

Inauguraldissertation

Study and exploitation of the direct
DNA repair protein
O6-MethylGuanine
MethylTransferase – MGMT

Zur Erlangung der Würde eines Doktors der Philosophie
vorgelegt der Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

eingereicht von

Malgorzata Marta Murawska

Basel, 2021

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Erstbetreuer: Prof. Dr. Dennis Gillingham

Zweitbetreuer: Prof. Dr. Thomas R. Ward

Externe Expertin: Prof. Dr. Anna K. H. Hirsch

Basel, den 02. März 2021

Prof. Dr. Marcel Mayor

Doctoral dissertation

Study and exploitation of the direct
DNA repair protein
O6-MethylGuanine
MethylTransferase - MGMT

Dissertation
for the Degree of Doctor of Philosophy
presented to the Faculty of Science at the University of Basel

submitted by

Malgorzata Marta Murawska

Basel, 2021

Pooh and Piglet walked home thoughtfully together in the golden evening, and for a long time they were silent.

"When you wake up in the morning, Pooh," said Piglet at last, "what's the first thing you say to yourself?"

"What's for breakfast?" said Pooh. "What do you say, Piglet?"

"I say, I wonder what's going to happen exciting today?" said Piglet.

Alan Alexander Milne
Winnie the Pooh

For Janusz, Mira and Arthur

Acknowledgements

First and foremost, I want to thank my supervisor **Prof. Dennis Gillingham** for all the support over the last four years. Thank you for the opportunity of working in your group and valuable discussions. I have only been in the group for a short time, but I can already see the changes in research focus and future directions, which are very exciting!

I also want to thank **Prof. Thomas Ward** as my secondary supervisor. Tom, I have not contributed much to your busy schedule, but when I did, you were always available and ready to help, for which I am very grateful.

Prof. Anna Hirsch, thank you for taking the role of an external expert. We know each other from the University of Groningen, where your incredible management style and kindness towards others made you famous. I am very happy to have you on my Doctorate Committee.

A big thank you to everybody that keeps the department running and that oftentimes could help solve problems on short notice: the **Werkstatt (Markus, Hisni, Andreas and Andres Koller)** for maintaining a great research atmosphere and fixing everything I broke within the last years. **Daniel Häussinger** and his entire group (**Thomas, Daniel, Raphael and Pascal**) for maintaining the NMRs. I want to thank **Oli Ilg** for all his help, for organizing all chemicals and materials and in particular for optimizing all the processes regarding the shop. I also want to thank the administrative staff, **Brigitte, Beatrice and Olaf**, and especially **Marina Mambelli Johnson**, for all the work they do in the background to keep the department running.

I would like to thank **Prof. Peter Kaiser**, for his incredible help and mentorship during my time at the University of California in Irvine. I am very lucky to have worked with you and your talented team. The friendly and welcoming group environment together with professionalism and expertise, helped during last year of my PhD. I would like to specifically acknowledge help of **Dr. Linda Lauinger, Dr. Geetha Durairaj and Anna Andronicos**.

I am especially grateful to **Dr. Seemon Coomar** for going with me through the “PhD experience” since my first day of work. Your friendship and support helped me getting through the most frustrating periods and you were always there to celebrate the happy ones, too. Thank you for endless discussions about the world, life, science, research and everyday gossips.

Dr. Mirjam Schreier – I am so lucky that I met you! You are an amazing scientist, teacher, friend and roommate. Your incredible hospitality and trust in people should lead everyone by example. I am forever grateful for all your help in general and during the most critical moments towards the end of my doctorate, in particular. Thank you for letting me stay with your family and for always being around.

Dr. Christina Wegeberg – thank you for being at the right place at the right time. You came to Basel during my last year and you made it wonderful. Our connections to University of Groningen and similar personalities helped us “click” at the first instant. You are a wonderful person and I will most certainly miss being around you. Thank you!

I want to acknowledge the **Gillingham group** in general, and **lab 306** in particular. Thank you for your constant help, at the end we spent most of our time together and your kindness and terrible taste in music make it count. Thanks for the great time on many retreats and happy hours. Specifically, I would like to mention my

Acknowledgements

colleagues **Dr. Basilius, Lucas, Caspar** and **Werner. Basilius**, apart from being so incredibly helpful and willing to spend hours solving other people's problems, you had also become my friend. Thank you for teaching me everything you knew about getting familiar in a lab, helping with early biology aspects of the project as well as endless computer support.

Special thanks to **Daniel Joss** for proofreading my thesis and for helping with any life / personal difficulties. You are an extremely talented and hard-working person and I am glad for everything you did for me.

Joining PhD Chemistry Community (**PCC**) was a wonderful experience. Apart from meeting great people it is truly a great community of talented scientists. I would like to thank its members: **Serena, Jaicy, Daniel, Anja, Fadri, Ivan, Ksenia, Dima, Patrick, Nadja, Vitto, Pascal** and **Alain**.

I would like to help my Stranger People friends: **Holly, Katrin, Joan** and **Cristina**. I remember looking forward to our weekly meetings with all of your hospitality, especially Joan and Holly. I miss those afternoons and am incredibly glad for your friendship.

I wish to thank my close friends: **Patrick, Lena, Zlatkov** and **Dragan** for many, many evenings, BBQs, beers, river swims, hikes and any outside activities imaginable.

My dear friends: **Cristina** and **Sasha**. We spent the best time together in Berkeley, and I was extremely happy we had a chance to all live in Switzerland at the same time. I am incredibly grateful for your support and all the great memories.

To my flatmates: **Carla, Sahitya** and **Mirj**. Sharing a place with a stranger can be a stressful experience, but you made it wonderful. I had never enjoyed being home more than when living with you.

Thank you **Michał** and **Madzia**, for bringing home closer by being so welcoming and supportive. I can always count on you and I feel like our friendship grew bigger and stronger. Thank you for your hospitality and advice throughout the last couple of years.

To my besties **Pyśka** and **Hannah** for endless patience and advice, never-ending discussions and support, I am truly thankful.

Brett, you have been at the very core of my success. Despite living across the globe and worrying a lot for the success of my PhD you never stopped supporting me. I am forever grateful for your love and commitment, especially in those difficult last few months. You are a dream come true, the most talented scientist and the love of my life.

Ogromnie dziękuję moim wspaniałym **rodzicom** i **braciom** za nieskończone wsparcie, poświęcenie i cierpliwość okazaną przez ostatnie 30 lat, które doprowadziły do tego momentu. Nie ma słów, które mogłyby zgrabnie opisać moją wdzięczność, chcę zatem podkreślić, że nie byłabym tu gdzie jestem bez Was.

Finally, I would like to acknowledge the **artistic influence** on my PhD and the stimulating creativity obtained by musical support of Aurora Aksnes, Stefani Germanotta, Torbjørn Brundtland, Svein Berge and Robin Carlsson; as well as the creators of The Good Place, How I built This and The Office Ladies podcasts.

Table of Content

Chapter 1 “Theoretical Background”	1
1 Introduction to DNA damage and repair	1
1.1 DNA damage	1
1.2 Different DNA repair mechanism pathways	4
1.3 DNA repair summary	8
2 Protein degradation pathways	9
2.1 Lysosomal degradation pathway	9
2.2 Proteasomal pathway	10
2.3 Summary:.....	14
3 Bibliography:	14
Chapter 2 "Substrate Scope"	27
4 Introduction	27
5 HPLC-based assay for MGMT activity screen	28
6 Fluorescence-based assay for MGMT activity screen	29
6.1 Evaluation of assay sensitivity.....	29
6.2 Single and double strand oligonucleotides as substrates for MGMT	30
7 Conclusions	32
8 Bibliography	32
Chapter 3 "MGMT with bifunctional molecules"	35
9 Introduction	35
10 Non-covalent MGMT-POI interaction	36
10.1 An MGMT bifunctional targeting Bcr-Abl	36
10.2 An MGMT bifunctional targeting CDK4 and CDK6.....	37
10.3 An MGMT bifunctionals targeting BRD4.....	39

11	Covalent MGMT - POI interaction	42
11.1	Small molecule with unspecific target.....	42
11.2	Study of MGMT – POI degradation by protein fusion.....	46
12	Conclusions	48
13	Bibliography	48
	Chapter 4 "Mechanistic Study"	51
14	Introduction	51
15	Degradation and de novo synthesis.....	51
15.1	Protein degradation rate and de novo synthesis.....	51
15.2	Source of alkylation.....	52
16	The inhibitory study of ubiquitin signaling proteasomal pathway.....	53
17	Proximity labeling of MGMT interacting partners	55
17.1	Overexpressed BioID-MGMT.....	55
17.2	Engineered HEK293T cell line stably expressing TurboID-MGMT	56
17.3	The inhibitory study of UPP in engineered HEK293T cell line, stably expressing TurboID-MGMT.....	60
18	siRNA inhibition.....	61
18.1	siRNA Cul1, CUL4A, CUL4B.....	61
18.2	siRNA RAD18.....	61
19	Conclusions	62
20	Bibliography	62



Abbreviations

%	Percentage	Cul	Cullin
°C	Celsius centigrade	CuSO₄	Copper (II) Sulfate
6-4PP	pyrimidine (6-4) pyrimidone PhotoProduct	Cys	Cysteine
A	Adenine	d	Doublet
AAVS1	Adeno-Associated Virus integration Site 1	D₂O	Deuterated water
Ac₂O	Diacetyl ether	DCM	Dichloromethane
ACN	Acetonitrile	DDB1	DNA Damage-Binding protein 1
Ada	O6-methylguanine-DNA- methyltransferase	DDB2	DNA Damage-Binding protein 2
aEJ	Alternative End Joining	DIEA	N,N-DiisopropylEthylAmine
AGT	O ⁶ -MethylGuanine-DNA methylTransferase	DMEM	Dulbecco's Modified Eagle Medium
AID	Auxin Inducible Degron	DMF	DiMethylFormamide
AlkA	3-methyladenine-DNA-glycosylase	DMSO	Dimethyl sulfoxide
AlkB	1-methyladenine-DNA-dioxygenase	DMSO-d₆	Deuterated DMSO
APC	Anaphase - Promoting Complex	DNA	Deoxyribonucleic acid
ATL	AlkylTransferase-Like protein	dsDNA	Double-standed DNA
BCNU	Bis-chloroethylnitrosourea	DTT	Dithiothreitol
BER	Base-excision repair	E	Elution
br	Broad	E. coli	Escherichia coli
BRD4	BRomoDomain-containing protein 4	E1	Ubiquitin-activating enzyme
C	Cytosine	E2	Ubiquitin conjugating enzyme
CBC	Cul2–elongin B–elongin C	E3	Ubiquitin ligase enzyme
CCD	Catalytic Cysteine Domain	EDC·HCl	N-(3-Dimethylaminopropyl)-N'- ethylcarbodiimide HydroChloride
CD₃CN	Deuterated Acetonitrile	EDTA	EthyleneDiamineTetraacetic Acid
CDCl₃	Deuterated Chloroform	EGF	Epidermal Growth Factor
CDK	Cell Division Kinase	eq.	Equivalent
cDNA	Complementary DNA	Et₃N	Triethylamine
CETN2	Centrin 2	EtOAc	Ethyl Acetate
Ch-Q	ChloroQuine	EtOH	Ethanol
CH₂Cl₂	Dichloromethane	FA	Formic Acid
CMT	ChloroMethyl Triazole	FBS	Fetal Bovine Serum
CMV	CytoMegaloVirus	FT	Flow Through
cNHEJ	Classical NHEJ	G	Guanine
Cp*³Ru(PP)	Chloro(cyclopentadienyl)bis(triphenylp hosphine)ruthenium (II)	GAPDH	GlycerAldehyde-3-Phosphate DeHydrogenase
h₃)₂Cl	hosphine)ruthenium (II)	GSF	Genomic Safe Harbours
CPD	Cyclobutane Pyrimidine Dimers	h	Hour
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats	H2A	Histone H2A
CTRL	Control Reaction	H₂O	Water
CuI	Copper Iodide	HATU	Hexafluorophosphate Azabenzotriazole Tetramethyl Uronium
		HCl	Hydrogen Chloride
		HECT	Homologous to E6-AP Carboxyl Terminus

Abbreviations

HEK	Human Embryonic Kidney	NEDD8	Neuronal Precursor cell-expressed Developmentally Downregulated protein 8
HeLa	Henrietta Lacks	NER	Nucleotide-excision repair
HEPES	(4-(2-HydroxyEthyl)-1-PiperazineEthaneSulfonic acid)	NH₄Cl	Ammonium chloride
His-	Histidine tag	NHEJ	NonHomologous End-Joining
HR	Homologues recombination	Ni	Nickel
HRMS	High Resolution Mass Spectrometry	NI	Not Induced
HSPB1	Heat Shock Protein Beta 1	mM	Milli-Molar
HUH6	Hepatoblastoma carcinoma	nm	Nano-Meter
IAA	IodoAcetAmide	nM	Nano-Molar
IB	ImmunoBlotting	NMR	Nuclear Magnetic Resonance
IGBP1	ImmunoGlobulin-Binding Protein 1	NMU	N-methyl-N'-nitrosourea
IgG	Immunoglobulin G	NO	Nitric Oxide
Indi	Indisulam	NP-40	Nonyl Phenoxy polyethoxy ethanol-40
J	Coupling constant	O6-BG	O6-BenzylGuanine
K	Kelvin	ON	Overnight
K₂CO₃	Potassium Carbonate	P	Pellet
kDa	kilo-Dalton	p	Pentad
KEAP	Kelch-like ECH-Associated Protein	PAGE	PolyAcrylamide Gel Electrophoresis
LC	Liquid Chromatography	PARP	PolyAdenosine Diphosphate-Ribose Polymerase
LiOH	Lithium Hydroxide	PBS	Phosphate-Buffered Saline
Log	Logarithm	PCLAF	PCNA-Associated Factor
m	Multiplet	PCR	Polymerase Chain Reaction
m/z	Mass-to-charge Ratio	Pd(PPh₃)₂	Bis(triphenylphosphine)palladium(II)
iPrOH	Isopropanol	Cl₂	dichloride
MeOD	Duterated Methanol	PEG	PolyEthylene Glycol
MeOH	Methanol	PhMe	Toluene
MES	2-(N-Morpholino)EthaneSulfonic acid	PL	Proximity Labeling
MGMT	Human O ⁶ -MethylGuanine-DNA MethylTransferase	POI	Protein Of Interest
MgSO₄	Magnesium Sulfate	PyBOP	Benzotriazol-1-yl- oxytripyrrolidinophosphonium hexafluorophosphate
MHz	Mega Hertz	q	Quartet
Min	Minute / Minutes	RBM39	RNA-Binding Motif protein 39
MMR	Mismatch repair	RBR	RING-Between-RING
MMS	Methyl methanesulfonate	RING	Really Interesting New Gene
MS	Mass Spectrometry	RIPA	Radioimmunoprecipitation assay
MsCl	Methanesulfonyl chloride	RNA	Ribonucleic acid
NaAc	Sodium Acetate	RNAi	RNA Interference
NaI	Sodium Iodide	ROS	Reactive Oxygen Species
NaN₃	Sodium Azide	RP	Reverse Phase
NaOH	Sodium Hydroxide	rpm	Revolutions Per Minute
ⁿBu₄NCl	N-tetrabutyl ammonium chloride	RPMI	Roswell Park Memorial Institute
NCBI	National Center for Biotechnology Information		

Abbreviations

RT	Room temperature	TOF	Time Of Flight
s	Singlet	TRIS	Trisaminomethane
SAM	S-AdenosylMethionine	tRNA-6G	Tyrosine tDNA with six modified guanines
SCF	Skp1–Cullin–F-box	UBC	Ubiquitin Conjugation domain
SDS	Sodium Dodecyl Sulfate	UPLC-MS	Ultra Performance Liquid Chromatography – Tandem Mass Spectrometry
shRNA	Short Hairpin RNA	UPP	Ubiquitin-mediated Proteasomal Pathway
siRNA	Small Interfering RNA	UV	Ultra violet
SPRF	Single - Polypeptide RING-Finger	W	Wash
ssDNA	Single-Stranded DNA	WT	Wild Type
Strep	Streptavidin	XPC	DNA repair protein complementing XP-C cells
T	Thymine	δ	Chemical shifts
t	Triplet	μm	Micro-Meter
T3P	Propylphosphonic anhydride	μM	Micro-Molar
TBST	Tris-Buffered Saline with Tween-20		
TE	Tris and EDTA		
TEA	TriEthylAmine		
TET	TETracycline		
TFA	TriFluoroAcetic acid		
THF	TetraHydroFuran		
TMT	Tandem Mass Tag		

Abstract

O-6-methylguanine-DNA methyltransferase (MGMT) is a human DNA repair protein that removes O6-guanine alkylation through covalent and irreversible transalkylation from the guanine O6-alkyl group onto its reactive site cysteine. Following these events, alkylated MGMT is targeted for degradation via the ubiquitin-dependent proteasomal degradation pathway. The current study is categorized by three approaches.

In the first approach, the breadth of alkyl groups tolerated by MGMT was explored and it was found that modified guanine on single stranded DNA were recognized, showing for the first time the ability of MGMT to utilize substrates other than an alkylated DNA duplex. That data indicates that MGMT could play a role in more than just in direct DNA repair but could also be involved in RNA repair and maintenance.

The second approach focused on kinetics of MGMT degradation, shows that *de novo* protein synthesis, depends on the source of the alkylation. Additionally, alkylated MGMT seems to interact with more than one E3 ligase involved in protein proteasomal degradation, most likely from Cullin-RING family.

In the third and final approach, MGMT was used as a tool for targeted protein degradation. For the first time, the results show new application of MGMT as a titratable and reversible pharmacologically controlled tag, evaluated as a fusion to GFP, Luciferase and biotin ligase protein, such as Turbo ID and BioID.

Chapter 1

Theoretical background

1 Introduction to DNA damage and repair

Deoxyribonucleic acid (DNA) is the universal building block of life – its transcription and translation gives an organism the functions of growth, metabolism, reproduction and adaptation to the environment¹. All the information needed for an organism to function are encrypted in the DNA using four conserved nitrogenous bases (nucleobases): thymine (T), cytosine (C), adenine (A) and guanine (G)² as shown in Figure 1A. They can be divided in two groups: pyrimidines and purines³. Pyrimidines (thymine and cytosine) are six-membered heterocycles containing four carbon atoms and two nitrogen atoms, placed at 1 and 3 position of the ring^{4,5}. Purines (adenine and guanine) consist of the same pyrimidine moiety which is fused with an imidazole ring. The nucleobases together with a deoxyribose moiety form nucleosides. Nucleosides are linked to phosphate groups at the 3' and 5' position of the ribose, forming nucleotides (Figure 1, A and B). Finally, the nucleotides can assemble to larger oligonucleotides. The two most important naturally occurring oligonucleotides are the single-stranded RNA (with ribose as the sugar moiety) and the single – or double-stranded DNA (with deoxyribose as the sugar). A double helix is formed when adenine and guanine from one strand pair through hydrogen bonds with thymine and cytosine from the second strand, respectively (Figure 1C)⁶.

The base sequence and pairing within the helix are essential for an organism's ability to access, store and transcribe the encoded information. Therefore, it is crucial for the structure to remain unchanged. Environmental impact can cause a variety of transformations leading to mutations such as deamination, missing bases, chemical modification of bases, formation of pyrimidine dimers and strand breaks, which are discussed below.

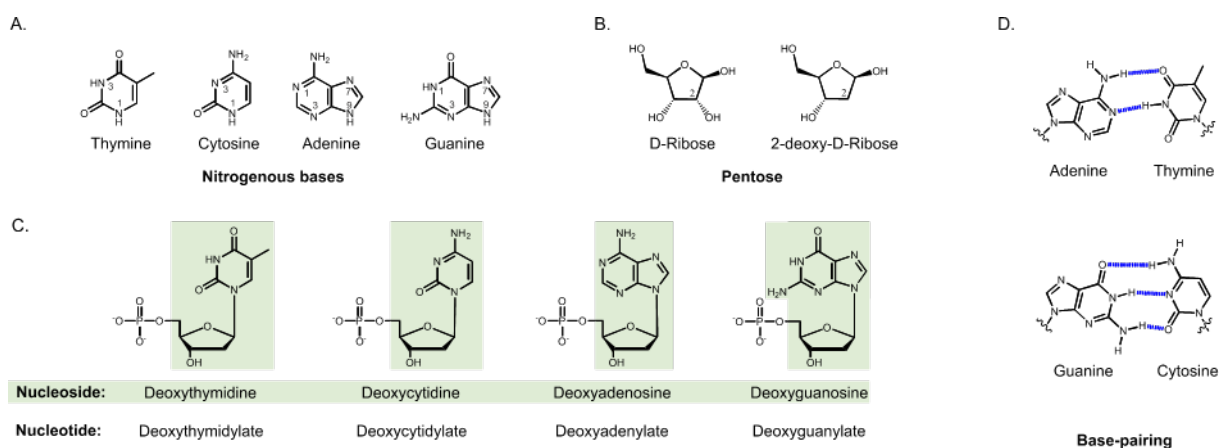


Figure 1. DNA building blocks. Four nucleobases (A) and pentoses (B); 2-deoxy-nucleosides and nucleotides (C); Purine and pyrimidine base pairing (D).

1.1 DNA damage

Modifications of nucleobase structures can be caused by endogenous or exogenous sources. Endogenous factors are either errors in DNA biosynthesis or metabolic processes affording toxic products^{7,8}. Mistakes in DNA synthesis do not change the chemical structure of a nucleobase but rather introduce a base-pair mismatch. The mismatch incorporates non-complementary Watson-Crick nucleotides, causing misalignment in the double helix. Exogenous factors cause

Chapter 1

changes in the nucleobase structure and they include: a) UV damage, b) radioactive radiation and c) active chemical damage: oxidation, deamination or alkylation⁹⁻¹¹.

1.1.1 UV damage:

The Sun irradiates the Earth with both UVA (320-340 nm) and UVB light (short UVB: 280-295 nm and long UVB: 295-320 nm). Both UVA and UVB are a major source of mutations in mammalian cells^{12,13}. Long UVB irradiation penetrates the skin and gets absorbed by the cellular DNA. After UVB-excitation, adjacent thymines in the DNA strand can be crosslinked to form amongst others either cyclobutane pyrimidine dimers (CPDs) or pyrimidine (6-4) pyrimidone photoproducts (6-4PP)¹⁴⁻¹⁶ shown in Figure 2.

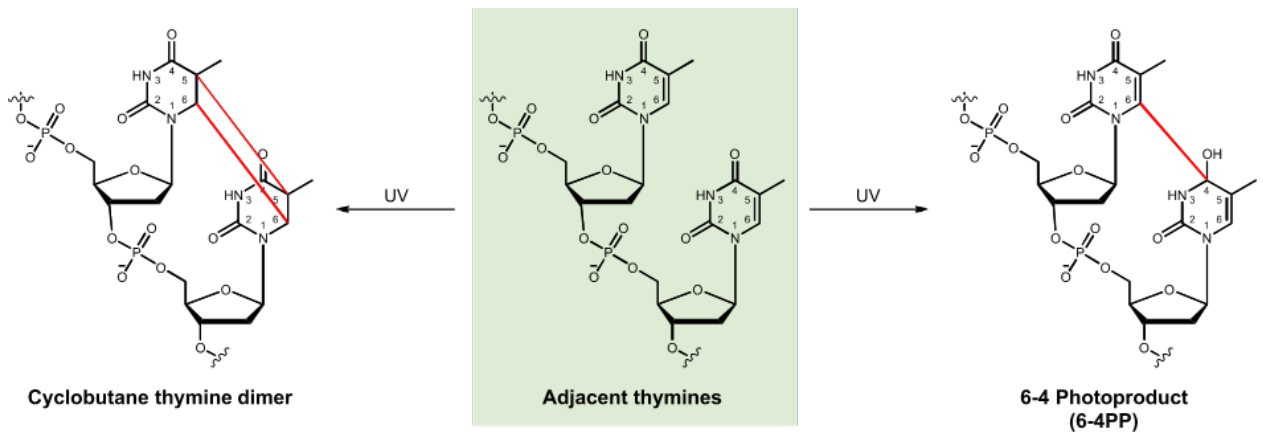


Figure 2. Lesion caused by UV irradiation affecting adjacent thymines. Formed cyclobutane dimers (left) or 6-4 photoproduct (right) is subject for repair.

1.1.2 High-energy radiation:

DNA exposure to high-energy radiation such as x-rays and gamma rays results in many lesions from single- or double-strand breaks are the most lethal for cells¹⁷⁻¹⁹ (Figure 3).

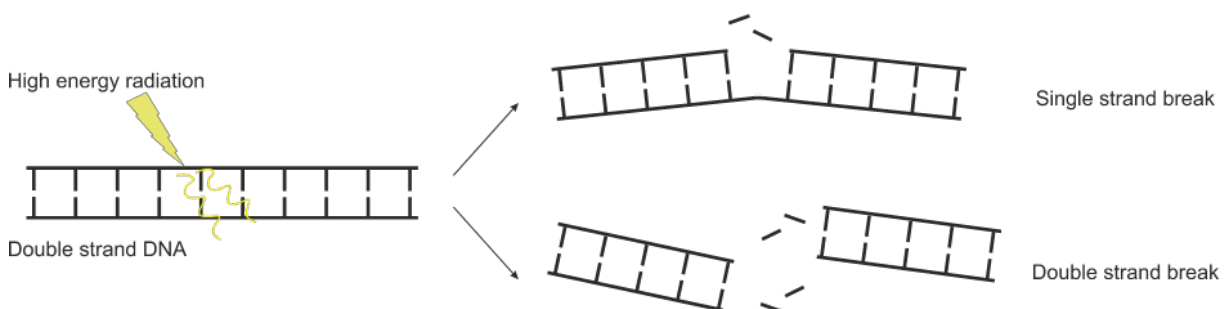


Figure 3. Radiation damage resulting in DNA single- or double- strand break as well as covalent crosslinking.

1.1.3 Active chemical damage:

- **Oxidation:**

Oxidative damage can be endogenous or exogenous, making it one of the most important and most common alterations of DNA^{20,21}. It is caused by reactive oxygen species (ROS) such as hydrogen peroxide, hydroxyl radicals and superoxide radicals, which are involved in epigenetic regulation. Of those, hydroxyl radicals are the most damaging ROS. A typical damage product caused by ROS is 8-oxo-7,8-dihydroguanine (8-oxoG)^{22,23} shown in Figure 4.

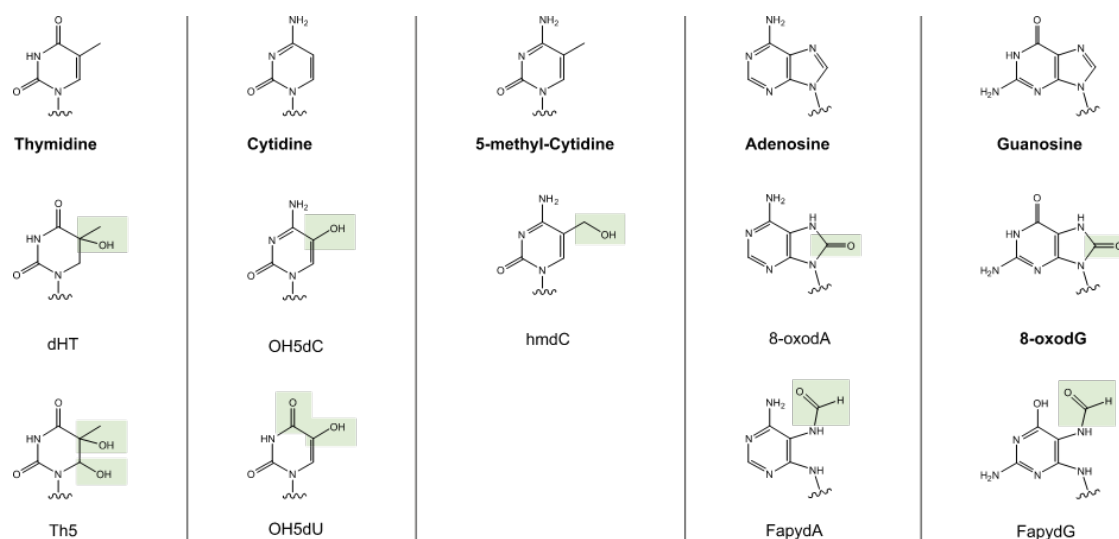


Figure 4. Most common oxidized nucleobases adapted from Svilar et al²⁴. dTH – 5,6-dihydrothymine; Th5 - 5-hydroxy-5,6-dihydrothymine; OH5dC - 5-hydroxy-2'-deoxycytidine; OH5dU - 5-hydroxy-2'-deoxyuridine; hmdC - 5-(hydroxymethyl)-2'-deoxycytidine; 8-oxoA - 8-oxo-7,8-dihydro-2'-deoxyadenosine; FapydA - 4,6-diamino-5-formamidopyrimidine; 8-oxodG - 8-oxo-7,8-dihydro-2'-deoxyguanosine; FapydG - 2,6-diamino-4-hydroxy-5-formamidopyrimidine.

- **Deamination:**

The process in which an amino group is removed from a molecule is called deamination. Spontaneous deamination affects the nucleobases; for example, uracil is formed from cytosine and thymine is formed upon deamination of 5-methylcytosine (Figure 5). The deamination of adenine to hypoxanthine results in a change in base pairing, as hypoxanthine no longer forms hydrogen bonds with thymine but is forming base pairs to cytosine instead.²⁵ Certain agents such as nitric oxide (NO) or peroxyxynitrite can catalyze the deamination processes^{26,27}.

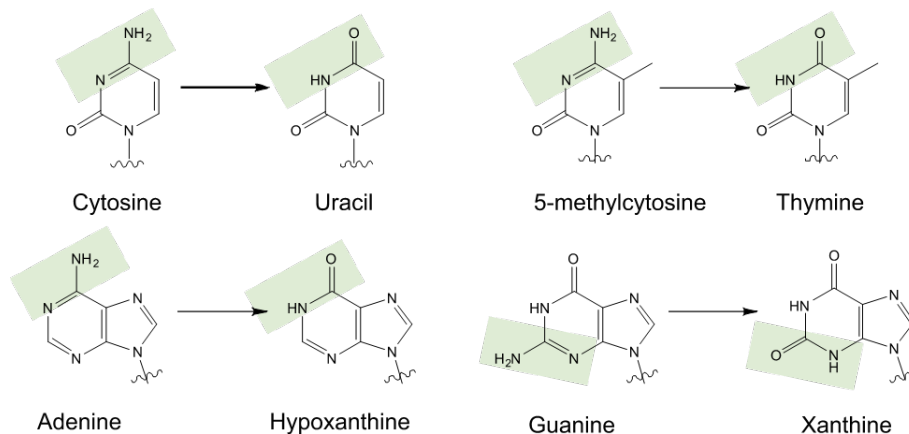


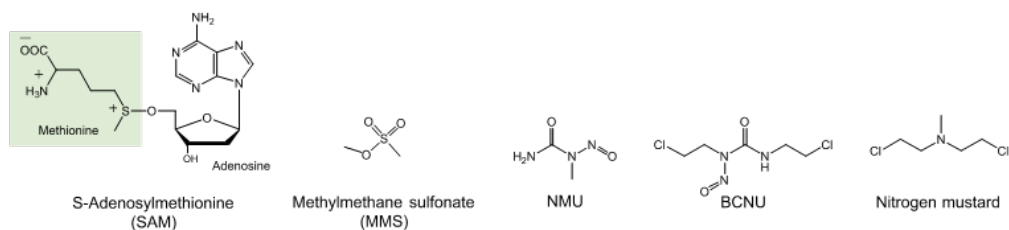
Figure 5. Deamination of cytosine, 5-methylcytosine, adenine and guanine.

Chapter 1

- **Alkylation:**

DNA alkylation refers to the covalent addition of alkyl groups to the nucleobases. If not repaired, those lesions can be toxic or mutagenic²⁸. The alkylation source can either be environmental (e.g. nitrogen mustard), part of a certain anti-cancer therapy (BCNU, MMS, NMU)²⁹ or it can come from the endogenous metabolic pathway (SAM: S-adenosylmethionine)^{30,31}. Alkylation modifications proceed either via S_N1 or via S_N2 mechanism³⁷. And independently of the source, alkylation can occur on different sites of the DNA, targeting either the heterocyclic bases or the DNA backbone^{32–36}. The most common alkylation agent following the S_N1 pathway is *N*-methyl-*N*'-nitrosourea (MNU)³⁵, while methylmethane sulfonate (MMS) and methyl halides (methyl chloride, nitrogen mustard or BCNU) pursue the S_N2 pathway^{35,38} (Figure 6C and D).

A.



B.

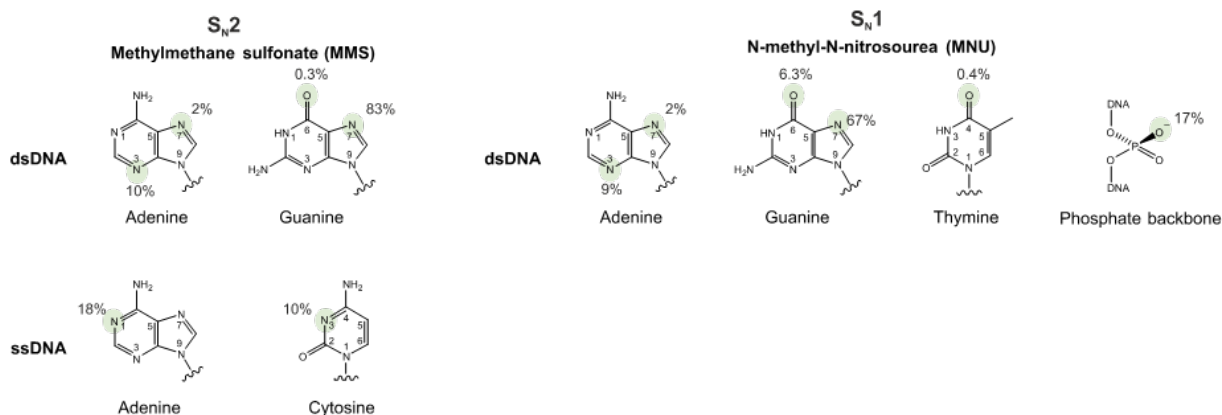


Figure 6. DNA alkylation. Common alkylating compounds (A); Frequency and location of S_N2 -type alkylating agent, on the example of MMS on dsDNA and ssDNA (B); Frequency and location of S_N1 -type alkylating agent, on the example of MNU on dsDNA and ssDNA. Adapted from^{28–38}.

1.2 Different DNA-repair mechanism

If DNA damage is left unrepaired, the damages can accumulate and consequently lead to cell apoptosis and possibly to oncogenesis^{39,40}. Given these consequences, cells have developed various DNA-repair mechanisms to preserve genome integrity. These include Mismatch repair (MMR), Base-excision repair (BER), Nucleotide-excision repair (NER), Strand break and direct repair - summarized in Table 1.

Source of damage	Type of damage	Repair mechanism
Replication error Mistakes during metabolism	Mismatch	Mismatch repair (MMR)
UV lesion	Oxidation: non-bulky damage. Deamination of cytosine to uracil	Base Excision repair (BER)
UV lesion	Oxidation: Bulky damage CPD, 6-4PP	Nucleotide excision repair (NER)
Replication error, Chromosome breakage, Telomere deprotection, Reactive Oxygen Species	Double strand break	Strand break repair
Alkylating agents, Endogenous metabolic pathway	Methylation Alkylation	Direct repair

Table 1. Summary of sources and types of DNA damage and corresponding repair mechanisms.

1.2.1 Mismatch repair

Mismatch lesions are caused during normal DNA metabolism, such as replication⁴¹. The thereby generated non-Watson – Crick base pairs as well as the misaligned strands are detected by the mismatch-recognition complex⁴². MMR is highly conserved between species and DNA strand specific^{42–44}. Defective MMR genes have been found to be responsible for the majority of hereditary non-polyposis colon cancers^{45,46}. During its mechanism of action MMR mediates the removal of single mispairing or mismatches involving several bases, and subsequently the resynthesis of the removed part from the freshly synthesized DNA strand⁷. To do so, the mismatch-recognition complex (in *E. coli*: MutSa, MutSb and MutLa, in humans their heterodimeric homologues^{44,47,48}) first recruits MutLb and MutLd to the mispair site. Secondly, the freshly recruited proteins activate the endonuclease which recognized the freshly synthesized strand based on its hemi-methylated status. Exonuclease cleaves few nucleobases surrounding the mismatch on the non-methylated, freshly synthesized strand containing the mismatch, initiating DNA resynthesis. Finally, DNA polymerase δ ⁴⁹ and ligase I to resynthesize the missing DNA fragment⁵⁰ and seal the nick⁵¹.

1.2.2 Excision repair

- **Base excision repair**

In humans there are two main excision pathways responsible for the removal of the majority of damaged bases in DNA: BER and NER⁵². BER is the primary pathway to remove non-bulky modifications caused by deamination of cytosine to uracil, oxidation caused by reactive oxygen species from the oxidative metabolism (8-oxoG thymine glycol), alkylation caused by normal cellular metabolites (SAM) as well as errors in DNA replication (misincorporation of dUTP or 8-oxo-dGTP)^{24,53,54}. In contrast, bulky lesions caused by UV-damage, as well as other crosslinks within the DNA structure are removed by NER^{8,55}.

In the BER process, four main steps can be identified for the excision and replacement of the modified nucleotide: lesion recognition, backbone cleavage, processing of loose ends and gap filling. Step 1 (lesion recognition) recruits DNA glycosylases to hydrolyze the *N*-glycosidic bond of the damaged base, leaving an apurinic or apyrimidinic site. During the second step the DNA backbone is cleaved by an AP endonuclease or by the DNA glycosylase bearing an AP lyase activity. Because of wide variety of endonucleases Step 3 can be processed via short or long pathway, called

Chapter 1

short-patch BER or long-patch BER, respectively. Short-patch repairs the gap of the size of one nucleotide, while the long-patch results in the gap of the size of 2-8 nucleotides⁵⁶⁻⁵⁸. Regardless the path the repair is followed by 5' to 3' DNA synthesis of the missing sequence. In the final step, the nick is sealed by the DNA ligase (Figure 7A).

- **Nucleotide excision repair (NER)**

NER is the most versatile DNA repair system facilitated by a machinery consisting of over twenty proteins⁵⁹, that acts on distorted DNA helices. As aforementioned, conformational changes in the helical structure can be induced by chemical adducts caused by cisplatin, benzopyrene or aflatoxin as well as by pyrimidine dimers induced by UV irradiation^{60,61}. In humans, NER's most important function⁶⁰ is to remove DNA damage caused by UV light^{8,55}. In order for the lesion to be recognized by NER the lesion must not only distort the DNA conformation but also covalently modify the DNA structure⁶²⁻⁶⁵. Similarly to BER, NER consists of four steps: damage recognition, formation of a stable precision complex around the damage site, excision of the damaged nucleotide and completion of DNA synthesis followed by ligation^{59,66}. The recognition step RAD 23B and cenatrin 2 (CENT2) initiate a multiprotein complex that marks the damaged strand on each side of the DNA^{67,68} (Figure 7B). Two different endonucleases associated with the multiprotein complex cleave DNA strand on 5' and 3' sides of the damage releasing fragment that is 24-32 nucleotides long^{59,69}. The gap is filled by either polymerase Pol δ , Pol ϵ or Pol κ and in the final step sealed by DNA ligase 1 or 3^{60,70,71} (Figure 7B).

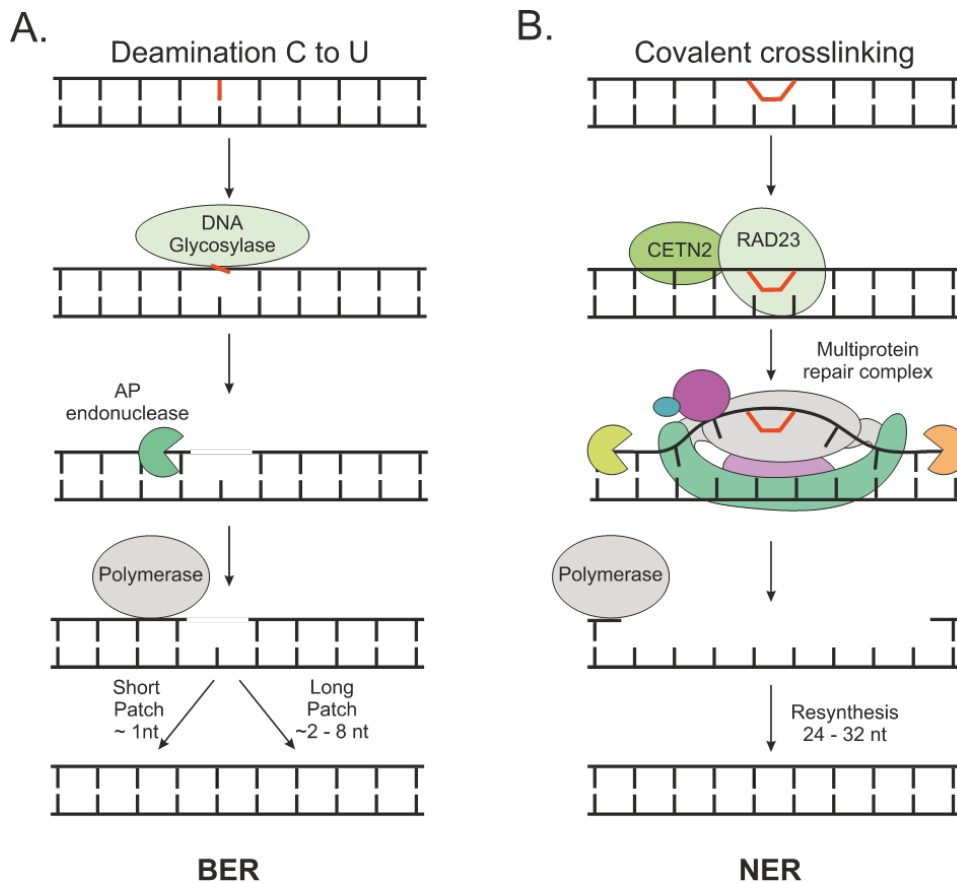


Figure 7. Simplified scheme of a base excision repair (BER) mechanism exemplified for a DNA damage caused by deamination of C to U. B) Simplified scheme of a nucleotide excision repair (NER) mechanism exemplified for a DNA damage caused by covalent crosslinking of two nucleotides. While BER is the primary pathway to remove short nucleotide patches (up to 8 nucleotides (nt)), NER involves in the repair of longer DNA fragments (24 to 32 nucleotides).

1.2.3 Strand-break repair

Chromosome breakage, dysfunctional replication-fork processing or telomere deprotection can lead to one- or two-DNA strand breaks. ROS are also contributing to DNA backbone breaks by destroying the deoxyribose residues. If the break occurs only on one strand the repair is conducted by BER. In contrast, if the break occurs on both strands, Double-Strand Break (DSB) repair pathway take place. Two main pathways are responsible for DSB repair: nonhomologous end-joining (NHEJ) or homologues repair (HR) (Figure 8)^{69,72–74}. The balance between these pathways depends on the phase of the cell cycle: HR is favored during S and G2 phase while NHEJ is dominant during G1 and G2 phase^{75–78}. In both cases the double-strand break is recognized by a Ku70–Ku80 heterodimer protein, which initiates the repair by either rapid NHEJ or longer HR⁷⁹ (Figure 8).

In humans, two NHEJ pathways can be distinguished: ‘classical’ NHEJ (cNHEJ) and alternative End Joining (aEJ) pathway, which is activated in the absence of cNHEJ⁸⁰. cNHEJ can form repair joints with up to 4bp of homology between two strands, while aEJ rapidly joins two ends of DNA strands without any homology^{74,81}. In contrast, HR proceeds via template-driven DNA synthesis^{73,82}.

During NHEJ, Ku70-Ku80 recruits DNA-PsCs to the damage site, which help with loose end processing and further recruitment of DNA ligase IV resulting in sealing the break⁸³. In contrast, HR involves exonucleases which digest fragment of 5’ strand of the DNA, resulting in long 3’ single-stranded DNA tail^{84,85}. To protect loose ends from recombination event, nuclear protein PARP1 (polyadenosine diphosphate-ribose polymerase-1) plays a role in temporary shielding DNA single-strand breaks and recruiting PRA proteins to further stabilize DNA^{86,87}. In the final step, template-driven DNA resynthesis takes place, facilitated by polymerase θ ⁸⁸ (Figure 8).

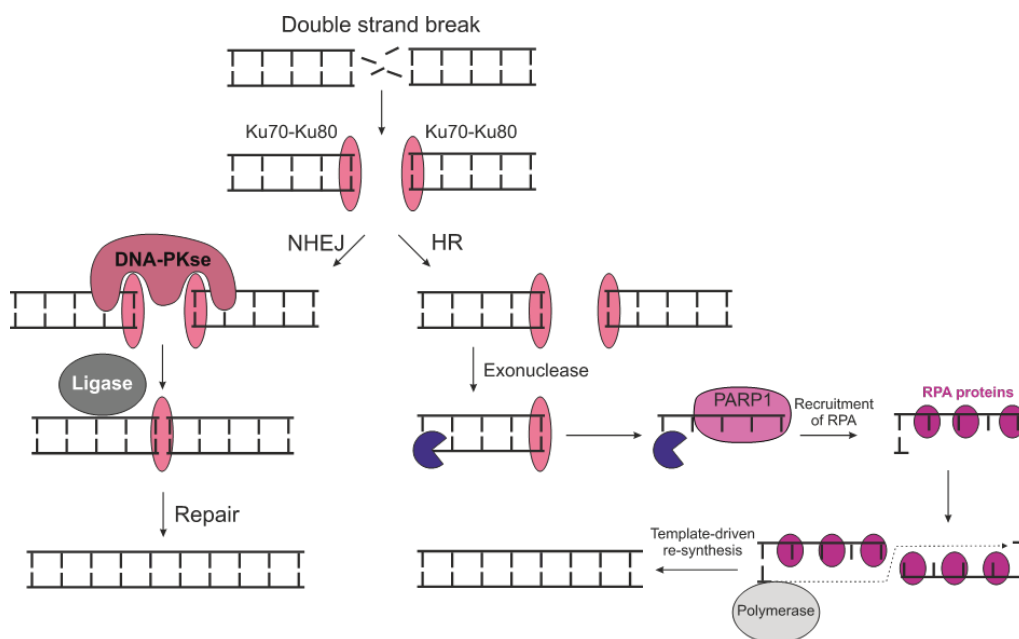


Figure 8. Simplified scheme of double-strand break repair that can either proceed via NHEJ (left) or HR pathway (right). Adapted from Scully et al.⁸⁵

1.2.4 Direct repair

Most DNA lesions are repaired via the excision-repair pathway, while only a small number of repairs are facilitated by direct damage repair. Glycosylases⁸⁶ and a family of AlkB proteins present in yeasts and bacteria facilitate the direct repair of alkylation damage of methylated nitrogen atoms^{87–90}, while the Ada protein in *E. coli* repairs

Chapter 1

alkylation on the phosphate backbone^{91,92}. Homologous proteins hABH2 and hABH3 were identified in humans and are known to repair 1-methyladenine and 3-methylcytosine^{93,94} (Figure 9). Significantly more mutagenic alkylation can arise from methylation of oxygen atoms, causing mispairing with thymine during replication. The alkyltransferase enzyme family, called O⁶-methylguanine-DNA methyltransferase (AGT, MGMT) conserved amongst yeast, bacteria and humans, repairs alkylation of the guanine O⁶-position^{95-98,99} (Figure 9). MGMT is particularly interesting due to its “suicide” character since it transfers the alkyl chain onto its own reactive cysteine (Cys145) that is part of an extremely conserved active site motif: Pro-Cys-His-Arg^{100,101}. After the alkyl transfer, MGMT is no longer active and undergoes rapid ubiquitin-mediated degradation¹⁰²⁻¹⁰⁵.

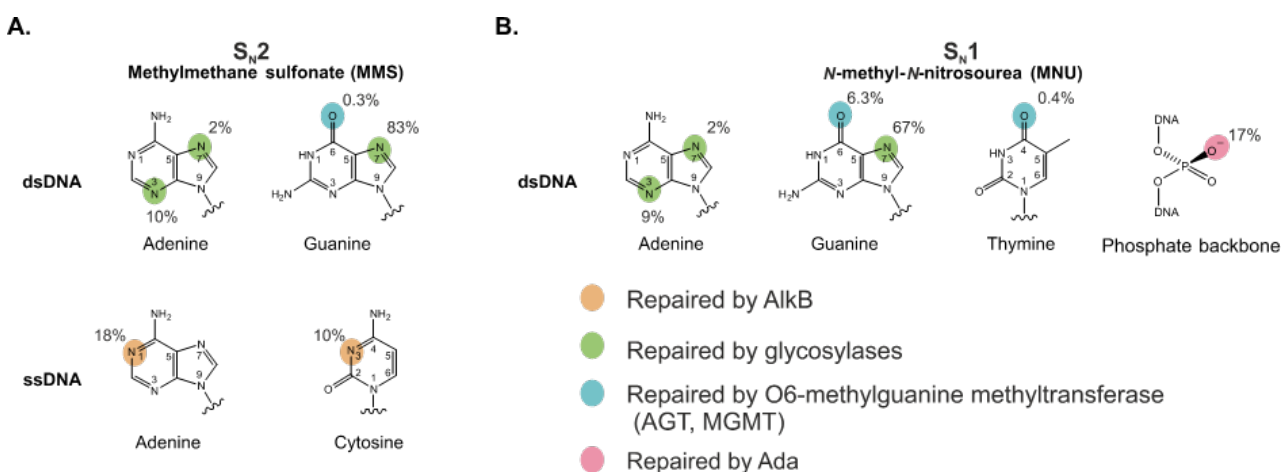


Figure 9. Typical sites of alkylation caused by S_N2 (A) and S_N1 (B) alkylation mechanism (see section 1) and enzymes involved in their direct repair.

1.3 DNA repair summary

All living organisms have developed conserved signaling networks that help keep their DNA structure intact. As the DNA is being exposed to many different damaging factors, a variety of repair pathways has evolved, each specializing in a particular type of damage repair. MMR mechanisms can either repair single mispairings or several bases, both of which are introduced by spontaneous mutation during normal DNA metabolism. BER as well as NER pathways recognize most of the damages caused by environmental factors. Both of those pathways remove a piece of DNA on one strand and re-synthesize it using the second strand as a template. BER leads to an abasic site after excision, while NER causes a gap of a few nucleotides. In the event of double-strand breaks, the repair cannot occur via template-repair and the two ends are brought back together by homologous recombination or by nonhomologous end-joining. HR is a slow process assisted by multi protein complex, while NHEJ fused strands without taking advantage of homology. Finally, direct repair is carried out by one protein that directly removes alkylation from either the nitrogen (AlkB, glycosylases), the oxygen of nucleobases (AGT) or removes alkylation from the phosphate backbone (Ada). AlkB-mediated repair results in recycling the protein, while AGT is itself methylated resulting in degradation and its *de novo* synthesis.

2 Protein degradation pathways

Healthy cells maintain a certain, well balanced intracellular protein level, called proteostasis^{109,110}, which is a dynamic equilibrium of protein synthesis and protein degradation. Changes between the two rates allow the cells to alter the protein levels in order to adapt to new environments, enable new functions or to restore proteostasis. An important aspect of proteostasis is the control over unnecessary, misfolded or unwanted proteins, which provides the control over gene expression^{111–113}, cell cycle and cellular differentiation^{114,115}, intracellular traffic of proteins^{116–118}, antigen presentation^{119,120} and furthermore provides the cell with energy and nutrients in starvation-stress environments^{121–123}.

Proteins can follow one of two degradation pathways in eukaryotic cells – proteasomal or lysosomal. Both pathways can have either a “clearance” function or a regulatory function, depending on cell environment¹²⁴. It has been suggested that proteins that undergo the proteasomal degradation pathway are relatively “short-lived” (several minutes), while proteins that are degraded via the lysosomal degradation pathway are usually “long-lived” (several days)^{121,125,126}. Additionally, it has been generally stated that proteasome-mediated degradation is controlled by highly specific interaction with inhibitors and activators^{127,128}, while lysosomal degradation is rather non-selective in nature.

2.1 Lysosomal degradation pathway

Lysosomes contain the highest concentration of proteases inside the cell¹²⁹ and play the role of a degradative and nutrient-recycling site for a wide variety of intracellular and extracellular proteins¹³⁰. Different mechanisms apply depending on type or origin of the protein:^{131,132}

- Extracellular, plasma membrane^{133–135} and secretory proteins^{135,136} are being transported through endocytosis. (Figure 10).
- Cytosolic proteins are transported to the lysosomes based on three main pathways: macro-autophagy (autophagy), micro-autophagy, and a chaperone-mediated direct transport^{131,132,136}, whereby the latter is only occurring in mammalian cells (Figure 11).

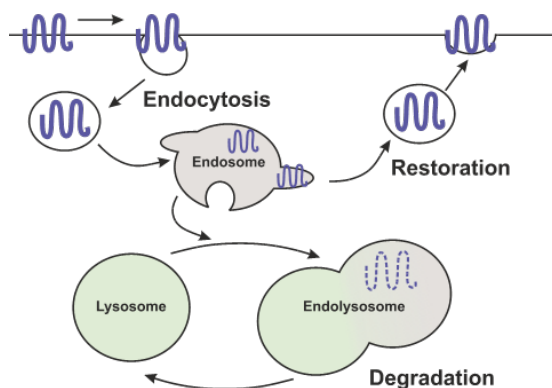


Figure 10. Endolysosomal degradation is based on the fusion of endosomes with the lysosome

Endolysosomal Degradation: Endolysosomal degradation is initiated on the cell surface by clathrin-dependent and clathrin-independent endocytosis^{137,138}. The newly formed endosome matures and is either retrieved back to the cell membrane or fused with the lysosome, where the proteins are degraded via proteolysis (Figure 10)^{139–141}.

Chapter 1

Macro-autophagy (Autophagy) (Figure 11), is operating under starvation conditions, during which bulk fragments of cytoplasm containing misfolded proteins are being recycled^{142,143}. During that process a distinguished intracellular double membrane vessel is formed, which surrounds a portion of the cytosol including proteins labeled for degradation.

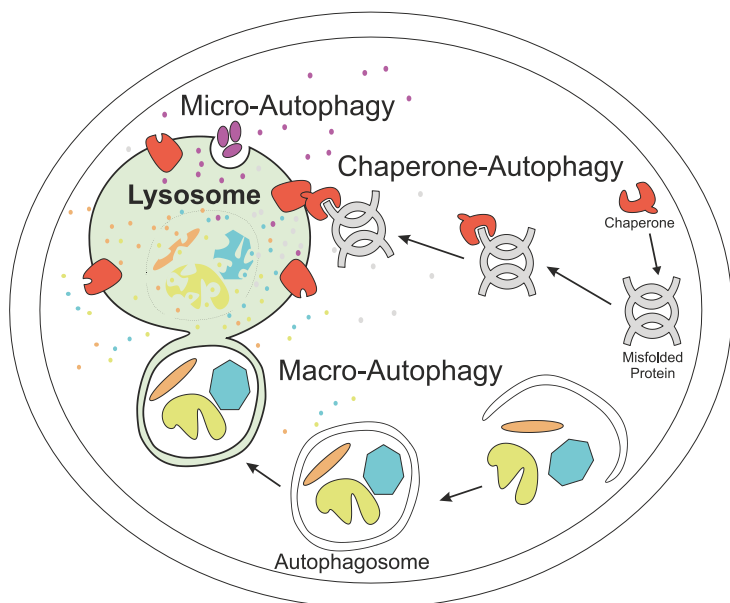


Figure 11. Mechanism of endolysosome formation. Membrane protein or endogenous protein (in purple) is captured in vesicle and transported to endosome, which can either restore it or degrade by fusion with the lysosome.

Such vessels called autophagosome transport its cargo to the lysosome. After fusion of the autophagosome with the lysosome, its content is being released and degraded by the proteases in the lysosome^{131,144,145}. Finally, digested protein building blocks (amino acids) diffuse through the lysosomal membrane back to the cytosol.

Micro-autophagy is responsible for the slow but continuous degradation of cytosolic proteins during normal nutritional conditions. In micro-autophagy, cargos are directly taken up via in-folding of the lysosome's outer membrane, resulting in the incorporation of small regions of cytosol and its degradation upon exposure to vacuolar hydrolases¹⁴⁶ (Figure 11).

2.2 Proteasomal pathway

Chaperone-Mediated Autophagy (Figure 11) has so far only been identified in mammalian cells. The pathway is activated by nutrient deprivation^{147,148} and is highly selective towards individual proteins, based on their amino acid sequence^{132,149}. Proteins are identified by a molecular chaperone in a one-by-one fashion and are delivered to the surface of the lysosomes¹⁵⁰. Membrane receptors recognize the target proteins¹⁵¹, unfold them and translocate them across the lysosomal membrane inside the lysosomal lumen where they are degraded¹⁵²⁻¹⁵⁴.

Proteins can either be degraded by the 20S or the 26S proteasome¹⁵⁵ (Figure 12A). The majority of cytosolic proteins are degraded by the 26S proteasome¹⁵⁶ in a highly selective, ubiquitin tagged and ATP-dependent manner¹⁵⁷⁻¹⁶⁰. While 20S-mediated degradation is not controlled by neither ubiquitin nor ATP, it recognizes partially or fully unfolded proteins caused by oxidation, mutations or aging¹⁶¹⁻¹⁶⁵.

The structure of the 26S proteasome shown in Figure 12B reveals a complex multi-protein structure, composed of a 20S doughnut-like proteasome core particle capped on one or both ends with the regulatory particle 19S. The 19S particle recognizes proteins that are labeled for degradation, opens a gate for substrate entry in the 20S core, unfolds the substrate and translocates the protein into 20S catalytic chamber, where it gets degraded¹⁶⁶⁻¹⁶⁹. The doughnut-like structure consists of four heptameric rings, which are stacked on top of each other and arranged in an α - β - β - α sequence¹⁷⁰. The N-termini of the α - subunit form a 'gate' by folding over the central pore and hindering the access to the proteolytic site, located on the β - subunit lumen¹⁷¹. Passage through this gate is the rate-limiting step that prevents degradation of non-specific proteins¹⁷². In contrast, 20S proteasome does not contain 19S regulatory subunits, allowing for more rapid degradation.

Irregularities in proteasome function directly affect the cell homeostasis. In the event of low proteasome efficiency, misfolded proteins accumulate forming toxic aggregates, leading to neurodegenerative diseases such as Parkinson's, Alzheimer's and Huntington's disease^{173,174}. On the other hand, too rapid degradation can cause the degradation of vital regulatory proteins, like tumor suppressors and cell cycle inhibitors, which can cause cancer genesis and progression¹⁷⁵ as well as autoimmune diseases or neurodegeneration¹⁷⁶.

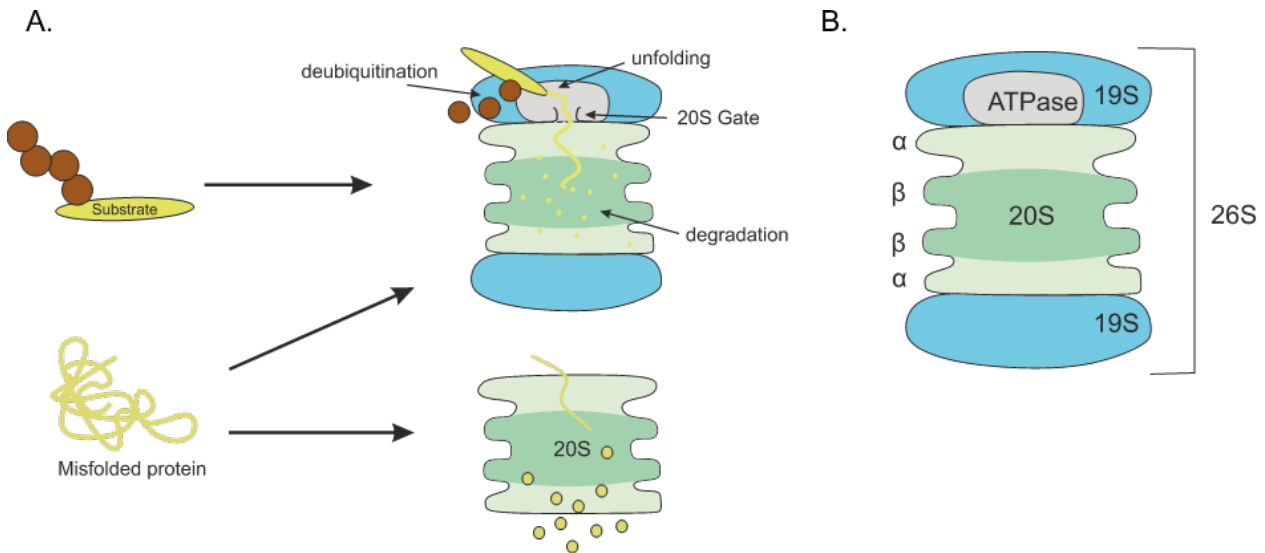


Figure 12. Protein degradation performed by either the 26S or the 20S proteasome (A). 26S proteasome containing the 20S core and two 19S regulatory particles (B). Adapted from Thibaudau et al.¹⁷⁷

2.2.1 Ubiquitin signaling

Ubiquitin is a small, highly conserved among species (Figure 13A), 8kD protein that act as a post-translational modifier¹⁷⁸ and regulates multiple cellular processes like the cell cycle, metabolic pathways, DNA stability as well as cellular trafficking involved in proteasomal protein degradation and autophagy^{179–182}. Ubiquitin gets covalently linked to proteins by forming an isopeptide bond between its C-terminal carboxylic acid and a lysine-residue of the target protein in a highly regulated cascade of enzymatic reactions involving the ubiquitin-activating enzyme (E1), ubiquitin conjugating enzyme (E2), and ubiquitin ligase enzyme (E3)^{183–185}. Ubiquitin itself consists of seven lysine residues: K6, K11, K27, K29, K33, K48 and K63 (Figure 13A) each of which can form an isopeptide bond with C-termini of another protein or another ubiquitin molecule. That way proteins can either be modified with a single ubiquitin moiety or with a chain of interconnected ubiquitin molecules. The sequence in which the interconnected ubiquitin chain is formed influences the fate of the labeled protein^{186–188}. Most common and best characterized are chains linked through K11, K48 and K63¹⁸⁹. Polyubiquitin chains linked through K48 are most often a signal that induces proteasomal degradation^{179,190,191} (Figure 13B). Chains linked through K63 are characteristic for inducing an autophagy degradation^{192–195} (Figure 13B), DNA repair and signal transduction¹⁹⁶. Single ubiquitin labeling (monoubiquitination) has been observed to play a role in endocytosis and transcriptional regulation¹⁸⁶. Additionally, non-lysine ubiquitin chains as well as mixed chain linkages have been observed, yet their role in cell signaling is poorly understood^{197,198}.

In the first step of the ubiquitin conjugation cascade, the cysteine residue in the active site of E1 covalently binds to ubiquitin via a high-energy thioester linkage (Figure 13C). That process is preceded by the formation of a ubiquitinyl adenylate intermediate from ubiquitin carboxyl group of Gly-76 and ATP^{199,200}. During the second step the activated ubiquitin is transferred to the cysteine residue of the E2 enzyme in a transacylation reaction (Figure 13C).

Chapter 1

Finally, E2 can either transfer the ubiquitin-residue directly to a protein substrate via an intermediate complex with E3 ligase or first transfer the residue on to E3 ligase (Figure 13C), which in turn will pass on the ubiquitin to the corresponding ϵ -amino group of lysine on the protein substrate²⁰¹. There are thousands of E2- and hundreds of E3-ligases in eukaryotic cells, each with a precise substrate activity. Different combinations of E2 and E3 enzymes result in a very selective protein tagging based on the/involving the formation of specific types ubiquitin intermolecular chains, which directly translates to the fate of each individual protein²⁰².

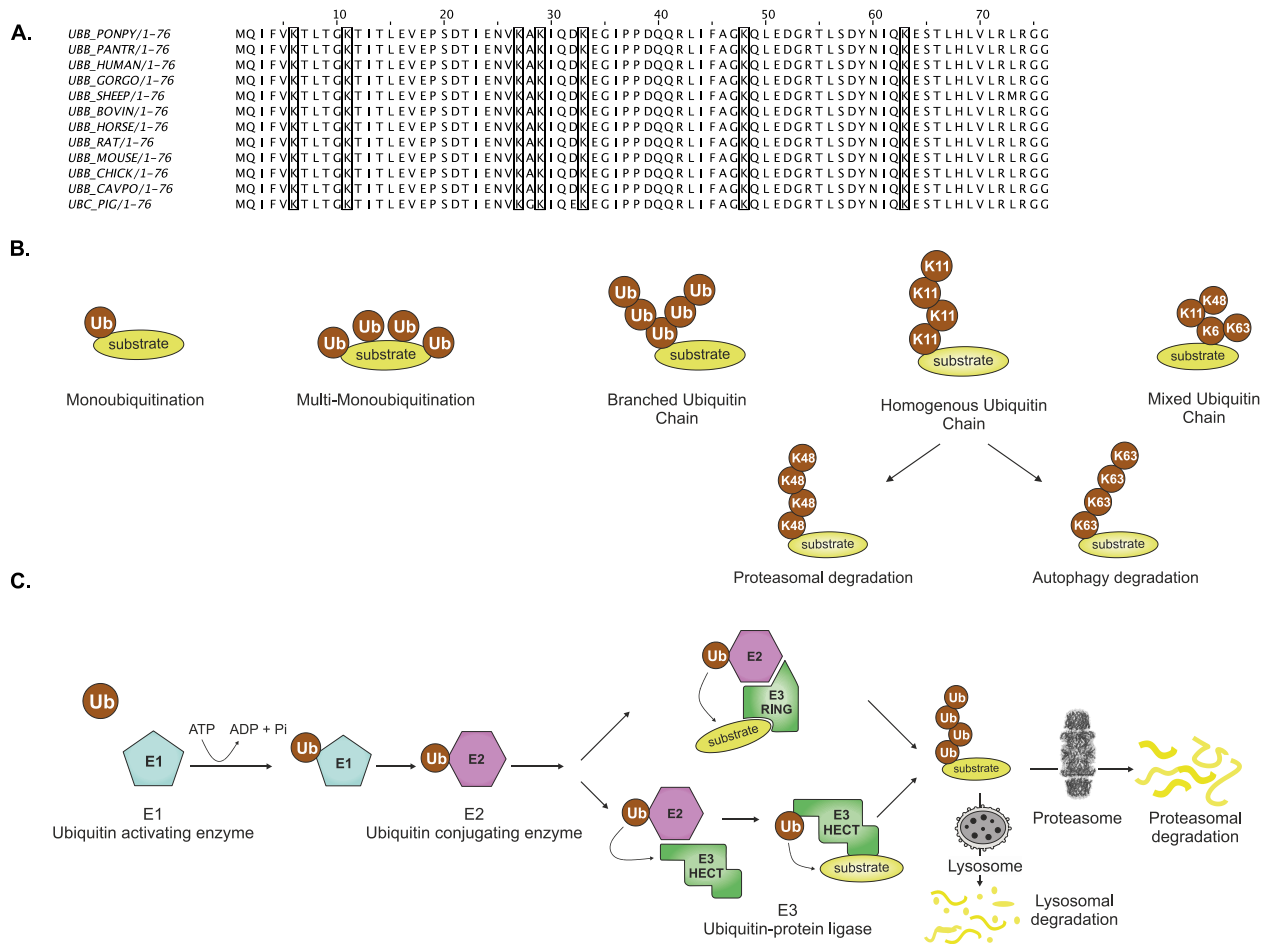


Figure 13. Ubiquitin-signaling. E1-E2-E3 ubiquitin conjugation pathway, can be a signal towards proteasomal or lysosomal degradation (A). Multiple alignments of reviewed eukaryotic ubiquitin protein sequences. [Source: UniProt®]. Highly conserved lysines K6, K11, K27, K29, K33, K48 and K63 relevant for ubiquitination are highlighted in boxes (B). Types of ubiquitin labeling and their signaling pathway (C).

2.2.2 Ubiquitin conjugating cascade

• Ubiquitin-Activating Enzyme – E1

E1 plays a crucial role in the ubiquitin signaling pathway, since only one functional E1 ligase has been identified, whose deletion is lethal^{203–206}. Apart from being involved in ubiquitin conjugating cascade, E1 interacts with so called ubiquitin-like family such as SUMO, NEDD8, ISG15, APG12, and UFM1²⁰⁷, increasing its role in cell signaling pathways. Structurally, three domains have been characterized in E1 enzyme:

- 1 - an adenylation domain with two ThiF-homology motifs. This domain is known to bind ATP and the corresponding E2 enzyme^{208–212}.
- 2- the catalytic cysteine domain (CCD), where the acyl group undergoes trans-acetylation with ubiquitin^{213,214}.
- 3- the C-terminal ubiquitin-folding domain which recruits E2 enzymes^{211,215–217}.

- **Ubiquitin-Conjugating Enzyme – E2**

The human genome encodes for more than forty E2 enzymes^{185,218,219}, which are further divided in seventeen subfamilies²²⁰. All E2s share a 130-150 amino acid long conserved domain in the catalytic core (or Ubiquitin Conjugation domain - UBC), which is the minimal unit sufficient for enzymatic activity. The UBC contains the catalytic cysteine, which forms a thioester bond with the ubiquitin that was previously activated by E1 enzyme. Especially crucial for the catalytic activity is the conserved negatively charged aspartic acid or serine, whereby studies have shown that phosphorylation of that serine increases E2 activity^{221–223} while its mutation abolishes UBC activity^{224,225}.

- **Ubiquitin-Protein Ligase – E3**

There are over 600 human E3 ligase genes which have been categorized into 5 main types based on their catalytic domains: *N*-end rule, Homologous to E6-AP Carboxyl Terminus (HECT), Really Interesting New Gene (RING), U-Box, or Ring-Between-Ring (RBR)²²⁶. Regardless the type, all E3 ligases can recognize characteristic motifs of E2 enzymes and catalyze the transfer of the ubiquitin molecule from the enzyme onto the target protein^{128,227}.

N-end rule family. E3 α was the first identified E3 ligase responsible for the degradation of two protein types²²⁸. The first type of substrates contains basic amino acids side chains such as Arg, Lys, and His at their *N*-termini^{229,230}, while the second type contains bulky and hydrophobic residues, like Leu, Ile, Phe, Trp, and Tyr at *N*-termini^{231–233}.

Homologous to E6-AP Carboxy Terminus (HECT) domain family (Figure 13C). The HECT family of proteins consists of at least 20 identified members²³⁴. Characteristic for the HECT domain is a conserved catalytic cysteine, which forms a ubiquitin-thioester intermediate and directly catalyzes substrate ubiquitination²³⁵.

Really Interesting New Gene (RING) finger family (Figure 13C)

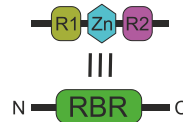
The RING family of E3 ligases serves as a platform to bring the ubiquitinated E2 ligase and the substrate in close proximity^{183,236}. All RING family members contain eight cysteine and histidine residues that can coordinate two zinc ions and are classified as either RING-H2 (with histidine at positions 4 and 5) or RING-HC (only one histidine at position 4)^{237–239}. Four main subfamilies can be distinguished: the (Skp1–Cullin–F-box) (SCF), Cul2–Elongin B–Elongin C (CBC), Anaphase - Promoting Complex (APC), and Single - Polypeptide RING-Finger (SPRF)¹⁸³. Many oncogenic E3 ligases like Cbl, Mdm2, inhibitors of apoptosis (IAPs) and Parkin belong to the SPRF type. Cbl mediates ubiquitylation of the epidermal growth factor (EGF) receptor^{240–242} and Mdm2 regulates tumor suppressor p53 turnover^{243–245}. Meanwhile IAPs are regulating programmed cell death^{246–248} and Parkin, one of the largest genes in human genome, plays a central role in mitophagy and mediates proteasomal degradation. It's loss of function causes dopaminergic neuronal death^{249–251}.

U-Box. U-box is structurally similar to the RING-finger family but lacks zinc-finger binding sites²⁵². Six mammalian U-box proteins have been isolated and shown to mediate ubiquitination in presence of E1 and E2 enzymes but in the absence of other E3 ligases, proving that the metal-chelating function is not essential for E3 activity²⁵². Additionally, deletion or point mutation of the conserved amino acids within this domain show loss of E3 activity²⁵³.

Chapter 1

RING-In-Between-RING (RBR) E3 Ubiquitin Ligases. RBR E3 ligases form a large family, schematically represented in Figure 14. Structurally, RBR proteins are characterized by two RING finger domains (RING1 and RING2) and a central zinc-binding domain, which is located between RING1 and RING2²⁵⁴⁻²⁵⁶. The RING domain serves as recruitment platform for E2 interaction, which then transfers ubiquitin from E2 enzyme to a cysteine in the RING-like domain²⁵⁴. Examples of RBR E3 ligases are Parkin, Parc and Dorfin (Figure 14)²⁵⁷.

RBR nomenclature:



RBR E3 ligases:

Parkin



Parc



Dorfin



Figure 14. Nomenclature and important examples of RBR E3 ligases adapted from reference²⁵⁷. R1- RING1, Zn – Zinc-binding domain, R2 – RING2.

2.3 Summary:

Cell proteostasis is a balance between protein synthesis and degradation, which allows for controlled gene expression, metabolism and growth. There are two main protein degradation pathways: lysosomal and proteasomal. The lysosomal pathway can be carried out via the endolysosome – for extracellular and membrane proteins, or via autophagy – for cytosolic proteins. Proteasomal degradation is carried out by 20S or 26S multi protein subunit. 26S proteasome consists of 20S as well as 19S domain, and it's ATP dependent. Both lysosomal and proteasomal pathways are controlled by ubiquitin signaling – a small protein containing seven lysine amino acids residues, which can form covalent modifications between each other or lysines of other proteins. These modifications are carried out by signaling cascades which include E1, E2 and E3 enzymes. Broad families of each enzyme type allow for a plethora of diversity in modification patterns. From these, polyubiquitin chains between ubiquitin K68 lysines are known to target proteins to lysosome, while polyubiquitin chains at lysine K48 target proteins to the proteasome.

3 Bibliography:

1. Mifflin, H. H. American heritage dictionary of the English language 5th edition. *New York* (2011).
2. Crick, F. The double helix: a personal view. *Nature* **248**, 766–769 (1974).
3. Hartman, S. C. & Buchanan, J. M. Nucleic acids, purines, pyrimidines (nucleotide synthesis). *Annu. Rev. Biochem.* **28**, 365–410 (1959).
4. Jardetzky, C. D. & Jardetzky, O. Investigation of the Structure of Purines, Pyrimidines, Ribose Nucleosides and Nucleotides by Proton Magnetic Resonance. III. *Journal of the American Chemical* (1960).
5. Jardetzky, C. D. & Jardetzky, O. Investigation of the Structure of Purines, Pyrimidines, Ribose Nucleosides and Nucleotides by Proton Magnetic Resonance. III. *Journal of the American Chemical* (1960).
6. Watson, J. D. & Crick, F. H. Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. *Nature* **171**, 737–738 (1953).
7. Jiricny, J. Postreplicative mismatch repair. *Cold Spring Harb. Perspect. Biol.* **5**, a012633 (2013).

8. Cadet, J. & Wagner, J. R. DNA base damage by reactive oxygen species, oxidizing agents, and UV radiation. *Cold Spring Harb. Perspect. Biol.* **5**, a012559 (2013).
9. Thomas, S., Lowe, J. E., Knowles, R. G., Green, I. C. & Green, M. H. L. Factors affecting the DNA damaging activity of superoxide and nitric oxide. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **402**, 77–84 (1998).
10. Kovacic, P. & Jacintho, J. D. Mechanisms of carcinogenesis: focus on oxidative stress and electron transfer. *Curr. Med. Chem.* **8**, 773–796 (2001).
11. Ridnour, L. A. *et al.* Nitric oxide regulates angiogenesis through a functional switch involving thrombospondin-1. *Proc. Natl. Acad. Sci. USA* **102**, 13147–13152 (2005).
12. You, Y. H. *et al.* Cyclobutane pyrimidine dimers are responsible for the vast majority of mutations induced by UVB irradiation in mammalian cells. *J. Biol. Chem.* **276**, 44688–44694 (2001).
13. Tewari, A., Sarkany, R. P. & Young, A. R. UVA1 induces cyclobutane pyrimidine dimers but not 6-4 photoproducts in human skin in vivo. *J. Invest. Dermatol.* **132**, 394–400 (2012).
14. Besaratinia, A. *et al.* Wavelength dependence of ultraviolet radiation-induced DNA damage as determined by laser irradiation suggests that cyclobutane pyrimidine dimers are the principal DNA lesions produced by terrestrial sunlight. *FASEB J.* **25**, 3079–3091 (2011).
15. Sugasawa, K. Molecular mechanisms of DNA damage recognition for mammalian nucleotide excision repair. *DNA Repair (Amst.)* **44**, 110–117 (2016).
16. Osakabe, A. *et al.* Structural basis of pyrimidine-pyrimidone (6-4) photoproduct recognition by UV-DDB in the nucleosome. *Sci. Rep.* **5**, 16330 (2015).
17. Gulston, M., de Lara, C., Jenner, T., Davis, E. & O'Neill, P. Processing of clustered DNA damage generates additional double-strand breaks in mammalian cells post-irradiation. *Nucleic Acids Res.* **32**, 1602–1609 (2004).
18. Núñez, M. I., McMillan, T. J., Valenzuela, M. T., Ruiz de Almodóvar, J. M. & Pedraza, V. Relationship between DNA damage, rejoining and cell killing by radiation in mammalian cells. *Radiother. Oncol.* **39**, 155–165 (1996).
19. Hsiao, Y. & Stewart, R. D. Monte Carlo simulation of DNA damage induction by x-rays and selected radioisotopes. *Phys. Med. Biol.* **53**, 233–244 (2008).
20. Breen, A. P. & Murphy, J. A. Reactions of oxyl radicals with DNA. *Free Radic. Biol. Med.* **18**, 1033–1077 (1995).
21. Beckman, K. B. & Ames, B. N. Oxidative decay of DNA. *J. Biol. Chem.* **272**, 19633–19636 (1997).
22. Valavanidis, A., Vlachogianni, T. & Fiotakis, C. 8-hydroxy-2'-deoxyguanosine (8-OHdG): A critical biomarker of oxidative stress and carcinogenesis. *J. Environ Sci Health C Environ Carcinog Ecotoxicol Rev* **27**, 120–139 (2009).
23. Kasai, H. Analysis of a form of oxidative DNA damage, 8-hydroxy-2'-deoxyguanosine, as a marker of cellular oxidative stress during carcinogenesis. *Mutat. Res.* **387**, 147–163 (1997).
24. Svilar, D., Goellner, E. M., Almeida, K. H. & Sobol, R. W. Base excision repair and lesion-dependent subpathways for repair of oxidative DNA damage. *Antioxid. Redox Signal.* **14**, 2491–2507 (2011).
25. Neddermann, P. *et al.* Cloning and expression of human G/T mismatch-specific thymine-DNA glycosylase. *J. Biol. Chem.* **271**, 12767–12774 (1996).
26. Wink, D. A. *et al.* DNA deaminating ability and genotoxicity of nitric oxide and its progenitors. *Science* **254**, 1001–1003 (1991).
27. Burney, S., Caulfield, J. L., Niles, J. C., Wishnok, J. S. & Tannenbaum, S. R. The chemistry of DNA damage from nitric oxide and peroxynitrite. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **424**, 37–49 (1999).
28. Drabløs, F. *et al.* Alkylation damage in DNA and RNA--repair mechanisms and medical significance. *DNA Repair (Amst.)* **3**, 1389–1407 (2004).
29. Lundin, C. *et al.* Methyl methanesulfonate (MMS) produces heat-labile DNA damage but no detectable in vivo DNA double-strand breaks. *Nucleic Acids Res.* **33**, 3799–3811 (2005).

Chapter 1

30. Rydberg, B. & Lindahl, T. Nonenzymatic methylation of DNA by the intracellular methyl group donor S-adenosyl-L-methionine is a potentially mutagenic reaction. *EMBO J.* **1**, 211–216 (1982).
31. Barrows, L. R. & Magee, P. N. Nonenzymatic methylation of DNA by S-adenosylmethionine in vitro. *Carcinogenesis* **3**, 349–351 (1982).
32. Saini, N. *et al.* Mutation signatures specific to DNA alkylating agents in yeast and cancers. *Nucleic Acids Res.* **48**, 3692–3707 (2020).
33. Shooter, K. V., Howse, R., Shah, S. A. & Lawley, P. D. The molecular basis for biological inactivation of nucleic acids. The action of methylating agents on the ribonucleic acid-containing bacteriophage R17. *Biochem. J.* **137**, 303–312 (1974).
34. Singer, B. & Grunberger, D. *Molecular biology of mutagens and carcinogens.* (2012).
35. Singer, B. & Grunberger, D. in *Molecular biology of mutagens and carcinogens* 45–96 (Springer US, 1983). doi:10.1007/978-1-4613-3772-0_4
36. Beranek, D. T. Distribution of methyl and ethyl adducts following alkylation with monofunctional alkylating agents. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **231**, 11–30 (1990).
37. Singer, B. & Grunberger, D. in *Molecular biology of mutagens and carcinogens* 45–96 (Springer US, 1983). doi:10.1007/978-1-4613-3772-0_4
38. Jaiswal, A. S. & Narayan, S. SN2 DNA-alkylating agent-induced phosphorylation of p53 and activation of p21 gene expression. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **500**, 17–30 (2002).
39. Mitchell, J. R., Hoeijmakers, J. H. J. & Niedernhofer, L. J. Divide and conquer: nucleotide excision repair battles cancer and ageing. *Curr. Opin. Cell Biol.* **15**, 232–240 (2003).
40. Hoeijmakers, J. H. Genome maintenance mechanisms for preventing cancer. *Nature* **411**, 366–374 (2001).
41. Modrich, P. & Lahue, R. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* **65**, 101–133 (1996).
42. Li, G.-M. Mechanisms and functions of DNA mismatch repair. *Cell Res.* **18**, 85–98 (2008).
43. Kunkel, T. A. & Erie, D. A. DNA mismatch repair. *Annu. Rev. Biochem.* **74**, 681–710 (2005).
44. Drummond, J. T., Li, G. M., Longley, M. J. & Modrich, P. Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. *Science* **268**, 1909–1912 (1995).
45. Lynch, H. T. & de la Chapelle, A. Genetic susceptibility to non-polyposis colorectal cancer. *J. Med. Genet.* **36**, 801–818 (1999).
46. Peltomaki, P. Role of DNA mismatch repair defects in the pathogenesis of human cancer. *Journal of clinical oncology* (2003).
47. Li, G. M. & Modrich, P. Restoration of mismatch repair to nuclear extracts of H6 colorectal tumor cells by a heterodimer of human MutL homologs. *Proc. Natl. Acad. Sci. USA* **92**, 1950–1954 (1995).
48. Papadopoulos, N. *et al.* Mutation of a mutL homolog in hereditary colon cancer. *Science* **263**, 1625–1629 (1994).
49. Longley, M. J., Pierce, A. J. & Modrich, P. DNA polymerase delta is required for human mismatch repair in vitro. *J. Biol. Chem.* **272**, 10917–10921 (1997).
50. Zhang, Y. *et al.* Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell* **122**, 693–705 (2005).
51. Kadyrov, F. A., Dzantiev, L., Constantin, N. & Modrich, P. Endonucleolytic function of MutLalpha in human mismatch repair. *Cell* **126**, 297–308 (2006).
52. Braithwaite, E., Wu, X. & Wang, Z. Repair of DNA lesions: mechanisms and relative repair efficiencies. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **424**, 207–219 (1999).
53. Otterlei, M. *et al.* Post-replicative base excision repair in replication foci. *EMBO J.* **18**, 3834–3844 (1999).
54. Krokan, H. E., Standal, R. & Slupphaug, G. DNA glycosylases in the base excision repair of DNA. *Biochem. J.*

325 (Pt 1), 1–16 (1997).

55. Ding, S. *et al.* Structural, energetic and dynamic properties of guanine(C8)-thymine(N3) cross-links in DNA provide insights on susceptibility to nucleotide excision repair. *Nucleic Acids Res.* **40**, 2506–2517 (2012).
56. Pascucci, B., Stucki, M., Jónsson, Z. O., Dogliotti, E. & Hübscher, U. Long patch base excision repair with purified human proteins. DNA ligase I as patch size mediator for DNA polymerases delta and epsilon. *J. Biol. Chem.* **274**, 33696–33702 (1999).
57. Matsumoto, Y. *et al.* Reconstitution of proliferating cell nuclear antigen-dependent repair of apurinic/apyrimidinic sites with purified human proteins. *J. Biol. Chem.* **274**, 33703–33708 (1999).
58. Krokan, H. E., Nilsen, H., Skorpen, F., Otterlei, M. & Slupphaug, G. Base excision repair of DNA in mammalian cells. *FEBS Lett.* **476**, 73–77 (2000).
59. Wood, R. D. Nucleotide excision repair in mammalian cells. *J. Biol. Chem.* **272**, 23465–23468 (1997).
60. Marteijn, J. A., Lans, H., Vermeulen, W. & Hoeijmakers, J. H. J. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat. Rev. Mol. Cell Biol.* **15**, 465–481 (2014).
61. Zhang, Y., Rohde, L. H. & Wu, H. Involvement of nucleotide excision and mismatch repair mechanisms in double strand break repair. *Curr. Genomics* **10**, 250–258 (2009).
62. Moggs, J. G., Szymkowski, D. E., Yamada, M., Karran, P. & Wood, R. D. Differential human nucleotide excision repair of paired and mispaired cisplatin-DNA adducts. *Nucleic Acids Res.* **25**, 480–491 (1997).
63. Bessho, T., Mu, D. & Sancar, A. Initiation of DNA interstrand cross-link repair in humans: the nucleotide excision repair system makes dual incisions 5' to the cross-linked base and removes a 22- to 28-nucleotide-long damage-free strand. *Mol. Cell. Biol.* **17**, 6822–6830 (1997).
64. Geacintov, N. E. *et al.* Thermodynamic and structural factors in the removal of bulky DNA adducts by the nucleotide excision repair machinery. *Biopolymers* **65**, 202–210 (2002).
65. Hess, M. T., Schwitter, U., Petretta, M., Giese, B. & Naegeli, H. Bipartite substrate discrimination by human nucleotide excision repair. *Proc. Natl. Acad. Sci. USA* **94**, 6664–6669 (1997).
66. de Laat, W. L., Jaspers, N. G. & Hoeijmakers, J. H. Molecular mechanism of nucleotide excision repair. *Genes Dev.* **13**, 768–785 (1999).
67. Masutani, C. *et al.* Purification and cloning of a nucleotide excision repair complex involving the xeroderma pigmentosum group C protein and a human homologue of yeast RAD23. *EMBO J.* **13**, 1831–1843 (1994).
68. Nishi, R. *et al.* Centrin 2 stimulates nucleotide excision repair by interacting with xeroderma pigmentosum group C protein. *Mol. Cell. Biol.* **25**, 5664–5674 (2005).
69. Lindahl, T. & Wood, R. D. Quality control by DNA repair. *Science* **286**, 1897–1905 (1999).
70. Kuraoka, I. *et al.* Removal of oxygen free-radical-induced 5',8-purine cyclodeoxynucleosides from DNA by the nucleotide excision-repair pathway in human cells. *Proc. Natl. Acad. Sci. USA* **97**, 3832–3837 (2000).
71. Spampinato, C. P. Protecting DNA from errors and damage: an overview of DNA repair mechanisms in plants compared to mammals. *Cell Mol. Life Sci.* **74**, 1693–1709 (2017).
72. Lindahl, T., Satoh, M. S., Poirier, G. G. & Klungland, A. Post-translational modification of poly(ADP-ribose) polymerase induced by DNA strand breaks. *Trends Biochem. Sci.* **20**, 405–411 (1995).
73. Le Rhun, Y., Kirkland, J. B. & Shah, G. M. Cellular responses to DNA damage in the absence of Poly(ADP-ribose) polymerase. *Biochem. Biophys. Res. Commun.* **245**, 1–10 (1998).
74. Pâques, F. & Haber, J. E. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **63**, 349–404 (1999).
75. Sung, P. & Klein, H. Mechanism of homologous recombination: mediators and helicases take on regulatory functions. *Nat. Rev. Mol. Cell Biol.* **7**, 739–750 (2006).
76. Hartlerode, A. J. & Scully, R. Mechanisms of double-strand break repair in somatic mammalian cells. *Biochem. J.* **423**, 157–168 (2009).

Chapter 1

77. Takata, M. *et al.* Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J.* **17**, 5497–5508 (1998).
78. Aylon, Y., Liefshitz, B. & Kupiec, M. The CDK regulates repair of double-strand breaks by homologous recombination during the cell cycle. *EMBO J.* **23**, 4868–4875 (2004).
79. Ira, G. *et al.* DNA end resection, homologous recombination and DNA damage checkpoint activation require CDK1. *Nature* **431**, 1011–1017 (2004).
80. Huertas, P., Cortés-Ledesma, F., Sartori, A. A., Aguilera, A. & Jackson, S. P. CDK targets Sae2 to control DNA-end resection and homologous recombination. *Nature* **455**, 689–692 (2008).
81. Chang, H. H. Y., Pannunzio, N. R., Adachi, N. & Lieber, M. R. Non-homologous DNA end joining and alternative pathways to double-strand break repair. *Nat. Rev. Mol. Cell Biol.* **18**, 495–506 (2017).
82. Pannunzio, N. R., Watanabe, G. & Lieber, M. R. Nonhomologous DNA end-joining for repair of DNA double-strand breaks. *J. Biol. Chem.* **293**, 10512–10523 (2018).
83. Symington, L. S. Role of RAD52 Epistasis Group Genes in Homologous Recombination and Double-Strand Break Repair. *Microbiol. Mol. Biol. Rev.* **66**, 630–670 (2002).
84. Britton, S., Coates, J. & Jackson, S. P. A new method for high-resolution imaging of Ku foci to decipher mechanisms of DNA double-strand break repair. *J. Cell Biol.* **202**, 579–595 (2013).
85. Scully, R., Panday, A., Elango, R. & Willis, N. A. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat. Rev. Mol. Cell Biol.* **20**, 698–714 (2019).
86. Wyatt, M. D., Allan, J. M., Lau, A. Y., Ellenberger, T. E. & Samson, L. D. 3-methyladenine DNA glycosylases: structure, function, and biological importance. *Bioessays* **21**, 668–676 (1999).
87. Sedgwick, B. Repairing DNA-methylation damage. *Nat. Rev. Mol. Cell Biol.* **5**, 148–157 (2004).
88. Kataoka, H., Yamamoto, Y. & Sekiguchi, M. A new gene (*alkB*) of *Escherichia coli* that controls sensitivity to methyl methane sulfonate. *J. Bacteriol.* **153**, 1301–1307 (1983).
89. Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T. & Sedgwick, B. Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature* **419**, 174–178 (2002).
90. Falnes, P. Ø., Johansen, R. F. & Seeberg, E. AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature* **419**, 178–182 (2002).
91. Lindahl, T., Sedgwick, B., Sekiguchi, M. & Nakabeppu, Y. Regulation and expression of the adaptive response to alkylating agents. *Annu. Rev. Biochem.* **57**, 133–157 (1988).
92. Sedgwick, B., Robins, P., Totty, N. & Lindahl, T. Functional domains and methyl acceptor sites of the *Escherichia coli* *ada* protein. *J. Biol. Chem.* **263**, 4430–4433 (1988).
93. Aas, P. A. *et al.* Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature* **421**, 859–863 (2003).
94. Duncan, T. *et al.* Reversal of DNA alkylation damage by two human dioxygenases. *Proc. Natl. Acad. Sci. USA* **99**, 16660–16665 (2002).
95. Pegg, A. E. Repair of O6-alkylguanine by alkyltransferases. *Mutation Research/Reviews in Mutation Research* **462**, 83–100 (2000).
96. Margison, G. P. & Santibáñez-Koref, M. F. O6-alkylguanine-DNA alkyltransferase: role in carcinogenesis and chemotherapy. *Bioessays* **24**, 255–266 (2002).
97. Kyrtopoulos, S. A. DNA adducts in humans after exposure to methylating agents. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **405**, 135–143 (1998).
98. Kleibl, K. Molecular mechanisms of adaptive response to alkylating agents in *Escherichia coli* and some remarks on O6-methylguanine DNA-methyltransferase in other organisms. *Mutation Research/Reviews in Mutation Research* **512**, 67–84 (2002).
99. Dumenco, L. L., Allay, E., Norton, K. & Gerson, S. L. The prevention of thymic lymphomas in transgenic mice

- by human O6-alkylguanine-DNA alkyltransferase. *Science* **259**, 219–222 (1993).
100. Tubbs, J. L., Pegg, A. E. & Tainer, J. A. DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O6-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair (Amst.)* **6**, 1100–1115 (2007).
 101. Demple, B. *et al.* Active site and complete sequence of the suicidal methyltransferase that counters alkylation mutagenesis. *Proc. Natl. Acad. Sci. USA* **82**, 2688–2692 (1985).
 102. Pegg, A. E., Dolan, M. E. & Moschel, R. C. Structure, function, and inhibition of O6-alkylguanine-DNA alkyltransferase. *Prog Nucleic Acid Res Mol Biol* **51**, 167–223 (1995).
 103. Lindahl, T., Demple, B. & Robins, P. Suicide inactivation of the E. coli O6-methylguanine-DNA methyltransferase. *EMBO J.* **1**, 1359–1363 (1982).
 104. Xu-Welliver, M. & Pegg, A. E. Degradation of the alkylated form of the DNA repair protein, O(6)-alkylguanine-DNA alkyltransferase. *Carcinogenesis* **23**, 823–830 (2002).
 105. Srivenugopal, K. S., Yuan, X. H., Friedman, H. S. & Ali-Osman, F. Ubiquitination-dependent proteolysis of O6-methylguanine-DNA methyltransferase in human and murine tumor cells following inactivation with O6-benzylguanine or 1,3-bis(2-chloroethyl)-1-nitrosourea. *Biochemistry* **35**, 1328–1334 (1996).
 106. Schimke, R. Regulation of protein degradation in mammalian tissues, 177-228. *MUNRO HN*
 107. Labbadia, J. & Morimoto, R. I. The biology of proteostasis in aging and disease. *Annu. Rev. Biochem.* **84**, 435–464 (2015).
 108. Kaneko, M. *et al.* Genome-wide identification and gene expression profiling of ubiquitin ligases for endoplasmic reticulum protein degradation. *Sci. Rep.* **6**, 30955 (2016).
 109. O'Brien, R. M. & Granner, D. K. Regulation of gene expression by insulin. *Biochem. J.* **278 (Pt 3)**, 609–619 (1991).
 110. Hochstrasser, M. Ubiquitin, proteasomes, and the regulation of intracellular protein degradation. *Curr. Opin. Cell Biol.* **7**, 215–223 (1995).
 111. Finley, D. V. Chau. *Ubiquitination*
 112. Kristensen, A. R., Gsponer, J. & Foster, L. J. Protein synthesis rate is the predominant regulator of protein expression during differentiation. *Mol. Syst. Biol.* **9**, 689 (2013).
 113. Doherty, F. J. & Mayer, R. J. Intracellular protein degradation. (1992).
 114. Zuzarte, M. *et al.* Intracellular traffic of the K⁺ channels TASK-1 and TASK-3: role of N- and C-terminal sorting signals and interaction with 14-3-3 proteins. *J. Physiol. (Lond.)* **587**, 929–952 (2009).
 115. Shao, D., Okuse, K. & Djamgoz, M. B. A. Protein-protein interactions involving voltage-gated sodium channels: Post-translational regulation, intracellular trafficking and functional expression. *Int. J. Biochem. Cell Biol.* **41**, 1471–1481 (2009).
 116. Blum, J. S., Wearsch, P. A. & Cresswell, P. Pathways of antigen processing. *Annu. Rev. Immunol.* **31**, 443–473 (2013).
 117. Dudziak, D. *et al.* Differential antigen processing by dendritic cell subsets in vivo. *Science* **315**, 107–111 (2007).
 118. Mortimore, G. E. & Pösö, A. R. Intracellular protein catabolism and its control during nutrient deprivation and supply. *Annu. Rev. Nutr.* **7**, 539–564 (1987).
 119. Kamada, Y., Sekito, T. & Ohsumi, Y. Autophagy in yeast: a TOR-mediated response to nutrient starvation. *Curr. Top. Microbiol. Immunol.* **279**, 73–84 (2004).
 120. Cuervo, A. M. & Dice, J. F. Lysosomes, a meeting point of proteins, chaperones, and proteases. *J. Mol. Med.* (1998).
 121. Ding, W.-X. & Yin, X.-M. Sorting, recognition and activation of the misfolded protein degradation pathways through macroautophagy and the proteasome. *Autophagy* **4**, 141–150 (2008).
 122. Fuertes, G., Villarroya, A. & Knecht, E. Role of proteasomes in the degradation of short-lived proteins in human

Chapter 1

- fibroblasts under various growth conditions. *Int. J. Biochem. Cell Biol.* **35**, 651–664 (2003).
123. Fuertes, G., Martín De Llano, J. J., Villarroja, A., Rivett, A. J. & Knecht, E. Changes in the proteolytic activities of proteasomes and lysosomes in human fibroblasts produced by serum withdrawal, amino-acid deprivation and confluent conditions. *Biochem. J.* **375**, 75–86 (2003).
 124. Chau, V. *et al.* A multiubiquitin chain is confined to specific lysine in a targeted short-lived protein. *Science* **243**, 1576–1583 (1989).
 125. Ciechanover, A. The ubiquitin-proteasome proteolytic pathway. *Cell* **79**, 13–21 (1994).
 126. Bohley, P. & Seglen, P. O. Proteases and proteolysis in the lysosome. *Experientia* **48**, 151–157 (1992).
 127. Mortimore, G. E. Mechanism and regulation of induced and basal protein degradation in liver. *Lysosomes: their role in protein* (1987).
 128. Dunn, W. A. Autophagy and related mechanisms of lysosome-mediated protein degradation. *Trends Cell Biol.* **4**, 139–143 (1994).
 129. Dice, J. F. Peptide sequences that target cytosolic proteins for lysosomal proteolysis. *Trends Biochem. Sci.* **15**, 305–309 (1990).
 130. Steinman, R. M., Mellman, I. S., Muller, W. A. & Cohn, Z. A. Endocytosis and the recycling of plasma membrane. *J. Cell Biol.* **96**, 1–27 (1983).
 131. Hare, J. F. Mechanisms of membrane protein turnover. *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes* **1031**, 71–90 (1990).
 132. Cullen, P. J. & Steinberg, F. To degrade or not to degrade: mechanisms and significance of endocytic recycling. *Nat. Rev. Mol. Cell Biol.* **19**, 679–696 (2018).
 133. Marzella, L. Autophagy, microautophagy and crinophagy as mechanisms for protein degradation. *Lysosomes: Their role in protein breakdown* (1987).
 134. Kirchhausen, T., Owen, D. & Harrison, S. C. Molecular structure, function, and dynamics of clathrin-mediated membrane traffic. *Cold Spring Harb. Perspect. Biol.* **6**, a016725 (2014).
 135. Mayor, S., Parton, R. G. & Donaldson, J. G. Clathrin-independent pathways of endocytosis. *Cold Spring Harb. Perspect. Biol.* **6**, (2014).
 136. Mayor, S., Presley, J. F. & Maxfield, F. R. Sorting of membrane components from endosomes and subsequent recycling to the cell surface occurs by a bulk flow process. *J. Cell Biol.* **121**, 1257–1269 (1993).
 137. Mellman, I. Endocytosis and molecular sorting. *Annu. Rev. Cell Dev. Biol.* **12**, 575–625 (1996).
 138. Jovic, M., Sharma, M., Rahajeng, J. & Caplan, S. The early endosome: a busy sorting station for proteins at the crossroads. *Histol Histopathol* **25**, 99–112 (2010).
 139. Klionsky, D. J. The molecular machinery of autophagy: unanswered questions. *J. Cell Sci.* **118**, 7–18 (2005).
 140. Yorimitsu, T. & Klionsky, D. J. Autophagy: molecular machinery for self-eating. *Cell Death Differ.* **12 Suppl 2**, 1542–1552 (2005).
 141. Kopitz, J., Kisen, G. O., Gordon, P. B., Bohley, P. & Seglen, P. O. Nonselective autophagy of cytosolic enzymes by isolated rat hepatocytes. *J. Cell Biol.* **111**, 941–953 (1990).
 142. Mortimore, G. E. & Reeta Pösö, A. The lysosomal pathway of intracellular proteolysis in liver: Regulation by amino acids. *Adv. Enzyme Regul.* **25**, 257–276 (1986).
 143. Mortimore, G. E., Lardeux, B. R. & Adams, C. E. Regulation of microautophagy and basal protein turnover in rat liver. Effects of short-term starvation. *J. Biol. Chem.* **263**, 2506–2512 (1988).
 144. Neff, N. T., Bourret, L., Miao, P. & Dice, J. F. Degradation of proteins microinjected into IMR-90 human diploid fibroblasts. *J. Cell Biol.* **91**, 184–194 (1981).
 145. Dice, J. F. Altered degradation of proteins microinjected into senescent human fibroblasts. *J. Biol. Chem.* **257**, 14624–14627 (1982).
 146. Chiang, H. L. & Dice, J. F. Peptide sequences that target proteins for enhanced degradation during serum

- withdrawal. *J. Biol. Chem.* **263**, 6797–6805 (1988).
147. Chiang, H. L., Terlecky, S. R., Plant, C. P. & Dice, J. F. A role for a 70-kilodalton heat shock protein in lysosomal degradation of intracellular proteins. *Science* **246**, 382–385 (1989).
 148. Cuervo, A. M. & Dice, J. F. A receptor for the selective uptake and degradation of proteins by lysosomes. *Science* **273**, 501–503 (1996).
 149. Terlecky, S. R. & Dice, J. F. Polypeptide import and degradation by isolated lysosomes. *J. Biol. Chem.* **268**, 23490–23495 (1993).
 150. Agarraberes, F. A., Terlecky, S. R. & Dice, J. F. An intralysosomal hsp70 is required for a selective pathway of lysosomal protein degradation. *J. Cell Biol.* **137**, 825–834 (1997).
 151. Cuervo, A. M., Knecht, E., Terlecky, S. R. & Dice, J. F. Activation of a selective pathway of lysosomal proteolysis in rat liver by prolonged starvation. *Am. J. Physiol.* **269**, C1200-8 (1995).
 152. Layfield, R., Lowe, J. & Bedford, L. The ubiquitin–proteasome system and neurodegenerative disorders. *Essays Biochem* **41**, 157–171 (2005).
 153. McKinnon, C. & Tabrizi, S. J. The ubiquitin-proteasome system in neurodegeneration. *Antioxid. Redox Signal.* **21**, 2302–2321 (2014).
 154. Dou, Q. P., Smith, D. M. & Daniel, K. G. Interruption of tumor cell cycle progression through proteasome inhibition: implications for cancer therapy. *PROGRESS IN CELL* (2003).
 155. Schmidt, M. & Finley, D. Regulation of proteasome activity in health and disease. *Biochim. Biophys. Acta* **1843**, 13–25 (2014).
 156. Jang, H. H. Regulation of protein degradation by proteasomes in cancer. *J. Cancer Prev.* **23**, 153–161 (2018).
 157. Glickman, M. H. & Ciechanover, A. The ubiquitin-proteasome proteolytic pathway: destruction for the sake of construction. *Physiol. Rev.* **82**, 373–428 (2002).
 158. Desautels, M. & Goldberg, A. L. Liver mitochondria contain an ATP-dependent, vanadate-sensitive pathway for the degradation of proteins. *Proc. Natl. Acad. Sci. USA* **79**, 1869–1873 (1982).
 159. Duque-Magalhães, M. C. & Gualberto, J. M. Regulation of mitochondrial proteolysis. *FEBS Lett.* **210**, 142–146 (1987).
 160. Lippincott, J. Degradation from the endoplasmic reticulum: disposing of newly synthesized proteins. *Schwartz*
 161. Simpson, M. V. The release of labeled amino acids from the proteins of rat liver slices. *J. Biol. Chem.* **201**, 143–154 (1953).
 162. Pickering, A. M. & Davies, K. J. A. Degradation of damaged proteins: the main function of the 20S proteasome. *Prog Mol Biol Transl Sci* **109**, 227–248 (2012).
 163. Aiken, C. T., Kaake, R. M., Wang, X. & Huang, L. Oxidative stress-mediated regulation of proteasome complexes. *Mol. Cell Proteomics* **10**, R110.006924 (2011).
 164. Ben-Nissan, G. & Sharon, M. Regulating the 20S proteasome ubiquitin-independent degradation pathway. *Biomolecules* **4**, 862–884 (2014).
 165. Raynes, R., Pomatto, L. C. D. & Davies, K. J. A. Degradation of oxidized proteins by the proteasome: Distinguishing between the 20S, 26S, and immunoproteasome proteolytic pathways. *Mol. Aspects Med.* **50**, 41–55 (2016).
 166. Davies, K. J. Degradation of oxidized proteins by the 20S proteasome. *Biochimie* **83**, 301–310 (2001).
 167. DeMartino, G. N. *et al.* Identification, purification, and characterization of a PA700-dependent activator of the proteasome. *J. Biol. Chem.* **271**, 3112–3118 (1996).
 168. Adams, G. M., Crotchet, B., Slaughter, C. A., DeMartino, G. N. & Gogol, E. P. Formation of proteasome-PA700 complexes directly correlates with activation of peptidase activity. *Biochemistry* **37**, 12927–12932 (1998).
 169. DeMartino, G. N. & Slaughter, C. A. The proteasome, a novel protease regulated by multiple mechanisms. *J. Biol. Chem.* **274**, 22123–22126 (1999).

Chapter 1

170. Voges, D., Zwickl, P. & Baumeister, W. The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu. Rev. Biochem.* **68**, 1015–1068 (1999).
171. Groll, M. *et al.* Structure of 20S proteasome from yeast at 2.4 Å resolution. *Nature* **386**, 463–471 (1997).
172. Groll, M. *et al.* A gated channel into the proteasome core particle. *Nat. Struct. Biol.* **7**, 1062–1067 (2000).
173. Köhler, A. *et al.* The axial channel of the proteasome core particle is gated by the Rpt2 ATPase and controls both substrate entry and product release. *Mol. Cell* **7**, 1143–1152 (2001).
174. Thibaut, T. A. & Smith, D. M. A practical review of proteasome pharmacology. *Pharmacol. Rev.* **71**, 170–197 (2019).
175. Ciechanover, A., Elias, S., Heller, H., Ferber, S. & Hershko, A. Characterization of the heat-stable polypeptide of the ATP-dependent proteolytic system from reticulocytes. *J. Biol. Chem.* **255**, 7525–7528 (1980).
176. Finley, D. Recognition and processing of ubiquitin-protein conjugates by the proteasome. *Annu. Rev. Biochem.* **78**, 477–513 (2009).
177. Grumati, P. & Dikic, I. Ubiquitin signaling and autophagy. *J. Biol. Chem.* **293**, 5404–5413 (2018).
178. Chen, R.-H., Chen, Y.-H. & Huang, T.-Y. Ubiquitin-mediated regulation of autophagy. *J. Biomed Sci* **26**, 80 (2019).
179. Le Guerroué, F. & Youle, R. J. Ubiquitin signaling in neurodegenerative diseases: an autophagy and proteasome perspective. *Cell Death Differ.* (2020). doi:10.1038/s41418-020-00667-x
180. Deshaies, R. J. & Joazeiro, C. A. P. RING domain E3 ubiquitin ligases. *Annu. Rev. Biochem.* **78**, 399–434 (2009).
181. Schulman, B. A. & Harper, J. W. Ubiquitin-like protein activation by E1 enzymes: the apex for downstream signalling pathways. *Nat. Rev. Mol. Cell Biol.* **10**, 319–331 (2009).
182. Ye, Y. & Rape, M. Building ubiquitin chains: E2 enzymes at work. *Nat. Rev. Mol. Cell Biol.* **10**, 755–764 (2009).
183. Haglund, K., Di Fiore, P. P. & Dikic, I. Distinct monoubiquitin signals in receptor endocytosis. *Trends Biochem. Sci.* **28**, 598–603 (2003).
184. Hicke, L. & Dunn, R. Regulation of membrane protein transport by ubiquitin and ubiquitin-binding proteins. *Annu. Rev. Cell Dev. Biol.* **19**, 141–172 (2003).
185. Krappmann, D. & Scheidereit, C. A pervasive role of ubiquitin conjugation in activation and termination of IkappaB kinase pathways. *EMBO Rep.* **6**, 321–326 (2005).
186. Hicke, L., Schubert, H. L. & Hill, C. P. Ubiquitin-binding domains. *Nat. Rev. Mol. Cell Biol.* **6**, 610–621 (2005).
187. Thrower, J. S., Hoffman, L., Rechsteiner, M. & Pickart, C. M. Recognition of the polyubiquitin proteolytic signal. *EMBO J.* **19**, 94–102 (2000).
188. van Nocker, S. & Vierstra, R. D. Multiubiquitin chains linked through lysine 48 are abundant in vivo and are competent intermediates in the ubiquitin proteolytic pathway. *J. Biol. Chem.* **268**, 24766–24773 (1993).
189. Ferreira, J. V., Soares, A. R., Ramalho, J. S., Pereira, P. & Girao, H. K63 linked ubiquitin chain formation is a signal for HIF1A degradation by Chaperone-Mediated Autophagy. *Sci. Rep.* **5**, 10210 (2015).
190. Kirkin, V. *et al.* A role for NBR1 in autophagosomal degradation of ubiquitinated substrates. *Mol. Cell* **33**, 505–516 (2009).
191. McKeon, J. E., Sha, D., Li, L. & Chin, L.-S. Parkin-mediated K63-polyubiquitination targets ubiquitin C-terminal hydrolase L1 for degradation by the autophagy-lysosome system. *Cell Mol. Life Sci.* **72**, 1811–1824 (2015).
192. Tan, J. M. M. *et al.* Lysine 63-linked ubiquitination promotes the formation and autophagic clearance of protein inclusions associated with neurodegenerative diseases. *Hum. Mol. Genet.* **17**, 431–439 (2008).
193. Chan, N. L. & Hill, C. P. Defining polyubiquitin chain topology. *Nature structural biology* (2001).
194. Vosper, J. M. D. *et al.* Ubiquitylation on canonical and non-canonical sites targets the transcription factor neurogenin for ubiquitin-mediated proteolysis. *J. Biol. Chem.* **284**, 15458–15468 (2009).
195. McDowell, G. S., Kucerova, R. & Philpott, A. Non-canonical ubiquitylation of the proneural protein Ngn2 occurs

in both *Xenopus* embryos and mammalian cells. *Biochem. Biophys. Res. Commun.* **400**, 655–660 (2010).

196. Gehrke, P. P. & Jennissen, H. P. ATP-dependent proteolysis and the role of ubiquitin in rabbit cardiac muscle. *Biol. Chem. Hoppe. Seyler.* **368**, 691–708 (1987).
197. Haas, A. L., Warms, J. V., Hershko, A. & Rose, I. A. Ubiquitin-activating enzyme. Mechanism and role in protein-ubiquitin conjugation. *J. Biol. Chem.* **257**, 2543–2548 (1982).
198. Hershko, A., Heller, H., Elias, S. & Ciechanover, A. Components of ubiquitin-protein ligase system. Resolution, affinity purification, and role in protein breakdown. *J. Biol. Chem.* **258**, 8206–8214 (1983).
199. Ciechanover, A. Intracellular protein degradation: from a vague idea through the lysosome and the ubiquitin-proteasome system and onto human diseases and drug targeting. *Bioorg. Med. Chem.* **21**, 3400–3410 (2013).
200. McGrath, J. P., Jentsch, S. & Varshavsky, A. UBA 1: an essential yeast gene encoding ubiquitin-activating enzyme. *EMBO J.* **10**, 227–236 (1991).
201. Lambert-Smith, I. A., Saunders, D. N. & Yerbury, J. J. The pivotal role of ubiquitin-activating enzyme E1 (UBA1) in neuronal health and neurodegeneration. *Int. J. Biochem. Cell Biol.* **123**, 105746 (2020).
202. Moudry, P. *et al.* Ubiquitin-activating enzyme UBA1 is required for cellular response to DNA damage. *Cell Cycle* **11**, 1573–1582 (2012).
203. Groen, E. J. N. & Gillingwater, T. H. UBA1: at the crossroads of ubiquitin homeostasis and neurodegeneration. *Trends Mol. Med.* **21**, 622–632 (2015).
204. Jentsch, S. & Pyrowolakis, G. Ubiquitin and its kin: how close are the family ties? *Trends Cell Biol.* **10**, 335–342 (2000).
205. Lake, M. W., Wuebbens, M. M., Rajagopalan, K. V. & Schindelin, H. Mechanism of ubiquitin activation revealed by the structure of a bacterial MoeB-MoaD complex. *Nature* **414**, 325–329 (2001).
206. Duda, D. M., Walden, H., Sfoudouris, J. & Schulman, B. A. Structural analysis of *Escherichia coli* ThiF. *J. Mol. Biol.* **349**, 774–786 (2005).
207. Lehmann, C., Begley, T. P. & Ealick, S. E. Structure of the *Escherichia coli* ThiS-ThiF complex, a key component of the sulfur transfer system in thiamin biosynthesis. *Biochemistry* **45**, 11–19 (2006).
208. Walden, H., Podgorski, M. S. & Schulman, B. A. Insights into the ubiquitin transfer cascade from the structure of the activating enzyme for NEDD8. *Nature* **422**, 330–334 (2003).
209. Lois, L. M. & Lima, C. D. Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. *EMBO J.* **24**, 439–451 (2005).
210. Szczepanowski, R. H., Filipek, R. & Bochtler, M. Crystal structure of a fragment of mouse ubiquitin-activating enzyme. *J. Biol. Chem.* **280**, 22006–22011 (2005).
211. Schäfer, A., Kuhn, M. & Schindelin, H. Structure of the ubiquitin-activating enzyme loaded with two ubiquitin molecules. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **70**, 1311–1320 (2014).
212. Bencsath, K. P., Podgorski, M. S., Pagala, V. R., Slaughter, C. A. & Schulman, B. A. Identification of a multifunctional binding site on Ubc9p required for Smt3p conjugation. *J. Biol. Chem.* **277**, 47938–47945 (2002).
213. Huang, D. T. *et al.* Structural basis for recruitment of Ubc12 by an E2 binding domain in NEDD8's E1. *Mol. Cell* **17**, 341–350 (2005).
214. Huang, D. T. *et al.* Basis for a ubiquitin-like protein thioester switch toggling E1-E2 affinity. *Nature* **445**, 394–398 (2007).
215. van Wijk, S. J. L. & Timmers, H. T. M. The family of ubiquitin-conjugating enzymes (E2s): deciding between life and death of proteins. *FASEB J.* **24**, 981–993 (2010).
216. Wenzel, D. M., Stoll, K. E. & Klevit, R. E. E2s: structurally economical and functionally replete. *Biochem. J.* **433**, 31–42 (2011).
217. Michelle, C., Vourc'h, P., Mignon, L. & Andres, C. R. What was the set of ubiquitin and ubiquitin-like conjugating enzymes in the eukaryote common ancestor? *J. Mol. Evol.* **68**, 616–628 (2009).

Chapter 1

218. Sarcevic, B., Mawson, A., Baker, R. T. & Sutherland, R. L. Regulation of the ubiquitin-conjugating enzyme hHR6A by CDK-mediated phosphorylation. *EMBO J.* **21**, 2009–2018 (2002).
219. Wood, A., Schneider, J., Dover, J., Johnston, M. & Shilatifard, A. The Bur1/Bur2 complex is required for histone H2B monoubiquitination by Rad6/Bre1 and histone methylation by COMPASS. *Mol. Cell* **20**, 589–599 (2005).
220. Valimberti, I., Tiberti, M., Lambrugh, M., Sarcevic, B. & Papaleo, E. E2 superfamily of ubiquitin-conjugating enzymes: constitutively active or activated through phosphorylation in the catalytic cleft. *Sci. Rep.* **5**, 14849 (2015).
221. Sung, P., Prakash, S. & Prakash, L. Stable ester conjugate between the *Saccharomyces cerevisiae* RAD6 protein and ubiquitin has no biological activity. *J. Mol. Biol.* **221**, 745–749 (1991).
222. Sommer, T. & Jentsch, S. A protein translocation defect linked to ubiquitin conjugation at the endoplasmic reticulum. *Nature* **365**, 176–179 (1993).
223. Zheng, N. & Shabek, N. Ubiquitin ligases: structure, function, and regulation. *Annu. Rev. Biochem.* **86**, 129–157 (2017).
224. Hershko, A. & Ciechanover, A. The ubiquitin system. *Annu. Rev. Biochem.* **67**, 425–479 (1998).
225. Kwon, Y. T., Xia, Z., Davydov, I. V., Lecker, S. H. & Varshavsky, A. Construction and analysis of mouse strains lacking the ubiquitin ligase UBR1 (E3 α) of the N-end rule pathway. *Mol. Cell. Biol.* **21**, 8007–8021 (2001).
226. Kwon, Y. T., Lévy, F. & Varshavsky, A. Bivalent inhibitor of the N-end rule pathway. *J. Biol. Chem.* **274**, 18135–18139 (1999).
227. Varshavsky, A. The N-end rule: functions, mysteries, uses. *Proc. Natl. Acad. Sci. USA* **93**, 12142–12149 (1996).
228. Reiss, Y., Kaim, D. & Hershko, A. Specificity of binding of NH₂-terminal residue of proteins to ubiquitin-protein ligase. Use of amino acid derivatives to characterize specific binding sites. *J. Biol. Chem.* **263**, 2693–2698 (1988).
229. Gonda, D. K. *et al.* Universality and structure of the N-end rule. *J. Biol. Chem.* **264**, 16700–16712 (1989).
230. Bartel, B., Wüning, I. & Varshavsky, A. The recognition component of the N-end rule pathway. *EMBO J.* **9**, 3179–3189 (1990).
231. Schwarz, S. E., Rosa, J. L. & Scheffner, M. Characterization of human hect domain family members and their interaction with UbcH5 and UbcH7. *J. Biol. Chem.* **273**, 12148–12154 (1998).
232. Rotin, D. & Kumar, S. Physiological functions of the HECT family of ubiquitin ligases. *Nat. Rev. Mol. Cell Biol.* **10**, 398–409 (2009).
233. Plechanovová, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. & Hay, R. T. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115–120 (2012).
234. Lorick, K. L. *et al.* RING fingers mediate ubiquitin-conjugating enzyme (E2)-dependent ubiquitination. *Proc. Natl. Acad. Sci. USA* **96**, 11364–11369 (1999).
235. Lovering, R. *et al.* Identification and preliminary characterization of a protein motif related to the zinc finger. *Proc. Natl. Acad. Sci. USA* **90**, 2112–2116 (1993).
236. Borden, K. L. & Freemont, P. S. The RING finger domain: a recent example of a sequence-structure family. *Curr. Opin. Struct. Biol.* **6**, 395–401 (1996).
237. Joazeiro, C. A. *et al.* The tyrosine kinase negative regulator c-Cbl as a RING-type, E2-dependent ubiquitin-protein ligase. *Science* **286**, 309–312 (1999).
238. Levkowitz, G. *et al.* c-Cbl/Sli-1 regulates endocytic sorting and ubiquitination of the epidermal growth factor receptor. *Genes Dev.* **12**, 3663–3674 (1998).
239. Wang, S. *et al.* E3 ubiquitin ligases Cbl-b and c-Cbl downregulate PD-L1 in EGFR wild-type non-small cell lung cancer. *FEBS Lett.* **592**, 621–630 (2018).
240. den Besten, W., Kuo, M.-L., Tago, K., Williams, R. T. & Sherr, C. J. Ubiquitination of, and sumoylation by, the Arf tumor suppressor. *Isr Med Assoc J* **8**, 249–251 (2006).
241. Honda, R., Tanaka, H. & Yasuda, H. Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53.

FEBS Lett. **420**, 25–27 (1997).

242. Konopleva, M. *et al.* MDM2 inhibition: an important step forward in cancer therapy. *Leukemia* **34**, 2858–2874 (2020).
243. Rothe, M., Pan, M. G., Henzel, W. J., Ayres, T. M. & Goeddel, D. V. The TNFR2-TRAF signaling complex contains two novel proteins related to baculoviral inhibitor of apoptosis proteins. *Cell* **83**, 1243–1252 (1995).
244. Roy, N. *et al.* The gene for neuronal apoptosis inhibitory protein is partially deleted in individuals with spinal muscular atrophy. *Cell* **80**, 167–178 (1995).
245. Deveraux, Q. L., Takahashi, R., Salvesen, G. S. & Reed, J. C. X-linked IAP is a direct inhibitor of cell-death proteases. *Nature* **388**, 300–304 (1997).
246. Imai, Y. *et al.* An unfolded putative transmembrane polypeptide, which can lead to endoplasmic reticulum stress, is a substrate of Parkin. *Cell* **105**, 891–902 (2001).
247. Kitada, T. *et al.* Mutations in the parkin gene cause autosomal recessive juvenile parkinsonism. *Nature* **392**, 605–608 (1998).
248. Shimura, H. *et al.* Immunohistochemical and subcellular localization of Parkin protein: absence of protein in autosomal recessive juvenile parkinsonism patients. *Ann. Neurol.* **45**, 668–672 (1999).
249. Aravind, L. & Koonin, E. V. The U box is a modified RING finger - a common domain in ubiquitination. *Curr. Biol.* **10**, R132–4 (2000).
250. Hatakeyama, S., Yada, M., Matsumoto, M., Ishida, N. & Nakayama, K. I. U box proteins as a new family of ubiquitin-protein ligases. *J. Biol. Chem.* **276**, 33111–33120 (2001).
251. Wenzel, D. M., Lissounov, A., Brzovic, P. S. & Klevit, R. E. UBCH7 reactivity profile reveals parkin and HHARI to be RING/HECT hybrids. *Nature* **474**, 105–108 (2011).
252. Marín, I. & Ferrús, A. Comparative genomics of the RBR family, including the Parkinson's disease-related gene parkin and the genes of the ariadne subfamily. *Mol. Biol. Evol.* **19**, 2039–2050 (2002).
253. Morett, E. & Bork, P. A novel transactivation domain in parkin. *Trends Biochem. Sci.* **24**, 229–231 (1999).
254. Spratt, D. E., Walden, H. & Shaw, G. S. RBR E3 ubiquitin ligases: new structures, new insights, new questions. *Biochem. J.* **458**, 421–437 (2014).

Chapter 2

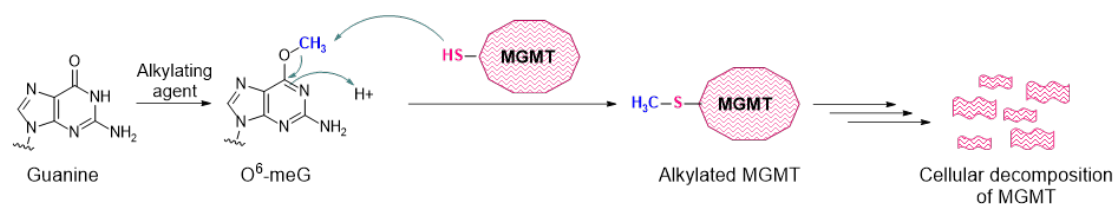
Substrate scope of the repair protein MGMT

4 Introduction

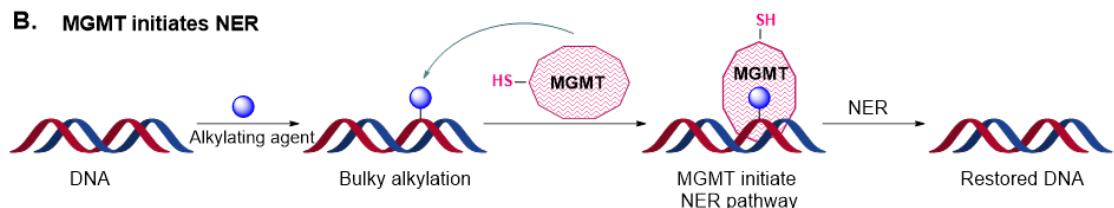
First characterized in *Escherichia coli* (*E. coli*)¹ and later in humans², O⁶-alkylguanine-DNA alkyltransferase (AGT) removes an alkyl group from modified DNA. Specifically, AGT removes the alkyl group from the guanine O⁶ position, by transferring it to its reactive cysteine (C145)^{3,4}. Alkylated AGT is no longer active after the transfer of the alkyl group and undergoes cellular decomposition⁵. Interestingly, AGT has not been found in other well characterized bacteria, as for example *D. radiodurans*⁶ or *T. thermophilus*⁷. A family of alkyltransferase-like (ATL) proteins has been identified in bacteria and yeast^{8,9}. The proteins of the ATL family exhibit close sequence homology to the AGT family, however, the cysteine in the active site is replaced by alanine, isoleucine or tryptophan¹⁰. It has been shown that ATL proteins are involved in DNA repair by binding to the methylation on O⁶-guanine in both single- and double-stranded oligonucleotides, despite the absence of the reactive cysteine and its inability to directly remove the methyl alkylation from the substrate^{9,11-17}. Since *E. coli* contains both AGT and ATL proteins, it was suggested that AGT repairs methylated and small alkylated adducts on DNA, while ATL gets involved when alkylation cannot be removed by AGT and instead initiates the nucleotide excision repair (NER) pathway^{12,18,19}.

Humans lack the ATL protein and, therefore, I hypothesize that human AGT (MGMT) plays both the roles of AGT and ATL proteins. Depending on the size of the alkyl group, MGMT can detect alkylated DNA damage and either directly repair it or initiate the NER pathway (Scheme 1). In this context, I wanted to investigate whether MGMT can directly repair guanine O⁶-alkylations in RNA. The goal of this project is to test the substrate scope of MGMT by screening different types of alkylation damages on either single- or double-stranded oligonucleotides.

A. MGMT in direct repair pathway



B. MGMT initiates NER



Scheme 1. Hypothesis about MGMT in the role of AGT and AGL proteins. If the alkylation is not sterically demanding (e.g. a methyl group), the reactive cysteine of MGMT will transfer the alkyl group onto itself and gets degraded, whereby the guanine will be restored. (A). If the alkylation is too bulky, MGMT will bind to the lesion site triggering the NER pathway, ultimately removing the damaged base pairs and restoring the DNA structure (B).

5 HPLC-based assay for MGMT activity screen

I decided to establish an *in vitro* assay, in which the recombinant MGMT protein interacts with the previously reported pseudosubstrate O⁶-benzyl guanine (O⁶-BG)²⁰ (Figure 15A). The reaction progress was monitored with an Ultra Performance Liquid Chromatography–Mass Spectrometer (UPLC-MS) at 288 nm (Figure 15B and C). As the reaction progresses, MGMT removes the benzyl moiety from O⁶-BG forming guanine as a product of the reaction. As a result, the absorbance intensity of O⁶-BG decreases, and the absorbance intensity of guanine increases. To predict the retention time of the formed guanine moiety a 1:1 ratio of O⁶-BG and guanine was used as a reference (Figure 15D). The areas of the O⁶-BG peak were integrated and compared to a control reaction at time $t = 0$ hours. The reaction was performed in triplicate and the reaction progress was measured at three different time points: 1, 3 and 5 hours (Figure 15E). The data showed such inconsistency between replicates that I needed alternative solutions. I therefore turned to an assay that used fluorescence detection of small dyes, which, in the case of cyanine-5 (Cy5), can be detected in picomolar concentrations (see experimental section).

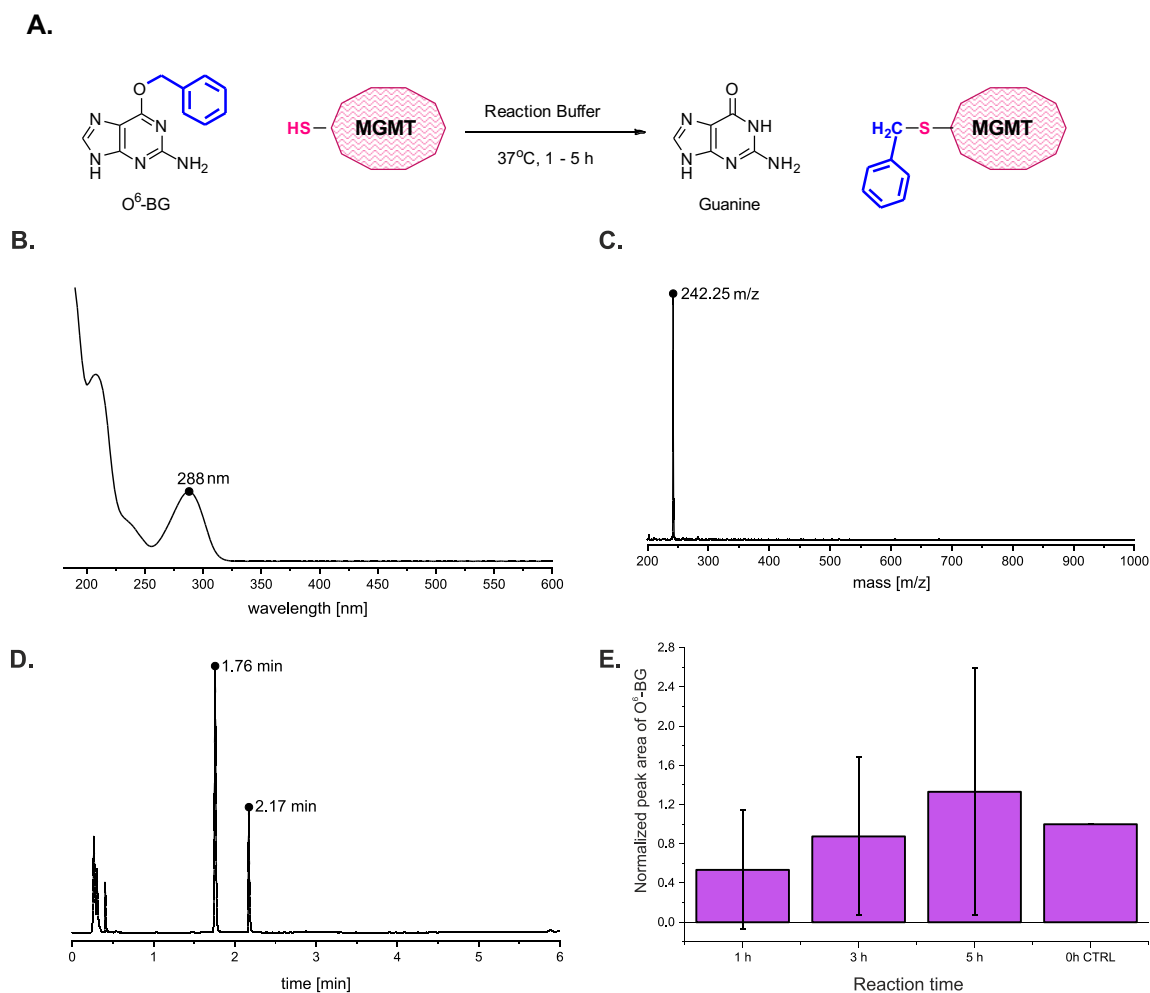
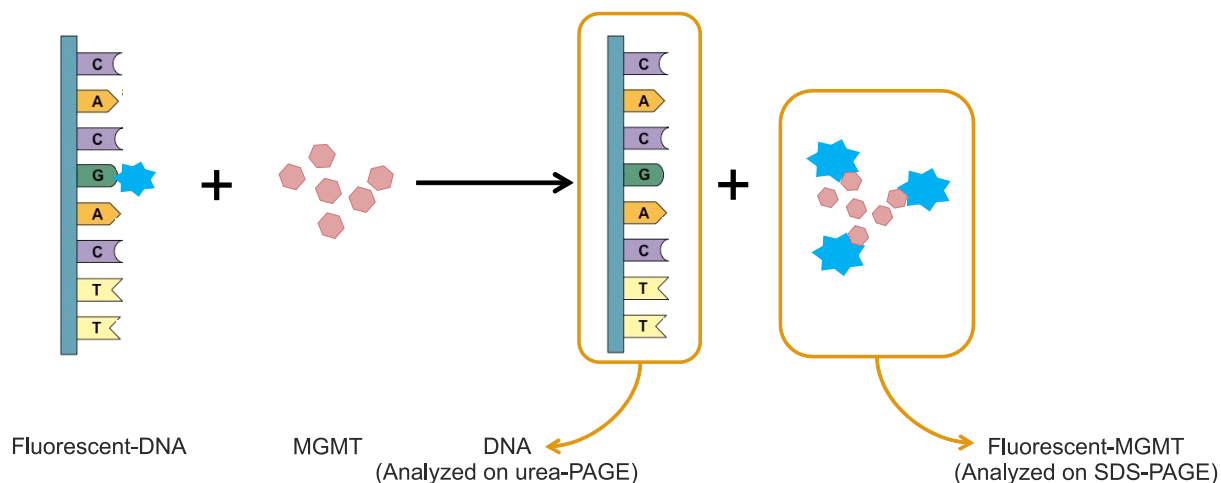


Figure 15. UPLC-MS assay. Reaction scheme illustrating the MGMT activity measurement based on monitoring the absorption of O⁶-BG at 288 nm (A); UV absorption spectrum of O⁶-BG (B); MS spectrum of O⁶-BG. Expected $[M+1]$: 242.25 m/z, found 242.25 m/z (C). UPLC trace of a reference sample with O⁶-BG and guanine at 1:1 ratio. Guanine has a retention time of 1.765 min and O⁶-BG of 2.175 min (D); Integrated peak areas of O⁶-BG from UPLC trace after 1, 3 or 5 hours of incubation with MGMT normalized to the control reaction (CTRL) at $t=0$ hours. SD of three replicates (E).

6 Fluorescence-based assay for MGMT activity screen

In order to overcome the low detection limit of the UPLC-based assay, I developed a fluorescence-based assay to screen the activity of MGMT. I used a double-stranded DNA (dsDNA) containing a guanine that was modified with a fluorescent moiety at the O⁶-position (Scheme 2), using methods established in my group (see experimental Procedures). MGMT would alkylate itself by transferring the fluorophore from the DNA onto itself. The reaction mixture would then be evaluated under the Cy5 channel (695/55 nm) both by urea-PAGE (to detect fluorescent DNA) and SDS-PAGE (to detect the fluorescent protein). With progressing reaction time, I expected to see a decreasing fluorescence signal on the DNA gel and an increasing fluorescence signal on the protein gel.



Scheme 2. Schematic representation of the fluorescence-based activity assay. For simplicity only one of the two DNA strands is depicted.

6.1 Evaluation of assay sensitivity

Data presented in Figure 16 show that the fluorescence intensity on the DNA gel decreases with longer reaction times (Figure 16A), while it is increasing on the protein-gel, as expected (Figure 16C). The DNA gel was stained with SYBR gold to confirm that the position of the fluorescent signals corresponded to DNA (Figure 16B). The same was true for the protein gel, which was stained with Coomassie Blue confirming that the fluorescent signal originates from the 25 kD MGMT protein (Figure 16E). Free Cy5 did not show any unspecific interaction with MGMT, indicating that the change in signal intensities arises solely from MGMT's repair of alkylated DNA. Additionally, the reaction seems to be completed within 30 minutes (Figure 16C). Interestingly, MGMT does not show any specificity towards Cy5-modified-O⁶-deoxyguanosine (**GM-1**) (Figure 16C-E), which suggests that MGMT activity requires the binding to an oligonucleotide. Finally, these data not only show the activity of MGMT, but it also gives information about the time course of direct repair progression, as well as, that the enzyme is able to repair guanine derivatives with large alkyl groups. That finding suggests, that the reason for the lack of ATL proteins in human cells might come from the evolution of MGMT to higher flexibility in terms of substrate scope, rather than playing a role in NER signaling pathway.

To prove the stoichiometric character of the MGMT - substrate interaction, a titration experiment was performed with increasing ratio of fluorescent DNA with respect to MGMT (5:1, 10:1 and 15:1, Figure 16F and G). The resulting data showed an increasing fluorescent signal on the DNA gel but did not show an increasing fluorescence of MGMT on the protein gel, supporting the stoichiometric reaction of MGMT with the substrate.

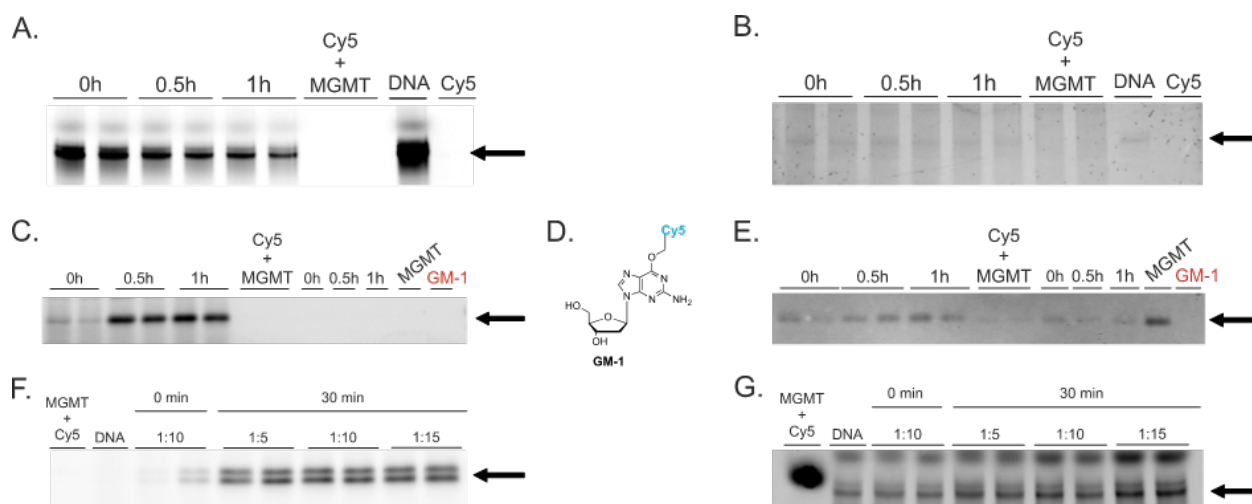


Figure 16. MGMT in vitro activity determined with a PAGE gel fluorescent assay. 15% PAGE, 7 M urea DNA gel analyzed under the Cy5-channel of the fluorescence reader (A), and the same gel stained with SYBR gold for DNA detection (B). 18% SDS-PAGE protein gel, fluorescence detected with the Cy5-channel (C); the same gel stained with Coomassie Blue for protein detection (E). GM-1 structure (D). MGMT titration with increasing amount of fluorescent DNA, 18% SDS-PAGE protein gel, fluorescence detected with the Cy5-channel (F) and 15% PAGE, 7 M urea DNA gel analyzed under the Cy5-channel of the fluorescence reader (G). Arrows pointing at bands corresponding to the expected product.

6.2 Single- and double-stranded oligonucleotides as substrates for MGMT

MGMT activity towards other substrates, as for example single-stranded DNA or tDNA, has not been previously thoroughly investigated. Apart from substrate diversity, I also wanted to assess the activity of MGMT from different expression systems. Therefore, I used MGMT expressed in insect cells (*Spodoptera frugiperda* Sf21 cells), HeLa cell lysate with endogenous MGMT, HeLa cell lysate with transiently expressed MGMT, as well as MGMT expressed in *E. coli*. Each sample was tested with O⁶-fluorescently modified guanine of double-stranded DNA and tDNA. Due to the low tRNA stability at room temperature, a DNA sequence mimicking tyrosine tRNA *E. coli* was used, where uracil was replaced with a thymine nucleotide (see experimental procedures). The different DNA samples contained either one or three modified guanines (G or 3G, respectively) and the fluorescent substrates had been normalized to the same Relative Fluorescence Unit (RFU) values, meaning that the same fluorescence would be expected in case of full conversion of the substrate. The protein gel was blotted and analyzed with a housekeeping protein (tubulin) antibody confirming that the equal amount of lysate was used for both HeLa cell lysates. Similarly, staining with MGMT antibody shows the overall amount of MGMT protein used in the study (Figure 17A middle and last panel, respectively). The obtained data were analyzed based on the fluorescence intensity of the observed bands, which is expected to correlate with the activity of MGMT (see section 6).

The results presented in Figure 17 show that regardless of the MGMT source, double-stranded DNA is the most favorable substrate for MGMT. Interestingly, single-stranded DNA with three alkylated guanines is a better substrate for MGMT compared to the single alkylated one (Figure 17B). Finally, the presented results show that the alkylated guanine in tDNA can be repaired by MGMT.

The resulting data show that the activity of MGMT differs depending on its source. One has to be careful drawing conclusions, since the amount of protein is not the same across the samples. By measuring the fluorescence intensity and comparing it to the amount of MGMT in the sample (see experimental section), one can observe that despite a 5 times higher amount of MGMT is expressed in bacteria compared to one expressed in insect cells, the

fluorescence intensity of the Cy5-modified protein is only twice as high. This would suggest that Sf9-expressed MGMT is more active compared to MGMT expressed in *E. coli* (Figure 17A and B).

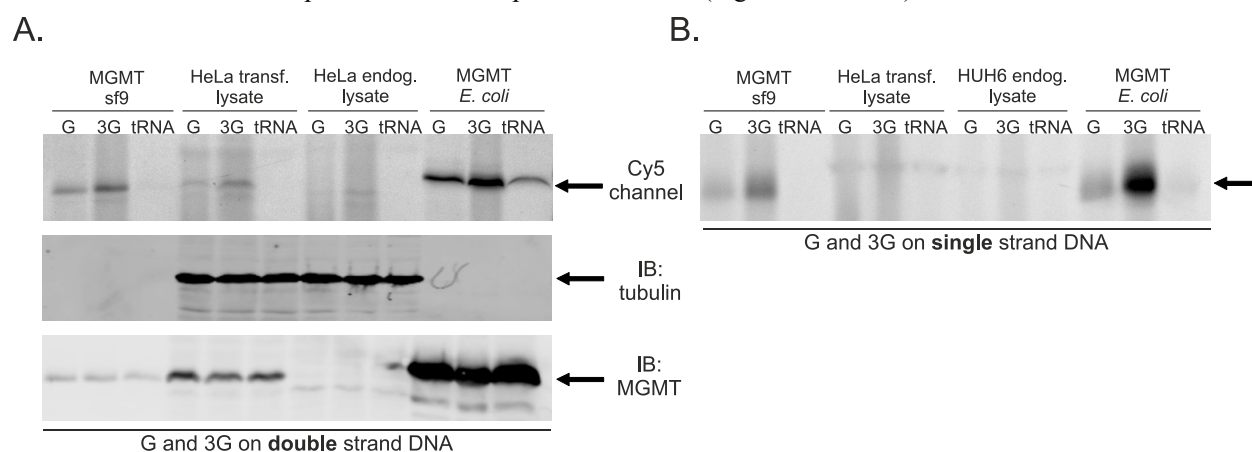


Figure 17. MGMT expressed in insect cells, HeLa cell lysate, HUH cell lysate or MGMT (*E. coli*) with Cy5-modified *O*⁶-guanine on tDNA or double-stranded DNA with one (G) or three modifications (3G). 15% SDS-PAGE gel imaged under a fluorescence channel at 695/55nm (A, top panel); The same gel after immunoblotting stained with anti-tubulin confirming equal loading of the HeLa cell lysate (A, middle panel) the same blot stained with anti-MGMT antibody to indicate the amount of MGMT protein present in each sample (A, bottom panel); Similarly, the same assay was performed with single-stranded DNA. 18% SDS-PAGE gel imaged under a fluorescence reader at 695/55nm (B). G: DNA with one guanine modified at *O*⁶-position with Cy5. 3G: DNA with three guanines modified at *O*⁶-position with Cy5. tDNA: tDNA with six guanines modified at *O*⁶-position with Cy5. Black arrows are pointing at the expected product bands

Apart from studying MGMT's ability to repair oligonucleotide-based substrates, I was interested in using MGMT pseudosubstrates, such as chloromethyl triazole (CMT)²¹. This application is a first step towards using MGMT as a tool for targeted protein degradation (Chapter 3). Targeted protein degradation has recently emerged as novel strategy for tumor therapy^{22–24}. In order to evaluate the possible differences in MGMT activity with its natural substrate (dsDNA) compared to its activity with CMT, a Cy5-modified chloromethyl triazole probe was synthesized (**GM-2**) and used in a fluorescent-based assay (Figure 18A). I used MGMT expressed in *E. coli* (Figure 18B and C) and *hepatocellular carcinoma* (HUH6) cell lysates (Figure 18D and E) as well as DNA with one- or three- fluorescently modified guanines (G, 3G) for this studies. The used fluorescent substrates were normalized to 370 RFU values.

Data obtained with the recombinant MGMT (Figure 18B and C) show that GM-2 is a better substrate for MGMT than double-stranded DNA. When using the whole cell lysate from HUH6 cells, the opposite was observed: DNA showed a greater specificity towards MGMT, while GM-2 bound with off-target proteins (Figure 18D and E). This result is expected, since the highly reactive chloromethyl group can form covalent bonds with the nucleophilic amino acids of other proteins.

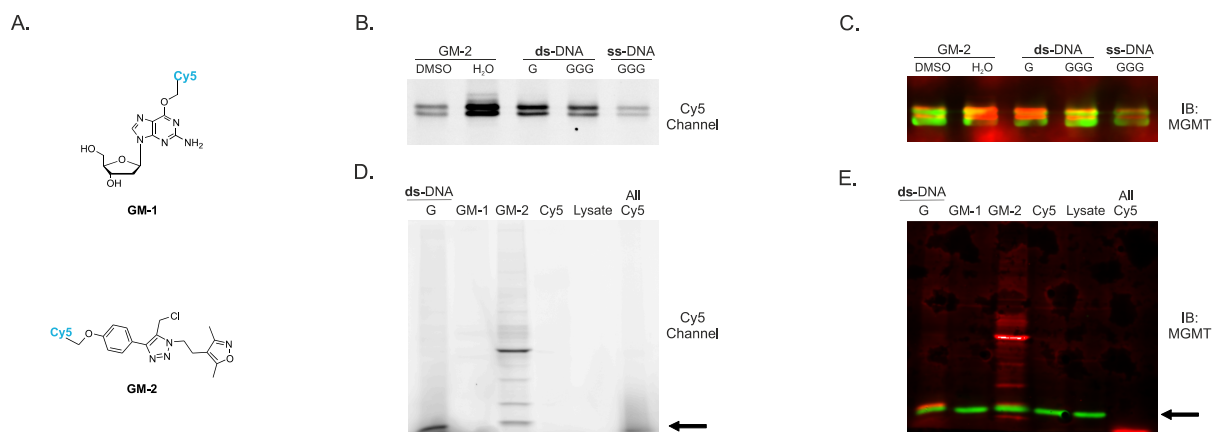


Figure 18. Substrate screen with purified MGMT expressed in *E. coli* (B and C) or HUH6 whole cell lysate (D and E). Chemical structures of GM-1 and GM-2 (A). Western blot imaged under the Cy5 channel using his-MGMT (B). The same blot stained with anti-tubulin antibody (green) and anti-MGMT antibody (red) overlay (C). Western blot analysis using a whole cell lysate of HUH6 cells imaged under the Cy5 channel (D). The same blot stained with anti-tubulin antibody (green) and anti-MGMT antibody (red) overlay (E).

7 Conclusions

The fluorescent electrophoresis assay is a very sensitive technique for studying MGMT's reactivity with dsDNA, ssDNA, tDNA or small molecules. I was able to confirm the stoichiometric behavior of MGMT in terms of alkylation repair as well as show its interaction with different substrates. Presented data show for the first time that MGMT is capable to directly repair O⁶-alkylguanine lesions on single-stranded DNA. Findings from this study indicate that MGMT could play a role in more than just in direct DNA repair but could also be involved in RNA repair and maintenance.

The NER pathway is facilitated by a multi-protein machinery, which requires a lot of energy from the cell to be assembled. MGMTs ability of removing large substituents from guanine suggest its improved function as a result of cellular adaptation to alkylation damage. By improving the substrate scope, the NER pathways do no longer need to be initiated.

Finally, a small molecule can be used as a substrate for MGMT and its application could be expanded to use MGMT as a tool for targeted protein degradation (investigated in Chapter 3). This preliminary data, however, suggest that careful optimization needs to be done before choosing a substrate for efficiently targeting MGMT.

8 Bibliography

1. Demple, B., Jacobsson, A., Olsson, M., Robins, P. & Lindahl, T. Repair of alkylated DNA in *Escherichia coli*. Physical properties of O⁶-methylguanine-DNA methyltransferase. *J. Biol. Chem.* **257**, 13776–13780 (1982).
2. Tubbs, J. L., Pegg, A. E. & Tainer, J. A. DNA binding, nucleotide flipping, and the helix-turn-helix motif in base repair by O⁶-alkylguanine-DNA alkyltransferase and its implications for cancer chemotherapy. *DNA Repair (Amst.)* **6**, 1100–1115 (2007).
3. Lindahl, T., Sedgwick, B., Sekiguchi, M. & Nakabeppu, Y. Regulation and expression of the adaptive response to alkylating agents. *Annu. Rev. Biochem.* **57**, 133–157 (1988).
4. Pegg, A. E. Repair of O⁶-alkylguanine by alkyltransferases. *Mutation Research/Reviews in Mutation Research* **462**, 83–100 (2000).
5. Pegg, A. E. Repair of O⁶-alkylguanine by alkyltransferases. *Mutation Research/Reviews in Mutation Research* **462**, 83–100 (2000).
6. Fang, Q., Kanugula, S. & Pegg, A. E. Function of domains of human O⁶-alkylguanine-DNA alkyltransferase.

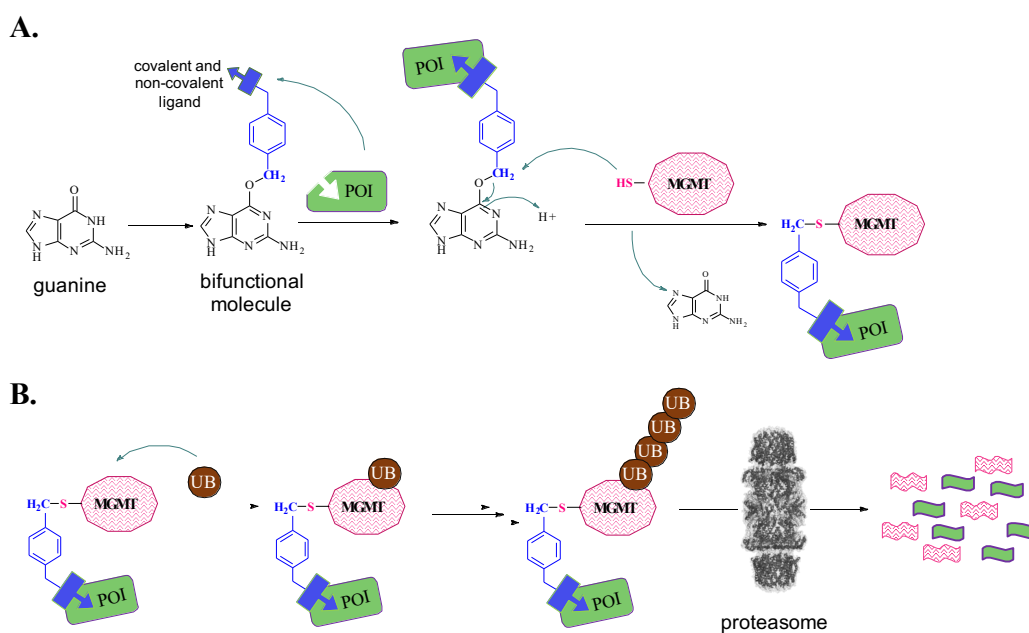
- Biochemistry* **44**, 15396–15405 (2005).
7. Morita, R., Nakagawa, N., Kuramitsu, S. & Masui, R. An O6-methylguanine-DNA methyltransferase-like protein from *Thermus thermophilus* interacts with a nucleotide excision repair protein. *J. Biochem.* **144**, 267–277 (2008).
 8. Pearson, S. J., Ferguson, J., Santibanez-Koref, M. & Margison, G. P. Inhibition of O6-methylguanine-DNA methyltransferase by an alkyltransferase-like protein from *Escherichia coli*. *Nucleic Acids Res.* **33**, 3837–3844 (2005).
 9. Margison, G. P. *et al.* Alkyltransferase-like proteins. *DNA Repair (Amst.)* **6**, 1222–1228 (2007).
 10. Pearson, S. J., Ferguson, J., Santibanez-Koref, M. & Margison, G. P. Inhibition of O6-methylguanine-DNA methyltransferase by an alkyltransferase-like protein from *Escherichia coli*. *Nucleic Acids Res.* **33**, 3837–3844 (2005).
 11. Morita, R., Nakagawa, N., Kuramitsu, S. & Masui, R. An O6-methylguanine-DNA methyltransferase-like protein from *Thermus thermophilus* interacts with a nucleotide excision repair protein. *J. Biochem.* **144**, 267–277 (2008).
 12. Tubbs, J. L. *et al.* Flipping of alkylated DNA damage bridges base and nucleotide excision repair. *Nature* **459**, 808–813 (2009).
 13. Margison, G. P., Povey, A. C., Kaina, B. & Santibáñez Koref, M. F. Variability and regulation of O6-alkylguanine-DNA alkyltransferase. *Carcinogenesis* **24**, 625–635 (2003).
 14. Pearson, S. J., Ferguson, J., Santibanez-Koref, M. & Margison, G. P. Inhibition of O6-methylguanine-DNA methyltransferase by an alkyltransferase-like protein from *Escherichia coli*. *Nucleic Acids Res.* **33**, 3837–3844 (2005).
 15. Pearson, S. J. *et al.* A novel DNA damage recognition protein in *Schizosaccharomyces pombe*. *Nucleic Acids Res.* **34**, 2347–2354 (2006).
 16. Morita, R., Nakagawa, N., Kuramitsu, S. & Masui, R. An O6-methylguanine-DNA methyltransferase-like protein from *Thermus thermophilus* interacts with a nucleotide excision repair protein. *J. Biochem.* **144**, 267–277 (2008).
 17. Chen, C.-S. *et al.* A proteome chip approach reveals new DNA damage recognition activities in *Escherichia coli*. *Nat. Methods* **5**, 69–74 (2008).
 18. Mazon, G., Philippin, G., Cadet, J., Gasparutto, D. & Fuchs, R. P. The alkyltransferase-like *ybaZ* gene product enhances nucleotide excision repair of O(6)-alkylguanine adducts in *E. coli*. *DNA Repair (Amst.)* **8**, 697–703 (2009).
 19. Tubbs, J. L. *et al.* Flipping of alkylated DNA damage bridges base and nucleotide excision repair. *Nature* **459**, 808–813 (2009).
 20. Sedgwick, B. & Lindahl, T. Recent progress on the Ada response for inducible repair of DNA alkylation damage. *Oncogene* **21**, 8886–8894 (2002).
 21. Falnes, P. Ø., Johansen, R. F. & Seeberg, E. AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. *Nature* **419**, 178–182 (2002).
 22. Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T. & Sedgwick, B. Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. *Nature* **419**, 174–178 (2002).
 23. Aas, P. A. *et al.* Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. *Nature* **421**, 859–863 (2003).
 24. Pegg, A. E. Mammalian O6-alkylguanine-DNA alkyltransferase: regulation and importance in response to alkylating carcinogenic and therapeutic agents. *Cancer Res.* **50**, 6119–6129 (1990).
 25. Wang, C., Abegg, D., Hoch, D. G. & Adibekian, A. Chemoproteomics-Enabled Discovery of a Potent and Selective Inhibitor of the DNA Repair Protein MGMT. *Angew. Chem. Int. Ed. Engl.* **55**, 2911–2915 (2016).

Chapter 3

MGMT probes with bifunctional molecules

9 Introduction

Having established that small-molecule can serve as covalent substrate for MGMT and knowing that the protein is relatively flexible in terms of substrate scope (Chapter 2), the ultimate goal and aim of this project is to use MGMT as a tool for targeted protein degradation. The design of a double degrader, where O⁶-benzylguanine would be modified at the para-position on the benzene ring (blue) with a short linker at the end of which a substrate for protein of interest (POI: green), is outlined in Scheme 3A. This bifunctional molecule incorporates an activity against two targets: POI and MGMT (pink) and allows thereby the selection of a specific target as well as the initiation of its degradation. The reactive cysteine residue of MGMT will form a covalent bond with the α -carbon of O⁶-guanine alkylation, targeting the POI for subsequent ubiquitination and proteasomal degradation (Scheme 3B).



Scheme 3. MGMT as a tool for targeted protein degradation. MGMT inhibition by treatment with bifunctional molecule targeting POI leading to two protein interaction (A). Polyubiquitination and subsequent proteasomal degradation of MGMT and POI (B).

I planned to use the novel concept of MGMT as a degrader using several approaches. First, I chose three targets with known non-covalent binders. Those proteins have been studied in cancer-progression pathways for which small molecule inhibitors had been reported: **dasatinib** for Bcr-Abl proteins^{1,2}, **palbociclib** for CDK4/6³ and (+)-**JQ1** for BRD4 protein^{4,5}. The reported compounds were modified with a reactive handle for chemical modification in order to make them suitable for the formation of bifunctional compounds. Despite the elegance of this novel approach there are a few expected challenges:

Chapter 3

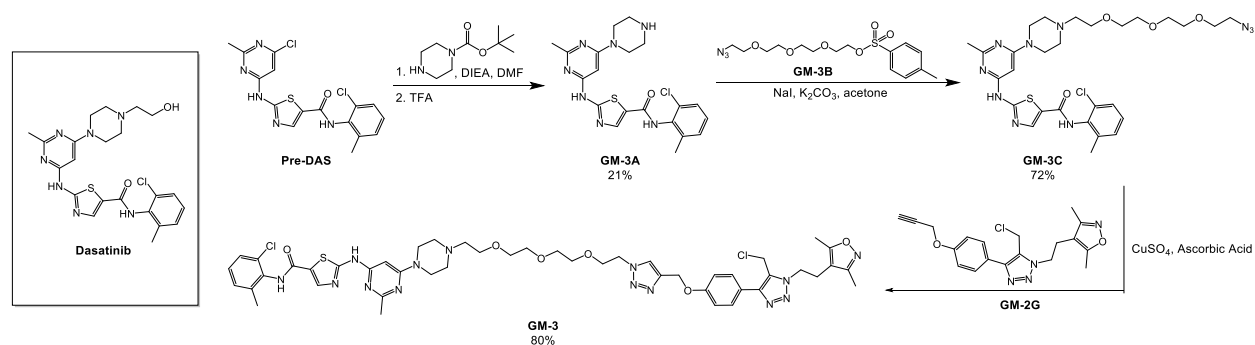
- piggy-backing degradation of a POI on the native MGMT ubiquitination pathway would rely on engagement of both proteins to the shared substrate. If MGMT bound to the shared substrate significantly faster, it would undergo degradation without binding the POI.
- the proteasome contains a ubiquitin signaling check point, which could potentially detach non-covalently bound POI from the complex.

Therefore, in addition to non-covalent engagement of the POI, I envisioned another approach where the POIs would be targeted with covalent ligands. In this way I could precisely match kinetics and lower the risk that deubiquitination would rescue the tagged protein.

10 Non-covalent MGMT-POI interaction

10.1 An MGMT bifunctional targeting Bcr-Abl

Bcr-Abl is a tyrosine kinase – a protein fusion, which results from a reciprocal translocation between long arms of chromosomes 9 (Abl) and 22 (Bcr)^{6,7,8}. That specific translocation is called Philadelphia Chromosome and is associated with chronic myelogenous leukemia^{6,9}. Inhibiting Bcr-Abl kinase activity had been used as a treatment of Bcr-Abl positive leukemias^{10,11}. One of such inhibitors – dasatinib – is particularly interesting, since it is able to cross blood-brain barrier, and it is widely used for Bcr-Abl driven cancers ($IC_{50} = 3.3$ nM in K562 cells)². Therefore, I decided to use dasatinib with a modification introduced on the hydroxy group of the ethyl chain, which had been previously reported not to affect dasatinib activity^{12–14}. As shown in synthetic Scheme 2, starting with precursor Pre-DAS, containing a reactive chlorine moiety in the pyrimidine ring, a piperidine building block was introduced (**GM-3A**) that was subsequently connected to a PEG3-azide moiety (**GM-3C**). The introduced PEG3-azide linker can then be attached to an MGMT substrate (**GM-2G**) using copper catalyzed click reaction (Scheme 4). That particular alkyne-bearing chloromethyl triazole molecule had previously been used in the MCF7 breast cancer cell line and the Caco-2 colon cancer cell line, as a pull-down probe¹⁵.



Scheme 4. Synthetic scheme of GM-3. Dasatinib structure shown in a square box - $IC_{50} = 3.3$ nM in K562 cells².

Cell treatment experiments were performed as follows: K562 cells were incubated with vehicle (DMSO), negative control (GM-2G) or bifunctional molecule (GM-3) at different concentrations for 24 hours. Expected results would show disappearance of Bcr-Abl band after treatment with GM-3, but not with DMSO or GM-2G. Data presented in Figure 19 indicate poor protein transfer as well as poor antibody staining, making quantification of bands impossible. Additionally, bands corresponding to Bcr-Abl in lanes treated with GM-3 seems to be

stronger compared to negative controls. MGMT evaluation was not possible due to big differences in protein size – Bcr-Abl (140-250 kDa) and MGMT (25 kDa).

Lack of GM-3 activity could arise from low levels of active endogenous MGMT in K562 cell line. In order to evaluate that hypothesis five different cell lines had been compared by lysing similar number of cells with RIPA buffer and using same protein amount (25 µg) for Western-Blot analysis (Figure 19B). Apart from K562, other commonly used cell lines were used, such as HeLa, HEK293T, HT-29 and hepatocellular line HUH6. Data shown in Figure 19B and quantified in Figure 19C, clearly indicate that there is a very low level of endogenous MGMT in K562 cells. That information helped in moving forward choosing the right cell line for the next experiments.

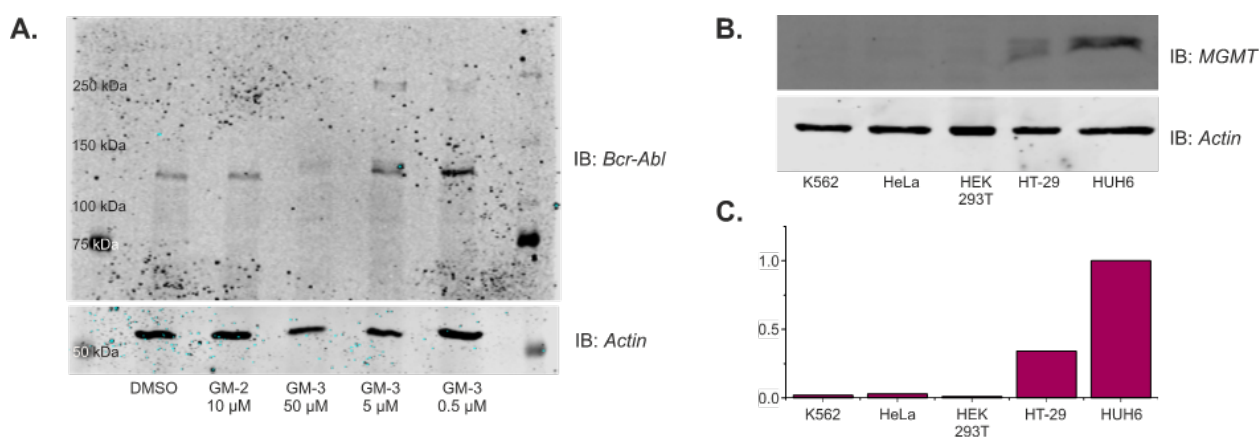


Figure 19. Western blot from 24h treatment of K562 cells with vehicle (DMSO), negative control (GM-2G) or bifunctional molecule (GM-3) (A). Western blot analysis of MGMT endogenous levels from lysates of different cell lines: K562, HeLa, HEK293T, HT-29

The level of MGMT activity varies between species and cell types^{16,17}. Human liver cancer cells (*Hepatocellular carcinoma*) are known to have high levels of MGMT activity¹⁶. Ten different hepatocarcinoma cell lysates had been used for MGMT endogenous protein level screen (Figure 20) from which HUH6 had been shown to have the highest. Therefore, HUH6 was selected for further studies as a model system.

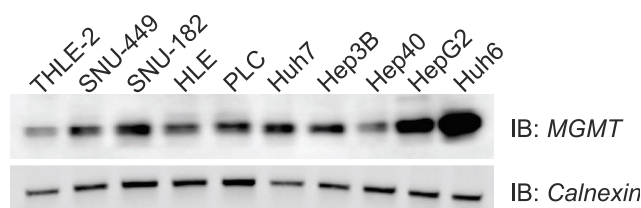


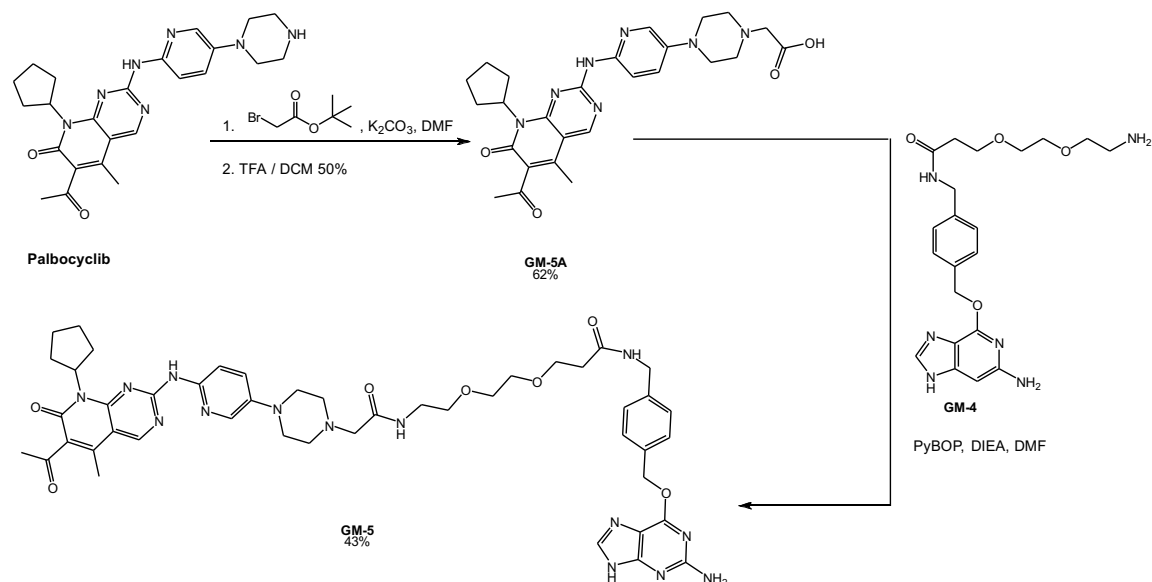
Figure 20. Western blot analysis of ten different Hepatocellular carcinoma lines, stained against MGMT and housekeeping protein calnexin.

10.2 An MGMT bifunctional targeting CDK4 and CDK6

Cell division protein kinase 4 and 6 (CDK4 and CDK6, CDK4/6) are playing an important role in cell proliferation serving as an activator checkpoint during cell cycle G_1 to S phase^{18,19}. The loss of control over that checkpoint results in accumulation of CDK4 and CDK6 proteins. That affects growth-factor depending signal transduction, yielding in uncontrollable cell proliferation resulting in cancer growth²⁰⁻²². One way to overcome accumulated CDK4/6 activity is using highly selective small-molecule inhibitors, that would target only CDK4/6 without suppressing other CDK activities. Such drug molecules have been developed by Pfizer, Eli Lilly &

Chapter 3

Company and Novartis named palbociclib, abemaciclib, ribociclib, respectively. For investigating the MGMT degradation approach, I choose palbociclib ($IC_{50} \cong 20 \mu\text{M}$)³. The bifunctional degrader was synthesized by an acetate modification on secondary amine in piperazine moiety with the use of *tert*-butyl protected bromoacetate and subsequent attachment of O⁶-BG-PEG2-NH₂ via a peptide coupling step resulting in **GM-5** (Scheme 5).



Scheme 5. Synthetic route to the bifunctional degrader GM-5

To show the concentration dependence of MGMT-induced degradation initiated by **GM-5**, HUH6 cells had been treated with one-fold of magnitude increments in concentration of **GM-5** (100 μM , 10 μM , 1 μM , 0.1 μM). Palbociclib and DMSO were used as negative controls for CDK4/6 as well as O⁶-BG as positive control for MGMT. Expected results would show MGMT and CDK4/6 dose-dependent degradation. From the results, presented in Figure 21, it is clear that MGMT levels decrease with increasing concentration of **GM-5** as expected, but CDK4/6 remains unaffected. The results indicate that MGMT binds to bifunctional molecule and remains active. However, lack of CDK4/6 degradation might arise from two factors. Firstly, MGMT binds to the shared substrate significantly more rapidly than CDK4/6 and undergoes the subsequent degradation without being attached to CDK4/6. Secondly, the non-covalent attachment of two proteins is not sufficient for targeted protein degradation and the proteasome degrades the complex of the bifunctional molecule and MGMT without the non-covalently bound CDK4/6.

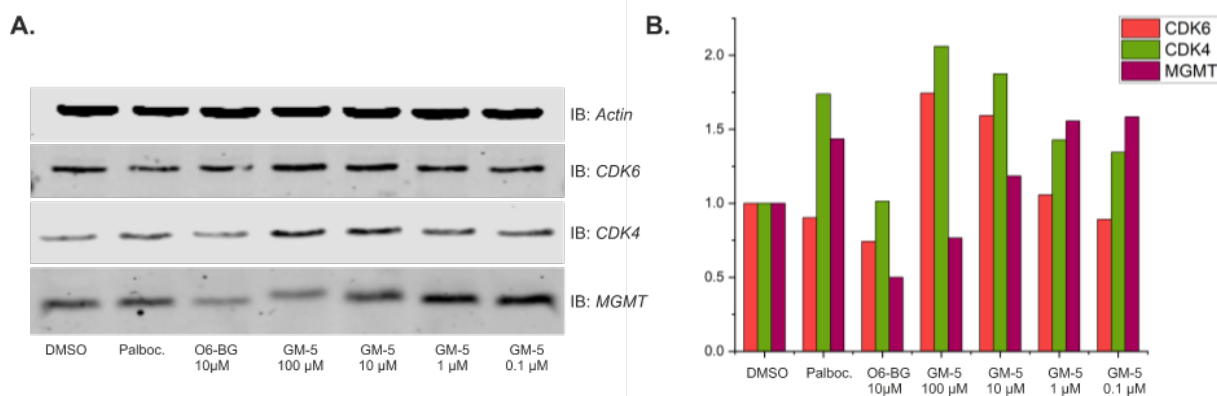
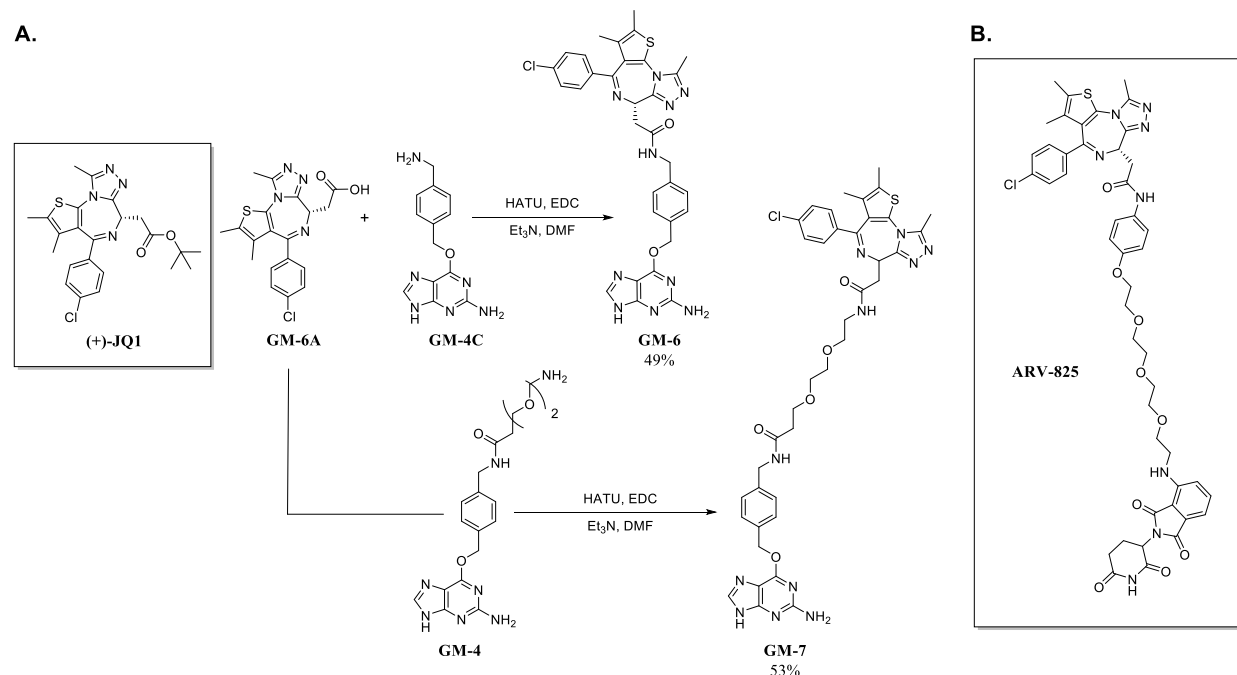


Figure 21. Western blot analysis of 24 hours treatments in HUH6 cells with different concentrations of GM-5 and O⁶-BG and Palbociclib as a control (A). Data quantification relative to protein content in DMSO treated sample (B).

10.3 An MGMT bifunctionals targeting BRD4

Bromodomain proteins are associated with acetylated chromatin acting as transcription factors which makes them a very important biological target²³. BRD4 has been a target for cancer therapy since it was discovered that the protein control gene expression involved in cells mitotic progression, mediating positive transcription elongation complex P-TEFb – a validated target in chronic lymphocytic leukaemia^{23–26}. BRD4 inhibitor (+)-JQ1 had been developed in order to disrupt protein binding with acetylated chromatin and therefore inhibit cell growth, cause cell-cycle arrest in G1 phase without inducing apoptosis^{4,5}.

I decided to prepare two compounds: one, where (+)-JQ1 is directly attached to O⁶-BG (**GM-6**) or via PEG2 linker (**GM-7**). The synthetic route presented on Scheme 6A shows the first step of deprotection of carboxylic acid on (+)-JQ1 molecule followed by peptide coupling with primary amine from **GM-4C** or **GM-4** forming **GM-6** and **GM-7**, respectively.



Scheme 6. Synthetic routes towards **GM-6** and **GM-7** (A). Structure of **ARV-825** with binding constant to BRD4 $K_d=50-90$ nM⁵ (B).

To show the concentration dependence of MGMT-induced degradation initiated by **GM-7**, HUH6 cells were treated with various concentrations of **GM-7** (10 μ M, 1 μ M and 0.1 μ M). DMSO was used as a negative control for both proteins. The bifunctional degrader **ARV-825** (Scheme 6B) had been shown to target BRD4 to degradation²⁷, and was therefore used as positive control for BRD4 loss. O⁶-BG was used as a positive control for MGMT degradation. Figure 22A shows corresponding western blot and quantified data that indicate protein-to-drug response in dose-dependent manner for MGMT. The used anti-BRD4 antibody shows three bands in total, out of which two bands correspond to two BRD4 isomers, the long and short one. From those, (+)-JQ1 and **ARV-825** targets the short isomer. The middle band seems to be an unspecific target of the antibody. Following short isomer-BRD4 band intensity, it is clearly visible that the protein response is not consistent across two data sets (Figure 22A vs D). Similarly to previous protein targets, BRD4 is a large protein (152 kDa), therefore it is critical to resolve both MGMT and BRD4 on SDS-PAGE, since high molecular weight proteins at the periphery of the gel will transfer with lesser efficiency compared to small molecular weight proteins (MGMT). This factor could

Chapter 3

lead to ambiguous and imprecise conclusions in data validation and quantification and was therefore circumvented as described in the following.

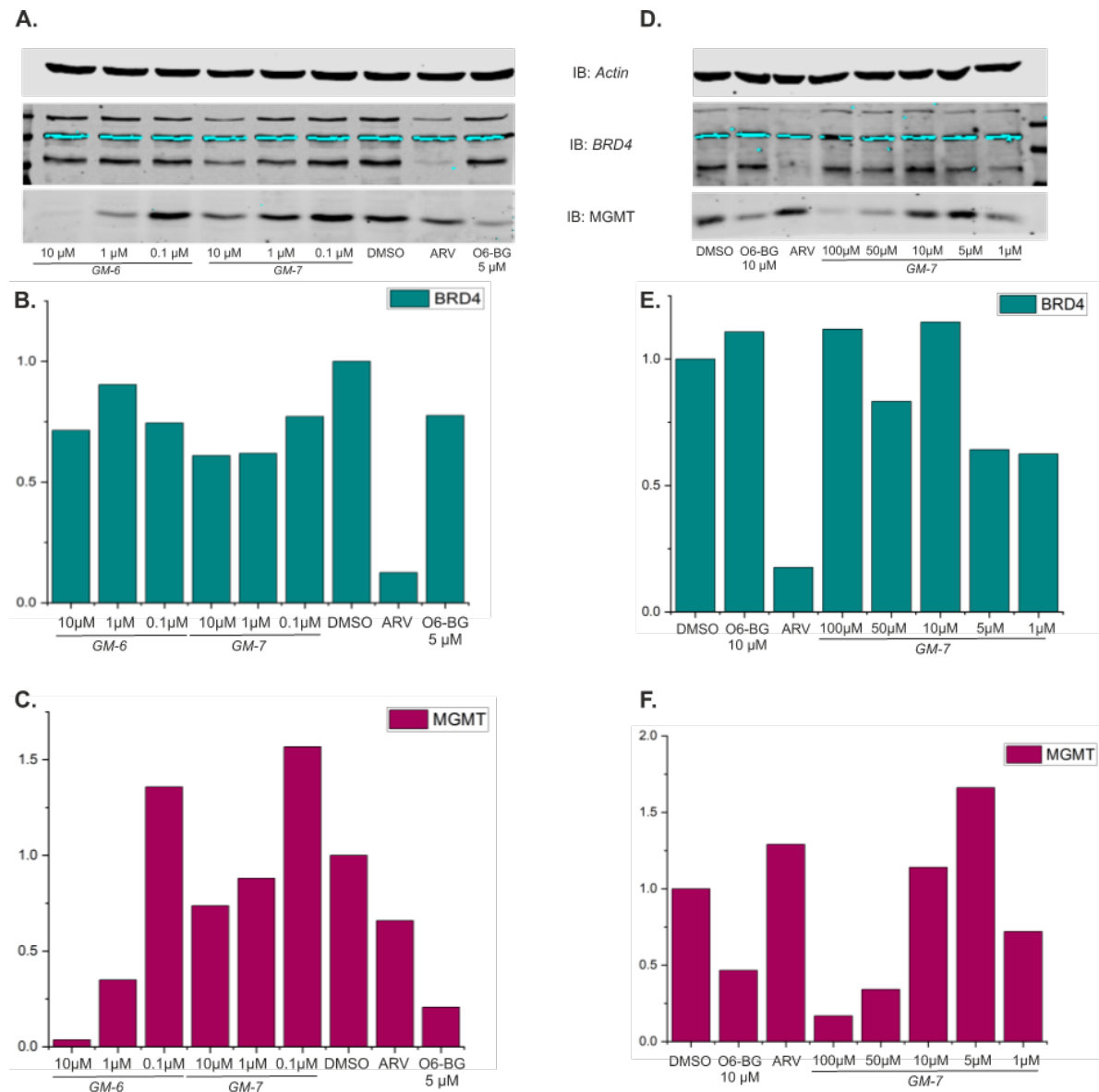


Figure 22. Concentration dependent protein response study. Western blot analysis of HUH6 cells treatment for 24h with different concentration of GM-6 (A) and GM-7 (A and D); DMSO, O⁶-BG, ARV-825 (ARV) used as controls. Data quantification: relative protein concentration to DMSO treated sample (B, C, E, F).

To address the concern regarding poor transfer efficiency, I decided to lower SDS-PAGE gel concentration, allowing BRD4 bands to migrate further. Doing so, I will not be able to visualize MGMT, however with use of appropriate control data should be more apparent to assess. Three time points had been chosen (6 hours, 12 hours and 24 hours) for HUH6 cell treatments with ARV-825 (0.5 μ M), O⁶-BG (5 μ M) and GM-6 (10 μ M). Data presented in Figure 24 show that in all cases BRD4 is being depleted after treatment with GM-6 compared to both the DMSO and O⁶-BG treated samples.

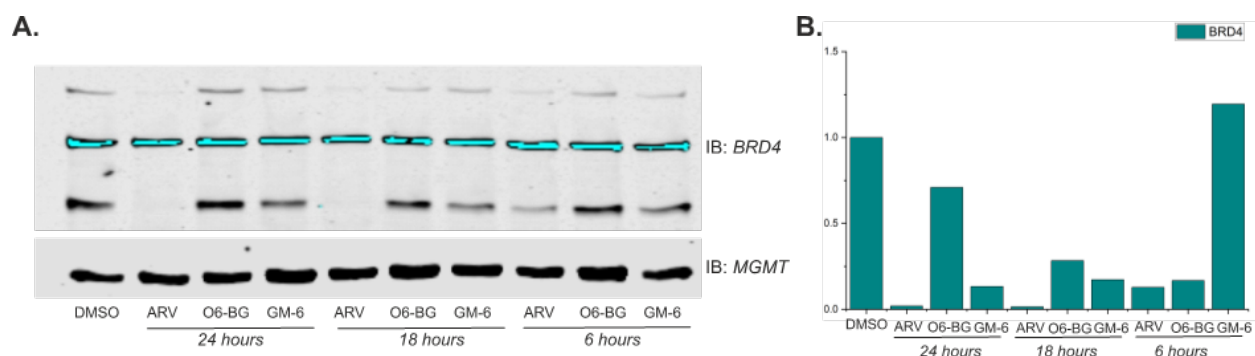


Figure 23. Analysis of time-dependent degradation of BRD4 treated with DMSO, ARV-825, O⁶-BG or GM-6, shown by western blot (A). Data quantification relative to protein content in DMSO treated sample (B).

To further evaluate MGMT-depending mechanism, I decided to use siRNA to knock-down MGMT expression in HUH6 cells. By removing endogenous MGMT, I expected BRD4 levels to remain unchanged upon treatment with GM-6. Similarly, DMSO, ARV-825, O⁶-BG were used as a control. Additionally, I used **GM-8** as a negative control for BRD4 (Figure 24C). A representative western blot is shown in Figure 24A. In the absence of MGMT, BRD4 levels are not rescued as compared to degradation in the presence of MGMT in the cell. Interestingly, **GM-8** lowers BRD4 levels to an equal extent as **GM-6** indicating that both BRD4 and MGMT interact with **GM-6**, but independently from each other. Quantified data from at least three biological replicate show lack of significant statistical difference in BRD4 levels upon treatment with bifunctional molecule **GM-6** (Figure 24B).

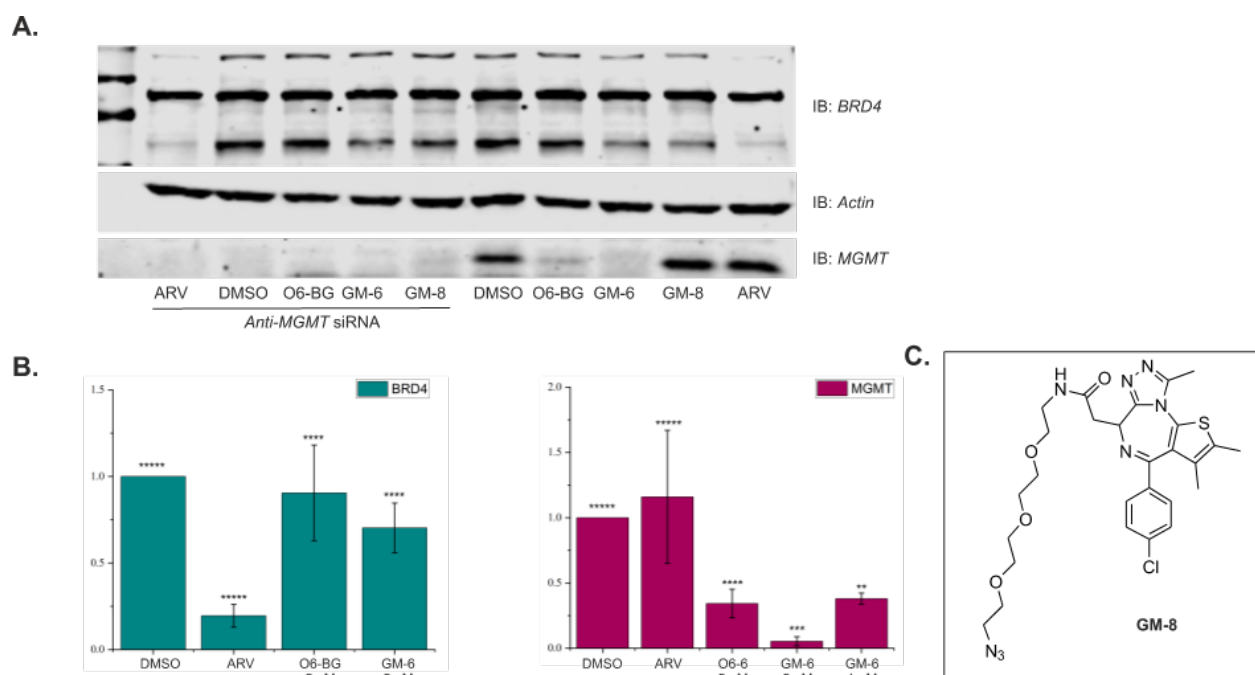


Figure 24. HUH6 cell treatment evaluation by western blot (A). Data quantification relative to protein content in DMSO treated sample (B). Structure of GM-8 (C). * represents number of biological replicates used for quantification.

Chapter 3

Presented data show ambiguous results. One has to be careful drawing conclusions. BRD4 is a large protein which compared to MGMT carries difficulties with visualization method, e.g. in western blot, where one has to rely on homogenous protein transfer that is affected by the protein size. Data obtained by researcher in my laboratory – Dr Coomar, indicate that 10 μM – range concentrations of (+)-JQ1 are sufficient to induce BRD4 degradation *via* hydrophobic tagging²⁸. The effect caused exposing a hydrophobic region of protein causing its proteasomal degradation, had been described before as a degradation signal for unfolded proteins^{29,30}.

11 Covalent MGMT - POI interaction

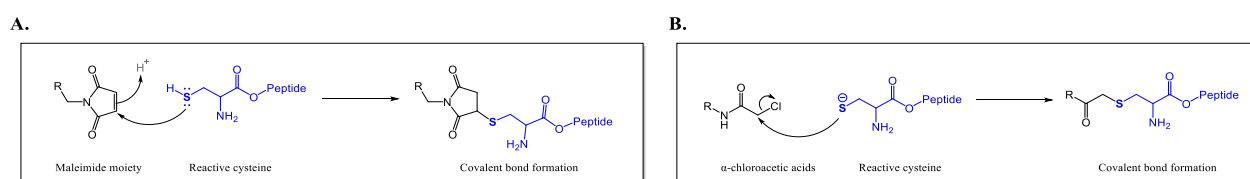
In order to prove that MGMT is in fact able to co-degrade a POI and to circumvent the issues with non-covalent POI-double degrader interactions, I decided to investigate a doubly covalent approach. It is known that MGMT binds covalently to its substrate. However, it is possible that despite bringing two proteins in close proximity via a bifunctional molecule, the degradation of non-covalently attached POI is not occurring because of lack of ubiquitin transfer.

I followed two strategies:

- 1) Small molecule with unspecific target,
- 2) Fusion protein with MGMT.

11.1 Small molecule with unspecific target

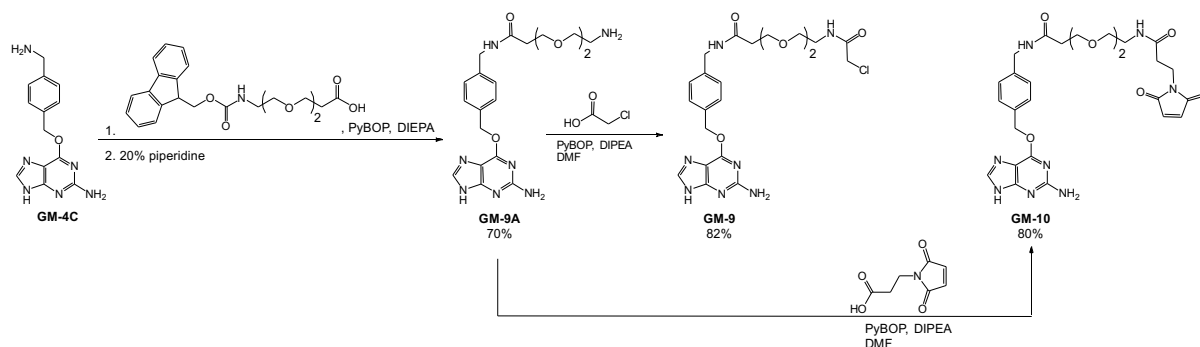
The first strategy uses the nucleophilic character of cysteine – one of the most reactive proteogenic amino acids. Reactive cysteines are responsible for the activity of different classes of proteins, like proteases, oxidoreductases, kinases and acyltransferases³¹. To target reactive cysteines, I decided to rely on two established electrophilic moieties in bioconjugation: maleimide and α -chloroacetamide^{32–35}. Maleimides are Michael acceptors and react with cysteine residues by forming a thiosuccinimide bond³⁶ (Scheme 7A), while α -haloacetic acids react with cysteines via nucleophilic substitution (Scheme 7B). Both had been extensively used as a cysteine-tagging reagents^{32–35}.



Scheme 7. Mechanism of nucleophilic attack of cysteine thiolate residue on maleimide (A) or α -chloroacetamide (B).

Both electrophiles were attached to an O⁶-BG derivative bearing a PEG2 linker terminated with a primary amine (**GM-9A**). An amide coupling step between the primary amine of a PEG2 linker and the carboxylic acid of appropriately functionalized electrophiles (Scheme 8) delivered the bifunctionals – **GM-9** and **GM-10**, respectively.

MGMT probes with bifunctional molecules



Scheme 8. Synthetic route towards **GM-9** and **GM-10**.

Dose-dependent study was performed on HUH6 cells in order to evaluate MGMT response towards electrophilic compounds. Treatments were visualized using western blot (Figure 25A) and quantified (Figure 25C). Data show that MGMT undergoes degradation as expected. The last lane seems to show **GM-10** effect on MGMT levels of 1 μ M treatments to be as high as 50 μ M, however, I do expect that data point to be caused by poor protein transfer, that often shows problems at the peripheries of the SDS-PAGE gel. Similarly, tubulin band intensity in the lane with cell treated with 100 μ M of **GM-9** had been affected by the air bubble present between the SDS-PAGE gel and nitrocellulose membrane during protein transfer. In order to correctly quantify MGMT levels, MGMT band intensity had been compared to the average intensity of all tubulin bands (Figure 25B).

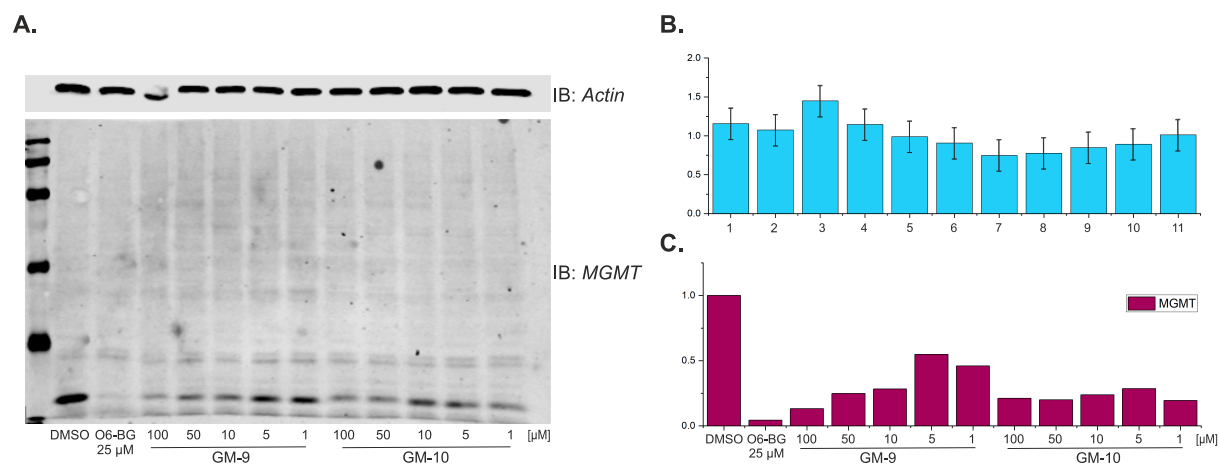


Figure 25. Doubly covalent bifunctional molecules. Western blot of HUH6 cell treatments with **GM-9** and **GM-10**, stained against MGMT and β -actin (A). Tubulin quantification: bands intensity showing standard deviation from average across all bands. $STDV \pm 20\%$ (B). Quantified western blot data showing MGMT degradation in dose dependent manner upon treatments with **GM-9** and **GM-10** relative to DMSO (C).

If MGMT were covalently bound to another protein, that might be visible on western blot in a series of bands of different molecular weight, upon staining with anti-MGMT antibody. However, no such observation had been made (Figure 8A). One explanation is that western-blot immune staining has too low sensitivity. Another being based on MGMT kinetics: MGMT-double degrader complex is removed in the cell faster than binding of the double degrader to the POI.

I chose proteomics as the most sensitive assay to evaluate double-covalent character of both bifunctional compounds. The advantage of this technique is that not only I will be able to determine whether MGMT can be covalently attached to another protein via the electrophilic handle, but also identify interacting unknown proteins.

Chapter 3

The assay includes HUH6 cell lysate treatments with **GM-9** and **GM-10** for 12 hours followed by immunoprecipitation with an anti-MGMT antibody by protein A coated magnetic beads.

To verify sample quality before submission to Proteomics Core Facility as well as evaluate immunoprecipitation efficiency, I used western blot. The same volume of all samples was used and stained with anti-MGMT antibody (Figure 26). Unfortunately, MGMT antibody would also stain the antibody used for immunoprecipitation, showing 50 kDa and 25 kDa of dissociated IgG antibody heavy and light chain (Figure 26, marked with arrows). Despite that, there are visible bands in elution samples of both **GM-9** and **GM-10** (Figure 26, marked with stars), which indicates MGMT covalent attachment with another protein. The prepared samples were then submitted in triplicate to Proteomics Core Facility at the University of Basel, Biozentrum.

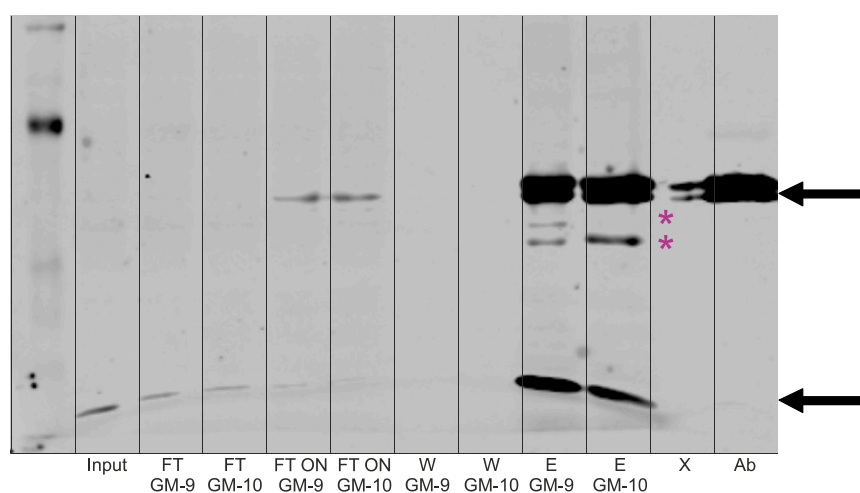


Figure 26. Western blot analysis of HUH6 immunoprecipitation efficiency evaluation. Input – sample before adding beads. FT – Flow Through, sample not-bound to beads after 5 minutes incubation. FT ON – Flow Through Overnight, sample not-bound to beads after overnight incubation. W – Wash. E – Elution. X – empty well. Ab – Antibody used for immunoprecipitation. Pink arrows are pointing on dissociated IgG antibody heavy and light chain – upper arrow correspond to 50 kDa, lower to 25 kDa. Pink star – marking additional bands visible after staining with anti-MGMT antibody, indicating covalent attachment to reactive cysteine of unknown target.

Data showing protein levels had been plotted as a log₂ function of ratio between samples treated with **GM-9** or **GM-10** vs DMSO or O⁶-BG (Figure 27). Negative log₂ values shows all proteins that were negatively enriched in covalent probe treatments compared to DMSO or O⁶-BG treated samples. Y-axis values indicate significance in difference between samples. The higher q-value, the higher significance of data point in the sample. The α -chloroacetamide did not show covalent modification of the protein target. Furthermore, it also failed to bind MGMT, which should be visible in upper left quarter of the plot shown in Figure 27B. Maleimide treatment, on the other hand, shows MGMT in upper, left quarter, as expected. Interestingly, RBM39 was found as a hit in GM-10 treated sample (Figure 27E). These data show differences in protein levels between covalent bifunctional molecules and DMSO samples. More relevant analysis is represented by α -chloroacetamide or maleimide samples compared to O⁶-BG treatments, where along enriched proteins PCLAF, KEAP, HSPB1 and IGBP1 can be seen (Figure 27C and F). PCLAF is a PCNA-associated factor, a known interactor of MGMT^{37–39}. KEAP1 protein expected for strong electrophiles^{40–43}, HSPB1 responds to oxidative stress or DNA damage⁴⁴ and IGF1P proteins contribute to DNA repair⁴⁵, providing indication that the developed assay is able to detect proteins interacting with MGMT.

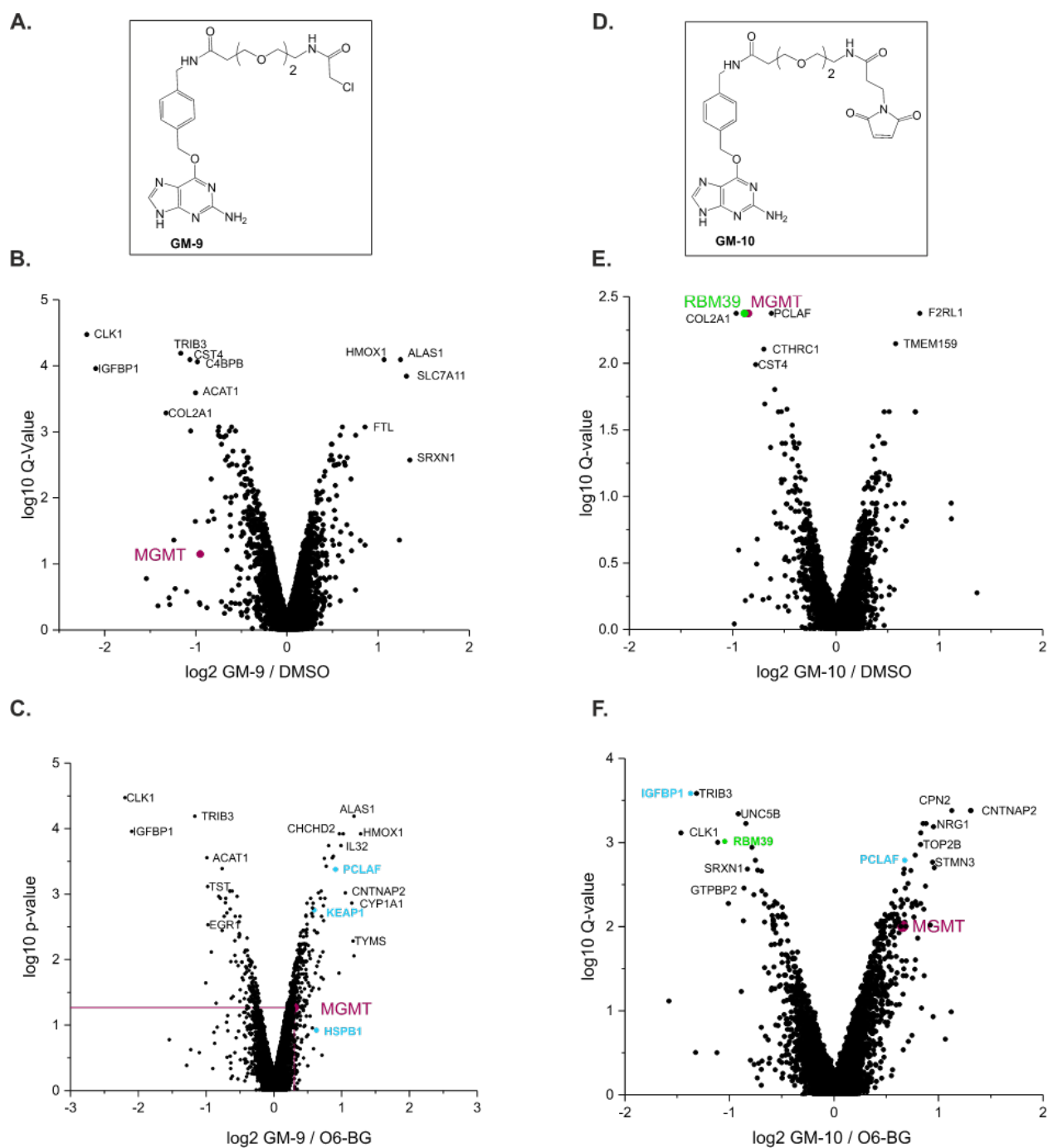


Figure 27. Proteomics data using the potential double degraders **GM-9** and **GM-10**. Chemical structure of **GM-9** (A) and **GM-10** (D). Protein level of TMT-labeled samples treated with **GM-9** and **GM-10** compared to treatments with DMSO (B and E) or O⁶-BG (C and F).

To evaluate proteomics data, specifically an RBM39 protein, HUH6 cells had been treated with **GM-10** as well as negative control compound **GM-11**, that lacks guanine (Figure 28B). As a positive control for RBM39 degradation, Indisulam (Ind.) was used^{46–48}. Lysed cells were used for western blot analysis staining with anti-RBM39 and anti-MGMT, antibodies (Figure 28A). Quantified protein levels are shown in Figure 28C.

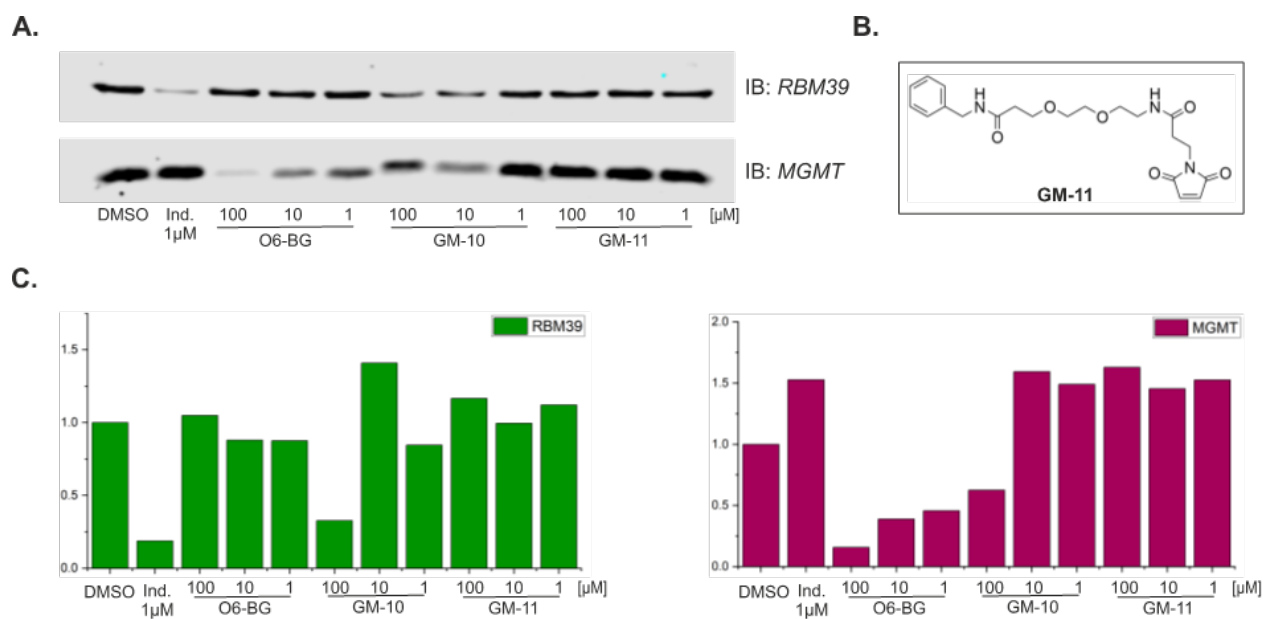


Figure 28. Data analysis of HUH6 cell treatment with **GM-10** and **GM-11**. Western blot with Anti-MGMT, Anti-RBM39 (A). Chemical structure of control compound **GM-11** (B). Quantified signal intensity of RBM39 and MGMT relative to protein content in DMSO treated sample (C).

As much as data presented in Figure 28 show promising results, it seems almost unlikely that such unspecific electrophile is so potent against one, very specific target such as RBM39. One possible explanation is the contamination coming from indisulam sample that could occur during pipetting.

Freshly synthesized **GM-10** compound had been used (**repGM-10**) and the experiment was repeated, this time showing no effect on RBM39 protein levels, (Figure 29). The results obtained suggest that future studies should continue. It remains unclear whether it is the assay sensitivity or sample preparation that restricts data evaluation.

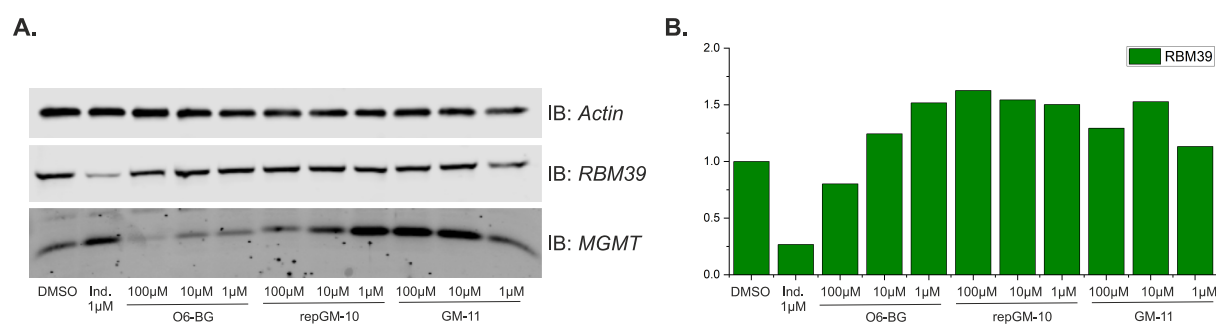


Figure 29. Western blot analysis of HUH6 cells treated with either DMSO, Indisulam (Ind.), *O*⁶-benzylguanine (*O*⁶-BG), re-synthesized **GM-10** (**repGM-10**) or negative control **GM-11** (A). Quantified data shown as relative protein concentration to vehicle treated (DMSO) sample (B).

11.2 Study of MGMT – POI degradation by protein fusion

Degron tags are frequently used as mechanistic tools in biology research, as a loss-of-function method for reducing levels of target protein⁴⁹. Two main techniques are being widely used: Auxin Inducible Degron (AID) system⁵⁰ and dTAG⁵¹. The disadvantage of the AID system is the requirement of introducing two proteins in studied system, while dTAG requires only one. AID, on the other hand, is triggered by addition of small molecule – auxin (Mw = 175), while dTAG requires large synthetic adducts of molecular weight over 1000. Using MGMT

MGMT probes with bifunctional molecules

as a degron tag would bring advantages of using one protein system and small organic molecule of molecular weight of 240 for O⁶-BG or 326 for lomeguatrib.

To validate the concept of covalent MGMT-triggered degradation, I decided to test whether MGMT-based fusion proteins could be degraded. In this setup the degradation will solely depend on MGMT active site modification. I chose Luciferase (60.6 kDa), GFP (29.4 kDa) and two generations of biotin ligase proteins BioID2 (26.4 kDa) and TurboID (35.4 kDa) as fusion targets in order to analyze the influence of protein size on MGMT degradation abilities. Almost all protein fusions had been cloned under the CMV promoter, transfected and overexpressed in HEK293T cells. Since CMV is a very strong promoter, I had to optimize transfection conditions in order to observe the correlation between MGMT degradation and amount of used plasmid (Figure 30).

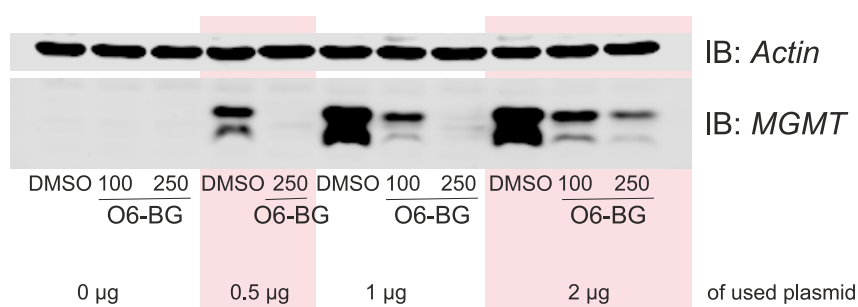


Figure 30. Western blot analysis of HEK293T cells transfected with either 0.5, 1 or 2 µg of plasmid per 3.5 cm². Transfected wells were treated with 100- or 250 µM of O⁶-BG. Bands intensity is validated by actin bands. Not transfected cells do not stain against MGMT antibody.

I settled on 0.5 µg of plasmid for transfection and 250 µM of O⁶-BG for MGMT degradation. Cells had been transfected with mentioned plasmids using control fusion of GFP-MGMT where a point mutation was introduced to the reactive cysteine in the MGMT active site: **GFP-MGMT(C145A)**. I was expecting all fusions to be degraded by MGMT upon treatment with O⁶-BG, except for GFP-MGMT(C145A). Results presented in Figure 31 prove that MGMT is capable to degrade proteins over twice its own size such as Luciferase.

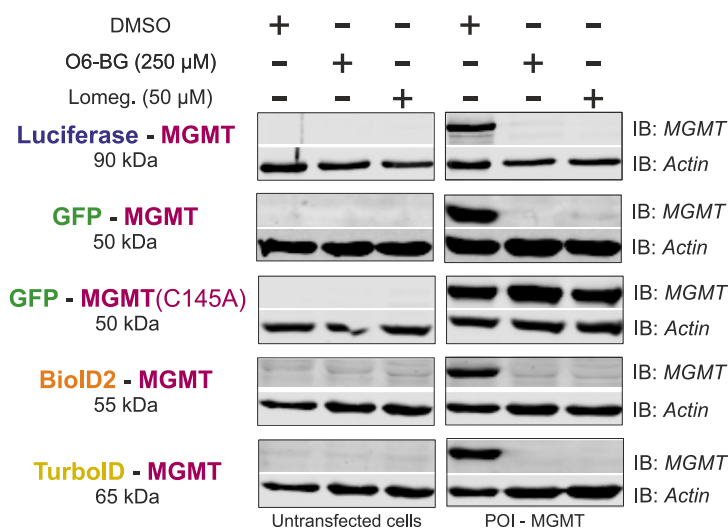


Figure 31. Western blot analysis of fusion-MGMT degradation.

12 Conclusions

Using MGMT as a tool for targeted protein degradation is challenging and results are often misleading. I presented data demonstrating MGMT capability of degrading other proteins via a covalent linker. Results show that MGMT activity is dependent on cysteine C145, where mutation leads to loss of interactions with its substrates - O⁶-benzylguanine and lomeguatrib. However, at this stage of understanding, MGMT is not able to degrade POI via non-covalent interactions. It remains unclear, whether the nature of this limitation arises from MGMT kinetics or carefully controlled proteasomal substrate recognition. This question should be addressed in future studies. A possible experiment could include cell-free study with all interacting partners – MGMT, POI, bifunctional molecule – as well as ubiquitin-dependent proteasomal machinery, such as E1, E2, E3 ligase, ubiquitin and 26S proteasome. Such controlled environment should provide conditions that allow clear readout. Alternatively, similar cell-free study with 20S proteasome could be carried out. Unlike 26S, 20S proteasome lacks the regulatory subunit which controls recognition of ubiquitinated substrates, but still contains the active proteolytic core⁵². If a POI would follow MGMT-mediated degradation upon treatment with non-covalent bifunctional molecule, the obvious conclusion would point to limitations caused by the 26S proteasomal regulatory subunit.

Presented results show successful application of MGMT as degron tag in human cells, adding to a toolbox of target validation methods. Although not relevant for bifunctional degraders, the mechanism of MGMT degradation was also a prominent question. Inspired by this result, in a later chapter (Chapter 4) I describe how I developed a biotin-ID fusion as a more effective system for pull-downs to probe the MGMT degradation mechanism.

13 Bibliography

1. Porkka, K. *et al.* Dasatinib crosses the blood-brain barrier and is an efficient therapy for central nervous system Philadelphia chromosome-positive leukemia. *Blood* **112**, 1005–1012 (2008).
2. Dalgıç, C. T. *et al.* Investigating the Role of JAK/STAT Pathway on Dasatinib-Induced Apoptosis for CML Cell Model K562. *Clin. Lymphoma Myeloma Leuk.* **15 Suppl**, S161-6 (2015).
3. Fry, D. W. *et al.* Specific inhibition of cyclin-dependent kinase 4/6 by PD 0332991 and associated antitumor activity in human tumor xenografts. *Mol. Cancer Ther.* **3**, 1427–1438 (2004).
4. Alghamdi, S. *et al.* BET protein inhibitor JQ1 inhibits growth and modulates WNT signaling in mesenchymal stem cells. *Stem Cell Res Ther* **7**, 22 (2016).
5. Filippakopoulos, P. *et al.* Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).
6. Rowley, J. D. A New Consistent Chromosomal Abnormality in Chronic Myelogenous Leukaemia identified by Quinacrine Fluorescence and Giemsa Staining. *Nature* **243**, 290–293 (1973).
7. Heisterkamp, N. *et al.* Localization of the c-abl oncogene adjacent to a translocation break point in chronic myelocytic leukaemia. *Nature* **306**, 239–242 (1983).
8. Bartram, C. R., de Klein, A., Hagemeijer, A. & van Agthoven, T. Translocation of c-abl oncogene correlates with the presence of a Philadelphia chromosome in chronic myelocytic leukaemia. *Nature* (1983).
9. Nowell, P. & Hungerford, D. A minute chromosome in human chronic granulocytic leukemia. *Landmarks in Medical Genetics* (2004).
10. Oda, T., Tamura, S., Matsuguchi, T., Griffin, J. D. & Druker, B. J. The SH2 domain of ABL is not required for factor-independent growth induced by BCR-ABL in a murine myeloid cell line. *Leukemia* **9**, 295–301 (1995).

11. Lugo, T. G., Pendergast, A. M., Muller, A. J. & Witte, O. N. Tyrosine kinase activity and transformation potency of bcr-abl oncogene products. *Science* **247**, 1079–1082 (1990).
12. Lai, A. C. *et al.* Modular PROTAC Design for the Degradation of Oncogenic BCR-ABL. *Angew. Chem. Int. Ed. Engl.* **55**, 807–810 (2016).
13. Yang, Y. *et al.* Global PROTAC Toolbox for Degrading BCR-ABL Overcomes Drug-Resistant Mutants and Adverse Effects. *J. Med. Chem.* **63**, 8567–8583 (2020).
14. Ottis, P. *et al.* Assessing different E3 ligases for small molecule induced protein ubiquitination and degradation. *ACS Chem. Biol.* **12**, 2570–2578 (2017).
15. Wang, C., Abegg, D., Hoch, D. G. & Adibekian, A. Chemoproteomics-Enabled Discovery of a Potent and Selective Inhibitor of the DNA Repair Protein MGMT. *Angew. Chem. Int. Ed. Engl.* **55**, 2911–2915 (2016).
16. Gerson, S. L., Trey, J. E., Miller, K. & Berger, N. A. Comparison of O⁶-alkylguanine-DNA alkyltransferase activity based on cellular DNA content in human, rat and mouse tissues. *Carcinogenesis* (1986).
17. Citron, M. *et al.* O⁶-methylguanine-DNA methyltransferase in human normal and tumor tissue from brain, lung, and ovary. *Cancer Res.* **51**, 4131–4134 (1991).
18. Kato, J., Matsushime, H., Hiebert, S. W., Ewen, M. E. & Sherr, C. J. Direct binding of cyclin D to the retinoblastoma gene product (pRb) and pRb phosphorylation by the cyclin D-dependent kinase CDK4. *Genes Dev.* **7**, 331–342 (1993).
19. Harbour, J. W., Luo, R. X., Dei Santi, A., Postigo, A. A. & Dean, D. C. Cdk phosphorylation triggers sequential intramolecular interactions that progressively block Rb functions as cells move through G1. *Cell* **98**, 859–869 (1999).
20. Meyerson, M. & Harlow, E. Identification of G1 kinase activity for cdk6, a novel cyclin D partner. *Mol. Cell. Biol.* **14**, 2077–2086 (1994).
21. Matsushime, H., Roussel, M. F., Ashmun, R. A. & Sherr, C. J. Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle. *Cell* **65**, 701–713 (1991).
22. Pardee, A. B. G1 events and regulation of cell proliferation. *Science* **246**, 603–608 (1989).
23. Dey, A., Nishiyama, A., Karpova, T., McNally, J. & Ozato, K. Brd4 marks select genes on mitotic chromatin and directs postmitotic transcription. *Mol. Biol. Cell* **20**, 4899–4909 (2009).
24. Yang, X.-J. Multisite protein modification and intramolecular signaling. *Oncogene* **24**, 1653–1662 (2005).
25. Yang, Z. *et al.* Recruitment of P-TEFb for stimulation of transcriptional elongation by the bromodomain protein Brd4. *Mol. Cell* **19**, 535–545 (2005).
26. Phelps, M. A. *et al.* Clinical response and pharmacokinetics from a phase 1 study of an active dosing schedule of flavopiridol in relapsed chronic lymphocytic leukemia. *Blood* **113**, 2637–2645 (2009).
27. Lu, J. *et al.* Hijacking the e3 ubiquitin ligase cereblon to efficiently target BRD4. *Chem. Biol.* **22**, 755–763 (2015).
28. Blake, R. A. GNE-0011, a novel monovalent BRD4 degrader. (2019).
29. Agashe, V. R., Shastry, M. C. & Udgaonkar, J. B. Initial hydrophobic collapse in the folding of barstar. *Nature* **377**, 754–757 (1995).
30. Lins, L. & Brasseur, R. The hydrophobic effect in protein folding. *FASEB J.* **9**, 535–540 (1995).
31. Giles, N. M., Giles, G. I. & Jacob, C. Multiple roles of cysteine in biocatalysis. *Biochem. Biophys. Res. Commun.* **300**, 1–4 (2003).
32. Dickens, F. Interaction of halogenacetates and SH compounds: The reaction of halogenacetic acids with glutathione and cysteine. The mechanism of iodoacetate poisoning of *Biochemical Journal* (1933).
33. Cole, R. D., Stein, W. H. & Moore, S. On the cysteine content of human hemoglobin. *J. Biol. Chem.* **233**, 1359–1363 (1958).
34. Barglow, K. T. & Cravatt, B. F. Discovering disease-associated enzymes by proteome reactivity profiling.

Chapter 3

- Chem. Biol.* **11**, 1523–1531 (2004).
35. Gunnoo, S. B. & Madder, A. Chemical Protein Modification through Cysteine. *Chembiochem* **17**, 529–553 (2016).
 36. Moore, J. E. & Ward, W. H. Cross-linking of Bovine Plasma Albumin and Wool Keratin. *J. Am. Chem. Soc.* **78**, 2414–2418 (1956).
 37. Mostofa, A., Punganuru, S. R., Madala, H. R. & Srivenugopal, K. S. S-phase Specific Downregulation of Human O6-Methylguanine DNA Methyltransferase (MGMT) and its Serendipitous Interactions with PCNA and p21cip1 Proteins in Glioma Cells. *Neoplasia* **20**, 305–323 (2018).
 38. Niture, S. K., Doneanu, C. E., Velu, C. S., Bailey, N. I. & Srivenugopal, K. S. Proteomic analysis of human O6-methylguanine-DNA methyltransferase by affinity chromatography and tandem mass spectrometry. *Biochem. Biophys. Res. Commun.* **337**, 1176–1184 (2005).
 39. Mostofa, A., Punganuru, S. R., Madala, H. R. & Srivenugopal, K. S. S-phase Specific Downregulation of Human O6-Methylguanine DNA Methyltransferase (MGMT) and its Serendipitous Interactions with PCNA and p21cip1 Proteins in Glioma Cells. *Neoplasia* **20**, 305–323 (2018).
 40. Dinkova-Kostova, A. T., Kostov, R. V. & Canning, P. Keap1, the cysteine-based mammalian intracellular sensor for electrophiles and oxidants. *Arch. Biochem. Biophys.* **617**, 84–93 (2017).
 41. Hong, F., Sekhar, K. R., Freeman, M. L. & Liebler, D. C. Specific patterns of electrophile adduction trigger Keap1 ubiquitination and Nrf2 activation. *J. Biol. Chem.* **280**, 31768–31775 (2005).
 42. Uruno, A. & Motohashi, H. The Keap1-Nrf2 system as an in vivo sensor for electrophiles. *Nitric Oxide* **25**, 153–160 (2011).
 43. Dinkova-Kostova, A. T., Kostov, R. V. & Canning, P. Keap1, the cysteine-based mammalian intracellular sensor for electrophiles and oxidants. *Arch. Biochem. Biophys.* **617**, 84–93 (2017).
 44. Ye, H. *et al.* HSPB1 Enhances SIRT2-Mediated G6PD Activation and Promotes Glioma Cell Proliferation. *PLoS One* **11**, e0164285 (2016).
 45. Baxter, R. C. IGF binding proteins in cancer: mechanistic and clinical insights. *Nat. Rev. Cancer* **14**, 329–341 (2014).
 46. Uehara, T. *et al.* Selective degradation of splicing factor CAPER α by anticancer sulfonamides. *Nat. Chem. Biol.* **13**, 675–680 (2017).
 47. Han, T. *et al.* Anticancer sulfonamides target splicing by inducing RBM39 degradation via recruitment to DCAF15. *Science* **356**, (2017).
 48. Bussiere, D. E. *et al.* Structural basis of indisulam-mediated RBM39 recruitment to DCAF15 E3 ligase complex. *Nat. Chem. Biol.* **16**, 15–23 (2020).
 49. Housden, B. E. *et al.* Loss-of-function genetic tools for animal models: cross-species and cross-platform differences. *Nat. Rev. Genet.* **18**, 24–40 (2017).
 50. Li, S., Prasanna, X., Salo, V. T., Vattulainen, I. & Ikonen, E. An efficient auxin-inducible degron system with low basal degradation in human cells. *Nat. Methods* **16**, 866–869 (2019).
 51. Nabet, B. *et al.* The dTAG system for immediate and target-specific protein degradation. *Nat. Chem. Biol.* **14**, 431–441 (2018).
 52. Unno, M. *et al.* The structure of the mammalian 20S proteasome at 2.75 Å resolution. *Structure* **10**, 609–618 (2002).

Chapter 4

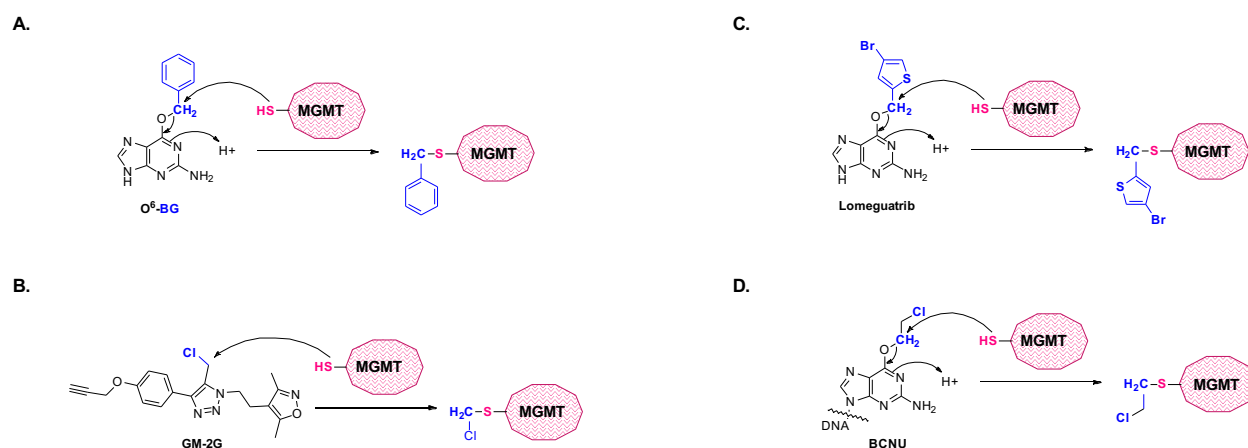
Mechanistic study of MGMT

14 Introduction

In order to use MGMT as a tool for targeted protein degradation (Chapter 3), it is important to understand the mechanism of its degradation. In this chapter, I will focus on MGMT degradation and resynthesis kinetics as well as interacting partners in the ubiquitin-dependent degradation pathway.

15 Degradation and de novo synthesis

Previous studies on MGMT kinetics were carried out in HT-29 human colon cancer cells¹⁻³. However, there are no reports on MGMT degradation studies in HUH6 cells. The substrate scope of MGMT was shown to be not restrained to O⁶ guanine lesions in DNA and that pseudosubstrates (small molecules that resemble DNA-based O⁶-G lesions) can also deactivate the protein^{1,2,4}. Reported studies suggest that the recovery of MGMT activity is slow and results from de novo protein synthesis⁵⁻⁸. I decided to perform a detailed mechanistic study concerning the MGMT degradation and de novo synthesis kinetics in HUH6 cell line. Both an O⁶-guanine alkylating agent (BCNU) as well as pseudosubstrates, such as O⁶-BG, lomeguatrib and **GM-2G** were used (Scheme 9).



Scheme 9. MGMT alkylation small molecule pseudo-substrates (A-C) or DNA alkylation (D).

15.1 Protein degradation rate and de novo synthesis

In order to determine the MGMT degradation rate, a time-point experiment was performed in HUH6 cells with either 5 or 250 μM of O⁶-BG. MGMT levels were normalized to a DMSO sample (Figure 32A). Presented data show that MGMT levels differ depending on O⁶-BG concentration. It seems like it takes over 6 hours for 5 μM O⁶-BG to start degrading MGMT, while 250 μM O⁶-BG initiate immediate degradation.

To understand whether MGMT degradation is cell-line dependent, I also examined degradation in a cervical cancer cell line (HeLa) as a comparison. Resulting data show a slower, but consistent degradation rate confirming the hypothesis that the MGMT activity is characteristic for specific cell type (Figure 32B).

To determine *de novo* protein synthesis rate, cells were depleted from MGMT protein levels by treatments with O⁶-BG. 24 hours of treatments were found to be sufficient for 95% of MGMT depletion (Figure 32A). After that time point, cell media was replenished, and cells collected after 0.5, 1.5, 4, 7.5, 19.5 21.5 and 23.5 hours. MGMT protein levels were normalized to a DMSO-treated sample (Figure 32C). Obtained results confirm that MGMT is a relatively stable protein with half-life greater than 24 hours⁹.

O⁶-BG was further used to confirm its specificity towards MGMT in a TMT-labeled proteomics experiment, showing significant protein downregulation (Figure 32D).

15.2 Source of alkylation

To further determine the influence of different class of MGMT degraders I looked into DNA alkylating agent and pseudosubstrates. I expect the DNA alkylating agent BCNU to act slower on MGMT degradation, since the compound has to first alkylate DNA, then recruit MGMT to the modified site and finally alkylate MGMT triggering its degradation. Unlike BCNU, pseudosubstrates are directly modifying MGMT's reactive cysteine. In order to test that hypothesis, HUH6 cells had been treated with one-fold of magnitude increments in concentration of all four compounds (150 μ M, 15 μ M, 1.5 μ M). DMSO was used as a negative control for MGMT degradation (Figure 32E and F).

The presented data support the hypothesis that alkylating agents require significantly more time or higher concentrations to degrade more than 50% of endogenous MGMT compared to pseudosubstrates. From all three substrates, the most potent is **GM-2G**, followed by lomeguatrib and O⁶-BG, respectively (Figure 32E, F). Knowing that 6 hours treatments with pseudosubstrates for MGMT in HUH6 cells are sufficient to observe more than 50% changes in MGMT levels (Figure 32A), those conditions were chosen for further study of the degradation mechanism.

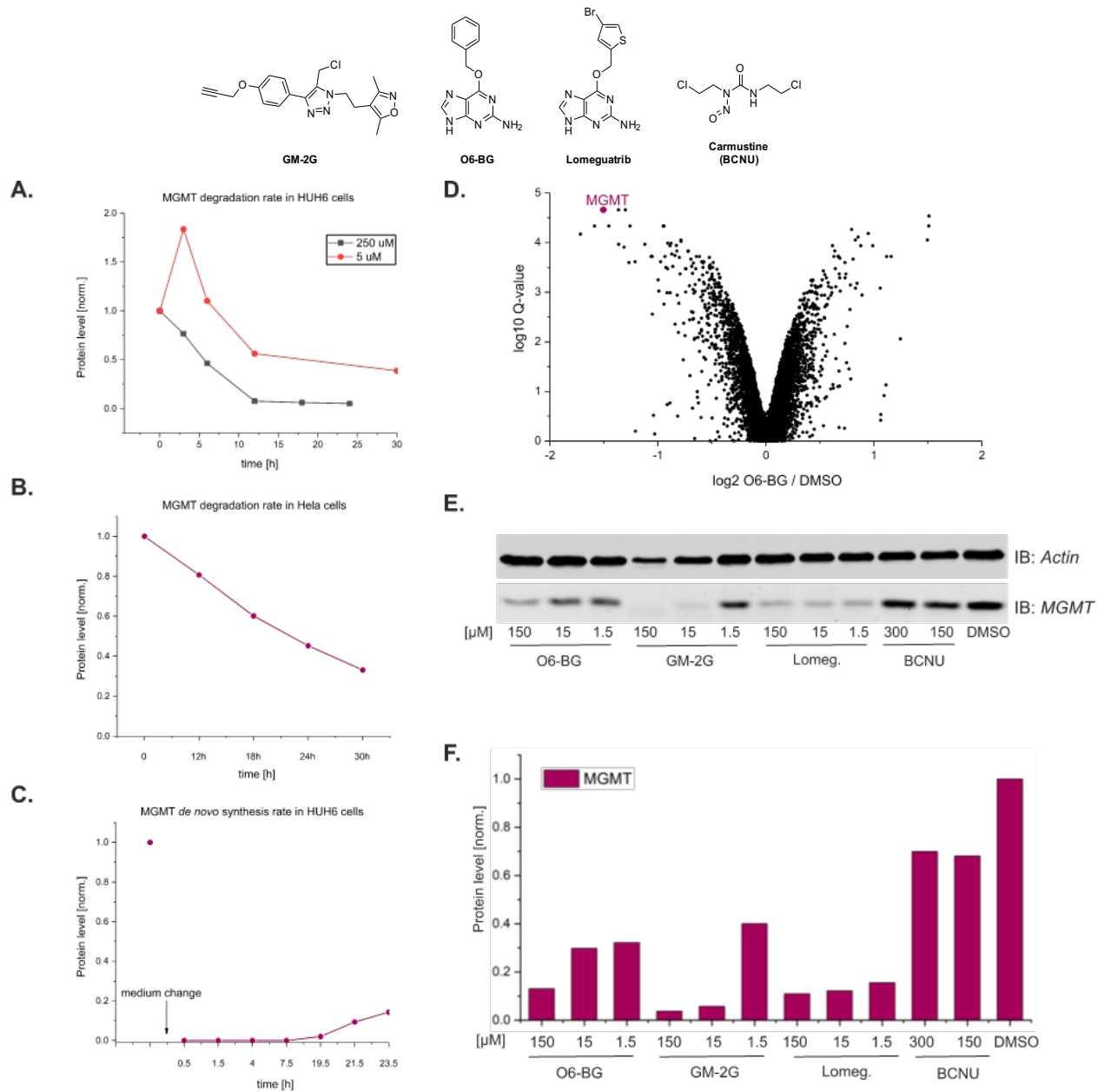


Figure 32. MGMT protein level in HUH cells after treatment with 250 μM of O^6 -BG (squares) and 5 μM (circles) (A); MGMT protein level in HeLa after treatment with 5 μM of O^6 -BG (B); MGMT protein level in HUH6 cells depleted with O^6 -BG 250 μM for 24 hours followed by media replenishment (C); TMT proteomics data establish the selectivity of MGMT inhibitor O^6 -BG 250 μM with MGMT identified as a significantly down-regulated protein (purple dot) (D). MGMT degradation rates influenced by different alkylation sources during 6 hours treatments represent by the western blot analysis (E) and quantified data with relative protein content in DMSO treated sample (F).

16 The inhibitory study of ubiquitin signaling proteasomal pathway

MGMT had been reported to undergo ubiquitin-mediated proteasome pathway (Figure 33A) in human colorectal cancer (HT-29) and leukemia (CEM) cell lines¹⁰. To confirm that the same pathway is involved in *Hepatocellular carcinoma* (HUH6), I decided to incubate cells with UPP pathway inhibitors, such as ubiquitin-activating enzyme (E1) inhibitor - **TAK243**, proteasome inhibitor - **bortezomib** and **MG-132** as well as neddylation inhibitor (**MLN4924**) for four hours, followed by 6- or 20- hours treatments with O^6 -BG (Figure 33B

Chapter 4

and C, respectively). If MGMT undergoes ubiquitin-mediated proteasomal pathway (UPP), inhibiting proteasomal as well as E1 activity should result in accumulation of MGMT in the cell despite treatments with O⁶-BG. Majority of protein degradation in cell is controlled by either proteasomal or lysosomal degradation pathway^{11–14}. Therefore, as a control I used lysosome inhibitors – Chloroquine (Ch-Q) or ammonium chloride (NH₄Cl). If MGMT will not be lysed by UPP, it would most likely be degraded by lysosome, which will be observed after using lysosome inhibitors. Finally, a few DNA repair proteins had been reported to be substrates for Cullin-RING 4 family (CUL4) of E3 ligase, such as: PCNA^{15,16}, XPC^{17–19}, DDB2^{20–22}, H2A^{23–25}, SET8²⁶ or H4²⁷. To verify whether MGMT is a substrate of Cullin-RING E3 ligase, I decided to use highly selective NEDD8-activating enzyme inhibitor (MLN4924)²⁸. MLN4924 prevents entire neddylation modification cascade, leading to deactivation of all Cullin-RING ligases²⁸.

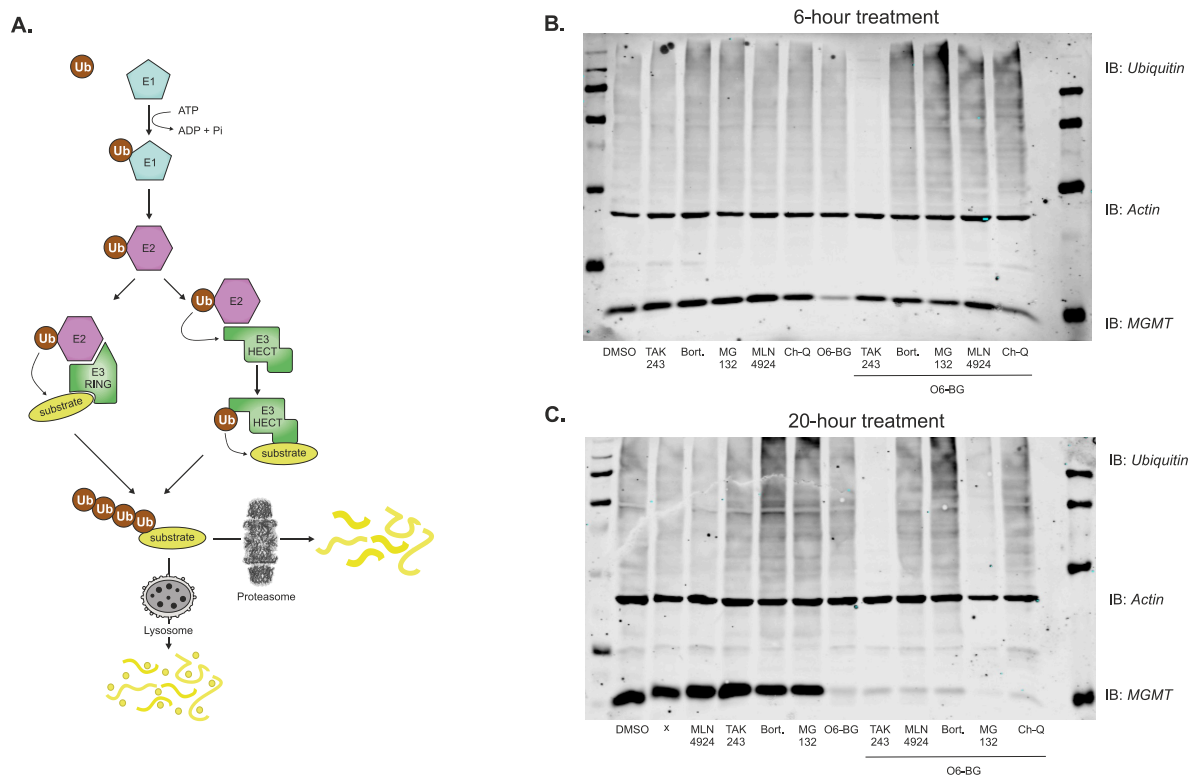


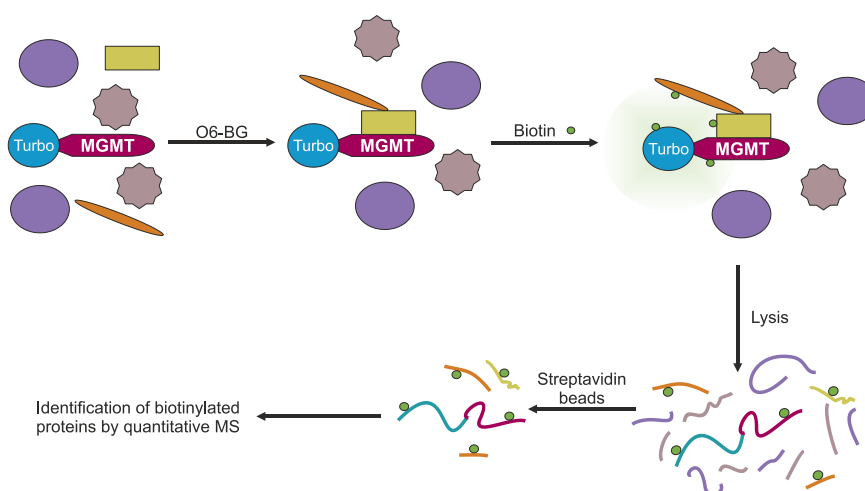
Figure 33. Inhibition strategy of MGMT degradation pathway study. Ubiquitin mediated degradation pathway leading to either proteasomal or lysosomal degradation (A). Western blot analysis of inhibitory study with 4-hour incubation with E1, E3, proteasome or lysosome inhibitor followed by 6-hour (B) or 20-hour (C) treatment with 250 μ M O⁶-BG. All experiments were performed in triplicate, images show representative sample.

The results show clear differences in treatment time, indicating that during the longer time points specificity of inhibitors might be compromised and data interpretation became unreliable. After 6 hours of O⁶-BG treatment, MGMT does not get degraded while UPP pathway machinery is inhibited (E1 and proteasome). Blocking lysosomes with Ch-Q (for NH₄Cl inhibition see Experimental Procedures, Figure S3) MGMT levels decreased after treatment with O⁶-BG, indicating that MGMT is not degrading via lysosomal pathway. Interestingly, MGMT seems to undergo degradation via Cullin-RING family of E3 ligases, which had not been previously reported. Blocking neddylation with MLN4924 prevents MGMT from degradation, proving its role in UPP in HUH6 cells.

17 Proximity labeling of MGMT interacting partners

A highly valuable tool for study the spatial interaction of protein in living cells based on enzyme-catalyzed proximity labeling (PL) had been recently developed²⁹. Used as an alternative to immunoprecipitation, the advantage of PL is the covalent tagging of diffusible endogenous biomolecules in a proximity-dependent manner as well as operating in living cells, preserving native spatial relationships³⁰. Tagged biomolecules can be enriched on anti-tag beads and further characterized by high-sensitivity, quantitative mass spectrometry (MS)^{31–33}.

To probe MGMT interacting partners, I decided to use biotin ligase as PL enzyme^{34–38} (Scheme 10). When fused with MGMT, biotin ligase will covalently biotinylate proteins in close proximity to MGMT, which will be further identified by MS after cell lysis and enrichment on streptavidin beads. Two pathways were used for PL approach: first, using BioID2 as biotin ligase³⁹, which was fused to *N*-terminus of MGMT protein and transiently expressed under CMV promoter. Second, TurboID⁴⁰, fused to *N*-terminus of MGMT protein and stably expressed under TET-inducible promoter.



*Scheme 10 TurboID as a PL enzyme for MGMT interactive partners study. TurboID is fused to *N*-terminal of MGMT and upon treatment with MGMT substrate (*O*⁶-BG) protein starts interaction with its corresponding ubiquitin-mediated proteasome pathway members. After addition of biotin TurboID covalently tags proteins in close proximity, that can be further analyzed by quantitative MS after lysis and enrichment on streptavidin beads.*

17.1 Overexpressed BioID-MGMT

BioID2 (26.4 kDa) was fused to the *N*-terminus of MGMT and transiently expressed in HEK293T cells. Biotin was supplied to cell media (50 μ M) to facilitate biotin ligase activity and *O*⁶-BG was added (200 μ M) in order to initiate MGMT degradation. Cells were incubated for 24 hours and subsequently, the lysed samples were enriched on streptavidin beads and analyzed by quantitative MS (Figure 34). Expected results would show MGMT interacting partners during degradation signaling pathway.

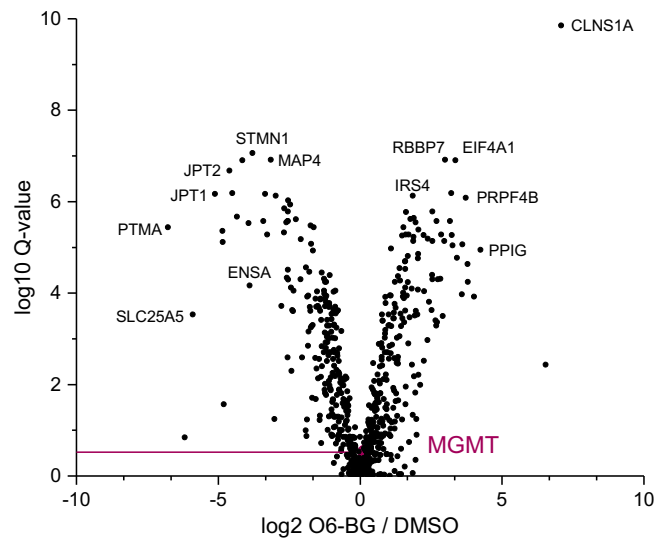


Figure 34. Protein Level analysis of BioID2-MGMT pulldown after treatment with 50 μ M of O6-BG in relevance to DMSO treatment.

The results presented in Figure 34 do not deliver an immediate answer to the stated question regarding protein-protein interactions during MGMT degradation pathway, since none of the strongly downregulated or upregulated proteins are part of proteasomal degradation pathway. BioID2 requires 16 hours for biotin tagging in order to produce a sufficient amount of labeling that directly translates to the amount of sample material required for proteomic analysis. That aspect of BioID2 unable the study of rapid and dynamic processes that occur on the timescale of minutes to few hours. Additionally, the transient expression under a strong promoter resulted in high off-target labeling.

To address those limitations, I decided to use a new, more efficient generation of biotin ligase – TurboID, which requires as little as 10 minutes for sufficient biotinylation activity³⁹. Additionally, stable expression under an inducible promoter is preferable over transient expression since it offers more precise control over protein levels.

17.2 Engineered HEK293T cell line stably expressing TurboID-MGMT

I decided to use CRISPR-Cas9 technique to incorporate TurboID-MGMT protein fusion in HEK293T genome. Randomly inserted genes might suffer from silencing, making their expression unpredictable^{41,42}. To overcome this issue, a few validated human genomic safe-harbors (GSHs) have been reported allowing for predictable and stable expression without negative effects on the host cell. I chose *AAVSI* gene locus on chromosome 19^{43,44}, which has been shown in HEK293T cells^{45,46} to show no disruption of cell cycle. I used a two plasmid approach: one with incorporated TurboID-MGMT to inducible tet-promoter pMK243 that contains left- and right- homology arms from *AAVSI* integration site⁴⁷ and pX330 DL plasmid bearing the Cas9 gene along with a coding sequence for the guide RNA to target the *AAVSI* locus⁴⁷ (Figure 35A).

Successful integration was confirmed by agarose gel analysis (Figure 35B) and Sanger sequencing (see Experimental). In order to validate activity of both proteins – MGMT and TurboID – I performed two separate experiments. Firstly, to confirm MGMT activity, cells were treated with O⁶-BG (250 μ M) over 3 – 24 hours. Degrading TurboID-MGMT fusion would confirm MGMT activity (Figure 35C). Secondly, I used three different biotin concentrations and three treatment time points to evaluate TurboID biotinylation activity (Figure 35D).

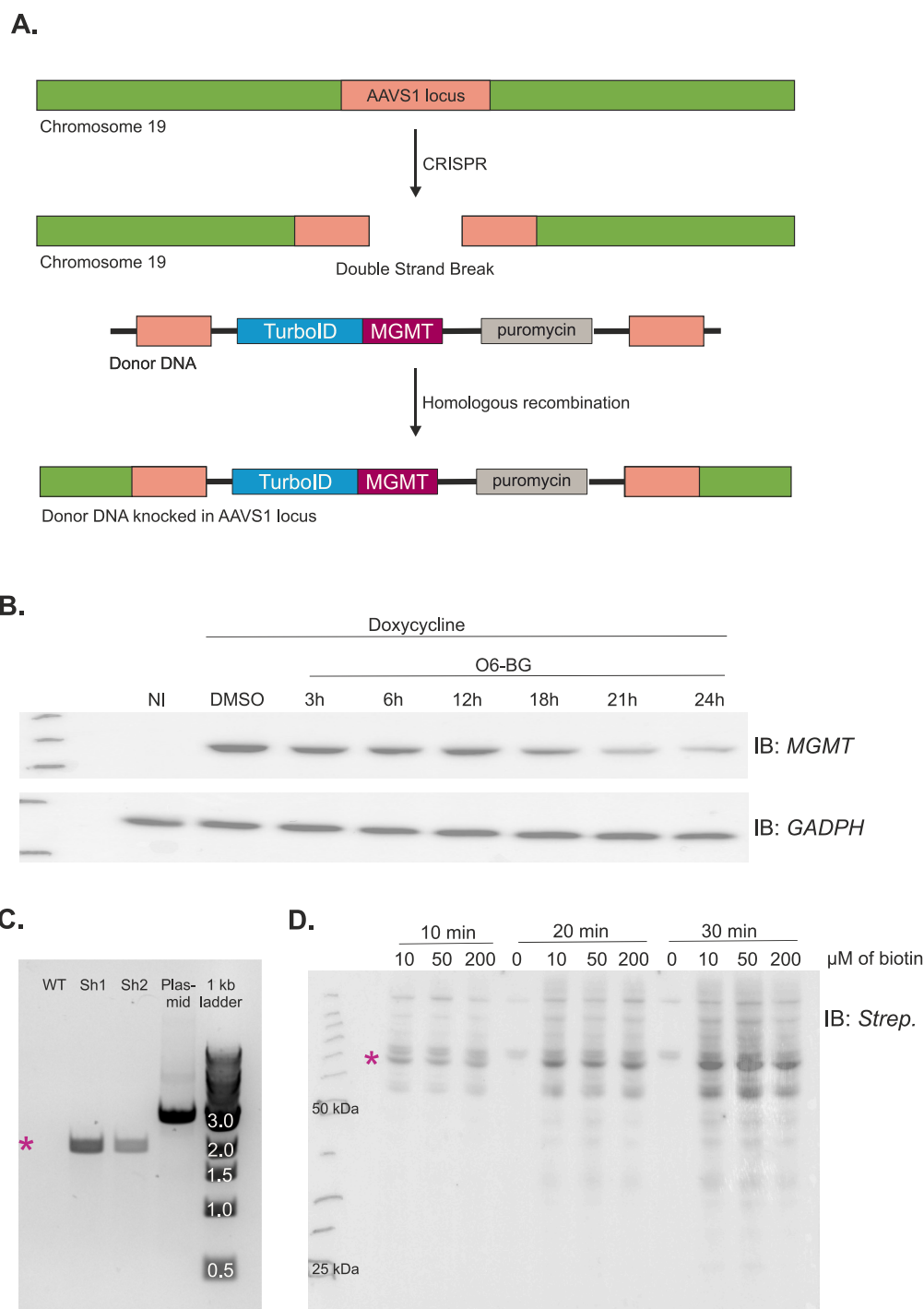


Figure 35. Evaluation of generically incorporated TurboID-MGMT protein fusion. Schematic representation of CRISPR-Cas9 mediated TurboID-MGMT knock-in into AAVS1 locus (A); Western blot analysis of induced TurboID-MGMT fusion indicating MGMT activity after treatment with O⁶-BG. NI – not induced (B); 1% agarose gel showing PCR amplified products with forward primer located outside AAVS1 safe harbor locus, but on Chromosome 19 and reverse primer inside the TurboID-MGMT cassette within AAVS1 locus. Only properly inserted sequence would result in positive readout of a 2 kbp product. Wild type (WT) cells as well as plasmid used for transfection (Plasmid) would not show 2 kbp band. Pink star corresponds to desired product location on the gel (C); Western blot analysis of TurboID activity. Induced TurboID-MGMT fusion was incubated with 0, 10, 50 or 200 μM of biotin for either 10, 20 or 30 minutes (D); Streptavidin-HRP was used for visualization. Pink star corresponds to desired product location on the blot.

Figure 35B shows the tightly controlled TET-inducible promoter, which suppresses expression of TurboID-MGMT without addition of doxycycline. Additionally, resulting protein fusion contains active MGMT, as evidenced by its degradation in response to O⁶-BG treatments (Figure 35B). Interestingly, TurboID shows little

Chapter 4

dependence on biotin concentration but strong dependence on incubation time (Figure 35D), with 10 minutes as the optimal for sufficiently detectable labeling with minimum off-target events (Figure 35D).

With the system engineered and the optimal labeling conditions in hand, I turned to task of labeling and isolating MGMT interacting partners. In order to see MGMT's interacting partners in response to O⁶-BG treatments, I choose a relatively early 6-hour treatment time to make sure not all MGMT is degraded. Longer incubation with O⁶-BG could cause depletion in TurboID-MGMT levels before biotin-staining would be completed. First, TurboID-MGMT expression was induced for 16 hours, and this was followed by a medium change into O⁶-BG-containing medium. After six hours in the presence of 250 μM O⁶-BG, the cells were treated for a further 10 min with 10 μM biotin. Cells were collected, lysed (Input sample collection) and incubated on streptavidin beads for 4 hours. Unbound cell lysate was collected as a flow through (FT) sample. Beads were further washed three times and 10% of beads were used for elution and compared with the same volume of sample as input and FT indicating successful on-bead enrichment (Figure 36A). Remaining 90% of beads were used for proteomic analysis showing the difference in peptide count between O⁶-BG treated samples compared to DMSO (Figure 36B).

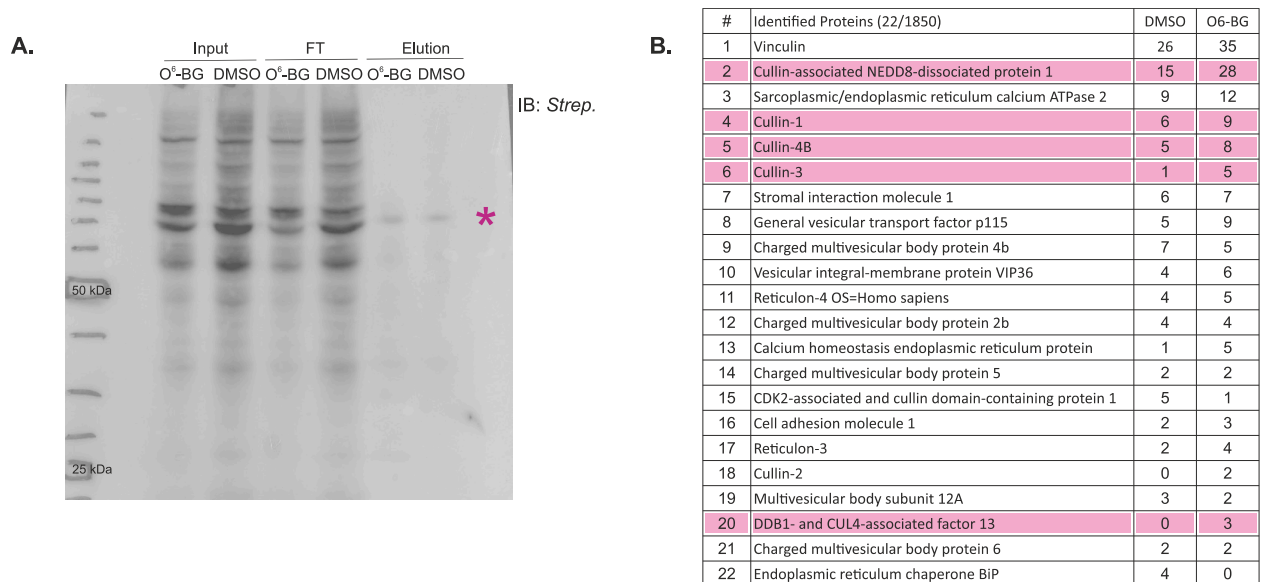


Figure 36. MGMT interacting partners in UPP pathway. Immuno-precipitation of TurboID-MGMT showing cell lysate in three stages of sample on bead enrichment: input, Flow through (FT) and elution for O⁶-BG and DMSO treated samples, respectively (A). Pink star corresponds to TurboID-MGMT protein bands on the blot; Proteomics data showing top scored total peptide count between DMSO and O⁶-BG treated samples (B). UPP relevant proteins are highlighted in pink.

Data shown in Figure 36B represent single experiment, therefore statistical relevance of obtained results cannot be evaluated. Results presented in form of a top 5% score of total peptide count between DMSO and O⁶-BG, indicates that the excess of peptide count seen in O⁶-BG sample compared to DMSO sample show proteins involved in MGMT degradation pathway. In the pool of identified proteins, three are Cullin-RING E3 ligases (CUL1, CUL3, CUL4B), one is a Cullin-associated NEDD8-dissociated protein 1 and one in DDB1 – and CUL4 – associated factor, all in support of previously observed involvement of Cullin-RING family of E3 ligase in MGMT degradation pathway.

Further biological replicates are needed to increase the significance of these results. Apart from MGMT pseudosubstrate treatment, I decided to include also DNA-alkylating agent samples. All samples had been performed in triplicate and following previously described conditions, the enrichment efficiency was evaluated by

western blot analysis (Figure 37A). Proteomics data for the BCNU sample (Figure 37B) do not point directly at any UPP partners, making its interpretation difficult. The sample for lomeguatrib is more promising showing many proteasome components in the most significant hits (Figure 37C, positive side of the x-axis). A troubling reality though is that there was little overlap between the first dataset (Figure 36) and the next one (Figure 37). I believe, however, that the enrichment step can still be optimized and that this will improve the signal for the enriched proteins. TurboID-MGMT is the most highly biotinylated protein (intramolecular reaction for labeling) and the streptavidin beads are saturated mostly with self-biotinylated protein fusion. This hypothesis is consistent with the fact that the flow-through of the enrichment protocol (Figure 37A) still shows substantial biotinylated protein. A potential solution would be to preclean the cell lysate with anti-MGMT antibody incorporated on agarose beads. Remaining cell lysate lacking TurboID-MGMT would then be subjected to the streptavidin bead enrichment. Through this two-step purification method, the sample will be enriched in low-level biotin-labeled proteins, which might otherwise not bind to the surface of strep-coated beads.

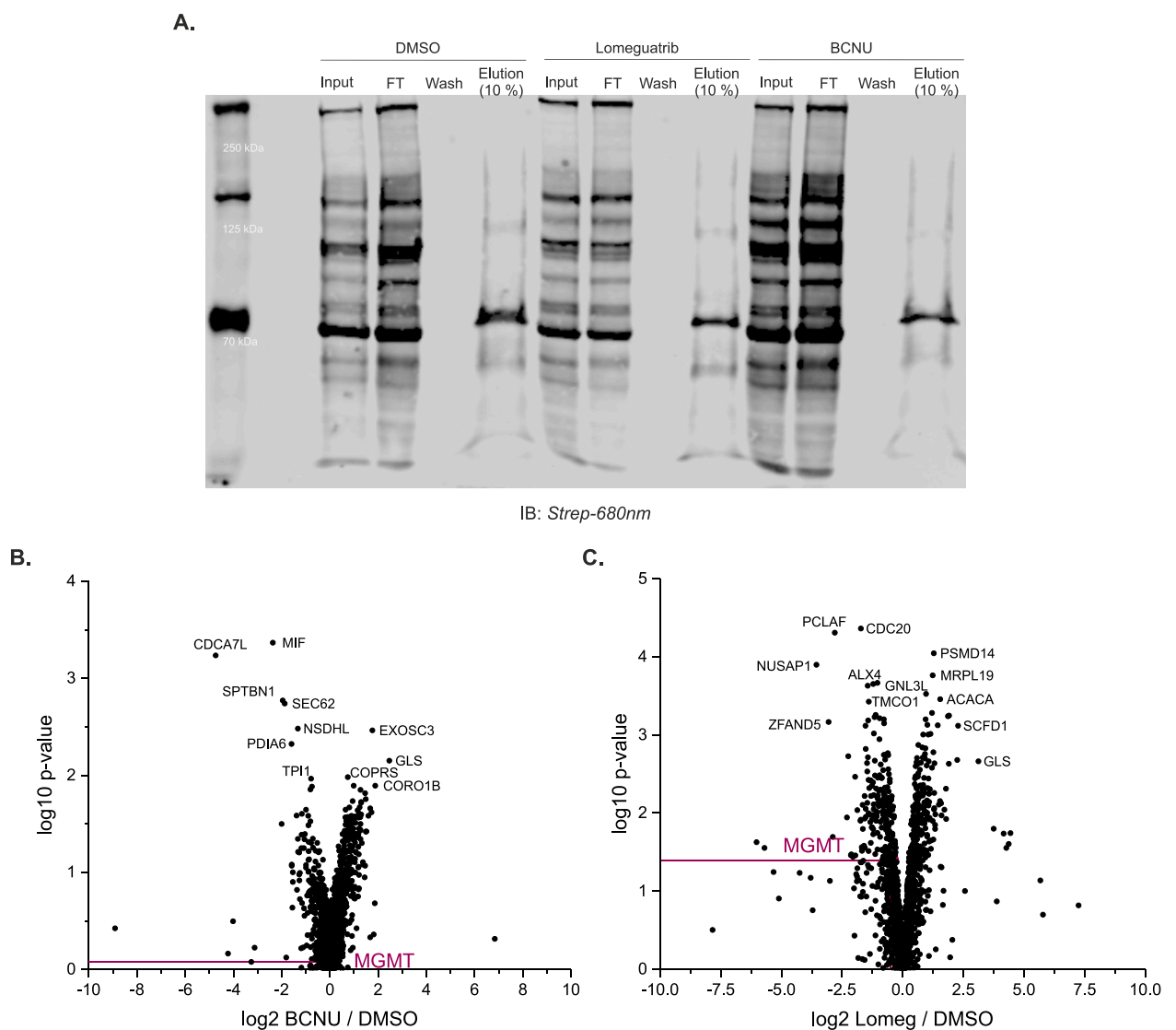


Figure 37. IP evaluation replicates performed in Basel (A). Proteomics results showing protein level after treatments with BCNU (B) or lomeguatrib (C).

17.3 The inhibitory study of UPP in engineered HEK293T cell line, stably expressing TurboID-MGMT

An interesting aspect of MGMT degradation pathway is whether MGMT degradation is affected by the presence of another protein. In order to study differences in the UPP pathway in the presence of other proteins, I decided to choose the engineering cell line stably expressing TurboID-MGMT fusion. At first, I decided to establish conditions in which the expression of MGMT will not be too high – high concentration of doxycycline in cell medium will cause constant protein expression, masking the degradation efficiency. Hence first experimental set up was a doxycycline titration analysis, where 1125, 280, 70, 36, 19, 9 nM of doxycycline was used, followed by 6-hour treatment with lomeguatrib (50 μ M). The resulting data show that \sim 1 μ M of doxycycline is a minimal concentration required for TET-promoter to induce protein expression and 50 μ M lomeguatrib degrades only 60% of MGMT (Figure 38A and B). With these conditions in hand, I performed an inhibitory study as described in Section 16, changing O⁶-BG to lomeguatrib.

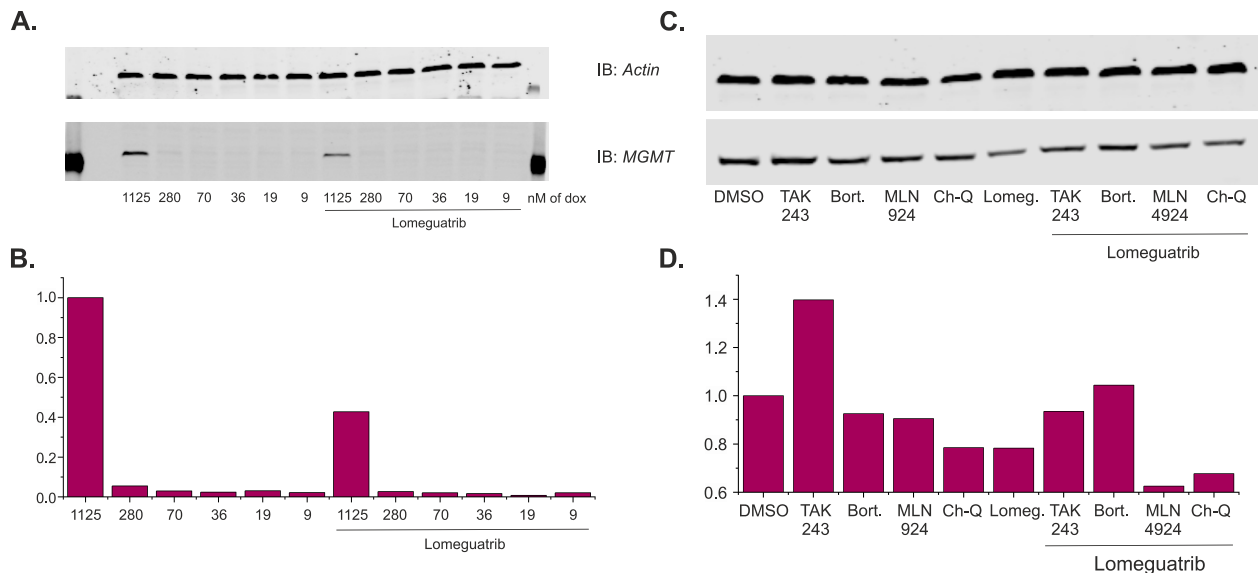


Figure 38. Probing MGMT-fusion degradation pathway. Doxycycline titration – determining optimal doxycycline concentration (A, B). Inhibitory mechanistic analysis of UPP pathway (C, D).

Results presented in Figure 38 C and D show much smaller dynamic range of lomeguatrib degradation abilities, making it difficult to form clear conclusions. However, it seems like E1 inhibitor (TAK243) as well as proteasome inhibitor (Bortezomib) are both being able to prevent MGMT from degradation, indicating its involvement in MGMT degradation pathway. Surprisingly, neddylation inhibitor - MLN924 does not prevent MGMT from degradation, indicating that Cullin-RING family of E3 ligases are not recognizing MGMT-fusion substrate. There are two possible explanation for observed data. One, suggesting that there is more than one signaling pathway responsible for MGMT utilization, addressing the importance of MGMT in DNA maintenance and cell proliferation. Alternatively, UPP partners responsible for MGMT degradation, just like MGMT activity, are cell line dependent.

18 siRNA inhibition

One of the great tools used for protein knockdown by gene silencing is incorporation of RNA interference (RNAi)⁴⁸. During this process the expression of a target gene is suppressed by the selective inactivation of its corresponding mRNA using double-stranded RNA (dsRNA), such as small interfering RNAs (siRNAs) or short hairpin RNAs (shRNAs)^{49,50}. Once delivered to the cell cytoplasm, dsRNA activates RNAi leading to degradation of targeted gene mRNA. One of many delivery methods of the dsRNA to the cell is based on lentiviral transduction, which stably integrates the shRNA into the cell's genome, allowing for persistent expression.

18.1 siRNA Cul1, CUL4A, CUL4B.

To further evaluate Cullin-RING family of E3 ligases involved in MGMT degradation pathway, I decided to start with Cullin-RING 1 (CUL1) and Cullin-RING 4 (CUL4) subfamilies first. Other DNA repair proteins, e.g. PCNA^{15,16}, XPC¹⁷⁻¹⁹, DDB2²⁰⁻²², H2A²³⁻²⁵ had been reportedly regulated by those E3 ligases¹¹⁻²². I used lentiviral transduction to deliver shRNA to HEK293T cells stably expressing TurboID-MGMT protein fusion under inducible promoter in order to silence CUL1, CUL4A and CUL4B E3 ligase within the cell. TurboID-MGMT fusion expression was then induced with doxycycline and treated with 250 μ M of O⁶-BG for 24 hours in order to induce the degradation (Figure 39). Samples in which MGMT degradation would be prevented by silencing Cullin-RING ligase would indicate their involvement in mediating MGMT degradation.

Presented data show ambiguous results since the experimental set up calls for silencing of a crucial cell survival protein (CUL). The balance between silencing CUL1, CUL4A and CUL4B just enough to observe its effect on MGMT degradation but not too strong to cause cell apoptosis was challenging. Nevertheless, the western blot analysis of MGMT levels (Figure 39A) show an involvement of CUL4A in MGMT degradation pathway. Unfortunately, any visible changes in CUL4A levels cannot be seen (Figure 39C).

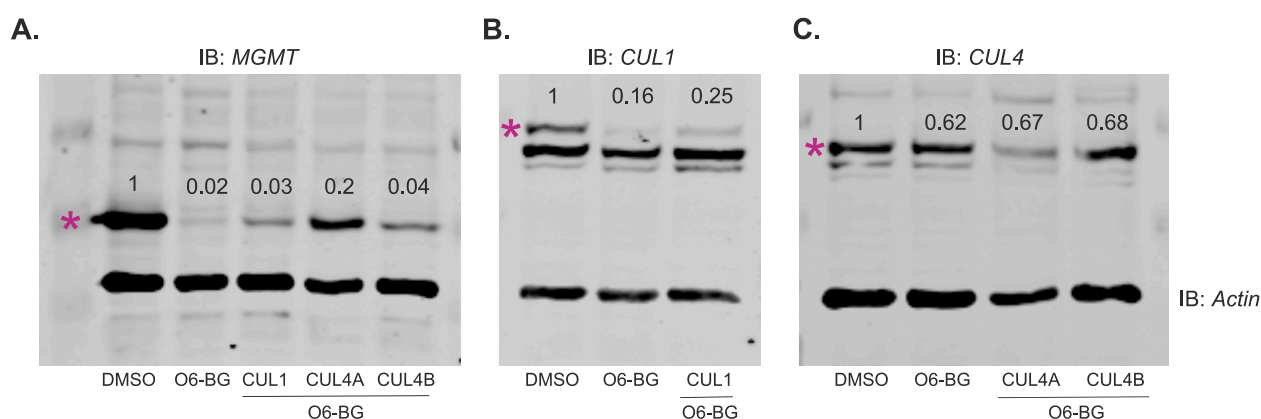


Figure 39. siRNA treatments of generically incorporated TurboID-MGMT protein fusion analyzed by western blot and stained with MGMT antibody (A), Cullin1 antibody (B) or Cullin4 antibody (C). Pink star corresponds to desired protein bands on the blot. Quantified and normalized to DMSO sample had been noted above the analyzed band on each blot.

18.2 siRNA RAD18

One recent study on the MGMT degradation pathway in human *Nasopharyngeal carcinoma* cells suggested involvement of the RAD18 E3 ligase is mediating MGMT ubiquitination⁵¹. RAD18 is a DNA damage-activated, RING-type E3 ubiquitin ligase that functions as a key regulator in homology-directed repair of the DNA

Chapter 4

damage signal as well as post-replication repair⁵²⁻⁵⁶. To evaluate those results, siRNA had been used to silence RAD18 expression in HUH6 cells, followed by O⁶-BG (100 μ M) and lomeguatrib (1.5 nM) treatments (Figure 40). Presented results do not imply that RAD18 is involved in MGMT utilization pathway, since silencing of RAD18 does not prevent MGMT degradation after treatments with O⁶-BG and lomeguatrib. However, the study should include more biological replicates as well as verification of siRNA efficiency, for example by using anti-RAD18 antibody. RAD18-mediated MGMT degradation study is yet another example, how MGMT degradation pathway is cell line specific.

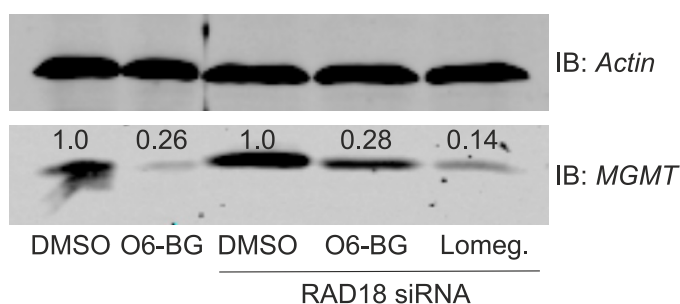


Figure 40. RAD18-mediated MGMT degradation study through siRNA protein silencing.

19 Conclusions

MGMT degradation and *de novo* protein synthesis rates depend on the source of alkylation. Direct DNA-alkylating agents require higher concentrations or longer incubation times in comparison to pseudosubstrates (O⁶-BG, lomeguatrib, **GM-2G**) to elicit MGMT degradation. Alkylated MGMT undergoes UPP degradation and there seem to be several E3 ligases involved in the process. I identified two candidate E3 ligases, homing in on the specific substrate receptors directing the degradation is the next challenge. The inexorable degradation of MGMT is supported by the earlier study of Pegg *et al.*, where, in an attempt to find one specific ubiquitination site of MGMT, all of the twelve lysines were mutated to arginine⁵⁷. However, neither a single mutant, nor two combinations of triple mutants were able to prevent MGMT from ubiquitin labeling⁵⁷. This implies a strong pressure for ubiquitination of MGMT on any available lysine. I am currently refining the TurboID labeling/proteomics protocol in an attempt to clarify the precise UPP pathways and the associated substrate receptors involved in MGMT degradation.

20 Bibliography

1. Moschel, R. C., McDougall, M. G., Dolan, M. E., Stine, L. & Pegg, A. E. Structural features of substituted purine derivatives compatible with depletion of human O⁶-alkylguanine-DNA alkyltransferase. *J. Med. Chem.* **35**, 4486–4491 (1992).
2. Dolan, M. E., Moschel, R. C. & Pegg, A. E. Depletion of mammalian O⁶-alkylguanine-DNA alkyltransferase activity by O⁶-benzylguanine provides a means to evaluate the role of this protein in protection against carcinogenic and therapeutic alkylating agents. *Proc. Natl. Acad. Sci. USA* **87**, 5368–5372 (1990).
3. Srivenugopal, K. S., Yuan, X. H., Friedman, H. S. & Ali-Osman, F. Ubiquitination-dependent proteolysis of O⁶-methylguanine-DNA methyltransferase in human and murine tumor cells following inactivation with O⁶-benzylguanine or 1,3-bis(2-chloroethyl)-1-nitrosourea. *Biochemistry* **35**, 1328–1334 (1996).
4. Pegg, A. E. Mammalian O⁶-alkylguanine-DNA alkyltransferase: regulation and importance in response to alkylating carcinogenic and therapeutic agents. *Cancer Res.* **50**, 6119–6129 (1990).

5. Pieper, R. O. Futscher. BW, Dong. Q.. and Erickson, LC Effects of streptozotocin/bis-chloroethylnitrosourea combination therapy on O6-methylguanine DNA *Cancer Res.* (1991).
6. Marathi, U. K., Kroes, R. A., Dolan, M. E. & Erickson, L. C. Prolonged depletion of O6-methylguanine DNA methyltransferase activity following exposure to O6-benzylguanine with or without streptozotocin enhances 1,3-bis(2-chloroethyl)-1-nitrosourea sensitivity in vitro. *Cancer Res.* **53**, 4281–4286 (1993).
7. Marathi, U. K., Dolan, M. E. & Erickson, L. C. Anti-neoplastic activity of sequenced administration of O6-benzylguanine, streptozotocin, and 1,3-bis(2-chloroethyl)-1-nitrosourea in vitro and in vivo. *Biochem. Pharmacol.* **48**, 2127–2134 (1994).
8. Fritz, G. & Kaina, B. Stress factors affecting expression of O6-methylguanine-DNA methyltransferase mRNA in rat hepatoma cells. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression* **1171**, 35–40 (1992).
9. Fritz, G., Tano, K., Mitra, S. & Kaina, B. Inducibility of the DNA repair gene encoding O6-methylguanine-DNA methyltransferase in mammalian cells by DNA-damaging treatments. *Mol. Cell. Biol.* **11**, 4660–4668 (1991).
10. Srivenugopal, K. S., Yuan, X. H., Friedman, H. S. & Ali-Osman, F. Ubiquitination-dependent proteolysis of O6-methylguanine-DNA methyltransferase in human and murine tumor cells following inactivation with O6-benzylguanine or 1,3-bis(2-chloroethyl)-1-nitrosourea. *Biochemistry* **35**, 1328–1334 (1996).
11. Eskelinen, E. L. & Saftig, P. Autophagy: a lysosomal degradation pathway with a central role in health and disease. *Biochimica et Biophysica Acta (BBA)*
12. Ohsumi, Y. Historical landmarks of autophagy research. *Cell Res.* **24**, 9–23 (2014).
13. Efeyan, A., Comb, W. C. & Sabatini, D. M. Nutrient-sensing mechanisms and pathways. *Nature* **517**, 302–310 (2015).
14. Kaur, J. & Debnath, J. Autophagy at the crossroads of catabolism and anabolism. *Nat. Rev. Mol. Cell Biol.* **16**, 461–472 (2015).
15. Havens, C. G. & Walter, J. C. Mechanism of CRL4(Cdt2), a PCNA-dependent E3 ubiquitin ligase. *Genes Dev.* **25**, 1568–1582 (2011).
16. Higa, L. A. *et al.* L2DTL/CDT2 interacts with the CUL4/DDB1 complex and PCNA and regulates CDT1 proteolysis in response to DNA damage. *Cell Cycle* **5**, 1675–1680 (2006).
17. Luijsterburg, M. S. *et al.* Dynamic in vivo interaction of DDB2 E3 ubiquitin ligase with UV-damaged DNA is independent of damage-recognition protein XPC. *J. Cell Sci.* **120**, 2706–2716 (2007).
18. Mu, H., Geacintov, N. E., Broyde, S., Yeo, J.-E. & Schärer, O. D. Molecular basis for damage recognition and verification by XPC-RAD23B and TFIIH in nucleotide excision repair. *DNA Repair (Amst.)* **71**, 33–42 (2018).
19. Shell, S. M. *et al.* Xeroderma pigmentosum complementation group C protein (XPC) serves as a general sensor of damaged DNA. *DNA Repair (Amst.)* **12**, 947–953 (2013).
20. Cavadini, S. *et al.* Cullin-RING ubiquitin E3 ligase regulation by the COP9 signalosome. *Nature* **531**, 598–603 (2016).
21. Fischer, E. S. *et al.* The molecular basis of CRL4DDB2/CSA ubiquitin ligase architecture, targeting, and activation. *Cell* **147**, 1024–1039 (2011).
22. Groisman, R. *et al.* The ubiquitin ligase activity in the DDB2 and CSA complexes is differentially regulated by the COP9 signalosome in response to DNA damage. *Cell* **113**, 357–367 (2003).
23. Guerrero-Santoro, J. *et al.* The cullin 4B-based UV-damaged DNA-binding protein ligase binds to UV-damaged chromatin and ubiquitinates histone H2A. *Cancer Res.* **68**, 5014–5022 (2008).

Chapter 4

24. Kapetanaki, M. G. *et al.* The DDB1-CUL4ADDB2 ubiquitin ligase is deficient in xeroderma pigmentosum group E and targets histone H2A at UV-damaged DNA sites. *Proc. Natl. Acad. Sci. USA* **103**, 2588–2593 (2006).
25. Lan, L. *et al.* Monoubiquitinated histone H2A destabilizes photolesion-containing nucleosomes with concomitant release of UV-damaged DNA-binding protein E3 ligase. *J. Biol. Chem.* **287**, 12036–12049 (2012).
26. Jørgensen, S. *et al.* SET8 is degraded via PCNA-coupled CRL4(CDT2) ubiquitylation in S phase and after UV irradiation. *J. Cell Biol.* **192**, 43–54 (2011).
27. Oda, H. *et al.* Regulation of the histone H4 monomethylase PR-Set7 by CRL4(Cdt2)-mediated PCNA-dependent degradation during DNA damage. *Mol. Cell* **40**, 364–376 (2010).
28. Soucy, T. A. *et al.* An inhibitor of NEDD8-activating enzyme as a new approach to treat cancer. *Nature* **458**, 732–736 (2009).
29. Fernández-Suárez, M., Chen, T. S. & Ting, A. Y. Protein-protein interaction detection in vitro and in cells by proximity biotinylation. *J. Am. Chem. Soc.* **130**, 9251–9253 (2008).
30. Kim, D. I. & Roux, K. J. Filling the Void: Proximity-Based Labeling of Proteins in Living Cells. *Trends Cell Biol.* **26**, 804–817 (2016).
31. Michalski, A. *et al.* Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer. *Mol. Cell Proteomics* **10**, M111.011015 (2011).
32. Eliuk, S. & Makarov, A. Evolution of orbitrap mass spectrometry instrumentation. *Annu. Rev. Anal. Chem. (Palo Alto, Calif.)* **8**, 61–80 (2015).
33. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).
34. Rees, J. S., Li, X.-W., Perrett, S., Lilley, K. S. & Jackson, A. P. Protein neighbors and proximity proteomics. *Mol. Cell Proteomics* **14**, 2848–2856 (2015).
35. Chen, C.-L. & Perrimon, N. Proximity-dependent labeling methods for proteomic profiling in living cells. *Wiley Interdiscip Rev Dev Biol* **6**, (2017).
36. Li, P., Li, J., Wang, L. & Di, L.-J. Proximity labeling of interacting proteins: application of bioID as a discovery tool. *Proteomics* **17**, (2017).
37. Choi-Rhee, E., Schulman, H. & Cronan, J. E. Promiscuous protein biotinylation by Escherichia coli biotin protein ligase. *Protein Sci.* **13**, 3043–3050 (2004).
38. Roux, K. J., Kim, D. I., Raida, M. & Burke, B. A promiscuous biotin ligase fusion protein identifies proximal and interacting proteins in mammalian cells. *J. Cell Biol.* **196**, 801–810 (2012).
39. Kim, D. I. *et al.* An improved smaller biotin ligase for BioID proximity labeling. *Mol. Biol. Cell* **27**, 1188–1196 (2016).
40. Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–887 (2018).
41. Rivella, S. & Sadelain, M. Genetic treatment of severe hemoglobinopathies: the combat against transgene variegation and transgene silencing. *Semin Hematol* **35**, 112–125 (1998).
42. Bestor, T. H. Gene silencing as a threat to the success of gene therapy. *J. Clin. Invest.* **105**, 409–411 (2000).
43. Kotin, R. M., Linden, R. M. & Berns, K. I. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. *EMBO J.* **11**, 5071–5078 (1992).

44. Tan, I., Ng, C. H., Lim, L. & Leung, T. Phosphorylation of a novel myosin binding subunit of protein phosphatase 1 reveals a conserved mechanism in the regulation of actin cytoskeleton. *J. Biol. Chem.* **276**, 21209–21216 (2001).
45. DeKolver, R. C. *et al.* Functional genomics, proteomics, and regulatory DNA analysis in isogenic settings using zinc finger nuclease-driven transgenesis into a safe harbor locus in the human genome. *Genome Res.* **20**, 1133–1142 (2010).
46. Mali, P. *et al.* RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826 (2013).
47. Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M. T. Rapid Protein Depletion in Human Cells by Auxin-Inducible Degron Tagging with Short Homology Donors. *Cell Rep.* **15**, 210–218 (2016).
48. Setten, R. L., Rossi, J. J. & Han, S.-P. The current state and future directions of RNAi-based therapeutics. *Nat. Rev. Drug Discov.* **18**, 421–446 (2019).
49. Nellen, W. & Lichtenstein, C. What makes an mRNA anti-sense? *Trends Biochem. Sci.* **18**, 419–423 (1993).
50. Fire, A. *et al.* Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811 (1998).
51. Hsu, S.-H., Chen, S.-H., Kuo, C.-C. & Chang, J.-Y. Ubiquitin-conjugating enzyme E2 B regulates the ubiquitination of O6-methylguanine-DNA methyltransferase and BCNU sensitivity in human nasopharyngeal carcinoma cells. *Biochem. Pharmacol.* **158**, 327–338 (2018).
52. Huang, J. *et al.* RAD18 transmits DNA damage signalling to elicit homologous recombination repair. *Nat. Cell Biol.* **11**, 592–603 (2009).
53. Kobayashi, S. *et al.* Rad18 and Rnf8 facilitate homologous recombination by two distinct mechanisms, promoting Rad51 focus formation and suppressing the toxic effect of nonhomologous end joining. *Oncogene* **34**, 4403–4411 (2015).
54. Szüts, D., Simpson, L. J., Kabani, S., Yamazoe, M. & Sale, J. E. Role for RAD18 in homologous recombination in DT40 cells. *Mol. Cell. Biol.* **26**, 8032–8041 (2006).
55. Ulrich, H. D. Two-way communications between ubiquitin-like modifiers and DNA. *Nat. Struct. Mol. Biol.* **21**, 317–324 (2014).
56. Helchowski, C. M., Skow, L. F., Roberts, K. H., Chute, C. L. & Canman, C. E. A small ubiquitin binding domain inhibits ubiquitin-dependent protein recruitment to DNA repair foci. *Cell Cycle* **12**, 3749–3758 (2013).
57. Xu, M. Degradation of the alkylated form of the DNA repair protein, O6 -alkylguanine-DNA alkyltransferase. *Welliver*

Conclusions

The present study of MGMT has afforded numerous insights into the unique biochemistry of the protein and its turnover in the cell. The substrate scope of MGMT was shown to extend beyond duplex DNA, demonstrating activity on single stranded DNA and small molecules with high molecular weight (Chapter 2). It was revealed for the first time that MGMT can be used as a degron-tag in human cells for chemically-driven targeted protein degradation (Chapter 3). Finally, a detailed mechanistic study of the MGMT biosynthetic pathway showed that protein degradation is reversible and titratable. Several E3 ligases have been detected as potential interacting partners that facilitate MGMT degradation, showing the complexity of the ubiquitin-mediated signaling pathway. Among these, data points towards the Cullin-RING family of E3 ligases as the main interacting partner (Chapter 4). These findings have important implications in areas including cellular protein homeostasis, pharmacologic chaperones and targeted protein degradation.

Experimental Procedures

Table of Contents

Chapter 2 “Substrate Scope”	73
1 Fluorescent-based activity assay:.....	73
2 DNA modification procedure	76
3 pET19b 6xHis-MGMT plasmid generation	77
4 6xHis-MGMT purification	77
5 Chemical synthesis	78
5.1 Synthesis of GM-1.....	79
5.2 Synthesis of GM-2.....	82
Chapter 3 “MGMT with bifunctional molecules”	85
6 Cell culture.....	87
7 Immunoblotting.....	87
8 Generation of plasmids	87
9 Generation of cell pools expressing TurboID-13x-MGMT	88
10 Genomic DNA extraction.....	89
11 Immunoprecipitation	89
12 TMT labelling and LC MS/MS analysis	89
13 Chemical synthesis	90
13.1 Synthesis of MGMT- Bcr-Abl substrate: GM-3.	91
Synthesis of GM-3B’	92
Synthesis of GM-3B.....	93
Synthesis of GM-3C.....	93
Synthesis of GM-3	94
13.2 Synthesis of O ⁶ -BG analogues: GM-4.	96
13.3 Synthesis of BCR-Abl substrates: GM-5.	99
Synthesis of GM-5A’.....	99

Synthesis of GM-5	100
13.4 Synthesis of MGMT-BRD4 substrates: GM-6, GM-7, GM-8.	102
Synthesis of GM-6A	102
Synthesis of GM-6	103
Synthesis of GM-7	104
Synthesis of GM-8	106
13.5 Synthesis of covalent probes: GM-9, GM-10, GM-11	107
Synthesis of GM-9	107
Synthesis of GM-10	108
Synthesis of GM-11	109
Chapter 4 " Mechanistic Study"	109
14 Cell culture.....	111
15 Immunoblotting.....	111
16 Genomic DNA extraction.....	111
17 Generation of plasmids	112
18 Generation of cell pools expressing TurboID-13x-MGMT in AAVS1 locus	112
18.1 Amplified region:	113
18.2 Sequenced data:	114
19 Inhibitory study	114
20 MGMT recovery and chase study.....	115
21 TurboID Pulldowns.....	115
22 TMT labelling and LC MS/MS analysis	116
23 shRNA sequence	118

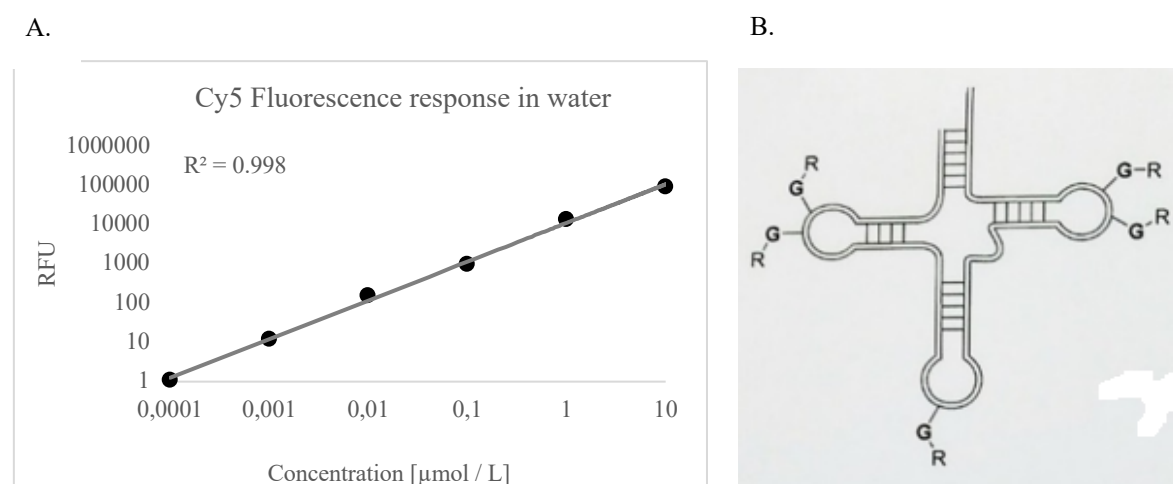
Chapter 2

Substrate scope of the repair protein MGMT

1 Fluorescent-based activity assay:

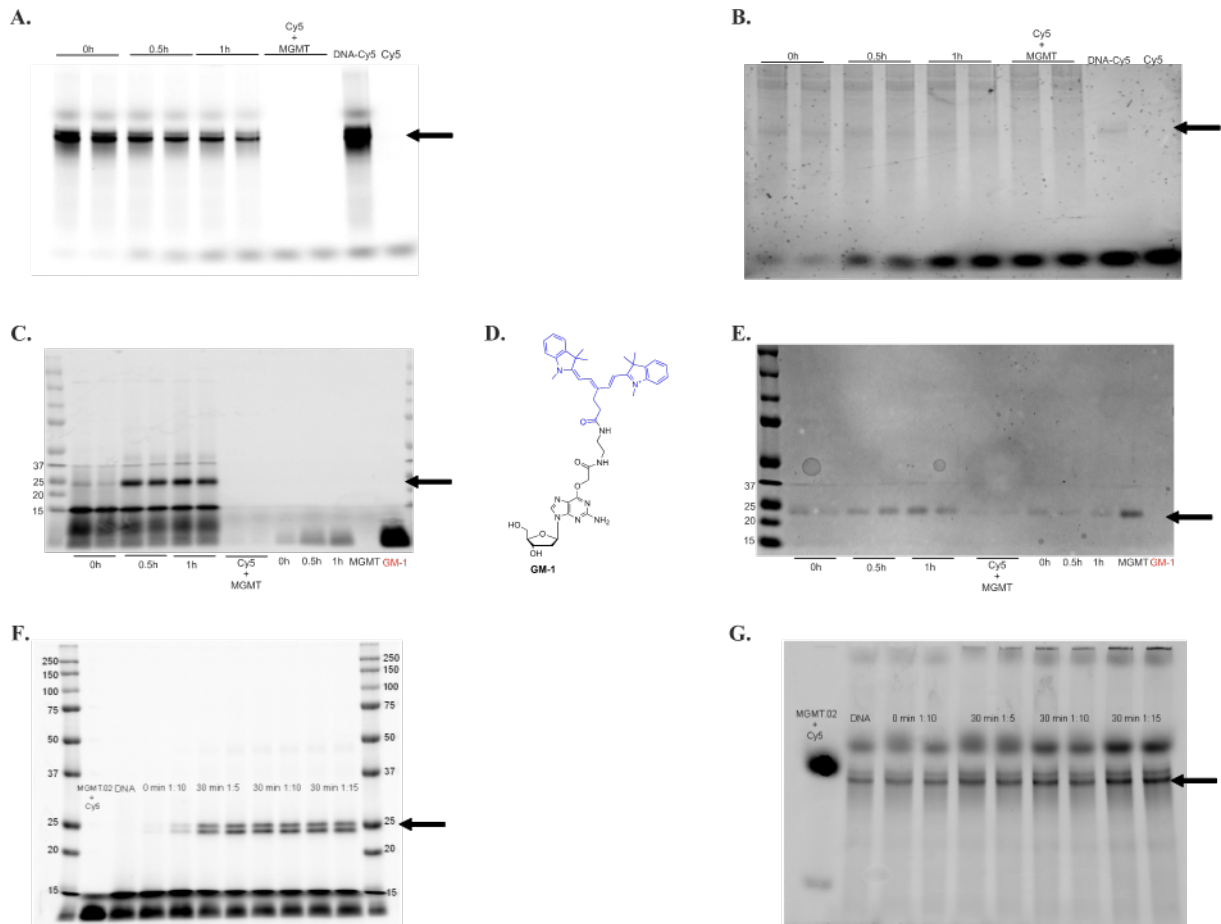
Recombinant MGMT: Sf9 (Abcam, ab136378) (1 μg) or *E. coli* MGMT (3 μg) were used in reaction buffer (50 mM TRIS-HCl pH 7.5, 0.1 mM EDTA pH 7.5, 5% glycerol, 1 mM DTT). Fluorescent MGMT substrates were added with fluorescence normalized to 370 RFU and the reaction mixture was incubated at 37°C for 1 hour. Gels were analyzed by BioRAD Universal Hood III under Cy5-filter 695/55. Instrument sensitivity towards Cy5 had been evaluated by series of dilutions in water (Figure S1A). Immunoblotting had been performed using nitrocellulose membrane (Amersham Protran 0.45 NC) using BioRAD Trans-Blot Turbo transfer system. Blots were blocked in Odyssey Blocking Buffer in TBS for 1h at room temperature, washed three times for 5 minutes with TBST (TBS with 0.1% Tween 20) and incubated with the primary antibody overnight at 4°C. After incubation blots were washed three times for 5 minutes with TBST and incubated with secondary antibody (Li-COR: IRDye 680RD Goat anti-Rabbit IgG or IRDye 800CW Goat anti-Mouse IgG, diluted 1:10000 in TBST) for 1 hour at room temperature. After secondary incubation blots were washed three times for 5 minutes with TBST and imaged with Odyssey CLx Imaging System. Primary antibodies used: Anti-MGMT (Abcam, ab80513, diluted as recommended to 1 $\mu\text{g}/\text{mL}$ in TBST). Anti-Actin (Santa Cruz Biotechnology). Cy5-fluorescent bands (Figure S3A) were evaluated in Image Lab software and MGMT bands (Figure S3C) were evaluated using Image Studio Lite software. Analyzed intensities are presented on Figure S4 and Table TS1.

Sequence of tDNA-6G mimicking tyrosine tRNA from *E. coli* where thymine was used instead of uracil. Purchased from Microsynth : GgtgggggtcccagcggccaaggagcagactgtaaatctgccGtcatcGacttcGaaGGttcGaatcctccccaccacca (Figure S1B)



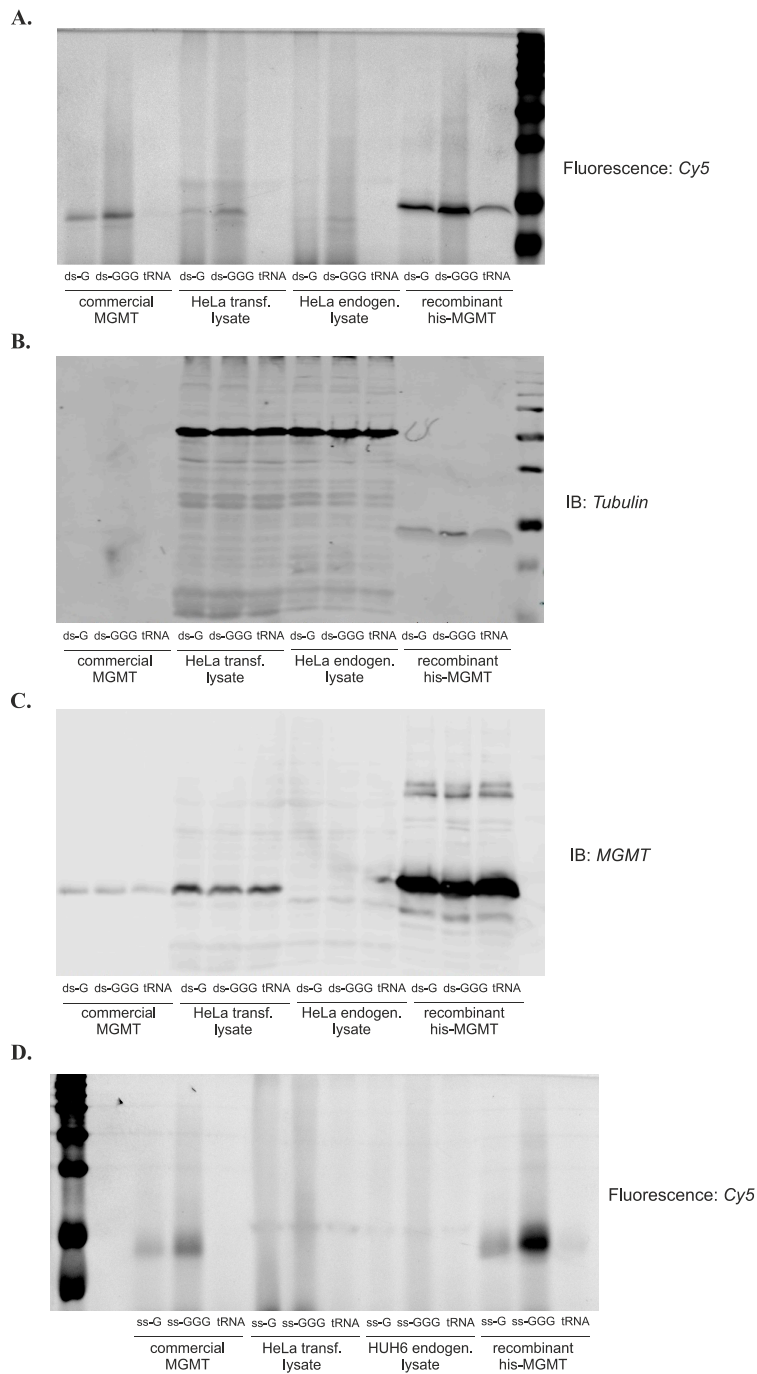
S 1. BioRAD Universal Hood Instrument sensitivity evaluation. Measured Cy5 absorption in water using Cy5 filter 695/55 nm (A). Schematic representation of tRNA modification on guanines (B).

Chapter 2



S 2. MGMT in vitro activity in PAGE gel fluorescent assay. 15% PAGE, 7 M urea DNA gel under Cy5-channel of fluorescence reader (A), and the same gel stained with SYBR gold for DNA detection (B). 18% SDS-PAGE protein gel detected with the Cy5-channel (C); the same gel stained with Coomassie Blue for protein detection (E). GM-1 structure (D). MGMT titration with increasing amount of fluorescent DNA, 18% SDS-PAGE protein gel detected with the Cy5-channel (F) and 15% PAGE, 7 M urea DNA gel under Cy5-channel of fluorescence reader (G). Arrows pointing at bands corresponding to expected product.

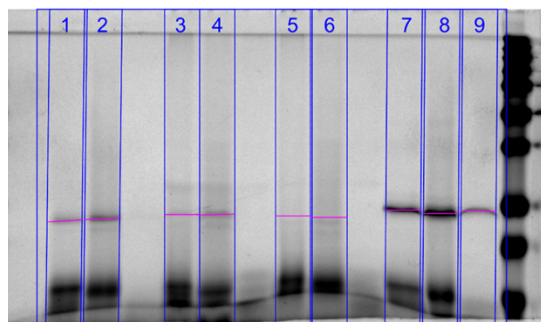
Experimental Procedures



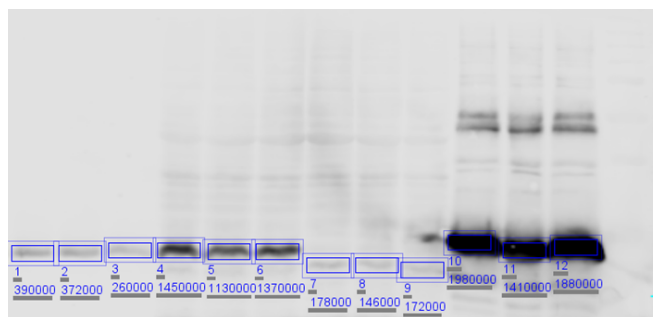
S 3. Different sources of MGMT reacting with dsDNA, ssDNA and tRNA substrates: MGMT expressed in insect cells, HeLa cell lysate, HUH cell lysate or MGMT (E. coli) with Cy5-modified O⁶-guanine on tRNA, double strand DNA with one or three modifications. 15% SDS-PAGE gel imaged under fluorescence reader: 695/55nm (A, top panel); The same gel after immunoblotting stained with anti-tubulin confirming equal loading of HeLa cell lysate (A, middle panel) the same blot stained with Anti-MGMT antibody to indicate amount of MGMT protein in each sample (A, bottom panel); Similarly, the same assay was performed with single-strand DNA. 18% SDS-PAGE gel under fluorescence reader: 695/55nm (B). dsG: double strand DNA with one modified guanine at O⁶-position with Cy5. dsGGG: double strand DNA with three modified guanines at O⁶-position with Cy5. tRNA-6G: tRNA with one modified guanine at O⁶-position with Cy5. ssG: single strand DNA with one modified guanine at O⁶-position with Cy5; ssGGG: single strand DNA with three modified guanines at O⁶-position with Cy5.

Chapter 2

A



B



S 4. Protein activity evaluation. Fluorescence measurement in Image Lab software (A) and MGMT level quantified in Image Studio Lite software (B).

Cy5 channel	Lane	G	GGG	tDNA	Lane	G	GGG	tDNA
	MGMT sf9	2247552	3871800		E.coli / Sf9	2.8	3.0	
	Hela transf	2422656	2549736		E.coli / Hela transf	2.6	4.5	
	Hela endog.	926136	901368		E.coli / Hela endog.	6.8	12.9	
	MGMT E coli	6258888	11593368	3559968				

WB (MGMT)	Lane	G	GGG	tDNA	Lane	G	GGG	tDNA
	MGMT sf9	390000	372000	260000	E.coli / Sf9	5.1	3.8	7.2
	Hela transf	1450000	1130000	1370000	E.coli / Hela transf	1.4	1.2	1.4
	Hela endog.	178000	146000	172000	E.coli / Hela endog.	11.1	9.7	10.9
	MGMT E coli	1980000	1410000	1880000				

TS 1. Fluorescence and MGMT quantification obtained from S4

2 DNA modification procedure

Adopted from *Chem. Commun.*, 2018,54, 9174-9177

General procedure: The MES buffer was placed in a 500 μ L tube, the DNA, CuSO₄ and the N-(but-3-yn-1-yl)-2-diazoacetamide were added and the reaction was mixed. The sodium ascorbate was added, and the reaction was mixed again. After 3 hours the reaction mixture was precipitated with ethanol, supernatant removed and pellet washed twice with 70% EtOH, air dried and reconstituted in H₂O.

40n GGG oligo: ttttttGtttttGtttttGtttttcttccccaccacca

40n 12TG11T oligo: ttttttttttGtttttttttcttccccaccacca

Alkylation Cu1 eq: MES (100 mM), DNA oligo (140 μ M), CuSO₄ (140 μ M), sodium ascorbate (0.7 mM) and N-(but-3-yn-1-yl)-2-diazoacetamide (17.5 mM).

Alkylation Cu10 eq: Cu1 eq: MES (100 mM), DNA oligo (140 μ M), CuSO₄ (1.4 mM), sodium ascorbate (7 mM) and N-(but-3-yn-1-yl)-2-diazoacetamide (17.5 mM).

Cy-5 fylation: oligo DNA (40 μ M), Cy5-N₃ (60 μ M), sodium ascorbate (200 μ M), CuSO₄ (200 μ M).

Experimental Procedures

All the component of the reaction (without CuSO₄) were mixed and degassed under nitrogen for 1 minute. The CuSO₄ was added and the reaction was incubated for 2 hours at room temperature. DNA was precipitated with ethanol, washed twice with 70% ethanol and purified using Nucleospin Gel and PCR Purification kit (Macherey-Nagel®) with buffer for ssDNA.

Modification efficiency was evaluated by PCR reaction (Figure S 2A) using HiDi DNA Polymerase (obtained from MyPols Biotch) in 6 extension cycles.

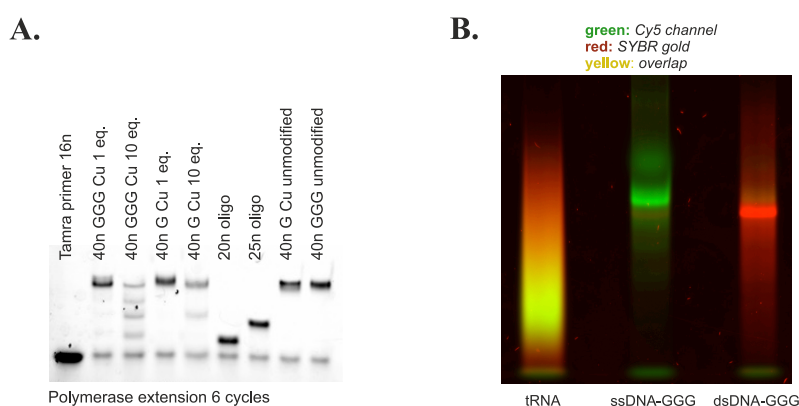
Primers used:

20n oligo: tttccttccccaccacca

25n oligo: tttGtttccttccccaccacca

16n Tamra primer: 5'-TAMRA-tggtggtggggaagg.

Additionally, 8% urea-PAGE gel was used to evaluate Cy5 modification of tRNA, ssDNA-GGG and dsDNA-GGG modification (Figure S 2B).



S 5. Agarose gel of PCR analysis DNA modification efficiency (A); 8% urea-PAGE gel showing Cy5 labelling efficiency or tRNA, ssDNA-GGG and dsDNA-GGG (B).

3 pET19b 6xHis-MGMT plasmid generation

RNA was extracted from HeLa cells (kind gift from Maier – Palivan-Meier laboratory, University of Basel) using TRI reagent (Invitrogen) following manufacture instructions. cDNA was prepared from extracted RNA using reverse transcriptase SuperScript III (Invitrogen) and MGMT coding sequence (obtained from NCBI database: <https://www.ncbi.nlm.nih.gov/nucore/M29971.1>) was amplified with primers:

1 with NheI restriction site: 5'- taacgctagcATGGACAAGGATTGTGAA-3'

2 with XhoI restriction site: 5'- taactcgagTCAGTTTCGGCCAGCAGG-3'

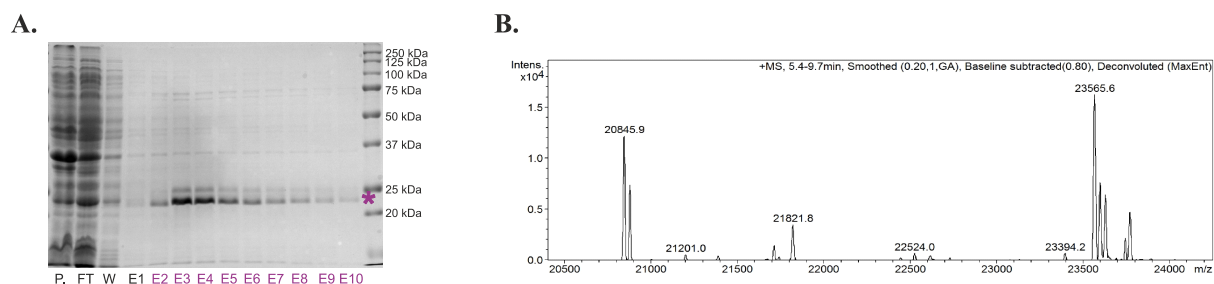
And cloned into pET19b plasmid backbone using XhoI and NdeI restriction sites.

4 6xHis-MGMT purification

250 mL of bacterial cell culture expressing 6xHis-MGMT was spun down at 4°C, 17568 x g for 20 minutes, supernatant was discarded, and pellet transferred to 50 mL Falcon Tube. Cell pellet was re-suspended in 20 mL of bacterial lysis buffer (20mM Tris-HCl pH 8, 200 mM NaCl, 10% glycerol) and sonicated three times for 4 minutes cycle (duty cycle 50%, output control 4). Lysed cells were centrifuged for 40 minutes at 4°C at 3724 x

Chapter 2

g and supernatant was transferred to new Falcon tube where 150 μ L of Ni-beads (Qiagen Ni-NTA beads) were added. Remaining pellet was dissolved in 1mL of 8M urea and Beads were rotated at 4°C for 1 hour and loaded on a column. Flow-through (FT), wash (with 10 mL of bacterial lysis buffer with 20 mM imidazole) and ten elutions (with 200 μ L of lysis buffer with 250mM imidazole) were collected and equal volumes were loaded on 12% SDS-PAGE gel (Figure S 3A). Combined fractions containing His-MGMT (E2-E10) were dialyzed in ZelluTrans dialysis bags (ROTH, Nominal molecular weight 6000 – 8000 Da) in 50 mM TRIS pH 7.5, 50 mM NaCl at 4°C for two days. Purified sample was evaluated by LC/MS-TOF performed on a Agilent 1100 Series instrument coupled to a micrOTOF ESI-TOF from Bruker, using a Jupiter C4 300A column (50 x 2 mm, 5 Microns, Phenomenex) (Figure S 3B).



S 6. His-MGMT purification. 12% SDS-PAGE gel with P-pellet, FT – Flow Through, E – Elusion (A). Pink star indicated MGMT size 23.5 kDa; HRMS spectra from combined fractions E2-E10. $[M+1]$ Expected 23.5 kDa. Found: 23.5651 kDa

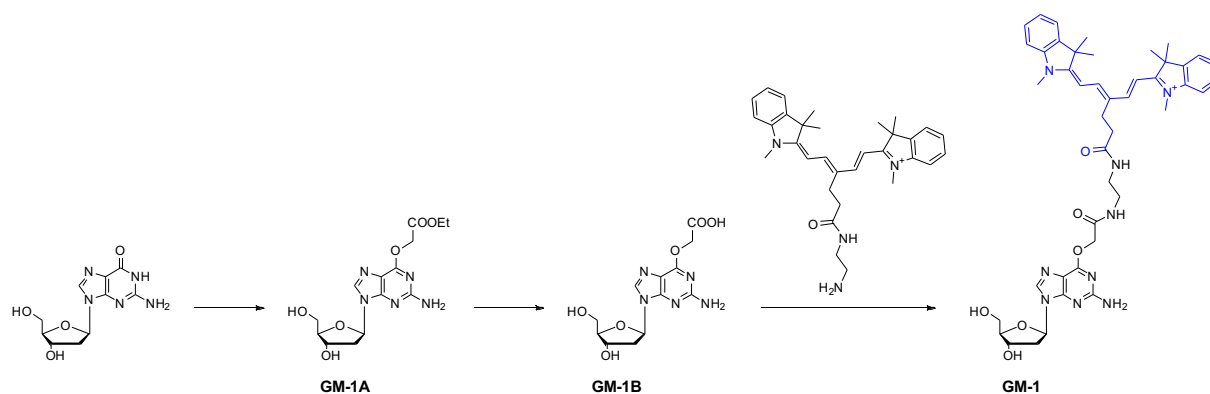
5 Chemical synthesis

^1H NMR, ^{13}C NMR and HMBC NMR spectra were recorded on BrukerAvance (400, 500 or 600 MHz proton frequency) spectrometer at 298.15 K. ^1H and $^{13}\text{C}\{^1\text{H}\}$ NMR chemical shifts (δ) are reported in ppm relative to TMS (0.00 ppm), with the solvent resonance used as internal reference, and coupling constants (J) are in Hertz (Hz). The following abbreviations were used to explain the multiplicities: s = singlet, d = doublet, t = triplet, q = quartet, p = pentad, m = multiplet, br = broad. The mass spectrometric data were obtained at the mass spectrometry facility of the University of Basel

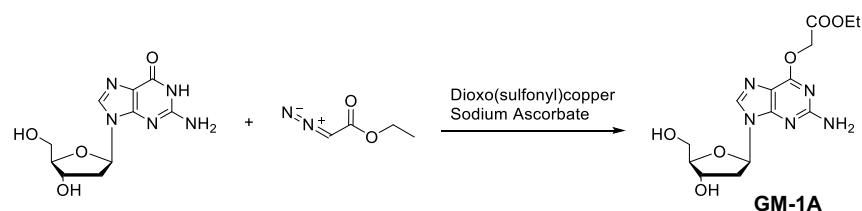
UPLC-MS : Agilent Technologies 1290 Infinity, Column Aligent EclipsePlusC18 RRHD 1.8 μm 2.1 x 50 mm. Solvents: Buffer A: water + 0.1 % Formic Acid, Buffer B: Acetonitrile + 0.1% Formic acid. Gradient: 0 - 1 minute at 2% of Buffer B, 1 – 3.5 minutes with linear gradient to 60% of Buffer B, 3.5 - 4.5 minutes with linear gradient to 90% Buffer B held for 0.5 minute and 0.5 minute equilibration at 2% of Buffer B.

RP-HPLC: UFLC Shimadzu; Column: Gemini® 5 μm NX-C18, LC column size 110 g 250 x 21.2 mm AX. Solvent system: Buffer A: Water, Buffer B: Acetonitrile; Gradient: 0 – 5 minutes at 1% of solvent B, 5 – 7 minutes with linear gradient to 7% of Buffer B, 7 – 27 minutes with linear gradient to 70% of Buffer B, 27 – 31 minutes with linear gradient to 99% of buffer B held for 1 minute, flow rate: 20 mL/minute.

5.1 Synthesis of GM-1

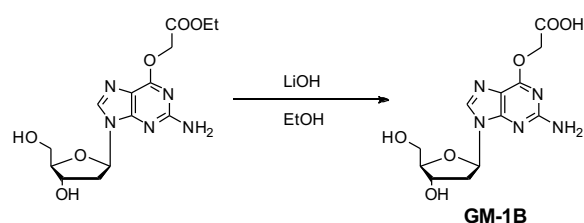


Synthesis of GM-1A



To a solution of $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ (70 mg) in water (235 mL) was added MES buffer (0.5 M, pH adjusted to 6 with concentrated NaOH, 56.1 mL). The reaction mixture was degassed, after which a solution of deoxyguanosine monohydrate (400 mg, 1.4 mmol, 1 eq.) in DMSO (1 mL), ethyl diazoacetate (1.7 mL, 14 mmol, 10 eq.) and sodium ascorbate (278 mg, 1.4 mmol, 1 eq.) were added. The reaction mixture was stirred at room temperature overnight. The reaction mixture was purified with RP-chromatography giving 231 mg of white solid (44% yield). ¹H NMR (500 MHz, MeOD) δ 8.14 (s, 1H), 6.29 (t, $J = 6.9$ Hz, 1H), 5.01 (s, 2H), 4.52 (dt, $J = 5.7, 2.6$ Hz, 1H), 4.15 (q, $J = 7.1$ Hz, 2H), 4.00 (q, $J = 2.9$ Hz, 1H), 3.79 (dd, $J = 12.2, 3.0$ Hz, 1H), 3.69 (dd, $J = 12.2, 3.4$ Hz, 1H), 2.78 – 2.68 (m, 1H), 2.32 (ddd, $J = 13.6, 6.0, 2.5$ Hz, 1H), 1.19 (t, $J = 7.1$ Hz, 3H).

Synthesis of GM-1B

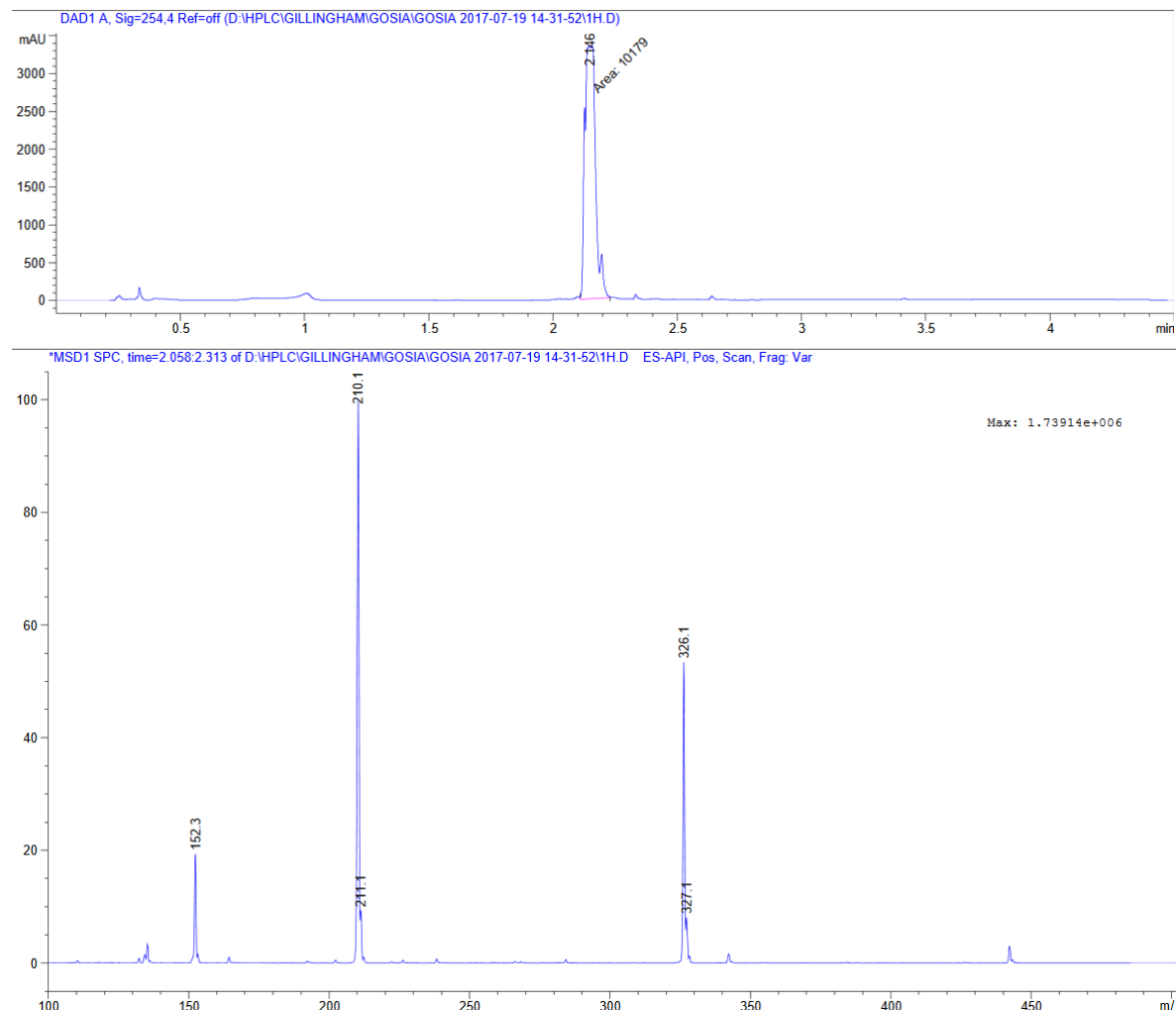


GM-1A (50 mg, 0.14 mmol, 1 eq) was dissolved in ethanol and LiOH (1.4 mL, 0.28 mmol, 2 eq) was added. Reaction was stirred at room temperature for 1 hour. Solvent was evaporated and product used in next step without further purification.

Chapter 2

^1H NMR (400 MHz, MeOD) δ 7.90 (s, 1H), 6.21 (d, $J = 6.9$ Hz, 1H), 4.70 (s, 2H), 4.50 – 4.41 (m, 1H), 3.94 (q, $J = 2.9$ Hz, 1H), 3.74 (dd, $J = 12.2, 3.0$ Hz, 1H), 3.63 (dd, $J = 12.3, 3.3$ Hz, 1H), 2.75 – 2.63 (m, 1H), 2.23 (ddd, $J = 13.5, 6.1, 2.6$ Hz, 1H).

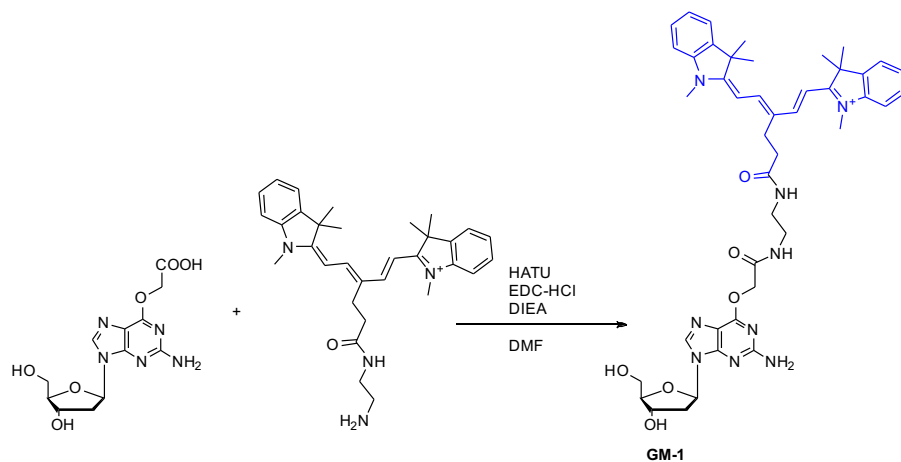
MS for $\text{C}_{12}\text{H}_{15}\text{N}_5\text{O}_6$ expected $[\text{M}+1]$: 326.10 m/z. Found: 326.1 m/z



S 7. Upper panel: UPLC chromatogram at 254 nm, indicating formation of the product at $rt = 2.146$ min. Lower panel: MS of the same peak confirming product formation $[\text{M}+1]$: 326.1 m/z.

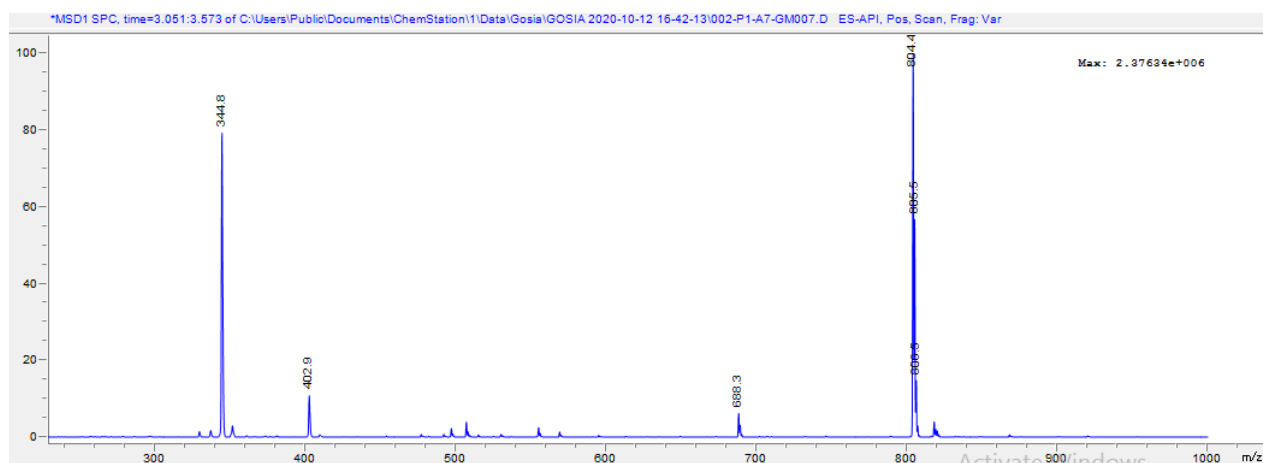
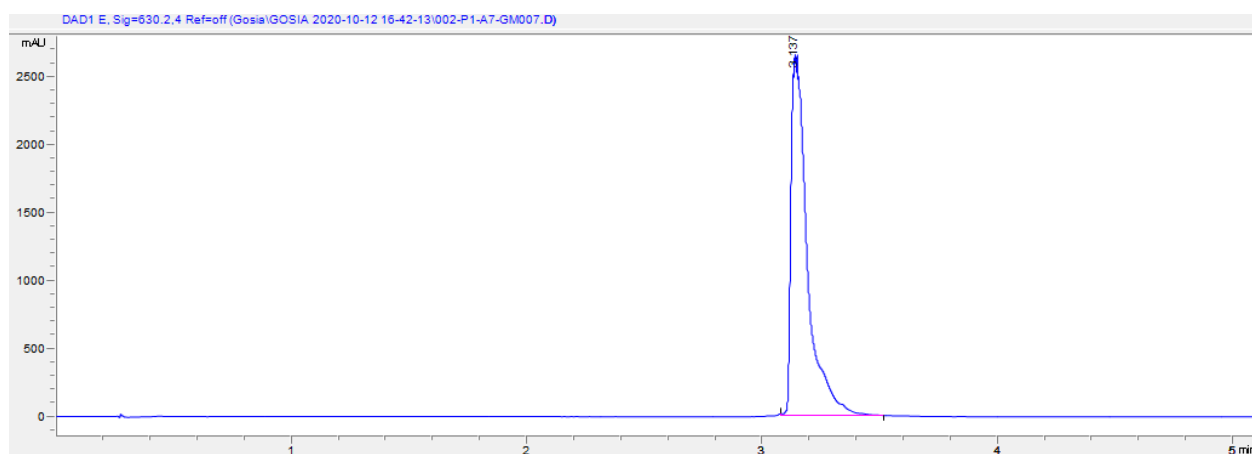
Experimental Procedures

Synthesis of GM-1.



GM-1B, DIPEA, EDC-HCl and HATU were mixed together with DMF until the solution was clear. Cy5-NH₂ was added and the reaction was stirred at room temperature for 1 hour. Solvent had been removed, remaining residue dissolved in DMSO and purified on RP-HPLC. The reaction was monitored at 630 nm.

MS for C₄₄H₅₄N₉O₆⁺ expected [M+1]: 805.42 m/z. Found: 805.5 m/z



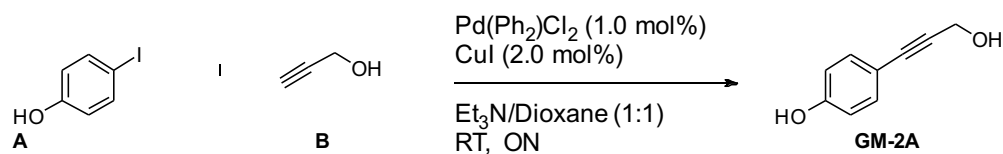
S 8. Upper panel: UPLC chromatogram at 254 nm, indicating formation of the product at *rt* = 2.146 min. Lower panel: MS of the same peak confirming product formation.

Chapter 2

5.2 Synthesis of GM-2.

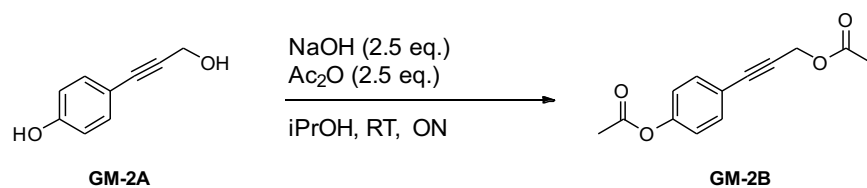
GM-2 was synthesized according to a published protocol [Angew. Chem. Int. Ed. 2016, 55, 2911–2915] (AA-CW538).

Synthesis of GM-2A:



A 25 mL Schlenk-flask was nitrogen fumed and evacuated *in vacuo* three times. The flask was equipped with a magnetic stir bar and 4-iodophenol (**A**, 825 mg, 3.75 mmol, 1.0 eq.) was diluted in dry triethylamine/dioxane (5 mL, 1/1). Subsequently, propargyl alcohol (**B**, 262 μL , 4.50 mmol, 1.2 eq.), $\text{Pd}(\text{PPh}_3)_2\text{Cl}_2$ (26.3 mg, 0.037 mmol, 1.0 mol%) and copper iodide (14.3 mg, 0.075 mmol, 2.0 mol%) were added. The dark yellow reaction mixture was stirred overnight at room temperature. The reaction mixture was purified by column chromatography (silica plug with $\text{EtOAc}/\text{Cyclohexane}$: 2/3). The product (**GM-2A**, 144 mg, 26%) was obtained as a white-yellow solid. $^1\text{H NMR}$ (400 MHz, MeOD) δ 7.23 – 7.21 (m, 2H), 6.71 – 6.69 (m, 2H), 4.34 (s, 2H).

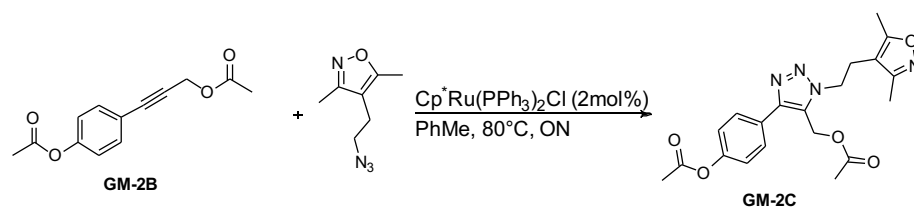
Synthesis of GM-2B:



The starting material (**GM-2A**, 125 mg, 0.840 mmol, 1.0 eq.) was added to a 25 mL round-bottom flask and dissolved in isopropanol. Sodium hydroxide (1 mL, 2M, 2.4 eq.) was added and stirred for 5 minutes. Diacetyl ether (300 μL , 3.15 mmol, 3.75 eq.) was added by syringe, whereby the reaction was activated by heating it for 15 min at 40 °C and then the solution was stirred overnight at room temperature. The reaction mixture was extracted with CH_2Cl_2 and purified by column chromatography ($\text{EtOAc}/\text{Cyclohexane}$: 1:2). The product (**GM-2B**, 49.0 mg, 25%) was obtained as a yellow oil.

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.47 – 7.45 (m, 2H), 7.06 – 7.04 (m, 2H), 4.89 (s, 2H), 2.30 (s, 3H), 2.13 (s, 3H).

Synthesis of GM-2C:



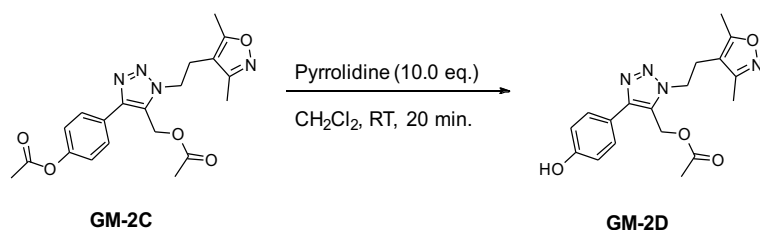
The reaction flask was equipped with a stirring bar and closed with a septum. It was filled with nitrogen and evacuated three times with vacuum. Dry toluene (3.0 mL) and **GM-2B** (42.0 mg, 0.180 mmol, 1.2 eq.) were added as well as the azide (14 μL , 0.150 mmol, 1.0 eq.) and $\text{Cp}^*\text{Ru}(\text{PPh}_3)_2\text{Cl}$ (2.40 mg, 0.003 mmol, 2.0 mol%). The

Experimental Procedures

reaction mixture was stirred overnight at 80°C in an oil bath. The resulting solution was concentrated, and the oily mixture was purified by column chromatography (gradient 0 - 65% EtOAc/Cyclohexane). The product (**GM-2C**, 17.0 mg, 28%) was obtained as a yellow oil.

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.76 – 7.74 (m, 2H), 7.21 – 7.19 (m, 2H), 4.96 (s, 2H), 4.54 (t, $J = 6.7$ Hz, 2H), 3.00 (t, $J = 6.7$ Hz, 2H), 2.32 (s, 3H), 2.14 (s, 3H), 2.08 (s, 3H), 2.06 (s, 3H).

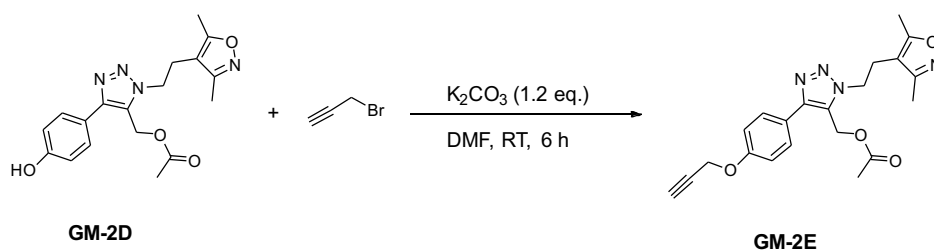
Synthesis of GM-2D:



To a nitrogen fumed 25 mL round-bottom flask was added **GM-2C** (35.0 mg, 0.090 mmol, 1.0 eq.) and dissolved in CH_2Cl_2 (3.0 mL). Then, pyrrolidine (77 μL , 0.920 mmol, 10 eq.) was added and stirred for 1 hour. The reaction was quenched with HCl (0.5 mL, 1M). The product (**GM-2D**, 14.0 mg, 44%) was obtained after purification by column chromatography (EtOAc/Cyclohexane, gradient 0 - 65%) as a yellow solid.

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.54 (d, $J = 8.5$ Hz, 2H), 6.95 (d, $J = 8.5$ Hz, 2H), 4.95 (s, 2H), 4.54 (t, $J = 6.7$ Hz, 2H), 3.00 (t, $J = 6.7$ Hz, 2H), 2.14 (s, 3H), 2.09 (s, 3H), 2.08 (s, 3H).

Synthesis of GM-2E:

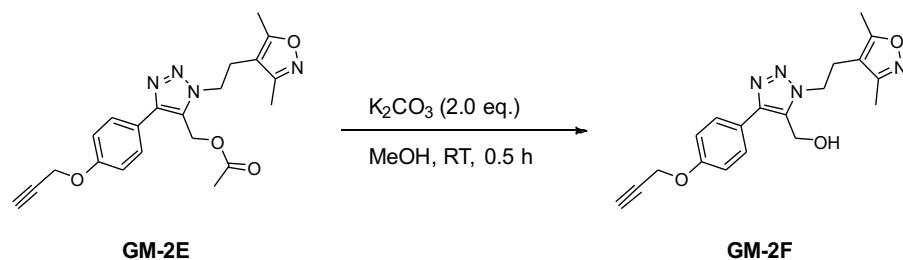


GM-2D (14.0 mg, 0.040 mmol, 1.0 eq.) was dissolved in dry DMF (1.0 mL) and potassium carbonate (7.00 mg, 0.051 mmol, 1.2 eq.) as well as propargyl bromide (4.4 μL , 80 wt.% in PhMe, 0.051 mmol, 1.2 eq.) were added. The reaction mixture was stirred overnight at room temperature, diluted in heptane and concentrated *in vacuo*. The product (**GM-2E**, 10.0 mg, 67%) was obtained as a yellow oil after column chromatography (EtOAc/Cyclohexane, gradient 0 - 60%).

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.67 – 7.64 (m, 2H), 7.09 – 7.07 (m, 2H), 4.94 (s, 2H), 4.74 (d, $J = 2.4$ Hz, 2H), 4.54 (t, $J = 6.7$ Hz, 2H), 2.55 (t, $J = 2.4$ Hz, 1H), 2.15 (s, 3H), 2.09 (s, 3H), 2.07 (s, 3H).

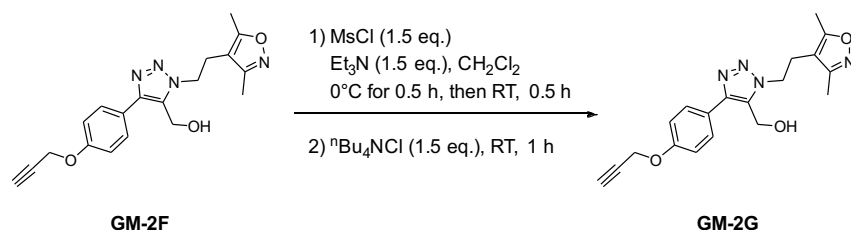
Chapter 2

Synthesis of GM-2F:



The starting material **GM-2E** (10.0 mg, 0.025 mmol, 1.0 eq.) was stirred for an hour in methanol (0.5 mL) after addition of potassium carbonate (10.0 mg, 0.070 mmol, 2.8 eq.). After completion of the reaction, which was controlled by UPLC-MS, the reaction mixture was filtered and concentrated. The remaining liquid was used directly for the following last step.

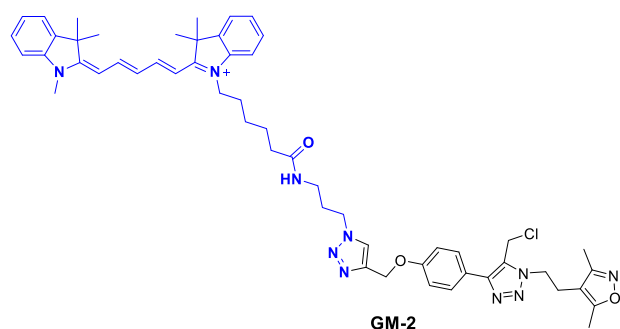
Synthesis of GM-2G:



To a nitrogen-fumed round-bottom flask which was cooled in an ice bath mesyl chloride (4.4 μL , 0.057 mmol, 1.5 eq.), trimethylamine (7.9 μL , 0.057 mmol, 1.5 eq.) and **GM-2F** were dissolved in dry CH_2Cl_2 (2.0 mL). The reaction mixture was stirred for 30 minutes at 0°C , then the ice bath was removed and the reaction was stirred for another 30 minutes. *N*-tetrabutyl ammonium chloride (15.8 mg, 0.057 mmol, 1.5 eq.) was added and the reaction was stirred for 1 hour. The solution was concentrated, and the residue was purified by RP-HPLC ($\text{H}_2\text{O} + 0.1\%$ TFA/ACN + 0.1% TFA, gradient 0 - 80%). The solvent was removed using a lyophilizer. The product (**GM-2G**, 3.00 mg, 34%) was obtained as white powder.

$^1\text{H NMR}$ (400 MHz, CDCl_3) δ 7.64 – 7.62 (m, 2H), 7.12 – 7.09 (m, 2H), 4.75 (d, $J = 2.4$ Hz, 2H), 4.49 (s, 2H), 4.44 (t, $J = 7.1$ Hz, 2H), 3.08 (t, $J = 7.1$ Hz, 2H), 2.55 (t, $J = 2.4$ Hz, 1H), 2.22 (s, 3H), 2.17 (s, 3H).

Synthesis of GM-2:

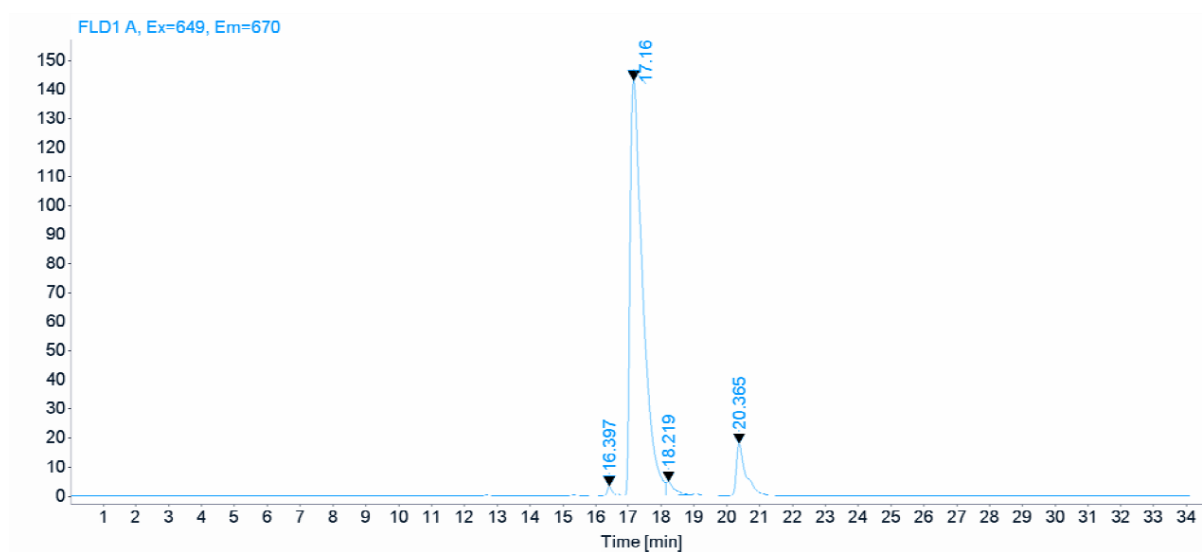


GM-2G (60 μM , 6 nmol, 1 eq.), Cy5-azide (Lumiprobe) (90 μM , 9 nmol, 1.5 eq.) and CuSO_4 (300 μM , 30 nmol, 5 eq.) were mixed and degassed with nitrogen for 1 minute. Sodium ascorbate was added (420 μM , 42 nmol, 7

Experimental Procedures

eq.), flushed once with nitrogen gas and the reaction mixture was incubated for 3 hours at room temperature with stirring. Crude product was purified on HPLC monitored at 670 nm, eluted product was at 20.365 minutes.

MS for $C_{54}H_{64}ClN_{10}O_3^+$ expected $[M+1]$: 935.48 m/z. Found: 1094 m/z



S 9. HPLC chromatogram at emission wavelength 670 nm showing obtained product: $rt = 17.16$ min and free Cy5: $rt = 20.36$ min.

Chapter 2

Chapter 3

MGMT probes with bifunctional molecules

6 Cell culture

HUH6 were gifts from the Hall laboratory and maintained at 37°C with 5% CO₂ in RPMI containing 10% FCS and 1% penicillin–streptomycin.

HEK293T ACC635 cells were obtained from the DSMZ and maintained at 37°C with 5% CO₂ in DMEM containing 10% FCS and 1% penicillin–streptomycin.

7 Immunoblotting

Cells were harvested, washed three times with PBS and lysed in freshly prepared RIPA buffer (50 mM Tris pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplied with 1x Complete Mini Protease Inhibitor Cocktail (Roche) and 10mM sodium orthovanadate. Protein concentration was quantified using Bradford Assay Reagents A, B and S (BioRAD #5000113, #5000114 and #5000115, respectively) measuring A_{750 nm} on TECAN SPARK® reader. 25 µg of protein lysate was developed on 8% SDS-PAGE gel and proteins were transferred to a nitrocellulose membrane (Amersham Protran 0.45 NC) using BioRAD Trans-Blot Turbo transfer system. Blots were blocked in Odyssey Blocking Buffer in TBS for 1h at room temperature, washed three times for 5 minutes with TBST (TBS with 0.1% Tween 20) and incubated with the primary antibody overnight at 4°C. After incubation blots were washed three times for 5 minutes with TBST and incubated with secondary antibody (Li-COR: IRDye 680RD Goat anti-Rabbit IgG or IRDye 800CW Goat anti-Mouse IgG, diluted 1:10000 in TBST) for 1h at room temperature. After secondary incubation blots were washed three times for 5 minutes with TBST and imaged with Odyssey CLx Imaging System.

Antibodies used in this study were as follow: Anti-MGMT (Abcam, ab80513, diluted as recommended to 1 µg/mL in TBST). Anti-Actin (Santa Cruz Biotechnology). Anti-BRD4 (Abcam, ab128874, working concentration 1:1000 in TBST buffer). Anti-c-ABL (Cell Signalling Technology, 2862T, working concentration 1:1000 in TBST buffer)

8 Generation of plasmids

The pMyc-BioID2-13xLinker-MCS vector was obtained from Addgene (#92308) and used as a backbone to which hMGMT sequence was cloned at the end of 13xLinker using *XhoI* and *KpnI* restriction site. Resulting vector was named pBioID2-13x-hMGMT. Plvx-TurboID-v5-Ires-puro (a kind gift from Davide Eletto, Hale laboratory, UZH Medicinal Virology, Addgene #) was used as a backbone to insert 13xlinker-hMGMT sequence (*EcoRI* /*PmeI* fragment of pBioID2-13xlinker-hMGMT) using *EcoRI* and *BamHI* restriction sites. pFLuc-13x-hMGMT was created by replacing myc-BioID2 sequence in the pBioID2-13x-hMGMT vector (*NheI* and *EcoRI* restriction sites) by FLuc sequence. GFP-MGMT fusion had been obtained by cloning cDNA hMGMT into commercial vector pEGFP-C1 (Clontech) using *Sall* and *KpnI* restriction sites.

Chapter 3

pTet-TurboID-13x-MGMT was generated by cloning TurboID-13x-MGMT sequence into the pMK243 vector (Tet-OsTIR1-puro, Addgene #72835) replacing OsTIR1 gene using HiFi DNA Assembly kit from New England Biolabs. Primers used:

5'-gatcctctagacatatgctgcagagatcttcagtttcggccagcaggc-3';

5'-gcctgctggccgaaactgaagatctctgcagcatatgtctagaggatc-3';

5'-gcacagtattgtctttgctagccatggtggcacgcgtgcg-3';

5'-cgcacgcgtgccaccatggctagcaagacaataactgtgcc-3'.

Sanger sequencing was provided by either Microsynth® facility in Basel, Switzerland or Genewiz® facility in Irvine, United States.

9 Generation of cell pools expressing TurboID-13x-MGMT

Tet-TurboID-MGMT and AAVS1 T2 CRIPR in pX330 DL (7 µg) (Addgene #72833) were used to transfect 293T cells. 10 µg of Tet-TurboID-MGMT and 7 µg of AAVS1 T2 CRIPR in pX330 DL were added to 1.5 mL of Opti-MEM media. 36 µL of TurboFect transfecting agent was added and the mixture was incubated for 15 minutes at room temperature and added dropwise to the target cells. After 24 hours the medium was changed and after 48 hours 1 µg/mL of puromycin was used for selection.

Survived population had been pulled together and validated by genomic DNA extraction and PCR amplification.

Primers used:

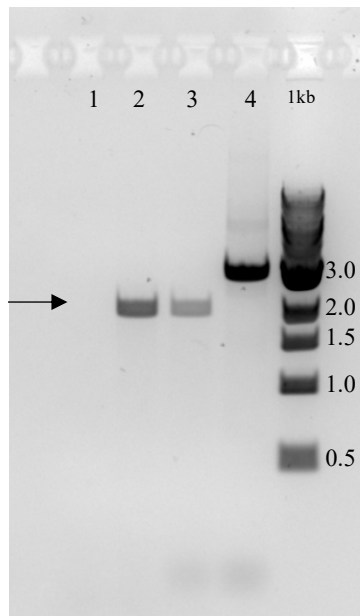
5'-gatcctctagacatatgctgcagagatcttcagtttcggccagcaggc-3';

5'-gcctgctggccgaaactgaagatctctgcagcatatgtctagaggatc-3';

5'-gcacagtattgtctttgctagccatggtggcacgcgtgcg-3';

5'-cgcacgcgtgccaccatggctagcaagacaataactgtgcc-3'.

Sanger sequencing was provided by either Microsynth® facility in Basel, Switzerland or Genewiz® facility in Irvine, United States



S 10. S 1. PCR products of amplified: 1 Genomic DNA WT; 2 DNA extracted from 293T ACC635 integrated in AAVS1 locus; 3 DNA extracted from 293T integrated in AAVS1 locus; 4 Tet-TurboID-MGMT plasmid used for transfection. Expected band size: 2020 bp.

10 Genomic DNA extraction

Cells had been suspended in serum containing media using trypsin and spun down for 5 minutes at 500 *x g*. Supernatant was discarded and cell pellet was washed three times with 10 mL of ice-cold PBS and centrifuged for 5 minutes at 500 *x g*. Cells were re-suspended in digestion buffer (100 mM NaCl, 10 mM TrisCl pH 8, 25 mM EDTA pH 8, 0.5% SDS and freshly added 0.1 mg/ml proteinase K. For $< 3 \times 10^7$ 0.3 ml digestion buffer was used) and incubated for 16 hours at 50°C with shaking. DNA was extracted by adding an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) and spinning down the mixture at 1700 *x g* for 10 minutes in swinging bucket. Top layer had been transferred to the new tube and 1/10 volume of 3M NaAc had been added. DNA immediately precipitated and was recovered by centrifugation at 1700 *x g* for 2 minutes, wash with 70% ethanol, air drying the pellet and dissolving it TE buffer.

11 Immunoprecipitation

50 μ L of DynabeadsTM Protein A beads (Invitrogen) were mixed with recombinant Anti-MGMT antibody [EPR4398] (ab108989) and incubated overnight at 4°C while rotating. Precleaned twice for 2 minutes with 1 mL of lysis buffer (20 mM Tris HCl pH 8, 137 mM NaCl, 1% Nonidet P-40 (NP-40), 2 mM EDTA), cell lysate was added and incubate ON at 4°C rotating. Washed three times 5 minutes with lysis buffer followed by addition of either GM-9 or GM-10, respectively and incubated at 37°C for 1 hour. Beads were washed twice with lysis buffer and eluted with 40 μ L of 2x SDS loading dye (125 mM Tris-HCl pH 6.8, 20% glycerol, 10% 2-mercaptoethanol, 4% SDS) and loaded on SDS-PAGE gel.

12 TMT labelling and LC MS/MS analysis

Sample aliquots comprising 25 μ g of peptides were labeled with isobaric tandem mass tags (TMT 10-plex, Thermo Fisher Scientific) as described previously (PMID:27345528). Shortly, peptides were resuspended in 20 μ L labeling buffer (2 M urea, 0.2 M HEPES, pH 8.3) and 5 μ L of each TMT reagent were added to the individual peptide samples followed by a 1 h incubation at 25°C, shaking at 500 rpm. To control for ratio distortion during quantification, a peptide calibration mixture consisting of six digested standard proteins mixed in different amounts was added to each sample before TMT labeling (for details see PMID:27345528). To quench the labelling reaction, 1.5 μ L aqueous 1.5 M hydroxylamine solution was added and samples were incubated for another 10 minutes at 25°C shaking at 500 rpm followed by pooling of all samples. The pH of the sample pool was increased to 11.9 by adding 1 M phosphate buffer (pH 12) and incubated for 20 min at 25°C shaking at 500 rpm to remove TMT labels linked to peptide hydroxyl groups. Subsequently, the reaction was stopped by adding 2 M hydrochloric acid until a pH < 2 was reached. Finally, peptide samples were further acidified using 5% TFA, desalted using Sep-Pak Vac 1cc (50 mg) C18 cartridges (Waters) according to the manufacturer's instructions and dried under vacuum.

TMT-labeled peptides were fractionated by high-pH reversed phase separation using a XBridge Peptide BEH C18 column (3,5 μ m, 130 Å, 1 mm x 150 mm, Waters) on an Agilent 1260 Infinity HPLC system. Peptides were loaded on column in buffer A (20 mM ammonium formate in water, pH 10) and eluted using a two-step linear gradient from 2% to 10% in 5 minutes and then to 50% buffer B (20 mM ammonium formate in 90% acetonitrile, pH 10) over 55 minutes at a flow rate of 42 μ L/min. Elution of peptides was monitored with a UV detector (215 nm, 254 nm) and a total of 36 fractions were collected, pooled into 12 fractions using a post-concatenation strategy as previously described (PMID:21500348) and dried under vacuum.

Chapter 3

Dried peptides were resuspended in 0.1% aqueous formic acid and subjected to LC–MS/MS analysis using a Orbitrap Fusion Lumos Mass Spectrometer fitted with an EASY-nLC 1200 (both Thermo Fisher Scientific) and a custom-made column heater set to 60°C. Peptides were resolved using a RP-HPLC column (75µm × 36cm) packed in-house with C18 resin (ReproSil-Pur C18–AQ, 1.9 µm resin; Dr. Maisch GmbH) at a flow rate of 0.2 µLmin⁻¹. The following gradient was used for peptide separation: from 5% B to 15% B over 9 min to 30% B over 90 min to 45 % B over 21 min to 95% B over 2 min followed by 18 min at 95% B. Buffer A was 0.1% formic acid in water and buffer B was 80% acetonitrile, 0.1% formic acid in water.

The mass spectrometer was operated in DDA mode with a cycle time of 3 seconds between master scans. Each master scan was acquired in the Orbitrap at a resolution of 120,000 FWHM (at 200 m/z) and a scan range from 375 to 1600 m/z followed by MS2 scans of the most intense precursors in the Orbitrap at a resolution of 30,000 FWHM (at 200 m/z) with isolation width of the quadrupole set to 1.1 m/z. Maximum ion injection time was set to 50ms (MS1) and 54 ms (MS2) with an AGC target set to 1e6 and 1e5, respectively. Only peptides with charge state 2 – 7 were included in the analysis. Monoisotopic precursor selection (MIPS) was set to Peptide, and the Intensity Threshold was set to 5e4. Peptides were fragmented by HCD (Higher-energy collisional dissociation) with collision energy set to 38%, and one microscan was acquired for each spectrum. The dynamic exclusion duration was set to 45s.

The acquired raw-files were converted to the mascot generic file (mgf) format using the msconvert tool (part of ProteoWizard, version 3.0.4624 (2013-6-3)) and searched using MASCOT against a human database (consisting of 40832 forward and reverse protein sequences downloaded from Uniprot on 20190307), the six calibration mix proteins (PMID:27345528) and 392 commonly observed contaminants. The precursor ion tolerance was set to 10 ppm and fragment ion tolerance was set to 0.02 Da. The search criteria were set as follows: full tryptic specificity was required (cleavage after lysine or arginine residues unless followed by proline), 3 missed cleavages were allowed, carbamidomethylation (C) and TMT6plex (K and peptide N-terminus) were set as fixed modification and oxidation (M) as a variable modification. Next, the database search results were imported into the Scaffold Q+ software (version 4.3.2, Proteome Software Inc., Portland, OR) and the protein false discovery rate was set to 1% based on the number of decoy hits. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. Acquired reporter ion intensities in the experiments were employed for automated quantification and statistical analysis using a modified version of our in-house developed SafeQuant R script (v2.3, PMID:27345528). This analysis included adjustment of reporter ion intensities, global data normalization by equalizing the total reporter ion intensity across all channels, summation of reporter ion intensities per protein and channel, calculation of protein abundance ratios and testing for differential abundance using empirical Bayes moderated t-statistics. Finally, the calculated p-values were corrected for multiple testing using the Benjamini–Hochberg method.

13 Chemical synthesis

¹H NMR, ¹³C NMR spectra were recorded on BrukerAvance (400, 500 or 600 MHz proton frequency) spectrometer at 298.15 K. ¹H and ¹³C{¹H} NMR chemical shifts (δ) are reported in ppm relative to TMS (0.00 ppm), with the solvent resonance used as internal reference, and coupling constants (J) are in Hertz (Hz). The following abbreviations were used to explain the multiplicities: s = singlet, d = doublet, t = triplet, q =

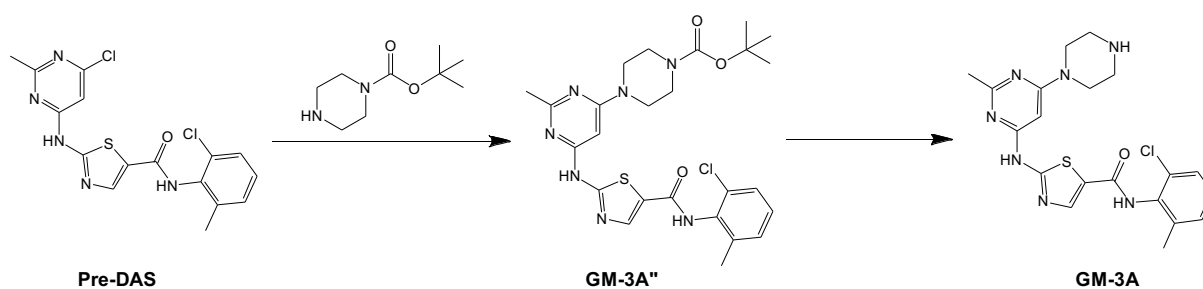
Experimental Procedures

quartet, m = multiplet, br = broad. The mass spectrometric data were obtained at the mass spectrometry facility of the University of Basel

UPLC-MS: Agilent Technologies 1290 Infinity, Column Aligent EclipsePlusC18 RRHD 1.8 μm 2.1 x 50 mm. Solvents: Buffer A: water + 0.1 % Formic Acid, Buffer B: Acetonitrile + 0.1% Formic acid. Gradient: 0 - 1 minute at 2% of Buffer B, 1 – 3.5 minutes with linear gradient to 60% of Buffer B, 3.5 - 4.5 minutes with linear gradient to 90% Buffer B held for 0.5 minute and 0.5 minute equilibration at 2% of Buffer B.

RP-HPLC: UFLC Shimadzu; Column: Gemini® 5 μm NX-C18, LC column size 110 g 250 x 21.2 mm AX. Solvent system: Buffer A: Water, Buffer B: Acetonitrile; Gradient: 0 – 5 minutes at 1% of solvent B, 5 – 7 minutes with linear gradient to 7% of Buffer B, 7 – 27 minutes with linear gradient to 70% of Buffer B, 27 – 31 minutes with linear gradient to 99% of buffer B held for 1 minute, flow rate: 20 mL/minute.

13.1 Synthesis of MGMT- Bcr-Abl substrate: GM-3.

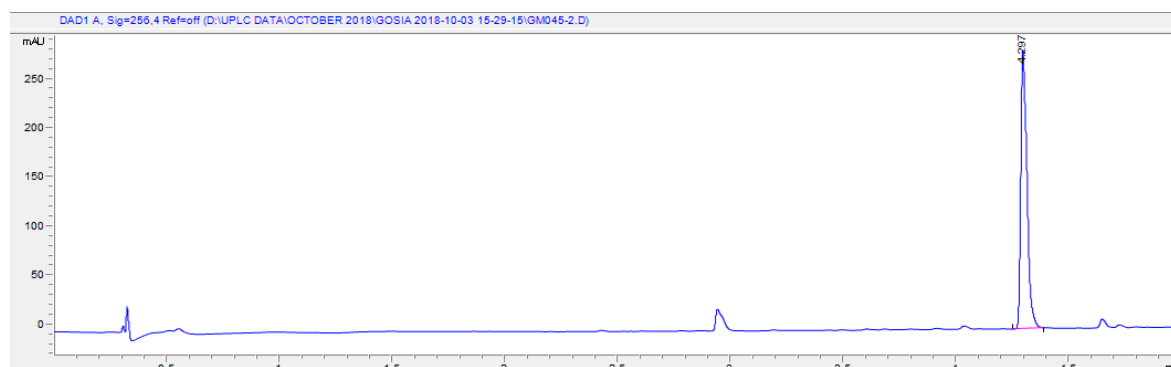


Synthesis of GM-3A''

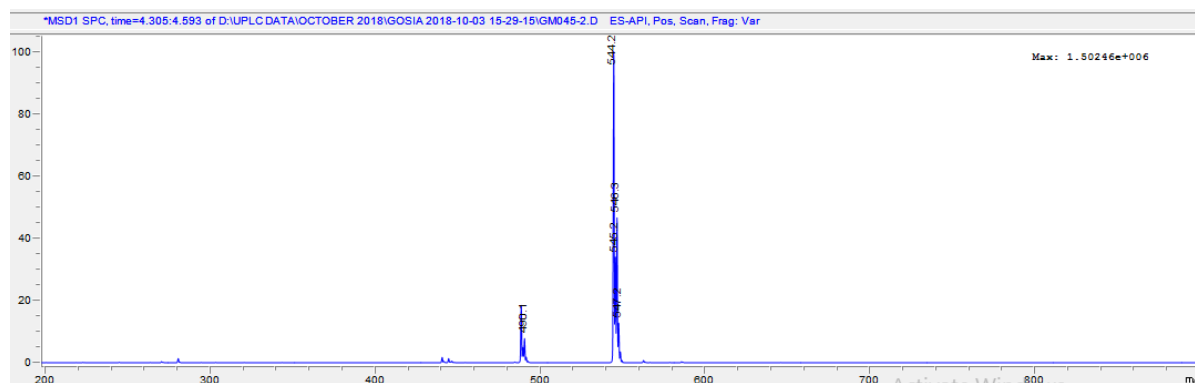
N-(2-chloro-6-methyl-phenyl)-2-[(6-chloro-2-methyl-pyrimidin-4-yl)amino]thiazole-5-carboxamide (200 mg, 1 eq.) was mixed with tert-Butyl 1-piperazinecarboxylate (151 mg, 1.1 eq.) and DIEA (500 μL , 4 eq.) and stirred for 16 hours at 100°C. Reaction mixture was concentrated under reduced pressure and purified on RP-HPLC giving white solid (276 mg, 21%).

$^1\text{H NMR}$ (400 MHz, DMSO- d_6) δ 11.51 (s, 1H), 9.88 (s, 1H), 8.22 (s, 1H), 7.40 (dd, $J = 7.5, 2.0$ Hz, 1H), 7.31 – 7.21 (m, 2H), 6.06 (s, 1H), 3.40 (d, $J = 18.8$ Hz, 8H), 2.42 (s, 3H), 2.24 (s, 3H), 1.42 (s, 9H).

MS for $\text{C}_{25}\text{H}_{30}\text{ClN}_7\text{O}_3\text{S}$ expected $[\text{M}+1]$: 544.2 m/z. Found: 544.2 m/z



Chapter 3



S 11. Upper panel: UPLC chromatogram at 256 nm, indicating formation of the product at $rt = 2.297$ min. Lower panel: MS of the same peak confirming product formation $[M+1]: 544.2$ m/z.

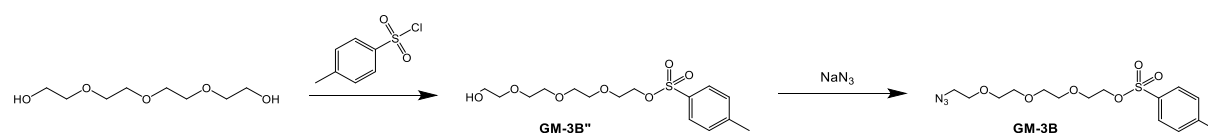
Synthesis of GM-3A

GM-3A (73 mg, 1 eq.) was dissolved in 1 mL of CH_2Cl_2 , cool down on ice to 0°C and 1 mL of TFA was added dropwise. The reaction was allowed to warm up to room temperature and stirred for 1 hour. The solvents had been removed under the reduced pressure and the product used for next step without further purification.

$^1\text{H NMR}$ (500 MHz, MeOD) δ 8.11 (s, 1H), 7.27 (dd, $J = 7.4, 1.7$ Hz, 1H), 7.20 – 7.08 (m, 2H), 6.25 (s, 1H), 3.88 (t, $J = 5.13$, 4H), 3.25 (t, $J = 5.13$, 4H), 2.44 (s, 3H), 2.24 (s, 3H).

$^{13}\text{C NMR}$ (126 MHz, MeOD) δ 164.12, 163.79, 162.00, 161.48, 161.20, 161.08, 160.91, 160.62, 155.77, 138.90, 132.74, 132.72, 128.81, 128.31, 126.96, 125.80, 83.40, 42.70, 41.07, 23.07, 17.35.

MS for $\text{C}_{20}\text{H}_{22}\text{ClN}_7\text{OS}$ expected $[M+1]: 443.13$ m/z. Found: 443.1



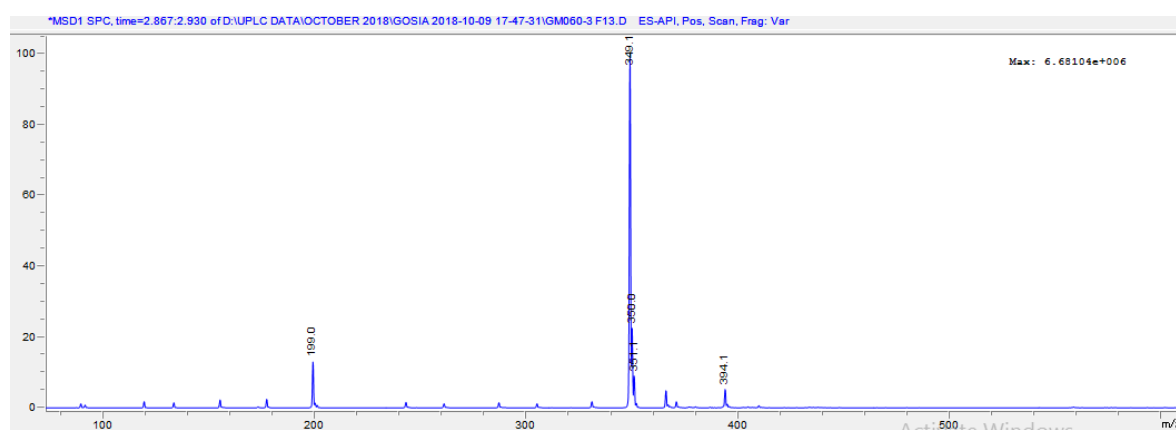
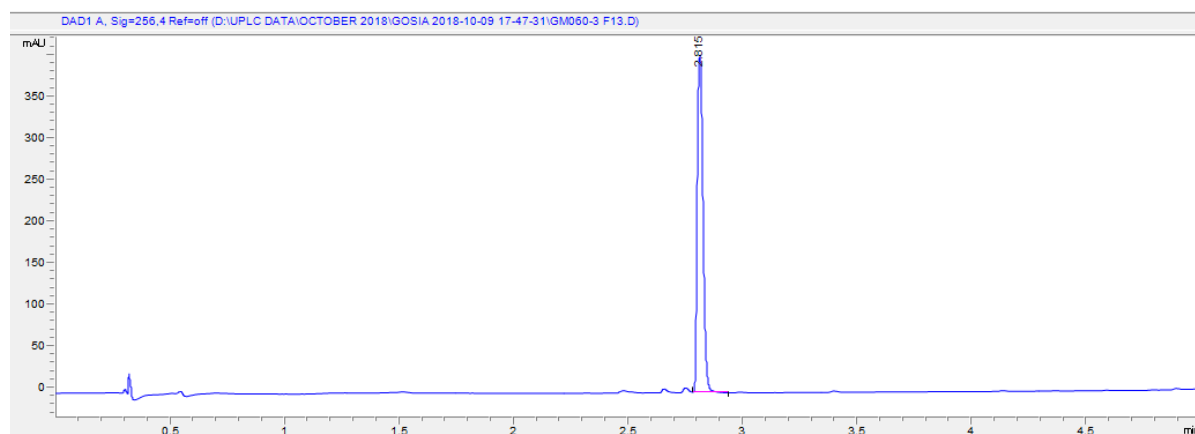
Synthesis of GM-3B" (Tos-PEG3-OH)

2-[2-[2-(2-hydroxyethoxy)ethoxy]ethoxy]ethanol (1 g, 1 eq.) and N,N-diethylethanamine (2.2 mL, 3 eq.) were dissolved in CH_2Cl_2 (5 mL) and cool to 0°C on ice bath. 4-methylbenzenesulfonyl chloride (1.96 mg, 2 eq.) was added slowly and the reaction mixture was allowed to warm up to room temperature by removing the ice bath and stirred for 20 hours. Solvents were removed under reduced pressure and product was purified on RP-HPLC giving 1.79 g of product (72.5% yield)

$^1\text{H NMR}$ (400 MHz, DMSO- d_6) δ 7.79 (d, $J = 8.3$ Hz, 2H), 7.48 (d, $J = 8.0$ Hz, 2H), 4.11 (s, 2H), 3.65 – 3.55 (m, 2H), 3.52 – 3.36 (m, 12H), 2.42 (s, 3H).

MS for $\text{C}_{15}\text{H}_{24}\text{O}_7\text{S}$ expected $[M+1]: 349.1$ m/z. Found: 349.1 m/z

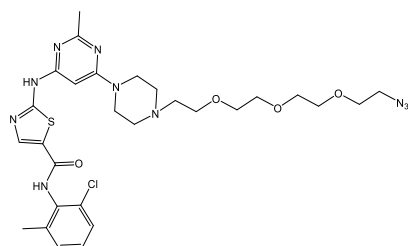
Experimental Procedures



S 12. Upper panel: UPLC chromatogram at 256 nm, indicating formation of the product at $rt = 2.815$ min. Lower panel: MS of the same peak confirming product formation $[M+1]$: 349.1 m/z .

Synthesis of GM-3B

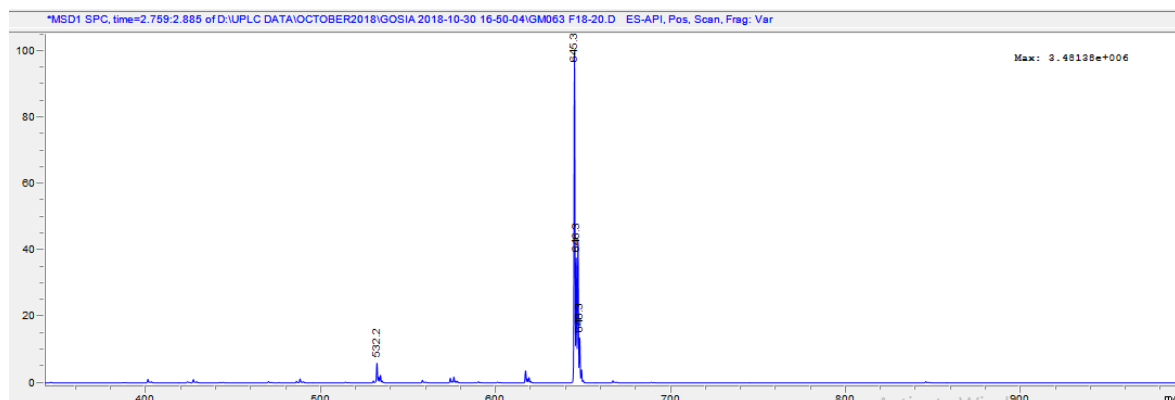
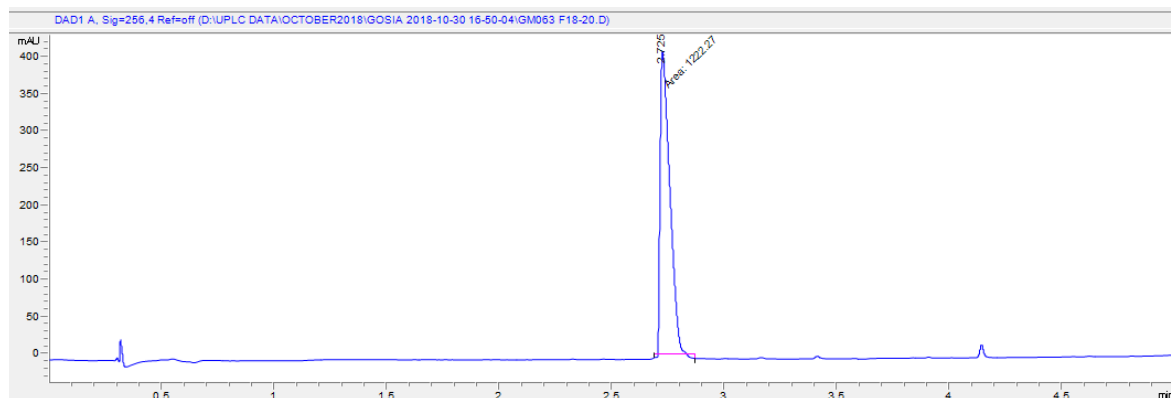
GM-3B (1.3 g, 1 eq.) was dissolved in ACN (5 mL) and NaN_3 was added (364 mg, 1.5 eq.) and the reaction was refluxed at 60°C for 20 hours. Solvent was removed under the reduced pressure and dissolved in H_2O to dissolve NaN_3 and extracted three times with CH_2Cl_2 . CH_2Cl_2 layer was dried over MgSO_4 , filtered and concentrated under reduced pressure. The product (**GM-3B**, 785 mg, 63% yield, transparent liquid) was used for next step without further purification.



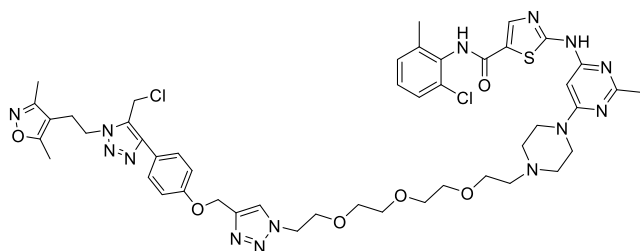
Synthesis of GM-3C (Pre-DAS-PEG3-N₃)

Chapter 3

GM-3A (20 mg, 1 eq.) was added to **GM-3B** (18.5 mg, 1.1 eq.) and K_2CO_3 (12 mg, 2 eq.) and dissolved in acetone (5 mL). NaI was added (1 mg, 0.15 eq.) and the reaction mixture was refluxed for 20 hours. Solvents were removed under the reduced pressure and the crude product purified on RP-HPLC giving 21 mg of white solid (72% yield). 1H NMR (400 MHz, DMSO- d_6) δ 11.65 (s, 1H), 10.40 (s, 1H), 9.91 (s, 1H), 7.39 (d, $J = 7.5, 2.1$ Hz, 1H), 7.27 (dd, 2H), 6.18 (s, 1H), 4.34 (s, 2H), 3.81 (s, 2H), 3.71 – 2.97 (m, 20H), 2.45 (s, 3H), 2.25 (s, 3H). MS for $C_{28}H_{37}ClN_{10}O_4S$ expected $[M+1]$: 645.24 m/z. Found: 645.3 m/z



S 13. Upper panel: UPLC chromatogram at 256 nm, indicating formation of the product at $rt = 2.725$ min. Lower panel: MS of the same peak confirming product formation $[M+1]$: 645.23 m/z.

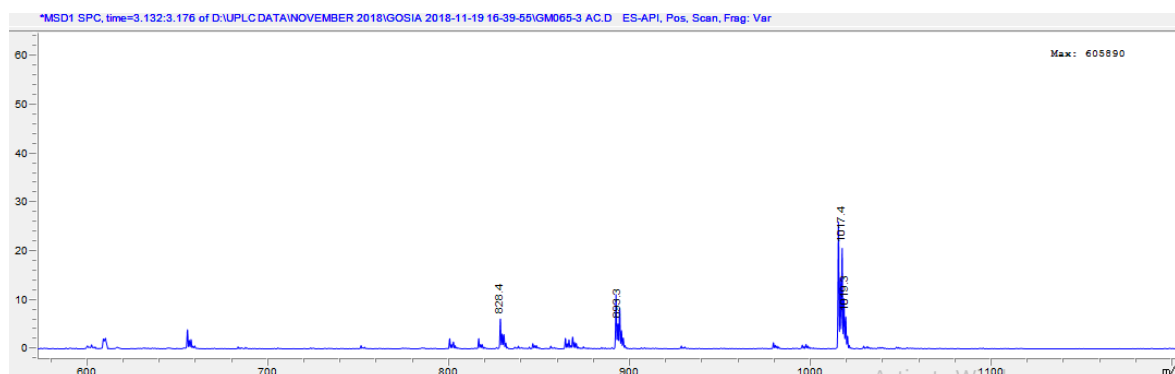
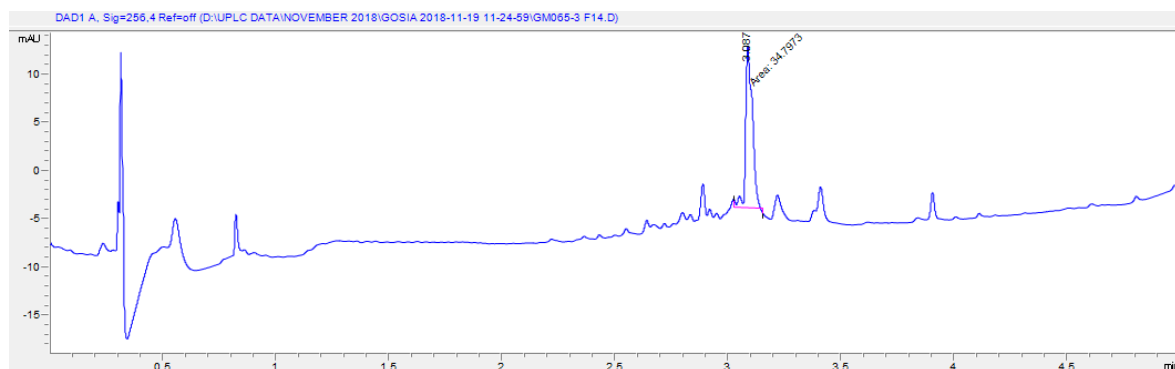


Synthesis of GM-3 (Pre-DAS+GM-2G)

GM-2G (64 μ mol, 400 μ L, 1 eq.) was mixed with **GM-3C** (80.08 μ mol, 455 μ L, 1.1 eq.), $CuSO_4$ (256 μ mol, 320 μ L, 5 eq.) and degassed for 1 minute. Then sodium ascorbate (502 μ mol, 448 μ L, 7 eq.) was added and the reaction mixture was flushed with nitrogen, sealed and incubated at room temperature for 3 hours.

Experimental Procedures

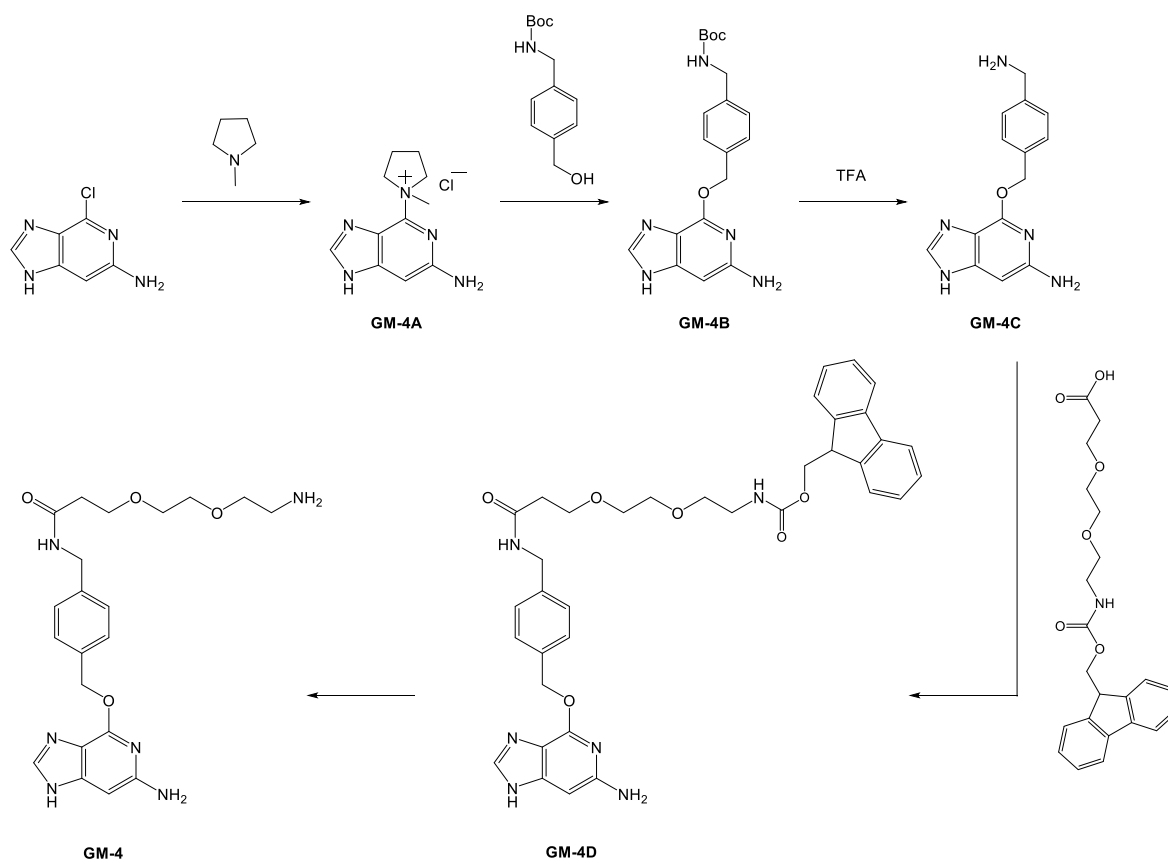
MS for $C_{47}H_{56}Cl_2N_{14}O_6S$ expected $[M+1]$: 1016.36 m/z. Found: 1016.4 m/z



S 14. Upper panel: UPLC chromatogram at 256 nm, indicating formation of the product at $rt = 3.087$ min. Lower panel: MS of the same peak confirming product formation $[M+1]$: 1016.4 m/z.

Chapter 3

13.2 Synthesis of O⁶-BG analogues: GM-4.



Synthesis of GM-4A

Adopted from *Nature Biotechnology*, 21, 86–89 (2003)

6-chloro-9H-purin-2-amine (1 g, 5.9 mmol, 1 eq.) was suspended in DMF and stirred for 35 minutes. 1-methylpyrrolidine (1.88 ml, 17.7 mmol, 3 eq.) was added. After two days, 10 mL acetone was added to precipitate the product as a pale yellow solid. The solution was filtered, the solid was washed with acetone and dried in vacuo to obtain the pure product (1.5 g, 5.6 mmol, 99%).

¹H NMR (500 MHz, D₂O) δ 8.15 (s, 1H), 4.66 – 4.55 (m, 2H), 3.96 – 3.83 (m, 2H), 3.59 (s, 3H), 2.35 – 2.19 (m, 2H), 2.18 – 2.02 (m, 2H).

Synthesis of GM-4B (O⁶-BG-NH₂-Boc)

2-Amino-6-chloropurine (100 mg, 1 eq.) and tert-butyl N-[[4-(hydroxymethyl)phenyl]methyl]carbamate (140 mg, 1 eq.) were dissolved in THF and tert-butyl N-[[4-(hydroxymethyl)phenyl]methyl]carbamate (235 mg, 2 eq.) was added. The reaction mixture was stirred at room temperature for 16 hours and the reaction was quenched with 1N HCl and reaction mixture was purified on RP-HPLC giving slightly yellow solid product with 20% yield.

¹H-NMR (500 MHz, Aceton-d₆) δ 8.08 (s, 1H), 7.34 – 7.39 (d, J = 7.9 Hz, 2H), 7.20 -7.24 (d, J = 7.9 Hz, 2H), 5.45 (s, 2H), 4.16 (s, 2H), 1.29 (s, 9H).

MS for C₁₈H₂₂N₆O₃ expected [M+1]: 371.2 m/z. Found: 371.2 m/z

Synthesis of GM-4C (O⁶-BG-NH₂)

GM-4B (20 mg) was dissolved in CH₂Cl₂:TFA 1:1 (1 mL) and stirred for 5 hours. Reaction volume had been reduced under reduced pressure and resuspended in dioxane (2 mL). Precipitate was filtered and washed twice with CH₂Cl₂ resulting in white solid (70% yield).

¹H-NMR (500 MHz, D₂O): 8.26 (s, 1H), 7.57 – 7.64 (d, 3J= 7.9 Hz, 2H), 7.49 -7.53 (d, 3J= 7.9 Hz, 2H), 5.64, (s, 2H), 4.22 (s, 2H).

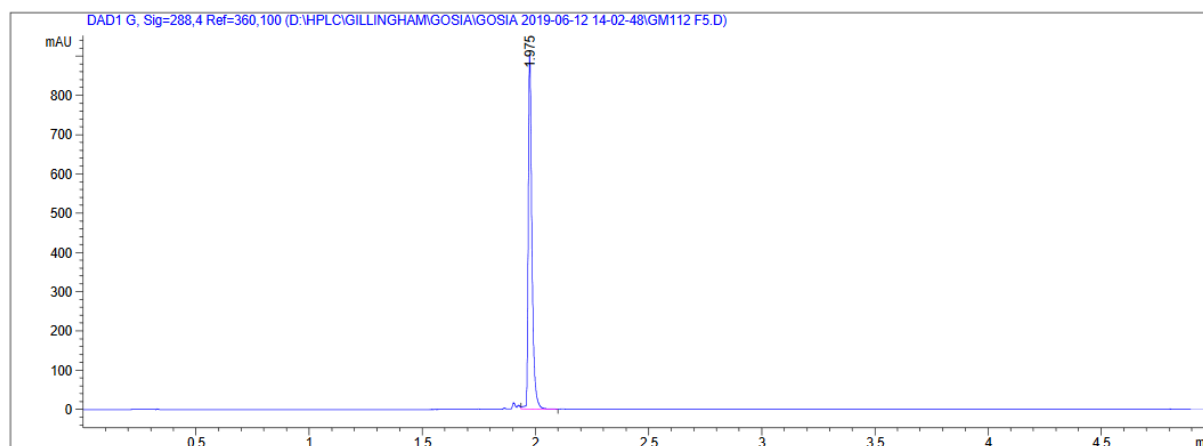
MS for C₁₃H₁₄N₆O expected [M+1]: 271.29 m/z. Found: 271.3 m/z

Synthesis of GM-4D (O⁶-BG-PEG2-NH₂-Fmoc)

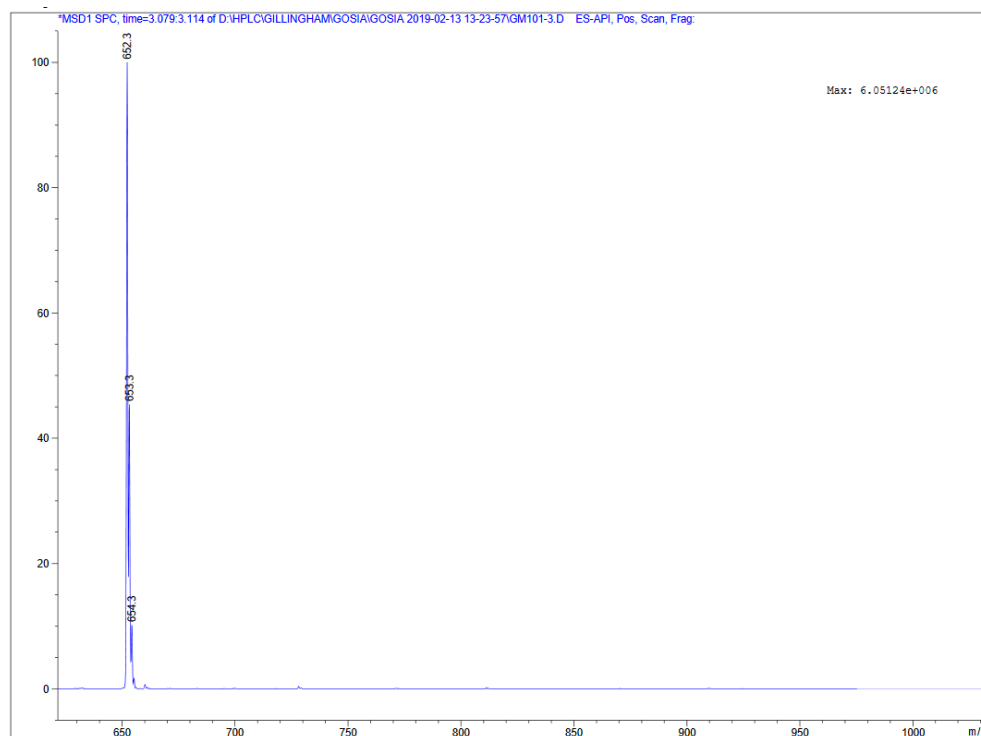
3-[2-[2-(9H-fluoren-9-ylmethoxycarbonylamino)ethoxy]ethoxy]propanoic acid (32.5 mg, 1.1 eq.) and PyBOP (42 mg, 1.1 eq.) were stirred in 2 mL of DMF for 45 minutes at room temperature. **GM-4C** (20 mg, 1 eq.) and DIEA (40 μL, 4 eq.) were added, stirred at 50°C for 5 minutes, cool down to room temperature and stirred for another hour. Solvents were removed under reduced pressure and reaction mixture purified on RP-HPLC resulting in white product with 80% yield.

¹H NMR (400 MHz, MeOD) δ 7.85 (s, 1H), 7.75 (d, J = 7.5 Hz, 2H), 7.60 (d, J = 7.5 Hz, 2H), 7.44 (d, J = 7.9 Hz, 2H), 7.36 (t, J = 7.4 Hz, 2H), 7.32 – 7.22 (m, 4H), 5.49 (s, 2H), 4.41 – 4.34 (m, 2H), 4.32 (d, J = 6.8 Hz, 2H), 4.15 (t, J = 6.9 Hz, 1H), 3.75 (t, J = 6.0 Hz, 2H), 3.57 (td, J = 5.5, 3.4 Hz, 4H), 3.46 (t, J = 5.5 Hz, 2H), 3.23 (t, J = 5.5 Hz, 2H), 2.48 (t, J = 6.0 Hz, 2H)

MS for C₃₆H₃₈N₆O₆ expected [M+1]: 651.29 m/z. Found: 652.3 m/z



Chapter 3



S 15. Upper panel: UPLC chromatogram at 288 nm, indicating formation of the product at $rt = 1.975$ min. Lower panel: MS of the same peak confirming product formation $[M+1]$: 652.3 m/z.

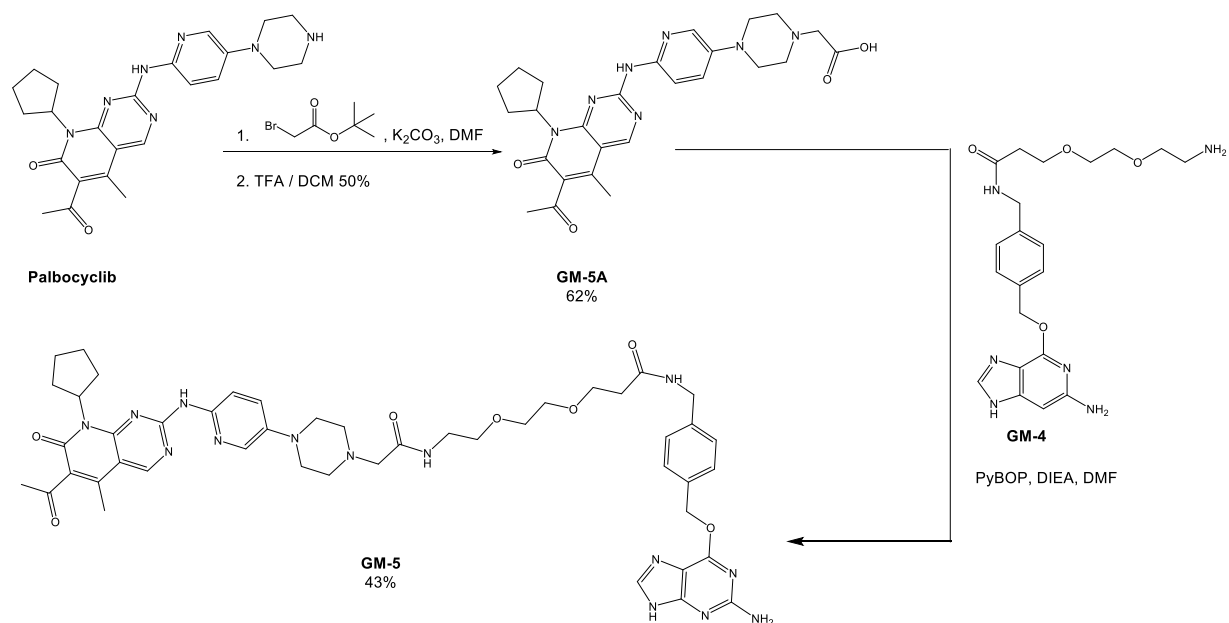
Synthesis of GM-4 (O⁶-BG-PEG2-NH₂)

GM-4D (50 mg) was dissolved in piperidine (4 mL) and stirred for 1 hour. Solvent was removed under the reduced pressure and sample purified on RP-HPLC giving white product (100% yield)

¹H NMR (400 MHz, DMSO-d₆) δ 8.43 (t, $J = 5.9$ Hz, 1H), 7.92 (s, 1H), 7.89 – 7.81 (m, 4H), 7.46 (d, $J = 7.9$ Hz, 2H), 7.28 (d, $J = 7.9$ Hz, 2H), 6.46 (s, 2H), 5.47 (s, 2H), 4.28 (d, $J = 5.9$ Hz, 2H), 3.65 (t, $J = 6.3$ Hz, 2H), 3.61 – 3.48 (m, 5H), 2.95 (q, $J = 5.4$ Hz, 2H), 2.40 (t, $J = 6.3$ Hz, 2H), 2.35 (s, 1H).

HRMS for C₂₀H₂₇N₇O₄ expected $[M+1]$: 430.22 m/z. Found $[M+1]$ 430.2202 m/z

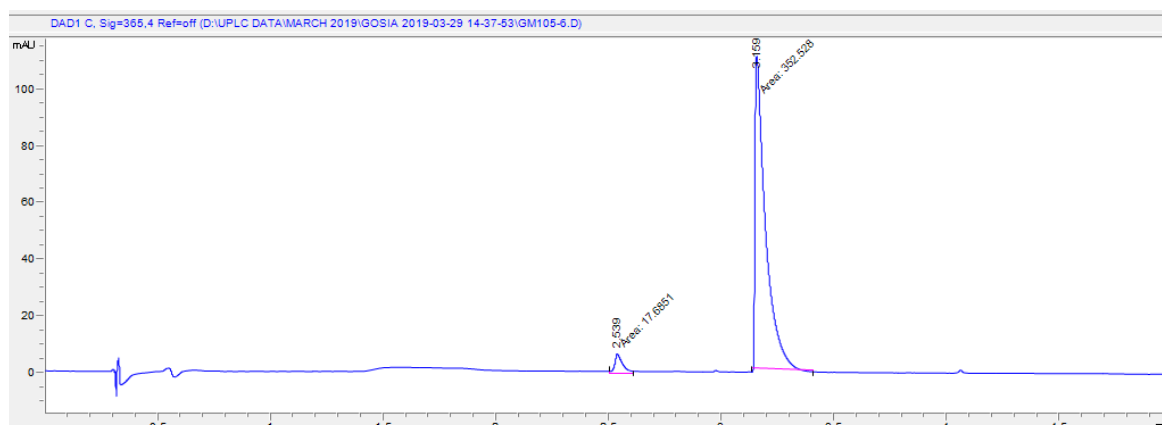
13.3 Synthesis of BCR-Abl substrates: GM-5.


Synthesis of GM-5A (Palbociclib + t-Bu-bromoacetate)

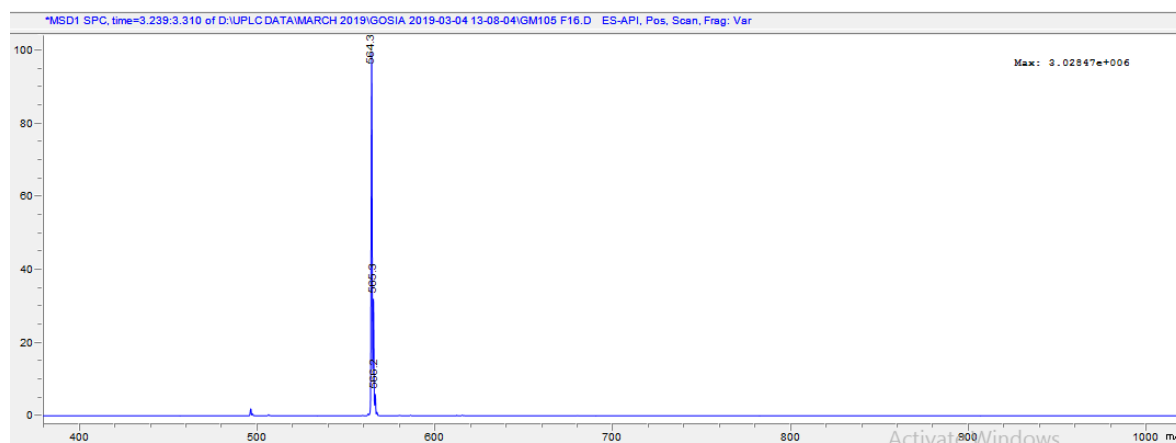
Palbociclib (10 mg, 1 eq.) was mixed with tert-butyl 2-bromoacetate (4.4 mg, 1 eq.), K₂CO₃ (15.4 mg, 5 eq.) and dissolved in DMF (5 mL) and stirred for 20 hours at the room temperature. Crude was purified by PR-HPLC (H₂O-ACN with 0.1% TFA) giving 10 mg of white solid (62% yield).

¹H NMR (500 MHz, DMSO-d₆) δ 10.31 (s, 1H), 8.97 (s, 1H), 8.11 (s, 1H), 7.91 (d, *J* = 9.0 Hz, 1H), 7.59 (dd, *J* = 9.1, 3.1 Hz, 1H), 5.84 (p, *J* = 8.8 Hz, 1H), 4.21 (s, 2H), 2.30 – 2.19 (m, 1H), 1.94 – 1.88 (m, 2H), 1.83 – 1.73 (m, 2H), 1.64 – 1.56 (m, 2H), 1.50 (s, 9H).

MS for C₃₀H₄₃N₇O₄ expected [M+1]: 566.34 m/z. Found: 566.3 m/z

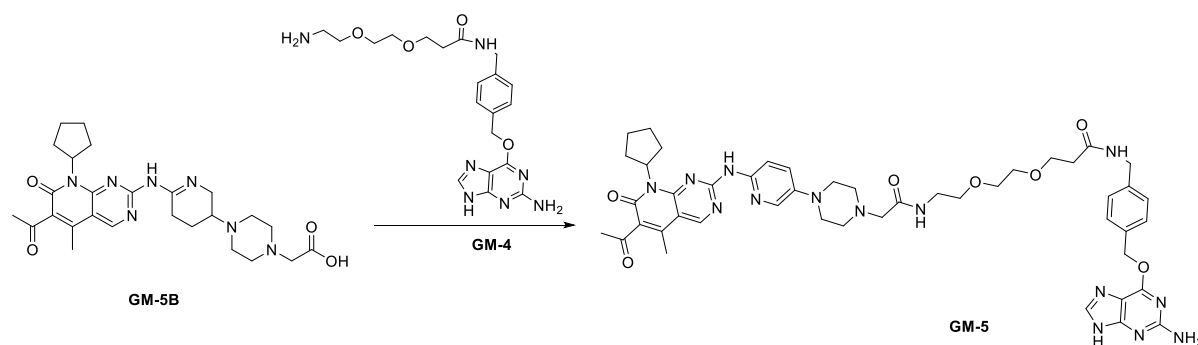


Chapter 3



S 16. Upper panel: UPLC chromatogram at 365 nm, indicating formation of the product at $rt = 3.159$ min. Lower panel: MS of the same peak confirming product formation $[M+1]: 566.3$ m/z.

GM-5A was dissolved in $\text{CH}_2\text{Cl}_2/\text{TFA}$ 1:1 (5 mL) and stirred at room temperature for 1 hour. Solvent was removed under reduced pressure and crude product (**GM-5A**) was used directly in next step.



Synthesis of GM-5

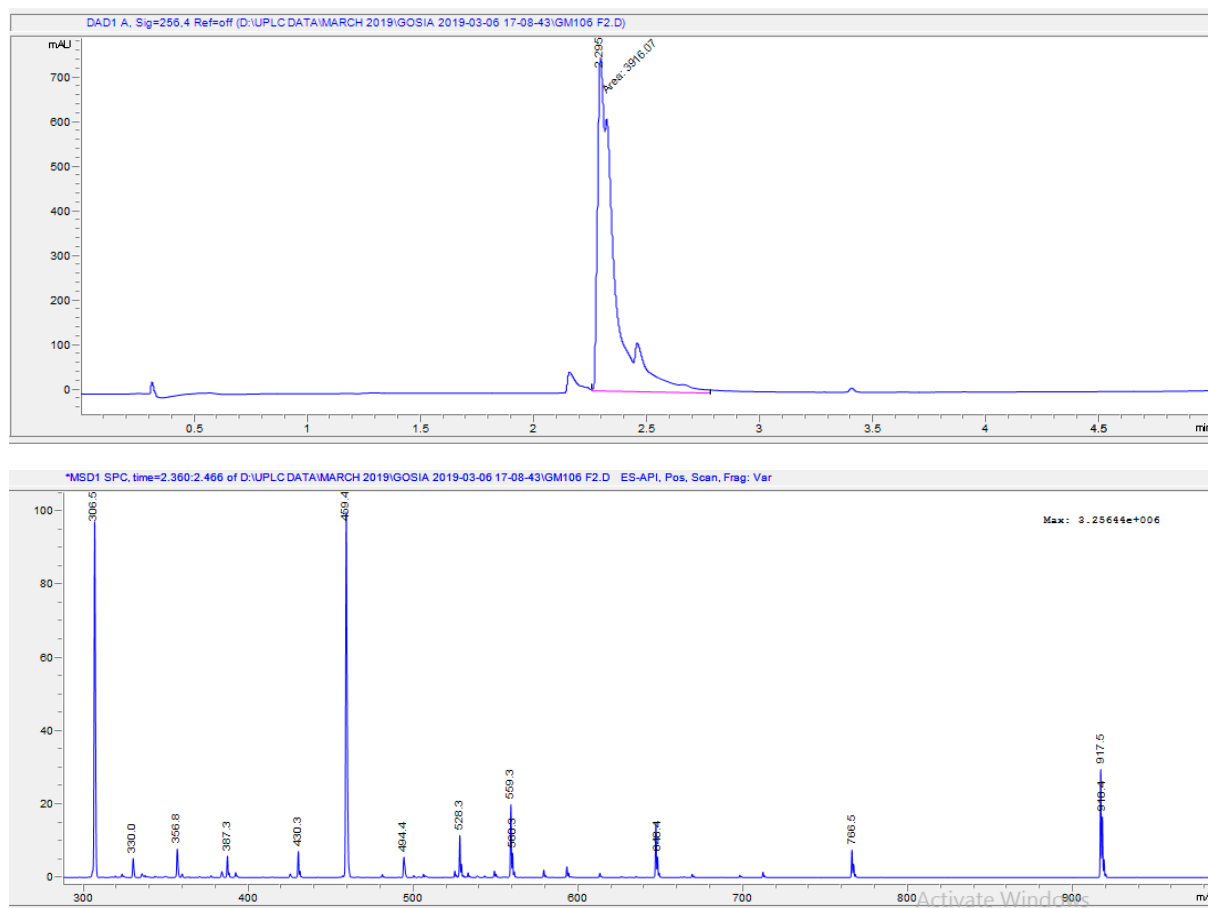
GM-5A (4.5 mg, 1 eq.) was mixed with PyBOP (5.1 mg, 1.1 eq.) and dissolved in DMF (1 mL) and stirred for 45 minutes. **GM-4** (3.82 mg, 1 eq.), DIPEA (6.4 μL , 4 eq.) were added and the reaction mixture was stirred at 50°C for 5 minutes, followed by 1 hour at room temperature. Solvent was removed under the reduced pressure and sample purified on RP-HPLC yielding 3.5 mg of white powder (43%).

$^1\text{H-NMR}$ (600 MHz, DMSO-d_6): 10.26 (s, 1H), 8.96 (s, 1H), 8.64 (br t, $J = 5.40$, 1H), 8.40 (br t, 1H, $J = 5.93$), 8.21 (s, 1H), 8.10 (br d, $J = 2.94$, 1H), 7.90 (d, $J = 9.06$, 1H), 7.56 (dd, $J = 9.10$, $J = 2.99$, 1H), 7.47 (d, $J = 8.07$, 2H), 7.28 (d, $J = 8.11$, 2H), 5.83 (q, $J = 8.95$, 1H), 5.49 (s, 2H), 4.28 (d, $J = 5.78$, 2H), 4.01 (s, 2H), 3.65 (t, $J = 6.39$, 2H), 3.52 (br s, 2H), 3.49 (br, 2H), 3.47 (t, $J = 5.51$, 2H), 3.31 (br, 8H), 2.42 (s, 3H), 2.39 (t, $J = 6.44$, 2H), 2.32 (s, 3H), 2.25 (br, 2H), 1.89 (br, 2H), 1.77 (br, 2H), 1.58 (br, 2H).

$^{13}\text{C-NMR}$ (151 MHz, DMSO-d_6): 202.39, 170.16, 164.12, 160.69, 159.06, 158.40, 158.15, 158.06, 157.85, 154.75, 154.54, 144.84, 141.96, 141.48, 139.57, 135.43, 134.43, 129.42, 128.69, 127.23, 125.79, 115.09, 106.56, 69.54, 69.42, 68.73, 67.43, 66.83, 56.15, 52.93, 51.32, 45.36, 41.84, 36.08, 31.31, 27.59, 25.12, 13.63.

MS for $\text{C}_{46}\text{H}_{56}\text{N}_{14}\text{O}_7$ expected $[M+1]: 917.45$ m/z. Found: 917.5 m/z

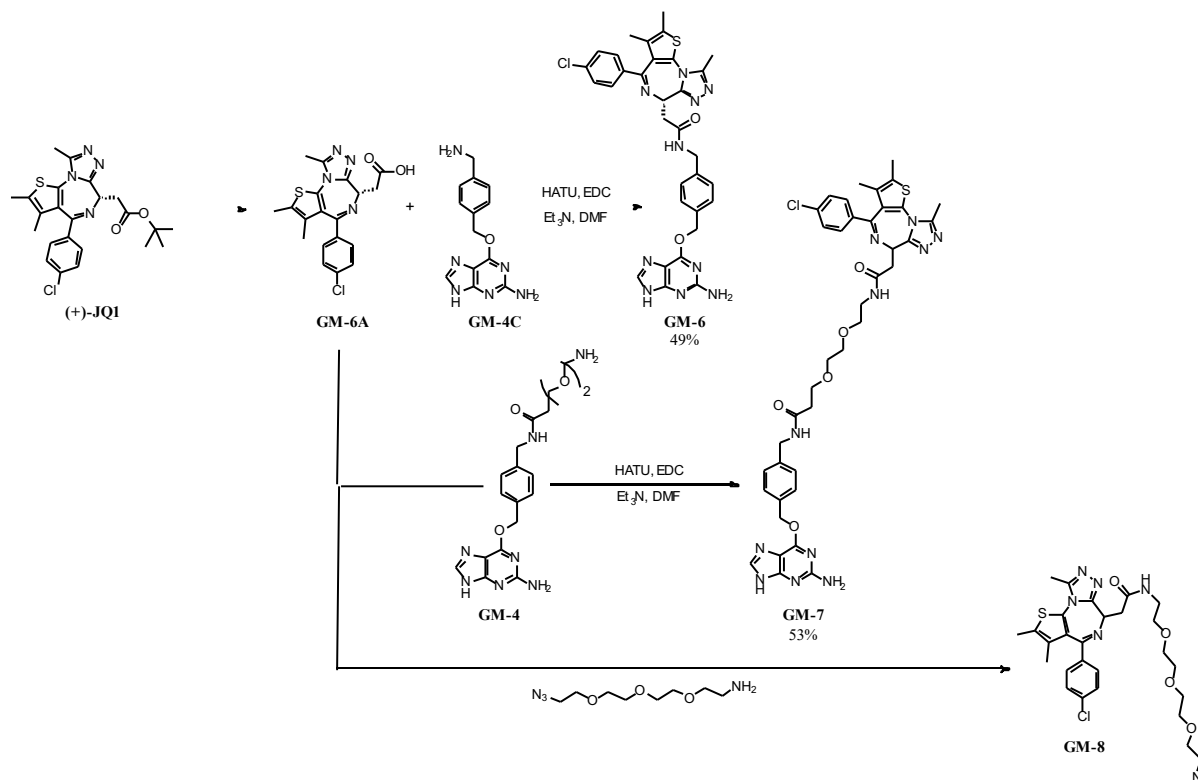
Experimental Procedures



S 17. Upper panel: UPLC chromatogram at 256 nm, indicating formation of the product at $rt = 2.295$ min. Lower panel: MS of the same peak confirming product formation $[M+1]$: 917.5 m/z.

Chapter 3

13.4 Synthesis of MGMT-BRD4 substrates: GM-6, GM-7, GM-8.

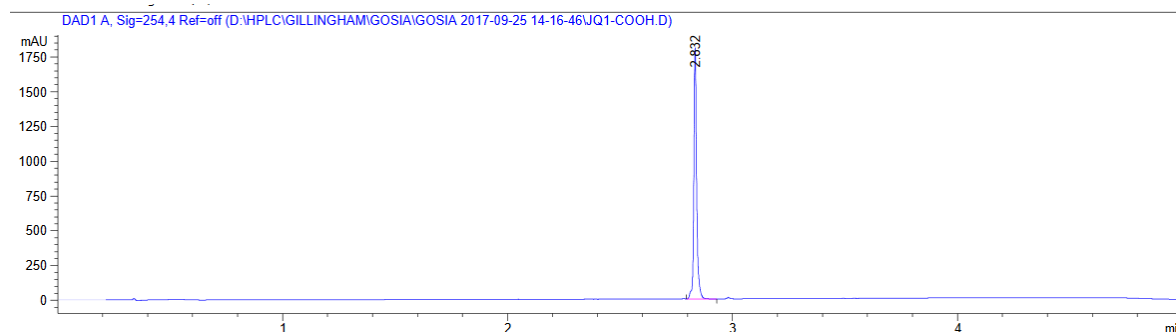


Synthesis of GM-6A

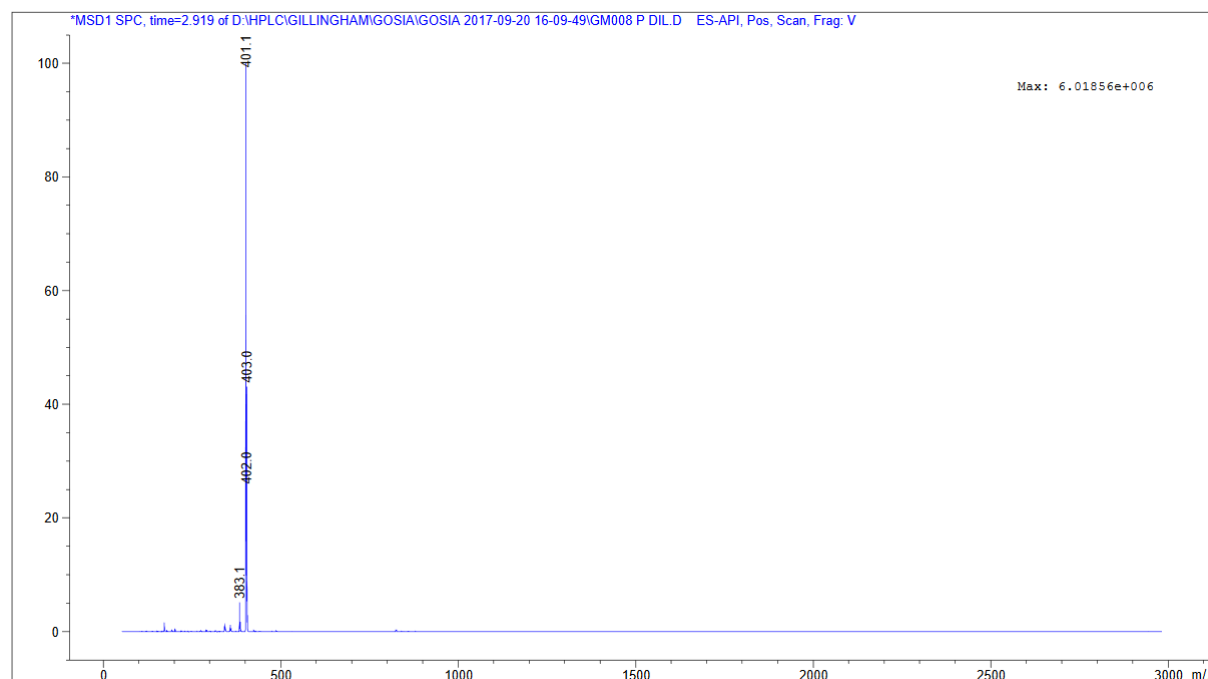
Adopted from patent WO/2016/146985.

(+)-JQ-1 (15 mg, 0.33 mmol) was dissolved in formic acid and stirred at room temperature for 16 hours and monitored with UPLC-MS. After full conversion the solvent was removed under reduced pressure and resulting white solid used for next step without further purification.

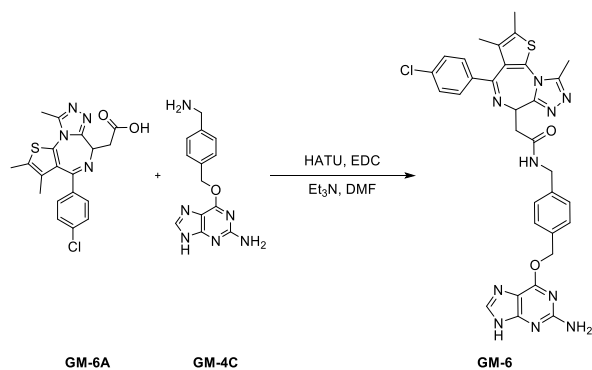
MS for $C_{19}H_{17}ClN_4O_2S$ expected $[M+1]$: 401.1 m/z. Found: 401.1 m/z



Experimental Procedures



S 18. Upper panel: UPLC chromatogram at 254 nm, indicating formation of the product at $rt = 2.832$ min. Lower panel: MS of the same peak confirming product formation $[M+1]: 401.1$ m/z.



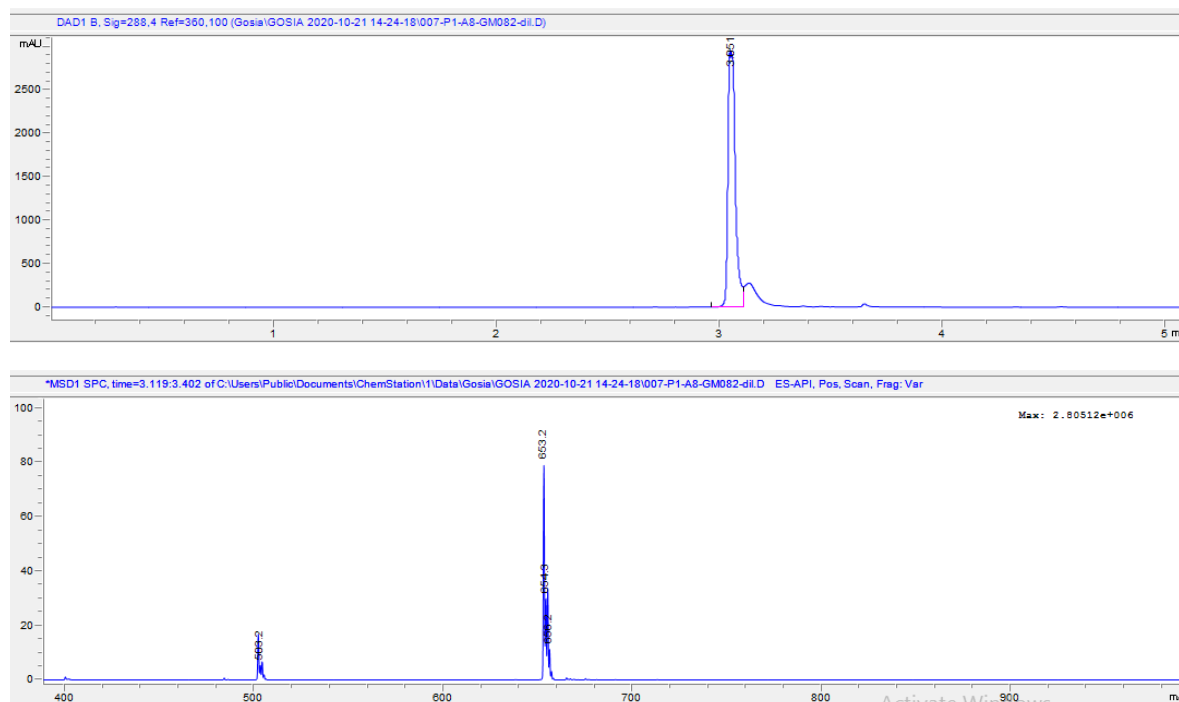
Synthesis of GM-6

GM-6A (4 mg, 1.1 eq.) was mixed with PyBOP (5.7 mg, 1.1 eq.) in DMF (1 mL) and stirred for 45 minutes. GM-4C (2.7 mg, 1 eq.), DIPEA (5.6 μL , 4 eq.) were added and the reaction mixture was stirred at 50°C for 5 minutes, followed by 1 hour at room temperature. Solvent was removed under the reduced pressure and sample purified on RP-HPLC yielding 2.3 mg of white powder (49%).

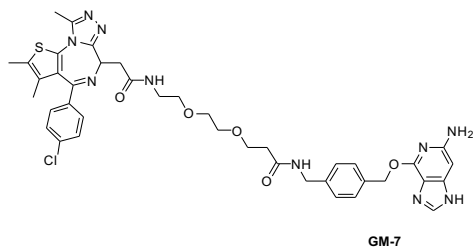
$^1\text{H NMR}$ (400 MHz, DMSO- d_6) δ 8.76 (s, 1H), 7.82 (s, 1H), 7.47 (dd, 4H), 7.40 (d, $J = 8.6$ Hz, 2H), 7.34 (d, $J = 8.2$ Hz, 2H), 6.22 (s, 2H), 5.48 (s, 2H), 2.60 (s, 3H), 2.55 (s, 2H), 2.46 (s, 1H), 2.41 (s, 3H), 1.62 (s, 3H), 1.24 (s, 1H).

HRMS for $\text{C}_{32}\text{H}_{29}\text{ClN}_{10}\text{O}_2\text{S}$ expected $[M+1]: 653.19$ m/z. Found $[M+1] 653.2$ m/z

Chapter 3



S 19. Upper panel: UPLC chromatogram at 288 nm, indicating formation of the product at $rt = 3.051$ min. Lower panel: MS of the same peak confirming product formation $[M+1]$: 653.2 m/z.



Synthesis of GM-7

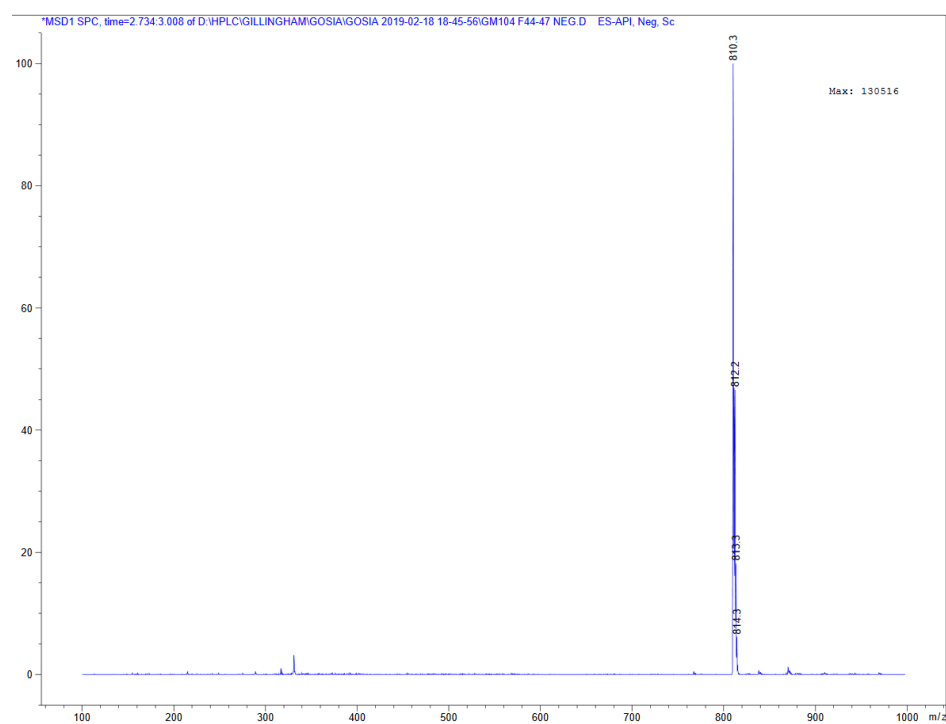
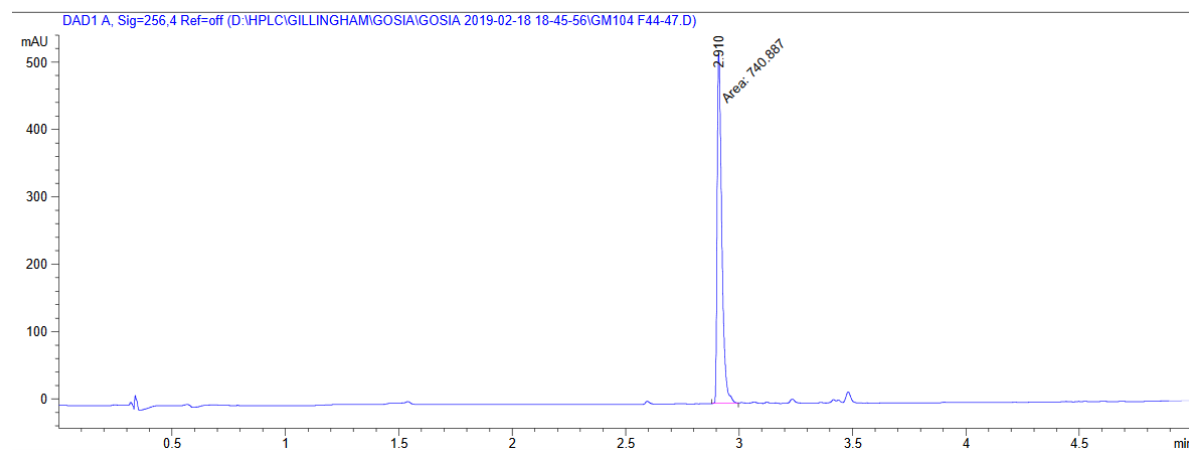
GM-6A (5.13 mg, 1.1 eq.) was mixed with PyBOP (6.66 mg, 1.1 eq.) in DMF (1 mL) and stirred for 45 minutes. **GM-4** (5 mg, 1 eq.), DIPEA (6.7 μ L, 4 eq.) were added and the reaction mixture was stirred at 50°C for 5 minutes, followed by 1 hour at room temperature. Solvent was removed under the reduced pressure and sample purified on RP-HPLC yielding 5 mg of white powder (53%).

$^1\text{H-NMR}$ (600 MHz, DMSO- d_6): 12.78 (broad, 1H), 8.37 (t, 1H, $J = 5.97$), 8.28 (t, 1H, $J = 5.70$), 8.08 (broad, 2H), 7.48 (d, 2H, $J = 8.70$), 7.45 (d, 2H, $J = 8.07$), 7.42 (d, 2H, $J = 8.50$), 7.27 (d, 2H, $J = 8.16$), 6.65 (broad, 1H), 5.47 (s, 2H), 4.51 (dd, 1H, $J = 8.07$, $J = 7.94$), 4.27 (d, 2H, $J = 5.99$), 3.64 (t, 2H, $J = 6.53$), 3.52 (m, 4H), 3.44 (t, 2H, $J = 5.85$), 3.27 (m, 2H), 3.23 (m, 2H), 2.59 (s, 3H), 2.40 (s, 3H), 2.38 (t, 2H, $J = 6.44$), 1.61, 1.61 (s, 3H).

$^{13}\text{C-NMR}$ (151 MHz, DMSO- d_6): 170.00, 169.60, 162.90, 159.04, 154.96, 149.67, 139.50, 136.62, 135.10, 134.51, 130.58, 130.03, 130.02, 130.00, 129.71, 128.52, 128.34, 126.94, 69.44, 69.09, 67.02, 66.75, 53.71, 41.70, 38.51, 37.39, 36.04, 13.93, 12.56, 11.18.

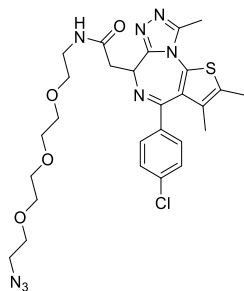
MS for $\text{C}_{39}\text{H}_{42}\text{ClN}_{11}\text{O}_5\text{S}$ expected $[M-1]$: 810.3 m/z. Found: 810.3 m/z

Experimental Procedures



S 20. Upper panel: UPLC chromatogram at 256 nm, indicating formation of the product at $rt = 2.910$ min. Lower panel: MS of the same peak confirming product formation $[M+1]$: 810.3 m/z.

Chapter 3

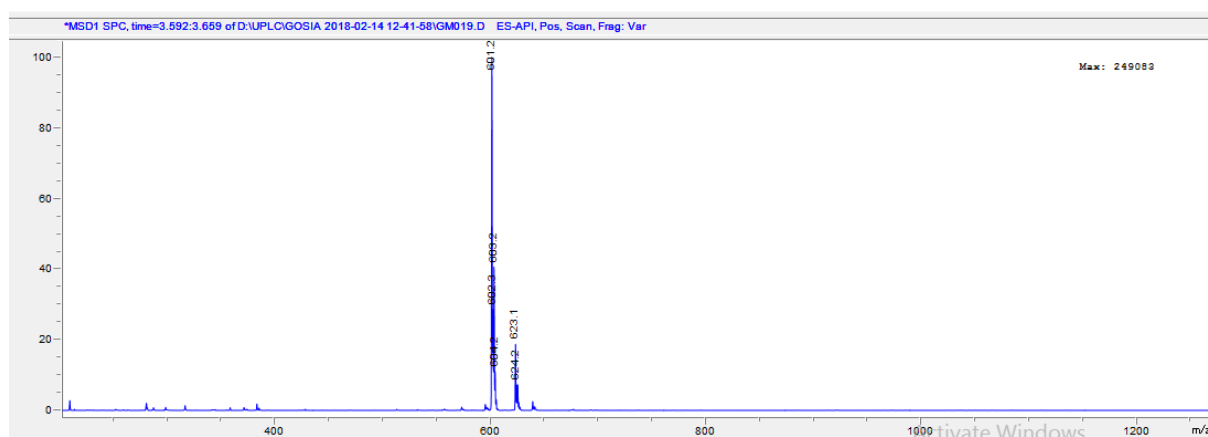
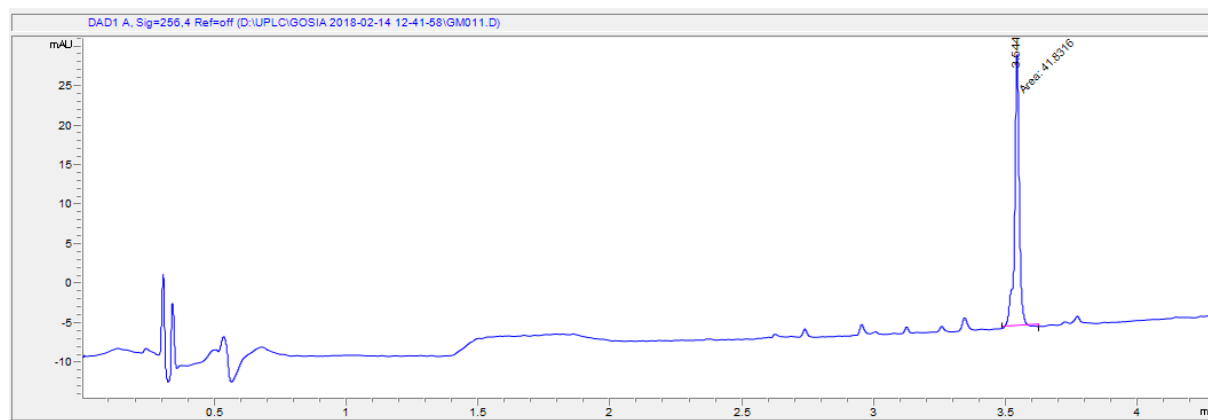


Synthesis of GM-8

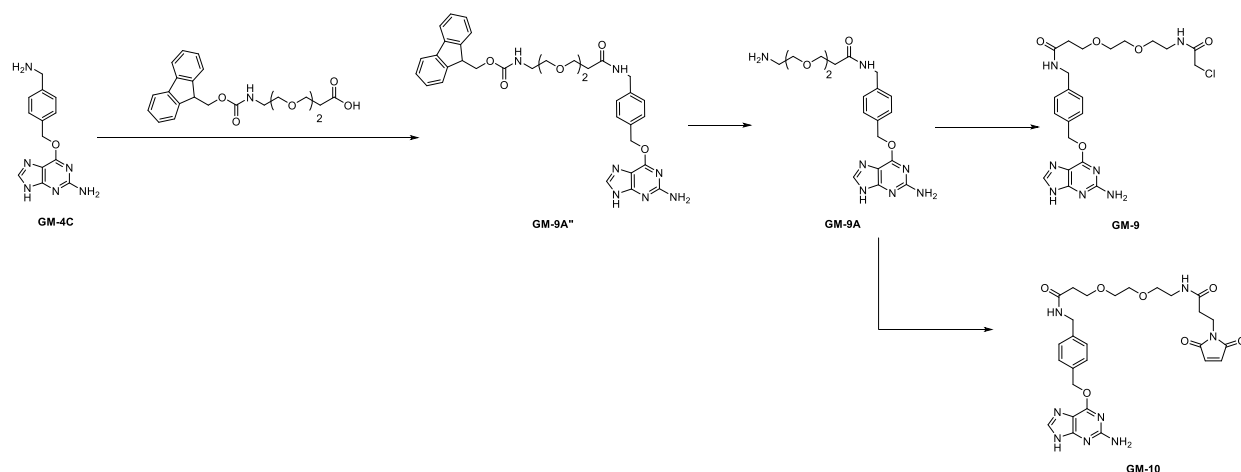
GM-6A (1 eq.) was mixed with PEG-3 diamine (1.5 eq.), HATU (4 eq.), DIEA (5 eq.) and dissolved in DMF. Reaction mixture was stirred at room temperature for 1 hour and purified on RP-HPLC (H₂O/ACN) resulting in white product with 56% yield.

¹H NMR (500 MHz, CD₃CN) δ 7.46 (d, J = 8.7 Hz, 1H), 7.42 (d, J = 8.7 Hz, 2H), 6.99 (s, 1H), 4.58 (t, J = 6.9 Hz, 1H), 3.64 – 3.54 (m, 10H), 3.52 (td, J = 5.4, 1.1 Hz, 2H), 3.37 – 3.34 (m, 3H), 3.29 (d, J = 6.9 Hz, 2H), 2.64 (s, 3H), 2.41 (s, 2H), 1.66 (s, 2H).

MS for C₂₇H₃₃ClN₈O₄S expected [M+1]: 601.2 m/z. Found: 601.1 m/z



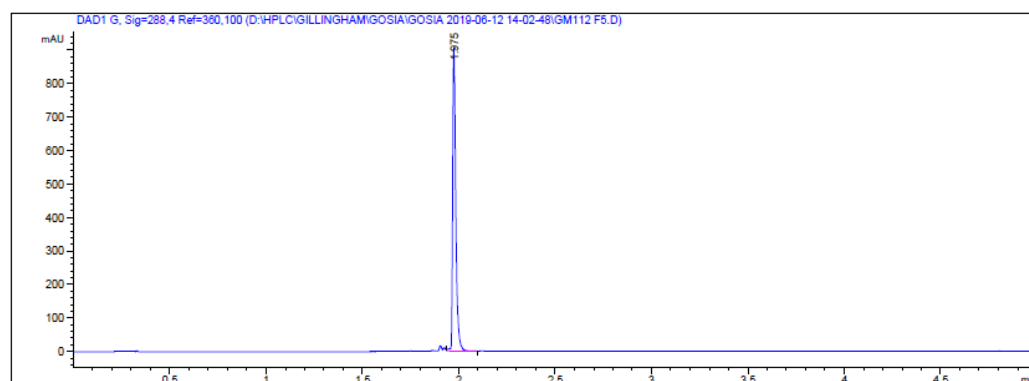
13.5 Synthesis of covalent probes: GM-9, GM-10, GM-11


Synthesis of GM-9

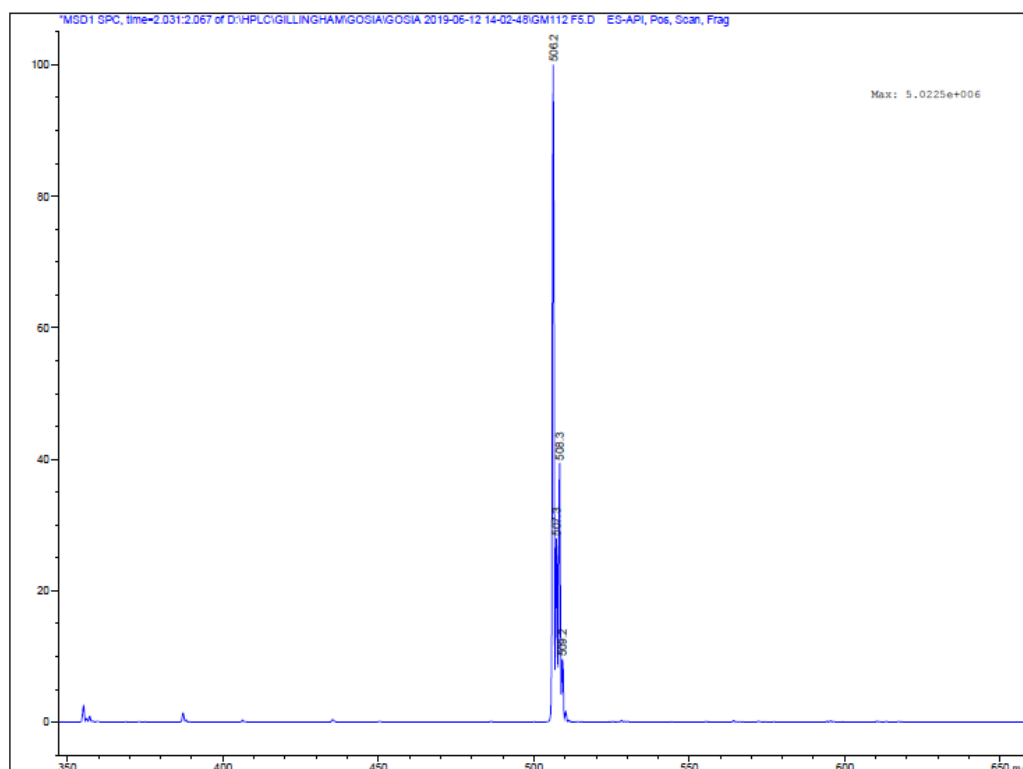
2-chloroacetic acid (7.0 mg, 1.06 eq.) was mixed with PyBOP (40.0 mg, 1.1 eq.) and dissolved in DMF (1 mL) and stirred for 45 minutes. **GM-9A** (30.0 mg, 1 eq.), DIEA (45 μ L, 5 eq.) were added and the reaction mixture was stirred at 50°C for 5 minutes, followed by 1 hour at room temperature. Solvent was removed under the reduced pressure and sample purified on RP-HPLC yielding 28.9 mg of white powder (82%).

$^1\text{H NMR}$ (500 MHz, DMSO- d_6) δ 8.44 (s, 1H), 8.39 (d, $J = 5.9$ Hz, 1H), 8.25 (d, $J = 5.4$ Hz, 1H), 7.49 (d, $J = 4.1$ Hz, 2H), 7.29 (d, $J = 4.0$ Hz, 2H), 5.53 (s, 2H), 4.28 (d, $J = 3.0$ Hz, 2H), 4.05 (s, 2H), 3.63 (t, $J = 6.4$ Hz, 2H), 3.50 (s, 4H), 3.43 (t, $J = 5.8$ Hz, 2H), 3.23 (d, $J = 5.9$ Hz, 2H), 2.38 (d, $J = 6.3$ Hz, 2H).

MS for C₂₂H₂₈ClN₇O₅ expected [M+1]: 506.2 m/z. Found: 506.2 m/z



Chapter 3



S 22. Upper panel: UPLC chromatogram at 288 nm, indicating formation of the product at $rt = 1.975$ min. Lower panel: MS of the same peak confirming product formation $[M+1]$: 506.2 m/z.

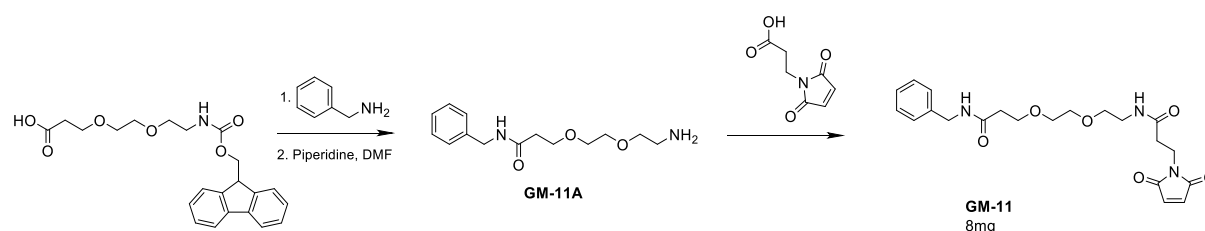
Synthesis of GM-10

3-(2,5-dioxopyrrol-1-yl)propanoic acid (5.2 mg, 1.1 eq.) was mixed with PyBOP (16.0 mg, 1.1 eq.) and dissolved in DMF (1 mL) and stirred for 45 minutes. **GM-9A** (12.0 mg, 1 eq.), DIEA (16 μ L, 4 eq.) were added and the reaction mixture was stirred at 50°C for 5 minutes, followed by 1 hour at room temperature. Solvent was removed under the reduced pressure and sample purified on RP-HPLC yielding 13.0 mg of white powder (80.1%).

$^1\text{H NMR}$ (500 MHz, DMSO- d_6) δ 8.40 (d, $J = 2.9$ Hz, 2H), 8.38 (s, 1H), 8.02 (t, $J = 5.6$ Hz, 1H), 7.49 (d, $J = 4.2$ Hz, 2H), 7.29 (d, $J = 4.1$ Hz, 2H), 6.99 (s, 2H), 5.52 (s, 2H), 4.28 (d, $J = 3.0$ Hz, 2H), 3.63 (t, $J = 6.4$ Hz, 2H), 3.59 (t, $J = 7.1$ Hz, 2H), 3.52 – 3.44 (m, 4H), 3.36 (t, $J = 5.9$ Hz, 2H), 3.14 (q, $J = 5.8$ Hz, 2H), 2.39 (t, $J = 6.3$ Hz, 2H), 2.33 (d, $J = 4.9$ Hz, 2H).

$^{13}\text{C NMR}$ (126 MHz, DMSO- d_6) δ 171.20, 170.67, 169.98, 159.37, 158.95, 158.68, 154.19, 141.07, 140.38, 135.00, 134.56, 129.29, 127.69, 69.95, 69.91, 69.46, 68.43, 67.30, 65.38, 42.27, 38.94, 36.60, 34.52, 34.39, 15.63.

HRMS for $\text{C}_{27}\text{H}_{32}\text{N}_8\text{O}_7$ expected $[M+1]$: 581.24 m/z. Found: 581.2467 m/z

Synthesis of GM-11

Synthesis of GM-11A

3-[2-[2-(9H-fluoren-9-ylmethoxycarbonylamino)ethoxy]ethoxy]propanoic acid (20 mg, 1 eq.) was mixed with T3P (40 mg, 2.5 eq.) in DMF (2 mL) and stirred for 15 minutes. Phenylmethanamine (5.9 mg, 1.1 eq.) was added with TEA and the reaction was stirred for 1 hour. Solvents were removed under the reduced pressure and crude product was dissolved in 20% piperidine in DMF and stirred for 1 hour. Crude product was purified on RP-HPLC yielding with 10.1 mg of white solid (74% yield).

¹H NMR (500 MHz, DMSO-d₆) δ 8.41 (t, *J* = 6.0 Hz, 1H), 7.31 (t, *J* = 7.7 Hz, 2H), 7.27 – 7.16 (m, 3H), 4.28 (d, *J* = 6.0 Hz, 2H), 3.65 (t, *J* = 6.3 Hz, 2H), 3.58 (t, *J* = 5.2 Hz, 2H), 3.56 – 3.48 (m, 4H), 2.95 (s, 2H), 2.40 (t, *J* = 6.3 Hz, 2H).

¹³C NMR (126 MHz, DMSO-d₆) δ 170.65, 158.66, 158.41, 139.96, 128.70, 127.62, 127.18, 70.07, 69.81, 67.27, 67.14, 42.47, 39.05, 36.53.

HRMS for C₁₄H₂₂N₂O₃ expected [M+1]: 267.16 m/z. Found: 267.1703 m/z

Synthesis of GM-11

3-(2,5-dioxopyrrol-1-yl)propanoic acid (7.0 mg, 1 eq.) was mixed with T3P (30 mg, 2.5 eq.) in DMF (2 mL) and stirred for 15 minutes. GM-11A (10 mg, 1 eq.) was added with TEA and the reaction was stirred for 1 hour. Solvents were removed under the reduced pressure and crude product was purified on RP-HPLC yielding with 8 mg of white solid (52% yield).

¹H NMR (500 MHz, DMSO-d₆) δ 8.35 (d, *J* = 5.5 Hz, 1H), 8.01 (t, *J* = 5.7 Hz, 1H), 7.30 (dd, 2H), 7.26 – 7.19 (m, 3H), 7.00 (s, 2H), 4.27 (d, *J* = 3.0 Hz, 2H), 3.64 (t, *J* = 6.4 Hz, 2H), 3.59 (t, *J* = 8.3 Hz, 2H), 3.48 (s, 4H), 3.36 (t, *J* = 5.9 Hz, 2H), 3.14 (q, *J* = 5.8 Hz, 2H), 2.38 (t, *J* = 6.3 Hz, 2H), 2.33 (t, *J* = 7.9, 6.7 Hz, 2H).

¹³C NMR (126 MHz, DMSO-d₆) δ 171.21, 170.62, 169.96, 139.96, 135.02, 128.68, 127.60, 127.15, 69.96, 69.91, 69.48, 67.32, 42.45, 38.94, 36.61, 34.53, 34.39.

HRMS for C₂₁H₂₇N₃O₆ expected [M+Na]: 440.2 m/z. Found: 440.1792 m/z

Chapter 3

Chapter 4

Mechanistic study of MGMT

14 Cell culture

HUH6 were gifts from the Hall laboratory and maintained at 37°C with 5% CO₂ in RPMI containing 10% FCS and 1% penicillin–streptomycin.

HEK293T ACC635 cells were obtained from the DSMZ , together with HeLa cells (gift from Palivan-Meier group at University of Basel), maintained at 37°C with 5% CO₂ in DMEM containing 10% FCS and 1% penicillin–streptomycin.

15 Immunoblotting

Cells were harvested, washed three times with PBS and lysed in freshly prepared RIPA buffer (50 mM Tris pH 8, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS) supplied with 1x Complete Mini Protease Inhibitor Cocktail (Roche) and 10 mM sodium orthovanadate. Protein concentration was quantified using Bradford Assay Reagents A, B and S (BioRAD #5000113, #5000114 and #5000115, respectively) measuring A_{750 nm} on TECAN SPARK® reader. 25 µg of protein lysate was developed on 8% SDS-PAGE gel and proteins were transferred to a nitrocellulose membrane (Amersham Protran 0.45 NC) using BioRAD Trans-Blot Turbo transfer system. Blots were blocked in Odyssey Blocking Buffer in TBS for 1 hour at room temperature, washed three times for 5 minutes with TBST (TBS with 0.1% Tween 20) and incubated with the primary antibody overnight at 4°C. After incubation blots were washed three times for 5 minutes with TBST and incubated with secondary antibody (Li-COR: IRDye 680RD Goat anti-Rabbit IgG or IRDye 800CW Goat anti-Mouse IgG, diluted 1:10000 in TBST) for 1 hour at room temperature. After secondary incubation blots were washed three times for 5 minutes with TBST and imaged with Odyssey CLx Imaging System.

Antibodies used in this study were as follow: Anti-MGMT (Abcam, ab80513, diluted as recommended to 1 µg/mL in TBST). Anti-Actin (Santa Cruz Biotechnology). Anti-BRD4 (Abcam, ab128874, working concentration 1:1000 in TBST buffer). Anti-c-ABL (Cell Signalling Technology, 2862T, working concentration 1:1000 in TBST buffer)

16 Genomic DNA extraction

Cells had been suspended in serum containing media using trypsin and spun down for 5 minutes at 500 x g. Supernatant was discarded and cell pellet was washed three times with 10 mL of ice-cold PBS and centrifuged for 5 minutes at 500 x g. Cells were re-suspended in digestion buffer (100 mM NaCl, 10 mM TrisCl pH 8, 25 mM EDTA pH 8, 0.5% SDS and freshly added 0.1 mg/mL proteinase K. For < 3 x 10⁷ 0.3 mL digestion buffer was used) and incubated for 16 hours at 50°C with shaking. DNA was extracted by adding an equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) and spinning down the mixture at 1700 x g for 10 minutes in swinging bucket. Top layer had been transferred to the new tube and 1/10 volume of 3M NaAc had been added. DNA immediately precipitated and was recovered by centrifugation at 1700 x g for 2 minutes, wash with 70% ethanol, air drying the pellet and dissolving it TE buffer.

Chapter 4

17 Generation of plasmids

pTet-TurboID-13x-MGMT was generated by cloning TurboID-13x-MGMT sequence into the pMK243 vector (Tet-OsTIR1-PURO, Addgene #72835) replacing OsTIR1 gene using HiFi DNA Assembly kit from New England Biolabs.

Primers used:

5'-gatcctctagacatatgctgcagagatctTCAGTTTCGGCCAGCAGGC-3';

5'-GCCTGCTGGCCGAAACTGAagatctctgcagcatatgtctagaggatc-3';

5'-gcacagtattgtctttgctagccatggtggcacgcgtgcg-3';

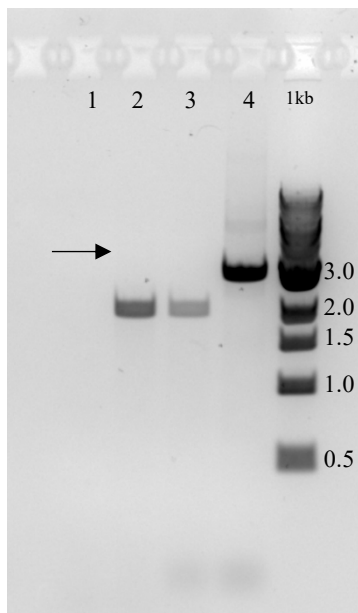
5'-cgcacgcgtgccaccatggctagcaagacaataactgtgcc-3'.

Sanger sequencing was provided by either Microsynth® facility in Basel, Switzerland or Genewiz® facility in Irvine, United States.

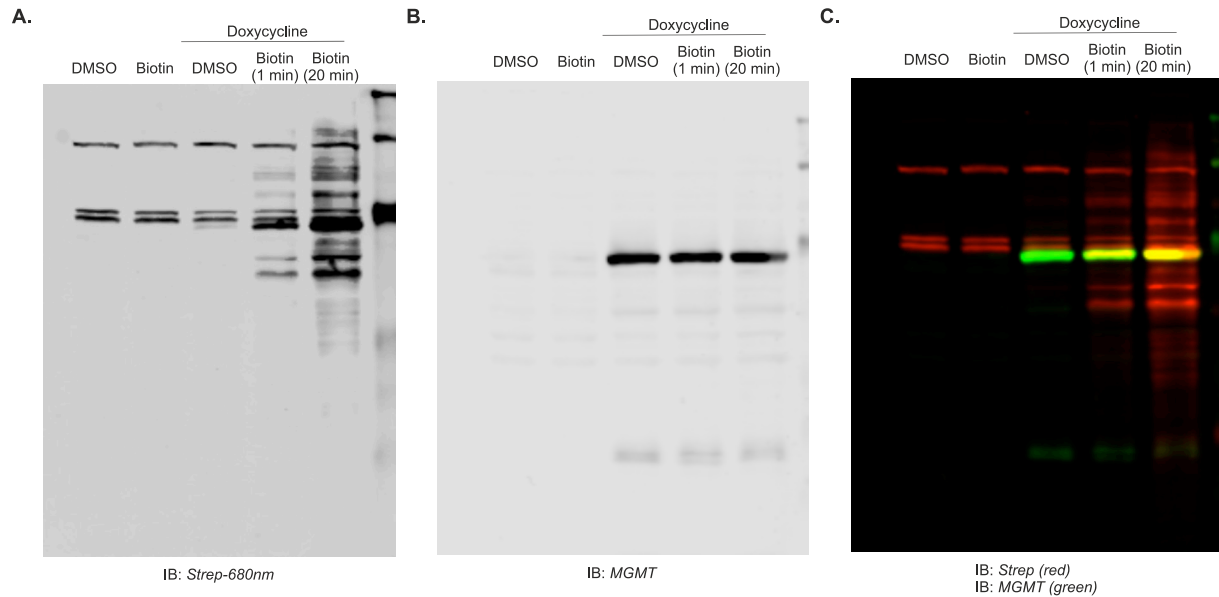
18 Generation of cell pools expressing TurboID-13x-MGMT in AAVS1 locus

Tet-TurboID-MGMT and *AAVS1 T2 CRISPR* in *pX330 DL* (7 µg) (Addgene #72833) were used to transfect 293T cells. 10 µg of Tet-TurboID-MGMT and 7 µg of AAVS1 T2 CRIPR in pX330 DL were added to 1.5 mL of Opti-MEM media. 36 µL of TurboFect transfecting agent was added and the mixture was incubated for 15 minutes at room temperature and added dropwise to the target cells. After 24 hours the medium was changed and after 48 hours 1 µg/mL of puromycin was used for selection.

Survived population had been pulled together and validated by genomic DNA extraction, PCR amplification and sequencing. Primers used: 5'-ggcaaacagcataagctggtc-3' and 5'-gagcttaccatgaccgagtac-3', generated fragment: 2020 bp.



S 1. PCR products of amplified: 1 Genomic DNA WT; 2 DNA extracted from 293T ACC635 integrated in AAVS1 locus; 3 DNA extracted from 293T integrated in AAVS1 locus; 4 Tet-TurboID-MGMT plasmid used for transfection. Expected band size: 2020 bp.



S. 2. TurboID-MGMT expression and BioID activity evaluation by western blot analysis. A: Strained with Strep-680nm, B: stained with anti-MGMT antibody. C: overlaped.

18.1 Amplified region:

5'ggcaaacagcataagctggtcacccacaccagacctgacccaaccagctcccctgcttctggccacgtaacctgagaagggaatccctcctctctgaac
cccagcccacccaatgctccaggcctcctgggataccccgaagagtgagttgccaagcagtcaccccacagttggaggagaatccacccaaaaggcagcctg
gtagacagggtcggggtggcctctctgtggggtccaggccaagtaggtggcctggggcctctgggggatgcaggggaagggggatgcaggggaacggggatg
caggggaacggggctcagctgaagagcagagccaggaacccctgtagggaaggggcaggagagccaggggatgagatggtggacagggaaagggggac
agggaaagcctgagcgcctcctgggcttccaaggactcaaccagaagcccagagcagggccttagggaaagcgggacccctgctctggcgagggaatat
gtcccagatagcactgggactctttaaggaaagaagtaggagaagaagaggagtagagcggccacgacctggtgaacacctaggacgcaccattctc
acaaaggagtttccacacggacacccccctcctcaccacagccctgccaggacggggctggctactggccttatctcacaggtaaaaactgacgcacggagga
acaataaattggggactagaaggtgaagagccaaagttagaactcaggaccaacttattctgattttgttttccaaactgcttctccttgggaagtgaaggaag
ctgcagcaccaggatcagtgaaacgcaccagacagccgcgtcagagcagctcaggttctgggagagggtagcgcaggggtggcactgagaaccgggcaggtc
acgcatcccccttccctccccccctgccaagctctccctcccaggatcctctctggtccatcgtaaacacctaagaggttctggcaaggagagatggc
tccaggaaatgggggtgtgtcaccagataaggaatctgcctaacaggaggtgggggttagaccaatcagggagactaggaaggaggagcctaaggatggg
gcttttctgtaccaatcctgtccctagtaaagcttggtacccatagagcccaccgcacccccagcatcctgctattgtcttccaatcctccccctgctgctgcc
ccacccccccccagaatagaatgacacctactcagacaatgcgatgcaatttctcattttattaggaaggacagtgaggagtgccacctccagggtcaaggaa
ggcacgggggaggggcaacaacagatggctggcaactagaaggcacagggtacctcaggcaccgggcttgcgggtcatgcaccaggtgcgggtccttcg
ggcacctcgcagctggcggtgacgggtgaagccgagccgctcgtagaaggggaggttgcggggcgggaggtctccaggaaggcggcaccggcgct
cggccgctccactccggggagcagcagcggcgtccccagacccttgcctgggtgctggggcagacgccgacgggtggccaggaaccacgcgggctccttg
ggcgggtgcggcgccaggagccttccatctgtgctgcggccagccgggaaccgctcaactcggccatgcgcgggcccgatctcggcgaacaccgcccc
gcttcgacgctctccggcgtgtccagaccgccaccgcggcggcgtcgtccgcgaccacaccttccgatgctgagcccagcgcgctgaggaagagttctg
cagctcgggtgaccgctcagatgtggcggtccggatcgacggtgtggcgcgtggcggggtagtcggcgaacgcggcggcaggggtcgtacggccctgggga
cgtcgtcgcgggtggcagggcgcaccgtgggctgtactcggctatgtaagctcggccccactgtggggtggaggggacag-3'

Chapter 4

18.2 Sequenced data:

Using Primer: 5'-ggcaaacagcataagctggtc-3'

Nnncaaacccagctccctctgcttcttggccacgtaacctgagaagggaatccctctctctgaaccccagcccacccaatgctccaggcctctctgggatacccc
gaagagtgagtttgccaagcagtcaccccacagttggaggagaatccacccaaaaggcagcctgtagacagggctggggtggcctctctggtgggtccaggcc
aagtaggtggcctggggcctctgggggatgcaggggaaggggatgcaggggaacggggatgcaggggaacggggctcagctctgaagagcagagccagg
aacccctgtagggaagggcagagagccagggcatgagatggtggacgaggaagggggacagggaaacctgagcgcctctctgggcttccaaggact
caaaccagaagcccagagcagggccttagggaagcgggacctgctctggcggggaataatgtcccagatagcactggggactcttfaaggaaagaaggat
ggagaaagagaaagggagtagaggcggccacgacctggtgaacacctaggacgaccattctcacaagggagtttccacagcgacacccccctctacca
cagccctgccaggacgggctggtactgacctatctcacagtaaaactgacgcacggggaacaataataattggggactagaaggtgaagagccaaagt
agaactcaggaccaactattctgatttgttttccaaactgcttctctctgggaagtgaaggaagctgcagcaccaggatcagtgaaacgcaccagacagccgc
gtcagagcagctcaggttctgggagaggtagcgcagggtggccactgagaaccggcaggtcacgcacccccctctcccccacccctccaagctctcc
ctccaggatcctctctggtccatgtaagcaaaccttagaggttctggcaaggagagagatggctccaggaaatgggggtgtgtaccagataaggaatctgcc
taacaggaggtgggggttagaccaataatcaggagactaggaagggaggcctaaggatggggctttctgtcacaatcctgtccctagtaaagcttggtacc
atagagcccaccgcatccccagcatgcctgctwtg.

Using Primer 5'-gagcttaccatgaccgagtac-3'

Tccccctgcaagtctcctcccaggatcctctctggtccatcgtgaagcaaaccttagaggttctggcaaggagagagatggctccaggaaatgggggtgtgt
caccagataaggaatctgcctaacaggaggtgggggttagaccaatcagggagactaggaaggaggagcctaaggatggggctttctgtcaccatcctgt
ccctagtaaagcttggtacccatagagcccaccgcatccccagcatgcctgctattgttcccaatcctcccccttctgtctgccccacccccccccagaat
agaatgacacctactcagacaatgcgatgcaatttctcattttattaggaaggacagtgagggtggcacctccagggtcaaggaaggcacgggggaggggca
aacaacagatggctggcaactagaaggcacagggtacctcaggcaccgggctgctgggctatgcaccaggtgctgctcctcgggcacctcagctcggcgg
tgacggtgaagccagccgctcgtagaaggggaggttgcggggcggaggtctccaggaaggcgggacccccggcgcgctcggccgctccactccggg
gagcacgagggcgtgccagacccttgcctggtgctggggcagacgccgaggtggccaggaaccacgcccgtccttggccgggtgcccggcagga
ggccttccatctgttctgctgcccagccgggaaccgctcaactcggccatgctgcccggccgatctcggcgaacaccgccccgctcagctctccggcgtg
gtccagaccgccaccgcccgtctccgacccacacctgcccgatgctgagcccagcgcgtgaggaagagttctgacgtcggtagccgctcgtat
gtggcggtccggatcagcgggtggtgctgctggcgggtagtcggcgaacgcggcggcaggggtgctacggccctggggang.

19 Inhibitory study

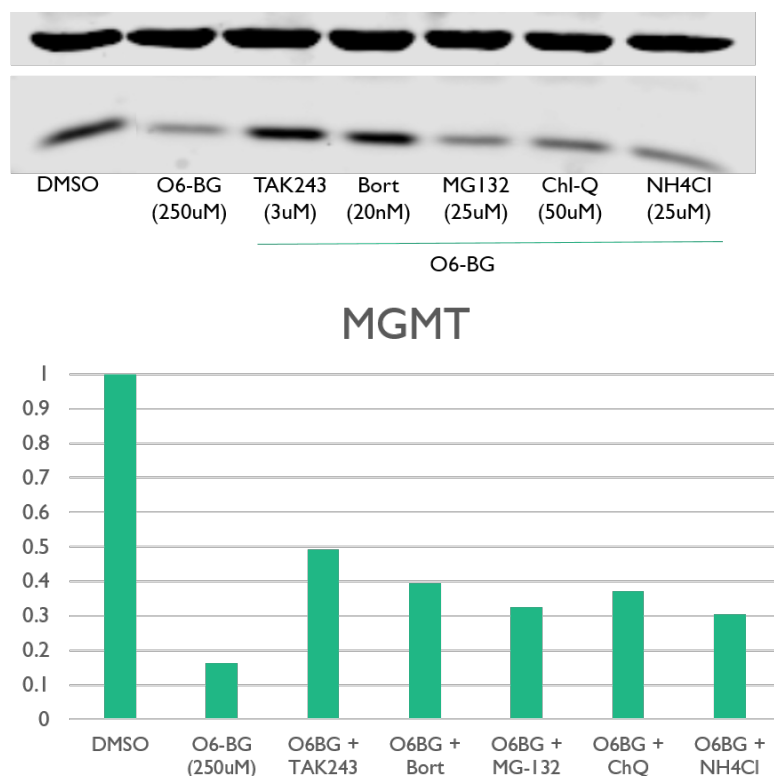
HUH6 cells or HEK293T cells with TurboID-13x-MGMT in AAVS1 were seeded in 12-well plates and treated at 80% confluency with either vehicle (DMSO), Bortezomib (20 nM, Activate Scientific), MG-132 (25 μ M, Adipogen), Chloroquine (50 μ M, Sigma Aldrich), O⁶-BG (250 μ M, Sigma Aldrich), TAK243 (3 μ M, MedKoo Biosciences, Inc.) or MLN-4924 (500 nM, MedKoo Biosciences, Inc.) for 4 hours. HEK cells were additionally treated with doxycycline (1 μ M) 16 hours prior to treatments. After that time O⁶-BG was added (250 μ M) and incubated for 6 more hours. Cells were washed three times with PBS, lysed with RIPA buffer and used for immunoblotting.

HUH6 cells were transfected with Silencer™ Pre-Designed siRNA RAD18 (Thermo Scientific, AM16708; siRNA ID: 28198) using GenJet transfecting agent (SignaGen Laboratories). Transfection efficiency was checked after 48 hours.

293T cells with stable expression of inducible TurboID-MGMT protein fusion had been infected with shRNA lentiviral constructs containing Cullin1, Cullin4A or Cullin4B shRNA, respectively (**Table TS2**). 48 hours post-

Experimental Procedures

infection cells were treated with doxycycline to induce expression of TurboID-MGMT. 12 hours post induction cells were treated with O⁶-BG (250 μM) for 24 hours, washed, lysed and used for immunoblotting.



S 23. MGMT degradation study by inhibition of UPP pathway (TAK243, Bort, MG132) and lysosomal pathway (Ch-Q and NH₄-Cl).

20 MGMT recovery and chase study

HUH6 cells had been seeded in 12-well plates and treated with vehicle or O⁶-BG (250 μM) for 24 hours, medium had been replenished in all wells and cells collected after 0.5, 1.5, 4, 7.5, 19.5, 21.5 or 23.5 hours. Cells had been lysed in RIPA buffer and used for immunoblotting.

21 TurboID Pulldowns

Protocol adopted from *Sears, R.M.; May, D.G.; Roux, K.J. Methods Mol. Biol. 2019, 2012, 299–313.*

Cells had been seeded in two 10 cm plates per condition. When reached 80% confluency 10 μM biotin had been added dropwise to target cells and placed back in the incubator for 10 minutes. Serum-containing medium had been removed and cells were washed twice with 5 mL of PBS and 540 μL of urea lysis buffer (8 M urea in 50 mM Tris pH 7.4 with Protease inhibitor and 1 mM DTT) was added to each plate. Cells from two plates were combined in 5 mL tubes and 20% Triton X-100 was added to final concentration of 1%. Cells were moved on ice and sonicated three times for 30s operation – 1 minute rest cycle (Sonicator: 30% duty, output level 3). After three cycles 1260 μL of lysis buffer was added, cells were sonicated for 1 minute and centrifuged for 10 minutes at 16500 x g at 4°C. Supernatant was transferred to new tube and incubated with Streptavidin Sepharose High Performance beads for 4 hours. Streptavidin beads were washed four times with 8 M urea 50 mM Tris pH 7.4 wash buffer and re-suspended in 50 mM ammonium bicarbonate with 1 mM biotin. DTT and iodoacetamide (IAA) was added to 20 mM and 15 mM final concentration, respectively. Beads were incubated for 1 hour with shaking,

Chapter 4

peptides were digested by adding 10 µg/mL trypsin and incubated at 37°C for 12 hours. Samples were centrifuged and supernatant was transferred to new tube. Beads were washed once with 50 mM ammonium bicarbonate with 1 mM biotin and combined supernatants were purified on Pierce C18 -100 µl Tips (ThermoFisher Scientific) following the manufacturer guide. Eluted sample was dried under vacuum and re-suspended in 0.1% aqueous formic acid and subjected to LC–MS/MS analysis.

22 TMT labelling and LC MS/MS analysis

Sample aliquots comprising 25 µg of peptides were labeled with isobaric tandem mass tags (TMT 10-plex, Thermo Fisher Scientific) as described previously (PMID:27345528). Shortly, peptides were resuspended in 20 µL labeling buffer (2 M urea, 0.2 M HEPES, pH 8.3) and 5 µL of each TMT reagent were added to the individual peptide samples followed by a 1 hour incubation at 25°C, shaking at 500 rpm. To control for ratio distortion during quantification, a peptide calibration mixture consisting of six digested standard proteins mixed in different amounts was added to each sample before TMT labeling (for details see PMID:27345528). To quench the labelling reaction, 1.5 µL aqueous 1.5 M hydroxylamine solution was added and samples were incubated for another 10 minutes at 25°C shaking at 500 rpm followed by pooling of all samples. The pH of the sample pool was increased to 12 by adding 1 M phosphate buffer (pH 12) and incubated for 20 minutes at 25°C shaking at 500 rpm to remove TMT labels linked to peptide hydroxyl groups. Subsequently, the reaction was stopped by adding 2 M hydrochloric acid until a pH < 2 was reached. Finally, peptide samples were further acidified using 5% TFA, desalted using Sep-Pak Vac 1cc (50 mg) C18 cartridges (Waters) according to the manufacturer's instructions and dried under vacuum.

TMT-labeled peptides were fractionated by high-pH reversed phase separation using a XBridge Peptide BEH C18 column (3,5 µm, 130 Å, 1 mm x 150 mm, Waters) on an Agilent 1260 Infinity HPLC system. Peptides were loaded on column in buffer A (20 mM ammonium formate in water, pH 10) and eluted using a two-step linear gradient from 2% to 10% in 5 minutes and then to 50% buffer B (20 mM ammonium formate in 90% acetonitrile, pH 10) over 55 minutes at a flow rate of 42 µl/min. Elution of peptides was monitored with a UV detector (215 nm, 254 nm) and a total of 36 fractions were collected, pooled into 12 fractions using a post-concatenation strategy as previously described (PMID:21500348) and dried under vacuum.

Dried peptides were resuspended in 0.1% aqueous formic acid and subjected to LC–MS/MS analysis using a Orbitrap Fusion Lumos Mass Spectrometer fitted with an EASY-nLC 1200 (both Thermo Fisher Scientific) and a custom-made column heater set to 60°C. Peptides were resolved using a RP-HPLC column (75 µm × 36 cm) packed in-house with C18 resin (ReproSil-Pur C18–AQ, 1.9 µm resin; Dr. Maisch GmbH) at a flow rate of 0.2 µLmin⁻¹. The following gradient was used for peptide separation: from 5% B to 15% B over 9 minutes to 30% B over 90 minutes to 45 % B over 21 minutes to 95% B over 2 minutes followed by 18 minutes at 95% B. Buffer A was 0.1% formic acid in water and buffer B was 80% acetonitrile, 0.1% formic acid in water.

The mass spectrometer was operated in DDA mode with a cycle time of 3 seconds between master scans. Each master scan was acquired in the Orbitrap at a resolution of 120,000 FWHM (at 200 m/z) and a scan range from 375 to 1600 m/z followed by MS2 scans of the most intense precursors in the Orbitrap at a resolution of 30,000 FWHM (at 200 m/z) with isolation width of the quadrupole set to 1.1 m/z. Maximum ion injection time was set to 50 ms (MS1) and 54 ms (MS2) with an AGC target set to 1e6 and 1e5, respectively. Only peptides with charge state 2 – 7 were included in the analysis. Monoisotopic precursor selection (MIPS) was set to Peptide, and the Intensity Threshold was set to 5e4. Peptides were fragmented by HCD (Higher-energy collisional dissociation)

Experimental Procedures

with collision energy set to 38%, and one microscan was acquired for each spectrum. The dynamic exclusion duration was set to 45 seconds.

The acquired raw-files were converted to the mascot generic file (mgf) format using the msconvert tool (part of ProteoWizard, version 3.0.4624 (2013-6-3)) and searched using MASCOT against a human database (consisting of 40832 forward and reverse protein sequences downloaded from Uniprot on 20190307), the six calibration mix proteins (PMID:27345528) and 392 commonly observed contaminants. The precursor ion tolerance was set to 10 ppm and fragment ion tolerance was set to 0.02 Da. The search criteria were set as follows: full tryptic specificity was required (cleavage after lysine or arginine residues unless followed by proline), 3 missed cleavages were allowed, carbamidomethylation (C) and TMT6plex (K and peptide N-terminus) were set as fixed modification and oxidation (M) as a variable modification. Next, the database search results were imported into the Scaffold Q+ software (version 4.3.2, Proteome Software Inc., Portland, OR) and the protein false discovery rate was set to 1% based on the number of decoy hits. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters. Acquired reporter ion intensities in the experiments were employed for automated quantification and statistical analysis using a modified version of our in-house developed SafeQuant R script (v2.3, PMID:27345528). This analysis included adjustment of reporter ion intensities, global data normalization by equalizing the total reporter ion intensity across all channels, summation of reporter ion intensities per protein and channel, calculation of protein abundance ratios and testing for differential abundance using empirical Bayes moderated t-statistics. Finally, the calculated p-values were corrected for multiple testing using the Benjamini–Hochberg method.

23 shRNA sequence

Vector	Vector Type	Full Hairpin Sequence	Sense Sequence	Human Gene IDs	Human Gene Descriptions
plKO.1	Lentiviral	CCGGCCATGATATGTGGTCTAAGAAGCTCGAGTCTTTAGACCACATATCATGGTTTTT	CCATGATATGTGGTCTAAGAA	8451	culin 4A
plKO.1	Lentiviral	CCGGGCAGGTGTATAAAGATTCATTTCTCGAGAAATGAATCTTTATACACCTGGCTTTT	GCAGGTGTATAAAGATTCAAT	8451	culin 4A
plKO.1	Lentiviral	CCGGGCCAAAGGTTAATGCGAGGAAACCTCGAGTTTCCCTGCATTAACCTTTGGCTTTT	GCCAAAGGTTAATGCGAGGAAA	8451	culin 4A
plKO.1	Lentiviral	CCGGGCAGAACTGATCGCAAAGCATCTCGAGATGGCTTTGGCATCAGTTCTGCTTTT	GCAGAAGCTGATCGCAAAAGCAT	8451	culin 4A
plKO.1	Lentiviral	CCGGCCAGAATATTCCTTAACCATGTACTCGAGTACATGGTTAAGATATTTCTGGTTTTT	CCAGAATATCTTAACCATGTA	8451	culin 4A
plKO.1	Lentiviral	CCGGCCATGAAAGAAGCATTTGAACCTCGAGTTCAAAATGCTTCTTTCAATGGCTTTTT	GCCATGAAAGAAGCATTTGAA	8450	culin 4B
plKO.1	Lentiviral	CCGGCCTGAATATCTACATCATGTTCTCTCGAGAAACATGATGTAGATATTCAGGTTTTT	CCTGAATATCTACATCATGTT	8450	culin 4B
plKO.1	Lentiviral	CCGGCCGGACTACATGGAAAGAGATCTCGAGATCTCTTTCCATGTAGTCCCGTTTTT	CGGACTACATGGAAAGAGAT	8450	culin 4B
plKO.1	Lentiviral	CCGGGCAGAAITTTAAAGAGGGTAAACCTCGAGTTTACCCCTTTAAATTCCTGCTTTTT	GCAGAAITTTAAAGAGGGTAAA	8450	culin 4B
plKO.1	Lentiviral	CCGGCAATTTCTTCAGAAAGGTTTACTCGAGTAAACCCTTTCTGAAGAATTGCTTTTT	GCAATTTCTCAGAAAGGTTTA	8450	culin 4B
plKO.1	Lentiviral	CCGGGCACACAAAGATGAATTAGCAACTCGAGTTGCTAATTCATCTTGTGTGCTTTTT	GCACACAAAGATGAATTAGCAA	8454	culin 1
plKO.1	Lentiviral	CCGGCGTGTATATCAAGTTGTCTACTCGAGTAGACAACCTGATATAACCAACGTTTTT	CGTGGTATATCAAGTTGTCTA	8454	culin 1
plKO.1	Lentiviral	CCGGGCCACAGCATGATCTCCAAAGTTACTCGAGTAACTTGGAGATFCATGCTGGCTTTTT	GCCACAGCATGATCTCCAAGTTA	8454	culin 1
plKO.1	Lentiviral	CCGGGATTTGATGAGATGAGAGTGTACTGAGTAGACTCTCATCCAAATCTTTTT	GATTTGATGAGATGAGAGTGTA	8454	culin 1
plKO.1	Lentiviral	CCGGCCCCGACGCAAAATAGTTTCATGTTCTCGAGACATGAACCTAFTTGTGGGGTTTTT	CCCCGACGCAAAATAGTTTCATGT	8454	culin 1

TS 2. shRNA sequences of *CUL1*, *CUL4A*, *CUL4B*

Malgorzata (Gosia) Murawska

Curriculum Vitae



Personal Information

Email: m.m.h.murawska@gmail.com

ORCID: [0000-0003-1822-9607](https://orcid.org/0000-0003-1822-9607)

Twitter: [@mu_gosia](https://twitter.com/mu_gosia)

Education

- 04.2017 – 04.2021 **PhD, Chemical Biology.** University of Basel; Chemistry Department, Basel, Switzerland.
Supervisor: Prof. Dennis Gillingham
Project: Targeted protein degradation using small molecule inducers in mammalian cells.
Acquired competences and techniques: cloning, protein expression and purification, bacterial and mammalian cell culture handling, synthetic chemistry, small molecule activity screen.
- 03.2020 – 09.2020 **Scientific Collaboration,** University of California at Irvine; School of Medicine, Irvine, California, United States.
Supervisor: Prof. Peter Kaiser
Project: Chemical Tailoring of RNA and Vesicles to Create New Drugs.
Acquired competences and techniques: CRISPR-Cas9 gene knock-ins, shRNA gene silencing, E3-ligase screening.
- 05.2016 **MSc, Organic Chemistry.** Stratingh Institute, University of Groningen, The Netherlands.
Supervisor: Prof. Ben L. Feringa
Project: (I) Synthesis of overcrowded alkene switch for modification of peptides; (II) Synthesis of a photo-switchable AT-hook useful for photo-controlling DNA-peptide binding interactions.
Acquired competences and techniques: multi-step organic synthesis, solid phase peptide synthesis, photochemistry, isothermal titration calorimetry, fluorescence displacement titrations assays, NMR, HPLC, UV-Vis spectroscopy, fluorescence spectroscopy.
- 08.2014 **BSc, Chemistry.** University of Warsaw, Poland.
Project: Isolation and characterization of humic acids of natural and chemical origin.
Acquired competences and techniques: organic synthesis, electron microscopy, titrimetry.
- 06.2013 – 09.2013 **Internship.** University of Groningen, The Netherlands.
Supervisors: Prof. W. Szymanski, Prof. D. Janssen.
Project: Process optimization for a haloalkane dehalogenase-catalyzed kinetic resolution of racemic haloamides.
Acquired Competences and techniques: HPLC.

Chapter 4

Employment History

- 04.2017 – 04.2021 **PhD student, Chemical Biology.** University of Basel; Chemistry Department, Basel, Switzerland.
Supervisor: Prof. Dennis Gillingham
- 01.2016 – 02.2017 **Research Assistant, Plant Biology.** Lawrence Berkeley National Laboratory; Joint BioEnergy Institute Feedstocks Division; Berkeley, California, United States.
Supervisor: Dr. Jenny Mortimer
Project: Investigating the role of lipid glycosylation in cell wall biosynthesis.
Acquired competences and techniques: cell membrane lipid extraction and purification, MALDI spectroscopy, analytical thin layer chromatography, HPLC.
- 07.2015 – 11.2015 **Internship, Organic Chemistry.** Novartis Institute for BioMedical Research, Basel, Switzerland.
Supervisor: Dr. Siem Veenstra.
Project: Synthesis of drug precursor and enzyme inhibitors.
Acquired competences and techniques: organic synthesis, liquid chromatography - mass spectroscopy, thin layer chromatography, NMR.

Approved Research Projects

I have been awarded a mobility grant by SNSF in my PhD project 200021_172521 “*Chemical Tailoring of RNA and Vesicles to Create New Drugs*”, which allowed me to establish and execute a successful collaboration at the University of California in Irvine, USA.

Supervision of junior researchers

- 04.2019 – 05.2019 Chemistry Bachelor student from University of Basel: Seyid Colak
- 10.2018 – 01.2019 Pharmacology Master Student from University of Antwerp: Jyotika Bhatta.
- 05.2018 – 08.2018 Chemistry Master Student from Université de Haute-Alsace: David Vidal
- 01.2018 – 03.2018 Chemistry Bachelor student from University of Basel: Werner Jauslin

Teaching activities

- 06.2019 **Teaching** experimental work of a group of 10 students in 2-weeks intensive Organic Chemistry practical course, University of Basel.
Daily: guiding and supervising students work, checking reports and lab journal, grading.
- 09.2018 – 02.2019 **Teaching** experimental work of a group of 12 students in 4 months Organic Chemistry practical course, University of Basel.
Daily: guiding and supervising students work, checking reports and lab journal, grading.
Weekly: conducting seminars and exams, exam grading.

- 07.2018 **Teaching** experimental work of a group of 12 students in 2-weeks intensive Organic Chemistry practical course, University of Basel.
Daily: guiding and supervising students work, checking reports and lab journal, grading.
- 09.2015 – 11.2015 **Student – assistant.** Assisting, experimental work of a group of 6 students in 6 weeks intensive Organic Chemistry practical course, University of Groningen.
Daily: guiding and supervising students work, checking and grading reports and lab journals.
- 05.2013 – 07.2013 **Student – assistant.** Assisting, experimental work of a group of 6 students in 6 weeks intensive Organic Chemistry practical course, University of Groningen.
Daily: guiding and supervising students work, checking and grading reports and lab journals.

Prizes, awards, fellowships

- | | | |
|-------------------|---|------------|
| 10.2019 | Swiss National Science Foundation SNSF, <i>Mobility in project</i> , Basel, Switzerland | (8250 CHF) |
| 01.2016 | Groninger Universiteits Fonds Scholarship, Groningen, The Netherlands | (400 €) |
| 11.2015 | Backer Foundation Scholarship, Groningen, Netherlands | (2000 €) |
| 09.2015 | Marco Polo Scholarship, Groningen, The Netherlands | (1200 €) |
| 05.2014 | Fundusz Stypendialny KIK, Warsaw, Poland | (470 €) |
| 09.2013 – 07.2014 | European Union <i>Lifelong Learning</i> Program – Erasmus, Warsaw, Poland | (2700 €) |

Personal skills

Languages: Fluent in English and Polish. Basic proficiency in German.

Transferable Skills, complete courses: *Writing Productivity, Negotiation Skills for Scientists, Critical Reasoning and Logic, Intuitive Decision Making*

Digital competences: proficient in *Origin, ChemDraw, CorelDRAW, PyMOL* programs as well as the whole *Microsoft Office Suite*.

Publication list

- B. Jing, T. Ishikawa, N. Soltis, N. Inada, Y. Liang, **Gosia M. Murawska**, L. Fang, F. Anderbehan, V. Pidatala, X. Yu, E. Baidoo, M. Kawai-Yamada, D. Kliebenstein, D. Loque, P. Dupree, J. Mortimer, "Plants with altered sphingolipid glycosylation have increased resistance to biotrophic but not necrotrophic pathogens", *Plant Direct*, **2021**, 5, e00309. [DOI: 10.1002/pld3.309](https://doi.org/10.1002/pld3.309)
- **Gosia M. Murawska**, C. Poloni, N. A. Simeth, W. Szymanski and B. L. Feringa, "Comparative Study of Photoswitchable Zinc-Finger Domain and AT-Hook Motif for Light-Controlled Peptide-DNA Binding", *Chem. Eur. J.*, **2019**, 25, 4965–4973. [DOI: 10.1002/chem.201900090](https://doi.org/10.1002/chem.201900090).
- D. Yan, S. R. Yadav, A. Paterlini, W. Nicolas, J. Pettit, L. Brocard, I. Belevich, M. Grison, A. Vatén, L. Karami, S. el-Shawk, J.-Y. Lee, **Gosia M. Murawska**, J. Mortimer, M. Knoblauch, E. Jokitalo, J. Markham, E. Bayer and Y. Helariutta, "Sphingolipid biosynthesis modulates plasmodesmal ultrastructure and phloem unloading", *Nat. Plants*, **2019**, 5, 604–615. [DOI: 10.1038/s41477-019-0429-5](https://doi.org/10.1038/s41477-019-0429-5)

Chapter 4

- L. Fang, T. Ishikawa, E. A. Rennie, **Gosia M. Murawska**, J. Lao, J. Van, A. YL Tsai, E. E. K. Baidoo, J. Xu, J. D. Keasling, M. Kawai-Yamada, T. Demura, H. V. Scheller and J. C. Mortimer, “Loss of inositol phosphorylceramide sphingolipid mannosylation induces plant immune responses and reduces cellulose content in Arabidopsis”, *The Plant Cell*, **2016**, 28, 2991–3004. [DOI:10.1105/tpc.16.00186](https://doi.org/10.1105/tpc.16.00186).
- Acknowledged for analytical and quantitative method development: R. J. Floor, H. J. Wijma, D. I. Colpa, A. Ramos-Silva, P. A. Jekel, W. Szymanski, B. L. Feringa, S. J. Marrink and D. B. Janssen, “Computational Library Design for Increasing Haloalkane Dehalogenase Stability”, *ChemBioChem*, **2014**, 15, 1660–1672. [DOI: 10.1002/cbic.201402128](https://doi.org/10.1002/cbic.201402128)
- B. Jing, T. Ishikawa, N. Soltis, N. Inada, Y. Liang, **Gosia M. Murawska**, F. Andeberhan, R. Pidatala, X. Yu, E. Baidoo, M. Kawai-Yamada, D. Loque, D. J. Kliebenstein, P. Dupree and J. C. Mortimer, «GONST2 transports GDP-Mannose for sphingolipid glycosylation in the Golgi apparatus of Arabidopsis», *bioRxiv* **2018**, 346775. [DOI: 10.1101/346775](https://doi.org/10.1101/346775).
- D. Joss, J. Vallapurackal, A. Baiyoumy, V. Chimisso, F. Christoffel, P. Herr, C. Meyer, D. Miarzlou, **Gosia M. Murawska**, N. Niggli, K. Reznikova, I. Urosev, P. van Gerwen “Christmas Symposium Basel December 6, 2019”, *Chimia (Aarau)*, **2020**, 74 (1), 63. [DOI: 10.2533/chimia.2020.63](https://doi.org/10.2533/chimia.2020.63).
- D. Joss, S. Rigo, F. Christoffel, P. Herr, **Gosia M. Murawska**, B. Sauter, A. Stampfli, I. Urosev, J. Vallapurackal “Conference Report”, *Chimia (Aarau)*, **2019**, 73 (1), 98. [DOI: 10.2533/chimia.2019.98](https://doi.org/10.2533/chimia.2019.98).