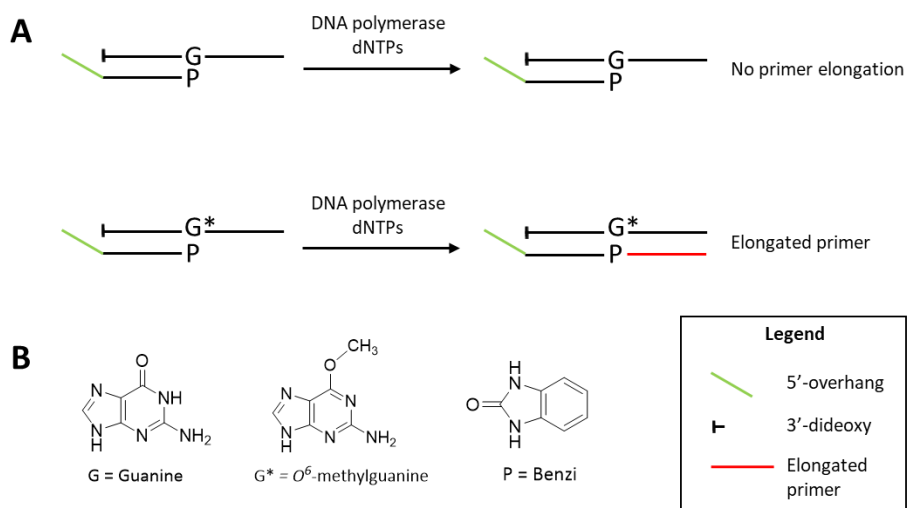
Detecting DNA adducts in cancer genes is important for understanding cancer etiology. This study reports a strategy to identify the mutagenic DNA adduct *O*⁶-methylguanine in K-Ras. The strategy involves selective replication past a synthetic primer when placed opposite *O*⁶-methylguanine. Future work can apply this approach to other cancer-relevant genes.

2.2 Introduction, results and discussion

DNA adducts contribute to cancer initiation and can be biomarkers of carcinogenesis as well as cancer therapy efficacy. Strategies to detect DNA damage often describe global damage levels, e.g. mass spectrometry, immunological assays and ³²P-postlabeling, and information regarding their genomic sequence context is missing.¹⁻⁵ In recent years, various methods have been developed to report on damage in a sequence-specific context, for instance single-molecule approaches like nanopore and SMRT sequencing as well as whole genome sequencing.⁶⁻¹⁴ Recently, the emergence of epigenetic factors being important in driving biological responses to chemicals^{15, 16} emphasizes the need for methods to define the presence of DNA damage in particular genes. However, DNA damage sequencing approaches are only recently emerging and have not been generalized to many forms of DNA damage. *O*⁶-methylguanine (*O*⁶-MeG) is a mutagenic DNA adduct that is formed upon exposure to endogenous and exogenous alkylating agents, e.g. *N*-nitroso-*N*-methylurea and *N*-nitrosodimethylamine.^{17, 18} The K-Ras gene is mutated in 43% of colon cancer cases,¹⁹ for which nitrosamine exposure may be a contributing factor, but lack of strategies to investigate the formation of *O*⁶-MeG in K-Ras or other mutation hotspot loci limits the capacity to relate mutagenesis with specific DNA methylation events. This work describes a proof of principle approach to detect the mutagenic adduct *O*⁶-MeG in the K-Ras gene sequence.

In previous studies addressing structural requirements for extension from a modified primer terminus, we reported that Dpo4, a Y-family DNA polymerase, elongated past *O*⁶-alkylguanine DNA adducts when paired opposite a synthetic nucleoside analogue.^{20, 21} This process was effective for elongation past *O*⁶-MeG when paired opposite the hydrophobic nucleoside analogue Benzi, designed to pair preferentially opposite *O*⁶-alkylguanines.²⁰⁻²³ When Benzi was paired opposite G, however, Dpo4 stalled and the primer was not elongated.²¹ We envisioned that the selective extension past *O*⁶-MeG as compared to G could be the basis of a strategy to detect *O*⁶-MeG. However, preliminary studies with oligonucleotides in the K-Ras sequence (Table 1) revealed that Dpo4 elongated past Benzi paired opposite both *O*⁶-MeG and G similarly (Fig. S4). On the other hand, replacing Dpo4 with either KlenTaq or a previously reported mutant of KlenTaq

M474K showed selective extension, i.e. Benzi primer was only enzymatically elongated when paired opposite O^6 -MeG and not G (Figs. 1, S6). In this study, we have developed a strategy to detect O^6 -MeG in the target gene sequence of K-Ras by using a 3'-Benzi-modified primer that hybridizes opposite damaged and undamaged templates and an engineered DNA polymerase that is selective and efficient for extension of modified primers (Scheme 1).



Scheme 1. (A) KlenTaq elongates specifically past O^6 -MeG:Benzi termini and does not elongate past G:Benzi in the K-Ras gene sequence. This discrimination provides a mean to detect O^6 -MeG by linear DNA amplification. Extension products are separated by polyacrylamide gel electrophoresis and visualized by SYBR Gold staining. The elongated primer (red region) is distinguishable from the O^6 -MeG- and G-containing templates by the presence of a 5' overhang (green region) on the primer. The template contains a 3'-dideoxy modification to prevent undesired extension of the template. (B) Chemical structures of guanine (G), O^6 -MeG (G*) and Benzi (P).

Table 1. Sequences for DNA templates and primers

Name	DNA Sequence (5'-3') ^[a]
G template	GTA GTT GGA GCT <u>GGT</u> GGC GTA GGC AAG
O^6 -MeG template	GTA GTT GGA GCT <u>GG*T</u> GGC GTA GGC AAG
Benzi primer 1	CTT GCC TAC GCC AP
Benzi primer 2	NCTT GCC TAC GCC AP

^[a] G* = O^6 -MeG, underlined bases = codon 12 of K-Ras, P = Benzi, N = (T)₂₁ that is represented as a 5'-overhang (green line) in Scheme 1A.

The potential of selective enzymatic extension past *O*⁶-MeG in a target gene was tested in the sequence of codon 12 of the K-Ras gene, a mutational hotspot locus which is frequently mutated in colorectal cancers.¹⁹ We performed primer extension experiments with 3'-Benzi-modified primers and templates with the sequence of K-Ras gene where the second base of codon 12 contains a G or *O*⁶-MeG (Table 1). Dpo4 elongated past Benzi paired opposite both *O*⁶-MeG and G similarly, suggesting a gap in the potential to extend the strategy to other gene-related sequences (Fig. S4). We tested the influence of the enzyme on this process, replacing Dpo4 with bypass-proficient engineered enzymes, KlenTaq and its M747K mutant.^{23, 24} The change in the enzyme resulted in selective extension, whereby Benzi primer was only enzymatically elongated when paired opposite *O*⁶-MeG but not G (Figs. 1, S6).

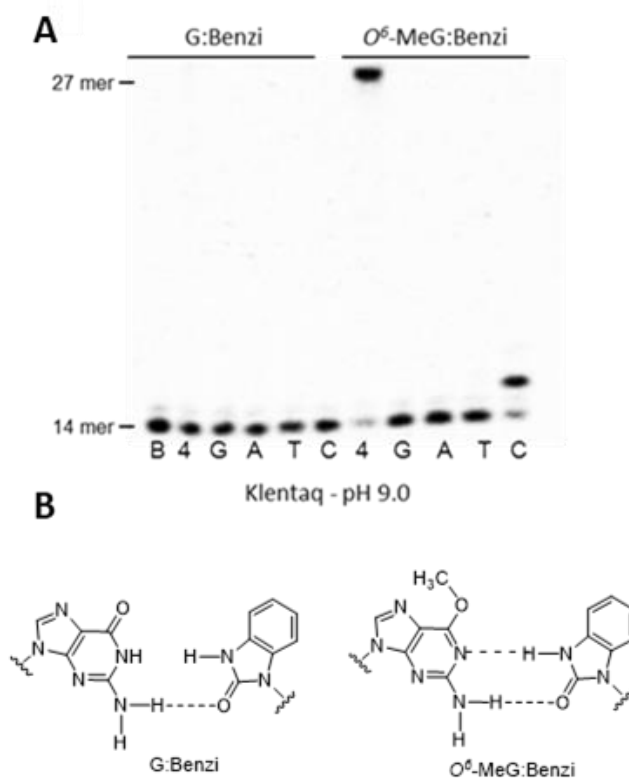


Figure 1. (A) KlenTaq-mediated elongation past G:Benzi (left) and *O*⁶-MeG:Benzi (right) at 55 °C for 10 min at pH 9.0 (Lanes in the gel are designated as follows: B = 14mer blank, 4 = all 4 canonical dNTPs, and the single nucleotide conditions include: G = only dGTP, A = only dATP, T = only dTTP, and C = only dCTP). (B) Proposed base pairing interactions for Benzi opposite G and *O*⁶-MeG.

The ability for extension by KlenTaq past Benzi-containing primers paired opposite G and *O*⁶-MeG was evaluated by further primer extension analysis. KlenTaq elongated

selectively past Benzi paired opposite O^6 -MeG, and not G, in the K-Ras DNA sequence (Table 1). For these experiments, KlenTaq DNA polymerase (25 units), DNA (10 nM) and dNTPs (10 μ M) were incubated at 55 °C for 10 min. For O^6 -MeG:Benzi DNA, KlenTaq extended in an error-free manner and only dCTP was inserted under single nucleotide conditions (Fig. 1). In the case of all four dNTPs, KlenTaq performed full-length extension for O^6 -MeG:Benzi and no extension was observed for G:Benzi. KlenTaq efficiently elongated past O^6 -MeG:Benzi DNA in the presence of all four dNTPs (90% elongation, Fig. 1A). Analogous primer extension assays were performed with a lower concentration of KlenTaq (2.5 units) and less elongation was observed for the O^6 -MeG:Benzi DNA (~65% extension, Fig. S5).

Elongation past Benzi-containing DNA was also tested with the KlenTaq mutant M747K, reported to have an increased proclivity for lesion bypass. This engineered polymerase is thought to increase favourable electrostatic interactions between the DNA backbone and the enzyme, thus promoting damage bypass.^{25, 26} For the primer extension reactions, KlenTaq M747K (20 nM), DNA (10 nM) and dNTPs (10 μ M) were incubated at 55 °C for 10 min (Fig. S6). For G:Benzi DNA, KlenTaq M747K did not catalyze DNA synthesis (Fig. S6). Conversely, for O^6 -MeG:Benzi DNA, KlenTaq M747K elongated by correctly inserting dCMP (Fig. S6). Also, for O^6 -MeG:Benzi, KlenTaq M747K elongated the primer to the full length of the template in the presence of all 4 dNTPs (~100%, Fig. S6). KlenTaq M747K-mediated elongation was also tested at 72 °C, however, no full-length extension products were observed for either G:Benzi or O^6 -MeG:Benzi (Fig. S7). A small percentage of dCMP (~10%) was inserted for O^6 -MeG:Benzi (Fig. S7). It is likely that at 72 °C the DNA is not hybridized, as the theoretical T_m is ~50 °C. Therefore, temperatures below 72 °C are necessary for optimal DNA polymerase-mediated elongation. A slight reduction in selectivity past O^6 -MeG:Benzi vs. G:Benzi was observed for KlenTaq M747K (Fig. S6) compared to wild type KlenTaq (Fig. 1). Therefore, wild type KlenTaq was used in further experiments.

Having established wild type KlenTaq as the preferred enzyme for selective extension from DNA templates containing O^6 -MeG, we tested if the Benzi primer could be linearly amplified in the presence of G and O^6 -MeG templates. For these studies, a modified Benzi oligonucleotide (Benzi primer 2, Table 1) was synthesized (Fig. S3) with a 21-nt 5'-overhang to enable resolution of amplicons (48-nt extended primer) from the initial template (27-nt). To ensure that the observed 48-nt oligo signal was exclusively a result of elongation of the Benzi primer, templates with the 3'-dideoxy ends were used. To achieve the highest selectivity and enzyme efficiency, we optimized temperature and time of each reaction step, cycle number, concentration of primer and templates. In addition, we evaluated the effect of varying concentrations of Mg^{2+} ions (Fig. S8). Final conditions for the amplification reactions consisted of Benzi primer 2 (200 nM) incubated

with either *O*⁶-MeG or G template (2 nM each template), KlenTaq (20 nM) and in the presence of all four dNTPs (250 μM). Alternating cycles (35) of DNA melting (95°C, 30 s), annealing (40°C, 30 s) and extension step (55°C, 30 s) were performed and products were separated on polyacrylamide gels and visualized by staining with SYBR gold. Under these linear amplification conditions, selective elongation for *O*⁶-MeG:Benzi DNA was observed (Fig. 2A), corroborating findings from the primer extension data (Fig. 1).

To test if the selective elongation from *O*⁶-MeG:Benzi could be a viable strategy for identifying *O*⁶-MeG, we performed linear amplification studies in a mixture of damaged and undamaged DNA. Here, the concentration of G template was constant (2 nM) while increasing the concentrations of *O*⁶-MeG (0, 0.5, 1, 1.5, 2 nM). In the presence of a mixture of DNA templates, the signal increased as a function of increasing concentrations of *O*⁶-MeG (Fig. 2B). The ratio of *O*⁶-MeG to G plotted against the corresponding amount of extension showed a linear relationship (Fig. 2B). In a 50:50 mixture of *O*⁶-MeG:G, 15% of extension was observed and extension doubled (30%) with a 2-fold increase in the ratio of *O*⁶-MeG:G. These results demonstrate Benzi as an effective chemical marker to detect *O*⁶-MeG in a mixture of undamaged and damaged DNA.

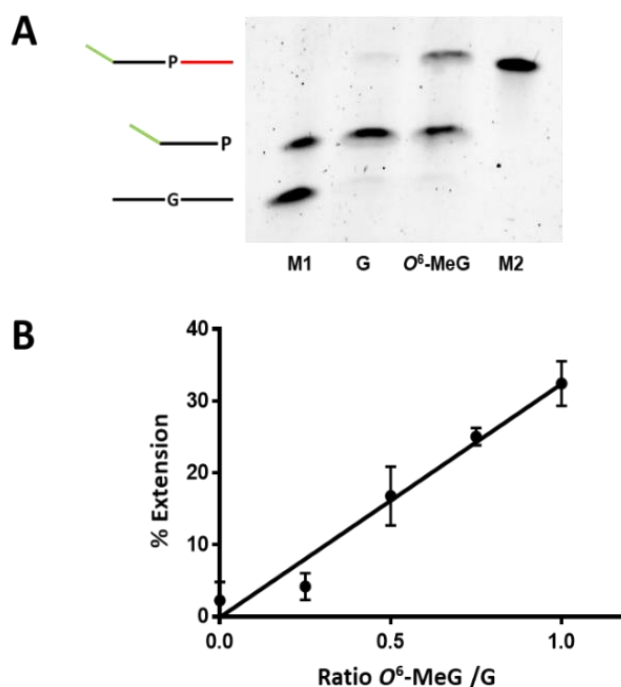


Figure 2. KlenTaq-mediated elongation past G:Benzi and *O*⁶-MeG:Benzi with Benzi primer 2 (Table 1). Linear amplification conditions comprised of (1) denaturation at 95°C for 2 min; (2) 35 cycles alternating steps of denaturation at 95°C for 30 s, annealing at 40°C for 30 s, extension at 55°C for 30 s; (3) final extension at 55°C for 3 min; pH 9, 1XKTQ buffer. (A) Gel separated amplicons (M1 = marker lane with 27mer G template and 35mer Benzi

primer 2; G = reaction with G template; O^6 -MeG = reaction with O^6 -MeG template; M2 = marker lane with 48mer extended primer). Note: Initial templates are not detected by SYBR Gold staining as they are below the limit of detection. (B) Linear amplification was performed in the presence of G and O^6 -MeG templates with increasing concentrations of O^6 -MeG template and a constant concentration of G template (2 nM). All reactions were performed in three independent technical replicates ($y = 32.49x - 0.11$; $R^2 = 0.97$).

To understand structural interactions that contribute to the preferred extension of Benzi paired opposite O^6 -MeG by KlenTaq, we performed computational modelling studies. Previous X-ray data with Benzi as an incoming triphosphate (BenziTP) paired opposite O^6 -MeG (i.e. ternary complex; Fig. 3A) showed BenziTP pairs opposite O^6 -MeG with two hydrogen bond interactions.²⁶ Starting from the ternary complex crystal structure (PDB ID: 3RTV)²⁶ with wild type KlenTaq, undamaged primer:template DNA, and an incoming dCTP we (1) removed the incoming dCTP (2) replaced the terminal 3' nucleotide on the primer strand with Benzi, and (3) inserted G and O^6 -MeG in the template (Fig 3B) to reflect the configuration of the experiments described above.

We evaluated the distance and orientation of the terminal base pairs in the active site of KlenTaq. First, we compared deoxyribose distances for O^6 -MeG:Benzi and the natural G:C base pairs. We found that the C1'-C1' distance for the G:C base pair (10.5 Å) is comparable to the O^6 -MeG:Benzi base pair (10.4 Å and 10.2 Å for the methyl group in a distal and proximal conformation, respectively, Fig. S9). These distances are similar to what was reported for the X-ray structure; O^6 -MeG:Benzi with a slightly longer C1'-C1' distance of 11.0 Å and G:C of 10.6 Å.²⁶ We speculate that the closer distance of O^6 -MeG:Benzi in our modelled data may result from the slight torsion that Benzi imposes when covalently bound to the DNA primer (binary complex, Fig. 3B). Next, we compared the orientation of Benzi in the enzymatic pocket when paired opposite O^6 -MeG and G. We observed minor rotations on amino acid residues; however, the overall orientation of Benzi was not sufficiently different when paired opposite damaged and undamaged DNA.

Overall energies and hydrogen bond interactions were determined for terminal base pairs in the active site of KlenTaq. We determined the computed energies of the terminal O^6 -MeG:Benzi and G:Benzi base pairs. These calculated energies take into account both Van der Waals and hydrogen bond interactions. Computed energies of -0.3 kcal/mol for O^6 -MeG:Benzi and +26 kcal/mol for G:Benzi were determined. Further, we found that the hydrogen bond pattern in our model corroborates the previous X-ray data.²⁶ For example, Benzi and O^6 -MeG interact by two hydrogen bonds; one hydrogen bond between the N1 on O^6 -MeG and the -NH donor on Benzi (2.19 Å, Fig. 3C) and another between the NH₂ donor on O^6 -MeG and the carbonyl group on Benzi (1.79 Å when the

methyl group is in the distal conformation, Fig. 3C; 1.96 Å in the proximal conformation, Fig. S9). Further, G:Benzi displays a potential steric clash between the -NH moiety on Benzi and the -NH at the N1 position on G (Fig. 3C), possibly impeding KlenTaq catalysis. On the contrary, *O*⁶-MeG:Benzi is more analogous to the natural G:C Watson-Crick base pair, but with two hydrogen bonds. This more favourable configuration of *O*⁶-MeG:Benzi vs. G:Benzi may explain the observed preferential elongation by KlenTaq.

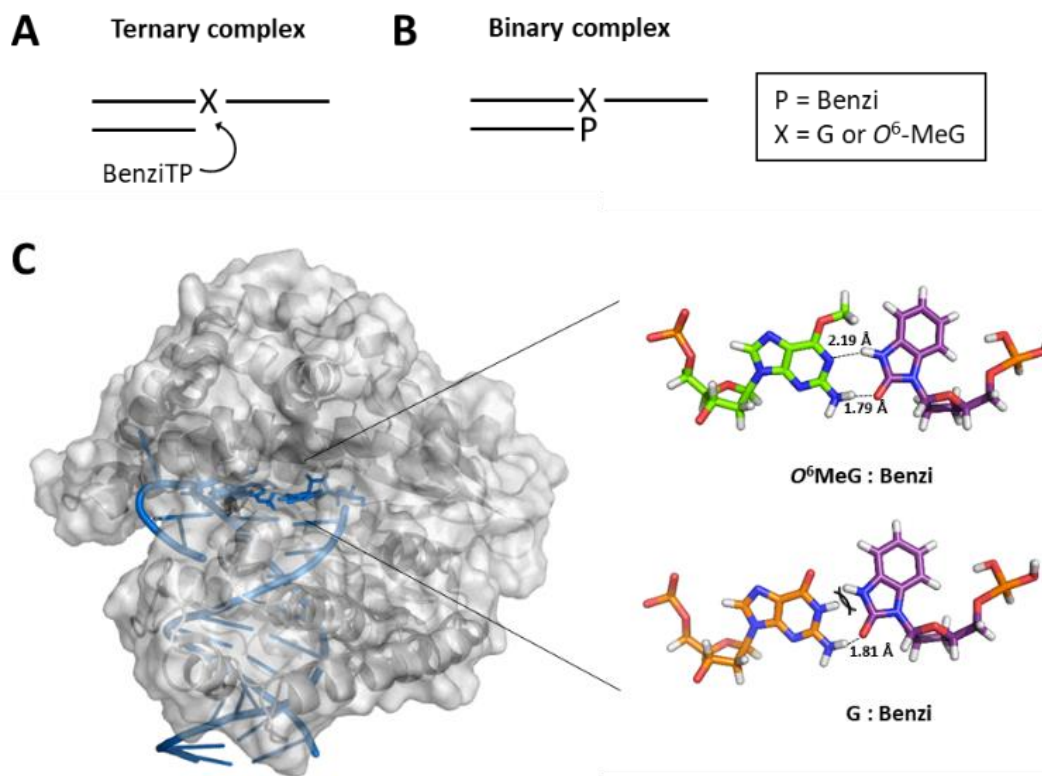


Figure 3. (A) Scheme of the ternary complex previously crystallized for X-ray analysis (X = G or *O*⁶-MeG). (B) Scheme of the binary complex with Benzi covalently bound to the 3' end of the primer (P= Benzi). (C) KlenTaq (PDB ID: 3RTV) was energy minimized with imidicOperating Environment software. The active site of KlenTaq (zoomed-in region) shows Benzi paired opposite *O*⁶-MeG (top) and G (bottom). Two hydrogen bond interactions were computed for *O*⁶-MeG:Benzi and one hydrogen bond for G:Benzi.

In conclusion, this work presents a first enzyme-based strategy reporting on the presence of *O*⁶-MeG in the cancer-relevant gene sequence of K-Ras by means of synthetic DNA-based amplification. This strategy works by identifying *O*⁶-MeG in a sequence-specific manner and, therefore, could be useful for linking modifications in

particular genes with specific mutational outcomes, which is not possible using typical DNA adduct detection methods that generally lose information about adduct sequence. In this strategy, damage is detected in a pre-defined locus. Compared to genome-wide damage mapping strategies, this approach provides a targeted strategy that could be adapted to other cancer genes, thereby providing defined chemical resolution of damage in hotspot regions of genes. Future work will require assessing adduct selectivity, as well as addressing the need for increased sensitivity to detect damage at biologically relevant levels. In addition, this approach could be adapted to detecting DNA damage in biological samples; however, this will require an enrichment strategy to assess low abundance forms of damage. This approach could be expanded to detect O^6 -MeG in other cancer-relevant gene sequences, thereby leading to a better understanding of the relationship between DNA damage and disease initiation and progression.

2.3 Supplementary Information (Experimental Section)

Phosphoramidite synthesis

Benzi phosphoramidite was synthesized as previously described.⁸¹

Oligonucleotide synthesis

Modified DNA was synthesized by solid-phase DNA synthesis on a Mermade 4 DNA synthesizer (Bioautomation Corporation) in trityl-off mode. Universal Q SynBase™ CPG support columns were used as well as natural nucleotide phosphoramidites from Link Technologies Ltd. (Lanarkshire, Scotland). O^6 -MeG phosphoramidite was purchased from Glen Research (Sterling, VA, USA). Oligonucleotide deprotection was carried out by treatment with 30% aqueous ammonium hydroxide at 55 °C for 10 hours, except in the case for the O^6 -MeG DNA where a 1:1 solution of aqueous ammonium hydroxide and *N*-methylamine was used. The oligonucleotides were dried under reduced pressure, re-suspended in DNase/RNase free water and filtered with 0.45 μM nylon filters (Merck Millipore) before HPLC purification. The oligonucleotide was purified by reverse phase HPLC using an Agilent eclipse XDB C-18 5 μM 4.6 x 150 mm column. The mobile phase consisted of 50 mM triethylammonium acetate (TEAA) and acetonitrile (ACN). The flow rate was set to 1.0 mL/min with a mobile phase starting at from 0-10% ACN over 18 min, increasing to 50% ACN from 18.1-20 min, and equilibrated from 20.1-25 min at 10% ACN. Fractions corresponding to the desired oligonucleotides were collected, lyophilized and re-suspended in DNase/RNase free water. The Benzi Primer 1 and Benzi Primer 2 DNA (DNA sequences see Table 1) were purified by polyacrylamide gel electrophoresis. The DNA concentrations were measured by UV absorbance on a Varian Cary 100 Bio UV/Vis spectrophotometer.

Primer extension analysis

The Benzi Primer 1 (Table 1) DNA was 5'-end labelled using ³²P-ATP (Perkin Elmer) and T4 polynucleotide kinase (Promega Corp, Madison, WI, USA). The labelled Benzi DNA was annealed to complementary templates containing G or *O*⁶-MeG, Codon-12 G and Codon-12 *O*⁶-MeG in Table 1, respectively by heating the duplexes at 95 °C for 5 min and slow cooling overnight. KlenTaq M747K (a gift from Andreas Marx, University of Konstanz, Germany) primer extension reactions (10 µl) contained 20 nM enzyme, 10 nM DNA, 10 µM of each dNTP, 50 mM Tris HCl (pH 9.2), 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 0.1% Tween 20. The reactions were incubated at either 55 or 72 °C for 10 min. For Dpo4-mediated reactions (10 µl), 20 nM enzyme (Trevigen), 10 nM DNA, 10 µM of each dNTP, 50 mM Tris-HCl (pH 8.0), 2.5 mM MgCl₂, 50 mM NaCl, 5 mM DTT, 100 µg/ml bovine serum albumin, 5% glycerol was used and reactions were incubated at 37 °C for 30 min. For KlenTaq 1 (DNA Polymerase Technologies, St. Louis, MO, USA) reactions (10 µl), 25 units enzyme, 10 nM DNA, 10 µM of each dNTP, 250 mM Tris-HCl (pH 7.9), 15 mM MgCl₂, 0.13% Brij 58, 40 mM (NH₄)₂SO₄ was used and reactions were incubated at 55 °C for 10 min. Reactions were terminated by adding 10 µl of stop buffer (95% formamide, 20 mM EDTA and 0.5% xylene cyanol and bromophenol blue). Reaction mixtures (4 µl) were loaded onto 15% polyacrylamide/7 M urea gels and extension products were visualized from a phosphorimaging screen with a phosphorimager (BioRad, Hercules, CA, USA).

Molecular modelling studies

Structures were computed with the Molecular Operating Environment software (Chemical Computing Group). Crystal structure of WT KlenTaq with incoming ddCTP opposite template G (G1) was used (PDB ID: 3RTV).²⁶ PDB file was loaded as Molecular Assembly. The ternary complex of crystal structure KlenTaq:template:primer was prepped by Quick Prep option. Base pair before G1 in the crystal structure was G:C (G0:C0, template:primer). For modelling studies, C0 was replaced by Benzi by using the software Molecular Builder; G0 was kept as G for modelling studies on unmodified template, or replaced by *O*⁶-MeG for modelling studies on *O*⁶-MeG template. After bases replacement, for energy minimizations, the potential energy of the protein was fixed at distance 4.5 to 9 Å from Benzi-base pairs, and tethered at distance >9 Å, by applying Amber 10 ETH force field, with 0.1 gradient. Visualization was performed in the PyMol software (Schrodinger) and images were rendered with ray 2400.

Linear amplification

Reaction mixtures (25 µl) contained 20 nM of KlenTaq (myPols), 1x KTQ buffer (50 mM Tris HCl (pH 9.1), 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, 0.1% Tween 20), 250 µM of each dNTP, 200 nM of the Benzi primer 2, and 2 nM of 27mer template (Fig. 3). Reactions with increasing ratio of *O*⁶-MeG/G (Fig. 4) contained 2 nM of G 27mer template and x nM of *O*⁶-MeG 27mer template (x = 0; 0.5; 1; 1.5; 2 nM). Reactions were performed on a T3000 Thermocycler (Biometra). An initial

denaturation step at 95 °C for 2 min was performed, followed by 35 cycles of 95 °C for 30 s, 40 °C for 30 s and 55 °C for 30 s, followed by a final step of elongation at 55°C for 3 min. For imaging, 10 µl of each reaction were added to 10 µl of 2% SDS in formamide, boiled for 10 min at 100 °C, then immediately placed on ice, and loaded on a home-made 15% acrylamide / 7 M urea gel, which was previously equilibrated for 30 min with 1x TBE (100 mM Tris base, 100 mM boric acid, 2 mM EDTA). Gels were run at 300 V, at rt for 30 min, and at 4 °C for further 2-2.5 h. Gels were imaged on a ChemiDoc MP Imaging System (BioRad), after 10 min incubation in 1x SYBR gold (Invitrogen) in TBE.

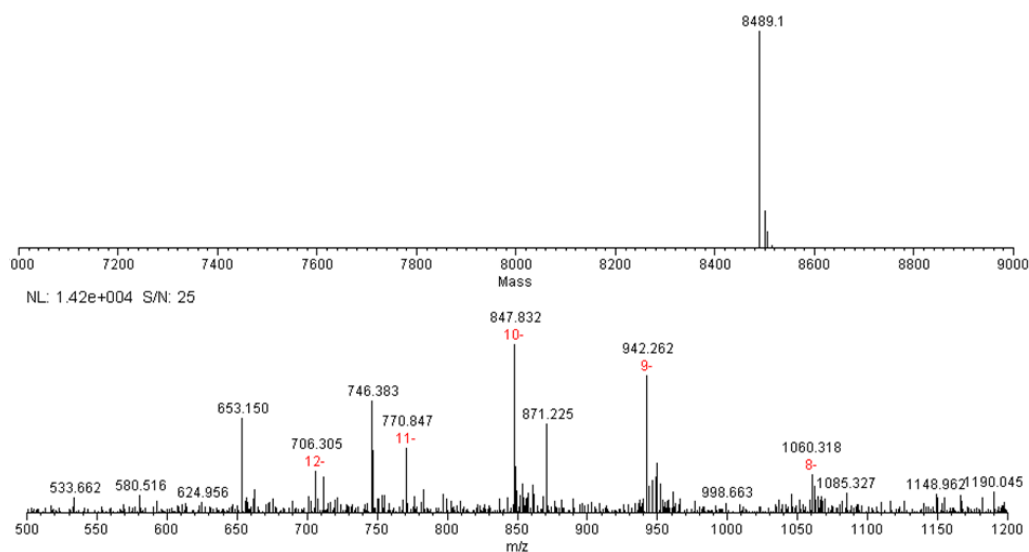


Figure S1. Mass spectra of codon-12 O^6 -MeG DNA; sequence: 5'- GTA GTT GGA GCT GXT GGC GTA GGC AAG-3' (X = O^6 -MeG, parent mass 8489)

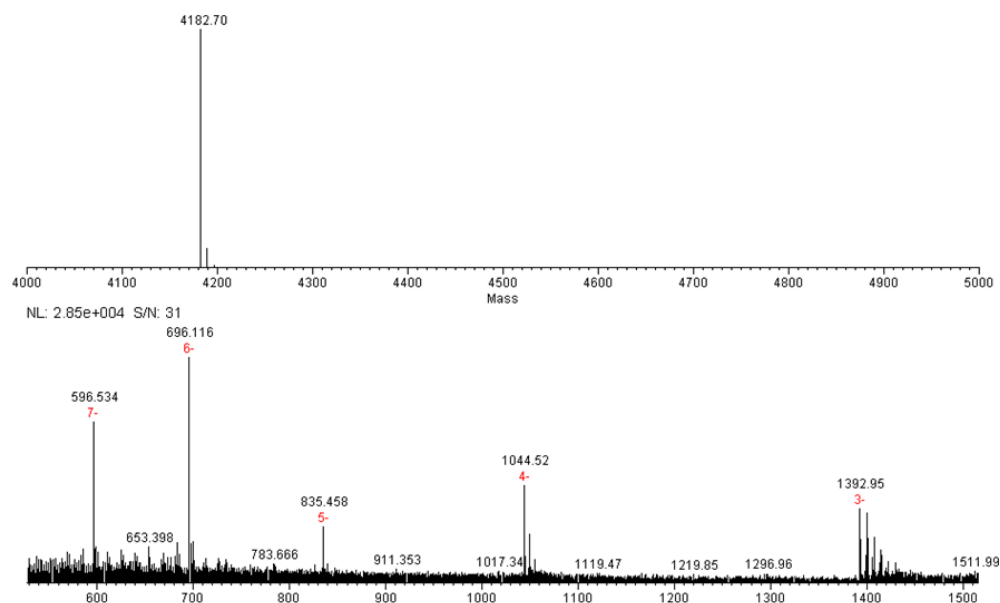


Figure S2. Mass spectra of codon-12 Benzi DNA; sequence: 5'- CTT GCC TAC GCC AP-3' (P = Benzi, parent mass 4182)

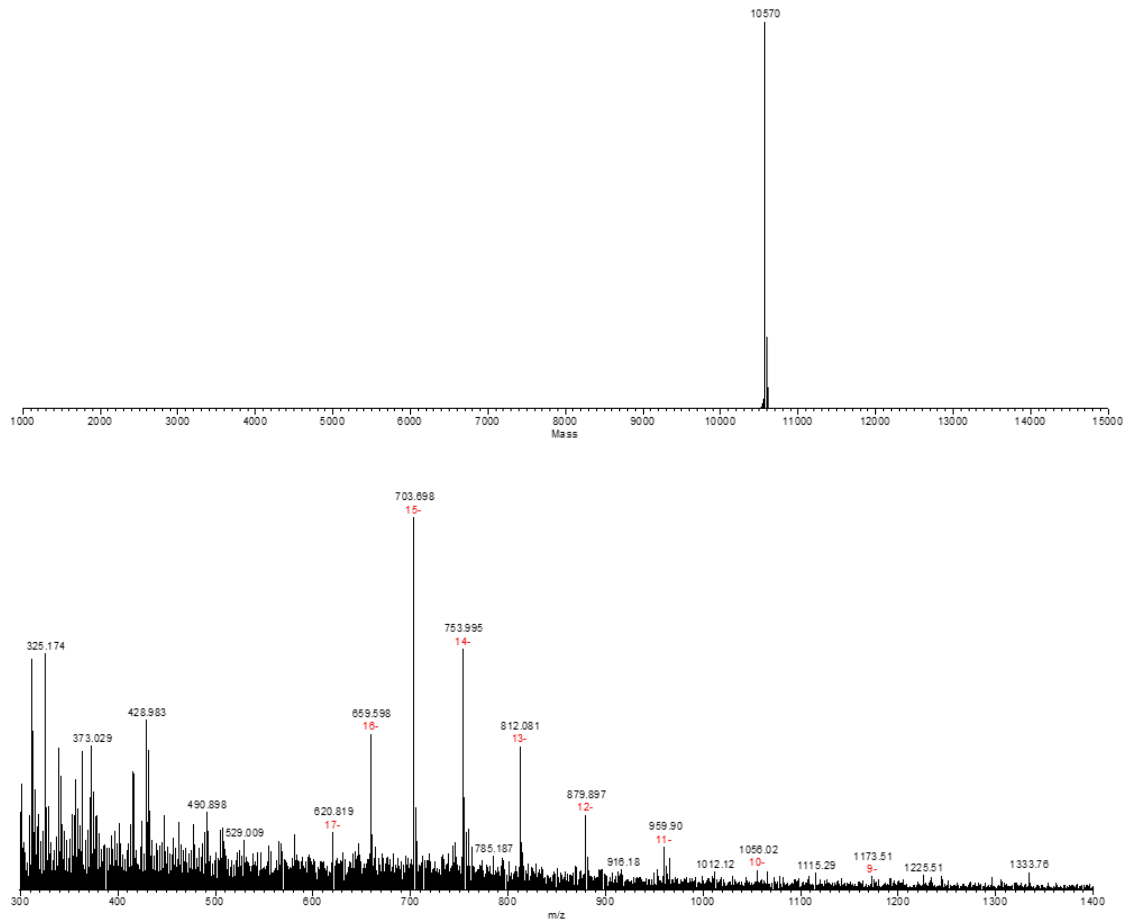


Figure S3. Mass spectra of Benzi-dT₂₁ DNA; sequence: 5'- T₂₁ CTT GCC TAC GCC AP-3' (P = Benzi, parent mass 10570)

A) Dpo4 primer extension for K-Ras template (26 °C)

B) Dpo4 primer extension for K-Ras template (37 °C)

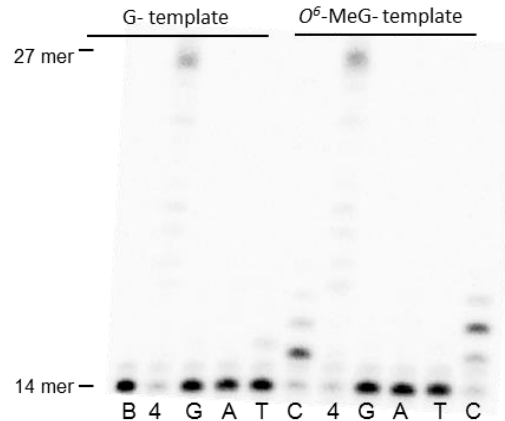
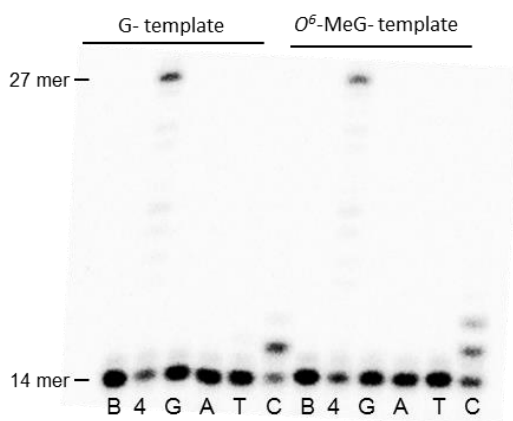


Figure S4. Dpo4-mediated elongation past Benzi:G and Benzi: O^6 -MeG at (A) 26 °C (B) 37 °C.

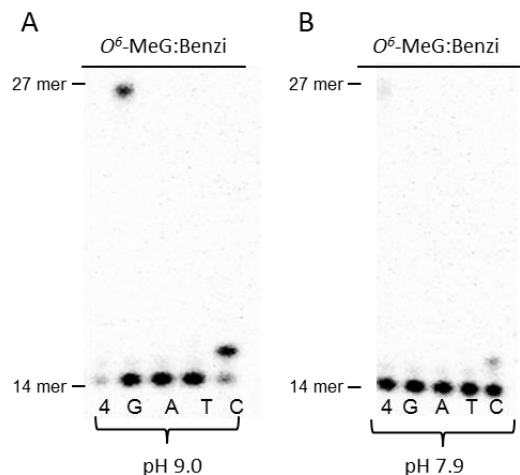


Figure S5. KlenTaq (2.5 units) elongation past Benzi:*O*⁶-MeG at 55 °C (A) pH 9 and (B) and pH 7.9.

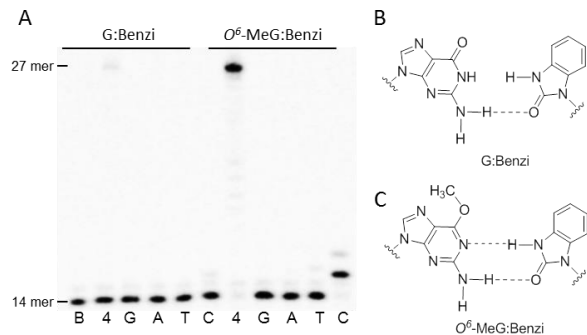


Figure S6. (A) KlenTaq M747K-mediated elongation past G:Benzi and *O*⁶-MeG:Benzi primer template DNA at 55 °C for 10 min (B = 14mer blank, 4 = 4 canonical dNTPs, G = dGTP, A = dATP, T = dTTP, and C= dCTP) and (B) proposed hydrogen bonding between *O*⁶-MeG and Benzi.

-KTQ primer extension for K-Ras template (72 °C)

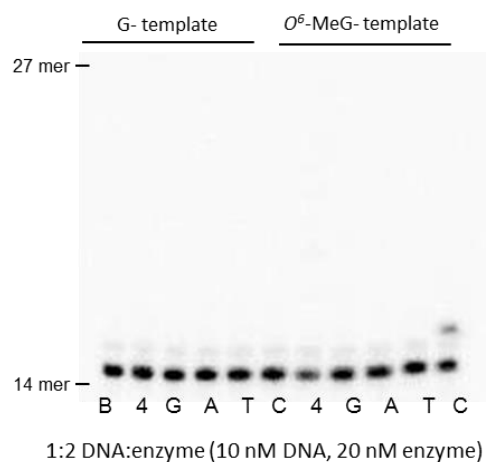


Figure S7. KlenTaq M747K elongation past Benzi:G and Benzi:*O*⁶-MeG at 72 °C.

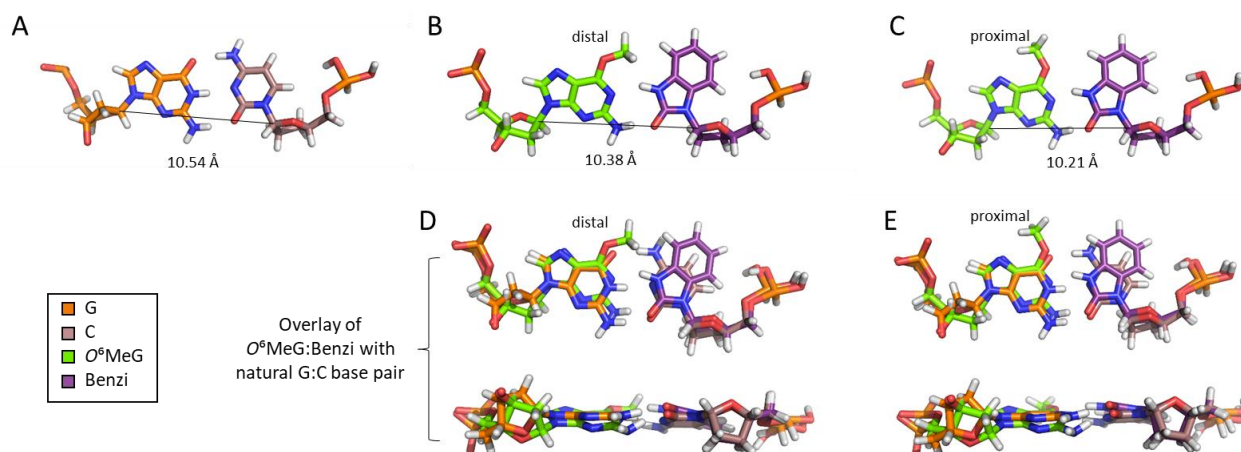


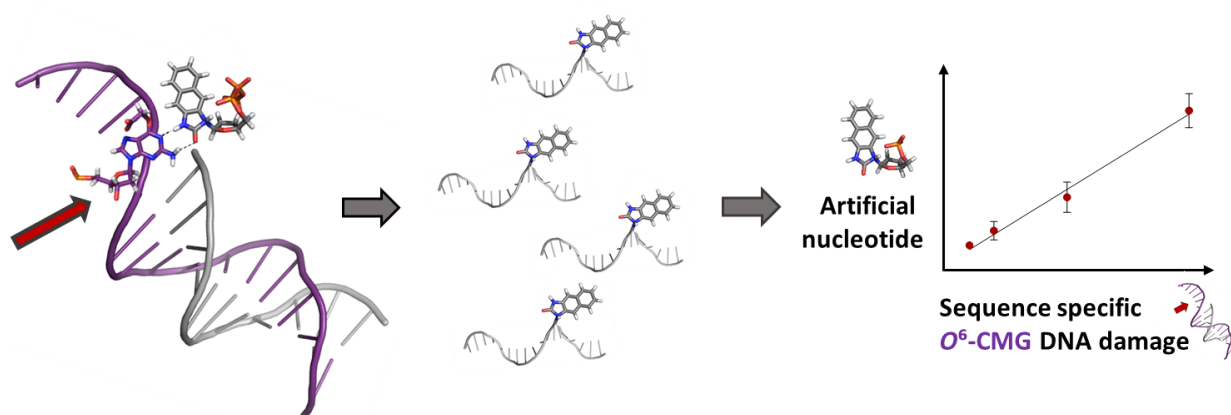
Figure S8. Modelled structures with wild type KlenTaq with C1'-C1' distances shown in Å for terminal template:primer base pairs (A) canonical G:C (B) O^6 -MeG:Benzi with the methyl group in distal conformation and (C) O^6 -MeG:Benzi with the methyl group in proximal conformation. In addition, overlay structures comprised of G:C and O^6 -MeG:Benzi where determined for (D) the methyl in distal conformation and (E) the methyl in the proximal conformation.

2.4 References

1. Guo, J.; Villalta, P. W.; Turesky, R. J., Data-Independent Mass Spectrometry Approach for Screening and Identification of DNA Adducts. *Analytical chemistry* **2017**, *89* (21), 11728-11736.
2. Singh, R.; Farmer, P. B., Liquid chromatography-electrospray ionization-mass spectrometry: the future of DNA adduct detection. *Carcinogenesis* **2006**, *27* (2), 178-196.
3. Brown, K., Methods for the detection of DNA adducts. *Methods in Mol. Biol.* **2012**, *817*, 207-230.
4. Jones, N. J., 32 P-postlabelling for the sensitive detection of DNA adducts. *Methods in Mol. Biol.* **2012**, *817*, 183-206.
5. Phillips, D. H.; Farmer, P. B.; Beland, F. A.; Nath, R. G.; Poirier, M. C.; Reddy, M. V.; Turteltaub, K. W., Methods of DNA adduct determination and their application to testing compounds for genotoxicity. *Environ.Mol. Mutagen.* **2000**, *35*, 222-233.
6. Hu, J.; Adar, S.; Selby, C. P.; Lieb, J. D.; Sancar, A., Genome-wide analysis of human global and transcription-coupled excision repair of UV damage at single-nucleotide resolution. *Genes Dev.* **2015**, *29* (9), 948-60.
7. Li, W.; Hu, J.; Adebali, O.; Adar, S.; Yang, Y.; Chiou, Y.-Y.; Sancar, A., Human genome-wide repair map of DNA damage caused by the cigarette smoke carcinogen benzo[a]pyrene. *Proceedings of the National Academy of Sciences of the United States of America* **2017**, *114* (26), 6752-6757.
8. Wu, J.; McKeague, M.; Sturla, S. J., Nucleotide-Resolution Genome-Wide Mapping of Oxidative DNA Damage by Click-Code-Seq. *J. Am. Chem. Soc.* **2018**, *140* (31), 9783-9787.
9. Ding, J.; Taylor, M. S.; Jackson, A. P.; Reijns, M. A. M., Genome-wide mapping of embedded ribonucleotides and other noncanonical nucleotides using emRiboSeq and EndoSeq. *Nat. Protocols* **2015**, *10*, 1433.

10. Hu, J.; Lieb, J. D.; Sancar, A.; Adar, S., Cisplatin DNA damage and repair maps of the human genome at single-nucleotide resolution. *Proceedings of the National Academy of Sciences* **2016**, *113* (41), 11507.
11. An, N.; Fleming, A. M.; White, H. S.; Burrows, C. J., Nanopore detection of 8-oxoguanine in the human telomere repeat sequence. *ACS nano* **2015**, *9* (4), 4296-307.
12. Trantakis, I. A.; Nilforoushan, A.; Dahlmann, H. A.; Stäuble, C. K.; Sturla, S. J., In-Geno Quantification of O^6 -Methylguanine with Elongated Nucleoside Analogues on Gold Nanoprobes. *Journal of the American Chemical Society* **2016**, *138* (27), 8497-8504.
13. Hu, J.; Adebali, O.; Adar, S.; Sancar, A., Dynamic maps of UV damage formation and repair for the human genome. *Proceedings of the National Academy of Sciences of the United States of America* **2017**, *114* (26), 6758-6763.
14. Flusberg, B. A.; Webster, D. R.; Lee, J. H.; Travers, K. J.; Olivares, E. C.; Clark, T. A.; Korlach, J.; Turner, S. W., Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat. Methods* **2010**, *7*, 461.
15. Fleming, A. M.; Ding, Y.; Burrows, C. J., Oxidative DNA damage is epigenetic by regulating gene transcription via base excision repair. *Proc. Natl. Acad. Sci. U S A* **2017**, *114* (10), 2604-2609.
16. Pfeifer, G. P., Defining Driver DNA Methylation Changes in Human Cancer. *Int. J. Mol. Sci.* **2018**, *19* (4).
17. Likhachev, A. J.; Ivanov, M. N.; Bresil, H.; Planche-Martel, G.; Montesano, R.; Margison, G. P., Carcinogenicity of single doses of N-nitroso-N-methylurea and N-nitroso-N-ethylurea in Syrian golden hamsters and the persistence of alkylated purines in the DNA of various tissues. *Cancer Res.* **1983**, *43* (2), 829-833.
18. Arimoto-Kobayashi, S.; Kaji, K.; Sweetman, G. M. A.; Hayatsu, H., Mutation and formation of methyl- and hydroxylguanine adducts in DNA caused by N-nitrosodimethylamine and N-nitrosodiethylamine with UVA irradiation. *Carcinogenesis* **1997**, *18* (12), 2429-2433.
19. Jones, R. P.; Sutton, P. A.; Evans, J. P.; Clifford, R.; McAvoy, A.; Lewis, J.; Rousseau, A.; Mountford, R.; McWhirter, D.; Malik, H. Z., Specific mutations in KRAS codon 12 are associated with worse overall survival in patients with advanced and recurrent colorectal cancer. *Br. J. Cancer* **2017**, *116* (7), 923-929.
20. Gahlon, H. L.; Schweizer, B. W.; Sturla, S. J., Tolerance of base pair size and shape in postlesion DNA synthesis. *J. Am. Chem. Soc.* **2013**, *135* (17), 6384-6387.
21. Gahlon, H. L.; Boby, M. L.; Sturla, S. J., O^6 -Alkylguanine Postlesion DNA Synthesis Is Correct with the Right Complement of Hydrogen Bonding. *ACS Chem. Biol.* **2014**, *9* (12), 2807-2814.
22. Gahlon, H. L.; Sturla, S. J., Hydrogen bonding or stacking interactions in differentiating duplex stability in oligonucleotides containing synthetic nucleoside probes for alkylated DNA. *Chem. Eur. J.* **2013**, *19* (33), 11062-11067.
23. Wyss, L. A.; Nilforoushan, A.; Williams, D. M.; Marx, A.; Sturla, S. J., The use of an artificial nucleotide for polymerase-based recognition of carcinogenic O^6 -alkylguanine DNA adducts. *Nucleic Acids Res.* **2016**, *44* (14), 6564-6573.
24. Obeid, S.; Schnur, A.; Gloeckner, C.; Blatter, N.; Welte, W.; Diederichs, K.; Marx, A., Learning from directed evolution: *Thermus aquaticus* DNA polymerase mutants with translesion synthesis activity. *ChemBioChem* **2011**, *12*, 1574-1580.
25. Gloeckner, C.; Sauter, K. B. M.; Marx, A., Evolving a Thermostable DNA Polymerase That Amplifies from Highly Damaged Templates. *Angew. Chem. Int. Edit.* **2007**, *46* (17), 3115-3117.
26. Betz, K.; Nilforoushan, A.; Wyss, L. A.; Diederichs, K.; Sturla, S. J.; Marx, A., Structural basis for the selective incorporation of an artificial nucleotide opposite a DNA adduct by a DNA polymerase. *Chemical Communications* **2017**, *53* (94), 12704-12707.

Chapter 3: Sequence-specific Quantitation of Mutagenic DNA damage via Polymerase Amplification with an Artificial Nucleotide



Claudia M.N. Aloisi, Arman Nilforoushan, Nathalie Ziegler, and Shana J. Sturla*

Reproduced with permission from ACS, *manuscript submitted*

Unpublished work copyright © 2019 American Chemical Society

C.M.N. Aloisi synthesized ExBenziTP, performed primer extension assays, steady-state kinetics, molecular modeling studies, linear amplification experiments, DNA sample hydrolysis and purification for mass spectrometry, and mass spectrometry analysis, interpreted data, and wrote the manuscript. A. Nilforoushan synthesized ExBenziTP and interpreted data. N. Ziegler created the mass spectrometry method, performed mass spectrometry analysis, and wrote the corresponding experimental section. S.J. Sturla interpreted data, and wrote the manuscript.

3.1 Abstract

DNA mutations can result from replication errors due to different forms of DNA damage, including low abundance DNA adducts induced by reactions with electrophiles. The lack of strategies to measure DNA adducts within genomic loci, however, limits our understanding of chemical mutagenesis. The use of artificial nucleotides incorporated opposite DNA adducts by engineered DNA polymerases offers a potential basis for site-specific detection of DNA adducts, but the availability of effective artificial nucleotides that insert opposite DNA adducts is extremely limited, and furthermore, there has been no report of a quantitative strategy for determining how much DNA alkylation occurs in a sequence of interest. In this work, we synthesized an artificial nucleotide triphosphate that is selectively inserted opposite *O*⁶-carboxymethyl-guanine DNA by an engineered polymerase and is required for DNA synthesis past the adduct. We characterized the mechanism of this enzymatic process and demonstrated that the artificial nucleotide is a marker for the presence and location in the genome of *O*⁶-carboxymethyl-guanine. Finally, we established a mass spectrometric method for quantifying the incorporated artificial nucleotide and obtained a linear relationship with the amount of *O*⁶-carboxymethyl-guanine in the target sequence. In this work, we present a strategy to identify, locate and quantify a mutagenic DNA adduct, advancing tools for linking DNA alkylation to mutagenesis and for detecting DNA adducts in genes as potential diagnostic biomarkers for cancer prevention.

3.2 Introduction

DNA integrity is continuously threatened by endogenous and exogenous DNA-reactive chemicals. The resulting chemical adducts to DNA can initiate adverse biological consequences including cell death and mutation. *O*⁶-alkyl-guanines (*O*⁶-alkylGs) are mutagenic DNA adducts that have been linked to carcinogenesis.¹⁻⁴ They can form from anticancer drugs,⁵ antibiotics,⁶ and environmental exposures such as cigarette smoke or red and processed meat consumption.⁷⁻¹¹ In particular, *O*⁶-carboxymethyl-guanine (*O*⁶-CMG; Fig. 1B) was significantly higher in exfoliated colonocytes of people who eat meat vs. vegetarians,¹⁰ and has been detected upon *in vitro* gastrointestinal digestion of meat,¹¹ suggesting its formation as a molecular initiating event in colorectal carcinogenesis (CRC) linked to meat consumption.¹¹⁻¹³

Mutation spectra from colon cells of people with CRC resembles the spectrum induced by the carboxymethylating agent potassium diazoacetate (KDA).¹⁴ Spectra measured in the cancer-relevant gene *p53* of CRC tissues and in a KDA-treated *p53* gene-containing plasmid were rich in C>T transition mutations, potentially arising from the misincorporation of T opposite *O*⁶-CMG

during replication, as suggested by *in vitro* studies interrogating the fidelity of replication past O^6 -CMG.¹⁵⁻¹⁷ When O^6 -CMG was placed in a DNA template and replicated by translesion DNA synthesis (TLS) polymerases (Pols), it was found that O^6 -CMG promotes the misincorporation of bases by Y- and B-family TLS Pols.¹⁵ In another study, by transfecting a DNA plasmid containing O^6 -CMG in human cells, it was concluded that O^6 -CMG moderately blocks DNA replication and induces mutations at substantial frequencies.¹⁶ While there is evidence regarding the mutagenic potential of O^6 -CMG, there is however a lack of data concerning the presence and accumulation of O^6 -CMG in CRC hotspot regions of the genome.

To establish a cause-effect relationship between O^6 -CMG and CRC, the detection and quantification of DNA damage in the genome is critically needed. O^6 -CMG is typically detected by P32 postlabeling,¹⁸ immunoblotting,¹⁹ affinity chromatography,⁸ and mass spectrometry.²⁰⁻²² The most sensitive O^6 -CMG measurement was by ESI-MS³ with a LOQ of 73.4 amol,²² allowing detection of 0.05 O^6 -CMG per 10^7 nucleotides in human cells. However, for all approaches, only the total level of damage can be determined and information on the genomic location of O^6 -CMG is lost.

A nanopore-based O^6 -CMG sequencing approach has been recently applied to study the behavior of O^6 -CMG during replication by Phi29 DNA Pol.²³ In this system, the presence of the adduct is associated with an alteration of current signal only for pre-defined arrangements of 4-base contexts. Additionally, the characteristic signal was obtained by using a low throughput nanopore that is not commercially available.

Earlier, DNA adduct-directed artificial nucleotides have been developed as a basis for interrogating damaged DNA in duplex hybridization or polymerase-mediated synthesis contexts.²⁴ We have reported a variety of nucleoside analogues that stabilize damaged DNA vs. undamaged DNA.^{25, 26} The heterocyclic imidic nucleoside analogues Benzi and ExBenzi (Fig. 1B) consist of a conjugated π -system, a hydrogen bond accepting carbonyl group and an imino-based hydrogen bond donor and⁴, when paired opposite O^6 -alkylG adducts, form a stabilized duplex.²⁷⁻³¹ When ExBenzi was included in a short oligonucleotide attached to gold nanoparticles, the presence of O^6 -alkylG adducts in target DNA strands disrupted nanoparticle aggregation and induced a color change, indicative of the presence of the adduct.²⁷ These hybridization probes provide an excellent quantitative read-out, however, do not involve any DNA amplification, and it is restricted in its applicability by the limit of detection of 138 fmol of modified DNA measured in the presence of 6 pmol of DNA.

Polymerase-mediated amplification has revolutionized the life sciences, and the first report of its use to amplify a DNA adduct involved the bypass-proficient Pol mutant KlenTaq M747K,^{29, 30, 32} and the heterocyclic imidic nucleotide triphosphate (TP) Benzi. Benzi was selectively incorporated opposite O^6 -alkylG and DNA complementary to the damaged templates could be amplified by sequential repeats of primer extension reactions, with the presence of Benzi

marking the damage location. However, the approach lacked any quantitative analysis capacity meaning that the amount of O⁶-alkylG in the DNA sample could not be determined.

To fulfil both the requirement of quantification of low-occurrence O⁶-CMG and the need to retain information regarding the DNA sequence, we present herein a chemical-biochemical-analytical combined strategy for the quantitative detection of O⁶-CMG in a sequence-targeted manner. We synthesized (Scheme S1) the triphosphate of the aromatic heterocyclic imide nucleoside ExBenzi and showed for the first time that it is specifically incorporated opposite O⁶-CMG by an engineered Pol. We characterized the incorporation rates of ExBenziTP vs. natural bases opposite O⁶-CMG, and used molecular modelling based on the KlenTaq M747K crystal structure to identify a structural basis for ExBenzi selectivity. Finally, we developed a mass spectrometric method to quantify the incorporated ExBenzi nucleoside, which reflects the amount of initial O⁶-CMG. With this combination of chemical probe, polymerase-mediated synthesis and mass spectrometric analysis, we were able to detect and quantify O⁶-CMG adduct in a specific DNA sequence context. With the capacity to target any DNA sequence, this strategy is anticipated to help elucidate how DNA damage in hotspot regions of the genome impacts the mutagenesis and carcinogenesis processes.

3.3 Results and discussion

3.3.1 Artificial nucleotide analogue is specifically incorporated opposite O⁶-CMG DNA adduct

To characterize the efficiency of incorporation of ExBenzi opposite O⁶-CMG, we performed primer extension experiments using the engineered DNA Pol KlenTaq M747K, which has an established capacity to bypass and incorporate chemically modified nucleotides.^{29, 30, 32} Thus, a 5'-end radiolabeled 23-nucleotide (nt) primer was annealed to a 28 nt template (SI, material and methods) with either G or O⁶-CMG at nt position 24. The ability of KlenTaq M747K to incorporate ExBenzi opposite G or O⁶-CMG was tested by incubating the annealed primer-template DNA with ExBenziTP and KlenTaq M747K at 55 °C for 10 min (Fig. 1A). For comparison, the reaction was conducted with other single nucleotides as controls, i.e. the previously reported artificial nucleotide BenziTP and dTTP, which is the canonical base most frequently incorporated opposite O⁶-CMG by KlenTaq M747K.²⁹ The percentage of incorporated product was calculated as the ratio of the amount of n+1 extension product to the initial amount of primer (Fig. 1B). We found that opposite O⁶-CMG, ExBenzi is more efficiently incorporated (59%) than T (12%), and that it is incorporated at a higher level than the previously reported Benzi (41%). Opposite a template containing G, there was almost no evidence for nucleotide incorporation (4% for both Benzi and ExBenzi vs. 12% for T).

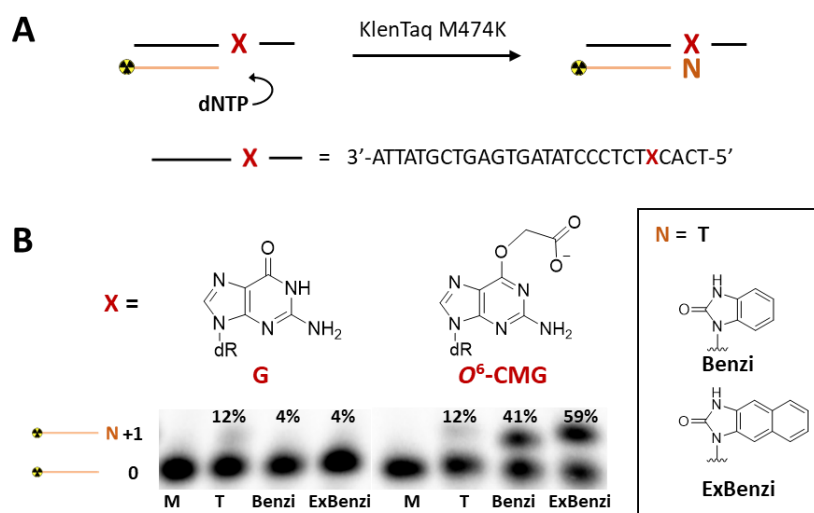


Figure 1 Single nucleotide incorporation opposite O^6 -CMG (A) Scheme of KlenTaq M474K-mediated single nucleotide-incorporation and DNA template sequence used in this study **(B)** Single nucleotide-incorporation of T, Benzi and ExBenzi opposite X = G or O^6 -CMG, catalyzed by KlenTaq M474K, analyzed by gel shift assay on denaturing polyacrylamide gels and visualized by autoradiography (M = marker lane with 23 nt primer).

To quantitatively assess the efficiencies of nucleotide incorporation and characterize how replication over O^6 -CMG by KlenTaq M474K depends on the presence of ExBenziTP, steady-state kinetic analyses of these primer extension processes were performed. Thus, the amount of incorporation of single nucleotides was measured over time (1 min) at increasing concentration of each nucleotide (10-500 μM ; Fig. S1). Oligonucleotides resulting from single incorporation of nucleotides were separated on polyacrylamide gel and visualized by autoradiography. The equilibrium constant for binding affinity (K_M) and the catalytic turnover (k_{cat}) were derived for incorporation by KlenTaq M474K of Benzi, ExBenzi or T, opposite G or O^6 -CMG (Table 1). Not surprisingly, for ExBenzi incorporation, K_M values for incorporation template by either damaged or undamaged DNA templates were very similar (52 and 59 μM , respectively). Most interestingly, a higher binding affinity to the enzyme was calculated for both ExBenzi and Benzi when compared to T. Indeed, the K_M for insertion of ExBenzi opposite O^6 -CMG was almost 5-fold lower than the K_M for insertion of T (59 and 242 μM respectively). Conversely, k_{cat} for the incorporation of ExBenzi opposite O^6 -CMG was around 2-fold higher than for the incorporation of Benzi or T in the same context. As a result, the catalytic efficiency (k_{cat}/K_M) for the incorporation of ExBenzi opposite O^6 -CMG was 7-fold higher than for T opposite O^6 -CMG (Table 1). Remarkably, ExBenzi displayed the highest selectivity for incorporation opposite damaged vs.

undamaged DNA, being incorporated opposite G five times less efficiently when compared to Benzi, resulting in a final 70-fold higher template-selective incorporation of ExBenzi opposite *O*⁶-CMG vs. G.

Table 1 Steady-state kinetic parameters for nucleotide incorporation by KlenTaq M747K DNA polymerase

X*	dNTP	K_M [μM]	k_{cat} [min⁻¹]	k_{cat}/K_M [μM⁻¹ min⁻¹]	Relative^a k_{cat}/K_M
<i>O</i> ⁶ -CMG	T	242 ± 24	2.6	0.011	0.15
	Benzi	43 ± 6	2.3	0.054	0.76
	ExBenzi	59 ± 8	4.1	0.071	1
G	Benzi	48 ± 8	0.24	0.005	0.07
	ExBenzi	52 ± 7	0.06	0.001	0.01

^a Relative k_{cat}/K_M equals catalytic efficiency (k_{cat}/K_M) relative to that of ExBenziTP incorporation opposite *O*⁶-CMG

3.3.2 Structural basis for ExBenzi selective incorporation opposite *O*⁶-CMG by KlenTaq M747K

We were interested to understand the physical basis for the remarkably high selectivity of incorporation of ExBenzi opposite *O*⁶-CMG, despite its large size, both when compared to other nucleotides and in terms of template selectivity. Thus, we performed molecular modelling studies with ExBenziTP, BenziTP or dTTP opposite *O*⁶-CMG or G in DNA bound to KlenTaq M747K. By performing molecular mechanics-based computational modeling, we built a model starting from the crystal structure of KlenTaq M747K in a ternary complex with double-stranded DNA and an incoming dCTP (PDB ID: 5O7T). We then replaced incoming dCTP with ExBenziTP, BenziTP or dTTP. The template strand had either G (original structure) or *O*⁶-CMG opposite the incoming base. Upon energy minimization to identify the most stable and high-occupancy conformer of the constructed DNA-enzyme complex (Fig. 2A), ExBenzi and *O*⁶-CMG interacted by two hydrogen bonds (Fig. 2B, top): one between the N1 of *O*⁶-CMG and the –NH donor of ExBenzi (1.9 Å), and one between the NH₂ donor of *O*⁶-CMG and the carbonyl group of ExBenzi (2.1 Å). These interactions are consistent with previous models^{28, 30} and crystallographic analysis³³ of a similar construct with Benzi. In contrast, the G:ExBenziTP structure predicted only one hydrogen bond (2.0 Å) and a potential steric clash between the –NH moiety on ExBenzi and the –NH at the N1 position on *O*⁶-CMG (Fig. 2B, middle), expected to hinder catalysis. As a result, the complex with *O*⁶-CMG:ExBenziTP is predicted to be more stable (computed free energy of -8.0 kcal/mol) than that of G:ExBenziTP (+32.0 kcal/mol). Furthermore, by overlapping the above computed base pairs, we observed a different orientation for ExBenzi depending on whether it is paired with *O*⁶-CMG or G (Fig. 2C). When paired with G, the steric clash imposed a slight rotation to ExBenzi that resulted in reduced planarity. The computationally derived

geometry and interactions of ExBenzi with O^6 -CMG vs. G help explain the experimentally derived catalytic parameters (Table 1 and graph on Fig. 2D).

Having established a structural basis for template-selective incorporation of ExBenzi, we were interested to understand the selective incorporation of ExBenziTP over dTTP by KlenTaq M747K. Thus, we modelled O^6 -CMG:TTP in the active site of the enzyme and found a similar planarity (Fig. 2C) and hydrogen bond pattern (Fig. 2B, bottom) for O^6 -CMG:TTP and O^6 -CMG:ExBenziTP, suggesting that in this instance interactions between the paired bases do not determine the selectivity of ExBenzi incorporation, or that differences in the interactions are not evident in these models. We speculated that the interactions of the incoming base, ExBenziTP vs. dTTP, with the enzyme might account for the highly favorable incorporation of ExBenzi over T opposite O^6 -CMG, also based on the value of K_M calculated for ExBenziTP and dTTP incorporation, which suggests a stronger binding affinity for ExBenziTP (Table 1). Indeed, the energy calculated for the interaction of KlenTaq M747K with ExBenziTP is -225.2 kcal/mol vs. -204.5 kcal/mol with dTTP. Interactions between the triphosphate groups of the incoming nucleotide and the enzyme made the largest contribution to this difference (Fig. 2E). For example, when Arg659, predicted to interact strongly with the phosphate of ExBenziTP, was mutated to Ala, the computed energy difference was +43 kcal/mol, whereas, when Phe667, predicted to interact with T base, was mutated to Ala, there was not a significant impact on computed energy (Fig. S5). The interaction with the enzyme was predicted to be slightly more favorable for ExBenziTP when compared to BenziTP as well (not shown), in agreement with kinetic measurements (Table 1).

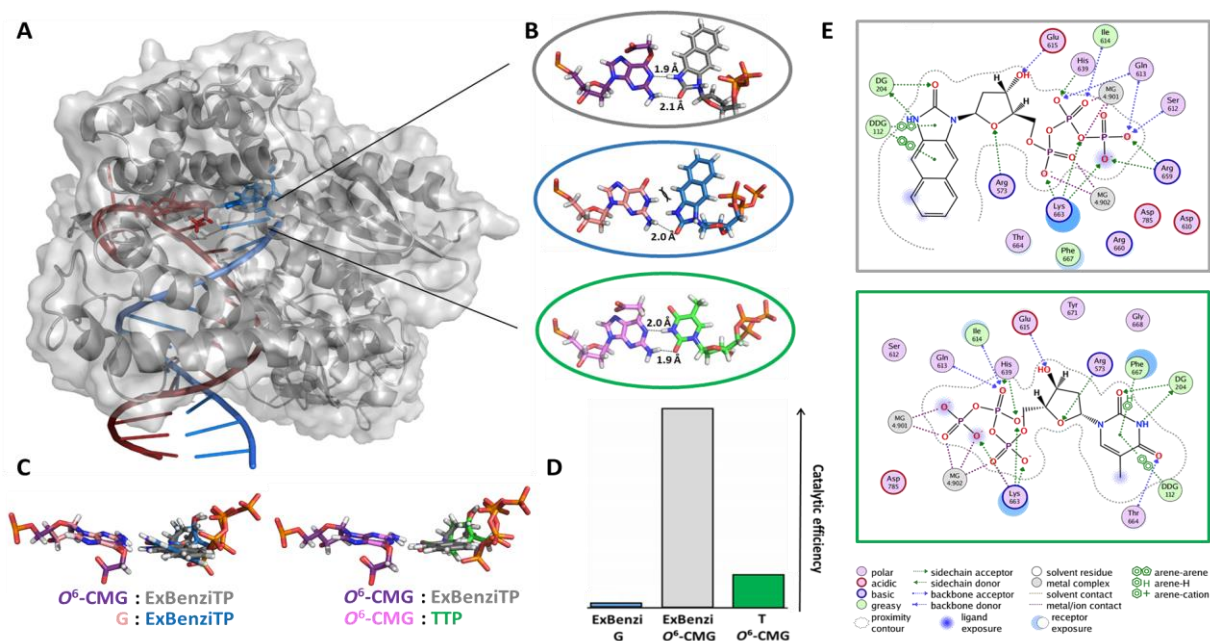


Figure 2 Molecular modelling of interactions in the active site of KlenTaq M747K (A) KlenTaq M747K (PDB ID: 5O7T) was energy minimized with Molecular Operating Environment software. **(B)** In its active site (zoomed-in region), ExBenziTP paired opposite O^6 -CMG (top) via two hydrogen bonds and opposite G (middle) via one hydrogen bond. dTTP paired opposite O^6 -CMG (bottom) via two hydrogen bonds. **(C)** Overlay of base pairs O^6 -CMG:ExBenziTP and O^6 -CMG:TTP (similarly planarity, bottom). Opposite G, ExBenzi is tilted (top). **(D)** Relative catalytic efficiency (Table 1) of incorporation of ExBenziTP opposite G (blue bar), opposite O^6 -CMG (grey) and of TTP opposite O^6 -CMG (green). **(E)** 2D ligand-protein interactions are shown between KlenTaq M747K and ExBenziTP (top) and TTP (bottom).

The modeling studies performed herein suggest that a combination of interactions of the incoming nucleotide with the polymerase and between the bases of the nascent pair drives the selective reaction. The interaction of an incoming base (e.g. TTP and ExBenziTP) with the enzyme represents a first energetic barrier in the incorporation process.³⁴ We speculate that, once in the active site, the interaction of the incoming TP with the templating base promotes the rate-limiting conformational changes of the enzyme required for insertion. Therefore, the stabilization of the system caused by the lowest energy base pair is hypothesized to be pivotal in driving the base insertion (e.g. ExBenziTP opposite O^6 -CMG vs. G). Our findings provide a structural and energetic basis for the rational development of artificial nucleotides for DNA adduct detection, and imply that the artificial nucleotides should be designed to compensate the altered hydrogen bonding capacity of the DNA adduct and to favorably interact with the Pol enzyme responsible for DNA synthesis. Overall, modelling studies corroborated results of primer extension (Fig. 1) and kinetics experiments (Table 1), supporting the further investigation of ExBenzi as a marker for O^6 -CMG.

3.3.3 ExBenzi is required for efficient amplification of a damaged DNA primer

To test whether the polymerase can be extended from ExBenzi once incorporated opposite O^6 -CMG, we performed DNA primer extension studies under conditions suitable for synthesis of a full-length complement of the damaged DNA template. Reaction conditions were the same as for single nucleotide incorporation experiments, except that the reaction mixture was supplemented with all four natural dNTPs, with or without BenziTP or ExBenziTP (Fig. 3A). Primer elongation products were analyzed by gel shift assay on denaturing polyacrylamide gels and visualized by autoradiography. In the presence of all four natural dNTPs, replication was stalled at O^6 -CMG, resulting in little (17%) fully extended primer. However, when either BenziTP or ExBenziTP was additionally added, O^6 -CMG was effectively bypassed (44% and 63% full-

length products, respectively). Finally, we confirmed that replication of undamaged DNA (Fig. 3B, X=G) in the presence of the four natural dNTPs occurs with full primer extension around 90% (band +5 and +6, Fig. 3B), regardless of the presence of BenziTP or ExBenziTP. Thus, BenziTP and especially ExBenziTP promote full-length DNA synthesis past O^6 -CMG, allowing replication of DNA containing O^6 -CMG.

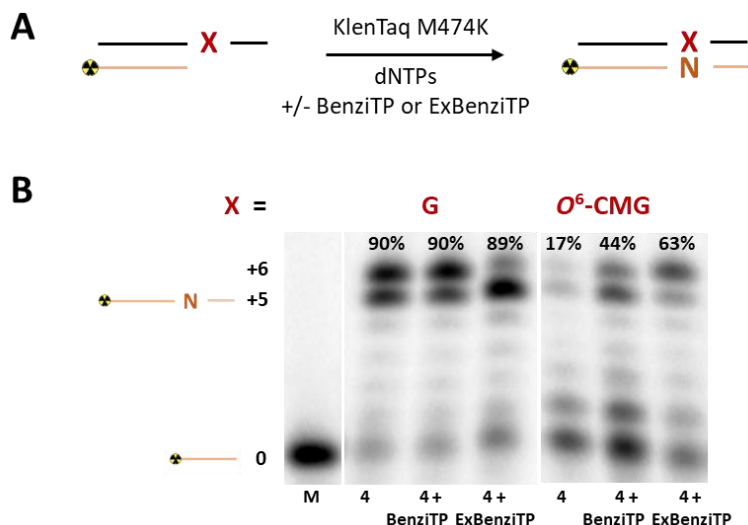


Figure 3 Primer extension past O^6 -CMG (A) Polymerase-mediated DNA synthesis **(B)** Full extension of DNA primer complementary to undamaged (X = G) or damaged (X = O^6 -CMG) DNA template, in the presence of all four natural dNTPS (4), and of BenziTP (4 + BenziTP) or of ExBenziTP (4 + ExBenziTP). Reactions were catalyzed by KlenTaq M474K, analyzed by gel shift assay on denaturing polyacrylamide gels and visualized by autoradiography (M = marker lane with 23 nt primer).

Having established ExBenzi as an effective complement to O^6 -CMG, both when inserted and extended by KlenTaq M747K, we tested whether a primer annealed to a template containing O^6 -CMG could be amplified in the presence of ExBenzi. For these studies, DNA primer and template were not pre-annealed: steps of annealing, primer extension and duplex melting were repeated to allow that the 28 nt DNA could template several rounds of primer extension. To maximize selectivity and enzyme efficiency, we optimized amplification cycles (Fig. S2), temperature (Fig. S3), DNA and TP concentration (Fig. S4). Optimal conditions consisted of 2 nM of 28 nt template, 300 nM of 23 nt primer, 10 μ M of each dNTPs and of ExBenziTP, 50 nM KlenTaq M747K (Fig. 4A). Alternating cycles (50) of DNA melting (95 $^{\circ}$ C, 30 s), annealing (42 $^{\circ}$ C, 30 s) and extension steps (55 $^{\circ}$ C, 30 s) were performed. Products were separated on a polyacrylamide gel and visualized by staining with SYBR gold fluorescent dye (Fig. 4B). Under

ExBenzi nucleoside was quantified by LC-MS/MS in SRM mode, by monitoring the transition m/z 301→185 (Fig. S6). To correct for normal instrumental variation and account for sample loss during sample processing, we evaluated artificial nucleosides that are structural analogues of ExBenzi, such as Benzi, BIM, Per and ExBIM, as internal standards (IS) that could avoid the need for synthesizing isotopically modified ExBenzi.²⁷⁻³¹ ExBIM nucleoside (Fig. S6) was optimal in having similar chemical properties to ExBenzi nucleoside, eluting at a similar retention time, and having a unique mass transition of m/z 285→169. We optimized the preparation of samples to achieve a similar recovery for ExBenzi nucleoside and for the IS ExBIM nucleoside of around 90%.

To quantify ExBenzi nucleoside, calibration curves were prepared in the presence of sample matrix, consisting of nucleosides, enzyme and salts (1xKTQ, SI), as matrix effect was observed for the ionization of both nucleosides. Standard samples in matrix were prepared analogously to what described above for the experimental samples. Briefly, linear amplification reactions were run using unmodified DNA template under the usual conditions but in the absence of ExBenzi. Samples were spiked with internal standard ExBIM (10 nM) and with ExBenzi (5-500 nM). By relating the signal of ExBenzi to that of ExBIM, we could quantify ExBenzi nucleoside in a matrix with a LOQ of 20 fmol (Fig. S7). Analysis of ExBenzi released from amplicons that were templated with DNA containing O^6 -CMG showed a linear increase of ExBenzi signal with increasing concentrations of O^6 -CMG in the template. ExBenzi could be detected also when reactions were performed with unmodified DNA, probably due to unspecific incorporation or non-covalent binding of ExBenzi, such as intercalation.³⁵ We could reduce the unspecific signal by disrupting such non-covalent interactions by varying pH or salt concentration; however, these additional sample processing steps significantly reduced recovery rates. Therefore, the background levels of ExBenzi in samples arising from amplification of unmodified DNA was used for background subtraction for the analysis. Thus, we found a linear increase of ExBenzi with increasing initial O^6 -CMG levels in the template (Fig. 5).

The approach described here establishes for the first time a chemical basis for amplification and quantitation of DNA damage in a sequence-target manner, achieved by mean of hybridization with a DNA primer. This approach is useful for the study of isolated oligonucleotides, however, the sensitivity is insufficient for sequence-specific adduct detection of samples from cells. In cultured cells exposed to azaserine (0 - 450 μ M), O^6 -CMG levels were 0.3-9.1 lesions/ 10^7 nucleotides.²² The low occurrence of O^6 -CMG is even more relevant in the case of sequence-specific detection; for example, there are only 5 guanines per genome (i.e. 6.6×10^9 nucleotides) belonging to codons 12 and 13 of the *k-ras* gene, which are commonly mutated in CRC.³⁶ With the herein presented combined strategy, even assuming that all *k-ras* Gs are carboxymethylated, an increase of sensitivity of a few orders of magnitude is needed. Nonetheless, for cell-based applications, there are several aspects that can be developed in

future studies to address this limitation, including polymerase enzyme engineering/evolution,^{37, 38} higher mass spectrometric sensitivity,³⁹ and sample enrichment.⁴⁰

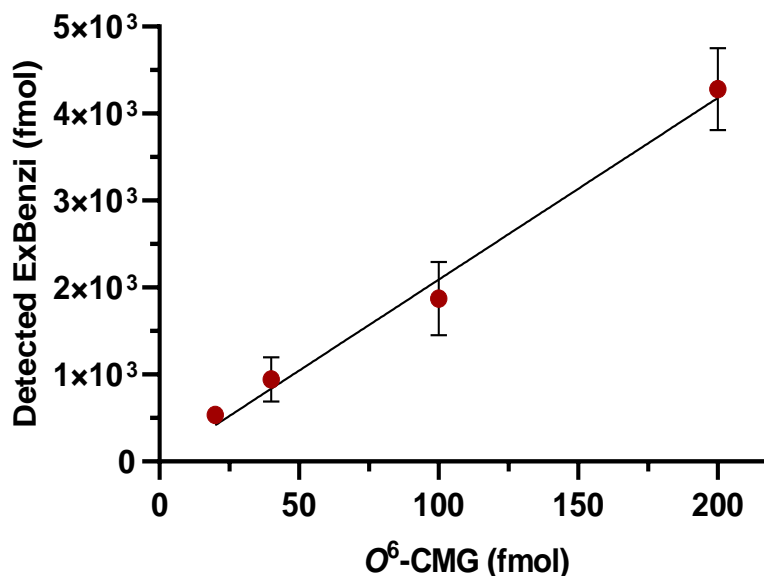


Figure 5 Quantitation of ExBenzi in amplicons from DNA containing *O*⁶-CMG. ExBenzi incorporated by linearly-amplified primer reactions from DNA with varying amounts of *O*⁶-CMG by mass spectrometry, based on the calibration curve in Fig. S7. Data is average from four independent replicates, each made of two technical replicates. $R^2 = 0.95$.

3.4 Conclusion

Interest in understanding the genomic location of DNA adducts is growing in medical sciences,⁴¹⁻⁴⁵ as some are linked to mutagenicity and to increased risk of cancer,⁴⁶ as well as drug efficacy.⁴⁷ The *O*⁶-CMG DNA adduct focused on in this study is mutagenic and hypothesized to be linked with CRC development.¹¹⁻¹³ We developed a strategy combining sequence-specific amplification and quantitation of *O*⁶-CMG DNA adducts in a DNA oligonucleotide. The artificial synthetic nucleotide ExBenzi is the most specifically and efficiently incorporated opposite *O*⁶-CMG reported to date. Using the engineered Pol KlenTaq M747K, we replicated damaged DNA strands such that the amount of ExBenzi incorporated during repeated DNA synthesis is a linear amplification of *O*⁶-CMG in DNA. We elucidated a structural basis for the high selectivity and specificity involving complementarity of H bonds and planarity of the *O*⁶-CMG:ExBenziTP pair in the KlenTaq M747K active site. Finally, we developed a mass spectrometric method for quantification of ExBenzi nucleoside and showed that the amount of ExBenzi incorporated during DNA synthesis correlates linearly with the initial *O*⁶-CMG DNA, allowing therefore its

quantitation. The detection of O^6 -CMG with retention of DNA sequence information presented in this study lays the foundation to the study of the relationship between DNA damage and mutations in hotspot regions of the genome, and to the development of strategies for predictive or pre-diagnostic purposes.

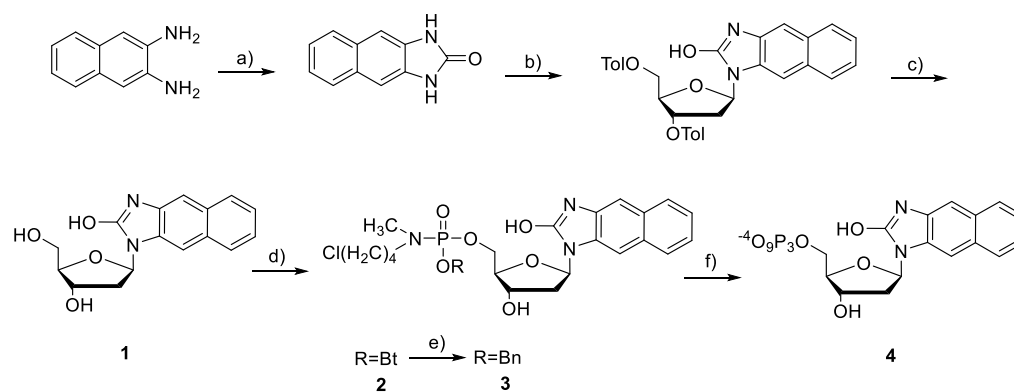
3.5 Supplementary Information (Experimental Section)

Chemicals and Oligonucleotides

T4 polynucleotide kinase was purchased from Promega Corp (Madison, USA). KlenTaq M747K was kindly donated by Prof. Marx Andreas (University of Konstanz, Germany). [γ - 32 P] ATP was purchased from PerkinElmer Life Sciences (Schwerzenbach, Switzerland). 1-(α -Chloro-3,5-di-*O*-(*p*-toluoyl)-2-deoxy-D-ribose was purchased from Carbosynth (Compton, UK). Reagents for solid-phase DNA synthesis and all natural dNTPs were obtained from Link technologies Ltd. (Lanarkshire, Scotland). SYBR gold was purchased from Invitrogen (Carlsbad, CA, USA). All other reagents were purchased from Sigma Aldrich (Buchs, Switzerland). Thin layer chromatography was carried out with silica gel 60 F254 plates with aluminum backing. A Biotage system with pre-packed SNAP ultra SiO₂ cartridges was used for flash chromatography. ¹H, ¹³C, ³¹P NMR spectra were recorded on a Bruker Biospin 400 MHz NMR machine, and chemical shifts are reported in parts per million (ppm, δ) relative to the chemical shift of the respective solvent.

Undamaged oligonucleotides (28mer template: 5'- ACT CGT CTC CCT ATA GTG AGT CGT ATT A - 3'; 23mer primer 5'-TAA TAC GAC TCA CTA TAG GGA GA-3'; and 5'-FAM-labelled 28mer template and 23mer primer) were purchased from Eurogentec (Seraing, Belgium). The O^6 -CMG 28mer template (5'- ACT CXT CTC CCT ATA GTG AGT CGT ATT A -3'; X= O^6 -CMG) was kindly donated by Prof. David M. Williams (University of Sheffield, UK).

Synthesis of artificial triphosphates ExBenziTP (4)

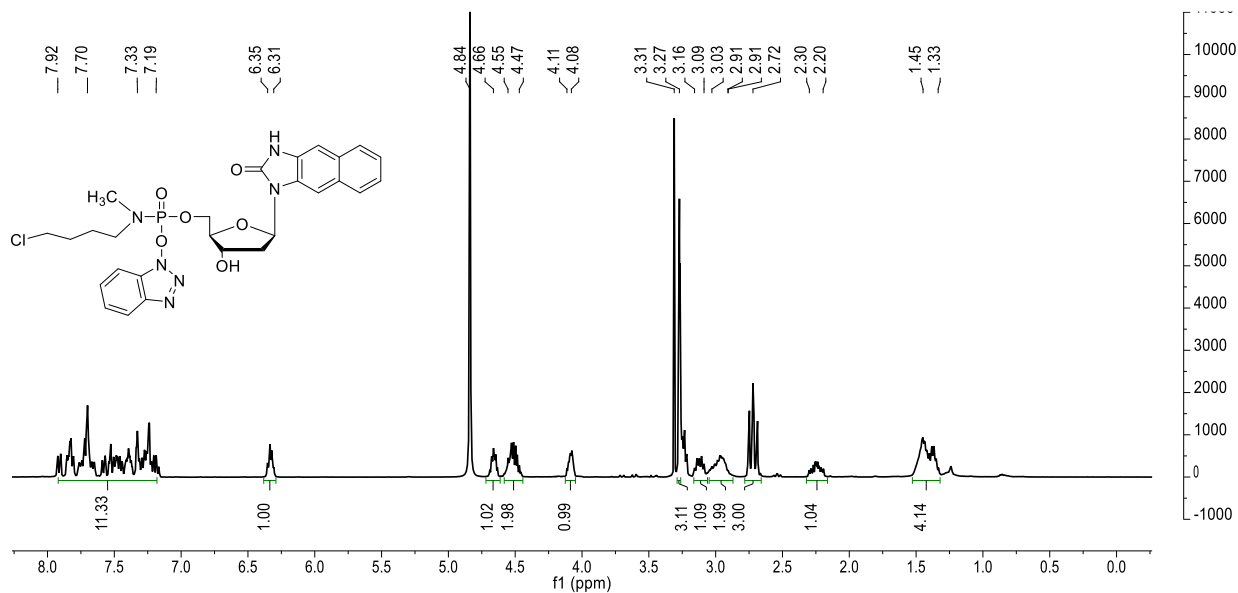


Scheme S1: Synthesis of artificial nucleotide ExBenziTP. a) urea, microwave irradiation.²⁷ b) (i) NaH, THF, 0°C; (ii) 1-(α -chloro-3,5-di-*O*-p-toluoyl-2-deoxy-D-ribose, 25°C. c) NaH, MeOH:THF (1:9), 25°C.[2] d) (i) (4-chlorobutyl)(methyl)phosphoramidic chloride, *N*-Hydroxybenzotriazole (HOBt), Pyr, THF, 25°C; (ii) NMI, THF, 25°C. e) BnOH, DMAP, THF, 25°C. f) (i) H₂, 10% Pd/C, DMF, 25°C; (ii) PPI, 25°C.

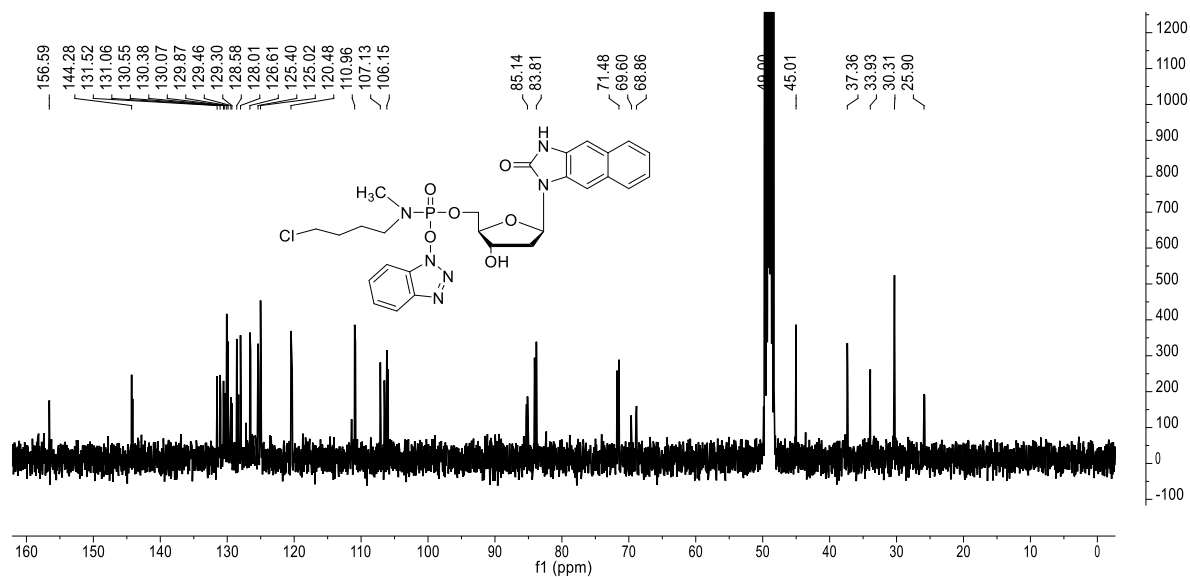
(2R,3S,5R)-(1H-benzo[d][1,2,3]triazol-1-yl ((3-hydroxy-5-(2-oxo-2,3-dihydro-1H-naphtho[2,3-d]imidazol-1-yl)tetrahydrofuran-2-yl)methyl) (4-chlorobutyl)(methyl)phosphoramidate (2)

To a solution of *N*-Hydroxybenzotriazole (0.529 g, 3.920 mmol, azeotropically dried twice in benzene) in pyridine (0.25 mL) and THF (2.5 mL) (4-chlorobutyl)(methyl)phosphoramidic dichloride (0.466 g, 1.961 mmol) in THF (1 mL) was added dropwise. The resulting mixture was stirred under nitrogen atmosphere for 4 hours at 25 °C, whereas a white precipitate formed. The white precipitate was filtered and the filtrate was added to a solution in pyridine (0.5 mL) of **1** (0.147 g, 0.490 mmol), which was synthesized from 2,3-diaminonaphthalene as previously reported.²⁷ Then, *N*-methylimidazole (previously dried over MS) (0.161 g, 1.960 mmol) was added and the mixture stirred at 25 °C for 1 hour under nitrogen atmosphere. THF was removed under reduced pressure and the residue was extracted with CH₂Cl₂ (2x) and washed with saturated aqueous NH₄Cl (1x) and saturated aqueous NaHCO₃ (1x). Combined organic layers were dried over Na₂SO₄, filtered and concentrated. The resulting residue was purified by flash chromatography on a Biotage system using a DCM:MeOH gradient (3-10% MeOH in 15 CV, 1 CV = 75 mL) the resulting fractions were not pure, so a second purification by flash chromatography was carried out a Biotage system using a DCM:MeOH gradient(3-8% MeOH in 20 CV, 1 CV = 15 mL) yielding **2** (0.060 g, 20%) as a white solid. R_f (SiO₂, DCM:MeOH 97:3) = 0.28; ¹H NMR (400 MHz, CD₃OD) δ 7.92 – 7.19 (m, 11H), 6.35 – 6.31 (m, 1H), 4.66 (dt, J = 11.8, 5.1 Hz, 1H), 4.55 – 4.47 (m, 2H), 4.11 – 4.08 (m, 1H), 3.27 (dt, J = 3.3, 1.6 Hz, 3H), 3.16 – 3.09 (m, 1H), 3.03 – 2.91 (m, 2H), 2.76 – 2.69 (m, 3H), 2.30 – 2.20 (m, 1H), 1.45– 1.33 (m, 4H); ¹³C NMR (100 MHz, CD₃OD) δ 156.59, 144.28, 131.52, 131.06, 130.55, 130.38, 130.07, 129.87, 129.46, 129.30, 128.58, 128.01, 126.61, 125.40, 125.02, 120.48, 110.96, 107.13, 106.15, 85.14, 83.81, 71.48, 69.60, 68.86, 45.01, 37.36, 33.93, 30.31, 25.90; ³¹P NMR (160 MHz, CD₃OD) relative to phosphoric acidstandard: δ 13.07 (d, J = 14.6 Hz, 1P); HRMS (ESI) calculated for C₂₇H₃₀CIN₆O₆P: [M+Na⁺] m/z 623.1545, found: 623.1548.

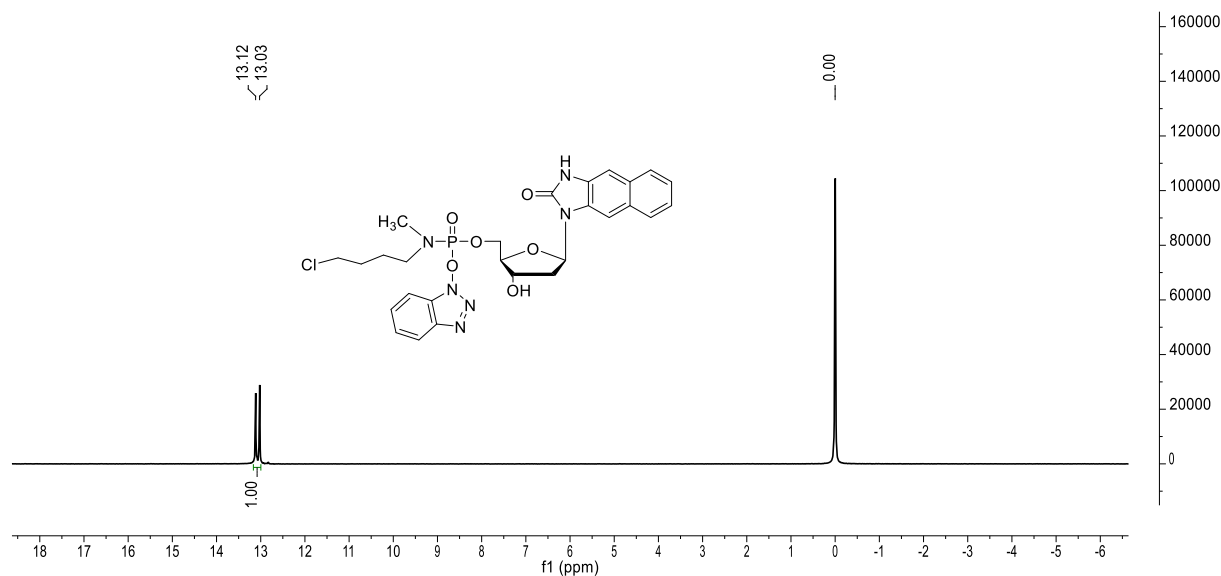
^1H NMR (2R,3S,5R)-(1H-benzo[d][1,2,3]triazol-1-yl ((3-hydroxy-5-(2-oxo-2,3-dihydro-1H-naphtho[2,3-d]imidazol-1-yl)tetrahydrofuran-2-yl)methyl) (4-chlorobutyl)(methyl) phosphoramidate (2) (CD_3OD)



^{13}C NMR (2R,3S,5R)-(1H-benzo[d][1,2,3]triazol-1-yl ((3-hydroxy-5-(2-oxo-2,3-dihydro-1H-naphtho[2,3-d]imidazol-1-yl)tetrahydrofuran-2-yl)methyl) (4-chlorobutyl)(methyl) phosphoramidate (2) (CD_3OD)



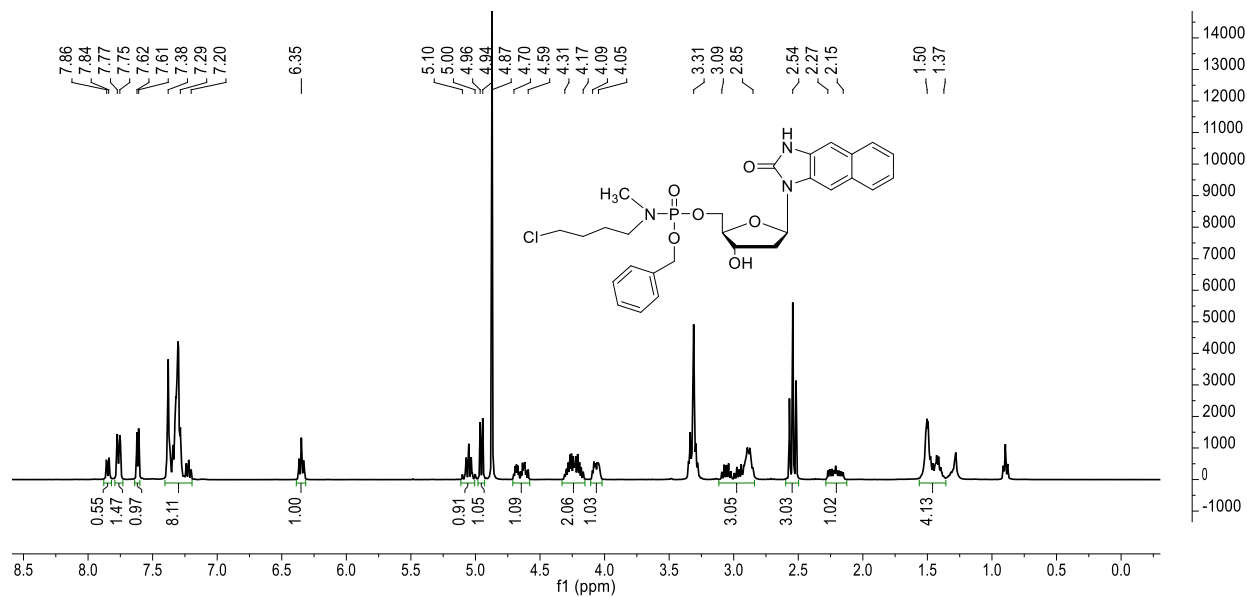
³¹P NMR (2*R*,3*S*,5*R*)-(1*H*-benzo[*d*][1,2,3]triazol-1-yl ((3-hydroxy-5-(2-oxo-2,3-dihydro-1*H*-naphtho[2,3-*d*]imidazol-1-yl)tetrahydrofuran-2-yl)methyl) (4-chlorobutyl)(methyl) phosphoramidate (2) (CD₃OD)



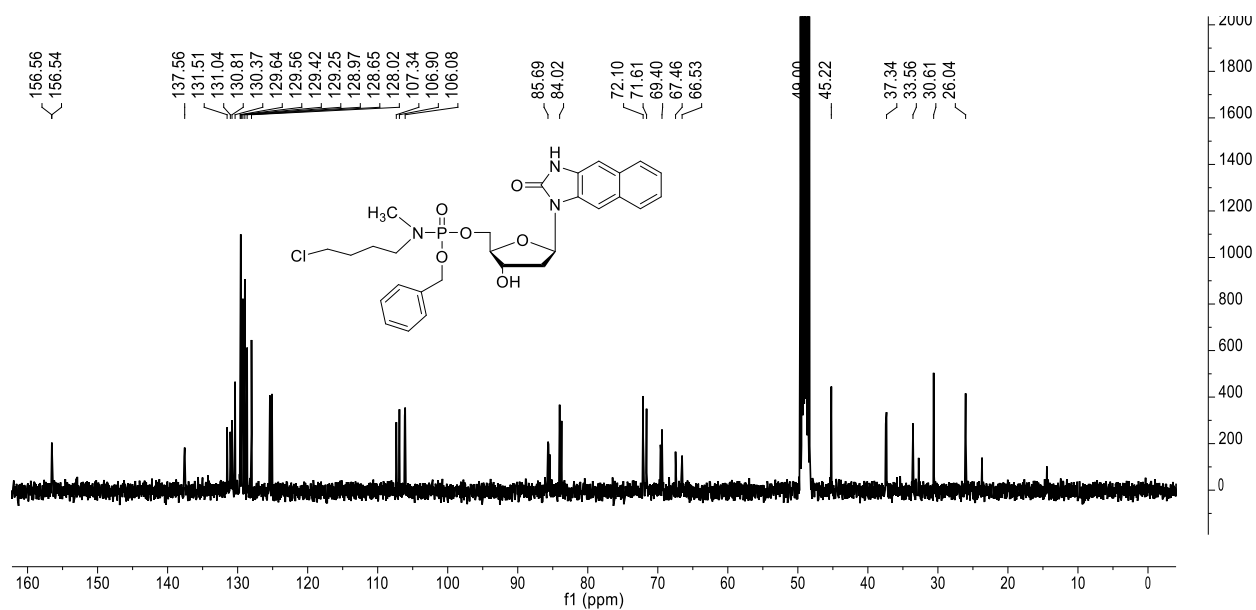
(2*R*,3*S*,5*R*)benzyl ((3-hydroxy-5-(2-oxo-2,3-dihydro-1*H*-naphtho[2,3-*d*]imidazol-1-yl)tetrahydrofuran-2-yl)methyl) (4-chlorobutyl)(methyl)phosphoramidate (3)

Compound **2** (0.060 g, 0.102 mmol) was dissolved in THF (1 mL) and Benzyl alcohol (0.552 g, 5.111 mmol) was added in one portion. To the solution, 4-Dimethylaminopyridine (0.050 g, 0.409 mmol) was added and the reaction was stirred under nitrogen atmosphere at 25 °C for 96 hours. The reaction was concentrated under reduced pressure and purified by flash chromatography on a Biotage system using a DCM:MeOH gradient (0-10% MeOH over 30 CV, 1 CV = 15 ml), yielding product **3** (0.030 g, 52%) as a white foamy solid. *R*_f (SiO₂, DCM:MeOH 95:5) = 0.27; ¹H NMR (400 MHz, CD₃OD) δ 7.85 (d, *J* = 6.2 Hz, 0.5H), 7.76 (d, *J* = 7.8 Hz, 1.5H), 7.61 (d, *J* = 5.2 Hz, 1H), 7.39 – 7.20 (m, 8H), 6.35 (t, *J* = 7.3 Hz, 1H), 5.10 – 5.00 (m, 1H), 4.95 (d, *J* = 8.0 Hz, 1H), 4.70 – 4.59 (m, 1H), 4.31 – 4.17 (m, 2H), 4.09 – 4.05 (m, 1H), 3.09 – 2.85 (m, 3H), 2.54 (t, *J* = 10.2 Hz, 3H), 2.27 – 2.15 (m, 1H), 1.50 – 1.37 (m, 4H); ¹³C NMR (100 MHz, CD₃OD) δ 156.56, 156.54, 137.56, 131.51, 131.04, 130.81, 130.37, 129.64, 129.56, 129.42, 129.25, 128.97, 128.65, 128.02, 107.34, 106.90, 106.08, 85.69, 84.02, 72.10, 71.61, 69.40, 67.46, 66.53, 49.00, 45.22, 37.34, 33.56, 30.61, 26.04; ³¹P NMR (160 MHz, CD₃OD) relative to phosphoric acid standard: δ 11.69 (d, *J* = 69.5 Hz, 1P); HRMS (ESI) calculated for C₂₈H₃₃ClN₃O₆P: [M+Na⁺] *m/z* 596.1688, found: 596.1686.

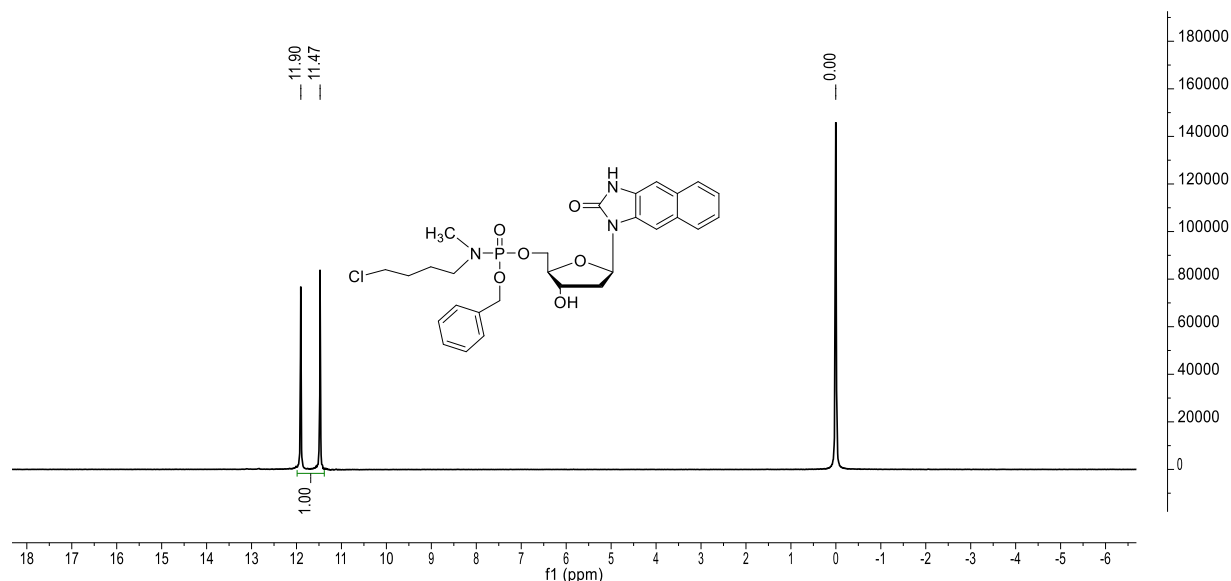
^1H NMR (2*R*,3*S*,5*R*)benzyl ((3-hydroxy-5-(2-oxo-2,3-dihydro-1*H*-naphtho[2,3-*d*]imidazol-1-yl)tetrahydrofuran-2-yl)methyl) (4-chlorobutyl)(methyl)phosphoramidate (3) (CD_3OD)



^{13}C NMR (2*R*,3*S*,5*R*)benzyl ((3-hydroxy-5-(2-oxo-2,3-dihydro-1*H*-naphtho[2,3-*d*]imidazol-1-yl)tetrahydrofuran-2-yl)methyl) (4-chlorobutyl)(methyl)phosphoramidate (3) (CD_3OD)



³¹P NMR (2*R*,3*S*,5*R*)benzyl ((3-hydroxy-5-(2-oxo-2,3-dihydro-1*H*-naphtho[2,3-*d*]imidazol-1-yl)tetrahydrofuran-2-yl)methyl) (4-chlorobutyl)(methyl)phosphoramidate (**3**) (CD₃OD)

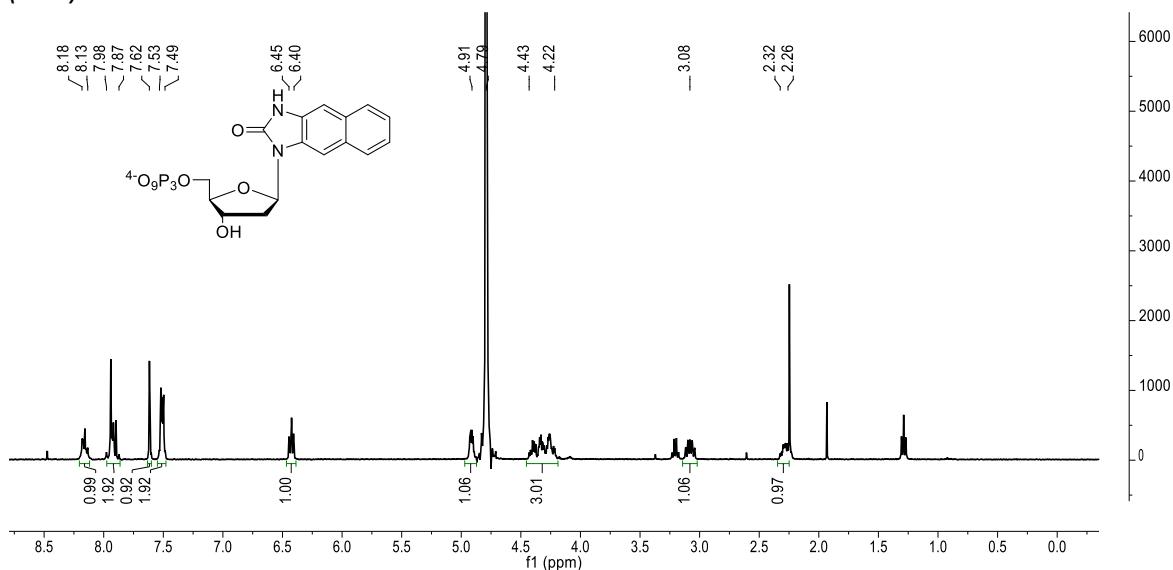


(2*R*,3*S*,5*R*)-(3-hydroxy-5-(2-oxo-2,3-dihydro-1*H*-naphtho[2,3-*d*]imidazol-1-yl)tetrahydrofuran-2-yl)methyl triphosphate (4**)**

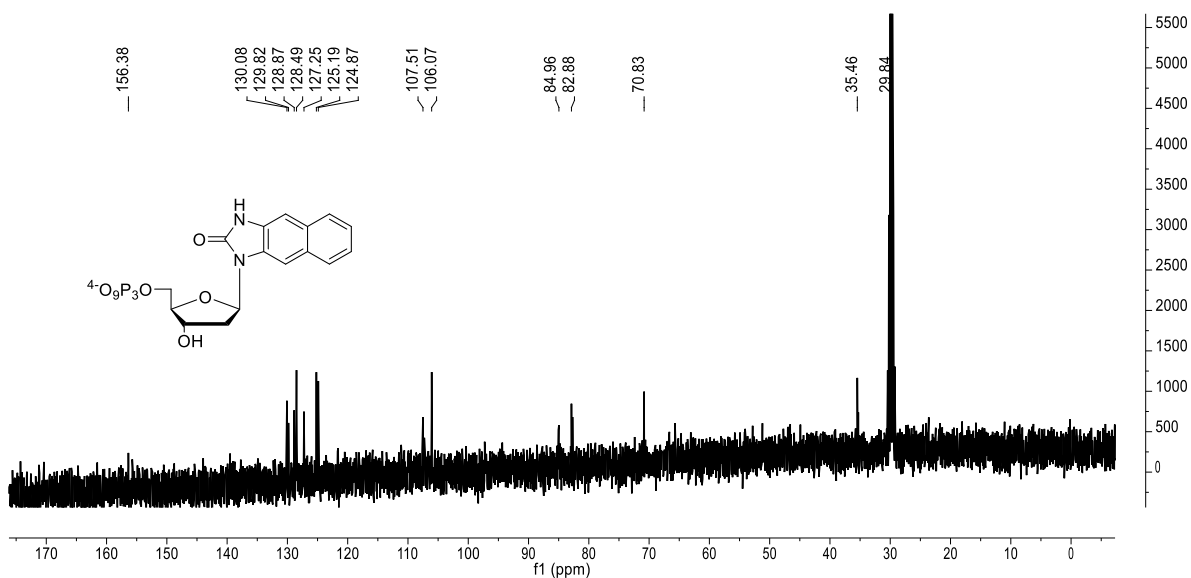
Compound **3** (0.030 g, 0.052 mmol) was dissolved in anhydrous DMF (2 mL) and given to a flask set under H₂. A tip of palladium catalyst (10% Pd/C) was added and the mixture was stirred under an H₂ atmosphere at 25 °C for 2 hours. The reaction mixture was filtered through a 0.45 μm syringe filter to remove the catalyst. Tris(tetra-*n*-butylammonium) hydrogen pyrophosphate (0.029 g, 0.052 mmol) was added to the filtrate, and the mixture was stirred for 4 hours at 25 °C under nitrogen atmosphere. DMF was removed under reduced pressure and residual was diluted in 1 mL MilliQ water. The resulting solution was purified twice by reverse-phase HPLC: first on a Phenomenex LUNA C18 column (250 x 10 mm) with an acetonitrile in 50 mM TEAA (pH 7) gradient (2-60% in 25 min, flow rate 2 ml/min); the collected fractions were subjected to a second purification using the same column and solvents (20-60% ACN in 50 mM TEAA (pH 7) in 25 min) yielding product **4** (0.006 g, 21%) with triethylammonium (TEA⁺) as counter ion. To exchange the counter ion to the sodiated form, product **4** (0.006 g, 0.011 mmol) was dissolved in 0.25 mL MilliQ water, then NaClO₄ (0.040 g, 0.330 mmol) was added and the mixture was stirred for 2 hours at 25 °C. Then, Acetone (1.75 mL) was added and the mixture centrifuged for 3 minutes at 14'000 g. The supernatant was carefully discarded and the pellet was dissolved in 0.5 mL water and lyophilized to give final sodiated triphosphate **4** (0.0012 g). ¹H NMR (400 MHz, D₂O) δ 8.18 – 8.13 (m, 1H), 7.98 – 7.87 (m, 2H), 7.62 (s, 1H), 7.53 – 7.49 (m, 2H), 6.45 – 6.40 (m, 1H), 4.91 (dd, *J* = 7.3, 3.7 Hz, 1H), 4.43 – 4.22 (m, 3H), 3.08 (dt, *J* = 16.2, 8.2 Hz, 1H), 2.32 – 2.26 (m, 1H); ¹³C NMR (100 MHz, D₂O, referenced to acetone) δ 156.38, 130.08, 129.82, 128.87,

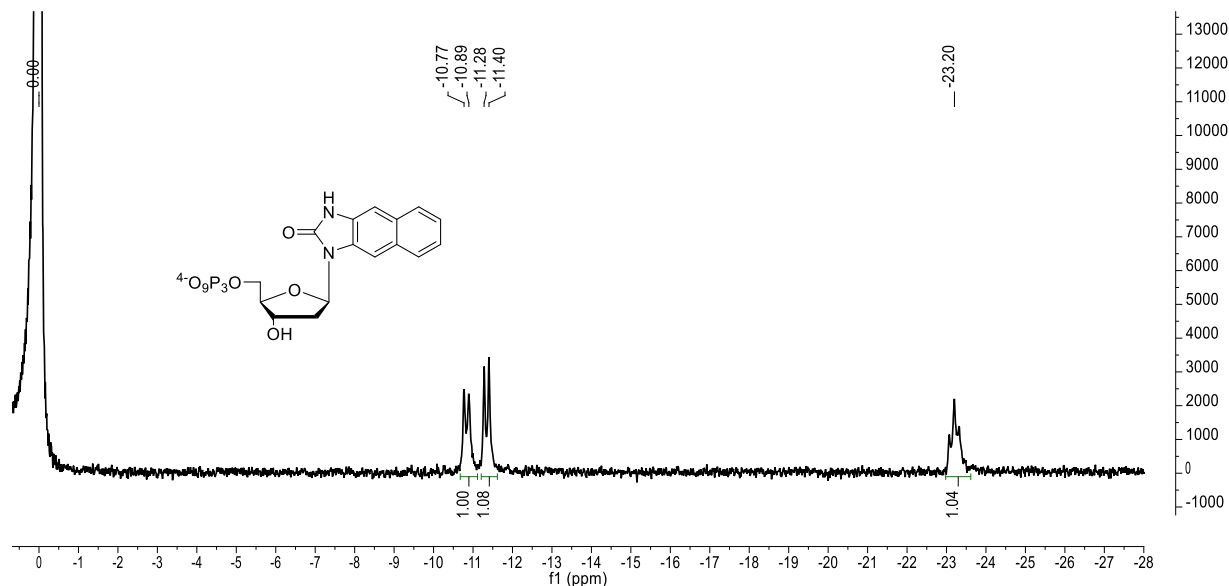
128.49, 127.25, 125.19, 124.87, 107.51, 106.07, 84.96, 82.88, 70.83, 35.46; ^{31}P NMR (160 MHz, D_2O) relative to phosphoric acid standard: δ -10.83 (d, $J = 19.8$ Hz, 1P), -11.34 (d, $J = 20.1$ Hz, 1P), -23.20 (t, $J = 20.5$ Hz, 1P); HRMS (ESI) calculated for $\text{C}_{16}\text{H}_{19}\text{N}_2\text{O}_{13}\text{P}_3$: $[\text{M}-\text{H}]^+$ m/z 539.0016, found: 539.0024.

^1H NMR (2*R*,3*S*,5*R*)-(3-hydroxy-5-(2-oxo-2,3-dihydro-1*H*-naphtho[2,3-*d*]imidazol-1-yl)tetrahydrofuran-2-yl)methyl triphosphate (4) (D_2O)



^{13}C NMR (2*R*,3*S*,5*R*)-(3-hydroxy-5-(2-oxo-2,3-dihydro-1*H*-naphtho[2,3-*d*]imidazol-1-yl)tetrahydrofuran-2-yl)methyl triphosphate (4) (D_2O)



³¹P NMR (2*R*,3*S*,5*R*)-(3-hydroxy-5-(2-oxo-2,3-dihydro-1*H*-naphtho[2,3-*d*]imidazol-1-yl)tetrahydrofuran-2-yl)methyl triphosphate (4) (D₂O)**General conditions for primer extension assays**

Radioactive labelling of the 23mer primer strand at the 5' end was carried out using T4 polynucleotide kinase and [γ -³²P] ATP and following the manufacturer protocol. The labeled primer strand was annealed to either the unmodified 28mer or the *O*⁶-CMG 28mer template by heating at 95 °C for 5 min a mixture of primer:template (1.5:1) and slow cooling down overnight. Standard primer extension reactions (10 μ l) contained 1 \times KTQ reaction buffer, 5 nM KlenTaq M747K, 15 nM DNA (15 nM primer and 22.5 nM template), and 10 μ M dNTPs. In full-length DNA synthesis experiments, reactions contained all four natural dNTPs (10 μ M total) with or without either BenziTP or ExBenziTP (10 μ M). Primer/template, nucleotides and DNA polymerase were incubated at 55 °C for 10 min. 1 \times KTQ reaction buffer contained 50 mM Tris-HCl (pH 9.2), 16 mM (NH₄)₂SO₄, 2.5 mM MgCl₂, and 0.1% Tween 20. Reactions were quenched by adding 20 μ l PAGE gel loading buffer (80% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanole FF). Reaction mixtures (4 μ l) were loaded onto 15% polyacrylamide / 7 M urea gel, which was previously equilibrated for 30 min with 1 \times TBE (100 mM Tris base, 100 mM boric acid, 2 mM EDTA). Extension products were visualized by autoradiography using a phosphorimager (BioRad, Hercules, CA, USA) by incubating the gel containing radioactive products on a phosphorimaging screen for 4h. Product band intensities were quantified by densitometry using Quantity One Software (Bio-Rad).

Steady-state kinetic assays

Steady-state kinetics parameters for single nucleotide incorporation by DNA polymerase KlenTaq M747K were determined under single completed hit conditions.^{48,49} For varying dNTP concentrations, the quantity of $n + 1$ product formed by performing the reaction at 55 °C was measured. Reaction mixtures included 5 nM KlenTaq M747K, 100 nM primer, 150 nM template, and 1× KTQ reaction buffer. Reactions were initiated by adding pre-warmed enzyme and DNA mix to pre-warmed dNTPs. Reactions were quenched by adding PAGE loading buffer. Reaction mixtures (4 μ L) were separated on a 15% polyacrylamide / 7 M urea denaturing gel, visualized by autoradiography, and quantified with Quantity One Software (Bio-Rad). To obtain kinetic parameters v_{max} , K_M and k_{cat} , the intensities of $n + 1$ bands (quantified on the Quantity One Software, Bio-Rad) were fit to a Michaelis-Menten hyperbola curve using GraphPad Prism 8 (GraphPad Software). Reactions were performed at least in triplicate and the averaged Michaelis-Menten curves are shown in S1. For K_M values, means (\pm standard deviations) are reported (Table 1).

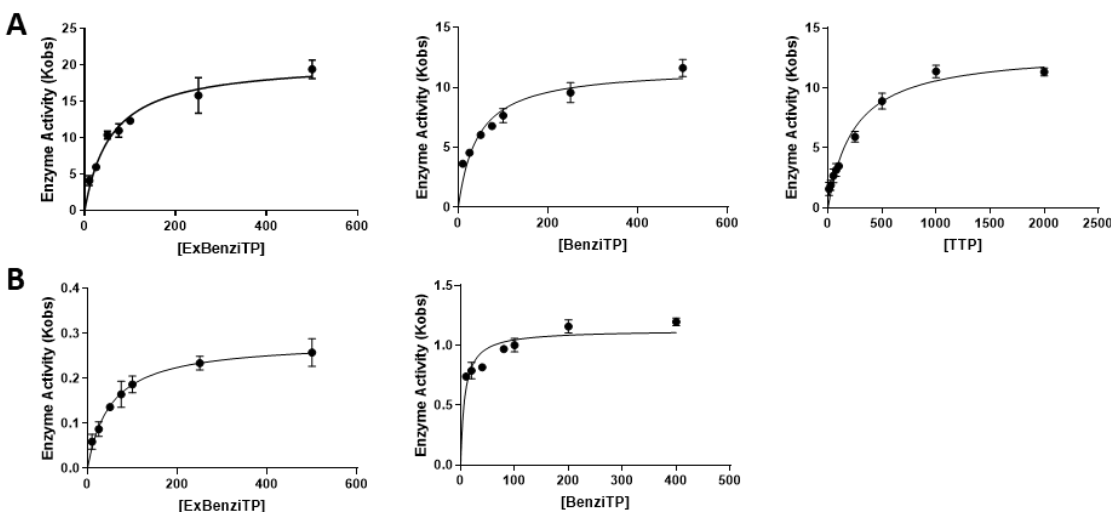


Figure S1 Michaelis-Menten hyperbola curves. The observed rate constants (Kobs) of nucleotide incorporation were derived from the $n + 1$ bands of PAGE gel relative to the single incorporation of ExBenziTP, BenziTP and TTP opposite O^6 -CMG (**A**) and of ExBenziTP and BenziTP opposite G (**B**). Kobs are plotted against the nucleotide concentration and data were fit to Michaelis-Menten curves on GraphPad Prism 8. Error bars indicate standard deviations among independent replicates.

Linear amplification of O^6 -CMG DNA

Reaction mixtures (20 μ l) contained 20 nM of KlenTaq M747K, 1x KTQ, 250 μ M of each dNTP, 300 nM of the 23mer primer and 10 nM of O^6 -CMG 28mer template. Primer was 5'-FAM-labelled were indicated (5'-FAM-labelled 28mer template was only used as a marker for

imaging). Reactions were performed on a T3000 Thermocycler (Biometra). Final conditions included: initial denaturation step at 95 °C for 2 min, followed by 50 cycles of 95 °C for 30 s, 42 °C for 30 s and 55 °C for 30 s, followed by a final step of elongation at 55 °C for 3 min. For imaging, 10 µl of each reaction were added to 10 µl of 2% SDS in formamide and loaded on a 15% acrylamide / 7 M urea gel. Gels were run at 300 V, at rt for 2-2.5 h. Gels were imaged on a ChemiDoc MP Imaging System (BioRad), either after staining with 1x SYBR gold in TBE (Ex/Em 495/537 nm) or directly when the 23mer primer was FAM-labelled (Ex/Em 490/520 nm).

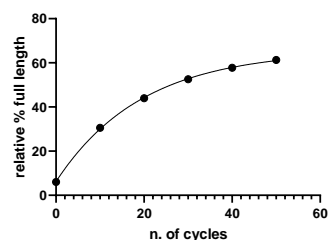


Figure S2: Amplification cycles. Reactions of linear amplifications of the 23mer primer complementary to *O*⁶-CMG 28mer template were performed by alternating 0, 10, 20, 30, 40 or 50 cycles of melting, annealing and extension. Extension products were quantified from PAGE gels, normalized to the initial amount of primer (i.e. unreacted primer + extension product) and plotted against the number of cycles to an exponential fit with one-phase decay ($R^2 = 0.99$) on GraphPad Prism 8.

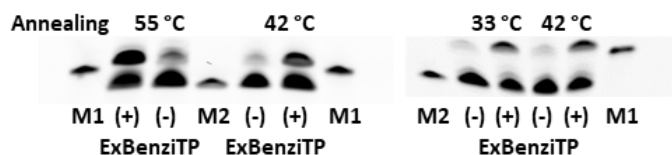


Figure S3 Temperature. Reactions of linear amplifications of a FAM-labelled 23mer primer complementary to *O*⁶-CMG 28mer template were tested at varying temperatures of annealing (33°C, 42°C and 55°C) and extension (not shown) to achieve the highest bypass of *O*⁶-CMG in the presence of ExBenziTP (+) and the lowest bypass in the absence of without ExBenziTP (-). (M1= marker line for 28mer; M2= marker line for 23mer)

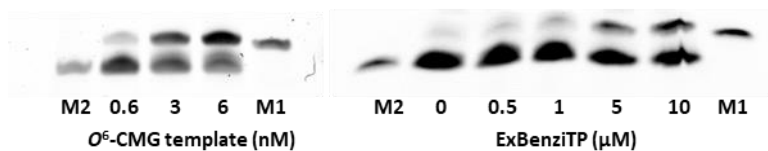


Figure S4 Concentration of reagents. Reactions of linear amplifications of a FAM-labelled 23mer primer complementary to *O*⁶-CMG 28mer template were tested at varying concentrations of initial *O*⁶-CMG template (0.6 – 6 nM) and of ExBenziTP (0 – 10 µM). (M1= marker line for 28mer; M2= marker line for 23mer)

DNA hydrolysis and sample preparation

Amplified DNA was purified by size exclusion filtration (Bio-Spin[®] 6, BioRad), following the manufacturer protocol and enzymatically hydrolyzed following a previously reported one-step protocol.⁵⁰ The amount of enzymes was scaled down to the size of our samples and optimized to our system, resulting in the final concentrations of 30 mU/ml of phosphodiesterase, 50 U/ml of Benzonase, 40 U/ml of alkaline phosphatase. Samples were hydrolyzed in the presence of the IS, ExBIM nucleoside (10 nM). No removal of enzymes by filtration was needed: the commonly performed step of enzyme filtration did not improve nor affect in any way the intensity of the ion signals during the final mass spectrometric analysis, while reduced the final recovery of both ExBenzi and ExBIM nucleosides. Hydrolyzed nucleosides and IS were purified by solid-phase extraction on a Strata-X cartridge (30 μ m, Phenomenex, Torrance, CA). The samples were loaded on the preconditioned cartridge, washed with 1 mL of 20% methanol in H₂O, and eluted with 1 mL of 80% methanol + 0.1% formic acid. The eluted fractions were evaporated to dryness and stored at -20 °C for no longer than one week before mass spectrometric analysis. Samples were dissolved in 10% acetonitrile and injected in the ESI-LC-MS/MS system.

Mass spectrometric method for the detection of ExBenzi nucleoside

Samples were analyzed by LC-MS/MS using a nanoAcquity UPLC system (Waters; California, USA) coupled to a triple quadrupole mass spectrometer (TSQ Vantage, Thermo Fisher; Massachusetts, USA) with an electrospray ionization source. An Atlantis dC18 column (150 x 0.3 mm, 3 μ m – Waters; California, USA) was used at a flow rate of 10 μ L/min and kept at a temperature of 40 °C. Prior analysis, the column was washed for 4 min and re-equilibrated for 9 min. Samples were run with a gradient starting at 15% B and increasing to 90% B in 16 min (mobile phase A was 3% acetonitrile in water with 0.1% formic acid, phase B was acetonitrile with 0.1% formic acid). The Mass spectra were recorded in positive ionisation SRM mode with following parameters: capillary temperature, 270 °C; spray voltage, 3000 V; sheath gas pressure, 10; ion sweep gas pressure, 0; aux gas pressure, 5; Q2 CID gas pressure, 1.5 mTorr; collision gas, argon; scan width, m/z 0.01; scan time, 0.1 s. The transitions used were m/z 301 \rightarrow 185 for ExBenzi nucleoside and m/z 285 \rightarrow 169 ExBIM nucleoside, both with Collision Energy 13 V. Xcalibur software (Thermo) was used for data acquisition and processing.

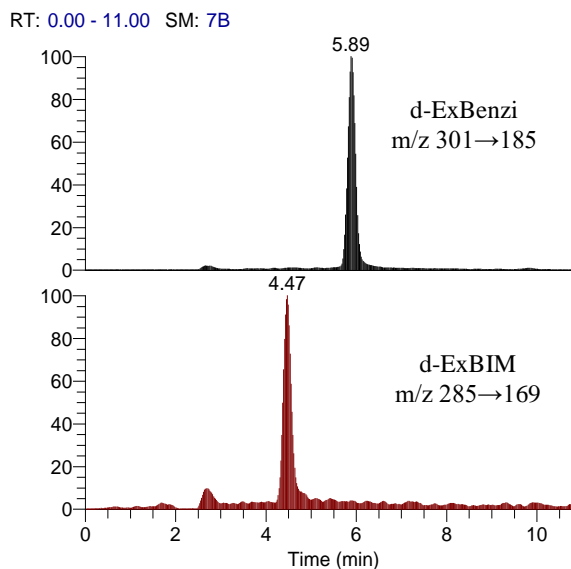


Figure S6 Mass chromatograms of ExBenzi and ExBIM nucleoside transition ions. DNA from linear amplification reactions was hydrolyzed, purified and the resulting nucleosides were separated by UPLC chromatography and ionized in a triple quadrupole mass spectrometer with electrospray ionization source and collision energy of 13 V. ExBenzi nucleoside eluted at 5.9 min and was identified via the transition m/z 301 \rightarrow 185 (top). ExBIM nucleoside eluted at 4.5 min and was identified via m/z 285 \rightarrow 169 transition (bottom).

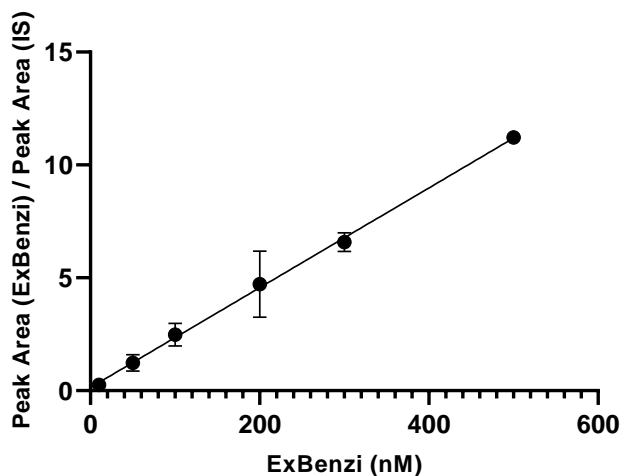


Figure S7 Calibration curve. DNA from linear amplification reactions performed in the absence of ExBenziTP were spiked with increasing concentrations of ExBenzi nucleoside (0 – 500 nM) and a fixed amount of ExBIM nucleoside (10 nM), hydrolysed, purified and analysed by mass spectrometry. The peak area of ExBenzi nucleoside was normalized to the one of ExBIM nucleoside and plotted against ExBenzi nucleoside concentration on GraphPad Prism 8 ($R^2 = 0.99$).

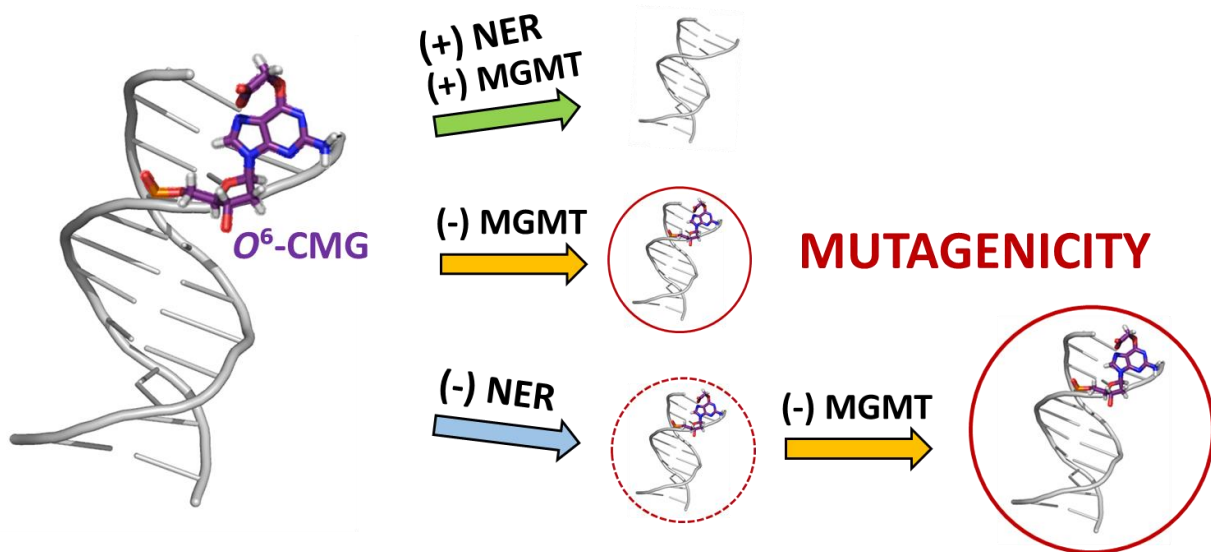
3.6 References

1. Du, H.; Wang, P.; Li, L.; Wang, Y., Repair and translesion synthesis of O⁶-alkylguanine DNA lesions in human cells. *J Biol Chem* **2019**, *294* (29), 11144-11153.
2. Ezerskyte, M.; Paredes, J. A.; Malvezzi, S.; Burns, J. A.; Margison, G. P.; Olsson, M.; Scicchitano, D. A.; Dreij, K., O⁶-methylguanine-induced transcriptional mutagenesis reduces p53 tumor-suppressor function. *Proc Natl Acad Sci* **2018**, *115* (18), 4731-4736.
3. Wang, P.; Wang, Y., Cytotoxic and mutagenic properties of O⁶-alkyl-2'-deoxyguanosine lesions in Escherichia coli cells. *J Biol Chem* **2018**, *293* (39), 15033-15042.
4. Margison, G. P.; Santibanez Koref, M. F.; Povey, A. C., Mechanisms of carcinogenicity/chemotherapy by O⁶-methylguanine. *Mutagenesis* **2002**, *17* (6), 483-7.
5. Roos, W. P.; Batista, L. F.; Naumann, S. C.; Wick, W.; Weller, M.; Menck, C. F.; Kaina, B., Apoptosis in malignant glioma cells triggered by the temozolomide-induced DNA lesion O⁶-methylguanine. *Oncogene* **2007**, *26* (2), 186-97.
6. Bennett, R. A.; Pegg, A. E., Alkylation of DNA in rat tissues following administration of streptozotocin. *Cancer Res* **1981**, *41* (7), 2786-90.
7. Montesano, R., Alkylation of DNA and tissue specificity in nitrosamine carcinogenesis. *J Supramol Struct Cell Biochem* **1981**, *17* (3), 259-73.
8. Harrison, K. L.; Jukes, R.; Cooper, D. P.; Shuker, D. E., Detection of concomitant formation of O⁶-carboxymethyl- and O⁶-methyl-2'-deoxyguanosine in DNA exposed to nitrosated glycine derivatives using a combined immunoaffinity/HPLC method. *Chemical research in toxicology* **1999**, *12* (1), 106-11.
9. Vanden Bussche, J.; Hemeryck, L. Y.; Van Hecke, T.; Kuhnle, G. G.; Pasmans, F.; Moore, S. A.; Van de Wiele, T.; De Smet, S.; Vanhaecke, L., O⁶-carboxymethylguanine DNA adduct formation and lipid peroxidation upon in vitro gastrointestinal digestion of haem-rich meat. *Mol Nutr Food Res* **2014**, *58* (9), 1883-96.
10. Lewin, M. H.; Bailey, N.; Bandaletova, T.; Bowman, R.; Cross, A. J.; Pollock, J.; Shuker, D. E.; Bingham, S. A., Red meat enhances the colonic formation of the DNA adduct O⁶-carboxymethyl guanine: implications for colorectal cancer risk. *Cancer research* **2006**, *66* (3), 1859-65.
11. Hemeryck, L. Y.; Rombouts, C.; Hecke, T. V.; Van Meulebroek, L.; Bussche, J. V.; De Smet, S.; Vanhaecke, L., In vitro DNA adduct profiling to mechanistically link red meat consumption to colon cancer promotion. *Toxicol Res* **2016**, *5* (5), 1346-1358.
12. Cross, A. J.; Ferrucci, L. M.; Risch, A.; Graubard, B. I.; Ward, M. H.; Park, Y.; Hollenbeck, A. R.; Schatzkin, A.; Sinha, R., A large prospective study of meat consumption and colorectal cancer risk: an investigation of potential mechanisms underlying this association. *Cancer Res* **2010**, *70* (6), 2406-14.
13. Steinberg, P., Red Meat-Derived Nitroso Compounds, Lipid Peroxidation Products and Colorectal Cancer. *Foods* **2019**, *8* (7).
14. Gottschalg, E.; Scott, G. B.; Burns, P. A.; Shuker, D. E., Potassium diazoacetate-induced p53 mutations in vitro in relation to formation of O⁶-carboxymethyl- and O⁶-methyl-2'-deoxyguanosine DNA adducts: relevance for gastrointestinal cancer. *Carcinogenesis* **2007**, *28* (2), 356-62.
15. Raz, M. H.; Dexter, H. R.; Millington, C. L.; van Loon, B.; Williams, D. M.; Sturla, S. J., Bypass of Mutagenic O⁶-Carboxymethylguanine DNA Adducts by Human Y- and B-Family Polymerases. *Chem Res Toxicol* **2016**, *29* (9), 1493-503.
16. Wu, J.; Wang, P.; Li, L.; Williams, N. L.; Ji, D.; Zahurancik, W. J.; You, C.; Wang, J.; Suo, Z.; Wang, Y., Replication studies of carboxymethylated DNA lesions in human cells. *Nucleic Acids Res* **2017**, *45* (12), 7276-7284.

17. Raz, M. H.; Sandell, E. S.; Patil, K. M.; Gillingham, D. G.; Sturla, S. J., High Sensitivity of Human Translesion DNA Synthesis Polymerase kappa to Variation in O⁶-Carboxymethylguanine Structures. *ACS Chem Biol* **2019**, *14* (2), 214-222.
18. Terasaki, M.; Totsuka, Y.; Nishimura, K.; Mukaisho, K.; Chen, K. H.; Hattori, T.; Takamura-Enya, T.; Sugimura, T.; Wakabayashi, K., Detection of endogenous DNA adducts, O-carboxymethyl-2'-deoxyguanosine and 3-ethanesulfonic acid-2'-deoxycytidine, in the rat stomach after duodenal reflux. *Cancer Sci* **2008**, *99* (9), 1741-6.
19. Cupid, B. C.; Zeng, Z.; Singh, R.; Shuker, D. E., Detection of O⁶-carboxymethyl-2'-deoxyguanosine in DNA following reaction of nitric oxide with glycine and in human blood DNA using a quantitative immunoslot blot assay. *Chem Res Toxicol* **2004**, *17* (3), 294-300.
20. Vanden Bussche, J.; Moore, S. A.; Pasmans, F.; Kuhnle, G. G.; Vanhaecke, L., An approach based on ultra-high pressure liquid chromatography-tandem mass spectrometry to quantify O⁶-methyl and O⁶-carboxymethylguanine DNA adducts in intestinal cell lines. *J Chromatogr A* **2012**, *1257*, 25-33.
21. Da Pieve, C.; Sahgal, N.; Moore, S. A.; Velasco-Garcia, M. N., Development of a liquid chromatography/tandem mass spectrometry method to investigate the presence of biomarkers of DNA damage in urine related to red meat consumption and risk of colorectal cancer. *Rapid Commun Mass Spectrom* **2013**, *27* (21), 2493-503.
22. Yu, Y.; Wang, J.; Wang, P.; Wang, Y., Quantification of Azaserine-Induced Carboxymethylated and Methylated DNA Lesions in Cells by Nanoflow Liquid Chromatography-Nanoelectrospray Ionization Tandem Mass Spectrometry Coupled with the Stable Isotope-Dilution Method. *Anal Chem* **2016**, *88* (16), 8036-42.
23. Wang, Y.; Patil, K.M.; Yan, S.; Zhang P.; Guo, W.; Wang, Y.; Chen, H.Y.; Gillingham, D.; Huang, S, Nanopore Sequencing Accurately Identifies the Mutagenic DNA Lesion O⁶-Carboxymethyl Guanine and Reveals Its Behavior in Replication. *Angewandte Chemie* **2019**.
24. Raz, M. H.; Aloisi, C. M. N.; Gahlon, H. L.; Sturla, S. J., DNA Adduct-Directed Synthetic Nucleosides. *Acc Chem Res* **2019**, *52* (5), 1391-1399.
25. Gong, J.; Sturla, S. J., A synthetic nucleoside probe that discerns a DNA adduct from unmodified DNA. *J Am Chem Soc* **2007**, *129* (16), 4882-3.
26. Gahlon, H. L.; Sturla, S. J., Hydrogen bonding or stacking interactions in differentiating duplex stability in oligonucleotides containing synthetic nucleoside probes for alkylated DNA. *Chemistry* **2013**, *19* (33), 11062-7.
27. Trantakis, I. A.; Nilforoushan, A.; Dahlmann, H. A.; Stauble, C. K.; Sturla, S. J., In-Gene Quantification of O⁶-Methylguanine with Elongated Nucleoside Analogues on Gold Nanoprobes. *J Am Chem Soc* **2016**, *138* (27), 8497-504.
28. Aloisi, C. M. N.; Sturla, S. J.; Gahlon, H. L., A gene-targeted polymerase-mediated strategy to identify O⁶-methylguanine damage. *Chem Commun* **2019**, *55* (27), 3895-3898.
29. Wyss, L. A.; Nilforoushan, A.; Eichenseher, F.; Suter, U.; Blatter, N.; Marx, A.; Sturla, S. J., Specific incorporation of an artificial nucleotide opposite a mutagenic DNA adduct by a DNA polymerase. *J Am Chem Soc* **2015**, *137* (1), 30-3.
30. Wyss, L. A.; Nilforoushan, A.; Williams, D. M.; Marx, A.; Sturla, S. J., The use of an artificial nucleotide for polymerase-based recognition of carcinogenic O⁶-alkylguanine DNA adducts. *Nucleic Acids Res* **2016**, *44* (14), 6564-73.
31. Dahlmann, H. A. B., F.D.; Kung, R.W.; Wyss L.A.; Gubler, I.; McKeague M.; Wetmore S.D.; Sturla S.J., Fluorescent Nucleobase Analogues with Extended Pi Surfaces Stabilize DNA Duplexes Containing O⁶-Alkylguanine Adducts *Helvetica Chimica Acta* **2018**, *101* (7).
32. Gloeckner, C.; Sauter, K. B.; Marx, A., Evolving a thermostable DNA polymerase that amplifies from highly damaged templates. *Angew Chem Int Ed Engl* **2007**, *46* (17), 3115-7.

33. Betz, K.; Nilforoushan, A.; Wyss, L. A.; Diederichs, K.; Sturla, S. J.; Marx, A., Structural basis for the selective incorporation of an artificial nucleotide opposite a DNA adduct by a DNA polymerase. *Chem Commun* **2017**, 53 (94), 12704-12707.
34. Hubscher, U., DNA Polymerases: Discovery, Characterization, and Functions in Cellular DNA Transactions. *World Scientific* **2010**.
35. Kowal, E. A.; Lad, R. R.; Pallan, P. S.; Dhummakupt, E.; Wawrzak, Z.; Egli, M.; Sturla, S. J.; Stone, M. P., Recognition of O⁶-benzyl-2'-deoxyguanosine by a perimidinone-derived synthetic nucleoside: a DNA interstrand stacking interaction. *Nucleic Acids Res* **2013**, 41 (15), 7566-76.
36. Yoon, H. H.; Tougeron, D.; Shi, Q.; Alberts, S. R.; Mahoney, M. R.; Nelson, G. D.; Nair, S. G.; Thibodeau, S. N.; Goldberg, R. M.; Sargent, D. J.; Sinicrope, F. A., KRAS codon 12 and 13 mutations in relation to disease-free survival in BRAF-wild-type stage III colon cancers from an adjuvant chemotherapy trial (N0147 alliance). *Clin Cancer Res* **2014**, 20 (11), 3033-43.
37. Gloeckner, C.; Kranaster, R.; Marx, A., Directed evolution of DNA polymerases: construction and screening of DNA polymerase mutant libraries. *Curr Protoc Chem Biol* **2010**, 2 (2), 89-109.
38. Wang, Y.; Ngor, A. K.; Nikoomanzar, A.; Chaput, J. C., Evolution of a General RNA-Cleaving FANA Enzyme. *Nat Commun* **2018**, 9 (1), 5067.
39. Tretyakova, N.; Goggin, M.; Sangaraju, D.; Janis, G., Quantitation of DNA adducts by stable isotope dilution mass spectrometry. *Chemical research in toxicology* **2012**, 25 (10), 2007-35.
40. Szychowski, J.; Mahdavi, A.; Hodas, J. J.; Bagert, J. D.; Ngo, J. T.; Landgraf, P.; Dieterich, D. C.; Schuman, E. M.; Tirrell, D. A., Cleavable biotin probes for labeling of biomolecules via azide-alkyne cycloaddition. *J Am Chem Soc* **2010**, 132 (51), 18351-60.
41. Sloan, D. B.; Broz, A. K.; Sharbrough, J.; Wu, Z., Detecting Rare Mutations and DNA Damage with Sequencing-Based Methods. *Trends Biotechnol* **2018**, 36 (7), 729-740.
42. Wu, J.; McKeague, M.; Sturla, S. J., Nucleotide-Resolution Genome-Wide Mapping of Oxidative DNA Damage by Click-Code-Seq. *J Am Chem Soc* **2018**, 140 (31), 9783-9787.
43. Hu, J.; Adebali, O.; Adar, S.; Sancar, A., Dynamic maps of UV damage formation and repair for the human genome. *Proceedings of the National Academy of Sciences of the United States of America* **2017**, 114 (26), 6758-6763.
44. Shu, X.; Xiong, X.; Song, J.; He, C.; Yi, C., Base-Resolution Analysis of Cisplatin-DNA Adducts at the Genome Scale. *Angew Chem Int Ed Engl* **2016**, 55 (46), 14246-14249.
45. Li, W.; Hu, J.; Adebali, O.; Adar, S.; Yang, Y.; Chiou, Y. Y.; Sancar, A., Human genome-wide repair map of DNA damage caused by the cigarette smoke carcinogen benzo[a]pyrene. *Proceedings of the National Academy of Sciences of the United States of America* **2017**, 114 (26), 6752-6757.
46. Basu, A. K.; Nohmi, T., Chemically-Induced DNA Damage, Mutagenesis, and Cancer. *Int J Mol Sci* **2018**, 19 (6).
47. Stornetta, A.; Zimmermann, M.; Cimino, G. D.; Henderson, P. T.; Sturla, S. J., DNA Adducts from Anticancer Drugs as Candidate Predictive Markers for Precision Medicine. *Chem Res Toxicol* **2017**, 30 (1), 388-409.
48. Boosalis M.S., Petruska J., Goodman M.F., DNA polymerase insertion fidelity. Gel assay for site-specific kinetics., *J. Biol. Chem.* **1987**, 262, 14689-14696
49. Gahlon H.L., Sturla S.J, Determining Steady-State Kinetics of DNA Polymerase Nucleotide Incorporation, *Methods Mol Biol.* **2019**, 1973, 299-311
50. Quinlivan, E. P.; Gregory, J. F., DNA digestion to deoxyribonucleoside: a simplified one-step procedure. *Anal Biochem* **2008**, 373, 2, 383-385

Chapter 4: MGMT and NER alleviate O^6 -Carboxymethyl-Guanine DNA Adduct Mutagenicity in Cells



O^6 -CMG = O^6 -Carboxymethyl-G; NER = nucleotide excision repair; MGMT = O^6 -Methyl-G DNA methyltransferase

Claudia M.N. Aloisi, Nora A. Escher, Susanne M. Geisen, Hyun S. Kim, Jung E. Yeo, Orlando D. Schärer, and Shana J. Sturla*

C.M.N Aloisi devised experiments, synthesized O^6 -CMG oligonucleotide, produced O^6 -CMG-sscDNA, performed cellular assays, EMSA, RNA and DNA extractions, qPCR, DNA amplification for sequencing, interpreted data and wrote the manuscript. N.A. Escher produced sscDNA, performed cellular assays and HPRT assays, RNA and DNA extractions, qPCR, DNA amplification for sequencing, and interpreted data. S.M. Geisen performed mass spectrometric measurements. J.E. Yeo produced sscDNA, purified O^6 -CMG-dscDNA. H.S. Kim performed NER assay. O.D. Schärer advised on experiments and interpreted data. S.J. Sturla interpreted data and wrote the manuscript.

4.1 Abstract

*O*⁶-alkylguanines are some of the most mutagenic damage to the DNA and have been associated with carcinogenesis. The occurrence of *O*⁶-carboxymethylguanine (*O*⁶-CMG) in particular has been linked to meat consumption and hypothesized to contribute to the development of colorectal cancer. However, the cellular fate of *O*⁶-CMG is poorly characterized and the literature on repair mechanisms that protect cells from *O*⁶-CMG mutagenicity is elusive and contradictory. With the present work, we unravelled a longstanding question in DNA repair concerning the identification of how cells target and remove *O*⁶-CMG. We found that cells deficient in either *O*⁶-methylguanine-DNA methyltransferase enzyme or nucleotide excision repair machinery are more sensitive to carboxymethylation. Furthermore, we demonstrated that impairment in any of those pathways and the insufficient repair of *O*⁶-CMG give rise to mutations in cells. Future studies will focus on defining the mechanisms of repair, and characterizing the induced mutational spectra, especially in colorectal cancer relevant genes. The findings and the planned experiments of this work will advance our understanding of the relationship between *O*⁶-CMG DNA damage and mutation, paving the way for the further study of mutagenicity and cancer etiology.

4.2 Introduction

DNA integrity is continuously endangered by exogenous alkylating agents or as a result of endogenous cellular events. To prevent adverse biological consequences, the cellular DNA damage response (DDR)¹ is activated, consisting of an entangled network of pathways to sense, signal and repair the damage, or mechanisms of tolerance that allows for overcoming the replication blockage of the damage and evade its lethal consequences. However, processes of DNA damage tolerance can introduce mutations in the sequence of DNA, increasing the likelihood of carcinogenesis when cancer driver genes are affected.

*O*⁶-alkyl-guanines (*O*⁶-alkylGs) are mutagenic DNA adducts that have been linked to the development of cancer.^{2, 3} They are formed by exposure to endogenous and exogenous N-nitroso compounds (NOCs)⁴ or to other chemicals that also decompose into alkyl-diazonium cations, such as the anticancer agent temozolomide.⁵ Common exogenous sources of NOCs and of their precursors that humans are exposed to are drugs,⁶ tobacco smoke,⁷ red and processed meat.⁷ NOCs precursors include nitrogenous dietary compounds that can be converted to NOCs through nitrite-derived nitrosation,⁸ expected to happen particularly in the acidic gastric environment.⁹ In addition to exogenous sources and *in vivo* nitrite-induced formation, NOCs can be endogenously produced in human cells, via mammalian cell-mediated nitrosation¹⁰ and bacterial nitrosation.¹¹

The adducts O^6 -methylguanine (O^6 -MeG) and O^6 -carboxymethylguanine (O^6 -CMG) in particular are relevant for human health as they have been found in human DNA blood samples of cancer patients and exfoliated colonocytes of meat-eaters.^{7, 12-14} Since O^6 -CMG has been detected in human colorectal DNA,¹⁴ it has been hypothesized as a potential initiator in the development of colorectal cancer (CRC) linked to diets high in meat.¹⁵ Additionally, the mutation spectrum observed in human CRC resembles the one induced by the carboxymethylating agent potassium diazoacetate (KDA).¹⁶ Spectra measured in the cancer-relevant gene *p53* of CRC tissues and in KDA-treated *p53* gene-containing plasmid were similarly rich in G>A mutation, that could arise from the misincorporation of T opposite O^6 -CMG during replication.¹⁷ Studies interrogating the fidelity of replication past O^6 -CMG performed *in vitro*¹⁷⁻¹⁹ and *in vivo* in bacteria²⁰ support the incorporation of wrong bases opposite O^6 -CMG as a basis for its mutagenicity, providing a potential cause-effect relationship with carcinogenesis. While the structural and mechanistic basis of O^6 -CMG mutagenicity have been investigated, there is a substantial lack of knowledge regarding the repair of O^6 -CMG and how this could alleviate the rise of mutations. The repair of O^6 -MeG has been extensively studied,^{21, 22} however, the role of different repair pathways in targeting O^6 -CMG remains elusive.

In a study from 1999, the sensitivity to the carcinogenic carboxymethylating agent azaserine of human cell lines with defined defects in DNA repair pathways was investigated.²³ Cells deficient in nucleotide excision repair (NER) factors were sensitive to azaserine, consistent with previous studies conducted in NER-deficient bacteria and yeast cells^{24, 25} and with the finding that azaserine is mutagenic in *E.Coli*.²⁶ By contrast, deficiencies in mismatch repair (MMR) pathway and O^6 -methylguanine-DNA methyltransferase (MGMT) repair protein did not influence cell survival.²³ The observation that cells deficient in MGMT were not more sensitive to azaserine corroborated studies by Shuker and Margison from 1997, where the transfer of the carboxymethyl from O^6 -CMG-oligonucleotides to purified alkyltransferases was not observed by mass spectrometry for any of the tested enzymes.²⁷

MGMT is an alkyl transferase enzyme highly proficient at targeting and removing O^6 -MeG from DNA.²⁸ Individuals deficient in MGMT are particularly sensitive to methylating agents,²⁹ and MGMT chemically-induced depletion is currently exploited in combination with alkylating anticancer drugs for cancer therapy.^{30, 31} Given the well-established role of MGMT in repairing O^6 -MeG, the activity of MGMT towards O^6 -CMG was readdressed in 2013 with a similar experimental set-up of that from 1997, and O^6 -CMG was a competent substrate of the purified human MGMT; however, experiments to confirm that the human MGMT repairs O^6 -CMG in cells have not been performed.³² In a recent work, MGMT has been suggested to repair other O^6 -alkyGs, especially those equipped with a small alkyl group, while bulkier alkyGs were targeted by the NER repair machinery.³³ But O^6 -CMG was not tested. To our knowledge, the role of NER and MGMT in repairing O^6 -CMG in human cells was only assayed by O'Driscoll et al. in 1999, and

results did not support any significant role of MGMT in the removal of O⁶-CMG.²³ More studies addressing the cellular fate of O⁶-CMG DNA damage are needed, with particular focus on the MGMT and NER pathways, considering (1) that there is contradictory evidence regarding MGMT activity towards O⁶-CMG,^{23, 32} (2) the established role of NER in the repair of O⁶-alkylG DNA damage in lower organisms,^{25, 34-36} and (3) the recent findings supporting a mutual contribution of NER and MGMT in repairing other O⁶-alkylGs DNA adducts in human cells.³³ Nevertheless, the contribution of other repair pathways, especially mismatch repair, should be tested.

The objective of this study was to resolve the relative contribution of NER and MGMT to O⁶-CMG persistence and mutagenicity in cells. We found that the carboxymethylating agent azaserine was more toxic and significantly more mutagenic in cells depleted of MGMT, probably due to the accumulation of O⁶-CMG adducts. NER-deficient cells also showed decreased cell survival and increased mutagenicity in response to carboxymethylation when compared to their repair proficient counterpart. While the combined impairment of NER and MGMT did not affect cell survival measurably, mutagenicity was significantly higher in comparison to cells proficient in at least one of the two pathways, raising the question on whether the activity of NER and MGMT in the removal of O⁶-CMG in the cell might be synergistic or otherwise inter-dependent. Preliminary results of cell-free NER assay contradicted findings in cells. Thus, we devise and suggest here studies to further characterize the role of NER in the repair of O⁶-CMG, including testing the hypothesis of a putative crosstalk between repair pathways.

4.3 Results

4.3.1 Reduced MGMT function sensitizes cells to carboxymethylating agent azaserine

Given the lack of consensus in the literature concerning whether MGMT repairs carboxymethylated DNA, we aimed to test the cellular impact of MGMT inhibition on cellular phenotype induced by azaserine. Previously, cell survival in response to azaserine has been employed to identify DNA repair pathways responsible to remove azaserine-induced DNA damage, and particularly O⁶-CMG.²³ However, azaserine toxicity can also result from other activities of the drug in the cell, such as the inhibition of nucleotide synthesis.³⁷ Therefore, to further assess the cellular response to azaserine-induced DNA damage as a function of DNA repair capacity, we were also interested in characterizing the contribution of MGMT repair in azaserine mutagenicity, which is expected to be more directly linked to DNA adducts. Thus, for our studies, we employed V79-4 hamster cells that allow the measurement of mutagenicity via the HPRT assay, which has been extensively performed on rodent cells³⁸⁻⁴¹ and could not be easily translated to human cells (not shown). To assess the contribution of MGMT, we chemically depleted V79-4 cells of MGMT by exposure to O⁶-benzyguanine (O⁶-BnG), which is a

well-established, potent MGMT inhibitor.⁴² We identified sub-toxic concentrations of O^6 -BnG to be below 60 μ M (Fig. S1A), and therefore used doses up to 60 μ M in our further studies in V79-4 cells.

To measure azaserine cytotoxicity in MGMT impaired cells, we exposed V79-4 cells to azaserine (0-5 μ M) and to O^6 -BnG (0, 40 μ M) for 24 h and measured cell viability after 2 days of incubation in drug-free media. We found that cell survival of V79-4 cells was lower when cells were also exposed to O^6 -BnG (Fig. 1A, top). The 2 days of post-incubation were selected according to results of cell proliferation obtained from V79-4 cells treated with azaserine (1 μ M) or co-treated with azaserine and O^6 -BnG (40 μ M) for 24 h. Cells were counted directly after treatment and every 24 h for 3 days. By allowing cells to grow after exposure to azaserine, the difference in cell number associated with MGMT proficiency was enhanced and better measurable. The resulting curve of cell growth was lower for cells co-exposed to O^6 -BnG, suggesting that MGMT inhibition effectively diminished the number of cells by inducing cell death, and possibly slowed down growth of remaining viable cells (Fig. 1A, bottom).

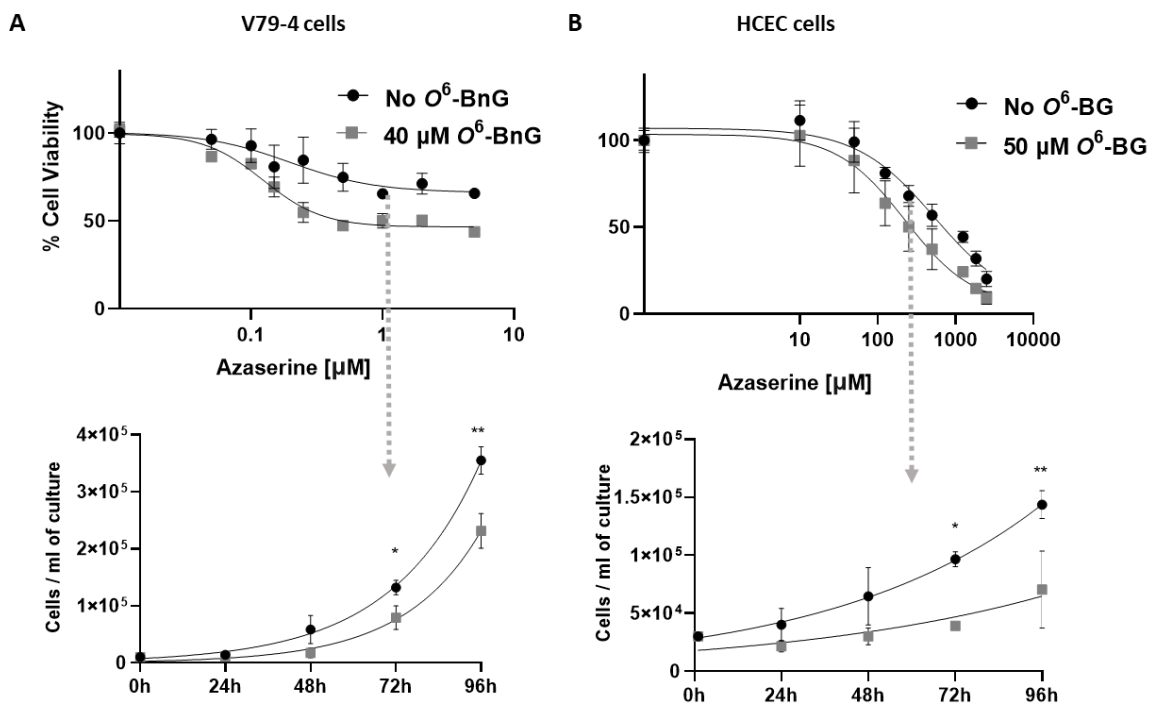


Figure 1. Cell viability and cell proliferation upon azaserine treatment. V79-4 (A), HCEC (B) cells were exposed for 24 h to increasing concentrations of azaserine (0-5 μ M for V79-4; 0-5 mM for HCEC) and to O^6 -BnG (40 μ M for V79-5; 50 μ M for HCEC). Cell viability (top graphs) was measured by CellTiter-Glo assay 2 or 3 days after the 24 h-exposure. To measure cell proliferation (bottom graphs), cells were counted immediately after treatment (24 h) with azaserine (1 μ M for V79-5 and 200 μ M for HCEC; concentration indicated by dashed arrow) and O^6 -BnG, or after 1 day (48 h), 2 days (73 h) or 3 days (96 h). Data were represented using

GraphPad Prism 8. Cell viability data were fit in a nonlinear regression curve, using a dose-response function; cell proliferation data were fit in an exponential curve, using a growth equation. Error bars indicate standard deviation. Experiments were performed three independent times, each with three technical replicates. Two-way ANOVA with Sidak's multiple comparison test was performed to compare cell proliferation between +/- O⁶-BnG at each time point (*p<0.05, **p<0.01).

Having observed that MGMT inhibition does increase the sensitivity of V79-4 cells to azaserine, we were interested in whether a similar phenotype could be observed in human cells. The MGMT human gene shares homology with other species including rodents, and, analogously to humans, mice deficient in MGMT were hypersensitive to alkylating agents and showed increased tumorigenesis.^{29, 43-45} Thus, we measured azaserine cytotoxicity upon MGMT inhibition in HCEC cells, immortalized but normal diploid cell line, derived from health human colon epithelium^{46, 47} and previously employed as model for healthy colon tissue.⁴⁸ We exposed HCEC cells to azaserine (0-5 mM) and to O⁶-BnG (50 μM) for 24 h and measured cell viability after 3 days of incubation in drug-free media. We found that cell viability was lower when cells were depleted of MGMT by exposure to O⁶-BnG (Fig. 1 B, top). The 3 days of post-incubation were selected upon measurement of cell proliferation over 4 days in HCEC cells exposed to azaserine (200 μM) or to azaserine and O⁶-BnG (50 μM). Similarly to what observed for V79-4 cells, the number of cells counted overtime was significantly lower for in those samples where MGMT was inhibited (Fig. 1 B, bottom), possibly as a consequence of cell death or reduced cellular growth.

4.3.2 Azaserine is more mutagenic in cells when MGMT is inhibited

Having established the influence of MGMT on the sensitivity of cells to azaserine, we anticipated a potentially more significant and relevant impact on mutagenicity due to the cause-effect relationship of DNA adducts with mutation. We performed the HPRT mutagenicity assay in V79-4 rodent cells exposed to azaserine and depleted of MGMT activity by exposure to O⁶-BnG. Ethyl methansulfonate (EMS) was employed as a positive control (Fig. S2). Preliminary tests were performed to identify the concentration of azaserine to test (Fig. S2) and the time of exposure. Final conditions consisted in 24 h-treatment, one week-incubation to express mutations, and one week-incubation in selection media to select mutants generated by the treatment.

We found that azaserine was mutagenic in V79-4 cells, and that mutagenicity was dose-dependent (Fig. 2A, black bars: 0; 0.5; 2 μM azaserine). O⁶-BnG instead was not mutagenic at the tested concentrations (Fig.2B, black bars), but significantly increased azaserine mutagenicity (Fig. 2B). The enhanced mutagenicity of azaserine upon MGMT inhibition suggests that the

depletion of MGMT could result in the accumulation and persistence of azaserine-induced O^6 -CMG DNA adducts, and give rise to mutations.

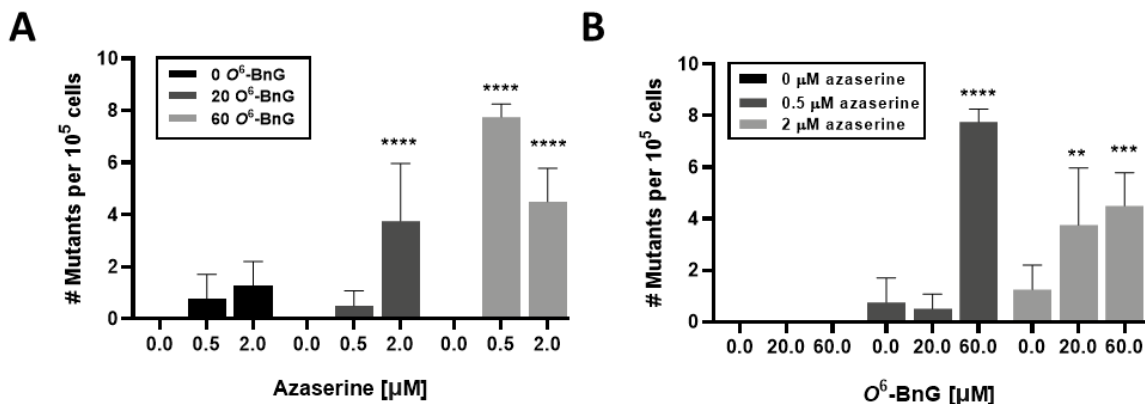


Figure 2. Mutagenicity in V79-4 cells exposed to azaserine and O^6 -BnG. V79-4 cells were exposed to azaserine (0, 0.5, 2 μ M) and of O^6 -BnG (0, 20, 60 μ M). HPRT assay was performed as described and mutant colonies were counted manually after two weeks. Data were represented using GraphPad Prism 8. Graphs A and B represent the same data, with significance values calculated by comparing to azaserine-untreated (A) or to O^6 -BnG-untreated (B) samples. Statistical significance was obtained by performing two-way ANOVA with a Tukey post hoc test (* p <0.05, ** p <0.01, *** p <0.001, **** p <0.0001). Significance is relative to untreated control for each set of data.

4.3.3 Sequencing data for *hprt* gene in MGMT-depleted V79-4 cells

To characterize findings regarding azaserine mutagenicity in cells depleted of MGMT and obtain the resulting mutational spectra, we sequenced the *hprt* gene of V79-4 cells exposed to azaserine and O^6 -BnG under the same conditions of the HPRT assay. DNA was extracted from cells, amplified via PCR reaction using Q5 polymerase and sent for Sanger sequencing. If the background obtained from Sanger sequencing will be too high, we will consider a more sensitive sequencing approach, such as pyrosequencing. Results from sequencing analysis are currently underway. Anticipated outcomes are that mutations in the *hprt* gene will increase with the dose of azaserine and with the dose of O^6 -BnG, i.e. increase upon loss of MGMT capacity. Additionally, we will characterize the mutational spectra induced by azaserine.

Furthermore, to test the clinical relevance of MGMT-dependent increase of azaserine mutagenicity, future studies will include the sequencing of genes that are linked to the development of CRC, namely *k-ras*, *p53* and *apc*, from HCEC human cells upon azaserine and O^6 -BnG exposure. In addition to the relative number of mutations across conditions (i.e. among doses of drugs), the location and the type of mutation will be analysed, and the rate of mutations across the three gene will be determined. From these studies, we expect to gain

important information to address questions concerning cancer development and the contribution of *O*⁶-CMG accumulation in cells.

4.3.4 *O*⁶-CMG adduct levels induced by azaserine in MGMT-depleted V79-4 cells

To test the hypothesis that the increase in mutagenicity observed upon MGMT inhibition is due to the persistence of *O*⁶-CMG adducts, adduct levels in azaserine-exposed cells as a function of *O*⁶-BnG exposure are being measured by mass spectrometry. V79-4 cells were treated with azaserine (0 or 2 μM) and with *O*⁶-BnG (60 μM) for 24 h; DNA was extracted from cells and purified for mass spectrometry analysis. Untreated samples were prepared as controls and to define the background level of *O*⁶-CMdG. Although azaserine can induce both *O*⁶-MeG and *O*⁶-CMG, it was previously reported that *O*⁶-MeG is formed at around 10 or 20-fold lower level than *O*⁶-CMG, depending on the cell line.⁴⁹ Furthermore, the phenotype (e.g. mutagenicity) induced by azaserine or by its analogue diazoacetate is usually ascribed to *O*⁶-CMG, as it differs greatly from the phenotype of *O*⁶-MeG or from the one induced by methylating agents.^{16, 23} Thus, to test whether *O*⁶-CMG is a major contributor of azaserine phenotypes in V79-4 cells too, we will also measure the level of *O*⁶-MedG.

Preliminary data obtained from other cells exposed to azaserine for 24 h (not shown) suggested an accumulation of *O*⁶-CMG upon MGMT inhibition, and confirmed the significantly lower occurrence of *O*⁶-MeG induced by azaserine, which has been previously reported. Adduct level measurements on V79-4 cells are currently underway. Anticipated results are that the level of adducts induced by azaserine will mirror the level of mutations obtained from the HPRT assay (Fig. 2) and will be dose and repair proficiency-dependent. Because the kinetics of *O*⁶-CMG formation by azaserine and repair is not well-characterized, we also anticipate that time course measurements might need to be performed.

Overall, with the presented studies on MGMT inhibition, we established that MGMT has a significant impact on the cellular response to azaserine. We found that the inhibition of MGMT sensitizes mammalian cells to carboxymethylation, and increases the rise of mutations. Further characterization of azaserine phenotype in association with MGMT should be carried out (e.g. MGMT expression level, preliminary tests on Fig. S3), as well as a more direct characterization of the MGMT role in the repair of *O*⁶-CMG (e.g. experiments to assay *O*⁶-CMG removal or MGMT binding to *O*⁶-CMG).

Considered the general lack of studies regarding the repair of *O*⁶-CMG damage, and that our findings on MGMT are contradictory with a previous work where MGMT deficiency had no significant impact on human cell sensitivity to azaserine,²³ future effort should focus on testing the role of other repair pathways in removing *O*⁶-CMG from the DNA. In particular, NER is known to repair *O*⁶-alkylG DNA damage in lower organisms,^{25, 34-36} and was previously suggested

to have a protective role in response to carboxymethylation in human cells.²³ Therefore, we were interested in confirming and further characterizing the role of NER in response to azaserine and to *O*⁶-CMG DNA damage.

4.3.5 Cells deficient in NER are more sensitive to carboxymethylating agent azaserine

To test the impact of cellular deficiencies in NER repair on azaserine phenotype, we characterized the relationship between responses of cells exposed to azaserine with the expression of XPA, XPC or CSB (HAP1 human cells), and ERCC1 (CHO rodent cells) NER proteins. In contrast to MGMT repair, which is mediated by a single enzyme, NER involves the concerted function of several factors, and it is made of two sub-pathways, global genome (GG) and transcription coupled (TC)-NER, which differ in the step of lesion recognition. GG-NER is initiated by XPC-RAD23B that recognizes DNA duplex distorting lesions, and TC-NER is initiated by the stalling of transcribing RNA polymerase with the help of CSB. XPA is a central NER factor, shared by both sub-pathways, which establishes a multitude of interactions with other NER proteins, probably recruiting and directing them. Finally, ERCC1-XPF is a NER endonuclease, responsible for the incision 5' to the lesion.

To assess the contribution of NER to azaserine cytotoxicity, we employed CHO rodent cells: CHO UV20 (ERCC1-deficient) cells and CHO UV20 cells transduced with a lentiviral vector expressing wild-type *hERCC1*, therefore "WT" in the NER function. We exposed CHO cells to increasing concentrations of azaserine (0-10 μ M) for 24 h and measured cell viability after 2 days of incubation in drug-free media. Cell survival in response to azaserine treatment was lower in CHO UV20 repair deficient cells, when compared to CHO WT cells (Fig. 3A). Noteworthy, no significant difference between CHO WT and CHO UV20 survival upon exposure to azaserine was recorded when measuring cell viability directly after the 24h-treatment (Fig. S4), suggesting that the time of measurement is critical when investigating the impact of DNA damage and the role of cellular repair pathways, as observed for other cells (Fig. 1, top graphs).

Having established that cellular deficiency of the NER protein ERCC1 sensitizes to azaserine CHO rodent cells, chosen for further mutagenicity testing, we were interested to test whether we could measure a similar phenotype in human cells. We tested cytotoxicity of azaserine in HAP1 cells deficient in XPA, and in HAP1 cells deficient in XPC or CSB, to dissect the contribution of GG-NER and TC-NER sub-pathways. Additionally, to assess whether azaserine affects the capacity of cells to form colonies, we performed clonogenic assays in HAP1 cells. However, HAP1 cells were resistant to azaserine treatment, displaying a plateau of cell viability at 50 % or above (Fig. S5A-B) and recovering throughout the clonogenic assays, even upon one-week exposure to high concentrations (Fig. S5C-D). Therefore, results from HAP1 cells were not conclusive, and HAP1 cells were not further tested. To test sensitivity of human cells to azaserine, as well as confirm that the phenotype observed for CHO cells is related to NER and

not specifically to the ERCC1 factor, future studies will include testing azaserine cytotoxicity in immortalized human cells derived from patients affected by other defined NER deficiency.

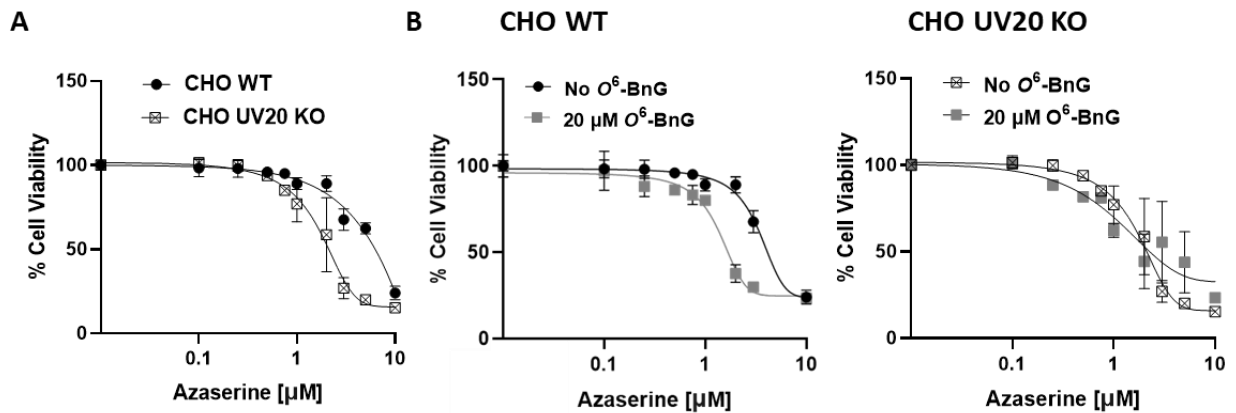


Figure 3. Cell viability and cell proliferation upon azaserine treatment. (A) CHO WT and CHO UV20 KO cells were treated for 24 h with increasing concentrations of azaserine (0-10 μM) or (B) co-treated with O^6 -BnG (20 μM). Cell viability was measured by CellTiter-Glo assay after two days. Data were represented using GraphPad Prism 8. Cell viability data were fit in a nonlinear regression curve, using a dose-response function. Error bars indicate standard deviation. Experiments were performed three independent times, each with three technical replicates.

4.3.6 Cells deficient in NER are not sensitized to carboxymethylation when depleted of MGMT

Having measured significant changes in response to azaserine of cells depleted of MGMT or NER separately, we were interested to test whether the simultaneous impairment of both NER and MGMT could have an additive effect on increased cellular sensitivity to azaserine. Therefore, we exposed CHO cells, WT and UV20, to increasing concentrations of azaserine (0-10 μM) and to a sub-toxic (Fig. S6) dose of O^6 -BnG (20 μM). While CHO WT cells were more sensitive to azaserine when also O^6 -BnG was included, viability of CHO UV20 cells with or without O^6 -BnG was not significantly different (Fig. 4B). The observation that the depletion of MGMT had no measurable effect on azaserine cytotoxicity in NER-deficient cells might be due to experimental limitations (e.g. high variability of the signal, wrong time point), or to saturation of azaserine toxic properties, or indicate no further dependence of azaserine toxicity on repair proficiency. Further studies need to be conducted to establish more directly what is the combined impact of NER and MGMT in response to azaserine and azaserine-induced DNA damage.

4.3.7 Azaserine is more mutagenic in cells deficient in NER

Given the higher the sensitivity of NER deficient CHO cells to azaserine toxicity, we anticipated a potentially similar impact on mutagenicity. To test the contribution of NER in azaserine mutagenicity, HPRT assay was performed in CHO WT and CHO UV20 cells exposed to

azaserine for 24 h. We found that azaserine was mutagenic in CHO cells at the highest dose tested (2 μM), and that NER-deficient CHO cells were significantly more mutated than the CHO WT cells upon exposure to azaserine (Fig. 4A), despite the variability in the level of mutants counted per dish that can be a limitation intrinsic of the HPRT assay. Noteworthy, the mutagenic concentration of 2 μM azaserine corresponded to around 50% viability in CHO cells (Fig. 3A), which is comparable to the toxicity of doses employed for mutagenicity testing in V79-4 cells and that were also mutagenic. Instead, the sub-toxic dose of 0.75 μM of azaserine, which causes only 10% reduction of viability in CHO cells (Fig. 3A), was not mutagenic.

4.3.8 Azaserine mutagenicity is significantly higher when both NER and MGMT are depleted in cells

In this work, we have established an increase of azaserine mutagenicity induced separately by MGMT inhibition and by NER deficiency in cells, which is likely caused by the accumulation of unrepaired DNA damage. Despite having obtained non-conclusive results from cell survival experiments where cells were impaired in both NER and MGMT, we were interested to assess whether assaying azaserine mutagenicity could give a more significant and relevant measurement of a mutual or additive impact of repair pathways on the phenotype of azaserine. Therefore, we exposed CHO cells, WT and UV20, to azaserine and co-treated with O^6 -BnG (20 μM). We found that when MGMT was inhibited by O^6 -BnG, azaserine was generally more mutagenic in CHO cells, inducing mutants also at the lower dose of 0.75 μM . Remarkably, the increase induced in NER-deficient CHO UV20 cells was significantly higher than the increase induced in CHO WT cells. Additionally, the difference between CHO UV20 and CHO WT *hprt* mutants is less than 3-fold when cells are exposed to 2 μM azaserine and no O^6 -BnG (Fig. 4A, light grey bars), but it reaches a difference of almost 10-fold when MGMT is inhibited (Fig. 4B, dark grey bars). Thus, we wondered whether MGMT and NER repair pathways could have a synergistic protective role towards azaserine-induced mutagenicity.

Noteworthy, some albeit low mutagenicity was measured for the first time in the absence of azaserine (negative control) when MGMT was inhibited in CHO NER deficient cells. Due to the absence of any mutagen, we reasoned that mutations might have arisen from mutagenic endogenous damage, from which the cell is left particularly unprotected when both NER and MGMT repair pathways are absent or defective. Because neither MGMT nor NER separate depletion was sufficient to give rise to mutations in CHO cells, we reasoned that either (1) MGMT- or NER-dependent mutagenicity arising from endogenous damage is individually below the limit of detection of the HPRT assay, or (2) that MGMT and NER might have a synergistic role towards endogenous damage. Although the former hypothesis is likely, especially given the low frequency of mutants observed in the untreated CHO UV20 KO + O^6 -BnG samples, more effort

should be invested to interrogate a putative relationship between MGMT and NER in DNA damage repair and specifically for the case of *O*⁶-CMG DNA adduct.

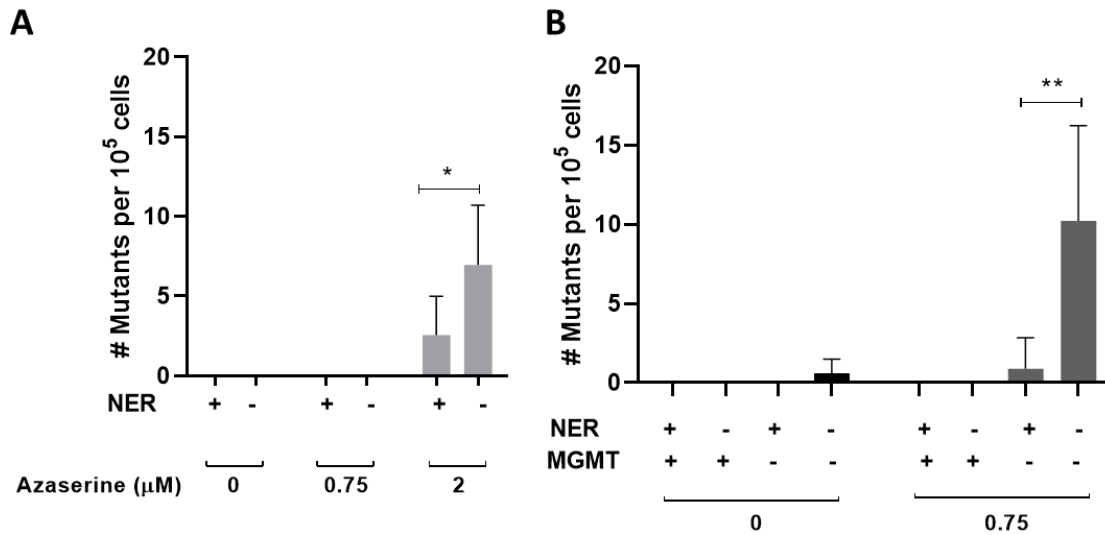


Figure 4. Mutagenicity in CHO WT (WT) and CHO UV20 KO (KO) cells exposed to azaserine and *O*⁶-BnG. CHO WT (NER +) and CHO UV20 KO (NER -) cells were exposed to two doses of azaserine (0, 0.75, 2 μM) (A). Azaserine-treated (0.75 μM; grey bars) and untreated (0 μM; black bar) CHO WT (NER +) and CHO UV20 KO (NER -) cells were exposed to 20 μM *O*⁶-BnG (MGMT -) or 0 μM *O*⁶-BnG (MGMT +) (B). HPRT assay was performed as described and mutant colonies were counted manually after two weeks. Data were represented using GraphPad Prism 8. Statistical significance was calculated by performing two-way ANOVA with a Tukey post hoc test (**p*<0.05, ***p*<0.01, ****p*<0.001, *****p*<0.0001). Significance is shown among selected bars.

4.3.9 Sequencing data for *hprt* gene in NER CHO cells

To characterize mutations induced by azaserine when NER is deficient, we are sequencing the *hprt* gene of CHO WT and UV20 cells exposed to azaserine under the same conditions of the HPRT assay. Results from sequencing analysis of the *hprt* gene from CHO cells are underway. Anticipated outcomes are that the level of mutations in the *hprt* gene will parallel findings of the HPRT assay. Additionally, we will analyse the mutational spectrum of NER-deficient cells and compare with that of proficient cells. We anticipate gaining important information regarding the impact of NER on alleviating azaserine-induced mutations.

Finally, to validate and characterize the combined role of MGMT and NER, we will sequence also mutations arising upon exposure to azaserine and to *O*⁶-BnG in NER-deficient CHO cells. We will compare the outcome of sequencing to results obtained from HPRT assay and contradictory findings of cell survival.

4.3.10 O^6 -CMG adduct levels induced by azaserine in NER-deficient CHO cells

To test the hypothesis that the increase in mutagenicity observed in NER-deficient cells (Fig. 5) is due to the persistence of O^6 -CMG adducts, adduct levels in azaserine-exposed cells as a function of O^6 -BnG exposure will be measured by mass spectrometry and experiments are underway. CHO cells were treated with azaserine (0, 2 μ M). Untreated samples were prepared as controls and to define the background level of O^6 -CMdG and O^6 -MeG. Anticipated results are that the level of adducts induced by azaserine will parallel the level of mutations obtained from the HPRT assay (Fig. 4).

Data collected from cell survival and mutagenicity in this work so far hint at NER and particularly at MGMT as repair pathways responsible to remove O^6 -CMG and protect against its toxic and mutagenic properties. However, more direct evidence of MGMT and NER role are needed. Future experiments to study the role of MGMT and NER are suggested in the discussion section. Preliminary assays performed so far are presented in the following paragraphs.

4.3.11 *In vitro* NER biochemical assays on O^6 -CMG DNA are contradictory regarding the role of NER in O^6 -CMG recognition and repair

To test more directly the role of NER on O^6 -CMG repair, and assess whether the NER machinery has the potential to recognize and excise O^6 -CMG from a DNA strand, we performed biochemical assays with NER proteins and DNA containing O^6 -CMG.

To test the binding to O^6 -CMG DNA of the purified XPC-RAD23B protein (damage recognition factor of a NER sub-pathway) we performed EMSA experiments. We synthesized a 24mer oligonucleotide containing O^6 -CMG and annealed it to a fluorescently labelled complementary strand (Fig. S7). As negative and positive controls for binding, we prepared respectively undamaged (G) dsDNA, and dsDNA containing a well-established NER target, namely acetylaminofluorene (AAF).⁵⁰ We incubated the different samples of dsDNA with increasing concentrations of XPC-RAD23, and we assessed the binding of the protein to the dsDNA by resolving samples on a denaturing gel and imaging for the fluorescent DNA (Fig. S7). When dsDNA was bound to XPC-RAD23B, a shift of signal towards the top of the gel is recorded (as DNA bound to a protein runs slower through a gel). We found that dsDNA containing O^6 -CMG bound to XPC-RAD23B at lower concentrations of the protein compared to the undamaged DNA, and at slightly lower concentrations when compared to the positive control AAF (Fig. 5). However, because XPC-RAD23B binds strongly to DNA even when not containing NER targets, the difference in signal shift is mild. For this reason, when performing EMSA analysis with XPC-RAD23B, unlabelled DNA is normally introduced as competitor for binding.⁵⁰ By engaging with XPC-RAD23B and subtracting it from binding to other unmodified DNA, the presence of a

competitor makes the difference between XPC-RAD23B binding to non-NER targets and to potential NER targets more easily detectable. In future studies, we propose to include a competitor when performing EMSA experiments.

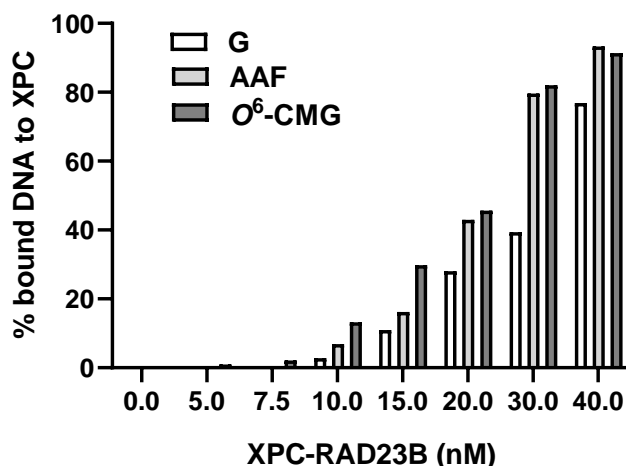


Figure 5. XPC-RAD23B binding to damaged and undamaged DNA. XPC-RAD23B at increasing concentrations (0-40 nM) was incubated with Cy5-labeled DNA: undamaged DNA (G, negative control) or DNA containing AAF adduct (positive control) or *O*⁶-CMG (here, oligo-*O*⁶-CMG-12 was used). Reactions were resolved on a denaturing gel and the shift of Cy5-labeled DNA signal (upper bands in Fig. S7 C) from untreated samples (0 nM XPC-RAD23B) was an indication of binding of the dsDNA to XPC-RAD23B. Band intensity was quantified on Image Lab software. The relative fraction of unbound DNA (lower bands, Fig. S7 C) was normalized as follows: $100 \times [\text{unbound DNA}] / ([\text{unbound DNA}] + [\text{bound DNA}])$ and plotted on GraphPad Prism 8.

To test the capability of the NER machinery to recognize, bind and excise a region of DNA containing *O*⁶-CMG, we performed a cell-free NER assay, by using purified NER proteins or cellular extracts, incubated with DNA containing *O*⁶-CMG, and assessing the excision of a short oligonucleotide containing *O*⁶-CMG via [α -³²P] DNA end-labelling. We produced circular dsDNA containing *O*⁶-CMG by extending 24mer *O*⁶-CMG oligonucleotides (oligo-*O*⁶-CMG-10 and oligo-*O*⁶-CMG-12) after annealing to circular ssDNA, which was produced in XL1 blue *E. Coli* cells infected with helper phage R408 (Fig. S8). We used dsDNA containing AAF-G adducts as positive control for NER excision. Samples containing either AAF-G or *O*⁶-CMG were incubated with HeLa extracts, or with NER proteins +/- XPA, which is pivotal for the proceeding of repair after damage recognition. We found that under none of the tested conditions DNA containing *O*⁶-CMG was excised. Instead, the positive control AAF-G was excised in all cases, except when XPA was not present (Fig. 6).

Despite binding of XPC-RAD23B to *O*⁶-CMG-DNA is not *per se* an indication of NER repair happening in the cell, and further testing is required, the preliminary data obtained from the

NER assay disagree with both EMSA and cell-based studies, and suggest no involvement of the NER proteins in removing O^6 -CMG. In future experiments, the NER assay will be repeated, and factors such as MGMT or others proteins that might be crucial for NER activation or recognition will be included in the mix of purified proteins. In the case of employing the HeLa extracts, Western Blot of MGMT and other relevant factors should be performed to confirm their presence (the level of mRNA for MGMT in HeLa cells is on the average;⁵¹ however, the actual amount of translated MGMT is not known). While being a powerful and robust tool to test the excision of damaged DNA by NER, the NER assay is based on a cell-free system and is restricted to the factors and conditions tested. Thus, more direct evidence of NER role in cells should be collected.

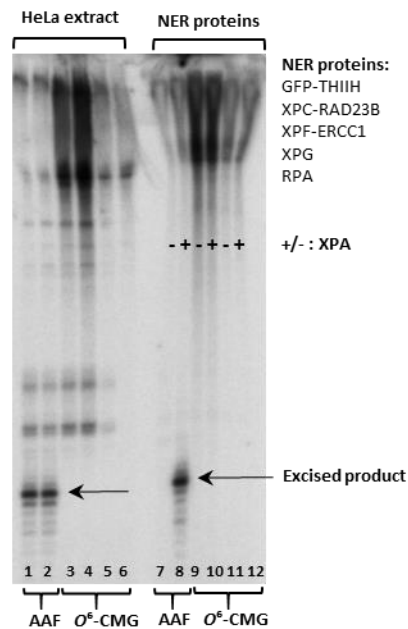


Figure 6. NER assay on O^6 -CMG plasmid. Samples 1-2, 7-8 consisted of G-AAF; samples 3-4, 9-10 of oligo- O^6 -CMG-10 and samples 5-6, 11-12 of oligo- O^6 -CMG-12. Samples 1-6 were incubated with HeLa extract for 45 min at 30°C; samples 7-12 were incubated with NER proteins (except XPA for samples 7, 9 and 11) for 90 min at 30 °C. After incubation with either HeLa extract or NER proteins, all samples were detected by annealing to a complementary oligonucleotide with a 5'-GpGpGpG overhang, which served as a template for end-labelling with [α - 32 P]-dCTP with sequenase. The reaction products were resolved on a denaturing polyacrylamide gel and imaged using a Phosphorimager.

4.4 Discussion and Conclusion

Mechanisms of DNA damage repair are pivotal to ensure cellular stability, and their disruption has been associated with adverse biological outcomes, particularly cell death,⁵²⁻⁵⁴ rise of mutations to DNA, and cancer.⁵⁵⁻⁵⁷ Thus, the contribution to mutagenesis and carcinogenesis of DNA damage that can be replicated depends on the cellular capacity to repair that type of damage. However, the identity and the cellular impact of repair that targets some mutagenic DNA adducts is not known or well-characterized, hindering our understanding of how mutations arise and our ability to assess the increased risk associated with the exposure to certain chemicals.

O⁶-CMG is a mutagenic DNA adduct to which humans can be exposed upon red and processed meat consumption,^{7, 12-14} and that has been hypothesized to drive the development of colorectal cancer.¹⁵ However, what are the cellular pathways specialized at repairing O⁶-CMG, therefore preventing its accumulation and alleviating O⁶-CMG-induced mutagenicity, is a longstanding question that awaits a clear answer. Few studies have been carried out to address this question, and only one was performed in human cells. Furthermore, results were contradictory, making our current knowledge of O⁶-CMG repair elusive.

In the present study, we investigated how the phenotype of O⁶-CMG-inducing drug azaserine in cells is affected by NER and MGMT repair pathways, for which contradictory evidences have been obtained regarding O⁶-CMG repair.^{23, 32} We found that azaserine was significantly more toxic in mammal cells depleted of MGMT. We measured mutagenicity via a phenotypic assay and observed that azaserine was mutagenic in rodent cells in a dose-dependent manner. In particular, mutagenicity was low at sub-toxic concentrations but the number of induced mutants significantly increased when MGMT was inhibited. Similarly, to confirm and further characterize NER repair of O⁶-CMG, we employed cells with defined deficiencies in NER and measured cell viability and mutant occurrence upon azaserine exposure. Impaired NER capacity sensitized cells to both azaserine toxicity and mutagenicity, hinting at a role of NER in the removal of O⁶-CMG. However, preliminary results from cell-free assays show that NER proteins were not proficient at excising O⁶-CMG, although the NER lesion recognition factor XPC bound to O⁶-CMG more strongly than to undamaged DNA. We reasoned that some factors that are not NER but needed for O⁶-CMG repair might have been missing in the experimental set-up of the NER assay and that some more testing is required. Finally, when we measured mutagenicity in cells depleted of both MGMT and NER, we obtained an increased rate of mutants, significantly higher than when MGMT and NER were depleted individually, and higher than their sum. Although more concentrations of azaserine and O⁶-BnG need to be tested (experiments ongoing) to define whether the combined effect of MGMT and NER depletion on mutagenicity is simply additive or synergistic, we speculated that the increased mutagenicity might derive from a cooperative or cross-dependent relationship between MGMT and NER.

Previous work performed with yeast and bacterial alkyltransferase-like (ATL) proteins, which are homologous of MGMT but lack transferase activity,^{25, 34-36} showed how ATL proteins serve as signalling factors of O^6 -alkylG to the NER pathway, by binding to O^6 -alkylGs and making the adducts easily detectable by NER. Although ATLs have not been found in higher eukaryotes, studies on ATLs signalling function raises the questions of whether alkyltransferases such as MGMT might have a similar role of cross-communication between repair pathways in the presence of O^6 -CMG in higher organisms, in addition to its transferase activity.

In the presented work, regardless of whether repair pathways function independently or cross-talk, we have unveiled the important contribution of both NER and MGMT in alleviating O^6 -CMG detrimental consequences in cells. Our findings shed light on some controversial evidence reported in the literature. Furthermore, data presented herein are consistent with a study published this year suggesting that MGMT and NER are the major cellular pathways proficient at repairing O^6 -alkylGs.³³ O^6 -alkylGs with increasing alkyl chain length were systematically tested and interrogated for their replication-blocking and mutagenic properties in cells with defined protein deficiencies. It has been found that both NER and MGMT depletion resulted in an increase of G>A mutations, but to different extents depending on the alkyl group.³³ However, O^6 -CMG was not addressed in this study, and it is of emerging interest due to its association to processes of carcinogenesis.

To further characterize the phenotype of O^6 -CMG, adduct measurement and sequencing analysis are underway. The major open questions remaining regard the exact mechanism of NER and MGMT on targeting O^6 -CMG, and whether other factor or repair pathways, such as mismatch repair, might be involved too. To gain additional evidence on MGMT and NER repair of O^6 -CMG and on their mode of action, the binding and interaction of proteins around the O^6 -CMG adduct site will need to be assayed. Some of the remaining experiments and open questions will be addressed with research concerning:

- 1) Sensitivity to azaserine of human NER-deficient cells, e.g. immortalized human cells derived from patients affected by NER deficiency
- 2) Sequencing of CRC hotspots upon azaserine treatment, in combination with MGMT and NER depletion
- 3) MGMT inhibition testing by mass spectrometric measurement of cells exposed to well-established MGMT targets, such as MNNG and MNU methylating agents, and to O^6 -BnG (experiments underway, preliminary results on HCEC cells)
- 4) Sensitivity to azaserine of cells deficient in mismatch repair, e.g. HCT116+chr3 cells⁵⁸
- 5) EMSA assay as described in the relevant section, additionally using a DNA competitor, and +/- purified MGMT
- 6) NER assay as described in the relevant section using different cell lysates and with NER proteins +/- purified MGMT

7) Removal of O⁶-CMG placed on the restriction site in a plasmid by +/- MGMT +/- NER proteins

With the present work, despite pending experimentation, we have already unravelled a longstanding question in DNA repair concerning the identification of how cells remove O⁶-CMG. Furthermore, we demonstrated that inadequate removal of this adduct gives rise to mutations. These findings advance an understanding of the relationship between O⁶-CMG DNA damage and mutation, paving the way to the study of mutagenicity and cancer etiology. The outcome of this research is expected to be of interest in the field fundamental research in DNA repair, and for the risk assessment to the exposure to certain drugs or dietary products, especially for those individuals with impaired repair due to genetic conditions or drug treatment.

4.5 Material and Methods

Synthesis of O⁶-CMG phosphoramidite and oligonucleotide

Oligonucleotide sequences were (24mer) 5'- CTATTACCGXCYCCACATGTCAGC -3'; where X, Y = G in unmodified oligo-G; X = O⁶-CMG, Y = G in O⁶-CMG-modified oligo-O⁶-CMG-10; X = G, Y = O⁶-CMG in O⁶-CMG-modified oligo-O⁶-CMG-12. Unmodified oligo-G and its sequence-complementary oligonucleotide were purchased by IDT (Iowa, USA). O⁶-CMG-modified oligos were synthesized by solid-phase DNA synthesis on a Mermade 4 DNA synthesizer (Bioautomation Corporation; Texas, USA) in trityl-on mode on a 1- μ mol scale. Universal Q SynBase™ CPG support columns and natural nucleotide phosphoramidites from Link Technologies Ltd. (Lanarkshire, Scotland, UK) were used, and O⁶-CMG phosphoramidite was obtained as previously reported.¹⁹ Oligonucleotide deprotection was carried out by treatment with 0.4 M NaOH (200 μ L) at RT overnight, and with 33% aqueous ammonium hydroxide (500 μ L) at RT for 24 hours. Residual ammonium was let evaporate, trityl-cleavage and oligonucleotide desalting were performed on reverse-phase C18 Sep-Pak cartridges (Waters; Massachusetts, USA) following the manufacturer's protocol and using 4 % trifluoroacetic acid for trityl cleavage. The oligonucleotides were dried under reduced pressure and purified by reverse phase HPLC using an Agilent eclipse XDB C-18 5 μ M 4.6 x 150 mm column. The mobile phase consisted of 50 mM triethylammonium acetate (TEAA) and acetonitrile (ACN). The flow rate was set to 1.0 mL/min with a mobile phase starting at from 10-40% ACN over 25 min, increasing to 50% ACN from 25.1 to 29 min, and equilibrated from 29.1-33 min at 10% ACN. Fractions corresponding to the desired oligonucleotides were collected, lyophilized and re-suspended in DNase/RNase free water. The DNA concentrations were measured by UV absorbance on Nanodrop spectrophotometer (ThermoFisher; Massachusetts, USA).

Electrophoretic Mobility Shift Assay (EMSA)

EMSA assay was performed on oligo-*O*⁶-CMG-10 and oligo-*O*⁶-CMG-12, annealed with a Cy5-labeled complementary strands and incubated with NER protein XPC-RAD23B (0 - 40 nM) as previously reported.⁵⁹ Oligo-G and oligo-AAF-12 (5'-CTATTACCGXCYCCACATGTCAGC-3'; X= G, Y = AAF) were used respectively as negative and positive control for binding. AAF-modified oligo was synthesized as previously reported.⁵⁹ The reaction mixtures were loaded onto a native 5% polyacrylamide gel pre-equilibrated with 0.5× TBE buffer and run at 4 °C for 1 h at 20 mA. Gels were imaged using a Typhoon 9400 Imager (Amersham Biosciences; Little Chalfont, UK). Band intensity was quantified on Image Lab software (Bio-Rad; California, USA). The relative fraction of unbound DNA was normalized as follows: $100 * [\text{unbound DNA}] / ([\text{unbound DNA}] + [\text{bound DNA}])$ and plotted on GraphPad Prism 8.

***O*⁶-CMG-single stranded circular DNA production and NER assay**

Single stranded circular DNA was generated from p98 (with a *Nar*I site engineered into pBlusscript II SK+) as previously described.⁵⁰ The oligonucleotides oligo-*O*⁶-CMG-10 and oligo-*O*⁶-CMG-12 (100 pmol) were 5'-phosphorylated by incubation with 20 units of T4 PNK and 2 mM ATP at 37 °C for 2 h. After annealing with 30 pmol of single-stranded p98, the mixture was incubated with dNTPs (800 μM), T4 DNA polymerase (90 units), and T4 DNA ligase (100 units) to generate the covalently closed circular DNA containing a single *O*⁶-CMG adduct. The closed circular DNA was purified by cesium chloride/ethidium bromide density gradient centrifugation, followed by consecutive butanol extractions to remove the ethidium bromide, and finally concentrated on a Centricon YM-30 (Millipore, Massachusetts, USA). The in vitro NER assay was performed as previously described,^{50, 59} using either HeLa cell extract (2 μL of 21 mg/mL), or NER purified proteins. The reaction products were visualized using a Phosphorimager, Typhoon 9400 (Amersham Biosciences; Little Chalfont, UK). The relative repair efficiency of each substrate was normalized to a nonspecific band and quantified by the Image Quant TL program from Amersham Biosciences.

Cell strains and culture conditions

V79-4 cells were purchased from ATCC (Virginia, USA). HCEC-1CT cell lines were kindly provided by Prof J.W. Shay (University of Texas, USA).^{46, 47} CHO UV20 (ERCC1-deficient) and CHO UV20 transduced with a lentiviral vector expressing wild-type *hERCC1* were donated by Prof R. Kanaar (Erasmus Medical Center, Netherlands). HAP1 cells were purchased from Horizon Discovery (Waterbeach, UK). All cells were maintained as monolayers in 10-cm dishes in a humidified, 5% CO₂ atmosphere at 37 °C. All culture media except the one for HCEC cells were supplemented with 10% FBS and 1% penicillin-streptomycin and were: DMEM (V79-4, CHO cells), IMDM (HAP1 cells). Media for HCEC cells consisted in 80% DMEM and 20% M199 Earle, supplemented with

2% of Hyclone serum, 25 ng/L epidermal growth factor, 1 µg/L hydrocortisone, 10 µg/L insulin, 2 µg/L transferrin, 50 µg/L gentamycin, 0.9 ng/L sodium selenite. All materials for cell culture were purchased from Invitrogen (California, USA).

Cytotoxicity assay

Cells were seeded in 96-well plates (2 x 10³ per well for V79-4 cells, 3 x 10³ per well for CHO cells, 4 x 10³ per well for HCEC cells, 10 x 10³ per well for HAP1 cells) and treated with solutions of test compounds (compound and concentrations indicated in the relevant figures) with final concentration of 0.1% of DMSO. Cells were treated with TritonX (Sigma) for positive control, and with 0.1% DMSO for negative control. After 24 h-exposure, cells were washed and drug-free medium was added and further incubated for 72 h. Cell survival was measured using CellTiter-Glo assay (Promega; Wisconsin, USA) following the manufacturer's instructions.⁶⁰ Each data set was acquired in triplicate for each of three or more experiments and the averaged luminescence values of the drug-treated cells were normalized to those of DMSO-treated cells (assigned as 100% viability). Data were analysed and plotted using GraphPad Prism (GraphPad Software Inc.; California, USA).

Clonogenic Survival assay

Clonogenic assays were conducted as previously reported.⁶¹ Exponentially growing HAP1 cells were plated in 6 cm dishes in duplicates at a density of 100, 500 or 2.000 cells/plate. Cells were exposed to azaserine (0 – 5 µM) for 24 h and colonies let grow for one week, or cells were exposed to azaserine (0 – 10 µM) for one week. When colonies are visible to the naked eye, the cells were fixed with 4 % paraformaldehyde for 1 h, and stained with 50% methanol, 7% acetic acid and 0.1% Coomassie Brilliant Blue for 1 h. Colonies were counted manually. The data were normalized to the number of colonies on untreated plates.

Mutagenicity (HPRT) assay

HPRT assay was performed on V79-4 and CHO cells with a protocol adapted from elsewhere.⁶² Cells were seeded in 30-cm dishes (5 x 10⁵ cells per dish). Pre-existing *hprt*⁻mutants were cleansed by incubating cells in medium supplemented with hypoxanthine, aminopterin and thymidine (HAT). After three days, HAT medium was removed and replaced by normal DMEM medium and left for 24h. Cells were reseeded in 10-cm dishes (10⁵ cells per dish) and one 10-cm dish was prepared for each condition to test. Cells were left to attach overnight. Cells were pre-incubated for 2 h with O⁶-BnG (0 – 60 µM for V79-4, 0 – 20 µM for CHO) and then added with azaserine (0 – 2 µM for V79-4, 0 – 1.25 µM for CHO) and incubated for additional 24 h. DMSO final concentration was 0.06%. Cells for positive control were treated with 1.7 or 3.5 mM ethyl methane sulfonate (EMS). After treatment, culture medium was changed with fresh drug-free

DMEM medium. Cells were let grow for 7 days, and medium was changed every three days. Per each condition, cells were reseeded in four 10-cm dishes (2×10^5 cells per dish) in selective medium containing 10 $\mu\text{g}/\text{ml}$ of 6-thioguanine (6-TG), and in two 6-cm dishes (300 cells per dish) in normal DMEM medium for plating efficiency (PE) calculations. All cells were incubated for 14 days, and medium was changed every 4 days (for the four selection dishes, fresh medium contained always 6-TG). After 14-days of *hprt*⁻ mutants selection, cells were stained with 50 % Giemsa staining solution (v/v) in methanol. Clones were counted manually and normalized to the PE in the case of CHO WT and KO, for which PE was the most different. Data were analyzed and plotted on GraphPad Prism (GraphPad Software Inc.; California, USA). Statistical significances of differences relative to control were calculated using two-way ANOVA with a Tukey post hoc test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$).

Sequencing of *hprt* and CRC-relevant genes

Cells were exposed for 24 h to azaserine (0 or 2 μM for V79-4 and CHO cells) and to *O*⁶-BnG (0 or 60 μM for V79-4; 0 or 20 μM for CHO), and incubated for one week, to mimic the expression period length of the HPRT assay. DNA was extracted using QIAamp DNA Mini Kit (Qiagen; Germany) and following manufacturer's instructions.⁶³ *hprt* gene was amplified by PCR using previously reported primers (forward 5'-TGACACTGGCAAAACAATGCA-3' and reverse 5'-GGTCCTTTTCACCAGCAAGCT-3').⁶⁴ PCR reactions consisted of 10 ng of genomic DNA, 0.5 μM of forward primer and 0.5 μM of reverse primers, 1x Q5 MasterMix already including Q5 Polymerase (Thermofisher; Massachusetts, USA). Reaction conditions were initial melting step at 98 °C for 30 s, 25 cycles of melting at 98°C for 10 s and annealing at 54 °C for 30s and extension at 72 °C for 30s, and final extension a at 72 °C for 5 min. Amplified DNA was purified using Monarch PCR & DNA Cleanup Kit (NEB, UK) and sent for Sanger sequencing to Microsynth company (Switzerland).

For future studies, *k-ras*, *p53* and *apc* genes will be sequenced from HCEC cells exposed to azaserine, using primers designed and reported elsewhere. All primers were checked for gene recognition and specificity using blast.ncbi.nlm.nih.gov online platform. Forward and reverse primers selected for sequencing of HCEC DNA are respectively:

5'-ACTGAATATAAACTTGTGGTCCATGGAGCT-3'; 5'-TTATCTGTATCAAAGAATGGTCCTGCACCA-3' (*k-ras*, codon 12 and 13);⁶⁵ 5'-GAGACGACAGGGCTGGTT-3'; 5'- CCACTGACAACCACCTT-3' (*p53*);⁶⁶ 5'-GCAGATTCTGCTAATACCCTGCAAATAGCATTAA-3'; 5'-CCTGAAGAAAATTCAACAGCTTTGTGCCTG-3' (*apc*).⁶⁷

Quantitative PCR (qPCR) of MGMT

HCEC cells were treated with azaserine (0 - 1825 μM) for 4 r or 24 h. Cell pellets were collected by centrifugation at 2000 x g for 5 min. RNA was extracted using RNeasy MinElute Cleanup Kit

(Qiagen; Germany) and following manufacturer's instructions.⁶⁸ Extracted RNA was quantified by UV absorbance on Nanodrop spectrophotometer (ThermoFisher; Massachusetts, USA). Aliquots of 500 ng of RNA were retrotranscribed using High-Capacity cDNA Reverse Transcription Kit (ThermoFisher; Massachusetts, USA) and following manufacturer's instructions.⁶⁹ MGMT portions were amplified and quantified by qPCR from the retrotranscribed cDNA. Primers to amplify MGMT were selected from commercially available ones (forward primer 5'-CACCGTTTGC GACTTGGTA-3'; reverse primer 5'-CTCCAGCTTCCCCAAAGG-3'),⁷⁰ or designed (forward 5'-CCTGGCTGAATGCCTATTTCCAC-3'; reverse 5'-GCAGCTTCATAACACCTGTCTG-3') on esemble.org, checked for gene recognition on Primer Blast online software (NIH) and for secondary structures on Oligo Calculation online software (ThermoFisher; Massachusetts, USA), and purchased from Eurogentec (Belgium). Primers to amplify negative control β -actin (forward 5'-TCCTCCTGGGCATGGAGT-3'; reverse 5'-AGCACTGTGTTGGCGTACAG-3') were designed as above and were purchased from Eurogentec (Belgium). qPCR reactions consisted of 12 ng of cDNA, 1 μ M of forward primer and 1 μ M of reverse primers, 1x PCR MasterMix including Polymerase and SYBR green dye (ThermoFisher; Massachusetts, USA). qPCR conditions consisted of an initial melting step at 95 °C for 5 min and of 45 cycles of melting at 95°C for 30 s and annealing and extension at 60 °C for 1 min, each cycle followed by fluorescence measurement of SYBR green. Melting temperatures were measured at the end of each qPCR experiments to ensure the presence of one kind of amplicons in each sample. Threshold cycle (Ct) values were derived by setting an arbitrary threshold for the fluorescence signal of 0.2 for all samples. Ct values indicating the expression level of MGMT gene were normalized to the GAPDH level and plotted on GraphPad Prism (GraphPad Software Inc.; California, USA).

Mass Spectrometric quantitation of O⁶-MedG and O⁶-CMdG

O⁶-MedG and O⁶-CMdG adduct level are measured in cells treated for 24 h at the same concentrations of azaserine (0 or 2 μ M for both V79-4 and CHO cells) and O⁶-BnG (0 or 60 μ M for V79-4; 0 or 20 μ M for CHO) employed for the HPRT assay. To check for MGMT inhibition, O⁶-MedG adduct level is measured in cells treated with MNNG (30 μ M for V79-4 and CHO; 10 μ M for HCEC) +/- O⁶-BnG to check for MGMT inhibition. DNA is extracted using QIAamp DNA Mini Kit (Qiagen; Germany) and following manufacturer's instructions.⁶³ O⁶-MedG and O⁶-CMdG are enzymatically hydrolyzed following a previously reported one-step protocol.⁷¹ Samples are hydrolyzed in the presence of isotopically labelled internal standards, D₃-O⁶-MedG and ¹⁵N₅-O⁶-CMdG (100 fmol and 400 fmol respectively). Hydrolyzed nucleosides are purified by solid-phase extraction on a SepPak C18 cartridge (Waters Co., Milford, MA) following the manufacturer's instructions, and dried. Dried samples are reconstituted in 10 μ l of MS grade water and 2 μ l analyzed by an Acquity UPLC M-Class (Waters Co., Milford, MA) equipped with an 5 μ l injection

loop and coupled to an Orbitrap Fusion Lumos (Thermo Scientific, Waltham, MA). PRM scanning in positive ionization mode is performed using nano-electrospray as described previously.⁴⁹ The following transitions are monitored for O^6 -MedG, D_3 - O^6 -MedG and O^6 -CMdG, $^{15}N_5$ - O^6 -CMdG respectively: m/z 282.1195 to 166.0722, m/z 285.1384 to 169.0911 and m/z 326.1096 to 210.0622, m/z 331.0946 to 215.0472. MS instrument control and data acquisition are performed using Xcalibur (version 4.0, ThermoFisher Scientific).

Acknowledgments

We acknowledge and thank Xuan Li for preliminary testing of HPRT assay, Dr. Hyun Suk Kim for introducing to and helping with EMSA experiments, Jiyoung Park for support with the cell culture, Dr. Gabriele Fontana for introducing to and patiently advising with qPCR.

4.6 Supplementary Figures

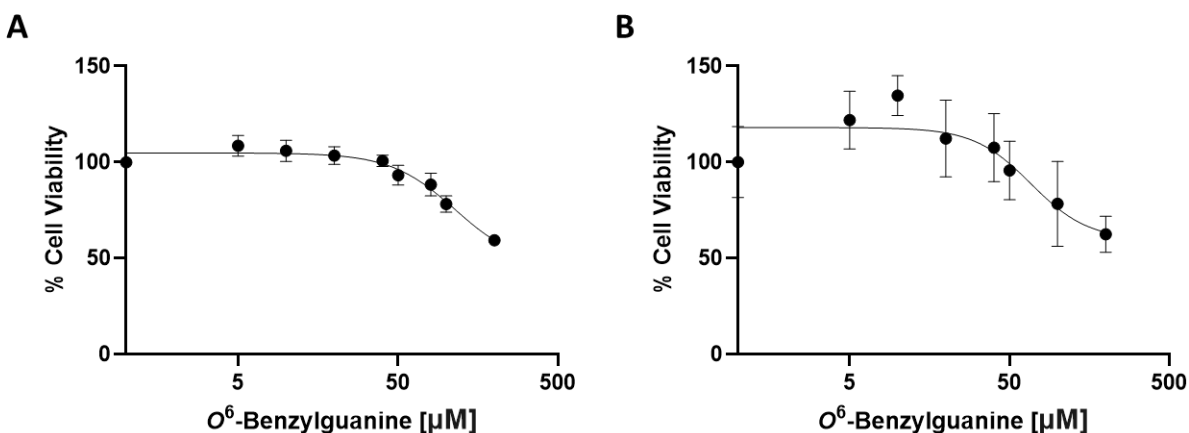


Figure S1. O^6 -BnG toxicity in V79-4 and HCEC cells. V79-4 (A) and HCEC cells (B) were exposed for 24 h to increasing concentrations of O^6 -BnG. Cell viability was measured by CellTiter-Glo. Data were represented using GraphPad Prism 8. Cell viability data were fit in a nonlinear regression curve, using a dose-response function. Error bars indicate standard deviation. Experiments were repeated twice, each time in technical triplicates.

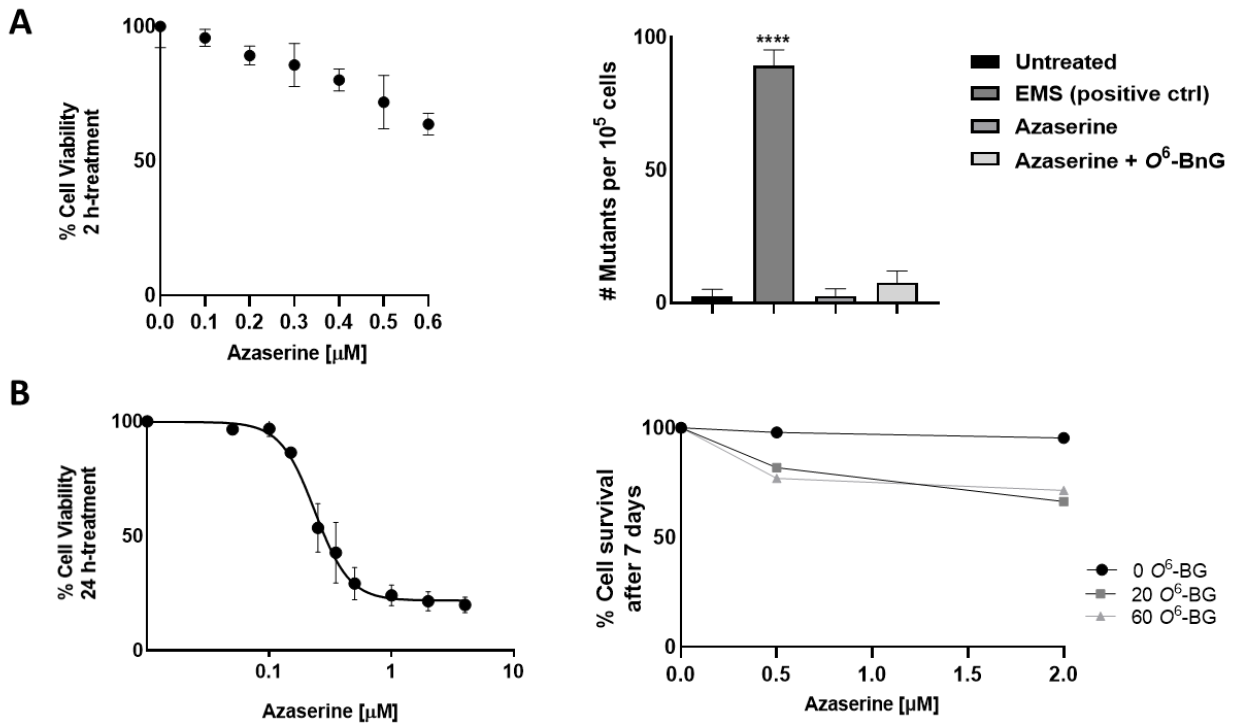


Figure S2. Preliminary measurements of viability and mutagenicity in V79-4 cells. (A) Cell viability upon 2 h-exposure to increasing concentrations (0-0.6 μ M) azaserine was measured by CellTiter-Glo assay (left). Preliminary measurements of HPRT assay (right) were performed on cells exposed to azaserine for 2 h at the sub-toxic concentration of 0.4 μ M. Colonies were counted and normalized to PE. Data were analyzed on GraphPad Prism 8 and one-way ANOVA test was performed (**** p <0.0001). (B) Viability of V79-4 cells was measured upon 24 h-treatment (left). Cells exposed for 24 h to increasing concentrations of azaserine and O^6 -BnG were let 7 days in normal media and counted (right) before reseeding in selection media.

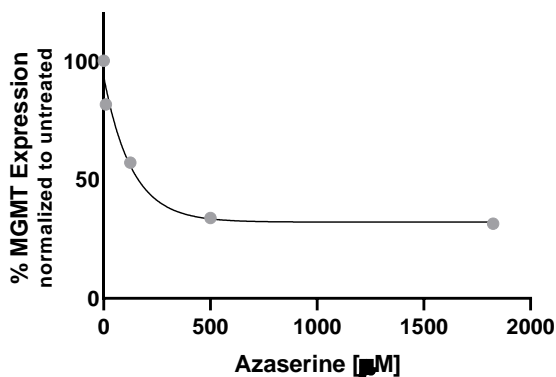


Figure S3. Relative expression level of MGMT in HCEC cells upon exposure to azaserine. HCEC cells were exposed to increasing concentrations of azaserine (0-2 mM), RNA was extracted and

retro-transcribed in cDNA. The cDNA was amplified in triplicates by qPCR, using DNA primers for MGMT gene. From the reactions of qPCR, values of Ct were obtained by setting an arbitrary threshold of 0.2. Δ Cts were calculated by subtracting the values of Ct obtained under the same condition from the amplification of a housekeeping gene (normalizer). Values of $\Delta\Delta$ Ct were calculated by subtracting the Δ Ct of the untreated sample. From the $\Delta\Delta$ Cts, the relative expression levels of MGMT were obtained via the formula: relative expression level = $100 * 2^{-\Delta\Delta Ct}$. Data was fitted in a one-phase decay exponential curve using GraphPad Prism 8.

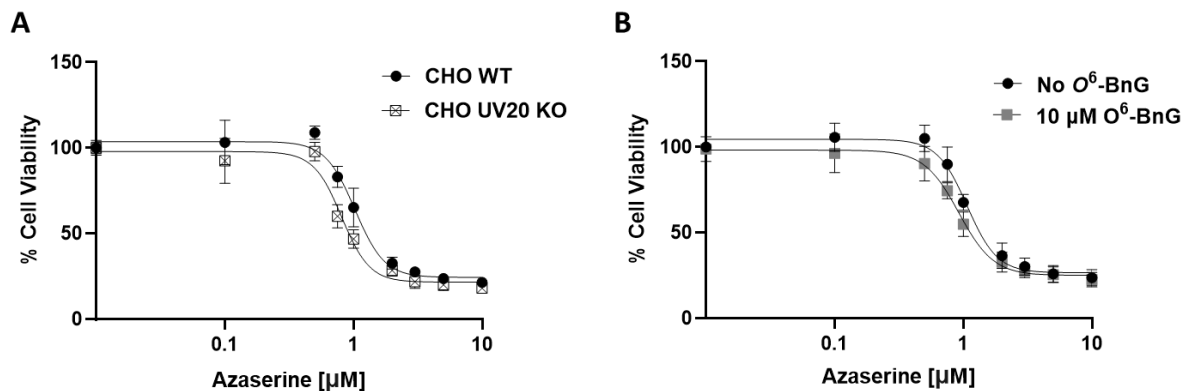


Figure S4. Viability of CHO cells after 24 h exposure to azaserine. CHO WT and UV20 KO cells were exposed to increasing concentrations of azaserine (A) CHO WT cells are exposed to increasing concentrations of azaserine or co-treated with 10 μ M O^6 -BnG (B) Viability was measured directly after 24 h-exposure using CellTiter-Glo assay. Data from at least three independent experiments, each performed in triplicates, were averaged and fit in a nonlinear regression curve, using a dose-response function on GraphPad Prism 8. Error bars indicate standard deviation.

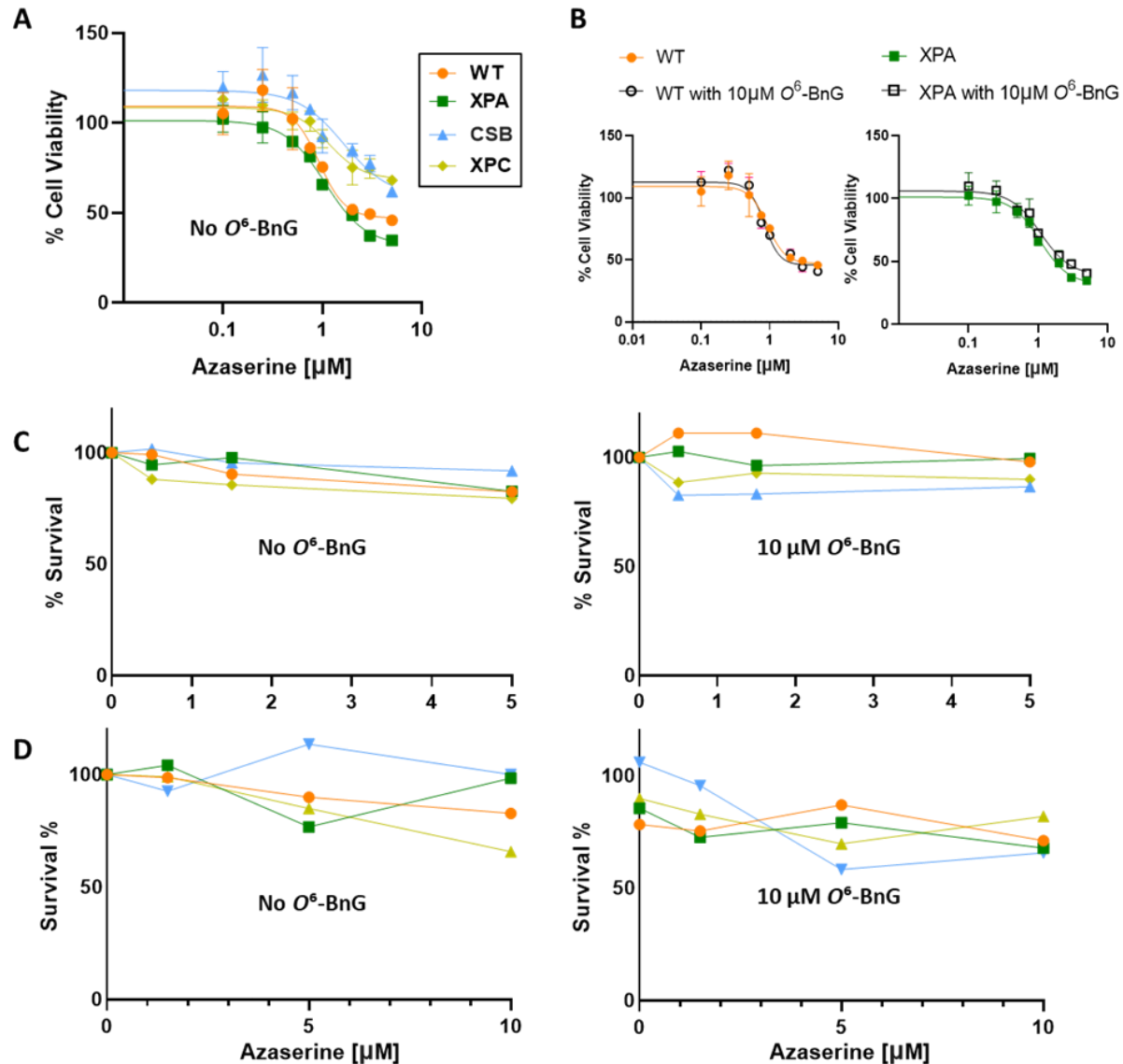


Figure S5. Cell viability and Clonogenic assay on HAP1 cells upon treatment with azaserine. (A) Viability was measured on HAP1 cells (WT and deficient in NER factors XPA, XPC, CSB) exposed for 24 h to increasing concentrations of azaserine, using CellTiter-Glo assay. (B) All HAP1 cells were also co-treated with 10 μM O^6 -BnG and assayed for viability using CellTiter-Glo assay. As examples, WT HAP1 and XPA cells are shown. Data were represented using GraphPad Prism 8. Cell viability data were fit in a nonlinear regression curve, using a dose-response function. Error bars indicate standard deviation. (C-D) Clonogenic assay was performed on HAP1 cells (WT and deficient in NER factors XPA, XPC, CSB) exposed to increasing concentrations of azaserine (left) and co-treated with 10 μM O^6 -BnG (right). Cells were exposed to drugs for 24 h and incubated for one week (C) or exposed to drugs for one week (D). Colonies were fixed, stained and counted. Each experiment was performed in duplicate. Error bars are not shown for clarity.

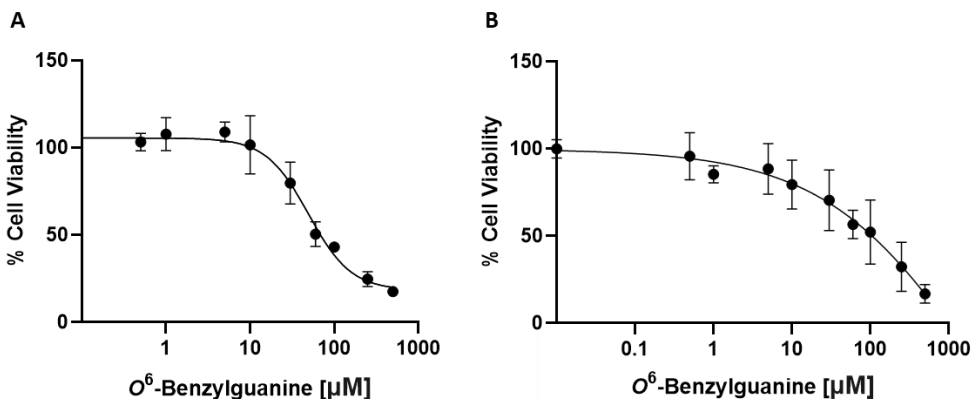


Figure S6. O^6 -BnG toxicity in HAP1 and CHO cells. (A) HAP1 and (B) CHO WT were exposed for 24 h to increasing concentrations of O^6 -BnG. Cell viability was measured by CellTiter-Glo assay. Data were represented using GraphPad Prism 8. Cell viability data were fit in a nonlinear regression curve, using a dose-response function. Error bars indicate standard deviation. Experiments were repeated twice, each time in technical triplicates.

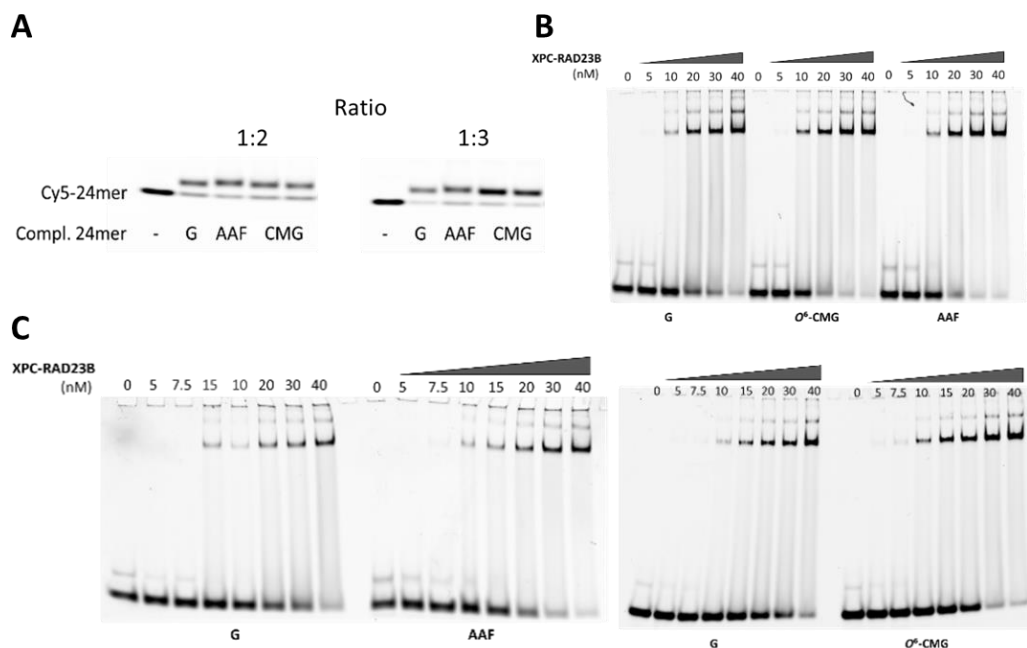


Figure S7. Annealing of oligonucleotides and EMSA for binding of XPC-RAD23B to O^6 -CMG. (A) To prepare dsDNA for EMSA, undamaged (G) and damaged (oligo-AAF-12, oligo- O^6 -CMG-10 and oligo- O^6 -CMG-12) 24mer oligonucleotides were annealed to a complementary Cy5-labeled 24mer strand, in a ratio 1:2 or 1:3. Homogenous annealing through the four samples was evaluated by imaging of samples run on a denaturing gel. (B-C) Undamaged (G) and damaged (oligo-AAF-12 and oligo- O^6 -CMG-12) 24mer oligonucleotides annealed to Cy5-labelled 24mer

were incubated with increasing concentrations of XPC-RAD23B protein. Reactions were resolved on a denaturing gel, and binding of the protein to the dsDNA was evaluated by quantification of the intensity of the upper bands using ImageLab software. Data obtained from gels in (C) are plotted in Fig. 4.

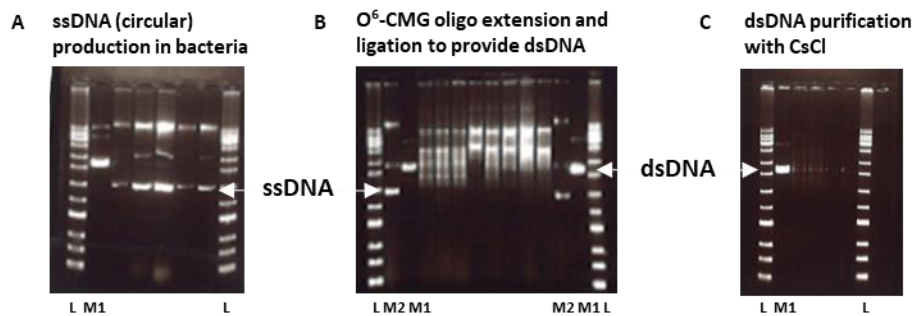


Figure S8 Production of dsDNA containing *O*⁶-CMG. (A) ssDNA was produced from XL1 blue cells infected with helper phage R408 in the presence of plasmid p98. (B) Oligo-*O*⁶-CMG-12 was annealed to the ssDNA, extended and ligated, resulting in a dsDNA (plasmid) containing *O*⁶-CMG. (C) The *O*⁶-CMG-dsDNA plasmid was purified by cesium chloride (CsCl) density gradient centrifugation. After each step, DNA samples were run on a 1.5 % agarose gel, stained with 6x gel loading dye purple and imaged by fluorescence (L = ladder lane; M1 = marker lane containing dsDNA (p98); M2 = marker lane containing ssDNA).

4.7 References

1. Giglia-Mari, G.; Zotter, A.; Vermeulen, W., DNA damage response. *Cold Spring Harb Perspect Biol* **2011**, *3* (1), a000745.
2. Yarosh, D. B., The role of *O*⁶-methylguanine-DNA methyltransferase in cell survival, mutagenesis and carcinogenesis. *Mutat Res* **1985**, *145* (1-2), 1-16.
3. Margison, G. P.; Santibanez Koref, M. F.; Povey, A. C., Mechanisms of carcinogenicity/chemotherapy by *O*⁶-methylguanine. *Mutagenesis* **2002**, *17* (6), 483-7.
4. Harrison, K. L.; Jukes, R.; Cooper, D. P.; Shuker, D. E., Detection of concomitant formation of *O*⁶-carboxymethyl- and *O*⁶-methyl-2'-deoxyguanosine in DNA exposed to nitrosated glycine derivatives using a combined immunoaffinity/HPLC method. *Chemical research in toxicology* **1999**, *12* (1), 106-11.
5. Roos, W. P.; Batista, L. F.; Naumann, S. C.; Wick, W.; Weller, M.; Menck, C. F.; Kaina, B., Apoptosis in malignant glioma cells triggered by the temozolomide-induced DNA lesion *O*⁶-methylguanine. *Oncogene* **2007**, *26* (2), 186-97.
6. Bennett, R. A.; Pegg, A. E., Alkylation of DNA in rat tissues following administration of streptozotocin. *Cancer Res* **1981**, *41* (7), 2786-90.
7. Lewin, M. H.; Bailey, N.; Bandaletova, T.; Bowman, R.; Cross, A. J.; Pollock, J.; Shuker, D. E.; Bingham, S. A., Red meat enhances the colonic formation of the DNA adduct *O*⁶-carboxymethyl guanine: implications for colorectal cancer risk. *Cancer research* **2006**, *66* (3), 1859-65.

8. Mirvish, S. S., Role of N-nitroso compounds (NOC) and N-nitrosation in etiology of gastric, esophageal, nasopharyngeal and bladder cancer and contribution to cancer of known exposures to NOC. *Cancer Lett* **1995**, *93* (1), 17-48.
9. Leaf, C. D.; Wishnok, J. S.; Tannenbaum, S. R., Mechanisms of endogenous nitrosation. *Cancer Surv* **1989**, *8* (2), 323-34.
10. Feelisch, M.; Rassaf, T.; Mnaimneh, S.; Singh, N.; Bryan, N. S.; Jourdain, D.; Kelm, M., Concomitant S-, N-, and heme-nitros(yl)ation in biological tissues and fluids: implications for the fate of NO in vivo. *FASEB J* **2002**, *16* (13), 1775-85.
11. Calmels, S.; Ohshima, H.; Henry, Y.; Bartsch, H., Characterization of bacterial cytochrome cd(1)-nitrite reductase as one enzyme responsible for catalysis of nitrosation of secondary amines. *Carcinogenesis* **1996**, *17* (3), 533-6.
12. Vanden Bussche, J.; Hemeryck, L. Y.; Van Hecke, T.; Kuhnle, G. G.; Pasmans, F.; Moore, S. A.; Van de Wiele, T.; De Smet, S.; Vanhaecke, L., O⁶-carboxymethylguanine DNA adduct formation and lipid peroxidation upon in vitro gastrointestinal digestion of haem-rich meat. *Mol Nutr Food Res* **2014**, *58* (9), 1883-96.
13. group, E. s., O⁶-methylguanine in blood leucocyte DNA: an association with the geographic prevalence of gastric cancer and with low levels of serum pepsinogen A, a marker of severe chronic atrophic gastritis. *Carcinogenesis* **1994**, *15* (9), 1815-20.
14. Hall, C. N.; Badawi, A. F.; O'Connor, P. J.; Saffhill, R., The detection of alkylation damage in the DNA of human gastrointestinal tissues. *British journal of cancer* **1991**, *64* (1), 59-63.
15. Cross, A. J.; Ferrucci, L. M.; Risch, A.; Graubard, B. I.; Ward, M. H.; Park, Y.; Hollenbeck, A. R.; Schatzkin, A.; Sinha, R., A large prospective study of meat consumption and colorectal cancer risk: an investigation of potential mechanisms underlying this association. *Cancer Res* **2010**, *70* (6), 2406-14.
16. Gottschalg, E.; Scott, G. B.; Burns, P. A.; Shuker, D. E., Potassium diazoacetate-induced p53 mutations in vitro in relation to formation of O⁶-carboxymethyl- and O⁶-methyl-2'-deoxyguanosine DNA adducts: relevance for gastrointestinal cancer. *Carcinogenesis* **2007**, *28* (2), 356-62.
17. Wu, J.; Wang, P.; Li, L.; Williams, N. L.; Ji, D.; Zahurancik, W. J.; You, C.; Wang, J.; Suo, Z.; Wang, Y., Replication studies of carboxymethylated DNA lesions in human cells. *Nucleic Acids Res* **2017**, *45* (12), 7276-7284.
18. Raz, M. H.; Dexter, H. R.; Millington, C. L.; van Loon, B.; Williams, D. M.; Sturla, S. J., Bypass of Mutagenic O⁶-Carboxymethylguanine DNA Adducts by Human Y- and B-Family Polymerases. *Chem Res Toxicol* **2016**, *29* (9), 1493-503.
19. Raz, M. H.; Sandell, E. S.; Patil, K. M.; Gillingham, D. G.; Sturla, S. J., High Sensitivity of Human Translesion DNA Synthesis Polymerase kappa to Variation in O⁶-Carboxymethylguanine Structures. *ACS Chem Biol* **2019**, *14* (2), 214-222.
20. Wang, P.; Leng, J.; Wang, Y., DNA replication studies of N-nitroso compound-induced O⁶-alkyl-2'-deoxyguanosine lesions in Escherichia coli. *J Biol Chem* **2019**, *294* (11), 3899-3908.
21. Hermisson, M.; Klumpp, A.; Wick, W.; Wischhusen, J.; Nagel, G.; Roos, W.; Kaina, B.; Weller, M., O⁶-methylguanine DNA methyltransferase and p53 status predict temozolomide sensitivity in human malignant glioma cells. *Journal of neurochemistry* **2006**, *96* (3), 766-76.
22. Martin, L.; Marples, B.; Coffey, M.; Lawler, M.; Hollywood, D.; Maignol, L., Recognition of O⁶MeG lesions by MGMT and mismatch repair proficiency may be a prerequisite for low-dose radiation hypersensitivity. *Radiation research* **2009**, *172* (4), 405-13.
23. O'Driscoll, M.; Macpherson, P.; Xu, Y. Z.; Karran, P., The cytotoxicity of DNA carboxymethylation and methylation by the model carboxymethylating agent azaserine in human cells. *Carcinogenesis* **1999**, *20* (9), 1855-62.

24. Aramini, J. M.; Tubbs, J. L.; Kanugula, S.; Rossi, P.; Ertekin, A.; Maglaqui, M.; Hamilton, K.; Ciccocanti, C. T.; Jiang, M.; Xiao, R.; Soong, T. T.; Rost, B.; Acton, T. B.; Everett, J. K.; Pegg, A. E.; Tainer, J. A.; Montelione, G. T., Structural basis of O⁶-alkylguanine recognition by a bacterial alkyltransferase-like DNA repair protein. *J Biol Chem* **2010**, *285* (18), 13736-41.
25. Wilkinson, O. J.; Latypov, V.; Tubbs, J. L.; Millington, C. L.; Morita, R.; Blackburn, H.; Marriott, A.; McGown, G.; Thorncroft, M.; Watson, A. J.; Connolly, B. A.; Grasby, J. A.; Masui, R.; Hunter, C. A.; Tainer, J. A.; Margison, G. P.; Williams, D. M., Alkyltransferase-like protein (At1) distinguishes alkylated guanines for DNA repair using cation-pi interactions. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109* (46), 18755-60.
26. Kubitschek, H. E.; Sepanski, R. J., Azaserine: survival and mutation in Escherichia coli. *Mutation research* **1982**, *94* (1), 31-8.
27. Shuker, D. E.; Margison, G. P., Nitrosated glycine derivatives as a potential source of O⁶-methylguanine in DNA. *Cancer research* **1997**, *57* (3), 366-9.
28. Kawate, H.; Ihara, K.; Kohda, K.; Sakumi, K.; Sekiguchi, M., Mouse methyltransferase for repair of O⁶-methylguanine and O⁴-methylthymine in DNA. *Carcinogenesis* **1995**, *16* (7), 1595-602.
29. Shiraishi, A.; Sakumi, K.; Sekiguchi, M., Increased susceptibility to chemotherapeutic alkylating agents of mice deficient in DNA repair methyltransferase. *Carcinogenesis* **2000**, *21* (10), 1879-83.
30. van Niftherik, K. A.; van den Berg, J.; van der Meide, W. F.; Ameziane, N.; Wedekind, L. E.; Steenbergen, R. D.; Leenstra, S.; Lafleur, M. V.; Slotman, B. J.; Stalpers, L. J.; Sminia, P., Absence of the MGMT protein as well as methylation of the MGMT promoter predict the sensitivity for temozolomide. *British journal of cancer* **2010**, *103* (1), 29-35.
31. Bobola, M. S.; Alnoor, M.; Chen, J. Y.; Kolstoe, D. D.; Silbergeld, D. L.; Rostomily, R. C.; Blank, A.; Chamberlain, M. C.; Silber, J. R., O⁶-methylguanine-DNA methyltransferase activity is associated with response to alkylating agent therapy and with MGMT promoter methylation in glioblastoma and anaplastic glioma. *BBA Clin* **2015**, *3*, 1-10.
32. Senthong, P.; Millington, C. L.; Wilkinson, O. J.; Marriott, A. S.; Watson, A. J.; Reamtong, O.; Evers, C. E.; Williams, D. M.; Margison, G. P.; Povey, A. C., The nitrosated bile acid DNA lesion O⁶-carboxymethylguanine is a substrate for the human DNA repair protein O⁶-methylguanine-DNA methyltransferase. *Nucleic Acids Res* **2013**, *41* (5), 3047-55.
33. Du, H.; Wang, P.; Li, L.; Wang, Y., Repair and translesion synthesis of O (6)-alkylguanine DNA lesions in human cells. *J Biol Chem* **2019**, *294* (29), 11144-11153.
34. Tubbs, J. L.; Latypov, V.; Kanugula, S.; Butt, A.; Melikishvili, M.; Kraehenbuehl, R.; Fleck, O.; Marriott, A.; Watson, A. J.; Verbeek, B.; McGown, G.; Thorncroft, M.; Santibanez-Koref, M. F.; Millington, C.; Arvai, A. S.; Kroeger, M. D.; Peterson, L. A.; Williams, D. M.; Fried, M. G.; Margison, G. P.; Pegg, A. E.; Tainer, J. A., Flipping of alkylated DNA damage bridges base and nucleotide excision repair. *Nature* **2009**, *459* (7248), 808-13.
35. Latypov, V. F.; Tubbs, J. L.; Watson, A. J.; Marriott, A. S.; McGown, G.; Thorncroft, M.; Wilkinson, O. J.; Senthong, P.; Butt, A.; Arvai, A. S.; Millington, C. L.; Povey, A. C.; Williams, D. M.; Santibanez-Koref, M. F.; Tainer, J. A.; Margison, G. P., At1 regulates choice between global genome and transcription-coupled repair of O⁶-alkylguanines. *Molecular cell* **2012**, *47* (1), 50-60.
36. Tomaszowski, K. H.; Aasland, D.; Margison, G. P.; Williams, E.; Pinder, S. I.; Modesti, M.; Fuchs, R. P.; Kaina, B., The bacterial alkyltransferase-like (eATL) protein protects mammalian cells against methylating agent-induced toxicity. *DNA repair* **2015**, *28*, 14-20.
37. Sartorelli, A. C.; Booth, B. A., Inhibition of the synthesis of thymine nucleotides by azaserine. *Mol Pharmacol* **1967**, *3* (1), 71-80.
38. Nestmann, E. R.; Brillinger, R. L.; Gilman, J. P.; Rudd, C. J.; Swierenga, S. H., Recommended protocols based on a survey of current practice in genotoxicity testing laboratories: II. Mutation in

Chinese hamster ovary, V79 Chinese hamster lung and L5178Y mouse lymphoma cells. *Mutation research* **1991**, 246 (2), 255-84.

39. Schweikl, H.; Schmalz, G., Glutaraldehyde-containing dentin bonding agents are mutagens in mammalian cells in vitro. *J Biomed Mater Res* **1997**, 36 (3), 284-8.

40. Schweikl, H.; Schmalz, G.; Rackebrandt, K., The mutagenic activity of unpolymerized resin monomers in Salmonella typhimurium and V79 cells. *Mutation research* **1998**, 415 (1-2), 119-30.

41. Follmann, W.; Lucas, S., Effects of the mycotoxin ochratoxin A in a bacterial and a mammalian in vitro mutagenicity test system. *Arch Toxicol* **2003**, 77 (5), 298-304.

42. Dolan, M. E.; Pegg, A. E., O⁶-benzylguanine and its role in chemotherapy. *Clin Cancer Res* **1997**, 3 (6), 837-47.

43. Tsuzuki, T.; Sakumi, K.; Shiraishi, A.; Kawate, H.; Igarashi, H.; Iwakuma, T.; Tominaga, Y.; Zhang, S.; Shimizu, S.; Ishikawa, T.; et al., Targeted disruption of the DNA repair methyltransferase gene renders mice hypersensitive to alkylating agent. *Carcinogenesis* **1996**, 17 (6), 1215-20.

44. Iwakuma, T.; Sakumi, K.; Nakatsuru, Y.; Kawate, H.; Igarashi, H.; Shiraishi, A.; Tsuzuki, T.; Ishikawa, T.; Sekiguchi, M., High incidence of nitrosamine-induced tumorigenesis in mice lacking DNA repair methyltransferase. *Carcinogenesis* **1997**, 18 (8), 1631-5.

45. Sakumi, K.; Shiraishi, A.; Shimizu, S.; Tsuzuki, T.; Ishikawa, T.; Sekiguchi, M., Methylnitrosourea-induced tumorigenesis in MGMT gene knockout mice. *Cancer research* **1997**, 57 (12), 2415-8.

46. Zhang, L.; Kim, S.; Jia, G.; Buhmeida, A.; Dallol, A.; Wright, W. E.; Fornace, A. J.; Al-Qahtani, M.; Shay, J. W., Exome Sequencing of Normal and Isogenic Transformed Human Colonic Epithelial Cells (HCECs) Reveals Novel Genes Potentially Involved in the Early Stages of Colorectal Tumorigenesis. *BMC Genomics* **2015**, 16 Suppl 1, S8.

47. Roig, A. I.; Eskiocak, U.; Hight, S. K.; Kim, S. B.; Delgado, O.; Souza, R. F.; Spechler, S. J.; Wright, W. E.; Shay, J. W., Immortalized epithelial cells derived from human colon biopsies express stem cell markers and differentiate in vitro. *Gastroenterology* **2010**, 138 (3), 1012-21.e1-5.

48. Erzinger, M. M.; Bovet, C.; Hecht, K. M.; Senger, S.; Winiker, P.; Sobotzki, N.; Cristea, S.; Beerenwinkel, N.; Shay, J. W.; Marra, G.; Wollscheid, B.; Sturla, S. J., Sulforaphane Preconditioning Sensitizes Human Colon Cancer Cells towards the Bioreductive Anticancer Prodrug PR-104A. *PLoS One* **2016**, 11 (3), e0150219.

49. Yu, Y.; Wang, J.; Wang, P.; Wang, Y., Quantification of Azaserine-Induced Carboxymethylated and Methylated DNA Lesions in Cells by Nanoflow Liquid Chromatography-Nanoelectrospray Ionization Tandem Mass Spectrometry Coupled with the Stable Isotope-Dilution Method. *Anal Chem* **2016**, 88 (16), 8036-42.

50. Gillet, L. C.; Alzeer, J.; Scharer, O. D., Site-specific incorporation of N-(deoxyguanosin-8-yl)-2-acetylaminofluorene (dG-AAF) into oligonucleotides using modified 'ultra-mild' DNA synthesis. *Nucleic Acids Res* **2005**, 33 (6), 1961-9.

51. (CCLE), B. I. C. C. L. E., <https://portals.broadinstitute.org/ccle/page?gene=MGMT>. **2019**.

52. Surova, O.; Zhivotovsky, B., Various modes of cell death induced by DNA damage. *Oncogene* **2013**, 32 (33), 3789-97.

53. Bhupana Padu Sunkesula, S. R.; Swain, U.; Babu, P. P., Cell death is associated with reduced base excision repair during chronic alcohol administration in adult rat brain. *Neurochem Res* **2008**, 33 (6), 1117-28.

54. Borrás-Fresneda, M.; Barquintero, J. F.; Gomolka, M.; Hornhardt, S.; Rossler, U.; Armengol, G.; Barrios, L., Differences in DNA Repair Capacity, Cell Death and Transcriptional Response after Irradiation between a Radiosensitive and a Radioresistant Cell Line. *Sci Rep* **2016**, 6, 27043.

55. Chatterjee, N.; Walker, G. C., Mechanisms of DNA damage, repair, and mutagenesis. *Environ Mol Mutagen* **2017**, 58 (5), 235-263.

56. Foresta, M.; Izzotti, A.; La Maestra, S.; Micale, R.; Poggi, A.; Vecchio, D.; Frosina, G., Accelerated repair and reduced mutagenicity of DNA damage induced by cigarette smoke in human bronchial cells transfected with E.coli formamidopyrimidine DNA glycosylase. *PLoS One* **2014**, *9* (1), e87984.
57. Moore, J. M.; Correa, R.; Rosenberg, S. M.; Hastings, P. J., Persistent damaged bases in DNA allow mutagenic break repair in Escherichia coli. *PLoS Genet* **2017**, *13* (7), e1006733.
58. Marra, G.; Iaccarino, I.; Lettieri, T.; Roscilli, G.; Delmastro, P.; Jiricny, J., Mismatch repair deficiency associated with overexpression of the MSH3 gene. *Proceedings of the National Academy of Sciences of the United States of America* **1998**, *95* (15), 8568-73.
59. Yeo, J. E.; Khoo, A.; Fagbemi, A. F.; Scharer, O. D., The efficiencies of damage recognition and excision correlate with duplex destabilization induced by acetylaminofluorene adducts in human nucleotide excision repair. *Chemical research in toxicology* **2012**, *25* (11), 2462-8.
60. <https://ch.promega.com/-/media/files/resources/protocols/technical-bulletins/0/celltiter-glo-luminescent-cell-viability-assay-protocol.pdf>. (*Promega online protocol*).
61. Su, Y.; Orelli, B.; Madireddy, A.; Niedernhofer, L. J.; Scharer, O. D., Multiple DNA binding domains mediate the function of the ERCC1-XPF protein in nucleotide excision repair. *J Biol Chem* **2012**, *287* (26), 21846-55.
62. Klein, C. B.; Broday, L.; Costa, M., Mutagenesis assays in mammalian cells. *Curr Protoc Toxicol* **2001**, *Chapter 3*, Unit3.3.
63. <https://www.qiagen.com/ch/products/discovery-and-translational-research/dna-rna-purification/dna-purification/genomic-dna/qiaamp-dna-mini-kit/#orderinginformation>. (*Qiagen online protocol*).
64. Curtis, K. M.; Gomez, L. A.; Rios, C.; Garbayo, E.; Raval, A. P.; Perez-Pinzon, M. A.; Schiller, P. C., EF1alpha and RPL13a represent normalization genes suitable for RT-qPCR analysis of bone marrow derived mesenchymal stem cells. *BMC Mol Biol* **2010**, *11*, 61.
65. Chretien, A. S.; Harle, A.; Meyer-Lefebvre, M.; Rouyer, M.; Husson, M.; Ramacci, C.; Harter, V.; Genin, P.; Leroux, A.; Merlin, J. L., Optimization of routine KRAS mutation PCR-based testing procedure for rational individualized first-line-targeted therapy selection in metastatic colorectal cancer. *Cancer Med* **2013**, *2* (1), 11-20.
66. Defavery, R.; Lemos, J. A.; Kashima, S.; Bernardes, J. E.; Scridelli, C. A.; Covas, D. T.; Tone, L. G., Analysis of the p53 gene by PCR-SSCP in ten cases of Wilms' tumor. *Sao Paulo Med J* **2000**, *118* (2), 49-52.
67. Figer, A.; Irmin, L.; Geva, R.; Flex, D.; Sulkes, A.; Friedman, E., Genetic analysis of the APC gene regions involved in attenuated APC phenotype in Israeli patients with early onset and familial colorectal cancer. *British journal of cancer* **2001**, *85* (4), 523-6.
68. <https://www.qiagen.com/ie/products/discovery-and-translational-research/dna-rna-purification/rna-purification/rna-clean-up/rneasy-minelute-cleanup-kit/#orderinginformation>. (*Qiagen online protocol*).
69. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/cms_042557.pdf. (*Thermofisher online protocol*).
70. https://www.origene.com/catalog/gene-expression/qpcr-primer-pairs/hp206100/mgmt-human-qpcr-primer-pair-nm_002412. (*Origene primers*)
71. Quinlivan, E. P.; Gregory, J. F., 3rd, DNA digestion to deoxyribonucleoside: a simplified one-step procedure. *Anal Biochem* **2008**, *373* (2), 383-5.

Chapter 5: Summary and Outlook

The overarching goal of the work presented in this thesis was to investigate and devise tools to unravel mechanisms of mutagenesis induced by DNA damage. The studies described herein concerned addressing gaps in the technology of detection of mutagenic DNA adducts, as well as in the knowledge of cellular responses to the adducts, which together limit our capacity to study the cause-effect relationship between adducts and mutations, and our understanding on how mutations arise in cells.

In the last study of this thesis (*Chapter 4*), we presented ongoing experimentation aimed to address the longstanding question of what are the cellular pathways proficient at repairing the mutagenic O^6 -CMG DNA adduct. O^6 -CMG has been hypothesized to be a molecular initiator of colorectal cancer linked to meat consumption,¹⁻³ and its mutagenicity properties have been extensively investigated in cell-free studies.⁴⁻⁶ However, contradictory evidence regarding the repair pathways responsible to target O^6 -CMG hinders our understanding of how O^6 -CMG might drive mutagenicity in cells. Studies to address this gap are needed, and to elucidate in particular (1) what are the cellular pathways able to remove O^6 -CMG, (2) how they affect O^6 -CMG-induced toxicity and (3) mutagenicity, and (4) what is their mode of action. These aspects are extensively discussed in *Chapter 4*, where preliminary results and future directions are presented. Although more experiments are required or underway, findings of this work have already addressed an important missing piece of knowledge regarding the cellular phenotype of O^6 -CMG DNA damage, and how this might be connected to mutations *in vivo*.

While the studies described in *Chapters 2 and 3* were completed works either published or submitted, there are several further directions for development of these studies, which are the focus of the *Summary and Outlook* presented here.

5.1 Artificial nucleotides as markers of presence, location and amount of mutagenic DNA damage

DNA damage that interferes with the canonical base pairing and that can be replicated have the potential to induce mutations in the primary structure of the DNA. The lack of methods to measure specific DNA damage within genomic loci limits our capacity to relate mutagenesis with specific DNA damage events.

In the first two studies of this work, *Chapters 2 and 3*, we developed strategies for the detection of mutagenic O^6 -alkyl-guanine (O^6 -alkylG)⁷⁻¹⁰ DNA adducts, namely O^6 -methylguanine (O^6 -MeG) and O^6 -carboxymethyl-guanine (O^6 -CMG), in a sequence-targeted manner. Strategies were based on the use of artificial nucleotides that were specifically paired with the target

damage. The resulting base pairs were bypassed by an engineered polymerase, allowing therefore signal amplification. By iteratively performing reactions of DNA synthesis of a DNA primer hybridized to a target damaged sequence in the presence of the artificial nucleotide (either as the 3'-end nucleotide of the primer, or as TP), we obtained amplified signals that linearly increased with the amount of damage, marking both its presence and DNA sequence location. In *Chapter 2*, we qualitatively detected O^6 -MeG to the single base resolution in a mixture of damaged and undamaged DNA sequences corresponding to the hotspot region of a cancer relevant gene. In *Chapter 3*, we coupled the artificial nucleotide-polymerase amplification-based strategy to an analytical measurement, and generated a quantitative signal for trace amount of O^6 -CMG in a sequence context.

5.2 Future directions of artificial nucleotide-based strategies for the detection of rare DNA damage events

To expand the application of strategies presented in *Chapter 2* and *3* of this thesis, and allow the detection of DNA adducts in biological samples, three key topics should immediately be addressed: (1) generality of target DNA sequences, (2) selectivity with regards to DNA adduct structures, and (3) greater sensitivity, for detection of DNA adducts present at low cellular levels.

5.2.1 Generalization of target DNA sequences

A strength of the detection strategies herein presented is that DNA adducts can be sensed in sequences that are targeted by simply designing a complementary DNA primer. The intrinsic simplicity of the sequence targeting provides an opportunity for generalization of detection to any sequence context. Despite having developed each strategy on one DNA sequence, diverse sequences have been employed with Benzi and ExBenzi nucleotides in other reported or unpublished work, and the pairing between the artificial nucleotide and DNA adduct is sequence-independent according to molecular modelling studies. Therefore, we expect that any sequence can be easily targeted. However, the base context might affect the overall enzymatic catalytic efficiency, and, for example, the calibration curve obtained for the quantification of O^6 -CMG as a function of detected ExBenzi (Fig. 5, Chapter 3) might be primer-template dependent and might need to be obtained before quantitation in a new target sequence. Future studies include testing the detection of DNA adducts in more CRC-relevant sequences, such as hotspot regions of *p53*,¹¹ *apc*¹² and *k-ras*¹³ genes, as well as sequences from genes that are not highly mutated in CRC and in non-hotspot regions of the aforementioned CRC-relevant genes. Ensuring that detection can be expanded to these targets will allow in the future to employ these strategies on biological samples, for mutagenicity and carcinogenicity studies, or for the development as pre-diagnostic tools.

5.2.2 Selectivity towards the DNA adduct structure

To detect DNA adducts in biological samples by means of the presented strategies, some further aspects will be developed and tested, such as selectivity to the target DNA adduct. Based on the nucleotide chemistry and on preliminary experiments performed in our laboratory, we expect no or low pairing of the artificial nucleotides opposite DNA adducts that are not O^6 -alkyGs. However, both Benzi and ExBenzi can pair, in principle, with all O^6 -alkyGs. Benzi-modified primer employed for the sensing of O^6 -MeG in *Chapter 2* might pair to other O^6 -alkyGs tightly enough to allow replication past the modified base pair. ExBenzi (*Chapter 3*) might be incorporated opposite both O^6 -CMG (target) and O^6 -MeG (non-target), as previously observed for its analogue Benzi.¹⁴ At the same time, however, we speculatively expect that DNA adducts already targeted by strategies presented in *Chapter 2* and *3* will give the highest signal to noise ratio and the highest template selectivity vs. other O^6 -alkyGs. Indeed, replication past O^6 -MeG (*Chapter 2*) will be most likely the fastest and prevalent process in a putative sample containing several types of O^6 -alkyGs. Similarly, the undesired incorporation of ExBenzi (*Chapter 3*) opposite O^6 -MeG might account only for a small fraction of the total O^6 -MeG and not be representative of its amount, as O^6 -MeG is efficiently bypassed during DNA synthesis even in the presence of only natural dNTPs. Nevertheless, we do acknowledge that non-target O^6 -alkyGs are expected to be targeted and that their contribution needs to be characterized. We envision at least two possible scenarios to tackle such limitation: (1) the partial contribution of non-target O^6 -alkyGs to the signal could be quantified and subtracted, or, (2) more effectively, polymerase enzymes could be evolved^{15, 16} to be highly and exclusively specific for the incorporation/replication of the target DNA adduct.

5.2.3 Strategies to increase sensitivity of detection

To increase the sensitivity of detection for application to biological samples, particularly of O^6 -CMG via ExBenziTP-mass spectrometry quantitative strategy, aspects concerning the biochemical process and the sample preparation for the analytical measurement can be improved. (1) Screening or evolution of polymerases for a specific activity is an approach that could also be undertaken to reach higher catalytic efficiencies and rate of incorporation, apart from achieving higher template selectivity as discussed in the previous subparagraph. (2) Strategies of sample enrichment¹⁷ could be developed to concentrate the target and to remove the sample matrix, which also implies improving the ratio signal to noise in the case of the analytical measurement. For example, DNA primers could be equipped with a biotin tag at the 5'-end, extended and amplified, and captured via interaction with immobilized streptavidin.¹⁸ Such approach implies that only DNA containing the primer sequence and potentially ExBenzi will be hydrolyzed and analyzed by mass spectrometry, removing both the target, template DNA and all the rest of the matrix (e.g. non target genomic DNA). Additionally, by resolving the enriched samples on denaturing gel, extended primers could be separated from non-extended,

unreacted primers, which is employed in large excess, further concentrating and purifying sample that contains ExBenzi. An alternative, more powerful strategy would consist in directly tagging the base, by functionalizing its aromatic rings with an aliphatic chain linked to an alkyne, which could then be coupled to an azide-biotin molecule¹⁷ and enriched via streptavidin binding. However, while the functional group would be placed to avoid interference with its hydrogen-bond active atoms, changing the structure of the base will most likely affect the entire catalytic process, making this strategy less viable than tagging the primer. Nevertheless, molecular modelling could be performed to predict such disruption and help choose the most favorable position on ExBenzi base and the structure of the functional group.

The data presented in *Chapter 2* and *3* entails significant advances in the detection of mutagenic DNA adducts at target locations in a DNA sequence, which will enable the study of potential cause-effect associations between DNA damage and DNA mutations at specific DNA regions. Following future research aimed to improve sensitivity and selectivity of detection with regard to the DNA adduct, artificial nucleotide-based strategies could be tested for the detection of DNA adducts in biological samples, such as cells in culture to model and study processes of mutagenicity, or tissue samples to develop pre-diagnostic tools.

Furthermore, in *Chapter 4*, we uncovered mechanisms of mutagenesis induced by DNA damage in cells, i.e. reduced DNA repair capacity and accumulation of DNA damage in cells. This ongoing work will provide important information that could be integrated in the *in vitro* model of mutagenicity enabled in the future by the developed detection strategies.

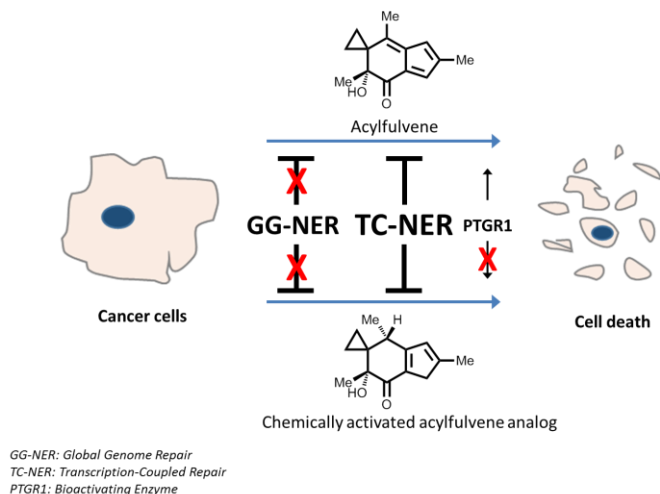
With the research of this thesis, we developed tools and advanced knowledge aimed to elucidate mechanisms of mutagenesis induced by DNA damage that are of high biological and clinical relevance. The information gained from this work and from future studies will have wide implications on toxicology and cancer research, particularly on the prediction of mutagenesis and carcinogenesis processes linked to environmental exposure and genetic deficiencies.

5.3 References

1. Cross, A. J.; Ferrucci, L. M.; Risch, A.; Graubard, B. I.; Ward, M. H.; Park, Y.; Hollenbeck, A. R.; Schatzkin, A.; Sinha, R., A large prospective study of meat consumption and colorectal cancer risk: an investigation of potential mechanisms underlying this association. *Cancer Res* **2010**, *70* (6), 2406-14.
2. Steinberg, P., Red Meat-Derived Nitroso Compounds, Lipid Peroxidation Products and Colorectal Cancer. *Foods* **2019**, *8* (7).

3. Hemeryck, L. Y.; Rombouts, C.; Hecke, T. V.; Van Meulebroek, L.; Bussche, J. V.; De Smet, S.; Vanhaecke, L., In vitro DNA adduct profiling to mechanistically link red meat consumption to colon cancer promotion. *Toxicol Res* **2016**, *5* (5), 1346-1358.
4. Gottschalg, E.; Scott, G. B.; Burns, P. A.; Shuker, D. E., Potassium diazoacetate-induced p53 mutations in vitro in relation to formation of O⁶-carboxymethyl- and O⁶-methyl-2'-deoxyguanosine DNA adducts: relevance for gastrointestinal cancer. *Carcinogenesis* **2007**, *28* (2), 356-62.
5. Raz, M. H.; Sandell, E. S.; Patil, K. M.; Gillingham, D. G.; Sturla, S. J., High Sensitivity of Human Translesion DNA Synthesis Polymerase kappa to Variation in O⁶-Carboxymethylguanine Structures. *ACS Chem Biol* **2019**, *14* (2), 214-222.
6. Wang, P.; Leng, J.; Wang, Y., DNA replication studies of N-nitroso compound-induced O⁶-alkyl-2'-deoxyguanosine lesions in Escherichia coli. *J Biol Chem* **2019**, *294* (11), 3899-3908.
7. Du, H.; Wang, P.; Li, L.; Wang, Y., Repair and translesion synthesis of O⁶-alkylguanine DNA lesions in human cells. *J Biol Chem* **2019**, *294* (29), 11144-11153.
8. Ezerskyte, M.; Paredes, J. A.; Malvezzi, S.; Burns, J. A.; Margison, G. P.; Olsson, M.; Scicchitano, D. A.; Dreij, K., O⁶-methylguanine-induced transcriptional mutagenesis reduces p53 tumor-suppressor function. *Proc Natl Acad Sci* **2018**, *115* (18), 4731-4736.
9. Wang, P.; Wang, Y., Cytotoxic and mutagenic properties of O⁶-alkyl-2'-deoxyguanosine lesions in Escherichia coli cells. *J Biol Chem* **2018**, *293* (39), 15033-15042.
10. Margison, G. P.; Santibanez Koref, M. F.; Povey, A. C., Mechanisms of carcinogenicity/chemotherapy by O⁶-methylguanine. *Mutagenesis* **2002**, *17* (6), 483-7.
11. Li, X. L.; Zhou, J.; Chen, Z. R.; Chng, W. J., P53 mutations in colorectal cancer - molecular pathogenesis and pharmacological reactivation. *World J Gastroenterol* **2015**, *21* (1), 84-93.
12. Fodde, R., The APC gene in colorectal cancer. *Eur J Cancer* **2002**, *38* (7), 867-71.
13. Porru, M.; Pompili, L.; Caruso, C.; Biroccio, A.; Leonetti, C., Targeting KRAS in metastatic colorectal cancer: current strategies and emerging opportunities. *J Exp Clin Cancer Res* **2018**, *37* (1), 57.
14. Wyss, L. A.; Nilforoushan, A.; Williams, D. M.; Marx, A.; Sturla, S. J., The use of an artificial nucleotide for polymerase-based recognition of carcinogenic O⁶-alkylguanine DNA adducts. *Nucleic Acids Research* **2016**, *44* (14), 6564-6573.
15. Gloeckner, C.; Kranaster, R.; Marx, A., Directed evolution of DNA polymerases: construction and screening of DNA polymerase mutant libraries. *Curr Protoc Chem Biol* **2010**, *2* (2), 89-109.
16. Wang, Y.; Ngor, A. K.; Nikoomanzar, A.; Chaput, J. C., Evolution of a General RNA-Cleaving FANA Enzyme. *Nat Commun* **2018**, *9* (1), 5067.
17. Szychowski, J.; Mahdavi, A.; Hodas, J. J.; Bagert, J. D.; Ngo, J. T.; Landgraf, P.; Dieterich, D. C.; Schuman, E. M.; Tirrell, D. A., Cleavable biotin probes for labeling of biomolecules via azide-alkyne cycloaddition. *J Am Chem Soc* **2010**, *132* (51), 18351-60.
18. Dundas, C. M.; Demonte, D.; Park, S., Streptavidin-biotin technology: improvements and innovations in chemical and biological applications. *Appl Microbiol Biotechnol* **2013**, *97* (21), 9343-53.

Chapter 6: Appendix A: Modulation of cytotoxicity by transcription-coupled nucleotide excision repair is independent of the requirement for bioactivation of acylfulvene



Reproduced with permission from

Claudia Otto, Graciela Spivak, Claudia M.N. Aloisi, Mirco Menigatti, Hanspeter Naegeli, Philip C. Hanawalt, Marina Tanasova, Shana J. Sturla*, Modulation of cytotoxicity by transcription-coupled nucleotide excision repair is independent of the requirement for bioactivation of acylfulvene, *Chemical Research in Toxicology* **2017**, 30, 3, 769-77

Copyright © 2017 American Chemical Society

C.M.N. Aloisi synthesized acylfulvene, performed cytotoxicity assays, siRNA transfection and Western Blot experiments, interpreted data and wrote the manuscript. C. Otto performed preliminary siRNA experiments, RT-PCR and Western Blot experiments, and interpreted data. G. Spivak performed cell survival and RNA synthesis assays, and interpreted data. M. Menigatti performed RT-PCR experiments. H. Naegeli conceived siRNA experiments and provided material. P.C. Hanawalt interpreted data and wrote the manuscript. M. Tanasova synthesized acylfulvene and iso-MO, interpreted data and wrote the manuscript. S.J. Sturla devised experiments, interpreted data and wrote the manuscript.

6.1 Abstract

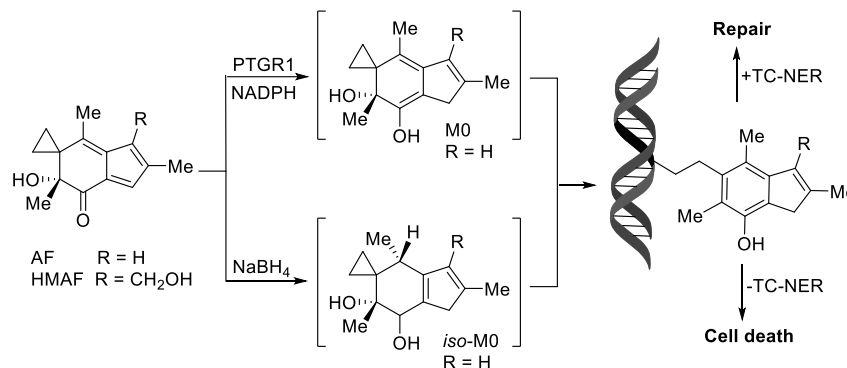
Bioactivation as well as DNA repair affect the susceptibility of cancer cells to the action of DNA-alkylating chemotherapeutic drugs. However, information is limited with regard to the relative contributions of these processes to the biological outcome of metabolically activated DNA alkylating agents. We evaluated the influence of cellular bioactivation capacity and DNA repair on cytotoxicity of the DNA alkylating agent, acylfulvene (AF). We compared cytotoxicity and RNA synthesis inhibition by AF and its synthetic activated analog *iso*-M0 in a panel of fibroblast cell lines with deficiencies in transcription-coupled (TC-NER) or global genome nucleotide excision repair (GG-NER). We related these data to the inherent bioactivation capacity of each cell type on the basis of mRNA levels. We demonstrated that specific inactivation of TC-NER by siRNA had the largest positive impact on AF activity in a cancer cell line. These findings establish that transcription-coupled DNA repair reduces cellular sensitivity to AF, independent of the requirement for bioactivation.

6.2 Introduction

DNA repair contributes significantly to the responses of mammalian cells to DNA-alkylating drugs.¹ In addition, it is well-established that xenobiotic-metabolizing enzymes can strongly influence the cellular responses, either by bioactivation or through detoxification.² Reliable interpretation of biological endpoints requires consideration of possible relationships between these two processes, i.e. repair and metabolism, especially in the context of metabolically-activated DNA alkylating agents. For the DNA-alkylating antitumor acylfulvenes (AFs) and their natural product precursor illudin S,^{3, 4} both DNA repair and bioactivation have been implicated in cytotoxicity, however, their potentially competing influences have not been evaluated. Understanding the relative contributions of repair and metabolism to drug cytotoxicity may lead to more reliable predictions of toxic outcomes and could facilitate the development of more selective chemotherapies. Furthermore, this information may provide predictive markers to aid in selection of patients who may respond more effectively to chemotherapies with DNA alkylating drugs.

The AF class of compounds, which includes the hydroxymethylated analog HMAF, an experimental clinical drug called irofulven (Scheme 1), are DNA-alkylating anticancer agents derived from illudin natural products. AFs have several favorable characteristics for applications in chemotherapy.⁵ In cell-based cytotoxicity studies, AFs were shown to induce apoptosis in tumor cells but were cytostatic in normal cells.⁶ In tumor xenografts, HMAF was a more effective tumor growth inhibitor than cisplatin, doxorubicin, irinotecan, or mitomycin C, and it also sensitized multidrug-resistant cells to some conventional chemotherapeutic drugs.⁷⁻⁹ Despite encouraging pre-clinical data,⁴ however, the results from early clinical trials were

ambiguous, and a Phase III clinical trial was terminated.¹⁰ Although the clinical outcomes were insufficient for promoting the drug, certain individuals exhibited outstanding results. Thus, there is renewed interest in the clinical potential of acyfulvenes¹¹ in the context of biomarker-driven personalized medicine strategies, in which patients can be stratified on the basis of molecular markers, such that who are most likely benefit from the therapy may be identified.¹²



Scheme 1. Enzymatic (top) or chemical (bottom) reduction of AF yields isomeric reactive intermediates that form the same AF-DNA adduct.

The DNA alkylation and cytotoxicity of AF and HMAF have been shown to depend upon cellular drug-bioactivating capacity, specifically the capacity for enone reduction.¹³⁻¹⁷ It has been shown that reductive bioactivation of AF is mediated by NADPH-dependent prostaglandin reductase 1 (PTGR1), also referred to as alkenal/one oxidoreductase (AOR), and that this activity promotes its cytotoxicity.^{15, 17, 18} AF alkylates purine bases through interactions in the minor groove of DNA, leading to depurination and subsequent strand breaks.¹³ Increased levels of AF-specific DNA adducts correlate with higher levels of PTGR1.^{15, 18} Cells that overexpress PTGR1 by 10-fold suffered 4-fold higher levels of AF-induced DNA adducts than control cells. However, adduct levels decreased as a function of the duration of treatment,¹⁸ possibly due to repair.

In addition to the well-established association of AF toxicity with reductive bioactivation, cell resistance to HMAF and illudin S was found to be dependent upon nucleotide excision repair (NER).¹⁹⁻²² NER can remove a wide variety of mainly bulky DNA lesions such as UV-induced photoproducts and chemical adducts. NER is divided in two sub-pathways: Transcription-coupled NER (TC-NER) removes DNA lesions exclusively from the template strands in actively transcribed DNA,^{23, 24} whereas global genome NER (GG-NER) repairs lesions throughout the entire genome regardless of transcriptional activity. TC- and GG-NER operate by similar mechanisms, except for the initial damage recognition step. For GG-NER, damage recognition is performed by XPC in complex with hRAD23b, and in some cases by XPE (DDB2); TC-NER requires translocating RNA polymerase II as well as the CSB, CSA and UVSSA proteins but not XPC and XPE. When the cytotoxicity of illudin S and HMAF was evaluated in NER deficient cells, both compounds were more toxic than to normal cells, while deficient base excision repair had no

significant effect.^{4, 20} Interestingly, the cytotoxicity was specifically reduced by effective TC-NER, but not impacted by GG-NER.²¹

Although it is well-established that PTGR1-mediated reductive bioactivation enhances AF cytotoxicity, it is not known whether this process may confound the interpretation of AF cytotoxicity data from human fibroblasts with variable repair capacities.²¹ Therefore, we evaluated reductive bioactivation in the context of cellular TC-NER-proficiency, employing a combination of strategies. A synthetic analog of AF that is chemically activated and therefore does not require PTGR1 for activation (*iso-M0*, Scheme 1)¹⁷ was tested in fibroblasts with varying NER capacities to determine the relative contributions of bioactivation and DNA repair to the cellular response to AF. *Iso-M0* is isomeric to the reactive intermediate formed during PTGR1-mediated activation, and it alkylates DNA, to yield the same DNA adduct profile as AF in the presence of PTGR1. However, the potency of *iso-M0* is independent of the cellular reductive bioactivation capacity.¹⁷

We compared cytotoxicity and RNA synthesis inhibition induced by AF, HMAF and *iso-M0*, in human fibroblasts that were proficient or deficient in GG-NER, TC-NER, or both. Levels of the bioactivating enzyme PTGR1 were also compared amongst these human fibroblast strains. Furthermore, to compare the AF sensitivity of cells with modulated bioactivation and repair capacity, but an otherwise isogenic background, SW480 cells and SW480 cells engineered to overexpress the AF-bioactivating enzyme PTGR1 (SW480-PTGR1), were treated with siRNA for selected NER factors, and modulation of cytotoxicity was evaluated.

6.3 Material and Methods

Cell strains and culture conditions

Primary fibroblast cell lines were purchased from Coriell Institute for Medical Research. WT GM00037 are non-fetal primary fibroblasts from a repair proficient donor. XP and CS cells are from unexposed skin biopsies of patients with xeroderma pigmentosum (XP) or Cockayne syndrome (CS). SW-480 cells were provided by Dr. Giancarlo Marra (University of Zurich). SW-480-PTGR1 cells were generated as described by Pietsch et al.¹⁸ Fibroblasts were maintained as monolayers in 75 cm² flasks in GIBCO minimum essential medium with Earle's salts and glutamax, supplemented with 15% FBS and 1% penicillin-streptomycin in a humidified, 5% CO₂ atmosphere at 37 °C. SW-480 control cells and SW-480 cells stably transfected to overexpress PTGR1 were maintained in RPMI 1640 medium with 10% fetal calf serum (v/v) and 1% pen/strep (v/v) and incubated at 37 °C in a humidified incubator containing 5% CO₂.

Chemicals

(-)-Illudin S, isolated from *Omphalotus* species,²⁵ was provided by MGI Pharma (Bloomington, MN). (-)-Acylfulvene ((-)-AF) was synthesized from illudin S by a previously published

procedure.⁵ The chemically activated AF analog (iso-M0) was synthesized from AF following a previously published procedure.¹⁷ 3-[4, 5-dimethyl-2-yl]-2, 5-diphenyltetrazolium bromide (MTT) was from SIGMA. Stock 100x solutions in DMSO were stored at -20°C.

Cytotoxicity assay

Cells were plated in 96-well plates (~2 x 10³ per well) and treated with solutions of test compounds that were diluted with culture medium (0.1% final concentration of DMSO). DMSO-treated cells were used as control and backgrounds from cell-free wells were subtracted. After 1 h, cells were washed and the medium was replaced with fresh drug-free medium and further incubated for 72 h. Cell survival was measured by adding 50 µl medium containing 0.5 mg/ml MTT per well and incubating for 4 h at 37 °C; the intensity of the developed color was assessed by monitoring absorbance at 540 nm. Each data set was acquired in triplicate for each of 3 or more experiments and the averaged absorption values of the drug-treated cells were normalized to those of DMSO-treated cells (assigned as 100% viability). Treated cells were examined microscopically to confirm cytotoxicity at the highest doses used.

RNA synthesis recovery

RNA synthesis levels in AF- and iso-M0-treated cells were determined by measuring incorporation of radiolabeled uridine into RNA as a function of post-incubation period. Fibroblasts were grown in 75 cm² flasks with complete medium containing 0.02 µCi [2-¹⁴C]thymidine (Moravek) until semi-confluency (2-3 days). Cells were then transferred to 12-well plates, as follows: the [2-¹⁴C] thymidine-containing medium was collected, cells were treated with 2 mL of 0.25% trypsin-EDTA at 37 °C for 5 min and transferred to centrifuge tubes. Cells were then pelleted by centrifugation (5 min, 5000g), resuspended in the collected [2-¹⁴C] thymidine-containing medium, plated in 12-well plates and incubated overnight. Two 12-well plates were set up for each cell line, with 3 wells for each treatment and incubation time. For drug treatments, the [2-¹⁴C] thymidine-containing medium was collected, the wells rinsed with PBS, and 0.5 ml HBSS containing AF (20 ng/ml), iso-M0 (20 ng/ml) or no drug (control) were added to each well. After 1-hour incubation at room temperature, the drug was removed and 0.5 ml of the reserved [2-¹⁴C] thymidine-containing medium was added to each well. The “24 hour” plates were immediately placed in the incubator.

To pulse-label nascent RNA, 0.1 ml of medium containing [5-³H]uridine (Moravek) was added to each well to achieve a 3 µCi/ml final concentration. The cells were incubated for 1 h at 37 °C, the medium removed, and the wells were rinsed twice with PBS. Lysis solution (0.5 ml; 150 mM NaCl, 10 mM Tris HCl pH 8.0, 1 mM EDTA, 0.5% SDS, and 10 mg/ml proteinase K) was added to each well. The plates were incubated at 37 °C for 1 h, the lysates were transferred to centrifuge tubes, the wells were rinsed with 0.5 ml PBS and rinses were added to the corresponding samples. Lysates were incubated at 54 °C for 1 hour. Trichloroacetic acid (20 %, 0.5 ml) was

added to each tube, and samples were chilled at 4 °C for at least two hours. Precipitated nucleic acids were collected on Millipore HA 0.45 µm filters, which were dried under a heat lamp. Radioactivity was measured by scintillation counting. Three biological experiments were done, each in triplicate. Samples labeled with only ¹⁴C or ³H were used to calculate and correct overlaps. Ratios of ³H to ¹⁴C radioactivity were used to normalize the amount of transcription to the number of cells, and were averaged for each set of triplicate measurements. Recoveries were calculated relative to the ratios in untreated cells.

Reverse-transcription real time PCR analysis of PTGR1 expression

Untreated cells were collected from two flasks at 90-95% confluence as follows: 2 mL of 0.25 % trypsin-EDTA was added and the cells were incubated for 5 min at 37 °C. Total mRNA was extracted using the QIAGEN miRNeasy® mini kit and kept on dry ice until cDNA synthesis (see Supplementary Information for primer sequences and PCR conditions). Synthesis of first-strand cDNA was performed with the Transcriptor First Strand cDNA Synthesis kit (Roche) according to manufacturers' instructions with random hexamer primers. Expression of prostaglandin reductase 1 (PTGR1, Gene ID: 22949) and of the reference gene porphobilinogen deaminase (PBGD, Gene ID: 3145) was measured with the LightCycler® 480 real-time PCR system and a LightCycler® 480 SYBR Green I master kit (Roche). The analysis was performed in duplicate and relative PTGR1 expression in the different cell lines was calculated by the method of Pfaffl et al.²⁶

siRNA of NER factors in SW480 and PTGR1-overexpressing cells

SW-480 cells were maintained in RPMI 1640 medium with 10% fetal calf serum (v/v) and 1% pen/strep (v/v) and incubated at 37 °C in a humidified incubator containing 5% CO₂. siRNA sequences (Table S2, Supplementary Information) were purchased from Microsynth. Transfections were performed with Lipofectamine RNAiMAX according to the manufacturer's protocol (Invitrogen). The siRNA concentrations were 10 nM for non-coding siRNA and the silencing of XPA, XPC and CSB targets (siRNA sequences are listed in Supplementary Material). Protein levels were analyzed 72 h after siRNA transfections by Western blot. For treatment with Acylfulvene, cells were seeded 48 h after siRNA transfection in 96-well plates at a density of 3.0×10^3 cells/well and were allowed to attach overnight. Cell viability experiments were initiated by replacing the maintenance media with media containing AF (0, 10, 100, 150, 300, 500, 1000, and 5000 nM final concentration, 0.1-0.5% DMSO). Cell viability was measured 48 h later with the CellTiter-Glo luminescent cell viability assay according to the manufacturer's protocol (Promega). With this assay, the number of viable cells in culture is determined on the basis of ATP quantification as an indicator of metabolically active cells. Three independent experiments were performed, and each condition was analyzed three times (i.e. three biological replicates, each with three technical replicates). Reduction of protein expression was confirmed by

western blot analysis whereby after 72 h of siRNA treatment, cells were lysed in RIPA buffer (Pierce, Thermo Scientific) containing protease inhibitors (Complete Tablets EDTA-free EASYpack, Roche). Protein samples (40 µg) were loaded onto an SDS PAGE gel (4-12 % NuPAGE bis-tris gels, Invitrogen), electrophoresed at 200 V for 180 min in 1XMOPS running buffer (NuPAGE MOPS SDS Running buffer 20X, Life technologies) and transferred onto PVDF membranes (Amersham Hybond-P) at 30 V for 50 min. Staining was performed with the following primary antibodies and respective secondary HRP antibodies: Mouse-anti-XPC (Abcam, 1:500), and rabbit-anti-CSB (Abcam, 1:500). All primary antibodies were diluted in 3% BSA, TBS-Tween. Western blots were developed with Pierce ECL western blotting substrate (Thermo Scientific) and exposed to X-ray film.

6.4 Results

6.4.1 Cytotoxicity of AF and iso-M0 in human fibroblasts with different DNA repair proficiencies

In previous studies, CSB deficiency in human fibroblasts, associated with lack of TC-NER function, was shown to correlate with higher toxicity of illudin S and HMAF^{20, 21}. In contrast, XPC deficiency, associated with lack of GG-NER function but proficiency in TC-NER, had no impact on HMAF or illudin S cytotoxicity,^{20, 21} indicating that HMAF and illudin S-induced DNA adducts are selectively processed by TC-NER. However, in those studies, AF was not tested and the bioactivation potential of the cells was not measured. Therefore, to first confirm that TC-NER proficiency influences AF cytotoxicity in the same manner as for HMAF and illudin S, we evaluated the toxicity of AF in human fibroblast cells competent for GG- and/or TC-NER, or neither. As anticipated, AF was most active in NER-deficient XPA cells, inducing 70 % cell death at 2 ng/ml (Fig. 1A). Somewhat lower cytotoxicity (60 % cell death) was detected in TC-NER-deficient CSB cells. The GG-NER-deficient XPC and NER-proficient WT cells were resistant to low AF concentrations (1-5 ng/ml), but responded with 20% and 40% lethality, respectively, at higher AF concentration. The higher susceptibility of XPA and CSB cells ($p < 0.039$) implicates NER (and specifically TC-NER) as a factor controlling AF cytotoxicity. Furthermore, AF cytotoxicity was enhanced specifically when TC-NER was compromised. As anticipated, these data for AF are consistent with those previously reported for HMAF and illudin S.^{20, 21}

After confirming that AF cytotoxicity was enhanced in XPA or CSB cells, we evaluated the survival of the same cells treated with the AF analog *iso-M0*, in order to control for any possible alteration of this profile due to potential differences in cellular bioactivating capacity (Fig. 1B). We found that NER-deficient XP-A and TC-NER deficient CS-B cells were significantly susceptible to *iso-M0*, while WT and GG-NER-deficient XP-C cells were resistant. Thus AF and *iso-M0* elicited

similar responses in XP cells, with TC-NER-deficient cells being susceptible, and TC-NER-proficient cells being resistant towards AF or iso-M0.

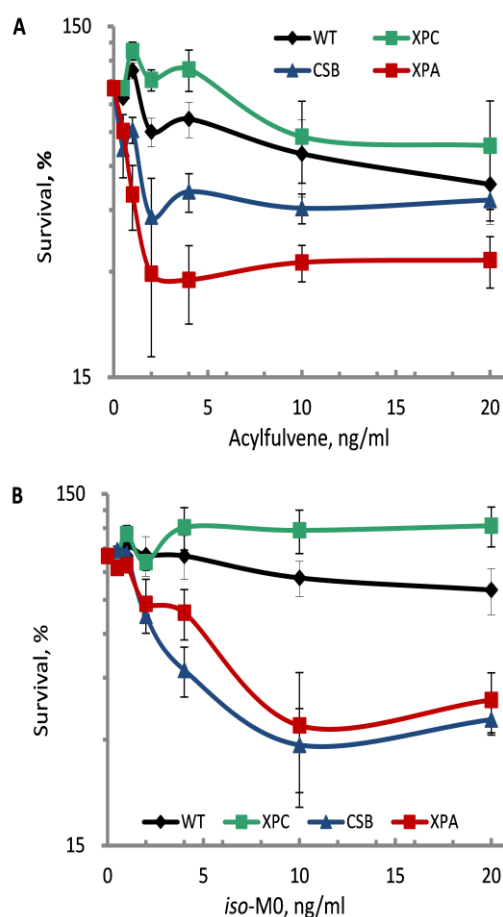


Figure 1. Cytotoxicities of AF and *iso*-M0 in human fibroblast cells. Cells were exposed to AF (A) or *iso*-M0 (B) for 1 h, followed by 72 h postincubation in drug-free medium and assayed for cell survival.

6.4.2 AF and its activated analog *iso*-M0 inhibit RNA synthesis

Reduction in cellular transcriptional level in response to a DNA damaging agent reflects the proficiency of the cells in TC-NER, because of the link between DNA damage-induced arrest of RNA polymerase II and initiation of TC-NER.^{24, 27} We analyzed the propensity of AF and *iso*-M0 to inhibit RNA synthesis during 1-h treatments with the drugs and monitored its recovery in WT, XP and CS cells. Nascent RNA levels were measured both immediately after treatment with AF or *iso*-M0 (0 h), or following 24 h post-incubation of the treated cells in drug-free medium (24 h). Data immediately after treatment were collected to assess inhibitory activities of AF and *iso*-M0, while the changes in RNA levels after 24 h post-incubation should reflect the potential of the

cells to recover from the damage, thus providing a measure of TC-NER or GG-NER involvement in AF and *iso*-M0 induced lesion recognition and repair.

Immediately after AF treatment, RNA synthesis was inhibited in all cell lines, although to different extents (Fig. 2; AF). Thus, RNA synthesis was inhibited by 80 % in WT, by 65 % in CSB cells, and by 55 % ($p < 0.039$) in XPA or XPC cells. Treating cells with *iso*-M0 resulted in relatively similar levels of RNA synthesis inhibition, ranging within 47-55 % (Fig. 2, *iso*-M0).

After 24 h post-incubation in drug-free medium, cells proficient in TC-NER almost completely recovered from the drug-induced damage (Fig. 2; AF, 24 h and *iso*-M0, 24 h). Recovery was 74 % and 87 % in AF-treated WT and XPC cells, respectively, and complete in *iso*-M0-treated WT and XPC cells. In contrast, RNA synthesis did not recover in drug-treated NER-deficient XPA or TC-NER-deficient CSB cells. Although AF completely suppressed RNA synthesis in TC-NER-deficient cells 24 h after treatment, *iso*-M0 had no further inhibitory effect on RNA synthesis in XPA and CSB cells during post-incubation (Fig. 2, AF, 24 h and *iso*-M0, 24 h).

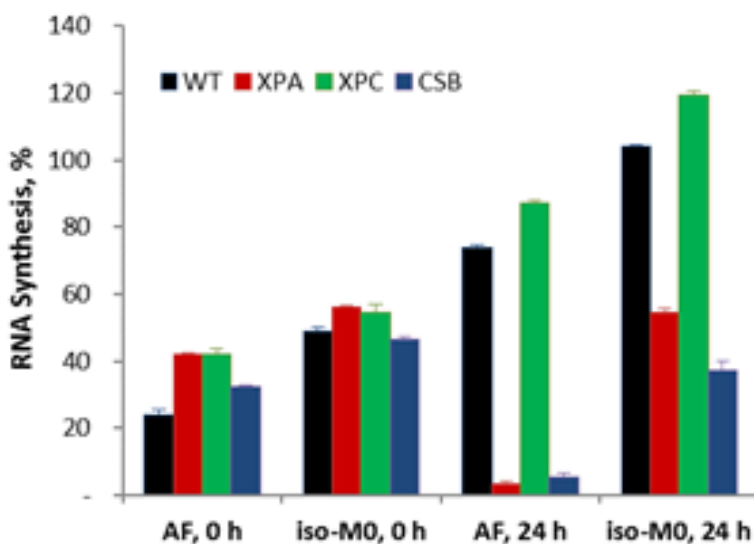


Figure 2. RNA levels in cells treated with 20 ng/ml AF or *iso*-M0 immediately after 1 h treatment (AF, 0 h and *iso*-M0, 0 h) and after 24 h post-incubation in drug free medium (AF, 24 h and *iso*-M0, 24 h). Values were normalized to those in mock-treated cells.

6.4.3 PTGR1 levels in WT, XP and CSB cells

AF cytotoxicity and levels of AF-specific DNA adducts have been shown to correlate with the abundance of PTGR1 in cell-based and cell-free systems; the mechanism of cellular bioactivation involves reduction of the unsaturated enone by this NADPH-dependent reductase.^{13-17, 28} We examined basal PTGR1 levels in cell lysates of WT, XPA, XPC and CSB cells by real time RT-PCR analysis. The WT fibroblasts expressed PTGR1 the least. In XPA and XPC cells, the average levels of PTGR1 mRNA appeared to be 2.5-fold higher than in WT cells, and in CSB cells, 1.8-fold higher

(Fig. S1, Supplementary Information). These data suggest that the differences observed for the susceptibility of the fibroblasts to AF (Figure 1), do not correlate with bioactivation potential. Moreover, PTGR1 levels may be slightly lower in CSB than XPA cells, yet the CSB cells are more sensitive, substantiating the overriding importance of DNA repair differences in the response to the drug.

6.4.4 Transient depletion of certain NER factors in cells can enhance AF cytotoxicity

To compare AF sensitivity in cells with modulated bioactivation and repair capacity, but an otherwise isogenic background, SW480 human cancer cells engineered to overexpress the AF-bioactivating enzyme PTGR1 (SW480-PTGR1),¹⁸ were treated with siRNA for XPC and CSB and the impact on cytotoxicity was evaluated. Cells were transfected with siRNA for down regulation of XPC or CSB separately (Table S2, Supplementary Information). Transfected cells were re-seeded 48 h later and cellular sensitivity towards AF was evaluated after another 48 h period; XPC and CSB knock-down was confirmed to be effective by western blot (Fig. S2, Supplementary Information). The SW480-PTGR1 cells, which had 17-fold higher levels of PTGR1 protein,¹⁸ were more sensitive to the drug as compared to control cells. For non-targeted control transfections, EC₅₀ values were 141 nM in SW480-PTGR1 cells vs. 255 nM in SW480 cells (Figure 3). These results are consistent with a previous report in which significant increases in DNA adduct levels and cytotoxicity were detected in SW480-PTGR1 vs. SW480 cells.¹⁸

For both cell lines SW480 and SW480-PTGR1, significantly higher sensitivity towards AF was observed when CSB was transiently silenced: the EC₅₀ went from 255 nM (control) to 29 nM (CSB) in SW480 cells, and from 141 nM (control) to 66 nM (CSB) in SW480-PTGR1 (Figure 3). Furthermore, silencing XPC had no significant impact on toxicity relative to the non-coding controls. These data further corroborate the conclusion from experiments in primary fibroblasts, with a significant increase in susceptibility to the drug demonstrated when CSB, but not XPC was silenced, confirming the role of TC-NER in AF repair.

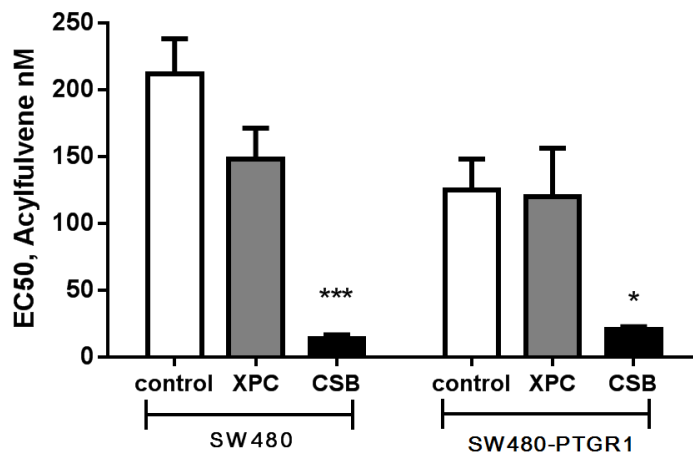


Figure 3. Cytotoxicity of AF in SW480 wild type (SW480) and PTGR1-overexpressing cells (SW480-PTGR1) transfected with siRNA to knock-down GG-NER (XPC) or TC-NER (CSB). Control cells were transfected with a non-coding sequence of siRNA (control). Cell viability was measured with two technical replicates from each of three independent experiments. Mean values and standard errors of the mean are indicated. Statistical significance of differences relative to control was calculated using two-way ANOVA with a Dunnett post hoc test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

6.5 Discussion

The experimental anticancer drug HMAF is not toxic to TC-NER-proficient cells,^{19, 20} whereas cells with high capacity for enone reduction are particularly susceptible to the drug.^{13-17, 28} Thus, roles for DNA repair and drug bioactivation in the cytotoxicity of AFs have been independently studied, but the potentially confounding interplay between these factors has not been addressed in concert.

Using human fibroblasts deficient in TC-NER activity (CSB), GG-NER activity (XPC), both repair pathways (XPA) or neither (WT), we found that AF was most potent toward XPA and CSB cells. These cellular response patterns are very similar to those that were previously reported for the AF analog HMAF by Koepfel²¹ and Jaspers²⁰. We also determined that the synthetic analog iso-MO, which does not require bioactivation, elicits toxic responses in cells with defective TC-NER. In addition to the assessment of cell sensitivity, we evaluated recovery of RNA synthesis in cells immediately following treatment with AF and after a 24 h recovery period; this assay reveals the ability of the cells to remove lesions from transcribed DNA strands. Furthermore, we characterized mRNA levels of the AF-bioactivating enzyme PTGR1 in the fibroblast strains. Considering the results of these three studies together, a picture that emerges is that despite

slightly lower PTGR1 in CSB cells, cell viability is decreased and cells fail to recover RNA synthesis capacity (Table 1). These data for AF, and those of Koepfel et al.²¹ and Jaspers et al.²⁰ reported previously for HMAF, support the concept that cellular resistance to AFs is governed by TC-NER proficiency.

Table 1. Relative direction and magnitude of changes in bioactivation protein levels, cell survival and recovery of RNA synthesis

	Relative to response of WT fibroblasts ^a		
	XPA	XPC	CSB
PTGR1-associated mRNA levels	↑↑	↑↑	↑
Cell survival ^b	↓	↔	↓
Recovery of RNA Synthesis ^c	↓	↔	↓

^a Red arrows pointing up indicate increased levels relative to WT; green arrows pointing down indicate decreased response relative to WT; orange horizontal arrows indicate lack of effect. Number of red arrows reflects magnitude of increase. ^b 2 h treatment, 20 ng/ml AF, 72 h re-growth. ^c 1 h treatment, 20 ng/ml AF, 24 h re-growth

The relative impacts of bioactivation and repair was further established by a chemical probe approach using *iso*-M0, a synthetic analog of AF that does not require bioactivation to alkylate DNA in the same manner as AF.¹⁸ Thus, *iso*-M0 was equally cytotoxic to XPA and CSB cells, whereas AF cytotoxicity differed, possibly due to the slightly higher levels of PTGR1 in the XPA cells. These data are consistent with previous findings that increased bioactivation activity increases levels of DNA modification and the cytotoxic response in NER-deficient cells.^{13-17, 28} Moreover, this process does not appear to diminish with importance of TC-NER in cellular resistance to AF or its bioactive forms. Interestingly, the CSB cells seem to be more sensitive to *iso*-M0 than to AF at higher doses (20 ng/mL, Fig. 1), yet RNA synthesis recovery in the same cells was greater for *iso*-M0 than AF (Fig. 2). The reason for this observation could be a more promiscuous reactivity of *iso*-M0, modifying a wider range of biomolecules also contributing to cytotoxicity, but reducing the contribution of the impedance of RNA synthesis.

Acylfulvene alkylates DNA primarily from minor-groove binding and modification of the 3-position of Adn. While these adducts are quite unstable in naked DNA (half-life 2-8 hours), they appear to persist significantly longer in NER proficient cells (1 d). The depurination half-life of AF-adducts in cells treated with a G2 checkpoint inhibitor that reduces cellular DNA repair

function was extended to 12 days.²² Taken together with the results presented here, the model that emerges is that AF adducts persist long enough and are sufficiently large to stall RNA Pol II, however, it appears they orient in the duplex such that they are non-helix-distorting and thus are not recognized by XPC. Another DNA alkylating agent for which a similar model has been established is the genotoxin aristolochic acid (AA).²⁹ AA reacts upon bioactivation by cellular nitroreductases with purine bases in DNA to form aristolactam-DNA adducts. Unlike acylfulvene adducts, these are non-depurinating and their persistence in the genome and their marked strand bias towards the non-transcribed DNA strand suggests their selective repair by TC-NER.³⁰ A mechanistic and structural basis for the lack of repair of AA-adducts by GG-NER is that AA-adducts cause only minor destabilization of the DNA duplex and intercalate in the helix, avoiding recognition by the XPC-HR23B complex of the GG-NER pathway to initiate repair.³¹

Silencing NER factors in the human colon cancer cell line SW480 with standard or elevated PTGR1 levels in SW480 cells stably transfected with Myc-DDK tagged PTGR1 cDNA¹⁸ corroborated that also in a cancer cell line, TC-NER has a significant influence on AF toxicity independent of bioactivation and moreover supports beyond a doubt that the differences in repair-deficient fibroblasts are not due to inherent differences between the individuals from whom cells were obtained (Figure 3). Silencing CSB by siRNA led to a 2-9 fold increase in toxicity, a change with a magnitude similar to what was observed previously (2-4 fold increase) when the same experiment was performed with cisplatin in a different cancer cell line.³² Furthermore, the association between AF cytotoxicity and PTGR1 expression observed here and previously^{13-17, 28} is consistent with the positive correlation between HMAF cytotoxicity and enone reduction capacity characterized previously in a panel of 60 human cancer cell lines, where DNA repair proficiency differences were not considered.¹⁴

In light of the current results, this correlation should be further considered by also integrating knowledge of the DNA repair proficiencies in those cells. For example, leukemia-derived HL-60 cells are TC-NER proficient,³³ and were relatively insensitive to HMAF. In contrast, ovarian and prostate cancer cells are TC-NER-deficient^{32, 34} and were 100-fold more susceptible to HMAF.¹⁴ It should be noted, however, that ovarian and prostate cancer cells have 10-fold higher enone reductase capacity than leukemia cells,¹⁴ and that enhanced bioactivating capacity, in addition to TC-NER-deficiency, could contribute to the strong activity of HMAF in these cells. Finally, TC-NER-proficient breast cancer cells (MCF7 and MBA-MB-231) were 10-fold more resistant than ovarian and prostate cancers to HMAF, despite being equally efficient in enone-reduction.¹⁴ These relationships suggest that for bioactivated DNA alkylating drugs, establishing both bioactivation capacity and pathway-specific repair proficiency of tumors may be an effective combination of biomarkers for predicting outcome of chemotherapy.

6.6 Conclusion

We established that TC-NER proficiency significantly reduces AF cytotoxicity, and that the synthetic AF analog *iso*-M0, which induces toxicity independent of cellular bioactivation capacity, exhibits a cellular toxicity profile analogous to AF. These data corroborate the role of repair while also considering potential bioactivation influences, demonstrating a chemistry-based approach (i.e. chemically activated drug analog) for dissecting the role of bioactivation from other biochemical factors influencing drug activity. Finally, using a siRNA approach, we show in human cancer cell lines that knock-down of CSB, but not XPC, has a large positive effect on AF activity. This study suggests that future development of AF-related clinical candidates should account for DNA repair and bioactivation-related biomarkers, as well as a possible role for combined therapeutic strategies involving inhibition of TC-NER.

6.7 Supplementary Information

Reverse-transcription real time PCR analysis of PTGR1 expression in WT and XP cells

Untreated WT and XP cells were collected from two flasks of 90-95% confluence as follows: 2 mL of 0.25% trypsin-EDTA was added and the cells were incubated for 5 min at 37 °C. Total mRNA was extracted from using the QIAGEN miRNeasy® mini kit and kept on dry ice until cDNA synthesis (Table S1). Synthesis of first-strand cDNA was performed with the Transcriptor First Strand cDNA Synthesis kit (Roche) according to manufacturer's instructions with random hexamer primers. Expression of prostaglandin reductase 1 (PTGR1, Gene ID: 22949) and of the reference gene porphobilinogen deaminase (PBGD, Gene ID: 3145) was measured with the LightCycler® 480 real-time PCR system and a LightCycler® 480 SYBR Green I master kit (Roche). The analysis was performed in duplicate and relative PTGR1 expression in the different cell lines was calculated by the method of Pfaffl et al.²⁶ The obtained data was normalized with respect to WT cells and is summarized in Figure S1.

Table S1. Gene sequences for real-time PCR analysis^a

Gene	Primer 5' → 3'	T, °C ^b	bp ^c
PTGR1	FW: GGCTCCTGAGCTTCAGGATG	56	140
	RV: GCTTCAAGCAGGACCTCTCCA		
PBGD	FW: CAACGGCGGAAGAAAACAG	56	195
	RV: TCTCTCCAATCTTAGAGAGTG		

^aPTGR1, prostaglandin reductase, PBGD, porphobilinogen deaminase;

^bannealing temperature; ^cproduct length, base-pairs (bp); FW, forward; RV, reverse.

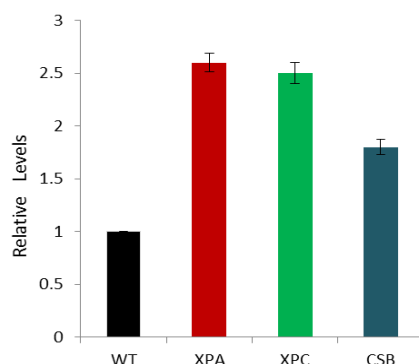


Figure S1. Relative PTGR1 levels established by correlating expression of the PTGR1-coding mRNA in a real-time reverse-transcription PCR analysis. The data is normalized to WT.

Knock down of NER factors in SW480 control and PTGR1 overexpressing cells and treatment with Acylfulvene

The siRNA sequences were purchased from Microsynth. Transfections were performed with Lipofectamine RNAiMAX according to the manufacturer's protocol (Invitrogen). The siRNA concentrations were 10 nM for non-coding siRNA and the silencing of XPC and CSB targets (siRNA sequences are listed in Supplementary Material). Cells were analyzed 72 h after siRNA transfections by Western blot. For treatment with Acylfulvene, cells were seeded 48 h after siRNA transfection in 96-well plates at a density of 3.0×10^3 cells/well and were allowed to attach overnight. Cell viability experiments were initiated by replacing the maintenance media with media containing AF (0, 10, 100, 150, 300, 500, 1000, and 5000 nM final concentration, 0.1-0.5 % DMSO). Cell viability was measured 48 h later with the CellTiter-Glo[®] Luminescent Cell Viability Assay according to the manufacturer's protocol (Promega). This assay determines the number of viable cells in culture based on quantitation of the present ATP as an indicator of metabolically active cells.

Table S2. siRNA sequences

Target	Sequence 5' -> 3'	Source / Reference
Non-coding	AAU UCU CCG AAC GUG UCA CGU	Qiagen, Cat. No. SI03650325
XPC	GCA AAU GGC UUC UAU CGA ATT	Qiagen, Cat. No. SI00066227
CSB	GAA GCA AGG UUG UAA UAA ATT	Microsynth

XPC and CSB levels in SW480 control and SW480 PTGR1 cells after siRNA treatment

Cells were collected after 72 hours of siRNA treatment and lysed in RIPA buffer (Pierce, Thermo Scientific) containing protease inhibitors (Complete Tablets EDTA-free EASYpack, Roche). After cell lysis 40 µg protein of each sample were loaded onto an SDS PAGE gel (4-12 % NuPAGE bis-tris gels, Invitrogen), electrophoresed at 200V for 180min in 1x MOPS running buffer (NuPAGE MOPS SDS Running buffer 20x, Life Technologies) and transferred onto PVDF membranes (Amersham Hybond-P) at 30V for 50min. Staining of NER factors was performed with the following primary antibodies and respective secondary HRP antibodies: Mouse-anti-XPC (Abcam, 1:500), and rabbit-anti-CSB (Abcam, 1:500). All primary antibodies were diluted in 3 % BSA, TBS-Tween. Western blots were developed with Pierce ECL western blotting substrate (Thermo Scientific) and exposure to X-ray film.

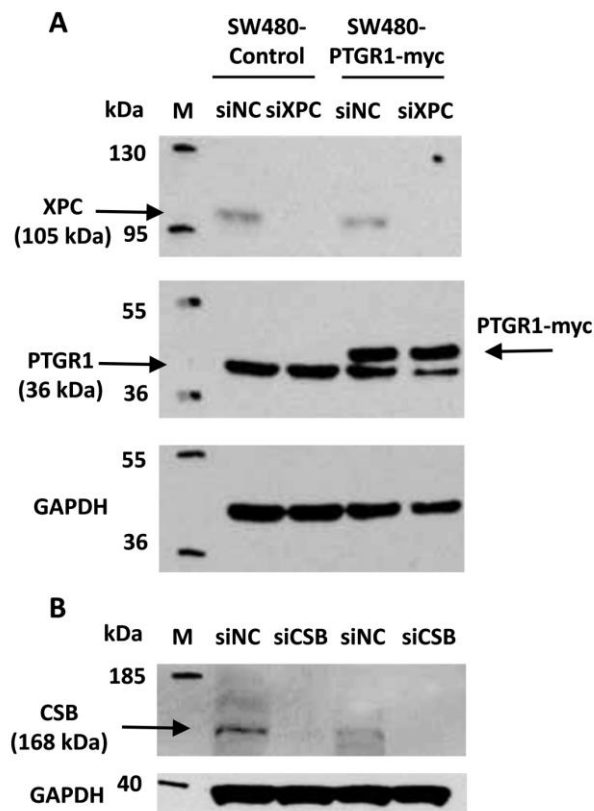


Fig. S2. Silencing of NER factors XPC (A) and CSB (B) in SW480 and SW480-PTGR1 cells. M: Marker; siNC: siRNA with non-coding sequence, siXPC: siRNA targeting XPC, siCSB: siRNA targeting CSB. GAPDH served as loading control.

6.8 References

1. *DNA Repair in Cancer Therapy: Molecular Targets and Clinical Applications*. **2011**.
2. Ripp, S. L., *Advances in Bioactivation Research*. Springer New York: **2008**.
3. Kelner, M. J.; McMorris, T. C.; Estes, L.; Starr, R. J.; Rutherford, M.; Montoya, M.; Samson, K. M.; Taetle, R., Efficacy of Acylfulvene Illudin analogues against a metastatic lung carcinoma MV522 xenograft nonresponsive to traditional anticancer agents: retention of activity against various mdr phenotypes and unusual cytotoxicity against ERCC2 and ERCC3 DNA helicase-deficient cells. *Cancer Res.* **1995**, *55* (21), 4936-40.
4. Tanasova, M.; Sturla, S. J., Chemistry and Biology of Acylfulvenes: Sesquiterpene-Derived Antitumor Agents. *Chem Rev* **2012**, *112* (6), 3578-3610.
5. McMorris, T. C.; Kelner, M. J.; Wang, W.; Diaz, M. A.; Estes, L. A.; Taetle, R., Acylfulvenes, a new class of potent antitumor agents. *Experientia* **1996**, *52* (1), 75-80.
6. Woynarowska, B. A.; Woynarowski, J. M.; Herzig, M. C.; Roberts, K.; Higdon, A. L.; MacDonald, J. R., Differential cytotoxicity and induction of apoptosis in tumor and normal cells by hydroxymethylacylfulvene (HMAF). *Biochem. Pharmacol.* **2000**, *59* (10), 1217-1226.
7. Sato, Y.; Kashimoto, S.; MacDonald, J. R.; Nakano, K., In vivo antitumour efficacy of MGI-114 (6-hydroxymethylacylfulvene, HMAF) in various human tumour xenograft models including several lung and gastric tumours. *Eur J Cancer* **2001**, *37* (11), 1419-1428.
8. Woo, M. H.; Peterson, J. K.; Billups, C.; Liang, H.; Bjornsti, M. A.; Houghton, P. J., Enhanced antitumor activity of irifulven in combination with irinotecan in pediatric solid tumor xenograft models. *Cancer Chemoth Pharm* **2005**, *55* (5), 411-419.
9. Serova, M.; Calvo, F.; Lokiec, F.; Koepfel, F.; Poindessous, V.; Larsen, A. K.; Van Laar, E. S.; Waters, S. J.; Cvitkovic, E.; Raymond, E., Characterizations of irifulven cytotoxicity in combination with cisplatin and oxaliplatin in human colon, breast, and ovarian cancer cells. *Cancer Chemoth Pharm* **2006**, *57* (4), 491-499.
10. MGI PHARMA Stops Phase 3 Irofulven Clinical Trial for Refractory Pancreatic Cancer Patients. In *MGI Pharma* 2002.
11. Institute, M. P., MPI's drug development arm Oncology Venture and Lantern Pharma announce partnership to advance Irofulven for metastatic prostate cancer. **2015**.
12. Beckman, R. A.; Chen, C., Efficient, Adaptive Clinical Validation of Predictive Biomarkers in Cancer Therapeutic Development. *Adv Exp Med Biol* **2015**, *867*, 81-90.
13. Dick, R. A.; Yu, X.; Kensler, T. W., NADPH alkenal/one oxidoreductase activity determines sensitivity of cancer cells to the chemotherapeutic alkylating agent irifulven. *Clin. Cancer Res.* **2004**, *10* (4), 1492-9.
14. Neels, J. F.; Gong, J.; Yu, X.; Sturla, S. J., Quantitative correlation of drug bioactivation and deoxyadenosine alkylation by acylfulvene. *Chem. Res. Toxicol.* **2007**, *20* (10), 1513-1519.
15. Gong, J.; Neels, J. F.; Yu, X.; Kensler, T. W.; Peterson, L. A.; Sturla, S. J., Investigating the role of stereochemistry in the activity of anticancer acylfulvenes: synthesis, reductase-mediated bioactivation, and cellular toxicity. *J. Med. Chem.* **2006**, *49* (8), 2593-2599.
16. Gong, J.; Vaidyanathan, V. G.; Yu, X.; Kensler, T. W.; Peterson, L. A.; Sturla, S. J., Depurinating acylfulvene-DNA adducts: characterizing cellular chemical reactions of a selective antitumor agent. *J. Am. Chem. Soc.* **2007**, *129* (7), 2101-2111.
17. Pietsch, K. E.; Neels, J. F.; Yu, X.; Gong, J.; Sturla, S. J., Chemical and enzymatic reductive activation of acylfulvene to isomeric cytotoxic reactive intermediates. *Chemical research in toxicology* **2011**, *24* (11), 2044-54.

18. Pietsch, K. E.; van Midwoud, P. M.; Villalta, P. W.; Sturla, S. J., Quantification of acylfulvene- and illudin S-DNA adducts in cells with variable bioactivation capacities. *Chemical research in toxicology* **2013**, *26* (1), 146-55.
19. Escargueil, A. E.; Poindessous, V.; Soares, D. G.; Sarasin, A.; Cook, P. R.; Larsen, A. K., Influence of irofulven, a transcription-coupled repair-specific antitumor agent, on RNA polymerase activity, stability and dynamics in living mammalian cells. *Journal of cell science* **2008**, *121* (Pt 8), 1275-1283.
20. Jaspers, N. G.; Raams, A.; Kelner, M. J.; Ng, J. M.; Yamashita, Y. M.; Takeda, S.; McMorris, T. C.; Hoeijmakers, J. H., Anti-tumour compounds illudin S and Irofulven induce DNA lesions ignored by global repair and exclusively processed by transcription- and replication-coupled repair pathways. *DNA repair* **2002**, *1* (12), 1027-38.
21. Koepfel, F.; Poindessous, V.; Lazar, V.; Raymond, E.; Sarasin, A.; Larsen, A. K., Irofulven cytotoxicity depends on transcription-coupled nucleotide excision repair and is correlated with XPG expression in solid tumor cells. *Clin. Cancer Res.* **2004**, *10* (16), 5604-5613.
22. van Midwoud, P. M.; Sturla, S. J., Improved efficacy of acylfulvene in colon cancer cells when combined with a nuclear excision repair inhibitor. *Chem Res Toxicol* **2013**, *26* (11), 1674-82.
23. Svejstrup, J. Q., Mechanisms of transcription-coupled DNA repair. *Nat. Rev. Mol. Cell Biol.* **2002**, *3* (1), 21-29.
24. Hanawalt, P. C.; Spivak, G., Transcription-coupled DNA repair: two decades of progress and surprises. *Nature reviews. Molecular cell biology* **2008**, *9* (12), 958-70.
25. McCloud, T. G.; Klueh, P. A.; Pearl, K. C.; Cartner, L. K.; Muschik, G. M.; Poole, K. K., Isolation of illudin S from the mature basidiocarp of *Omphalotus illudens*, and isolation of illudins S and M from the fermentation broth of *Omphalotus olearis*. *Nat Prod Lett* **1996**, *9* (2), 87-95.
26. Pfaffl, M. W., A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic acids research* **2001**, *29* (9), e45.
27. Spivak, G.; Hanawalt, P. C., Host cell reactivation of plasmids containing oxidative DNA lesions is defective in Cockayne syndrome but normal in UV-sensitive syndrome fibroblasts. *DNA repair* **2006**, *5* (1), 13-22.
28. Pietsch, K. E.; van Midwoud, P. M.; Villalta, P. W.; Sturla, S. J., Quantification of acylfulvene- and illudin S-DNA adducts in cells with variable bioactivation capacities. *Chem. Res. Toxicol.* **2012**, *in revision*.
29. Sidorenko, V. S.; Yeo, J. E.; Bonala, R. R.; Johnson, F.; Scharer, O. D.; Grollman, A. P., Lack of recognition by global-genome nucleotide excision repair accounts for the high mutagenicity and persistence of aristolactam-DNA adducts. *Nucleic Acids Res* **2012**, *40* (6), 2494-505.
30. Kucab, J. E.; Phillips, D. H.; Arlt, V. M., Linking environmental carcinogen exposure to TP53 mutations in human tumours using the human TP53 knock-in (Hupki) mouse model. *The FEBS journal* **2010**, *277* (12), 2567-83.
31. Lukin, M.; Zaliznyak, T.; Attaluri, S.; Johnson, F.; de Los Santos, C., Solution structure of duplex DNA containing a beta-carba-Fapy-dG lesion. *Chem Res Toxicol* **2012**, *25* (11), 2423-31.
32. Stubbert, L. J.; Smith, J. M.; McKay, B. C., Decreased transcription-coupled nucleotide excision repair capacity is associated with increased p53- and MLH1-independent apoptosis in response to cisplatin. *BMC cancer* **2010**, *10*, 207.
33. Hsu, P. H.; Hanawalt, P. C.; Nospikel, T., Nucleotide excision repair phenotype of human acute myeloid leukemia cell lines at various stages of differentiation. *Mutation research* **2007**, *614* (1-2), 3-15.
34. Takebayashi, Y.; Nakayama, K.; Kanzaki, A.; Miyashita, H.; Ogura, O.; Mori, S.; Mutoh, M.; Miyazaki, K.; Fukumoto, M.; Pommier, Y., Loss of heterozygosity of nucleotide excision repair factors in sporadic ovarian, colon and lung carcinomas: implication for their roles of carcinogenesis in human solid tumors. *Cancer Lett* **2001**, *174* (2), 115-125.

Chapter 7 - Appendix B: Mechanism of RNA polymerase II stalling by DNA alkylation

Stefano Malvezzi^{a‡}, Lukas Farnung^{b‡}, Claudia M.N. Aloisi^a, Todor Angelov^a, Patrick Cramer^b, Shana J. Sturla^{a*}

^aDepartment of Health Sciences and Technology, ETH Zurich, Zurich, Switzerland

^bMax Planck Institute for Biophysical Chemistry, Am Fassberg 11, 37077 Göttingen, Germany

Manuscript published in PNAS 2017, 46, 114, 12172-12177

C.M.N devised and performed cell studies, interpreted data and wrote the relevant experimental and result section. S. Malvezzi synthesized DNA oligonucleotides, performed enzymatic experiments, interpreted data and wrote the manuscript. L. Farnung purified RNA Pol II, determined the crystal structure and interpreted data. T. Angelov synthesized DNA oligonucleotides. P. Cramer interpreted data. S.J. Sturla interpreted data and wrote the manuscript.

7.1 Abstract

Several anticancer agents form DNA adducts in the minor groove, interfering with DNA replication and transcription, and inducing apoptosis. Therapeutic resistance can occur, however, when cells are proficient in the removal of drug-induced damage. Acylfulvenes are a class of experimental anticancer agents with a unique repair profile suggesting their capacity to stall RNA polymerase (Pol) II and trigger transcription-coupled nucleotide excision repair. Here we show how different forms of DNA alkylation impair transcription by RNA Pol II in cells and with the isolated enzyme, and unravel a new mode of RNA Pol II stalling that is due to alkylation of DNA in the minor groove. We incorporated a model for acylfulvene adducts, the stable 3-deaza-3-methoxynaphtylethyl-adenosine analog (3d-Napht-A), and smaller 3-deaza-adenosine analogs, into DNA oligonucleotides to assess RNA Pol II transcription elongation *in vitro*. RNA Pol II was strongly blocked by a 3d-Napht-A analog but bypassed smaller analogs. Crystal structure analysis revealed that a DNA base containing 3d-Napht-A can occupy the +1 templating position and impair closing of the trigger loop in the Pol II active center and polymerase translocation into the next template position. These results show how RNA Pol II copes with minor groove DNA alkylation and establish a mechanism for drug resistance.

7.2 Introduction

The catalysis of transcription by RNA polymerase is fundamental to the viability of growing cells. In conventional cancer chemotherapy, DNA damage products that interfere with genomic processes initiate cell death. However, therapeutic resistance of cancer cells to DNA alkylating agents may result when cells are proficient in the repair of drug-induced damage.¹⁻³ Acylfulvenes (AFs) are a class of experimental anticancer drugs that appear to be selectively repaired by the transcription-coupled sub-pathway of nucleotide excision repair (TC-NER).⁴⁻⁶ Thus, AF-induced DNA adducts are preferentially removed from actively transcribed regions of the genome but are largely ignored by global-genome repair. These findings suggest that the AF-DNA adduct does not perturb the duplex structure, yet impedes the progress of RNA polymerase. Understanding the chemical basis of how DNA alkylation stalls RNA polymerase to initiate repair can provide key insight for the design of therapeutics less prone to failure due to resistance.

AFs are semi-synthetic derivatives of the fungal sesquiterpene illudin S and alkylate DNA in the minor groove,⁷ a general profile for several anticancer agents such as distamycins, lexitropsins, duocarmycins, and ecteinascidin 743, as well as peptide-based minor groove binders.⁸ Hydroxymethylacylfulvene (HMAF, irofulven) was tested previously in clinical trials for several cancers and is anticipated to re-enter clinical trials with biomarker-driven patient

stratification strategies.⁹ Acylfulvenes are pro-drugs that are reductively activated by prostaglandin reductase 1 (PTGR1) to alkylate primarily position 3 of adenine (Figure 1a).⁷ Inhibition of DNA synthesis and capacity to induce cell cycle arrest of minor-groove alkylating agents is attributed primarily to alkylation of DNA, and levels of 3-AF-A in AF-treated cancer cells are proportional to cytotoxicity.^{7, 10-13}

Cells have evolved various DNA repair functions, and large helix-distorting adducts are often removed by nucleotide excision repair (NER). NER is divided into two sub-pathways: transcription coupled- (TC-) and global genome- (GG-) NER.¹⁴ TC-NER-deficient human fibroblast cells are more sensitive to illudin S, AF and HMAF compared to GG-NER-deficient human fibroblast cells, suggesting that TC-NER selectively repairs AF adducts (Figure 1a).⁴⁻⁶ Moreover, siRNA-mediated down-regulation of TC-NER in a cancer cell line greatly increased their sensitivity to AF, whereas down-regulation of GG-NER did not (Figure 1a).⁶ Since AF adducts are selectively repaired by TCR and ignored by GG-NER, a large portion of adducts that are not repaired in surviving cells are expected to be a basis of the increased mutation frequency observed in cell lines.¹⁵ Indeed, error-prone bypass of minor groove modifications has been characterized for certain translesion synthesis DNA polymerases.¹⁶ Finally, exposing a cancer cell line to UCN-01, a compound that inhibits Chk1, enabling checkpoint evasion and reduction of NER function, increased sensitivity to AF and AF adducts persisted longer.^{12, 17} The repair and persistence profiles for AF-DNA adducts thus suggests that the basis of selective removal is associated with the DNA damage recognition step.

The main difference between the TC- and GG-sub-pathways of NER is the DNA damage recognition step: GG-NER is activated by XPC-RAD23B factors that sense the altered helix conformation induced by DNA damage, whereas TC-NER is activated by the stalling of RNA polymerase II (Pol II).^{14, 18} Therefore, a working hypothesis is that 3-AF-A stalls Pol II but does not induce significant helix distortion.⁶ Pol II efficiently transcribes over small DNA lesions like *O*⁶-methyl-guanine (*O*⁶-Me-G), 8-oxo-guanine (8-oxo-G), N2-1-carboxyethyl-guanine and N3-carboxymethyl-thymine, among others,¹⁹⁻²² however, larger modifications can induce it to stall. Upon Pol II stalling, downstream factors are recruited to the damage site, and the repair process then progresses in the same manner as for GG-NER.^{14, 18}

The mechanism of Pol II stalling at several DNA lesions has been elucidated. Cyclobutane pyrimidine dimers (CPDs), cisplatin 1,2-d(GPG) intrastrand crosslinks, monofunctional pyrriplatin-guanine adducts and 8,5'-cyclo-2'-deoxyadenosine strongly inhibit Pol II progression and crystal structures of yeast Pol II bound to DNA containing these lesions were solved to elucidate the basis of polymerase stalling or bypass.²³⁻²⁸ Two flexible elements of the Pol II active site, the bridge helix and the trigger loop, are required for nucleotide addition and nucleic acid translocation to the next template position.²⁹⁻³² Wang and co-workers reported the blockage of Pol II by minor groove binding pyrrole-imidazole polyamides and identified their interaction with

the residues Arg1386 and His1387 in the conserved switch 1 region of Rpb1 by molecular modelling.³³ Despite these insights, there is no direct observation of Pol II encountering a DNA template with minor groove adducts, and it is unknown how Pol II deals with minor-groove-binding and -modifying agents.

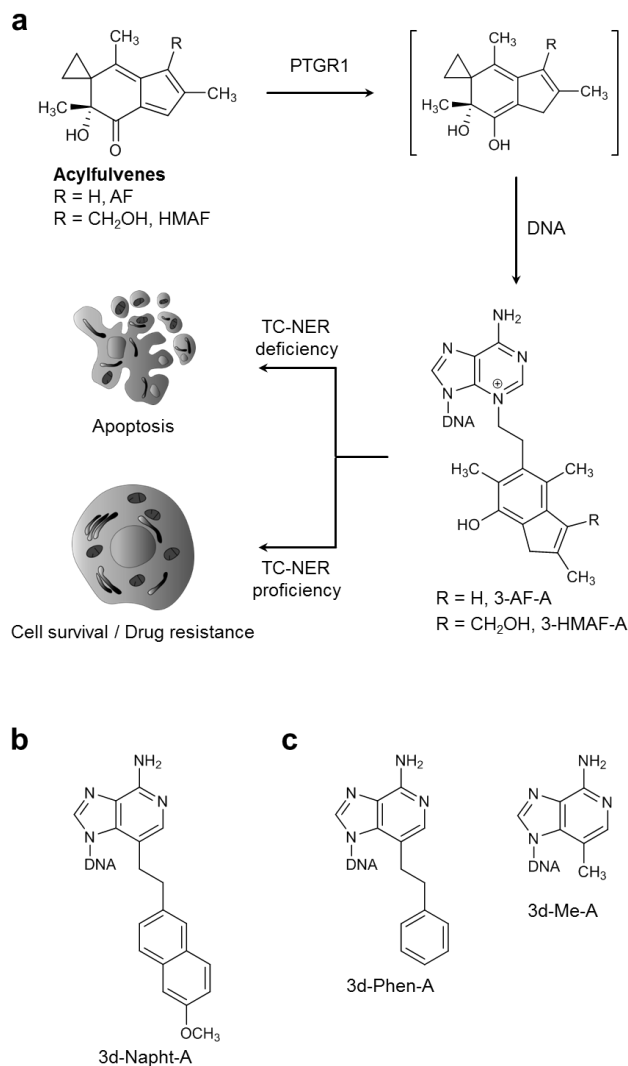


Figure 1 Mechanism of minor groove DNA adduct formation from acylfulvenes and structures of the minor groove adduct models used in this study. a) Acylfulvenes are bioactivated in the cytosol by the reductase PTGR1 to form a reactive intermediate that preferentially alkylates position 3 of adenine in genomic DNA. The cytotoxicity of acylfulvenes is dependent on the DNA repair proficiency of exposed cells, with cells deficient in TC-NER being more susceptible than cells proficient in TC-NER. b) Structure of the 3-deaza-3-methoxynaphtylethyl-adenosine analog (3d-Napht-A) incorporated in oligonucleotides and used as model of acylfulvene-derived DNA adducts in biochemical and crystallographic studies. c) Structure of 3-deaza-adenosine analogs

of decreasing size, 3-deaza-3-phenethyl-adenosine (3d-Phen-A) and 3-deaza-3-methyl-adenosine (3d-Me-A) used to enable structure-activity analysis of Pol II.

In this study, the question of how minor groove DNA alkylation impedes RNA synthesis was addressed. Evaluating transcription activity in a human cancer cell confirmed the inhibited RNA synthesis. Due to the chemical instability of the native AF adduct,^{7, 12} we characterized the behavior of Pol II in transcription over its stable analog 3-deaza-3-methoxynaphtylethyl-adenosine (3d-Napht-A) (Figure 1b).¹⁶ We tested the capacity of 3d-Napht-A to stall purified RNA Pol II with DNA constructs mimicking the transcription bubble and ternary elongation complexes, and tested the limits of Pol II's size tolerance of minor groove alkyl adducts by carrying out the same investigation with other 3-deaza-alkyl-adenosine analogs of systematically smaller size, i.e. 3-deaza-3-phenethyl-adenosine (3d-Phen-A) and 3-deaza-3-methyl-adenosine (3d-Me-A) (Figure 1c). The mechanism of stalling was furthermore investigated by crystallographic analysis of the stalled Pol II in the ternary elongation complex with bound DNA template and RNA transcript. These data reveal a previously unobserved mode of Pol II stalling and provide insights into the chemical topology of minor groove modifications required to stall transcription in cancer cells.

7.3 Results

7.3.1 DNA alkylation impairs RNA synthesis in cells

It has been established that cells proficient in TC-NER are more resistant to cytotoxicity induced by AFs.⁴⁻⁶ To evaluate how AF impacts RNA synthesis in cancer cells relative to a simple methylating agent, we compared RNA synthesis activity following treatment with HMAF vs. MMS. MMS can form various DNA methylation adducts, mainly the non-cytotoxic 7-Me-G, but also 3-Me-A, and to a lesser extent *O*⁶-Me-G.^{34, 35} A significant reduction of RNA synthesis in colon adenocarcinoma SW480 cells overexpressing PTGR1 (SW480-PTGR1), an enzyme required for metabolic bioactivation of AFs, was observed after treatment with 1 μ M HMAF, but not 1 μ M MMS (Supplementary Figure 2), concentrations at which all cells remained viable (Supplementary Figure 1). Because MMS is much less potent than HMAF, we also evaluated RNA synthesis at doses that reduce cell viability equally with both compounds (Supplementary Figure 3). Again, RNA synthesis was significantly reduced in cells treated with 0.5 μ M HMAF, but not in cells treated with 500 μ M MMS. These results provide evidence for Pol II stalling after minor groove DNA alkylation in cancer cells.

7.3.2 Pol II transcription can stall at minor-groove alkylation adducts

With strong indication that AF-DNA damage stalls transcription in cells, we assessed directly the capacity of the major 3-AF DNA adduct to impede the progress of RNA Pol II using purified

yeast Pol II and DNA that contains 3d-Napht-A, a chemically stable analog of the adduct,¹⁶ in constructs of a transcription bubble with 3d-Napht-A four bases downstream of the growing RNA 3'-end (Supplementary Figure 4a).³⁶ When RNA was extended in the presence of four NTPs, Pol II stalled after nucleotide incorporation opposite the 3d-Napht-A, whereas it transcribed efficiently over unmodified A or the smaller methylated analog 3d-Me-A (Supplementary Figure 4b). RNA extension of a transcription bubble with 3d-Napht-A nine bases downstream of the RNA 3'-end also behaved similarly (Supplementary Figure 4c, d).

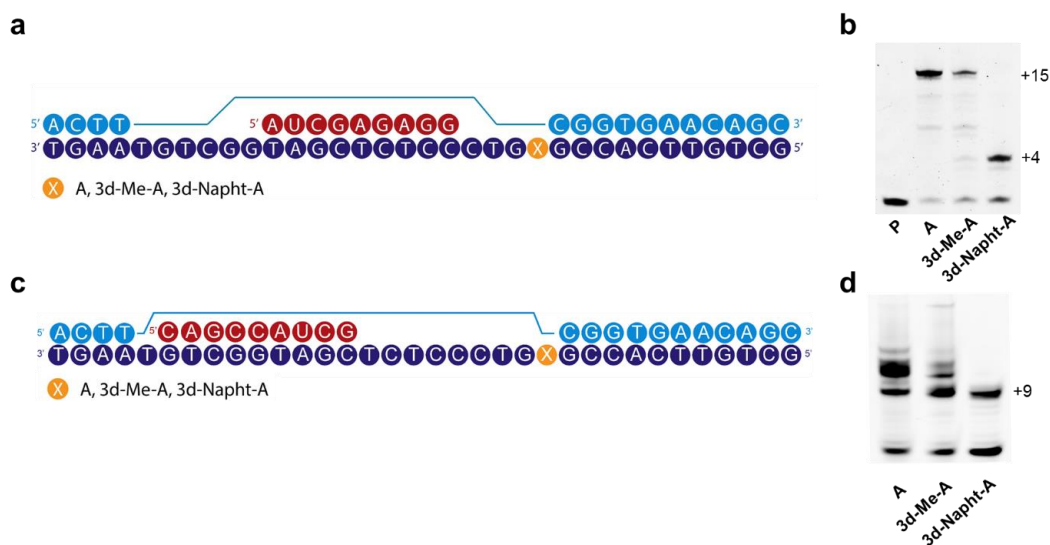


Figure 2 Pol II transcription with DNA constructs mimicking the transcription bubble. a, c) DNA constructs used for testing the impact of 3d-alkyl-A analogs on the propensity of Pol II to form run-off transcripts. DNA template (dark blue) and RNA primer (red) were annealed, then allowed to react with Pol II (10 min at 25°C). The non-templating DNA strand (light blue) was then annealed to create the transcription bubble. Alkylated adenine analogs were placed four nucleotides (a, b) or nine nucleotides (c, d) downstream of the transcription start on the DNA template. b, d) Denaturing gel electrophoresis representing Pol II transcription over A, 3d-Me-A and 3d-Napht-A in presence of all four NTPs (1mM). Reactions were quenched after 25 min. The site of the modified base is indicated with an x.

We next characterized RNA extension with ternary elongation complexes composed of an RNA annealed to a DNA template containing the analog at the nucleotide incorporation site +1 (Figure 2a). In the presence of all four NMPs, Pol II stalled at 3d-Napht-A after incorporating one NMP (Figure 2b) and no bypass was observed even after 60 min. When the same reactions were carried out with individual nucleotides, Pol II incorporated UMP and misincorporated CMP opposite 3d-Napht-A but not GMP and AMP (Figure 2c).

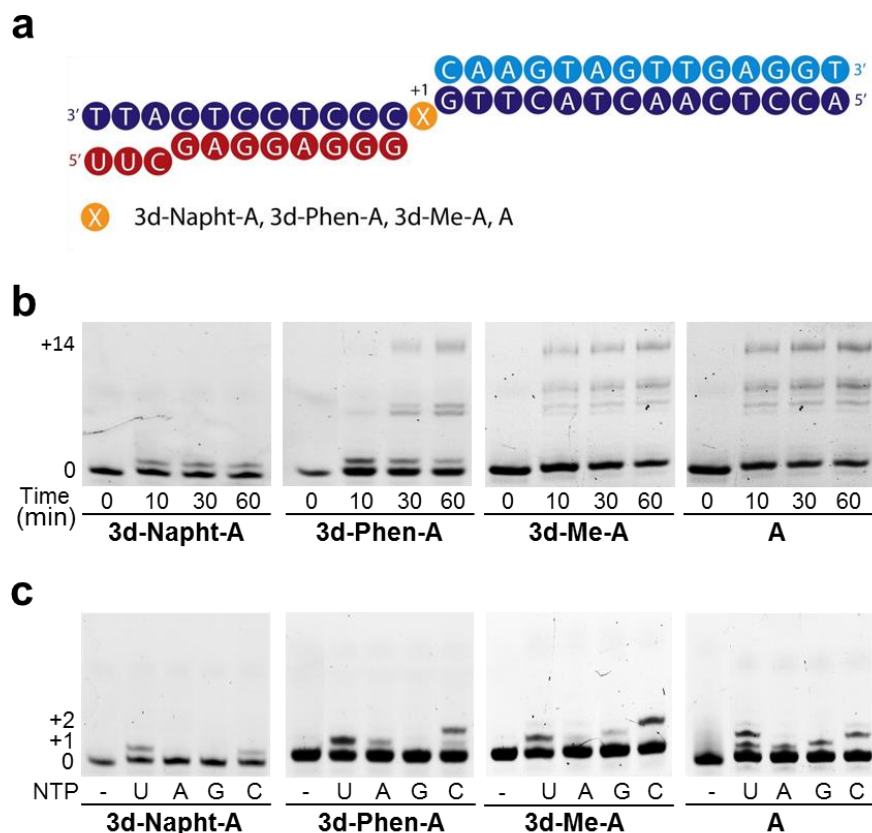


Figure 3 Pol II full-length synthesis and single nucleotide incorporation. a) Ternary elongation complexes containing A, 3d-Me-A, 3d-Phen-A or 3d-Napht-A at the underlined position, used for primer extension reactions. DNA template is dark blue, non-template DNA is light blue, RNA primer is red. b) Denaturing gel electrophoresis of Pol II transcription products in the presence of all four NTPs (1mM). Extension of the RNA primer to the end of the DNA template forms the indicated run-off product (+14). c) Single nucleotide incorporation opposite A, 3d-Me-A, 3d-Phen-A and 3d-Napht-A. Reactions were performed in presence of a single NTP (1mM) and were quenched after 20 min.

7.3.3 Pol II is tolerant to smaller modifications but errors arise

To derive the maximum adduct size tolerated by Pol II, RNA extension was evaluated for templates containing smaller alkyl groups. Pol II efficiently extended RNA annealed with templates containing A or 3d-Me-A, forming full-length products (Figure 2b). 3d-Phen-A was bypassed, but product accumulated after incorporation of the first nucleotide, suggesting the intermediate size adduct slows Pol II but does not impede it completely.

In the presence of individual nucleotides, UMP was very efficiently incorporated, leading to a +2 band resulting from insertion of UMP opposite A plus the next G. For modified substrates

with 3d-Me-A or 3d-Phen-A, however, only the +1 band was observed for UMP, suggesting that the 3-alkyl groups reduced transcription fidelity. A similar degree of CMP misincorporation was observed in all cases, regardless of the DNA being modified or not (Supplementary Figure 5). Moreover, Pol II efficiently misincorporated CMP opposite the next templating G forming a +2 band product, whereas in the case of 3d-Napht-A, only one nucleotide was inserted opposite the analog.

7.3.4 DNA alkylation alters Pol II incorporation and extension efficiency

Having established that Pol II incorporates UMP and CMP to varying extents depending on the identity of DNA modification, we measured apparent reaction rates for their incorporation by Pol II (Figure 3). For insertion of UTP opposite A or 3d-Me-A, incorporation rates were similar ($9.0 \pm 1.4 \text{ min}^{-1}$ and $8.0 \pm 1.2 \text{ min}^{-1}$ respectively), whereas rates opposite 3d-Phen-A and 3d-Napht-A were significantly reduced ($2.5 \pm 0.2 \text{ min}^{-1}$ and $1.2 \pm 0.2 \text{ min}^{-1}$, respectively). The CMP misincorporation profile was quite different, however, with more than 10-fold greater incorporation opposite 3d-Me-A vs. A ($5.8 \pm 0.4 \text{ min}^{-1}$ vs. $0.5 \pm 0.1 \text{ min}^{-1}$), whereas 3d-Phen-A and 3d-Napht-A were similarly low as A ($0.28 \pm 0.03 \text{ min}^{-1}$ and $0.39 \pm 0.05 \text{ min}^{-1}$) (Figure 3a). The increased misincorporation of CMP opposite the 3-alkyl-adenosine analogs relative to A (Figure 3a) suggests a potential basis of transcriptional mutagenesis (TM) in cells.^{21, 37}

To understand the impact of 3-deaza-3-alkyl-adenosine analogs on Pol II activity during the extension steps following incorporation opposite an adduct, we evaluated RNA synthesis using DNA constructs with the RNA transcript containing either U or C opposite 3-deaza-3-alkyl-adenosine analogs and measured rates of CMP incorporation opposite the next templating G (Figure 3b). Extension after 3d-Me-A was similar to after A, but after 3d-Phen-A, extension from U or C was 170- or 250-fold lower, respectively, and after 3d-Napht-A, extension by Pol II was totally blocked, with no incorporation even after 60 min. This observation was consistent with the initially observed stalling after insertion of one nucleotide opposite the modified base. From these data, it could be concluded that Pol II is stalled by the larger analog 3d-Napht-A, whereas it bypasses smaller analogs with the efficiency decreasing in step with the size of the 3-alkyl group size. The efficient bypass of 3d-Me-A was consistent with the observation that exposing cells to MMS did not reduce transcription.

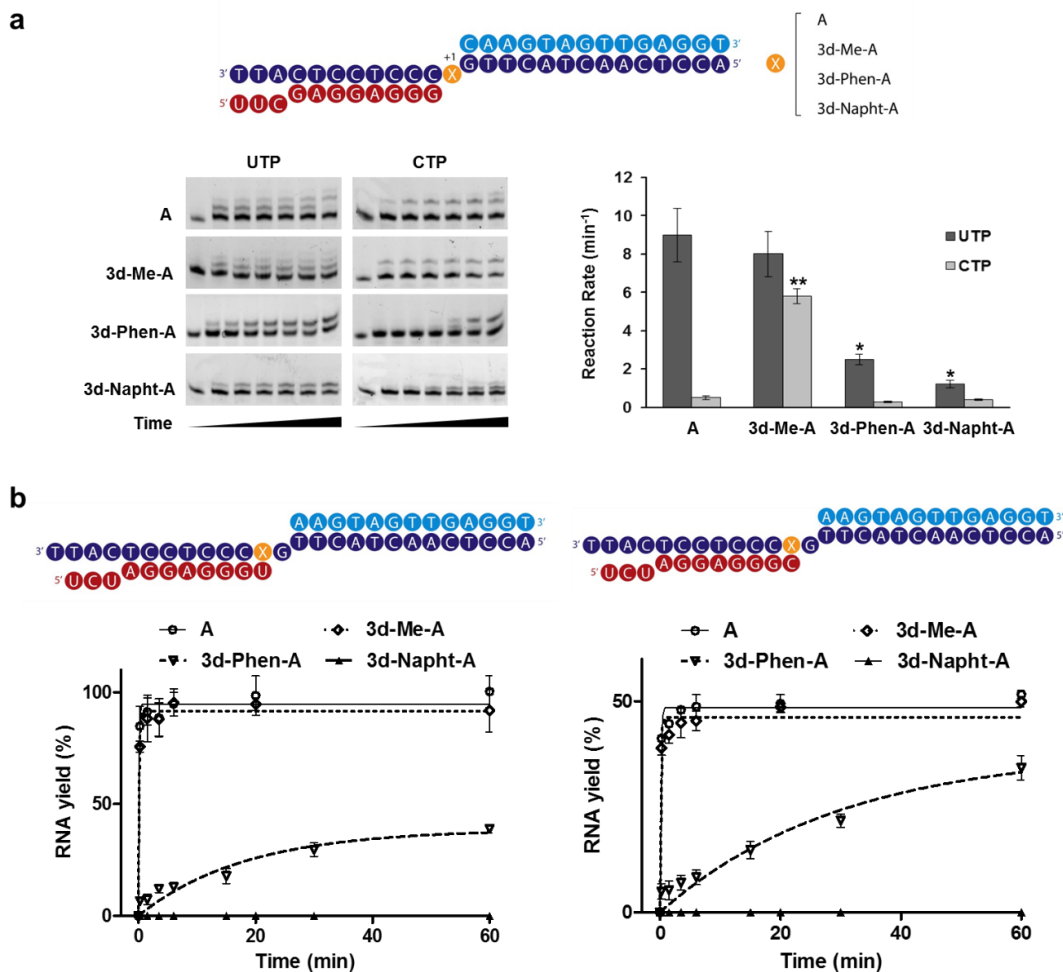


Figure 4 Single nucleotide incorporation and extension kinetics. a) Gel denaturing electrophoresis of primer extension kinetics by Pol II for UMP and CMP incorporation opposite unmodified adenine or 3-deaza-adenosine analogs (left). Reactions contained 1 mM NTP and were stopped at various time points between 0 and 60 min. Schematic representation of apparent rate constants for UMP and CMP incorporation kinetics (right). Error bars represent standard deviation (N=3) and statistical significance for NMP incorporation opposite the 3-deaza-3-alkyl-adenosine adducts compared to A was determined by unpaired Student t-test with Welch's correction (* $p < 0.05$, ** $p < 0.01$). b) Rates (min^{-1}) of RNA primer extension from the correct base pair $A_{\text{analog}}:U$ (left) and the mismatched base pair $A_{\text{analog}}:C$ (right). The DNA template contained either A, 3d-Me-A, 3d-Phen-A or 3d-Napht-A at the indicated position. RNA yield is plotted as a function of time (100% RNA yield corresponds to 50% of primer extension). Error bars represent standard deviation (N=3).

7.3.5 Structural investigation of Pol II stalling by DNA alkylation

To investigate the mechanism of Pol II stalling opposite an alkylated template base, we determined the structure of a Pol II elongation complex in the presence of 3d-Napht-A. We reconstituted a 12-subunit *S. cerevisiae* Pol II elongation complex (EC) with a modified nucleic acid scaffold (Figure 4a). The resulting complex was crystallized, and the structure was determined using molecular replacement. Electron density was observed for 8 bp of downstream DNA, for the DNA template strand up to upstream position -10, and for the entire RNA except for the 5'-terminal base at the upstream end of the DNA-RNA hybrid (Figure 4b). The register of nucleic acids was unambiguously defined by bromine labelling of the DNA template strand at position -4 and anomalous diffraction (Figure 4c). The DNA modification was clearly observed in the electron density (Figure 4c). The structure was refined to a free R factor of 22% at a resolution of 3.2 Å, resulting in an atomic model with very good stereochemistry (Supplementary Table 1).

The structure revealed a previously unobserved mode of Pol II interaction with a modified DNA nucleotide. The EC adopts the post-translocation state with the modified adenine base being accommodated in the templating position opposite of the +1 site that binds the nucleoside triphosphate substrate (Figure 4b). The 3-deaza-3-methoxynaphtylethyl group contacts the central bridge helix in the Pol II active center from underneath. In particular, the second aromatic ring of 3d-Napht-A forms van der Waals contacts with the side chain methyl group of bridge helix residue Thr831. The trigger loop adopts an open conformation, similar to that observed in a previous EC structure that adopts the post-translocation state (PDB 1Y1W).

The structure immediately suggests the mechanism of impaired Pol II progression and stalling at an alkylated template base and explains the structure-activity profiles observed herein. Superposition of a Pol II EC structure containing a closed trigger loop (PDB 2E2H)³⁸ showed that the closed state cannot be accommodated in the presence of the alkylated nucleotide due to clashes between trigger loop residues Thr1080 and Leu1081 and the 3-deaza-3-methoxynaphtylethyl group (Figure 4d). Inefficient nucleotide incorporation can apparently still take place, because catalysis can occur in a trigger loop-independent, low-fidelity fashion,³⁹ explaining the observed misincorporations. After a single (mis)incorporation event, translocation is strongly impaired, however, as this would result in major clashes of the 3d-Napht-A moiety with Pol II residues that very intimately interact with the minor groove of the hybrid base pair in the post-translocation position -1. In addition, movement of the bridge helix, also required for translocation,⁴⁰ is likely impaired by the observed alkyl-bridge helix contact.

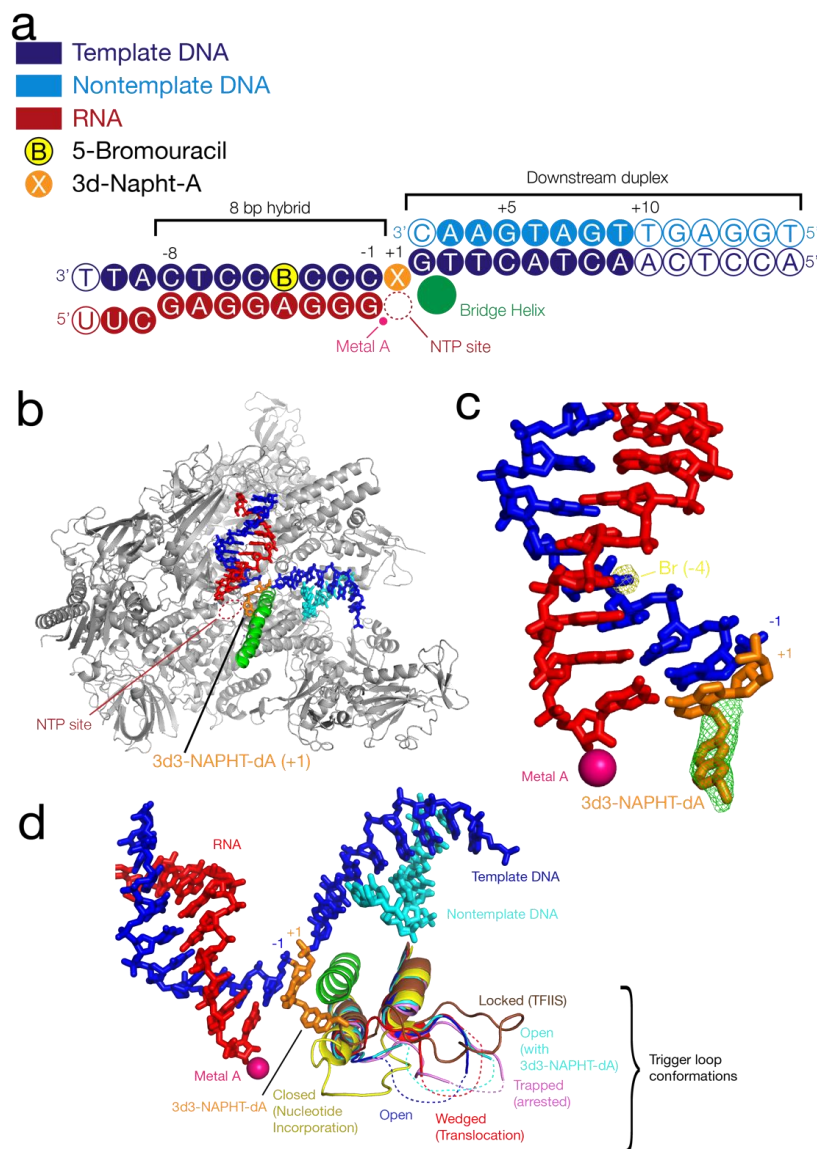


Figure 5 Structure of a Pol II elongation complex in the presence of 3d-Napht-A in the +1 site. (a) The nucleic acid scaffold used in the crystal structure is depicted schematically with respect to the +1 site containing the 3d-Napht-A base modification. The colour scheme is used throughout the figure. (b) Overview of the Pol II elongation complex structure with the modified adenine present in the +1 site, the nucleotide addition site. Pol II is shown as a silver ribbon model in a side view. The bridge helix is coloured in green. Nucleic acids are presented as stick models. 3d-Napht-A is highlighted in orange. (c) Simulated annealing omit map after removing the 3d-Napht moiety at the +1 site from the model ($F_o - F_c$ map, contoured at 3σ , positive density is coloured in green). A peak in the anomalous difference Fourier map (density is coloured in yellow, contoured at 6σ) reveals the exact position of the bromine atom at position -4, allowing

unambiguous assignment of the post-translocated state. (d) Trigger loop conformation. Comparison of the conformation of the trigger loop observed here (cyan) with alternative conformations observed previously for Pol II. The closed trigger loop (PDB 2E2H), open trigger loop in the post-translocation state (PDB 1Y1W), wedged trigger loop (PDB 2VUM), trapped trigger loop (PDB 3PO2), and locked trigger loop (PDB 3PO3) are depicted in yellow, blue, red, violet, and brown, respectively. Metal A is shown as a magenta sphere.

7.4 Discussion

Here we report on the mechanism of Pol II stalling in response to minor groove DNA alkylation relevant to minor-groove-alkylating anticancer agents. A reduction in RNA synthesis was observed when cancer cells were treated with the drug HMAF, but not when treated with a methylating agent, suggesting the capacity of the minor groove HMAF modification, but not a methyl group, to stall RNA Pol II. Thus, RNA extension reactions were performed with DNA constructs containing stable synthetic analogs that model drug alkylation products, a novel approach that overcomes the limitation of chemical instability that has precluded previous detailed evaluation of such modification. The large analog 3d-Napht-A, corresponding to the main adduct from HMAF, was found to block the progress of Pol II after insertion of a nucleotide (Figure 2b). The rate of UMP incorporation was reduced, with background CMP misincorporation being maintained (Figure 3a). Moreover, further extension of RNA primers was completely inhibited. Finally, structural studies indicated that the mechanism of stalling differs from those defined previously at other DNA lesions, revealing a combination of interactions with the base pair at position -1 and the second ring of the modification with the trigger loop of RNA Pol II as the basis.

TC-NER-deficient cells have increased sensitivity to AFs, suggesting the capacity of the adducts to stall transcription.⁴⁻⁶ HMAF exposure of SW480 cells overexpressing PTGR1 led to a significant reduction in RNA synthesis capacity, consistent with previous observations concerning reduced RNA synthesis in human leukemia and HeLa cells, as well as increased degradation of stalled Pol II.^{4, 13, 41} Moreover, we found that the methylating agent MMS did not reduce RNA synthesis in cells (Supplementary Figure 2), consistent with the observation of transcription over 3d-Me-A.⁴² While MMS does not form 3-Me-A exclusively, its toxicity is mainly attributed to this adduct, and the use of an extremely high MMS dose, i.e. 500 μ M, still led to no reduction in RNA synthesis. The 3-Me-A adduct is also formed by the minor groove-binding drug methyl lexitropsin (Me-lex), for which higher drug sensitivity was observed in human glioma cells deficient in BER rather than TC-NER.⁴³⁻⁴⁶

The stalling of isolated Pol II at 3d-Napht-A observed in this study supports the assertion that in cells treated with AFs, transcription is stalled and the stalled Pol II initiates TC-NER and

reduces drug action. The chemical basis of this cellular process appears to be associated with the alkyl adduct preventing translocation of the nascent +1 base pair to the next template position due to clashes with Pol II residues binding the minor groove edge of the DNA-RNA base pair at position -1, and by a clash of the second ring on the methoxynaphtyl group with two residues of the mobile trigger loop (Figure 4). These interactions highlight a different Pol II stalling mechanism than predicted by Dervan and Wang in the case of polyamide minor-groove binding agents, wherein Pol II was stalled upstream of the polyamide-bound DNA by interactions between the Switch 1 region and the minor-groove binder impeding translocation.³³ The structure analysis and structure-activity relationships elucidated in the current study suggests the smaller phenethyl group of 3d-Phen-A does not prevent the trigger loop from closing, allowing Pol II to bypass analogs with only one aromatic ring. Finally, Pol II preferentially incorporated UMP opposite the smaller analogs 3d-Me-A and 3d-Phen-A, and RNA could be extended.

The mechanism of Pol II stalling reported herein should inform the design of improved anticancer alkylating agents that inhibit DNA synthesis without impeding the progression of RNA polymerase, thus evading TC-NER. Based on our results, 3-adenosine adducts with only one ring are tolerated by Pol II but do block the replicative DNA polymerase hPol α .¹⁶ Moreover, the chemical modifications did not destabilize DNA duplexes, suggesting that 3-AF-A adducts do not induce helix distortion and evade GG-NER.¹⁶ Thus, 3-adenosine adducts containing one phenyl ring may have a desired balance of properties in order to inhibit DNA replication but avoid TC-NER. The detailed insight derived from this study reveals how clashes between DNA adducts and residues binding the DNA-RNA minor groove edge and the mobile trigger loop of RNA Pol II prevent translocation after insertion of a nucleotide opposite an adduct. Knowledge concerning structural characteristics that impede Pol II improves our understanding of how cells initiate the repair of damaged DNA and may also support the design of more effective cancer therapeutics.

7.5 Material and Methods

Material

Chemicals, reagents, and electrophoresis supplies were from Sigma-Aldrich (St. Louis, MO, USA) and Bio-Rad Laboratories (Hercules, CA, USA). Phosphoramidites of natural nucleosides and 3-deaza-3-methyl-adenine were purchased from Glen Research (Sterling, VA, USA). Phosphoramidites of 3-deaza-3-phenethyl-adenine and 3-deaza-3-methoxynaphtalenethyl-adenine were synthesized as described previously.¹⁵ Unmodified DNA oligonucleotides and fluorescein-labeled RNA primers were obtained from VBC Biotech, Vienna. The transcription buffer was 20 mM HEPES pH 7.6, 60 mM $(\text{NH}_4)_2\text{SO}_4$, 8 mM MgSO_4 , 10 μM ZnCl_2 , 10% glycerol, 10 mM DTT. SW480 cells overexpressing PTGR1 were obtained as previously described.¹¹

RNA synthesis recovery

Cells seeded at a density of 18,000 cells/well were washed twice with PBS solution (Gibco) and incubated with HMAF (0.5, 1 μ M final concentration, in 0.1% DMSO medium), MMS (1, 500 μ M final concentration, in 0.1% DMSO medium), or in 0.1% DMSO medium as negative control, for 6 h at 37 °C. Transcription activity was evaluated with an RNA synthesis recovery assay (RRS) adapted from a previously published procedure.¹⁵ Full experimental details concerning cell culture conditions and RRS analysis appear in the Supplementary Information.

Synthesis of oligonucleotides containing 3-deaza-3-alkyl-adenosine adducts

Oligonucleotides with a modified adenine analog were synthesized by solid phase chemical DNA synthesis on a Mermade 4 DNA synthesizer (Bioautomation). DNA templates for Pol II primer extension reactions were a 26mer (5'-ACCTCAACTACTTGACCCTCCTCATT-3') and a 34mer (5'-GCTGTTACCGAGGTCCCTCTCGATGGCTGTAAGT-3'), adapted from previous studies.³⁴ Synthesized oligonucleotides were purified by HPLC (Agilent 1100 Series) using an Agilent Luna 25mm C18 column. The chromatographic mobile phases were 50 mM triethylammonium acetate and acetonitrile (ACN) and the gradient used was 10-15% ACN over 35 min. The eluted fractions containing DNA templates were concentrated to dryness in a MiVac centrifugal evaporator (GeneVac), re-suspended in 100 μ L deionized water and checked for purity by direct injection into an Agilent MSD SL ion trap mass spectrometer with electrospray ionization.

Transcription with Pol II

Yeast Pol II was purified as described previously.²² Transcription reactions to assess the efficiency of Pol II transcription in the presence of 3-adenine adducts were performed with DNA scaffolds composed of a 26-mer DNA template containing Ade or modified Ade at the position indicated with an underline (5'-ACCTCAACTACTTGACCCTCCTCATT-3'), a 14-mer non-template DNA (VBC Biotech, Vienna, Austria) (5'-CAAGTAGTTGAGGT-3') and an 11-mer RNA primer labeled with fluorescein at the 5' end (FAM-5'-UUCGAGGAGGG-3'). Oligonucleotides (final concentration 11 μ M for template and non-template DNA; 10 μ M for RNA primer) were annealed by heating to 90 °C in TE buffer with 50 mM NaCl and allowing to slowly cool to 25 °C. ECs were formed by incubating 5 pmol Pol II with 10 pmol DNA scaffold at 20 °C for 20 min in the transcription buffer (total volume 10 μ L). Transcription reactions were initiated by adding NTPs (1 μ L, 10 mM). Reactions were allowed to proceed for 20 min, then quenched by addition of aqueous EDTA (5 μ L, 50 mM). The formation of extended primers was monitored by denaturing gel electrophoresis (7 M urea, 20% acrylamide) followed by visualization with Molecular Imager Gel Doc XR+Imaging System from Bio-Rad. Gel band intensities were quantified with Image J Lab 3.0 and percentage of primer extension calculated with the equation $\sum I_{(n+x)} / (\sum I_{(n+x)} + I_n)$, where $I_{(n+x)}$ represents the intensity of all bands above the

primer and I_n represents the band intensity of the non-extended primer. Primer extension reactions were performed in the presence of a two-fold molar excess of DNA scaffold compared to the enzyme, leaving a theoretical 50% unbound portion of primers un-processed. These residual RNA primers were used as an internal reference band.

Pol II bypass of DNA adducts placed four bases downstream to the transcription start was performed with ECs of 34-mer DNA template (5'-GCTGTTACCGAGTCCCTCTCGATGGCTGTAAGT-3'), 34-mer non-template DNA (5'-ACTTACAGCCATCGAGAGGGACTCGGTGAACAGC-3') and 9-mer RNA primer labeled with fluorescein at the 5' end (5'-AUGGAGAGG-3'). ECs were assembled as described previously.²² Briefly, 200 pmol of RNA primer and DNA template were annealed by heating to 90 °C in TE buffer with 50 mM NaCl and allowing to cool slowly to 25 °C. 5 pmol of RNA:DNA duplex then was incubated with 15 pmol of Pol II in the transcription buffer for 10 min at room temperature. A 50-fold molar excess of 34-mer non-template DNA was added and the reaction was allowed to continue for 10 min at room temperature in order to form a DNA construct mimicking the transcription bubble. The transcription reaction was performed at 37 °C and initiated by addition of NTPs (10 mM, 1 µL). After 30 min, the reaction was quenched with addition EDTA solution (50 mM, 5 µL). RNA products were visualized with a Molecular Imager Gel Doc XR+Imaging System after PAGE electrophoresis.

For steady-state kinetic experiments, primer extension reactions were performed at varying time points (between 0 and 60 min). Extension bands visualized after PAGE electrophoresis and primer extension was quantified with the equation $\sum I_{(n+x)} / (\sum I_{(n+x)} + I_n)$, where $I_{(n+x)}$ represents the intensity of all bands above the primer and I_n represents the band intensity of the non-extended primer. Resulting values were plotted using GraphPad Prism and non-linear regression analysis was carried out using the equation $Y = Y_m * (1 - e^{-(k*X)})$, where Y_m is the percentage of product formed, k is the apparent reaction rate and X is time.

Crystal structure analysis

12-subunit Pol II (4.1 mg/mL) was incubated with a 1.5-fold molar excess of nucleic acid scaffold containing the 3d-Napht-A analog and 5-Br-U. The sample was incubated for 20 min on ice before crystallization by hanging drop vapor diffusion using as reservoir solution 4-7% PEG 6000, 200 mM ammonium acetate, 300 mM sodium acetate, 50 mM HEPES pH 7.0 and 5 mM TCEP. Crystals were grown for 4-8 days, cryo-protected in mother solution supplemented with 22% glycerol and nucleic acid scaffold. Crystals were incubated overnight at 8 °C before they were harvested and flash-cooled in liquid nitrogen. Diffraction data were collected at beamline X06SA of the Swiss Light Source (Villigen, Switzerland) or beamline P13 at the Deutsches Elektronensynchrotron (DESY, Hamburg, Germany). The native data set was collected at a wavelength of 0.976 Å. The Bromine peak data set was collected at a wavelength of 0.919 Å. Diffraction images were processed with XDS.⁴⁴ The structure was solved with molecular

replacement using a 12-subunit Pol II structure (PDB 3HOX) without nucleic acids. Refinement was performed using Phenix.Refine.⁴⁵ Refinement statistics are summarized in Supplementary Table 1. The final model was analyzed using MolProbity.⁴⁶ 92 % of residues were in Ramchandran plot favored regions, 7.2 % in allowed regions. 0.7 % Ramachandran outliers were observed. Figures were generated using PyMOL.⁴⁷

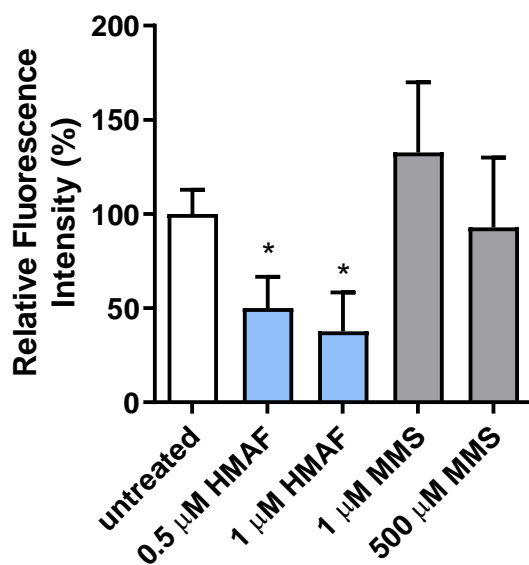
7.6 Supplementary Information

Cell culture

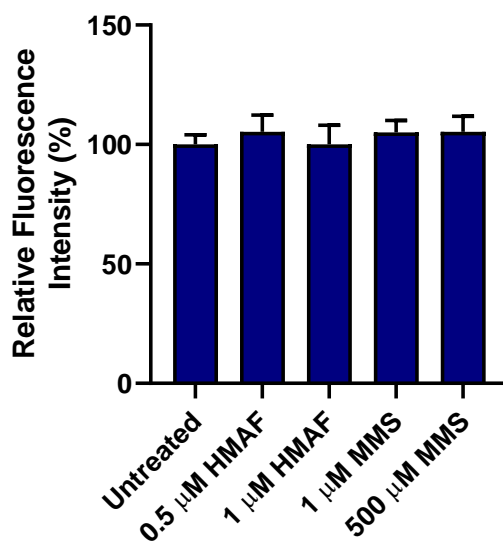
Cells were maintained in RPMI medium 1640 + GlutaMax (Gibco) with 10% fetal calf serum, 1% pen/strep and 0.1 mg/ml geneticin (Gibco), and incubated at 37 °C in a humidified incubator containing 5% CO₂. Cells were seeded in 96-well plate (Corning Costar, black clear flat bottom) at a density of 18,000 cells/well.

RNA Synthesis Recovery

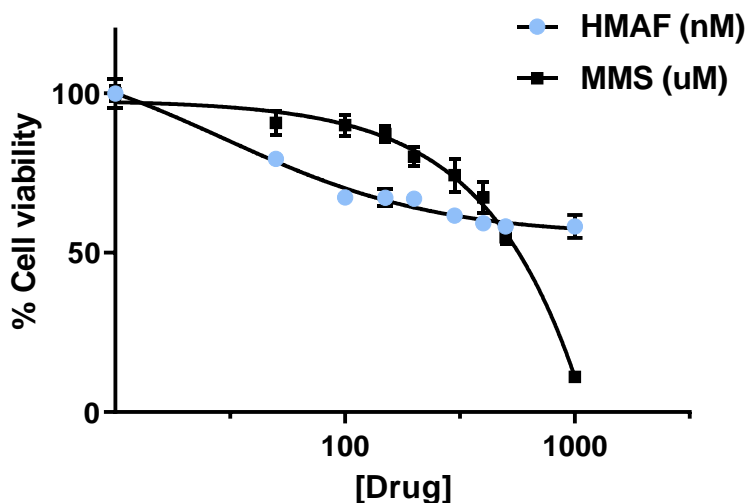
After drug exposure, cells were washed twice with PBS solution and incubated with 500 μM 5-ethynil uridine (EU, Life Technologies Lot 1711751) in serum-free medium at 37 °C for 2 h. Cells were washed twice with PBS, fixed and permeabilized at room temperature for 20 min with a freshly-prepared buffer containing 300 mM sucrose, 2% formaldehyde and 0.5% Triton X-100 in PBS. Cells were washed three times with PBS (5 min/wash) and exposed to the freshly-prepared azide-coupling solution (10 μM AlexaFluor488 azide (Life Technologies), 50 mM Tris-HCl (pH 7.3), 4 mM CuSO₄ and 10 mM sodium ascorbate in nuclease-free water) at room temperature for 1 h. Control cells were incubated in the same coupling solution, containing DMSO instead of the azide. After the coupling, cells were washed three times with PBS containing 0.05% Tween 20 (PBST). Cells were stained with 100 μl of 20 ng/ml DAPI (Dojindo) in PBS at room temperature for 20 min and washed three times with PBST. Fluorescence intensity was recorded in PBS with a Tecan Infinite 200 PRO series reader (Männedorf, Switzerland). The wavelengths used were 490 nm (ex) and 525 nm (em) for AlexaFluor488, and 358 nm (ex) and 461 nm (em) for DAPI. Fluorescence was measured twice. The experiment was carried out three times with 0.5 μM HMAF and 500 μM MMS, and twice with 1 μM HMAF and 1 μM MMS. Each condition was tested with four or five replicates per experiment. Statistical analysis was performed with GraphPad Prism 6 using One-way ANOVA with Dunnett post hoc test.



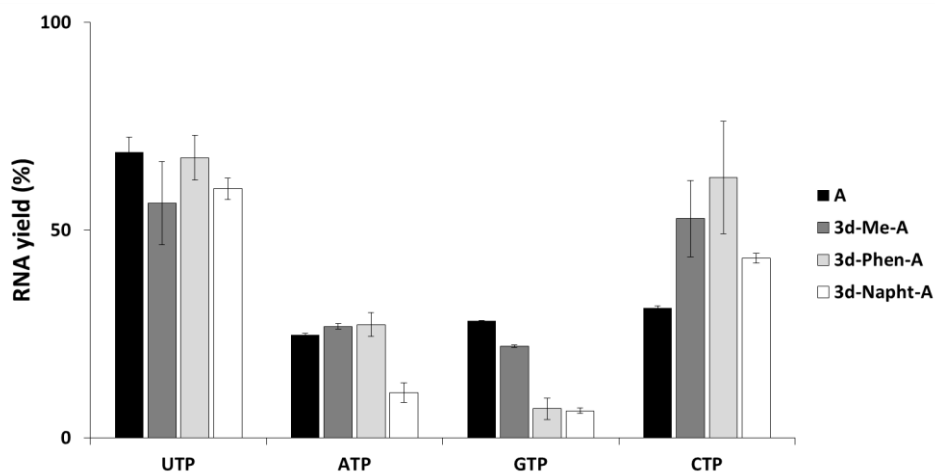
Supplementary Figure 1. Levels of RNA synthesis in SW480-PTGR1 cells treated with HMAF and MMS. RNA synthesis was quantified on the basis of emitted fluorescence indicating incorporation of 5-ethynyl uridine in RNA during transcription, followed by coupling with the fluorescent azide dye Alexa488. Fluorescence signal was normalized to the signal from untreated cells. Error bars represent standard deviation ($N \geq 8$). Statistical significance of treated cells compared to untreated cells was determined by one-way ANOVA with Dunnett post hot test (* $p < 0.0001$).



Supplementary Figure 2. DAPI fluorescence signal in SW480-PTGR1 cells after 6 h of exposure to HMAF (0.5 and 1 μM) or MMS (500 and 1 μM). The average of fluorescence is represented with standard deviation ($N=3$).



Supplementary Figure 3. Cell viability of SW480-PTGR1 cells after 6 h of exposure to increasing concentration of HMAF (0-1000 nM) or MMS (0-1000 μM) and 24 of post-incubation. Cell viability was determined with the CellTiter Glow assay. Average of cell viability is represented with standard deviation (N=3).



Supplementary Figure 4. Quantification of single nucleotide incorporation opposite the 3-alkyl adenosine analogs. Reactions were performed in presence of a single NTP (1mM) and were quenched after 20 min. The average of RNA yield is represented with standard deviation (N=2).

Table 1. Structural data collection and refinement statistics.

EC I with 3-deaza-3-methoxynaphtylethyl group	EC II with 3-deaza-3-methoxynaphtylethyl group –
---	--

	(PDB code)	Bromine peak
Space group	C222 ₁	C222 ₁
<i>a, b, c</i> (Å)	221.4, 394.9 283.7	220.4, 394.2, 283.5
Resolution (Å)	49.88 - 3.2 (3.31 - 3.2)	49.7-3.61 (3.73-3.61)
<i>R</i> _{merge}	0.137 (2.29)	0.403 (2.103)
<i>R</i> _{meas}	0.142 (2.383)	0.41 (2.18)
<i>I</i> / σ (<i>I</i>)	16.38 (1.4)	12.9 (1.5)
<i>CC</i> _{1/2}	0.999 (0.53)	0.992 (0.711)
Completeness (%)	100 (100)	100 (100)
Redundancy	13.5 (13.3)	38.9 (14.2)
Refinement		
Resolution (Å)	49.88 - 3.2 (3.314 - 3.2)	
No. reflections	202,994 (20,129)	
<i>R</i> _{work} / <i>R</i> _{free}	0.18 (0.33) / 0.22 (0.36)	
No. atoms	31,670	
Protein/DNA	31,626	
Ligand/ion	44	
Protein	139.6	
Ligand/ion	140.8	
R.m.s. deviations		
Bond lengths (Å)	0.033	
Bond angles (°)	1.03	

^a Values in parentheses are for highest-resolution shell.

7.7 References

1. Salehan, M. R.; Morse, H. R., DNA damage repair and tolerance: a role in chemotherapeutic drug resistance. *British journal of biomedical science* **2013**, *70* (1), 31-40.

2. Cirauqui, B.; Margeli, M.; Quiroga, V.; Quer, A.; Karachaliou, N.; Chaib, I.; Ramirez, J. L.; Munoz, A.; Pollan, C.; Planas, I.; Drozdowsky, A.; Rosell, R., DNA repair pathways to regulate response to chemoradiotherapy in patients with locally advanced head and neck cancer. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine* **2016**, *37* (10), 13435-13443.
3. Yu, W. K.; Wang, Z.; Fong, C. C.; Liu, D.; Yip, T. C.; Au, S. K.; Zhu, G.; Yang, M., Chemoresistant lung cancer stem cells display high DNA repair capability to remove cisplatin-induced DNA damage. *British journal of pharmacology* **2017**, *174* (4), 302-313.
4. Jaspers, N. G.; Raams, A.; Kelner, M. J.; Ng, J. M.; Yamashita, Y. M.; Takeda, S.; McMorris, T. C.; Hoeijmakers, J. H., Anti-tumour compounds illudin S and Irofulven induce DNA lesions ignored by global repair and exclusively processed by transcription- and replication-coupled repair pathways. *DNA repair* **2002**, *1* (12), 1027-38.
5. Koepfel, F.; Poindessous, V.; Lazar, V.; Raymond, E.; Sarasin, A.; Larsen, A. K., Irofulven cytotoxicity depends on transcription-coupled nucleotide excision repair and is correlated with XPG expression in solid tumor cells. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2004**, *10* (16), 5604-13.
6. Otto, C.; Spivak, G.; Aloisi, C. M.; Menigatti, M.; Naegeli, H.; Hanawalt, P. C.; Tanasova, M.; Sturla, S. J., Modulation of Cytotoxicity by Transcription-Coupled Nucleotide Excision Repair Is Independent of the Requirement for Bioactivation of Acylfulvene. *Chemical research in toxicology* **2017**, *30* (3), 769-776.
7. Gong, J.; Vaidyanathan, V. G.; Yu, X.; Kensler, T. W.; Peterson, L. A.; Sturla, S. J., Depurinating acylfulvene-DNA adducts: characterizing cellular chemical reactions of a selective antitumor agent. *Journal of the American Chemical Society* **2007**, *129* (7), 2101-11.
8. Hargrove, A. E.; Martinez, T. F.; Hare, A. A.; Kurmis, A. A.; Phillips, J. W.; Sud, S.; Pienta, K. J.; Dervan, P. B., Tumor Repression of VCaP Xenografts by a Pyrrole-Imidazole Polyamide. *PLoS one* **2015**, *10* (11), e0143161.
9. Oncology Venture, MPI's drug development arm and Lantern Pharma receive ICIP grant to advance Irofulven for metastatic prostate cancer. <https://globenewswire.com/news-release/2015/03/04/712201/10123208/en/Oncology-Venture-MPI-s-drug-development-arm-and-Lantern-Pharma-receive-ICIP-grant-to-advance-Irofulven-for-metastatic-prostate-cancer.html> **May 29, 2015**, last access 06.10.2016.
10. Neels, J. F.; Gong, J.; Yu, X.; Sturla, S. J., Quantitative correlation of drug bioactivation and deoxyadenosine alkylation by acylfulvene. *Chemical research in toxicology* **2007**, *20* (10), 1513-9.
11. Pietsch, K. E.; van Midwoud, P. M.; Villalta, P. W.; Sturla, S. J., Quantification of acylfulvene- and illudin S-DNA adducts in cells with variable bioactivation capacities. *Chemical research in toxicology* **2013**, *26* (1), 146-55.
12. van Midwoud, P. M.; Sturla, S. J., Improved efficacy of acylfulvene in colon cancer cells when combined with a nuclear excision repair inhibitor. *Chemical research in toxicology* **2013**, *26* (11), 1674-82.
13. Woynarowski, J. M.; Napier, C.; Koester, S. K.; Chen, S. F.; Troyer, D.; Chapman, W.; MacDonald, J. R., Effects on DNA integrity and apoptosis induction by a novel antitumor sesquiterpene drug, 6-hydroxymethylacylfulvene (HMAF, MGI 114). *Biochemical pharmacology* **1997**, *54* (11), 1181-93.
14. Hanawalt, P. C., Subpathways of nucleotide excision repair and their regulation. *Oncogene* **2002**, *21* (58), 8949-56.
15. Glatt, H.; Pietsch, K. E.; Sturla, S. J.; Meinel, W., Sulfotransferase-independent genotoxicity of illudin S and its acylfulvene derivatives in bacterial and mammalian cells. *Archives of toxicology* **2014**, *88* (1), 161-9.

16. Malvezzi, S.; Angelov, T.; Sturla, S. J., Minor Groove 3-Deaza-Adenosine Analogues: Synthesis and Bypass in Translesion DNA Synthesis. *Chemistry* **2017**, *23* (5), 1101-1109.
17. Jiang, H.; Yang, L. Y., Cell cycle checkpoint abrogator UCN-01 inhibits DNA repair: association with attenuation of the interaction of XPA and ERCC1 nucleotide excision repair proteins. *Cancer research* **1999**, *59* (18), 4529-34.
18. Ljungman, M.; Lane, D. P., Transcription - guarding the genome by sensing DNA damage. *Nature reviews. Cancer* **2004**, *4* (9), 727-37.
19. Dimitri, A.; Burns, J. A.; Broyde, S.; Scicchitano, D. A., Transcription elongation past O^6 -methylguanine by human RNA polymerase II and bacteriophage T7 RNA polymerase. *Nucleic acids research* **2008**, *36* (20), 6459-71.
20. Tornaletti, S.; Maeda, L. S.; Kolodner, R. D.; Hanawalt, P. C., Effect of 8-oxoguanine on transcription elongation by T7 RNA polymerase and mammalian RNA polymerase II. *DNA repair* **2004**, *3* (5), 483-94.
21. You, C.; Dai, X.; Yuan, B.; Wang, J.; Wang, J.; Brooks, P. J.; Niedernhofer, L. J.; Wang, Y., A quantitative assay for assessing the effects of DNA lesions on transcription. *Nature chemical biology* **2012**, *8* (10), 817-22.
22. You, C.; Wang, J.; Dai, X.; Wang, Y., Transcriptional inhibition and mutagenesis induced by N-nitroso compound-derived carboxymethylated thymidine adducts in DNA. *Nucleic acids research* **2015**, *43* (2), 1012-8.
23. Brueckner, F.; Hennecke, U.; Carell, T.; Cramer, P., CPD damage recognition by transcribing RNA polymerase II. *Science* **2007**, *315* (5813), 859-62.
24. Damsma, G. E.; Alt, A.; Brueckner, F.; Carell, T.; Cramer, P., Mechanism of transcriptional stalling at cisplatin-damaged DNA. *Nature structural & molecular biology* **2007**, *14* (12), 1127-33.
25. Walmacq, C.; Wang, L.; Chong, J.; Scibelli, K.; Lubkowska, L.; Gnatt, A.; Brooks, P. J.; Wang, D.; Kashlev, M., Mechanism of RNA polymerase II bypass of oxidative cyclopurine DNA lesions. *Proceedings of the National Academy of Sciences of the United States of America* **2015**, *112* (5), E410-9.
26. Wang, D.; Zhu, G.; Huang, X.; Lippard, S. J., X-ray structure and mechanism of RNA polymerase II stalled at an antineoplastic monofunctional platinum-DNA adduct. *Proceedings of the National Academy of Sciences of the United States of America* **2010**, *107* (21), 9584-9.
27. Shin, J. H.; Xu, L.; Wang, D., Mechanism of transcription-coupled DNA modification recognition. *Cell & bioscience* **2017**, *7*, 9.
28. Shin, J. H.; Xu, L.; Wang, D., RNA polymerase II acts as a selective sensor for DNA lesions and endogenous DNA modifications. *Transcription* **2016**, *7* (3), 57-62.
29. Xu, L.; Butler, K. V.; Chong, J.; Wengel, J.; Kool, E. T.; Wang, D., Dissecting the chemical interactions and substrate structural signatures governing RNA polymerase II trigger loop closure by synthetic nucleic acid analogues. *Nucleic acids research* **2014**, *42* (9), 5863-70.
30. Fouqueau, T.; Zeller, M. E.; Cheung, A. C.; Cramer, P.; Thomm, M., The RNA polymerase trigger loop functions in all three phases of the transcription cycle. *Nucleic acids research* **2013**, *41* (14), 7048-59.
31. Brueckner, F.; Ortiz, J.; Cramer, P., A movie of the RNA polymerase nucleotide addition cycle. *Current opinion in structural biology* **2009**, *19* (3), 294-9.
32. Da, L. T.; Pardo-Avila, F.; Xu, L.; Silva, D. A.; Zhang, L.; Gao, X.; Wang, D.; Huang, X., Bridge helix bending promotes RNA polymerase II backtracking through a critical and conserved threonine residue. *Nature communications* **2016**, *7*, 11244.
33. Xu, L.; Wang, W.; Gotte, D.; Yang, F.; Hare, A. A.; Welch, T. R.; Li, B. C.; Shin, J. H.; Chong, J.; Strathern, J. N.; Dervan, P. B.; Wang, D., RNA polymerase II senses obstruction in the DNA minor groove

via a conserved sensor motif. *Proceedings of the National Academy of Sciences of the United States of America* **2016**, *113* (44), 12426-12431.

34. Beranek, D. T., Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutation research* **1990**, *231* (1), 11-30.
35. Brink, A.; Schulz, B.; Stopper, H.; Lutz, W. K., Biological significance of DNA adducts investigated by simultaneous analysis of different endpoints of genotoxicity in L5178Y mouse lymphoma cells treated with methyl methanesulfonate. *Mutation research* **2007**, *625* (1-2), 94-101.
36. Gaykalova, D. A.; Kulaeva, O. I.; Pestov, N. A.; Hsieh, F. K.; Studitsky, V. M., Experimental analysis of the mechanism of chromatin remodeling by RNA polymerase II. *Methods in enzymology* **2012**, *512*, 293-314.
37. Burns, J. A.; Dreij, K.; Cartularo, L.; Scicchitano, D. A., *O*⁶-methylguanine induces altered proteins at the level of transcription in human cells. *Nucleic acids research* **2010**, *38* (22), 8178-87.
38. Wang, D.; Bushnell, D. A.; Westover, K. D.; Kaplan, C. D.; Kornberg, R. D., Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **2006**, *127* (5), 941-54.
39. Touloukhonov, I.; Zhang, J.; Palangat, M.; Landick, R., A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Molecular cell* **2007**, *27* (3), 406-19.
40. Brueckner, F.; Cramer, P., Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nature structural & molecular biology* **2008**, *15* (8), 811-8.
41. Escargueil, A. E.; Poindessous, V.; Soares, D. G.; Sarasin, A.; Cook, P. R.; Larsen, A. K., Influence of irifolven, a transcription-coupled repair-specific antitumor agent, on RNA polymerase activity, stability and dynamics in living mammalian cells. *Journal of cell science* **2008**, *121* (Pt 8), 1275-83.
42. Monti, P.; Broxson, C.; Inga, A.; Wang, R. W.; Menichini, P.; Tornaletti, S.; Gold, B.; Fronza, G., 3-Methyl-3-deazaadenine, a stable isostere of N3-methyl-adenine, is efficiently bypassed by replication in vivo and by transcription in vitro. *DNA repair* **2011**, *10* (8), 861-8.
43. Sweet, J. M.; Carda, B.; Small, G. D., Repair of 3-methyladenine and 7-methylguanine in nuclear DNA of *Chlamydomonas*: requirement for protein synthesis. *Mutation research* **1981**, *84* (1), 73-82.
44. Metz, A. H.; Hollis, T.; Eichman, B. F., DNA damage recognition and repair by 3-methyladenine DNA glycosylase I (TAG). *The EMBO journal* **2007**, *26* (9), 2411-20.
45. Bobola, M. S.; Kolstoe, D. D.; Blank, A.; Chamberlain, M. C.; Silber, J. R., Repair of 3-methyladenine and abasic sites by base excision repair mediates glioblastoma resistance to temozolomide. *Frontiers in oncology* **2012**, *2*, 176.
46. Bobola, M. S.; Varadarajan, S.; Smith, N. W.; Goff, R. D.; Kolstoe, D. D.; Blank, A.; Gold, B.; Silber, J. R., Human glioma cell sensitivity to the sequence-specific alkylating agent methyl-lexitropsin. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2007**, *13* (2 Pt 1), 612-20.
47. Kabsch, W., Xds. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (Pt 2), 125-32.
48. Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H., PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (Pt 2), 213-21.
49. Chen, V. B.; Arendall, W. B., 3rd; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C., MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* **2010**, *66* (Pt 1), 12-21.
50. Schrodinger, LLC, The PyMOL Molecular Graphics System, Version 1.8. 2015.

Curriculum Vitae

Claudia M. N. Aloisi

Department of Health Sciences and Technology, ETH Zürich
 Email: claudia.aloisi@hest.ethz.ch | Phone: +41446322699
 ORCID: 0000-0002-2279-7757



EDUCATION

-
- | | |
|----------------------------------|--|
| 05/2015-11/2019
(Anticipated) | PhD in Biochemical Toxicology
ETH Zurich (Swiss Federal Institute of Technology Zurich), Zurich, Switzerland
Molecular Life Science Doctoral Program Zurich |
| 10/2012-12/2014 | Master in Organic Chemistry (<i>cum Laude</i>)
University of Pisa, Pisa, Italy |
| 09/2009-07/2012 | Bachelor in Chemistry (<i>cum Laude</i>)
University of Catania, Catania, Italy |

RESEARCH EXPERIENCE

-
- | | |
|-------------------|--|
| 05/2015 – Present | Graduate Research Assistant, ETH Zurich (Department of Health Science and Technology)
<i>Gene-targeted identification of DNA lesions to unravel mechanisms of mutagenesis and carcinogenesis</i>
Advisor: Prof. Dr. Shana J. Sturla |
| 06/2019 – 07/2019 | Visiting graduate researcher, UNIST Ulsan (Scholarship: Young Researchers' Exchange Programme)
Advisor: Prof. Dr. Orlando D. Schärer |
| 02/2014 – 09/2014 | Master Thesis Intern, King's college of London UK (Scholarship: Erasmus Outgoing Consortia Placement)
<i>Biotechnological production of <i>Isd11</i> and study of its state of aggregation</i>
Advisor: Prof. Dr. Annalisa Pastore |
| 01/2012 – 06/2012 | Bachelor Thesis Intern, CNR-IBC Italy (National Research Council-Institute of Biomolecular Chemistry)
<i>Synthesis of a folic acid-calix[4]arene conjugate as a potential imaging agent</i>
Advisor: Dr. Corrada Geraci |

AWARDS AND FELLOWSHIPS

-
- | | |
|---------|--|
| 03/2019 | Fellowship, Exchange Programme between Korea and Switzerland |
| 07/2018 | Travel grant, Swiss Society of Toxicology |
| 07/2018 | Travel grant, Molecular Life Science doctoral program Zurich |
| 06/2018 | Travel support for session chairs, Gordon Research Conference |
| 09/2017 | Travel grant, Autumn School Chemical Biology & ISBOC-11, Konstanz, Germany |

11/2016	Award best oral presentation, DNA damage response Symposium, Mainz, Germany
11/2016	Travel grant, DNA damage response Symposium, Mainz, Germany
01/2014	Scholarship for Master Thesis internship, Erasmus Outgoing Consortia Placement

SUPERVISION OF RESEARCHERS

02/2019 – Present	Emma Sandell, Doctoral thesis in Biochemical Toxicology, ETH Zurich
02/2019 – Present	Nora Escher, Master internship and thesis in Biology, ETH Zurich
05/2018 – 10/2018	Xuan Li, Master thesis in Food Science, ETH Zurich

TEACHING EXPERIENCE

04/2017 – Present	Teaching Assistant , Toxicology and Nutrition Laboratory Bachelor Course, ETHZ
11/2018	Teaching Assistant , Tutorial in <i>Transcription-blocking agents in cancer therapy</i>
03/2017	Teaching certificate , ETHZ
08/2016	Teaching Assistant , Molecular Disease Mechanism, ETHZ
05/2016	Teaching Assistant , Introduction to Molecular Toxicology, ETHZ

SERVICE AND OUTREACH

06/2018-06/2020	Chair of 2020 Gordon Research Seminar on Mutagenesis
05/2019	Scientific group retreat “Collaboration and communication” organizer, Malaga, Spain
06/2018	Discussion leader, 2018 Gordon Research Seminar on Mutagenesis, Newry, ME, US
01/2017-05/2018	Molecular Life Science Lecture Series organizing committee
09/2017	Scientific group retreat “Project synergies” organizer, Barcelona, Spain

PUBLICATIONS

1. Aloisi C.M.N., Escher N.A., Geisen S.M., Kim H.S., Yeo J.E., Schärer O.D., Sturla S.J., MGMT and NER alleviate O⁶-Carboxymethyl-Guanine DNA Adduct mutagenicity in cells, *Manuscript in preparation*
2. Aloisi C.M.N., Nilforoushan A., Ziegler N., Sturla S.J., Sequence-specific Quantitation of Mutagenic DNA damage via Polymerase Amplification with an Artificial Nucleotide, *Manuscript submitted*
3. Aloisi C.M.N., Sturla S.J., Gahlon H.L., A gene-targeted polymerase amplification strategy to identify O⁶-methylguanine damage, *Chem. Comm.* **2019**, *55*, 3895
4. Raez M.H., Aloisi C.M.N., Gahlon H.L., Sturla S.J., DNA adduct-directed synthetic nucleosides, *Acc. Chem. Res.* **2019**, *52*, 5, 1391-1399
5. Malvezzi S., Farnung L., Aloisi C.M.N., Angelov T., Cramer P., Sturla S.J., Mechanism of RNA polymerase II stalling by DNA alkylation, *PNAS* **2017**, *46*, 114, 12172-12177
6. Otto C., Spivak G., Aloisi C.M.N., Menigatti M., Naegeli H., Hanawalt P.C., Tanasova M., Sturla S.J.; Modulation of Cytotoxicity by Transcription-Coupled Nucleotide Excision Repair Is Independent of the Requirement for Bioactivation of Acylfulvene, *Chem Res Toxicol* **2017**, *30*, 3, 769-77
7. Yan R., Friemel M., Aloisi C.M.N., Huynen M., Taylor I.A., Leimkühler S., Pastore A.; The Eukaryotic-Specific ISD11 Is a Complex-Orphan Protein with Ability to Bind the Prokaryotic IscS, *PLoS ONE* **2016**

PRESENTATIONS

1. Swiss Chemical Society Fall Meeting 2019: [Aloisi C.M.N.](#), Sturla S. J., *Gene-targeted identification of DNA lesions to unravel mechanisms of mutagenesis*, Zurich, September **2019** (oral presentation)
2. German Society for Environmental Mutation Research (GUM) Conference 2019,: [Aloisi C.M.N.](#), Sturla S. J., *Gene-targeted identification of DNA lesions to unravel mechanisms of mutagenesis*, Basel, September **2019** (oral presentation)
3. Mutagenesis Gordon Research Conference (GRC): [Aloisi C.M.N.](#), Sturla S. J., *Sequence-specific detection of rare DNA damage events*, Newry, Maine, US, June **2018** (poster and flash oral presentation)
4. Mutagenesis Gordon Research Seminar (GRS): [Aloisi C.M.N.](#), Sturla S. J., *Sequence-specific detection of rare DNA damage events*, Newry, Maine, US, June **2018** (poster and oral presentation)
5. Genome Dynamics-RRR meeting: [Aloisi C.M.N.](#), Sturla S. J., *How cancer cells resist a minor groove alkylating agent*, Zurich, October **2017** (oral presentation)
6. International Symposium on Bioorganic Chemistry: [Aloisi C.M.N.](#), Nilforoushan A., Sturla S. J., *Strategies for sequence-specific detection and quantitation of pro-mutagenic O⁶-alkylguanine DNA adducts*, Konstanz, Germany, September **2017** (poster)
7. Symposium on DNA damage response, genetic instability and cancer: [Aloisi C.M.N.](#), Sturla S. J., *Transcription-coupled nucleotide excision repair as a basis for cellular resistance to acylfulvene anticancer alkylating agents*, Mainz, Germany, November **2016** (Award best oral presentation)
8. 13th Molecular Life Science retreat: [Aloisi C.M.N.](#), Sturla S. J., *Transcription-coupled nucleotide excision repair as a basis for cellular resistance to acylfulvene anticancer alkylating agents*, Kiental, Switzerland, September **2016** (poster)
9. Latsis Symposium on Personalized Medicine: [Aloisi C.M.N.](#), Sturla S. J., *Transcription-coupled nucleotide excision repair as a basis for cellular resistance to acylfulvene anticancer alkylating agents*, Zurich, Switzerland, June **2016** (poster and flash oral presentation)

PERSONAL SKILLS

Languages: Italian: Native Speaker; English: Fluent; German, French, Spanish: Basic communication

Digital competences: MOE (molecular simulation and modeling), PyMol (molecular visualization), Adobe illustrator (graphics), GraphPad Prism (data analysis), Image Lab (gel analysis), Xcalibur (chromatography-mass spectrometry data visualization and analysis)

REFERENCES

Prof. Dr. Shana Sturla, ETH Zurich, Zurich, Switzerland shana.sturla@hest.ethz.ch

Prof. Dr. Orlando Schärer, UNIST, Ulsan, South Korea orlando.scharer@ibs.re.kr

Prof. Dr. Julian Sale, University of Cambridge, Cambridge, UK jes@mrc-lmb.cam.ac.uk

Prof. Dr. Maureen McKeague, McGill University, Montreal, Canada maureen.mckeague@mcgill.ca

Dr. Nancy Ryan Gray, Gordon Research Conference president and CEO ngray@grc.org

