

University of Oxford



Nuffield Department of Medicine

Target Discovery Institute
Structural Genomics Consortium

**Investigating strategies to modify PARP14 function
through macrodomain inhibition**

Marion Schuller

St Edmund Hall
Systems Approaches to Biomedical Science
Industrial Doctorate Centre

A thesis submitted for the degree of
Doctor of Philosophy

Michaelmas Term 2017

Investigating strategies to modify PARP14 function through macrodomain inhibition

Marion Schuller, St Edmund Hall
D.Phil. Thesis, Michaelmas 2017

Abstract

Macrodomains are conserved protein interaction modules that are present in all domains of life and mediate recognition of sequence motifs harbouring adenosine diphosphate ribose (ADPR) modifications. In addition, some of them are able to control the turnover of ADPR signalling through their catalytic activity removing these modifications. Macrodomains are hence implicated in a variety of cellular processes as well as in diseases including cancer and viral pathogenesis. The polyadenosine-diphosphate-ribose polymerase (PARP) family member PARP14 is one of twelve human macrodomain-containing proteins; it contains three macrodomains in addition to its catalytic PARP domain. PARP14 was shown to be involved in several cellular processes linked to cancer development in for example B-cell lymphoma and hepatocellular carcinoma. Therefore PARP14 including its macrodomains has emerged as a potential therapeutic target. However, the lack of specific small molecule inhibitors has hampered domain-specific target validation studies so far. Current approaches focus on inhibitor development for its PARP domain, yet attaining selectivity of these inhibitors over other PARP enzymes has been challenging.

The aim of the work described in this thesis was therefore to evaluate the possibility of targeting PARP14 via inhibition of its macrodomains with small molecule inhibitors as an alternative to PARP domain inhibition. These studies revealed that displacing PARP14 from its target sites requires inhibitor development for both its second and third macrodomain. Druggability of both macrodomains has been demonstrated by the highly selective allosteric macrodomain 2 inhibitor GeA-69 and by several fragment hits targeting the ADPR binding site of macrodomain 3. Finally, PARP14 was confirmed to be implicated in DNA repair mechanisms protecting cells against replication stress, suggesting that a dual PARP14 macrodomain inhibitor may provide the possibility to potentiate genotoxic chemotherapy for cancer treatment.

This thesis and the work presented in it were completed at the University of Oxford (Nuffield Department of Medicine) between October 2014 and December 2017 under supervision of Professor Dr. Stefan Knapp, Professor Dr. Benedikt M. Kessler and Dr. Jonathan M. Elkins, with Dr. Sandra Jacobs and Dr. Clemens Scheufler (NIBR, Basel, Switzerland) as industrial supervisors. The 45,000 compound library screen against PARP14 macrodomain 2 was performed at Novartis Institute for BioMedical Research (NIBR, Basel, Switzerland). All contributions made by others have been clearly stated in the text. This work has not been submitted in full or partial fulfilment of any other degree.

Oxford, 20 December 2017

Marion Schuller

I would like to express special appreciation and thanks to

Professor Dr. Stefan Knapp and **Professor Dr. Benedikt Kessler**

for providing me with the opportunity to undertake my doctoral research under their supervision. I am very grateful for all their support in any academic matter, their patience in developing my scientific skills, and their guidance throughout my studies. I greatly appreciate my freedom conceptualising parts of this study and their continuous support of my ideas, which gave me confidence in my work and myself.

Furthermore, I would very much like to thank **Dr. Jonathan Elkins** for his constant support as well as guidance and encouragement throughout my time at the SGC. I am grateful for his generous time for my project including providing training and proof-reading of this thesis. It is the numerous discussions that were educational and often inspiring, that I will truly miss.

I also very much like to thank **Dr. Sandra Jacob** and in particular **Dr. Clemens Scheufler** from Novartis Institute for BioMedical Research (NIBR, Basel) for supporting my D.Phil. studies. I was very fortunate to get the possibility to perform a part of my project at Novartis and to gain insights into the pharmaceutical industry which was a great experience. I am also grateful to **Dr. Hans Widmer** for coordinating my project between Novartis and the University of Oxford.

I am further indebted to the Engineering and Physical Sciences Research Council (**EPSRC**) and **Novartis Pharma AG** for providing through the Systems Approaches to Biomedical Science – Industrial Doctorate Centre (**SABS-IDC**) financial support of the project and particularly my personal funding which without would have been impossible for me to pursue these studies.

Very special thanks goes to everyone including the facility teams in the **Target Discovery Institute** and **Structural Genomics Consortium** for providing such a friendly, helpful and inspiring atmosphere which made it an exceptional place I very much enjoyed working over the few years. In particular, I like to thank all **members of the Kessler group** for welcoming me to their group and being my scientific family, with special thanks to Dr. Adan Pinto Fernandez and Alia Komsany for their helpful discussions and assistance in the metabolomic studies. I am very grateful to the **Knapp group**, in particular to Chela Nunez Alonso and Eidarus Salah, for all their help and support – personally as well as scientifically. I also like to thank **Dr. Kilian Huber** and

group members for providing the lab space and being so wonderful and helpful colleagues, with special thanks to Karin Oleg, Paulina Siejka, and Dr. Sarah Blair-Reid for their training in CRISPR cell line editing. Furthermore, I like to take this opportunity to thank **Dr. Oleg Fedorov and group members** for their support, helpful discussions, and assistance in all kind of screening matters. I am also very grateful to **Dr. Jörg Trappe**, Sandra Hodel, and Inga Galuba (NIBR, Basel) for their great help and efforts performing the NDDI compound library screen against PARP14 MD2. Moreover, I would very much like to thank **Professor Dr. Franz Bracher** and **Professor Dr. Paul Brennan** for their collaborations, in particular **Dr. Kerstin Riedel** and **Moses Moustakim** for synthesising the PARP14 MD2 inhibitors with such great efforts. Further, I thank Dr. Marina Tintelnot (NIBR, Basel) for her time on PARP14 MD2 inhibitor design and Dr. Vincent Fagan for his help with characterisation of *iso*-ADPR. Very special thanks goes to the SGC crystallography **group of Professor Dr. Frank v. Delft**, in particular **Dr. Romain Talon**, for their training, help, and efforts for performing the crystallographic fragment screen. The screening day at Diamond Light Source is one of the most unforgettable days of my D.Phil. time. Unforgettable is also the adventure of cloning full-length PARP14 for which I like to thank **Dr. Pavel Savitsky**. It was a pleasure and honour to learn the art of cloning from you! Moreover, I would like to thank the Cell-based HTS group of **Dr. Daniel Ebner**, especially **Dr. Val Millar**, for all their support and efforts developing image analysis algorithms along with assistance in microscopy matters, which was essential for establishing the migration and adhesion assays.

I am also very thankful to **Dr. Ivan Ahel** and **Dr. Ian Gibbs-Seymour** for performing the laser micro-irradiation experiments and their helpful discussions. I further owe great thanks to **Professor Dr. Madalena Tarsounas and group members**, in particular **Dr. Xianning Lai**, for their training and guidance in performing the DNA damage experiments.

Finally, I like to express my sincere gratitude to my parents and my sister Karin for their constant support, encouragement, and understanding over all the years of my studies. I am further very grateful to my other family members and friends who have supported me along the way. This work is also in memory of my grandparents who could never see its completion, and of Herbert Marcinek and his family.

ABSTRACT	i
DECLARATION	ii
ACKNOWLEDGEMENTS	iii
TABLE OF CONTENTS	v
TABLE OF ABBREVIATIONS	ix
I. INTRODUCTION	1
1. Protein ADP-ribosylation – an ancient and versatile post-translational modification	1
2. Macrodomains and WWE domains as readers of protein ADP-ribosylation	5
2.1. WWE domains	5
2.2. Macrodomains	6
3. The human macrodomain-containing protein family and its diverse biological functions	8
3.1. The macrodomain-containing histone variant macroH2A	9
3.2. MacroD-type macrodomains	10
3.3. ALC1, PARG and TARG1	10
3.4. MacroPARPs	11
3.5. PARP14 – a triple macrodomain-containing protein with multiple functions	12
4. Human macrodomain-containing proteins in disease	16
5. The therapeutic potential of macrodomain targeting and current progress	17
6. Aim of this thesis	19
II. RESULTS AND DISCUSSION	20
1. PARP14 sequence recognition by its macrodomains and WWE domain	20
1.1. PARP14 is recruited to ADP-ribosylated sites mainly through its macrodomains	20
1.2. All three PARP14 macrodomains contribute to PARP14 recruitment in cells	26
2. Targeting PARP14 macrodomains	38
2.1. Inhibitor development for PARP14 MD2	38
2.1.1. GeA-69 – a new class of macrodomain inhibitors targeting PARP14 MD2	38
2.1.2. GeA-69 selectivity over PARP15 MD1	50
2.1.3. SAR of the GeA-69 lead compound	57
2.1.4. Finding new PARP14 MD2 inhibitors – screening experience with Novartis using a public 45,000 compound library	67
2.2. A FBDD approach using X-ray crystallography for targeting PARP14 MD3	81
2.2.1. Performance of a crystallographic fragment screen against PARP14 MD3	82
2.2.2. Identification and validation of PARP14 MD3 fragment hits	91

3. Exploring pharmacological targeting strategies for PARP14.....	104
3.1. PARP14 knock-down and knock-out in HeLa and U-2 OS cell lines.....	104
3.2. Role of PARP14 in cell adhesion and migration processes.....	114
3.2.1. Testing PARP14 importance for cell adhesion.....	114
3.2.2. Characterisation of the migratory potential of PARP14 KO cells.....	117
3.3. PARP14 and the Warburg effect – a brief study on PARP14 KO HeLa cells.....	122
3.4. Role of PARP14 in DNA damage response.....	130
III. FUTURE DIRECTIONS.....	142
IV. EXPERIMENTAL PROCEDURES.....	143
1. General reagents.....	143
2. Oligonucleotides.....	143
3. Plasmids.....	146
4. Antibodies.....	151
5. Work with nucleic acids.....	152
5.1. DNA preparation.....	152
5.2. mRNA preparation.....	152
5.3. Agarose gel electrophoresis.....	152
5.4. Cloning.....	153
5.4.1. Gateway® cloning of mammalian expression constructs.....	153
5.4.2. Ligation-independent cloning for prokaryotic expression constructs.....	153
5.4.3. Site-directed mutagenesis.....	154
5.4.4. Cloning of sgRNAs into vector for co-expression with Cas9.....	154
6. Work with eukaryotic cells (cell culture).....	155
6.1. Eukaryotic cell lines.....	155
6.2. General cell culture techniques.....	156
6.3. Transient transfection.....	157
6.3.1. siRNA transfection.....	157
6.3.2. Transfection of mammalian expression constructs.....	157
6.4. Generation of PARP14 KO cell lines using the CRISPR-Cas9 system.....	158
6.4.1. PARP14 sgRNA design for CRISPR genome editing.....	158
6.4.2. Functional validation of PARP14 sgRNAs in HEK293 cells.....	158
6.4.3. Transfection, selection and clonal expansion of HeLa and U-2 OS cells.....	158
6.4.4. Validation of PARP14 edits by gDNA analysis.....	159
6.4.5. Validation of PARP14 edits by mRNA analysis.....	159

7. Work with prokaryotic cells.....	160
7.1. Bacteria strains.....	160
7.2. Preparation of competent cells.....	160
7.3. Bacterial transformation.....	161
7.4. Expression of recombinant proteins.....	161
8. Work with proteins.....	162
8.1. Purification of His ₆ -, GST- or biotin fusion proteins.....	162
8.2. Denaturing discontinuous polyacrylamide gel electrophoresis (SDS-PAGE).....	162
8.3. Mass spectrometric analysis of purified proteins.....	162
8.4. Western blotting and immunodetection of proteins.....	163
8.5. Macrodomein pull-down assay procedure.....	163
8.6. Cell fractionation.....	164
8.7. Protein crystallisation.....	165
9. PARP14 MD3 fragment screen.....	165
9.1. PARP14 MD3 <i>apo</i> protein crystallisation.....	165
9.2. Fragment soaking and data collection.....	166
9.3. PARP14 fragment screen - data processing.....	166
10. Biophysical assays for compound screening and characterisation.....	166
10.1. Macrodomein AlphaScreen assay.....	166
10.2. HTRF (Homogenous Time-Resolved Fluorescence) assay.....	167
10.3. Differential scanning fluorimetry (T _m -shift assay).....	168
10.4. NanoDSF (nano differential scanning fluorimetry).....	168
10.5. Biolayer interferometry (BLI).....	168
10.6. Isothermal titration calorimetry (ITC).....	169
10.7. Dynamic light scattering (DLS).....	169
11. Cellular assays.....	170
11.1. Cellular thermal shift assay (CETSA).....	170
11.2. Cell adhesion assay.....	171
11.3. Scratch migration assay.....	171
11.4. Cellular proliferation assays.....	172
11.5. Clonogenic survival assay.....	173
11.6. Seahorse experiments.....	173
11.7. Immunofluorescence imaging.....	174
11.8. Laser micro-irradiation assay.....	175
12. Generation of <i>iso</i> -ADPR.....	175

V. APPENDIX	177
1. Supplementary information chapter 1	177
2. Supplementary information chapter 2	180
3. Supplementary information chapter 3	204
4. Supplementary experimental details	218
4.1. Vector maps	218
4.2. Supplementary experimental procedures	219
VI. REFERENCES	223

Å	Ångström (10^{-10} metre)	Da	Dalton
A ₂₅₈	absorption at 258 nm	DAPI	4',6-diamidino-2-phenylindole
aa	amino acid	DDR	DNA damage response
ADB	antibody dilution buffer	2-DG	2-deoxyglucose
ADP	adenosine diphosphate	DLBCL	diffuse large B cell lymphoma
ADPR	adenosine diphosphate ribose	DLS	dynamic light scattering
ALC1	amplified in liver cancer 1	DMEM	Dulbecco's modified Eagle's medium
ALPHA	amplified luminescent proximity homogeneous assay	DMSO	dimethylsulphoxide
AMP	adenosine monophosphate	DNA	deoxyribonucleic acid
ANOVA	analysis of variance	DR-GFP	direct repeat GFP
Ara-C	cytosine β -D-arabinofuranoside	DSB	double-strand break
ATCC	American Type Culture Collection	DSF	differential scanning fluorimetry
ATM	ataxia telangiectasia mutated	DTT	dithiothreitol
ATP	adenosine triphosphate		
AU	absorbance unit		
		EC ₅₀	half maximal effective concentration
53BP1	p53-binding protein 1	ECAR	extracellular acidification rate
BAL	B-aggressive lymphoma	ECL	enhanced chemoluminescence
Bis-Tris	2-[Bis(2-hydroxyethyl)amino]-2-(hydroxymethyl)propane-1,3-diol	<i>E. coli</i>	Escherichia coli
BLAST	basic local alignment search tool	EDTA	ethylenediaminetetraacetic acid
BLI	biolayer interferometry		
bp	base pair(s)	EGTA	ethyleneglycol-bis(2-aminoethylether)- <i>N,N,N',N'</i> -tetraacetic acid
BRD9	Bromodomain containing 9		
BSA	bovine serum albumin	EMEM	Eagle's Minimum Essential Medium
		ESI-TOF MS	electron spray ionization time-of-flight mass spectrometry
°C	degree celsius	<i>et al.</i>	et alii
c	concentration		
cal	calorie, 1 cal = 4.184 J		
Cas9	CRISPR associated protein 9	FBS	fetal bovine serum
cDNA	complementary deoxyribonucleic acid	FBDD	fragment-based drug discovery
CETSA	cellular thermal shift assay	FL	full-length
		FLAG	DYKDDDDK epitope
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-2-hydroxy-1-propanesulfonate		
clog P	calculated log P	g	gram
CMFDA	5-chloromethylfluorescein diacetate	g	gravitation constant ($g = 9.81 \text{ m/s}^2$)
CPT	camptothecin	gDNA	genomic DNA
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats	GF	gel filtration
		GFP	green fluorescence protein
		GSH	glutathione
		GST	glutathione-S-transferase
		Gy	Gray; unit of absorbed dose of IR
ΔH	change in binding enthalpy		
ΔP	difference in potential energy	h	hour(s)
ΔS	change in entropy	HEK	human embryo kidney

HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid	m	milli; 10^{-3}
		M	molar (mol)
HPLC	high-performance liquid chromatography	M	MW marker for SDS-PAGE
		<i>m/z</i>	mass to charge ratio
HR	homologous recombination	MARylation	mono-ADP-ribosylation
HRP	horseradish peroxidase	MD	macrodomain
HTRF	homogenous time-resolved fluorescence	MES	2-(<i>N</i> -morpholino)ethanesulfonic acid
HTS	high-throughput screen	min	minute(s)
HU	hydroxyurea	MOPS	3-(<i>N</i> -morpholino)propanesulfonic acid
		mRNA	messenger RNA
IC ₅₀	half maximal inhibitory concentration	MS	mass spectrometry
i.e.	id est; that is	MW	molecular weight
IF	immunofluorescence	MWCO	molecular weight cut-off
IFN γ	interferon γ		
IgG	immunoglobulin G	<i>n</i>	binding stoichiometry
IL-4	interleukin 4	n	nano; 10^{-9}
IMAC	immobilised metal ion affinity chromatography	n/a	not available
indels	insertion/deletions	NAD	nicotinamide adenine dinucleotide
IPTG	isopropyl- β -d-1-thiogalactopyranoside	NDDI	Novartis Drug Discovery Incubator
IR	ionising radiation	NES	nuclear export signal
ITC	isothermal titration calorimetry	NF κ B	nuclear factor kappa-light-chain-enhancer of activated B cells
		NLS	nuclear localisation signal
		NMR	nuclear magnetic resonance
		NP-40	Nonidet P-40
JNK	Jun N-terminal kinase	ns	not significant
k	kilo; 10^3	OAADPR	O-Acetyl-ADP-ribose
<i>K</i>	association constant	OCR	oxygen consumption rate
<i>k_a</i>	association rate constant	OD ₆₀₀	optical density at 600 nm
<i>k_d</i>	dissociation rate constant	O/N	overnight
<i>K_D</i>	dissociation constant		
KD	knock-down		
KO	knock-out	p	pico; 10^{-12}
		p.a.	pro analysi
		PAGE	polyacrylamide gel electrophoresis
λ	wavelength	PAMPA	parallel artificial membrane permeation assay
l	litre		
LB	Lysogeny Broth (LB miller)	PanDDA	Pan-Dataset Density Analysis
LC-MS	liquid chromatography–mass spectrometry	PAR	poly-ADP-ribose
		PARG	poly-ADP-ribose glycohydrolase
log <i>P</i>	partition coefficient between <i>n</i> -octanol and water	PARP	poly-ADP-ribose polymerase
		PARylation	poly-ADP-ribosylation
		PBM	PAR binding motif
		PBS	phosphate-buffered saline
μ	micro; 10^{-6}	PBZ	PAR-binding zinc finger

PCNA	proliferating cell nuclear antigen	TLCK	tosyl-l-lysyl-chloromethane hydrochloride
PCR	polymerase chain reaction	T _m	transition midpoint
P _e	effective permeability	Tris	tris(hydroxymethyl)aminomethane
PEG	polyethylene glycol	Triton X-100	polyethylene glycol <i>tert</i> -octylphenyl ether
PEG2kMME	polyethylene glycol monomethyl ether 2000	Tween-20	polyethylene glycol sorbitan monolaurate
PEI	polyethylenimine	V	volt
PEP	phosphoenolpyruvate	v/v	volume per volume
PKM2	pyruvate kinase isoform M2	Vol	volume
PMSF	phenylmethylsulfonyl fluoride	w/v	weight per volume
PTM	post-translational modification		
PVDF	polyvinylidene fluoride		
QC	quality control	WB	Western blot
		wt	wild-type
RING	Really Interesting New Gene		
RNA	ribonucleic acid		
RNF146	RING Finger Protein 146	x	times/ -fold
RRM	RNA recognition motif	XRCC1	X-ray repair cross-complementing protein 1
rpm	rounds per minute		
RT	room temperature		
		YFP	yellow fluorescence protein
SAR	structure–activity relationship		
SARS	severe acute respiratory syndrome		
SD	standard deviation		
SDS	sodium dodecyl sulphate		
sec	second		
SEM	standard error of the mean		
sgRNA	single guide RNA		
siRNA	small interfering RNA		
SSA	superstreptavidin		
STAT	signal transducer and activator of transcription		
t _{1/2}	half-life; residence time (1/k _d)		
TARG1	terminal ADP-ribose protein glycohydrolase 1		
TB	Terrific Broth		
TBS-T	tris-buffered saline with 0.1 % Tween-20		
TC	tissue culture		
TCA	trichloroacetic acid		
TCA (cycle)	tricarboxylic acid cycle		
TCEP	tris(2-carboxyethyl)phosphine		
TEV	Tobacco Etch virus		

I. INTRODUCTION

1. **Protein ADP-ribosylation – an ancient and versatile post-translational modification**

Post-translation modifications (PTMs) have a fundamental role in cell biology. PTMs modify the protein's surface and/or modulate protein conformation and thus protein activity and stability. Furthermore, they dictate cellular location and macromolecular interactions depending on the cell type, the tissue and the environmental conditions ¹. More than 200 different types of PTMs have been identified so far mainly through MS-based proteomic studies, with phosphorylation, acetylation and N-linked glycosylation being the most common ². PTMs are site-specific occurring either on a specific amino acid residue or on multiple sites, whereby the same or different modifications can be combined on the same polypeptide which substantially increases the heterogeneity of the gene products ^{1,3}. PTMs are thought to act independently, however, it is known that different types of PTMs crosstalk (e.g. phosphorylation-dependent ubiquitination of substrates) or act in combination (e.g. multiple phosphorylation sites in a protein resulting in a combinatorial effect on activity) ⁴. Consequently, PTMs are involved in a variety of diverse cellular functions. While phosphorylation cascades for instance mediate many signalling pathways, polyubiquitination can induce protein degradation, and acetylation marks on histones control (along with methylation and other marks) nuclear processes such as gene transcription ^{1,5,6}. Each of these modifications is regulated by a dynamic interplay between different enzymes that covalently attach PTMs onto proteins ('writers'), enzymes that revert these reactions and thus remove PTMs ('erasers') as well as proteins that recognise those modifications and thus have control over down-stream effects ('readers'). Such reader modules are either encoded within the same polypeptide as the catalytic domains of writers and erasers, or are present in protein partners that associate with them ⁷.

Protein ADP-ribosylation belongs to the array of post-translational modifications and was first described in the 1960s by Chambon *et al.* ^{8,9}. ADP-ribosylation is a phylogenetically ancient

nucleotide-based protein modification that is conserved in all organisms from eukaryotes to bacteria, viruses and archaea ¹⁰. Protein ADP-ribosylation is catalysed by ‘writers’ termed poly-ADP-ribose polymerases (PARPs) including mono- and poly-ADP-ribosyltransferases (ARTs) ^{11,12}. In humans, 17 PARP family members have been identified based on the homology to the PARP catalytic domain of PARP1, the founding member that has been the most extensively studied of this class of enzymes ^{13,14}. Of note is that the PARP family was recently organised under a new ‘ADP-ribosyltransferase, diphtheria toxin-like’ (ARTD) nomenclature due to the inaccuracy of the usage of the term ‘polymerase’ and the provided numbering ¹¹. However, in this work the old nomenclature for this protein family will be used. PARPs (with exception of a few inactive family members) transfer either a single adenosine diphosphate ribose (ADPR) unit (MARylation) or multiple units (PARylation) from β -NAD⁺ (nicotinamide adenine dinucleotide) onto acceptor proteins ^{12,15,16}. The majority of the PARP enzymes have been identified as mono-ADP-ribosyltransferases while the specificity of the reaction and thus the classification of the PARPs is based on secondary structural features of the PARP catalytic domain together with the presence of a conserved His-Tyr-Glu (H-Y-E) triad ¹⁶. The acidic residue is essential for PARylation activity while dispensable for MARylation reactions and often found to be replaced with neutral residues in the corresponding enzymes ¹⁷. In the PARP-catalysed reaction, the substrate β -NAD⁺ is cleaved into nicotinamide and ADPR that is subsequently covalently attached to target proteins with glutamate, aspartate, arginine, lysine and serine being the main well-known acceptor sites ^{15,18–21}. For protein PARylation, several ADPR units are coupled via *O*-glycosidic bonds between the adenine-proximal ribose of one unit and the distal ribose of another unit. In additional branching reactions on the polymer, ADPR units are joined between two distal ribose rings [Figure 1] ¹⁵. Poly-ADP-ribose (PAR) polymers can reach up to 200 ADPR units in length, with branching occurring every 20 to 50 units forming long and negatively charged PAR chains of highly dynamic and transient nature ^{22–25}. Efficient removal and degradation of PAR polymers and MARylation sites is catalysed by different hydrolases, ‘erasers’ of protein ADP-ribosylation. The break-down of PAR is predominantly mediated by poly(ADPR) glycohydrolase (PARG) which acts mainly as exo-glycohydrolase and hydrolyses the glycosidic bond between the ADPR units,

thereby generating free ADPR^{26,27}. Alternatively, PAR degradation can also be catalysed by the ADP-ribosylarginine hydrolase ARH3^{28,29}. However, both enzymes, PARG and ARH3, are incapable of removing the last, protein-bound ADPR units generating MARYlated proteins. This de-MARYlation reaction and consequently the reversal of protein ADP-ribosylation is performed by the ADP-ribosylarginine hydrolase ARH1 and in particular by mono-ADP-ribosyl hydrolases, enzymes that are members of the macrodomain family (MacroD1, MacroD2 and TARG1)^{29–33}.

Due to the size and diversity of the PARP family, protein ADP-ribosylation regulates a variety of fundamental cellular and biological processes. These include among others DNA repair and cell apoptosis, regulation of transcription, cytoplasmic stress responses, cell division and protein degradation^{15,34,35}. At the molecular level, ADP-ribosylation either directly affects the function of the target protein via the covalent modification or mediates non-covalent complex formation with proteins that are able to recognise these modifications. While MARYlation has been described to mediate specific protein-protein interactions in diverse signalling pathways e.g. of NF- κ B response³⁵, the major types of regulation by PARYlation are the inhibition of protein-nucleic acid interactions, the formation of scaffolds that alter protein localisation and promote protein interactions and the regulation of ubiquitination events whereby the latter two are driven by PAR-binding modules¹⁵. For instance in the event of DNA damage, PARP1 is strongly activated, resulting in extensive PARYlation of PARP1 itself as well as chromatin-associated proteins such as histones and other DNA repair factors. The generated PAR then provides a reversible protein binding scaffold for the recruitment of PAR binding proteins that are involved in chromatin remodelling (e.g. ALC1, APLF), DNA repair (e.g. XRCC1) or generalised DDR signalling (e.g. p53). Multiprotein complexes are formed that ultimately facilitate DNA repair^{15,36}.

Several ADPR recognising domains have been identified so far including the four following evolutionary conserved domains which are found in numerous proteins with diverse functions: PAR-binding motifs (PBMs), PAR-binding zinc fingers (PBZs), the WWE domains and the macrodomains [Figure 1]³⁷. PBMs were the first motifs described for their ability to

non-covalently interact with PAR and found in proteins involved in RNA metabolism, chromatin remodelling and DNA damage signalling and repair e.g. p53 and XRCC1^{38,39}. The PBZ is characterised as a specific and high affinity PAR-binding module due to its ability to simultaneously recognise up to two adenine rings of neighbouring ADPR units in a PAR chain. In humans, the PBZ is only present in very few proteins such as the DNA-repair protein APLF and the mitotic checkpoint protein CHFR^{37,40,41}. WWE domains and macrodomains are described in more detail below and are the focus of this work.

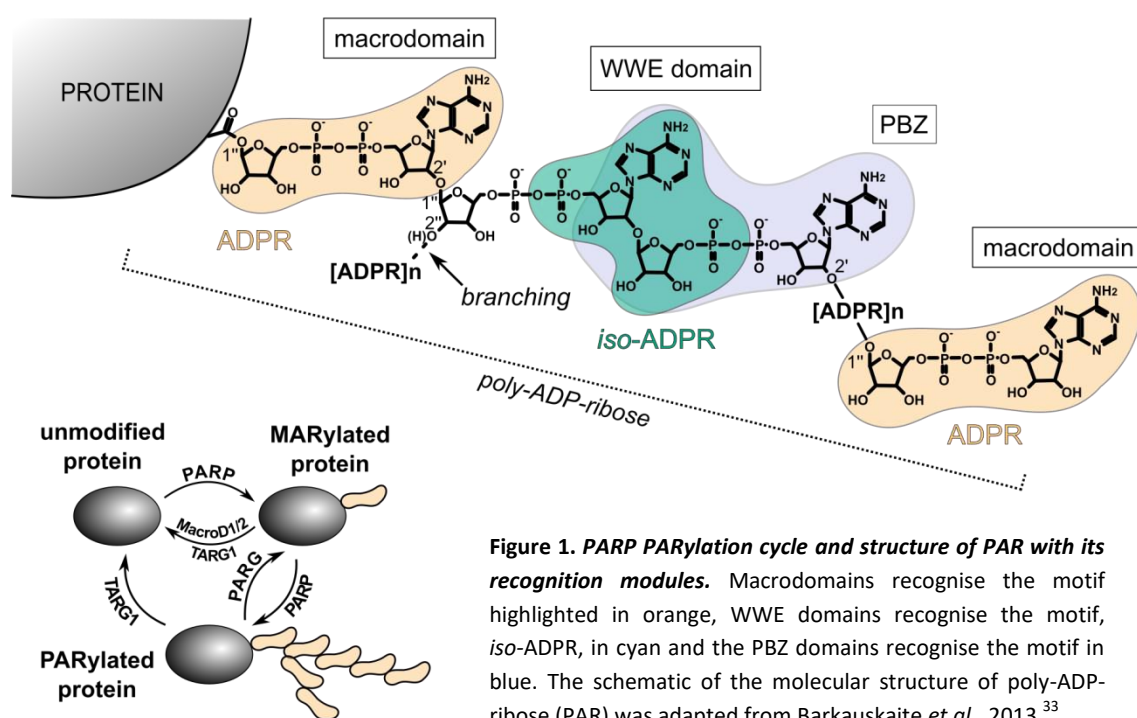


Figure 1. PARylation cycle and structure of PAR with its recognition modules. Macrodomains recognise the motif highlighted in orange, WWE domains recognise the motif, *iso*-ADPR, in cyan and the PBZ domains recognise the motif in blue. The schematic of the molecular structure of poly-ADP-ribose (PAR) was adapted from Barkauskaite *et al.*, 2013³³.

2. Macrodomains and WWE domains as readers of protein ADP-ribosylation

2.1. WWE domains

WWE domains are named after their most conserved three residues, two tryptophan residues and a glutamate residue that stabilise its fold. WWE domains have been shown to recognise a specific subunit of PAR, i.e. *iso*-ADPR^{42,43}. *Iso*-ADPR is defined as the smallest PAR structural unit containing the unique ribose-ribose glycosidic bond which is formed during PAR synthesis by PARP family members⁴³. Based on NMR solution and X-ray structures of the WWE domains in RNF146, PARP11 and PARP14, these domains

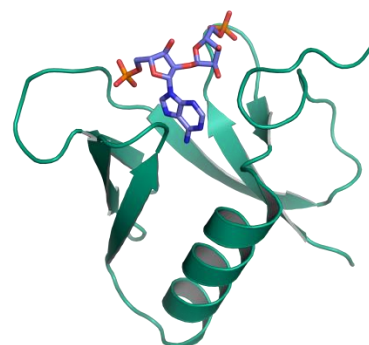


Figure 2. WWE domain of RNF146 in complex with *iso*-ADPR (PDB ID 3V3L)⁴³.

The cartoon structure of the WWE domain is shown in cyan, and the *iso*-ADPR ligand as atom-coloured stick model.

can be characterised as small, globular proteins that contain six β -strands, forming half of a β -barrel like structure which is covered by an α -helix on the other side^{43,44}. The *iso*-ADPR is bound on the top of the β -barrel structure with the adenine inserting into the pocket [Figure 2]. The two phosphate moieties are involved in several interactions with each of the positively charged edges of the WWE domain binding site, potentially explaining its specificity over ADPR which has the phosphate on the same side⁴³. So far, WWE domains have been identified in two functional protein subfamilies that are associated with either ubiquitination or PARYlation⁴². In particular, E3 ubiquitin ligases (RNF146, Deltex family members, HUWE1, Trip12) and the PARP family members (PARP7 and PARP11-14) have been reported to contain one or in some cases two WWE domains combined with other catalytic and regulatory domains^{12,42}. For the RING-type E3 ubiquitin ligase RNF146, a functional connection between its E3 ligase activity and PAR recognition could be shown. The PAR binding of RNF146 between its WWE and RING domain functions as an allosteric signal. It induces a conformational change in the RING domain which is then functional to bind and activate an ubiquitin-conjugating enzyme (E2)⁴⁵. Consequently, PAR binding to RNF146 activates its E3 ligase function and results in

polyubiquitination of several PARylated proteins involved in DNA damage response including PARP1, XCRR1, DNA ligase III for subsequent proteosomal degradation thereby regulating cell survival and DNA repair^{45,46}. In a similar way, RNF146 is involved in regulation of the Wnt signalling pathway by first interacting with ADP-ribosylated axin and then promoting its ubiquitination and degradation⁴⁷. The mechanisms, and in particular cellular downstream effects of PAR recognition by the WWE domains present in PARP family members, are largely uncharacterised. However, sequence alignments were performed and revealed that not all of the PARP WWE domains have conserved all of the fold-defining WWE residues as well as the four residues that seem to be crucial for *iso*-ADPR binding. An example is the WWE domain of PARP14 which has therefore been predicted to be incapable of PAR binding⁴³ [Figure S1].

2.2. Macrodomains

Macrodomains constitute a family of ancient evolutionary highly conserved protein domains that can be found in all of the phylogenetic kingdoms including eukaryotes, bacteria, archaea and ssRNA viruses⁴⁸. In humans, at least ten genes encode 12 macrodomain-containing proteins including splice variants^{48,49} of which the vertebrate histone variant macroH2A1 was the first identified member⁵⁰. Macrodomain-containing proteins may include up to three macrodomains often combined with other regulatory and signalling modules [Figure 3]^{51,52}.

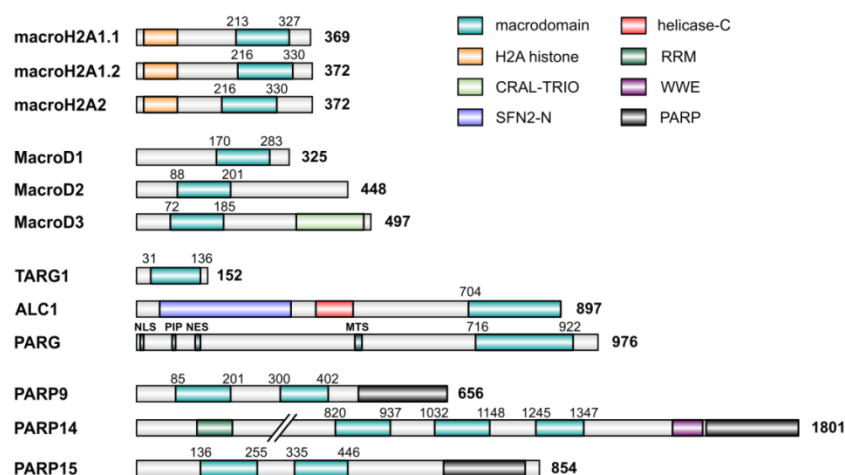


Figure 3. Domain structure of human macrodomain-containing proteins. For legend, see top of next page.

Legend to Figure 3: Domain structure of human macrodomain-containing proteins.

CRAL-TRIO: cellular retinaldehyde-binding protein-triple functional domain protein; RRM: RNA recognition motif; PARP: poly-ADP-ribose polymerase catalytic domain; NLS: nuclear localisation signal; NES: nuclear export sequence; MTS: mitochondrial targeting sequence; PIP: PCNA-interacting protein. The schematic was adapted from Posavec *et al.* ⁵⁰ and Feijs *et al.* ⁵² and prepared using IBS ⁵³.

Macrodomains are globular protein domains with 130-190 amino acids, thus approximately 25 kDa in size. Their structure consists of a mixed $\alpha/\beta/\alpha$ fold composed (for most human macrodomains) of central β -sheets with six β -strands in a mixed parallel/antiparallel arrangement which are surrounded by five α -helices as shown for instance for PARP14 MD2 [Figure 4] ⁵⁴⁻⁵⁷. Based on the first macrodomain structure of the thermophile *Archeoglobus fulgidus*, Af1521, the macrodomain fold was shown to bear resemblance to the N-terminal DNA-binding domain of the leucine aminopeptidases PepA as well as to the P-loop family of nucleotide hydrolases. Therefore, macrodomain functions associated with nucleotide binding and processing were suggested ⁵⁵. The majority of the 16 human macrodomains have since then been characterised as ligand-binding domains for NAD^+ -derived metabolites. These include in particular ADPR in its monomeric and polymeric form, PAR, as well as its analogues for instance ADP-ribose-1''-phosphate (Appr-1''-P) or O-acetyl-ADP-ribose (OAADPR) ^{54,58-61}. Appr-1''-P constitutes a tRNA splicing byproduct ⁶², while OAADPR is a metabolite produced through deacetylation of proteins by sirtuins such as SIRT1 ⁶³. The ADPR itself is bound slightly bent in a deep cleft on the crest of domain [Figure 4] and is stabilised by several conserved interactions within the binding pocket ⁵⁴. Furthermore, ADPR is recognised by macrodomains in its free form or as a post-translational modification attached to proteins, whereby the macrodomain recognises the entire ADPR moiety either directly at the protein attachment site or at the termini of PAR chains [Figure 1] ^{33,54,64}. This is in contrast to the previously mentioned ADPR binding modules such as the WWE or PBZ domain which bind within the PAR chain and recognise two sequential ADPR units. Four of the human macrodomains additionally possess catalytic activity. MacroD1, MacroD2 and TARG1 have been shown to deacetylate OAADPR ^{59,65} and more importantly to function as mono(ADP-ribosyl)

hydrolases, thus reversing protein ADP-ribosylation, as briefly mentioned previously^{30–32}. By hydrolysis of the ribose-protein acceptor bond, these three macrodomains remove the protein bound ADPR from MARYlated proteins, with TARG1 being additionally able to bind PAR and to remove the entire polymer from PARylated proteins²⁹. The macrodomain in PARG is together with an associated accessory domain responsible for the glycohydrolase activity of PARG, degrading PAR chains by cleavage of the ribose-ribose bonds²⁶. Of note is that (with exception of PARG) catalytically active macrodomains are stand-alone macrodomain proteins, however, reader macrodomains only occur in multidomain proteins, combining signal recognition and translation into various down-stream effects.

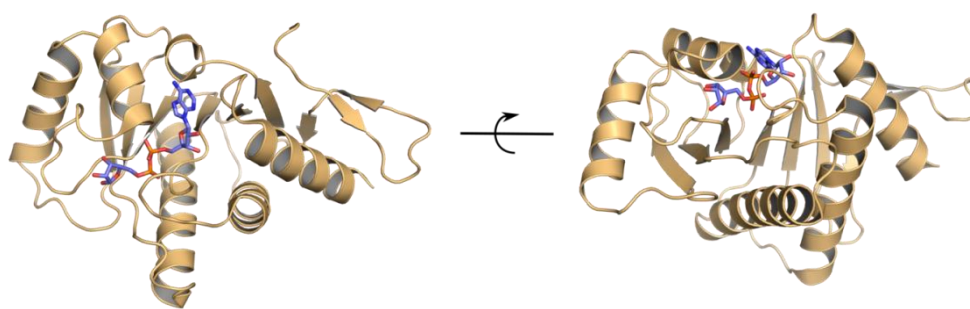


Figure 4. PARP14 MD2 in complex with ADPR (PDB ID 3Q71)⁶⁶. The cartoon structure of PARP14 MD2 is shown in yellow, and the ADPR ligand as atom-coloured stick model.

3. The human macrodomain-containing protein family and its diverse biological functions

The abilities of macrodomains of recognition, processing and turnover of ADPR signalling account for the variety of fundamental cellular processes in which macrodomains are involved. In the following, an overview of main functions of the human macrodomain-containing proteins [Figure 3, 5] shall be provided.

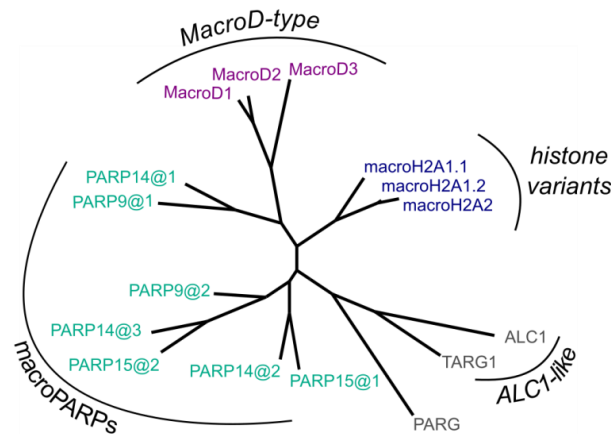


Figure 5. Phylogenetic representation of human macrodomains. The histone variants and the macroPARPs are shown as separate groups although both belong to the as ‘macroH2A-like’ defined branch ⁶⁷.

3.1. The macrodomain-containing histone variant macroH2A

MacroH2A is an H2A core histone variant that contains, apart from an N-terminal histone fold for integration into the chromatin, a C-terminal macrodomain ⁶⁸. Two genes encode the proteins macroH2A1 and macroH2A2, whereby two isoforms, macroH2A1.1 and macroH2A1.2, are generated from macroH2A1 through alternative splicing ⁶⁹. While macroH2A1.1 binds ADPR and derivatives, macroH2A2 and macroH2A1.2 are unable to bind these ligands, for the latter caused by the splicing event which alters the ADPR binding pocket ⁷⁰.

MacroH2A variants have been associated with the regulation of gene expression programmes during cellular differentiation and vertebrate development ⁷¹. MacroH2A1 was initially shown to be involved in gene transcription repression contributing to the silencing of the X chromosome in somatic cells of females ⁷². However, studies performed on embryonic stem cells suggested that macroH2A1 also facilitates transcriptional activation of genes important for their differentiation and development ⁷³. MacroH2A1.1 was also shown to be recruited to sites of PARP1 activation upon DNA damage resulting in transient chromatin compaction and reduction of the recruitment of DNA damage factors Ku70–Ku80, thus it was suggested to regulate DNA damage response ⁷⁴. However, macroH2A1 knockout mice showed metabolic rather than DNA damage phenotypes

which assumed a role for macroH2A1 in regulating gene expression programmes that control metabolic pathways^{61,75}.

3.2. MacroD-type macrodomains

MacroD1 and MacroD2 bind ADPR with high affinity in comparison to other human macrodomains and although their catalytic activities were well characterised *in vitro*, their biological roles are less understood. MacroD1 (also called LRP16) was reported to be involved in estrogen and androgen receptor signalling by acting as cofactor and enhancing the receptor-mediated transcription activities^{76,77}. Furthermore, it was suggested that MacroD1 integrates into the NF- κ B transcriptional complex through associating with its p65 component and is required for its functional activation⁷⁸. However, MacroD1 also possesses a mitochondrial targeting signal and shows mitochondrial localisation⁶⁰. It is unclear how this relates with its proposed nuclear functions. MacroD2 is present in the nucleus and cytoplasm⁶⁰ and was shown to activate glycogen synthase kinase 3 β (GSK3 β) *in vitro* and in cells by reversing its MARYlation-mediated inhibition catalysed by PARP10³⁰. The function of MacroD3 also known as GDAP2 (ganglioside-induced differentiation associated protein 2) is so far not known. Although it shows structural similarity with MacroD1 and MacroD2, MacroD3 does not bind ADPR or its derivatives⁶⁰. However, due to a incorporated lipid-binding SEC14 domain in the protein, a function in the integration of ADPR and lipid-mediated signalling has been suggested⁶⁷.

3.3. ALC1, PARG and TARG1

ALC1 ('amplified in liver cancer 1', also called CHD1L) contains, apart from its N-terminal macrodomain, a helicase domain that is also found in the Snf2 family of ATP-dependent chromatin remodelers^{58,79}. ALC1 has been reported to be involved in several functions including DNA damage repair, gene regulation, cell proliferation, embryonal development and

p53-independent apoptosis⁶⁷. In the context of DNA damage repair, ALC1 senses via its ADPR binding domain PARP1-generated PAR at DNA lesion sites. This PAR recognition and consequent ALC1 recruitment leads to formation of stable ALC1-nucleosome-PARP1 complexes promoting ALC1 chromatin-remodelling activity, interaction with further DNA damage repair factors and ultimately to DNA repair^{58,80}. Similarly, PARG performs a critical role in DNA damage repair. PARG was shown to be recruited to sites of DNA damage in a PARP1 and PAR-dependent manner but also by its interaction with PCNA. The latter is assumed to be important for a prolonged and constant recruitment of PARG which hydrolyses PAR, thereby facilitating regulation of later repair events and recycling of PAR^{67,81,82}. Finally, depletion of TARG1, the ‘terminal ADP-ribose protein glycohydrolase 1’ (C6orf130, LOC221443), from cells was also shown to result in defects in DNA repair as well as in proliferation defects and cellular senescence but the underlying molecular mechanism are less understood³².

3.4. MacroPARPs

The so called ‘macroPARP’ group comprises three proteins, PARP9, PARP14 and PARP15, with PARP9 being the first member identified in a genome-wide search for risk-related genes in diffuse large B-cell lymphoma. The proteins were therefore originally termed ‘BAL’ (B-Aggressive Lymphoma) family due to their overexpression and disease association in B-cell lymphoma^{83,84}. MacroPARPs are characterised by containing multiple macrodomains in combination with a C-terminal PARP domain. The PARP domains of PARP14 and PARP15 show MARylation activity, while the PARP domain of PARP9 seems to be catalytically inactive¹⁶. Both, PARP9 and PARP15, have two macrodomains and PARP14 has three consecutive macrodomains⁵¹. All three proteins are encoded on the same chromosomal locus, 3q21, within ~200 kb. It was shown that their genes have evolved under strong positive selection with the PARP15 gene being unique in its pattern of recurrent loss and gain. The PARP15 gene arises in early mammalian evolution from a partial duplication of PARP14, consisting of the second and

third macrodomain and the PARP domain⁸⁵. Positive selection is described as characteristic of critical host immunity genes and as an indication for host genes being involved in genetic conflict with viruses that counter-evolve to evade the host antiviral defences⁸⁶. Since strong evidence of positive selection was found for the macrodomains of PARP9, 14 and 15 and the PARP domain of PARP14, an antiviral activity of these proteins in host-virus conflicts was suggested⁸⁵. In earlier studies however, all three macroPARPs were described as transcriptional regulators with PARP9 and PARP15 repressing transcription when tethered to a promoter⁸⁴. PARP9 has additionally been shown to be recruited to DNA damage sites in a PARP1/PAR-dependent manner, thereby localising its binding partner BBAP (a DELTEX family member with ubiquitin E3 ligase activity) to DNA breaks. The subsequent histone H4 ubiquitination by BBAP promotes the recruitment of additional repair factors, including 53BP1 and BRCA1⁸⁷. Due to the relevance for this project, PARP14 functions are discussed more detailed in the following.

3.5. PARP14 – a triple macrodomain-containing protein with multiple functions

PARP14 is unique among the macrodomain-containing proteins in the way that it combines several ADP-ribosylation recognising modules of two different types (three macrodomains and one WWE domain) with a catalytically active PARP domain at the C-terminus and a RNA recognition motif (RRM) at the N-terminus. PARP14 can be described in general as a pro-survival protein with cytoplasmic and nuclear localisation. Several cellular functions of PARP14 were proposed ranging from transcriptional regulation, immune and stress response, cell morphology and motility to metabolic regulation and DNA repair which are summarised in the following.

PARP14 was initially shown to be involved in transcriptional regulation, modulating interleukin 4 (IL-4) induced STAT6-dependent gene expression, and was therefore also termed ‘CoaSt6’ (collaborator of STAT6)^{88–90}. Mechanistically, IL-4 interaction with its surface receptors results in receptor dimerisation and phosphorylation by Janus kinases and subsequently to STAT6

phosphorylation. STAT6 dimerises, translocates to the nucleus and initiates transcription after recruiting several coactivators including p100. IL-4 furthermore stimulates the catalytic activity of PARP14. PARP14 MARYlates histone deacetylase (HDAC) 2 and 3 which are consequently released from the IL-4 responsive promoter. This allows efficient promoter binding of STAT6 and its transcriptional coactivators, including p100 and p300. p100 interacts with RNA polymerase II⁹¹ and is also MARYlated by PARP14 which might help to bridge STAT6 to the transcription machinery^{52,92}. It was confirmed in lymphocytes that the induction of several B-cell survival factors, including Mcl-1 and Pim-1, by IL-4 is mediated by PARP14. Both the anti-apoptotic Bcl-2 family member Mcl-1 and the Pim kinase family member Pim-1, cooperate with oncogenes such as c-Myc and promote cell survival and lymphomagenesis⁹³⁻⁹⁵. In lymphoma cells, the IL-4-mediated transcriptional function of PARP14 was also linked to metabolic regulation by transcriptional induction of glycolytic genes. B-cells showed an increased glucose uptake associated with higher rates of glycolysis and glucose oxidation by a PARP14-STAT6-dependent mechanism⁹⁶. In support of a metabolic function of PARP14 is the finding that PARP14 stabilises and thus contributes to the secretion of phosphoglucose isomerase (PGI), a cytosolic enzyme essential for glycolysis and gluconeogenesis which additionally acts extracellularly as a cytokine termed ‘autocrine motility factor’, by inhibiting its ubiquitination-mediated degradation⁹⁷. Moreover, PARP14 has recently been shown to promote the Warburg effect in hepatocellular carcinoma cells by inhibiting the pro-apoptotic kinase JNK1. This in turn halts JNK1-mediated phosphorylation of PKM2 and the resulting low PKM2 activity positively regulates aerobic glycolysis, promoting cell survival and proliferation⁹⁸. In an earlier study, PARP14 was also identified as pro-survival factor acting in the JNK signalling pathway, however thereby promoting the pathogenesis of multiple myeloma cells as downstream effector of JNK2. Constitutively activation of JNK2 was found to be associated with increased expression levels of PARP14 in myeloma cells. Further, as in the previous study, PARP14 was shown to directly inhibit JNK1 kinase activity which was assumed to be crucial for the suppression of basal apoptosis and thus the survival of myeloma cells⁹⁹.

Apart from these PARP14 functions related to transcriptional regulation and cell metabolism, PARP14 was shown to play a role in immune and stress response. Like other PARP family members, such as PARP13, PARP14 was shown to be a component of stress granules and (similarly to PARP13) suggested to be involved in protection of mRNA degradation under stress conditions. It was hypothesised that its ADP-ribosylation activity (increased under stress) on miRNA binding Argonaute family members might affect functioning of the Argonaute/miRNA complex and thus effective miRNA silencing¹⁰⁰. Although PARP14 possesses a RRM domain, a more mechanistic link of PARP14 catalytic activity, mRNA binding and its protection from miRNA silencing was not provided. However in a following independent study, PARP14 was shown to form a complex with the mRNA-destabilising protein tristetraprolin (TTP) and the 3' untranslated region of the mRNA of macrophage TF (tissue factor, CD142) presumably with its RRM domain. This resulted in an increase in TF mRNA stability and consequently TF expression in PARP14-deficient cells. TF, a transmembrane cell surface glycoprotein, triggers the extrinsic coagulation cascade and thus the findings were in line with the phenotype of PARP14^{-/-} mice which appeared to be healthy, however, showed increased thrombus formation¹⁰¹. PARP14^{-/-} mouse models however also indicated that PARP14 plays a critical role in immunological processes which are linked to the IL-4-mediated STAT6 gene expression regulation by PARP14 as described above. PARP14 (and its catalytic activity) seems to promote IgA and the STAT6-dependent IgE antibody response and to be required for B- and T-cell development by mediating the IL-4 stimulated differentiation of naive T-cells into several T-helper cell types e.g. Th2, Th17^{93,102-104}. In allergic disease models with PARP14^{-/-} mice, this PARP14-promoted Th2 differentiation was associated with allergic airway disease - where Th2 responses and IgE are promoters of the asthmatic conditions - but contrarily could also limit the severity of allergic skin disease^{105,106}. It was concluded that although PARP14 has similar effects on T-cell cytokine production in several allergic disease models, the outcome of those effects is distinct, depending on the target organ of disease¹⁰⁶. Furthermore, a global proteomic analysis of macrophage cell lines treated with IFN γ or IL-4 revealed that PARP14 cross-regulates with PARP9 macrophage activation which could be linked to human coronary artery disease by network analysis¹⁰⁷. Apart

from its role in the anti-inflammatory IL-4–STAT6 pathway, PARP14 was additionally shown to suppress the pro-inflammatory IFN γ –STAT1 axis by inhibiting STAT1 activation (phosphorylation) through ADP-ribosylation. This is antagonised by PARP9 which interacts with PARP14 and thereby promotes IFN γ pro-inflammatory responses. In PARP14^{-/-} mouse models, PARP14's protective effects against arterial disease, i.e. its anti-atherogenic role by suppression of macrophage activation, could be confirmed ¹⁰⁷.

The plethora of PARP14 functions was expanded by further *in vitro* studies. A systematic study of the 17 PARP protein family members by analysis of their knock-down phenotypes as well as their cell cycle localisation after ectopic overexpression as GFP-tagged proteins indicated that PARP14 is involved in actin cytoskeleton regulation as a component of focal adhesion complexes ¹⁰⁸. PARP14 knock-down in HeLa cells resulted in severe defects in cell motility as well as in an atypical cell morphology. Moreover, a critical role of PARP14 in DNA damage repair and protection against replication stress was proposed ¹⁰⁹. PARP14 was shown to interact with the PCNA DNA replication machinery and to promote replication of DNA lesions and common fragile sites by homologous recombination. Thus, PARP14 is described as important factor for mitigating replication stress and providing genomic stability ¹⁰⁹.

Finally, it is possible that PARP14 may participate in the NF- κ B signalling cascade by its interaction with PARP10. PARP14 macrodomains were shown to bind MARYlated PARP10 *in vitro* as well as to co-localise with PARP10 in cells. PARP10 regulates the NF- κ B pathway by MARYlating NEMO (NF- κ B essential modulator), the key component in the NF- κ B pathway that assembles the IKK complex. MARYlation prevents polyubiquitination of NEMO and thus activation of the IKK complex, subsequently resulting in decreased NF- κ B signalling and gene transcription ^{66,110}. It needs however to be confirmed if and how this interaction between PARP14 and PARP10 is relevant for NF- κ B signalling.

4. Human macrodomain-containing proteins in disease

Dysregulation of macrodomain function by overexpression or mutations has been associated with several diseases including inflammation, developmental defects, neurodegeneration and in particular cancer^{49,67}. Several macrodomain-containing proteins have been shown to be overexpressed in different types of cancers and are linked to their progression and aggressiveness with poor prognostic outcome. For instance, MacroD1 is overexpressed in a range of human tumours including endometrial, gastric, colorectal and breast carcinoma and associated with tumour proliferation, invasion and metastasis^{49,111–115}. MacroD2 was also shown to be overexpressed in a subset of breast cancers and shown to mediate tamoxifen resistance and estrogen independent growth¹¹⁶. Furthermore, ALC1 is frequently amplified in particular in hepatocellular carcinoma and identified as an oncogene promoting tumorigenesis, i.e. cell survival, proliferation and metastasis^{117–119}. Apart from its role in DNA damage repair mentioned previously⁵⁸, ALC1 has been shown to accelerate G1/S phase transition and to inhibit apoptosis by activation of the AKT pathway as well as by binding to the apoptotic protein Nur77 with its macrodomain which blocks Nur77 translocation from the nucleus to the mitochondria, and thus its initiation of apoptosis^{117,120,121}. Also the macroPARP members play a role in cancer cell biology. PARP9 was identified as an oncogenic survival factor in diffuse large B-cell lymphoma and is associated with increased cell migration, proliferation, survival and chemo-resistance^{83,122,123}. Mechanistically, PARP9 has been shown to act in the IFN γ –STAT1–IRF1–p53 axis, blocking IRF1-mediated apoptosis and inhibition of growth. PARP9 forms a complex with STAT1 through its macrodomains in an ADP-ribosylation-dependent manner, and is then recruited by STAT1 to its target promoters and inhibits the expression of the tumour suppressor IRF1 and simultaneously up-regulates proto-oncogenes including IRF2, BCL2 and BCL6¹²³. Interestingly, PARP9 was also able to up-regulate expression of PARP14 in this context and a regulatory crosstalk between PARP9 and PARP14 was suggested¹²³. PARP14, found to be overexpressed in particular in lymphoid cell lines and hepatocellular carcinoma, promotes through its action in the IL-4–STAT6

and JNK signalling pathways not only the survival of cancer cells but also the ability of cancer cells to adapt their metabolism to cope with the demands of enhanced cell growth and proliferation. Its stabilisation of the autocrine motility factor, an enzyme involved in tumour progression and metastasis, as well as its knock-down phenotype displaying impaired motility and increased cellular adhesiveness of cancer cells, are further indications of the importance of PARP14 in cancer-related processes.

5. The therapeutic potential of macrodomain targeting and current progress

Given their pathobiological importance, macrodomain-containing proteins are seen as potential therapeutic targets. Macrodomain-containing proteins involved in DNA damage recognition and response and/or transcriptional regulation such as PARG, ALC1, macroH2A1.1 and the macroPARPs with PARP14 may be of particular interest. Beyond that, targeting macrodomain-containing proteins present in pathogens may also be a promising therapeutic approach. Apart from bacterial macrodomains, a range of viruses including (SARS-) Corona-, Alpha-, Togaviruses and the hepatitis E virus encode macrodomains that play a critical role for host cell infection and viral replication^{67,124–126}.

Targeting components of ADPR signalling has furthermore already been a successful strategy for cancer therapy where PARP1 inhibitors have been successful in clinical trials and have been approved for treatment^{127,128}. Two generations of PARP1 inhibitors were developed including AstraZeneca's first-in-class inhibitor olaparib which are in Phase III trials or FDA approved for treatment of ovarian and breast cancer^{128–130}. Targeting macrodomains with small molecule inhibitors may provide an attractive alternative to PARP inhibitors by disrupting ADPR signalling by a different mechanism and without the difficulties of obtaining inhibitors specific for individual PARP enzymes. So far, inhibitor development has focused on the PARG macrodomain. Only recently, a chemical probe targeting PARG hydrolase activity has been reported which shows convincing on-target pharmacology and selectivity in contrast to early inhibitors such as rhodamine, phenolic hydrazide hydrazones or salicylanilide-based compounds or the ADPR

analogue adenosine 5'-diphosphate (hydroxymethyl) pyrrolidinediol (ADP-HPD)^{131–136}. Furthermore, potential inhibitors for the nsP3 macrodomain present in the chikungunya alphavirus have been identified using virtual screening and molecular dynamics simulations; however, data confirming any *in vitro* activities of these inhibitors or of further inhibitor development has not been reported so far¹³⁷. Although the availability of macrodomain inhibitors for use as tool compounds for target validation and beyond would be highly desirable, efficient, selective and cell active inhibitors targeting other macrodomains than PARG are not published so far. The development may be hampered due to the lack of suitable high-throughput screening (HTS)-compatible biochemical assays to identify chemical starting points for drug discovery projects. Only recently a HTS HTRF screening assay was reported that was used for inhibitor discovery targeting the glycohydrolase activity of PARG¹³⁸. This assay is, however, specific for PARG and does not enable identification of inhibitors for other macrodomains that have reader domain function. Therefore, in my rotation project prior to this work, an AlphaScreen-based screening strategy suitable for high-throughput screening of most human macrodomains was developed, leading to the discovery of a macrodomain inhibitor, GeA-69,

[Figure 6] which targets PARP14 MD2 with micromolar affinity via an allosteric binding mechanism. PARP14 as an emerging therapeutic target is also currently in focus for inhibitor development by other research groups. However, all published approaches so far address targeting its catalytic PARP domain and not of any other PARP14 domain^{139–141}. Most PARP domain inhibitors show significant activity on catalytic domains of other PARP family members, in particular of PARP1, and the optimisation of their selectivity profile is challenging. Targeting PARP14 via its macrodomain(s) and thus interfering with recruitment and cellular localisation of PARP14 might therefore be an alternative strategy to prevent PARP14 MARYlation activity on its respective targets and its mediation of various down-stream effects. The same outcome would be to expect as with inhibition of the PARP14 catalytic domain. GeA-69 thus represents a new class of macrodomain inhibitors and moreover may provide a new strategy for targeting PARP14, a protein of established therapeutic value.

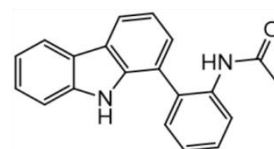


Figure 6. Structure of GeA-69.

6. Aim of this thesis

The first part of this work focuses on a further understanding of the sequence recognition of PARP14 by its three macrodomains and WWE domain. To target PARP14 by inhibition of its recruitment to ADP-ribosylated target sites, instead of inhibition of its catalytic domain, it is important to analyse which macrodomain(s) need to be inhibited to efficiently prevent PARP14 recruitment. It is unknown if the inhibition of only one reader domain could be compensated by the other three ADP-ribosylation recognising modules which would require inhibitor development for several, potentially different domains, making this strategy of PARP14 targeting less attractive. Furthermore, PARP14 macrodomains have been shown to bind ADPR *in vitro* with micromolar affinity, however, sequence specificity of its WWE is unknown. The aim was therefore to show whether the PARP14 WWE domain is still capable for PAR binding or if its function was replaced by the three macrodomains as predicted.

The aim of the second part of the project was to provide a better understanding of the druggability of PARP14 macrodomains. The new PARP14 MD2 inhibitor, GeA-69, was fully characterised concerning biophysical properties, selectivity over other macrodomain family members as well as target engagement in a cellular environment. Moreover, the performance of a medium-throughput screen in collaboration with Novartis (Basel, Switzerland) as well as a fragment-based screening approach by X-ray crystallography to allow the discovery of small molecules with new chemical scaffolds targeting PARP14 macrodomains was evaluated. The aim was to identify compounds that are able to mimic ADPR and thus are amenable as small competitive inhibitors for the ADPR binding site. This could potentially provide a basis for the development of macrodomain inhibitors.

In the third and final part of the project, pharmacological targeting strategies for PARP14 were explored. The aim was to confirm some of the multiple functions proposed for PARP14 by genetic and chemical inhibition of PARP14 function and to suggest or develop an assay system that would allow evaluating potential PARP14 (macrodomain) inhibitors in the future.

II. RESULTS AND DISCUSSION

1. **PARP14 sequence recognition by its macrodomains and WWE domain**

1.1. PARP14 is recruited to ADP-ribosylated sites mainly through its macrodomains

In order to efficiently target PARP14 via its ADPR-recognising modules, an understanding of the sequence specificity and recognition of its macrodomains and WWE domain is crucial. PARP14 macrodomains were previously shown to specifically recognise MARylated proteins⁶⁶. Although being in close proximity in the polypeptide, the macrodomains bound ADPR independently from each other, thus did not cooperate or have allosteric contact⁶⁶. While MD1 was shown to interact with ADPR only weakly, MD2 and MD3 recognised ADPR with significantly higher affinity (~5 μM). Interestingly, alternate binding affinities for MD2 and MD3 *in vitro* and in cells were suggested. MD3 apparently bound more efficiently *in vitro* and MD2 in cells to auto-modified PARP10⁶⁶. Due to the relevance for this project, the ADPR binding affinities of the individual macrodomains which were thought to be directly related to their efficiency in PARP14 recruitment to target sites were confirmed. Isothermal titration calorimetry (ITC) was used to measure these interactions in solution. ITC is a thermodynamic analysis method that allows determining the dissociation constant K_D , the binding stoichiometry n and binding enthalpy ΔH of macromolecular interactions, e.g. of protein-protein or protein-ligand interactions, without the need of chemical tagging or immobilisation of the interacting components¹⁴². By directly measuring either the absorbed or released heat, thus the enthalpy change, upon binding of two partners, ITC provides a thermodynamic profile of the binding event¹⁴³. Although being a method with rather low throughput that requires relatively large amounts of protein, ITC is described as one of the most reliable techniques for the quantitative measurement of ligand binding to a soluble protein¹⁴⁴ and was therefore used for studies throughout this work. The ITC measurements showed that ADPR indeed bound weakly to PARP14 MD1 with a K_D of $(137 \pm 7) \mu\text{M}$, and MD2 recognised ADPR with an higher affinity of $(6.8 \pm 0.1) \mu\text{M}$ than MD3 (K_D (ADPR) =

(15 ± 0.9) μM) [Figure 9B]. Of note is that, here as in the following, the errors given for K_D values obtained from ITC measurements are only errors of the fit of the data to an one-site binding isotherm, since errors of the K_D values themselves could not be measured by independent biological repeats. Furthermore, K_D values of weak micro-molar binding events are of significantly higher inaccuracy due to the low signal-to-noise ratio and therefore strongly depend on the method of fixing the baseline of the raw ITC data which was performed manually, but as consistently as possible. Thus, the binding affinities measured for MD2 and MD3 were broadly consistent with the previous study. Interestingly, binding of ADP towards MD2 could not be detected under these conditions despite the high similarity of ADPR and ADP [Figure 9B]. This suggests that the terminal ribose unit of ADPR significantly contributes to the recognition of ADPR either through direct interactions to the macrodomain or from its influence on the beta-phosphate which has a different charge state in ADP compared to ADPR. It can be assumed that it is of biological relevance that the macrodomain does not recognise ADP to avoid competition with this cofactor in the cell. Following the confirmation of ADPR binding of the PARP14 macrodomains, it was also tested whether the WWE domain of PARP14 is capable of PAR recognition. The sequence alignment of the WWE domain of PARP14 and RNF146 revealed that not all of the fold-defining WWE residues as well as the four conserved residues that seem to be important for *iso*-ADPR binding as defined by Wang *et al.*⁴³ are conserved in the WWE domain of PARP14 [Figure 7]. However, PAR recognition by the PARP14 WWE domain cannot be completely excluded since two of the conserved residues involved in *iso*-ADPR binding (Tyr1539, Tyr1576) are still present in this domain.

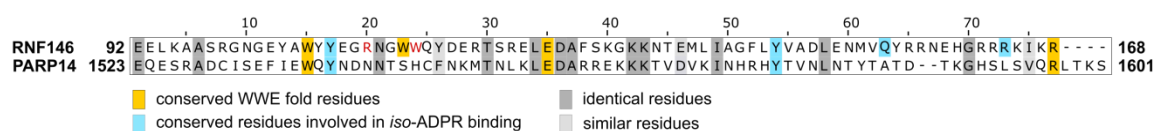


Figure 7. Sequence alignment of the WWE domains of PARP14 and RNF146. Most conserved residues important for WWE domain folding and binding of *iso*-ADPR defined by Wang *et al.*⁴³ are highlighted in yellow and cyan, respectively. Residues highlighted in red are non-conserved, however, interacting with *iso*-ADPR in RNF146.

The detection of PAR binding of PAR-recognising modules is typically performed by preparation of [^{32}P]-labelled PAR in a PARP1 auto-modification or substrate-assisted catalysed reaction whereby histones are generally used as acceptor proteins. [^{32}P]-PAR is then incubated either protein-bound or protein-free (after alkaline detachment) with the corresponding PAR-binding proteins that are immobilised on nitrocellulose membranes and detected by autoradiography after several washes of the membrane^{25,40,60,145}. Alternatively, a method was published for *in vitro* biosynthesis and purification of *iso*-ADPR, the unique internal structural unit of PAR, which is not commercially available⁴³ [Figure 8B]. The procedure included the digestion of the PARP1-generated PAR polymer to *iso*-ADPR by phosphodiesterase. Due to the unavailability of autoradiography and the possibility of crystallisation attempts in case of confirming PAR binding of the WWE domain, *iso*-ADPR was generated based on the latter protocol. Since PARP1 requires for efficient polymerase activity its DNA binding domain as well as the auto-modification domain (which has been confirmed by failure of initial attempts using the catalytic domain of PARP1 only) and is however challenging to purify as full-length protein¹⁴⁶, PAR was generated using the catalytic domain of PARP5 (TNKS1). Its suitability for PAR generation has been recently established^{146,147}. Although PARP5 has been shown to be ~6 fold less active than PARP1, the catalytic domain of PARP5 was prepared in good yield and it did not require additional activation of its auto-modification reaction. Furthermore, the catalytic property of PARP5 to polymerise $\beta\text{-NAD}^+$ to linear oligomers with ~20 ADPR units instead to long and branched chains obtained with PARP1¹¹ was thought to be beneficial for the planned enzymatic PAR digestion. The final workflow is outlined in Figure 8A, allowing the efficient preparation of *iso*-ADPR in quantities and purity sufficient for biochemical analysis [Figure 8]. Since it has been shown that for the PARP1 auto-PARylation reaction a catalytic dimer is required, where the two units simultaneously function as catalyst and acceptor molecules¹⁴⁸, only PARP5 purified as dimer was used for the reaction [Figure 8C]. It should be mentioned that the *iso*-ADPR batch generated for these studies contained BisTris salt as a contaminant due its use in the ion exchange buffer during the purification process [Figure S2]. The use of a volatile buffer would have enabled

easier separation of the *iso*-ADPR from the buffer component. Therefore, control binding experiments were performed in the presence of BisTris.

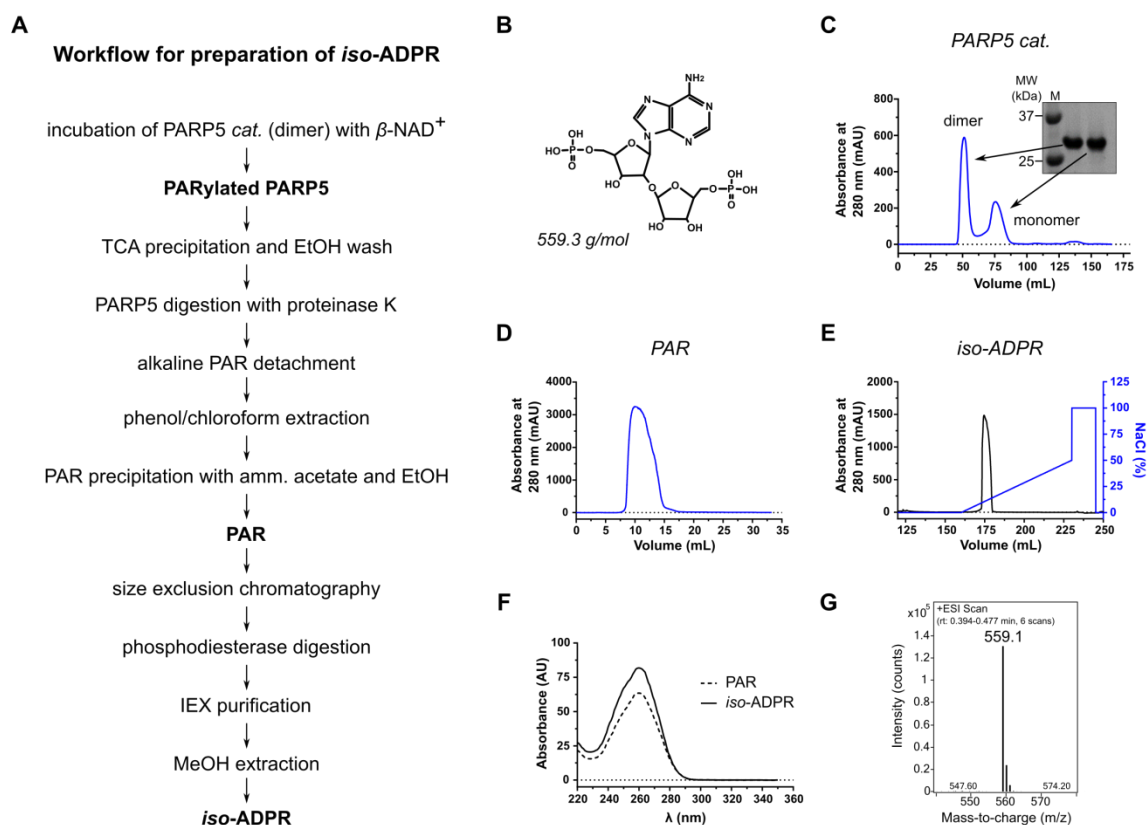


Figure 8. Preparation of *iso*-ADPR for PARP14 WWE binding studies. (A) Schematic workflow for *in vitro* biosynthesis and purification. (B) Molecular structure of *iso*-ADPR. (C) Preparation of the catalytic domain of PARP5. The gel filtration profile shows the dimer formation of cat. PARP5; SDS-PAGE of the final samples of the dimer and monomer fraction is shown next to it. (D) Gel filtration profile of PAR purified on a S75 10/300 GL column. (E) Anion exchange purification of *iso*-ADPR using a MonoQ 5/50 GL. The *iso*-ADPR eluted at ~100 mM NaCl. (F) Absorbance spectra of PAR and *iso*-ADPR for determining the concentrations of the stock solutions. (G) Mass spectrometry spectrum confirming the correct mass of the product, *iso*-ADPR.

The following ITC analysis demonstrated that the WWE domain of RNF146 bound to *iso*-ADPR with high affinity ($K_D = (508 \pm 12)$ nM), confirming previous results⁴³ and additionally validating the suitability of the prepared ligand to test PAR binding of the WWE domain of PARP14 [Figure 9C]. At concentrations of 0.5 mM and 1 mM *iso*-ADPR, no binding of the WWE domain of PARP14 was observed by ITC [Figure 9C]. Since ITC is however not the method of choice for direct detection of very weak binding, nano-Differential Scanning Fluorimetry (nanoDSF) experiments were additionally performed. Differential scanning fluorimetry allows monitoring the thermal stabilisation of the native protein structure upon ligand binding. The thermal stability is

described by the thermal unfolding transition midpoint T_m (°C), the temperature at which half of the protein is unfolded^{149,150}. NanoDSF is a label-free method which measures the intrinsic tryptophan fluorescence of proteins which relies on the close surrounding of amino acids and changes upon thermal unfolding. The T_m is calculated from the fluorescence wavelength ratio of 330/350 nm that describes the shift of tryptophan emission upon unfolding¹⁵¹. Since the WWE fold is stabilised through tryptophans in the core of the protein and *iso*-ADPR does not absorb light at this wavelength, which may cause quenching effects [Figure 8F], nanoDSF seemed to be a very suitable method for measuring if the domain can be stabilised through *iso*-ADPR binding. Indeed, in the presence of 0.5 mM or 1 mM *iso*-ADPR, the WWE domain of PARP14 showed a clear concentration-dependent thermal stabilisation with a T_m shift of 4.5 °C and 7 °C, respectively. No stabilisation of the WWE domain was observed at lower *iso*-ADPR concentrations or in its absence [Figure 9D]. This suggests that the PARP14 WWE domain does recognise *iso*-ADPR and is stabilised upon its binding. However since this effect was only observed at high ligand concentrations and no binding was detected by calorimetric experiments, the interaction between the PARP14 WWE domain and *iso*-ADPR is probably very weak and in the high micromolar to mM K_D range. As mentioned previously, two conserved residues shown to be involved in *iso*-ADPR binding by RNF146 are not present in the PARP14 WWE domain which explains the low binding affinity towards this ligand. It can therefore be concluded that the interaction of the WWE domain with PAR is probably too weak to efficiently recruit PARP14 to PARylated target sites and it can be hypothesised that the function of the WWE domain was evolutionarily taken over by the macrodomains. It should however also be considered that the experiments were performed with the smallest PAR structural unit and not with whole PAR chains. Additional PAR assays such as nitrocellulose membrane assays using [³²P]-PAR as previously mentioned would need to be performed to fully confirm the weak PAR-binding capabilities of the WWE domain. Of note is further that potentially in the context of the full-length protein the weak affinity may contribute to PAR binding by a resulting avidity effect of multiple weak interactions together with the PARP14 macrodomains, which also only show micromolar affinity towards ADPR. Nevertheless, inhibitor development should address the

domains which show the highest binding affinity for ADP-ribosylated targets, thus the macrodomains to most efficiently prevent PARP14 recruitment and function on its respective target sites.

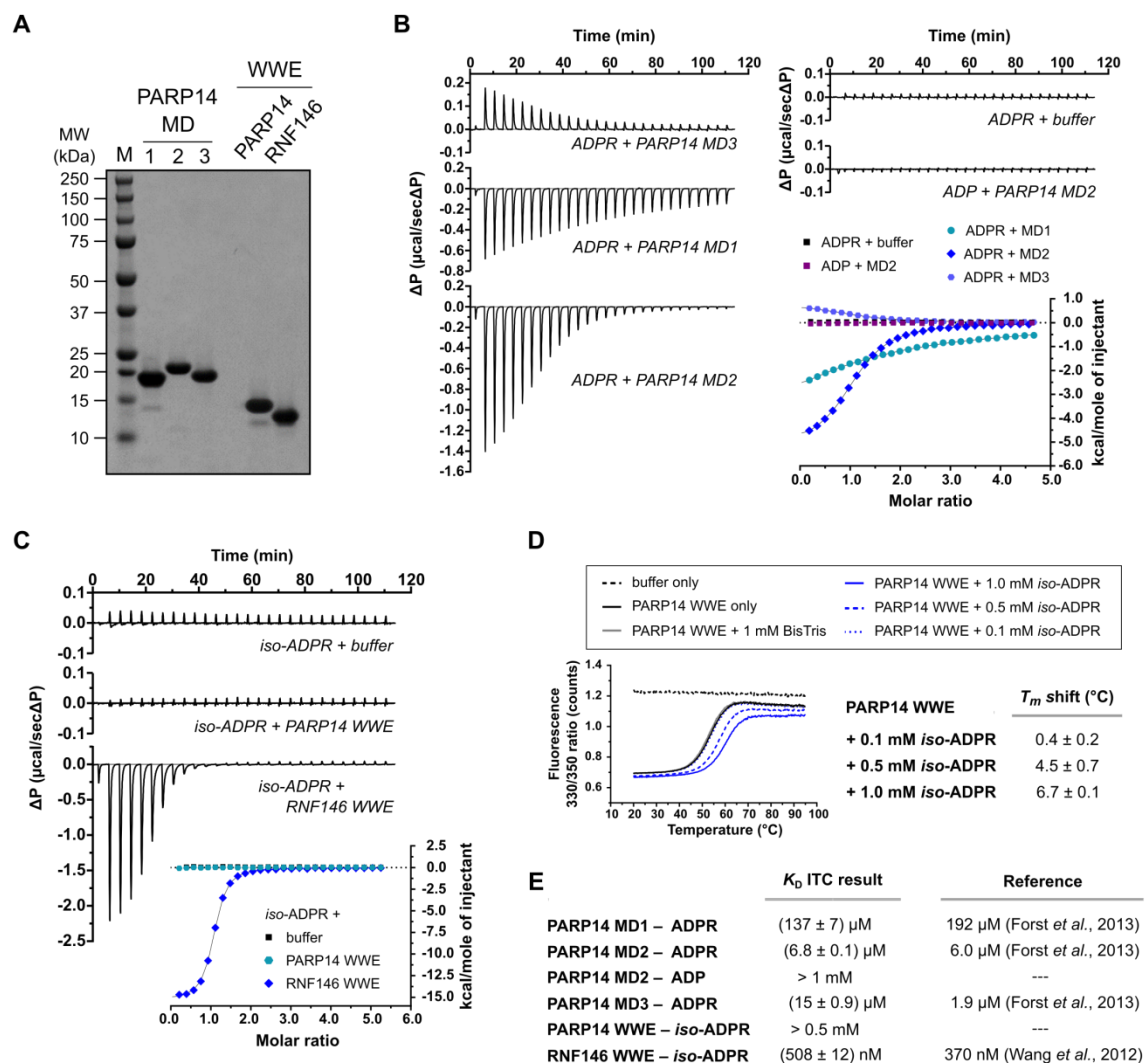


Figure 9. PARP14 macrodomain and WWE domain sequence recognition studies. (A) SDS-PAGE of purified PARP14 macrodomains and the WWE domains of PARP14 and RNF146. ~2 μg of each protein were loaded alongside a protein marker (M) and visualised with Coomassie Blue staining. **(B)** PARP14 macrodomains recognise ADPR with micromolar affinity, consistent with previous published results. Of note is that ADPR binding to PARP14 MD3 was endothermic in contrast to previous measurements by Forst *et al.*⁶⁶ where ADPR bound exothermically to PARP14 MD3 like MD1 and MD2. This could be due to the difference in the measurement temperature; the experiments shown here were carried out at 15 °C while Forst *et al.* performed the titrations at 25 °C. Titration curves show the stepwise addition of 1.0 mM ADPR or ADP into buffer only or buffer containing 40 μM of PARP14 MD1, 2 or 3. **(C)** Iso-ADPR binding of the WWE domain of PARP14 cannot be detected by ITC. Titration curves show the stepwise addition of 0.5 mM iso-ADPR into buffer only or buffer containing 25 μM of the WWE domain of PARP14 or RNF146. **(D)** The WWE domain of PARP14 recognises iso-ADPR. NanoDSF measurements reveal the thermal stabilisation of the PARP14 WWE domain in presence of iso-ADPR in a concentration-dependent manner. T_m shift data is mean $\pm$ SD, $n = 6$ of three independent experiments. **(E)** Summary of K_D ITC results from measurements shown in (B) and (C). Additional binding and fitting parameters are provided in Figure S3. *Note:* For all ITC graphs in this figure: The insets below the raw injection profiles show normalised binding enthalpies corrected for the heat of dilution as a function of binding site saturation (symbols as indicated above or in the inset). Solid lines represent a nonlinear least squares fit using a single-site binding model.

1.2. All three PARP14 macrodomains contribute to PARP14 recruitment in cells

Since PARP14 harbours three macrodomains, it is important to know which macrodomain(s) need to be inhibited to prevent PARP14 recruitment to its ADP-ribosylated target sites. It is conceivable that inhibition of only one macrodomain could be compensated by the other two macrodomains, in particular since MD2 and MD3 show similar binding affinities *in vitro* and are both independently binding to ADPR⁶⁶.

In order to address this question, an assay system was established that allowed studying PARP14 and its macrodomain recruitment to ADP-ribosylated target sites in cells. It is well known that DNA damage sites are hot spots for PAR generation catalysed by PARP1. Upon sensing and binding to DNA breaks, PARP1 is strongly activated, leading to extensive PARylation of chromatin-associated proteins such as histones and PARP1 itself³⁶. Macrodomain-containing proteins which are involved in DNA damage repair such as ALC1 and PARP9 have been shown to be recruited to these DNA damage sites through PAR binding with their macrodomains which can be visualised in laser micro-irradiation assays^{58,87}. In contrast to DNA damage assays using genotoxic chemicals or diffused irradiation, laser micro-irradiation assays induce DNA damage with a focused laser beam and thereby allow visualisation and measurement of kinetics and dynamics of repair factor recruitment to these laser defined target sites¹⁵². For performance of the assay, cells are typically sensitised with the halopyrimidine deoxynucleoside 5-bromo-2'-deoxyuridine (BrdU), which is incorporated into the DNA of replicating cells and is then photoactivated upon laser excitation, thereby inducing single- and double-strand breaks through released free bromine radicals. By additionally overexpressing the protein of interest fused to a fluorescent tag, recruitment of the target protein to the laser-induced DNA damage sites can be monitored under a microscope^{58,153,154}. The assay set-up even provides the possibility of monitoring PAR-dependent recruitment of macrodomains that are not necessarily directly associated with DNA damage response such as MacroD2³¹. The only requirement for this technique is a good nuclear localisation of the protein of interest. Since PARP14 macrodomains

have been reported to localise to sites of DNA damage¹⁰⁹, this assay system seemed to be suitable for studying the macrodomain-mediated PARP14 recruitment to ADP-ribosylated sites in cells.

In a first step, mutagenesis studies were performed on PARP14 MD2 to identify ADPR-binding-deficient mutants that could serve as controls for the assays. The crystal structure of PARP14 MD2 in complex with ADPR guided the design of the mutations [Figure 10]⁶⁶.

The ADPR is stabilised by several hydrogen bond interactions in a deep pocket of the protein. The adenine base forms a hydrogen bond with its nitrogen N6 to the non-conserved Gln1024 and undergoes interaction with His1172. The adenine proximal ribose unit is hydrogen

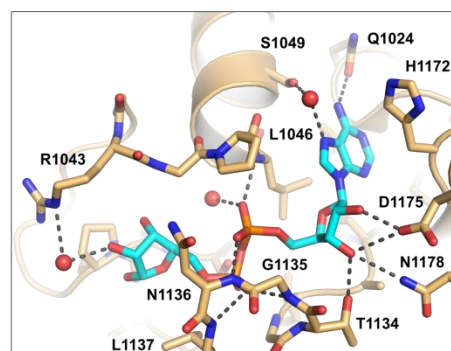


Figure 10. PARP14 MD2 in complex with ADPR (PDB ID 3Q71)⁶⁶. Critical residues of PARP14 MD2 (yellow) involved in direct interactions with ADPR (cyan) are shown as atom-coloured sticks. Hydrogen bonds between ligand and protein or structural waters are indicated by dotted lines.

bonded to Asp1175 and Thr1134. The pyrophosphates are held by loop regions, mainly interacting with the backbone NH of Leu1046, Gly1135, Asn1136 and Leu1137. The terminal ribose is involved in a complex with a water molecule together with Arg1043. Mutations were introduced either singly or combined into PARP14 MD2 that could sterically interfere with or significantly reduce binding of ADPR to the macrodomain and included Q1024A, H1172A, D1175A, D1175Y and G1044E [Figure 11A]. The latter mutation, G1044E, is the corresponding mutation of the amino acid Gly42 to glutamate in the macrodomain of Af1521 which has been previously reported to prevent ADPR binding⁶⁴. It is worth mentioning that at the corresponding position of Gly1044 most of the macrodomains, which have been shown to bind ADPR, have a glycine, an alanine (only LOC221443) or a serine (PARP15 MD2 and PARG). All of these (apart from PARG) have also at least another glycine or alanine next to it. Only macroH2A1.2, macroH2A2 and MacroD3 that don't bind ADPR have more sterically demanding or charged residues (aspartate, glutamate, asparagine) at this position, indicating that this site might be important for accommodating the ADPR. Additionally, a PARP14 MD2 mutant carrying the mutation G1055E was prepared since, although this residue is located at the surface of the macrodomain and not facing the ADPR binding site, it was published previously⁶⁶ that it prevents

ADPR binding of PARP14 MD2. It should be mentioned that all of these mutations are only designed mutations and do not correspond to any MD2 somatic mutation identified by genomic analysis which is provided in the COSMIC database for somatic mutations in cancer^{155,156}. Out of 17 generated PARP14 MD2 mutants, ten mutants were characterised in regard to their ADPR binding abilities by ITC [Figure 11B-D, S4]. The measurements showed that the substitution of Asp1175 with tyrosine had no significant effect on the ADPR binding affinity of MD2. The replacement of His1172 and in particular Glu1024 with alanine decreased the binding affinity 4- and 8- fold, respectively, due to reduced hydrogen bonding with the ligand. When Gly1044 was mutated into glutamate, no ADPR binding was detected under the assay conditions tested. It can be assumed that the glutamate side chain leads to a clash with the terminal ribose unit preventing ADPR binding to the macrodomain. The subsequently prepared PARP14 MD1 mutant with the G1044E-corresponding mutation G832E also did not show any ADPR binding under the conditions tested [Figure S4]. Finally, the PARP14 MD3_{G1257E} mutant (the MD2_{G1044E}-corresponding mutant) as well as MD2_{G1055E} were not soluble as purified proteins that would have allowed their testing for ADPR binding. Therefore, PARP14 MD1_{G832E}, MD2_{G1044E} and a MD2 mutant harbouring the two additional mutations, H1172A and D1175Y, as well as MD3_{G1257E} were selected as controls for the assays. Although MD3_{G1257E} could not be characterised *in vitro* as a purified protein, it was assumed that the mutation prevents ADPR binding of the macrodomain based on structural analysis.

Legend to Figure 11: PARP14 MD2 mutagenesis studies.

(A) Mutations introduced into the ADPR binding site of PARP14 MD2 (PDB ID 3Q71) (singly or combined) are highlighted in red. For Asp1175, only its mutation to tyrosine is shown. **(B)** SDS-PAGE of purified PARP14 MD2 mutants (1-10) that were characterised in regard to their ADPR binding ability by ITC. Alongside protein standards (M), ~2 µg of protein was loaded onto the gel and visualised with Coomassie blue staining. All mutants were purified by IMAC and gel filtration. Only for mutant 5 (G1044E) the hexahistidine tag was removed and an additional reverse IMAC purification step was carried out. Proper folding of the mutant proteins was assumed based on the observed gel filtration profiles, all showing a symmetric peak at the expected elution volumes. **(C)** Measurement of binding affinities of PARP14 MD2 wild-type and mutant proteins to ADPR by ITC. Representative titration curves show raw thermograms resulting from stepwise addition of 0.7 mM ADPR into a solution containing the corresponding macrodomain protein (35 µM) at 12 °C. The small graph in the lower part of the figure shows an overlay of the integrated thermograms (first data point corrected) that are fitted to an equilibrium binding isotherm for PARP14 MD2 wild-type and the H1172A mutant using a one-site model. **(D)** Summary of the ITC data. For the PARP14 MD2 mutants 7 and 8, 0.6 mM ADPR was titrated into a solution containing 30 µM of the macrodomain protein. For all other titrations, 0.7 mM ADPR was titrated against 35 µM of macrodomain protein. All measurements were performed at 12 °C. Additional binding and fitting parameters are provided in Figure S4.

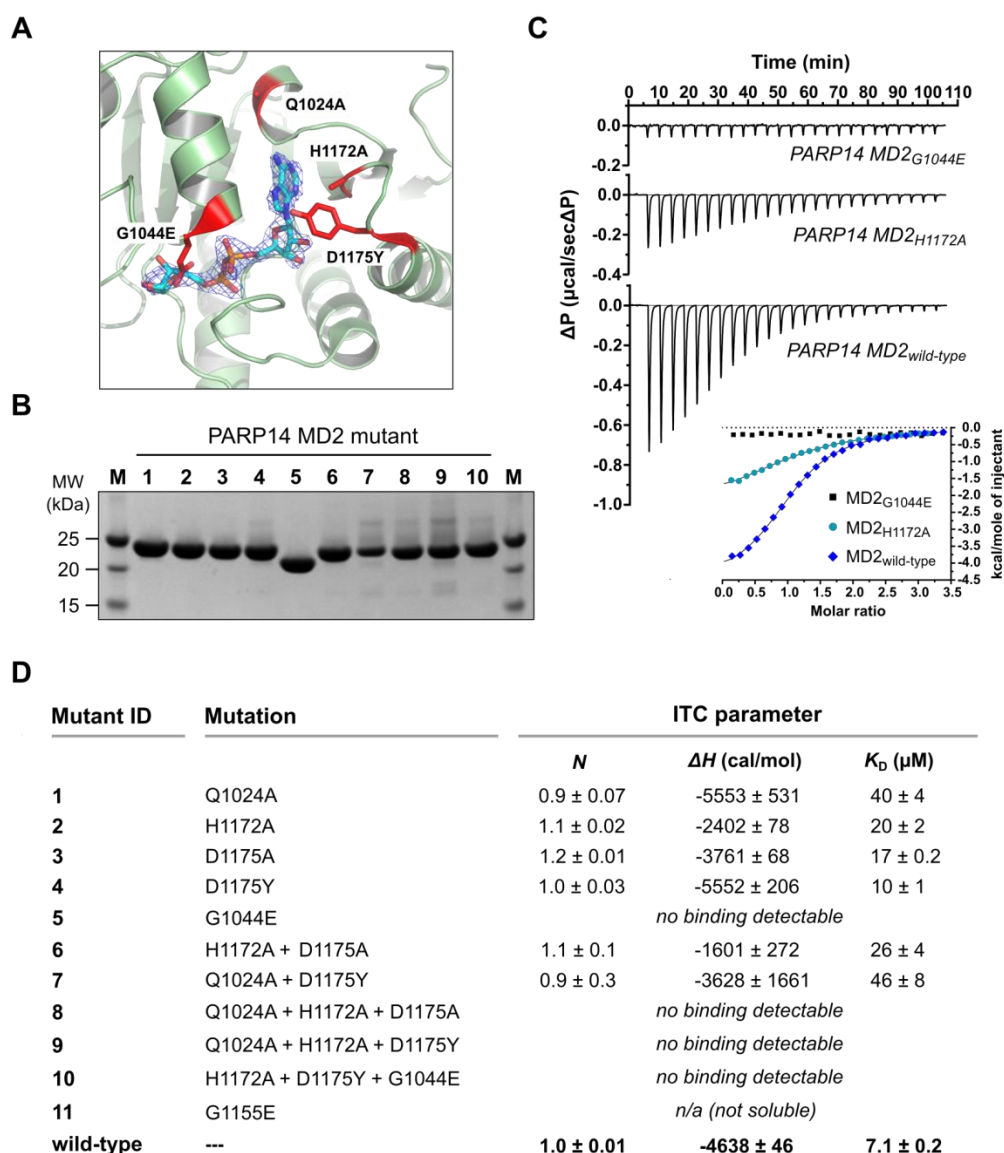


Figure 11. PARP14 MD2 mutagenesis studies. For legend, see bottom of previous page.

Following the mutagenesis studies, constructs for mammalian expression of PARP14 full-length (1-1801) and different PARP14 truncations with yellow fluorescent protein (YFP) tag were prepared by Gateway[®] cloning to enable the visualisation of the recruitment of the corresponding proteins to the laser-induced DNA damage sites [Figure 12A]. It should be mentioned that full-length PARP14 (1-1801) is demanding to transfer into new vector systems because of the highly recombinogenic nature of its first 1000 basepairs on the N-terminus, a problem known in the PARP field. Numerous cloning attempts including trying different cloning techniques and optimisations were necessary to finally obtain the mutation free YFP-PARP14 (1-1801) encoding plasmid [Figure 12B]. Since a strong nuclear localisation of the proteins is important for

performance of the assay, a nuclear localisation signal (NLS) was additionally attached to the C-terminus of the YFP-PARP14 truncations, the single macrodomain MD2 and the triple macrodomain constructs encoding MD1 to MD3, in order to promote their up-take into the nucleus. The nuclear transport of macromolecules is mediated by members of the karyopherin transport receptor family which either import ('importins') or export ('exportins') their cargo through the nuclear pore complex¹⁵⁷. The transport mechanisms are driven through recognition of nuclear localisation signals (NLS) or nuclear export signals (NES) by the respective karyopherin members¹⁵⁸. For promoting the up-take of the macrodomain proteins, a classical monopartite NLS (PKKKRKV) from the simian virus 40 (SV40) large T antigen was used, which is characterised by a single cluster of five consecutive basic amino acids and is recognised by importin- α ¹⁵⁹. It should be mentioned that analysis of full-length PARP14 using bioinformatics NLS and NES prediction tools did not identify any NES in PARP14, however, a NLS sequence was identified between MD3 and the WWE domain, predicting PARP14 localisation to the nucleus and cytoplasm as actually confirmed by experimental data. Furthermore, mammalian expression constructs encoding the ADPR-binding-deficient macrodomains MD1_{G832E}, MD2_{G1044E} or MD2_{G1044E, H1172A, D1175Y} and MD3_{G1257E} were prepared. All constructs were confirmed to result in correct protein expression [Figure 13A] and the localisation of the gene products was subsequently determined in U-2 OS cells transiently expressing full-length PARP14 or PARP14 truncations [Figure 13B]. U-2 OS cells were used for these studies since this cell line was to be used for the subsequent laser micro-irradiation assays. The single macrodomain, PARP14 MD2, localised almost exclusively to the nucleus, very likely due to its small size and the attached NLS promoting its nuclear up-take. The initial MD1-3 constructs with and without NLS resulted in rather poor nuclear localisation of the proteins. Therefore, the position of the NLS was optimised by introducing additional residues between the MD3 C-terminus and the NLS. A spacer of five amino acids was sufficient to provide good accessibility and recognition of the NLS, resulting in improved nuclear up-take of MD1-3 protein. Unfortunately, the PARP14 'macroPARP' truncation containing the three macrodomains, the WWE and PARP domain as well as full-length PARP14 did not localise strongly to the nucleus and showed additionally lower expression levels

compared to the other constructs [Figure 13A, B]. Furthermore, bright fluorescent bodies were observed in the cytoplasm indicating the formation of aggregates of the macroPARP and full-length PARP14 proteins. It could very likely be that the proteins, due to their size (141 kDa for YFP-macroPARP and 231 kDa for YFP-PARP14), misfolded or were partially unfolded upon overexpression, leading to protein aggregation and also consequently to dysfunction. The size and the associated protein aggregation were probably also the factors contributing to the poor nuclear uptake. Thus, only the constructs encoding PARP14 MD2 (with NLS) as well as MD1-3 (with spacer and NLS) showed sufficient nuclear localisation to allow analysis of the PARP14 macrodomain recruitment to laser-induced DNA damage sites.

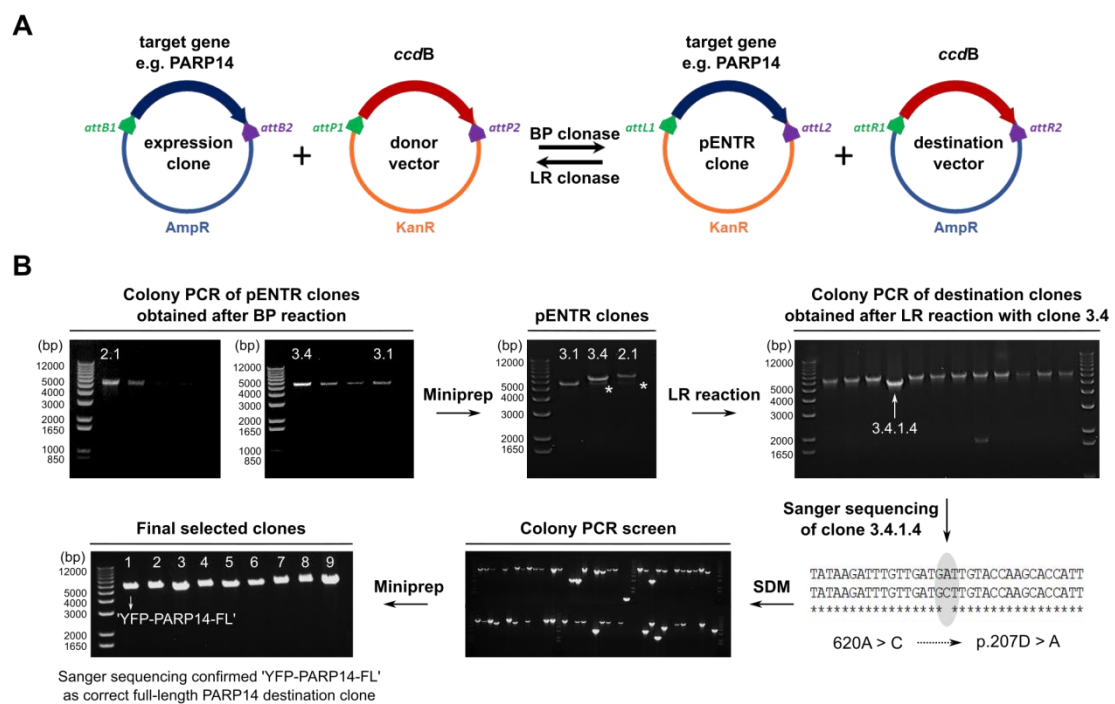


Figure 12. Preparation of PARP14 mammalian YFP-expression constructs using Gateway® cloning. (A) Schematic illustrating the Gateway® cloning technology. In brief, the PARP14 regions of interest were first PCR amplified with *attB* recombination sites (this step was omitted for full-length PARP14 cloning). PCR products (or vectors) were then recombined with the pDONR™221 vector containing the *attP* sites using the BP clonase mix where the *attB* and *attP* sites serve as binding (attachment) sites for the recombination enzymes in this mix. The recombination reaction results in pENTR clones containing *attL* sites which can be recombined with the *attR* sites present in the YFP-destination vector using the LR clonase enzymes, thereby creating the final expression clone in the destination vector. (B) Steps of final full-length PARP14 Gateway® cloning. Following the BP reaction and a colony PCR screen, PARP14 inserts of three positive pENTR clones were PCR amplified after plasmid isolation showing the highly instable and recombinogenic nature of the PARP14 gene. The LR reaction was set-up with all three clones (containing mixtures of the PARP14 genes (the additional weak bands in 3.4 and 2.1 are indicated with *). PARP14 inserts of destination clones obtained after the LR recombination reaction were then PCR amplified which revealed only one clone (3.4.1.4) originating from the 3.4 pENTR clone with correct size, however, with a mutation in the recombinogenic region of PARP14. Following several site-directed mutagenesis attempts, a correct clone (labelled 'YFP-PARP14-FL' on the gel) could be isolated and was confirmed by Sanger sequencing. All samples shown were run on a 1 % agarose gel and visualised using pegGreen DNA/RNA dye under UV light.

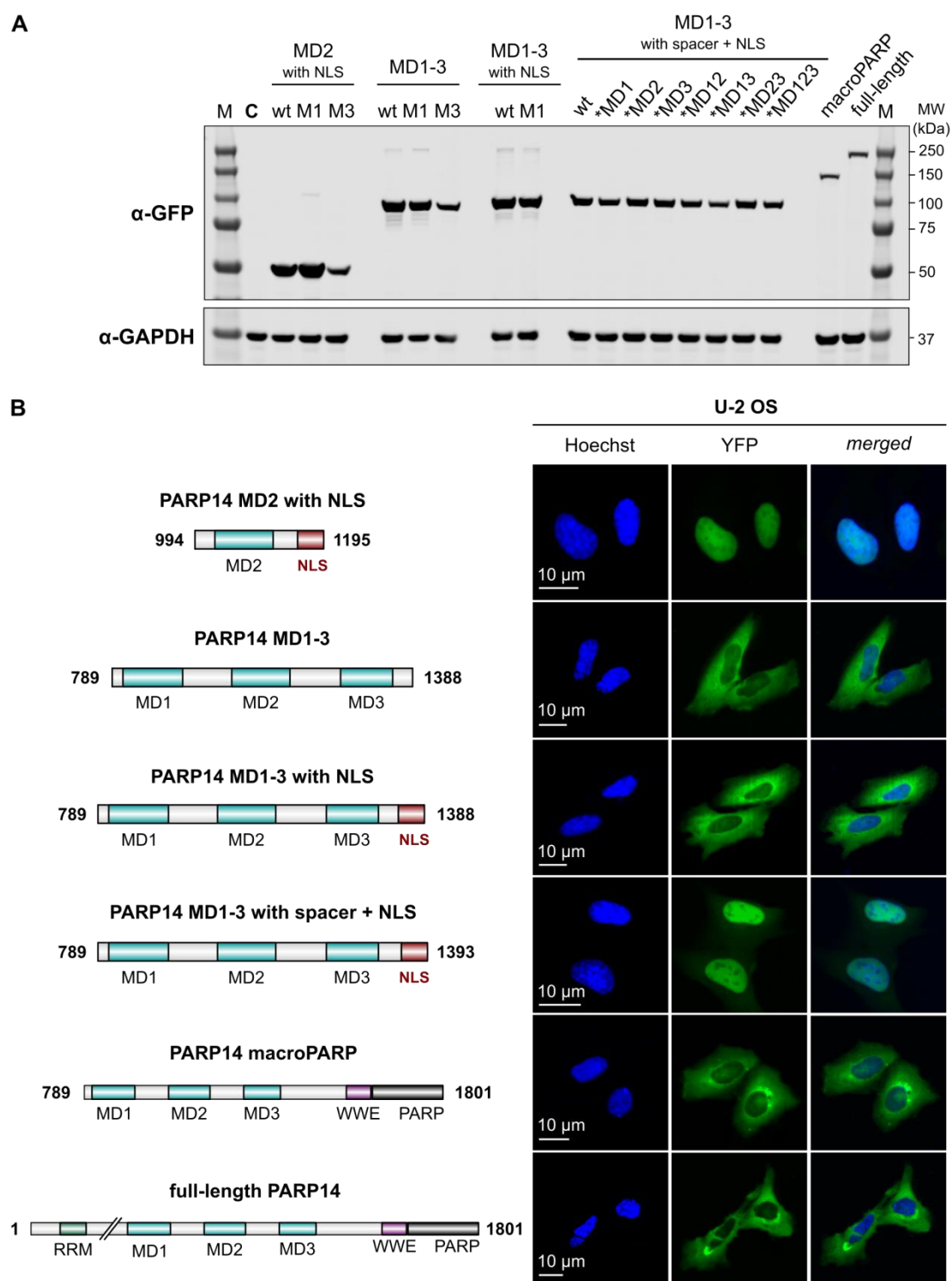


Figure 13. Test expression and localisation studies of YFP-PARP14 full-length and truncation constructs. (A) HEK293 cells were either left untransfected (control, C) or were transfected with the indicated YFP-PARP14 full-length and truncation constructs and analysed for expression products by Western blotting 48 h post transfection. All constructs resulted in expression of the proteins with the expected sizes, however, PARP14 macroPARP and full-length PARP14 showed low expression levels. Alongside of protein standards (M), 40 μg of total lysate were loaded for analysis of the macroPARP and full-length construct and 20 μg of total cell lysate for all other constructs. (B) The cellular localisation of the expression products was tested in U-2 OS cells. Representative images for each construct type are shown. Only the single MD2 construct including NLS and the triple macrodomain construct with spacer and NLS showed good nuclear localisation and could be used for laser micro-irradiation studies. The PARP14 macroPARP and full-length construct showed bright cytoplasmic bodies indicating protein aggregation. For imaging, cells were fixed with 4 % formaldehyde 48 h post transfection and nuclei were counterstained with Hoechst33342.

The following studies were performed in collaboration with the Dr. Ivan Ahel's Laboratory (Sir William Dunn School of Pathology, Oxford). In order to establish the laser micro-irradiation assays, experiments were first performed with the single macrodomain, PARP14 MD2 [Figure 14]. For this, BrdU sensitised U-2 OS cells were transfected with YFP-PARP14 MD2 and subjected to laser micro-irradiation coupled to live-cell imaging. It was observed that PARP14 MD2 wild-type is recruited to these laser-defined DNA lesions within seconds after induction of DNA damage in contrast to the corresponding ADPR-binding-deficient mutants, PARP14 MD2_{G1044E} and PARP14 MD2_{G1044E, H1172A, D1175Y} [Figure 14A]. This step also further validated the ADPR binding studies previously performed with the purified mutant proteins. Furthermore, the PARP14 MD2 recruitment was dependent on both the presence and activity of PARP1 [Figure 14B, C]. In the absence of PARP1, in PARP1^{-/-} cells, and after inhibition of PARP1 catalytic activity using the PARP domain inhibitor olaparib, recruitment of the macrodomain was not detected in comparison to the corresponding controls.

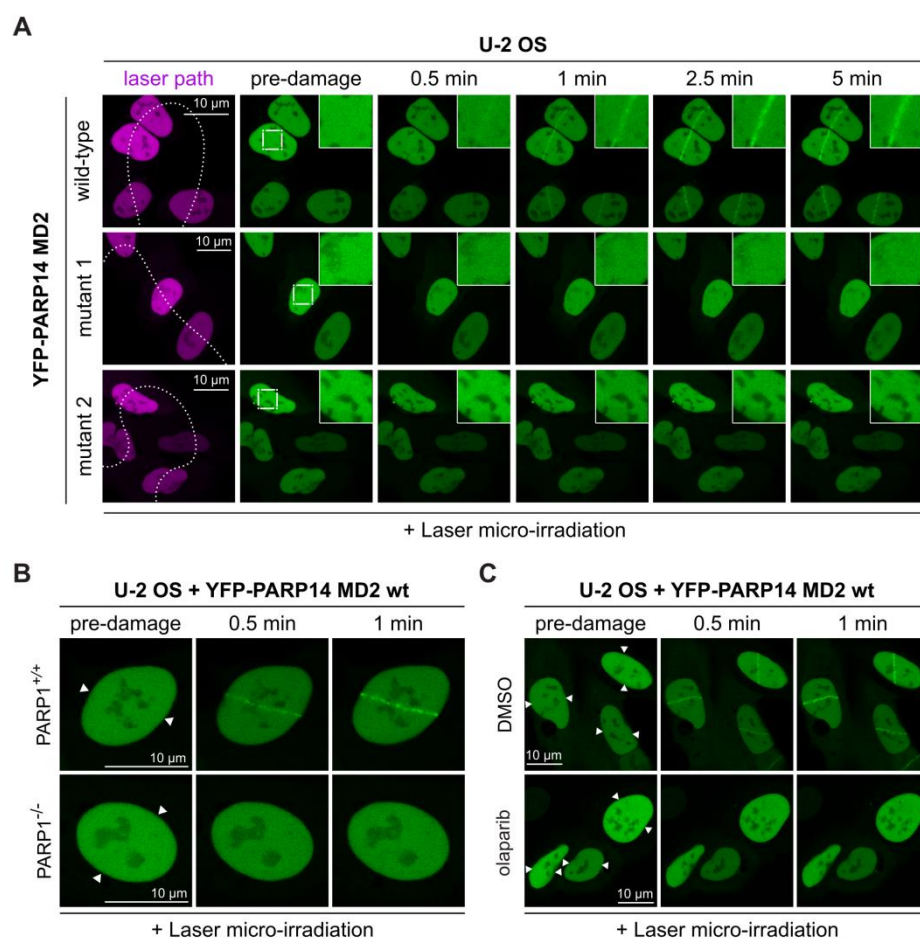


Figure 14. PARP14 MD2 is recruited to laser-induced DNA damage sites. For legend, see top of next page.

Legend to Figure 14: PARP14 MD2 is recruited to laser-induced DNA damage sites.

(A) PARP14 MD2 wild-type is recruited to sites of laser-induced DNA damage in contrast to the corresponding ADPR-binding-deficient mutants, mutant 1 (PARP14 MD2_{G1044E}) and mutant 2 (PARP14 MD2_{G1044E, H1172A, D1175V}). U-2 OS cells were transfected with YFP-PARP14 MD2 wt or mutant, subjected to laser micro-irradiation and imaged at the indicated times. **(B)** PARP14 MD2 recruitment depended on the presence of PARP1. U-2 OS/PARP1^{+/+} and U-2 OS/PARP1^{-/-} cells were transfected with YFP-PARP14 MD2 wt, subjected to laser micro-irradiation (white arrows) and imaged at the indicated times. **(C)** PARP14 MD2 recruitment depended on the activity of PARP1. U-2 OS cells transiently expressing YFP-PARP14 MD2 wt were pre-treated for 1 hour with DMSO or with 10 μ M of the PARP1 inhibitor olaparib, subjected to laser micro-irradiation (white arrows) and imaged at the indicated times. **Note:** These experiments were performed by Dr. Ian Gibbs-Seymour (Dr. Ivan Ahel's Laboratory, Sir William Dunn School of Pathology, Oxford).

Subsequently, laser micro-irradiation assays were performed with the YFP-PARP14 MD1-3 proteins summarised in Figure 15A. The wild-type triple macrodomain protein MD1-3 showed strong recruitment to laser-induced DNA damage sites. Furthermore, the introduction of a mutation resulting in macrodomain ADPR-binding deficiency only in a single macrodomain in MD1-3 did not significantly impair the recruitment of MD1-3. The ADPR-binding deficiency of MD1 or MD3 was presumably compensated by the other still functional macrodomains. If only MD2 was mutated, a very minor decrease in recruitment efficiency was observed. Next, the introduction of an additional mutation, thus resulting in only one functional macrodomain capable of ADPR binding in MD1-3, allowed directly studying which of the macrodomains contributes strongest to the MD1-3 recruitment. It was observed that MD2 had the strongest effect on MD1-3 recruitment followed by MD3 and then MD1. The ability of MD1 to recruit MD1-3 to DNA damages sites was very poor, consistent with its low ADPR binding affinity previously determined by calorimetric studies. It was seen that the simultaneous inhibition of MD2 and MD3 resulted in significant impairment of MD1-3 recruitment to ADP-ribosylated sites. The stronger recruitment of MD1-3 by MD2 in comparison to MD3 was also in line with the previous *in vitro* measurements, confirming the higher ADPR binding affinity of MD2 in comparison to MD3. These findings therefore also do not suggest alternate binding affinities of MD2 and MD3 *in vitro* and in cells as stated by Forst *et al.*⁶⁶. Finally, only inhibition of all three macrodomains completely prevented MD1-3 recruitment to the laser-induced DNA damage sites, which confirmed the consistency of the dataset and validated the MD3_{G1257E} mutation. Despite their poor

nuclear localisation, PARP14 macroPARP and full-length PARP14 were also tested in this set-up [Figure 15B]. Weak re-localisation of PARP14 macroPARP to laser-induced DNA damage sites was observed, however, no re-localisation was detected for full-length PARP14 as expected.

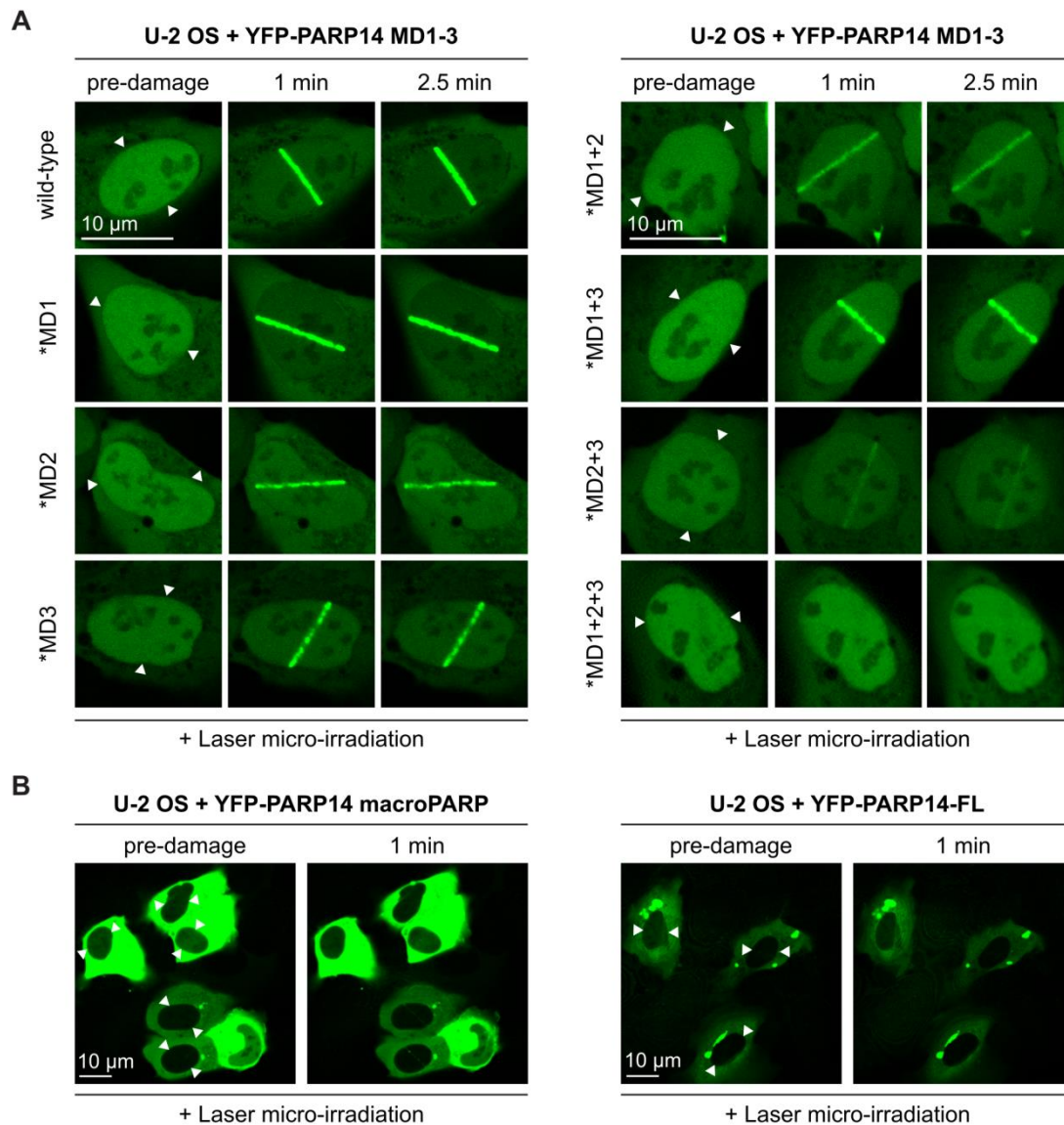


Figure 15. PARP14 recruitment studies to laser-induced DNA damage sites. (A) U-2 OS cells were transfected with YFP-PARP14 Macro1-3 wt or mutants, subjected to laser micro-irradiation (white arrows) and imaged at the indicated times. Macrodomains harbouring a mutation resulting in ADPR-binding deficiency are indicated with *. All macrodomains contribute to the recruitment of PARP14 to the DNA damage sites with MD2 having the strongest recruitment efficiency. Recruitment of PARP14 was only completely prevented upon mutation of all three macrodomains. (B) U-2 OS cells were transfected with YFP-PARP14 macroPARP (left) or YFP-PARP14-FL (right), subjected to laser micro-irradiation (white arrows) and imaged at the indicated times. The macroPARP protein showed very weak recruitment and no recruitment was observed of full-length PARP14. Green cytoplasmic bodies probably indicated protein aggregation resulting in poor nuclear localisation and dysfunction. **Note:** The experiments were performed by Dr. Ian Gibbs-Seymour (Dr. Ivan Ahel's Laboratory, Sir William Dunn School of Pathology, Oxford).

These laser micro-irradiation experiments gave rise to the following conclusions: PARP14 MD2 and therefore also PARP14 MD1-3 are recruited to laser-induced DNA lesions in a PAR/PARP1-dependent manner, thus recognising PARylated target sites. Yet, this result is in contrast to Forst *et al.*⁶⁶ who defined PARP14 macrodomains as readers of MARYlation and stated that the PARP14 macrodomains, in particular MD2 and MD3, recognise MARYlated targets such as PARP10 but not PARylated PARP1. It cannot be excluded that MARYlation events are involved in PARP14 macrodomain recruitment upon laser-induced DNA damage, however, these data showed that MD2 recruitment to DNA damage sites was not driven by MARYlation. No PARP14 MD2 recruitment was observed in U-2 OS/PARP1^{-/-} cells which still have functional PARP family members capable of MARYlation of respective target sites and PARP1 is well known for its PARylation, not MARYlation, activity in particular in the context of DNA damage. These results were additionally confirmed by pulldown experiments [Figure 16], where PARP14 MD2 was shown to pull-down PARylated PARP1, whose identity and modification was confirmed by mass spectrometry and anti-PAR Western blotting, respectively. Since PARP14 macrodomains were however shown to be unable to recognise PARylated PARP1 directly in *in vitro* pulldown assays performed with purified and automodified PARP1⁶⁶, it could be possible that the PARP14

macrodomain recognition of PARylated PARP1 in cells is only indirectly. It could be hypothesised that a complex is formed involving the macrodomain, PARP1 and a MARYlated target protein that is bound to PARP1. MARYlation by another PARP member of the intermediate target protein may have dependency on its prior PARylation by PARP1.

Consistent with the *in vitro* studies, MD2 contributes strongest to the PARP14 recruitment to ADP-ribosylated sites, followed by MD3. The recruitment efficiency of MD1 is rather negligible. PARP14 macrodomain inhibitor development should therefore address MD2, however since MD2

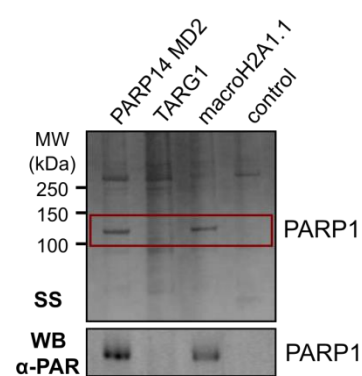


Figure 16. PARP14 MD2 interacts with PARylated PARP1. Whole cell lysates from H₂O₂-treated Jurkat cells were incubated with macrodomain baits immobilised on Dynabeads®. Proteins isolated by pull-down were eluted with 1 mM ADPR and visualised by silver staining (SS) and α-PAR Western blotting (WB). Control: no macrodomain bait immobilised. TARG1 hydrolases PAR from PARP1.

inhibition can be almost completely compensated by MD3, an additional inhibitor targeting MD3 (or a dual MD2/MD3 inhibitor) is required to significantly impair PARP14 recruitment and thus function on its target sites.

2. Targeting PARP14 macrodomains

As discussed in chapter 1, efficient PARP14 targeting via its macrodomains to prevent its recruitment to ADP-ribosylated target sites may require inhibitor development for both MD2 and MD3. The first part of this chapter therefore has its focus on the progress on inhibitor development for MD2 while the second part shows the first attempts for MD3 targeting.

2.1. Inhibitor development for PARP14 MD2

To understand PARP14 MD2 targeting in more detail, the new macrodomain inhibitor GeA-69 was firstly characterised concerning its biophysical properties, selectivity over human macrodomain family members, and its target engagement in cells. Secondly, GeA-69 analogues, which were designed and synthesised in collaboration with chemistry groups, were characterised to develop a detailed structure-activity-relationship (SAR) analysis of GeA-69 in order to improve its potency towards PARP14 MD2. Lastly, a medium-sized *in vitro* screen was performed against PARP14 MD2 in collaboration with Novartis (NIBR, Basel) with the hope of identifying small molecule inhibitors with new chemical scaffolds amenable for binding the ADPR site of this macrodomain.

2.1.1. GeA-69 – a new class of macrodomain inhibitors targeting PARP14 MD2

GeA-69 was discovered in my rotation project prior to this work by screening 3000 compounds of an in-house kinase inhibitor library against PARP14 MD2 along with three other macrodomains, MacroD1, MacroD2 and TARG1. The screen was performed using an AlphaScreen (amplified luminescence proximity homogenous assay) technology based assay that was developed in this project specifically for macrodomain screening. The assay principle is described in Figure 17.

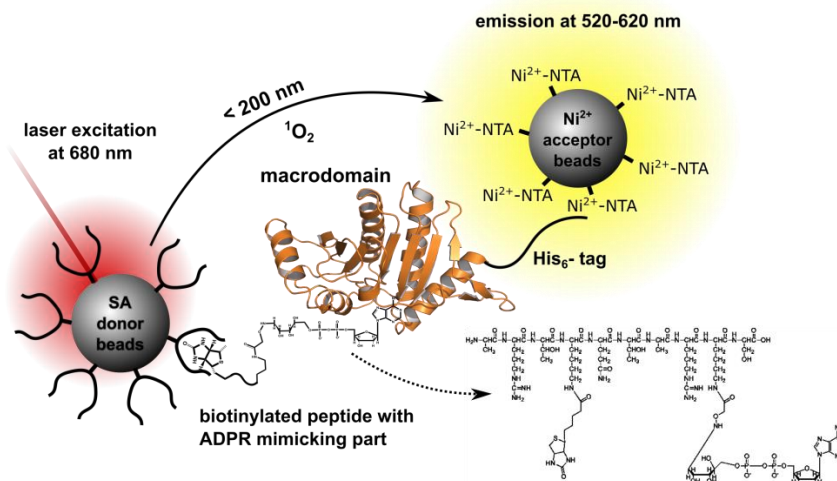


Figure 17. Principle of the AlphaScreen-based displacement binding assay developed for macrodomain screening.

The macrodomain AlphaScreen peptide (displayed in the right bottom corner) binds via its biotinylated lysine to streptavidin coated donor beads whereas the macrodomain protein binds to Ni²⁺-chelated acceptor beads with its hexahistidine tag. Binding of the macrodomain to the ADPR imitating part of the peptide enables an energy transfer from the donor to the acceptor beads that results in a measurable AlphaScreen signal. More specifically, donor beads convert ambient oxygen to singlet oxygen upon illumination at 680 nm which can diffuse up to 200 nm in solution and results in light emission at 520-620 nm from the thereby excited acceptor beads. Macrodomein inhibitors are able to displace the peptide from the macrodomain binding site resulting in a reduced emission. **Note:** The macrodomain AlphaScreen peptide, a biotinylated and ADP-ribosylated 11 residue peptide derived from histone H3, was designed by Dr. Jonathan Elkins and has the following sequence: ARTK(Bio)QTARK(Aoa-RADP)S.

The rationale behind a kinase inhibitor screen was that macrodomains as binding domains of ADPR, which is structurally closely related to ATP, may show activity towards ATP-competitive kinase inhibitors or fragments. GeA-69 is a carbazole-based compound with an acetaminophenyl substituent attached to ring A and only showed inhibitory activity towards PARP14 MD2 out of the four macrodomains screened. ITC experiments verified its binding to PARP14 MD2 with a K_D of (862 ± 40) nM in a 1:1 binding stoichiometry. In contrast, a closely related compound MnK2-68 (developed during the following SAR studies) did not bind and was used as a negative control in subsequent assays [Figure 18, S8].

Legend to Figure 18: GeA-69 as a new class of macrodomain inhibitors.

(A) Molecular structure of GeA-69 discovered in a kinase inhibitor library screen and its inactive analogue MnK2-68 developed during the following SAR studies. **(B)** ITC confirmed binding of GeA-69 to PARP14 MD2 with a K_D of (862 ± 40) nM in a 1:1 binding stoichiometry. Raw injection heats are shown for a blank titration of PARP14 MD2 into buffer (top), and titrations of PARP14 MD2 into solutions of the inactive analogue MnK2-68 (middle) and the active compound GeA-69 (bottom). The inset shows normalised binding enthalpies corrected for the heat of dilution as a function of binding site saturation. The ITC parameters for GeA-69 binding to PARP14 MD2 are provided in Figure S8. **(C)** Schematic representation of the PARP14 domain architecture. Residue numbers refer to the human PARP14 ortholog. The macrodomain targeted by GeA-69 is highlighted with a red star.

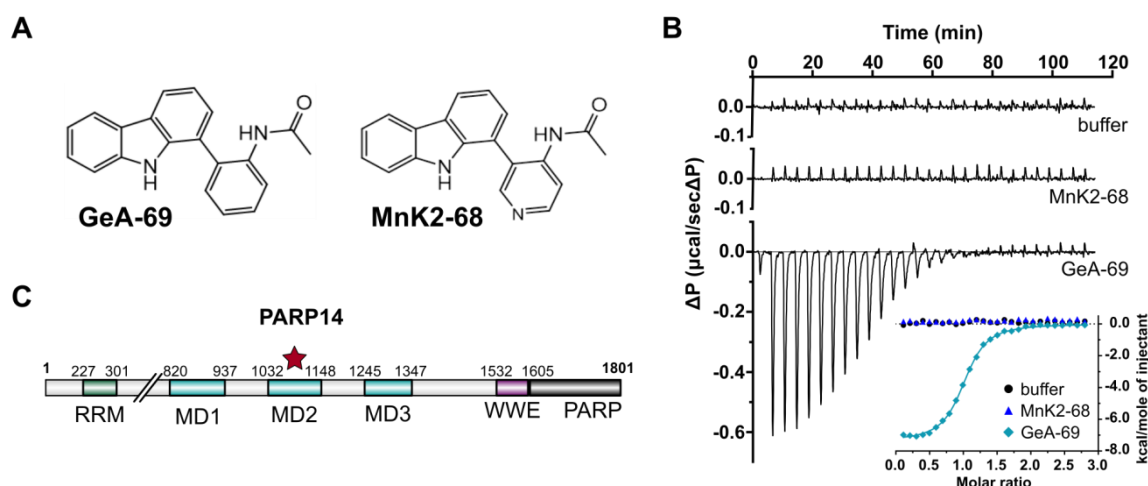


Figure 18. GeA-69 as a new class of macrodomain inhibitors. For legend, see bottom of previous page.

To establish the binding mode of GeA-69, the co-crystal structure of PARP14 MD2 with GeA-69 was determined by members of the group. Initial attempts at co-crystallisation of GeA-69 with wild-type PARP14 MD2 gave crystals that diffracted X-rays only to modest resolution. Therefore, five different surface-entropy reduction (SER) mutants of PARP14 MD2 were prepared. In these SER constructs, up to five lysine residues on the protein surface were truncated to serine residues, to improve the likelihood of forming tighter crystal contacts¹⁶⁰. Each of these mutant proteins was used in crystallisation trials with either GeA-69 or closely-related analogues. Finally, a construct of the PARP14 MD2_{K1048S, K1154S, K1158S, K1162S} mutant gave a crystal diffracting to 1.6 Å resolution in complex with compound MnK2-13 which was well resolved in the electron density map [Figure 19A, S5A]. MnK2-13 is the methanesulfonamide analogue of the acetamide GeA-69 and binds to PARP14 MD2 with a K_D of $(2.1 \pm 0.1) \mu\text{M}$ [Figure S5B]. Surprisingly MnK2-13 did not occupy the ADP-ribose binding site but was deeply buried within the macrodomain with no contact with the solvent phase [Figure 19A]. Ligand binding was stabilised by hydrophobic and hydrogen bonding interactions to the protein [Figure 19B]. The planar carbazole itself was bound in a hydrophobic pocket with several hydrophobic interactions including the T-shaped π - π interaction with ring C and Phe1129 and with an additional hydrogen bond of the NH of the carbazole to the proline carbonyl of Pro1130. The benzene ring of MnK2-13 was surrounded by

hydrophobic residues including Phe1182, Leu1171 and Phe1169. The sulfonamide moiety of MnK2-13 formed several hydrogen bonds to the macrodomain including to the amide of Ile1132 and the primary carboxamide of the Asn1178 side chain. Furthermore, the NH of the sulfonamide group was involved in a complex with a water molecule together with the Asn1178 carbonyl group and Lys1141. Superimposition of the ADPR-bound macrodomain structure (PDB ID 3Q71) and the MnK2-13-bound structure furthermore showed that Phe1144 was pushed aside by the carbazole of the inhibitor. Ring A of the carbazole may therefore be crucial for enabling the binding of the inhibitor to the protein. Using this MnK2-13 co-crystal structure, GeA-69 was then docked into PARP14 MD2 to rationalise its slightly higher binding affinity compared to MnK2-13. The docking model may suggest that the orientation of the carbonyl in the acetamide group in GeA-69 may be ideal (considering bond length and angle) to allow a stronger hydrogen bond interaction to Asn1178 in comparison to the sulfonamide oxygen of MnK2-13 [Figure S6]. The inactivity of the negative control, MnK2-68, is even more difficult to rationalise based on this model. Potentially the increase in polarity and differences in electron distribution in the pyridine ring system forms less favourable hydrophobic interactions with PARP14 MD2. Interestingly, MnK2-13 bound adjacent to but not in the ADPR binding site as would have been expected based on its carbazole scaffold that was used for the development of ATP mimetic inhibitors¹⁶¹. The structure superimposition revealed that upon inhibitor binding a loop from Pro1130 to Pro1140 was pushed into the ADPR binding site in the inhibitor complex displacing the bound ADPR [Figure 19C]. A portion of this loop consisting of residues Gly1133-Gly1135 showed weaker electron density and may thus be present in multiple conformations. ITC competition experiments confirmed GeA-69 as an inhibitor of PARP14 MD2 able to compete with its ADPR binding. In the presence of 80 μM GeA-69, binding of ADPR to the macrodomain was not detected while in absence of the inhibitor ADPR bound with a K_D of $(7.8 \pm 0.5) \mu\text{M}$ [Figure 19C, S17].

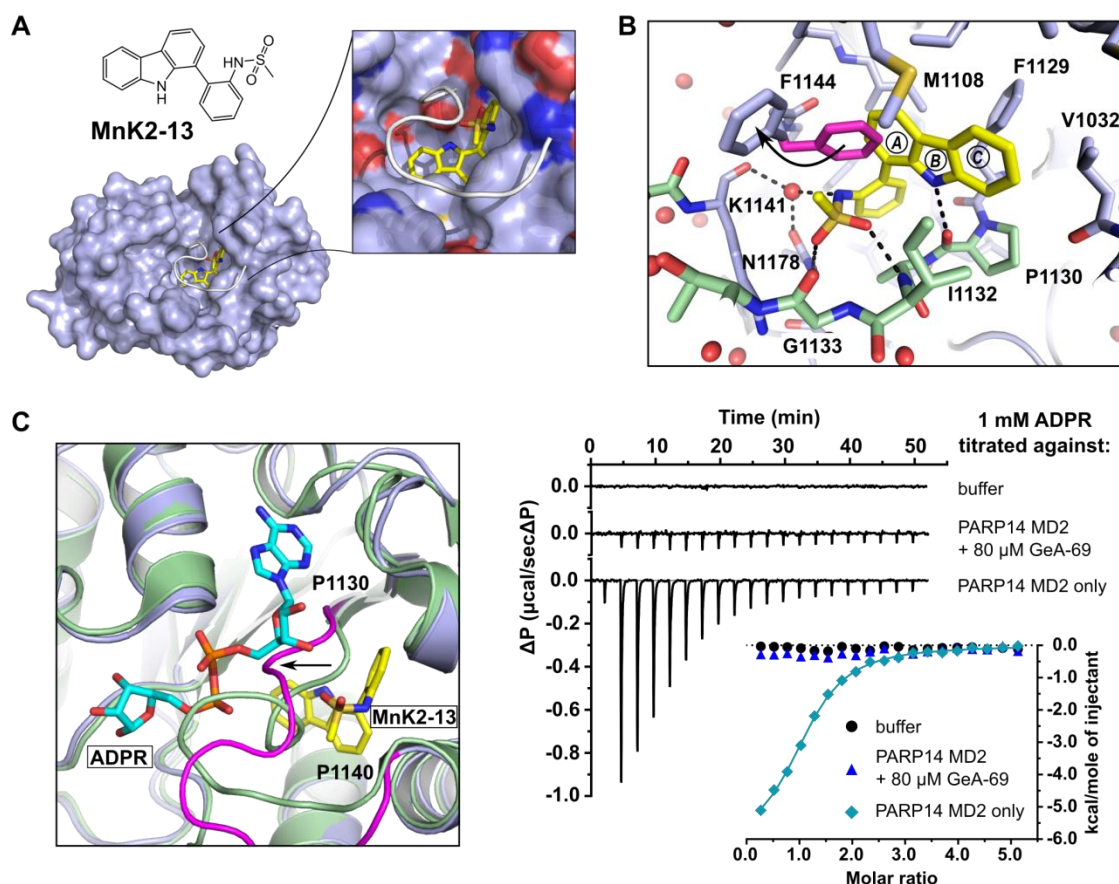


Figure 19. Crystal structure of PARP14 MD2 in complex with MnK2-13 reveals an allosteric binding mode. (A) Surface representation of PARP14 MD2 in complex with the macrodomain active GeA-69 analogue MnK2-13. The macrodomain inhibitor (atom-coloured sticks) is deeply bound within the macrodomain. The solvent accessible surface of the residues Ala1131-Pro1140 (white loop) was removed to visualise the inhibitor. **(B)** The binding site of MnK2-13 in PARP14 MD2. Residues involved in direct contacts with the inhibitor are shown in atom-coloured sticks. Hydrogen bonds between ligand and protein or structural waters are indicated by dotted lines. Superimposition of the ADPR-bound macrodomain structure (PDB ID 3Q71)⁶⁶ and inhibitor-bound macrodomain structure shows that Phe1144 (magenta) is pushed aside by the carbazole ring system of MnK2-13 (arrow). The part of the loop (Pro1130-Pro1140) that is rearranged by comparison to the ADPR-bound structure is shown in green. **(C)** GeA-69 is an allosteric inhibitor of ADPR binding to PARP14 MD2. **(Left)** Superimposition of the ADPR-bound macrodomain structure (PDB ID 3Q71, green)⁶⁶ with the inhibitor-bound macrodomain structure (blue) revealed GeA-69 (and MnK2-13) as allosteric inhibitor. Binding of MnK2-13 (macrodomain active analogue of GeA-69) to the macrodomain results in movement of the loop Pro1130-Pro1140 (highlighted in magenta) into the ADPR binding site displacing the bound ADPR. **(Right)** ITC experiments confirmed GeA-69 as inhibitor of PARP14 MD2 competing with its ADPR binding. Raw injection heats are shown resulting from titration of a 1 mM ADPR solution into the cell containing either buffer (top) or 40 μM PARP14 MD2 only (bottom) or 40 μM PARP14 MD2 in presence of 80 μM GeA-69 (middle). Only in absence of the inhibitor, the expected K_D for ADPR of $(7.8 \pm 0.5) \mu\text{M}$ was observed. Additional binding and fitting parameters are provided in Figure S17. **Note:** PARP14 MD2 co-crystallisation with MnK2-13 was performed by Kristin Uth¹⁶², structure determination by Dr. Jonathan Elkins (SGC).

GeA-69 selectivity over the other human macrodomain family members was then assessed. For the majority of the human macrodomains, purification systems were already established prior to this work and the proteins resulting from these purifications were directly used for selectivity studies. Only four macrodomain proteins, ALC1, PARG, PARP9 MD1 and PARP15 MD1, were

not available and therefore prepared, of which PARP15 MD1 was most demanding in terms of its purification as discussed in Chapter II.2.1.2. To perform the selectivity studies, the macrodomain AlphaScreen assay was considered as method of choice due to its simplicity, throughput, low protein consumption and the possibility to screen at higher compound concentrations. Therefore, all macrodomains were tested (if not done so before) for their suitability in the macrodomain AlphaScreen assay. Ideal assay concentrations were determined in dose-response experiments for all of the macrodomains [Figure S7]. Including PARP14 MD2, nine of the 16 human macrodomains gave a consistent and stable assay signal with a Z' on average of 0.85 and could thus be assayed by AlphaScreen using the macrodomain AlphaScreen peptide. Due to the design of the macrodomain AlphaScreen assay, macrodomains lacking ADPR binding capability such as MacroD3, macroH2A1.2 and macroH2A2 could not be characterised using this assay. Also the ADPR binding proteins PARP15 MD2, PARG and ALC1 could not be successfully established using this assay format. No assay signal was obtained for PARP15 MD2¹⁶³ and ALC1 [Figure S7], for which binding to the ADPR imitating part of the peptide may have been significantly decreased due to the method of conjugation to the peptide resulting in a ring-opened terminal ribose unit, or alternatively due to the presence of the attached peptide sequence. ALC1 showed slight instability during the purification process; therefore it cannot be excluded that the protein was partially unfolded preventing efficient peptide binding. In the case of PARG, an AlphaScreen signal could be obtained even in the absence of the macrodomain AlphaScreen peptide, potentially indicating non-specific binding of PARG to the streptavidin coated donor beads [Figure S7]. ITC or biolayer interferometry (BLI) were therefore considered as alternative screening strategies for macrodomains that could not be characterised using the AlphaScreen assay. Finally, eleven of the fifteen human macrodomain proteins were successfully tested for GeA-69 selectivity either by the macrodomain AlphaScreen assay at a compound concentration of 200 μ M, by BLI at a compound concentration of 50 μ M or by ITC at a compound concentration of 30 μ M [Figure 20A, S8]. Binding was only detected to PARP14 MD2 proving the very high target selectivity of GeA-69 over the macrodomain family. Although the macrodomain fold is highly conserved, at the amino acid level sequence divergence is high, and structural and

sequence comparison revealed a number of differences around the binding site that provide a rationale for the selectivity of GeA-69 [Figure 20B, C]. It should be mentioned that for three of the remaining macrodomains, PARG, ALC1 and macroH2A2, the selectivity could not be determined since the proteins were not available with tag suitable for BLI testing or were too unstable over the duration of the ITC experiment. Selectivity of GeA-69 over PARP15 MD1, the macrodomain with highest structural similarity to PARP14 MD2, was also tested and the results are discussed in Chapter II.2.1.2.

Since GeA-69 was part of a kinase inhibitor library, originally synthesised in a programme aimed at the optimisation of the activities of the alkaloid annomontine^{164,165}, the internal database was searched for experimental thermal shift assay data of kinases that were screened against this compound. 46 kinases representing different branches of the phylogenetic kinase tree were screened [Figure S9], none of them showing any significant stability shifts and thus no binding affinity for this inhibitor.

Legend to Figure 20: GeA-69 is a selective macrodomain inhibitor.

(A) GeA-69 shows target selectivity over human macrodomain family members. Selectivity was assessed for 12 out of 16 human macrodomains (shown by circles) by the macrodomain AlphaScreen assay at a compound concentration of 200 μ M (top right) or by biolayer interferometry (BLI) at a compound concentration of 50 μ M (bottom right). Binding was only detected with PARP14 MD2 (red circle). The numbers in the circles correspond to macrodomains shown on SDS-PAGE (bottom left). PARP14 MD1, 2 and 3 proteins were prepared with different tags. Details are provided in Figure S20. **(B)** Multiple sequence alignment of human macrodomains confirming no other human macrodomain has conserved all the residues that bind MnK2-13 and therefore also GeA-69. Residues of PARP14 MD2 involved in MnK2-13 interaction are highlighted in green, numbers on top of the alignment refer to the human PARP14 sequence. Box colour legend: (blue) identical amino acid, (light blue) similar amino acid, (brown) amino acid likely to sterically interfere with inhibitor binding, (light brown) amino acid not favouring the inhibitor binding like the corresponding PARP14 MD2 residue based on its physicochemical properties. The full alignment is shown in Figure S10. **(C)** GeA-69 selectivity over representative human macrodomain family members based on structural considerations. The MnK2-13-bound PARP14 MD2 structure (shown in blue) was superimposed to the corresponding macrodomain MacroD1 (PDB ID 2X47)⁵⁸, macroH2A2 (PDB ID 2XD7) or TARG1 (PDB ID 4J5R)³¹ structure. Residues probably causing a sterical clash and thus preventing GeA-69 or MnK2-13 binding, respectively, to the macrodomain are highlighted in magenta.

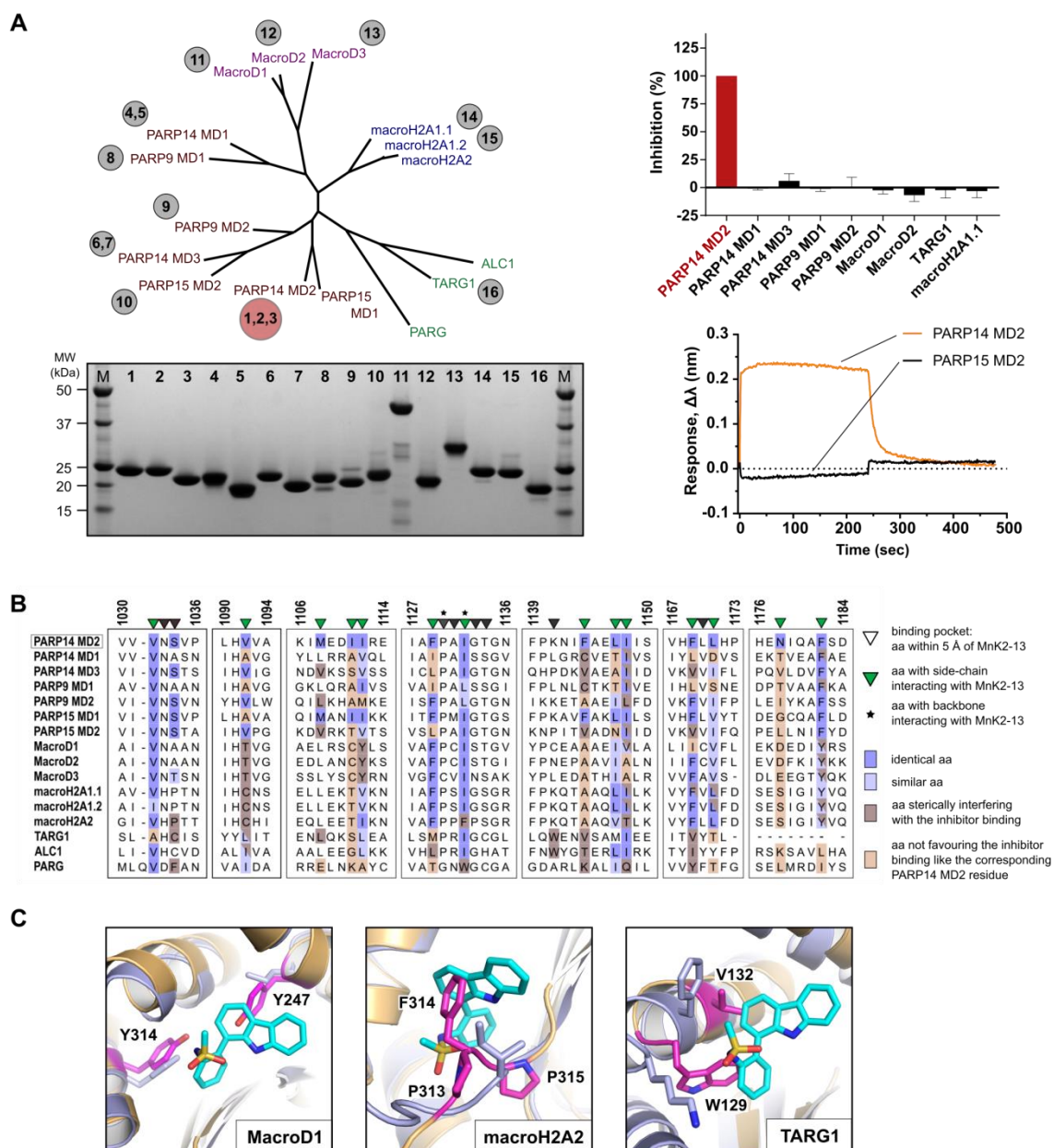


Figure 20. GeA-69 is a selective macrodomain inhibitor. For legend, see bottom of previous page.

Furthermore, GeA-69 was characterised in regard to its biophysical properties. BLI experiments showed that GeA-69 exhibited, despite its enclosed binding position, fast on/fast off kinetics when binding to PARP14 MD2, with a mean dissociation rate constant of $2.4 \times 10^{-2} \text{ s}^{-1}$ resulting in a half-life of binding ($t_{1/2}$, residence time) of 42 seconds [Figure 21A]. Dynamic light scattering (DLS) revealed that GeA-69 is soluble in cell culture medium up to 250 μM which was therefore set as the maximum concentration for following cellular assays [Figure 21B]. Moreover, performance of parallel artificial membrane permeability assays (PAMPA) in collaboration with

Novartis (Basel, Switzerland) showed that GeA-69 was highly cell permeable with a $\log P_e$ of -3.5 over three representative pH values (4.0, 6.8, 8.0) [Figure 21C]. Finally, cytotoxicity assays showed that GeA-69 has only moderate cytotoxicity with an EC_{50} of $\sim 50 \mu M$ on different cell lines tested (HeLa, HEK293 and U-2 OS) [Figure 21D]. This observed inhibitor cytotoxicity at this high concentration was probably not a specific effect of PARP14 MD2 inhibition, as shown by the negative control compound MnK2-68 [Figure S11A]. Thus altogether, GeA-69 demonstrated a biochemical profile to warrant the evaluation of cell-based activities. In an attempt to show GeA-69 target engagement to PARP14 in cells, cellular thermal shift assays (CETSA) were performed using U-2 OS cells which showed good levels of PARP14 expression^{109,166,167}. CETSA is based on the same principle as *in vitro* thermal shift assays, i.e. the ligand-induced thermodynamic stabilisation of the target protein, however, allows monitoring target engagement inside cells without the need to use labelled molecules (tracers) or engineered protein to follow the binding process¹⁶⁸. To perform the assay, intact live U-2 OS cells were treated with $250 \mu M$ GeA-69. After treatment, cells were transiently heated to thermally denature and precipitate the proteins and lysed which was followed by analysis of the soluble protein fraction by Western blotting to generate melting curves of endogenous PARP14 [Figure 21E, S11B]^{169,170}. Repeatedly, no thermal stabilisation of PARP14 by GeA-69 was detected under these conditions. However, this result is perhaps not surprising since PARP14 is a large protein (203 kDa) and stabilisation of only one small domain may have limited effect overall. This is one of the known limitations of the CETSA assay¹⁷⁰.

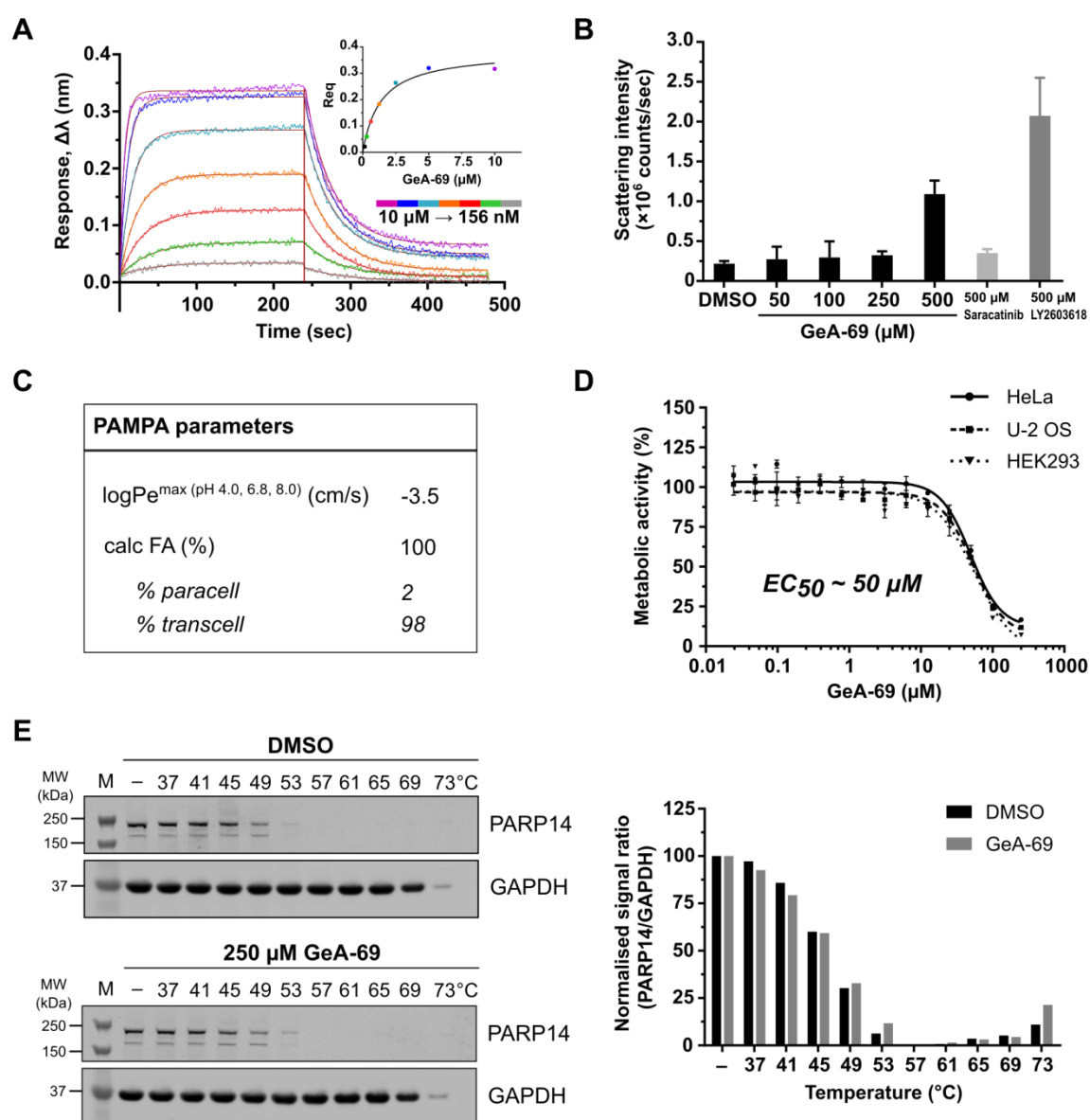


Figure 21. Characterisation of GeA-69 concerning biophysical and cellular properties. (A) Binding kinetics of GeA-69 to PARP14 MD2. The BLI sensorgram shows the kinetics of GeA-69 association and dissociation to PARP14 MD2. The inhibitor concentration range and steady state analysis are shown next to the traces. Mean k_a is $1.35 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ while the mean k_d is $2.4 \times 10^{-2} \text{ s}^{-1}$ resulting in a half-life of binding ($t_{1/2}$, residence time) of 42 seconds. Here, a K_D for GeA-69 of $(1.4 \pm 0.2) \mu\text{M}$ was calculated. (B) GeA-69 is soluble in cell culture medium up to 250 μM as shown by DLS experiments. Saracatinib was used as a compound control with high solubility, LY2603618 as control with low solubility. (C) GeA-69 is highly cell permeable. Cell permeability was assessed by PAMPA (parallel artificial permeability assay) performance. The highest effective permeability ($\log P_e$) of -3.5 is identical over the three representative pH values tested. The fraction of absorbed drug (FA) was calculated to 100 % with 2 % passing paracellular and 98 % transcellular over the membrane. Active transport mechanisms are not considered in this model. (D) GeA-69 shows moderate cytotoxicity. HeLa, U-2 OS and HEK293 cells were incubated for 72 h in the presence of different GeA-69 concentrations (25 nM - 250 μM) and cell viability was measured using the AlamarBlue™ reagent. EC_{50} values were calculated from three independent experiments with $EC_{50_HeLa} = (58 \pm 3.9) \mu\text{M}$, $EC_{50_U-2 OS} = (52 \pm 1.1) \mu\text{M}$, $EC_{50_HEK293} = (54 \pm 2.2) \mu\text{M}$. Results are expressed as mean $\pm$ SEM, $n = 3$. (E) CETSA experiments performed with intact U-2 OS cells treated with 250 μM GeA-69 for 2 h show that GeA-69 is not capable of thermal stabilisation of endogenous PARP14. The result is representative for three independent experiments. **Note:** The PAMPA data was generated by Norbert Damage (Novartis Institutes for BioMedical Research, Basel, Switzerland).

Therefore, laser micro-irradiation assays that were previously established for PARP14 MD2 were performed to assess GeA-69 target engagement in cells. In these assay it was evaluated if GeA-69 could interfere with PARP14 MD2 recruitment to laser-induced DNA damage sites. Indeed, it was observed that pre-treatment of U-2 OS cells with GeA-69 for 2 hours affected PARP14 MD2 re-localisation in this set-up [Figure 22A]. The macrodomain recruitment was significantly decreased in cells pre-treated with 50 μ M GeA-69 and completely prevented in cells exposed to 250 μ M GeA-69. As expected, the negative control MnK2-68 did not interfere with PARP14 MD2 recruitment and further, GeA-69 did not affect the recruitment of an alternative macrodomain-containing protein, YFP-ALC1. Although a high inhibitor concentration had to be applied to completely prevent PARP14 MD2 recruitment to DNA damage sites, these experiments proved that GeA-69 engages PARP14 MD2 in intact cells. It is difficult to rationalise why this high GeA-69 assay concentration is required to completely inhibit PARP14 MD2 recruitment considering the high cell permeability of GeA-69. It may be due to its rather hydrophobic carbazole scaffold which may be able to interact or intercalate with hydrophobic cellular components, significantly reducing the effective concentration on PARP14 MD2. It could furthermore be considered that at the sites of DNA damage, the local concentration of ADPR/ADP-ribosylated proteins is very high against which the inhibitor needs to compete, requiring these high GeA-69 concentrations.

Moreover, it was tested whether GeA-69 affected the recruitment of the triple macrodomain construct PARP14 MD1-3 [Figure 22B]. Consistent with previous results obtained with ADPR-binding-deficient macrodomains [Chapter II.1.2.], GeA-69 was not able to prevent MD1-3 recruitment to the laser-induced DNA damage sites. No difference in MD1-3 recruitment was observed between cells pre-treated with DMSO, the negative control MnK2-68 or GeA-69. Since GeA-69 was shown to engage PARP14 MD2 in cells, it can therefore be excluded that GeA-69 has an effect on MD1 or MD3 since these macrodomains must be responsible for the continued recruitment of MD1-3 in the presence of GeA-69. Consequently, it can also be concluded that GeA-69 is not able to prevent recruitment of full-length PARP14 to its respective target sites. In further agreement with the previous MD1-3 mutagenesis studies was that in presence of 250 μ M

GeA-69 the recruitment of MD1-3 could be significantly impaired by additionally blocking MD3 ADPR recognition by a mutation introduced in its binding site [Figure 22C]. This confirmed again the requirement of at least an additional MD3 inhibitor or a different, orthosteric, inhibitor able to bind both MD2 and MD3 to affect PARP14 recruitment to its ADP-ribosylated target sites and thus to affect PARP14 function.

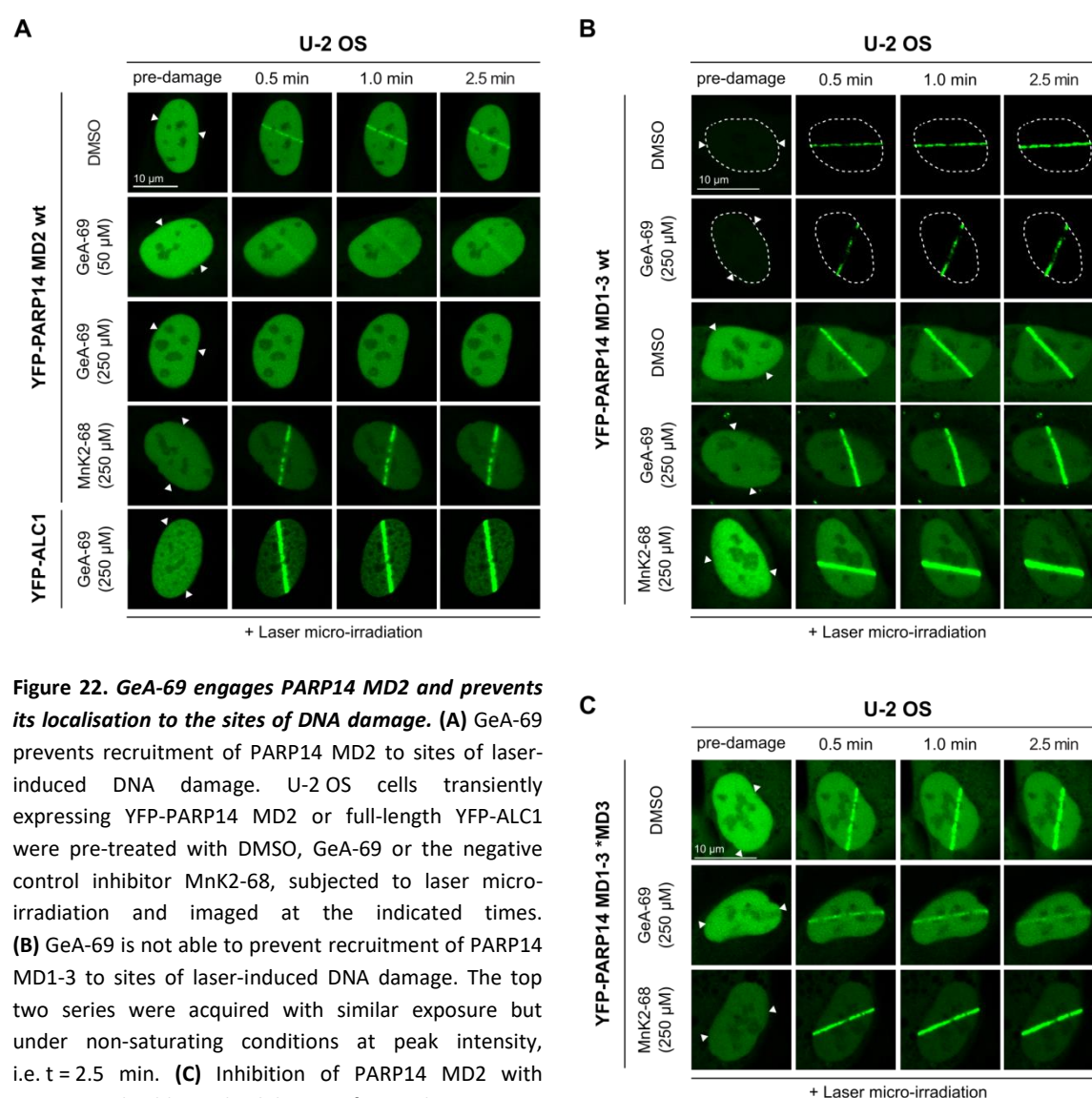


Figure 22. GeA-69 engages PARP14 MD2 and prevents its localisation to the sites of DNA damage. (A) GeA-69 prevents recruitment of PARP14 MD2 to sites of laser-induced DNA damage. U-2 OS cells transiently expressing YFP-PARP14 MD2 or full-length YFP-ALC1 were pre-treated with DMSO, GeA-69 or the negative control inhibitor MnK2-68, subjected to laser micro-irradiation and imaged at the indicated times. **(B)** GeA-69 is not able to prevent recruitment of PARP14 MD1-3 to sites of laser-induced DNA damage. The top two series were acquired with similar exposure but under non-saturating conditions at peak intensity, i.e. $t = 2.5$ min. **(C)** Inhibition of PARP14 MD2 with GeA-69 and additional inhibition of MD3 by mutation significantly reduces PARP14 MD1-3 recruitment to sites of laser-induced DNA damage. **Note:** The experiments were performed by Dr. Ian Gibbs-Seymour (Dr. Ivan Ahel's Laboratory, Sir William Dunn School of Pathology, Oxford).

2.1.2. GeA-69 selectivity over PARP15 MD1

PARP15 MD1 arose from gene duplication of PARP14 MD2⁸⁵ and is therefore the structurally closest macrodomain to PARP14 MD2. It was therefore of high interest to test whether GeA-69 is also selective over the first macrodomain of PARP15. Since no crystal structure for PARP15 MD1 is available, a homology model for PARP15 MD1 was build and used for the structural analysis. Sequence as well as structural comparison of PARP15 MD1 and PARP14 MD2 revealed that the residues around the binding site of GeA-69 are mostly identical in both proteins with exception of Val1092, Leu1171

and Asn1178 which are replaced in PARP15 MD1 by Ala90, Val171 and Gly178, respectively [Figure 23]. It can be assumed that Val171 similarly interacts with MnK2-13/GeA-69 like Leu1171 in PARP14 MD2 and the replacement of Val1092 with Ala90 in PARP15 MD1 reduces the binding affinity but certainly not prevents binding. However, Asn1178 may be important for binding and coordination of the sulfon-/acetamide moiety of MnK2-13/GeA-69 together with a structural water, potentially resulting in GeA-69 selectivity over PARP15 MD1 [Figure 23]. The aim was to obtain biophysical data confirming this hypothesis.

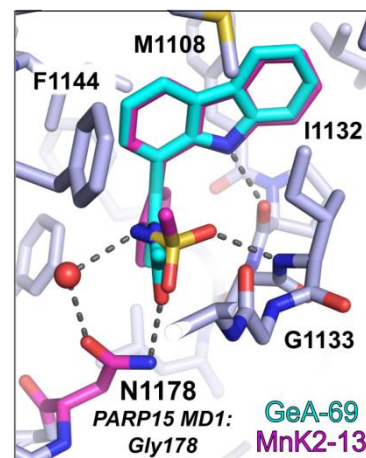


Figure 23. PARP14 MD2 in complex with MnK2-13 (magenta) and GeA-69 (cyan) which was docked into the binding site.

For performing this selectivity study, sufficient quantities of PARP15 MD1 protein had to be prepared in a first step. Since no expression construct was available that yielded soluble PARP15 MD1 protein, fifteen PARP15 MD1 truncations were designed and cloned into a vector for bacterial expression with hexahistidine tag [Figure 24A]. All constructs were transformed into *E. coli* and derivative clones were tested in a small expression scale for their ability to produce soluble PARP15 MD1 protein. PARP15 MD1 expression could be achieved with all constructs, however, none of them yielded a sufficient amount of soluble protein although high salt

concentrations and glycerol for protein stabilisation and prevention of aggregate formation¹⁷¹ were used during the purification [Figure 24B]. PARP15A-c017 with which some soluble protein was obtained was additionally expressed at a larger scale [Figure 24C]. Consistent with the test expression results, only a small amount of soluble PARP15 MD1 protein was isolated together with other impurities after the first Ni-IMAC step. Further purification by gel filtration (GF) only slightly improved the purity of the sample and additional purification steps could not be performed due to the low protein yield. GF fractions containing PARP15 MD1 protein were therefore pooled and it was tested if an AlphaScreen signal with the PARP15 MD1 present in the samples could be generated. In dose-response experiments, no AlphaScreen signal could be obtained even at the highest peptide/protein concentrations, indicating that either PARP15 MD1 was not able to recognise the macrodomain AlphaScreen peptide or more likely that an insufficient amount of correctly folded and thus fully functional PARP15 MD1 was present in the assay [Figure 24C]. Three other PARP15 MD1 constructs were tested similarly in large scale with the same outcome [Figure S12B]. During all protein purifications, it could be observed that PARP15 MD1 was highly expressed, however, was found after cell lysis in the pellet as insoluble fraction and not in the soluble fraction. It was concluded that the protein was probably expressed in insoluble inclusion bodies. Inclusion bodies can occur upon high protein expression levels in *E. coli* and their formation is attributed to a high local concentration of naive polypeptide chains, inefficient folding and the unavailability of chaperones assisting the protein folding process. The resulting partially folded or misfolded protein intermediates then accumulate via hydrophobic contacts to insoluble, mainly non-functional cytoplasmic protein aggregates^{172,173}. Strategies to solubilise proteins produced in *E. coli* include the recovery of these aggregated forms through isolation of the inclusion bodies followed by *in vitro* protein refolding or variation of the expression conditions to decrease the formation of inclusion bodies and consequently improve the yield of soluble protein. *In vitro* protein refolding is however often time consuming, challenging in optimisation and associated with low yields of active protein¹⁷⁴ and very few attempts of PARP15 MD1 expression with modified conditions were not successful [Figure S12C]. Instead, PARP15 MD1 was expressed in an eukaryotic expression host by setting-up a baculovirus

expression system in insect cells which, based on past experience, has a higher success rate than optimisation of the different methods of protein solubilisation. Additionally, PARP15 MD1 was expressed as fusion protein with glutathione-S-transferase (GST) tag to increase its solubility¹⁷⁵. Cloning and expression of respective PARP15 MD1 constructs were performed by the SGC Biotechnology group. From test expressions, two constructs were considered to be suitable for large scale expression of PARP15 MD1 (PARP15A-c024 and PARP15A-c039). Both encode the same PARP15 MD1 region; however, PARP15A-c039 contains an additional biotin tag for a potential application in biophysical assays such as BLI.

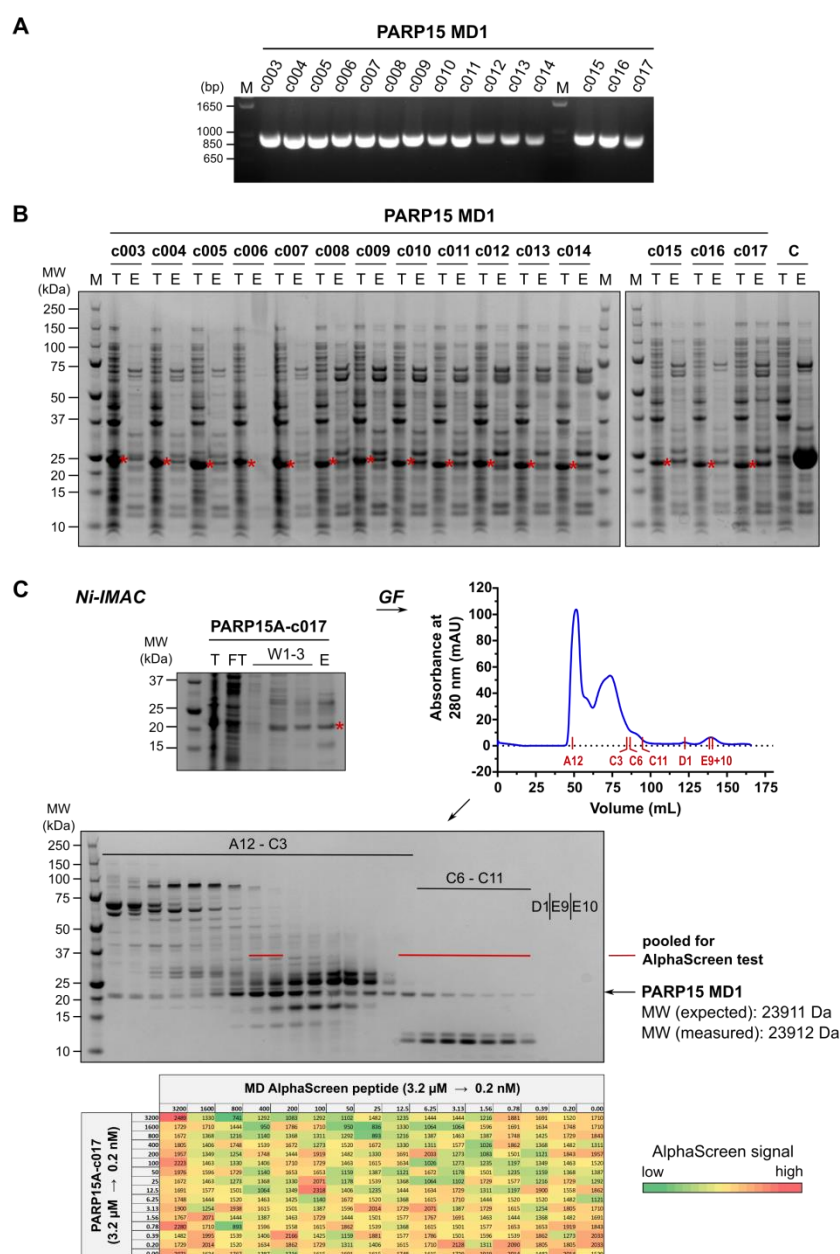


Figure 24. PARP15 MD1 preparation attempts by bacterial expression for selectivity studies. For legend, see top of next page.

Legend to Figure 24: PARP15 MD1 preparation attempts by bacterial expression for selectivity studies.

(A) Fifteen PARP15 MD1 constructs were cloned into pNIC-Bsa4 vectors for bacterial expression. PCR amplified PARP15 MD1 inserts of the corresponding expression constructs from a colony PCR screen are shown. (B) Small scale test expression of fifteen His₆-tagged PARP15 MD1 constructs. Although PARP15 MD1 expression (red star) was obtained from all constructs as seen in the total cell lysate fraction (T), all constructs yielded only in a very small amount of soluble PARP15 MD1 protein concluded from the eluate fraction (E). A PARP14 MD2 construct was used as control (C) for the test expression. (C) PARP15 MD1 large scale expression attempt from construct PARP15A-c017. Similar to the test expression result, only a very small amount of soluble PARP15 MD1 was obtained after cell lysis and Ni-IMAC purification. The eluate fraction was further purified on a S75 16/60 GF column and GF fractions analysed by SDS-PAGE. Fractions marked with a red line were pooled and tested in an AlphaScreen dose-response experiment shown below the SDS-PAGE gel. No AlphaScreen signal could be generated with the protein sample, indicating that the protein was either not fully functional or was unable to bind the macrodomain AlphaScreen peptide or both.

Following expression and cell lysis, PARP15 MD1 was successfully isolated from both expression constructs with good purity by GSH affinity purification [Figure 25A], yet the yield was not high, and indeed was not as high as the previous test expression results. Although similar amounts of protein were obtained from both constructs, only PARP15 MD1 from construct PARP15A-c024 could be further purified by gel filtration [Figure 25B]. The gel filtration profile showed that PARP15 MD1 eluted before the GST-tag with the expected elution volume in a symmetric peak in contrast to PARP15 MD1 from the construct PARP15A-c039 for which at the corresponding elution volume only a shoulder-like profile of low intensity was obtained. Intact mass and tryptic digest analysis by mass spectrometry of the samples confirmed the identity of PARP15 MD1 [Figure S13], the co-expressed GST-tag, which is commonly observed for stable, strongly expressed protein tags, as well as the presence of the heat shock protein Hsp70. Since Hsp70 is known for binding to partially folded or misfolded proteins to prevent aggregation and promote refolding¹⁷⁶, it is likely that not all PARP15 MD1 protein was completely folded. Gel filtration fractions were pooled from the PARP15A-c024 purification to the sample c024-S1 (that very likely contained only partially or misfolded PARP15 MD1 protein due its early and flat elution profile) and sample c024-S2 (that was assumed to contain at least to a certain fraction native PARP15 MD1 protein). The c024-S2 corresponding fractions from PARP15A-c039 were also pooled to the sample c039-S1.

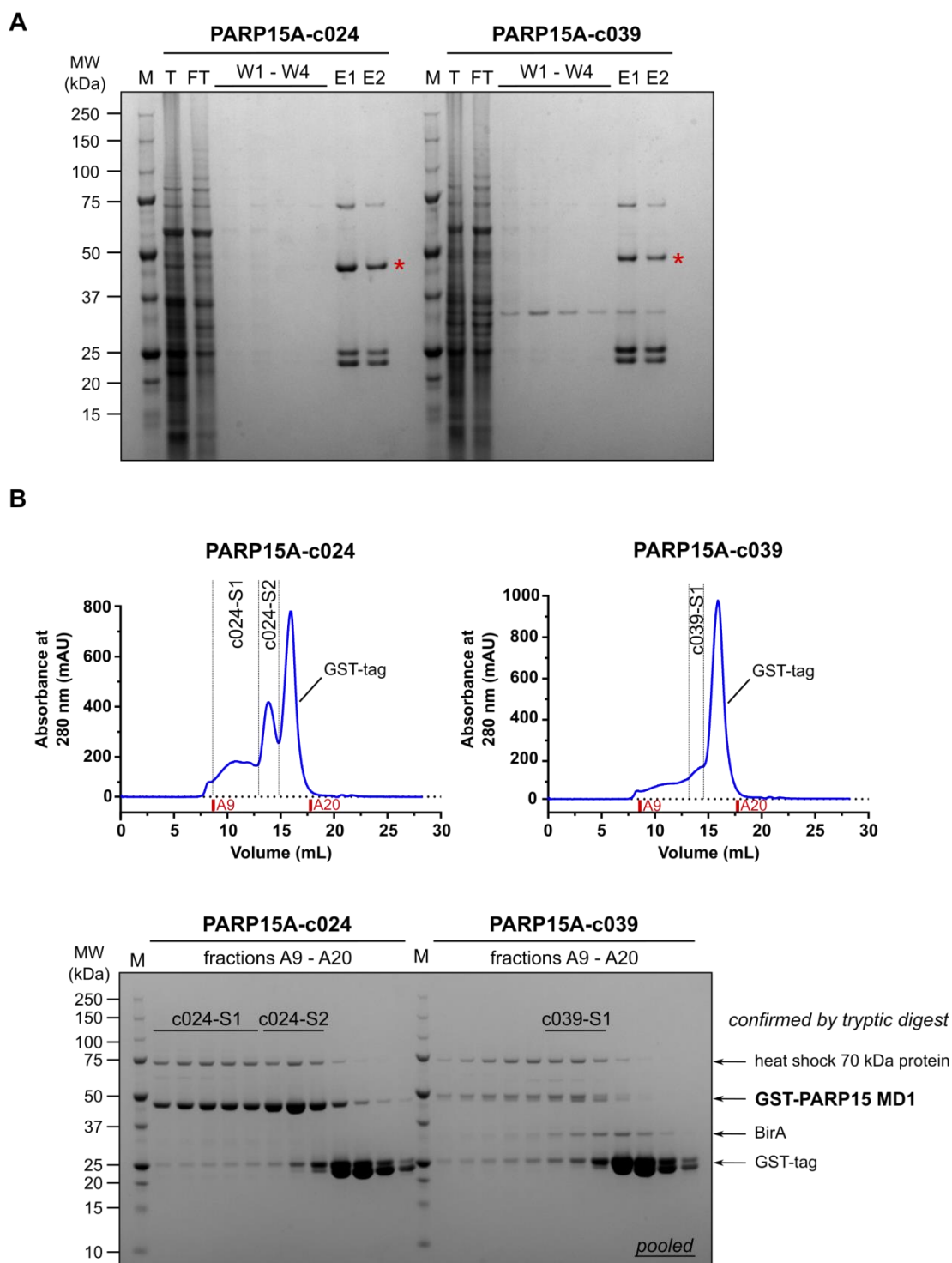


Figure 25. PARP15 MD1 preparation by expression from insect cells for selectivity studies. (A) GST-tagged PARP15 MD1 was expressed in insect cells from the constructs PARP15A-c024 and PARP15A-c039 in large scale. Soluble PARP15 MD1 protein (red star) could be obtained after cell lysis and affinity purification using glutathione sepharose. **(B)** Pooled eluate fractions from each construct (E1 and E2 from the affinity purification step shown in (A)) were further purified by gel filtration using a S75 10/300 GL column. Shown are the obtained gel filtration chromatograms (top) and the corresponding SDS-PAGE analysis of the GF fractions (bottom) for each construct. Fractions were pooled as indicated into the samples c024-S1, c024-S2 and c039-S1 for further characterisation. Proteins bands excised from the gel were analysed by tryptic digest and results are noted next to the gel.

These samples were then further characterised by differential scanning fluorimetry (DSF) in order to assess if the GST-tagged PARP15 MD1 proteins in the corresponding samples were correctly folded. The thermal profile clearly revealed two unfolding states for the sample c024-S2, the first transition being at 37 °C for PARP15 MD1 and the second being at 56 °C for the GST-tag [Figure 26A]. The latter is consistent with the literature¹⁷⁷ and was confirmed by using GST-tagged BRD9 as control. As expected, a two transition profile could not be observed for c024-S1 and c039-S2, confirming that the majority of the PARP15 MD1 was not in its native state. However, the low T_m of 37 °C of PARP15 MD1 in c024-S2 and the additionally higher fluorescence values at the beginning of the profile may nevertheless indicate partial protein unfolding. To obtain further evidence, an HTRF (Homogenous Time-Resolved Fluorescence) assay was set-up, similarly to the macrodomain AlphaScreen assay. Anti-GST-Eu³⁺ cryptate was used as energy donor fluorophore that could bind the GST-tag of PARP15 MD1 while Streptavidin-XL665 was chosen as the acceptor fluorophore binding to the biotin tag of the macrodomain AlphaScreen peptide. Provided PARP15 MD1 could recognise the ADPR imitating part of the peptide, donor and acceptor fluorophore would be brought into close proximity so that, upon excitation of Eu³⁺-cryptate, light emission from XL665 at 665 nm could be measured as HTRF signal¹⁷⁸. A wide range of peptide concentrations were titrated against several PARP15 MD1 protein concentrations. However, no HTRF signal could be generated in any condition with PARP15 MD1 that was significantly higher than that obtained with the buffer control, even at the highest peptide concentration that was compatible with the assay [Figure 26B]. This could either indicate that PARP15 MD1 is not able to bind the macrodomain AlphaScreen peptide with sufficient affinity (e.g. such as PARP15 MD2) or confirms that the protein was partially unfolded and thus not functional. ITC experiments testing for ADPR binding could not be performed since the amount of PARP15 MD1 required was not available. Finally, testing GeA-69 selectivity over PARP15 MD1 (c024-S2) by BLI through immobilisation via its GST-tag could not be realised due to the unavailability of the required anti-GST sensor tips.

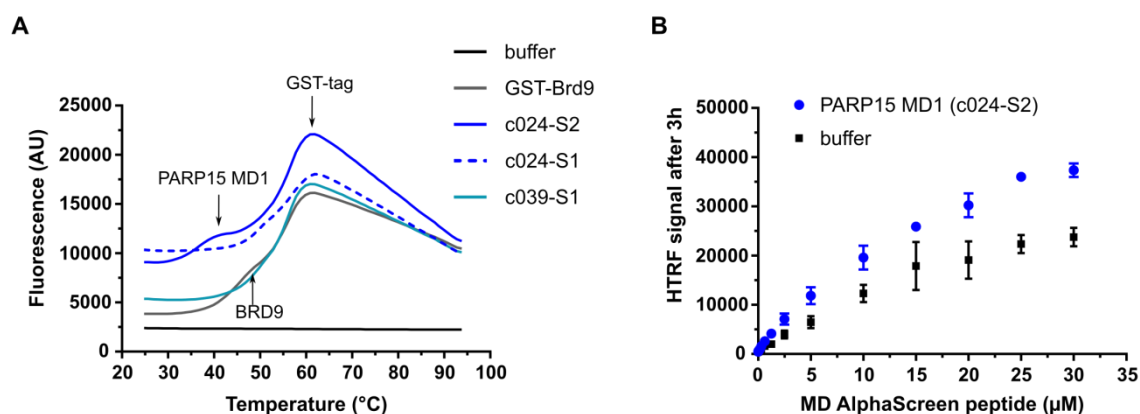


Figure 26. GST-PARP15 MD1 characterisation by DSF and in the HTRF assay. (A) GST-PARP15 MD1 samples (see Figure 25 for purification) were analysed by DSF. GST-BRD9 was used as GST-tagged control protein. For PARP15 MD1 (c024-S2) the two unfolding states could be observed (similar to the GST-BRD9 sample), the first transition being at 37 °C for PARP15 MD1 and the second being at 56 °C for the GST-tag. In contrast, the samples c024-S1 and c039-S1 did not show two transitions indicating that these PARP15 MD1 proteins were not properly folded. The interconnecting line of the data points (mean of $n = 4$) is shown. (B) GST-PARP15 MD1 (c024-S2) could not generate a HTRF signal (fluorescence signal ratio 665 nm/620 nm) using the macrodomain AlphaScreen peptide, indicating that the protein was either only partially properly folded or was unable to bind the macrodomain AlphaScreen peptide or both. Data points are mean $\pm$ SEM, $n = 3$.

In summary, GeA-69 is a new PARP14 MD2 inhibitor that showed good ligand efficiency and is highly selective over other human macrodomains, probably as a consequence of its allosteric binding mode. Its selectivity over PARP15 MD1 could not be clarified due to the difficulty of preparing PARP15 MD1 protein and to confirm its native state with confidence. *In vitro* profiling showed that GeA-69 shows a fast binding kinetic, is soluble to high micromolar concentrations and is highly cell permeable. More importantly, GeA-69 is cell-active and prevents PARP14 MD2 recruitment to laser-induced DNA damage sites. However, the flat aromatic structure of its carbazole which is often found in DNA-intercalating chemotypes¹⁷⁹ is not a very attractive template. Therefore, an extensive SAR study for GeA-69 was performed in collaboration with two chemistry laboratories to improve its scaffold and potency which is shown in the following.

2.1.3. SAR of the GeA-69 lead compound

In an attempt to efficiently establish the SAR study, database searches of in-house and commercially available compounds were performed. However, the hits showed low similarity to GeA-69 and were not considered as good GeA-69 derivatives. Therefore, a medicinal chemistry approach was directly started on GeA-69 to improve the scaffold template and in particular its potency. GeA-69 analogues were designed and synthesised first by Dr. Kerstin Riedel (Professor F. Bracher Laboratory, LMU, Munich) and then by Moses Moustakim (Professor P. Brennan Laboratory, SGC medicinal chemistry group, Oxford). All compounds were characterised for their binding affinity to PARP14 MD2 by determination of their IC_{50} values by AlphaScreen and/or K_D values by BLI and/or ITC. An expression and purification system that was already established in the rotation project prior to this work allowed the preparation of sufficient quantities of PARP14 MD2 with good purity that was required for the studies. PARP14 MD2 proteins were either prepared with hexahistidine tag for use in the macrodomain AlphaScreen assay, with biotin tag for performance of BLI experiments or without any tag for compound characterisation by ITC. A representative example of a PARP14 MD2 purification is shown in Figure 27.

Legend to Figure 27: Representative example of PARP14 MD2 purification for GeA-69 analogue screening and characterisation.

After recombinant overexpression in *E. coli* and cell lysis by high pressure homogenisation, PARP14 MD2 protein was purified by Ni-IMAC, followed by His₆-tag cleavage by TEV protease digestion overnight and gel filtration (GF) using a sepharose based S75 16/60 column (M = marker, T = total - whole cell lysate, FT = flow-through, W1-3 = wash fractions with increasing imidazole concentrations of 20-60 mM, E1-3 = eluate fractions with increasing imidazole concentrations of 250-350 mM). Alternatively, for usage of the protein in the AlphaScreen assay, the protein was directly purified by gel filtration after the Ni-IMAC step without His₆-tag cleavage to obtain purified His₆-tagged PARP14 MD2. GF fractions were analysed for purity by SDS-PAGE. Fractions with high target/background ratio were pooled (*here*: A3 – A10) and passed through a low volume of Ni-sepharose to remove any remaining impurities including the His₆-tagged TEV protease (29 kDa) ('Rebinding'; FT = flow-through fraction, W1-3 = wash fraction with increasing imidazole concentrations of 10-30 mM, E = eluate fraction with 250 mM imidazole). The flow-through fraction was finally concentrated to the desired concentration (typically ~20 mg/mL) and stored at -80°C. The protein identity was confirmed by determining the intact mass by mass spectrometry (expected mass: 21739 Da) at the end of the purification process and/or before the gel filtration purification for additionally confirming completeness of His₆-tag removal.

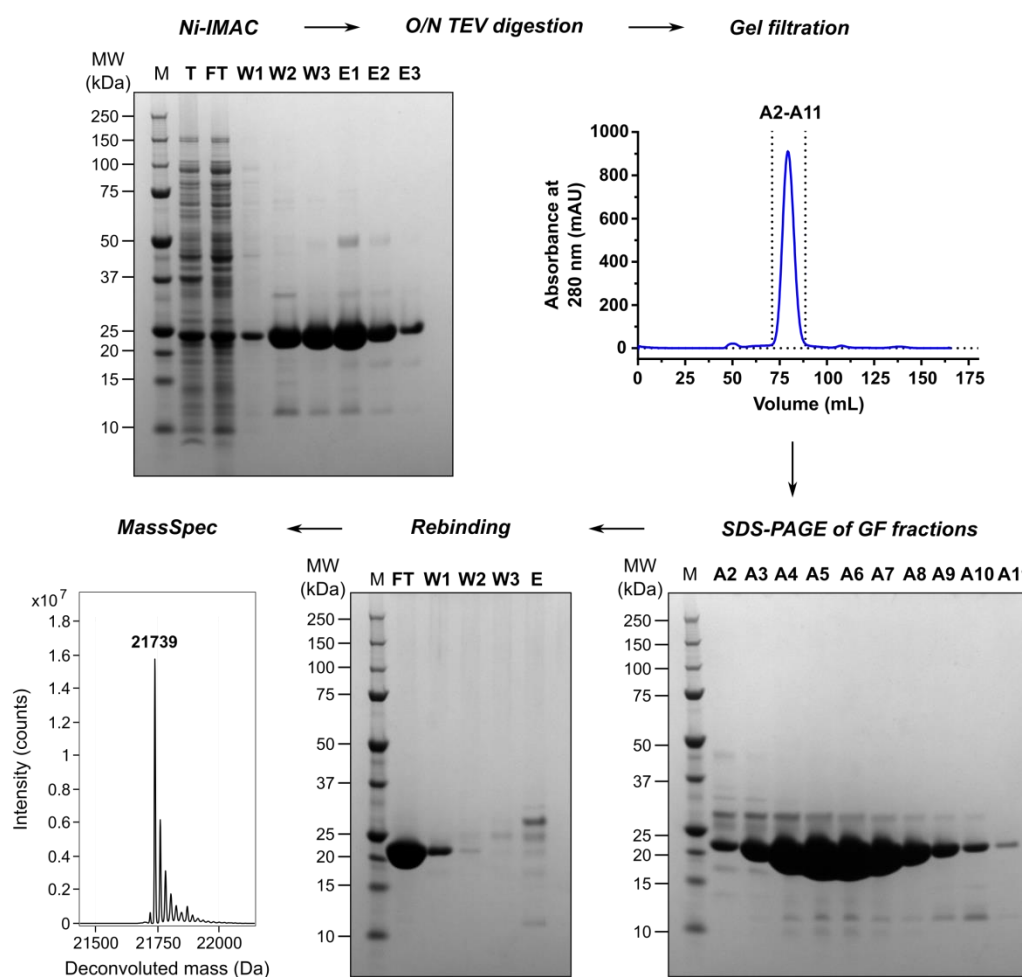


Figure 27. Representative example of PARP14 MD2 purification for GeA-69 analogue screening and characterisation. For legend, see bottom of previous page.

Various scaffold variations including removals, additions and substitutions of functional groups were made on the hit compound GeA-69 that allowed a detailed SAR analysis of binding of GeA-69 and its analogues to PARP14 MD2. The SAR results are schematically presented in Figure 28 while the screening results of the individual compounds are summarised in Table S2. Due to the unavailability of a co-crystal structure of PARP14 MD2 with GeA-69 at the beginning of the SAR studies, a binding mode of GeA-69 in the rather larger ADPR binding pocket of the macrodomain was initially assumed. Early SAR development focused to analyse whether and to what extent the functional groups present in GeA-69 interact with PARP14 MD2 as well as to explore the space around the GeA-69 binding site. For the latter, the carbazole and phenylacetamide moiety were rearranged and the scaffold was extended by variations of the

amide attaching aliphatic linear or cyclic residues or different (hetero-) aromatic cycles. The accessibility and ease of the amide for modification made this a very attractive approach in the early as well as in the later SAR study to improve the potency of GeA-69. Furthermore, systematic substitutions and variations of the functional groups, i.e. the secondary amine in the carbazole and acetamide as well as the carbonyl group of the latter, were made to assess their hydrogen bond donor or acceptor potential interacting with this macrodomain. The amide was also replaced with a urea function as well as a sulfonamide that provided an additional oxygen for interaction and changed the geometry of the amide from trigonal planar to tetrahedral. In this series, the bioisostere of GeA-69 was synthesised, i.e. MnK2-13, which was successfully co-crystallised with PARP14 MD2. The co-crystal structure then allowed retrospective rationalisation: Amide modifications are tolerated, yet with increasing size, the activity of the analogues on the macrodomain is reduced. All of these variations could not provide additional strong hydrophobic, cation- π or hydrogen bond interactions for macrodomain binding that would result in improved potency. Furthermore, all functional groups substantially contribute to the binding of the inhibitor, in particular the hydrogen bond donor function of the carbazole amine as well as the hydrogen bond acceptor function of the carbonyloxygen are essential for binding. The co-crystal structure then allowed a more rational design of GeA-69 derivatives. Since the inhibitor was deeply buried within the macrodomain, modifications were made that were not demanding in space, however should allow additional interactions with the macrodomain. In order to improve the carbazole to a more drug-like scaffold, the ring B of the carbazole was modified, resulting in an acridone or in a seco-derivative of which the latter was thought to enable an improved fitting into the binding site. However, binding of both derivatives could not be measured at a concentration of 50 μ M. Moreover, indole derivatives, thus replacement of ring C with an isopropyl moiety (MnK2-97) or a combined methyl/phenyl substitution (E62738a), resulted in significant decrease of the binding affinity. The characterisation of MnK2-97 is shown in Figure 29. The hydrophobic nature of the carbazole core is also essential for binding, and therefore increasing its polarity by introducing heterocycles, such as α - or β -carboline, or additions of polar substituents around the scaffold was less tolerated and often resulted in inactive

compounds. The phenyl ring was also in focus of the SAR studies. Due to space limitation in the binding site, either substitutions were made, resulting in pyridine analogues, or small groups, i.e. methyl, nitrile or trifluoromethyl groups, were attached in *ortho*, *meta* or *para* position [Figure 28C]. None of the modifications could improve the potency of GeA-69, and more interestingly, all pyridine analogues showed no measurable binding at 50 μM . It was assumed that the binding of the compound having the nitrogen in *ortho* position to the carbazole (MnK2-69) to the macrodomain might be significantly diminished because the pyridine nitrogen could slightly interact by hydrogen bonding with the secondary amine in the carbazole moiety. As a result the compound is not anymore in its active conformation for efficient PARP14 MD2 binding. From this series, MnK2-68 was chosen as negative control for GeA-69 for biophysical or cellular assays.

In particular during the course of the later SAR studies, numerous GeA-69 derivatives with a variety of amide substitutions were prepared. Almost all amide extensions were tolerated but could not improve the potency of GeA-69, showing binding affinities between 1-10 μM . Only E62881a, a sulfonamide analogue extended with a phenyl group having a nitrile in *meta* position, consistently performed better than GeA-69 in the AlphaScreen assay with an IC_{50} of 0.66 μM [Figure 30]. Since E62637a, the E62881a analogue missing the nitrile group, inhibited the macrodomain with only an IC_{50} of 3.6 μM , and furthermore the analogues having the nitrile group in *para* (E62892a) or *ortho* (E62895a) position showed only IC_{50} values of 6.2 μM and 5.4 μM , respectively, the nitrile group was assumed to specifically interact with the macrodomain. Therefore, the E62881a acetamide analogue E62942a and additional phenyl variations were synthesised in an attempt to enhance this interaction or to mimic this interaction with other electron-rich groups. Unfortunately, neither E62942a nor the other derivatives showed similar or improved potency than E62881a and GeA-69, respectively. A drawback of this compound series was the high clogP of the compounds of > 5.7 , resulting in poor solubility in aqueous solution. The IC_{50} values could be determined by AlphaScreen since the AlphaScreen assay buffer contained detergent (0.05 % CHAPS) which disrupts aggregates and thereby helps to solubilise the compounds¹⁸⁰. BLI experiments could also be performed in presence of 0.01 % Tween-20

enabling characterisation of the binding kinetics of E62881a and estimation of the K_D value of E62881a to (550 ± 220) nM which is in agreement with the AlphaScreen IC_{50} value of 660 nM. However, characterisation of E62881a and E62942a by ITC was therefore not possible since the compound solubility in the ITC assay buffer was too poor to detect binding [Figure 30D]. In an attempt to improve the water solubility of E62881a, this compound was prepared as salt protonating its carbazole nitrogen and using chloride as counterion. The solubility seemed to be improved, yet E62881a binding to PARP14 MD2 could again not be detected and analysed by ITC. Although these calorimetric measurements can be carried out in presence of detergent, the addition of detergent to the assay buffer was omitted since the determination of the K_D value by ITC was seen as a final characterisation step of the compounds that considered both, the binding affinity as well as biophysical properties such as compound solubility. Those compounds such as E62881a that could not be characterised by ITC using standard assay conditions shall maybe not be considered as lead compounds for further inhibitor development; they can only guide the compound optimisation process. Thus, despite all of these efforts providing a comprehensive SAR analysis of GeA-69 for its target PARP14 MD2, GeA-69 potency could not be improved. Its optimisation might be challenging since being buried inside the hydrophobic environment of the protein, the hydrophobic interactions need to be maintained, resulting in rather poor biophysical properties of the compounds, and thus a strategy maybe worth considering would be to grow the compound to the protein surface to enable more polar interactions.

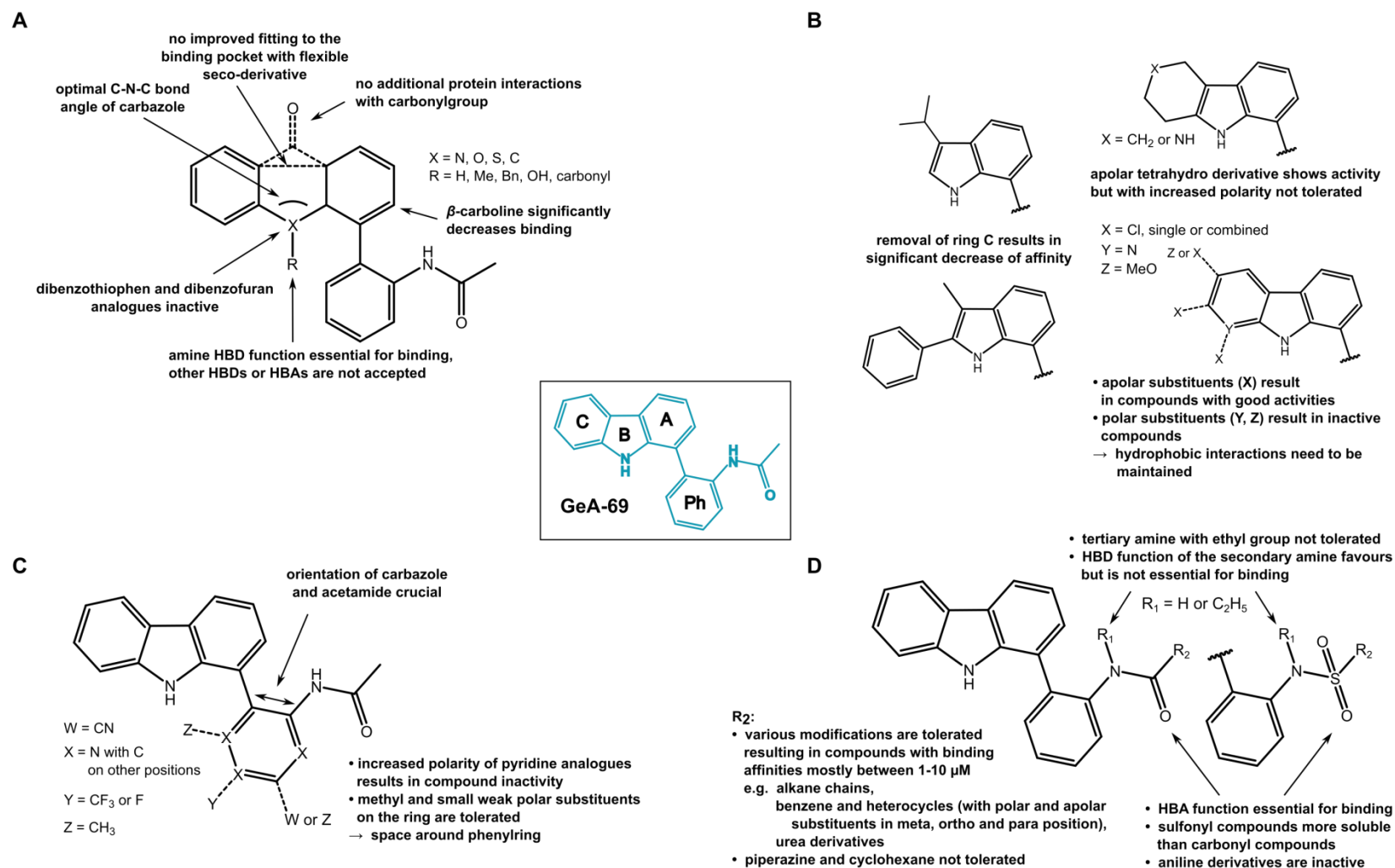


Figure 28. Structure-activity-relationship (SAR) study for GeA-69 and analogues for PARP14 MD2 binding developed in collaboration with Dr. Kerstin Riedel (Professor F. Bracher Laboratory, LMU, Munich) and Moses Moustakim (Professor P. Brennan, SGC, Oxford). (A) Summary of modifications of the carbazole rings A and B. (B) Summary of modifications of the carbazole ring C. (C) Summary of phenylring variations. (D) Overview of acetyl/sulfonamide variations. HBA: hydrogen bond acceptor, HBD: hydrogen bond donor.

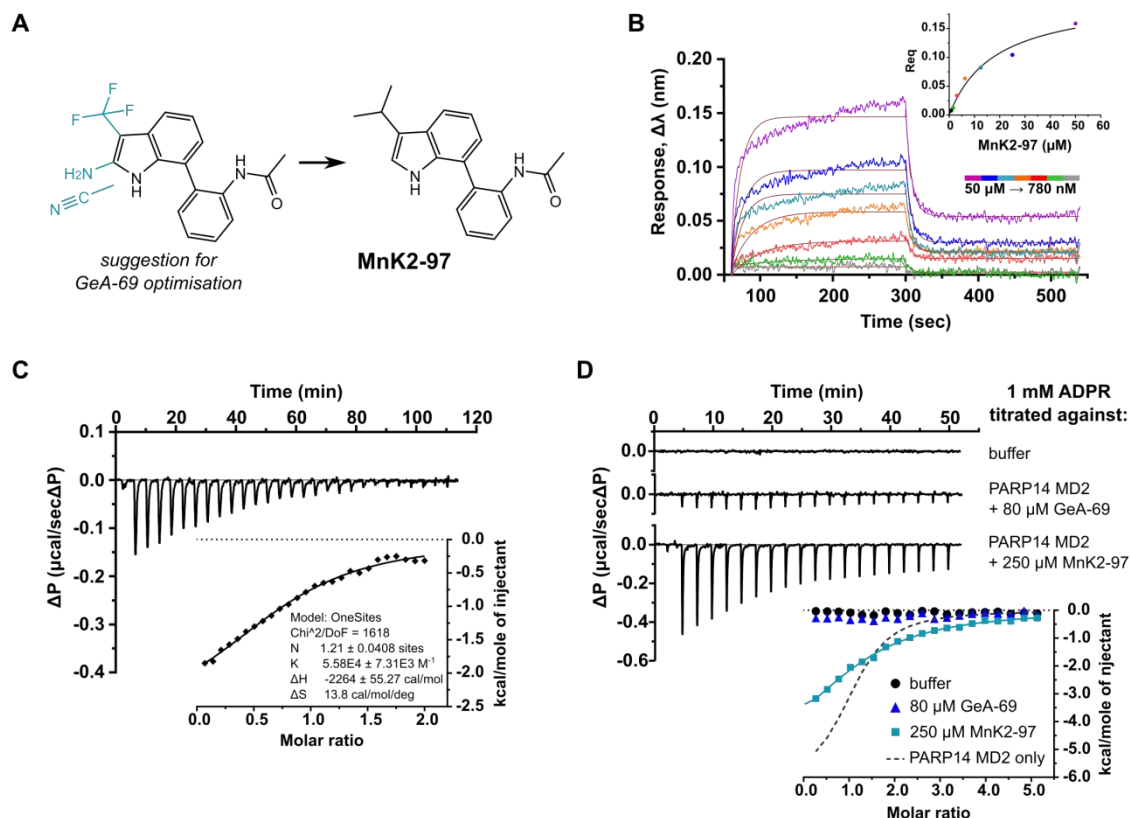


Figure 29. Example of GeA-69 analogue characterisation (I). (A) Molecular structure of MnK2-97 as the first GeA-69 analogue without carbazole scaffold. (B) MnK2-97 binds PARP14 MD2 with a K_D of $(18.0 \pm 3.9) \mu\text{M}$ with similar kinetics to GeA-69 as shown by BLI measurements. Mean k_a is $2.4 \times 10^4 \text{ M}^{-1}\text{s}^{-1}$ while mean k_d is $1.0 \times 10^{-1} \text{ s}^{-1}$ resulting in a half-life of binding ($t_{1/2}$, residence time) of 10 seconds (GeA-69: $t_{1/2} = 42 \text{ sec}$). (C) Binding of MnK2-97 to PARP14 MD2 was confirmed by ITC (also with a K_D of $(18 \pm 2.4) \mu\text{M}$) by titrating PARP14 MD2 into a $40 \mu\text{M}$ solution of MnK2-97. Shown are the raw injection heats of the titration and the normalised binding enthalpies as a function of binding site saturation with the calculated binding parameters employing a single-site binding model. (D) ITC competition experiments showed that MnK2-97 is able to compete but not to completely prevent ADPR binding to PARP14 MD2 at a concentration of $250 \mu\text{M}$. Raw injection heats are shown resulting from titration of a 1 mM ADPR solution into the cell containing either buffer (top) or $40 \mu\text{M}$ PARP14 MD2 in presence of $80 \mu\text{M}$ GeA-69 (middle) or $250 \mu\text{M}$ MnK2-97 (bottom). The normalised binding enthalpies are shown in the inset including the measurement of PARP14 MD2 only as a reference (K_D (ADPR) = $(7.8 \pm 0.5) \mu\text{M}$). The calculated K_D for ADPR binding of PARP14 MD2 in presence of $250 \mu\text{M}$ MnK2-97 was calculated to $(40 \pm 3.8) \mu\text{M}$. Additional binding and fitting parameters are provided in Figure S17.

Legend to Figure 30: Example of GeA-69 analogue characterisation (II).

(A) Molecular structure of E62881a and E62942a, with E62881a being the first GeA-69 analogue with extended scaffold and same binding affinity as GeA-69 towards PARP14 MD2. (B) Representative IC_{50} curves of GeA-69, E62881a and E62942a obtained from AlphaScreen assays performed with PARP14 MD2. (C) E62881a shows a slower on/off kinetic than GeA-69 to PARP14 MD2. Mean k_a is $6.6 \times 10^3 \text{ M}^{-1}\text{s}^{-1}$ while mean k_d is $5.4 \times 10^{-3} \text{ s}^{-1}$ resulting in a half-life of binding ($t_{1/2}$, residence time) of 185 seconds (GeA-69: $t_{1/2} = 42 \text{ sec}$). The calculated K_D (E62881a) is $(550 \pm 220) \text{ nM}$. (D) Binding of E62881a and E62942a cannot be confirmed by ITC measurements, probably due to the poor compound solubility in buffer without detergent. Raw injection heats are shown from titrations of $200 \mu\text{M}$ PARP14 MD2 into a solution containing $10 \mu\text{M}$ E62881a or E62942a. (Same results were obtained with higher protein and compound concentrations.)

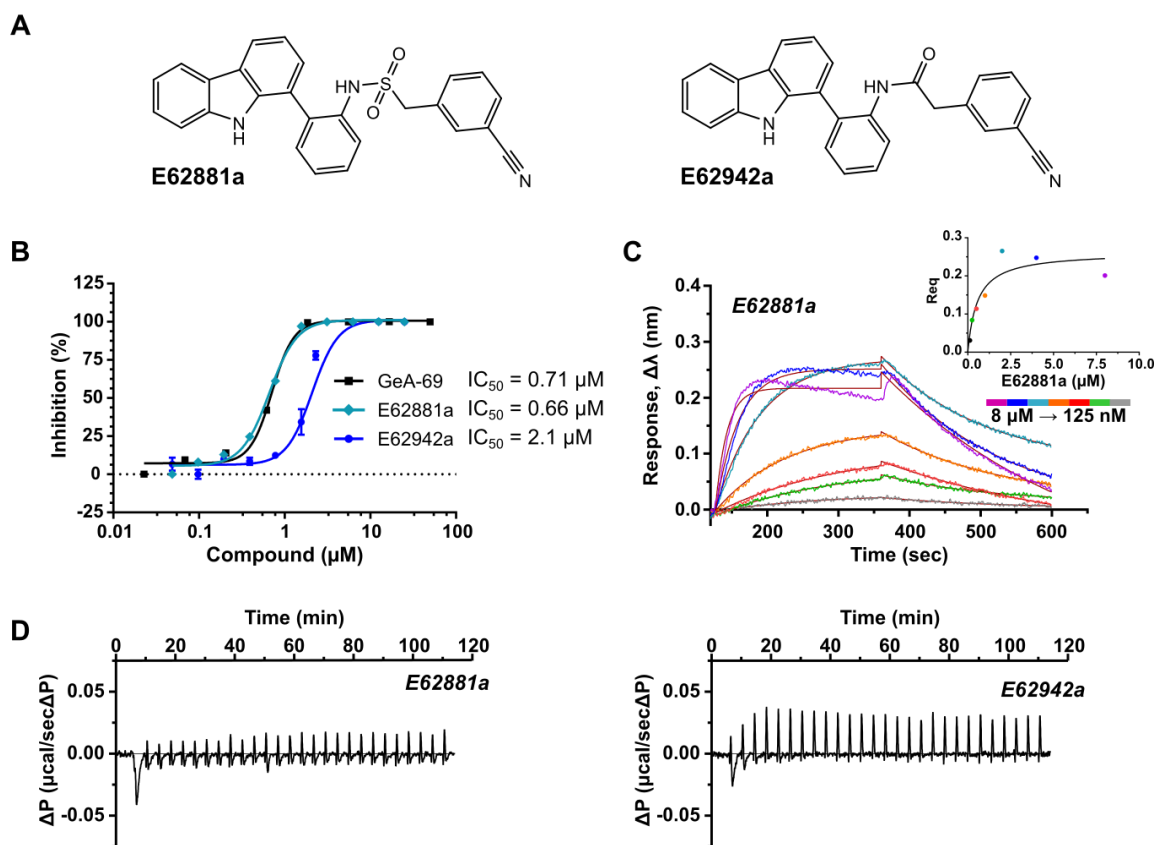


Figure 30. Example of GeA-69 analogue characterisation (II). For legend, see bottom of previous page.

To support and guide the SAR studies, it was attempted to co-crystallise GeA-69 derivatives with PARP14 MD2 to obtain structural information about the binding mode of the compounds. This present work begun with characterisation of compounds synthesised after the compound series of acetamide/sulfonamide variations including MnK2-13 which was successfully co-crystallised. These new compound series aimed to study modifications on the carbazole scaffold and phenyl ring of GeA-69. Almost all GeA-69 variations in these series showed no or very weak inhibitory activity towards PARP14 MD2 and were therefore not taken into co-crystallisation trials. Only MnK2-97 was a candidate considered for co-crystallisation. Although showing significantly lower potency than GeA-69 with a K_D of 18 μM , MnK2-97 was the first compound that had the carbazole replaced with an indole scaffold and was therefore far more drug-like. Information about the orientation of the indole and its attached isopropyl moiety in the protein could have guided the design of the indole based compound series. Therefore, the PARP14 MD2 SER mutants including the mutant MD2_{K1048S}, K1154S, K1158S, K1162S (PARP14A-c013) which gave the

crystal structure in complex with MnK2-13 were purified and set-up in crystallisation trials [Figure 31A]. Only the SER mutant resulting from PARP14A-c012 was excluded from crystallisation trials since SDS-PAGE and mass spectrometry analysis revealed two truncations of unexpected size of the macrodomain which were assigned after Sanger sequencing to a mutation of the stop codon in the corresponding SER mutant encoding plasmid. The crystallisation attempts can be summarised as follows: Crystalline structures were obtained for most of the constructs apart from PARP14A-c014. Their morphology was poor; either crystal clusters formed with flower-like structures or very flat, un-mountable crystals adhered to the surface of the drop [Figure 31A (a-f)]. Attempts to improve the morphology of the crystalline structures were not successful and included variation of the protein concentration, setting-up of follow-up screens of selected crystallisation conditions, seeding of the protein clusters into new screens or testing of additive screens. None of the very few mounted crystals diffracted X-rays. Since in particular disordered crystal clusters were obtained, it was considered if this is caused by the flexible N-terminus whose first nine amino acids were too disordered in the electron density map of the PARP14 MD2-MnK2-13 co-crystal structure to be resolved. Therefore, three N-terminal truncations of each SER mutant were prepared and set-up for crystallisation trials with MnK2-97 [Figure 31B, C]. Hits were mainly obtained with truncation 2 of the construct PARP14A-c011 but as before the crystals showed no diffraction of X-rays.

It seemed that PARP14 MD2 did not readily form well diffracting crystals as seen before and also observed by others (Prof. Herwig Schüler (Karolinska Institutet, Stockholm) personal communication). Since no crystal structure of PARP14 MD2 is available in its *apo* form, it can be assumed that the use of ligands that stabilise the macrodomain such as ADPR may be required to promote its crystallisation. Good ligand solubility may thereby also be beneficial. However, MnK2-97 shows neither high solubility in the protein buffer nor strong binding to PARP14 MD2 like GeA-69 or MnK2-13. At the compound concentrations (1-2 mM) which were necessary to ensure saturation of the macrodomain, i.e. at least $2 \times$ molar concentration of the protein with $> 10 \times K_D$, MnK2-97 showed precipitation that was also often associated with protein precipitation. Apart from MnK2-97, a co-crystal structure with the acetamide/sulfonamide GeA-69 analogues

E62942a and E62881a would have also been of interest to support the design of analogues with structural information. A few crystallisation trials were not successful either, very likely due to the very poor compound solubility that made even the characterisation of the compounds by ITC impossible. A possible future strategy would be to attempt the co-crystallisation of close analogues of E62881a with good solubility in combination with a MD1-MD2 double construct. This construct could be successfully crystallised in its *apo* form as well as in complex with ADPR⁶⁶, indicating that MD2 in conjugation with MD1 might be more crystallisable and does not require macrodomain stabilisation by a ligand to form crystals.

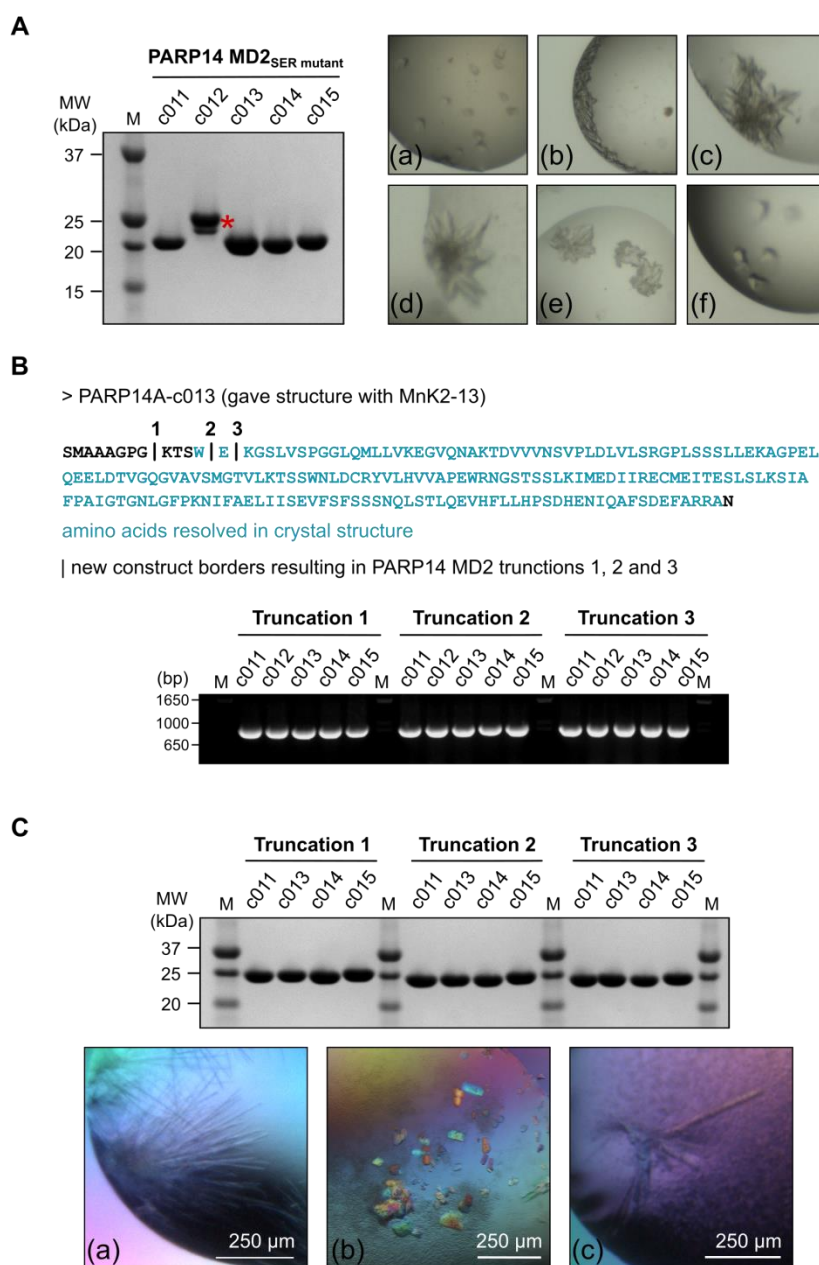


Figure 31. PARP14 MD2 co-crystallisation attempts with MnK2-97. For legend, see top of next page.

Legend to Figure 31: PARP14 MD2 co-crystallisation attempts with MnK2-97.

(A) Five SER mutants of PARP14 MD2 were purified (PARP14A-c011 to c015) (shown by SDS-PAGE on the right) for crystallisation trials with MnK2-97. Note that the double bands of larger size compared to the other constructs obtained with PARP14A-c012 (marked with red star) results from a mutation in the stop codon of the construct. Representative examples of crystallisation results with the other MD2 SER mutants are shown on the left and were obtained under the following conditions: **(a)** 0.3 M magnesium formate dehydrate, 0.1 M Bis-Tris (pH 5.5) with c011; **(b)** 20 % PEG6000, 10 % ethylene glycol, 0.1 M MES (pH 6.0), 0.2 M ammonium chloride with c011; **(c)** 1.6 M sodium citrate tribasic dihydrate (pH 6.5) with c013; **(d)** 3 M sodium chloride, 0.1 M Bis-Tris (pH 5.5) with c015; **(e)** 0.2 M sodium chloride, 25 % PEG3350, 0.1 M HEPES (pH 7.5) with c015; **(f)** 2.4 M sodium malonate with c015. **(B)** Design of new truncation constructs based on the PARP14 MD2 SER construct PARP14A-c013 which could be successfully crystallised with MnK2-13. Truncations were made on the N-terminus to obtain three constructs containing mainly the amino acids resolved in the crystal structure (shown in blue). PCR amplified inserts of the corresponding PARP14 MD2 constructs from a colony PCR screen are shown below the construct design. **(C)** The indicated PARP14 MD2 SER truncation constructs were purified (shown by SDS-PAGE) and set-up in crystallisation trials. Representative hits obtained are shown below. **(a)** PARP14A-c013 truncation 2 with 60 % v/v Tacsimate™ (pH 7.0); **(b)** PARP14A-c011 truncation 2 with 0.2 M lithium sulfate, 0.1 M sodium acetate (pH 4.5), 50 % w/v PEG400; **(c)** PARP14A-c011 truncation 2 with 20 % PEG3350, 10 % ethylene glycol, 0.1 M Bis-Tris-propane (pH 8.5), 0.2 M sodium sulfate.

2.1.4. Finding new PARP14 MD2 inhibitors – screening experience with Novartis using a public 45,000 compound library

It was also desired to obtain an orthosteric PARP14 MD2 inhibitor which might perhaps be easier to develop than GeA-69. Therefore, a medium-throughput screen was performed against PARP14 MD2 to identify new chemical starting points that could provide the basis for new macrodomain inhibitor development. The compound screen itself was carried out on the Novartis Institute for Biomedical Research (NIBR) site in Basel (Switzerland) in two rotation projects each lasting one month where all resources including compound library, materials and equipment were provided. The workflow of the screen is outlined in Figure 32A.

Prior to the screen, PARP14 MD2 was prepared in sufficient quantities of high purity with hexahistidine tag for its use in the macrodomain AlphaScreen assay as well as with biotin tag and without tag for the following hit validation by BLI and ITC, respectively [Figure 32B]. Next, experiments were performed to ensure that the macrodomain AlphaScreen method could be

directly translated to the Novartis screening systems. This included a quality control of the AlphaScreen components, the new PARP14 MD2 batch and the macrodomain AlphaScreen peptide. The peptide quality and stability after several harsh freeze-thaw cycles was verified by LC-MS and the new PARP14 MD2 batch was tested in AlphaScreen dose-response experiments together with the peptide confirming its ability to produce a good and stable AlphaScreen signal. Furthermore, the DMSO sensitivity of the AlphaScreen assay was assessed. The stock concentration of the Novartis screening compounds of 2 mM resulted in a final DMSO assay concentration of 1.25 %. The test confirmed that DMSO concentrations of up to 2.0 % were tolerated and did not influence the AlphaScreen signal by for instance interfering with the beads or the functionality of the protein. Furthermore, the composition of the AlphaScreen buffer was re-considered concerning its additives. Apart from the addition of the detergent CHAPS (0.05 %) to avoid compound aggregation that is often associated with non-specific protein inhibition and BSA as carrier protein to decrease protein loss on surfaces and to stabilise the macrodomain in the assay, the screening assay buffer conditions were optimised by including a reducing agent^{180–182}. Compounds can also act directly as an oxidising agent and the addition of reducing agents protects the target protein from either oxidation or compound adduct formation¹⁸³. As it has also been shown previously in the group¹⁶³ that the presence of a reducing agent in the assay buffer reduces the hit rate of false positive compounds in the macrodomain AlphaScreen assay, it was tested independently from this study whether the reducing agents TCEP and DTT are compatible with the macrodomain screening method. Titration of the reducing agents combined with IC₅₀ experiments for the macrodomain inhibitor GeA-69 confirmed that both reducing agents were well tolerated by the assay and therefore 0.5 mM TCEP was added to the screening assay buffer. Following these optimisation and quality controls, the NDDI (NIBR Drug Discovery Incubator) compound library comprising ~45,000 compounds was screened against PARP14 MD2 at a compound concentration of 25 µM. The NDDI library is a subset of the compound libraries offered by commercial vendors containing structurally diverse compounds with drug-like properties (based on Lipinski's 'rule-of-5') and without known unacceptable structural features including reactive groups and known assay interference motives. All compounds were screened in

a single-shot mode using GeA-69 as quality control on each assay plate. Although Z' values and the IC₅₀ obtained for GeA-69 confirmed the quality of each plate, some compounds were re-screened in a single-shot mode if screening (signal) patterns could be detected by manual inspection of the readout of each plate.

In the first 30,000 compound screen 664 compounds showed more than 50 % PARP14 MD2 inhibition, and 352 compounds were identified with this activity in the 15,000 compound screen, resulting in a consistent hit rate of 2.2 %. Due to the susceptibility of the AlphaScreen for assay interfering compounds¹⁸⁴, these initial hit compounds were then further validated by counterscreening (as far as the compound stock was available) using a biotinylated-His₆ peptide that enables to detect false positive hits, i.e. compounds that interfere with the assay itself by absorbing the excitation or emission light, quenching the singlet oxygen, mimicking biotin or being Nickel chelating competitors. Furthermore, dose-response assays were performed to obtain IC₅₀ values for these compounds. In summary, 0.08 %, i.e. 38 compounds, of the 45,000 compounds showed more than 80 % inhibition of the macrodomain and less than 25 % activity in the counterscreen and were considered as good hits [Figure 32C]. From each primary screen, 21 and 20 compounds, respectively, of different chemotypes were selected for further hit characterisation based on their PARP14 MD2 inhibition and counterscreen activity [Figure 33A, 35A]. As strategy to identify non-selective binders, more than one hit of a certain chemotype was selected in some cases since it was assumed that real binders would result in a certain SAR in contrast to compounds non-specifically interacting with the macrodomain. An example is the NVP-BNV410-NX-1 to NVP-BNV442-NX-1 series from the 15,000 compound screen. Based on throughput and sample requirements, BLI was considered as a suitable secondary screening platform of the compounds while ITC was intended only to be used as tertiary assay for final hit compound validation.

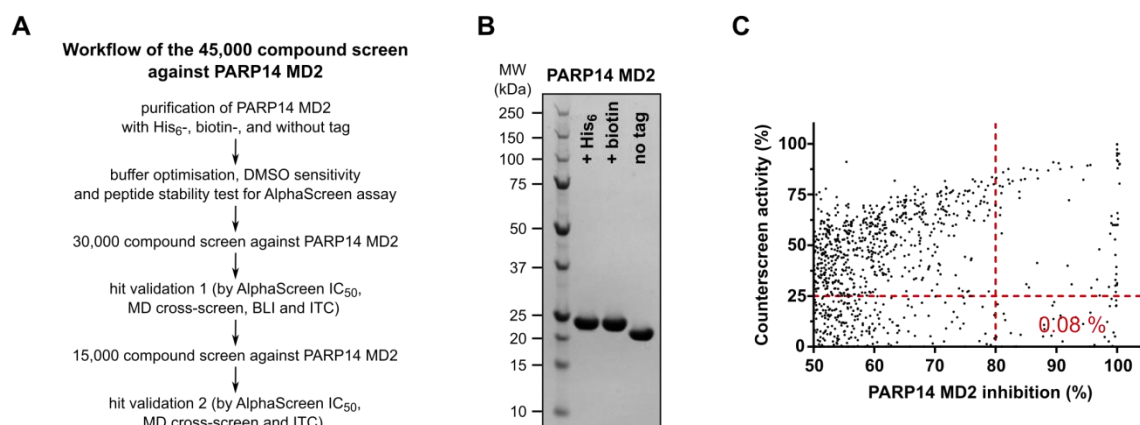


Figure 32. Performance of a 45,000 compound screen against PARP14 MD2 by AlphaScreen. (A) Workflow of the performance and the hit validation process of the 45,000 compound screen against PARP14 MD2. (B) SDS-PAGE of purified PARP14 MD2 proteins used for the primary screen by AlphaScreen (with His₆-tag) and for the secondary screens by BLI (with biotin tag) and ITC (no tag). (C) The graph shows the result of the primary screen of the NDDI library provided by Novartis (Basel, Switzerland) and displays the AlphaScreen activity of 953 compounds that showed more than 50 % macrodomain inhibition. 0.08 % of the 45,000 compounds showed more than 80 % inhibition of the macrodomain and less than 25 % activity in the counterscreen (red lines).

All compounds selected for the final hit validation were first screened in the macrodomain AlphaScreen assay at the SGC using the same PARP14 MD2 batch that was used for the primary screen at Novartis. The reproducibility of the screening results was poor, in particular for the compounds selected from the first 30,000 compound screen [Figure 33A]. The PARP14 MD2 activity could be reproduced for only four compounds of which two compounds could be identified as assay interfering compounds by a macrodomain cross-screen [Figure 33B]. Furthermore, no convincing BLI sensorgrams or ITC titration profiles were obtained, indicating specific binding of any hit to PARP14 MD2 [Figure S15]. The compound concentrations in the respective assays were based on the IC₅₀ values prior determined by the AlphaScreen assay and the compound solubility in the assay buffer. Therefore, for the hits selected from the 30,000 compound screen, new DMSO solutions from the powder stocks were prepared by Novartis which additionally underwent a quality control by LC-MS. For most of the compounds the identity and nominal concentration of 10 mM was confirmed [Figure S14]. Using the new stock solutions, the initial screening results could then be reproduced with broadly consistent PARP14 inhibition at 25 μ M and IC₅₀ values for 13 compounds [Figure 33A]. These compounds were then again validated together with the hits selected from the 15,000 compound screen by ITC.

A

	Compound ID	Novartis			SGC	New stocks	
		PARP14 MD2 inhibition (%)	Counterscreen activity (%)	IC ₅₀ (μM)	PARP14 MD2 inhibition (%)	PARP14 MD2 inhibition (%)	IC ₅₀ (μM)
1	NVP-BMM025-AA-1	100.0	---	7.1	89.4	99.7	6.3
2	NVP-BNP116-NX-1	99.7	---	15	8.6	99.3	12
3	NVP-BMF069-NX-1	99.3	---	6.2	97.0	99.9	4.1
4	NVP-BMW608-NX-1	99.2	0.3	>25	-2.1	-6.0	>100
5	NVP-BMD153-NX-1	98.4	12.3	5.5	1.2	6.48	>100
6	NVP-BMC079-NX-1	98.0	4.8	>25	-4.3	11.83	>100
7	NVP-BNP071-NX-1	96.2	-1.0	11	0.8	93.2	37
8	NVP-BNC609-AA-1	95.7	29.0	10	6.2	98.0	12
9	NVP-BMC081-NX-1	95.5	-1.9	>25	-3.8	-6.0	>100
10	NVP-BMB730-NX-1	91.0	2.7	>25	4.3	17.9	>100
11	NVP-BME758-NX-1	89.2	5.5	>25	-0.9	19.3	>100
12	NVP-BMW822-NX-1	88.0	10.7	22	4.7	92.7	19
13	NVP-BNH686-NX-1	87.5	13.8	14	7.6	-6.6	>100
14	NVP-BMJ791-NX-1	87.0	---	---	63.6	93.0	7.0
15	NVP-BNC605-AA-1	81.3	16.3	23	6.9	85.6	29
16	NVP-BMP569-NX-1	81.0	-0.7	0.6	48.2	91.1	9.3
17	NVP-BNC594-NX-1	80.3	20.9	14	9.5	91.2	17
18	NVP-BME919-NX-1	65.5	20.3	12	-4.3	67.6	11
19	NVP-BML846-NX-1	65.6	7.6	---	0.1	47.1	76
20	NVP-BMU249-NX-1	69.6	1.6	---	-1.9	2.9	>100
21	NVP-BNC592-NX-1	73.7	14.6	---	8.4	88.1	39

B

	Compound ID	(%) Macrodomain inhibition at 25 μM compound concentration						
		PARP14 MD2	PARP14 MD3	PARP9 MD2	MacroD1	MacroD2	TARG1	macroH2A1.1
1	NVP-BMM025-AA-1	86.6	33.2	3.7	38.3	-9.5	1.9	54.6
2	NVP-BNP116-NX-1	9.0	-9.0	-8.0	3.2	-8.0	-14.6	-4.2
3	NVP-BMF069-NX-1	94.8	58.6	16.2	75.1	-4.6	30.7	73.7
4	NVP-BMW608-NX-1	-5.1	-9.6	-6.5	-3.5	-2.4	-20.1	-11.2
5	NVP-BMD153-NX-1	-2.3	-10.5	-9.8	1.3	-4.7	-15.4	-15.8
6	NVP-BMC079-NX-1	-5.2	-12.6	-7.9	-2.8	-0.9	-10.7	-14.9
7	NVP-BNP071-NX-1	4.8	-7.8	-4.6	2.0	-3.3	-5.6	-6.4
8	NVP-BNC609-AA-1	6.5	0.2	4.6	11.8	11.3	-4.6	4.2
9	NVP-BMC081-NX-1	-3.1	-9.6	-8.2	-0.2	-0.7	-11.0	-10.6
10	NVP-BMB730-NX-1	-0.3	-3.7	5.5	-0.1	-1.6	-20.5	-12.3
11	NVP-BME758-NX-1	-0.3	-3.2	-2.3	-1.1	-0.3	-5.3	-5.7
12	NVP-BMW822-NX-1	3.3	-4.9	6.8	12.2	0.7	0.2	-5.0
13	NVP-BNH686-NX-1	5.0	-6.9	-0.1	9.5	0.9	-8.1	-4.0
14	NVP-BMJ791-NX-1	61.7	59.5	62.5	68.0	58.3	66.4	62.7
15	NVP-BNC605-AA-1	2.6	-5.7	-4.3	10.9	0.5	-10.8	2.0
16	NVP-BMP569-NX-1	46.8	44.1	49.9	58.6	42.7	43.7	52.4
17	NVP-BNC594-NX-1	5.8	1.3	1.9	9.7	9.5	-4.1	10.8
18	NVP-BME919-NX-1	-3.6	-10.9	-7.5	3.9	-1.9	-1.0	3.5
19	NVP-BML846-NX-1	-1.2	-8.0	-2.6	7.1	4.2	5.0	11.4
20	NVP-BMU249-NX-1	-1.9	-5.1	1.7	11.1	5.9	-4.5	2.8
21	NVP-BNC592-NX-1	7.7	10.5	7.2	24.8	10.4	3.3	7.2
C1	GeA-69	99.6	10.1	12.6	20.4	14.3	1.2	11.3
C0	DMSO	-0.4	-0.4	-0.3	-0.3	-0.3	-0.4	-8.3

Figure 33. PARP14 MD2 45, 000 compound screen – hit selection and validation by AlphaScreen (I). (A) Hit selection from 30,000 compound screen based on screening results obtained at Novartis. Due to the poor reproducibility of the screening results obtained at the SGC, new stock solution were obtained and characterised at Novartis for PARP14 MD2 inhibition. PARP14 MD2 inhibition at a compound concentration of 25 μM is shown. Compound macrodomain inhibition and counterscreen activity (%) were determined at Novartis in a single experiment (N=1, n=1) and at the SGC in a single experiment with two replicates (N=1, n=2). Compound macrodomain inhibition (%) measured with new compound stock solutions were determined at Novartis in a single experiment with two replicates (N=1, n=2). IC₅₀ values are mean of two replicates (N=1, n=2). Where indicated with '---', no counterscreen or IC₅₀ value could be obtained for the hit due to the unavailability of the compound stock. (B) Validation of hit compounds by macrodomain cross-screening using the old compound stock solutions. Values are mean of the macrodomain inhibition (%) in the AlphaScreen assay of two independent experiments with each two replicates (N=2, n=2). Note: The molecular structures of the compounds are shown in Table S3.

In summary, for all of the selected hit compounds from both screens, binding to PARP14 MD2 could not be confirmed. No ITC titration thermograms could be obtained for the compounds that verified specific binding to PARP14 MD2 [Figure S15 and S16]. From the first 30,000 compound screen, the most promising compounds were NVP-BMM025-NX-1 and NVP-BMF069-NX-1 due to their consistent activity on PARP14 MD2 (regardless of whether old or new compound stocks were used) and their mixed activity profile over different macrodomain proteins in the AlphaScreen assay. Of both compounds, NVP-BMF069-NX-1 showed slightly better BLI and ITC profiles and was therefore characterised in more detail [Figure 34]. NVP-BMF069-NX-1 showed good solubility in the assay buffer (probably also due to its carboxylic acid group) that allowed working with high compound concentrations which is beneficial for confirming weak binders. Indeed, after optimising ITC conditions regarding protein and compound concentrations, heat upon NVP-BMF069-NX-1 injection to PARP14 MD2 could be measured indicating a binding event [Figure 34B]. However, the integrated and normalised binding enthalpies could not be fitted to a function of binding site saturation with the expected 1:1 stoichiometry, even when significant variations of the protein and compound concentrations were considered. Moreover, in ITC competition experiments [Figure 34C], NVP-BMF069-NX-1 was unable to compete with ADPR binding to PARP14 MD2 even when present at a concentration of 2 mM which would have been expected considering its previously determined IC_{50} value of $\sim 5 \mu M$. Closer analysis of the BLI sensorgram showed furthermore that NVP-BMF069-NX-1 binds very likely non-specifically to the macrodomain since a very similar profile was obtained when testing NVP-BMF069-NX-1 against SPIN1, an unrelated tudor domain protein [Figure 34D]. This observation was subsequently confirmed for the majority of the other hit compounds showing highly similar BLI sensorgrams when screened against PARP14 MD2 and SPIN1. Finally, since NVP-BMF069-NX-1 showed inhibitory activity on MacroD1 in the macrodomain cross-screen, NVP-BMF069-NX-1 was soaked at a concentration of 10 mM into MacroD1 *apo* crystals which were provided by a group member. Analysis of the X-ray diffraction data showed that the compound was not present in the crystal. All of these observations indicated that NVP-BMF069-NX-1 does not specifically interact with PARP14 MD2 and might probably interfere with the AlphaScreen assay itself, leading to the

positive AlphaScreen result in contrast to the remaining assays. Counterscreening of NVP-BMF069-NX-1 revealed that this compound indeed showed some interference in the AlphaScreen assay, excluding it finally as real binder to PARP14 MD2.

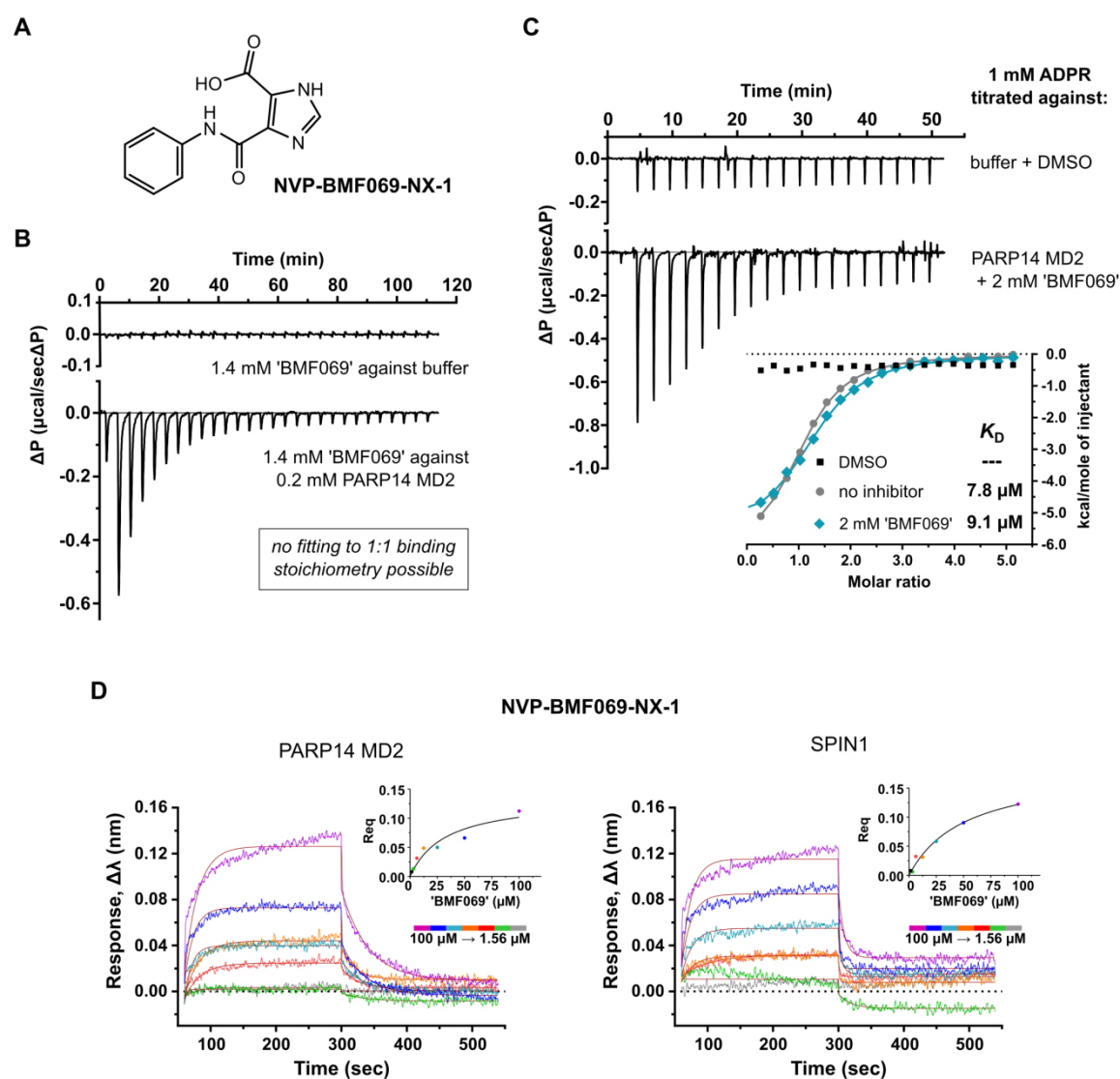


Figure 34. Final hit validation of the PARP14 MD2 screening hit NVP-BMF069-NX-1 resulting from 30,000 compound screen. (A) Molecular structure of NVP-BMF069-NX-1. (B) Raw injection heats resulting from titration of NVP-BMF069-NX-1 against PARP14 MD2 at the indicated concentrations are shown. Normalised binding enthalpies could not be fitted to an approximately 1:1 binding stoichiometry ($N = 0.02$) indicating non-specific binding of NVP-BMF069-NX-1 to PARP14 MD2. (C) NVP-BMF069-NX-1 is not able to compete with ADPR binding on PARP14 MD2 even if present at a concentration of 2 mM. Raw injection heats of the experiment are shown on the top, normalised binding enthalpies with the calculated K_D for ADPR in the inset. Additional binding and fitting parameters are provided in Figure S17. (D) Problematic of hit compound validation by BLI at high compound concentrations. Shown are BLI sensorgrams obtained for NVP-BMF069-NX-1 which was screened against PARP14 MD2 and SPIN1, a macrodomain unrelated tudor domain protein, at a concentration of 100 μM . For PARP14 MD2, a K_D of $(120 \pm 73) \mu\text{M}$ was calculated, for SPIN1 a K_D of $(52 \pm 11) \mu\text{M}$. The examples are representative for the majority of the hit compounds and illustrate that the PARP14 MD2 binding signal is probably due to non-specific binding events. The inhibitor concentration range and steady state analysis are shown next to the traces.

The validation process of the hit compounds selected from the 15,000 compound screen was performed similarly, however, without screening of the hit compounds by BLI due to the problems of non-specific binding at high assay concentrations that were observed previously. All selected hit compounds were first tested for their reproducibility of PARP14 MD2 inhibition in the AlphaScreen assay. As shown in Figure 35A, the reproducibility of the primary screening results generated at Novartis is higher than of the 30,000 compound screen although not all of the compounds showed still the same inhibitory activity on PARP14 MD2. Furthermore, the IC_{50} values obtained were decreased at least two-fold and up to eight-fold, with the majority between 10-50 μ M instead of 2-10 μ M as measured at Novartis. All hit compounds were also screened at a concentration of 25 μ M over several macrodomains by AlphaScreen [Figure 35B]. It is unclear why the PARP14 MD2 inhibition at 25 μ M in the cross-screen was repeatedly not in good agreement with the inhibitory activity of PARP14 MD2 previously determined at this concentration. In the macrodomain cross-screen, only NVP-BNU459-NX-1 showed high inhibition of PARP14 MD2, however, had also significant inhibitory activity on most of the other macrodomains. NVP-BMF352-NX-1 showed activity on all macrodomains with an apparent inhibition similar to its counterscreen activity, and was therefore assumed to be an assay interfering compound and excluded from further hit validation.

Legend to Figure 35: PARP14 MD2 45, 000 compound screen – hit selection and validation by AlphaScreen (II).

(A) Hit selection from 15,000 compound screen based on screening results obtained at Novartis. The result of the re-screening of the compounds at the SGC is shown next to the Novartis screening results. PARP14 MD2 inhibition at a compound concentration of 25 μ M is shown. Compound macrodomain inhibition and counterscreen activity (%) were determined at Novartis in a single experiment (N=1, n=1) and at the SGC with three replicates from two independent experiments (N=2, n=3). IC_{50} values were determined at Novartis in single dose-response experiments (N=1, n=1). IC_{50} values determined at the SGC are mean of three replicates from two independent experiments (N=2, n=3). **(B)** Validation of hit compounds by macrodomain cross-screening. Values are mean of the macrodomain inhibition (%) in the AlphaScreen assay of three independent experiments with each two duplicates (N=3, n=2). *Note:* The molecular structures of the compounds are shown in Table S3.

A

	Compound ID	Novartis			SGC	
		PARP14 MD2 inhibition (%)	Counterscreen activity (%)	IC ₅₀ (μM)	PARP14 MD2 inhibition (%)	IC ₅₀ (μM)
22	NVP-BNU459-NX-1	99.9	39.3	2.1	99.9	1.8
23	NVP-BNV428-NX-1	99.9	35.8	2.5	86.8	11
24	NVP-BNV427-NX-1	99.9	22.8	5.0	73.8	18
25	NVP-BNU457-NX-1	99.8	20.1	3.4	65.7	27
26	NVP-BNM150-NX-1	99.7	30.1	3.9	68.3	21
27	NVP-BNX029-NX-1	99.7	33.0	1.8	49.4	31
28	NVP-BNU542-NX-1	99.7	5.1	4.5	89.6	15
29	NVP-BNU540-NX-1	99.4	7.0	6.4	65.5	23
30	NVP-BNV433-NX-1	99.1	23.6	2.9	66.8	22
31	NVP-BNQ598-NX-1	98.8	20.0	22	28.4	66
32	NVP-BMX995-NX-1	96.7	33.1	13	47.8	56
33	NVP-BMY002-NX-1	96.3	0.9	19	35.9	>100
34	NVP-BNV425-NX-1	94.4	11.2	11	39.9	>100
35	NVP-BMF352-NX-1	91.9	39.9	6.7	63.4	44
36	NVP-BNV413-NX-1	88.0	8.6	7.9	55.1	37
37	NVP-BMY031-NX-1	85.5	29.2	2.7	77.8	1.1
38	NVP-BNV442-NX-1	82.5	6.7	8.3	41.5	55
39	NVP-BNV240-NX-1	81.6	13.1	14	25.3	47
40	NVP-BNV410-NX-1	79.8	4.1	18	31.5	>100
41	NVP-BNL153-NX-1	51.7	0.3	22	21.8	64

B

	Compound ID	(%) Macrodomain inhibition at 25 μM compound concentration						
		PARP14 MD2	PARP14 MD3	PARP9 MD2	MacroD1	MacroD2	TARG1	macroH2A1.1
22	NVP-BNU459-NX-1	99.8	95.9	81.3	65.9	31.4	97.7	92.6
23	NVP-BNV428-NX-1	55.9	33.6	2.1	17.1	-3.0	56.0	52.9
24	NVP-BNV427-NX-1	33.2	23.4	8.3	25.5	3.2	55.7	50.9
25	NVP-BNU457-NX-1	31.7	24.6	18.2	9.8	13.3	29.6	47.9
26	NVP-BNM150-NX-1	22.8	20.9	8.2	22.8	11.3	33.6	31.5
27	NVP-BNX029-NX-1	24.4	21.5	15.2	14.0	10.5	25.1	43.6
28	NVP-BNU542-NX-1	-1.3	5.1	-1.6	2.3	3.3	-7.3	9.7
29	NVP-BNU540-NX-1	0.5	11.0	3.3	4.6	3.7	29.6	30.1
30	NVP-BNV433-NX-1	34.7	30.3	7.5	18.8	6.9	44.4	51.2
31	NVP-BNQ598-NX-1	2.9	9.8	0.9	4.2	5.9	1.7	6.1
32	NVP-BMX995-NX-1	17.2	28.0	21.0	23.9	21.1	33.1	21.6
33	NVP-BMY002-NX-1	-0.8	5.0	-6.5	1.8	0.2	7.9	8.1
34	NVP-BNV425-NX-1	8.8	7.3	0.6	5.5	5.2	35.2	29.1
35	NVP-BMF352-NX-1	32.9	40.5	32.7	34.5	31.7	57.4	40.0
36	NVP-BNV413-NX-1	10.2	20.0	6.7	16.5	3.6	40.6	38.1
37	NVP-BMY031-NX-1	14.1	22.2	14.3	16.5	18.2	8.5	26.2
38	NVP-BNV442-NX-1	3.3	7.9	6.1	5.4	2.1	31.0	26.7
39	NVP-BNV240-NX-1	3.4	4.1	6.6	5.8	6.6	9.5	17.0
40	NVP-BNV410-NX-1	-1.0	2.3	-1.9	1.1	1.6	12.2	12.5
41	NVP-BNL153-NX-1	-2.8	5.0	1.5	1.6	5.3	-1.5	-7.7
C1	GeA-69	99.7	13.2	11.2	12.4	12.1	15.1	8.3
C0	DMSO	-0.2	-0.3	-0.2	-0.2	-0.2	-0.3	-8.7

Figure 35. PARP14 MD2 45, 000 compound screen – hit selection and validation by AlphaScreen (II). For legend, see bottom of previous page.

The majority of the compounds were then characterised for their binding affinities to PARP14 MD2 by ITC using the MicoCal200 which is slightly less sensitive compared to the VP-ITC

instrument, yet allows a higher sample throughput. Apart for NVP-BMX995-NX-1, no ITC titration thermogram was obtained for the other hit compounds that revealed an obvious binding event to PARP14 MD2 under the conditions tested [Figure S16]. As final validation step, NVP-BNU459-NX-1 and NVP-BMX995-NX-1 were tested in additional ITC experiments [Figure 36] due to their consistent AlphaScreen activity and a good initial ITC thermogram, respectively. Neither could binding of the hit compounds to PARP14 MD2 be detected in direct titrations of the compounds to the protein nor were both compounds able to compete with ADPR binding to PARP14 MD2 under the conditions tested. Due to the poor solubility of both compounds in assay buffer, their concentration in the competition experiments was limited to 100 μM . This concentration might be insufficient to completely prevent ADPR binding to PARP14 MD2, however, it was expected that a slight decrease in the binding affinity of ADPR could be measured. Therefore, NVP-BNU459-NX-1 and NVP-BMX995-NX-1 were not verified as real binders to PARP14 MD2.

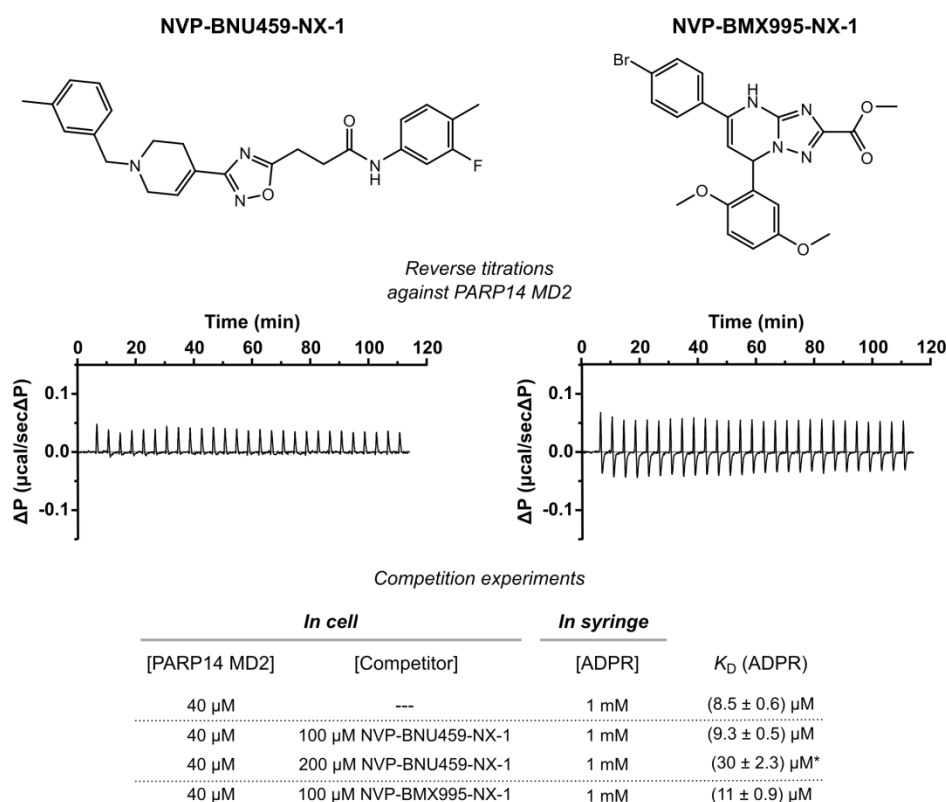


Figure 36. Final hit validation of the PARP14 MD2 screening hits NVP-BNU459-NX-1 and NVP-BMX995-NX-1 resulting from 15,000 compound screen. (top) Molecular structure of the hit compounds NVP-BNU459-NX-1 and NVP-BMX995-NX-1. **(middle)** Binding of NVP-BNU459-NX-1 and NVP-BMX995-NX-1 cannot be detected by ITC. Shown are raw injection heats resulting from titration of PARP14 MD2 into a solution with the corresponding compounds. **(bottom)** NVP-BNU459-NX-1 and NVP-BMX995-NX-1 are not able to compete with ADPR to PARP14 MD2 shown by ITC competition experiments. * Protein precipitation was observed in presence of 200 μM NVP-BNU459-NX-1. Additional binding and fitting parameters are provided in Figure S17.

Conclusions from the 45,000 compound screen

In vitro screening is one of several approaches to generate new chemical hit matter. Although being rather resource-intensive, it provides a relatively unbiased way to identify compounds with new chemical scaffolds and binding modes for targets such as macrodomains for which knowledge concerning its druggability is limited so that a focused screen is not possible^{185,186}. Therefore, a 45,000 compound screen against PARP14 MD2 was performed which was set-up in a high-throughput screening manner. Only a few hit compounds could be obtained from the screen that showed an acceptable initial profile concerning PARP14 MD2 inhibition and assay interference potential. 41 hit compounds were finally selected based on these criteria of which none could be verified as real binders to PARP14 MD2 in the following hit validation process by BLI, ITC or macrodomain cross-screening. The latter enabled in particular the identification of unselective, assay interfering compounds that were not filtered out by prior counterscreening. Since the identity and nominal concentration of all hit compound stock solutions were confirmed (with a few exceptions) by LC-MS [Figure S14], the quality of the compounds and an associated incorrect assay concentration can be excluded as reason for the outcome of the validation process. Closer retrospective analysis of the selective compounds however revealed that the majority of the hits were similar to compounds previously reported as aggregators¹⁸⁷ and/or had a high calculated logP between 3.8 and up to even 6.9. Although detergent was used in the AlphaScreen assay to prevent compound aggregation, the high lipophilicity of the compounds may be the reason for the limited reproducibility and the high false positive hit rate of the AlphaScreen results for several of the compounds. Compounds with high clogP are more likely to non-specifically interact with proteins which cannot always be detected by counterscreening. In several cases, the high compound lipophilicity can however be seen to be reflected in the increased counterscreen activity and in the ITC thermograms of the corresponding screening hits. The poor solubility in the assay buffer leads to formation of compound aggregates which in turn can promote protein instability and aggregation, explaining the significant noise in several of the ITC thermograms but

could also explain why binding of the compounds could not be confirmed. Since the NDDI library consisted of compounds selected for their structural and drug-like properties based on Lipinski's 'rule-of-5', compounds with a clogP in particular greater than 5 were not expected to be present. However, the clogP of the screening hits should in general have been considered for hit selection to improve the hit validation process.

Nevertheless, several hit compounds showed good (initial) PARP14 MD2 inhibition, acceptable AlphaScreen assay interference and were not previously reported as aggregators with a clogP between 1.7 and 2.3. Examples, for which also the primary screening results could be broadly reproduced, are compounds from the series NVP-BNV410-NX-1 to NVP-BNV442-NX-1 as well as NVP-BNV240-NX-1. The latter and two compounds of the series (NVP-BNV413-NX-1, NVP-BNV428-NX-1) were tested by ITC at a compound concentration of 50 μ M to 80 μ M and their binding to PARP14 MD2 could not be verified. NVP-BNV428-NX-1 showed an AlphaScreen IC_{50} of 11 μ M (which is probably slightly overestimated due to its increased AlphaScreen assay activity) and NVP-BNV413-NX-1 and NVP-BNV240-NX-1 of 37 μ M and 50 μ M, respectively. With a calculated Wiseman c -parameter of at least 1 (defined as $c = n \times [M] \times K$, with the binding stoichiometry n , the association constant K and the macromolecule concentration in the cell $[M]$ (here the compound concentration due reverse performed titration)), it was assumed that the ITC concentration should allow at least the detection of compound binding^{188,189}. However, it is also very likely that far higher compound concentrations would be required to confirm binding which could not be realised due to the limitation of the compound solubility.

NVP-BMF069-NX-1 was the most promising compound from the first hit validation process. NVP-BMF069-NX-1 is highly soluble with a clogP of 0.9 which can be attributed to its carboxylic acid attached to the imidazole based scaffold. Although NVP-BMF069-NX-1 showed a mixed activity profile in the AlphaScreen macrodomain cross-screen with high activity on PARP14 MD2, calorimetric experiments indicated that NVP-BMF069-NX-1 interacts only non-specifically with the macrodomain. It is possible that NVP-BMF069-NX-1 interferes with the AlphaScreen assay due to its delocalised electron system, thereby potentially absorbing singlet

oxygen in the AlphaScreen assay, and/or can electrostatically interact with its negative charge at neutral pH with some of the macrodomains. Also NVP-BNU459-NX-1 and NVP-BMX995-NX-1 showed good activity in the AlphaScreen assay or in initial calorimetric experiments but could not finally be verified as specific binders. It was noted that both compounds have an alkene next to an aromatic system, an oxadiazole in NVP-BNU459-NX-1 and a triazole in NVP-BMX995-NX-1, and it is very likely that the alkene could be functionalised through delocalisation with the aromatic system to non-specifically interact with the protein.

It cannot be excluded that true specific binders of PARP14 MD2 may have been missed either in the hit selection processes, due to the assay conditions or the assay format itself. As reported previously, it is rather unlikely that true hits were missed in the primary screening process itself because, although all compounds were screened just once at a single concentration, several compounds of a certain chemotype (in many cases found to show similar assay activity) are generally included in the compound library¹⁸⁶. This could be confirmed with the screening hits obtained. The majority of hits could be classified to only a few chemotype groups. During the hit selection process, it is however possible that compounds were excluded due to their counterscreen activity that would have been actually shown also true inhibitory activity on PARP14 MD2. Although those compounds might not be desirable hit compounds, they would nevertheless be valuable starting points for hit optimisation and templates for a second, scaffold focused screen. The screening results may have been also influenced by the assay conditions, where in particular the choice of reducing agent should be mentioned. It was shown that the use of reducing agents can produce both false positives and false negatives, and therefore have potential risks of missing positive hits¹⁸¹. The screening concentration of the compounds is another factor that could be considered. A slightly higher screening concentration of for instance 50 μ M would have included more of the compounds from below the 50 % inhibition cut-off into the hit validation process. Finally, it is also likely that an alternative assay, for instance an HTRF assay, results in slightly different screening hits including potentially a real binder. Although the HTRF assay has a rather limited dynamic range compared to the AlphaScreen assay, the advantage of time-resolved fluorescence assays lies in their sensitivity resulting from a very low background due to the use of

fluorophores such as lanthanides with long fluorescence decay time. Thus, by the time of the TRF assay signal detection, most of the fluorescence of organic fluorophores, i.e. from chemical compounds, is decayed, leaving as only source of signal the emission of the acceptor fluorophore by specific excitation of the donor fluorophore^{178,190}.

The outcome and in particular success of an *in vitro* screen depends on many factors, certainly on the druggability of a target, however, as well as on the size and quality of the compound library. With increasing size and diversity of a screening collection, the probability increases to identify a true hit for a new target. In order to discover the first-in-class PARG macrodomain inhibitor PDD00017273, a 1.4 million compound screen (conducted by AstraZeneca) was required from which just three structurally related hit compounds with an anthraquinone scaffold could be identified as true hits for PARG¹³¹. It is the limited ‘drug-like’ chemical space that is covered by the screening libraries that requires depending on the target druggability large screening campaigns to generate new hit matter. An alternative approach which allows covering a much wider chemical space without the need of performing a resource-intensive HTS is a fragment-based drug discovery approach (FBDD) which was therefore applied for hit identification attempts for PARP14 MD3.

2.2. A FBDD approach using X-ray crystallography for targeting PARP14 MD3

In contrast to traditional high-throughput screening, fragment-based drug discovery (FBDD) is characterised by screening of only a limited number of compounds (typically ~1000) of relatively low molecular weight (150-300 Da)^{185,191}. The concept behind this strategy is that small molecules (fragments) can sample the chemical space more efficiently than larger compounds of higher complexity. A fragment is representative for a large amount of compounds that can incorporate or be inspired by that fragment. In order to be identified as hit in a HTS, the compound needs to engage in several interactions with the target which is directly associated with a certain potency of the hit compound, typically in the 30 μ M – nM range but also with a higher screening effort to find a suitable candidate for the binding site. Due to their far lower complexity, fragments have limited interaction possibilities with the target but the probability of matching to a certain binding site is higher, resulting in a higher hit rate. Fragments are thus held by the target over very few contacts which is directly reflected in a low binding affinity of fragments, typically in the mM to 100 μ M range^{185,191}. Therefore, fragment libraries need to be screened with far more sensitive screening techniques compared to HTS libraries. Instead of *in vitro* biochemical or cellular assays, surface plasmon resonance (SPR), nuclear magnetic resonance (NMR) or X-ray crystallography are particularly used to identify fragment hits with the advantage of the latter to directly provide information of the fragment-protein binding mode^{192,193}. While for HTS hits the emphasis lies on their potency, fragment hits are selected in particular for high ligand efficiency, i.e. the binding energy of the ligand per non-hydrogen atom^{191,194}. The potency of the fragments is then improved through subsequent medicinal chemistry with three main strategies: fragment linking, in which two or more fragments that bind independently to distinct sites are linked together whereby the ideal linkage can result in an additivity effect additionally increasing the potency, fragment growing (more common), thus appending additional chemical groups to establish more interactions within the binding site, and fragment merging, where structural features of the fragments are combined^{191,195,196} [Figure 37]. Both HTS and the FBDD have their

advantages and limitations for finally generating new hit and later lead compounds, however, based on the prior *in vitro* screening experience, the latter seemed to be a worthwhile alternative approach for attempting the identification of new chemical starting points for PARP14 macrodomains.

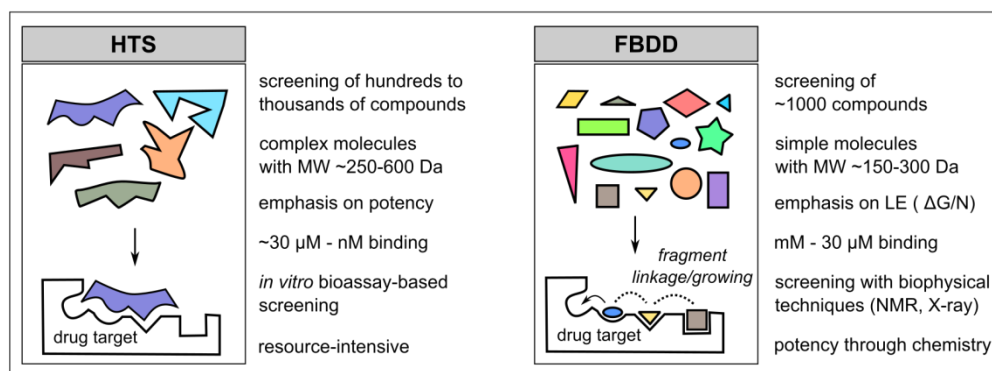


Figure 37. Comparison of HTS (high-throughput screening) and FBDD (fragment-based drug discovery) principle. Main characteristics of both screening approaches as reviewed by Rees *et al.*¹⁹¹ are listed on the side of the schematics. LE: ligand efficiency. Dashed lines: representative for fragment linkage. Arrow: representative for fragment growth. The figure was adapted from Park *et al.*¹⁹⁷.

2.2.1. Performance of a crystallographic fragment screen against PARP14 MD3

As shown previously, PARP14 macrodomain inhibitor development needs to address both MD2 and MD3 to significantly impair PARP14 recruitment to target sites. Since GeA-69 already provided in principle the possibility to selectively target PARP14 MD2 in cells, it was of high interest to also identify new chemical scaffolds targeting PARP14 MD3. The FBDD approach was therefore evaluated with PARP14 MD3, whereby X-ray crystallography was the preferred fragment screening strategy since it allows the identification of fragment hits directly together with the analysis of their binding mode due to the structural information obtained. This could directly guide evolving the fragments to larger, more potent compounds by follow-up chemistry. Furthermore, the accessibility to the recently established X-ray fragment screening pipeline of the SGC in collaboration with the XChem fragment-screening facility at Diamond Light Source could technically enable this approach. The only requirement for performing the screen was the ability

to reliably crystallise PARP14 MD3 in its *apo* form. PARP14 MD3 was previously crystallised by Forst *et al.* resulting in a high resolution structure of this macrodomain⁶⁶. In order to test the crystallisability of PARP14 MD3, crystallisation plates testing two course screens (HIN3, JCSG7) were set-up with pre-purified protein that was however prior buffer exchanged and concentrated to 40 mg/mL to follow the previous crystallisation procedure by Forst *et al.*⁶⁶. Only one hit consisting of a dense cluster of thin, plate-like crystals could be obtained in a similar condition that was previously reported. The hit was subsequently optimised to crystals diffracting to 1.2-1.4 Å, which confirmed the reproducibility of PARP14 MD3 *apo* crystallisation. For efficient performance of a fragment screen by X-ray crystallography, a very reliable crystallisation system for PARP14 MD3 had to be established. The aims were set high due to the very restricted time available at the synchrotron and its XChem facility for performing this screen. Therefore, this system should allow producing crystals in at least 50 % of the wells of a crystallisation plate within a few days with several crystals in each drop. Since each fragment is added to only one well and re-soaking could not be afforded due to time constraints, the latter should ensure that fragments are not lost for screening due to the loss of a crystal during the mounting procedure. The aim was to have crystals of 100-150 µm in size with good morphology growing loosely in the drop and also with no skin formation on the drop, which should allow quick crystal harvesting, with the latter being also important for efficient soaking of the compound into the crystal. The crystals should furthermore consistently diffract to high resolution, tolerate high solvent i.e. DMSO concentrations and not require complicated cryoprotection.

The establishment of this crystallisation system is shown in Figure 39. Following the X-ray diffraction testing of the PARP14 MD3 crystals, variations of crystallisation conditions were tested with the pre-purified protein in an attempt to obtain single crystals without a surrounding crystal cluster (similar to the initial hit). Furthermore, three follow-up conditions were selected to test the reproducibility of PARP14 MD3 crystallisation. Although conditions could be found to reduce the crystal cluster formation, the crystals obtained with these screens were of rather poor morphology and showed insufficient reproducibility (~20 %). Therefore, PARP14 MD3 was freshly purified on a large scale [Figure 38].

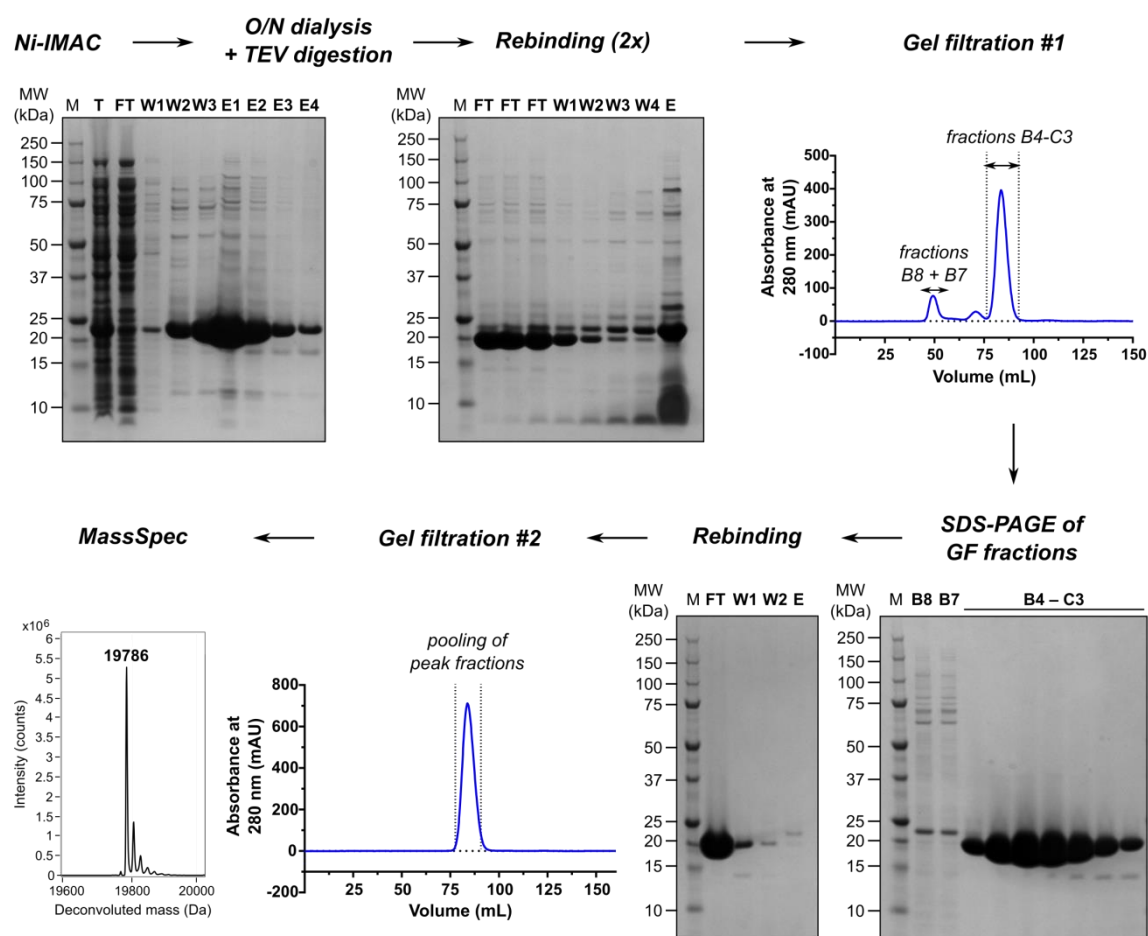


Figure 38. Purification of PARP14 MD3 for crystallisation for the crystallographic fragment screen. After recombinant overexpression in *E. coli* and cell lysis by high pressure homogenisation, PARP14 MD3 protein was purified by Ni-IMAC, followed by His₆-tag cleavage by TEV protease digestion and dialysis in imidazole-free buffer overnight of the eluate fractions E1-4. The dialysed protein samples (three batches, FT1-3) were purified again by Ni-IMAC ('Rebinding'; performed twice to remove impurities including the cut His₆-tags, the TEV protease and uncleaved PARP14 MD3). The pooled flow-through fractions were further purified by gel filtration (GF) using a sepharose based S75 16/60 column. GF fractions were analysed for purity by SDS-PAGE and fractions with high target/background ratio were pooled (here as example: B4 – C3) and passed through a low volume of Ni-sepharose to remove any remaining impurities. The flow-through was additionally purified by gel filtration and only the peak fractions (not including fractions from the peak tails) were concentrated to 40 mg/mL, flash frozen in liquid N₂ and stored at -80 °C. The protein identity was confirmed by measuring the intact mass by mass spectrometry (expected mass: 19786 Da). (T = total - whole cell lysate, FT = flow-through, W = wash fractions with increasing imidazole concentrations of 20-60 mM, E = eluate fractions with increasing imidazole concentrations of 250-350 mM).

Consequently, the PARP14 MD3 protein was set-up in a follow-up crystallisation screen designed around the initial hit conditions. Probably due to the purity of the protein sample and lack of spontaneous formation of nucleation sites, only one single crystal, however of good size, could be obtained. This crystal was subsequently used for seeding and by optimisation of the crystallisation

parameters and seeding procedure, high quality crystals could be reliably obtained with reproducibility > 85 %. The seeding procedure was crucial to establish this crystallisation system, since by using seeds, that provide the nucleation sites for crystal growth, the crystal formation did not depend on spontaneous nucleation anymore but could be well controlled to yield crystals of size and amounts as required additionally in only a few hours [Figure 40]. Thus, the crystal growth was completed within 24 hours and by using a dilution of 1:500 of the seed stock, 5-10 crystals grew in each drop of an average size of 100 μm . Higher seed concentrations resulted in too many and small crystals whereas with higher dilutions too few and large crystals were obtained. The previous crystallisation conditions including protein concentration, crystallisation solution and temperature had however to be adjusted with the seeding procedure. Initial seeding attempts using crystallisation conditions in which the seeding crystal was obtained with high protein concentrations at 20 °C resulted in uncontrolled fast nucleation and crystal growth. The contact of the pipetting robot needle with the drop (for adding the seeds) induced additional strong nucleation (as seen at the ring like structures in the drop) which used-up the protein preventing the formation of larger crystals. Therefore, screens at 20 °C and 4 °C titrating the protein and precipitant concentrations were performed. The decrease of the protein concentration to 30 mg/mL and the crystallisation temperature from 20 °C to 4 °C using crystallisation plates pre-equilibrated to 4 °C were found to be crucial to enable a controlled crystal growth initiated through nucleation by the seeds only. Lower protein concentrations significantly reduced crystal formation and lower precipitant concentrations in particular of the PEG polymer (PEG2kMME) resulted in unstable crystals. Although crystals were obtained at all PEG polymer concentrations (20-30 %) after 24 hours, only at high concentrations of 26-30 % the crystals were stable over 7 days and not prone to re-dissolving. The final crystallisation conditions are summarised in Figure 40C.

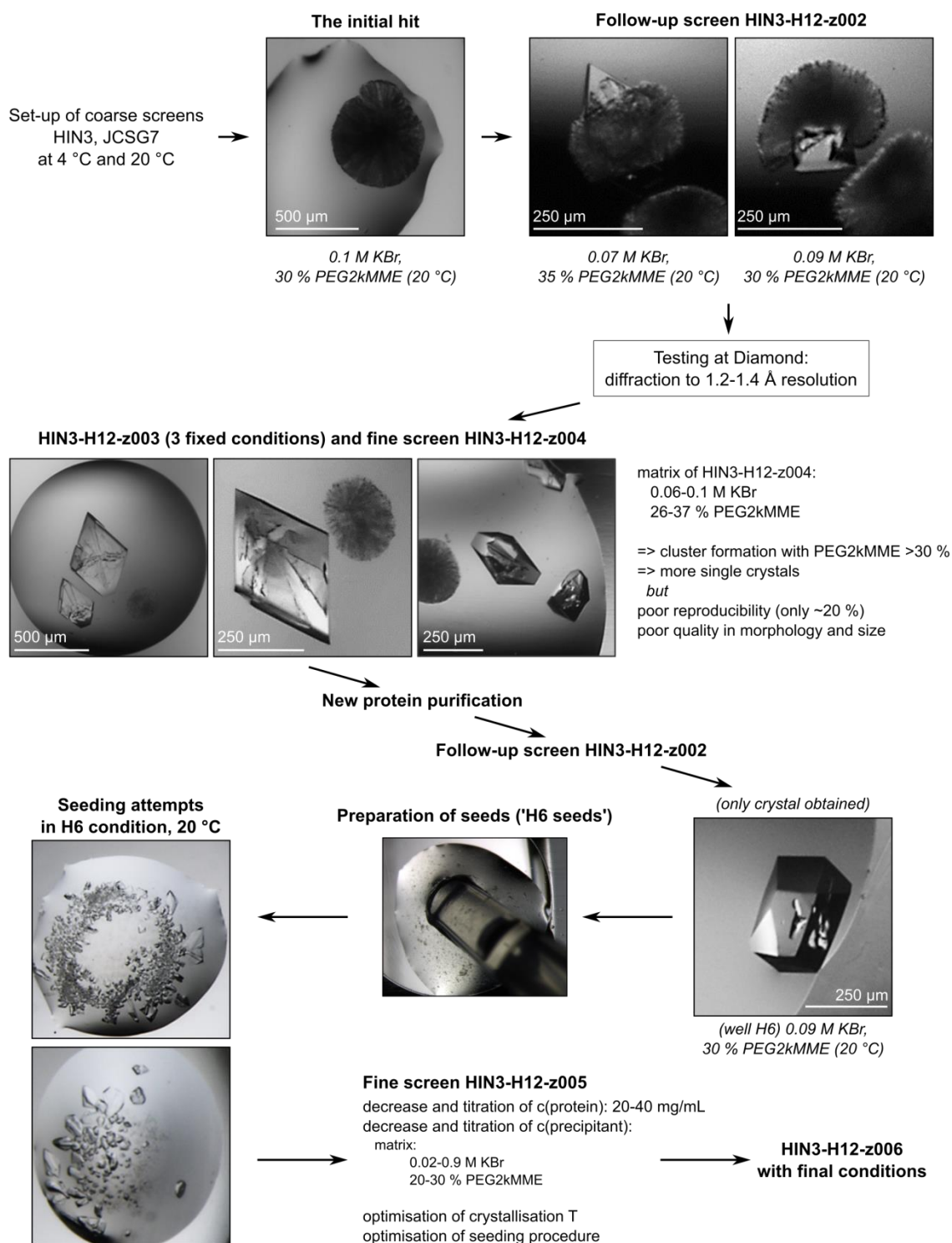


Figure 39. Optimisation of PARP14 MD3 apo crystallisation for performing a crystallographic fragment screen. Crystallisation attempts from the initial hit until HIN3-H12-z004 were performed with pre-purified and stored protein. The following crystallisation optimisation was performed with freshly purified protein (the 'new protein purification' is shown in Figure 38). The protein concentration used for obtaining the initial hit was 40 mg/mL. PEG2kMME: polyethylene glycol 2000 monomethyl ether.

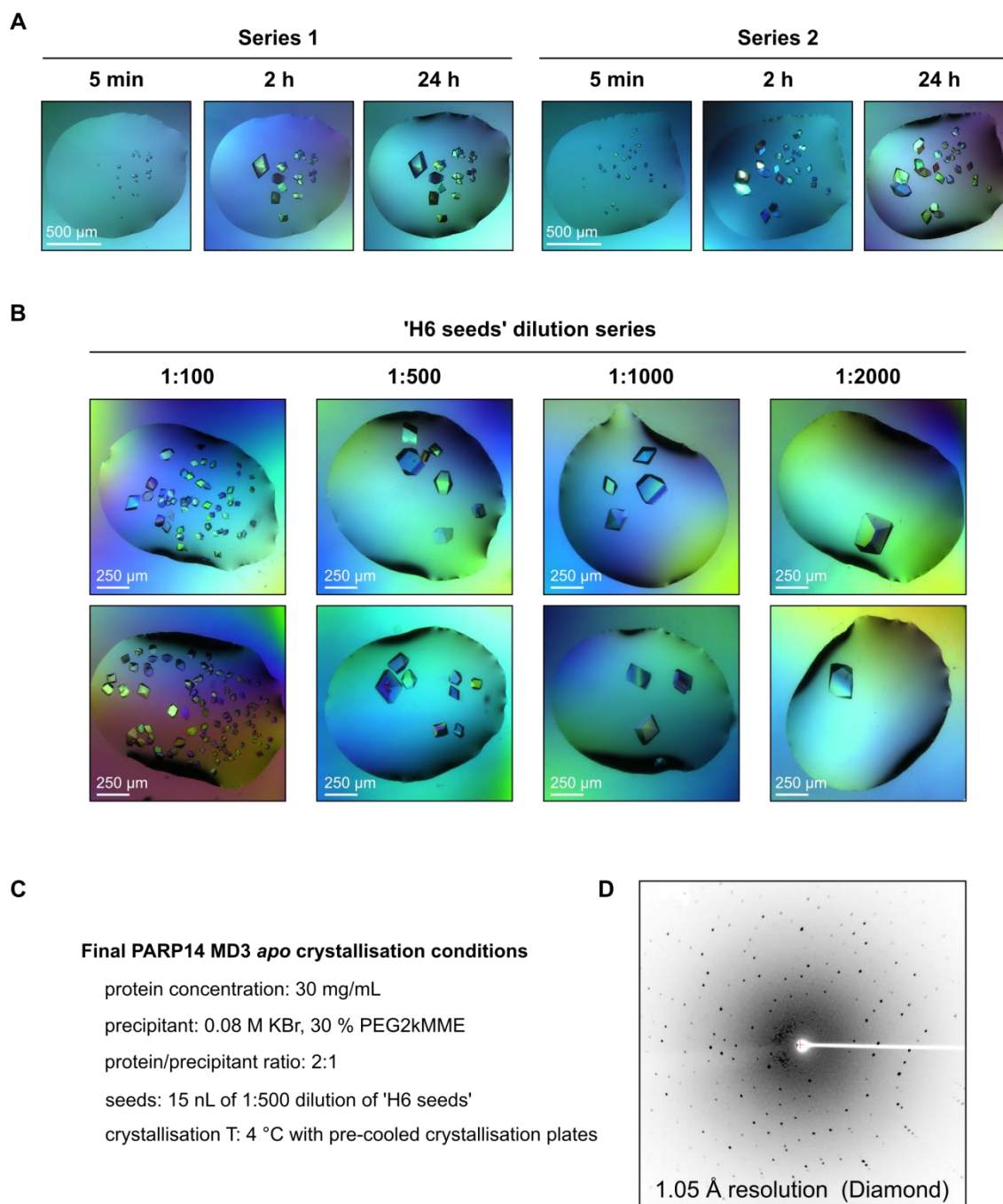


Figure 40. PARP14 MD3 *apo* crystallisation. **(A)** With the optimised crystallisation conditions, crystal growth is completed after 24 hours. Examples of two time series are shown. **(B)** 'H6 seeds' dilution series to determine the best seeding concentration aiming for several (5-10) crystals per drop with good size for mounting into D (100 μm) and E (150 μm) loops. **(C)** Final PARP14 MD3 *apo* crystallisation conditions. **(D)** Example of a diffraction pattern obtained from a PARP14 MD3 crystal from in-house screening (the whole diffraction pattern is shown with the edges of the detector being at 2.6 Å and the corners being at 2.0 Å). Diffraction to 1.05 Å resolution was confirmed at I0-4 (Diamond Light Source) for a series of 30 crystals in a solvent tolerance test.

Following the set-up of this crystallisation system for PARP14 MD3 *apo* and its validation for reliably yielding crystals with high X-ray diffraction by in-house screening, a solvent tolerance screen was performed at the XChem facility at Diamond Light Source to determine the amount of solvent that is tolerated by the crystals. Due to their in general weak binding affinity, fragments solubilised mainly in DMSO are soaked with rather high concentration to ensure a good ligand occupancy in the crystal. However, the resulting DMSO concentrations can affect the integrity of the crystal and thus the X-ray diffraction^{198,199}. Therefore, the crystals were tested in triplicates for X-ray diffraction after soaking for 4 hours at final DMSO concentrations of 5-30 % (v/v) or at final ethylene glycol concentrations of 10-30 % (v/v). The latter was tested since a subset of the fragments of the library was solubilised in this solvent. The crystal soaking was performed by acoustic dispensing as recently set-up at the XChem facility, where the solvents or fragments are transferred from the solvent- or fragment-library plate to the (inverted) sitting-drop crystallisation plate containing the crystals using an Echo 550 liquid handling system [Figure 41B]. The diffraction of the 33 crystals was not affected by any of the DMSO or ethylene glycol concentrations compared to the non-soaked control crystals, with an average diffraction of 1.06 Å. The crystals diffracted to 1.05 Å even when soaked at 30 % DMSO, allowing high fragment concentrations for the screen. Additionally, a test screen of 50 fragments was performed at a fragment concentration of 50 mM with average diffraction to 1.1 Å resolution. It should be mentioned that for the solvent tolerance screen as for the fragment screen, no cryoprotectant was used which confirmed previous experiences that by matching the mounting loop to the crystal size and limiting excessive solvent surrounding the crystal, cryoprotection is not required which allows quick crystal harvesting after the soaking procedure^{198,200}. Furthermore, the use of PEG2kMME at a concentration of 30 % may act slightly as a cryoprotectant. Altogether, the crystallisation system fulfilled all requirements set previously so that the fragment screen for PARP14 MD3 could be performed. The final workflow is outlined in Figure 41A.

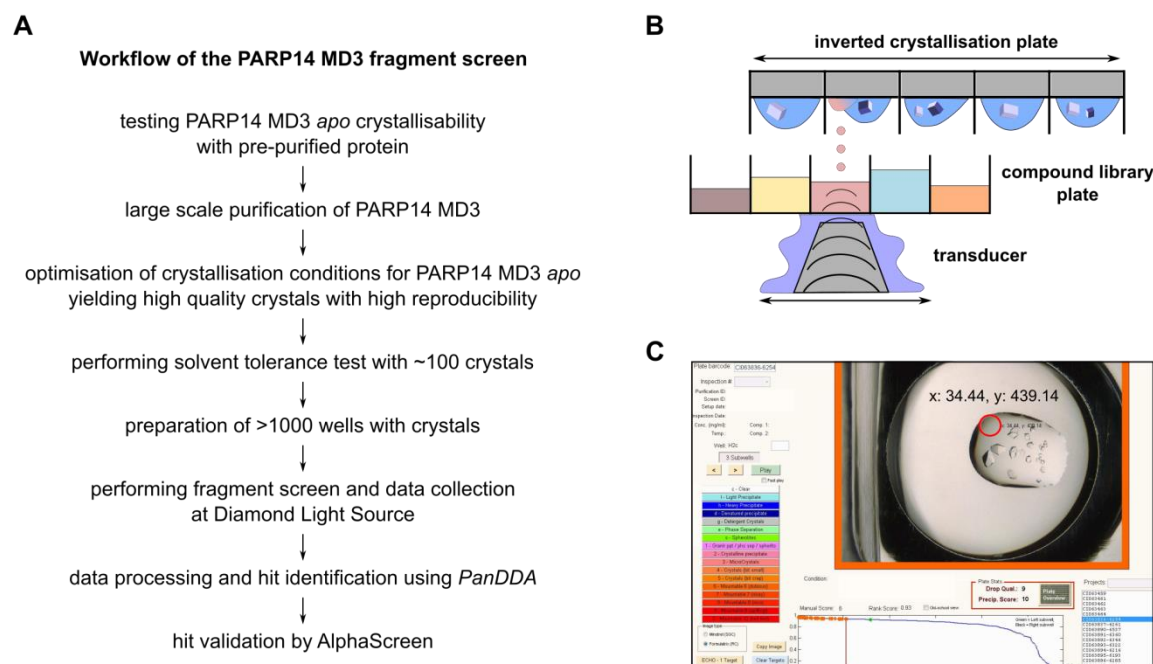


Figure 41. (A) Workflow for performing the crystallographic fragment screen with PARP14 MD3. **(B)** Schematic representation of acoustic droplet ejection for crystal soaking using an Echo 550 liquid handling system. The schematic was adapted from Collins *et al.*¹⁹⁸ **(C)** Echo targeting using *TeXRank*. An excerpt of the interface of *TeXRank* is shown. Clicking at the position highlighted with the red circle records the xy coordinates (written on top of the drop for clarity) for later acoustic dispensing of the fragment by the Echo 550 liquid handling system.

To perform the fragment screen, several crystallisation plates with PARP14 MD3 *apo* crystals (> 1000 wells) were prepared and processed for the Echo targeting using *TeXRank* [Figure 41C]. Three fragment libraries are provided by the XChem facility at Diamond Light Source including the ‘Edelris Keychemical Fragments’, the ‘Maybridge 1000’ and the ‘DSPL2 - Diamond-SGC Poised Library 2.0’ of which the latter was screened against PARP14 MD3. The DSPL2 library is a collection of structurally diverse fragments commercially available from different vendors that were defined as ‘poised’²⁰¹. Poised fragments contain at least one functional group that can be chemically derivatised using a robust and general synthetic reaction. This allows a rapid elaboration of the fragment hit into a library of analogues by parallel chemistry and thus the quick development of a SAR of the respective hits. Due to furthermore positive previous experiences with this library at the SGC²⁰¹, the DSPL2 library was selected for screening against PARP14 MD3. The library consisted of 776 fragments at a compound concentration of 500 mM

of which 645 fragments (due to time constraints) were screened by singleton soaking at a final concentration of 50 mM, calculated based on the initial drop volume. As shown previously, the resulting DMSO concentration of 10 % did not affect the integrity and thus X-ray diffraction of the crystals. All fragments were soaked in batches (usually half of a crystallisation plate was used for one batch) into the crystals for at least 4 hours which was extended for some batches to 6 hours upon delays in processing of the previous batches. The soaking concentration and time needed to reach good occupancy of the fragment in the crystals, and thus to be able to detect a fragment as hit, varies, with the latter ranging from a few minutes to overnight (or even longer) incubations of the crystals with the fragment. This is determined by factors such as the binding affinity of the fragment, its solubility in the crystallisation drop and the nature of the bulk-solvent channels through the crystal lattice^{199,202}. Based on previous experiences, a soaking time of 4 hours is sufficient for most fragments to guarantee their diffusion from the dispensing site through the crystallisation drop into the interior of the crystal over the lattice channels¹⁹⁸.

Following the fragment soaking, the crystals were harvested manually but robot-assisted that allowed an average (final) mounting time of 25 sec per crystal. The X-ray diffraction data were then collected in an automatic mode of the beamline which used diffraction based centring of the loop in the X-ray beam. With one duty cycle, i.e. mounting the crystal containing pin on the goniometer, loop centring, data collection and returning the pin back into the robot storage tank, lasting ~2 min, the diffraction data was collected in slightly less than 24 hours. Out of 645 mounted crystals, 622 workable data-sets were collected with an average resolution of 1.09 Å. Thus, screening data from 23 fragments (3.5 %) could not be obtained because either the crystal showed no diffraction anymore (11 cases) or diffraction could be seen, however, the auto-processing failed (7 cases) or due to crystal mounting mistakes in particular in the beginning of the screen (5 cases). Overall, the performance of the fragment screen for PARP14 MD3 was successful and allowed the following hit identification and validation process.

2.2.2. Identification and validation of PARP14 MD3 fragment hits

The 622 data sets were processed and analysed using the *XChemExplorer*, a data management and workflow tool developed for batch processing of multiple crystallographic data sets in particular in the context of structure-based ligand discovery projects such as fragment screenings²⁰³. *XChemExplorer* does not automate the structure determination process but provides a graphical interface that incorporates different software packages to process, solve and refine protein–ligand structures, thereby enabling their large-scale simultaneous analysis. In a first step, the data collections were reviewed and the log files generated from selected auto-processing pipelines. The difference map pipeline *DIMPLE* was then used for map calculation and initial refinement using a reference model of PARP14 MD3 to solve the structures by molecular replacement. To prepare the PARP14 MD3 *apo* reference model, a data set collected in the solvent tolerance test was processed. Next, *AceDRG*²⁰⁴ was used to generate the coordinates and geometry restraints of the ligands. Then, the hit identification process was performed using the *PanDDA* (Pan-Dataset Density Analysis) software recently developed for applications such as fragment screenings²⁰⁵. *PanDDA* allows reliable identification of binding events by calculating a so called ‘event map’ which is defined as $[\text{event map}] = [\text{dataset map}] - \text{BDC} \times [\text{ground state map}]$. The ground state map, representing the crystal without bound ligand, was obtained by averaging 200 independently measured crystals. Since fragment binding events are rare and only occur in a subfraction of the whole crystal, their signal contribution to the ground-state map is negligible. Using an averaged, instead of a single, ground-state map for subtraction further reduces experimental noise in the ground-state estimate and the event map. Multiplication of the ground-state map with the BDC, the ‘Background Density Correction’ factor, approximates the ground-state, i.e. background, density in fragment-bound crystals which is superimposed to the density of the ligand-bound state. Thus, the subtraction of the background density from the data sets identifies only regions that are unique to each data set, including the density resulting from a bound fragment²⁰⁵. 36 datasets were identified by *PanDDA* (5.8 % of the datasets) with statistically significant

difference signal, of which into 25 the corresponding ligand could be modelled with high confidence. All resulting models were exported and validated after one refinement step. 17 models in which the fragment was also partially or even well resolved in the crystallographic map, which may be an indirect measure of the binding strength of the fragments, were finally refined using *REFMAC* and *Coot* [Table 1].

Closer analysis of the fragment-bound structures showed that all fragments identified bound around the same site of the ADPR binding pocket and no other allosteric binding sites were identified [Figure 42 and S18]. In order to analyse the position of the fragments relative to the natural ligand, ADPR, the PARP14 MD3 fragment structures were superimposed with the PARP14 MD3 ADPR-bound structure (PDB ID 4ABK)⁶⁶. Prior comparison of the *apo* and ADPR-bound conformation of the macrodomain showed that ADPR binding induces only a very minor movement in the loop engaging with the pyrophosphates and the terminal ribose which allowed this analysis. All fragments bound around the adenine site, thus none were able to enter the diphosphate-terminal ribose site. The surface representations of the *apo* and ADPR-bound structure showed that this region is rather covered by this loop that needs to be pushed aside for ADPR binding. In contrast, the adenine-proximal ribose site is very accessible and open to the solvent phase in both conformations. It could be considered that either the screened fragments were too weak to induce this loop shift or the flexibility of the loop required is not provided in the crystal to this extent. Furthermore, it could be seen that the majority of the fragments bound rather close or almost in the crystal contact and only very few fragments were found deeper in the binding pocket, the latter being summarised in Figure 42. Interestingly, almost all fragments contained a saturated or aromatic (hetero) cycle that targeted Phe1371 in the macrodomain. In particular, all of the fragments identified in close proximity to the crystal contact showed only this interaction with the macrodomain and it may be possible that Phe1371 helps to hold the fragments by the hydrophobic interaction, i.e. π - π stacking, when passing through the crystal. Superimposition of the structures also showed the flexibility of Phe1371 to engage in particular with the aromatic ring system of the compounds [Figure 43A]. It is furthermore interesting that the recently published PARG inhibitor PDD00017273¹³¹ occupies the same site of the ADPR

binding pocket with roughly the same relative position to the ADPR. PDD00017273 moreover targets like the PARP14 MD3 fragments the corresponding phenylalanine residue (Phe902 in PARG) with its quinazolinedione core motif [Figure S19].

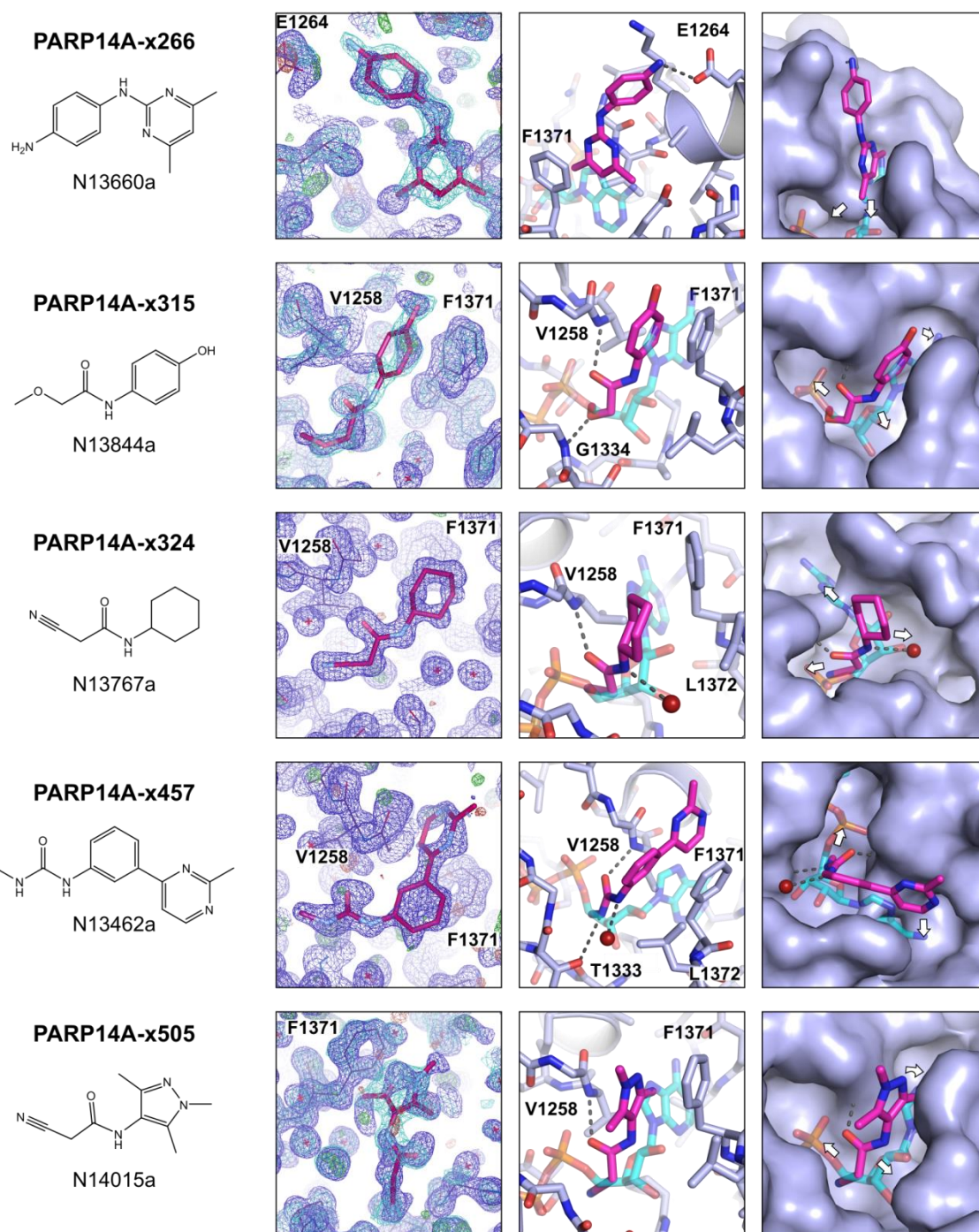


Figure 42 [continued on next page].

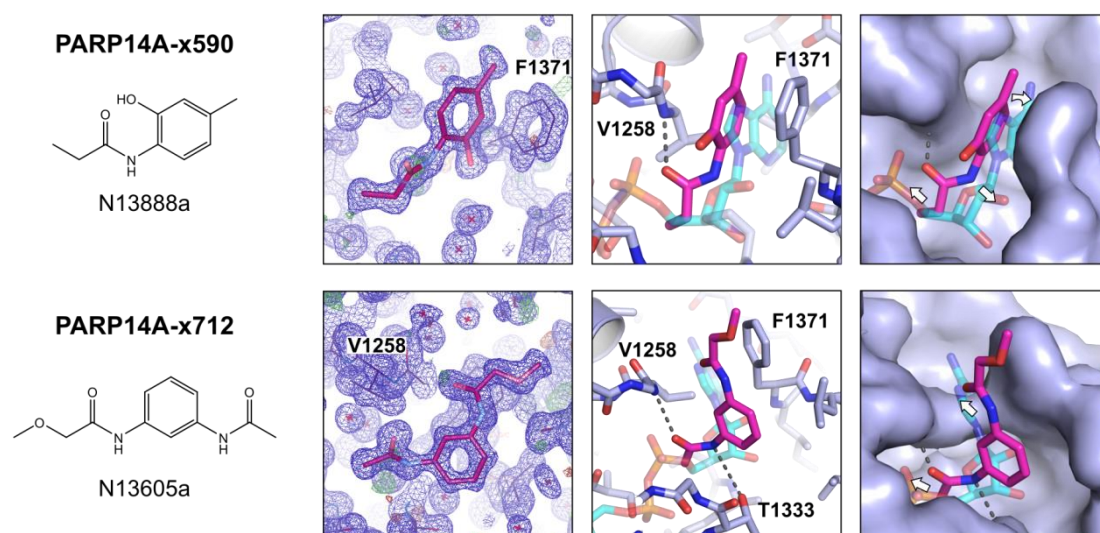


Figure 42. Fragment hits obtained in PARP14 MD3 fragment screen by X-ray crystallography. The molecular structure of the fragment identified in the corresponding crystal is shown on the left. The left images from the panel show the fragments in the crystallographic model. The $2F_o - F_c$ electron density map contoured at 1.0σ is shown in blue. For weaker fragments, the calculated *PanDDA* event map countoured at 2.0σ is additionally shown in cyan. The $F_o - F_c$ electron density map, green/red, is contoured at $\pm 3.0\sigma$. The images in the middle of the panel show the binding site of the fragments in PARP14 MD3 which residues are shown as atom-coloured sticks. Hydrogen bonds between ligand and protein or structural waters are indicated by dotted lines. As reference for the position of the fragment relative to the ADPR in the PARP14 MD3 binding site, PARP14 MD3 in complex with ADPR (PDB ID 4ABK)⁶⁶ was superimposed with the corresponding fragment structure and the ADPR shown as atom-coloured sticks in low transparency in the model. The images on the right show the surface representation of PARP14 MD3 in complex with the corresponding fragment. White arrows indicate in which directions the fragments could be expanded into the binding site.

However, a few fragments showed additional interactions with PARP14 MD3. The fragment in x266 was only partially resolved in the crystallographic map, however, was well defined in the *PanDDA* event map. Although being located rather on the protein surface, it seemed to target apart from Phe1371 the side chain of Glu1264 with its primary amine over an electrostatic interaction. Only in x266, Glu1264 was found in this conformation which supports this conclusion. Furthermore, all of the other fragments found deeper in the adenine binding site formed a hydrogen bond to the backbone amide of Val1258 via their amide or urea group. Only in x315, the angle of this bond was not ideal ($\angle (\text{C}=\text{O} \cdots \text{N}) \sim 113^\circ$) which would probably allow a favourable interaction. This might be due to the additional hydrogen bond interaction of the fragment with its ether oxygen to Gly1334 which restrains together with Phe1371 the fragment in its orientation. This might also explain why the x590 fragment, which is very similar to the fragment in x315 however lacks the ether group, was far better resolved in the crystallographic

map and therefore potentially bound stronger to the site. The fragments present in x324 and x505 are also both extended on the amide however by a nitrile group that does not seem to specifically target any residue in the binding site. The x505 fragment is less resolved in the crystallographic map (although being well defined in the *PanDDA* event map) than the x324 fragment. The latter seems to be held stronger in the binding site over additional van der Waals interactions between Leu1372 and its cyclohexane as well as an interaction with a structural water. The x457 fragment is the only fragment from these series with a urea group instead of an amide. This group seems to interact with Val1258, a structural water (the same that engages with the x324 fragment) as well as additionally with Thr1333. The phenylpyrimidine is held by van der Waals interactions with Leu1372 and in particular Phe1371 in the binding site. Due to the interactions of the polar urea group, the phenyl cannot be orientated for ideal π - π stacking with Phe1371 which could increase its binding strength and resolution in the crystallographic map. The fragment present in x712 which is well resolved in the electron density map was found rather close to the crystal contact and interacts like the x457 fragment with Thr1333. Its orientation in the binding site is interesting since although it contains the same structural motif as the x315 fragment (only with an additional amide group in *meta* position of the benzene) it is not binding in the same mode like the x315 fragment. The x712 fragment interacts with Val1258 via its amide instead of its methoxyacetamide although the latter would allow an additional hydrogen bond with the ether oxygen and Gly1334 and a better orientation of the benzene ring for Phe1371 targeting. Thus, the hydrogen bond interaction with Thr1333 together with the binding being less restrained and entropically favoured may result in this binding mode. The common structural core of these fragments and their targeting of the adenine-ribose binding site of PARP14 MD3 is summarised in Figure 43.

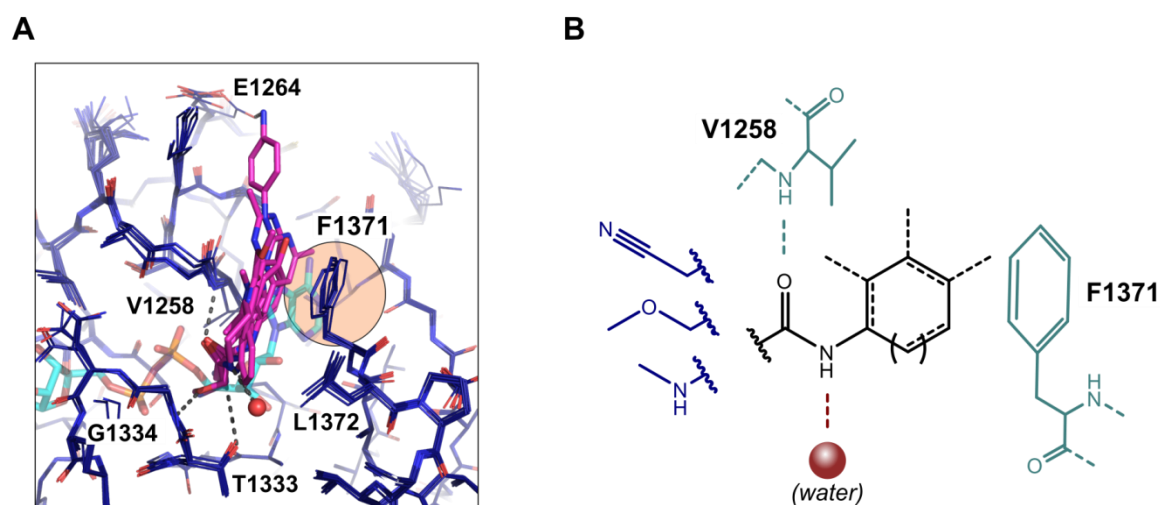


Figure 43. PARP14 MD3 targeting of the identified fragments. (A) Almost all fragments identified target Phe1371 with a saturated or aromatic (hetero) cycle. Several fragments furthermore show interaction with the amide in the backbone of Val1258 and some additionally with a structural water. The fragment bound in the crystal x266 interacts with E1264, the fragment bound in x315 with G1334. The PARP14 MD3 structures with the fragments shown in Figure 42 were aligned and superimposed to the PARP14 MD3-ADPR bound structure (PDB ID 4ABK)⁶⁶ to indicate the position of the fragments in the PARP14 MD3 binding site relative to ADPR. The residues of PARP14 MD3 are displayed as atom-coloured lines, the fragments and the ADPR as atom-coloured sticks. Hydrogen bonds are indicated by dotted lines. (B) Schematic representation of the interactions to PARP14 MD3 shown by at least two fragments. The common structural core is shown in black, examples for amide extensions in blue, the structural water in red and the main targeted PARP14 MD3 residues in green. Dotted lines around the ring system indicate its expansion in *meta*, *ortho* and/or *para* position.

Table 1. Data collection and refinement table for PARP14 MD3 fragment screen structures shown in Figure 42. The PARP14 MD3 *apo* structure (PDB ID 4ABL) ⁶⁶ was used to build the reference model by molecular replacement. The fragment RSCC (real-space correlation coefficient) describes the correlation between the model and the observed density (recommended value > 0.7). The B-factor ratio measures the consistency of the fragment model with surrounding protein residues, and is calculated from the B-factors of the fragment atoms and the fragment surrounding side-chain atoms within 4 Å (recommended value ~1.0). r.m.s.d.: coordinate root-mean-squared deviation.

^a Values within parentheses refer to the highest resolution shell.

	x266	x315	x324	x457	x505	x590	x712
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
No. of molecules in the asymmetric unit	1	1	1	1	1	1	1
Unit cell dimensions							
<i>a</i> , <i>b</i> , <i>c</i> (Å)	34.1, 41.4, 111.0	34.3, 41.7, 111.0	34.4, 41.6, 111.4	34.3, 41.7, 111.4	34.6, 41.6, 110.7	34.4, 41.5, 110.8	34.3, 41.8, 110.5
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Data collection							
Resolution range (Å) ^a	29.05 - 1.11 (1.15 - 1.11)	29.2 - 1.17 (1.21 - 1.17)	29.27 - 1.15 (1.20 - 1.15)	55.68 - 1.08 (1.12 - 1.08)	29.31 - 1.05 (1.09 - 1.05)	29.22 - 1.05 (1.09 - 1.05)	29.17 - 1.14 (1.18 - 1.14)
Unique observations ^a	60296 (4325)	54221 (4925)	56159 (4465)	62265 (2975)	55612 (991)	60506 (1180)	57368 (4566)
Average multiplicity ^a	1.9 (1.3)	1.9 (1.6)	1.9 (1.6)	1.9 (1.2)	2.0 (1.5)	1.9 (1.3)	1.9 (1.4)
Completeness (%) ^a	95.7 (69.6)	99.0 (91.4)	97.2 (78.7)	89.7 (43.7)	73.9 (13.4)	80.9 (16.0)	97.3 (78.4)
<i>R</i> _{merge} ^a	0.02 (0.23)	0.02 (0.33)	0.02 (0.36)	0.02 (0.11)	0.02 (0.21)	0.03 (0.11)	0.02 (0.35)
Mean (<i>I</i>)/σ(<i>I</i>) ^a	18.8 (1.3)	17.2 (1.6)	20.3 (1.6)	23.2 (2.4)	28.3 (2.5)	25.0 (5.1)	21.3 (1.3)
Mean CC(1/2) ^a	1 (0.921)	1 (0.888)	0.999 (0.703)	0.999 (0.973)	0.999 (0.861)	0.996 (0.975)	1 (0.818)
Refinement							
<i>R</i> -value, <i>R</i> _{free} (%)	21.4, 23.3	19.9, 22.3	18.7, 20.9	20.1, 21.9	19.3, 20.6	18.5, 20.2	21.2, 23.6
r.m.s.d. from ideal bond length (Å)	0.012	0.011	0.011	0.010	0.010	0.010	0.011
r.m.s.d. from ideal bond angle (°)	1.49	1.29	1.14	1.17	1.16	1.15	1.20
Ramachandran plot (favoured/allowed) (%)	98.3 / 100	98.3 / 100	96.7 / 100	97.8 / 100	98.9 / 100	98.9 / 100	95.0 / 100
Fragments							
	N13660a	N13844a	N13767a	N13462a	N14015a	N13888a	N13605a
Occupancy	0.45	0.37	0.78	0.55	0.71	0.80	0.57
r.m.s.d.	0.16	0.60	0.16	0.19	0.11	0.09	0.16
RSCC	0.83	0.64	0.95	0.84	0.88	0.97	0.92
B-factor ratio	1.11	1.50	1.00	1.13	1.35	0.94	0.91

Finally, in an attempt to validate the binding of the fragments also in a biophysical assay, five fragments were selected based on diversity in the scaffold, their resolution in the crystallographic map and the observed interactions with PARP14 MD3. The fragments were purchased together with a few structurally related compounds that were thought to potentially allow an initial SAR [Table 2, 3]. The final selection of SAR compounds was very limited and determined by their commercially availability, acceptable delivery time and price. Most of the compounds that would have matched the core of the fragments with additional variations required synthesis and were therefore not considered. The identified fragments in x315, x457 and x590 were used as templates for the searches and an aromatic ring system for targeting Phe1371 as well as an amide or urea group were defined as essential structural features of the SAR compounds.

Table 2. Selected fragment hits for characterisation of PARP14 MD3 binding in biophysical assays.

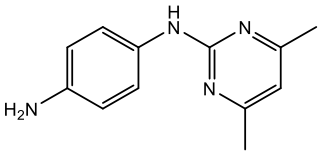
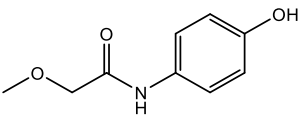
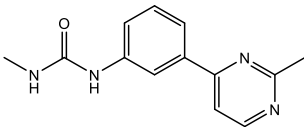
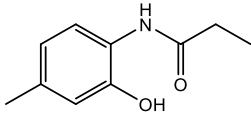
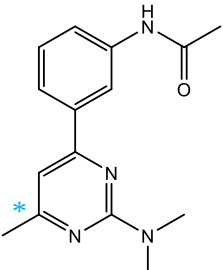
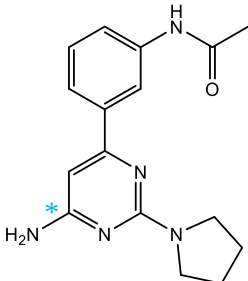
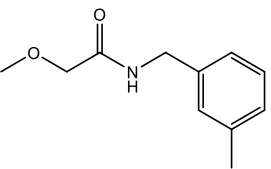
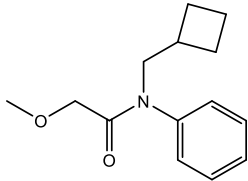
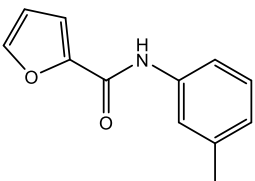
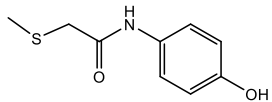
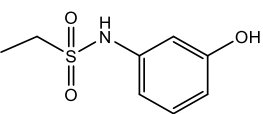
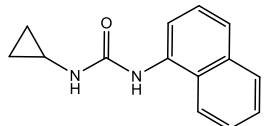
Selected fragment hits (crystal ID / compound ID of DSPL2 library)		
x266 / N13660a	x315 / N13844a	x457 / N13462a
		
x590 / 'N13888a' *	x712 / N13605a	comment * N13888a: (not affordable) 

Table 3. SAR compounds of fragment hits for characterisation of PARP14 MD3 binding in biophysical assays.

Compound structure	Comment	Compound structure	Comment
E63361a  • x457 fragment derivative • variation of head group to target Asp1235 • not available as urea analogue and with amide at *, thus requires flip of pyrimidine		E63362a  • x457 fragment derivative • variation of head group to explore space in binding pocket • not available as urea analogue and with pyrrolidine group at *, thus requires flip of pyrimidine	
E63363a  • x315 fragment derivative • with CH₂ spacer between aryl and amide group to allow more flexibility and better Phe1371 targeting • only available without hydroxyl group		E63364a  • x315 fragment derivative • amide variation with small ring system to explore space • not available without methoxy and with (preferred) cyclopropane group	
E63365a  • x315 fragment derivative • furan variation of the ether extension of the amide; more rigid; contrast to E63363a • only available without hydroxyl group		E63366a  • x315 fragment derivative • isostere replacement of ether	
E63367a  • x590 fragment derivative • bioisostere replacement of the amide with sulfonamide • hydroxyl group not in <i>ortho</i> position available		E63368a  • x457 fragment derivative • variation of the benzene for better Phe1371 targeting • cyclopropane extension of urea group to explore space	

All fragments were tested up to a concentration of 1 mM in macrodomain AlphaScreen dose-response assays for their ability to displace the peptide from PARP14 MD3. Due to the high compound concentrations applied, all fragments were counterscreened at the corresponding concentration. Subtraction of the counterscreen activity from the PARP14 MD3 AlphaScreen assay activity revealed that none of the fragments was able to displace the AlphaScreen peptide from the macrodomain. It is very likely that the binding of the fragments to PARP14 MD3 is far too weak to measure any activity on the macrodomain in competition experiments such as the AlphaScreen assay. Therefore, the fragments could potentially be validated by directly measuring their binding to PARP14 MD3 using sensitive techniques, for instance SPR, that allow determining K_D values in the high micromolar range. Furthermore, co-crystallisation of the SAR fragments to analyse their binding mode and occupancy in the crystal as a measure of binding strength may provide the starting points for future PARP14 MD3 inhibitor development projects. Since all fragments are poised, it should be possible together with the structural information to rather rapidly build up a comprehensive initial SAR library, potentially leading to compounds with higher binding affinities. Synthesis of analogues by amide couplings or by click reactions with the nitrile containing fragments are possible strategies to explore the sites of the binding pocket as indicated in Figure 42.

Conclusions of the PARP14 macrodomain hit discovery approaches

In summary, two different approaches for discovery of small molecules with new chemical scaffolds targeting PARP14 macrodomains were evaluated, i.e. traditional *in vitro* screening of a medium-sized compound library as well as structure-based fragment screening using X-ray crystallography. Both approaches have their merits and drawbacks as discussed previously and their success generating new hit matter depends among other factors certainly on the target and its druggability. This is in particular true for a crystallographic fragment screening approach for which a prerequisite is not only the reliable crystallisability of the target but also that the binding pocket of interest is accessible for the ligand by diffusion through solvent channels within the crystal lattice. Are these sites involved in crystal packing or otherwise obscured, then this approach is not feasible for hit discovery¹⁹⁸. It is speculative that only fragments could be obtained that target the adenine-ribose site and not the pyrophosphate-ribose site because of the coverage of the latter by a loop region of the protein. In *in vitro* screenings, the ligand binding in solution provides freedom for potential structural changes in the protein and thus other types of hits can be discovered such as the PARP14 MD2 inhibitor GeA-69. It is unlikely that compounds of this molecular size that bind deep within the protein, and therefore necessarily need to induce a structural change for binding can be identified in fragment screenings by X-ray crystallography. An alternative are fragment screenings by NMR, which the size of the macrodomains of ~25 kDa would allow, however, the final decision for the kind of drug discovery approach pursued is after all also determined by the availability of the facility and resources required as well as the expertise in performing and processing of the amounts of data generated.

In either case, in the *in vitro* screening as well as in the crystallographic fragment screening approach, the following hit validation process is a crucial but challenging part. It is not possible to avoid false positives hits and the hit validation is essential for filtering these hits out and confirming true binders. Due to the inherent characteristic of structure-based hit discovery, it is

rather easy to select interesting and true hits based on the structural information. Fragments binding in crystal contacts identified by hit identification methods such as *PanDDA* can be quickly excluded from further hit validation. However, due to the weak binding of the fragments, their characterisation and confirmation of binding in solution by biophysical assays is challenging. The crystallographic fragment screen was performed at a high concentration of ~50 mM which drives the binding of the fragments to the protein even with affinities in the high micromolar range, however, with which *in vitro* biophysical assays such as the AlphaScreen cannot be performed due to assay interference. Sensitive biophysical methods are required to assess the binding strength of the fragments and it was therefore not unexpected that no activity of the fragments in the AlphaScreen competition assay at a concentration of 1 mM could be observed. In contrast, hits obtained from the medium-sized *in vitro* screen could be characterised by determining IC₅₀ values with the AlphaScreen since their initial screening concentration was set to 25 µM and compounds binding as weak as the fragments would not have been identified as hits. The validation process focussed therefore in particular to confirm specific binding of the compounds. Counterscreening allows excluding the majority of assay interfering compounds, however, calorimetric binding measurements are seen as standard for reliable confirmation of binding. This proved to be difficult for many hits due to their poor water solubility. This might not only be the reason for the observed inconsistency of the screening results but furthermore decreased the probability of the compounds to be co-crystallisable with targets that itself do not readily crystallise like PARP14 MD2 to alternatively obtain structural information of the binding mode.

For PARP14 macrodomain targeting, both approaches could yield chemical hit matter yet with its limitations. From *in vitro* screening of in total ~50,000 small molecules from in-house and the academic compound library provided by Novartis, only one hit compound, GeA-69, targeting PARP14 MD2 could be identified. Despite its good ligand efficiency, availability of structural information and considerable medicinal chemistry efforts, GeA-69 could not be improved in its potency. Considering its molecular size and the fact that it is completely surrounded by the

protein, which enables interactions to the protein with all of its sites, indicates that the hydrophobic and the few hydrogen bond interactions between GeA-69 and PARP14 MD2 are not ideally positioned which would result in the desired nanomolar affinities. A significant change of the scaffold might be required which development may not be trivial.

Compared to the *in vitro* screen, the crystallographic fragment screen was more successful regarding the hit rate (as expected), providing several related scaffolds that are amenable for the ADPR binding site. However, all fragments are very weak binders, based on the fact that they are not characterisable in the AlphaScreen competition assay, and it will probably require iterative rounds of analogue synthesis, biophysical characterisation and structure determinations to obtain suitable hit and later lead compounds. Numerous fragments could nonetheless be successfully optimised even to final drugs entering the clinical phase ²⁰⁶, demonstrating its feasibility. The development of a PARP14 MD3 inhibitor would be essential for preventing PARP14 recruitment to target sites via its macrodomains and thus to gain insights into PARP14 macrodomain biology, for which the results obtained in this work may be valuable starting points.

3. Exploring pharmacological targeting strategies for PARP14

Although different cellular functions have been proposed for PARP14 ranging from transcriptional regulation, immune response, cell migration to metabolic regulation and DNA repair, the majority of the findings show evidence of the value of therapeutic targeting of PARP14. Due to the lack of PARP14-specific inhibitors, almost no data are available on phenotypic effects of PARP14 inhibition. Only Peng *et al.*¹³⁹ has been able to show that a chemical knock-down of endogenous PARP14 activities using PARP14 catalytic domain inhibitors induced (similarly to a genetic PARP14 knock-down) an apoptotic phenotype in cancer cells. The plethora of different PARP14 functions that have been proposed in the literature makes the assay development for evaluating PARP14 inhibitors a challenging task. Many PARP14 functions are not well understood on a molecular level and it is unclear if or how some of these PARP14 functions are linked to each other. The aim of the final part of this project was therefore to explore pharmacological targeting strategies for PARP14 and to develop an assay system suitable for characterisation of PARP14 macrodomain inhibitors. For this, three different PARP14 functions that were most recently reported were selected for evaluation: the role of PARP14 in cell migration and adhesion shown by Vyas *et al.*¹⁰⁸, PARP14 involvement in the Warburg effect shown by Iansante *et al.*⁹⁸ and the contribution of PARP14 in DNA repair shown by Nicolae *et al.*¹⁰⁹. All of these studies provided good experimental evidence and used cellular model systems, addressing both cytoplasmic as well as nuclear functions of PARP14.

3.1. PARP14 knock-down and knock-out in HeLa and U-2 OS cell lines

The PARP14 studies were initially planned to be performed with genetic and chemical interference of the PARP14 function using siRNAs and the PARP14 macrodomain inhibitor GeA-69, respectively. Since GeA-69 showed activity on MD2 only and did not prevent full-length PARP14 recruitment to target sites without additional inhibition of MD3, the latter approach

could not be realised. PARP14 knock-down phenotypes were induced using siRNAs which were previously reported for performing the PARP14 DNA damage response ('siPARP14 #1' and 'siPARP14 #2')¹⁰⁹ and motility studies ('siPARP14 #3')¹⁰⁸.

HeLa and U-2 OS cell lines were considered for the following PARP14 studies since HeLa cells were used in the PARP14 motility and DNA damage assays and U-2 OS cells were used also in the latter^{108,109} and in the established PARP14 macrodomain laser micro-irradiation assays. Both cell lines are furthermore immortalised and well characterised, fast growing in non-demanding culture conditions, amenable for gene editing and provide a cancer background. Cell lines reported with very high PARP14 expression such as lymphoid cell lines, fibroblasts and hepatocellular carcinoma cell lines^{84,98,166} were not considered due to either their unavailability, low transfection efficiency or unsuitability for the envisaged assays.

U-2 OS cells showed higher endogenous PARP14 expression levels than HeLa cells, which was consistent with the reported RNA expression levels¹⁶⁶, and for both cell lines siRNA-induced PARP14 knock-down was most efficient 72 hours post transfection [Figure 44A, S21A]. The HeLa cell phenotypes resulting from the siRNA-induced PARP14 knock-down were analysed in more detail, focusing on changes in cell viability and morphology. In agreement with previous observations^{98,108}, the cell viability was significantly reduced upon PARP14 knock-down, in particular when induced by siPARP14 #1 and siPARP14 #3 [Figure 44B]. Interestingly, the PARP14 knock-down using siPARP14 #3 furthermore resulted in strong alterations in cell morphology compared to control cells or cells with PARP14 knock-down induced by siPARP14 #1 and #2 which showed a normal cell morphology [Figure 44B]. This phenotype induced by siPARP14 #3 showing elongated processes extending from the cell body was also described by Vyas *et al.*¹⁰⁸ who therefore concluded a PARP14 function in actin cytoskeletal regulation. However, it was striking that this phenotype could only be observed using siPARP14 #3 and furthermore not in U-2 OS cells [Figure S21B]. Analysis of the siPARP14 #3 nucleotide sequence using BLAST® revealed vitrin (*gene*: VIT) as a hit apart from PARP14, with 16 (out of 19) adjacent nucleotides of the siRNA matching the vitrin transcript. Vitrin is an extracellular matrix protein which is assumed to play a role in matrix assembly, cell adhesion, and

migration due to its two VWA (Willebrand A) domains^{207,208}. In contrast, siPARP14 #1 and siPARP14 #2 did not show vitrin as BLAST[®] hit. Moreover, based on the RNA expression profile provided by the Human Protein Atlas²⁰⁹, vitrin is expressed in HeLa cells (similar to PARP14 expression levels) but not in U-2 OS cells and it is very tempting to assume that therefore siPARP14 #3 did not result in a corresponding phenotype in U-2 OS cells. All this may indicate that this visual PARP14 knock-down phenotype induced by siPARP14 #3 and thus the proposed PARP14 function in actin cytoskeletal regulation analysed by Vyas *et al.*¹⁰⁸ potentially results from vitrin silencing and not from PARP14 targeting. In order to confirm this hypothesis, HeLa cells should be further analysed in cell adhesion and migration assays. However, establishing in particular the planned scratch migration assays, where monolayers of cells are required, proved very difficult due to the decrease of the cell viability resulting from PARP14 knock-down by siPARP14 #1.

Instead of screening several PARP14 targeting siRNAs for their suitability, PARP14 was knocked-out in HeLa and U-2 OS cells using the CRISPR-Cas9 system. Based on PARP14 KO mice which were reported to appear healthy and a CRISPR KO screen^{101,210}, PARP14 was assumed to be a non-essential gene, theoretically enabling this approach. The possibility of performing various experiments with stable PARP14 KO cell lines without the need of a transfection procedure prior to the assay were the final motivation to engineer these HeLa and U-2 OS cell lines.

Legend to Figure 44: PARP14 expression and knock-down in HeLa and U-2 OS cell lines.

(A) HeLa and U-2 OS cells were left either untreated (-) or were transfected with siMOCK or PARP14 targeting siRNAs and analysed for PARP14 expression 72 hours post transfection by Western blotting (loading: 45 µg of total cell lysate for HeLa cells, 35 µg of total cell lysate for U-2 OS cells). **(B)** Phenotypic effects of PARP14 siRNA knock-down. Wild-type HeLa cells were transfected either with 20 nM siMOCK or 20 nM siPARP14 #1, #2 or #3 and analysed 72 hours post transfection. PARP14 knock-down induced by siPARP14 #1 resulted in significant decrease in cell viability as shown by nuclei co-staining with Hoechst33342 (blue) and propidium iodide (red) of transfected cells (top right). Furthermore, in contrast to siPARP14 #1 and #2, siPARP14 #3 results in a phenotype with extended cellular protrusions as also previously reported¹⁰⁸. The black rectangles in the representative bright field images (top) indicate the section that was enlarged and shown below the corresponding image.

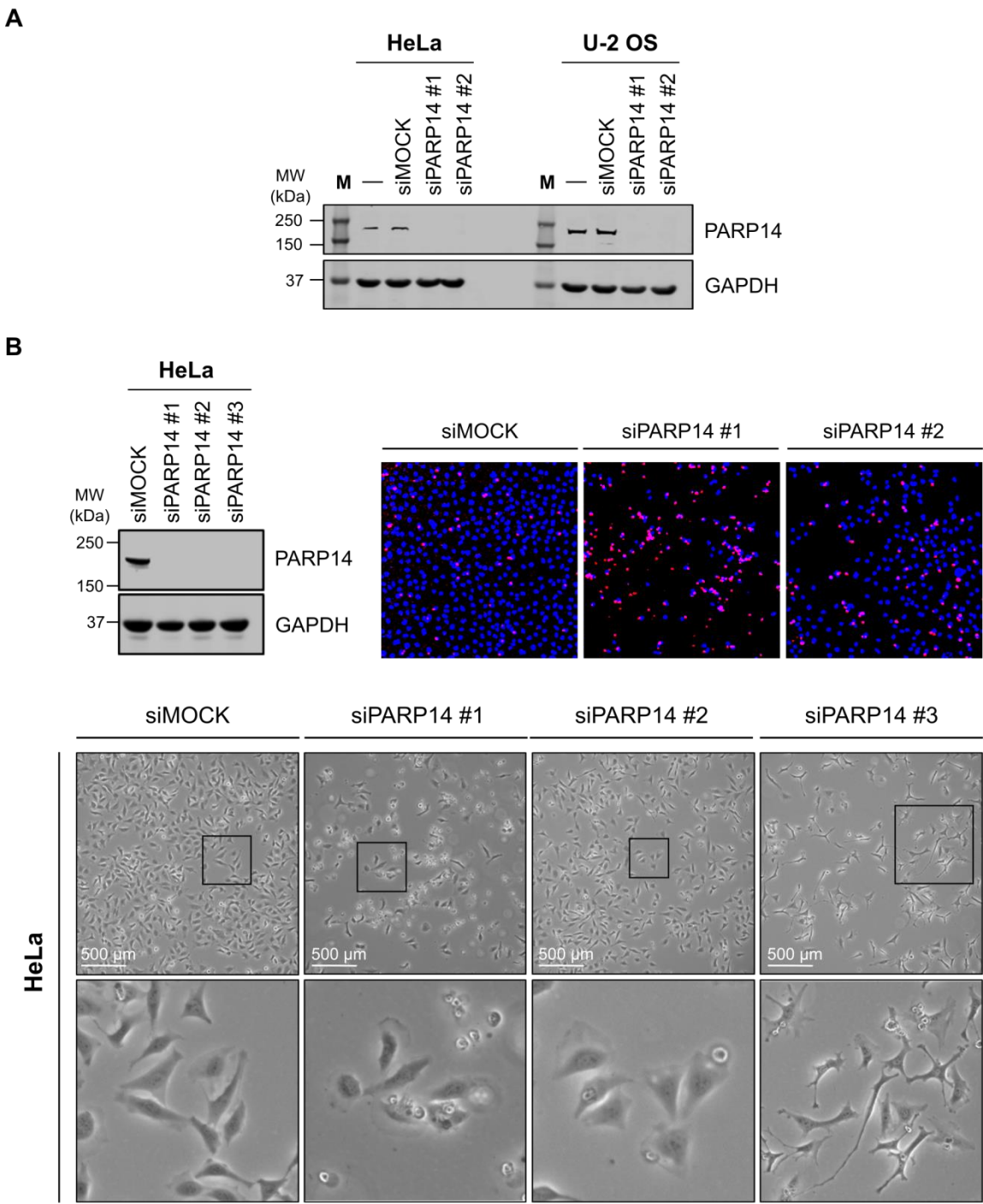


Figure 44. *PARP14 expression and knock-down in HeLa and U-2 OS cell lines.* For legend, see bottom of previous page.

The clustered, regularly interspaced, short palindromic repeat (CRISPR) technology using the RNA-guided nuclease ‘CRISPR-associated protein 9’ (Cas9) was derived from adaptable immune mechanisms used by bacteria and provides a tool for rapid and efficient targeted genome editing²¹¹. In contrast to ZFN and TALEN methods, which use protein-DNA interactions for targeting, the Cas9 nuclease is directed to genomic target sites by single-guide RNAs (sgRNAs)

where it mediates like other nucleases DNA double-strand breaks (DSB). These DSBs typically undergo one of two major pathways for DNA damage repair: nonhomologous end-joining (NHEJ) and homology-directed repair (HDR) (in presence of a repair template). NHEJ is error-prone and can lead to introduction of insertion/deletions (indels) of various lengths, which can result in frameshift mutations and premature stop codons within coding exons, thus generating gene knock-outs. Alternatively, DSBs are repaired by HDR which can be used to introduce specific mutations or to insert desired sequences through recombination of the target sites with an exogenously supplied DNA repair template^{211,212}. The principle and workflow for generating PARP14 KO HeLa and U-2 OS cell lines is outlined in Figure 45.

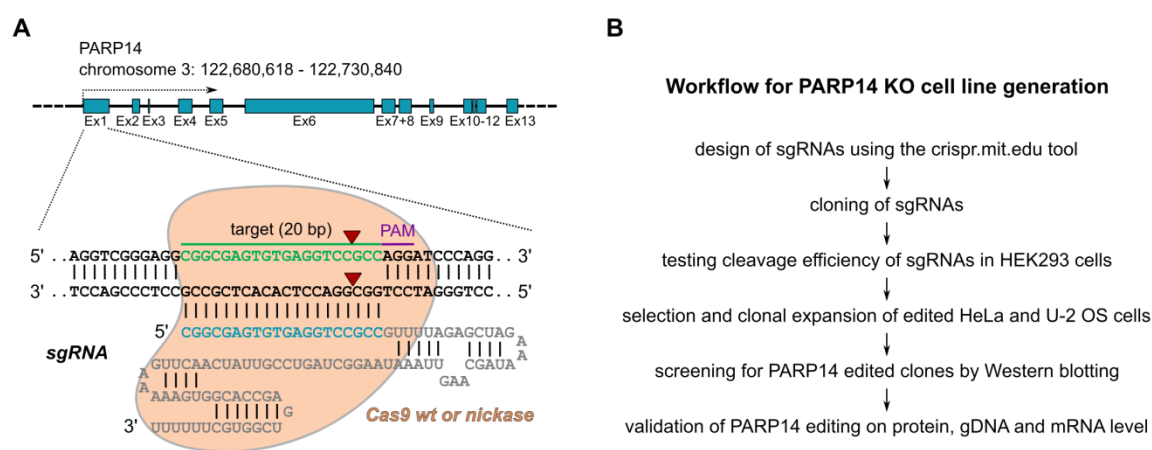


Figure 45. PARP14 KO cell line generation using the CRISPR-Cas9 system. (A) Principle of the PARP14 gene editing strategy (figure adapted from Ran *et al.*²¹²). SgRNAs were designed for targeting exon 1 of PARP14. PARP14 gene schematic (top) shows arrangement of 13 of 17 exons present in PARP14; for illustration, introns are shown in 10 times shorter size. As example, the blue nucleotide sequence cloned into the sgRNA pSpCas9(BB)-2A-Puro-PARP14-c002 targets the 20 nucleotide sequence in exon 1 shown in green which is located upstream of the requisite 5'-NGG protospacer adjacent motif (PAM). The sgRNA pairs with the target sequence and thus directs the Cas9 nuclease (co-expressed with the sgRNA vector system) to the genomic locus of PARP14. The Cas9 nuclease mediates double (Cas9 wild-type) or single (Cas9 nickase with D10A mutation) strand breaks ~3 bp upstream of the PAM (red triangle) resulting in PARP14 editing upon errors made by the DNA repair machinery. (B) Schematic workflow for HeLa and U-2 OS PARP14 KO cell line generation.

After designing and cloning of sgRNAs targeting the exon 1 of PARP14, their DNA cleavage efficiency was evaluated in HEK293 cells. PARP14 gene editing potential was confirmed for all sgRNAs of which the combination of the Cas9 nickase constructs c001 and c003 showed the highest cleavage efficiency [Figure 46A]. Following this functional validation of the sgRNAs, HeLa and U-2 OS cell lines were edited on their PARP14 gene sites. After optimisation of the

transfection, selection and clonal expansion procedure, 50 HeLa and 64 U-2 OS clones were derived from the corresponding parental cell lines and were screened for PARP14 KO editing by Western blot analysis [Figure 46B and S22].

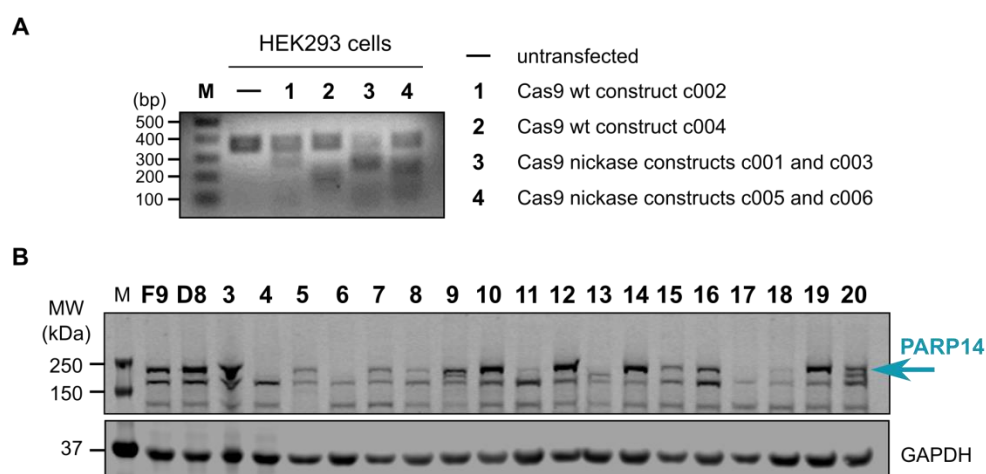


Figure 46. (A) Functional validation of PARP14 targeting sgRNAs by testing their PARP14 gene editing efficiency in HEK293 cells. Genomic DNA of HEK293 cells which were left either untransfected or were transfected with sgRNA constructs as indicated was extracted and PARP14 exon 1 was PCR-amplified. The resulting PCR products were re-annealed and analysed after T7 endonuclease I digestion. The sgRNA Cas9 nickase constructs c001 and c003 show the highest cleavage efficiency. **(B)** Example of screening for PARP14 edited HeLa clones by Western blotting (50 µg of total lysate were loaded and the blot was overexposed to analyse for PARP14 edits). Clones 3-20 were obtained from HeLa cells transfected with sgRNA Cas9 nickase constructs c001 and c003. Clone F9 and D8 result from cells transfected with GFP-targeting control sgRNAs. Clone 4, 6 and 17 are PARP14 KO clones shown by further validation.

Potential HeLa and U-2 OS PARP14 KO cell lines were then selected and validated by characterisation of their PARP14 edits on protein, gDNA and mRNA level [Figure 47]. Validation on the protein level excluded the U-2 OS clone 49 which probably expressed a truncated version of PARP14 [Figure 47A]. Next, gDNA analysis suggested the presence of two alleles encoding PARP14 in HeLa cells and three PARP14 alleles in the U-2 OS cells. Based on the reported karyotypes of these cell lines ^{213–215}, the observed PARP14 copy numbers were expected for U-2 OS cells but rather not for HeLa cells for which the presence of three chromosomes 3, yet with abnormalities, was described. Using CRISP-ID ²¹⁶, a web application tool developed to decode CRISPR-mediated indels by Sanger sequencing, HeLa and U-2 OS clones could be genotyped for PARP14 edits with high confidence, apart from U-2 OS cells with three differently edited alleles [Figure 47B, Table S4]. The HeLa clones showed various edits including insertions

and deletions of 8-22 bp, with two clones (21 and 50) showing the same gDNA edit on both alleles. The U-2 OS clone 30 showed different edits, one being an insertion of 240 bp which could be confirmed (as all other edits) by gel electrophoresis. Due to this significant and rather atypical edit, this clone was excluded from further analysis. The gDNA (as well as the later mRNA analysis) of the U-2 OS clone 49 revealed an in frame deletion of 9 bp. Since the resulting PARP14 product would be 202 kDa in size, the assumed PARP14 truncation previously detected by Western blotting cannot be rationalised based on this edit. Finally, all other PARP14 edited HeLa and U-2 OS clones were validated on the mRNA level [Figure 47C, Table S5]. The exons 1-4 of the PARP14 mRNA (~600 bp) could be PCR-amplified from all clones which indicated that PARP14 mRNA was present in all of them and did not undergo non-sense mediated decay due to any gDNA edit. Furthermore due to the use of different exon-crossing primers, it could be concluded that no intron sequences were present in this part of the PARP14 mRNAs, indicating that probably all PARP14 mRNA transcripts were correctly spliced. This required analysis of the mRNA products by Sanger sequencing. Interestingly, in most cases the mRNA edit of one allele was different to the gDNA edit, probably indicating different splicing events of the transcripts depending on the gDNA edit. Only the HeLa clones 21 and 50 showed the same gDNA and mRNA edit in exon 1, i.e. deletions of 22 and 16 bp, respectively. These HeLa clones as well as the U-2 OS PARP14 KO clones 23 and 65 were selected for the following PARP14 studies [Figure 47D]. For all of these clones, the gDNA and subsequent mRNA edit led to formation of a premature stop codon and thus to a disruption of the reading frame, resulting in no PARP14 protein products longer than 100 amino acids. In parallel to the PARP14 KO cell line generation, clonal wild-type PARP14 cell lines were prepared by transfecting the parental HeLa and U-2 OS cell lines with GFP-targeting instead of PARP14 targeting sgRNAs. These clonal wild-type cell lines (HeLa clone D8, U-2 OS clone E9) were used as control cell lines for the following PARP14 studies, which were considered to be more similar to the also clonal PARP14 KO cell lines than the parental, heterogeneous, cell lines due to their transfection, selection and clonal expansion background.

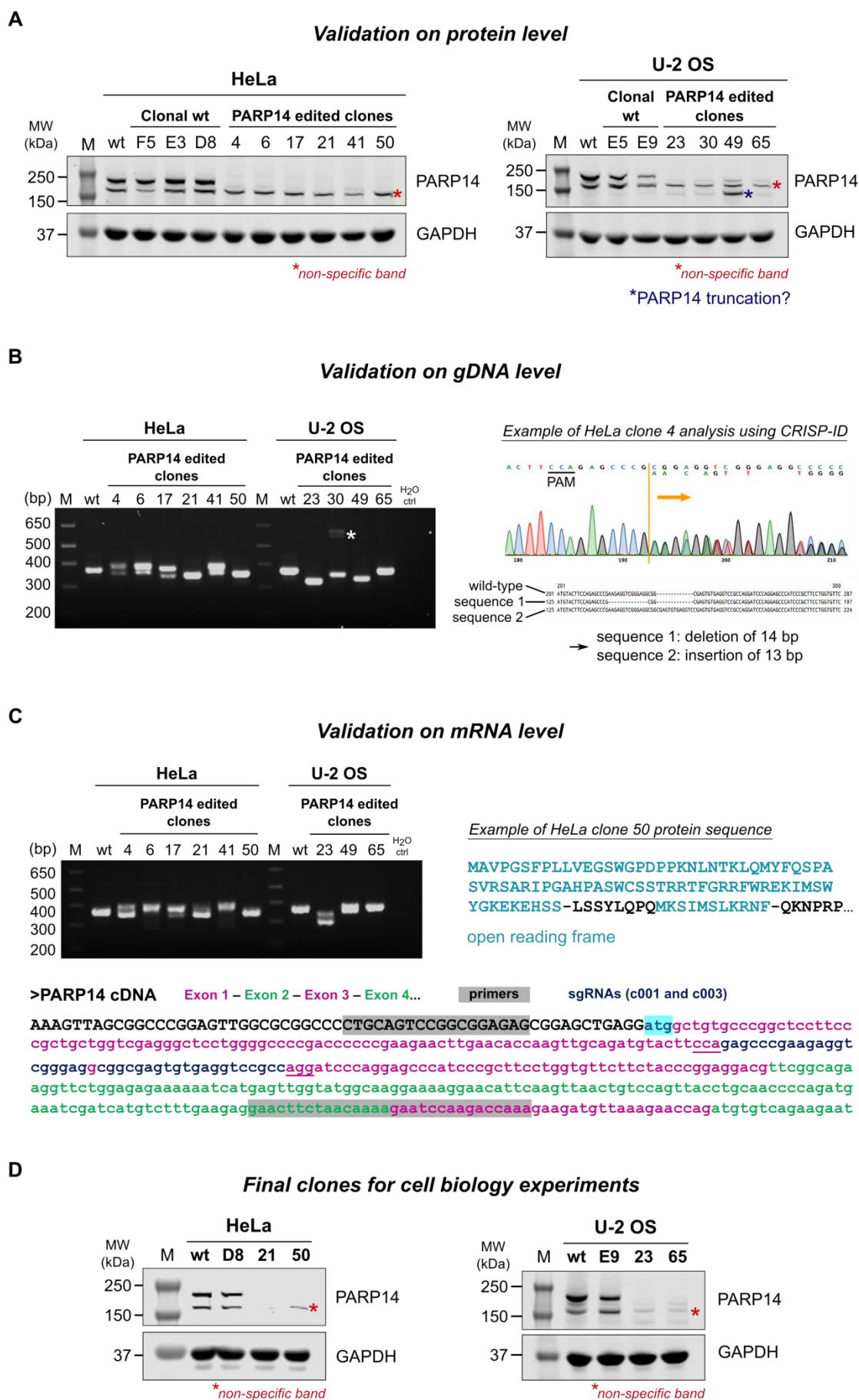


Figure 47. HeLa and U-2 OS PARP14 KO cell line validation on protein, gDNA and mRNA level. For legend, see top of next page.

Legend to Figure 47: HeLa and U-2 OS PARP14 KO cell line validation on protein, gDNA and mRNA level.

(A) Validation on protein level by analysing from screening selected clones by Western blotting. **(B)** Validation on gDNA level. For this, the exon 1 of PARP14 was PCR-amplified from the extracted genomic DNA and analysed by Sanger sequencing. Sequencing data obtained from clones with differently edited alleles were analysed using the CRISP-ID online software tool ²¹⁶ which extracts individual sequences from the sequencing data and detects indel size and location of the CRISPR-Cas9 targeted region. An example is shown for the HeLa clone 4. The yellow arrow in the chromatogram marks the start of the two different overlaying sequence traces resulting from different gDNA edits. Analysis results were confirmed by gel electrophoresis (left). The white star marks the 240 bp insert identified in the exon of the U-2 OS clone 30. **(C)** Validation on mRNA level. PARP14 cDNA was PCR-amplified with different primer pairs covering up to 600 bp (exon 1-4) and analysed by Sanger sequencing. Schematic below the agarose gel shows the design of one primer pair used for amplification and sequencing. All reverse primers were designed as exon-crossing primers so that in case the intron sequence is still present, i.e. the mRNA is not correctly processed, the primers cannot anneal at the PCR annealing temperature of 62 °C. The start codon of the PARP14 cDNA is highlighted in cyan; the PAM sequence before the sgRNAs is underlined. **(D)** Final selected clones used in the following PARP14 studies are shown.

Prior to the PARP14 studies, the validated HeLa and U-2 OS cell lines were characterised regarding their proliferation behaviour and visual phenotype compared to the clonal and parental wild-type cell lines [Figure 48]. HeLa wild-type and PARP14 KO cell lines showed no difference in growth behaviour with a doubling time of 24 h. Also the parental wild-type U-2 OS cell line showed the same proliferation rate as the PARP14 KO cell lines with a doubling time of 35 h. Only the doubling time of the clonal wild-type U-2 OS cell line E9 was higher with 50 h [Figure 48A]. Furthermore, brightfield and confocal images visualising the cytoskeleton of the HeLa and U-2 OS wild-type and PARP14 KO cell lines did not reveal any visual difference in the phenotype within each cell type. All generated PARP14 KO cell lines did not show a phenotype that would suggest any defects in actin cytoskeletal regulation and thus in functions associated with it, i.e. cell adhesion and motility [Figure 48B]. This may be another indication that the reported actin cytoskeletal misregulation in HeLa cells upon PARP14 siRNA knock-down using siPARP14 #3 is based on off-target effects. To obtain functional data confirming this hypothesis, the PARP14 KO cell lines were characterised regarding their cell adhesion and migration behaviour.

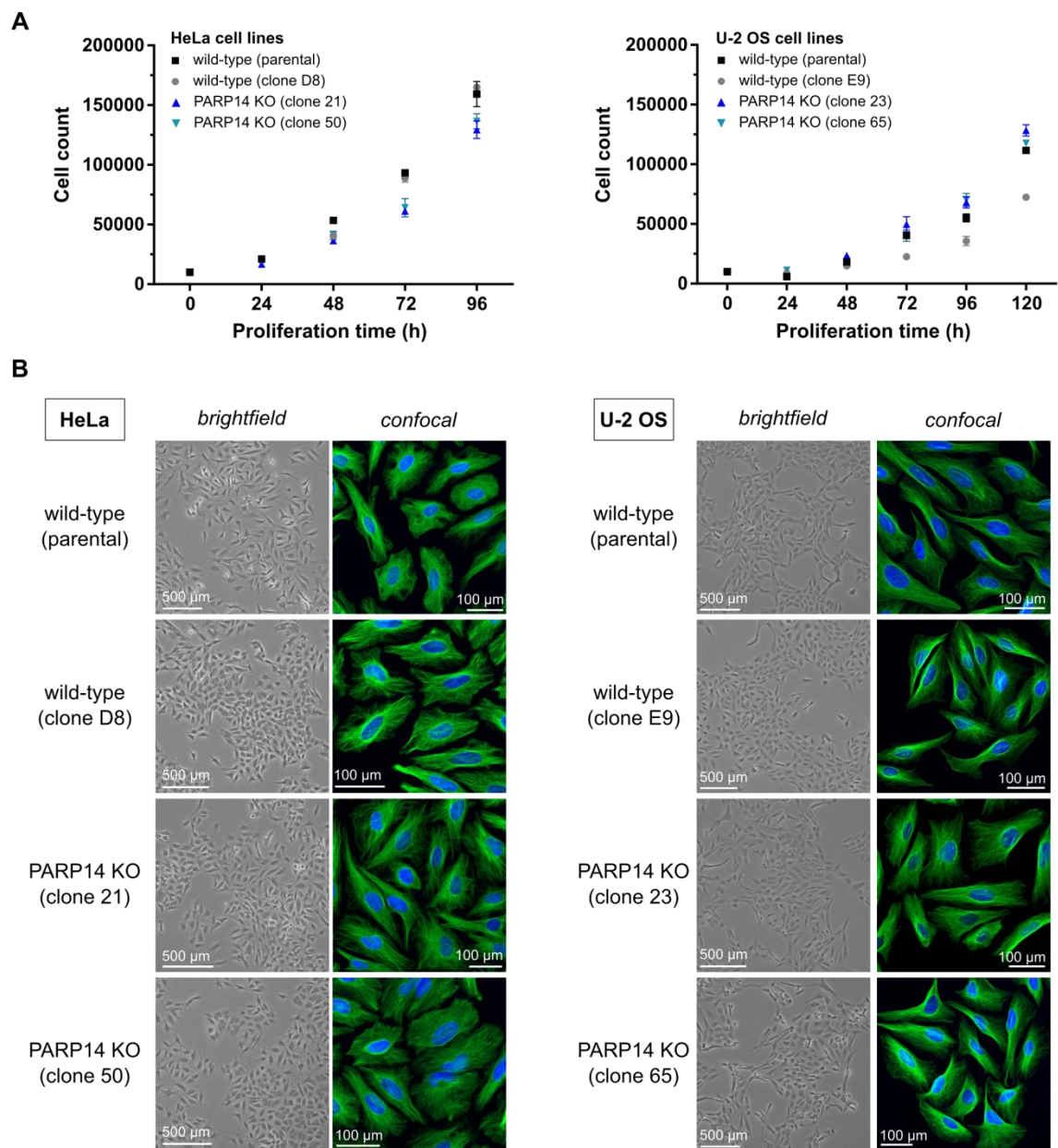


Figure 48. Proliferation behaviour and visual phenotype of HeLa and U-2 OS wild-type and PARP14 KO cell lines. (A) HeLa and U-2 OS cells were seeded at low density and proliferation was monitored by determining the cell count every 24 h (data points: mean $\pm$ SEM, $n = 3$). HeLa wild-type and PARP14 KO cell lines show no difference in growth behaviour with a doubling time of 24 h. The parental wild-type U-2 OS cell line shows the same proliferation rate as the PARP14 KO cell lines with a doubling time of 35 h. The clonal wild-type U-2 OS cell line shows a higher doubling time of 50 h. (B) Brightfield and confocal images of HeLa and U-2 OS wild-type and PARP14 KO cell lines do not reveal a visual difference in phenotype within each cell type. For confocal imaging, cells were fixed with 4 % formaldehyde and the cytoskeleton was visualised with α -tubulin staining (green). The nucleus was counterstained with Hoechst33342 (blue).

3.2. Role of PARP14 in cell adhesion and migration processes

Analysis of the PARP14 knock-down phenotype in HeLa cells performed by Vyas *et al.*¹⁰⁸ suggested PARP14 as a focal adhesion complex protein. Real-time imaging showed that migrating PARP14 knock-down cells appeared unable to retract membrane protrusion efficiently, resulting in elongated dendritic-like extensions and an atypical cell morphology. A process crucial for cell migration is the dynamic assembly and disassembly of focal adhesion complexes which mature from small focal complexes to large multi-protein complexes and function as anchor for the actin cytoskeleton to the substrate, the extracellular matrix^{217,218}. Balancing the formation and turnover of the focal adhesion complexes regulates the adhesive strength of the cell onto the substrate and allows the cell to pull itself and to move in the desired direction by then again releasing from the substrate. PARP14 was shown to localise to focal adhesions and seemed to be required for efficient focal adhesion turnover. The resulting increased adhesive force to the substrate due to PARP14 knock-down prevents proper retraction of these cellular protrusions, thereby impairing cell motility¹⁰⁸. Since the macroPARP family member PARP9 was also reported to promote migration in B-cell lymphomas⁸³, the HeLa PARP14 KO cell lines were characterised regarding their adhesive force to different substrates as well as their migration potential.

3.2.1. Testing PARP14 importance for cell adhesion

The adhesive strength of control and PARP14 knock-down HeLa cells has been examined by Vyas *et al.*¹⁰⁸ by performing a centrifugation-based assay (comparing the number of cells remaining adhered to the substrate after constant application of centrifugal force), a cell-spreading assay (measuring the area of cell spreading per unit time) and a trypsin-based detachment assay (measuring the time required to detach the cells from the substrate). The latter was established for characterisation of the HeLa PARP14 KO cell lines. The cell adhesion assay workflow is shown

in Figure 49A. In order to quantify the number of cells remaining attached to the substrate after a certain time of the trypsin treatment as accurately as possible, the trypsination reaction was stopped using the protease inhibitor PMSF and counts of adherent cells after treatment were determined by automated image analysis based quantitation of Hoechst33342 stained cells [Figure S24]. In several independent assays performed with standard TC treated plates, no significant difference in the adhesive strength between wild-type and PARP14 KO cell lines could be observed. Only the clonal wild-type cell line showed a slightly lower adhesive strength compared to the other cell lines [Figure 49C, S28B]. Therefore, to imitate the extracellular matrix and perform the assays under similar conditions as Vyas *et al.*¹⁰⁸, trypsin-based detachment assays were additionally performed on fibronectin and laminin coated plates [Figure 49D, E, S28B]. Compared to standard TC coated plates, HeLa PARP14 KO and wild-type cell lines adhered far stronger to the fibronectin coated plates, that required therefore higher trypsin concentrations for cell detachment, yet showed lower adherence to laminin coated plates. On the latter, no difference between the cell lines in their adhesion strength was observed. On fibronectin coated plates, the parental wild-type and PARP14 KO cell lines adhered similarly strong to the substrate. Only as previously observed, the clonal wild-type cell line D8 showed reduced adherence to the fibronectin coated plates compared to the other cell lines. Due to time constraints, other clonal wild-type HeLa cell lines could not be tested in this assay to verify if this observation was a cell line specific or indeed a PARP14 KO-specific effect. Since the parental wild-type cell line however consistently showed the same adhesive strength as the PARP14 KO cell lines and the visual atypical cell morphology resulting from inefficient retraction of cellular protrusions was not observed in PARP14 KO cells, this data suggests that the observed difference in the clonal cell lines is only a cell line-specific effect and PARP14 is not crucial as stated by Vyas *et al.*¹⁰⁸ for cell adhesion processes.

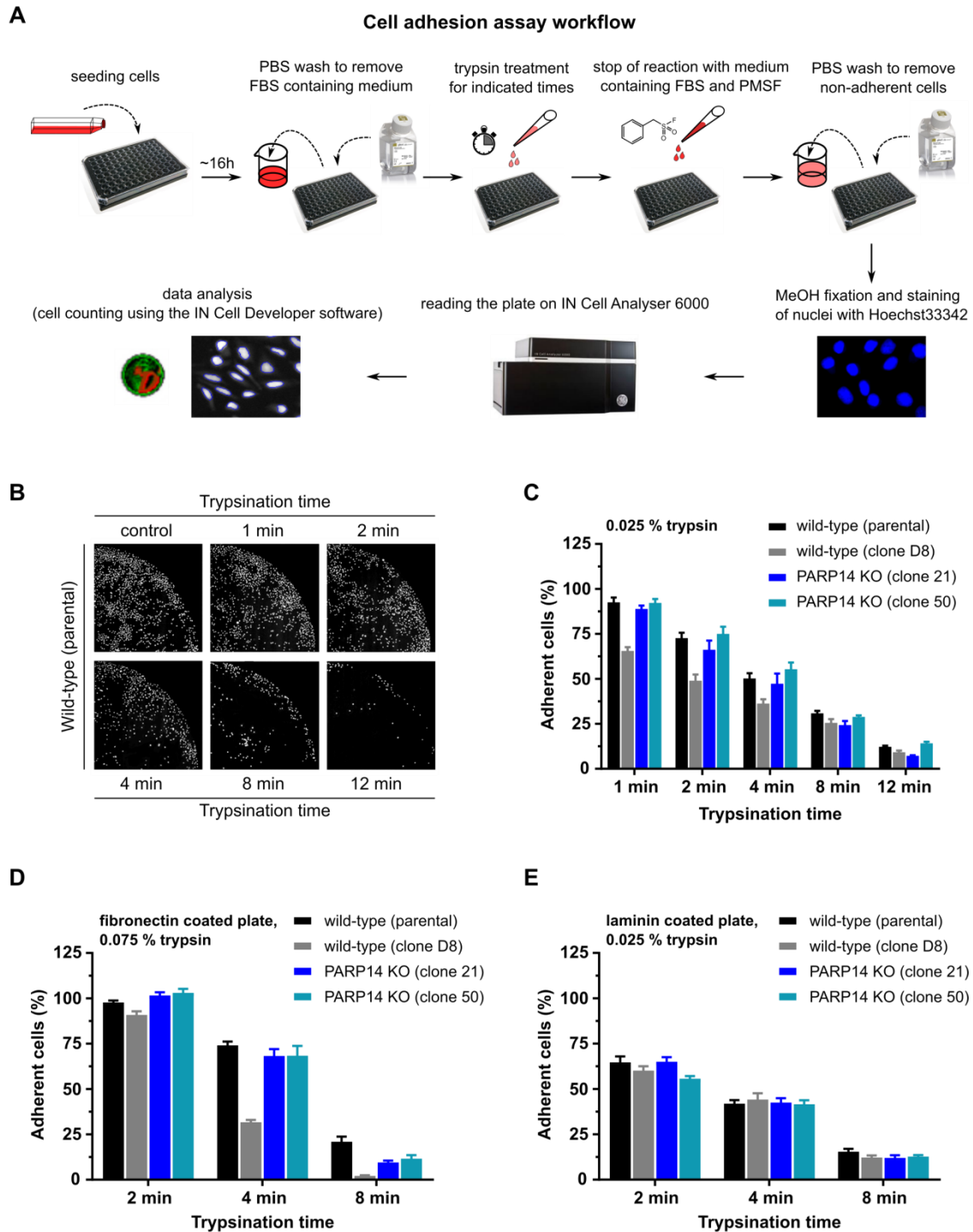


Figure 49. Testing the adhesive strength of wild-type and PARP14 KO HeLa cell lines in a trypsin-based detachment assay. (A) Workflow of the established cell adhesion assay. PMSF: phenylmethylsulfonyl fluoride. (B) Example images of a time series acquired after a cell adhesion assay performed with the wild-type (parental) HeLa cell line. Those images were further processed and cell counts were obtained using a customised protocol of the IN Cell Developer Toolbox software [Figure S24]. Control: trypsin-untreated cells. (C) PARP14 KO cell lines show similar adhesive strength as the parental wild-type HeLa cell line but are more adhesive than the clonal PARP14 wild-type cell line (D8). This cell adhesion assay was performed on a standard TC treated plate. (D and E) Cell adhesion assays performed with TC plates coated with different substrates as indicated. For all cell lines, cell adhesion was significantly increased on fibronectin coated plates and significantly decreased on laminin coated plates. On laminin coated plates, no difference in the adhesive strength of PARP14 wild-type and KO cell lines was observed. On fibronectin coated plates, PARP14 KO cell lines behave similarly to the wild-type (parental) cell line but are more adhesive to the substrate than the wild-type clonal cell line D8. For all graphs, the data was normalised to the corresponding trypsin-untreated controls, where data points are mean $\pm$ SEM, $n = 6$. Assay repeats are shown in Figure S28B.

3.2.2. Characterisation of the migratory potential of PARP14 KO cells

The migratory potential of the HeLa wild-type and PARP14 KO cell lines was assessed by establishing a scratch (wound-healing) migration assay. Scratch assays enable a kinetic assessment of the cell migration including recording of the phenotype of the moving cells ²¹⁹. In principle, a confluent monolayer of cells adherent to the plate is scratched ('wounded') by scraping off an area of cells, e.g. using a pipette tip. The closure of the scratch is microscopically monitored over time which allows conclusions about the motility of the cells ^{219,220}. Since scratch assays are also suitable for compound high-throughput screening ²²¹, a scratch assay was established that should allow both assessing the migratory potential of PARP14 KO cell lines compared to the control cell lines as well as potential screening of PARP14 (macrodomain) inhibitors. For this, the assay was set-up in a 96-well format using a 96 pin array to produce scratches of uniform shape and consistent placement within each well. Furthermore, the cells were stained with CellTracker™ Green CMFDA dye prior to the assay to enable monitoring of the closure of the scratch at regular time intervals (every 2 hours over 24 hours) using an automated fluorescence microscope. This procedure provided images of good resolution and of higher contrast than normal brightfield images which was greatly beneficial for the following data processing. The acquisition of these time-course series for ~100 wells required the development of an automated image analysis routine which finally allowed a reliable, rapid and consistent quantitation of the scratch areas at the respective time points to determine the migration rate of the cell lines [Figure S25, S26]. However, the scratch assay itself cannot distinguish the closure of the scratch area due to cell proliferation from closure through actual cell migration. Since the migration assays were performed over 24 hours, this factor had to be considered based on the previously determined doubling time of 24 hours of the HeLa cell lines. Therefore, cytosine arabinoside (Ara-C) was used as a proliferation control inhibitor in the scratch assay [Figure 50] ²²². For the HeLa PARP14 wild-type and KO cell lines an Ara-C IC₅₀ halting proliferation of 65-180 nM was determined [Figure 50C] and the performance of scratch assays in

presence of 5-20 μM Ara-C showed that Ara-C treatment did largely suppress cell proliferation over the duration of the assay, however, it neither affected cell migration behaviour nor cell viability as seen by counterstaining with propidium iodide [Figure 50A, B]. Of note is that a very minor increase of the cell count over the duration of the assay could nevertheless be observed, even after treatment of the cells with 20 μM Ara-C. This is likely due to the application of Ara-C to the cells just prior to the assay. In order to completely halt cell proliferation, either longer pre treatment of the cells with Ara-C or culturing the cells for 24-48 hours in medium with lower serum concentration prior to the assay could be considered. As further control for the following PARP14 migration studies and in respect to potential compound screening, the scratch assay was established with Cytochalasin D as migration control inhibitor which was previously reported to be suitable for this purpose²²¹. Treatment of the HeLa PARP14 wild-type and KO cell lines with 1 μM Cytochalasin D completely inhibited cell migration, yet it did not affect cell viability [Figure 51].

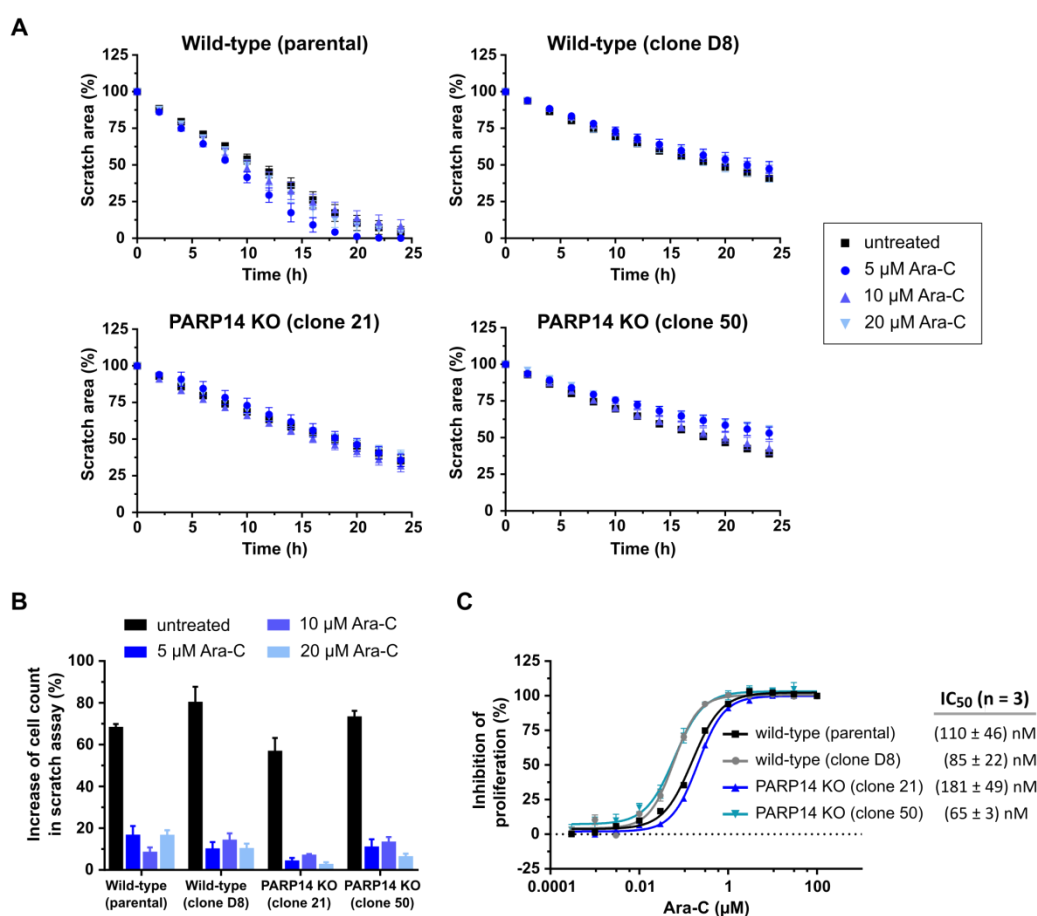


Figure 50. Establishing a scratch migration assay for wild-type and PARP14 KO HeLa cell lines using Ara-C as proliferation control inhibitor. For legend, see top of next page.

Legend to Figure 50: Establishing a scratch migration assay for wild-type and PARP14 KO HeLa cell lines using Ara-C as proliferation control inhibitor.

(A) Migration curves of HeLa cell lines in presence of Ara-C at three different concentrations are shown. For all cell lines, Ara-C treatment did not influence the migration behaviour. Data points are mean $\pm$ SEM, $n = 6$. (B) Determining the increase of the cell count over 24 h in the scratch assay shown in (A) confirmed efficiency of halting proliferation with Ara-C. Data points are mean $\pm$ SEM, $n = 6$. (C) IC_{50} values of Ara-C for HeLa cell lines from three independent experiments. A representative curve of one experiment is shown. Data points are mean $\pm$ SD, $n = 3$. Cells were treated with Ara-C at different concentrations (100 μ M to 300 pM) for 24 h (duration of the scratch assay) and then live stained with Hoechst33342 to count the nuclei for quantitation. For generating the graph, data was normalised with 0 % inhibition defined by the cell count of untreated wells and 100 % inhibition defined by the cell count of wells treated with 100 μ M Ara-C.

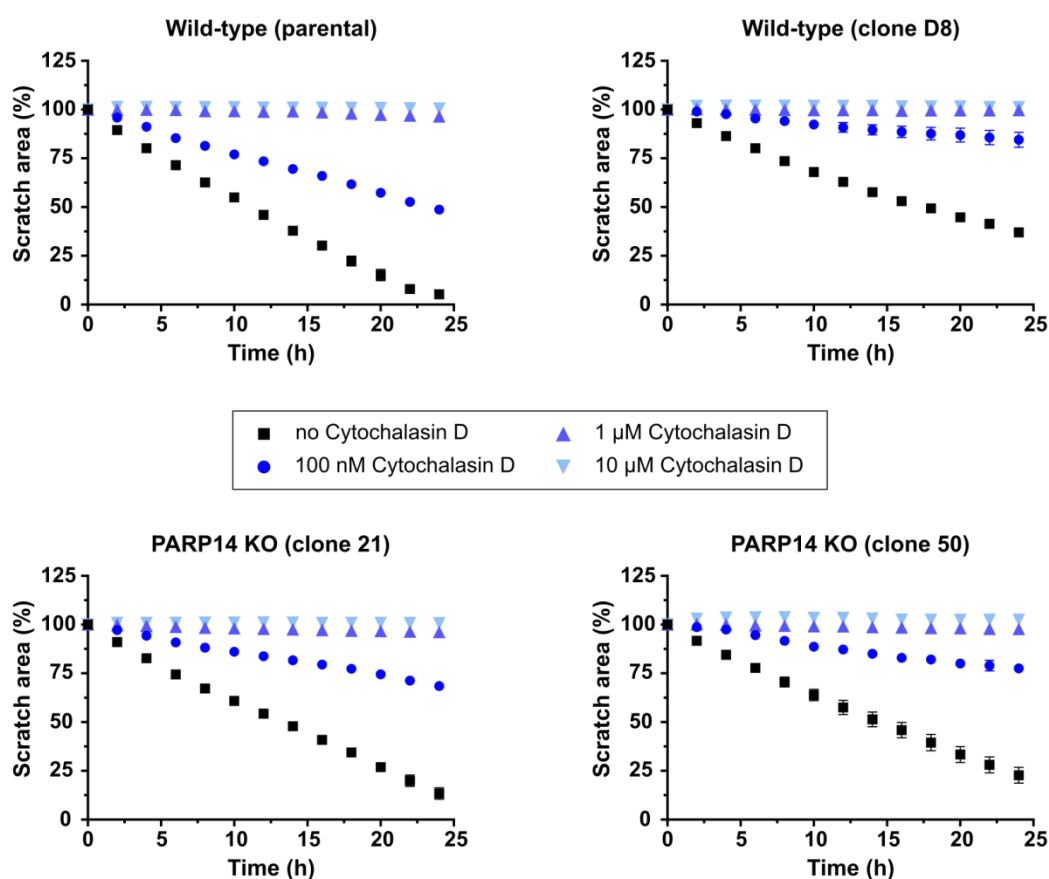


Figure 51. Establishing the scratch migration assay for wild-type and PARP14 KO HeLa cell lines with Cytochalasin D as migration control inhibitor. Migration curves of HeLa cell lines in presence of Cytochalasin D at three different concentrations are shown. For all cell lines, 1 μ M Cytochalasin D efficiently inhibited cell migration. Data points are mean $\pm$ SEM, $n = 6$.

Finally, the migration potential of the HeLa wild-type and PARP14 KO cell lines was evaluated in the established scratch assay. In several independent repeats, both PARP14 KO cell lines

migrated almost always with the same rate, however, slower than the parental wild-type cell line and with a higher or similar rate as the clonal wild-type cell line [Figure 52, S27, S28]. After 24 hours, the initial cell-free scratch area was reduced by the parental wild-type cell line to over 90 %, with the migration rate even increasing with the passage number. The clonal wild-type and PARP14 KO cell lines migrated slower, reducing the initial scratch area to 30-50 % over 24 hours [Figure 52C]. The reason for the different migration behaviour of the wild-type parental and clonal cell line is unknown as for the previously observed difference in cell adhesion. It seems the clonal cell line may engage less with the substrate, potentially resulting in a reduced adhesive strength which could be necessary for more efficient migration. Although the wild-type parental and clonal cell lines thus provided a rather large window, it was nevertheless expected that the clonal PARP14 KO cell lines with the hypothesised impaired cell migration function would show a significantly reduced rate of migration compared to the clonal wild-type cell line. Since this could not be observed in all assays performed, these data together with the prior acquired cell adhesion data indicates that PARP14 is not as crucial for cell migration processes as stated by Vyas *et al.*¹⁰⁸ – if at all. In support of this is the observation that HeLa cells – wild-type and PARP14 KO – treated with Cytochalasin D developed small dendritic-like structures extending from the cell bodies in all directions, resulting in a very similar phenotype observed after PARP14 knock-down using siPARP14 #3. This very clear visual phenotype showing inability or significant impairment of motility, respectively, could not be observed in all generated HeLa PARP14 KO cell lines, neither under normal culture conditions nor during migration in the scratch assay. No visual indication or functional experimental data thus allowed concluding an impaired motility of HeLa cell lines due to PARP14 depletion. Alternative assays, e.g. transwell assays which were performed for showing the role of PARP9 in cell migration⁸³ or 3D spheroid migration assays which are able to closer mimic the *in vivo* situation in tumours²¹⁹, could further validate these results, however, were for this reason and due to time constraints not further considered. Instead the PARP14 wild-type and KO cell lines were further characterised regarding their metabolic activity and their response to DNA damage.

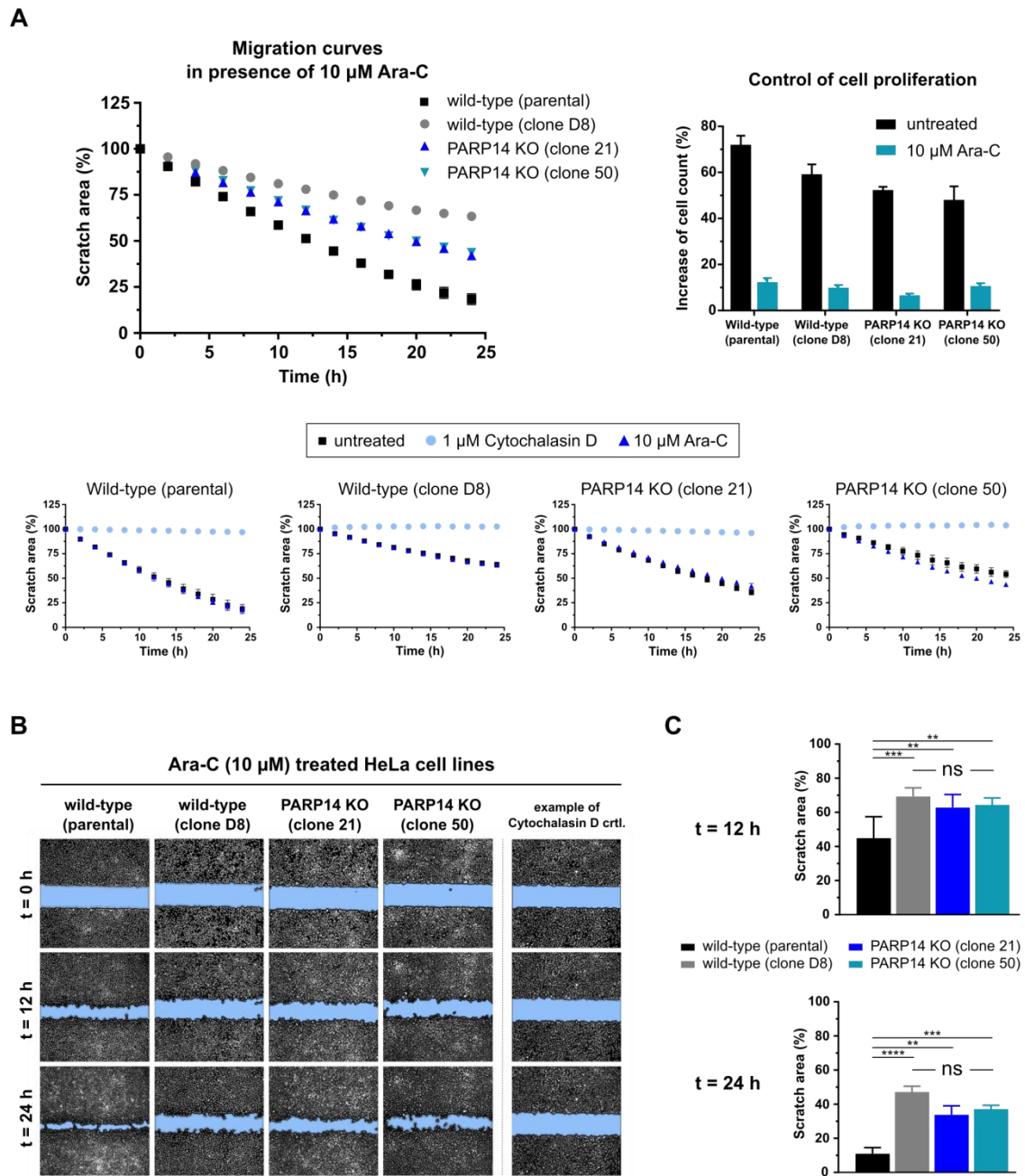


Figure 52. PARP14 wild-type and KO HeLa cell lines in the established scratch migration assay. (A) Migration curves of HeLa cell lines in presence of 10 μ M Ara-C are shown on the left, confirmation of efficiency of halting proliferation with Ara-C during the assay (over 24 h) on the right. Individual migration curves also including the corresponding Cytochalasin D controls for the different HeLa cell lines are shown in the panel below. Plots are representative for six independent experiments. Data points are mean $\pm$ SEM, $n > 4$ for untreated cells, $n > 8$ for Ara-C treated cells, and $n > 5$ for Cytochalasin D treated cells. **(B)** Representative images from the migration assay shown in (A) at the indicated time points which were further processed using a customised protocol of the IN Cell Developer Toolbox software [Figure S25]. It can be seen that the PARP14 KO cell lines migrate slower than the parental wild-type cell line, yet faster than the clonal wild-type cell line D8. **(C)** Graphs show the quantitation of the scratch area after 12 and 24 hours of six independent experiments. The assay repeats are shown in Figure S27 and S28A. Error bars represent SEM, $n = 6$, * adj. $p < 0.05$, ** adj. $p < 0.01$, *** adj. $p < 0.001$ (multiple comparison one-way anova test).

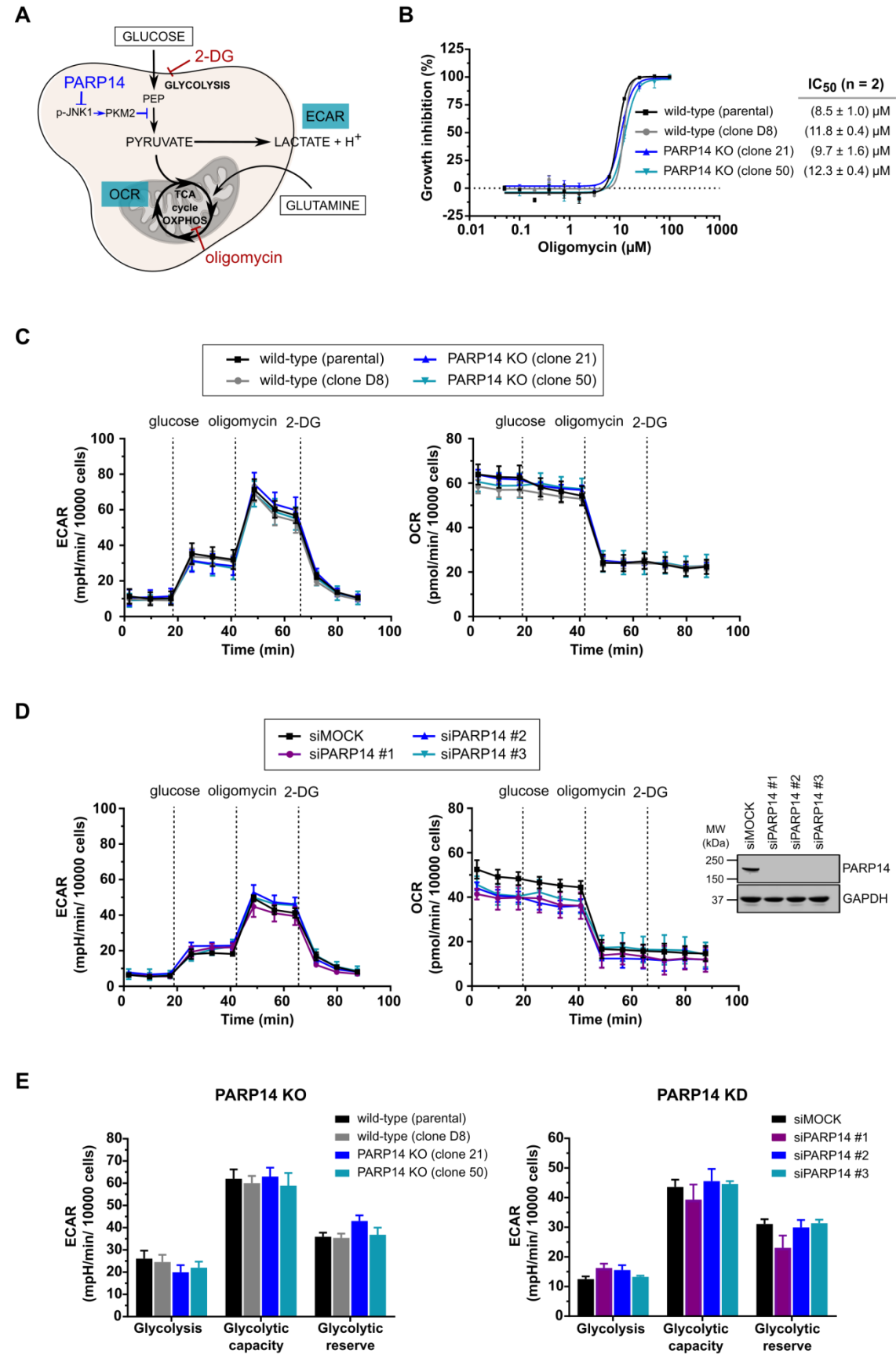
3.3. PARP14 and the Warburg effect – a brief study on PARP14 KO HeLa cells

Several studies suggested that PARP14 plays a critical role in regulation of cell metabolism either by inducing transcription of glycolytic genes or by contributing to the secretion of phosphoglucose isomerase, thereby increasing glycolysis, or by acting in the JNK signalling pathway, promoting the Warburg effect⁹⁶⁻⁹⁸. The Warburg effect is a characteristic of many cancer cells^{223,224} and describes the metabolic switch from ATP generation by mitochondrial oxidative phosphorylation to ATP generation by glycolysis only, regardless of the presence of oxygen. This metabolic process, referred to as ‘aerobic glycolysis’ is far less efficient in ATP and thus energy production, however, is thought to facilitate the use of intermediary glucose metabolites to generate reducing equivalents (NADPH) and consequently to enable the biosynthesis of macromolecules (carbohydrates, nucleotides, proteins and lipids). This change in metabolism allows doubling of the biomass required for cell proliferation^{225,226}. The pyruvate kinase isoform M2 (PKM2), which catalyses the final rate-limiting step in glycolysis, i.e. the conversion of phosphoenolpyruvate (PEP) to pyruvate, was shown to be a key mediator of the Warburg effect. Low PKM2 activity is assumed to slow down the passages of metabolites through glycolysis and consequently promotes the entering of these substrates into alternative metabolic pathways for macromolecular biosynthesis. In contrast, high PKM2 activity is associated with ATP synthesis over the TCA cycle and oxidative phosphorylation^{225,227,228}. Various factors have been reported to control PKM2 activity^{228,229} and recently, Iansante *et al.*⁹⁸ provided evidence that the low PKM2 activity, promoting aerobic glycolysis, is regulated by PARP14. Mechanistically, PARP14 has been shown to inhibit JNK1 which consequently prevented JNK1 binding and activation of PKM2 through phosphorylation [Figure 53A]. Whether JNK1 inhibition is mediated by PARP14 directly for instance by MARYlation or indirectly via upstream targets of JNK1 is unknown. Since PARP14 knock-down in several different cancer cell lines (HCC, RPMI-8226, U87, MCF7) impaired aerobic glycolysis metabolism, Iansante *et al.* concluded that PARP14 may generally promote the Warburg effect in human cancers⁹⁸. It was therefore tested

whether the HeLa PARP14 wild-type and KO cell lines indeed showed different metabolic behaviour, which would indicate a role for PARP14 in the Warburg effect. Compared to cells metabolically relying on oxidative phosphorylation thus showing increased mitochondrial respiration, cells with an aerobic glycolysis metabolism show increased glucose consumption and lactate production. The latter is converted from pyruvate by the lactate dehydrogenase and its excretion results in a higher rate of extracellular acidification^{98,225}. HeLa wild-type and PARP14 KO cell lines were therefore characterised in so-called seahorse assays that allow assessing the glycolysis and oxidative phosphorylation activity of the cells by simultaneously measuring in real-time the extracellular acidification rate (ECAR) and the oxygen consumption rate (OCR), respectively²³⁰ [Figure 53A]. According to the proposed model, PARP14-expressing cell lines were expected to show a higher ECAR and a lower OCR than the PARP14 KO cell lines. Prior to the characterisation of the cell lines, the seahorse assay was optimised in particular regarding cell number, oligomycin concentration to ensure to measure the maximum glycolytic capacity and the normalisation procedure of the data. Due to the sensitivity of the assay and unavoidable minor variabilities of the cell number in each assay well, the latter is crucial for obtaining consistent and reliable data and was performed by quantitation of the cell number in the respective wells by live staining of the cells with Hoechst33342 right after the assay and counting using an automated image analysis procedure [Figure S24]. In the following experiments, the HeLa wild-type and PARP14 KO cell lines were then sequentially exposed to saturating concentrations of glucose, oligomycin and 2-deoxyglucose and their metabolic reaction resulting in changes of the ECAR and OCR was recorded. A typical bioenergetic profile is shown in Figure S29A. In several independently performed assays, typical ECAR and OCR profiles were obtained for both HeLa wild-type and PARP14 KO cell lines, however, neither a significant difference in the ECAR nor in the OCR between the cell lines could be measured at any condition and time point in the assay [Figure 53C]. HeLa wild-type and PARP14 KO cell lines therefore showed the same rate of glycolysis as well as the same glycolytic capacity and reserve [Figure 53E]. No significantly different metabolic behaviour was observed regardless whether the cell lines were adapted to high glucose medium (DMEM) containing diabetic levels of glucose (25 mM) or to medium containing

approximately normal physiological glucose levels (5.5 mM, EMEM) or were starved of glucose prior to the assay [Figure 54, S29]. It was also noted by Iansante *et al.*⁹⁸ that PARP14 knock-down did not result in a significant change in oxygen consumption. Based on the metabolic model proposed by Iansante *et al.*⁹⁸, it could yet be considered that in HeLa PARP14 KO cell lines pyruvate converted from PEP through unsuppressed PKM2 activity could directly enter the TCA cycle^{98,225}, resulting in an increased OCR which perhaps cannot be detected in seahorse assays. Thus, it was hypothesised that the PARP14 KO cell lines might show higher sensitivity to oligomycin treatment due to their increased metabolic dependency on oxidative phosphorylation compared to HeLa PARP14 wild-type cell lines, which should gain most of their energy by aerobic glycolysis. However, IC₅₀ determinations indicated that oligomycin treatment affected the cell viability of all cell lines to the same extent [Figure 53B].

It was rather unexpected that no difference in the ECAR profiles representing the glycolytic activity of HeLa wild-type and PARP14 KO cell lines could be observed as it was in the seahorse assays performed by Iansante *et al.*⁹⁸. However, Iansante *et al.*⁹⁸ reported impaired aerobic glycolysis of cell lines in which PARP14 was only transiently knocked-down using shRNA. It is conceivable that the PARP14 KO cell lines developed mechanisms after gene editing which compensated for the constant loss of PARP14 and which can also promote aerobic glycolysis. Wild-type and PARP14 KO cell lines would consequently not show a difference in glycolytic activity. Therefore, HeLa cells with siRNA-induced PARP14 knock-down were characterised in the established seahorse assay. As shown in Figure 53D, the HeLa wild-type and PARP14 knock-down cells showed the same ECAR profiles resulting from the sequential exposure to glucose, oligomycin and 2-deoxyglucose. A difference in the glycolytic activity, indicating impaired aerobic glycolysis of PARP14 knock-down cells compared to control cells, was not measured [Figure 53E]. Only the OCR profiles revealed a reduced oxygen consumption of HeLa cells with PARP14 knock-down. However, a reduced oxidative phosphorylation activity through PARP14 knock-down is contradictory with the metabolic model suggested by Iansante *et al.*⁹⁸ and it is very likely that this effect was observed due to the cells being in a more quiescent state after PARP14 knock-down that is, as shown previously, associated with a decrease of cell viability.



Legend to Figure 53: Characterisation of the metabolic behaviour of PARP14 wild-type, KO and KD HeLa cells.

(A) Metabolic model adapted from Iansante et al.⁹⁸ showing the proposed role of PARP14 in the regulation of the Warburg effect. PARP14 is assumed to inhibit the conversion of PEP to pyruvate over the p-JNK–PKM2 regulatory axis which should be possible to confirm by measuring the extracellular acidification rate (ECAR) and oxygen consumption rate (OCR) in Seahorse experiments. Upon glucose injection (initiating glycolysis), pyruvate is generated which is consequently converted to lactate (resulting in an increased ECAR) as well as to acetyl-CoA. The latter enters the tricarboxylic acid (TCA) cycle that is linked to the oxidative phosphorylation pathway and the electron transport chain, resulting in an increased OCR. As control, 2-deoxyglucose (2-DG) is used to inhibit glycolysis and oligomycin which inhibits the ATP Synthase Complex V of the electron transport chain and thus shifts the energy production towards glycolysis (and consequently increases the ECAR). (B) Wild-type cells which should not depend on energy production over oxidative phosphorylation as strong as the PARP14 KO cell lines are not less sensitive to oligomycin treatment than PARP14 KO cell lines as assumed. A representative graph of two independent experiments is shown. Data points are mean $\pm$ SEM, $n = 3$. (C, D) Measurements of ECAR and OCR of PARP14 wild-type, KO and KD cells in Seahorse experiments before and after sequential injection of 10 mM glucose, 2 μ M oligomycin and 50 mM 2-DG. The ECAR and OCR profiles cannot reveal significant differences between PARP14 wild-type, KO and KD cells. Data points are mean $\pm$ SD, $n = 9$ for PARP14 wild-type and KO cells, $n = 6$ for siMOCK and siPARP14 cells. Graph for PARP14 wt and KO cells is representative for five independent experiments. The siRNA-induced PARP14 knock-down experiment was performed once. (E) Calculation of metabolic parameters (for equations see section IV.11.6.) for the experiments shown in (C and D) which are not significantly different for PARP14 wild-type, KO and KD cells. Data points are mean $\pm$ SD ($n = 9$ for characterisation of PARP14 KO cells and $n = 6$ for characterisation of PARP14 KD cells). Assay repeats are shown in Figure S29B and C.

In a further attempt to reveal potential differences in the metabolic behaviour of HeLa wild-type and PARP14 KO cell lines, it was tested how these cell lines were able to cope with removals of the growth medium supplements glutamine and glucose [Figure 54], the two major sources for cells in culture that are catabolised to free energy and utilised for macromolecular biosynthesis²²⁵. Additionally, pyruvate was removed to confirm that the increase in the ECAR results from conversion of the injected glucose to pyruvate and not by conversion of externally provided pyruvate. For this, cells were kept for one hour in the corresponding growth medium condition prior to the Seahorse assays. It should be mentioned that the change of the growth medium to the assay medium (containing no pyruvate, no glucose but optional glutamine, depending on the growth medium conditions) for one hour prior to the recording of the ECAR and OCR profiles was retained. The following observations were made: Firstly, removal of pyruvate for an additional hour prior to the assay had no measurable effect on the ECAR and OCR profiles of all cell lines and based on the control measurement it was assumed that one hour pre-incubation of

the cells in medium without pyruvate (as also recommended by the manufacturer) is sufficient to ensure accurate measurement of glycolysis. Secondly, additional removal of glutamine from the growth and assay medium significantly reduced the OCR and the typical spike of the ECAR after oligomycin injection, whereby HeLa wild-type and PARP14 KO reacted in the same way. As expected, both effects were reversed by injecting glutamine (together with glucose) just prior to the oligomycin injection. This experiment demonstrated that glutamine directly enters the TCA cycle and is utilised by HeLa wild-type and PARP14 KO cell lines for mitochondrial respiration. Its removal is not compensated by shuttling pyruvate converted from glucose via glycolysis into the TCA cycle. This is in agreement with a study showing that glutamine, not glucose, is the major source for energy production by oxidative phosphorylation in cultured HeLa cells²³¹. It furthermore provides an explanation for the previous observation that all cell lines showed the same sensitivity towards oligomycin. Of note is that PARP14 knock-down cells also showed the same metabolic reaction compared to control cells upon removal of glutamine and pyruvate from the growth and assay medium. Finally, upon additional removal of glucose from the growth medium, all HeLa wild-type and PARP14 KO cell lines showed a significant decrease of the basal glycolysis rate and a dramatic increase of the ECAR, thus of glycolysis, upon glucose injection. The former observation was expected for all cell lines, however based on the model, HeLa wild-type cell lines were thought to show a higher ECAR than PARP14 KO cell lines due to the PKM2 suppression by PARP14. Testing PARP14-depleted cells compared to control cells in the condition with pyruvate and glutamine removed was also thought to provide a possibility to observe any regulatory influence of PARP14 on PKM2. Hereby only glucose was provided for conversion to pyruvate which is - depending on the PKM2 activity - either converted to lactate (increasing the ECAR) or passed into the TCA cycle which is down-regulated due to the removal of glutamine and could be theoretically up-regulated through pyruvate on demand of energy production through oxidative phosphorylation (increasing the OCR). However, both experimental set-ups could again not provide any evidence of regulation of the glycolytic rate by PARP14.

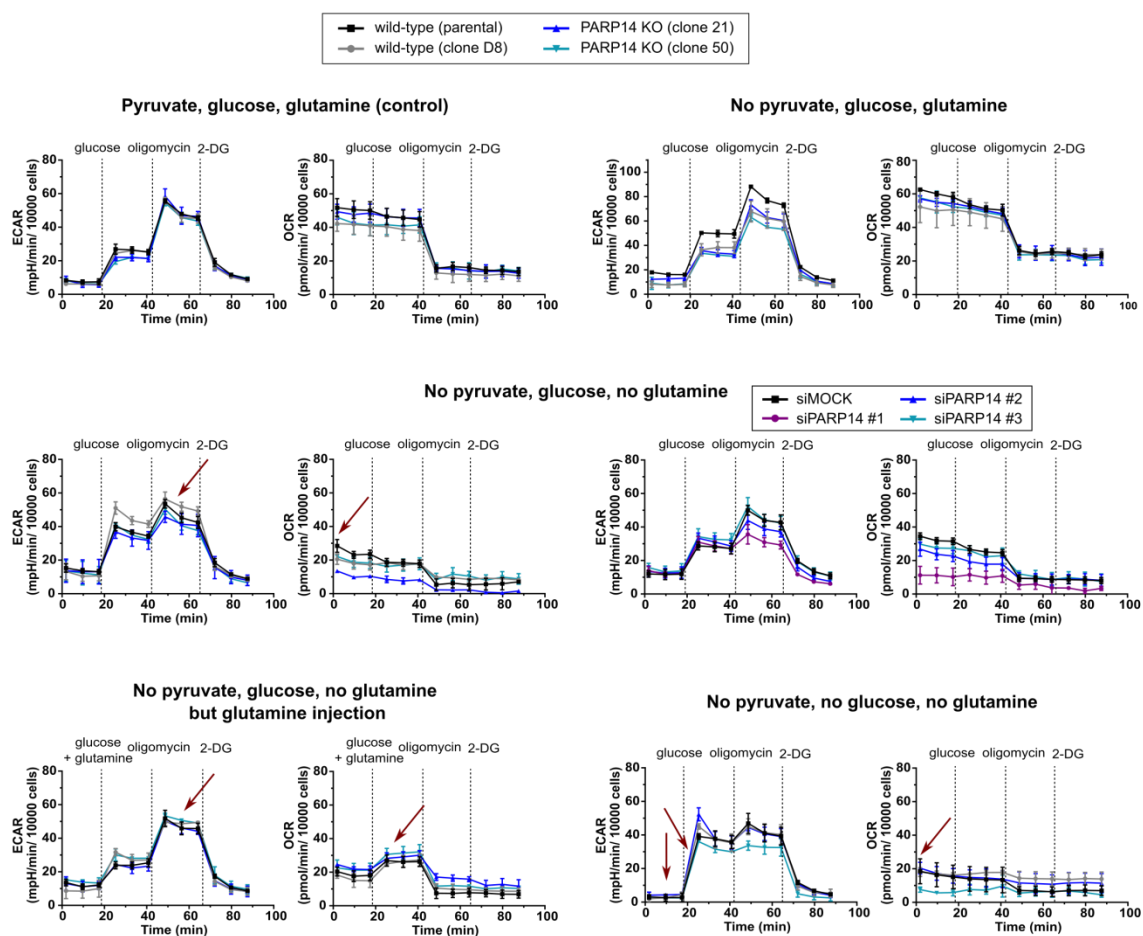


Figure 54. Variations of the seahorse assay performed with PARP14 KO, KD and wild-type HeLa cells. Prior to the assay measuring the ECAR and OCR before and after sequential injection of glucose, oligomycin and 2-DG, cells were kept for 1 h in growth medium containing no FBS and supplements as indicated over the graphs. In brief, removal of pyruvate has no effect on ECAR or OCR, removal of glutamine, that directly enters the TCA cycle, reduces the OCR, and removal of glucose decreases basal glycolysis and results in strong increase of the ECAR after injection of glucose in the seahorse assay (as indicated by red arrows). None of these variations could detect a different metabolic reaction of PARP14 KO, KD and wild-type cells. Data points are mean $\pm$ SD, $n \geq 3$. The ‘no pyruvate, no glucose, no glutamine’, ‘no pyruvate, glucose, glutamine’ and the siRNA experiment were performed once, graphs of the other sections are representative for two independent experiments.

In summary, throughout the performed seahorse assays, HeLa wild-type and PARP14-depleted cells consistently showed similar bioenergetic profiles. PARP14 depletion, through knock-out or siRNA-induced, did not result in a metabolic phenotype indicating impaired aerobic glycolysis which is in contrast to studies performed by Iansante *et al.*⁹⁸. Since PARP14 was proposed as universal regulator of the Warburg effect and HeLa cells moreover were previously characterised as cancer cells showing aerobic glycolysis²³², cell-type specificity may probably be excluded as reason for the inability to confirm the role of PARP14 in the Warburg effect. It would therefore be

necessary to investigate more thoroughly the proposed regulatory axis PARP14–JNK1–PKM2. This would include the characterisation of the JNK1 and PKM2 activity in HeLa wild-type and PARP14 KO cell lines as well as elucidating the mechanism of JNK1 inhibition by PARP14 which was not only associated with the Warburg effect but also with B-cell survival over the JNK2 signalling pathway⁹⁹. Probably most enlightening for understanding the role of PARP14 in cell metabolism would however be monitoring the glucose metabolism in HeLa wild-type and PARP14 KO cells by performing mass spectrometric metabolic flux analysis using ¹³C-labelled glucose. However due to the time constraint of this project, such an in-depth analysis of the PARP14 role in regulation of cancer cell metabolism could not be realised and instead the HeLa wild-type and PARP14 KO cell lines were characterised regarding their response to DNA damage.

3.4. Role of PARP14 in DNA damage response

Apart of the variety of cytoplasmic functions, nuclear functions including a role in transcriptional regulation and more recently in DNA repair were proposed for PARP14^{90,109}. In regard to the previously performed studies [Chapter II.1.2.] showing the recruitment of the PARP14 macrodomains to laser-induced DNA damage sites in U-2 OS cells, it was evaluated whether HeLa and U-2 OS PARP14 KO cell lines indeed showed a phenotype indicating a role of PARP14 in DNA damage response.

PARP14 was shown by Nicolae *et al.*¹⁰⁹ to be required for efficient repair of DNA strand breaks at replication forks by homologous recombination (HR). Together with translesion synthesis (TLS), HR is one of the major mechanisms essential for DNA repair at stalled replication forks. Stalling of the replication machinery can result from unrepaired DNA lesions at the replication site or can occur at common fragile sites, i.e. unstable genomic regions difficult to replicate due to their sequence properties. Since prolonged replication arrest can lead to fork disassembly and double-strand break formation, efficient repair mechanisms at these sites are critical for genomic stability^{233–235}. The recombinase RAD51 is an important mediator of the HR reaction, which is loaded by BRCA2 on the DNA strand end and catalyses the pairing of homologous DNA sequences²³⁶. PARP14 siRNA studies performed with U-2 OS and HeLa cells provided evidence that PARP14 associates with the DNA replication machinery component PCNA and is involved in HR at a step following RAD51 foci formation. Based on the interaction of PARP14 MD2 with RAD51, it was assumed that RAD51 is MARYlated after DNA damage. PARP14 was thereby suggested as primary enzyme responsible for RAD51 MARYlation and thus together with mono-ADP-ribosylation itself as a mediator of HR¹⁰⁹.

Cell fractionation assays performed with U-2 OS and HeLa cells showed that PARP14 localised primarily in the cytoplasm, however, was also present in the nucleus as a soluble form and bound to chromatin [Figure 55A, S30]. Furthermore, it could indeed be confirmed in U-2 OS cells that

the amount of chromatin-bound PARP14 increased upon treatment of the cells with the replication fork stalling agent hydroxyurea (HU) [Figure 55A]. HU induces replication stress by inhibiting the ribonucleotide reductase. Consequently, DNA synthesis is inhibited, slowing down initiation and progression of replication forks and finally creating DNA lesions and preventing the cell cycle passing through S phase^{237,238}. Of note is that in this, as in all following experiments, cells were treated with HU for 24 hours at a concentration of 2 mM. It was shown in U-2 OS cells that short HU treatment (< 2 h) only transiently stalls replication forks which are then restarted after release from the HU block. Only a longer HU treatment of 24 h or longer triggers the collapse of the fork resulting in double-strand breaks and activates HR repair²³⁹. Yet in contrast, PARP14 re-localisation to the nucleus after treatment of the cells with the topoisomerase inhibitor camptothecin (CPT) was not observed under the assay conditions tested. CPT also leads to replication stress by creating torsional strain within the DNA and thus induces replication-dependent double-strand breaks²³⁷. The effect could furthermore not be confirmed in HeLa cells, probably due to the far lower PARP14 levels in the nucleus [Figure S30B].

U-2 OS wild-type and PARP14 KO cell lines were then characterised for their response to replication stress induced by HU treatment. As shown in Figure 55B, PARP14 KO cell lines were indeed more sensitive to replication stress compared to the PARP14 wild-type cell lines confirming the reported PARP14 knock-down phenotype. In contrast to PARP14 wild-type cells, HU treatment led to strong induction of apoptosis in PARP14-depleted cells, as indicated by cleaved PARP1 which was detected up to 24 hours of recovery from the HU treatment. Furthermore, both PARP14 KO cell lines showed a far stronger activation of KAP1 which is phosphorylated in an ATM-dependent manner upon sensing of double-strand breaks by the MRN complex^{240,241}. pKAP1 levels were significantly attenuated in PARP14-expressing cells, suggesting their more efficient repair of DNA lesions and thus better abilities to cope with replication stress induced by HU. Similarly, the γ H2Ax signal was increased in the PARP14 KO cell lines, indicating together with the pKAP1 activation the presence of more unrepaired DNA lesions even 12 hours after release from the replication stress.

The high pKAP1 levels in PARP14 KO cell lines upon HU treatment indicated the activation of the ATM-dependent repair pathways and thus prompted analysis of the PARP14 wild-type and KO cell lines for 53BP1 foci formation upon induction of replication stress by HU. The tumour suppressor protein p53-binding protein 1 (53BP1) also functions in the ATM signalling cascade and localises to sites of DNA double-strand breaks, whereby KAP1 phosphorylation was shown to be critical for 53BP1-mediated repair^{242–244}. Interestingly even under normal replication conditions, HeLa and U-2 OS PARP14-depleted cells showed the formation of several large 53BP1 bodies in the nucleus, many of them co-localising with γ H2Ax foci, in comparison to PARP14-expressing cells [Figure 55C, 56]. Following HU treatment, these 53BP1 bodies became diffuse where PARP14 KO cell lines showed more cells with numerous distinct 53BP1 bodies than PARP14 wild-type cell lines. Moreover, the induction of replication stress elicited γ H2Ax foci formation in both PARP14 wild-type and KO cell lines as expected, however, γ H2Ax staining was far stronger in PARP14-depleted cell lines. The PARP14 KO cell lines showed more cells with very high γ H2Ax activation throughout the nucleus, which were assumed to be apoptotic. Thus, these results were in line with the previous observations¹⁰⁹ and confirmed the sensitivity of the PARP14 KO cell lines to HU-induced replication stress and DNA damage.

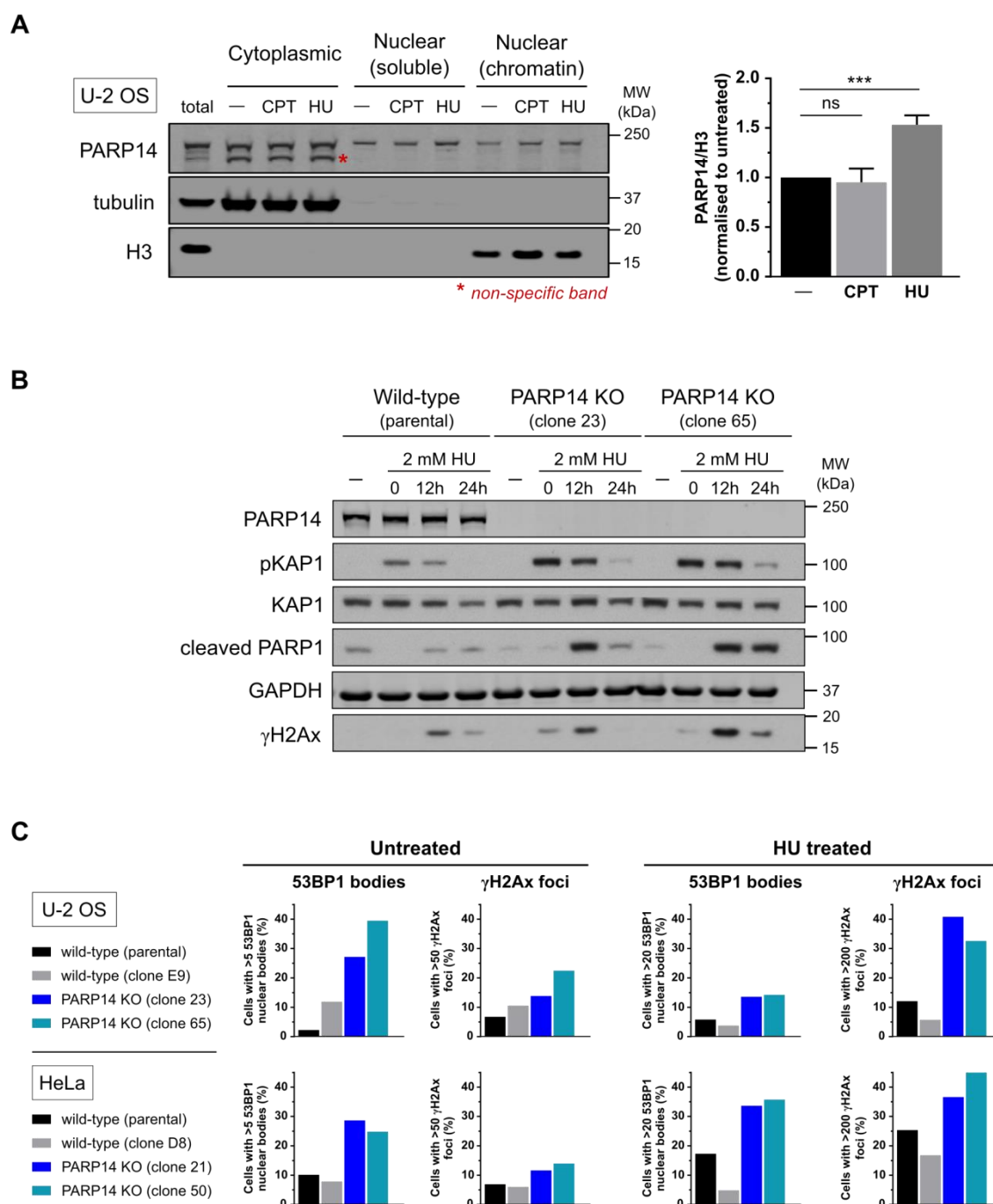


Figure 55. Indication for PARP14 role in DNA damage repair upon replication stress induced by HU. (A) Cell fractionation showed PARP14 localisation in cytoplasm and nucleus and confirmed increased PARP14 binding to chromatin in U-2 OS cells after HU treatment (2 mM for 24 h) reported by Nicolae *et al.*¹⁰⁹ in contrast after CPT treatment (1 μM, 2 h). α-tubulin and histone H3 were used as loading controls for cytoplasmic and nuclear fractions, respectively. Data in the graph is the densitometry analysis of three independent experiments, error bars represent SD, *** $p < 0.001$ (unpaired two-tailed *t*-test). Assay repeats are shown in Figure S30A. (B) Both PARP14 KO cell lines showed increased DNA damage response and a slower recovery from DNA damage induced by HU than the parental wild-type cell line. Wild-type and PARP14 KO U-2 OS cells were left untreated (-) or were treated with 2 mM HU for 24 h and analysed at $t = 0$ and after 12 h and 24 h recovery time by Western blotting with the indicated markers. GAPDH was used as loading control. The experiment was performed once. (C) Analysis of immunofluorescence experiments performed with U-2 OS and HeLa wild-type and PARP14 KO cell lines which were left untreated or were treated with 2 mM HU for 24 h. Representative images are shown in Figure S6. The number of 53BP1 nuclear bodies and γH2Ax foci (of 200-250 cells) was quantified for each condition which were increased in PARP14 KO cell lines compared to the wild-type cell lines. The experiment was performed once with one repeat for the U-2 OS cell lines and 53BP1 staining shown in Figure S31.

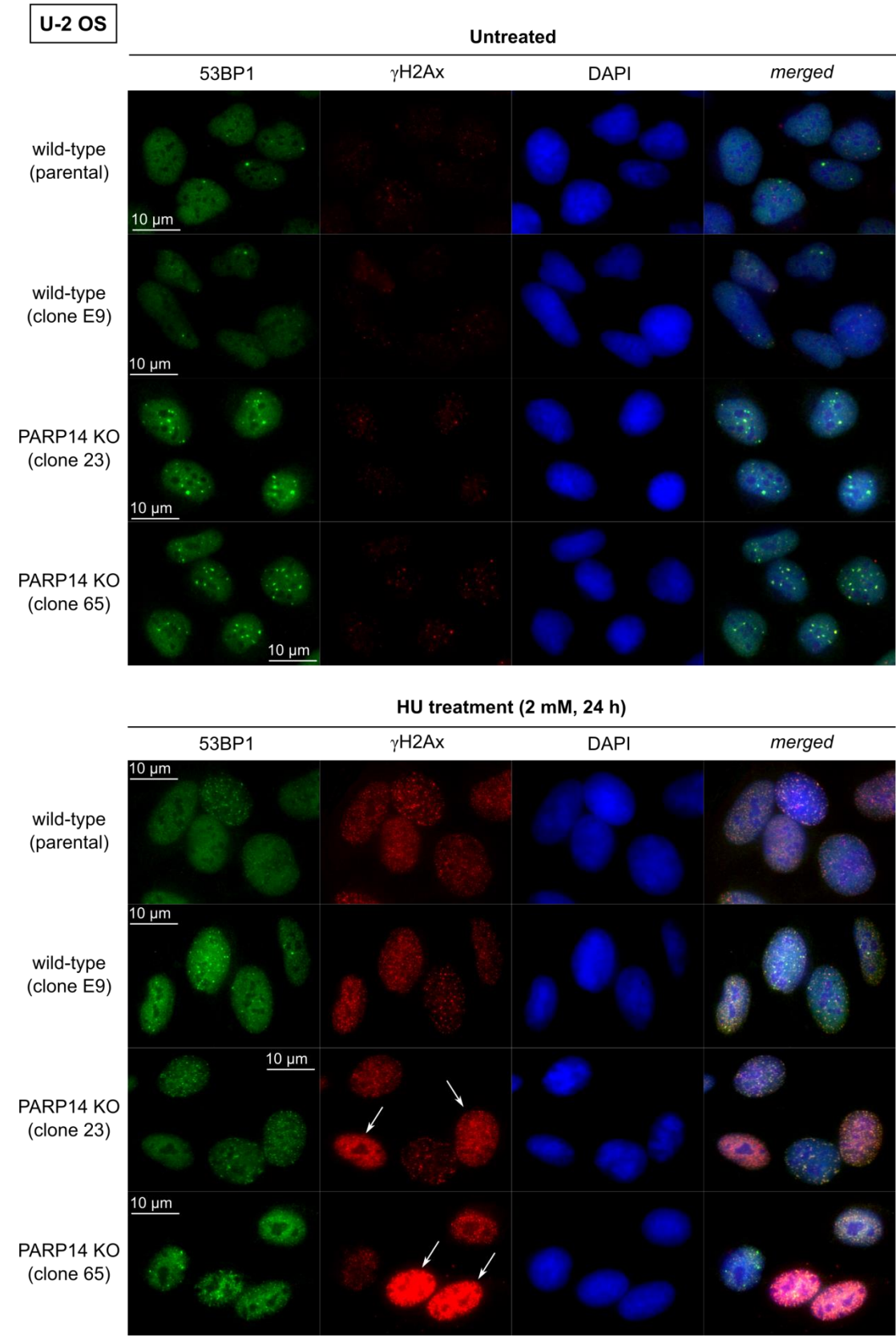


Figure 56. PARP14 KO U-2 OS and HeLa cell lines show stronger DNA damage response under normal replication conditions and under HU-induced replication stress compared to PARP14 wild-type cell lines. Cells were left untreated or were treated with 2 mM HU for 24 h and then processed for immunostaining with anti-53BP1 (green) and γ H2Ax (red) antibodies. DNA was counterstained with DAPI. White arrows indicate γ H2Ax positive (presumably apoptotic) cells which were counted as nuclei containing more than 200 foci. The quantitation of 53BP1 and γ H2Ax nuclear bodies is shown in Figure 55C.

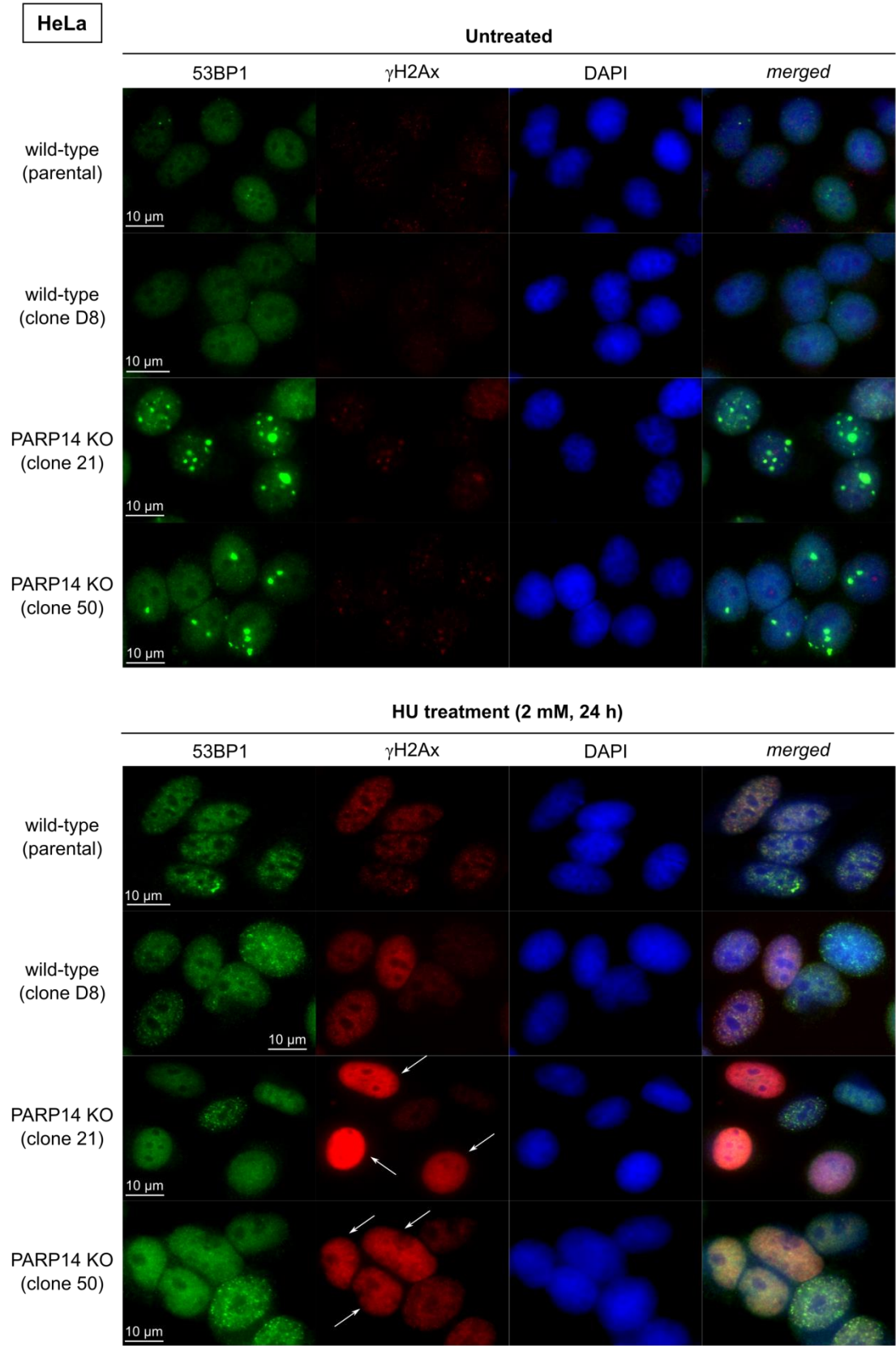


Figure 56 [continued].

Finally, it was tested whether the U-2 OS PARP14 KO cell lines also showed sensitivity to other replication stress and DNA damage (olaparib, CPT, cisplatin, doxorubicin) or apoptosis (staurosporine) inducing agents [Figure 57, 58]. Upon PARP1 inhibition, single-strand breaks are unrepaired and collapse replication forks into double-strand breaks in S phase²⁴⁵. Cells with HR deficiency are therefore sensitive to PARP1 inhibition which could be shown for PARP14 knock-down cells¹⁰⁹. Sensitivity of PARP14 KO cell lines to olaparib could yet not be confirmed either in cell viability or clonogenic survival assays. Besides, it was noted that the clonal cell lines were all significantly less sensitive to PARP1 inhibition than the parental wild-type cell line and it is tempting to correlate this observation to the selection background of these cell lines. Further, PARP14-depleted cells showed also no increased sensitivity to CPT as well as to staurosporine and doxorubicin. HepG2, MCF7 and HCC cells with induced PARP14 knock-down were previously reported to be more sensitive to the latter^{98,139}. Clonogenic survival assays could furthermore not reveal any increased sensitivity of PARP14 KO cells to ionising radiation (IR), that activates the ATM repair pathway²⁴⁶ in which PARP14 seemed to be involved, or cisplatin, which is known for inducing various structural DNA modifications creating severe DNA lesions²³⁷. However, hypersensitivity of the PARP14 KO cell lines to HU treatment was confirmed in clonogenic survival assays [Figure 58B, C] and partly in cell viability assays. It is unclear why in the latter the PARP14 KO cell line derived from clone 23 consistently showed the increased sensitivity to HU resulting in an IC_{50} of 120 μ M, however, the PARP14 KO cell line derived from clone 65 was as susceptible to the treatment as the parental wild-type cell line, with both showing an IC_{50} of ~240 μ M [Figure 57, S32]. Thus, among the different types of stress and DNA damage induced, only a sensitivity of the PARP14 KO cell lines to HU treatment was observed under the assay conditions tested.

These results may therefore allow the following conclusions: Under normal replication conditions, PARP14 may be required to suppress spontaneous DNA damage. This could explain the observation of several large 53BP1 nuclear bodies that co-localised with γ H2Ax foci in PARP14 KO cells in the absence of induced replication stress which probably marked the

presence of unrepaired DNA lesions activating ATM response. The cell cycle stage of these cells, of which the majority are to be assumed in G1 phase, could be revealed by cyclin A counterstaining, a marker for cells in S and G2 phase ²⁴⁷. In S phase, PARP14 may be indeed essential for protection against replication stress as suggested by Nicolae *et al.* ¹⁰⁹. Upon induction of replication stress and consequently DNA damage by HU, the PARP14-depleted cell lines showed increased DNA damage associated with a slower recovery rate from the treatment compared to the PARP14-expressing parental cell line. This observation together with the increased frequency of 53BP1 nuclear bodies and γ H2Ax foci in PARP14 KO cells after HU treatment may reflect their lower efficiency of replication stress induced DNA repair. Since 53BP1 is a mediator and activator of ATM in response to DNA damage ²⁴³, it further indicates that the PARP14 HR-mediated DNA repair involves activation of ATM signalling pathways. However, it is unclear why PARP14 KO cell lines seem to rather specifically show sensitivity to HU. PARP14 wild-type and KO cell lines reacted with same sensitivity to IR and other chemical agents inducing replication stress and DNA damage including in particular CPT, which rather may not suggest a general activity of PARP14 in HR-mediated repair but potentially a more specific mechanism elicited by HU treatment. Furthermore, this observed specificity to HU treatment may rather indicate that the previously performed laser micro-irradiation experiments [Chapter II.1.2.] showing the recruitment of PARP14 macrodomains to laser-induced DNA damage sites in a PARylation-dependent manner may not reflect physiological activities of endogenous PARP14 in all types of DNA repair. However, all this is speculative and a more in-depth analysis including repeating these observations and re-considering the assay conditions is required. Finally, the impairment of DNA repair and progression of the cells through S phase could directly result in the increased activation of apoptosis in PARP14-depleted cells, which can be concluded from the formation of cleaved PARP1 and agrees very well with the reported importance of PARP14 for cell survival ^{93,98,108,139}.

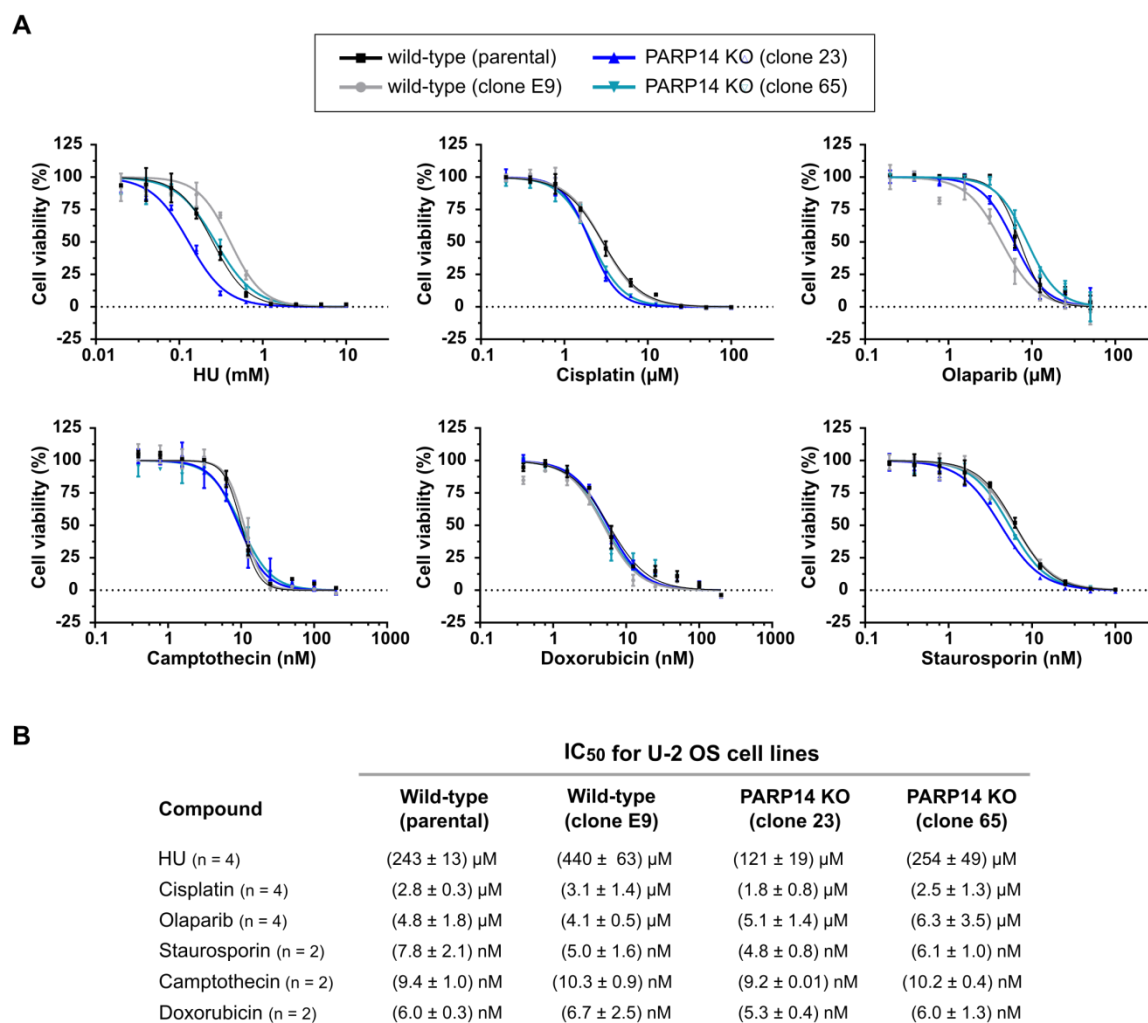


Figure 57. Wild-type and PARP14 KO U-2 OS cell lines tested for sensitivity to DNA damage and apoptosis inducing agents by resazurin viability assays. (A) U-2 OS cells were exposed to the indicated compounds at different concentrations and cell viability was measured after 6 days of treatment using the AlamarBlue™ reagent. Graphs show representative results for each compound where data points are mean ± SD, n = 3. **(B)** The IC₅₀ values calculated from these assays are provided in the table for each U-2 OS cell line, with the number of experiments given behind the corresponding compound. All viability assays performed are shown in Figure S32.

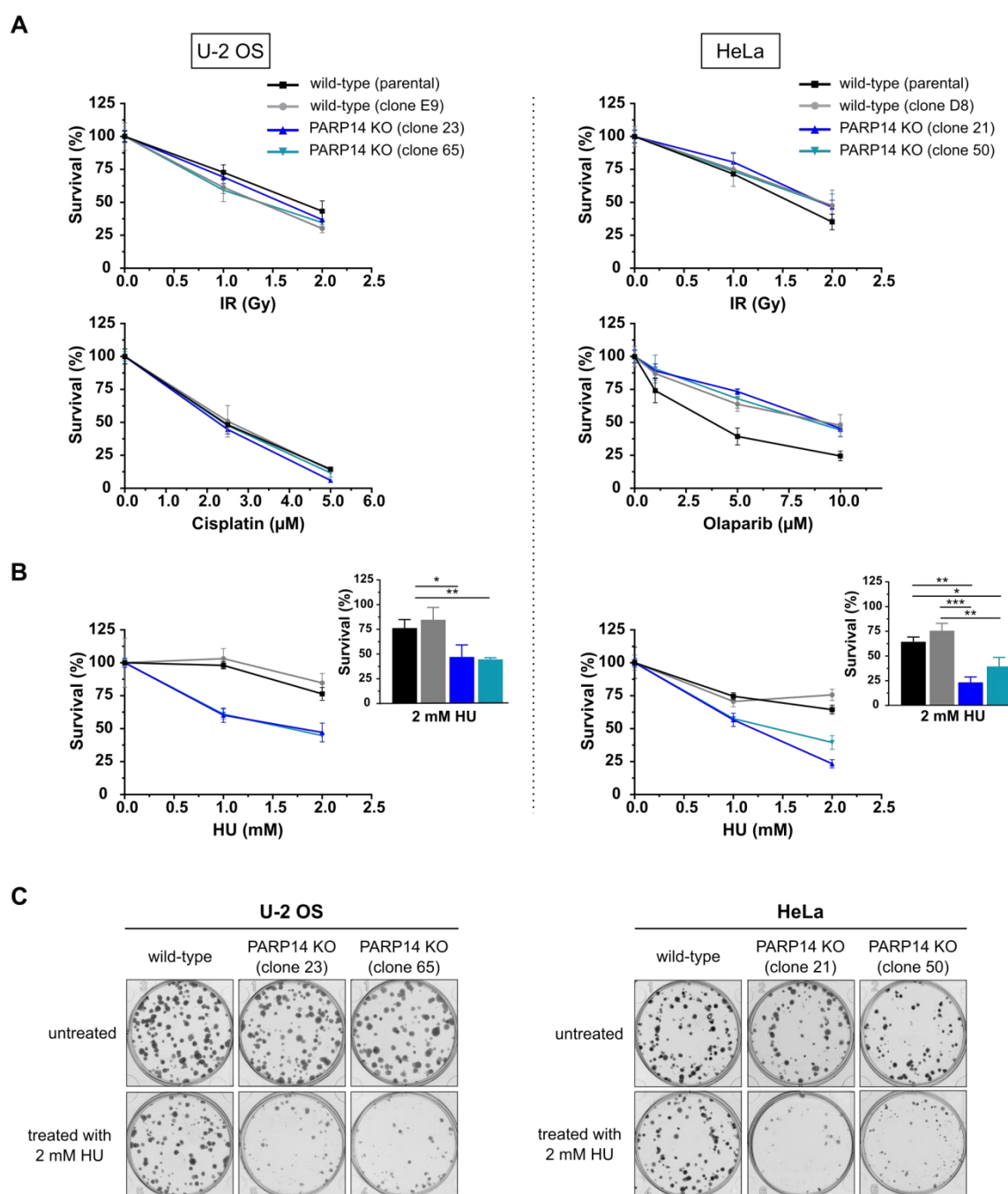


Figure 58. Wild-type and PARP14 KO HeLa and U-2 OS cell lines tested for sensitivity to DNA damage inducing agents by clonogenic survival assays. (A) Clonogenic experiments performed with IR and cisplatin or olaparib treated HeLa and U-2 OS cell lines did not reveal a difference in sensitivity to DNA damage of wild-type and PARP14 KO cells. Representative graphs of two independent experiments are shown for IR and olaparib. The clonogenic assay performed with cisplatin was performed once. Data points are mean $\pm$ SEM, $n = 3$. (B) Clonogenic assays showed sensitivity of PARP14 KO cell lines to HU treatment (2 mM, 24 h) compared to the corresponding wild-type cell lines. Graphs shown are representative for three independent experiments performed with the HeLa and U-2 OS cell lines. Error bars represent SEM, $n = 3$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (unpaired two-tailed t -test). (C) Representative images obtained from the clonogenic assays in (B) are shown. All clonogenic survival assays performed are shown in Figure S33.

Conclusions from the PARP14 cell biology studies

The focus of the last part of this work was to explore pharmacological targeting strategies for PARP14 with the aim to develop PARP14-specific assays to evaluate future PARP14 inhibitors. PARP14 is localised primarily in the cytoplasm but was found also in the nucleus, making (different) functions in both cellular compartments plausible. Out of the numerous functions proposed for PARP14, three functions derived from cellular models, with the potential to develop PARP14 cell assays for compound screening and evaluation, were analysed, i.e. the role of PARP14 in cell migration, in regulation of the Warburg effect and in DNA repair. These PARP14 functions were previously studied by comparing control cells with their corresponding PARP14 knock-down phenotype^{98,108,109}. Since siRNA-induced PARP14 knock-down, however, significantly reduced cell viability and resulted in visible off-target effects (siPARP14 #3) providing difficulties for assay development, HeLa and U-2 OS PARP14 KO cell lines were generated using the CRISPR-Cas9 system. The preparation of cell lines expressing PARP14 with ADPR-binding-deficient macrodomains as control cell lines for the assays would have been also highly desirable but could not be realised in the course of this work. Since off-target cleavages of sgRNAs cannot be excluded, two PARP14 KO clones of each cell line were selected for the following studies that were (apart from the U-2 OS clone 65) generated with a pair of nickase Cas9 sgRNAs, a strategy which was shown to greatly minimise any off-target mutagenesis^{248,249}. This should ensure that only PARP14-specific effects were observed. Since the PARP14-depleted cell lines displayed no visual difference compared to PARP14-expressing cell lines, assays were performed to functionally characterise these cell lines and which may allow the following conclusions: Firstly, PARP14 is not as essential for cell motility including cell adhesion and migration as reported¹⁰⁸. Secondly, a role of PARP14 in the Warburg effect was not confirmed by metabolically characterising PARP14 KO and knock-down cell lines. Thirdly, PARP14 may indeed play a role in protection against replication stress and DNA damage. Although a far more detailed analysis of the underlying mechanism of PARP14-mediated DNA repair assumed by HR

is required, this work provides further evidence for the hypothesis that PARP14 (macrodomain) inhibitors might prove useful to potentiate genotoxic chemotherapy by rendering cancer cells more susceptible to DNA damage¹⁰⁹. Therefore, the established migration and Seahorse assays cannot be considered for evaluating PARP14 inhibitors. Instead, specific DNA damage assays need to be set-up that allow characterising PARP14-mediated DNA repair in cancer cells. Immunofluorescence and Western blot analysis of DNA damage marker proteins of the HR and ATM signalling pathways including pKAP1, 53BP1, γ H2Ax, RPA in combination with apoptosis marker, e.g. cleaved PARP1, and general survival assays could be performed for initial characterisation of PARP14 inhibitors. Then, specific HR assays for instance using the DR-GFP reporter²⁵⁰ may be established for final validation of the efficiency of PARP14 (macrodomain) inhibitors to impair HR. Since compounds in the developmental phase may have an uncharacterised profile and can non-specifically induce apoptosis and DNA damage in particular if used in higher concentrations, resulting in a positive readout in general DNA damage assays, PARP/macro-domain inactive control compounds are seen as crucial controls for performing these assays. Finally, the established laser micro-irradiation assay monitoring the recruitment of PARP14 macrodomains to DNA damage sites [Chapter II.1.2.] is a valuable tool to initially prove target engagement of PARP14 macrodomain inhibitors and should therefore be routinely performed prior to any further evaluation.

III. FUTURE DIRECTIONS

Future directions for this work would need to include studying the role of the PARP14 macrodomains for PARP14 function at its respective target sites. For this, the generation of cell lines with ADPR-binding-deficient PARP14 macrodomains by introducing the corresponding point mutations validated in this thesis using CRISPR-Cas9 technology could be considered. Since the DNA damage role of PARP14 was the only role out of the previously published PARP14 studies that could be verified, using such CRISPR-Cas9-generated point mutants to study the DNA damage role of PARP14 would be an interesting avenue for future studies. Furthermore, future work may also be directed towards the development of more potent and selective PARP14 MD3 inhibitors using the structural information provided by the co-crystal structures of the fragment hits as well as optimising GeA-69 to a lead compound with desired low nanomolar affinity. Additional structural studies on dual domain constructs may help to clarify if a bivalent inhibitor that simultaneously binds MD2 and MD3 could be developed. Ideally, linkage of MD2 and MD3 targeting compounds would result in a very potent and specific PARP14 inhibitor which would have significant advantages over any PARP14 catalytic domain inhibitor. Yet, once available, both single and bivalent PARP14 macrodomain inhibitors would be highly useful tools for PARP14 target validation studies and perhaps beyond. Furthermore, the PARP14 MD3 fragment hits provide insight into chemical scaffolds that might be generally amenable for the ADPR binding site and may therefore serve as starting points for inhibitor development for other macrodomains. Ultimately, macrodomain inhibitors optimised in their biophysical properties as well as potency and selectivity will aid to elucidate the (patho-)physiological functions of macrodomains and macrodomain-containing proteins. A detailed understanding of their roles in the various ADP-ribosylation-regulated cellular processes and also of their so far rather poorly characterised signalling pathways and networks may in the future identify novel pharmacological targeting strategies.

IV. EXPERIMENTAL PROCEDURES

1. General reagents

General reagents met at least the criteria of the purity standard p.a. and were purchased from Sigma Aldrich with the following exceptions: Acros Organics (agar (powder), magnesium chloride, rubidium chloride, manganese(II) chloride), Alfa Aesar (Triton X-100), Apollo Scientific (HEPES, TCEP), ChemCruz (bicine), Fluka Analytical (polyethylenimine), Merck (LB broth (Miller), peptone from casein, glycerol, trichloroacetic acid), Melford Biolaboratories (yeast extract granulated), Thermo Fisher (chloramphenicol, dipotassium hydrogen phosphate, EDTA, ethanol (absolute), imidazole, IPTG, puromycin dihydrochloride, sodium dodecyl sulfate, sodium hydroxide) and VWR chemicals (sodium chloride). Buffers and solutions were prepared with ultra-pure water (18.2 MΩ) obtained from the water purification system Purelab Option DV25 and optionally sterile-filtered through 0.2 μm membrane filters (Pall Life Sciences).

2. Oligonucleotides

Unless stated otherwise, the following oligonucleotides were purchased from IDT (Integrated DNA Technologies (Belgium)).

PARP14 Gateway® cloning

Primer name	Primer sequence 5' → 3'	Comment
PARP14A-GW-f001 PARP14A-GW-r001	GACAAGTTTGTACAAAAAAGCAGGCTCCATGGGGCAGAAGTGCTTCTCTC GGACCACTTTGTACAAGAAAGCTGGGTTCACCCTTCTCTTTTCTTCATGT	PARP14 Macro1-3
PARP14A-GW-f002 PARP14A-GW-r002	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGCAGCAGCGGGGCGCTG GGAA GGGGACCACTTTGTACAAGAAAGCTGGGTTCAGACCTTTCGCTTCTTCTTC GGGACGAGATTCCATTAGCCC	PARP14 MD2 + NLS
PARP14A-GW-f003 PARP14A-GW-r003	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGGGCAGAAGTGCTTCT CTC GGGGACCACTTTGTACAAGAAAGCTGGGTTCAGACCTTTCGCTTCTTCTTC GGCCCTTCTCTTTTCTTCATGT	PARP14 Macro1-3 + NLS
PARP14A-GW-f004 PARP14A-GW-r004	GGGGACAAGTTTGTACAAAAAAGCAGGCTCCATGGGGCAGAAGTGCTTCT CTC GGGGACCACTTTGTACAAGAAAGCTGGGTTCATTTTCTAAACGTAATAAGG TACTCT	PARP14 macroPARP
PARP14A_ins_f005 PARP14A_ins_r005	GAAAAGAGAAGGGACTCAGCTTCTTCCCCGAAGAAGAAGCGAAAGGTCT GAACC GCTTCTTCTTCGGGAAGAAAGCTGAGTCCCTTCTCTTTTCTTCATGTGGC ATA	PARP14 Macro1-3 + spacer + NLS

PARP14 mutagenesis

Primer name	Primer sequence 5' → 3'	AA change	Comment
PARP14-M1-F PARP14-M1-R	TATGGTGA <u>ACT</u> GGCCGCTGCGCTCTCAAAAGCAGCTGGCC GGCCAGTTCACCATAATGCTTAAGGTCCTCATTAGATGCA	p.G832E	ADPR-binding deficiency in PARP14 MD1
PARP14-M2-F PARP14-M2-R	AGTAGAGAA <u>CC</u> TCTTTCTAAGTCCCTCTTGAAAAAGCTG AAGAGGTTCTCTACTAAGCACGAGATCCAAGGGAACGGAG	p.G1044E	ADPR-binding deficiency in PARP14 MD2
PARP14-M3-F PARP14-M3-R	AAAGCAGAA <u>GT</u> CTCCAAAGCAATTTTAGAATGTGCTGGAC GGAGACTTCTGCTTTGAGATTGAATGAGTTTGATGTTGA	p.G1257E	ADPR-binding deficiency in PARP14 MD3
PARP14-M4-F PARP14-M4-R	GGTGTGGCTAATGCTAAGACCGATGTTGTTGTCAACTCCGTT AGCATTAGCCACACCCCTCTTTCACCAACAGCATCTGCAGGCC	p.Q1024A	PARP14 MD2 binding site mutation
PARP14-M5-F PARP14-M5-R	GCTCCGAGTGCTCATGAAAATATTCAGGCATTTTCAGATGAA AGCACTCGGAGCCAGCAGAAAAGTGAACCTCTTGTAAGTTTT	p.H1172A, p.D1175A	PARP14 MD2 binding site mutation
PARP14-M6-F PARP14-M6-R	GCTCCGAGTTATCATGAAAATATTCAGGCATTTTCAGATGAA ATAACTCGGAGCCAGCAGAAAAGTGAACCTCTTGTAAGTTTT	p.H1172A, p.D1175Y	PARP14 MD2 binding site mutation
PARP14-M7-F2 PARP14-M7-R	CTGCTGGCTCCGAGTGATCATGAAAATATTCAGGCATTTTCA ACTCGGAGCCAGCAGAAAAGTGAACCTCTTGTAAGTTTCAG	p.H1172A	PARP14 MD2 binding site mutation
PARP14-M8-F PARP14-M8-R2	CCGAGTGCTCATGAAAATATTCAGGCATTTTCAGATGAATTT TTCATGAGCACTCGGGTGCAGCAGAAAAGTGAACCTCTTGTA	p.D1175A	PARP14 MD2 binding site mutation
PARP14-M9-F PARP14-M9-R2	CCGAGTTATCATGAAAATATTCAGGCATTTTCAGATGAATTT TTCATGATAACTCGGGTGCAGCAGAAAAGTGAACCTCTTGTA	p.D1175Y	PARP14 MD2 binding site mutation
PARP14-M10-F PARP14-M10-R	AAAGCTGA <u>ACC</u> AGAGCTCCAGGAGGAATTGGACACAGTTG CTCTGGTTCAGCTTTTTCCAAGAGGGACTTAGAAAAGAGGC	p.G1055E	PARP14 MD2 binding site mutation
PARP14-M11-F PARP14-M11-R	GTTGATGATTGTACCAAGCACCATTTCAATTAAACAACCTTC GGTACAATCATCAACAAATCTTATAGTATCTATGTGCTTT	p.D207A	YFP-PARP14 full-length mutagenesis

Note: The codons for mutation of the corresponding amino acid are underlined.

PARP14 MD2 ligation-independent cloning

Primer name	Primer sequence 5' → 3'	Comment
PARP14MD2-f001	TACTTCCAATCCATGAAAACATCATGGGAAAAAGGAAGCC	PARP14 MD2 truncation 1
PARP14MD2-f002	TACTTCCAATCCATGAAAAAGGAAGCCTGGTGTCCC	PARP14 MD2 truncation 2
PARP14MD2-f003	TACTTCCAATCCATGAAAGGAAGCCTGGTGTCCC	PARP14 MD2 truncation 3
PARP14MD2-r001	TATCCACCTTTACTGTCAATTAGCCCTTCTGGCAAATTCATC	for PARP14 MD2 truncations

PARP15 MD1 ligation-independent cloning

Primer name	Primer sequence 5' → 3'
PARP15A-f005	TACTTCCAATCCATGTCCAGAGACAACAAGTTCAGCA
PARP15A-f006	TACTTCCAATCCATGAACAAGTTCAGCAAGAAAGATTGTCTTTCA
PARP15A-f007	TACTTCCAATCCATGAGCAAGAAAGATTGTCTTTCAATCAGGAATG
PARP15A-f008	TACTTCCAATCCATGCTTTCAATCAGGAATGTTGTAGCTTCA
PARP15A-f009	TACTTCCAATCCATGGTTGTAGCTTCAATCCAAACCAAAGA
PARP15A-r006	TATCCACCTTTACTGCTGGCCTTGTGGGATTTATTCT
PARP15A-r007	TATCCACCTTTACTGTCAGGCCTTGTGGGATTTATTCTT
PARP15A-r008	TATCCACCTTTACTGTCATATTCTTGACCAGTTAGTGAATTCATCTAAAAATG

PARP14 gene editing using the CRISPR-Cas9 system

Guide ID	sgRNA (PARP14 target sequence)	On target locus	Target
PARP14_FW_01 / PARP14_N_b_FW_01	<i>CACCGCGGCGAGTGTGAGGTCCGCC</i>	Chr3: + 122680999	PARP14 transcript: ENST00000474629, exon 1
PARP14_RV_01 / PARP14_N_b_RV_01	<i>AAACGGCGGACCTCACAACCTCGCCGc</i>	Chr3: + 122680999	
PARP14_FW_02 / PARP14_N_a_FW_02	<i>CACCGTCAAGTCTTCGGGGGGTTCG</i>	Chr3: - 122680930	
PARP14_RV_02 / PARP14_N_a_RV_02	<i>AAACCGACCCCCCGAAGAACTTGAc</i>	Chr3: - 122680930	
PARP14_N_a_FW_01	<i>CACCGCTCCCGACCTCTTCGGGGCTC</i>	Chr3: - 122680975	
PARP14_N_a_RV_01	<i>AAACGAGCCCGAAGAGGTCGGGAGc</i>	Chr3: - 122680975	
PARP14_N_b_FW_02	<i>CACCGTGTACTTCCAGAGCCCGAAG</i>	Chr3: + 122680968	
PARP14_N_b_RV_02	<i>AAACCTTCGGGGCTCTGGAAGTACAc</i>	Chr3: + 122680968	

PARP14 editing analysis	Primer sequence 5' → 3'	Comment
PARP14_Ex1_Fwd1 PARP14_Ex1_Rev1 PARP14_Ex1_seq	GCGGGGGATGACTCTGCCATTCTGGGTCCGC TCACCCACCCCTCGCGCCCCCTACCGTCC GAAGGCAAAAGGAAACCAAG	PARP14 exon 1 gDNA amplification and sequencing
PARP14_mRNA_f PARP14_mRNA_r1 PARP14_mRNA_r2 PARP14_mRNA_r3	CTGCAGTCCGGCGGAGAG TTTGGTCTTGGATTCTTTTGTAGAAGTTC AATTCTTCTGACACATCTGGTTCTTTAACAT CAACAAATCTTATAGTATCTATGTGCTTTTGAAA	

Additional primers for screening and Sanger sequencing analysis

Primer name	Primer sequence 5' → 3'	Comment
PARP14-seq1 PARP14-seq2 PARP14-seq3 PARP14-seq4 PARP14-seq5 PARP14-seq6 PARP14-seq7-r	TGGAGAACATAAGTGGCCTG CCAATTAAAGTGGATCCAACAA GTAACTTCAGAAACCACAG GAAAGCTGCCCTACCACCAC TCACTCAAGATAATGGAAGA ACAACACCCAGATAAGGTTG TCTGGCAAATTCATCTGAA	for sequencing of PARP14 inserts in Gateway® vectors
M13 forward M13 reverse CMVF_pCDNA3 EGFP-C-F BGH Reverse	TGTAAAACGACGGCCAGT CAGGAAACAGCTATGACC CAACGGGACTTTCCAAAATG CATGGTCCTGCTGGAGTTCGTG TAGAAGGCACAGTCGAGG	
pLIC-F pLIC-R T7F T7R	TGTGAGCGGATAACAATTCC AGCAGCCAACCTCAGCTTCC TAATACGACTCACTATAGGG GCTAGTTATTGCTCAGCGG	for screening clones and sequencing of prokaryotic expression constructs with LIC and T7 promoter site
U6-Fwd	GAGGGCCTATTTCCCATGATTCC	
		sequencing sgRNA plasmids

siRNAs

siRNA ID	PARP14 targeting sequence 5' → 3'	Reference / Supplier
siMOCK	<i>n/a</i>	AllStars negative control siRNA (Qiagen)
siPARP14 #1	AGGCCGACTGTGACCAGATAGTGAA	reference: Nicolae <i>et al.</i> (Nucleic Acids Res., 2015) ¹⁰⁹ custom synthesis purchased from Dharmacon®
siPARP14 #2	CGGCACTACACAGTGAACCTGAACA	reference: Nicolae <i>et al.</i> (Nucleic Acids Res., 2015) ¹⁰⁹ custom synthesis purchased from Dharmacon®
siPARP14 #3	GGAAAGGGCUCACUCACAA	reference: Vyas <i>et al.</i> (Nat. Commun., 2013) ¹⁰⁸ purchased from Dharmacon®

3. Plasmids**Gateway® vectors and Gateway® cloning-derived plasmids**

Plasmid ID	Description (* = mutation)	Protein start	Protein end	Vector	Mutation(s)
PARP14A-c060	YFP-tagged MD2 wt with NLS	Ala994	Val1195	pDEST-N-YFP/FRT/TO	---
PARP14A-c061	YFP-tagged MD2 M1 with NLS	Ala994	Val1195	pDEST-N-YFP/FRT/TO	p.G1044E
PARP14A-c062	YFP-tagged MD2 M3 with NLS	Ala994	Val1195	pDEST-N-YFP/FRT/TO	p.G1044E, p.H1172A, p.D1175Y
PARP14A-c063	YFP-tagged MD1-3 wt no NLS	Gly789	Gly1388	pDEST-N-YFP/FRT/TO	---
PARP14A-c064	YFP-tagged MD1-3 M1 no NLS	Gly789	Gly1388	pDEST-N-YFP/FRT/TO	p.G1044E
PARP14A-c065	YFP-tagged MD1-3 M3 no NLS	Gly789	Gly1388	pDEST-N-YFP/FRT/TO	p.G1044E, p.H1172A, p.D1175Y
PARP14A-c066	YFP-tagged MD1-3 wt with NLS	Gly789	Gly1388	pDEST-N-YFP/FRT/TO	---
PARP14A-c067	YFP-tagged MD1-3 M1 with NLS	Gly789	Gly1388	pDEST-N-YFP/FRT/TO	p.G1044E
PARP14A-c068	YFP-tagged MD1-3 wt with spacer and NLS	Gly789	Ser1393	pDEST-N-YFP/FRT/TO	---
PARP14A-c069	YFP-tagged MD1-3 *MD1 with spacer and NLS	Gly789	Ser1393	pDEST-N-YFP/FRT/TO	p.G832E
PARP14A-c070	YFP-tagged MD1-3 *MD2 with spacer and NLS	Gly789	Ser1393	pDEST-N-YFP/FRT/TO	p.G1044E
PARP14A-c071	YFP-tagged MD1-3 *MD3 with spacer and NLS	Gly789	Ser1393	pDEST-N-YFP/FRT/TO	p.G1257E
PARP14A-c072	YFP-tagged MD1-3 *MD12 with spacer and NLS	Gly789	Ser1393	pDEST-N-YFP/FRT/TO	p.G832E, p.G1044E
PARP14A-c073	YFP-tagged MD1-3 *MD13 with spacer and NLS	Gly789	Ser1393	pDEST-N-YFP/FRT/TO	p.G832E, p.G1257E
PARP14A-c074	YFP-tagged MD1-3 *MD23 with spacer and NLS	Gly789	Ser1393	pDEST-N-YFP/FRT/TO	p.G1044E, p.G1257E
PARP14A-c075	YFP-tagged MD1-3 *MD123 with spacer and NLS	Gly789	Ser1393	pDEST-N-YFP/FRT/TO	p.G832E, p.G1044E, p.G1257E

PARP14A-c076	YFP-tagged macroPARP	Gly789	Lys1801	pDEST-N-YFP/FRT/TO	---
PARP14A-c077	YFP-tagged full-length PARP14	Met1	Lys1801	pDEST-N-YFP/FRT/TO	---

Note:

All constructs PARP14A-c060 to c077 were created by LR reaction of the pDEST-N-YFP/FRT/TO vector with the corresponding pDONRTM221 entry vectors.

Empty pDONR and destination vectors

Plasmid ID	Description
pDONR TM 221	Gateway [®] donor vector (Invitrogen, cat.no. 12536-017) for creating pENTR clones by recombination of <i>attB</i> site containing PCR products or vectors using BP Clonase TM enzyme mix
pDEST-N-YFP/FRT/TO	Gateway [®] destination vector for creating mammalian expression constructs for N-terminal YFP-tagged target proteins provided by the Dr. Ivan Ahel's Laboratory (Sir William Dunn School of Pathology, Oxford); reference: Ahel <i>et al.</i> (Science, 2009) ⁵⁸

Plasmids for PARP14 gene editing using the CRISPR-Cas9 system

Construct ID	sgRNA ID (fwd, 5' → 3')	Cas9 type	Vector
pSpCas9(BB)-2A-Puro-PARP14-c002	PARP14_FW_01	wild-type	pSpCas9(BB)-2A-Puro
pSpCas9(BB)-2A-Puro-PARP14-c004	PARP14_FW_02	wild-type	pSpCas9(BB)-2A-Puro
pSpCas9n(BB)-2A-Puro-PARP14-c001	PARP14_N_a_FW_01	nickase	pSpCas9n(BB)-2A-Puro
pSpCas9n(BB)-2A-Puro-PARP14-c003	PARP14_N_b_FW_01	nickase	pSpCas9n(BB)-2A-Puro
pSpCas9n(BB)-2A-Puro-PARP14-c005	PARP14_N_a_FW_02	nickase	pSpCas9n(BB)-2A-Puro
pSpCas9n(BB)-2A-Puro-PARP14-c006	PARP14_N_b_FW_02	nickase	pSpCas9n(BB)-2A-Puro

Control plasmids and empty vectors

Plasmid ID	Description/ sgRNA sequence (fwd, 5' → 3')	Cas9 type	Comment
(7583) pSpCas9(BB)2APuro (PX459)	original empty plasmid from addgene (#48139)	wild-type	for insertion of PARP14 sgRNAs
(7553) pSpCas9n(BB)2APuro (PX462) v2.0	original empty plasmid from addgene (#62987)	nickase	for insertion of PARP14 sgRNAs
(7582) pSpCas9(BB)-2A-GFP (PX458)	original empty plasmid from addgene (#48138)	wild-type	transfection control
pSpCas9(BB)-2A-Puro-GFP-C12	CACCGCTTGTCACACTTTTCGGTTA	wild-type	control plasmid, targets GFP
pSpCas9(BB)n-2A-Puro-GFP-C17	CACCGCCATCTAATTCAACAAGAAT	nickase	control plasmid, targets GFP, for co-transfection with pSpCas9(BB)n-2A-Puro-GFP-C18
pSpCas9(BB)n-2A-Puro-GFP-C18	CACCGCCATCTAATTCAACAAGAAT	nickase	control plasmid, targets GFP, for co-transfection with pSpCas9(BB)n-2A-Puro-GFP-C17

Prokaryotic and insect cell expression plasmids**Vectors**

Vector name	Gen Bank ID	Description
pNIC28-Bsa4	EF198106	pET expression vector with N-terminal His ₆ -tag, TEV protease cleavage site, kanamycin resistance
pNIC-Bio3	JN792439	pET expression vector with N-terminal His ₆ -tag, TEV protease cleavage site, biotinylation sequence in C-terminal fusion, kanamycin resistance
pNIC-ZB	GU452710	pET expression vector with N-terminal His ₆ - and Z-basic (ZB) tags, TEV protease cleavage site, kanamycin resistance
pNH-TrxT	GU269914	pET expression vector with N-terminal His ₆ -Trx fusion peptide, TEV protease cleavage site, kanamycin resistance
pNIC-CTHF	EF199844	pET expression vector with C-terminal His ₆ - and FLAG-tag, TEV protease cleavage site, kanamycin resistance
pFB-GST3C-LIC	n/a, SGC vector	baculovirus transfer vector with N-terminal GST-tag, 3C protease cleavage site, ampicillin resistance
pFB-GST3C-Bio5	n/a, SGC vector	baculovirus transfer vector with N-terminal GST-tag, 3C protease cleavage site, biotinylation sequence in C-terminal fusion, BirA expression under p10 promoter control for <i>in vivo</i> biotinylation, ampicillin resistance

Note: Vector maps are provided in supplementary experimental details, section V.4.1.

PARP14 MD2 plasmids

Plasmid ID	Description	Protein start	Protein end	Vector	Mutation(s)
PARP14A-c030	MD2 used for screening and as template for mutagenesis	Ala994	Asn1191	pNIC28-Bsa4	---
PARP14A-c002	MD2 for biotinylation	Ala994	Asn1191	pNIC-Bio3	---
PARP14A-c011	SER mutant of PARP14A-c030 (cloned by K. Uth) ¹⁶²	Ala994	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1106S, p.K1154S, p.K1158S, p.K1162S
PARP14A-c012	SER mutant of PARP14A-c030 (cloned by K. Uth) ¹⁶²	Ala994	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1106S mutation of stop codon
PARP14A-c013	SER mutant of PARP14A-c030 (cloned by K. Uth) ¹⁶²	Ala994	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1154S, p.K1158S, p.K1162S
PARP14A-c014	SER mutant of PARP14A-c030 (cloned by K. Uth) ¹⁶²	Ala994	Asn1191	pNIC28-Bsa4	p.K1106S, p.K1154S, p.K1158S, p.K1162S
PARP14A-c015	SER mutant of PARP14A-c030 (cloned by K. Uth) ¹⁶²	Ala994	Asn1191	pNIC28-Bsa4	p.K1048S
PARP14A-c034	truncation 1 of PARP14A-c011	Lys1000	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1106S, p.K1154S, p.K1158S, p.K1162S
PARP14A-c035	truncation 2 of PARP14A-c011	Glu1004	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1106S, p.K1154S, p.K1158S, p.K1162S
PARP14A-c036	truncation 3 of PARP14A-c011	Lys1005	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1106S, p.K1154S, p.K1158S, p.K1162S

PARP14A-c037	truncation 1 of PARP14A-c012	Lys1000	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1106S
PARP14A-c038	truncation 2 of PARP14A-c012	Glu1004	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1106S
PARP14A-c039	truncation 3 of PARP14A-c012	Lys1005	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1106S
PARP14A-c040	truncation 1 of PARP14A-c013	Lys1000	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1154S, p.K1158S, p.K1162S
PARP14A-c041	truncation 2 of PARP14A-c013	Glu1004	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1154S, p.K1158S, p.K1162S
PARP14A-c042	truncation 3 of PARP14A-c013	Lys1005	Asn1191	pNIC28-Bsa4	p.K1048S, p.K1154S, p.K1158S, p.K1162S
PARP14A-c043	truncation 1 of PARP14A-c014	Lys1000	Asn1191	pNIC28-Bsa4	p.K1106S, p.K1154S, p.K1158S, p.K1162S
PARP14A-c044	truncation 2 of PARP14A-c014	Glu1004	Asn1191	pNIC28-Bsa4	p.K1106S, p.K1154S, p.K1158S, p.K1162S
PARP14A-c045	truncation 3 of PARP14A-c014	Lys1005	Asn1191	pNIC28-Bsa4	p.K1106S, p.K1154S, p.K1158S, p.K1162S
PARP14A-c046	truncation 1 of PARP14A-c015	Lys1000	Asn1191	pNIC28-Bsa4	p.K1048S
PARP14A-c047	truncation 2 of PARP14A-c015	Glu1004	Asn1191	pNIC28-Bsa4	p.K1048S
PARP14A-c048	truncation 3 of PARP14A-c015	Lys1005	Asn1191	pNIC28-Bsa4	p.K1048S
PARP14A-c080	mutant 1	Ala994	Asn1191	pNIC28-Bsa4	p.Q1024A
PARP14A-c081	mutant 2	Ala994	Asn1191	pNIC28-Bsa4	p.H1172A
PARP14A-c082	mutant 3	Ala994	Asn1191	pNIC28-Bsa4	p.D1175A
PARP14A-c083	mutant 4	Ala994	Asn1191	pNIC28-Bsa4	p.D1175Y
PARP14A-c084	mutant 5	Ala994	Asn1191	pNIC28-Bsa4	p.G1044E
PARP14A-c085	mutant 6	Ala994	Asn1191	pNIC28-Bsa4	p.H1172A, p.D1175A
PARP14A-c086	mutant 7	Ala994	Asn1191	pNIC28-Bsa4	p.Q1024A, p.D1175Y
PARP14A-c087	mutant 8	Ala994	Asn1191	pNIC28-Bsa4	p.Q1024A, p.H1172A, p.D1175A
PARP14A-c088	mutant 9	Ala994	Asn1191	pNIC28-Bsa4	p.Q1024A, p.H1172A, p.D1175Y
PARP14A-c089	mutant 10	Ala994	Asn1191	pNIC28-Bsa4	p.H1172A, p.D1175Y, p.G1044E
PARP14A-c090	mutant 11	Ala994	Asn1191	pNIC28-Bsa4	p.G1155E

PARP15 MD1 plasmids

Plasmid ID	Description	Protein start	Protein end	Vector	Mutation(s)
PARP15A-c003	PARP15 MD1 bacterial expression	M62	A279	pNIC28-Bsa4	---
PARP15A-c004	PARP15 MD1 bacterial expression	M62	A274	pNIC28-Bsa4	---
PARP15A-c005	PARP15 MD1 bacterial expression	M62	I269	pNIC28-Bsa4	---
PARP15A-c006	PARP15 MD1 bacterial expression	N66	A279	pNIC28-Bsa4	---

PARP15A-c007	PARP15 MD1 bacterial expression	N66	A274	pNIC28-Bsa4	---
PARP15A-c008	PARP15 MD1 bacterial expression	N66	I269	pNIC28-Bsa4	---
PARP15A-c009	PARP15 MD1 bacterial expression	S69	A279	pNIC28-Bsa4	---
PARP15A-c010	PARP15 MD1 bacterial expression	S69	A274	pNIC28-Bsa4	---
PARP15A-c011	PARP15 MD1 bacterial expression	S69	I269	pNIC28-Bsa4	---
PARP15A-c012	PARP15 MD1 bacterial expression	L74	A279	pNIC28-Bsa4	---
PARP15A-c013	PARP15 MD1 bacterial expression	L74	A274	pNIC28-Bsa4	---
PARP15A-c014	PARP15 MD1 bacterial expression	L74	I269	pNIC28-Bsa4	---
PARP15A-c015	PARP15 MD1 bacterial expression	V79	A279	pNIC28-Bsa4	---
PARP15A-c016	PARP15 MD1 bacterial expression	V79	A274	pNIC28-Bsa4	---
PARP15A-c017	PARP15 MD1 bacterial expression	V79	I269	pNIC28-Bsa4	---
PARP15A-c024	PARP15 MD1 sf9 expression	Met62	Ala274	pFB-GST3C-LIC	---
PARP15A-c039	PARP15 MD1 sf9 expression	Met62	Ala274	pFB-GST3C-Bio5	---

Other macrodomain encoding plasmids used for selectivity studies

Plasmid ID	Description	Protein start	Protein end	Vector	Mutation(s)
PARP14A-c028	PARP14 MD1	Gly789	Lys979	pNIC28-Bsa4	---
PARP14A-c032	PARP14 MD3	Phe1208	Gly1388	pNIC28-Bsa4	---
PARP15A-c002	PARP15 MD2, purified by C. Sieg ¹⁶³	Gly287	Asn470	pNIC-Bio3	---
PARP9B-c001	PARP9 MD1	Gly102	Lys298	pNIC28-Bsa4	---
PARP9B-c005	PARP9 MD2, purified by C. Sieg ¹⁶³	Asn310	Asn493	pNIC28-Bsa4	---
MACROD1A-c010	MacroD1	Ser58	Ala325	pNH-TrxT	---
MACROD2A-c121	MacroD2	Thr69	Asn237	pNIC-CTHF	---
GDAP2A-c004	MacroD3, purified by C. Sieg ¹⁶³	Met1	Lys228	pNIC28-Bsa4	---
H2AFYA-c001	macroH2A1.1, purified by C. Sieg ¹⁶³	Thr178	Asn372	pNIC28-Bsa4	---
H2AFYB-c005	macroH2A1.2, purified by C. Sieg ¹⁶³	Gly177	Asn372	pNIC-Bio3	---
H2AFY2A-c020	macroH2A2	Lys167	Lys372	pNH-TrxT	---
LOC221443A-c000	TARG1	Met1	Leu152	pNIC28-Bsa4	---
CHD1LA-c023	ALC1	Lys630	Gln894	pNIC28-Bsa4	---
PARGA-c004	PARG	Ser448	Thr976	pNIC-ZB	---

WWE domain and PARP5 encoding plasmids

Plasmid ID	Description	Protein start	Protein end	Vector	Mutation(s)
PARP14A-c026	WWE of PARP14	Asp1503	Asp1611	pNIC28-Bsa4	---
RNF146A-c002	WWE of RNF146	Pro90	Asn190	pNIC28-Bsa4	---
PARP5A cat.	catalytic domain of PARP5	Gly1105	Thr1327	pNIC28-Bsa4	---

4. Antibodies

Target	Antibody ID/Ref. (company)	Description
α -(alpha) tubulin	EP1332Y (Abcam)	monoclonal rabbit IgG1 antibody, used 1:1000 in 3 % FBS/PBS (IF)
α -(alpha) tubulin	TAT1 (Cancer Research UK Monoclonal Antibody Service)	monoclonal mouse IgG1 antibody, used 1:15000 in 2 % BSA/TBS-T (WB)
α -cleaved PARP1	#9541 (Cell Signalling)	polyclonal rabbit IgG1 antibody, used 1:2000 in 2 % BSA/TBS-T (WB)
α -GAPDH	14C10 (Cell Signalling)	monoclonal rabbit IgG1 antibody, used 1:10000 in 5 % milk/TBS-T (WB)
α -GFP	B-2 (Santa Cruz)	monoclonal mouse IgG2a antibody, used 1:500 in 5 % milk/TBS-T (WB)
α - γ H2Ax (Ser139)	JBW301 (Millipore)	monoclonal mouse IgG1 antibody, used 1:2000 in 2 % BSA/TBS-T (WB) and 1:2000 in ADB (IF)
α -histone H3	865R2 (Invitrogen)	monoclonal mouse IgM antibody, used 1:1000 in 2 % BSA/TBS-T (WB)
α -KAP	A300-274A (Bethyl Lab.)	polyclonal mouse IgG1 antibody, used 1:5000 in 2 % BSA/TBS-T (WB)
α -pKAP1 (Ser824)	A300-767A (Bethyl Lab.)	polyclonal mouse IgG1 antibody, used 1:1000 in 2 % BSA/TBS-T (WB)
α -PARP14	C-1 (Santa Cruz Biotechnology)	monoclonal mouse IgG1 antibody, used 1:750 in 5 % milk/TBS-T (WB)
α -53BP1	NB100-304 (Novus)	polyclonal mouse IgG1 antibody, used 1:5000 in ADB (IF)
α -rabbit (HRP-conjugated)	P0447 (Dako)	secondary antibody, polyclonal goat anti-mouse IgG (HRP), used 1:5000 in 5 % milk/TBS-T (WB)
α -mouse (HRP-conjugated)	P0448 (Dako)	secondary antibody, polyclonal goat anti-rabbit IgG (HRP), used 1:5000 in 5 % milk/TBS-T (WB)
α -rabbit (fluorescent dye-labelled)	#A11034 (Invitrogen)	secondary antibody, Alexa Fluor [®] 488 polyclonal goat anti-mouse IgG, used 1:400 in 3 % FBS/PBS or ADB (IF)
α -rabbit (fluorescent dye-labelled)	#A21109 (Invitrogen)	secondary antibody, Alexa Fluor [®] 680 polyclonal goat anti-rabbit IgG, used 1:10000 in 5 % milk/TBS-T (WB)
α -mouse (fluorescent dye-labelled)	#A11003 (Invitrogen)	secondary antibody, Alexa Fluor [®] 546 polyclonal goat anti-mouse IgG, used 1:400 in 3 % FBS/PBS or ADB (IF)
α -mouse (fluorescent dye-labelled)	#A21037 (Invitrogen)	secondary antibody, Alexa Fluor [®] 750 polyclonal goat anti-mouse IgG, used 1:10000 in 5 % milk/TBS-T (WB)

Note: ADB = antibody dilution buffer containing 1 % goat-serum, 0.3 % BSA and 0.005 % Triton X 100 in PBS

5. Work with nucleic acids

5.1. DNA preparation

DNA from transformed bacteria was extracted using the GenElute™ Plasmid Miniprep Kit (Sigma) for small scale preparations and using the QIAfilter Plasmid Midi Kit (Qiagen) or QIAfilter Plasmid Maxi Kit (Qiagen) for larger scale preparations following the manufacturer's instructions. The GeneElute™ Gel Extraction Kit (Sigma) was used for DNA extraction and purification from agarose gels after gel electrophoresis. Genomic DNA from mammalian cells was extracted using the DNeasy® Blood & Tissue Kit (Qiagen) according to the manufacturer's instructions. Extracted DNA was stored at -20 °C.

5.2. mRNA preparation

Prior to RNA extraction, all surfaces and materials used were cleaned with 70 % EtOH and RNaseZap® RNase Decontamination Solution (Ambion). RNA was then extracted under sterile conditions using the PureLink® RNA MiniKit (Qiagen) following the manufacturer's instructions. Extracted RNA was stored at -80 °C.

5.3. Agarose gel electrophoresis

Gel electrophoresis of DNA was performed with 0.8-2 % agarose gels (concentration depending on application) using Iberose High Specification for Electrophoresis (WebScientific) stained with pegGREEN non-toxic DNA/RNA dye (peqlab) and run at 100 V for 30-45 min in TAE buffer (40 mM Tris-HCl (pH 8.6), 20 mM acetate, 1 mM EDTA). 10x BlueJuice™ Gel Loading Buffer (Invitrogen) was used at a concentration of 2x as sample loading buffer (were applicable) and the 1 Kb Plus DNA Ladder (Invitrogen) was used for sizing the DNA fragments.

5.4. Cloning

5.4.1. Gateway[®] cloning of mammalian expression constructs

Gateway[®] cloning was performed according to the manufacturer's instructions with minor modifications²⁵¹. For generation of pENTR clones, the corresponding PARP14 templates were diluted to < 1 ng/μL and the regions of interest were PCR amplified with the corresponding primer set (diluted to 5 μM) using Herculase II Fusion DNA Polymerase (Agilent). The resulting PCR products were purified using the QIAquick PCR Purification Kit (Qiagen) and set-up for the BP recombination reaction. For this, 1 μL of pDONR221[™] (~300 ng/reaction) was incubated with 1 μL of PCR product (~150 ng), 1 μL BP Clonase[™] enzyme mix (Invitrogen) and 2 μL H₂O for 3 hours or O/N at RT. For full-length PARP14 cloning, the PCR amplification step was omitted and instead 1 μL of the PARP14 mammalian expression vector (~300 ng/reaction) were directly set-up for the BP recombination reaction. In both cases, the reaction was stopped by adding 1 μL Proteinase K (Invitrogen) and after incubation for 10 min at 37 °C competent Stable *E. coli* (NEB) were transformed with 2 μL of the BP reaction mix. The following day, colonies were screened with MyTaq[™] DNA polymerase (Bioline) with M13F/R primers and positive clones propagated. Plasmid DNA was isolated and inserts analysed by Sanger sequencing. For generation of the YFP expression clones, 1 μL of the corresponding pENTR clone (100 ng/reaction) was incubated with 1 μL of YFP-destination vector (100 ng/reaction), 1 μL LR Clonase[™] enzyme mix (Invitrogen) and 2 μL H₂O for 2 hours or O/N at RT. Following the reaction, the same procedure as after the BP reaction was used, however, performing the colony PCR screen with the EGFP-C-F and BGH Reverse primer set.

5.4.2. Ligation-independent cloning for prokaryotic expression constructs

For generation of prokaryotic expression constructs, a protocol developed by the SGC Biotechnology team was followed. In brief, the DNA regions of interest were PCR amplified from the respective DNA templates with the corresponding primers using Herculase II Fusion DNA Polymerase (Agilent) by touch-down PCR (PARP15 MD1 construct) or conventional 3-step PCR (PARP14 MD2 truncation constructs). The resulting PCR products were DpnI (NEB) digested for 1 h at 37 °C, purified using the QIAquick PCR purification Kit (Qiagen) and then treated with T4 DNA polymerase (NEB) in presence of dCTP (NEB). The desired expression vectors were prepared for ligation by BsaI (NEB) digestion, followed by purification and treatment with T4 DNA polymerase (NEB) in presence of dGTP (NEB). Vector and PCR product treatments

with T4 DNA polymerase were carried out in a thermocycler, incubating for 30 min at 22 °C and then for 20 min at 75 °C. For ligation, the T4 treated vector was mixed with T4 treated insert in a 1:2 (v/v) ratio and incubated for 30 min at RT. The reaction mixture was directly transformed into chemically competent Mach1 *E. coli* cells by the heat-shock method. Positive transformed clones were then selected O/N on LB agar plates containing 5 % sucrose and 50 µg/mL kanamycin and screened for correct target DNA insertion by colony PCR using MyTaq™ DNA polymerase (Bioline) and the pLIC-F and pLIC-R primer set.

5.4.3. Site-directed mutagenesis

For introducing point mutations into expression plasmids or pENTR clones, the respective expression constructs were PCR amplified with the corresponding mutagenesis primers using Herculase II Fusion DNA Polymerase (Agilent). The samples obtained after the PCR reaction were digested with DpnI (NEB) at 37 °C, O/N and transformed into competent Stable *E. coli* (NEB) (pENTR clones only) or into Mach1 competent *E. coli*. Plasmids were prepared and correct amplification of the whole target sequence was confirmed by Sanger sequencing.

5.4.4. Cloning of sgRNAs into vector for co-expression with Cas9

The cloning procedure developed by Ran *et al.* (2013)²¹² was followed with minor modifications. In brief, top and bottom sgRNAs were diluted to 100 µM and then phosphorylated and annealed with T4 Polynucleotide Kinase (NEB) (37 °C, 30 min; 95 °C, 5 min; cooling to 25 °C at 5 °C min⁻¹). The phosphorylated and annealed sgRNAs were diluted 1:200 and ligated into pSpCas9(BB)-2A-puro or pSpCas9n(BB)-2A-puro, respectively, with FastDigest BbsI (NEB) and T4 ligase (NEB) in Tango buffer (incubation: 6 cycles with each 37 °C for 5 min and then 21 °C for 5 min). Competent Stable *E. coli* (NEB) were transformed with the products and selected on LB plates containing 100 µg/mL ampicillin. Plasmids were isolated and checked for sgRNA insertion by BbsI digestion. Insert verified plasmids were further analysed by Sanger sequencing from the U6 promoter using the U6-Fwd primer.

6. Work with eukaryotic cells (cell culture)

6.1. Eukaryotic cell lines

DMSO stocks of (parental) eukaryotic cell lines were provided by the SGC Chemical Biology group and were confirmed mycoplasma-free before usage for CRISPR editing and in cellular assays.

Cell line	ATCC ID	Description
HEK293	ATCC [®] CRL-1573 [™]	human epithelial cell line, origin: embryonic kidney, adherent
HeLa	ATCC [®] CCL-2 [™]	human epithelial cell line, origin: cervix adenocarcinoma, adherent
U-2 OS	ATCC [®] HTB-96 [™]	human epithelial cell line, origin: bone (osteosarcoma), adherent
Jurkat	ATCC [®] TIB-152 [™]	human T lymphocyte cell line, origin: peripheral blood (acute T cell leukemia), suspension
HeLa wild-type (clone E8)	<i>n/a</i>	clonal wild-type cell line derived from parental HeLa cell line after transfection with GFP targeting sgRNAs pSpCas9(BB)-2A-Puro-GFP-C18 and -C17
HeLa PARP14 KO (clone 21)	<i>n/a</i>	clonal PARP14 KO cell line derived from parental HeLa cell line after transfection with PARP14 targeting sgRNAs pSpCas9n(BB)-2A-Puro-PARP14-c001 and -c003
HeLa PARP14 KO (clone 50)	<i>n/a</i>	clonal PARP14 KO cell line derived from parental HeLa cell line after transfection with PARP14 targeting sgRNAs pSpCas9n(BB)-2A-Puro-PARP14-c001 and -c003
U-2 OS wild-type (clone E9)	<i>n/a</i>	clonal wild-type cell line derived from parental U-2 OS cell line after transfection with GFP targeting sgRNA pSpCas9(BB)-2A-Puro-GFP-C12
U-2 OS PARP14 KO (clone 23)	<i>n/a</i>	clonal PARP14 KO cell line derived from parental U-2 OS cell line after transfection with PARP14 targeting sgRNAs pSpCas9n(BB)-2A-Puro-PARP14-c001 and -c003
U-2 OS PARP14 KO (clone 65)	<i>n/a</i>	clonal PARP14 KO cell line derived from parental U-2 OS cell line after transfection with PARP14 targeting sgRNA pSpCas9(BB)-2A-Puro-PARP14-c002

6.2. General cell culture techniques

Culture conditions and method

Unless stated otherwise, HEK293, HeLa and U-2 OS cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Sigma) and Jurkat cells were cultured in RPMI-1640 media (Sigma) supplemented with 2 mM *L*-glutamine (Sigma), 1 % penicillin/streptomycin (Jurkat cells only) (Sigma) and 10 % (v/v) FBS (heat inactivated fetal bovine serum) (Gibco) under humidified atmosphere of 5 % CO₂ at 37 °C. Cells were typically passaged when reaching 80-90 % confluency or as required. For adherent cells, growth medium was removed, cells were briefly rinsed with Dulbecco's Phosphate Buffered Saline (Sigma) and then detached using 0.5 % trypsin/EDTA (Sigma) at 37 °C, 2-4 min. Trypsin was then inactivated by addition of fresh FBS containing growth medium. The cell suspension was transferred into a 15 mL Falcon tube and cells were spun down with 300×g, 3 min. Finally, the trypsin containing medium was aspirated, cells re-suspended in fresh growth medium and transferred into new the vessel, for general maintenance T-75 flasks (Corning®).

Cryo-conservation and thawing of cells

For cryo-conservation, $2-4 \times 10^6$ cells grown to log phase were re-suspended in 1.2 mL of growth medium containing 10 % FBS and 10 % (v/v) DMSO, transferred to cryo-vials (Greiner bio-one) and slowly cooled down over 24-48 hours to -80°C using a Mr. Frosty™ Freezing Container (ThermoFisher). Afterwards, cryo-vials were transferred to the -170 °C liquid nitrogen tank for long-term storage.

For recovering cells from frozen stocks, cryo-vials were quickly thawed in a water bath at 37 °C. After thawing, cells were instantly but slowly (~1 mL/min) transferred into 10 mL of pre-warmed full growth medium and then spun down (300×g, 3 min) for DMSO removal. Finally, cells were re-suspended in full growth medium and transferred into an appropriate cell culture flask.

Preparation of cell lysates

Cell lysates for Western blot analysis were prepared as follows: After generating the cell pellet by trypsination and centrifugation (300×g, 3 min), cell pellets were washed once with PBS and then either stored at -80 °C until further processing or were directly lysed in RIPA buffer

(50 mM Tris-HCl (pH 8.0), 150 mM NaCl, 1 % NP-40, 0.5 % Deoxycholate and 0.1 % SDS) containing cOmplete™ protease inhibitor cocktail (Roche) and PhosSTOP™ phosphatase inhibitor cocktail (Roche) for 30 min on ice. Cell debris was removed by centrifugation (10,000×g, 30 min, 4 °C) and the resulting supernatants were transferred to fresh Eppendorf tubes. Protein concentrations of the lysates were then determined using the Pierce™ BCA protein assay kit (ThermoFisher) according to the manufacturer's instructions. Following this assay, samples were either used immediately for further analysis or stored at -20 °C.

6.3. Transient transfection

6.3.1. siRNA transfection

Forward siRNA transfections were carried out with Lipofectamine® RNAiMAX reagent (Invitrogen) according to the manufacturer's instructions. In brief, PARP14 siRNA stock solutions of 50 µM were prepared. One day before transfection, cells were plated aiming for ~60 % confluency at time of transfection. siRNA-Lipofectamine RNAiMAX complexes were prepared in Opti-MEM® I Reduced Serum Medium (Gibco) and added to the cells with a final siRNA concentration of 20 nM per well. AllStars negative control siRNA ('siMOCK') (Qiagen) was used as siRNA transfection control.

6.3.2. Transfection of mammalian expression constructs

Transfections for test expression and imaging of PARP14 mammalian expression constructs were carried out using Lipofectamine® 3000 reagent (Invitrogen) according to the manufacturer's instructions. The day before transfection, cells were plated aiming for ~80 % confluency at time of transfection. DNA-Lipofectamine complexes were prepared in Opti-MEM® I Reduced Serum Medium (Gibco). Test expressions were performed in 6-well plates using 2500 ng DNA of each construct per well.

6.4. Generation of PARP14 KO cell lines using the CRISPR-Cas9 system

6.4.1. PARP14 sgRNA design for CRISPR genome editing

PARP14 sgRNAs for CRISPR genome editing were designed using the crispr.mit.edu online tool²¹² for exon 1 (bp region 63-246) of the PARP14 transcript ENST00000474629. Suitable sgRNAs for Cas9 wild-type and nickase constructs were selected based on the quality and off-target score. All sgRNAs selected had a high quality score (75-90) with more than three mismatches between the target sequence and the corresponding off-target.

6.4.2. Functional validation of PARP14 sgRNAs in HEK293 cells

Sufficient quantities of all PARP14 editing CRISPR plasmids were prepared and the day before transfection, HEK293 cells were plated in a 12-well plate at 200,000 cells/well. Following incubation overnight, confluency of 70 % was confirmed and transfection was carried out using Lipofectamine[®] 3000 (Invitrogen) and according to the manufacturer's instructions. The amount of DNA per well used for transfection was kept constant at 1000 ng/well (500 ng of plasmid for co-transfection of two nickase constructs) as well as of Lipofectamine[®] 3000 of 3 μ L and P3000 of 2 μ L. After 24 hours, GFP expression of the control plasmid confirmed transfection of over 90 %. 72 hours after transfection, the transfected HEK293 cells were harvested and genomic DNA was extracted. Exon 1 of PARP14 was PCR amplified using Herculase II Fusion DNA polymerase (Agilent) and the corresponding primer set (see section IV.2.) with prior optimised settings (95 °C, 10 min; 35 $\times$ (95 °C, 30 sec; 72 °C, 1.5 min); 72 °C, 5 min) and resulting PCR products purified using a QIAquick PCR purification Kit (Qiagen). 300 ng of each PCR product were re-annealed (gradient: 95 - 85 °C with -2 °C/sec, then 85 - 25 °C with 0.1 °C/sec), then digested with T7 Endonuclease I (NEB) and analysed on a 1 % agarose gel.

6.4.3. Transfection, selection and clonal expansion of HeLa and U-2 OS cells

The detailed procedures including transfection with PARP14 sgRNAs, selection and clonal isolation and expansion of HeLa and U-2 OS cells for PARP14 KO cell line generation are provided in the supplementary experimental details, section V.4.2.3.

6.4.4. Validation of PARP14 edits by gDNA analysis

Genomic DNA of U-2 OS and HeLa wild-type and PARP14 edited clones was extracted and exon 1 of PARP14 PCR amplified with the corresponding primer set (see section IV.2.) using Herculase II Fusion DNA Polymerase (Agilent) (95 °C, 10 min; 35 × (95 °C, 30 sec; 72 °C, 1.5 min); 72 °C, 5 min). The resulting PCR products were purified using the QIAquick PCR purification kit (Qiagen) and diluted to 1 ng/μl per 100 bp for Sanger sequencing analysis.

6.4.5. Validation of PARP14 edits by mRNA analysis

Total RNA of U-2 OS and HeLa wild-type and PARP14 edited clones was extracted and after cell lysis, 25-30 of 1.0 mm diameter Zirconia Beads (BioSpec) were added to the samples to homogenise the lysates in a bead beater (2×5 sec). Good mRNA quality was confirmed with a 260/280 absorbance measurement ratio of ~2.0. On the same day, mRNA was converted into cDNA using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) following the manufacturer's instructions. For cDNA analysis, PARP14 cDNA was PCR amplified with the corresponding primer set (see section IV.2.) using Herculase II Fusion DNA polymerase (Agilent) (95 °C, 10 min; 35 × (95 °C, 30 sec; 62 °C, 2 min; 72 °C, 1.0 min); 72 °C, 5 min). PCR products were analysed on 1 % agarose gel. Due to a co-amplified non-specific band, the PARP14 PCR products were recovered from the gel and PCR re-amplified. The resulting PCR products were purified with a QIAquick PCR purification kit (Quiagen) and diluted to 1 ng/μl per 100 bp for Sanger sequencing analysis.

7. Work with prokaryotic cells

7.1. Bacteria strains

Strain	Description
<i>E. coli</i> Rosetta™ BL21(DE3)-R3-pRARE2 (SGC)	with pRARE2 plasmid encoding rare codon tRNAs; chloramphenicol resistant; used for recombinant protein expression
<i>E. coli</i> Rosetta™ BL21(DE3)-R3-BirA (SGC)	co-expressing BirA (biotinylating enzyme) from a pACYC vector; with pRARE2 plasmid encoding rare codon tRNAs; chloramphenicol resistant; used for recombinant expression of biotinylated protein
<i>E. coli</i> Mach1™ (ThermoFisher)	used for cloning and plasmid propagation of prokaryotic expression constructs
<i>E. coli</i> Stable (NEB)	tetracycline and streptomycin resistant; used for cloning and plasmid propagation of eukaryotic expression constructs
<i>E. coli</i> ccdB survival™ 2 T1 ^R (Invitrogen)	resistant to the ccdB gene product; chloramphenicol resistant; used for ccdB vector propagation

7.2. Preparation of competent cells

Bacteria strains were made competent following the protocol developed by the SGC Biotechnology group. In brief, cells from frozen stocks were streaked out on LB agar plates with or without the appropriate antibiotics. 5 mL of O/N pre-cultures prepared from one colony were used to inoculate 500 mL of SOB medium (containing 2 % (w/v) peptone with casein, 0.5 % (w/v) bacto yeast extract, 85 mM NaCl, 2.5 mM KCl, 20 mM MgCl₂) and cultures were grown at 37 °C until OD₆₀₀ reached ~0.5. Then, medium was removed by centrifugation and cell pellets re-suspended in 160 mL ice-cold TFB1 buffer (30 mM potassium acetate, 100 mM RbCl, 10 mM CaCl₂, 50 mM MnCl₂, 15 % glycerol (v/v), adjusted to pH 5.85 with acetic acid) (160 mL per 500 mL culture). Cell pellets were then recovered by centrifugation (2600×g, 10 min, 4 °C) and after removal of the supernatant, cell pellets were re-suspended in 25 mL ice-cold TFB2 (10 mM MOPS free acid, 75 mM CaCl₂, 10 mM RbCl, 15 % (v/v) glycerol, adjusted to pH 6.5 with KOH). Cell suspensions were dispensed into pre-cooled Eppendorf tubes, flash frozen in liquid nitrogen and stored at -80 °C.

7.3. Bacterial transformation

Unless stated otherwise, 40 μ L chemically competent *E. coli* cells (thawed on ice) were transformed using 100-200 ng of expression plasmid by heat shock at 42 °C for 45 sec followed by 5 min incubation on ice. After transformation, 150 μ L LB (Miller) medium were added and bacteria incubated at 37 °C for 1 h. Positive transformants were selected overnight at 37 °C on LB agar plates containing the appropriate antibiotic(s).

7.4. Expression of recombinant proteins

The macrodomain proteins PARP14 MD1, MD2, MD2_{SERmut}, MD2_{SERmut-trunc} and MD3, PARP9 MD1, ALC1 and PARG were expressed with Tobacco Etch virus (TEV) protease-cleavable hexahistidine (His₆-) tags or ZB-tags (PARG only). Transformed BL21(DE3)-R3-pRARE cells were grown at 37 °C in TB (1.2 % tryptone, 2.4 % yeast extract, 17 mM KH₂PO₄, 72 mM K₂HPO₄, 0.45 % (v/v) glycerol) medium supplemented with appropriate antibiotics until OD₆₀₀ reached 1.0-1.2, then cooled to 18 °C and supplemented with 0.5 mM isopropyl- β -D-1-thiogalactopyranoside (IPTG) at an OD₆₀₀ of 1.2-1.5 to induce protein expression O/N. As an exception, cells transformed with MD2_{SERmut-trunc} constructs were grown in autoinduction TB medium (FroMedium) supplemented with 20 g/L glycerol, 10 % Antifoam 204 (Sigma) and appropriate antibiotics for 4 hours at 37 °C, followed by 48 hours at 18 °C. For expression of *in vivo* biotinylated macrodomain protein (PARP14 MD2), the construct was transformed into BL21 (DE3)-R3-BirA cells. Cells were grown at 37 °C in TB medium until OD₆₀₀ reached 1.0-1.2, then cooled to 18 °C and supplemented with 0.5 mM D-biotin (Sigma) dissolved in 10 mM bicine (pH 8.3) and 0.5 mM IPTG at an OD₆₀₀ of 1.2-1.5 to induce protein expression and biotinylation O/N. After protein expression, cells were harvested by centrifugation (6000 $\times$ g, 15 min, 4 °C) and pellets re-suspended in lysis purification buffer (see section V.4.2.1.) (with exception for PARP14 MD2_{SER-trunc} expression) and stored at -20 °C until further processing.

8. Work with proteins

Protein purifications and crystallisation were performed in general by following protocols and procedures developed at the SGC.

8.1. Purification of His₆-, GST- or biotin fusion proteins

The detailed experimental procedures are provided in the supplementary experimental details, section V.4.2.1. All proteins purified were characterised by SDS-PAGE as well as by mass spectrometric analysis (see section IV.8.2. and IV.8.3.).

8.2. Denaturing discontinuous polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins were analysed by SDS-PAGE using a 4-12 % Bis-Tris Gradient gel (Invitrogen) run in NuPAGE[®] MES SDS Running Buffers (Invitrogen) at 170 V for 1 h. Samples were loaded using loading dye (0.16 M Tris-HCl (pH 6.8), 4 % SDS, 20 % glycerol, 0.01 % bromophenol blue and 100 mM DTT) and visualised by InstantBlue[™] Coomassie blue staining (Expedeon) or alternatively by silver staining. For the latter, gels were first fixed with a solution containing 40 % EtOH and 10 % acetic acid, then washed twice in 30 % EtOH and once in water. Then, gels were washed in 0.02 % Na₂S₂O₃ for 1 min and after a brief wash in water incubated in 0.1 % AgNO₃ for 20 min at 4 °C. Afterwards, the gels were again washed in water, developed in 3 % Na₂CO₃/0.05 % formalin solution, briefly washed in water and finally fixed in 5 % acetic acid. Precision Plus Protein[™] Unstained Protein Standards (BioRad) were used for sizing the sample proteins.

8.3. Mass spectrometric analysis of purified proteins

Intact mass of the protein was determined using an Agilent ESI-TOF or ESI-Q-TOF mass spectrometer (Agilent) attached to an Agilent HPLC system. The protein sample was diluted to ~0.02 mg/mL in 0.1 % formic acid, injected on to a C3 reverse-phase column (2.1 mm x 12.5 mm Zorbax 5 µm 300SB-C3) and eluted with a water-methanol gradient, before injection into the mass detector. For some cases, the identity of proteins was additionally confirmed by tryptic

digestion and tandem MS (LC-MS/MS) from excised gel bands. The processing procedure of the latter is provided in the supplementary experimental details, section V.4.2.2.

8.4. Western blotting and immunodetection of proteins

Western blotting and immunodetection of proteins was carried out following the Western Blot Analysis guidelines by Li-COR Biosciences²⁵². After protein separation by SDS-PAGE, proteins were transferred electrophoretically onto Immobilon[®]-FL PVDF membranes (activated with MeOH (p.a.), pore size 0.45 µm) (Merck) for 70-80 min, 400 mA in WB transfer buffer containing 20 mM Tris-HCl (pH 8.3), 200 mM glycine, 20 % MeOH (wet western transfer). Alternatively (for DNA damage marker experiments), proteins were transferred onto nitrocellulose membranes (pore size 0.2 µm (GE Healthcare)) for 100 min, 30 V. Transfer efficiency was routinely controlled with Ponceau S staining (Sigma) for 1-2 min which was subsequently removed by several dH₂O washes. Afterwards, membranes were blocked with 5 % milk (BLOT-Quick Blocker[™] Reagent (Calbiochem)) in TBS-T (20 mM Tris-HCl (pH 7.6), 150 mM NaCl, 0.1 % Tween-20) for 1-2 hours at RT. For immunodetection, the membranes were first incubated with primary antibodies prepared in 5 % milk/TBS-T or 2 % BSA(Sigma)/TBS-T (for antibody dilutions see section IV.4.) O/N at 4 °C and then washed four times with TBS-T (5 min per wash) before incubation with secondary antibody for 45-60 min at RT, protected from light. Finally, membranes were washed again four times with TBS-T (5 min per wash) and protein detection was carried out using the Odyssey[®] CLx Imaging System (LI-COR) at the 700 and 800 nm channel or by detection on CL-XPosure[™] film (Thermo Scientific) with horseradish peroxidase enzyme (HRP) reaction using Pierce[™] ECL chemiluminescent substrate (Dako). Optionally, membranes were stripped using Restore[™] Plus Western Blot Stripping Buffer (Thermo Scientific). Precision Plus Protein[™] All Blue Prestained Protein Standards (BioRad) were used for sizing the sample proteins.

8.5. Macrodomain pull-down assay procedure

Biotinylated macrodomain proteins (600 pmol) were incubated with 12.5 µL in PBS equilibrated Dynabeads[®] MyOne[™] Streptavidin C1 magnetic beads (Life Technologies) in 487.5 µL PBS for 30 min at RT. Afterwards, the Dynabeads[®] were washed with PBS and incubated O/N at 4 °C with 1.0 mL of 4.0 mg cell lysate prepared from Jurkat cells that were prior solubilised in ice-cold lysis buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 20 mM MgCl₂, 1.0 % Digitonin and Complete[™] protease inhibitor cocktail (Roche). Dynabeads[®] were then washed four times in

buffer containing 50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 5 mM EDTA and proteins pulled-down were eluted sequentially for 20 min at 4 °C with 200 µL of 1 mM ADPR (Sigma), then with 400 µL of 100 mM glycine (pH 2.5) and finally by boiling of the Dynabeads® at 95 °C for 5 min. All elution samples were resolved by SDS-PAGE and visualised by silver staining and Western blotting and optionally further processed for mass spectrometric analysis.

8.6. Cell fractionation

Cell fractionation was performed based on the procedure developed by Méndez *et al.*²⁵³. Prior to cell fractionation, HeLa and U-2 OS wild-type and PARP14 KO cells were plated at $\sim 1.5 \times 10^6$ cells in 10 cm² TC dishes. After 18-24 h, cells were left either untreated or were treated with either 2 µM camptothecin (Sigma) for 1 hour or with 2 mM HU for 24 hours. Following the treatment, cells were collected by trypsinisation, pelleted by centrifugation (300×g, 3 min) and after a wash with cold PBS re-suspended in 80-150 µL (depending on the size of the pellet) of ice-cold buffer A (10 mM HEPES (pH 7.9), 10 mM KCl, 1.5 mM MgCl₂, 0.34 M sucrose, 10 % glycerol, 1 mM DTT, 0.1 % Triton X-100 and cOmplete™ protease inhibitor cocktail (Roche)). Cells were incubated on ice for 10 min and nuclei (fraction P1) were collected by centrifugation (5 min, 3500×g, 4 °C). The resulting supernatant (S1) was further clarified by high-speed centrifugation (5 min, 20000×g, 4 °C) and the supernatant (fraction S2) was collected for analysis by Western blotting and immunodetection. Fraction P1 was washed 3-4 times with 200 µL buffer A (centrifugation between washes: 5 min, 3500×g, 4 °C) until the pellet became slightly transparent. The nuclei were then lysed in ice-cold buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT, cOmplete™ protease inhibitor cocktail (Roche)) with the same volume used of buffer A, i.e. 80-150 µL. After 30 min incubation on ice, the insoluble chromatin (fraction P3) and soluble nuclear fraction (fraction S3) were separated by centrifugation (5 min, 3500×g, 4 °C). The P3 fraction was washed 3-4 times with 200 µL of buffer B until the pellet became fully transparent (centrifugation between washes: 5 min, 3500×g, 4 °C). SDS-PAGE loading buffer was then added to the P3 fraction (equal volume as used before for buffer A and buffer B) and the samples sonicated on ice (amplitude 35 %, 2×30 sec). Equal volumes of each fraction (S2, S3 and P3) were finally loaded with SDS-PAGE loading buffer onto a 4-12 % Bis-Tris Gradient gel and then further analysed by SDS-PAGE followed by Western blotting and immunodetection of PARP14, α-tubulin and histone H3.

8.7. Protein crystallisation

In general for crystallisation attempts of PARP14 MD2 proteins with MnK2-97, E62942a and E62881a, compounds were dissolved to 50 mM in DMSO and proteins concentrated to 20-40 mg/mL in buffer consisting of 25 mM HEPES (pH 7.4), 300 mM NaCl, (optional 5 % glycerol) and 0.5 mM TCEP. Compounds and proteins were mixed with c (compound) $\geq 1.5 \times c$ (protein) and $\gg 10 \times K_D$ of the compound for PARP14 MD2 and incubated on ice for approximately 30 min. In most cases, the sample was not centrifuged after incubation due to the low compound solubility and associated protein and/or compound precipitation. Samples were set-up in 150 nL volume drops on SWISSCI 3 Lens crystallisation sitting-drop plates at three ratios (2:1, 1:1, 1:2 of protein-inhibitor complex to crystallisation solution) using a Mosquito[®] crystallisation robot (TTP Labtech). JCSG, HIN, LFS and HCS sparse matrix crystallisation screens and individual components of the optimisation screens were obtained from Hampton Research and Molecular Dimensions, respectively. Drops were equilibrated at two temperatures (4 °C or 20 °C) and very few crystals were tested at I04 at Diamond Light Source (Harwell, UK) for diffraction. Follow-up screens were designed using the SGC screen designer and the MultiPROBE II Plus HT/Ex robotic liquid handling system (Perkin Elmer) to prepare matrix titrations of the original crystallant solution components, where optionally these new crystallant solutions were additionally supplemented with additives.

9. PARP14 MD3 fragment screen

The PARP14 MD3 fragment screen was performed with support of the SGC crystallography group (Professor Frank v. Delft).

9.1. PARP14 MD3 apo protein crystallisation

PARP14 MD3 was crystallised by mixing on pre-cooled sitting drop crystallisation plates 100 nL of 30 mg/mL protein in 20 mM HEPES (pH 7.4), 300 mM NaCl, 10 % glycerol, 0.5 mM TCEP with 50 nL of reservoir solution containing 80 mM KBr (Molecular Dimensions) and 30 % PEG2kMME (Molecular Dimensions) and adding 15 nL of 'H6 seeds' solution, diluted 1:500 from the stock in 90 mM KBr and 30 % PEG2kMME. 'H6 seeds' were prepared from a single crystal growing in 90 mM KBr and 30 % PEG2kMME which was smashed to seeds using a

3 mm PTFE seed bead (Fisher Scientific) using standard laboratory vortex at full speed in 100 μ L reservoir solution. Crystal growth was completed after 24 hours with incubation at 4 °C.

9.2. Fragment soaking and data collection

PARP14 MD3 *apo* crystals were soaked with compounds from the DSPL2.0 library consisting of 776 fragments at 500 mM in D₆ DMSO (and a subset in ethylene glycol). Soaking was performed by acoustically transferring 17.5 nL of compound solution to the crystallisation drop using an Echo 550 (Labcyte)¹⁹⁸ resulting in a final compound concentration of 50 mM with 10 % DMSO, calculated based on the initial drop volume. Crystals were incubated for 4-6 hours at 4 °C and then harvested (without cryoprotection) and cryocooled before X-ray diffraction data collection on the beamline I04-1 at Diamond Light Source (Harwell, UK).

9.3. PARP14 fragment screen - data processing

X-ray diffraction data was processed using the Diamond autoprocessing pipeline, which utilises *i.a.* *xia2*²⁵⁴ and *DIALS*²⁵⁵, and programs from the CCP4 suite²⁵⁶. Electron-density maps were generated using the *XChemExplorer*²⁰³ via *DIMPLE*²⁵⁷. Ligand restraints were generated with *AceDRG*²⁰⁴ and ligand binding was detected with *PanDDA*²⁰⁵, with ligands built into *PanDDA* event maps. Iterative refinement and manual model correction was performed using *REFMAC*²⁵⁸ and *Coot*²⁵⁹, respectively.

10. **Biophysical assays for compound screening and characterisation**

The following biophysical assays were performed in general following protocols developed by the SGC screening group (Dr. Oleg Fedorov) and the manufacturer's instructions of the respective kits and instruments.

10.1. Macrodomain AlphaScreen assay

Assays were performed with minor modifications from the manufacturer's protocol (PerkinElmer)^{260,261}. All reagents were diluted in buffer containing 25 mM HEPES (pH 7.4),

100 mM NaCl, 0.5 mM TCEP, 0.1 % BSA and 0.05 % CHAPS and allowed to equilibrate to RT before addition to plates. The assays were run in 20 μ L volumes in low-volume 384-well plates (ProxiPlate™-384 Plus, PerkinElmer) at RT. To determine ideal assay concentrations of the corresponding macrodomain protein and the macrodomain AlphaScreen peptide (a biotinylated and ADP-ribosylated 11 residue sequence), 4 μ L volumes of peptide (0-16 μ M; final assay concentration: 0-3.2 μ M) were incubated with 4 μ L volumes of His₆-tagged macrodomain protein (0-16 μ M; final assay concentration: 0-3.2 μ M) in 4 μ L buffer for 30 min at RT in foil-sealed plates. For compound screening and IC₅₀ characterisation, 12 μ L of a solution containing peptide and His₆-tagged macrodomain protein in the pre-determined assay concentrations (see Figure S7 in section V.2.) in assay buffer were incubated for 30 min at RT with 50 nL or 100 nL compound solution (pre-dispensed into the assay plate from 10 mM or 50 mM DMSO stocks using an Echo 550 (Labcyte)). Then, 8 μ L of streptavidin-coated donor beads (7 μ g/ml) and nickel chelate acceptor beads (7 μ g/ml) (Perkin Elmer AlphaScreen™ Histidine (Nickel Chelate) Detection Kit) were added under low light conditions and plates were incubated for 60 min at RT protected from light. Plates were read on a PHERAstar FS plate reader (BMG Labtech) using an AlphaScreen 680 excitation/570 emission filter set. Alternatively for counterscreening of the compounds, 12 μ L of 75 nM biotinylated and His₆-tagged linker peptide (PerkinElmer) was added to 50 nL or 100 nL of the compounds and plates were processed as described above. The IC₅₀ values were determined by nonlinear regression using GraphPad Prism v. 7.03 (GraphPad Software, CA, USA).

10.2. HTRF (Homogenous Time-Resolved Fluorescence) assay

HTRF assays were performed following the manufacturer's instructions (PerkinElmer)²⁶². The assays were performed in 20 μ L volumes in low-volume 384-well plates (ProxiPlate™-384 Plus, PerkinElmer) at RT. GST-tagged PARP15 MD1 protein was diluted to 20-2000 nM (final assay concentration 5-500 nM) and the macrodomain AlphaScreen peptide was diluted to 40 nM - 120 μ M (final assay concentration 10 nM - 30 μ M) in Binding Domain Diluent Buffer provided by the kit. The Streptavidin-XL665 was prepared in concentrations from 5 nM - 15 μ M (final assay concentrations 1.25 nM - 3.75 μ M) and the Anti-GST-Eu³⁺ Cryptate diluted to 1x from 50x stock (final concentration 0.25x) in Binding Domain Detection Buffer provided by the kit. 5 μ L of each component were then consecutively added to the assay plate. The plate was read at t0 and after 3 and 24 hours incubation at RT on a PHERAstar FS plate reader (BMG Labtech) measuring the fluorescence signals at 665 nm and 620 nm. Results were expressed as HTRF ratio (665 nm/620 nm) $\times 10^4$.

10.3. Differential scanning fluorimetry (T_m -shift assay)

Thermal melting experiments¹⁴⁹ were carried out using an Mx3005p Real Time PCR machine (Stratagene). Proteins were buffered in 10 mM HEPES (pH 7.4), 500 mM NaCl and assayed in a 96-well plate at a final concentration of 2.0 μ M in 20 μ L volume. SYPRO[®] Orange (Molecular Probes) was added as a fluorescence probe at a dilution of 1:1000. Excitation and emission filters for the SYPRO[®] Orange dye were set to 492 nm and 610 nm, respectively. The temperature was raised with a step of 1 $^{\circ}$ C per 21 sec from 25 $^{\circ}$ C to 96 $^{\circ}$ C and fluorescence readings were taken at each interval.

10.4. NanoDSF (nano differential scanning fluorimetry)

NanoDSF experiments were performed on the Nanotemper Prometheus NT.48 instrument²⁶³. PARP14 and RNF146 WWE domain proteins were diluted to 20 μ M and *iso*-ADPR to 100 μ M to 2.0 mM in assay buffer (10 mM HEPES (pH 7.4), 500 mM NaCl). 10 μ L of protein solution was added to 10 μ L of the corresponding *iso*-ADPR solution and incubated on ice for 20 min. The reaction mixtures were loaded into Nanotemper NT.48 Series NanoDSF Standard Capillaries which were inserted into the instrument. A discovery scan was performed to control the consistency of the tryptophan signal at 330 nm and 350 nm and for determining the excitation power applied for the following melting scan. The melting scan was then performed from 20 $^{\circ}$ C to 95 $^{\circ}$ C with a temperature slope of 5.0 $^{\circ}$ C/min and data analysed using the Nanotemper Prometheus NT analysis software provided by the manufacturer.

10.5. Biolayer interferometry (BLI)

Kinetic ligand-binding measurements were performed using an Octet RED384 BLI instrument (fortéBio)²⁶⁴. Superstreptavidin (SSA) biosensors were loaded with biotinylated macrodomain protein and equilibrated for 120 sec in assay buffer (25 mM HEPES (pH 7.4), 100 mM NaCl, 0.01 % Tween 20). Association and dissociation were monitored for 240 sec each in assay buffer at 25 $^{\circ}$ C. For compound characterisation, compounds were typically prepared as seven 1:1 serial dilutions starting from 10 μ M (GeA-69), 50 μ M (GeA-69 for selectivity screening and GeA-69 analogues) or 100–400 μ M for validation of HTS screening hits. Binding to the reference sensors (no protein attached) was subtracted before calculations and data was processed using the fortéBio analysis software provided by the manufacturer.

10.6. Isothermal titration calorimetry (ITC)

Binding experiments were carried out on either a VP-ITC microcalorimeter (MicroCal), iTC200 instrument (MicroCal) or Nano ITC microcalorimeter (TA Instruments)¹⁴³. For all measurements, proteins were dialysed O/N at 4 °C in 50 mM HEPES (pH 7.4), 300 mM NaCl, 5 % glycerol, 0.5 mM TCEP using D-tube™ Dialysis Midi MWCO 3.5 kDa (Novagen) dialysis tubes prior to the experiment. On the VP-ITC instrument, almost all of the experiments were performed at 12 °C or 15 °C, a reference power of 12 µCal/sec and a stirring speed of 307 rpm. The titrations were conducted using an initial injection of 2 µL followed by 28 identical injections of 10 µL with duration of 4 sec (per injection) and a spacing of 240 sec between injections. In particular compound screening and competition experiments were performed on the iTC200 instrument. The majority of the experiments were carried out at 20 °C, a reference power of 12 µCal/sec and 5 µCal/sec and a stirring speed of 1000 rpm. Following an initial injection of 0.2 µL, 20 identical injections of 2 µL were run with duration of 4 sec (per injection) and a spacing of 150 sec between injections. The Nano ITC microcalorimeter was alternatively used for the VP-ITC for compound characterisation. The experiments were performed at 25 °C and a stirring speed of 350 rpm. Following an initial injection of 3.7 µL, 30 identical injections of 7.96 µL were run with a spacing of 200 sec between injections. Data analysis was carried out using Origin software with the MicroCal plugin or with the *NanoAnalyse* software provided by the manufacturer for data acquired on the Nano ITC. Thermodynamic parameters were calculated using $\Delta G = \Delta H - T\Delta S = -RT\ln K$, where ΔG , ΔH , and ΔS are the changes in free energy, enthalpy, and entropy of binding, respectively. In all cases, a single binding site model was employed.

10.7. Dynamic light scattering (DLS)

DLS experiments were performed on an Avid Nano's W130i SABRe optics system. 5 µL of compound dilutions prepared in PBS or DMEM were loaded into the BladeCell™ and light scattering intensities were detected in an automatic mode at 20 °C with a 30.0 mW laser at 660 nm using a detector angle of 90° and a solvent refractive index of 1.333. The laser power was 100 %, and the integration time was 10 s. Data was filtered with the default settings of a minimum peak area of 5.0 % and minimum Rh of 0.2 nm. The scattering intensities were calculated by averaging three independent data sets, each with 10 measurements.

11. Cellular assays

Cellular thermal shift assays (CETSA)^{169,265} were performed following a protocol established by the SGC Chemical Biology group (Dr. Kilian Huber). Cell adhesion and migration protocols were developed in course of this work. Image analysis algorithms were developed in collaboration with the Cell-based High-Throughput Screening group at the TDI (Dr. Daniel Ebner). Seahorse experiments were performed following the manufacturer's instructions²⁶⁶. Immunofluorescence imaging for DNA damage marker detection was performed following protocols developed by the CR-UK/MRC Oxford Institute for Radiation Oncology group (Professor M. Tarsounas). Statistical analysis and EC₅₀ calculations were performed using GraphPad Prism v. 7.03 (GraphPad Software, CA, USA).

11.1. Cellular thermal shift assay (CETSA)

U-2 OS cells seeded into T75 flasks were treated at 80 % confluency with 250 μ M GeA-69 or DMSO for 2 hours. Following the treatment, cells were trypsinised, pelleted by centrifugation (300 $\times$ g, 3 min) and washed once with PBS. Cells were then re-suspended in 1.2 mL PBS and the resulting cell suspension was aliquoted over 11 PCR tubes (100 μ L each). Cells were pelleted in the PCR tubes by centrifugation (300 $\times$ g, 3 min) and the supernatant completely removed. Samples were then heated in a gradient PCR cycler with one of the following 10 temperatures: 37, 41, 45, 49, 53, 57, 61, 65, 69, 73 $^{\circ}$ C for 3 min, following 3 min at RT. For further analysis of the samples, cells were re-suspended in 30 μ L lysis buffer (50 mM Tris-HCl (pH 7.4), 100 mM NaCl, 1.5 mM MgCl₂, 5 % glycerol, 0.80 % NP-40, 25 mM NaF, 1 mM Na₃VO₄, 1 mM PMSF (Sigma), 1 mM DTT, 10 μ g/mL TLCK (Sigma)) containing protein inhibitor cocktail (leupeptin, aprotinin, soy bean trypsin) and lysed by three freeze-thaw cycles in liquid nitrogen. Insoluble cell debris and aggregated proteins were removed by centrifugation at 17,000 $\times$ g for 20 min at 4 $^{\circ}$ C. The supernatant was recovered of which 12 μ L were mixed with 4 μ L of 4 $\times$ SDS-PAGE loading dye and heated (2 min, 90 $^{\circ}$ C). A sample (10 μ L) was loaded for SDS-PAGE and was then further processed by Western blotting and immunodetection of PARP14 and GADPH for analysis of thermal stabilisation.

11.2. Cell adhesion assay

Cell adhesion assays were designed as trypsin-based detachment assays performed in 96-well plates (CellCarrier-96 Black, Optically Clear Bottom (PerkinElmer)) with either standard TC coating or additional laminin or fibronectin coating. Plates were coated with laminin (Sigma) or fibronectin (F1141, Sigma) following the manufacturer's instructions. For performing the assay, HeLa wild-type and PARP14 KO cells were plated on the 96-well assay plate at 5000 cells/well (2 rows for each cell line). The next day, confluency of 50-60 % was confirmed. Cells were washed twice with 120 μ L PBS and then treated for the indicated times (1-12 min) with 75 μ L of 0.025 % trypsin/EDTA which was prepared from 0.5 % stock trypsin/EDTA solution in PBS and warmed up to 37 °C directly prior to the use in the assay. 0.075 % trypsin/EDTA solution was used in assays performed with fibronectin coated plates. The reaction was stopped by addition of 75 μ L DMEM containing FBS and 2 mM PMSF (Sigma). In control wells, 75 μ L of PBS was added instead of trypsin/EDTA solution. All of those steps as well as the following steps were pipetted column-wise using a manual 8-channel pipette to apply the same pipetting force to each cell line type in the column. Following the trypsin/EDTA treatment, cells were washed twice with 100 μ L PBS to remove non-adherent cells, fixed with MeOH (10 min, 4 °C) and cell nuclei were stained with Hoechst33342 (1 μ M in PBS) (Molecular Probes). After 1 h at 4 °C, the stain was removed and replaced with 150 μ L PBS. The plate was protected from light and kept at 4 °C until reading on the automated confocal microscope IN Cell Analyser 6000 (GE Healthcare). The whole wells were imaged on the DAPI channel (excitation at 405 nm, emission at 455 nm) by imaging 4 fields per well with 4 $\times$ magnification and the number of nuclei/well was quantified using a customised protocol developed with the IN Cell Developer Toolbox software (GE Healthcare) [Figure S24].

11.3. Scratch migration assay

A cell migration assay was developed as scratch (wound healing) assay^{219,267}. For this, HeLa wild-type and PARP14 KO cells were plated at a density of 10,000 cells/well in 96-well plates (CellCarrier-96 Black, Optically Clear Bottom, TC treated (PerkinElmer)) (2 rows for each cell line, i.e. 24 wells) and the next day confluency of ~95 % was confirmed to perform the assay. First, cells were stained with 2.5 μ M CellTracker™ Green CMFDA Dye (Invitrogen) (diluted from 10 mM DMSO stock in serum-free DMEM and warmed-up to 37 °C before applying it to the cells). After incubation for 30 min at 37 °C, the stain was removed and cells washed once for 30 min with 120 μ L serum-free medium. Cells in wells for proliferation control (column 2 and

column 8) were additionally stained in this step with Hoechst33342 (2 μ M in serum-free medium). Then, cells were washed once for 30 min with 100 μ L medium containing additionally 10 μ M Ara-C (cytosine β -D-arabinofuranoside (Sigma)) for 12 sample wells, thereby removing the Hoechst33342 stain from the proliferation control wells. Following the staining and washing steps, the scratches into the cell layers were performed using the WoundMaker™ (EssenBioscience)²⁶⁷, a 96-pin mechanical device creating 700-800 μ m wounds. Loose cells were then removed from the scratch area by two consecutive washes with 120 μ L medium. Afterwards, 150 μ L of medium containing 10 μ M Ara-C was added to 12 sample wells, 150 μ L of medium containing 1 μ M Cytochalasin D (Sigma) was added to 6 wells for migration control and 150 μ L of medium only was added to the 6 remaining wells. The plate was sealed with gas permeable moisture barrier imaging seals (4titude) and placed into the pre-equilibrated IN Cell Analyser 6000 (GE Healthcare) providing a constant atmosphere of 37 °C and 5.0 % CO₂ around the assay plate. After imaging the proliferation control wells (whole well by imaging 4 fields with 4x magnification) on the DAPI channel (excitation at 405 nm, emission at 455 nm), the closure of the scratch was monitored by acquiring images every 2 h (over 24 h) on the FITC channel (excitation at 488 nm, emission at 525 nm) as well as brightfield images (one field centred on each well, 4x magnification). Following the time series acquisition, the plate was removed from the microscope, cell nuclei in all wells were counterstained with Hoechst33342 (1 μ M, 45 min) and imaged on the DAPI channel (whole well by imaging 4 fields with 4x magnification). Analysis of the scratch assay and quantification of the cell count was performed using customised protocols developed with the IN Cell Developer Toolbox software (GE Healthcare) [Figure S24, S25].

11.4. Cellular proliferation assays

Cellular proliferation was analysed by AlamarBlue™ assays or by live cell imaging of Hoechst33342 stained cells. For AlamarBlue™ assays, HeLa, U-2 OS or HEK293 cells were plated in 384-well plates at a density of 300 cells per well 18-24 h prior to compound treatment. Cells were treated with compound concentrations prepared by 1:1 serial dilution (top concentrations: 250 μ M for GeA-69 and MnK2-68, 10 mM for HU (Sigma), 100 μ M for cisplatin (Sigma) and olaparib (Sigma), 200 nM for camptothecin (Sigma) and doxorubicin (Sigma) and 100 nM for staurosporine (Sigma)). 72 h (GeA-69 and MnK2-68) or 6 days after compound treatment cells were exposed to AlamarBlue™ (Invitrogen) and the colorimetric change was measured at 530 nm and 590 nm using an Envision 2103 Multilabel Reader (Perkin Elmer). Alternatively, for measuring cell proliferation by quantitation of the nuclei count, cells were plated at a density of 2000-5000 cells 18-24 h prior to compound treatment in 96-well plates

(CellCarrier-96 Black, Optically Clear Bottom, TC treated (PerkinElmer)). Cells were treated with compound concentrations prepared by serial dilution (top concentrations: 100 μ M for Ara-C (Sigma) and oligomycin (Sigma)). After 24 h (Ara-C treatment) or 72 h (oligomycin treatment), compounds were removed and nuclei of cells were stained for 1 h with 1 μ M Hoechst33342 prepared in medium. Whole wells were imaged on the DAPI channel (4x magnification, 4 fields per well, excitation at 405 nm, emission at 455 nm) using the IN Cell Analyser 6000 (GE Healthcare). The number of nuclei/well was quantified using a customised protocol developed with the IN Cell Developer Toolbox software (GE Healthcare) [Figure S24].

11.5. Clonogenic survival assay

Cells were plated at a density of ~250 cells per well in 6-well plates. After 18-24 hours, cells were treated with different concentrations of compound (cisplatin, HU, olaparib) for 24 hours or with IR. Following the treatment, cells were incubated for 10-14 days to allow colony formation. Colonies were stained with 0.05 % (w/v) crystal violet (Sigma) in 50 % methanol. Cell survival was expressed relative to control cells.

11.6. Seahorse experiments

The extracellular acidification rate (ECAR) and the oxygen consumption rate (OCR) of cells were monitored in real-time with the Seahorse XF96 Analyzer (Agilent) using the Seahorse XF Glycolysis Stress Test Kit (Agilent), according to the manufacturer's instructions. The day prior to the assays, HeLa wild-type and PARP14 KO cells were seeded at a density of 8000 cells/well in a XF96 cell culture microplate containing complete growth medium and a XF96 sensor cartridge was equilibrated in XF calibrant solution O/N at 37 °C in a non-CO₂ incubator. Unless stated otherwise, cells were incubated in unbuffered DMEM recipe based XF Base Medium (without Phenol Red, 0 mM glucose) (Agilent) for 45 min in a non-CO₂ incubator prior to the assay. Cells were then placed into the pre-equilibrated XF96 Analyzer and ECAR and OCR were recorded under sequential injection of 10 mM glucose, 2 μ M oligomycin and 50 mM 2-deoxy-glucose. All measurements were normalised to the cell number determined after the assay by staining the cell nuclei with 2 μ M Hoechst33342, reading the plate on the IN Cell Analyser 6000 and then performing the quantification using a customised protocol developed with the IN Cell Developer Toolbox software (GE Healthcare) [Figure S24]. Seahorse data was analysed using the Wave Desktop 2.3.0 analysis software provided by Seahorse Bioscience. For calculating the metabolic parameters the following equations were used as recommended by the manufacturer:

(i) glycolysis = (maximum rate measurement before oligomycin injection) – (last rate measurement before glucose injection), (ii) glycolytic capacity = (maximum rate measurement after oligomycin injection) – (last rate measurement before glucose injection), (iii) glycolytic reserve = (glycolytic capacity) – (glycolysis).

11.7. Immunofluorescence imaging

Immunofluorescence imaging was performed on cells either plated in 96-well plates (CellCarrier-96 Black, Optically Clear Bottom, TC treated (PerkinElmer)) or grown on glass coverslips for DNA damage experiments. For the former, HeLa and U-2 OS wild-type and PARP14 KO cells were plated at ~5000 cells/well. After 18-24 h, the growth medium was aspirated, cells washed with PBS and fixed with 4 % formaldehyde in PBS for 20 min at RT. Following a PBS wash (150 μ L), cells were permeabilised with 0.5 % Triton X-100 in PBS for 10 min at RT and then washed with 150 μ L PBS. Afterwards it was blocked for 60 min with 3 % FBS in PBS and incubated with primary antibody (see table in section IV.4. for antibody dilutions) in 3 % FBS/PBS overnight at 4 °C. Following the incubation, cells were washed three times with PBS and it was incubated with secondary antibody at 1:500 in 3 % FBS/PBS for 1 hour at RT (protected from light). After a PBS wash, cell nuclei were counterstained with Hoechst33342 (2 μ M in PBS) for 10 min and then the cells were washed again with PBS and imaged on the IN Cell Analyser 6000 with the corresponding channels.

For performance of DNA damage experiments, HeLa and U-2 OS wild-type and PARP14 KO cells were grown on coverslips (diameter: 1.0 cm, thickness: 1.0 mm (VWR)) which were prior to usage autoclaved and for HeLa cells additionally coated with poly-D-lysine (Sigma). After 18 h of seeding the cells into 6-well plates with coverslips, cells were either left untreated or were treated with 2 mM HU for 24 h. Following the treatment, coverslips were mounted into 24-well plates and washed with PBS. Cells were swollen in hypotonic solution (85.5 mM NaCl/ 5 mM MgCl₂ (pH 7.0)) for 5 min at RT, fixed with 4 % paraformaldehyde (Electron Microscopy Sciences) in dH₂O (pH 8.5) for 10 min at RT and permeabilised by adding 0.03 % SDS to the fixative. After washing twice with 0.4 % Photo-Flo 200 (Kodak)/PBS and blocking with blocking buffer ADB (ADB = antibody dilution buffer containing 1 % goat-serum, 0.3 % BSA and 0.005 % Triton X-100 in PBS) for 15 min at RT, cells were incubated with primary antibody diluted in ADB overnight at RT (see table in section IV.4. for antibody dilutions). Cells were then subjected to three washes, each 5 min, with the first and third wash in 0.4 % Photo-Flo 200/ADB, the second wash with ADB plus 0.005 % Triton X-100. Afterwards, cells were incubated with secondary antibodies (see table in section IV.4. for antibody dilutions) in ADB for 1 hour at RT (protected from light). Finally, cells were washed with 0.4 % Photo-Flo 200/PBS for 10 min, then

with 0.4 % Photo-Flo 200/dH₂O for 1 min and the coverslips dried on filter paper (Whatman) (~60 min). In a last step, coverslips were mounted on microscope glass slides using the ProLong Antifade kit (Life Technologies) supplemented with 2 µg/mL DAPI. Samples were viewed with a Leica DMI6000B inverted microscope and fluorescence imaging workstation equipped with a HCX PL APO × 100/1.4–0.7 oil objective. Images were acquired using a Leica DFC350 FX R2 digital camera and LAS-AF software (Leica).

11.8. Laser micro-irradiation assay

Laser micro-irradiation was performed as described previously²⁶⁸ by Dr. Ian Gibbs-Seymour (Dr. Ivan Ahel's Laboratory, Sir William Dunn School of Pathology, Oxford). U-2 OS wild-type or *PARP1*^{-/-} cells were pre-sensitised for 24 h with 10 µM BrdU (Sigma), then the media was changed to Live Cell Imaging Solution (Thermo Fisher) before laser micro-irradiation and imaging. Laser micro-irradiation was performed on an Olympus Fluoview FV1200 confocal microscope equipped an inverted IX83 motorised stage with a 37 °C humidified chamber and 60x/1.40 oil UPlanSApo objective and 405 nm laser.

12. **Generation of iso-ADPR**

Based on the methods reported previously in particular by Wang *et al.*⁴³, Tan *et al.*¹⁴⁶ and Kiehlbauch *et al.*²⁵, iso-ADPR was prepared as follows. In a 5 mL reaction mixture, 6.5 mg PARP5 catalytic domain (from the dimer gel filtration fraction) was incubated with 20 mM β-NAD⁺ (Sigma) for 24 hours at RT in buffer containing 100 mM Tris-HCl (pH 8.0), 10 mM MgCl₂ and 1 mM DTT on a rotating device. Following the incubation, the reaction was stopped with 5 mL ice-cold 24 % trichloroacetic acid and the precipitate collected by centrifugation. The pellet was washed with ice-cold ethanol (99.9 %) and then dried using a speed vac evaporator. The pellet was further dissolved in 6 mL buffer containing 30 mM Tris-HCl (pH 8.0), 1.0 mM EDTA, 0.1 % SDS and 0.2 mg/mL proteinase K (Invitrogen) and the sample was digested for 48 hours at 37 °C. PAR polymer was detached by addition of 1 mL of 5 M NaOH and incubation of the reaction mixture at 60 °C for 1 h. The PAR polymer was recovered by phenol/chloroform/isoamyl alcohol (25:24:1) extraction (Acros organics) followed by ethanol/ammonium acetate (2 M final concentration) precipitation. The pellet was collected by centrifugation and washed with 70 % EtOH before air drying. The PAR polymer was further purified using a Superdex S75 10/300 GL column (GE Healthcare) running in dH₂O monitoring

UV absorbance at 255 nm. PAR containing fractions were pooled and digested over 48 hours at RT with 25 U snake venom phosphodiesterase (Worthington), (with an additional 25 U added after 24 hours), in buffer containing 100 mM Tris-HCl (pH 8.0), 100 mM NaCl and 15 mM MgCl₂. The product after the digestion was further purified by anion exchange chromatography using a MonoQ 5/50 GL (GE Healthcare) and a 25 mM Bis-Tris (pH 6.0) buffer system which is not recommended upon repeat of the protocol due to the impossibility of evaporating the buffer components. The product was then additionally purified on the MonoQ 5/50 GL using a 20 mM ammonium carbonate buffer system at pH 6.0. NaCl was removed from the sample by MeOH (MS grade) extraction. Finally, the sample was dried using a speed vac evaporator and dissolved in ddH₂O (MS grade) to a final concentration of 6 mM. *Iso*-ADPR concentrations were determined by UV absorbance measurement using a NanoDrop ND-1000 (Thermo Scientific) and the following equation: $[iso\text{-ADPR}] = [(A_{258})\text{cm}^{-1}]/[13,500 \text{ cm}^{-1}\text{M}^{-1}]^{146}$. LC-MS and NMR confirmed purity and identity of the product, *iso*-ADPR.

V. APPENDIX

1. Supplementary information chapter 1

The following figures are supplementary to chapter 1 of ‘Results and Discussion’ (section II.1.).

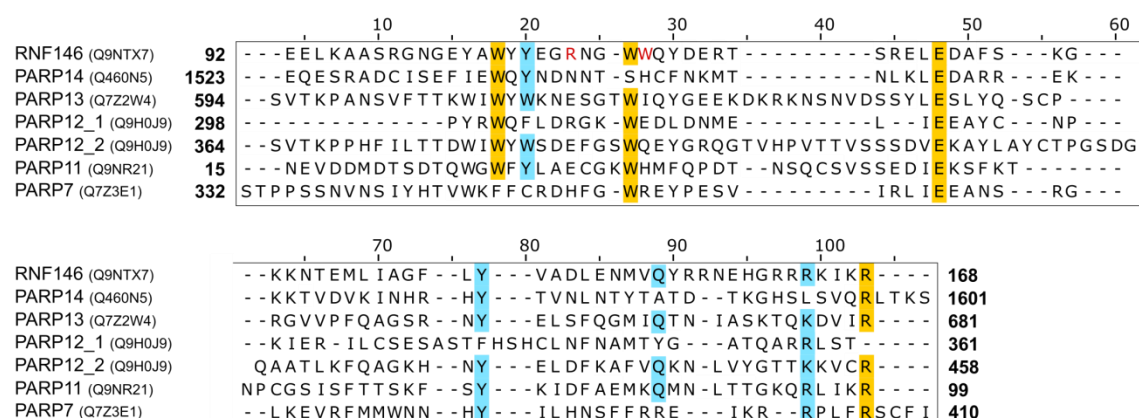


Figure S1. Multiple sequence alignment of known WWE domain sequences present in human PARP proteins using ClustalOmega. As defined by Wang *et al.*⁴³, most-conserved residues important for WWE domain folding are highlighted in yellow and conserved residues potentially involved in *iso*-ADPR binding in cyan. Residues highlighted in red are non-conserved but interact with *iso*-ADPR in RNF146. The UniProtKB IDs are given behind the proteins.

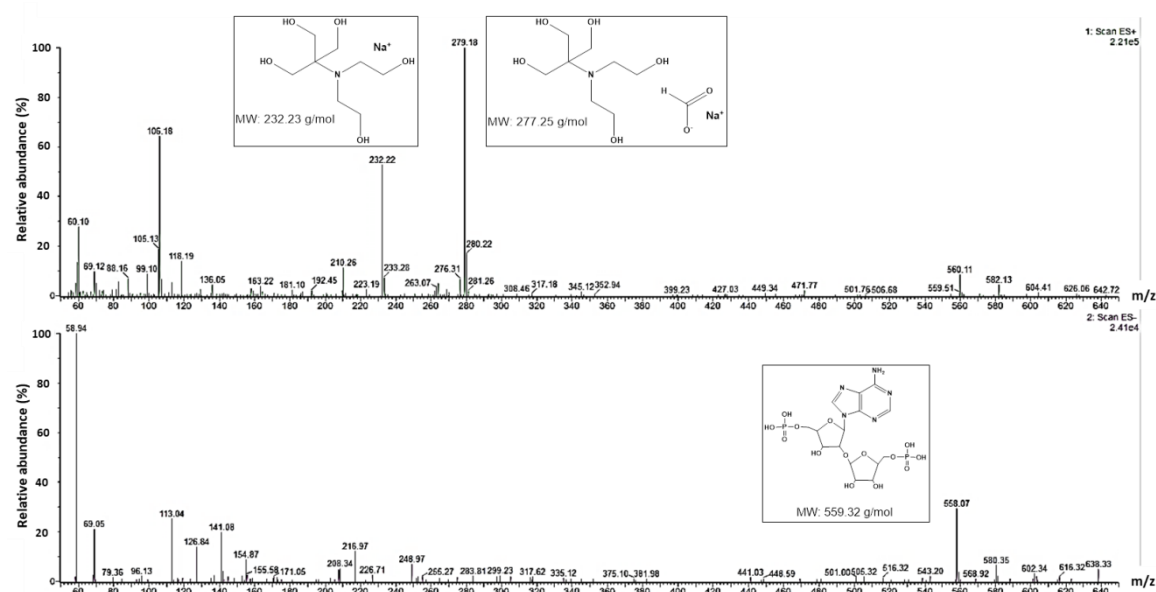


Figure S2 [related to Figure 8]. Mass spectrometric analysis of the *iso*-ADPR sample. Bis-Tris impurities from the purification process that are present in the *iso*-ADPR sample are shown on the top. Top: +ES scan, bottom –ES scan. *iso*-ADPR was detected as $[M-H]^-$ with m/z of 558.

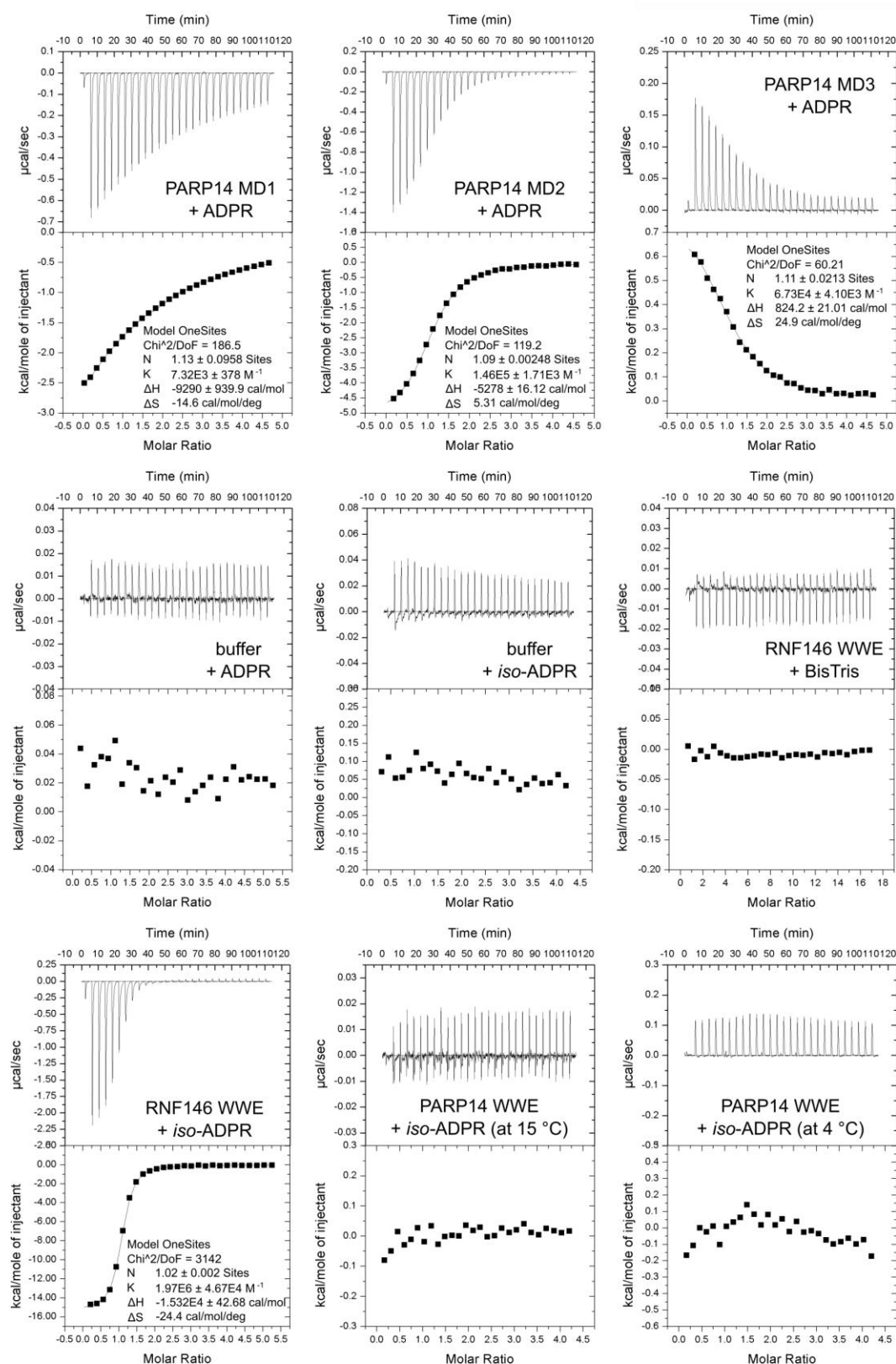


Figure S3 [related to Figure 9]. ITC data obtained from the Origin v7.0 software with the MicroCal plugin showing ITC results for macrodomain and WWE sequence recognition studies. The upper part of the panels show the raw titration curves resulting from stepwise addition of the ligand ADPR (1 mM), *iso*-ADPR (0.5 mM) or BisTris (2 mM) into a solution containing the macrodomain or WWE domain as indicated at a 25-fold (macrodomain) or 20-fold (WWE domain) lower concentration in the presence of 50 mM HEPES (pH 7.4), 300 mM NaCl, 5 % glycerol, 0.5 mM TCEP at 15 °C. The lower part of the panels show the normalised binding enthalpies corrected for the heat of dilution as a function of binding site saturation. Solid lines represent a nonlinear least squares fit using a single-site binding model.

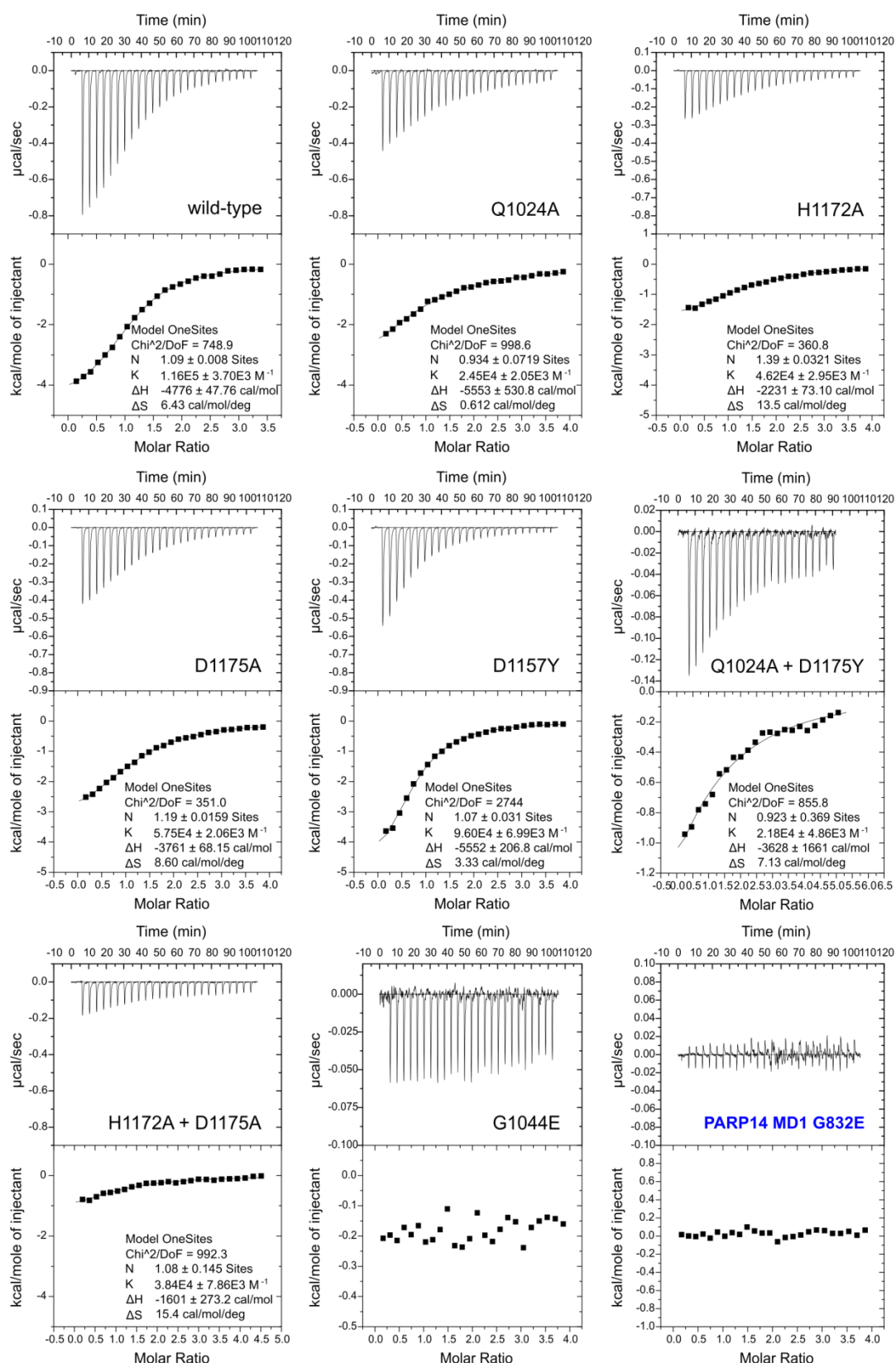


Figure S4 [related to Figure 11]. ITC data obtained from the Origin v7.0 software with the MicroCal plugin showing ITC results for macrodomain mutation studies. The upper part of the panels show the raw titration curves resulting from stepwise addition of the ADPR (0.7 mM or 0.6 mM) into a solution containing the corresponding PARP14 MD2 mutants (or PARP14 MD1 G832E were indicated) 20-fold less concentrated in the presence of 50 mM HEPES (pH 7.4), 300 mM NaCl, 5 % glycerol, 0.5 mM TCEP at 12 °C. The lower part of the panels shows the normalised binding enthalpies corrected for the heat of dilution as a function of binding site saturation. Solid lines represent a nonlinear least squares fit using a single-site binding model.

2. Supplementary information chapter 2

The following figures and tables are supplementary to chapter 2 of ‘Results and Discussion’ (section II.2.).

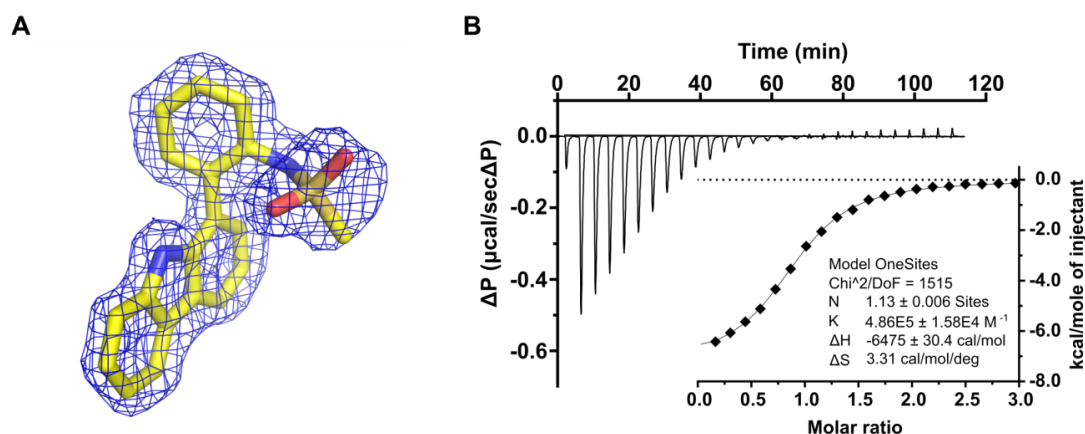


Figure S5 [related to Figure 19]. (A) The $2F_o - F_c$ electron density map contoured at 1.0σ around MnK2-13 as in the PARP14 MD2 co-crystal structure is shown in blue. **Note:** The image was prepared by Dr. Jonathan Elkins (SGC). (B) ITC experiment confirming binding of MnK2-13 to PARP14 MD2 with a K_D of $2.1 \mu\text{M}$. Raw injection heats are shown at the top, data fitting to a binding isotherm (one site binding model) is shown in the inset.

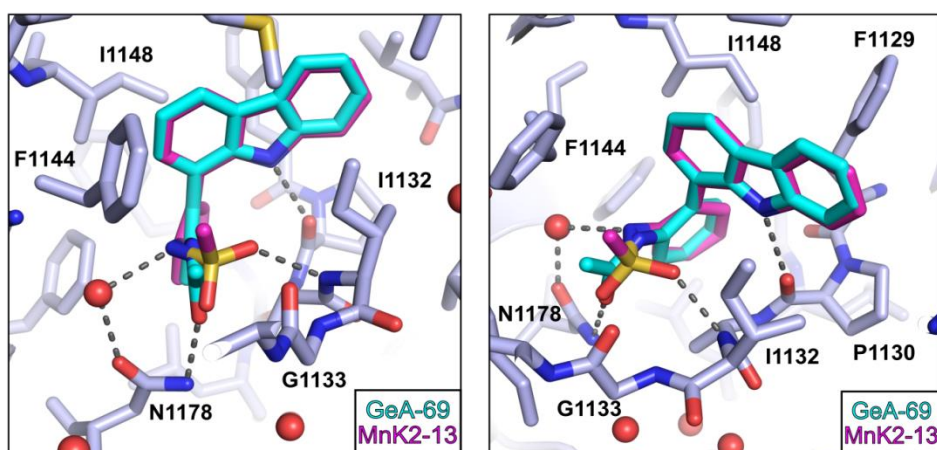
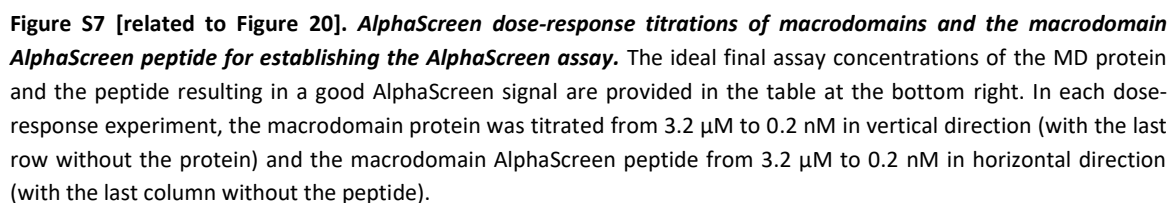


Figure S6 [related to Figure 19]. Docking of GeA-69 into PARP14 MD2. GeA-69 was docked into PARP14 MD2 using the PARP14 MD2-MnK2-13 co-crystal structure (PDB ID 5O2D) as template. The two views show the likely binding mode of GeA-69 which overlays almost exactly as expected with MnK2-13. Based on this model, the acetamide moiety of GeA-69 may interact slightly different with PARP14 MD2 compared to the sulfonamide moiety of MnK2-13. GeA-69 may probably form a stronger hydrogen bond (based on bond angle and length) to N1178 than MnK2-13 which additionally interacts with the amide of I1132 and an internally complexed water molecule. However, it is difficult to rationalise with confidence the slightly better binding affinity of GeA-69 (860 nM) compared to MnK2-13 ($2.1 \mu\text{M}$) with this model.



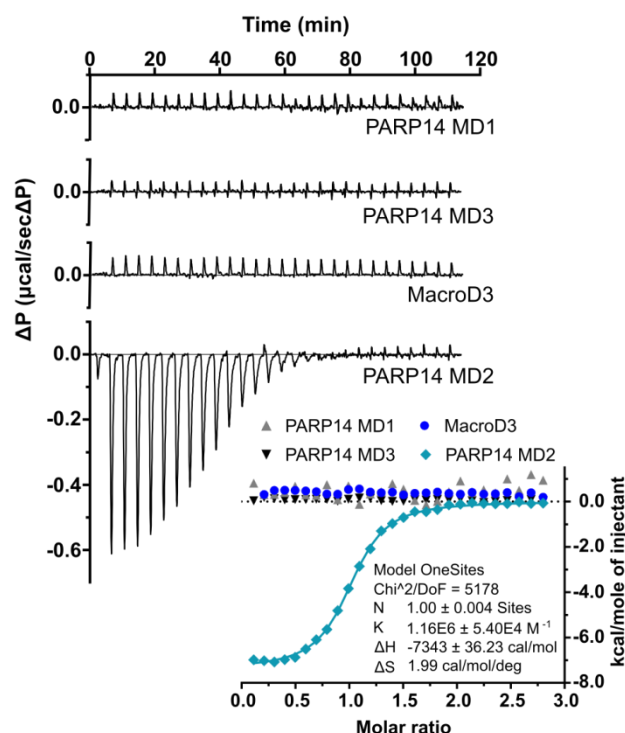


Figure S8 [related to Figures 18 and 20]. *GeA-69 binding to PARP14 MD2 and its selectivity over PARP14 MD1, PARP14 MD3 and MacroD3 tested by ITC.* Raw injection heats are shown resulting from stepwise addition of the corresponding macrodomain protein (300 μM) into a 10 times less concentrated solution containing 30 μM GeA-69 in the presence of 50 mM HEPES (pH 7.4), 300 mM NaCl, 5 % glycerol, 0.5 mM TCEP at 12 $^{\circ}\text{C}$. Binding could only be detected to PARP14 MD2. The inset shows the normalised binding enthalpies corrected for the heat of dilution as a function of binding site saturation, where the solid line represents a nonlinear least squares fit using a single-site binding model. The ITC binding parameters for the GeA-69 titration against PARP14 MD2 are provided in this inset.

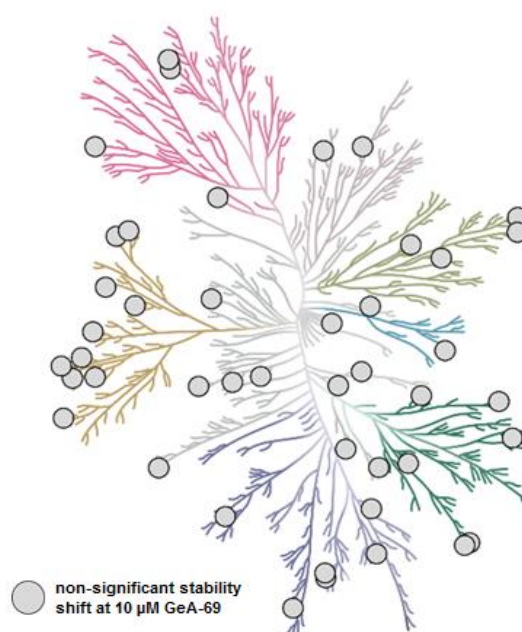


Figure S9. *Assessment of GeA-69 selectivity over the protein kinase family using differential scanning fluorimetry (DSF).* GeA-69 (with an assay concentration of 10 μM) did not reveal any significant temperature shift to any of the 46 screened kinases. The DSF screening data for GeA-69 was retrieved from the SGC internal database.

		10	20	30	40	50	60	70						
PARP14 MD2	1003	-WEKGS	LVSPGG	LQMLLVKE	-GVQNAK	---	-TDVV	-VNS	VPLDLV	---	-LSRGP	LSKSLLEK	AGPELQ	EEL
PARP14 MD1	791	-KCF	SRTVLAP	GVVLIVQQG	-DLARLP	---	-VDVV	-VNAS	NEDLK	---	-HYGG	LAALAA	SKAAGPEL	QADC
PARP14 MD3	1216	DSGVY	EMK-IGS	IIFQV	ASG-DITKEE	---	-ADV	I-VNST	SNSFN	---	-LKAGV	SKAILE	CAQGNV	EREC
PARP9 MD1	107	-LQV	FRKMLT	PRIELSV	WKD-DLTTHA	---	-VDAV	-VNA	ANEDLL	---	-HGGG	LALALV	KAGGFE	IQEES
PARP9 MD2	306	TPSF	NAMV-VNN	LTQIVQ	G-HIEWQT	---	-ADV	I-VNS	VNPHDI	---	-TVGP	VAKSIL	QQAGVEM	KSEF
PARP15 MD1	78	-NVV	ASIQTK	EGLNLKL	ISG-DVLYIW	---	-ADV	I-VNS	VPMNLQ	---	-LGGG	PLSRAFL	QKAGPML	QKEL
PARP15 MD2	293	CFTAY	EMK-IGA	ITFQV	ATG-DIATEQ	---	-VDV	I-VNST	ARTFN	---	-RKSGV	SRAILE	GAGQAVE	SEC
MacroD1	141	-EPY	KKDKQL	NEKISLLRS	-DITKLE	---	-VDA	I-VNA	ANSLL	---	-GGGV	VDCI	HRAAGPLL	TDEC
MacroD2	59	-QET	SQVKKSL	TEKVS	LYRG-DITLLE	---	-VDA	I-VNA	ANASLL	---	-GGGV	VDCI	HRAAGPCL	LAEC
MacroD3	43	-RSP	FLYNK	DVNGK	VVLWKG-DVALLN	---	-CTA	I-VNT	SNESLT	---	-DKNP	VSESI	FMLAGPDL	KEDL
macroH2A1.1	162	-TVL	STKSL	FLGQK	LQVQVQA-DIASID	---	-SDA	V-VHPT	NTDFY	---	-IGGE	VGNLT	EKKGGKEF	VEAV
macroH2A1.2	184	-TVL	STKSL	FLGQK	LNL IHS-EISN	LAGF-EVEAI	-INPT	NADID	---	-LKDD	LGNLT	EKKGGKEF	VEAV	
macroH2A2	184	-TIL	SSKSL	VLGQK	LSLTQS-DISH	IGSM-RVEGI	-VHPT	TAEID	---	-LKED	IGKALE	KAGGKEF	LETV	
TARG1	2	--A	SSLN	EDPE	GSRTYV	VKG-DL	FACPK	---	-TDSL	-AHC	ISED	CR---	-MGAG	IAVL
ALC1	704	-SAEL	DYQDP	DATSL	KYVSG-DVTH	PQAGA	EDAL	I-VHC	VDDSGH	---	-WGRG	GLFT	ALEKRS	-EPKRIY
PARG	714	-----	PLTRL	HVTY	EGTIEENG	---	-QGML	QVDF	ANRF	VGGG	VTSAGL	VQEEIR	FLINPEL	IISR
		80	90	100	110	120	130	140						
PARP14 MD2		DTVG	QG-VAV	SMGT	VLKTSSW	-----	-----	NLDC	R-----	YVLH	VVAPEWR			
PARP14 MD1		DQIV	KREG	RLLP	GNATISKAG	-----	-----	KLPH	Y-----	HVIH	AVGPRWS			
PARP14 MD3		SQQA	----	QQRK	NDYIITGGG	-----	-----	FLRC	K-----	NIHV	IGG---			
PARP9 MD1		KQFV	ARYG	KVSAGE	IAVTGAG	-----	-----	RLPC	K-----	QIIH	AVGPRWM			
PARP9 MD2		LATK	AK-QF	QSQL	VLVTGKF	-----	-----	NLFC	K-----	YIYH	VLWHSEF			
PARP15 MD1		DDRR	RE-T	EKVGN	IFMTSGC	-----	-----	NLDC	K-----	AVLH	AVAPYWN			
PARP15 MD2		AVLA	----	AQPH	RDFIITPGG	-----	-----	CLCK	K-----	IIHV	PGG---			
MacroD1		RTL	----	QSKT	GKAKITGGY	-----	-----	RLPA	K-----	YVHT	VGP IAY			
MacroD2		RNL	----	NGCD	TGHAKITCGY	-----	-----	DLPA	K-----	YVHT	VGP IAR			
MacroD3		QKL	----	KGCR	TGEAKLTGKF	-----	-----	NLAAR	-----	FIIH	TGPKYK			
macroH2A1.1		LELR	KKNGP	LEVAGA	AVSAGH	-----	-----	GLPA	K-----	FVHC	NSPVWG			
macroH2A1.2		LELR	KKNGP	LEVAGA	AVSAGH	-----	-----	GLPA	K-----	FVHC	NSPVWG			
macroH2A2		KELR	KSQGP	LEVAE	AAVSQSS	-----	-----	GLAA	K-----	FVHC	CHIPQWG			
TARG1		Q-----	-QKKS	GVEAV	LKRD	-----	-----	GR-----	-----	YIYL	ITKKRA			
ALC1		ELAG	KM-KD	LSLGG	VLLFPVD	-----	-----	-DKES	RNKGQ	DLAL	IVAQHRD			
PARG		LFTE	----	-VLDH	NECLIITGTE	QYSEY	TGYAET	YRWSR	SHEDG	SERDD	WQRRCT	----	-EIVA	
		150	160	170	180	190	200	210						
PARP14 MD2		NG----	STSSL	KIMED	II	IRECME	ITES	---	-LSLK	SI	FPA	IGTGNL	GFPK	NI
PARP14 MD1		GY----	EAPRC	VYLL	RRAV	QLSL	CLA	EK	---	-YKYR	SI	IPA	ISSGV	FGL
PARP14 MD3		-----	-NDV	KSSV	SSVL	QCEKE	---	-KNYS	SI	CLPA	IGT	GNAK	QHPD	KVA
PARP9 MD1		EW----	-DKQ	GCTG	KLQRA	IVSIL	NYVI	YKNTH	IKTVA	IPAL	SSG	IFQF	PLNL	CTKT
PARP9 MD2		P-----	-KPQ	ILK	HAMKE	CLEK	ECIE	---	-QNI	TSIS	FPAL	GTGN	MEIK	KETA
PARP15 MD1		NG-----	-AET	SWQ	IMANI	IKKCL	TTVE	-LSF	SSIT	FFPM	IGT	GS	LQFP	KAVFA
PARP15 MD2		-----	-KDV	RKT	VT	SVLE	ECEQ	---	-RKYS	TSVSL	PA	IGT	GNAG	KNP
MacroD1		GE----	-PSA	SQA	AELR	SCYL	SSLD	LLLE	---	-HRL	RSVAF	PCIST	GVFG	YPCE
MacroD2		GH----	-ING	SHKED	L	ANCY	KSSL	KLVE	---	-NNIR	SVAF	PCIST	GVFG	YPNE
MacroD3		SR----	-YRT	AAESS	LYSC	YRNVL	QLAKE	---	-QSM	SVGF	CV	INS	AKRG	PLED
macroH2A1.1		AD-----	-KCE	EL	EKT	VKNCL	ALADD	---	-KKL	KSIA	FP	SIG	SRNG	FPKQ
macroH2A1.2		AD-----	-KCE	EL	EKT	VKNCL	ALADD	---	-KKL	KSIA	FP	SIG	SRNG	FPKQ
macroH2A2		SD-----	-KCE	EQL	EET	IKNCL	SAAD	---	-KKL	KSIA	FP	SGR	NCFP	KQTA
TARG1		SH-----	-KPT	YENL	QKSLE	AMKSH	CLK	---	-NGV	DL	SMP	IGC	GLDRL	QWEN
ALC1		RS-NV	LSG	IKMA	ALEGL	KKIF	LA	AKK	---	-KKAS	VHL	PR	IGHAT	KGFN
PARG		RYLD	QFVPE	KMRRE	L	NKAY	CGFL	RP	GVSS	ENLSA	VATG	NWGC	GA	FGDARL
		220	230	240										
PARP14 MD2		-KTL	QEVH	FLL	HPS	DHEN	NIQAF	SDEF	FARRA	-----	1190			
PARP14 MD1		GHCL	KEIY	LVDVS	---	-EKT	VEAF	AEAV	KTVF	-----	978			
PARP14 MD3		-QSV	KKV	VVIFL	---	-PQV	LDV	FYAN	MKKRE	-----	1387			
PARP9 MD1		MSNL	KEI	HVSNE	---	-DPT	VAAF	KAASE	FIL	-----	296			
PARP9 MD2		-KHQ	LTVK	FVIF	PTDLE	IYKAF	SSEMA	KRS	-----	487				
PARP15 MD1		TSPL	QEVH	FV	YTNDD	EGCQ	AF	LE	FTNWS	-----	267			
PARP15 MD2		-PSL	KTVK	VVIFQ	---	-PELL	NIF	YDSM	KKRD	-----	464			
MacroD1		-DKV	DRL	IICV	FLL	---	-EKD	EDI	YRSRL	PHYF	-----	322		
MacroD2		-HEV	DRI	IFCV	FL	---	-EVD	FI	YKKM	NEFF	-----	240		
MacroD3		-ET	IEK	VVFAVS	---	-DLE	EET	YQKLL	PLYF	-----	223			
macroH2A1.1		-SS	IKTV	FVFLD	---	-SES	IGI	YVQ	EMAKLD	-----	369			
macroH2A1.2		-SS	IKTV	FVFLD	---	-SES	IGI	YVQ	EMAKLD	-----	370			
macroH2A2		-SSL	KNVY	FLLFD	---	-SES	IGI	YVQ	EMAKLD	-----	370			
TARG1		---	-DIK	ITYTL	-----	-----	-----	-----	-----	-----	152			
ALC1		---	-GIPT	YIYFP	---	-RSK	SAVL	HAQSS	SSSR	QLVP	-----	897		
PARG		---	-ERD	VVYFT	FG	-DSEL	MRDI	YSM	HIFL	TERK	-----	922		

Figure S10 [related to Figure 20]. Multiple sequence alignment by Promals3D of human macrodomains. Residues of PARP14 MD2 involved in GeA-69 interaction are highlighted in cyan, corresponding residues of other macrodomains in blue.

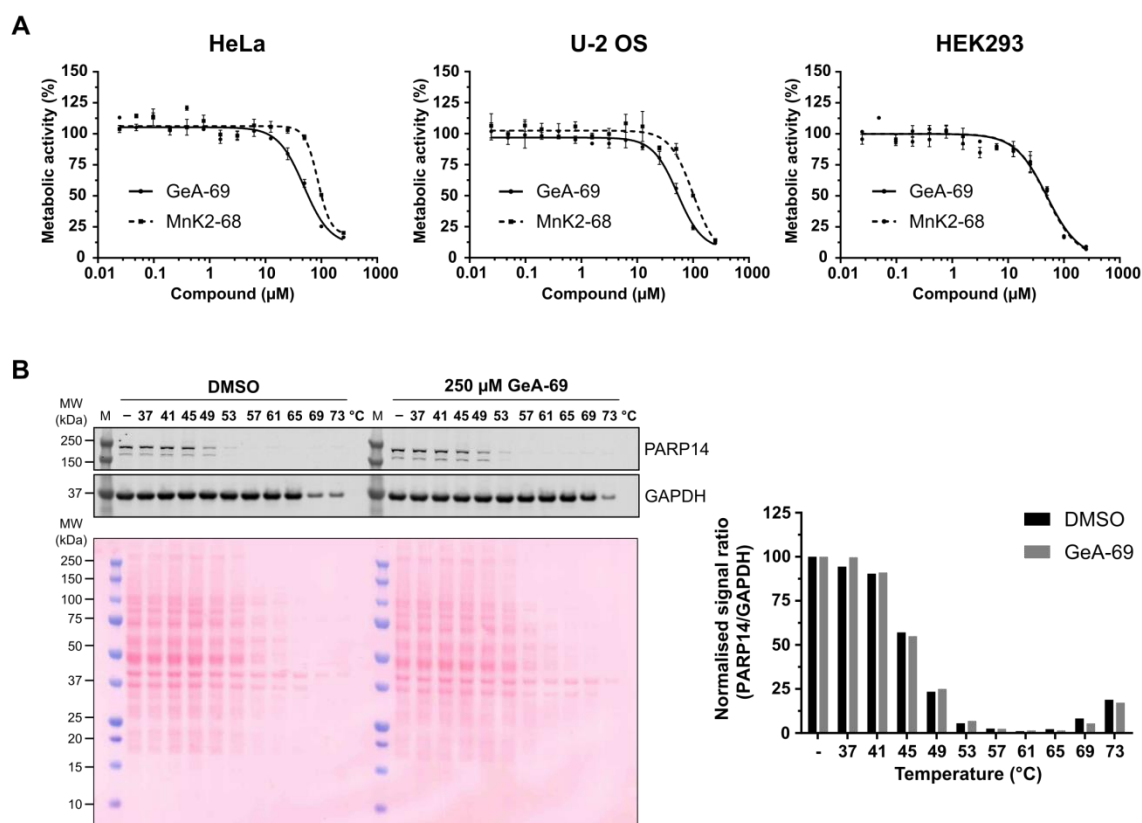


Figure S11 [related to Figure 21]. (A) GeA-69 and MnK2-68 show similar moderate cytotoxicity. HeLa, U-2 OS and HEK293 cells were incubated for 72 h in the presence of different GeA-69 or MnK2-68 concentrations (25 nM – 250 μM) and cell viability was measured using the AlamarBlue™ reagent. Representative graphs for each cell line are shown. Data points are mean $\pm$ SEM, $n = 3$. EC_{50} values were calculated from three (GeA-69) and two (MnK2-68) independent experiments: $\text{EC}_{50}(\text{GeA-69})_{\text{HeLa}} = (58 \pm 3.9) \mu\text{M}$, $\text{EC}_{50}(\text{GeA-69})_{\text{U-2 OS}} = (52 \pm 1.1) \mu\text{M}$, $\text{EC}_{50}(\text{GeA-69})_{\text{HEK293}} = (54 \pm 2.2) \mu\text{M}$, $\text{EC}_{50}(\text{MnK2-68})_{\text{HeLa}} = (97 \pm 10) \mu\text{M}$, $\text{EC}_{50}(\text{MnK2-68})_{\text{U-2 OS}} = (107 \pm 6.6) \mu\text{M}$, $\text{EC}_{50}(\text{MnK2-68})_{\text{HEK293}} = (48 \pm 0.5) \mu\text{M}$. Results are expressed as mean $\pm$ SEM. **(B)** Repeat of a CETSA experiment performed with GeA-69 showing additionally the Ponceau S staining after Western blot transfer of the proteins at the bottom left.

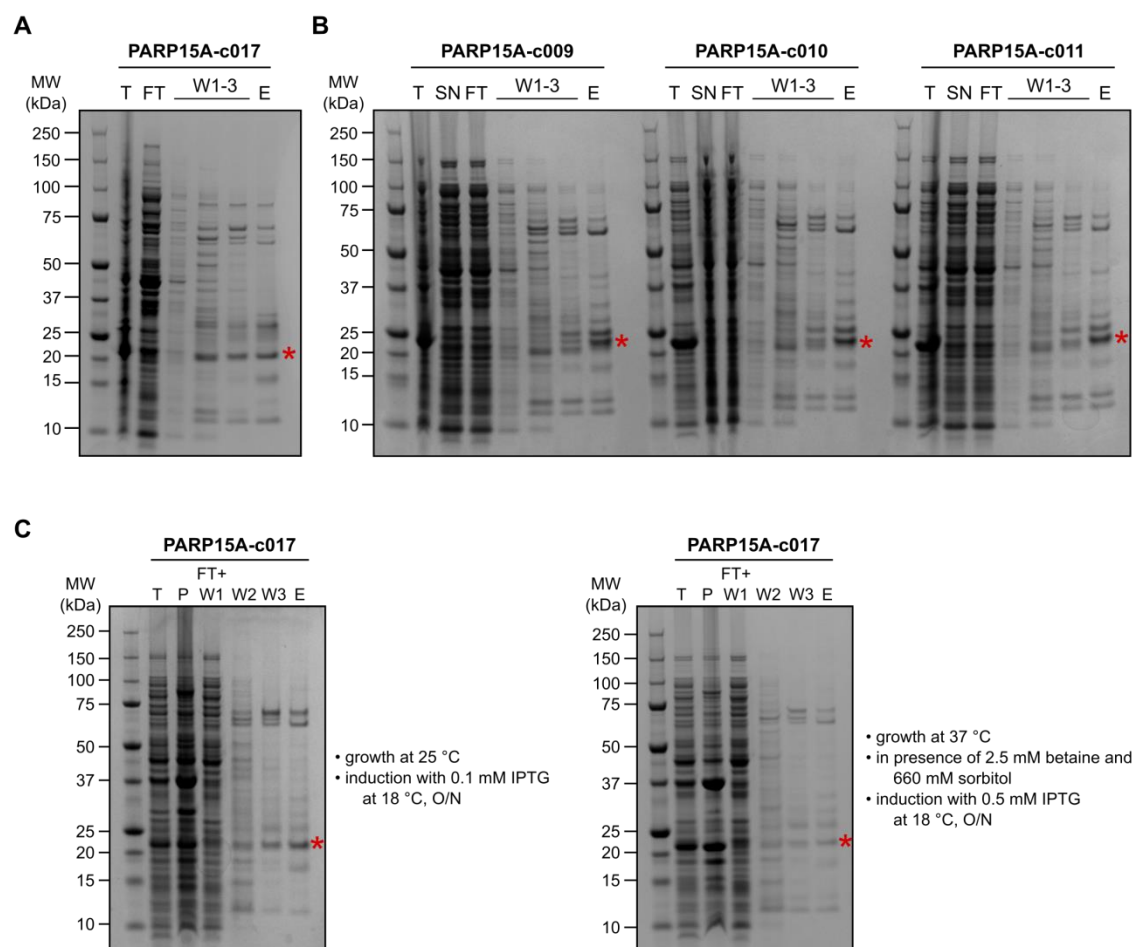


Figure S12 [related to Figure 24]. (A) and (B) Large scale purification attempts of PARP15 MD1 proteins expressed in *E. coli* from different expression constructs. All construct tested did not yield in sufficient amounts of soluble protein which could be purified further. On each gel (also in (C)), the red star marks the assumed protein band for PARP15 MD1. **(C)** Two variations of the standard expression conditions were tried in large scale with PARP15A-c017, none of them resulting in an increased yield of soluble PARP15 MD1. (T = total cell lysate, P = pellet obtained after centrifugation of the lysed cells, FT = flow-through, W1-3 = washes 1-3 with increasing imidazole concentrations, E = eluate)

Tryptic digest of PARP15 MD1 sample c024-S2

protein sequence coverage: 49 %

blue: identified peptides matching with PARP15 MD1

```

1  MGSPILGYWKIKGLVQPTRLLEYLEEKYEEHLYERDEGDKWRNKKFELG
51  LEFPNLPYYIDGDVKLTQSMAIIRYIADKHNMLGGCPKERAIEISMLEGAV
101 LDIRYGVSR IAYSKDFETLKVDFLSKLPMLKMFEDRLCHKTYLNGDHVT
151 HPDFMLYDALDVVLYMDPMCLDAFPKLVCFKKRIEAIPOIDKYLKSSKYI
201 AWPLQGWQATFGGGDHPKSDLEVLFGGPLSMDNRNKFSSKDCLSIIRNVV
251 ASIQTKEGNLKLIISGDVLYIWADVIIVNSVPMNLQGGGPLSR AFLQKAG
301 PMLQKELDDR RRETEEKVGNIFMTSGCNLDC KAVLHAVA PYWNGAETSW
351 QIMANI IKKCLTTVEVLSFSSITFPMIGTGS LQFPKAVFAKLILSEVFY
401 SSSTRPITSPLQEVHFLVYTNDDEGCQAFLEFTNWSRINPNKA

```

Intact mass of PARP15 MD1 sample c024-S2

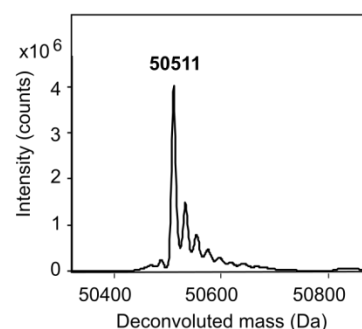


Figure S13 [related to Figure 25]. Tryptic digest result (left) and intact mass analysis (right) by mass spectrometry confirmed the identity of PARP15 MD1 in the sample c024-S2. The theoretical mass for PARP15 MD1 from construct PARP15A-c024 is 50600 Da. The mass difference of 89 Da indicates methionine loss with protein acetylation. This protein modification was also observed in the previous test expression of this construct performed by the SGC Biotechnology group.

	Compound ID	QC control (new stocks)	
		Identity confirmed	c (mM)
1	NVP-BMM025-AA-1	YES	10.4
2	NVP-BNP116-NX-1	YES	24.6
3	NVP-BMF069-NX-1	YES	4.3
4	NVP-BMW608-NX-1	YES	0.3
5	NVP-BMD153-NX-1	YES	8.0
6	NVP-BMC079-NX-1	YES	5.1
7	NVP-BNP071-NX-1	YES	12.8
8	NVP-BNC609-AA-1	NO	0.0
9	NVP-BMC081-NX-1	YES	12.5
10	NVP-BMB730-NX-1	---	---
11	NVP-BME758-NX-1	YES	32.3
12	NVP-BMW822-NX-1	YES	12.1
13	NVP-BNH686-NX-1	YES	2.1
14	NVP-BMJ791-NX-1	YES	10.9
15	NVP-BNC605-AA-1	YES	8.0
16	NVP-BMP569-NX-1	YES	1.2
17	NVP-BNC594-NX-1	YES	7.2
18	NVP-BME919-NX-1	YES	5.8
19	NVP-BML846-NX-1	YES	9.6
20	NVP-BMU249-NX-1	YES	29.9
21	NVP-BNC592-NX-1	NO	0.0

	Compound ID	QC control	
		Identity confirmed	c (mM)
22	NVP-BNU459-NX-1	YES	5.3
23	NVP-BNV428-NX-1	YES	9.1
24	NVP-BNV427-NX-1	YES	10.9
25	NVP-BNU457-NX-1	YES	6.4
26	NVP-BNM150-NX-1	YES	6.0
27	NVP-BNX029-NX-1	YES	9.4
28	NVP-BNU542-NX-1	YES	8.0
29	NVP-BNU540-NX-1	YES	10.9
30	NVP-BNV433-NX-1	YES	8.9
31	NVP-BNQ598-NX-1	YES	7.7
32	NVP-BMX995-NX-1	YES	5.1
33	NVP-BMY002-NX-1	NO	0.0
34	NVP-BNV425-NX-1	YES	9.1
35	NVP-BMF352-NX-1	YES	2.7
36	NVP-BNV413-NX-1	YES	8.5
37	NVP-BMY031-NX-1	YES	9.4
38	NVP-BNV442-NX-1	YES	9.4
39	NVP-BNV240-NX-1	YES	12.1
40	NVP-BNV410-NX-1	YES	11.3
41	NVP-BNL153-NX-1	YES	11.6

Figure S14 [related to Figures 33 and 35]. *MS QC controls of stock solution (nominal concentration: 10 mM) of selected hit compounds from 45,000 compound screen against PARP14 MD2. (Left)* Hit compounds from first (30,000) compound screen. The QCs measurements from the new (re-ordered) stock solutions are shown. *(Right)* Hit compounds from second (15,000) compound screen.

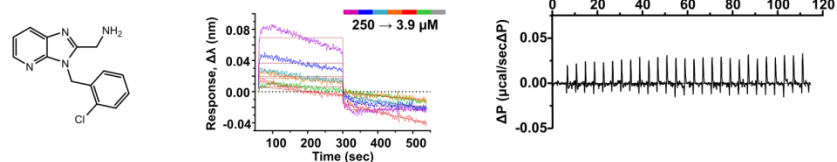
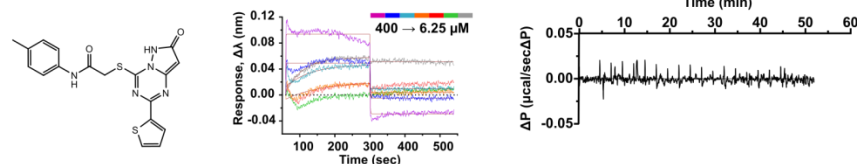
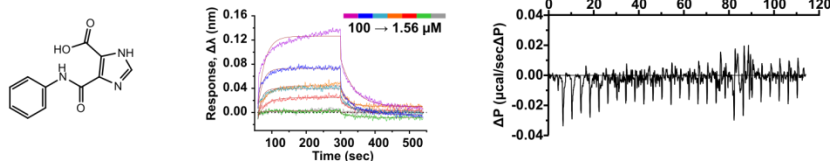
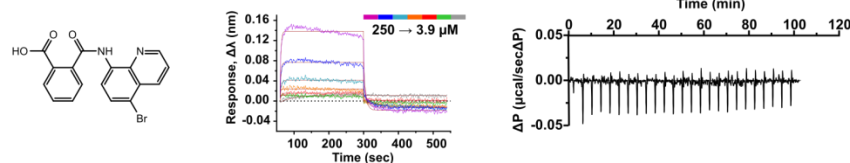
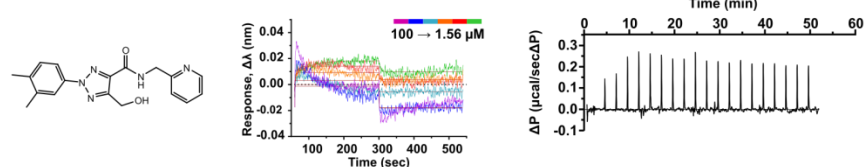
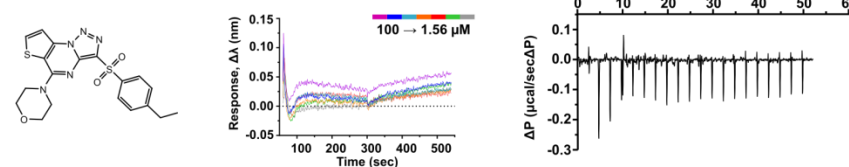
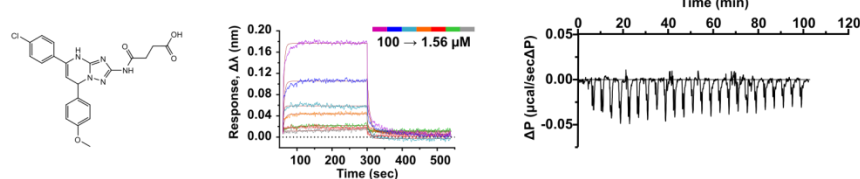
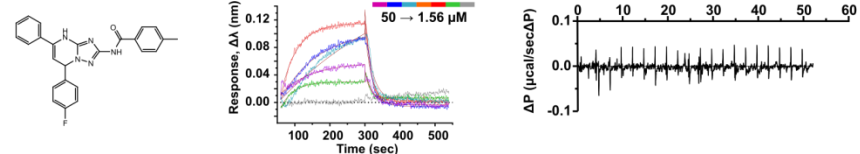
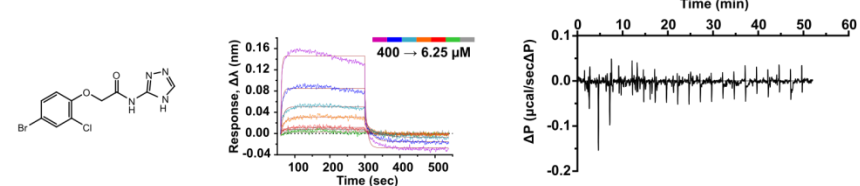
NVP-BMM025-AA-1**NVP-BNP116-NX-1****NVP-BMF069-NX-1****NVP-BMD153-NX-1****NVP-BMW822-NX-1****NVP-BNH686-NX-1****NVP-BNC592-NX-1****NVP-BNC594-NX-1****NVP-BME919-NX-1**

Figure S15. Hit validation by BLI and ITC from 30,000 compound screen. Compounds were characterised by BLI at top assay concentrations ranging from 50-400 μM as indicated and by ITC by titrating PARP14 MD2 into a 10 times less concentrated solution containing the corresponding hit compound in a concentration at 50 μM . Exceptions are NVP-BNC592-NX-1, NVP-BMD153-NX-1 which were screened at 40 μM and NVP-BMW822-NX-1 that was screened at 80 μM . All titrations were performed at 20 $^{\circ}\text{C}$. BLI measurements were only performed with old compound stock solutions. ITC titrations were performed with the new compound stock solutions.

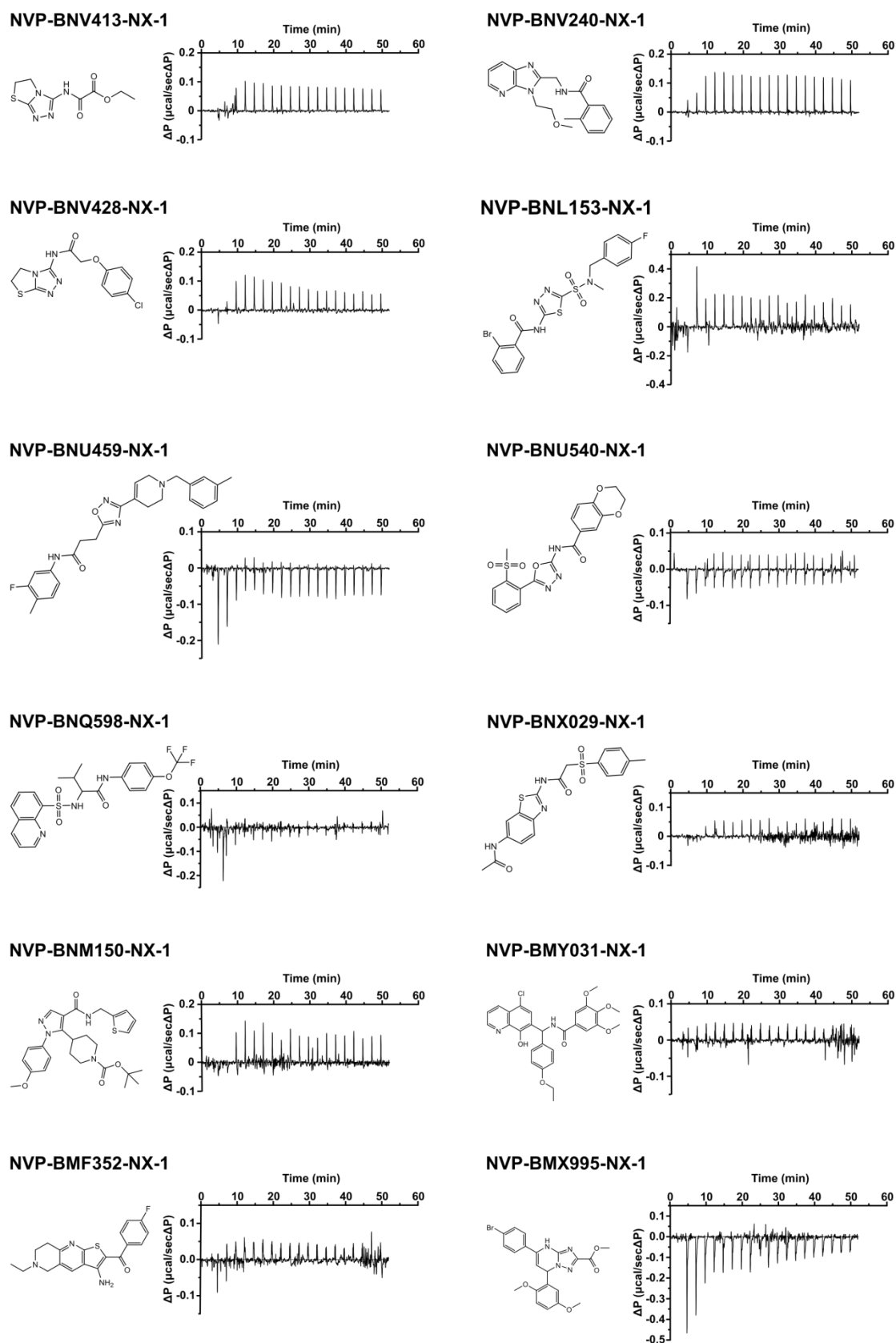


Figure S16. Hit validation by ITC from 15,000 compound screen. Compounds were characterised by ITC by titrating PARP14 MD2 into a 10 times less concentrated solution containing the corresponding hit compound at 50 μM . Exceptions are NVP-BNV240-NX-1 and NVP-BNL153-NX-1 which were screened at a concentration of 80 μM and NVP-BMX995-NX-1 that was screened at 70 μM . All titrations were performed at 20 $^{\circ}\text{C}$.

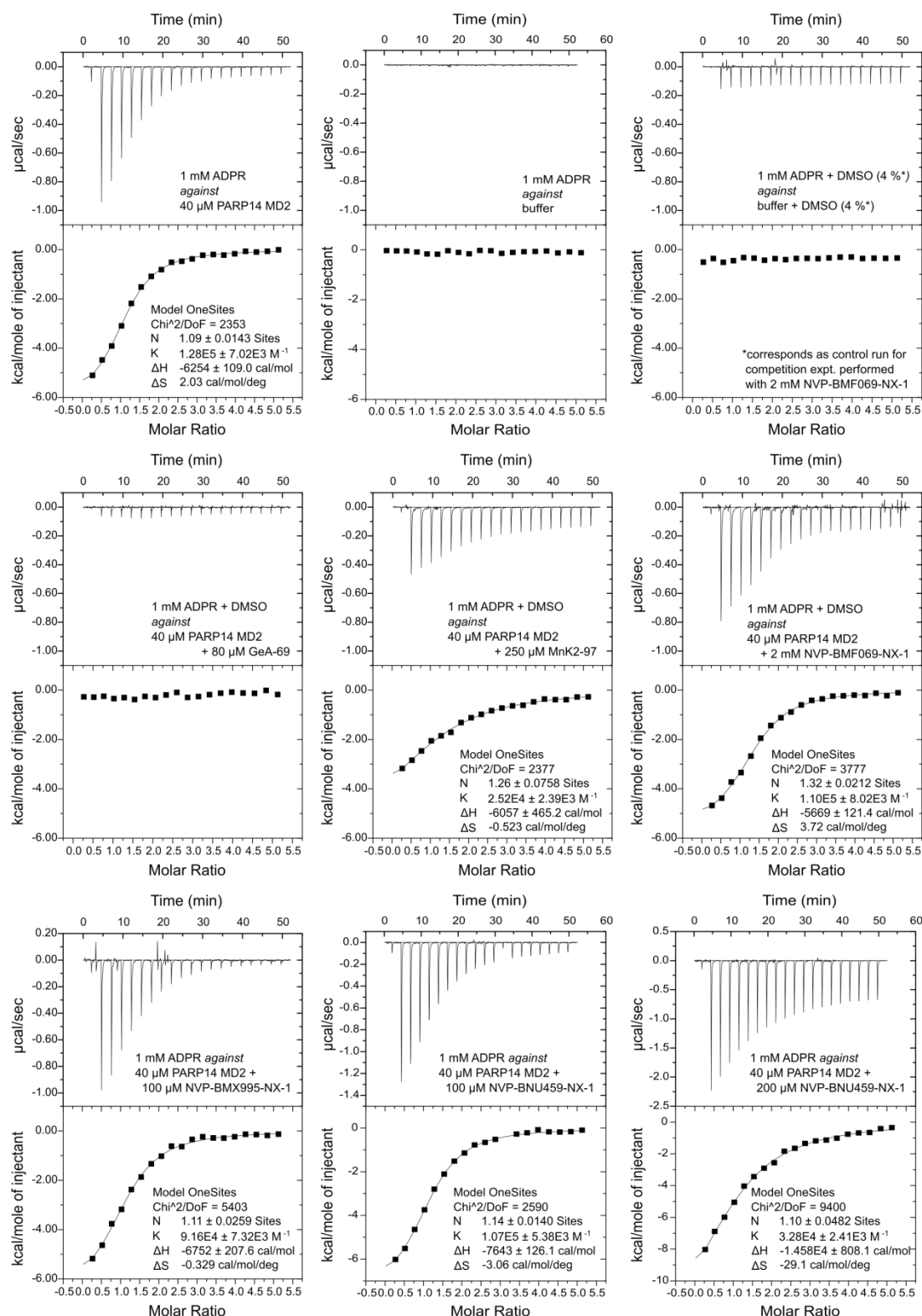


Figure S17 [related to Figures 19, 29, 34 and 36]. Summary of ITC competition experiments performed in this work.

The ITC data obtained from the Origin v7.0 software with the MicroCal plugin is shown for the competition experiments performed for characterisation of GeA-69, MnK2-97, NVP-BMF069-NX-1, NVP-BMX995-NX-1 and NVP-BNU459-NX-1 for their ability to compete with ADPR binding to PARP14 MD2. The upper part of the panels show the raw titration curves resulting from stepwise addition of the ligand ADPR (1 mM) into buffer or a solution containing PARP14 MD2 in a 25-fold lower concentration in the presence of 50 mM HEPES (pH 7.4), 300 mM NaCl, 5 % glycerol, 0.5 mM TCEP at 20 °C. The lower part of the panels show the normalised binding enthalpies corrected for the heat of dilution as a function of binding site saturation. Solid lines represent a nonlinear least squares fit using a single-site binding model.

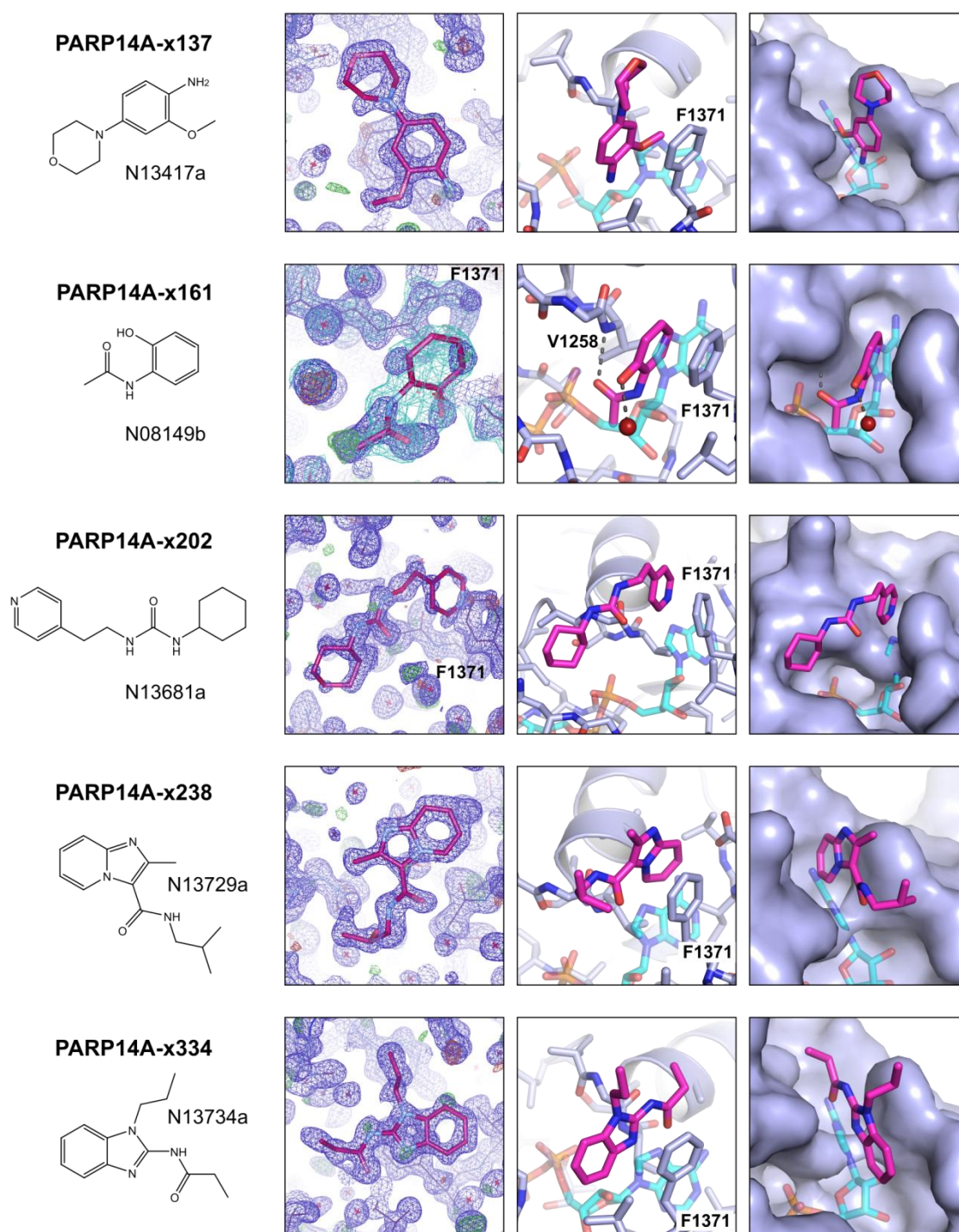


Figure S18-1 [related to Figure 42]. Fragment hits obtained in the crystallographic PARP14 MD3 fragment screen.

The molecular structure of the fragment identified in the corresponding crystal is shown on the left. The left images from the panel show the fragments in the crystallographic model. The $2F_o - F_c$ electron density map contoured at 1.0σ is shown in blue. For weaker fragments, the calculated *PanDDA* event map countoured at 2.0σ is additionally shown in cyan. The $F_o - F_c$ electron density map, green/red, is contoured at $\pm 3.0\sigma$. The images in the middle of the panel show the binding site of the fragments in PARP14 MD3 which residues are shown as atom-coloured sticks. Hydrogen bonds between ligand and protein or structural waters are indicated by dotted lines. As reference for the position of the fragment relative to the ADPR in the PARP14 MD3 binding site, PARP14 MD3 in complex with ADPR (PDB ID 4ABK)⁶⁶ was superimposed with the corresponding fragment structure and the ADPR shown as atom-coloured sticks in low transparency in the model. The images on the right show the surface representation of PARP14 MD3 in complex with the corresponding fragment.

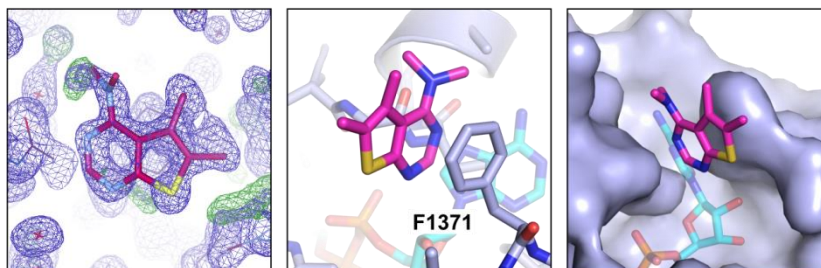
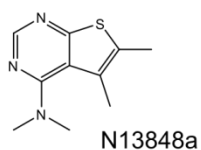
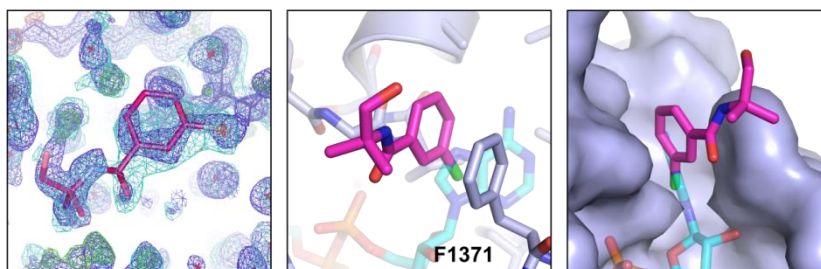
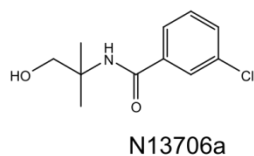
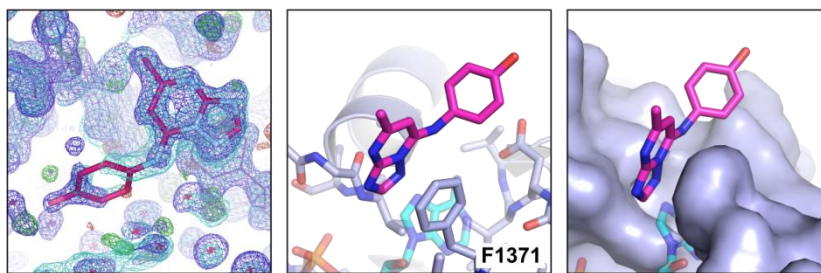
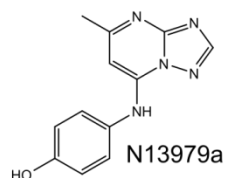
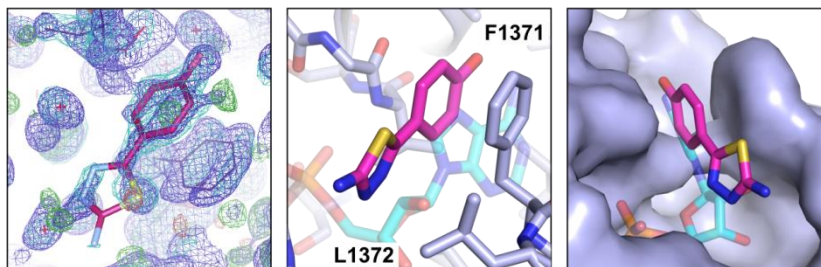
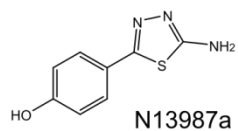
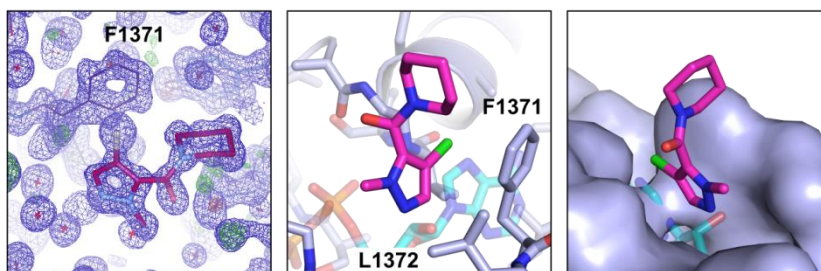
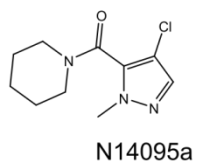
PARP14A-x484**PARP14A-x523****PARP14A-x567****PARP14A-x637****PARP14A-x690**

Figure S18-2 [Figure S18-1 continued, related to Figure 42]. *Fragment hits obtained in the crystallographic PARP14 MD3 fragment screen.*

Table S1. Data collection and refinement table for PARP14 MD3 fragment structures shown in Figure S18. The PARP14 MD3 *apo* structure (PDB ID 4ABL)⁶⁶ was used to build the reference model by molecular replacement. The fragment RSCC (real-space correlation coefficient) describes the correlation between the model and the observed density (recommended value > 0.7). The B-factor ratio measures the consistency of the fragment model with surrounding protein residues, and is calculated from the B-factors of the fragment atoms and the fragment surrounding side-chain atoms within 4 Å (recommended value ~1.0). r.m.s.d.: coordinate root-mean-squared deviation. ^a Values within parentheses refer to the highest resolution shell.

	x137	x161	x202	x238	x334
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
No. of molecules in the asymmetric unit	1	1	1	1	1
Unit cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	34.2, 41.4, 111.4	34.3, 41.5, 110.7	34.1, 41.7, 111.0	34.1, 41.4, 111.0	34.3, 41.7 111.2
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Data collection					
Resolution range (Å) ^a	55.7 - 1.05 (1.09 - 1.05)	55.4 - 1.05 (1.09 - 1.05)	29.0 - 1.05 (1.09 - 1.05)	29.0 - 1.12 (1.16 - 1.12)	39.0 - 1.05 (1.09 - 1.05)
Unique observations ^a	53220 (751)	61733 (1325)	63476 (1467)	55656 (2803)	64092 (1487)
Average multiplicity ^a	2.0 (1.6)	1.9 (1.1)	1.9 (1.1)	2.0 (1.9)	1.8 (1.2)
Completeness (%) ^a	71.3 (10.2)	82.8 (18.0)	85.2 (20.0)	90.8 (46.5)	85.3 (20.0)
<i>R</i> _{merge} ^a	0.02 (0.20)	0.02 (0.22)	0.02 (0.09)	0.02 (0.21)	0.02 (0.10)
Mean (<i>I</i>)/ σ (<i>I</i>) ^a	21.6 (2.5)	21.5 (1.3)	21.7 (2.6)	23.1 (2.3)	26.4 (3.1)
Mean CC(1/2) ^a	0.999 (0.939)	0.999 (0.889)	0.998 (0.983)	1 (0.944)	0.998 (0.967)
Refinement					
<i>R</i> -value, <i>R</i> _{free} (%)	19.6, 22.2	18.7, 20.6	20.3, 21.8	18.8, 21.7	19.5, 21.4
r.m.s.d. from ideal bond length (Å)	0.010	0.010	0.010	0.010	0.009
r.m.s.d. from ideal bond angle (°)	1.17	1.20	1.16	1.17	1.12
Ramachandran plot (favoured/allowed) (%)	97.8 / 100	97.8 / 100	96.7 / 100	97.8 / 100	96.7 / 100
Fragments					
	N13417a	N08149b	N13681a	N13729a	N13734a
Occupancy	0.57	0.54	0.64	0.52	0.74
r.m.s.d.	22.18	0.32	0.26	0.15	0.66
RSCC	0.83	0.8	0.89	0.91	0.93
B-factor ratio	1.18	1.44	1.25	0.89	1.17

Table S1 [continued].

	x484	x523	x567	x637	x690
Space group	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁ 2 ₁ 2 ₁
No. of molecules in the asymmetric unit	1	1	1	1	1
Unit cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	34.0, 41.5, 111.1	34.5, 41.6, 110.8	34.5, 41.5, 111.1	34.6, 41.7, 110.9	34.3, 41.6, 111.1
α , β , γ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Data collection					
Resolution range (Å) ^a	55.5 - 1.20 (1.25 - 1.20)	55.4 - 1.08 (1.12 - 1.08)	55.5 - 1.10 (1.14 - 1.10)	55.4 - 1.09 (1.13 - 1.09)	29.2 - 1.05 (1.09 - 1.05)
Unique observations ^a	49635 (4783)	61828 (2791)	58605 (2834)	62502 (3460)	61315 (1250)
Average multiplicity ^a	2.0 (1.9)	1.9 (1.2)	1.9 (1.3)	1.9 (1.3)	1.9 (1.2)
Completeness (%) ^a	99.2 (96.4)	89.0 (40.8)	89.0 (43.9)	92.1 (51.5)	81.6 (16.9)
<i>R</i> _{merge} ^a	0.02 (0.25)	0.02 (0.12)	0.01 (0.18)	0.02 (0.14)	0.02 (0.12)
Mean (<i>I</i>)/ σ (<i>I</i>) ^a	19.2 (2.2)	23.6 (2.6)	26.5 (2.4)	24.5 (2.3)	21.7 (3.0)
Mean CC(1/2) ^a	1 (0.923)	0.999 (0.981)	1 (0.93)	0.999 (0.97)	0.997 (0.949)
Refinement					
<i>R</i> -value, <i>R</i> _{free} (%)	19.0, 22.0	20.4, 22.3	19.4, 21.0	20.1, 22.2	19.9, 21.9
r.m.s.d. from ideal bond length (Å)	0.010	0.010	0.010	0.010	0.010
r.m.s.d. from ideal bond angle (°)	1.21	1.16	1.12	1.18	1.16
Ramachandran plot (favoured/allowed) (%)	98.3 / 100	97.8 / 100	97.8 / 100	97.8 / 100	97.8 / 100
Fragments					
	N13848a	N13706a	N13979a	N13987a	N14095a
Occupancy	0.48	0.53	0.51	0.56	0.76
r.m.s.d.	0.21	1.62	0.18	0.19	0.20
RSCC	0.82	0.71	0.84	0.77	0.97
B-factor ratio	1.12	1.81	1.07	1.63	1.26

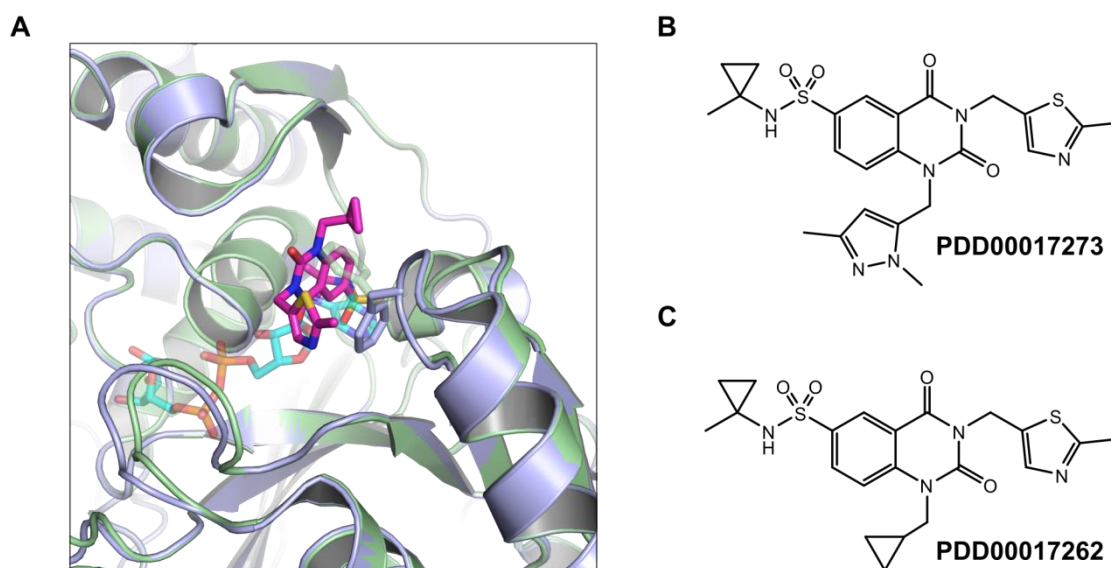


Figure S19. The PARG chemical probe PDD00017273 developed by James *et al.*¹³¹. PDD00017273 targets the same adenine-ribose site of the ADPR binding pocket as the fragments identified targeting PARP14 MD3. **(A)** Superimposition of the PARG macrodomain bound to a PARG inhibitor analogue (blue, PDB ID 5LHB)¹³¹ with the PARG macrodomain bound to ADPR (green, PDB ID 4B1H)²⁶⁹. The PARG inhibitor and the ADPR are shown as atom-coloured sticks as well as their targeted phenylalanine residue Phe902. The ADPR and Phe902 of the ADPR-bound structure are shown with low transparency for better illustration. **(B)** Molecular structure of the PARG inhibitor, PDD00017273. **(C)** Molecular structure of the PARG inhibitor PDD00017273 analogue that was co-crystallised with the PARG macrodomain shown in (A).

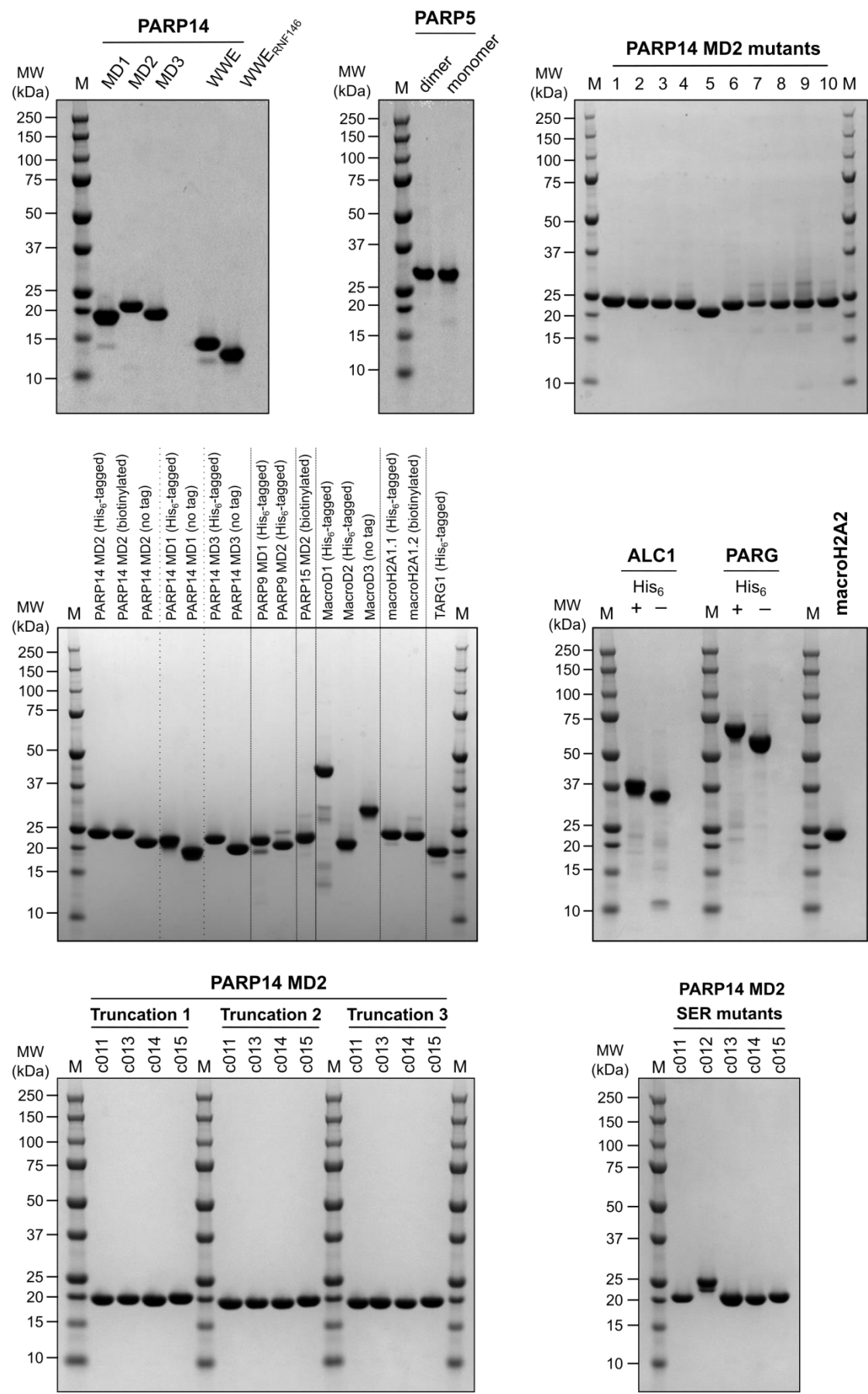
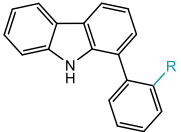


Figure S20. Un-cropped SDS-PAGE gel images showing proteins purified in this work.

Table S2 [related to Figure 28]. Summary of all synthesised GeA-69 derivatives with corresponding IC_{50} values (or K_D values where mentioned) for SAR development. All IC_{50} values were determined by AlphaScreen, only where indicated by BLI or ITC. Compounds marked with (*) were screened by the SGC screening team (Dr. O. Fedorov, SGC, Oxford). Compounds with supplier ID starting with 'MnK' were synthesised by Dr. Kerstin Riedel (Professor F. Bracher laboratory, LMU, Munich), compounds with supplier ID starting with 'MM' were synthesised by Moses Moustakim (Professor P. Brennan group, SGC, Oxford).

Variations of the amide group 			
SGC ID	Supplier ID	R	IC_{50} (AlphaScreen) (μ M)
K03724a-c	GeA-69 (reference)		0.7
K05675a	MnK2-26		1.0*
K05681a	MnK2-38		>50*
K04960a	MnK2-13		0.9
K05677a	MnK2-28		1.2*
K05684a	MnK2-42		2.9*
K05678a	MnK2-29		3.3*
K05256a	MnK2-19		1.0*
K04959a	MnK2-11		5.3*

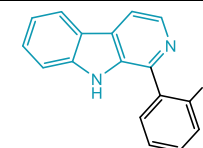
K04962a	MnK2-15		8.3*
K05676a	MnK2-27		1.2*
K04957a	MnK2-9		2.2*
K04953a	MnK2-5		>50*
K05679a	MnK2-36-2		10.5*
K05680a	MnK2-37		>50*
K05683a	MnK2-41		>50*
K05682a	MnK2-39		>50*
K05257a	MnK2-20		>50*
E60887a	MnK2-93		>100 (BLI) >40 (ITC)
E60888a	MnK2-94		>100 (BLI) >25 (ITC)
E62636a	MM_8013		7.6
E62637a	MM_8014		3.6
E62638a	MM_8015		9.0

E62683a	MM_8016		60	E62885a	MM8048H		8.1
E62684a	MM_8018		2.5	E62886a	MM8048L		4.3
E62685a	MM_8019		9.1	E62887a	MM8048W		2.8
E62686a	MM_8020		2.5	E62888a	MM8048D		6.2
E62687a	MM_8021		17	E62889a	MM8048R		7.1
E62688a	MM_8022		no inhibition at 50 μ M	E62890a	MM8048B		6.2
E62689a	MM_8023		9.1	E62891a	MM8048C		no inhibition at 50 μ M
E62690a	MM_8024		7.3	E62892a	MM8048X		6.2
E62880a	MM8048E		1.4	E62893a	MM8048P		8.7
E62881a	MM8048U		0.66, no binding at 40 μ M (ITC) even when prepared as salt; solubility problem	E62894a	MM8048Q		no inhibition at 50 μ M
E62882a	MM8048K		4.4	E62895a	MM8048Z		5.4
E62883a	MM8048ZB		10.9	E62896a	MM8048V		6.9
E62884a	MM8048A		6.6				

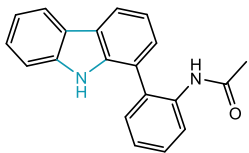
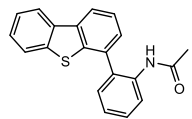
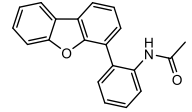
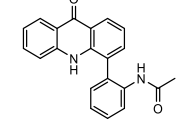
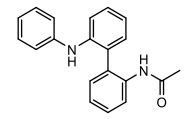
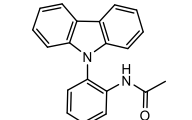
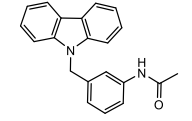
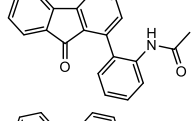
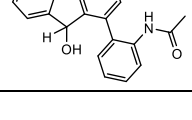
E62897a	MM8048N		6.3
E62898a	MM8048O		2.4
E62899a	MM8048T2		4.2
E62900a	MM8048I2		3.6
E62942a	MM8048ZC		2.1 no binding at 40 μ M (ITC), solubility problem
E62979a	8048ZD		1.2 1.6 (BLI)
E63015a	MM8048ZE		1.1 1.5 (ITC)
E63016a	MM8048ZF		1.0 2.7 (ITC)
E63018a	MM8048ZG		2.1 3.9 (ITC)
E63102a	8051A		2.4
E63103a	8051B		3.5
E63104a	8051C		6.6
E63105a	8051D		3.7
E63106a	8051E		5.2

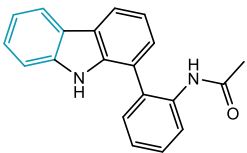
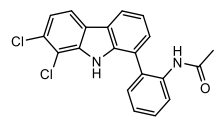
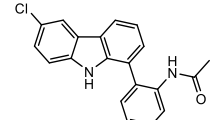
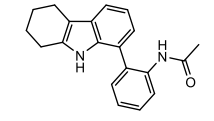
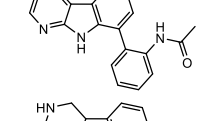
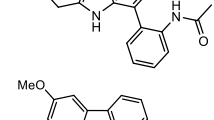
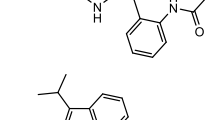
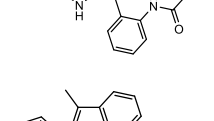
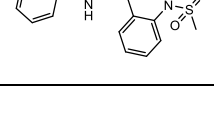
E63168a	8048-ZI		13.1
E63169a	8050-ZL		5.8
E63170a	8048-ZM		no inhibition at 50 μ M
E63171a	8051-F		5.4
E63172a	8051-G		no inhibition at 50 μ M
E63180a	8051-H		2.3
E63192a	8051ZO		no inhibition at 50 μ M
E63193a	8051ZN		no inhibition at 50 μ M
E63194a	8051ZP		no inhibition at 50 μ M

Compounds with β -carboline scaffold



K05259a	TT-V2-74		30*
K05258a	TT-V2-64		>50*

Variations on ring B		
K05743a	MnK2-55	 >50*
K05744a	MnK2-52	 >50*
K05674a	MnK2-25-2	 >50*
K05673a	MnK2-24-6	 >50*
K05260a	MnK HB S1P3	 >50*
K05261a	MnK HB S3P3	 >50*
K05994a	MnK CG6	 no inhibition at 50 μ M
K05995a	MnK CG7	 no inhibition at 50 μ M

Variations on ring C		
K05990a	MnK2-61	 1.9
K05991a	MnK2-64	 2.8
K05255a	MnK2-18	 1.1
K05742a	MnK2-47	 no inhibition at 50 μ M
E60062a	MnK2-79	 binding at 50 μ M not measurable (ITC and BLI)
E60061a	MnK2-75	 binding at 50 μ M not measurable (ITC and BLI)
E60889a	MnK2-97	 18 (ITC and BLI)
E62738a	MM_8039	 >40 (ITC)

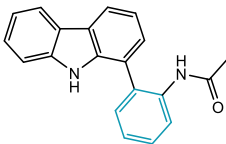
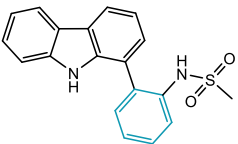
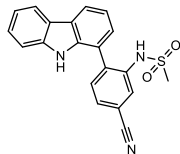
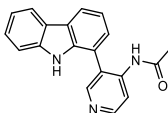
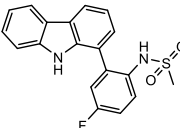
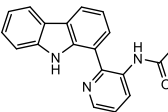
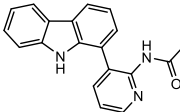
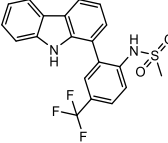
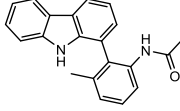
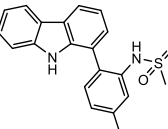
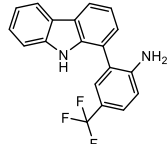
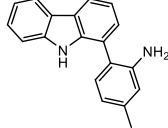
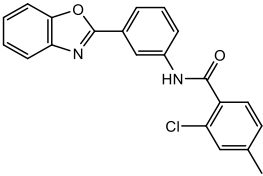
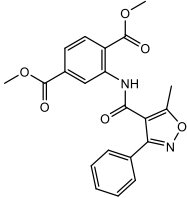
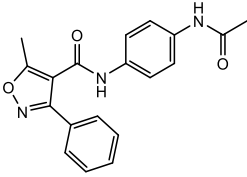
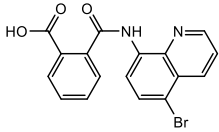
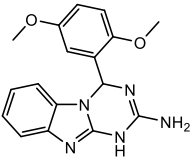
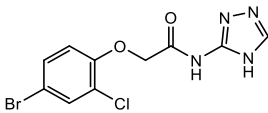
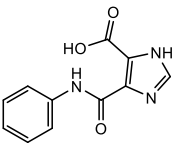
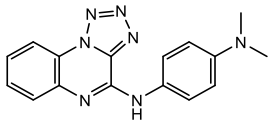
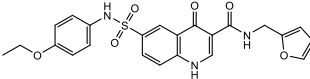
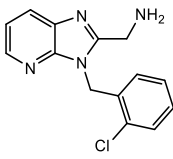
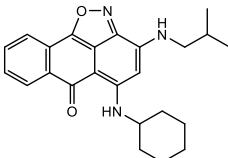
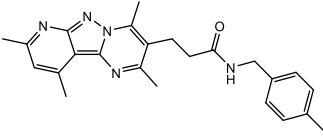
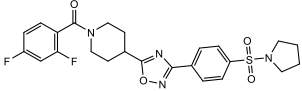
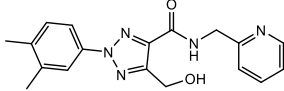
<i>Phenyl analogues</i>						
				E62737a	MM_8031	 5.2 (ITC)
K05992a	MnK2-68		no inhibition at 50 μ M	E62739a	MM_8040	 6.5 (ITC)
K05993a	MnK2-69		no inhibition at 50 μ M	E62777a	MM8041	 binding at 40 μ M not measurable (ITC)
E62701a	MM_8026		no inhibition at 50 μ M	E62810a	MM_8046	 1.7 1.6 (ITC)
E62702a	MM_8027		1.1			
E62703a	MM_8012		no inhibition at 50 μ M			
E62704a	MM_8025		no inhibition at 50 μ M			

Figure S3. Hit compounds from 45,000 compound screen performed at Novartis (NIBR, Basel) selected for further validation by ITC and BLI.

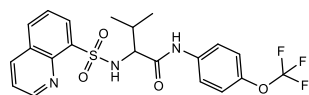
Hit selection (1) from 30,000 compound screen			
Compound ID	Structure	Compound ID	Structure
NVP-BMB730-NX-1		NVP-BMC079-NX-1	
NVP-BMC081-NX-1		NVP-BMD153-NX-1	
NVP-BME758-NX-1		NVP-BME919-NX-1	
NVP-BMF069-NX-1		NVP-BMJ791-NX-1	
NVP-BML846-NX-1		NVP-BMM025-AA-1	
NVP-BMP569-NX-1		NVP-BMU249-NX-1	
NVP-BMW608-NX-1		NVP-BMW822-NX-1	

NVP-BNC592-NX-1		NVP-BNC594-NX-1	
NVP-BNC605-AA-1		NVP-BNC609-AA-1	
NVP-BNH686-NX-1		NVP-BNP071-NX-1	
NVP-BNP116-NX-1			

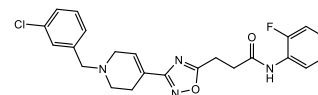
Hit selection (2) from 15,000 compound screen

Compound ID	Structure	Compound ID	Structure
NVP-BMF352-NX-1		NVP-BMX995-NX-1	
NVP-BMY002-NX-1		NVP-BMY031-NX-1	
NVP-BNL153-NX-1		NVP-BNM150-NX-1	

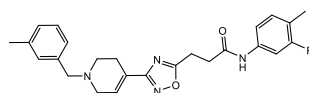
NVP-BNQ598-NX-1



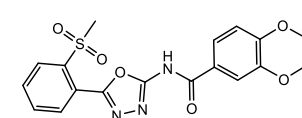
NVP-BNU457-NX-1



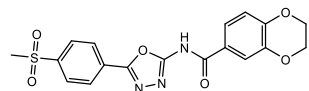
NVP-BNU459-NX-1



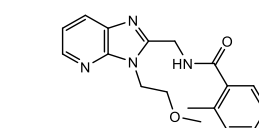
NVP-BNU540-NX-1



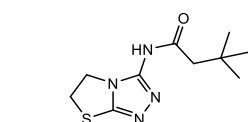
NVP-BNU542-NX-1



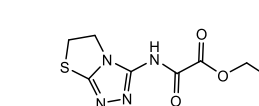
NVP-BNV240-NX-1



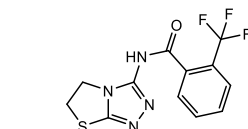
NVP-BNV410-NX-1



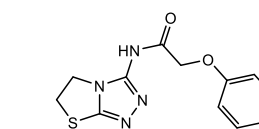
NVP-BNV413-NX-1



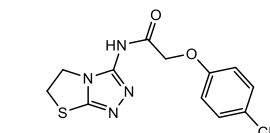
NVP-BNV425-NX-1



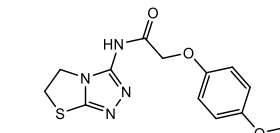
NVP-BNV427-NX-1



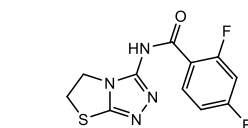
NVP-BNV428-NX-1



NVP-BNV433-NX-1



NVP-BNV442-NX-1



NVP-BNX029-NX-1

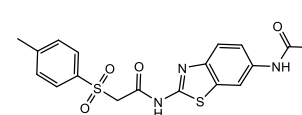


Figure S21 [related to Figure 44]. (A) Establishing siRNA-induced PARP14 knock-down conditions. Wild-type HeLa cells were transfected with siPARP14 #1 and siPARP14 #2 using Lipofectamine® 3000 and with siPARP14 #3 using Lipofectamine® RNAiMax. The latter transfection reagent was more effective and was used for following knock-down experiments. In all cases, PARP14 knock-down appeared 24 h post transfection. For following experiments, knock-down cells were analysed and used in assays 72 h post transfection. **(B)** PARP14 was knocked-down in U-2 OS cells with siPARP14 #3. The phenotype with cellular extended protrusion resulting of PARP14 knock-down with siPARP14 #3, which was observed in HeLa cells (shown as reference from experiment (A) on the right), could not be seen in U-2 OS cells. The black rectangles in representative bright field images at the top indicate the section that was enlarged and shown below the corresponding image.

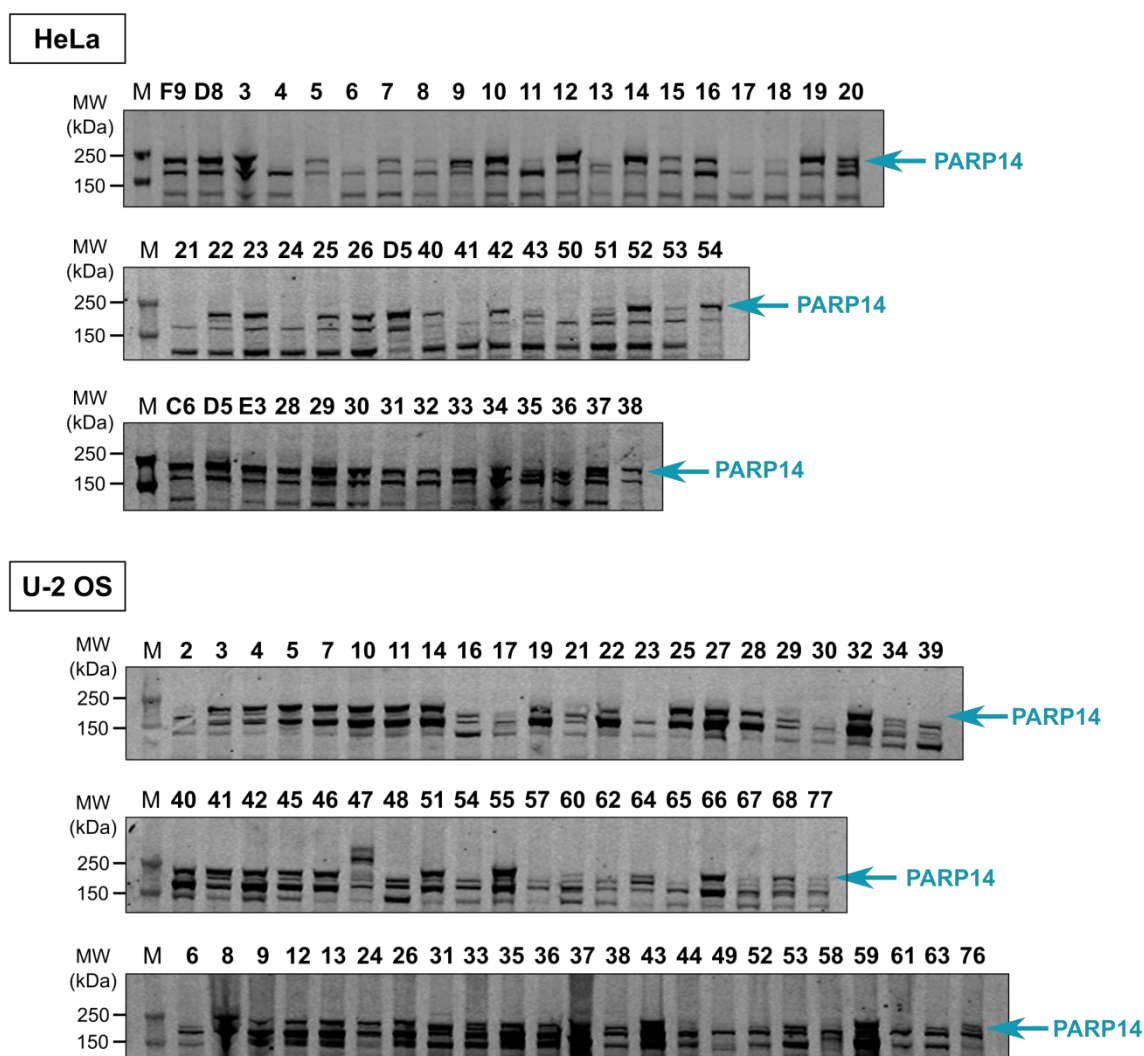


Figure S22 [related to Figure 46]. Screening of PARP14 edited HeLa and U-2 OS clones by Western blotting. 50 μ g of total cell lysate were loaded onto the gels and blots were overexposed to analyse for efficient PARP14 KO edits. Clones with alphabetical letter before a number are clonal wild-type cell lines. M = molecular weight marker.

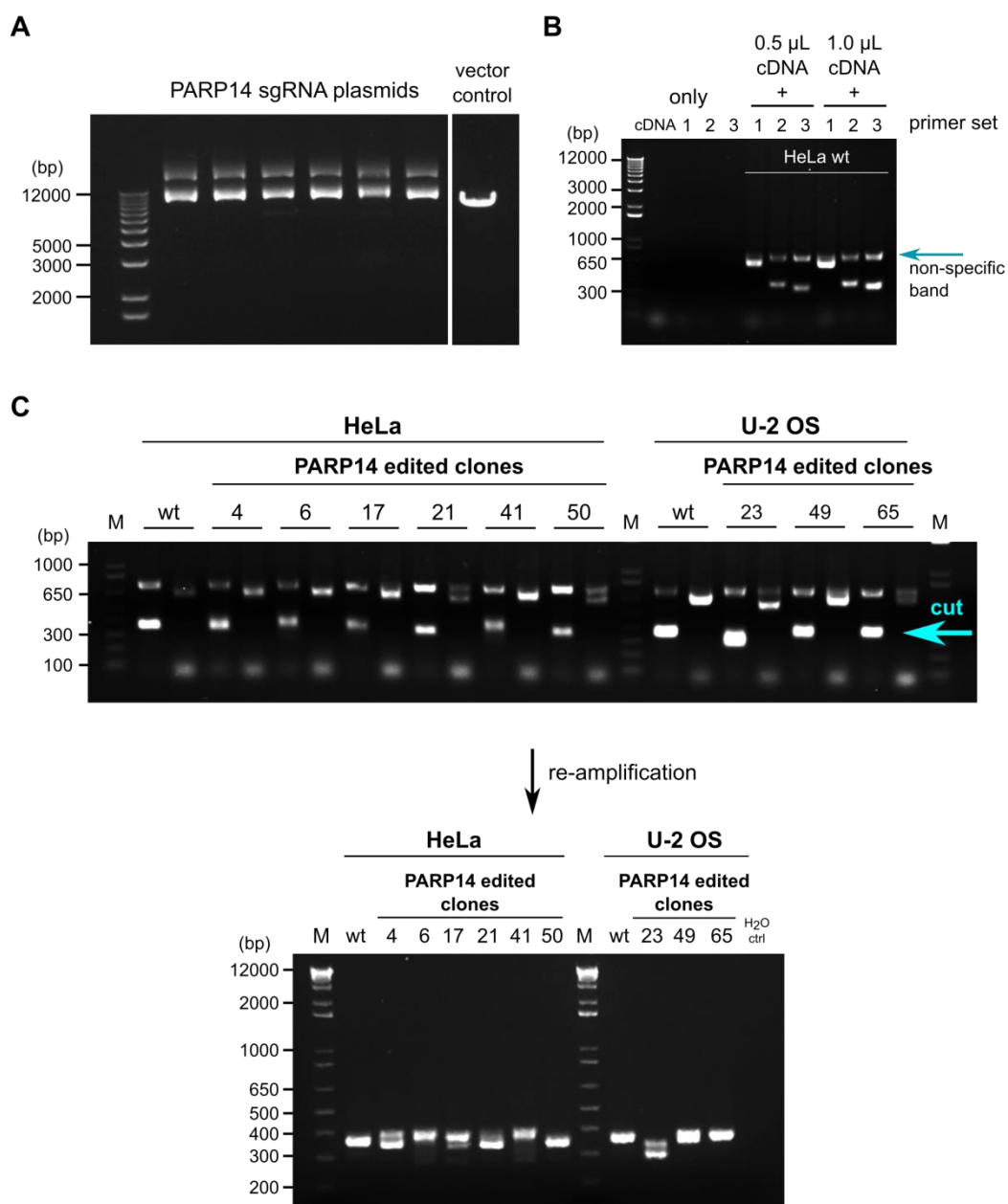


Figure S23 [related to Figure 47]. (A) sgRNA oligos were cloned into pSpCas9(BB)-2A-Puro or pSpCas9n(BB)-2A-Puro vectors. Prior to Sanger sequencing analysis, isolated plasmids were checked for sgRNA insertion by BbsI digestion followed by visualisation on a 1 % agarose gel. Since the BbsI restriction site is lost upon insertion of the sgRNA, vectors containing the inserts cannot be cleaved in comparison to the vector control. **(B)** For mRNA analysis, three different primer sets (1, 2 and 3) were tested for PCR amplification of PARP14 wild-type cDNA covering up to 600 bp (exons 1-4). For all primer sets, a non-specific band was obtained which could not be removed by alteration of the PCR conditions. **(C)** PARP14 mRNA was present in all selected PARP14 edited HeLa and U-2 OS clones. The cDNA of all clones was PCR amplified with two primer sets (primer set 1 (right) and primer set 3 (left)). To analyse the mRNA edit by Sanger sequencing, the PCR product obtained with primer set 3 was cut and re-amplified.

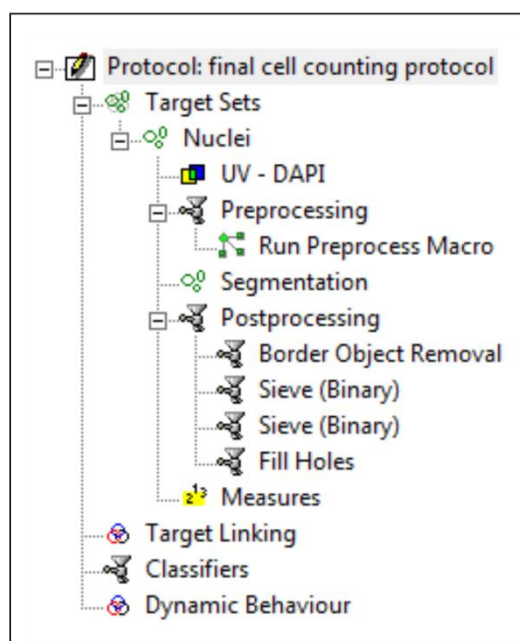
Table S4. Summary of PARP14 gDNA edits of generated HeLa and U-2 OS cell lines.

HeLa – PARP14 gDNA edits		
Clone ID	Indel location	Comment
clone 4	99_112 del 116_117 ins	allele1: 14 bp deletion, allele2: 13 bp insert
clone 6	95_104 del 116_117 ins	allele1: 10 bp deletion, allele2: 13 bp insert
clone 17	112_128 del 122_123 ins	allele1: 17 bp deletion, allele2: 8 bp insert
clone 21	102_123 del	allele1/2: 22 bp deletion
clone 41	95_104 del 126_127 ins	allele1: 10 bp deletion, allele2: 13 bp insert
clone 50	98_113 del	allele1/2: 16 bp deletion
U-2 OS – PARP14 gDNA edits		
Clone ID	Indel location	Comment
clone 23	87_132 del	allele1/2/3: 46 bp deletion
clone 30	101_102 ins 122_137 del	allele1: 240 bp insert, allele2: 15 bp deletion, allele3: ?
clone 49	125_133 del 129_135 del? 133_137 del?	main sequence: deletion of 9 bp (in frame), additional not analysable sequences
clone 65	127_130 del 130 del 131 del?	allele1: 4 bp deletion, allele2: 1 bp deletion, allele3: 1 bp deletion?

Table S5. Summary of PARP14 mRNA edits of generated HeLa and U-2 OS cell lines.

HeLa – PARP14 mRNA edits		
Clone ID	Indel location	Comment
clone 4	99_112 del 128_129 ins	one edit different to gDNA
clone 6	110_119 del 121_122 ins 126_127 ins	3 traces (1 × 10 bp deletion, 2 × 13 bp insertion), different edits to gDNA
clone 17	122_123 del ?	same main edit as gDNA (8 bp insert), additional not analysable sequences
clone 21	102_123 del	same edit as gDNA
clone 41	116_117 ins 126_127 ins 124_135 del	3 traces (2 × 13 bp insertion, 1 × 10 bp deletion), different edits to gDNA
clone 50	98_113 del	same edit as gDNA
U-2 OS – PARP14 mRNA edits		
Clone ID	Indel location	Comment
clone 23	87_132 del 95_162 del	1 × 46 bp deletion, 1 × 68 bp deletion
clone 49	125_133 del ?	same main edit as gDNA (9 bp deletion, in frame), additional not analysable sequences
clone 65	127_130 del 130 del 126_127 ins	3 traces (1 × 4 bp deletion, 1 × 1 bp deletion, 1 × 1 bp insertion), one edit different to gDNA

A



B

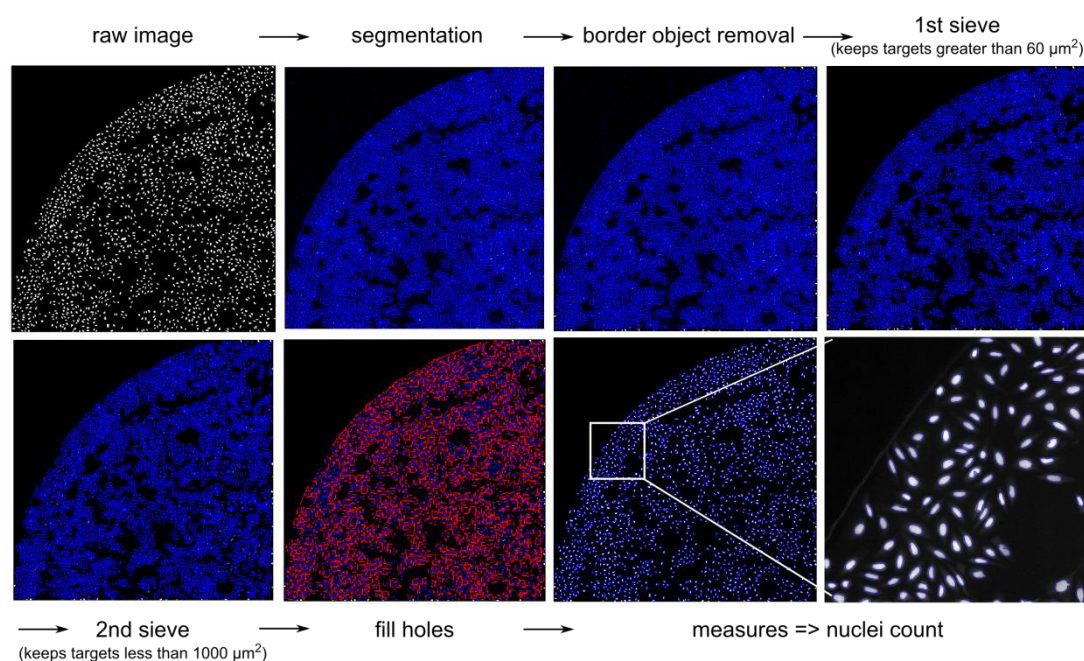


Figure S24. Final cell counting protocol for quantitation of number of Hoechst33342 stained nuclei. (A) The processing steps of the protocol established with the IN Cell Developer Toolbox software are shown. (B) Representative images resulting from the processing procedure are shown below. In a pre-processing step, the contrast of the image was enhanced using a Macro (comprising a detail extractor and smoothing filter). This was followed by the intensity-based segmentation of the enhanced image. In several post-processing steps, the mask was optimised and refined to fit the targets. Finally, nuclei were counted. **Note:** Image analysis algorithms were developed in collaboration with the Cell-based High-Throughput Screening group at the TDI.

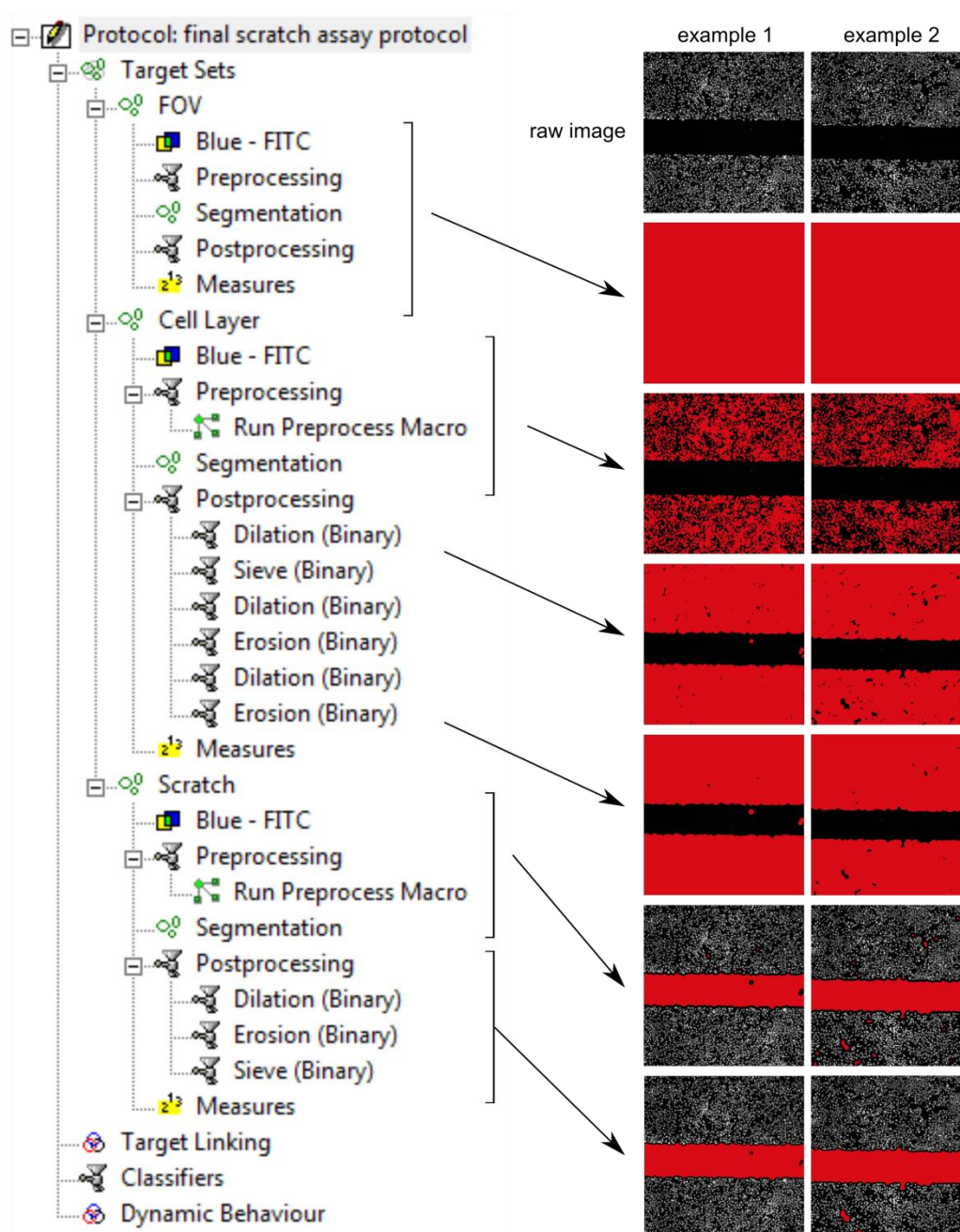


Figure S25. Final scratch assay protocol for analysis of scratch assays using the IN Cell Developer Toolbox software.

The steps of the protocol are shown on the left; images resulting from certain steps of the processing procedure are shown for two representative raw images (top) on the right (example 1: higher cell density, example 2: lower cell density). In a first step, the field of view (FOV) was determined to enable later the calculation of the scratch area in % of the whole field imaged. In a second step, the cell layer was segmented using the FITC channel. For this, the images were first pre-processed with a set-up Macro that enhanced the contrast of the image by stretching and then smoothening (median-based) the grey levels. This was followed by the fluorescence intensity-based segmentation. In a post-processing step, a sequence of processing steps was performed for optimising the mask to fit the target. In a third step, the scratch area was analysed. By first inverting the images and then converting them into binary images, the scratch area could be intensity-based segmented. A series of processing steps followed the segmentation to refine the mask to fit the target. Finally, the scratch area was measured and expressed as % of the FOV (in general for a high quality scratch ~20 %). As a QC step the number of 'holes' in the cell layer were counted, equalling 1 if scratch and cell layer could be perfectly segmented. **Note:** Image analysis algorithms were developed in collaboration with the Cell-based High-Throughput Screening group at the TDI.

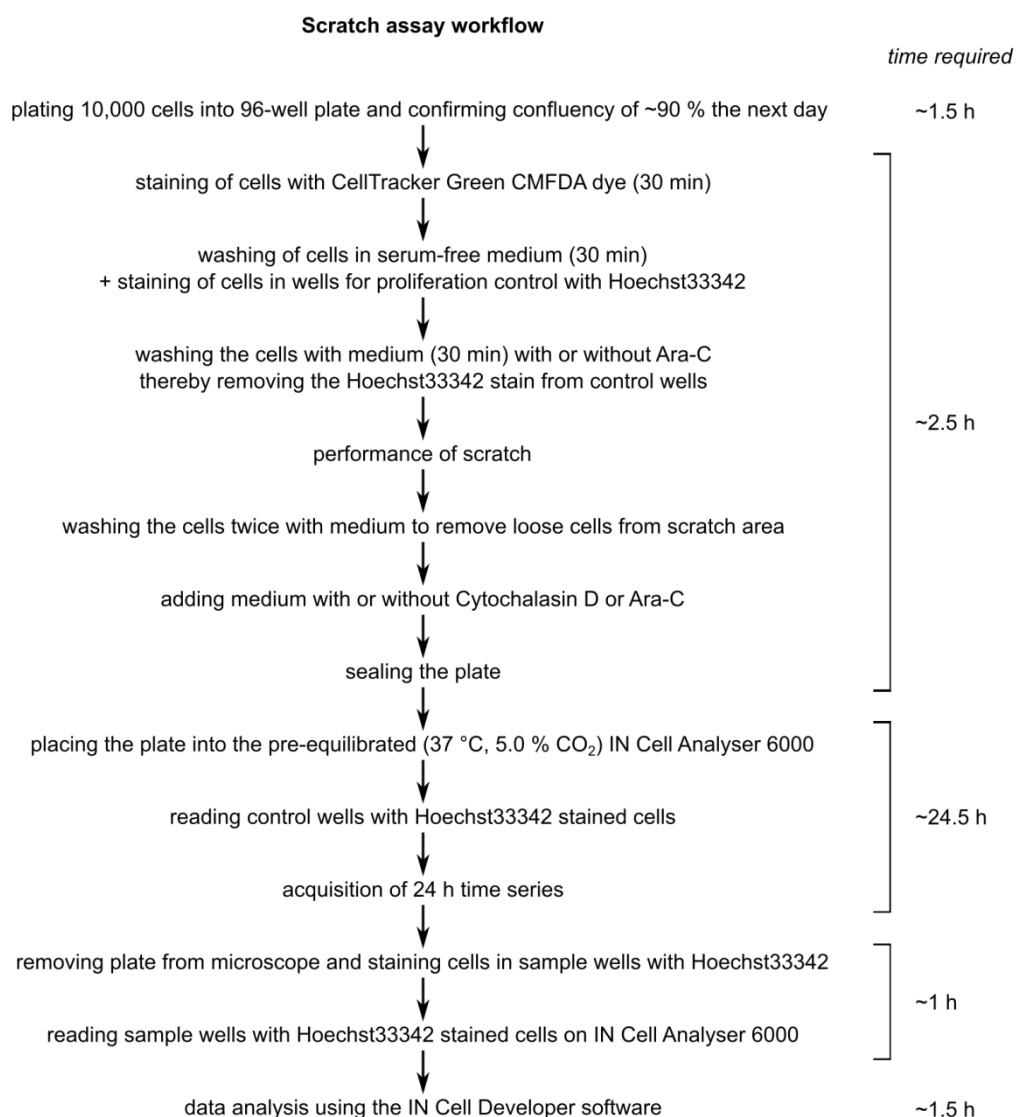
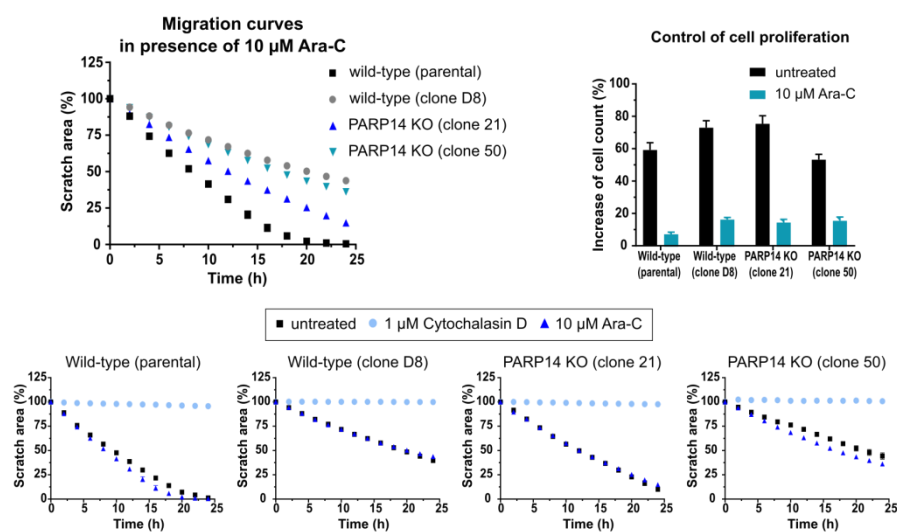
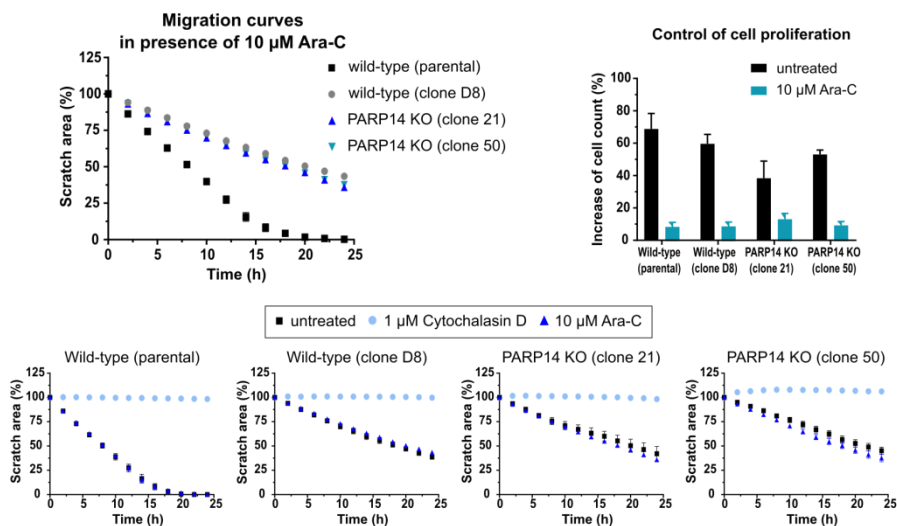


Figure S26. Schematic workflow of the established scratch migration assay. Please refer to section IV.11.3. for the detailed protocol.

A Repeat 2 of scratch assay including controls_higher passage number



B Repeat 3 of scratch assay including controls_higher passage number



C Repeat 4 of scratch assay including controls_low passage number

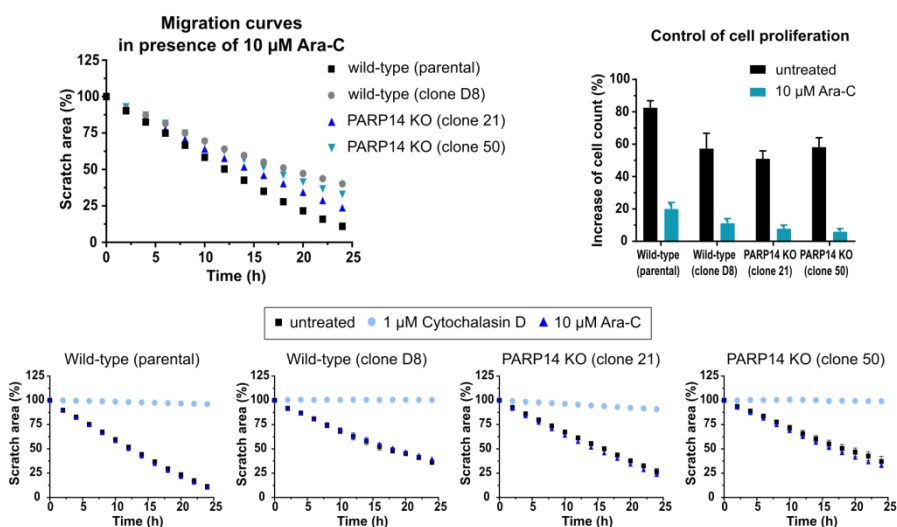


Figure S27 [related to Figure 52]. Repeats of scratch assays including Ara-C and Cytochalasin D controls. For legend, see top of next page.

Legend to Figure S27: Repeats of scratch assays including Ara-C and Cytochalasin D controls.

Wild-type (parental) cells with higher passage number migrated faster than at lower passage number. **(A-C)** Migration curves of HeLa cell lines in presence of 10 μ M Ara-C are shown on the left, confirmation of efficiency of halting proliferation with Ara-C during the assay (over 24 h) on the right. Individual migration curves also including the corresponding Cytochalasin D controls for the different HeLa cell lines is shown in the panel below. Data points are mean $\pm$ SEM, $n > 4$ for untreated cells, $n > 8$ for Ara-C treated cells, and $n > 5$ for Cytochalasin D treated cells.

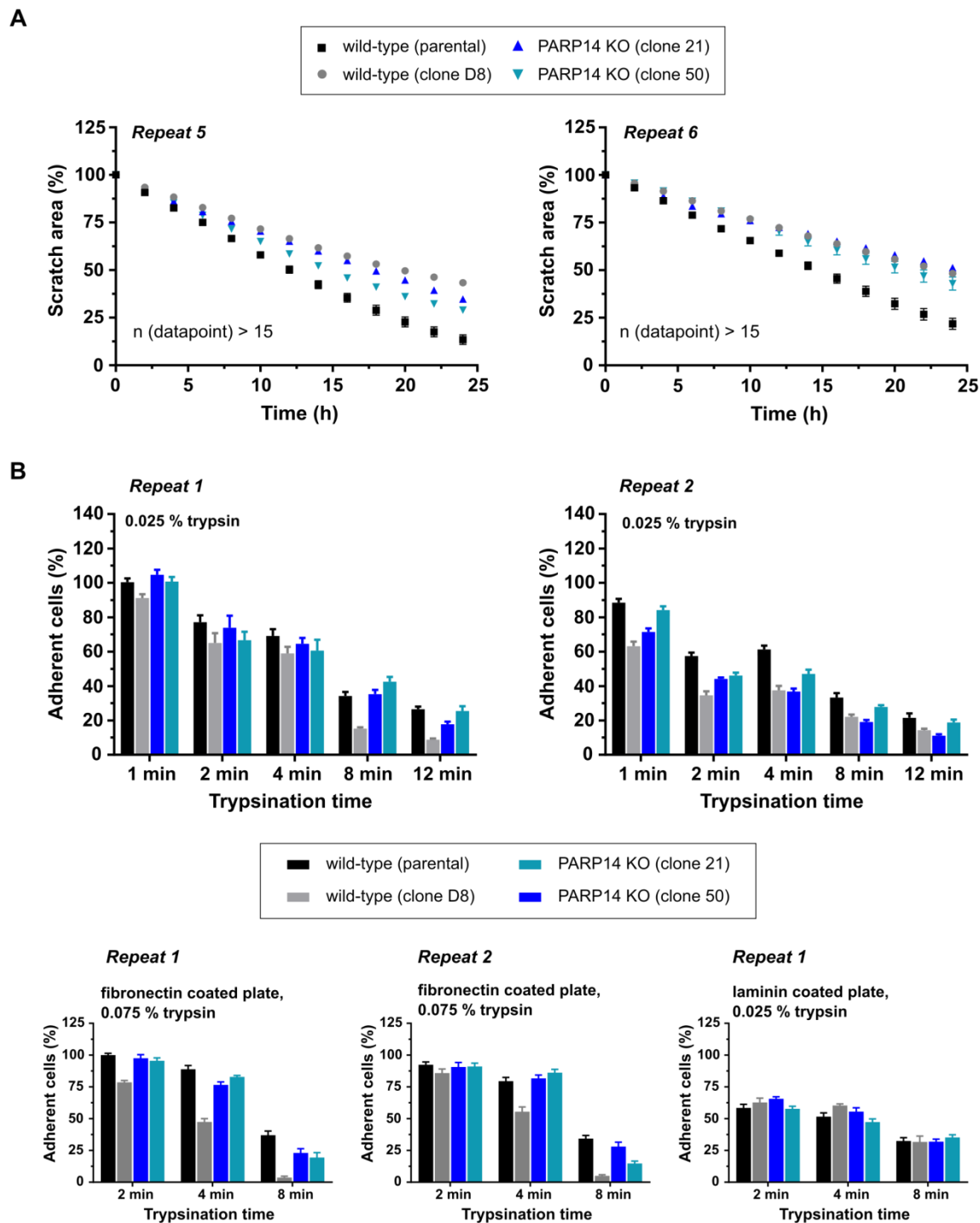


Figure S28 [related to Figures 49 and 52]. (A) Repeat of scratch assays without Ara-C and Cytochalasin D controls but with high number of replicates. Data points are mean $\pm$ SEM. **(B)** Repeats of cell adhesion assays using standard TC treated plates (top) and fibronectin or laminin coated plates (bottom). The data was normalised to the corresponding trypsin-untreated controls, where data points are mean $\pm$ SEM, $n = 6$.

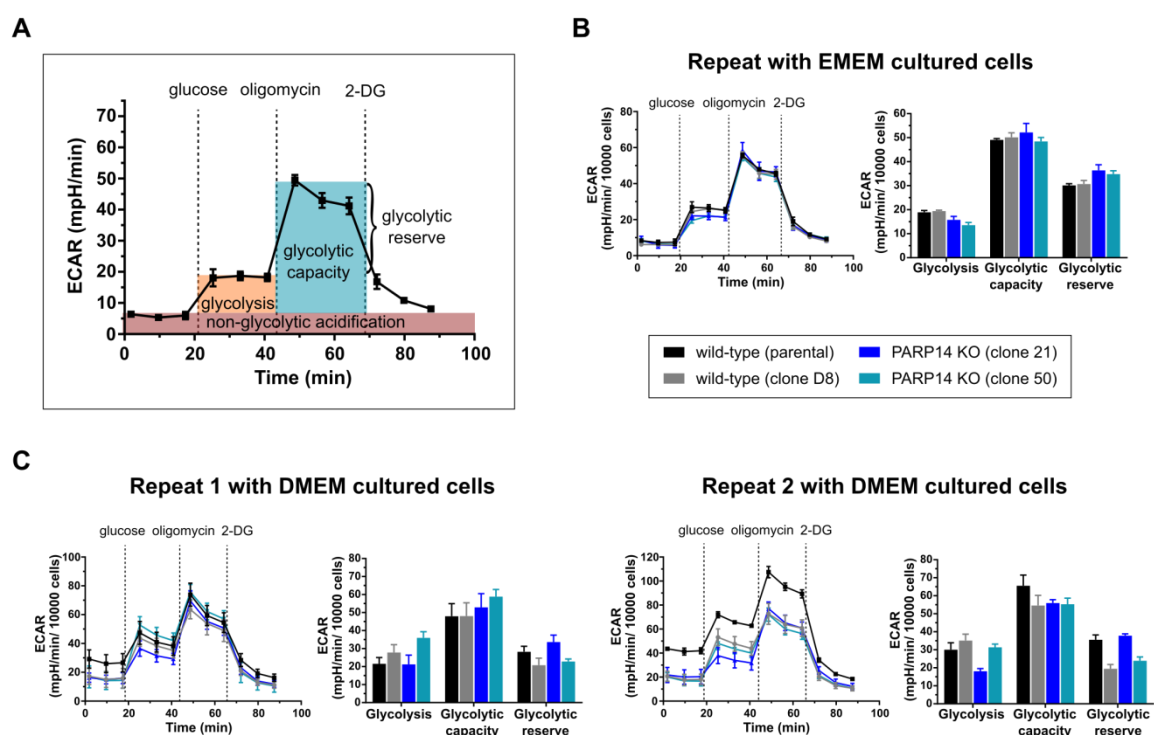


Figure S29 [related to Figures 53 and 54]. (A) A typical bioenergetic profile obtained in seahorse assays is shown. The glucose injection initiates glycolysis, resulting in production of lactate and an increase in the ECAR. Upon oligomycin injection, oxidative phosphorylation, thus mitochondrial ATP production and respiration, is blocked by inhibition of the ATP Synthase Complex V in the electron transport chain. Consequently, cells revert to glycolysis for ATP synthesis and therefore the ECAR increases (as the OCR decreases). Finally, the glucose analogue 2-deoxyglucose (2-DG) inhibits glycolysis through competitive binding to glucose hexokinase, reducing the ECAR back to the initial rate prior glucose injection. **(B)** and **(C)** Repeats of ECAR measurements and calculations of metabolic parameters of HeLa PARP14 wild-type and KO cells.

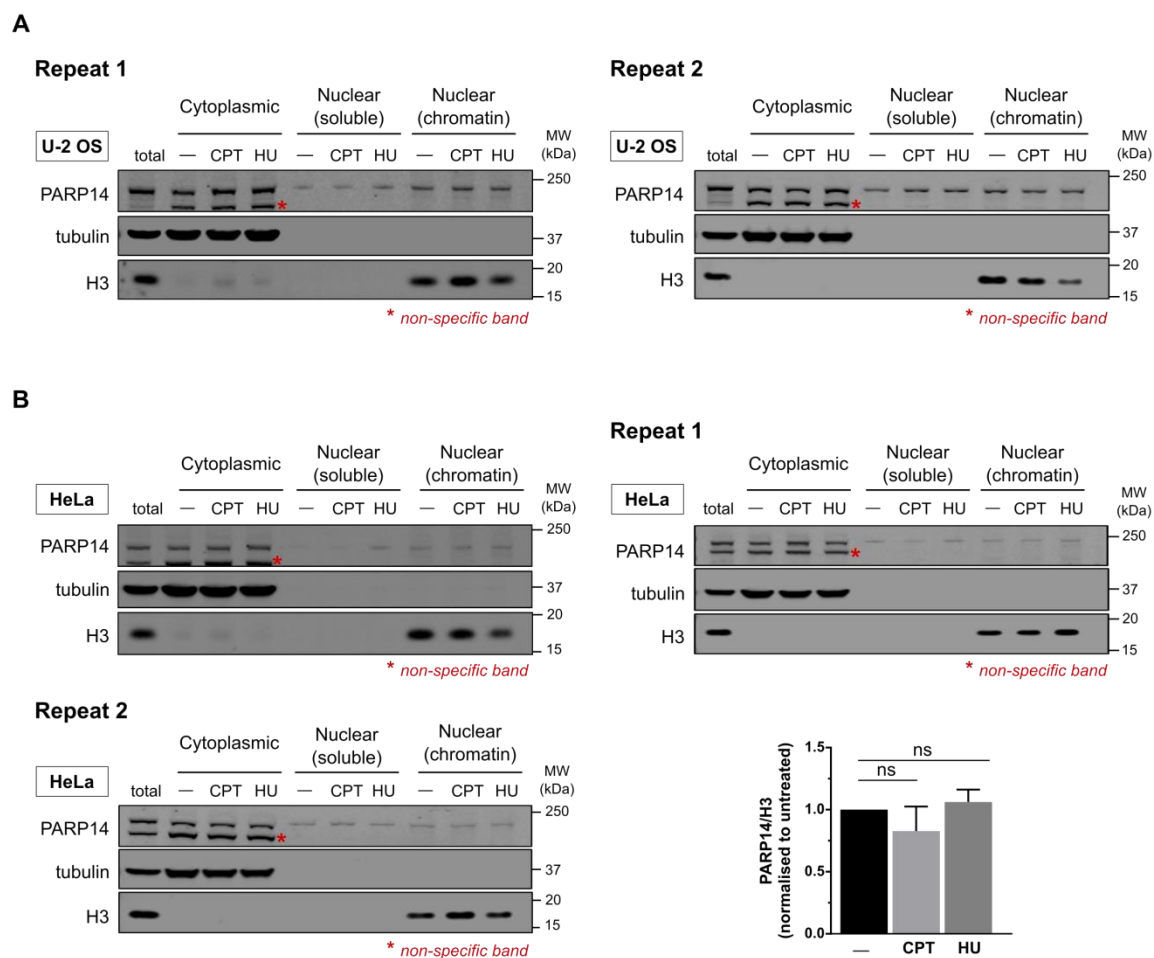


Figure S30 [related to Figure 55]. Cell fractionation assays performed with (A) U-2 OS and (B) HeLa cell lines. Cells were either left untreated (–) or treated with HU (2 mM for 24 h) and CPT (1 μ M, 2 h). α -tubulin and histone-H3 were used as loading controls for cytoplasmic and nuclear fractions, respectively. Data in the graph in (B) is the densitometry analysis of the three independent experiments, error bars represent SD, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (unpaired two-tailed t -test).

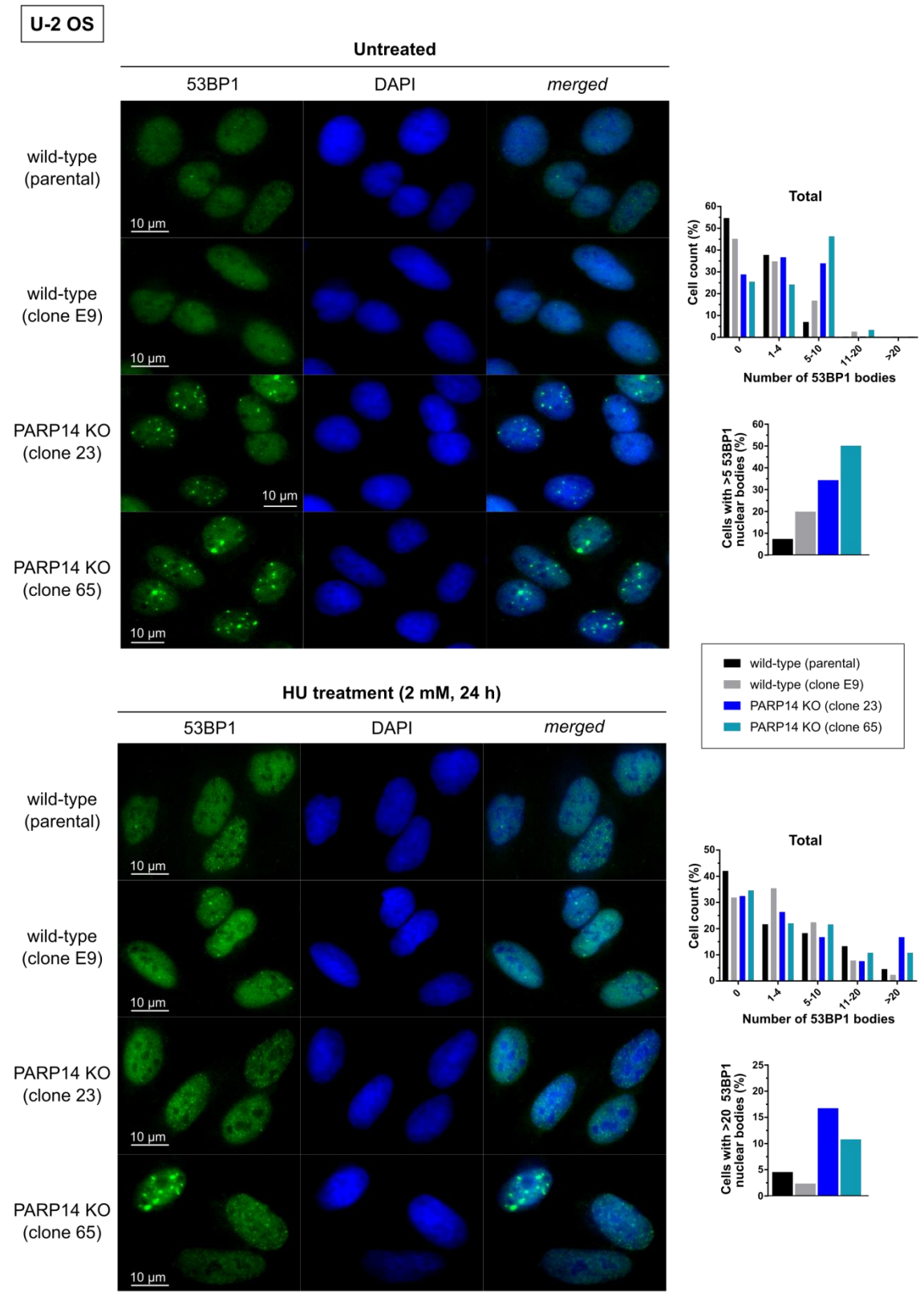


Figure S31 [related to Figure S6]. *PARP14 KO U-2 OS cell lines show stronger DNA damage response under normal replication conditions and after HU-induced DNA damage compared to the PARP14 wild-type cell lines.* Cells were left untreated or were treated with 2 mM HU for 24 h and then processed for immunostaining with anti-53BP1 (green) antibody. DNA was counterstained with DAPI. Quantitation of 53BP1 nuclear bodies (of 250-300 cells) is shown on the right of each panel of representative images for each condition.

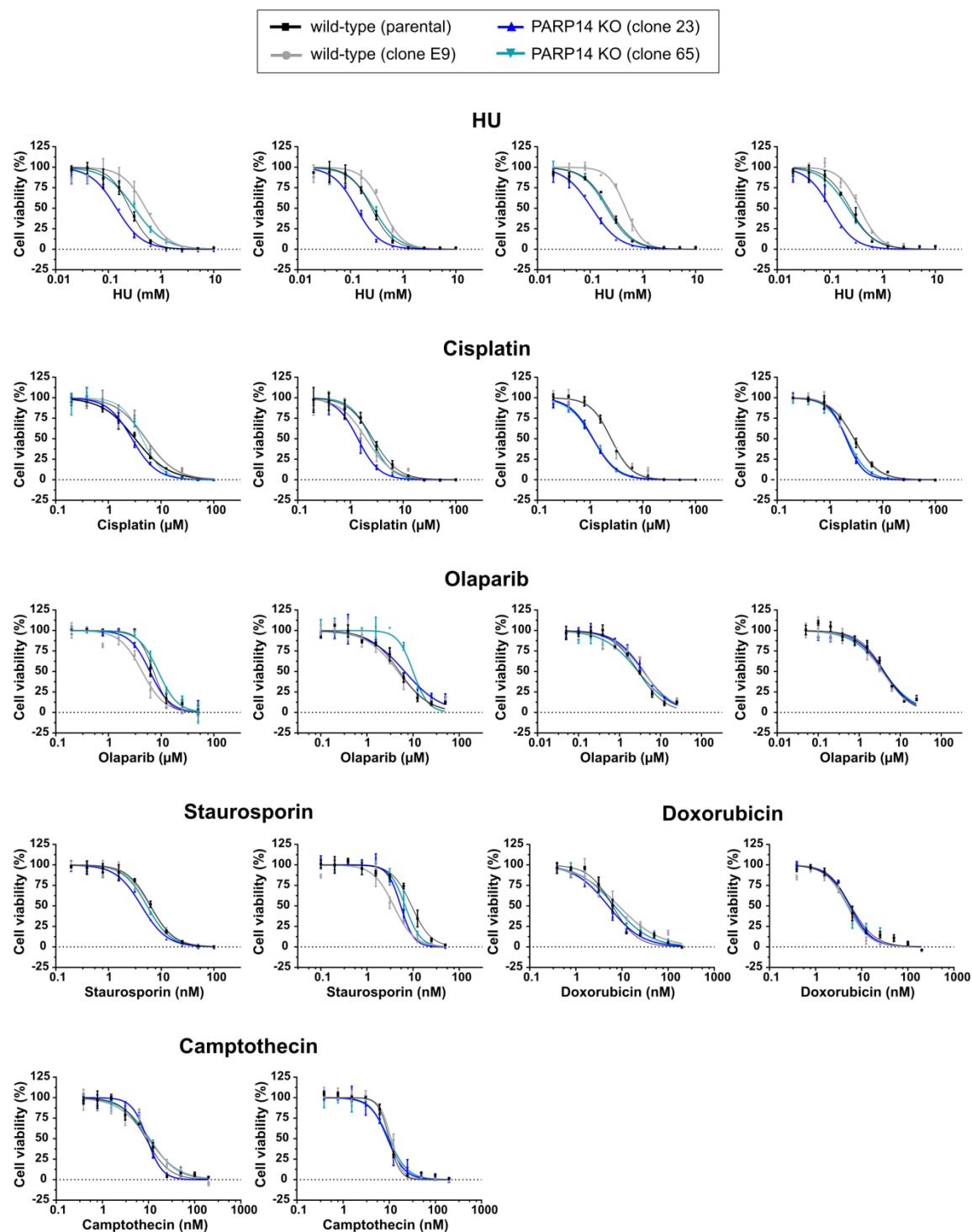
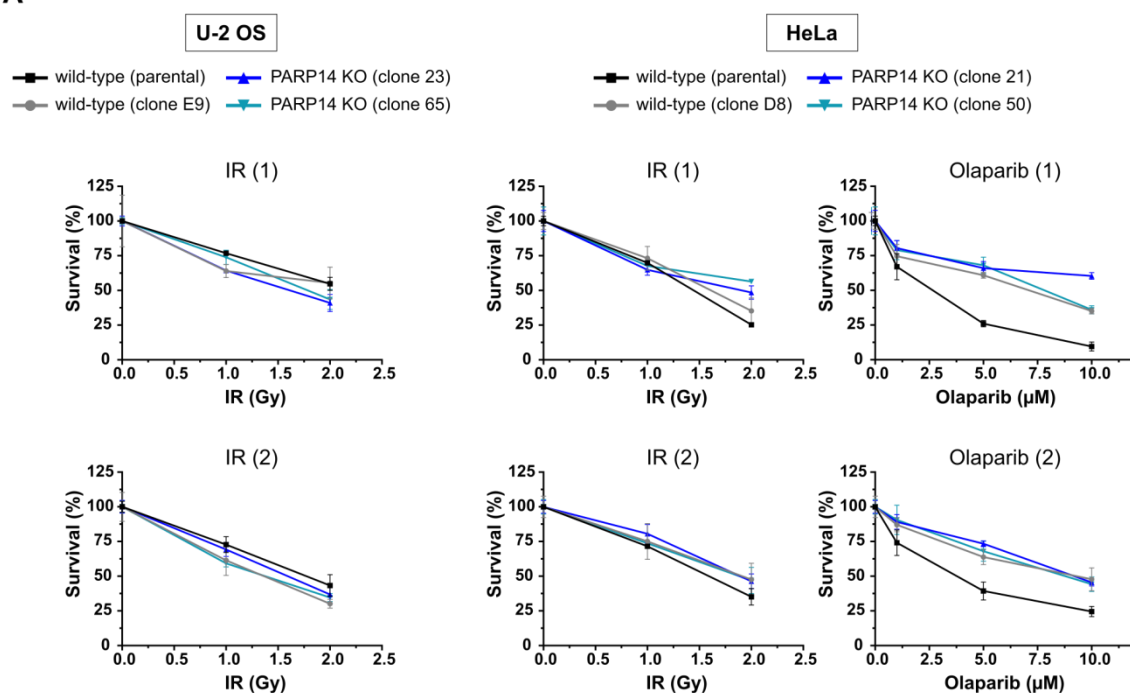


Figure S32 [related to Figure 57]. Repeats of resazurin viability assays for calculation of IC_{50} values. U-2 OS cells were exposed to the indicated compounds at different concentrations and cell viability was measured after 6 days of treatment using the AlamarBlue™ reagent. Data points are mean $\pm$ SD, n = 3.

A**B**

Treatment with 2 mM HU

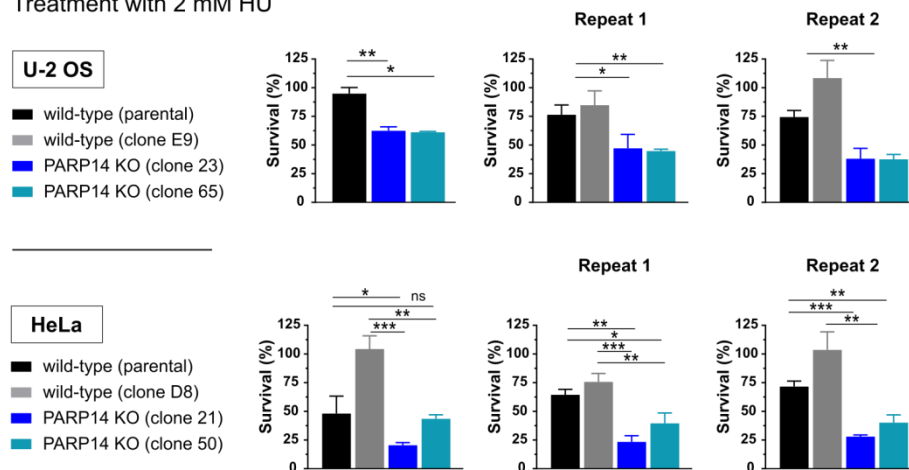
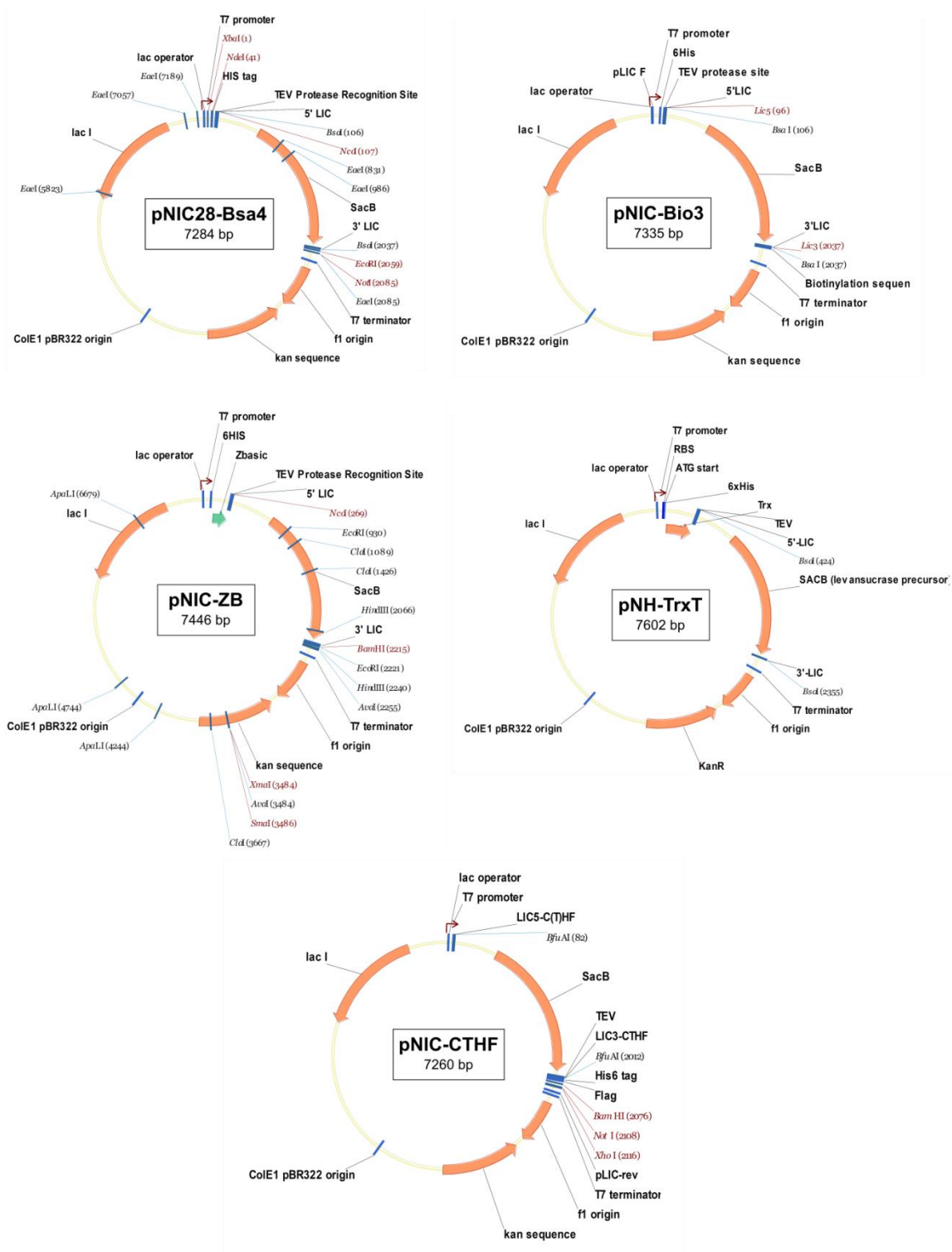
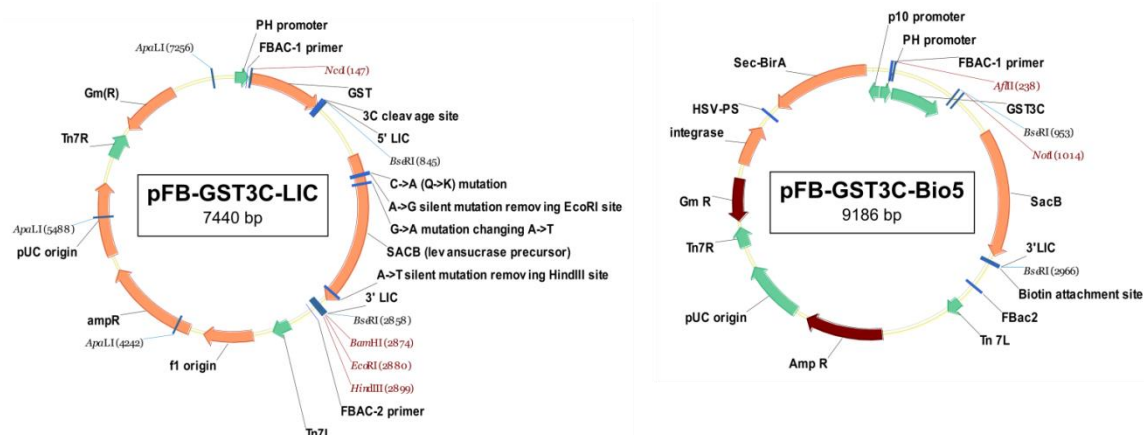


Figure S33 [related to Figure 58]. Summary of all clonogenic assays for (A) IR and olaparib treatment and (B) HU treatment. Cells were treated with olaparib and HU for 24 hours with the indicated concentrations. Results for U-2 OS cell lines are shown on the left, results for HeLa cell lines on the right. Error bars represent SEM, $n = 3$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$ (unpaired two-tailed t -test).

4. Supplementary experimental details

4.1. Vector maps





4.2. Supplementary experimental procedures

4.2.1. Purification of His₆-, GST- or biotin fusion proteins

Cell pellets re-suspended in lysis buffer (50 mM HEPES (pH 7.4), 500 mM NaCl, 20 mM imidazole, 5 % glycerol, 0.5 mM tris(2-carboxyethyl)phosphine [TCEP], 1:2,000 Calbiochem protease inhibitor cocktail set III) were quickly thawed on ice and lysed by high pressure homogenisation or sonication. Following cell lysis, DNA was precipitated with 0.15 % polyethyleneimine (PEI) and insoluble cell debris was removed by centrifugation (36000×g, 1 h, 4 °C). His₆- (-biotin) tagged proteins were then purified using Ni-Sepharose resin (GE Healthcare) and eluted stepwise in binding buffer with 40–250 mM imidazole. Typically, a high salt wash with 1 M NaCl was combined with the first elution step including 40 mM imidazole. As required, the His₆-tag or ZB-tag was removed from the macrodomain proteins at 4 °C overnight using recombinant TEV protease (used 1:40 -1:100 (w/w)) before gel filtration (Superdex 75 16/60, GE Healthcare) in GF buffer (25 mM HEPES (pH 7.4), 300 mM NaCl, 5 % glycerol, 0.5 mM TCEP). Proteins with removed His₆- tag were additionally passed over Ni-Sepharose resin as a final purification step. In the case of PARP14 MD3 purification for crystallisation, TEV His₆-tag cleavage after the first Ni-IMAC purification step was combined with overnight dialysis using SnakeSkin[®] Dialysis Tubing, 3500 MWCO (Thermo Scientific). The dialysed sample was passed over Ni-Sepharose resin before purification by gel filtration (Superdex 75 16/60) and further purification steps. For PARP14 MD3 the GF buffer contained 20 mM HEPES (pH 7.4), 300 mM NaCl, 10 % glycerol and 0.5 mM TCEP. All Ni-IMAC purification steps were performed by gravity-flow using glass columns (Sigma). For protein concentration steps,

Amicon® Ultra centrifugal filters (Merck) of appropriate MWCO size were used. Apart from Ni-IMAC rebinding, all other purification steps were carried out at 4 °C.

For purification of GST-tagged PARP15 MD1 expressed in insect cells by the SGC Biotechnology group, cell pellets were washed once with PBS and then re-suspended in lysis buffer (50 mM HEPES (pH 7.4), 500 mM NaCl, 5 % glycerol, 0.5 mM TCEP, 1:2,000 Calbiochem protease inhibitor cocktail set III). Cells were lysed by sonication at 35 % amplitude for 3 min effective time (pulsed 5 sec on, 10 sec off). Following cell lysis, DNA was precipitated with 0.15 % polyethyleneimine (PEI) and insoluble cell debris was removed by centrifugation (36000×g, 1 h, 4 °C). GST-tagged PARP15 MD1 proteins were then batch bound to Glutathione Sepharose 4B (GE Healthcare) by incubation for 1 hour at 4 °C on a rotating device. Beads were then collected in a gravity column and cleared with 50 mL washes with binding buffer (50 mM HEPES (pH 7.4), 500 mM NaCl, 5 % glycerol, 0.5 mM TCEP) until absorbance at 280 nm reached ~0. GSH bound proteins were then eluted twice with binding buffer containing 15 mM glutathione. The elution fractions were pooled and further purified by gel filtration (Superdex 75 10/300 GL) in GF buffer (20 mM HEPES (pH 7.4), 500 mM NaCl, 5 % glycerol, 0.5 mM TCEP). Peak fractions were pooled and concentrated to 3-5 mg/mL using a 30 kDa MWCO spin concentrator.

Nine PARP14 MD2_{SER-trunc} proteins were purified in parallel following the protocol developed by the SGC crystallography group. For this, cell pellets (~55 – 60 g bio mass) were quickly thawed, re-suspended in 250 mL binding buffer (50 mM HEPES (pH 7.4), 500 mM NaCl, 5 % glycerol, 0.5 mM TCEP) containing 0.5 mg/mL benzonase and 0.001 mg/mL lysozyme. Cell suspensions were aliquoted into 50 mL falcon tubes and then vortexed (5000 rpm, 5 min) and incubated for 30 min at RT. Afterwards, Triton X-100 was added to a final concentration of 0.8 % and tubes were stored at -80 °C overnight. The next day, cell suspensions were quickly thawed and imidazole was added to a final concentration of 20 mM. Cell debris was removed by centrifugation at 4000×g, 1 h, 4 °C and supernatants passed over His GraviTrap columns (GE Healthcare). Following washes with binding buffers containing 20 mM and 40 mM imidazole, respectively, proteins were eluted directly onto PD-10 desalting columns (GE Healthcare) with binding buffer containing 350 mM imidazole. Proteins were eluted from the PD-10 columns with binding buffer and TEV protease for O/N His₆-tag cleavage was added to the eluates (in a mass ratio of 1:40 for truncation constructs 1 and 2 and 1:20 for truncation constructs 3). The His₆-tagged removed proteins were again passed over His GraviTrap columns and purified by gel filtration (Yarra™ 5 µm SEC-2000 PrepLC column) as final purification step. All purification steps were carried out at 4 °C.

4.2.2. Confirmation of protein identities by tryptic digestion and LC-MS/MS

The identity of proteins in solution or in gel bands excised following SDS-PAGE was confirmed by tryptic digestion and tandem MS (LC-MS/MS). SDS-PAGE gel bands were excised as 1 x 4 mm slices using a gel cutting tip (GeneCatcher, Web Scientific) and stored in 10 % MeOH at 4 °C. The following sample processing was performed by the SGC Mass Spectrometry/Biotechnology group. Prior to digestion, the methanol solution was removed and replaced with 100 % acetonitrile for 2-5 minutes. The solution was then removed and replaced with 100 µL of 100 mM NH_4HCO_3 (pH 8.0). 1 µL of 1 M DTT was added and incubated at 56 °C for 40 minutes. 4 µL of 1 M iodoacetamide was then added and the reaction incubated at RT in the dark for 20 minutes. A further 1 µL of 1 M DTT, 200 µL of 100 mM NH_4HCO_3 and 1 µL of trypsin solution (sequencing grade, 1 mg/mL in 0.01 M HCl) was then added. Tryptic digestion proceeded at 37 °C for 16 hours and was terminated by addition of 3 µL of formic acid. LC-MS/MS was performed using a Dionex U3000 nano HPLC coupled to a Bruker Esquire HCT ion trap mass spectrometer. 1-5 µL of tryptic digest was loaded on to a 200 µm ID x 5 cm PS-DVB monolith column (Pepswift, Dionex Corp). A linear gradient of 0 % B to 15 % B was developed over 5 min, followed by a second linear gradient from 15 % B to 40 % B over 2 min. The column was washed at 90 % for 2 min and then equilibrated at 90 % B for a further 6 min. Solvent A was 2 % (v/v) acetonitrile, 0.1 % formic acid in water, solvent B was 80 % acetonitrile 0.1 % formic acid. The flow rate was 2.5 µL/min. The mass spectrometer was operated in positive ion, standard enhanced mode with a scan rate of 8100 m/z/sec and a scan range of 250 – 1800 m/z. The trap accumulation time was 200 msec and the accumulation target was 200,000 counts. Data-dependent peptide fragmentation was performed in Auto MS/MS mode. Compound extraction and peptide deconvolution was performed using the DA data analysis program (Bruker Daltonik). Database searching was performed using the Mascot 2.4 search algorithm (MatrixScience) with the following search parameters: Charge states +2, +3; MS tolerance 1.5 Da; MS/MS tolerance 1.3 Da; UniProt_SwissProt database without taxonomic restrictions.

4.2.3. HeLa and U-2 OS PARP14 KO cell line generation

For HeLa PARP14 KO cell line generation, HeLa wild-type cells were plated in a 24-well plate at 8000 cells/well in 500 µL of growth medium. When cells reached 90 % confluency, growth medium was replaced with 400 µL of fresh medium and transfection was carried out using Lipofectamine® 3000 (Invitrogen) according to the manufacturer's instructions. The amount of DNA per well used for transfection was kept constant at 500 ng/well (250 ng of plasmid for

co-transfection of two nickase constructs) as well as of Lipofectamine[®] 3000 of 1.5 μ L and P3000 of 1 μ L. 100 μ L of the prepared DNA-Lipofectamine complexes were added dropwise to the cells transfecting cells with either PARP14 editing pSpCas9 (wild-type/nickase) plasmids, or GFP targeting sSpCas9 (wild-type/nickase) plasmids or a GFP expressing transfection control plasmid. 48 hours post transfection, the medium was replaced, and 72 hours post transfection, cells were selected with medium containing 1 μ g/mL puromycin. After 48 hours of selection, the puromycin-containing medium was removed and replaced with growth medium containing 20 % FBS with following medium changes every 3 days. When cells reached 50 % confluency, clonal cell lines were isolated by single cell dilution in 96-well plates and kept in growth medium containing 20 % FBS. After 7 days, clonal cell lines were expanded by transferring into 24-well plates with growth medium containing 20 % FBS. When cell lines reached confluency over > 80 %, cells were further transferred into T25 flasks, expanded and then screened for PARP14 edits by Western blotting. Positive clones were further maintained reducing stepwise the FBS concentration to 10 % and then frozen down in 10 % DMSO.

For U-2 OS PARP14 KO cell line generation, U-2 OS wild-type cells were plated in a 12-well plate at 100,000 cells/well. The next day, confluency of 70 % was confirmed, growth medium was replaced with 900 μ L of fresh medium and transfection was carried out using Lipofectamine[®] 3000 (Invitrogen) according to the manufacturer's instructions. The amount of DNA per well used for transfection was kept constant at 1000 ng/well (500 ng of plasmid for co-transfection of two nickase constructs) as well as of Lipofectamine[®] 3000 of 3.0 μ L and P3000 of 2 μ L. 100 μ L of the prepared DNA-Lipofectamine complexes were added dropwise to the cells transfecting cells with either PARP14 editing pSpCas9 (wild-type/nickase) plasmids, or GFP targeting sSpCas9 (wild-type/nickase) plasmids or a GFP expressing transfection control plasmid. 48 hours post transfection, the medium was replaced, and 72 hours post transfection, cells were selected for 24 hours with medium containing 0.5 μ g/mL puromycin and for another 48 hours with 1.0 μ g/mL puromycin containing medium. After selection, cells were expanded in conditioned (50:50) growth medium containing 20 % FBS. When cells reached 50 % confluency, clonal cell lines were isolated by single cell dilution in 96-well plates and further kept in conditioned medium (50:50) and 20 % FBS. After 3 weeks, clonal cell lines were expanded by transferring first into 24-well plates in unconditioned medium but containing 20 % FBS and then into 6-well plates. When cell lines reached confluency over > 80 %, clones were screened for PARP14 edits by Western blotting. Positive clones were further maintained reducing stepwise the FBS concentration to 10 % and then frozen down in 10 % DMSO.

VI. REFERENCES

- [1] Jensen, O. N. (2006) Interpreting the protein language using proteomics. *Nat. Rev. Mol. Cell Biol.* 7, 391–403
- [2] Khoury, G. A., Baliban, R. C., and Floudas, C. A. (2011) Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Sci. Rep.* 1, 1–5
- [3] Prabakaran, S., Lippens, G., Steen, H., and Gunawardena, J. (2012) Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 4, 565–583
- [4] Hunter, T. (2007) The Age of Crosstalk: Phosphorylation, Ubiquitination, and Beyond. *Mol. Cell* 28, 730–738
- [5] Cosgrove, M. S., Boeke, J. D., and Wolberger, C. (2004) Regulated nucleosome mobility and the histone code. *Nat. Struct. Mol. Biol.* 11, 1037–1043
- [6] Karve, T. M., and Cheema, A. K. (2011) Small Changes Huge Impact: The Role of Protein Posttranslational Modifications in Cellular Homeostasis and Disease. *J. Amino Acids* 2011, 1–13
- [7] Torres, I. O., and Fujimori, D. G. (2015) Functional coupling between writers, erasers and readers of histone and DNA methylation. *Curr. Opin. Struct. Biol.* 35, 68–75
- [8] Kraus, W. L. (2015) PARPs and ADP-Ribosylation: 50 Years . . . and Counting. *Mol. Cell* 58, 902–910
- [9] Chambon, P., Weill, J. D., and Mandel, P. (1963) Nicotinamide mononucleotide activation of a new DNA-dependent polyadenylic acid synthesizing nuclear enzyme. *Biochem. Biophys. Res. Commun.* 11, 39–43
- [10] Perina, D., Mikoč, A., Ahel, J., Četković, H., Žaja, R., and Ahel, I. (2014) Distribution of protein poly(ADP-ribosyl)ation systems across all domains of life. *DNA Repair* 23, 4–16
- [11] Hottiger, M. O., Hassa, P. O., Lüscher, B., Schüler, H., and Koch-Nolte, F. (2010) Toward a unified nomenclature for mammalian ADP-ribosyltransferases. *Trends Biochem. Sci.* 35, 208–219
- [12] Schreiber, V., Dantzer, F., Ame, J.-C., and de Murcia, G. (2006) Poly(ADP-ribose): novel functions for an old molecule. *Nat. Rev. Mol. Cell Biol.* 7, 517–528
- [13] Amé, J.-C., Spenlehauer, C., and de Murcia, G. (2004) The PARP superfamily. *BioEssays* 26, 882–893
- [14] Otto, H., Reche, P. A., Bazan, F., Dittmar, K., Haag, F., and Koch-Nolte, F. (2005) *In silico* characterization of the family of PARP-like poly(ADP-ribosyl)transferases (pARTs). *BMC Genomics* 6, 139–161
- [15] Gibson, B. A., and Kraus, W. L. (2012) New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. *Nat. Rev. Mol. Cell Biol.* 13, 411–424

- [16] Vyas, S., Matic, I., Uchima, L., Rood, J., Zaja, R., Hay, R. T., Ahel, I., and Chang, P. (2014) Family-wide analysis of poly(ADP-ribose) polymerase activity. *Nat. Commun.* 5, 1–13
- [17] Barkauskaite, E., Jankevicius, G., and Ahel, I. (2015) Structures and Mechanisms of Enzymes Employed in the Synthesis and Degradation of PARP-Dependent Protein ADP-Ribosylation. *Mol. Cell* 58, 935–946
- [18] Rosenthal, F., Nanni, P., Barkow-Oesterreicher, S., and Hottiger, M. O. (2015) Optimization of LTQ-Orbitrap Mass Spectrometer Parameters for the Identification of ADP-Ribosylation Sites. *J. Proteome Res.* 14, 4072–4079
- [19] Chapman, J. D., Gagne, J.-P., Poirier, G. G., and Goodlett, D. R. (2013) Mapping PARP-1 Auto-ADP-ribosylation Sites by Liquid Chromatography–Tandem Mass Spectrometry. *J. Proteome Res.* 12, 1868–1880
- [20] Altmeyer, M., Messner, S., Hassa, P. O., Fey, M., and Hottiger, M. O. (2009) Molecular mechanism of poly(ADP-ribosylation) by PARP1 and identification of lysine residues as ADP-ribose acceptor sites. *Nucleic Acids Res.* 37, 3723–3738
- [21] Leidecker, O., Bonfiglio, J. J., Colby, T., Zhang, Q., Atanassov, I., Zaja, R., Palazzo, L., Stockum, A., Ahel, I., and Matic, I. (2016) Serine is a new target residue for endogenous ADP-ribosylation on histones. *Nat. Chem. Biol.* 12, 998–1000
- [22] Martello, R., Mangerich, A., Sass, S., Dedon, P. C., and Bürkle, A. (2013) Quantification of Cellular Poly(ADP-ribosylation) by Stable Isotope Dilution Mass Spectrometry Reveals Tissue-and Drug-Dependent Stress Response Dynamics. *ACS Chem. Biol.* 8, 1567–1575
- [23] Alvarez-gonzalez, R., and Jacobson, M. K. (1987) Characterization of Polymers of Adenosine Diphosphate Ribose Generated in Vitro and in Vivo. *Biochemistry* 26, 3218–3224
- [24] Alvarez-Gonzalez, R., and Althaus, F. R. (1989) Poly(ADP-ribose) catabolism in mammalian cells exposed to DNA-damaging agents. *Mutat. Res. - DNA Repair* 218, 67–74
- [25] Kiehlbauch, C. C., Aboul-Ela, N., Jacobson, E. L., Ringer, D. P., and Jacobson, M. K. (1993) High Resolution Fractionation and Characterization of ADP-ribose Polymers. *Anal. Biochem.* 208, 26–34
- [26] Slade, D., Dunstan, M. S., Barkauskaite, E., Weston, R., Lafite, P., Dixon, N., Ahel, M., Leys, D., and Ahel, I. (2011) The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase. *Nature* 477, 616–620
- [27] Barkauskaite, E., Brassington, A., Tan, E. S., Warwicker, J., Dunstan, M. S., Banos, B., Lafite, P., Ahel, M., Mitchison, T. J., Ahel, I., and Leys, D. (2013) Visualization of poly(ADP-ribose) bound to PARG reveals inherent balance between exo- and endo-glycohydrolase activities. *Nat. Commun.* 4, 1–8
- [28] Oka, S., Kato, J., and Moss, J. (2006) Identification and Characterization of a Mammalian 39-kDa Poly(ADP-ribose) Glycohydrolase. *J. Biol. Chem.* 281, 705–713
- [29] Steffen, J. D., and Pascal, J. M. (2013) New players to the field of ADP-ribosylation make the final cut. *EMBO J.* 32, 1205–1207
- [30] Rosenthal, F., Feijs, K. L. H., Frugier, E., Bonalli, M., Forst, A. H., Imhof, R., Winkler, H. C., Fischer, D., Caflisch, A., Hassa, P. O., Lüscher, B., and Hottiger, M. O. (2013) Macrod domain-containing proteins are new mono-ADP-ribosylhydrolases. *Nat. Struct. Mol. Biol.* 20, 502–507

- [31] Jankevicius, G., Hassler, M., Golia, B., Rybin, V., Zacharias, M., Timinszky, G., and Ladurner, A. G. (2013) A family of macrodomain proteins reverses cellular mono-ADP-ribosylation. *Nat. Struct. Mol. Biol.* 20, 508–514
- [32] Sharifi, R., Morra, R., Denise, C., Tallis, M., Chioza, B., Jankevicius, G., Simpson, M. A., Matic, I., Ozkan, E., Golia, B., Schellenberg, M. J., Weston, R., Williams, J. G., Rossi, M. N., Galehdari, H., Krahm, J., Wan, A., Trembath, R. C., Crosby, A. H., Ahel, D., Hay, R., Ladurner, A. G., Timinszky, G., Williams, R. S., and Ahel, I. (2013) Deficiency of terminal ADP-ribose protein glycohydrolase TARG1/C6orf130 in neurodegenerative disease. *EMBO J.* 32, 1225–1237
- [33] Barkauskaite, E., Jankevicius, G., Ladurner, A. G., Ahel, I., and Timinszky, G. (2013) The recognition and removal of cellular poly(ADP-ribose) signals. *FEBS J.* 280, 3491–3507
- [34] Hassa, P. O., Haenni, S. S., Elser, M., and Hottiger, M. O. (2006) Nuclear ADP-Ribosylation Reactions in Mammalian Cells: Where Are We Today and Where Are We Going? *Microbiol. Mol. Biol. Rev.* 70, 789–829
- [35] Feijs, K. L. H., Verheugd, P., and Lüscher, B. (2013) Expanding functions of intracellular resident mono-ADP-ribosylation in cell physiology. *FEBS J.* 280, 3519–3529
- [36] Tallis, M., Morra, R., Barkauskaite, E., and Ahel, I. (2014) Poly(ADP-ribosyl)ation in regulation of chromatin structure and the DNA damage response. *Chromosoma* 123, 79–90
- [37] Žaja, R., Mikoč, A., Barkauskaite, E., and Ahel, I. (2013) Molecular Insights into Poly(ADP-ribose) Recognition and Processing. *Biomolecules* 3, 1–17
- [38] Gagne, J., Isabelle, M., Lo, K. S., Bourassa, S., Hendzel, M. J., Dawson, V. L., Dawson, T. M., and Poirier, G. G. (2008) Proteome-wide identification of poly(ADP-ribose) binding proteins and poly(ADP-ribose)-associated protein complexes. *Nucleic Acids Res.* 36, 6959–6976
- [39] Pleschke, J. M., Kleczkowska, H. E., Strohm, M., and Althaus, F. R. (2000) Poly(ADP-ribose) Binds to Specific Domains in DNA Damage Checkpoint Proteins. *J. Biol. Chem.* 275, 40974–40980
- [40] Ahel, I., Ahel, D., Matsusaka, T., Clark, A. J., Pines, J., Boulton, S. J., and West, S. C. (2008) Poly(ADP-ribose)-binding zinc finger motifs in DNA repair/checkpoint proteins. *Nature* 451, 81–85
- [41] Oberoi, J., Richards, M. W., Crumpler, S., Brown, N., Blagg, J., and Bayliss, R. (2010) Structural Basis of Poly(ADP-ribose) Recognition by the Multizinc Binding Domain of Checkpoint with Forkhead-associated and RING Domains (CHFR). *J. Biol. Chem.* 285, 39348–39358
- [42] Aravind, L. (2001) The WWE domain: a common interaction module in protein ubiquitination and ADP ribosylation. *Trends Biochem. Sci.* 26, 273–275
- [43] Wang, Z., Michaud, G. A., Cheng, Z., Zhang, Y., Hinds, T. R., Fan, E., Cong, F., and Xu, W. (2012) Recognition of the iso-ADP-ribose moiety in poly(ADP-ribose) by WWE domains suggests a general mechanism for poly(ADP-ribosyl)ation-dependent ubiquitination. *Genes Dev.* 26, 235–240
- [44] He, F., Tsuda, K., Takahashi, M., Kuwasako, K., Terada, T., Shirouzu, M., Watanabe, S., Kigawa, T., Kobayashi, N., Güntert, P., Yokoyama, S., and Muto, Y. (2012) Structural insight into the interaction of ADP-ribose with the PARP WWE domains. *FEBS Lett.* 586, 3858–3864

- [45] DaRosa, P. A., Wang, Z., Jiang, X., Pruneda, J. N., Cong, F., Klevit, R. E., and Xu, W. (2014) Allosteric activation of the RNF146 ubiquitin ligase by a poly(ADP-ribosyl)ation signal. *Nature* 517, 223–226
- [46] Chul Kang, H., Lee, Y., Shin, J., Andrabi, S. A., Chi, Z., Gagné, J., Lee, Y., Seok Ko, H., Dae Lee, B., Poirier, G. G., Dawson, V. L., and Dawson, T. M. (2011) Iduna is a poly(ADP-ribose) (PAR)-dependent E3 ubiquitin ligase that regulates DNA damage. *PNAS* 108, 14103–14108
- [47] Callow, M. G., Tran, H., Phu, L., Lau, T., Lee, J., Sandoval, W. N., Liu, P. S., Bheddah, S., Tao, J., Lill, J. R., Hongo, J., Davis, D., Kirkpatrick, D. S., and Costa, M. (2011) Ubiquitin Ligase RNF146 Regulates Tankyrase and Axin to Promote Wnt Signaling. *PLoS One* 6, e22595
- [48] Li, X., Wu, Z., and Han, W. (2013) The macrodomain family: Rethinking an ancient domain from evolutionary perspectives. *Chinese Sci. Bull.* 58, 953–960
- [49] Han, W., Li, X., and Fu, X. (2011) The macro domain protein family: structure, functions, and their potential therapeutic implications. *Mutat. Res.* 727, 86–103
- [50] Pehrson, J. R., and Fried, V. A. (1992) MacroH2A, a Core Histone Containing a Large Nonhistone Region. *Science* 257, 1398–1400
- [51] Posavec, M., Timinszky, G., and Buschbeck, M. (2013) Macro domains as metabolite sensors on chromatin. *Cell. Mol. Life Sci.* 70, 1509–1524
- [52] Feijs, K. L. H., Forst, A. H., Verheugd, P., and Lüscher, B. (2013) Macrodomain-containing proteins: regulating new intracellular functions of mono(ADP-ribosyl)ation. *Nat. Rev. Mol. Cell Biol.* 13, 443–451
- [53] Liu, W., Xie, Y., Ma, J., Luo, X., Nie, P., Zuo, Z., Lahrmann, U., Zhao, Q., Zheng, Y., Zhao, Y., Xue, Y., and Ren, J. (2015) IBS: An illustrator for the presentation and visualization of biological sequences. *Bioinformatics* 31, 3359–3361
- [54] Karras, G. I., Kustatscher, G., Buhecha, H. R., Allen, M. D., Pugieux, C., Sait, F., Bycroft, M., and Ladurner, A. G. (2005) The macro domain is an ADP-ribose binding module. *EMBO J.* 24, 1911–1920
- [55] Allen, M. D., Buckle, A. M., Cordell, S. C., Löwe, J., and Bycroft, M. (2003) The Crystal Structure of AF1521 a Protein from *Archaeoglobus fulgidus* with Homology to the Non-histone Domain of MacroH2A. *J. Mol. Biol.* 330, 503–511
- [56] Egloff, M.-P., Malet, H., Putics, A., Heinonen, M., Dutartre, H., Frangeul, A., Gruez, A., Campanacci, V., Cambillau, C., Ziebuhr, J., Ahola, T., and Canard, B. (2006) Structural and Functional Basis for ADP-Ribose and Poly(ADP-Ribose) Binding by Viral Macro Domains. *J. Virol.* 80, 8493–8502
- [57] Feijs, K. L., Kleine, H., Braczynski, A., Forst, A. H., Herzog, N., Verheugd, P., Linzen, U., Kremmer, E., and Lüscher, B. (2013) ARTD10 substrate identification on protein microarrays: regulation of GSK3 β by mono-ADP-ribosylation. *Cell Commun. Signal.* 11, 5–15
- [58] Ahel, D., Hořejší, Z., Wiechens, N., Polo, S. E., Garcia-Wilson, E., Ahel, I., Flynn, H., Skehel, M., West, S. C., Jackson, S. P., Owen-Hughes, T., and Boulton, S. J. (2009) Poly(ADP-ribose)-Dependent Regulation of DNA Repair by the Chromatin Remodeling Enzyme ALC1. *Science* 325, 1240–1243

- [59] Chen, D., Vollmar, M., Rossi, M. N., Phillips, C., Kraehenbuehl, R., Slade, D., Mehrotra, P. V., Von Delft, F., Crosthwaite, S. K., Gileadi, O., Denu, J. M., and Ahel, I. (2011) Identification of Macrodomain Proteins as Novel O-Acetyl-ADP-Ribose Deacetylases. *J. Biol. Chem.* 286, 13261–13271
- [60] Neuvonen, M., and Ahola, T. (2009) Differential Activities of Cellular and Viral Macro Domain Proteins in Binding of ADP-Ribose Metabolites. *J. Mol. Biol.* 385, 212–225
- [61] Kraus, W. L. (2009) New functions for an ancient domain. *Nat. Struct. Mol. Biol.* 16, 904–907
- [62] Martzen, M. R., McCraith, S. M., Spinelli, S. L., Torres, F. M., Fields, S., Grayhack, E. J., and Phizicky, E. M. (1999) A Biochemical Genomics Approach for Identifying Genes by the Activity of Their Products. *Science* 286, 1153–1155
- [63] Tanner, K. G., Landry, J., Sternglanz, R., and Denu, J. M. (2000) Silent information regulator 2 family of NAD- dependent histone/protein deacetylases generates a unique product, 1-O-acetyl-ADP-ribose. *PNAS* 97, 14178–14182
- [64] Dani, N., Stilla, A., Marchegiani, A., Tamburro, A., Till, S., Ladurner, A. G., Corda, D., and Di Girolamo, M. (2009) Combining affinity purification by ADP-ribose-binding macro domains with mass spectrometry to define the mammalian ADP-ribosyl proteome. *PNAS* 106, 4243–4248
- [65] Peterson, F. C., Chen, D., Lytle, B. L., Rossi, M. N., Ahel, I., Denu, J. M., and Volkman, B. F. (2011) Orphan Macrodomain Protein (Human C6orf130) Is an O-Acyl-ADP-ribose Deacylase. *J. Biol. Chem.* 286, 35955–35965
- [66] Forst, A. H., Karlberg, T., Herzog, N., Thorsell, A. G., Gross, A., Feijs, K. L. H., Verheugd, P., Kursula, P., Nijmeijer, B., Kremmer, E., Kleine, H., Ladurner, A. G., Schüler, H., and Lüscher, B. (2013) Recognition of Mono-ADP-Ribosylated ARTD10 Substrates by ARTD8 Macrodomains. *Structure* 21, 462–475
- [67] Rack, J. G. M., Perina, D., and Ahel, I. (2016) Macrodomains: Structure, Function, Evolution, and Catalytic Activities. *Annu. Rev. Biochem.* 85, 431–454
- [68] Chakravarthy, S., Gundimella, S. K. Y., Caron, C., Perche, P., Pehrson, J. R., Khochbin, S., and Luger, K. (2005) Structural Characterization of the Histone Variant macroH2A. *Mol. Cell. Biol.* 25, 7616–7624
- [69] Rasmussen, T. P., Huang, T., Mastrangelo, M., Loring, J., Panning, B., and Jaenisch, R. (1999) Messenger RNAs encoding mouse histone macroH2A1 isoforms are expressed at similar levels in male and female cells and result from alternative splicing. *Nucleic Acids Res.* 27, 3685–3689
- [70] Kustatscher, G., Hothorn, M., Pugieux, C., Scheffzek, K., and Ladurner, A. G. (2005) Splicing regulates NAD metabolite binding to histone macroH2A. *Nat. Struct. Mol. Biol.* 12, 624–625
- [71] Buschbeck, M., Uribealago, I., Wibowo, I., Rué, P., Martin, D., Gutierrez, A., Morey, L., Guigó, R., López-Schier, H., and Di Croce, L. (2009) The histone variant macroH2A is an epigenetic regulator of key developmental genes. *Nat. Struct. Mol. Biol.* 16, 1074–1079
- [72] Nusinow, D. A., Herna, I., Fazzio, T., Shah, G., Kraus, W. L., and Panning, B. (2007) Poly(ADP-ribose) Polymerase 1 Is Inhibited by a Histone H2A Variant, MacroH2A, and Contributes to Silencing of the Inactive X Chromosome. *J. Biol. Chem.* 282, 12851–12859

- [73] Creppe, C., Janich, P., Cantariño, N., Noguera, M., Valero, V., Musulén, E., Douet, J., Posavec, M., Martin-Caballero, J., Sumoy, L., Di Croce, L., Benitah, S. A., and Buschbeck, M. (2012) MacroH2A1 Regulates the Balance between Self-Renewal and Differentiation Commitment in Embryonic and Adult Stem Cells. *Mol. Cell. Biol.* 32, 1442–1452
- [74] Timinszky, G., Till, S., Hassa, P. O., Hothorn, M., Kustatscher, G., Nijmeijer, B., Colombelli, J., Altmeyer, M., Stelzer, E. H. K., Scheffzek, K., Hottiger, M. O., and Ladurner, A. G. (2009) A macrodomain-containing histone rearranges chromatin upon sensing PARP1 activation. *Nat. Struct. Mol. Biol.* 16, 923–929
- [75] Changolkar, L. N., Costanzi, C., Leu, N. A., Chen, D., McLaughlin, K. J., and Pehrson, J. R. (2007) Developmental Changes in Histone macroH2A1-Mediated Gene regulation. *Mol. Cell. Biol.* 27, 2758–2764
- [76] Han, W.-D., Zhao, Y.-L., Meng, Y., Zang, L., Wu, Z., Li, Q., Si, Y., Huang, K., Ba, J., Morinaga, H., Nomura, M., and Mu, Y. (2007) Estrogenically regulated LRP16 interacts with estrogen receptor α and enhances the receptor's transcriptional activity. *Endocr. Relat. Cancer* 14, 741–753
- [77] Yang, J., Zhao, Y. L., Wu, Z. Q., Si, Y. L., Meng, Y. G., Fu, X. B., Mu, Y. M., and Han, W. D. (2009) The single-macro domain protein LRP16 is an essential cofactor of androgen receptor. *Endocr. Relat. Cancer* 16, 139–153
- [78] Wu, Z., Li, Y., Li, X., Ti, D., Zhao, Y., Si, Y., Mei, Q., Zhao, P., Fu, X., and Han, W. (2011) LRP16 Integrates into NF- κ B Transcriptional Complex and Is Required for Its Functional Activation. *PLoS One* 6, e18157
- [79] Ryan, D. P., and Owen-Hughes, T. (2011) Snf2-family proteins: Chromatin remodellers for any occasion. *Curr. Opin. Chem. Biol.* 15, 649–656
- [80] Gottschalk, A. J., Timinszky, G., Kong, S. E., Jin, J., Cai, Y., Swanson, S. K., Washburn, M. P., Florens, L., Ladurner, A. G., Conaway, J. W., and Conaway, R. C. (2009) Poly(ADP-ribosyl)ation directs recruitment and activation of an ATP-dependent chromatin remodeler. *PNAS* 106, 13770–13774
- [81] Mortusewicz, O., Fouquerel, E., Amé, J. C., Leonhardt, H., and Schreiber, V. (2011) PARG is recruited to DNA damage sites through poly(ADP-ribose)- and PCNA-dependent mechanisms. *Nucleic Acids Res.* 39, 5045–5056
- [82] Feng, X., and Koh, D. W. (2013) Roles of poly(ADP-Ribose) glycohydrolase in DNA damage and apoptosis. *Int. Rev. Cell Mol. Biol.* 304, 227–281
- [83] Aguiar, R. C. T., Yakushijin, Y., Kharbanda, S., Salgia, R., Fletcher, J. A., and Shipp, M. A. (2000) BAL is a novel risk-related gene in diffuse large B-cell lymphomas that enhances cellular migration. *Blood* 96, 4328–4334
- [84] Aguiar, R. C. T., Takeyama, K., He, C., Kreinbrink, K., and Shipp, M. A. (2005) B-aggressive lymphoma family proteins have unique domains that modulate transcription and exhibit poly(ADP-ribose) polymerase activity. *J. Biol. Chem.* 280, 33756–33765
- [85] Daugherty, M. D., Young, J. M., Kerns, J. A., and Malik, H. S. (2014) Rapid Evolution of PARP Genes Suggests a Broad Role for ADP-Ribosylation in Host-Virus Conflicts. *PLOS Genet.* 10, e1004403
- [86] Daugherty, M. D., and Malik, H. S. (2012) Rules of Engagement: Molecular Insights from Host-Virus Arms Races. *Annu. Rev. Genet.* 46, 677–700

- [87] Yan, Q., Xu, R., Zhu, L., Cheng, X., Wang, Z., Manis, J., and Shipp, M. A. (2013) BAL1 and Its Partner E3 Ligase, BBAP, Link Poly(ADP-Ribose) Activation, Ubiquitylation, and Double-Strand DNA Repair Independent of ATM, MDC1, and RNF8. *Mol. Cell. Biol.* 33, 845–857
- [88] Goenka, S., and Boothby, M. (2006) Selective potentiation of Stat-dependent gene expression by collaborator of Stat6 (CoaSt6), a transcriptional cofactor. *PNAS* 103, 4210–4215
- [89] Goenka, S., Sung, H. C., and Boothby, M. (2007) Collaborator of Stat6 (CoaSt6)-associated Poly(ADP-ribose) Polymerase Activity Modulates Stat6-dependent Gene Transcription. *J. Biol. Chem.* 282, 18732–18739
- [90] Mehrotra, P., Riley, J. P., Patel, R., Li, F., Voss, L., and Goenka, S. (2011) PARP-14 Functions as a Transcriptional Switch for Stat6-dependent Gene Activation. *J. Biol. Chem.* 286, 1767–1776
- [91] Yang, J., Aittoma, S., Pesu, M., Carter, K., Saarinen, J., Kalkkinen, N., Kieff, E., and Silvennoinen, O. (2002) Identification of p100 as a coactivator for STAT6 that bridges STAT6 with RNA polymerase II. *EMBO J.* 21, 4950–4958
- [92] Scarpa, E. S., Fabrizio, G., and Di Girolamo, M. (2013) A role of intracellular mono-ADP-ribosylation in cancer biology. *FEBS J.* 280, 3551–3562
- [93] Cho, S. H., Goenka, S., Henttinen, T., Gudapati, P., Reinikainen, A., Eischen, C. M., Lahesmaa, R., and Boothby, M. (2009) PARP-14, a member of the B aggressive lymphoma family, transduces survival signals in primary B cells. *Blood* 113, 2416–2425
- [94] Opferman, J. T., Letai, A., Beard, C., Sorcinelli, M. D., Ong, C. C., and Korsmeyer, S. J. (2003) Development and maintenance of B and T lymphocytes requires antiapoptotic MCL-1. *Nature* 426, 671–676
- [95] Van Lohuizen, M., Verbeek, S., Krimpenfort, P., Domen, J., Saris, C., Radaszkiewicz, T., and Berns, A. (1989) Predisposition to Lymphomagenesis in *pim-1* Transgenic Mice: Cooperation with *c-myc* and *N-myc* in Murine Leukemia Virus-Induced Tumors. *Cell* 56, 673–682
- [96] Cho, S. H., Ahn, A. K., Bhargava, P., Lee, C.-H., Eischen, C. M., McGuinness, O., and Boothby, M. (2011) Glycolytic rate and lymphomagenesis depend on PARP14, an ADP ribosyltransferase of the B aggressive lymphoma (BAL) family. *PNAS* 108, 15972–15977
- [97] Yanagawa, T., Funasaka, T., Tsutsumi, S., Hu, H., Watanabe, H., and Raz, A. (2007) Regulation of Phosphoglucose Isomerase/Autocrine Motility Factor Activities by the Poly(ADP-ribose) Polymerase Family-14. *Cancer Res.* 67, 8682–8689
- [98] Iansante, V., Choy, P. M., Fung, S. W., Liu, Y., Chai, J.-G., Dyson, J., Del Rio, A., D'Santos, C., Williams, R., Chokshi, S., Anders, R. A., Bubici, C., and Papa, S. (2015) PARP14 promotes the Warburg effect in hepatocellular carcinoma by inhibiting JNK1-dependent PKM2 phosphorylation and activation. *Nat. Commun.* 6, 1–15
- [99] Barbarulo, A., Iansante, V., Chaidos, A., Naresh, K., Rahemtulla, A., Franzoso, G., Karadimitris, A., Haskard, D. O., Papa, S., and Bubici, C. (2013) Poly(ADP-ribose) polymerase family member 14 (PARP14) is a novel effector of the JNK2-dependent pro-survival signal in multiple myeloma. *Oncogene* 32, 4231–4242

- [100] Leung, A. K. L., Vyas, S., Rood, J. E., Bhutkar, A., Sharp, P. A., and Chang, P. (2011) Poly(ADP-Ribose) Regulates Stress Responses and MicroRNA Activity in the Cytoplasm. *Mol. Cell* 42, 489–499
- [101] Iqbal, M. B., Johns, M., Cao, J., Liu, Y., Yu, S., Hyde, G. D., Laffan, M. A., Marchese, F. P., Cho, S. H., Clark, A. R., Gavins, F. N., Woollard, K. J., Perry, J., Mackman, N., Dean, J. L., Boothby, M., and Haskard, D. O. (2014) PARP-14 combines with tristetraprolin in the selective posttranscriptional control of macrophage tissue factor expression. *Blood* 124, 3646–3656
- [102] Cho, S. H., Raybuck, A., Wei, M., Erickson, J., Nam, K. T., Cox, R. G., Trochtenberg, A., Thomas, J. W., Williams, J., and Boothby, M. (2013) B cell-Intrinsic and -Extrinsic Regulation of Antibody Responses by PARP14, an Intracellular (ADP-Ribosyl)Transferase. *J. Immunol.* 191, 3169–3178
- [103] Riley, J. P., Kulkarni, A., Mehrotra, P., Koh, B., Perumal, N. B., Kaplan, M. H., and Goenka, S. (2013) PARP-14 Binds Specific DNA Sequences to Promote Th2 Cell Gene Expression. *PLoS One* 8, e83127
- [104] Mehrotra, P., Krishnamurthy, P., Sun, J., Goenka, S., and Kaplan, M. H. (2015) Poly-ADP-ribosyl polymerase-14 promotes T helper 17 and follicular T helper development. *Immunology* 146, 537–546
- [105] Mehrotra, P., Hollenbeck, A., Riley, J. P., Fang, L., Patel, R. J., Akhtar, N., and Goenka, S. (2013) PARP-14 and its enzyme activity regulates Th2 differentiation and allergic airway disease. *J. Allergy Clin. Immunol.* 131, 521–531
- [106] Krishnamurthy, P., Da-Silva-Arnold, S., Turner, M. J., Travers, J. B., and Kaplan, M. H. (2017) Poly-ADP ribose polymerase-14 limits severity of allergic skin disease. *Immunology* 152, 451–461
- [107] Iwata, H., Goettsch, C., Sharma, A., Ricchiuto, P., Goh, W. W. Bin, Halu, A., Yamada, I., Yoshida, H., Hara, T., Wei, M., Inoue, N., Fukuda, D., Mojcher, A., Mattson, P. C., Barabási, A.-L., Boothby, M., Aikawa, E., Singh, S. A., and Aikawa, M. (2016) PARP9 and PARP14 cross-regulate macrophage activation via STAT1 ADP-ribosylation. *Nat. Commun.* 7, 1–19
- [108] Vyas, S., Chesarone-Cataldo, M., Todorova, T., Huang, Y.-H., and Chang, P. (2013) A systematic analysis of the PARP protein family identifies new functions critical for cell physiology. *Nat. Commun.* 4, 1–13
- [109] Nicolae, C. M., Aho, E. R., Choe, K. N., Constantin, D., Hu, H. J., Lee, D., Myung, K., and Moldovan, G. L. (2015) A novel role for the mono-ADP-ribosyltransferase PARP14/ARTD8 in promoting homologous recombination and protecting against replication stress. *Nucleic Acids Res.* 43, 3143–3153
- [110] Verheugd, P., Forst, A. H., Milke, L., Herzog, N., Feijs, K. L. H., Kremmer, E., Kleine, H., and Lüscher, B. (2013) Regulation of NF- κ B signalling by the mono-ADP-ribosyltransferase ARTD10. *Nat. Commun.* 4, 1–11
- [111] Meng, Y. G., Han, W. D., Zhao, Y. L., Huang, K., Si, Y. L., Wu, Z. Q., and Mu, Y. M. (2007) Induction of the *LRP16* gene by estrogen promotes the invasive growth of Ishikawa human endometrial cancer cells through the downregulation of E-cadherin. *Cell Res.* 17, 869–880

- [112] Xi, H.-Q., Zhao, P., and Han, W.-D. (2010) Clinicopathological significance and prognostic value LRP16 expression in colorectal carcinoma. *World J. Gastroenterol.* 16, 1644–1648
- [113] Li, Y.-Z., Zhao, P., and Han, W.-D. (2009) Clinicopathological significance of LRP16 protein in 336 gastric carcinoma patients. *World J. Gastroenterol.* 15, 4833–4837
- [114] Yao, D.-J., Qiao, S., Zhang, Y., Zhao, Y.-T., and Yuan, C.-H.- (2017) Correlation between expression of LRP16 , Ki67 and EGFR and breast cancer clinical pathologic factors and prognosis. *Eur. Rev. Med. Pharmacol. Sci.* 21, 47–51
- [115] Zhao, P., Lu, Y., and Han, W. (2010) Clinicopathological significance and prognostic value of leukemia-related protein 16 expression in invasive ductal breast carcinoma. *Cancer Sci.* 101, 2262–2268
- [116] Mohseni, M., Cidado, J., Croessmann, S., Cravero, K., Cimino-Mathews, A., Wong, H. Y., Scharpf, R., Zabransky, D. J., Abukhdeir, A. M., Garay, J. P., Wang, G. M., Beaver, J. A., Cochran, R. L., Blair, B. G., Rosen, D. M., Erlanger, B., Argani, P., Hurley, P. J., Lauring, J., and Park, B. H. (2014) *MACROD2* overexpression mediates estrogen independent growth and tamoxifen resistance in breast cancers. *PNAS* 111, 17606–17611
- [117] Cheng, W., Su, Y., and Xu, F. (2013) *CHD1L*: a novel oncogene. *Mol.Cancer* 12, 170–177
- [118] Ma, N. F., Hu, L., Fung, J. M., Xie, D., Zheng, B. J., Chen, L., Tang, D. J., Fu, L., Wu, Z., Chen, M., Fang, Y., and Guan, X. Y. (2008) Isolation and Characterization of a Novel Oncogene, Amplified in Liver Cancer 1, within a Commonly Amplified Region at 1q21 in Hepatocellular Carcinoma. *Hepatology* 47, 503–510
- [119] Chen, L., Hon, T., Chan, M., Yuan, Y., Hu, L., Huang, J., Ma, S., Wang, J., Dong, S., Tang, K. H., Xie, D., Li, Y., and Guan, X. (2010) *CHD1L* promotes hepatocellular carcinoma progression and metastasis in mice and is associated with these processes in human patients. *J. Clin. Invest.* 120, 1178–1191
- [120] Chen, L., Hu, L., Chan, T. H. M., Tsao, G. S. W., Xie, D., Huo, K. K., Fu, L., Ma, S., Zheng, B. J., and Guan, X. Y. (2009) Chromodomain helicase/adenosine triphosphatase DNA binding protein 1-like (*CHD1L*) gene suppresses the nucleus-to-mitochondria translocation of Nur77 to sustain hepatocellular carcinoma cell survival. *Hepatology* 50, 122–129
- [121] Ji, X., Li, J., Zhu, L., Cai, J., Zhang, J., Qu, Y., Zhang, H., Liu, B., Zhao, R., and Zhu, Z. (2013) *CHD1L* promotes tumor progression and predicts survival in colorectal carcinoma. *J. Surg. Res.* 185, 84–91
- [122] Juszczynski, P., Kutok, J. L., Li, C., Mitra, J., Aguiar, R. C. T., and Shipp, M. A. (2006) *BAL1* and *BBAP* Are Regulated by a Gamma Interferon-Responsive Bidirectional Promoter and Are Overexpressed in Diffuse Large B-Cell Lymphomas with a Prominent Inflammatory Infiltrate. *Mol. Cell. Biol.* 26, 5348–5359
- [123] Camicia, R., Bachmann, S. B., Winkler, H. C., Beer, M., Tinguely, M., Haralambieva, E., and Hassa, P. O. (2013) *BAL1/ARTD9* represses the anti-proliferative and pro-apoptotic IFN γ -STAT1-IRF1-p53 axis in diffuse large B-cell lymphoma. *J. Cell Sci.* 126, 1969–1980
- [124] Park, E., and Griffin, D. E. (2009) The nsP3 macro domain is important for Sindbis virus replication in neurons and neurovirulence in mice. *Virology* 388, 305–314

- [125] Saikatendu, K. S., Joseph, J. S., Subramanian, V., Clayton, T., Griffith, M., Moy, K., Velasquez, J., Neuman, B. W., Buchmeier, M. J., Stevens, R. C., and Kuhn, P. (2005) Structural Basis of Severe Acute Respiratory Syndrome Coronavirus ADP-Ribose-1''-Phosphate Dephosphorylation by a Conserved Domain of nsP3. *Structure* 13, 1665–1675
- [126] Fehr, A. R., Athmer, J., Channappanavar, R., Phillips, J. M., Meyerholz, D. K., and Perlman, S. (2015) The nsp3 Macrodomain Promotes Virulence in Mice with Coronavirus-Induced Encephalitis. *J. Virol.* 89, 1523–1536
- [127] Sonnenblick, A., de Azambuja, E., Azim, H. A., and Piccart, M. (2014) An update on PARP inhibitors—moving to the adjuvant setting. *Nat. Rev. Clin. Oncol.* 12, 27–41
- [128] Lord, C. J., and Ashworth, A. (2017) PARP inhibitors: Synthetic lethality in the clinic. *Science* 355, 1152–1158
- [129] Garber, K. (2013) PARP inhibitors bounce back. *Nat. Rev. Drug Discov.* 12, 725–727
- [130] Mullard, A. (2017) PARP inhibitors plough on. *Nat. Rev. Drug Discov.* 16, 229
- [131] James, D. I., Smith, K. M., Jordan, A. M., Fairweather, E. E., Griffiths, L. A., Hamilton, N. S., Hitchin, J. R., Hutton, C. P., Jones, S., Kelly, P., McGonagle, A. E., Small, H., Stowell, A. I. J., Tucker, J., Waddell, I. D., Waszkowycz, B., and Ogilvie, D. J. (2016) First-in-class Chemical Probes against Poly(ADP-ribose) Glycohydrolase (PARG) Inhibit DNA Repair with Differential Pharmacology to Olaparib. *ACS Chem. Biol.* 11, 3179–3190
- [132] Finch, K. E., Knezevic, C. E., Nottbohm, A. C., Partlow, K. C., and Hergenrother, P. J. (2012) Selective Small Molecule Inhibition of Poly(ADP-ribose) Glycohydrolase (PARG). *ACS Chem. Biol.* 7, 563–570
- [133] Steffen, J. D., Coyle, D. L., Damodaran, K., Beroza, P., and Jacobson, M. K. (2011) Discovery and Structure-Activity Relationships of Modified Salicylanilides as Cell Permeable Inhibitors of Poly(ADP-ribose) Glycohydrolase (PARG). *J. Med. Chem.* 54, 5403–5413
- [134] Fauzee, N. J. S., Pan, J., and Wang, Y. L. (2010) PARP and PARG Inhibitors-New Therapeutic Targets in Cancer Treatment. *Pathol. Oncol. Res.* 16, 469–478
- [135] Islam, R., Koizumi, F., Kodera, Y., Inoue, K., Okawara, T., and Masutani, M. (2014) Design and synthesis of phenolic hydrazide hydrazones as potent poly(ADP-ribose) glycohydrolase (PARG) inhibitors. *Bioorg. Med. Chem. Lett.* 24, 3802–3806
- [136] Slama, J. T., Aboul-Ela, N., and Jacobson, M. K. (1995) Mechanism of Inhibition of Poly(ADP-ribose) Glycohydrolase by Adenosine Diphosphate (Hydroxymethyl) pyrrolidinediol. *J. Med. Chem.* 38, 4332–4336
- [137] Nguyen, P. T. V., Yu, H., and Keller, P. A. (2014) Discovery of in silico hits targeting the nsP3 macro domain of chikungunya virus. *J. Mol. Model.* 20, 1–12
- [138] Stowell, A. I. J., James, D. I., Waddell, I. D., Bennett, N., Truman, C., Hardern, I. M., and Ogilvie, D. J. (2016) A high-throughput screening-compatible homogeneous time-resolved fluorescence assay measuring the glycohydrolase activity of human poly(ADP-ribose) glycohydrolase. *Anal. Biochem.* 503, 58–64
- [139] Peng, B., Thorsell, A.-G., Karlberg, T., Schöler, H., and Yao, S. Q. (2016) Small Molecule Microarray Based Discovery of PARP14 Inhibitors. *Angew. Chemie Int. Ed.* 55, 1–7

- [140] Upton, K., Meyers, M., Thorsell, A.-G., Karlberg, T., Holechek, J., Lease, R., Schey, G., Wolf, E., Lucente, A., Schüler, H., and Ferraris, D. (2017) Design and synthesis of potent inhibitors of the mono(ADP-ribosyl)transferase, PARP14. *Bioorg. Med. Chem. Lett.* 27, 5–9
- [141] Yoneyama-Hirozane, M., Matsumoto, S., Toyoda, Y., Saikatendu, K. S., Zama, Y., Yonemori, K., Oonishi, M., Ishii, T., and Kawamoto, T. (2017) Identification of PARP14 inhibitors using novel methods for detecting auto-ribosylation. *Biochem. Biophys. Res. Commun.* 486, 626–631
- [142] Ladbury, J. E. (2004) Application of Isothermal Titration Calorimetry in the Biological Sciences: Things are heating up! *BioTechniques* 37, 885–887
- [143] Holdgate, G. A., and Ward, W. H. J. (2005) Measurements of binding thermodynamics in drug discovery. *Drug Discov. Today* 10, 1543–1550
- [144] Vedadi, M., Arrowsmith, C. H., Allali-Hassani, A., Senisterra, G., and Wasney, G. A. (2010) Biophysical characterization of recombinant proteins: A key to higher structural genomics success. *J. Struct. Biol.* 172, 107–119
- [145] Fahrner, J., Kranaster, R., Altmeyer, M., Marx, A., and Bürkle, A. (2007) Quantitative analysis of the binding affinity of poly(ADP-ribose) to specific binding proteins as a function of chain length. *Nucleic Acids Res.* 35, e143
- [146] Tan, E. S., Krukenberg, K. A., and Mitchison, T. J. (2012) Large scale preparation and characterisation of poly(ADP-ribose) and defined length polymers. *Anal. Biochem.* 428, 126–136
- [147] Li, C., Debing, Y., Jankevicius, G., Neyts, J., Ahel, I., Coutard, B., and Canard, B. (2016) Viral Macro Domains Reverse Protein ADP-Ribosylation. *J. Virol.* 90, 8478–8486
- [148] Mendoza-Alvarez, H., and Alvarez-Gonzalez, R. (1993) Poly(ADP-ribose) Polymerase Is a Catalytic Dimer and the Automodification Reaction Is Intermolecular. *J. Biol. Chem.* 268, 22575–22580
- [149] Fedorov, O., Niesen, F. H., and Knapp, S. (2012) Kinase Inhibitor Selectivity Profiling Using Differential Scanning Fluorimetry. *Methods Mol. Biol.* 795, 109–118
- [150] Niesen, F. H., Berglund, H., and Vedadi, M. (2007) The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* 2, 2212–2221
- [151] Schubert, T. (2015) nanoDSF-precisely revealing protein folding and stability. *Booklet by 2bind GmbH, Regensburg*, retrieved from <http://2bind.de/molecularinteractions/wp-content/uploads/2016/02/2bind-nanoDSF-booklet.pdf>
- [152] Ferrando-May, E., Tomas, M., Blumhardt, P., Stöckl, M., Fuchs, M., and Leitenstorfer, A. (2013) Highlighting the DNA damage response with ultrashort laser pulses in the near infrared and kinetic modeling. *Front. Genet.* 4, 1–8
- [153] Krasin, F., and Hutchinson, F. (1978) Double-strand breaks from single photochemical events in DNA containing 5-bromouracil. *Biophys. J.* 24, 645–656
- [154] Jakob, B., Rudolph, J. H., Gueven, N., Lavin, M. F., and Taucher-Scholz, G. (2005) Live Cell Imaging of Heavy-Ion-Induced Radiation Responses by Beamline Microscopy. *Radiat. Res.* 163, 681–690

- [155] COSMIC v83 (2017) COSMIC-Catalogue Of Somatic Mutations In Cancer [website], http://cancer.sanger.ac.uk/cosmic/gene/analysis?ln=PARP14_ENST00000474629, accessed December 2017
- [156] Forbes, S. A., Beare, D., Boutselakis, H., Bamford, S., Bindal, N., Tate, J., Cole, C. G., Ward, S., Dawson, E., Ponting, L., Stefancsik, R., Harsha, B., YinKok, C., Jia, M., Jubb, H., Sondka, Z., Thompson, S., De, T., and Campbell, P. J. (2017) COSMIC: Somatic cancer genetics at high-resolution. *Nucleic Acids Res.* 45, D777–D783
- [157] Pemberton, L. F., and Paschal, B. M. (2005) Mechanisms of Receptor-Mediated Nuclear Import and Nuclear Export. *Traffic* 6, 187–198
- [158] Macara, I. G. (2001) Transport into and out of the Nucleus. *Microbiol. Mol. Biol. Rev.* 65, 570–594
- [159] Lange, A., Mills, R. E., Lange, C. J., Stewart, M., Devine, S. E., and Corbett, A. H. (2007) Classical Nuclear Localization Signals: Definition, Function, and Interaction with Importin α . *J. Biol. Chem.* 282, 5101–5105
- [160] Longenecker, K. L., Garrard, S. M., Sheffield, P. J., and Derewenda, Z. S. (2001) Protein crystallization by rational mutagenesis of surface residues: Lys to Ala mutations promote crystallization of RhoGDI. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* 57, 679–688
- [161] Watterson, S. H., De Lucca, G. V., Shi, Q., Langevine, C. M., Liu, Q., Batt, D. G., Beaudoin Bertrand, M., Gong, H., Dai, J., Yip, S., Li, P., Sun, D., Wu, D. R., Wang, C., Zhang, Y., Traeger, S. C., Pattoli, M. A., Skala, S., Cheng, L., Obermeier, M. T., Vickery, R., Discenza, L. N., D'Arienzo, C. J., Zhang, Y., Heimrich, E., Gillooly, K. M., Taylor, T. L., Pulicicchio, C., McIntyre, K. W., Galella, M. A., Tebben, A. J., Muckelbauer, J. K., Chang, C., Rampulla, R., Mathur, A., Salter-Cid, L., Barrish, J. C., Carter, P. H., Fura, A., Burke, J. R., and Tino, J. A. (2016) Discovery of 6-Fluoro-5-(R)-(3-(S)-(8-fluoro-1-methyl-2,4-dioxo-1,2-dihydroquinazolin-3(4H)-yl)-2-methylphenyl)-2-(S)-(2-hydroxypropan-2-yl)-2,3,4,9-tetrahydro-1H-carbazole-8-carboxamide (BMS-986142): A Reversible Inhibitor of Bruton's Tyrosine Kinase (BTK) Conformationally Constrained by Two Locked Atropisomers. *J. Med. Chem.* 59, 9173–9200
- [162] Uth, K. (2014) Screening and Characterisation of Small Molecule Macro Domain Inhibitors for PARP14. Master thesis, University of Abertay Dundee
- [163] Sieg, C. (2013) Establishing an in vitro screening assay and initial evaluation of inhibitors for human macrodomains. Master thesis, University of Duisburg-Essen
- [164] Strödke, B., Gehring, A. P., and Bracher, F. (2015) Synthesis of Desaza Analogues of Annomontine and Canthin-4-one Alkaloids. *Arch. Pharm. Chem. Life Sci.* 348, 125–131
- [165] Kern, S., Agarwal, S., Huber, K., Gehring, A. P., Strödke, B., Wirth, C. C., Brügl, T., Abodo, L. O., Dandekar, T., Doerig, C., Fischer, R., Tobin, A. B., Alam, M. M., Bracher, F., and Pradel, G. (2014) Inhibition of the SR Protein-Phosphorylating CLK Kinases of *Plasmodium falciparum* Impairs Blood Stage Replication and Malaria Transmission. *PLoS One* 9, e105732
- [166] Human Protein Atlas (2017) The Human Protein Atlas v16.1 [website], <https://www.proteinatlas.org/ENSG00000173193-PARP14/cell>, accessed February 2017
- [167] Uhlen, M., Oksvold, P., Fagerberg, L., Lundberg, E., Jonasson, K., Forsberg, M., Zwahlen, M., Kampf, C., Wester, K., Hober, S., Wernerus, H., Björling, L., and Ponten, F. (2010) Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* 28, 1248–1250

- [168] Axelsson, H., Almqvist, H., Seashore-Ludlow, B., and Lundbäck, T. (2016) Screening for Target Engagement using the Cellular Thermal Shift Assay - CETSA. in *Assay Guidance Manual* (Sittampalam, G., Coussens, N., and Brimacombe, K. et al. eds), pp. 1–30, Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD)
- [169] Molina, D. M., Jafari, R., Ignatushchenko, M., Seki, T., Larsson, E. A., Dan, C., Sreekumar, L., Cao, Y., and Nordlund, P. (2013) Monitoring Drug Target Engagement in Cells and Tissues Using the Cellular Thermal Shift Assay. *Science* 341, 84–87
- [170] Jafari, R., Almqvist, H., Axelsson, H., Ignatushchenko, M., Lundbäck, T., Nordlund, P., and Molina, D. M. (2014) The cellular thermal shift assay for evaluating drug target interactions in cells. *Nat. Protoc.* 9, 2100–2122
- [171] Vagenende, V., Yap, M. G. S., and Trout, B. L. (2009) Mechanisms of Protein Stabilization and Prevention of Protein Aggregation by Glycerol. *Biochemistry* 48, 11084–11096
- [172] Villaverde, A., and Carrió, M. M. (2003) Protein aggregation in recombinant bacteria: Biological role of inclusion bodies. *Biotechnol. Lett.* 25, 1385–1395
- [173] Upadhyay, A. K., Murmu, A., Singh, A., and Panda, A. K. (2012) Kinetics of Inclusion Body Formation and Its Correlation with the Characteristics of Protein Aggregates in *Escherichia coli*. *PLoS One* 7, e33951
- [174] Yamaguchi, H., and Miyazaki, M. (2014) Refolding Techniques for Recovering Biologically Active Recombinant Proteins from Inclusion Bodies. *Biomolecules* 4, 235–251
- [175] Esposito, D., and Chatterjee, D. K. (2006) Enhancement of soluble protein expression through the use of fusion tags. *Curr. Opin. Biotechnol.* 17, 353–358
- [176] Mayer, M. P., and Bukau, B. (2005) Hsp70 chaperones: Cellular functions and molecular mechanism. *Cell. Mol. Life Sci.* 62, 670–684
- [177] Park, S. M., Ahn, K. J., Jung, H. Y., Park, J. H., and Kim, J. (2004) Effects of novel peptides derived from the acidic tail of synuclein (ATS) on the aggregation and stability of fusion proteins. *Protein Eng. Des. Sel.* 17, 251–260
- [178] Cisbio (2017) HTRF® theory in brief. *Product and Service Catalog*, retrieved from <http://auroscience.hu/docs/bros-htrf.pdf>
- [179] Sajewicz, W., and Dlugosz, A. (2000) Cytotoxicity of some Potential DNA Intercalators (Carbazole, Acridine and Anthracene Derivatives) Evaluated through Neutrophil Chemiluminescence. *J. Appl. Toxicol.* 20, 305–312
- [180] Feng, B. Y., and Shoichet, B. K. (2006) A Detergent-Based Assay for the Detection of Promiscuous Inhibitors. *Nat. Protoc.* 1, 550–553
- [181] Lee, H., Torres, J., Truong, L., Chaudhuri, R., Mittal, A., and Johnson, M. E. (2012) Reducing agents affect inhibitory activities of compounds: Results from multiple drug targets. *Anal. Biochem.* 423, 46–53
- [182] Ryan, A. J., Gray, N. M., Lowe, P. N., and Chung, C. W. (2003) Effect of Detergent on “Promiscuous” Inhibitors. *J. Med. Chem.* 46, 3448–3451

- [183] Thorne, N., Auld, D. S., and Inglese, J. (2010) Apparent activity in High-Throughput Screening: Origins of Compound-Dependent Assay Interference. *Curr. Opin. Chem. Biol.* 14, 315–324
- [184] Baell, J. B., and Holloway, G. A. (2010) New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* 53, 2719–2740
- [185] Hoelder, S., Clarke, P. A., and Workman, P. (2012) Discovery of small molecule cancer drugs: Successes, challenges and opportunities. *Mol. Oncol.* 6, 155–176
- [186] Macarron, R., Banks, M. N., Bojanic, D., Burns, D. J., Cirovic, D. A., Garyantes, T., Green, D. V. S., Hertzberg, R. P., Janzen, W. P., Paslay, J. W., Schopfer, U., and Sittampalam, G. S. (2011) Impact of high-throughput screening in biomedical research. *Nat. Rev. Drug Discov.* 10, 188–195
- [187] Irwin, J. J., Duan, D., Torosyan, H., Doak, A. K., Ziebart, K. T., Sterling, T., Tumanian, G., and Shoichet, B. K. (2015) An Aggregation Advisor for Ligand Discovery. *J. Med. Chem.* 58, 7076–7087
- [188] Wiseman, T., Williston, S., Brandts, J. F., and Lin, L. N. (1989) Rapid Measurement of Binding Constants and Heats of Binding Using a New Titration Calorimeter. *Anal. Biochem.* 179, 131–137
- [189] Turnbull, W. B., and Daranas, A. H. (2003) On the Value of c : Can Low Affinity Systems Be Studied by Isothermal Titration Calorimetry? *J. Am. Chem. Soc.* 125, 14859–14866
- [190] Glickman, J. F., Wu, X., Mercuri, R., Illy, C., Bowen, B. R., He, Y., and Sills, M. (2002) A Comparison of ALPHAScreen, TR-FRET, and TRF as Assay Methods for FXR Nuclear Receptors. *J. Biomol. Screen.* 7, 3–10
- [191] Rees, D. C., Congreve, M., Murray, C. W., and Carr, R. (2004) Fragment-Based Lead Discovery. *Nat. Rev. Drug Discov.* 3, 660–672
- [192] Congreve, M., Chessari, G., Tisi, D., and Woodhead, A. J. (2008) Recent Developments in Fragment-Based Drug Discovery. *J. Med. Chem.* 51, 3661–3680
- [193] Hubbard, R. E., and Murray, J. B. (2011) Experiences in fragment-based lead discovery. *Methods Enzymol.* 493, 509–531
- [194] Hopkins, A. L., Groom, C. R., and Alex, A. (2004) Ligand efficiency: A useful metric for lead selection. *Drug Discov. Today* 9, 430–431
- [195] Schulz, M. N., and Hubbard, R. E. (2009) Recent progress in fragment-based lead discovery. *Curr. Opin. Pharmacol.* 9, 615–621
- [196] Erlanson, D. A., Fesik, S. W., Hubbard, R. E., Jahnke, W., and Jhoti, H. (2016) Twenty years on: the impact of fragments on drug discovery. *Nat. Rev. Drug Discov.* 15, 605–619
- [197] Park, S., Mann, J., and Li, N. (2013) Targeted Inhibitor Design: Lessons from Small Molecule Drug Design, Directed Evolution, and Vaccine Research. *Chem. Eng. Process Tech.* 1, 1–6
- [198] Collins, P. M., Ng, T., Talon, R., Nekrosiute, K., Pearce, N. M., and Delft, F. Von (2017) Gentle, fast and effective crystal soaking by acoustic dispensing. *Acta Crystallogr. Sect. D: Struct. Biol.* 73, 246–255

- [199] Danley, D. E. (2006) Crystallization to obtain protein-ligand complexes for structure-aided drug design. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* 62, 569–575
- [200] Pellegrini, E., Piano, D., and Bowler, M. W. (2011) Direct cryocooling of naked crystals: are cryoprotection agents always necessary? *Acta Crystallogr. Sect. D: Biol. Crystallogr.* 67, 902–906
- [201] Cox, O. B., Krojer, T., Collins, P., Monteiro, O., Talon, R., Bradley, A., Fedorov, O., Amin, J., Marsden, B. D., Spencer, J., von Delft, F., and Brennan, P. E. (2016) A poised fragment library enables rapid synthetic expansion yielding the first reported inhibitors of PHIP(2), an atypical bromodomain. *Chem. Sci.* 7, 2322–2330
- [202] Patel, D., Bauman, J. D., and Arnold, E. (2014) Advantages of crystallographic fragment screening: Functional and mechanistic insights from a powerful platform for efficient drug discovery. *Prog. Biophys. Mol. Biol.* 116, 92–100
- [203] Krojer, T., Talon, R., Pearce, N., Collins, P., Douangamath, A., Brandao-Neto, J., Dias, A., Marsden, B., and Von Delft, F. (2017) The *XChemExplorer* graphical workflow tool for routine or large-scale protein–ligand structure determination research papers. *Acta Crystallogr. Sect. D: Struct. Biol.* 73, 267–278
- [204] Long, F., Nicholls, R. A., Emsley, P., Grazulis, S., Merkys, A., Vaitkus, A., and Murshudov, G. N. (2017) *AceDRG*: a stereochemical description generator for ligands research papers. *Acta Crystallogr. Sect. D: Struct. Biol.* 73, 112–122
- [205] Pearce, N. M., Krojer, T., Bradley, A. R., Collins, P., Nowak, R. P., Talon, R., Marsden, B. D., Kelm, S., Shi, J., Deane, C. M., and Delft, F. Von (2017) A multi-crystal method for extracting obscured crystallographic states from conventionally uninterpretable electron density. *Nat. Commun.* 8, 1–8
- [206] Hubbard, R. E. (2015) Fragment Based Lead Discovery. in *The Handbook of Medicinal Chemistry - principles and practice* (Davis, A., and Ward, S. E. eds), pp. 122–153, The Royal Society of Chemistry
- [207] Whittaker, C. A., and Hynes, R. O. (2002) Distribution and Evolution of von Willebrand/Integrin A Domains: Widely Dispersed Domains with Roles in Cell Adhesion and Elsewhere. *Mol. Biol. Cell* 13, 3369–3387
- [208] Tadayon, S. H., Vaziri-Pashkam, M., Kahali, P., Ansari Dezfouli, M., and Abbassian, A. (2016) Common Genetic Variant in VIT Is Associated with Human Brain Asymmetry. *Front. Hum. Neurosci.* 10, 1–8
- [209] Human Protein Atlas (2017) The Human Protein Atlas v17 [website], <https://www.proteinatlas.org/ENSG00000205221-VIT/cell>, accessed August 2017
- [210] Hart, T., Chandrashekhar, M., Aregger, M., Steinhart, Z., Brown, K. R., MacLeod, G., Mis, M., Zimmermann, M., Fradet-Turcotte, A., Sun, S., Mero, P., Dirks, P., Sidhu, S., Roth, F. P., Rissland, O. S., Durocher, D., Angers, S., and Moffat, J. (2015) High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* 163, 1515–1526
- [211] Sander, J. D., and Joung, J. K. (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32, 347–355
- [212] Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., and Zhang, F. (2013) Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* 8, 2281–2308

- [213] Landry, J. J. M., Pyl, P. T., Rausch, T., Zichner, T., Tekkedil, M. M., Stütz, A. M., Jauch, A., Aiyar, R. S., Pau, G., Delhomme, N., Gagneur, J., Korbel, J. O., Huber, W., and Steinmetz, L. M. (2013) The Genomic and Transcriptomic Landscape of a HeLa Cell Line. *G3: Genes, Genomes, Genetics* 3, 1213–1224
- [214] Janssen, A., and Medema, R. H. (2013) Genetic instability: tipping the balance. *Oncogene* 32, 4459–4470
- [215] Holkers, M., Maggio, I., Henriques, S. F. D., Janssen, J. M., Cathomen, T., and Gonçalves, M. A. F. V (2014) Adenoviral vector DNA for accurate genome editing with engineered nucleases. *Nat. Methods* 11, 1051–1057
- [216] Dehairs, J., Talebi, A., Cherifi, Y., and Swinnen, J. V. (2016) CRISP-ID: decoding CRISPR mediated indels by Sanger sequencing. *Sci. Rep.* 6, 1–5
- [217] Zaidel-Bar, R., Cohen, M., Addadi, L., and Geiger, B. (2004) Hierarchical assembly of cell-matrix adhesion complexes. *Biochem. Soc. Trans.* 32, 416–20
- [218] Burridge, K., Fath, K., Kelly, T., Nuckolls, G., and Turner, C. (1988) Focal Adhesions: Transmembrane Junctions Between The Extracellular Matrix And The Cytoskeleton. *Annu. Rev. Cell Dev. Biol.* 4, 487–525
- [219] Kramer, N., Walzl, A., Unger, C., Rosner, M., Krupitza, G., Hengstschläger, M., and Dolznig, H. (2013) *In vitro* cell migration and invasion assays. *Mutat. Res.* 752, 10–24
- [220] Liang, C.-C., Park, A. Y., and Guan, J.-L. (2007) *In vitro* scratch assay: a convenient and inexpensive method for analysis of cell migration *in vitro*. *Nat. Protoc.* 2, 329–333
- [221] Yarrow, J. C., Perlman, Z. E., Westwood, N. J., and Mitchison, T. J. (2004) A high-throughput cell migration assay using scratch wound healing, a comparison of image-based readout methods. *BMC Biotechnol.* 4, 1–9
- [222] Lengsfeld, A. M., and Maurer-Schultze, B. (1986) The effect of ara-C on survival and proliferation of HeLa cells. *J. Cancer Res. Clin. Oncol.* 111, 220–228
- [223] Hanahan, D., and Weinberg, R. A. (2011) Hallmarks of Cancer: The Next Generation. *Cell* 144, 646–674
- [224] Warburg, O. (1956) On the Origin of Cancer Cells. *Science* 123, 309–314
- [225] Vander Heiden, M. G., Cantley, L. C., and Thompson, C. B. (2009) Understanding the Warburg Effect: The Metabolic Requirements of Cell Proliferation. *Science* 324, 1029–1033
- [226] Cairns, R. A., Harris, I. S., and Mak, T. W. (2011) Regulation of cancer cell metabolism. *Nat. Rev. Cancer* 11, 85–95
- [227] Christofk, H. R., Vander Heiden, M. G., Harris, M. H., Ramanathan, A., Gerszten, R. E., Wei, R., Fleming, M. D., Schreiber, S. L., and Cantley, L. C. (2008) The M2 splice isoform of pyruvate kinase is important for cancer metabolism and tumour growth. *Nature* 452, 230–233
- [228] Dong, G., Mao, Q., Xia, W., Xu, Y., Wang, J., Xu, L., and Jiang, F. (2016) PKM2 and cancer: The function of PKM2 beyond glycolysis. *Oncol. Lett.* 11, 1980–1986
- [229] Wu, S., and Le, H. (2013) Dual roles of PKM2 in cancer metabolism. *Acta Biochim. Biophys. Sin.* 45, 27–35

- [230] TeSlaa, T., and Teitell, M. A. (2014) Techniques to Monitor Glycolysis. *Methods Enzymol.* 542, 91–114
- [231] Reitzer, L. J., Wice, B. M., and Kennel, D. (1979) Evidence That Glutamine, Not Sugar, Is the Major Energy Source for Cultured HeLa Cells. *J. Biol. Chem.* 254, 2669–2676
- [232] Xie, J., Wu, H., Dai, C., Pan, Q., Ding, Z., Hu, D., Ji, B., Luo, Y., and Hu, X. (2014) Beyond Warburg effect – dual metabolic nature of cancer cells. *Sci. Rep.* 4, 1–12
- [233] Branzei, D., and Foiani, M. (2010) Maintaining genome stability at the replication fork. *Nat. Rev. Mol. Cell Biol.* 11, 208–219
- [234] Franchitto, A., and Pichierri, P. (2011) Understanding the molecular basis of common fragile sites instability: Role of the proteins involved in the recovery of stalled replication forks. *Cell Cycle* 10, 4039–4046
- [235] Budzowska, M., and Kanaar, R. (2009) Mechanisms of Dealing with DNA Damage-Induced Replication Problems. *Cell Biochem. Biophys.* 53, 17–31
- [236] San Filippo, J., Sung, P., and Klein, H. (2008) Mechanism of Eukaryotic Homologous Recombination. *Annu. Rev. Biochem.* 77, 229–257
- [237] Vesela, E., Chroma, K., Turi, Z., and Mistrik, M. (2017) Common Chemical Inductors of Replication Stress: Focus on Cell-Based Studies. *Biomolecules* 7, 1–36
- [238] Krakoff, I. H., Brown, N. C., and Reichard, P. (1968) Inhibition of Ribonucleoside Diphosphate Reductase by Hydroxyurea. *Cancer Res.* 28, 1559–1565
- [239] Petermann, E., Orta, M. L., Issaeva, N., Schultz, N., and Helleday, T. (2010) Hydroxyurea-Stalled Replication Forks Become Progressively Inactivated and Require Two Different RAD51-Mediated Pathways for Restart and Repair. *Mol. Cell* 37, 492–502
- [240] Ziv, Y., Bielopolski, D., Galanty, Y., Lukas, C., Taya, Y., Schultz, D. C., Lukas, J., Bekker-Jensen, S., Bartek, J., and Shiloh, Y. (2006) Chromatin relaxation in response to DNA double-strand breaks is modulated by a novel ATM- and KAP-1 dependent pathway. *Nat. Cell Biol.* 8, 870–876
- [241] Shiloh, Y. (2006) The ATM-mediated DNA-damage response: taking shape. *Trends Biochem. Sci.* 31, 402–410
- [242] Noon, A. T., Shibata, A., Rief, N., Löbrich, M., Stewart, G. S., Jeggo, P. A., and Goodarzi, A. A. (2010) 53BP1-dependent robust localized KAP-1 phosphorylation is essential for heterochromatic DNA double-strand break repair. *Nat. Cell Biol.* 12, 177–184
- [243] Mochan, T. A., Venere, M., DiTullio, R. A., and Halazonetis, T. D. (2004) 53BP1, an activator of ATM in response to DNA damage. *DNA Repair* 3, 945–952
- [244] Lee, J.-H., Goodarzi, A. a, Jeggo, P. a, and Paull, T. T. (2010) 53BP1 promotes ATM activity through direct interactions with the MRN complex. *EMBO J.* 29, 574–585
- [245] Bryant, H. E., Schultz, N., Thomas, H. D., Parker, K. M., Flower, D., Lopez, E., Kyle, S., Meuth, M., Curtin, N. J., and Helleday, T. (2005) Specific killing of BRCA2-deficient tumours with inhibitors of poly(ADP-ribose) polymerase. *Nature* 434, 913–917
- [246] Pandita, T. K., Lieberman, H. B., Lim, D., Dhar, S., Zheng, W., Taya, Y., and Kastan, M. B. (2000) Ionizing radiation activates the ATM kinase throughout the cell cycle. *Oncogene* 19, 1386–1391

- [247] Lai, X., Broderick, R., Bergoglio, V., Zimmer, J., Badie, S., Niedzwiedz, W., Hoffmann, J.-S., and Tarsounas, M. (2017) MUS81 nuclease activity is essential for replication stress tolerance and chromosome segregation in BRCA2-deficient cells. *Nat. Commun.* 8, 1–10
- [248] Cho, S. W., Kim, S., Kim, Y., Kweon, J., Kim, H. S., Bae, S., and Kim, J.-S. (2014) Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* 24, 132–141
- [249] Ran, F. A., Hsu, P. D., Lin, C. Y., Gootenberg, J. S., Konermann, S., Trevino, A. E., Scott, D. A., Inoue, A., Matoba, S., Zhang, Y., and Zhang, F. (2013) Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. *Cell* 154, 1380–1389
- [250] Pierce, A. J., Johnson, R. D., Thompson, L. H., and Jasin, M. (1999) XRCC3 promotes homology-directed repair of DNA damage in mammalian cells. *Genes Dev.* 13, 2633–2638
- [251] Invitrogen (2003) Gateway® Technology: A universal technology to clone DNA sequences for functional analysis and expression in multiple systems. *User Guide MAN0000282*, retrieved from <https://tools.thermofisher.com/content/sfs/manuals/gatewayman.pdf>
- [252] LI-COR (2003) Western Blot Analysis. *Application Protocols Manual*, retrieved from <https://www.bio.purdue.edu/lab/kasinski/LabDocs/Licor%20Odyssey%20CLx%20Application%20Protocol%20Manual.pdf>
- [253] Méndez, J., and Stillman, B. (2000) Chromatin association of human origin recognition complex, Cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol. Cell. Biol.* 20, 8602–8612
- [254] Winter, G. (2010) xia2: an expert system for macromolecular crystallography data reduction. *J. Appl. Crystallogr.* 43, 186–190
- [255] Waterman, D. G., Winter, G., Gildea, R. J., Parkhurst, J. M., Brewster, A. S., Sauter, N. K., and Evans, G. (2016) Diffraction-geometry refinement in the DIALS framework. *Acta Crystallogr. Sect. D: Struct. Biol.* 72, 558–575
- [256] Winn, M. D., Charles, C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G. W., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* 67, 235–242
- [257] Wojdyr, M., Keegan, R., Winter, G. W., and Ashton, A. (2013) DIMPLE - a pipeline for the rapid generation of difference maps from protein crystals with putatively bound ligands. *Acta Crystallogr. Sect. A* 69, s299
- [258] Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F., and Vagin, A. A. (2011) REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* 67, 355–367
- [259] Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. Sect. D: Biol. Crystallogr.* 66, 486–501
- [260] PerkinElmer (2004) A Practical Guide to working with AlphaScreen. *User Guide*, retrieved from <https://www.urmc.rochester.edu/MediaLibraries/URMCMedia/hts/documents/AlphaScreenPracticalGuide.pdf>

- [261] Philpott, M., Yang, J., Tumber, T., Fedorov, O., Uttarkar, S., Filippakopoulos, P., Picaud, S., Keates, T., Felletar, I., Ciulli, A., Knapp, S., and Heightman, T. D. (2011) Bromodomain-peptide displacement assays for interactome mapping and inhibitor discovery. *Mol. Biosyst.* 7, 2899–2908
- [262] Cisbio (2014) EPIgeneous™ Binding Domain Kit B. *Product and Service Catalog*, document reference: 62BDBPEH Rev N°1, retrieved from https://www.cisbio.com/sites/default/files/ressources/cisbio_dd_pi_62BDBPEH.pdf
- [263] NanoTemper Technologies (2015) Prometheus NT.48. *Product Information*, retrieved from http://www.ac.nctu.edu.tw/uploads/bulletin_file/file/56ea73931d41c838b100000e/Product_Information_Prometheus_NT48.pdf
- [264] Tobias, R., and Kumaraswamy, S. (2013) Biomolecular Binding Kinetics Assays on the Octet Platform. *Appl. Note by fortéBio (Pall Life Sciences)*, retrieved from http://csbi.mit.edu/instrumentation/ForteBio_App_Note_14.pdf
- [265] Martinez Molina, D., and Nordlund, P. (2016) The Cellular Thermal Shift Assay: A Novel Biophysical Assay for In Situ Drug Target Engagement and Mechanistic Biomarker Studies. *Annu. Rev. Pharmacol. Toxicol.* 56, 141–161
- [266] Agilent Technologies (2017) Agilent Seahorse XF Glycolysis Stress Test Kit. *User Guide*, retrieved from https://www.agilent.com/cs/library/usermanuals/public/XF_Glycolysis_Stress_Test_Kit_User_Guide.pdf
- [267] EssenBioscience (2015) IncuCyte ZOOM® 96-Well Scratch Wound Cell Migration & Invasion Assays. *User Manual*, retrieved from https://www.essenbioscience.com/media/uploads/files/8400-0012-C00_-_Scratch_Wound_Cell_Migration__Invasion_User_Manual.pdf
- [268] Gibbs-Seymour, I., Fontana, P., Rack, J. G. M., and Ahel, I. (2016) HPF1/C4orf27 Is a PARP-1-Interacting Protein that Regulates PARP-1 ADP-Ribosylation Activity. *Mol. Cell* 62, 432–442
- [269] Tucker, J. A., Bennett, N., Brassington, C., Durant, S. T., Hassall, G., Holdgate, G., McAlister, M., Nissink, J. W. M., Truman, C., and Watson, M. (2012) Structures of the Human Poly (ADP-Ribose) Glycohydrolase Catalytic Domain Confirm Catalytic Mechanism and Explain Inhibition by ADP-HPD Derivatives. *PLoS One* 7, e50889
