

Structural and functional analysis of Serine-ADP-ribosylation



Pietro Fontana

**Lincoln College
University of Oxford**

A thesis submitted for the degree of Doctor of Philosophy

At the University of Oxford

Trinity Term, 2018

Structural and functional analysis of Serine-ADP-ribosylation

Pietro Fontana, Lincoln College

A thesis submitted for the degree of Doctor of Philosophy, Trinity Term 2018

Abstract

ADP-ribosylation is a protein post-translational modification involved in several biological pathways, including the DNA damage response. PARP1, one of the primary sensors of DNA breaks, and the main enzyme responsible for ADP-ribosylation has been the subject of study for several years. Despite the revelations from previous studies, we still lack a full understanding behind PARP1 regulation and activity. In this thesis, I will describe the identification and characterisation of a novel interactor/regulator of the PARP1 cycle, which we named Histone Parylation Factor 1 (HPF1), with particular emphasis of its role in DNA repair. In the second chapter, I will show a structural and biochemical characterisation of HPF1 and the discovery of HPF1 as a specificity factor inducing a novel post-translational modification named serine-ADP-ribosylation. In the third chapter, I will show that serine-ADP-ribosylation induced by HPF1 is a reversible post-translational modification. Further, I will show that the only human enzyme able to reverse serine-ADP-ribosylation is the previously poorly characterised glycohydrolase ARH3, and show an overall biochemical characterisation of ARH3. Finally, in the fourth and last experimental chapter, I will show that serine-ADP-ribosylation is a post-translational modification present also in the model organism *Drosophila melanogaster*. However, I'll show that in *Drosophila* the enzyme responsible for its removal is DmParg, as *Drosophila* do not possess an ARH3 homolog. Finally, I will show a biochemical characterisation of DmParg and preliminary data for its structural characterisation.

Acknowledgments

I would like to thank first of all my DPhil supervisor Dr. Ivan Ahel for allowing me to work on these projects and for leading me to become the researcher I am today. A big thank to all the members of the lab who have come and gone over the last four years, specially to Osamu Suyari for all the laugh we had together, to Edward Bartlett for all the good talks and reciprocal help, to Johannes Rack for always suggesting new techniques and giving me different options, to Antonio Ariza for all the help with crystallography and the countless synchrotron trips, to Kerryanne Crawford for all the great help with my future project and always having a smile, and to Marcin Suskiewicz for always being helpful from the very beginning. Finally, a big thank you to my dear Angela who helped me for many years to achieve all my goals and find my path, and her family Anna, Claudia and Ignazio who believed in me from the very first moment and always encouraged me.

Contents

Chapter 1. Introduction	12
1.1 PAR cycle	12
1.1.1 Writers of ADP-ribosylation	13
1.1.2 Readers of ADP-ribosylation	17
1.1.2.1 PBZ - PAR binding zinc finger	17
1.1.2.2 PBM - PAR binding motif	18
1.1.2.3 WWE domain	18
1.1.2.4 Macrodomain	19
1.1.2.5 Forkhead-associated (FHA)	20
1.1.2.6 Oligonucleotide/oligosaccharide-binding (OB fold)	21
1.1.3 Erasers of ADP-ribosylation	23
1.2 DNA damage repair	27
1.3 PAR dependent DNA damage repair	37
1.4 Therapeutic roles of PARPs inhibitor	39
1.5 Identification of ADP-ribosylation sites	41
1.6 Scope of the thesis	42
Chapter 2. Materials and methods	44
2.1 Standard PCR for cloning	44
2.2 Standard mutagenesis	45
2.3 Cloning, expression and purification of HsHPF1, DmHpf1 and NvHpf1	46
2.4 Poly-ADP-ribose binding assay	50
2.5 Crystallography data collection and processing	51
2.5.1 Data collection	51
2.5.2 Data processing	52
2.5.2.1 Spot finding and auto-indexing	52
2.5.2.2 Parameter refinement	52
2.5.2.3 Integration	53
2.5.2.4 Scaling and merging	53
2.5.3 Phasing	54
2.5.3.1 Molecular replacement	54
2.5.3.2 Isomorphous replacement	55
2.5.4 Phase improvement, model building and refinement	56
2.6 HPF1 Crystallisation, data collection, data processing, phasing and model building	57
2.7 Cloning, expression and purification of GST-tagged HPF1 and PARP1	60

2.7.1 GST pull-down assay	62
2.8 Cloning, expression and purification of HsPARP1 and DmParp1	63
2.9 PARP1 auto-modification assay	67
2.10 <i>In vitro</i> glycohydrolase assays	67
2.11 Cloning, expression and purification of PARP2	68
2.12 EMSA Electrophoretic mobility shift assay	71
2.13 Cloning, expression and purification of ARH1, ARH2 and ARH3	72
2.14 Purification of Ser-ADP-ribosylated histone peptide	76
2.15 ARH3 time course	76
2.16 Arginine-ADP-ribosylation of K562 and S2 cells extract	77
2.17 Lysine-ADP-ribosylation of H3 histone peptide	77
2.18 Recombinant histones and histones peptides	78
2.19 Cloning, expression and purification of DmParg, DmTarg1, DmTarg2 and DmTarg3	79
2.20 DmParg Δ 554-723 crystallisation and data collection	84
2.21 Proteins accession numbers	84

Chapter 3. Structural and biochemical characterisation of HPF1 and Ser-ADP-ribosylation	85
3.1 Identification of HPF1	86
3.2 HPF1 structural characterisation	89
3.3 Functional characterisation of HPF1	97
3.4 HPF1 interacts with PARP1 and PARP2 via catalytic domain and with recombinant histones	98
3.5 HPF1 suppresses PARP1 auto-modification	102
3.6 HPF1 promotes PARP1 trans-modification of histones	105
3.7 HPF1 regulates PARP1 and PARP2 activity	109
3.8 Serine ADP-ribosylation specificity	111
3.9 Minimal PARP1 region essential for Ser-ADP-ribosylation	115
3.10 HPF1 binds ssDNA and dsDNA (EMSA)	119
3.11 Conclusion on HPF1 characterisation	121

Chapter 4. Ser-ADP-ribosylation reversal by the hydrolase ARH3	122
4.1 Histone Ser-ADP-ribose modification is reversed by the ARH3 hydrolyse	123
4.2 Characterisation of ARH3 activities	131
4.3 Specificity of ARH3 activity	137
4.4 Conclusion on ARH3 characterisation	141

Chapter 5. Ser-ADP-ribosylation in <i>Drosophila melanogaster</i>	142
5.1 DmParp1/DmHpf1 complex in <i>Drosophila</i> can Ser-ADP-ribosylate histone proteins	144
5.2 DmParg hydrolyses Ser-ADP-ribosylation	148
5.3 Characterisation of the serine ADP-ribosylhydrolase activity of DmParg	155
5.4 Removal of mono-Ser-ADP-ribosylation by Parg is conserved in <i>Tetrahymena</i>	160
5.5 Structural characterisation of DmParg	163
5.6 Uncoupling mono-Ser-ADP-ribosylation and poly-ADP-ribosylation glycohydrolase activities	167
5.7 DmParg specificity assays	170
5.8 Conclusions on Ser-ADP-ribosylation in <i>Drosophila melanogaster</i>	174
Chapter 6. Discussion	175
6.1 Aims of the project	175
6.2 Identification of HPF1	176
6.3 Ser-ADP-ribosylation HPF1 dependent	178
6.4 Removal of Ser-ADP-ribosylation by ARH3	182
6.4 Ser-ADP-ribosylation in <i>Drosophila melanogaster</i>	184
References	188

Table of figures

Figure 1.1. Schematic representation of mono- and poly-ADP-ribosylation	15
Figure 1.2. Schematic representation of PARP family members	16
Figure 1.3. Structural representation of PAR binding domains.	22
Figure 1.4. PARG and ARH3 structural representation	26
Figure 1.5. Schematic BER representation.	31
Figure 1.6. Schematic NER representation.	32
Figure 1.8. Schematic SSB representation	34
Figure 1.9. Schematic HR representation	35
Figure 1.10. Schematic NHEJ representation.	36
Figure 2.1. HPF1 purification	49
Figure 2.2. PARP1 purification	66
Figure 2.3. PARP2 purification	70
Figure 2.4. ARH3 purification	75
Figure 2.5. DmParg purification	83
Figure 3.1. Schematic representation of HPF1 domain	88
Figure 3.2. NvHpf1 and DmHpf1 structural representation.	94
Figure 3.3. NvHpf1 and DmHpf1 superimpose structures	95
Figure 3.4. NvHpf1 linker loop, alignment and conservation	96
Figure 3.5. GST-PARP1 and N-His-HPF1 pull-down assay	99
Figure 3.6. GST-PARP1-CAT/GST-PARP2-CAT and N-His HPF1 pull-down assay	100
Figure 3.7. GST-HPF1 and histone tetramer H3/H4 pull-down assay	101
Figure 3.8. A - HPF1 suppress PARP1 auto-modification. B - HPF1 glycohydrolase assay	104
Figure 3.9. HPF1 promotes histone octamer ADP-ribosylation	106
Figure 3.10. HPF1 promotes histone tetramer ADP-ribosylation	107
Figure 3.11. HPF1 promotes histone H1 ADP-ribosylation	108
Figure 3.12. HPF1 promote Ser-ADP-ribosylation of histone tetramer	110

Figure 3.13. HPF1 promotes Ser-ADP-ribosylation of PARP1 BRCT domain	113
Figure 3.14. HPF1 promotes Ser-ADP-ribosylation of PARP1 FL	114
Figure 3.15. PARP1 ZnF is essential for Ser-ADP	117
Figure 3.16. Ser-ADP-ribosylation DNA independent by PARP1 Δ HD	118
Figure 3.17. EMSA, HPF1 interaction with ssDNA/dsDNA	120
Figure 4.1. ARH3 cleaves mono-Ser-ADPr from H3 histone peptide	126
Figure 4.2. ARH3 cleaves Ser-ADPr from recombinant H1 histone	127
Figure 4.3. PARG/ARH3 cleaves all ADPr from recombinant H1	128
Figure 4.4. ARH3 cleaves poly-ADPr and mono-Ser-ADPr from recombinant H1 histone	129
Figure 4.5. ARH3 cleaves poly-ADPr and mono-Ser-ADPr from PARP1 BRCT domain	130
Figure 4.6. ARH3 cleaves mono-Ser-ADPr from purified H3 histone peptide	133
Figure 4.7. ARH3 cleaves mono-Ser-ADPr from purified H2B histone peptide	134
Figure 4.8. ARH3 catalytic mutant and magnesium dependent	135
Figure 4.9. A - ARH3 concentration dependent assay. B - ARH3 time course	136
Figure 4.10. A - ARH3 specificity assay on mono-Glu-ADPr. B - ARH3 specificity assay on mono-Arg-ADPr	139
Figure 4.11. ARH3 specificity assay on mono-Lys-ADPr	140
Figure 5.1. DmParp1/DmHpf1 trans-ADPr of H3 histone peptide	146
Figure 5.2. DmParp1/DmHpf1 induce Ser-ADPr of H3 histone peptide	147
Figure 5.3. DmParg cleaves mono-Ser-ADPr from H3 histone peptide	151
Figure 5.4. DmParg cleaves poly-ADPr from auto-modified PARP1	152
Figure 5.5. DmParg cleaves mono-Ser-ADPr from recombinant histone tetramer H3/H4	153
Figure 5.6. DmParg cleaves mono-Ser-ADPr from H3 histone peptide and poly-ADPr from auto-modified PARP1	154
Figure 5.7. DmParg catalytic mutants and C-terminal truncations	158
Figure 5.8. Parg representative alignment	159
Figure 5.9. TtParg cleaves mono-Ser-ADPr	162
Figure 5.10. Schematic representation of PARG domain	165
Figure 5.11. Representative DmParg crystal	166

Figure 5.12. DmParg mutants, uncoupling mono-Ser-ADPr from poly-ADPr glycohydrolase activity	169
Figure 5.13. DmParg specificity assay on mono-Glu-ADPr	171
Figure 5.14. DmParg specificity assay on mono-Arg-ADPr	172
Figure 5.15. DmParg specificity assay on mono-Lys-ADPr	173
Figure 6.1. HPF1 regulation model	177

Tables

Table 2.1. PCR cloning condition.	44
Table 2.2. Mutagenesis condition.	45
Table 2.3. Primers for cloning HPF1s in pET28 vector and mutagenesis primers.	47
Table 2.4. Primers for cloning HPF1 in pGex-4T1 vector.	60
Table 2.5. Primers for cloning DmParp1 in pET28 vector.	64
Table 2.6. Primers for cloning ARH1 and ARH2 in gateway system and primers for ARH3 mutagenesis.	72
Table 2.7. Primers for cloning DmParg, DmTarg1, DmTarg2, DmTarg3, and primers for DmParg mutagenesis.	80
Table 3.1. HPF1s crystallisation attempts.	92
Table 3.2. Data Collection, phasing, and refinement statistics.	93
Table 4.1. DmParg crystallisation attempts.	164

List of abbreviation

ACN - Acetonitrile.

ADPr - ADP-ribose.

AIF - Apoptosis inducing factor.

ALC1 - Amplified liver cancer protein 1.

APE1 - Apurinic/aprimidinic endonuclease.

APLF - Aprataxin PNK-like protein.

ARH1 - ADP-ribosylhydrolase 1.

ARH2 - ADP-ribosylhydrolase 2.

ARH3 - ADP-ribosylhydrolase 3.

ART - ADP-ribosyl transferase.

ART2.2 - ADP-ribosyltransferase 2.2.

BER - Base excision repair.

BRCA1 - Breast cancer type 1 susceptibility protein.

BRCA2 - Breast cancer type 2 susceptibility protein.

BRCT - BRCA1 C-Terminus.

CAT - Catalytic domain

CHFR - Checkpoint with forkhead and ring finger domains.

DDR - DNA damage response.

DNA-PK - DNA protein kinase.

DSB - Double strand break.

dsDNA - Double strand DNA.

DTT – Dithiothreitol.

EDTA - Acido etilendiamminotetraacetico.

EXO1 - Exonuclease 1.

FHA - Forkhead-associated.

HPF1 Histone parylation factor 1.

HR - Homologous recombination.

hSSB1 - Human single strand binding protein.

IPTG - Isopropyl β -D-1-thiogalactopyranoside.

MACROD1 - MACRO domain-containing protein 1.

MACROD2 - MACRO domain-containing protein 2.

MAD - multi-wavelength anomalous diffraction.

MIR - multiple isomorphous replacement.

MMR - Mismatch repair.

NAD - Nicotinamide adenine dinucleotide.

NER - Nucleotide excision repair.

OB - Oligonucleotide/oligosaccharide-binding.

PAR - Poly-ADP-ribose.

PARG - PAR glycohydrolase.

PARP - PAR polymerase.

PARthanatos - Cell death PAR dependent.

PBM - PAR binding motif.

PBZ - PAR binding zinc finger.

PK - Protein kinase.

PNKP - Polynucleotide kinase 3'-phosphatase.

PTM - Post-translational modification.

RAD51 - Recombinase (DNA repair protein).

RAD52 - Recombinase (DNA repair protein).

RNF - Ring finger protein 146.

RPA - Replication protein A.

SAD - Single-wavelength anomalous diffraction.

SIR - single isomorphous replacement.

SNM1A - Sensitive to nitrogen mustard 1A.

SSB - Single strand break.

ssDNA - Single strand DNA.

TARG1 - Terminal ADP-ribose protein glycohydrolase-1.

TCEP - Tris(2-carboxyethyl)fosfina.

TFA - Trifluoroacetic acid.

TTC5 - Tetratricopeptide repeat domain 5.

WGR - Protein domain named after the three conserved residues tryptophan/glycine/arginine.

WWE - PAR binding motif named after the three conserved residues tryptophan/ tryptophan/glutamate.

XPA - Xeroderma pigmentosum group A-complementing protein (DNA damage recognition and repair factor).

XPC - Xeroderma pigmentosum group C-complementing protein (DNA damage recognition and repair factor).

XRCC1 - X-ray cross complementing protein-1.

ZnF - Zinc finger.

Chapter 1. Introduction

1.1 PAR cycle

ADP-ribosylation is one of the many protein post-translational modifications and is involved in several biological pathways, among which and not limited to are, transcription, cell cycle, cell death, mitosis and DNA damage response (D'Amours *et al.*, 1999; Barkauskaite *et al.*, 2013). Protein involved in ADP-ribosylation can be divided into three categories: writers or ADP-ribosyl transferases (ARTs), eraser or glycohydrolases and readers or PAR binding protein.

During the last decade, we have witnessed an increase in knowledge of the implications of ADP-ribosylation, with particular emphasis on the DNA Damage Response (DDR). ADP-ribosylation acts in the DDR in two-fold effect, first it acts as a post-translational modification regulating the activity of the modified substrate, and secondly it acts as a scaffold which can be bound by several PAR binding proteins or readers, allowing for the accumulation of these proteins around the DNA damage area (Barkauskaite *et al.*, 2015). ADP-ribosylation of histones, for instance, bring an increasing amount of negative charge inducing dissociation of the DNA from the histone protein, and leads to the recruitment of essential factors for DDR, and cell death by PARtanathos (a PAR-dependent type of cell death) (Kraus 2008; David *et al.*, 2009; Gibson and Kraus 2012; Barkauskaite *et al.*, 2013, Virag *et al.*, 2013; Zaja *et al.*, 2013, Tallis *et al.*, 2014).

1.1.1 Writers of ADP-ribosylation

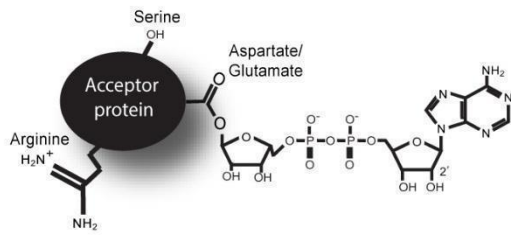
ADP-ribosylation is catalysed by the family of enzymes classified as ADP-ribosyl transferases (ARTs). ARTs use NAD⁺ as a substrate and transfer the ADP-ribose moiety onto its substrate acceptors, such as proteins or nucleic acid, and the reaction releases nicotinamide as a by-product (D'Amours *et al.*, 1999; Munnur and Ahel, 2017). ADP-ribose is linked by an ester bond with its target amino acid substrate, which is mainly aspartate/glutamate in the context of DDR. This process is defined as mono-ADP-ribosylation or MARYlation (**Figure 1.1**). A mono-ADP-ribosylated substrate can be further ADP-ribosylated by addition of ADP-ribose onto the first ADP-ribose through a 2',1''-O-glycosidic or 2'',1'''-O-glycosidic bond, defined as poly-ADP-ribosylation. Poly-ADP-ribose chains can be polymerised to the length of 200 ADP-ribose units, which can be either linear or branched, making it a very large modification with a negative charge double that of DNA (**Figure 1.1**) (Miwa *et al.*, 1979; Juarez-Salinas *et al.*, 1982; De Murcia *et al.*, 1983; Zaja *et al.*, 2013).

The human PARP family (Poly(ADP-ribose) polymerases) includes 17 members, among which PARP1, PARP2 and PARP3 are involved in the DDR. PARP1 as well as PARP4, PARP5a (Tankyrase-1) and PARP5b (Tankyrase-1) can catalysed poly-ADP-ribosylation. PARP3, PARP6, PARP8, PARP10, PARP12, PARP14 and PARP16, on the other hand, can catalysed just mono-ADP-ribosylation or very short poly-ADP-ribosylation (**Figure 1.2**) (Amé *et al.*, 1999; Otto *et al.*, 2005; Loseva *et al.*, 2010). Mono-ADP-ribosylation however, was initially discovered as a bacterial toxin mechanism (Tamura *et al.*, 1982; Hassa and Hottiger, 2008). ARTs, as well as PARPs, are evolutionary conserved from bacteria to eukaryotes (Slade *et al.*, 2011).

PARP1, the most characterized member of the PARP family, is activated upon induction of DNA damage. PARP1 activation will determine an increase of ADP-ribosylation up to 500-fold (D'Amors *et al.*, 1999). Mass spectrometry and *in vitro* experiments suggest that the primary substrates of PARP1 are PARP1 itself (auto-modification), and histones proteins (Gagné *et al.*, 2008; Isabelle *et al.*, 2010). PARP1 can also ADP-ribosylate many other proteins involved in DNA repair, and thus control a layer of regulation by this post-translational modification (Ogata *et al.*, 1981; Barkauskaite *et al.*, 2015).

ADP-ribosylation mediated by the PARP1 can then act as a recruitment scaffold for reader proteins or PAR-binding motifs. So far six different PAR-binding motifs have been identified; PAR binding zinc finger (PBZ), PAR binding motif (PBM), WWE (named after the three conserved residues tryptophan/ tryptophan/glutamate), macrodomain, forkhead-associated (FHA) and Oligonucleotide/oligosaccharide-binding (OB fold).

Mono-ADP-ribosylation



Poly-ADP-ribosylation

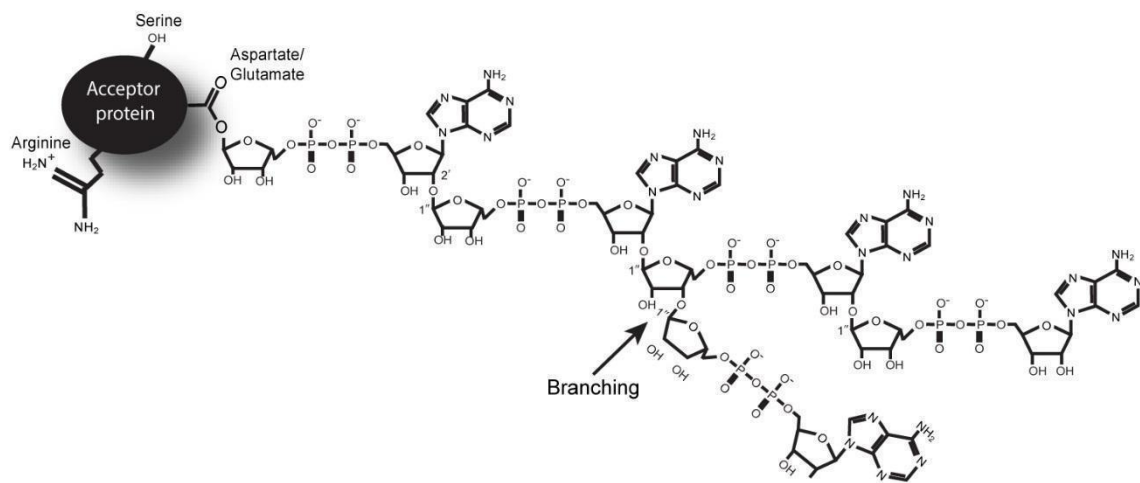


Figure 1.1. - Schematic representation of mono- and poly-ADP-ribosylation. Image adapted from Fontana et al., 2018)

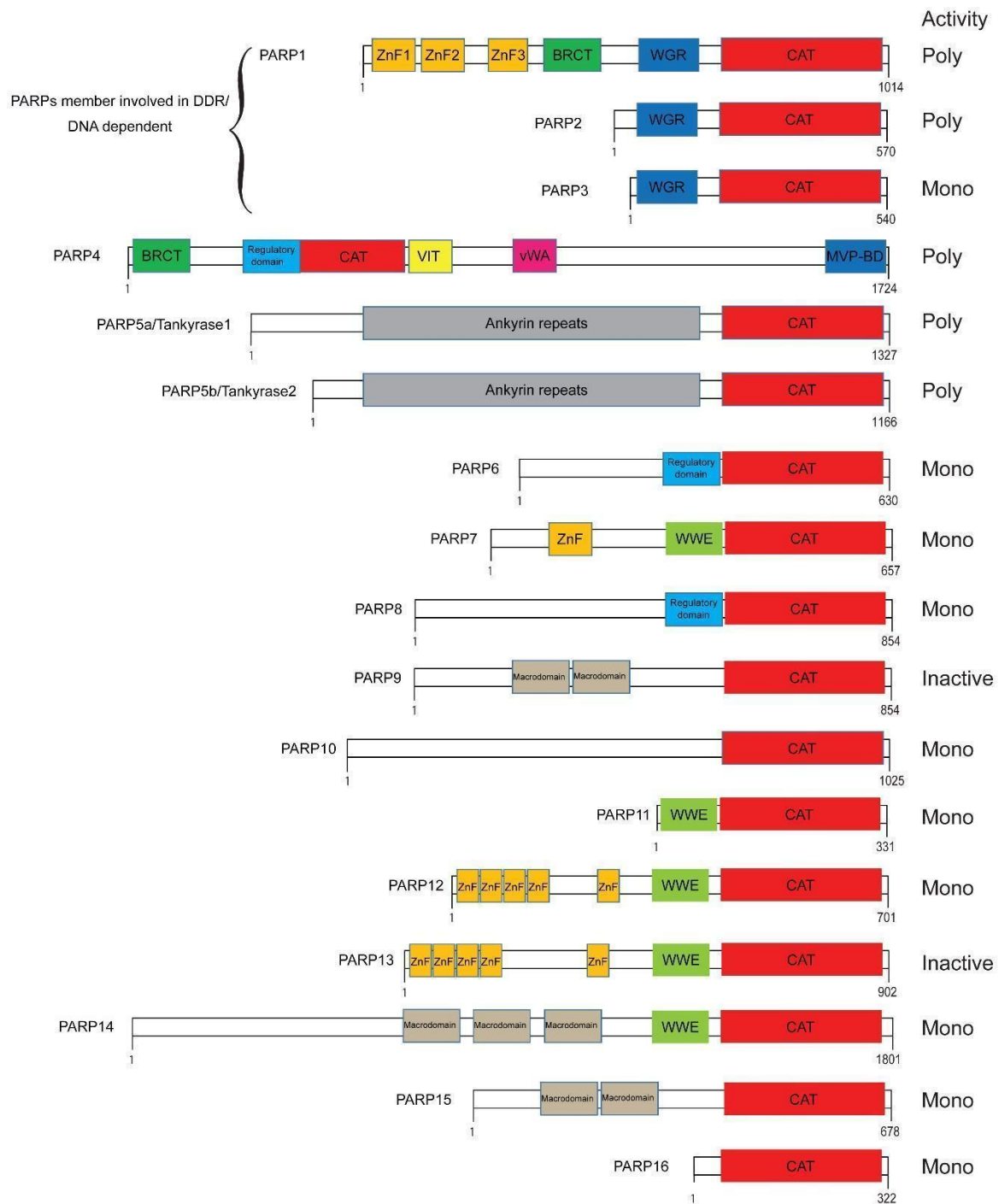


Figure 1.2. - Schematic representation of PARP family members.

1.1.2 Readers of ADP-ribosylation

1.1.2.1 PBZ - PAR binding zinc finger

PBZ is the PAR binding domain that is most relevant for the focus of this thesis as we identified a novel PARP1 regulator possessing such a domain. PBZ domains have been identified in all eukaryotes with the exclusion of yeast (Ahel *et al.*, 2008). Prior to the beginning of this study just three proteins containing PBZ were known, all of them were involved in DNA repair; APLF - aprataxin PNK-like protein, SNM1A - a DNA repair protein sensitive to nitrogen mustard 1A and CHFR - the checkpoint factor containing FHA and RING domain (Ahel *et al.*, 2008). Among these three factors, APLF is the most investigated in terms of PAR binding ability and also from a structural point of view by NMR and X-Ray crystallography (Eustermann *et al.*, 2010; Kim *et al.*, 2017). The APLF-PBZ domain lacks secondary structure with just a short α helix, and binds a zinc ion coordinated by two cysteines and two histidines. The APLF-PBZ domain can bind to a PAR-chain through the adenine and a proximal ribose from ADP-ribose and recognised the distal ADP-ribose (Eustermann *et al.*, 2010; Li *et al.*, 2010). Structural representation of APLF FHA domain is reported in **Figure 1.3A**.

1.1.2.2 PBM - PAR binding motif

The PBM is a very short domain, containing just eight amino acids with the consensus [HKR][X][X][AIQVY][KR][KR][AILV][FILPV] and it is predicted to be present in around 800 proteins (Gagné *et al.*, 2008). As the name suggests, PBM was the first identified among all of the PAR binding domains (Pleschke *et al.*, 2000). Unlike the other domains in this class, PBM can be easily found in many proteins such as AIF - Apoptosis inducing factor, XRCC1 - X-ray cross complementing gene 1, p53, Ku70 and many others (Malanga *et al.*, 1998, Pleschke *et al.*, 2000; Gagné *et al.*, 2008; Wang *et al.*, 2011).

1.1.2.3 WWE domain

WWE has been identified in the RNF146 protein as a PAR directed E3 ubiquitin ligase. RNF146 targets many PAR binding proteins involved in DDR, for instance, XRCC1, and directs them for ubiquitin dependent proteasome degradation (Aravind, 2001; Barkauskaite *et al.*, 2015; Callow *et al.*, 2011; Kang *et al.*, 2011). Structural characterisation by X-ray crystallography of the RNF146's WWE domain, either free or in complex with iso-ADP-ribose (the linkage between two ADP-ribosylation units), shows that WWE interacts specifically with the iso-ADP-ribose (Wang *et al.*, 2012). WWE is organised in half β -barrel composed by 6 β -strands with a side covered by a single α -helix. The WWE recognises the adenine ring of *iso*-ADPr through interaction with the β -barrel and the α -helix, at the same time the phosphate of the *iso*-ADPr interact extensively with the WWE domains, which is positively charged. Structural representation of RNF146 WWE domain is reported in **Figure 1.3B**.

1.1.2.4 Macrodomain

Macrodomains are the most well characterised PAR binding domain, and are relatively large in size (150-190 amino acids), and can be found in both prokaryotes and eukaryotes (**Figure 1.3C**) (Pehrson and Fried, 1992; Karras *et al.*, 2005). Macrodomains have been studied structurally, and the X-ray crystallographic analysis reveals a globular domain with a deep ligand binding pocket which interacts with both the adenine ring and pyrophosphate from the distal ribose (Allen *et al.*, 2003; Timinszky *et al.*, 2009).

Many macrodomains also possess catalytic activity, such as PARG and also MACROD1, MACROD1 and TARG1 (See section 1.1.3 Erasers of ADP-ribosylation). Different macrodomains possess different substrate specificities, for example, PARG's macrodomain cannot efficiently bind/cleave the terminal ADP-ribose. PARG's structural characterisation revealed the presence of a catalytic loop and additional essential non catalytic domain which likely reduce the accessibility to the catalytic site (Slade *et al.*, 2011). In contrast with other macrodomains which do not show such restraint such as MACROD1 and MACROD2 (Chen *et al.*, 2011; Jankevicius *et al.*, 2013).

These domains bind to PAR in different ways, and these subtle differences determine specificity and activity. The PARG macrodomain possesses a loop containing two conserved glutamic-acid residues which allow the interaction with, and catalysis of the PAR chain, leading to the release of free ADP-ribose (Slade *et al.*, 2011). Very differently, TARG1 possesses a conserved lysine in the catalytic site and forms an intermediate with the substrate (terminal glu-ADP-ribose) before the release of ADP-

ribose (Sharifi *et al.*, 2013). MACROD1 and MACROD2 possess an ulterior mechanism which forces the terminal-ADP-ribose to take a conformation optimal for the catalysis (Jankevicius *et al.*, 2013; Rosenthal *et al.*, 2013).

Interestingly, macrodomains can also be found in PARP enzymes, such as PARP14 (Hassa and Hottiger, 2008), PARP9 and also in ALC1 - amplified liver cancer protein 1 (Hakmé *et al.*, 2008; Ahel *et al.*, 2009), however the role of macrodomains in PARP enzymes are not well understood. Macrodomains are also present in viruses such as SARS-CoV and Chikungunya virus (Egloff *et al.*, 2006; Malet *et al.*, 2009).

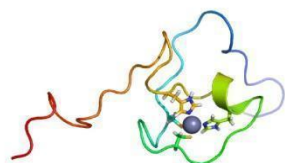
1.1.2.5 Forkhead-associated (FHA)

Forkhead-associated (FHA) domain containing proteins are proteins involved in DDR and are known for their ability to bind phospho-protein (Li *et al.*, 2002). In 2013, Li *et al.*, screened 19 FHA containing proteins for their ability to bind PAR and identified 2 of them, Aprataxin (APTX) (**Figure 1.3D**) and Polynucleotide Kinase 3'-Phosphatase (PNKP) as PAR binding protein. FHA showed high affinity for iso-ADP-ribose. Such interaction determined the recruitment of FHA on the DNA damage site (Li *et al.*, 2013).

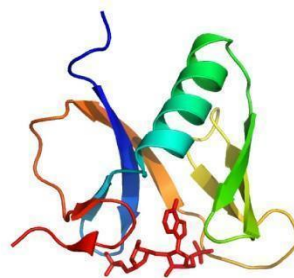
1.1.2.6 Oligonucleotide/oligosaccharide-binding (OB fold)

Oligonucleotide/oligosaccharide-binding (OB) fold was first discovered as a domain with affinity for ssDNA and RNA. In 2014, Zhang *et al.* reported the OB fold of human single strand binding protein (hSSB1) as a PAR binding domain. hSSB1 is recruited on the DNA damage sites, due to its interaction with the iso-ADP ribose, where it is important for early DNA damage repair. Moreover, they screened several other OB fold containing protein and found many of them able to interact with PAR confirming the existence of a new class of PAR binding protein (Zhang *et al.*, 2014). The OB fold shows 5 β -strands organised in a β -barrel covered at one extremity by an α -helice (Murzin, 1993). However, structural evidence of its interaction with ADPr is not yet available. Structural representation of human TTC5 C-terminal OB domain is shown in **Figure 1.3E**.

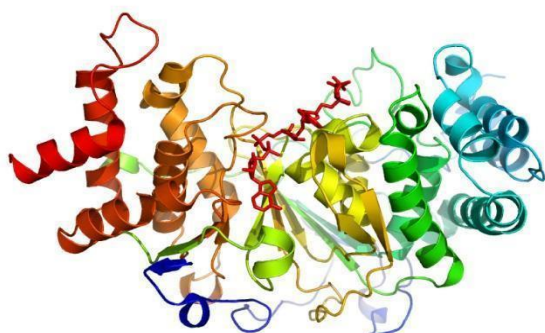
A



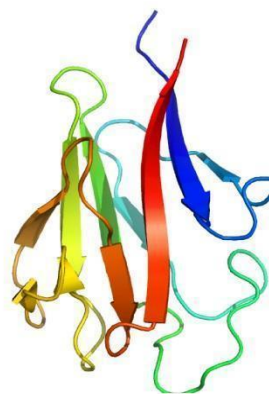
B



C



D



E

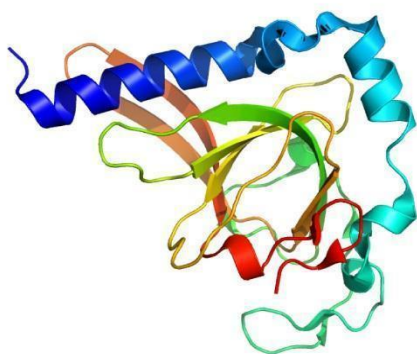


Figure 1.3. **A** - Structure of APLF-PBZ domain (PDB: 2KQB) (Eustermann *et al.*, 2010). **B** - Structure of RNF146 WWE domain in complex with *iso*-ADPr (PDB: 3V3L) (Wang *et al.*, 2012). **C** - Structure of PARG Macrodomain in complex with di-ADPr (PDB: 5A7R) (Lambrecht *et al.*, 2015). **D** - Structure of Aprataxin FHA domain (PDB: 3KT9) (Becherel *et al.*, 2010). **E** - Structure of TTC5 C-terminal OB domain (PDB: 2XVS) (Adams *et al.*, 2012).

1.1.3 Erasers of ADP-ribosylation

PARP activity is tightly regulated and, like most PTMs, poly-ADP-ribosylation is a reversible process (Slade *et al.*, 2011; Li *et al.*, 2014). The continuous activity of PARP enzymes, especially PARP1 given the efficiency of the enzyme and the ability to generate long PAR chains, would cause a critical energy collapse leading to cell death by necrosis (Hassa, 2009). The persistency of PAR-chains inhibits the first regulatory enzyme of the glycolysis hexokinase, which phosphorylates the glucose to glucose-6-phosphate. Hexokinase is a strong PAR binding protein which can be inhibited by the bind of PAR-chain determining energy collapse by blocking glycolysis (Andrabi *et al.*, 2014). Persistent activity of PARP1 would also lead to the formation of free PAR-chains, which exported into the cytoplasm, can cause cell death by PAR-thanatos (Ha and Snyder, 1999; Wang *et al.*, 2009; Tallis *et al.*, 2013). Poly-ADP-ribosylation is therefore rapidly degraded after just a few minutes, making the PTM quite a dynamic regulatory system (Alvarez-Gonzalez and Althaus, 1989). To date, several glycohydrolases or erasers involved in reversing poly-ADP-ribosylation have been identified, among which poly-ADP-ribose glycohydrolase (PARG) is the most thoroughly characterised (Lin *et al.*, 1997; Slade *et al.*, 2012; Barkauskaite *et al.*, 2013). Although the first observation of PAR hydrolysis was made in 1971, PARG was cloned some three decades later, in 1997 (Miwa and Sugimura, 1971; Lin *et al.*, 1997). Therefore PARG characterisation is a relatively recent addition to the PAR field (Hatakeyama *et al.*, 1986). PARG can be found in all eukaryotes, except yeast, and also in bacteria (Slade *et al.*, 2011). PARG possess a macrodomain fold structure (**Figure 1.4 A**), which is the most common among such glycohydrolases. PARG has both exo- and endo-glycohydrolase activity; the exo-glycohydrolase activity is

performed from the termini of the PAR-chain and releases ADP-ribose, while the endo-glycohydrolase activity is performed intra PAR-chain and leads to release of PAR fragments (Miwa and Sugimura, 1971; Ikejima and Gill 1988; Braun *et al.*, 1994; Brochu *et al.*, 1994). Whilst PARG is the most characterised ADP-ribosyl glycohydrolase to date, it can just remove PAR by cleaving the O-glycosidic ribose-ribose bond, and is unable remove the terminal ADP-ribose linked to glutamate/aspartate (Slade *et al.*, 2011; Barkauskaite *et al.*, 2013).

ARH3, a glycohydrolase containing a DraG-like fold domain, has been identified more recently as having similar activity to PARG (Oka *et al.*, 2006). However, its efficiency in removing the O-glycosidic ribose-ribose bond has been shown to be several fold less efficient than PARG (Oka *et al.*, 2006), and similarly to PARG, is unable to cleave the terminal ADP-ribose link to glutamate/aspartate (Oka *et al.*, 2006). ARH3 is also involved in the removal of O-acetyl ADP-ribose which is a product of the sirtuin enzymes (Ono *et al.*, 2006). Prior to the work of this thesis, the cellular role of ARH3 was not well understood, with the data suggesting a role in PARP1 dependent DNA damage response (Mashimo *et al.*, 2016). The X-ray crystal structure of ARH3 revealed a DraG-like fold domain being composed of only α -helices which is architecturally distinct from the macrodomain of PARG (**Figure 1.4 B**) (Mueller-Dieckmann *et al.*, 2006). The catalytic mechanism of ARH3 relies on two glutamate residues which coordinate two Mg^{2+} ions and lead to an acid/basic catalysis (Mueller-Dieckmann *et al.*, 2006). Interestingly, ARH3 can be found not just in the nucleus, where it plays an essential role in DNA damage repair regulation, but also in the mitochondria and cytosol where it is responsible for the degradation of PAR (Niere *et al.*, 2008). Although these studies suggest several possible cellular functions for ARH3 in PAR degradation, chapter 4 of this thesis provides convincing evidence for a new, main function of ARH3.

The removal of terminal ADP-ribose from glutamate/aspartate is carried out by a different glycohydrolase, Terminal-ADP-ribose glycohydrolase (TARG1). TARG1 is a macrodomain containing protein which is able to remove the terminal ADP-ribose linked to glutamate/aspartate, which in turn releases the PAR-chain, but it is unable to cleave O-glycosidic ribose-ribose bond between the ADP-ribose units (Rosenthal *et al.*, 2013; Sharifi *et al.*, 2013). Released PAR chains can then be degraded to individual ADP-ribose units by PARG or ARH3 activity. Two more glycohydrolases, MACRO domain-containing protein 1 (MACROD1) and MACRO domain-containing protein 2 (MACROD2), have also been found to be able to remove the terminal ADP-ribose linked to glutamate/aspartate, and are unable to cleave the O-glycosidic ribose-ribose bond (Wielckens *et al.*, 1982; Jankevicius *et al.*, 2013; Rosenthal *et al.*, 2013; Barkauskaite *et al.*, 2013). To add complexity, ARH1, a glycohydrolase containing a DraG-like fold domain with high similarity with ARH3, has been demonstrated to possess terminal ADP-ribose linked to arginine cleavage activity (Moss *et al.*, 1992). Arginine ADP-ribosylation is, however, a post-translational modification, mediated by ART2.2, involved in toxic response. ART2.2 is a murine enzyme expressed only in mature T cell and activates upon proteolytic cleavage followed by T cell activation. Activated ART2.2 acts as an inter-cellular regulator through Arginine ADP-ribosylation (Kahl *et al.*, 2000).

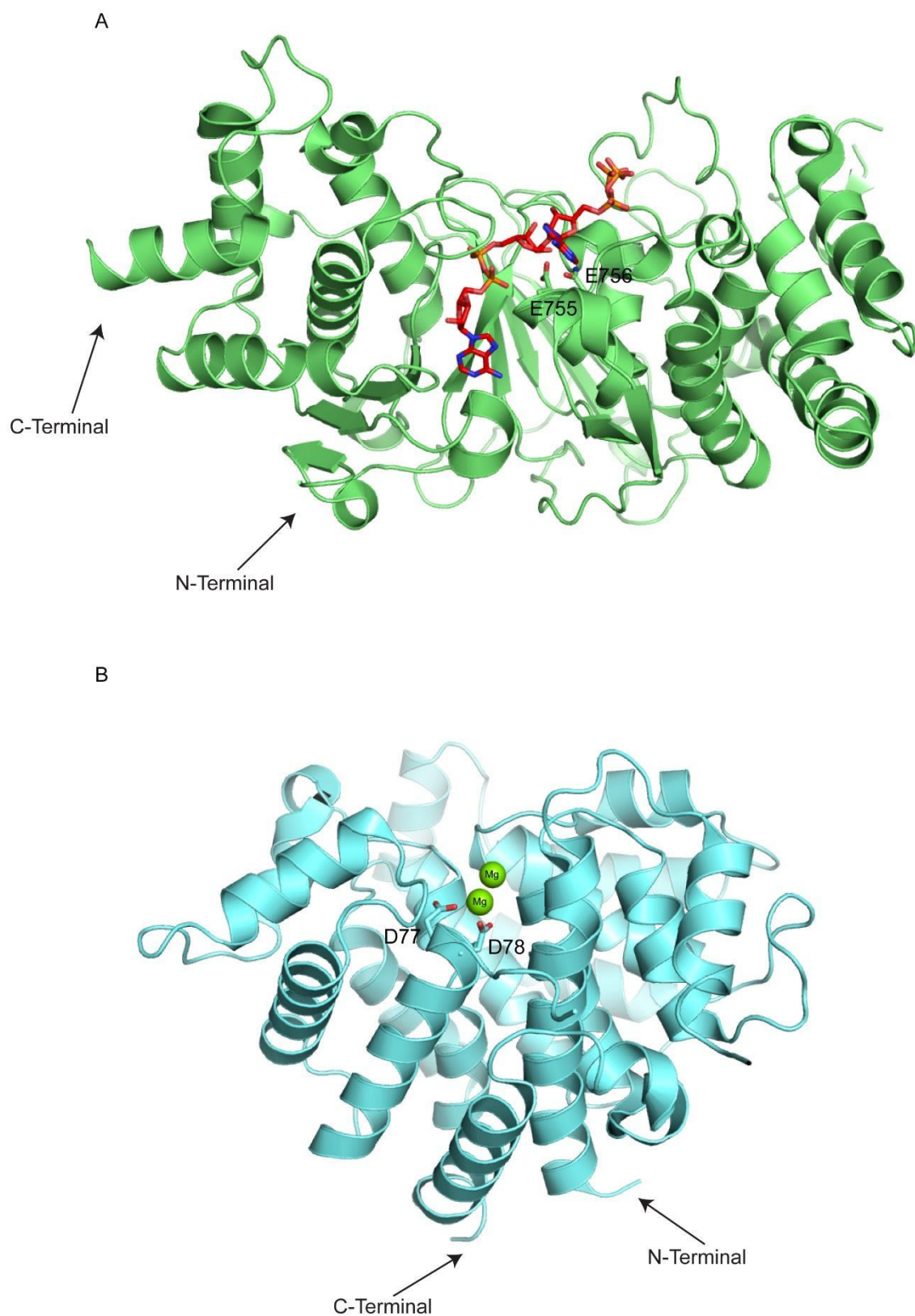


Figure 1.4. **A** - Structure of PARG Macrodomain in complex with di-ADPr, E756 is mutated in N (PDB: 5A7R) (Lambrecht *et al.*, 2015). **B** - Structure of ARH3 in with 2 magnesium ions (PDB: 2FP0) (Mueller-Dieckmann *et al.*, 2006).

1.2 DNA damage repair

Cells have evolved several different mechanisms of DNA repair which can be roughly separated into two classes, excision mechanism of DNA repair and repair of nick or broken DNA. The excision mechanism includes: base excision repair (BER), nucleotide excision repair (NER) and mismatch repair (MMR).

The BER mechanism functions by excising part of the DNA strand, including the damaged region, followed by new DNA synthesis by DNA polymerases (Lindahl, 1974). The initial step of recognition of backbone distortions is a low affinity process which follows a base flipping for high affinity recognition depends by the activity of a DNA glycosylase. The flipped bases will be oriented in the catalytic site of the glycosylase allowing glycosylation of the base forming an abasic (AP) site in the DNA. AP can also occur by direct damage. The endonuclease APE1 will then interact with the AP and create a nick and the base will be removed by the DNA polymerase β which contemporary fills the gap in the DNA together with the PCNA, using the undamaged strand as a template, and finally DNA ligase ligates the DNA (**Figure 1.5**) (Matsumoto *et al.*, 1995; Frosna *et al.*, 1996; Klungland *et al.*, 1997; Beard *et al.*, 2000).

NER is involved in repair of bigger distorted DNA region such as intrastrand cross link, for instance, pyrimidine dimers induced by UV. Usually a strand of 22-28 nucleotide is removed before the DNA polymerase can replicate the excised DNA (Marteijn *et al.*, 2014). The damage is recognised through its backbone by an excision repair complex. Initial recognition is mediated by the recognition factors RPA, XPA, and XPC (Mu *et al.*, 1997; Wakasugi *et al.*, 1998; Wakasugi *et al.*, 1999). Following the distorted DNA is removed by double nuclease digestion (Wood, 1997;

Wood, 1999; Petit et al., 1999). The complementary DNA is used as a template for DNA pol and finally, the DNA ligase ligates the newly synthesised DNA (**Figure 1.6**).

Different from the two previous mechanisms, MMR acts on mismatch repair which usually arises during DNA replication. Activation of MMR depends on the recognition by MutS α or MutS β to mismatches in DNA. Following DNA interaction MutS α will bind to ATP determining a conformational change in a clamp form state of MutS α which can move along the DNA. This active form of MutS α will recruit MutL α protein (Palombo *et al.*, 1995; Hombauer *et al.*, 2011). Following MutL α recruitment, the PCNA present in the replication complex will activate MutL α which will make a double DNA incision (Kadyrov *et al.*, 2006). Finally, MutS α activates the exonuclease EXO1 which will excise the incorrect DNA (Szankasi *et al.*, 1995; Tran *et al.*, 2001; Genschel *et al.*, 2002). The single-stranded DNA will be bound by RPA and finally, DNA pol δ or ϵ will replicate the excised DNA and the DNA ligase will ligate the new DNA (**Figure 1.7**) (Longley *et al.*, 1997; Zhang *et al.*, 2005).

The second class includes two different mechanisms: single strand break (SSB) and double strand break (DSB), both employ different mechanisms and do not utilise any excision. SSB depends on PARPs and on the activity of X-ray repair cross-complementing protein-1 (XRCC1) (Caldecott, 2003). PARP1 senses the DNA damage site, and its interaction with the DNA through the zinc finger exposes its catalytic site starting auto-modification. The formed PAR chain act in two fold; firstly as scaffold for recruitment of other proteins such as XRCC1, which can also be modified by ADP-ribosylation, and secondly the negative charge of PAR determines a release of PARP1 from the DNA leaving access to other factor for DNA repair such as polymerases, DNA ligase, APE1 and others (**Figure 1.8**) (Caldecott, 2003; Malanga *et al.*, 2005).

The most frequent occurrence of DSBs is usually during replication of a region of DNA with a SSB. There are two main mechanisms of DSB; homologous recombination (HR) and non-homologous end joining (NHEJ). These two mechanisms are usually dependent on the phase of the cell cycle.

HR required a sister chromatid as repair template, so it is only possible in S/G2 (Kass and Jasin, 2010). HR relies on DNA resection by the MRN complex (Mre11-Rad50-Nbs1) which initiates the resection from the both 5' ends of both the end of the DNA, leaving 3' overhangs single-stranded DNA (Sartori *et al.*, 2007). The 3' single-stranded DNA (ssDNA) are then bound by RPA protein, allowing the removal of secondary structures, followed replacement of RPA by RAD51 and other mediators such as RAD52 and BRCA2. This ssDNA-protein complex executes the DNA sequence homology finding. Finally, RAD51 will determine interaction with the complementary DNA and the DNA pol η will replicate the missing DNA fragment followed by ligation by DNA ligase I (**Figure 1.9**) (McIlwraith *et al.*, 2005; Ciccia and Elledge, 2010; Krejci *et al.*, 2012).

NHEJ is initiated by the recognition of the end of the broken DNA by the heterodimer Ku70/80. The heterodimer binds to the DNA in a ring conformation (Walker *et al.*, 2001). After the interaction, the heterodimer recruits DNA-PK holoenzyme, which induces a translocation of the heterodimer along the DNA and leaving the holoenzyme interacting with the DNA termini (Yoo *et al.*, 1999). In case DNA termini present an overhang, the missing part can be elongated by DNA pol μ or λ or alternatively the overhang can be excised by the nuclease Artemis (Jeggo *et al.*, 2002; Lieber *et al.*, 2008). Finally, the DNA ligase IV-XRCC4 complex ligates the extremity of DNA (**Figure 1.10**) (Ahnesorg *et al.*, 2006; Kass and Jasin, 2010). However, NHEJ is unable to repair replication-associated one-ended DSBs. More recently, an alternative NHEJ

which relies on PARP1 has been identified (Mansour *et al.*, 2010). NHEJ and HR are very different mechanisms, with the former leading to a deletion of the damaged DNA which eventually lead to the loss of genetic information (Kass and Jasin, 2010).

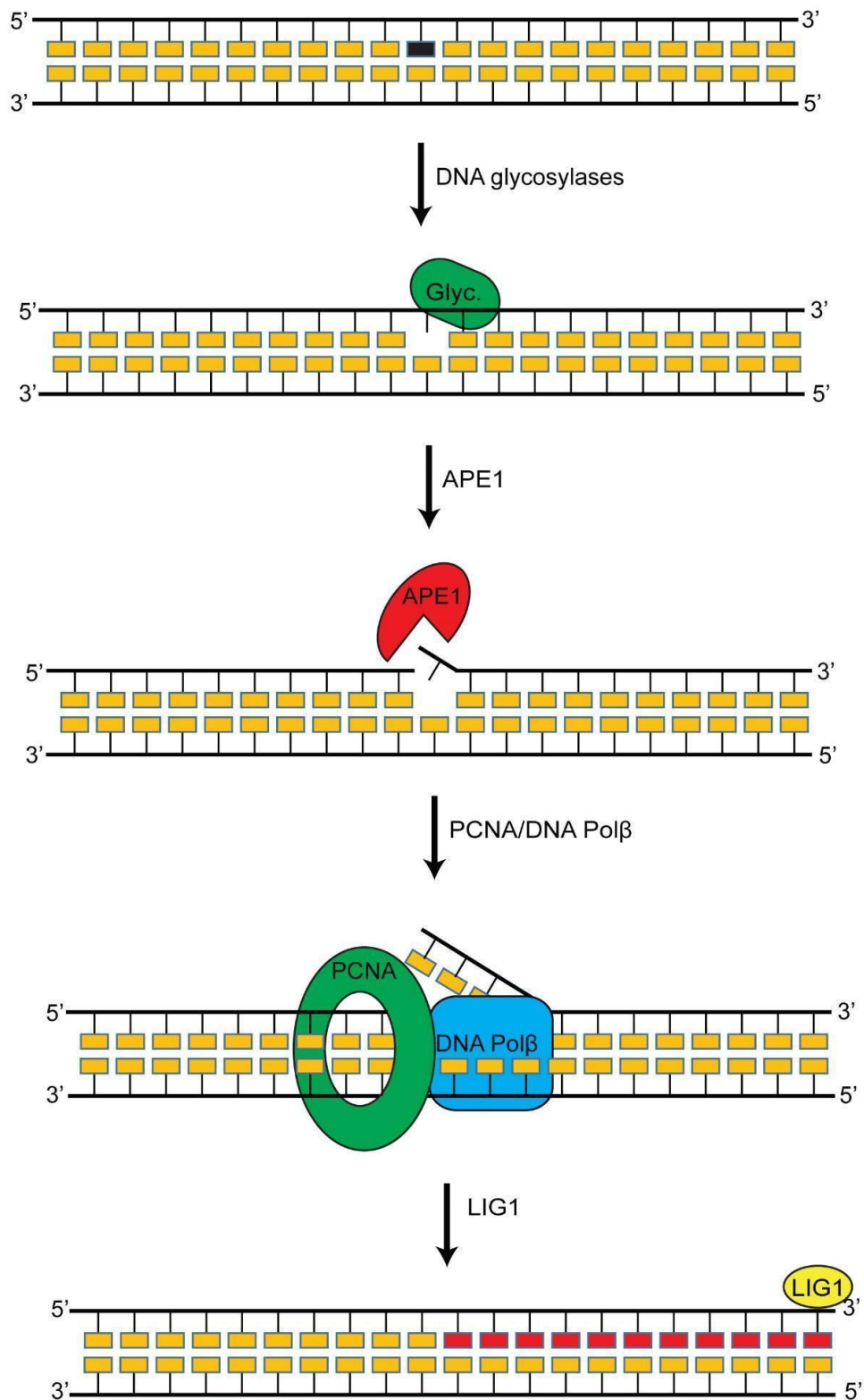


Figure 1.5. Schematic BER representation. DNA glycosylase forms an abasic (AP) site in the DNA. The P will be recognized by the endonuclease APE1 which creates a nick in the DNA. DNA polymerase β and PCNA will remove the DNA from the nick and at the same time fill the gap, finally, DNA ligase ligates the DNA.

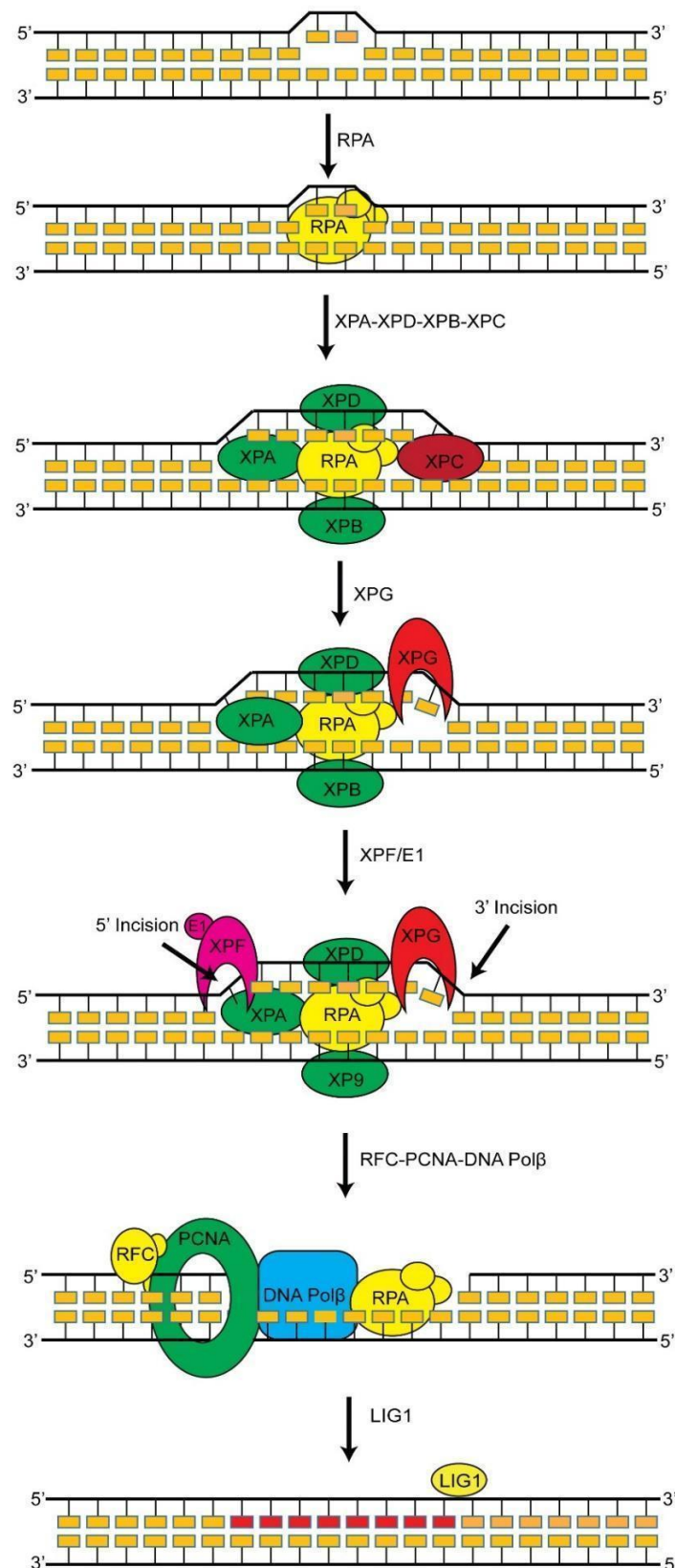


Figure 1.6. Schematic NER representation. The excision repair recognizes the DNA through its backbone. The distorted DNA is removed by double nuclease digestion (XPF and XPG). DNA polymerase and PCNA use undamaged strand as a template, finally the DNA ligase ligates the newly synthesised DNA.

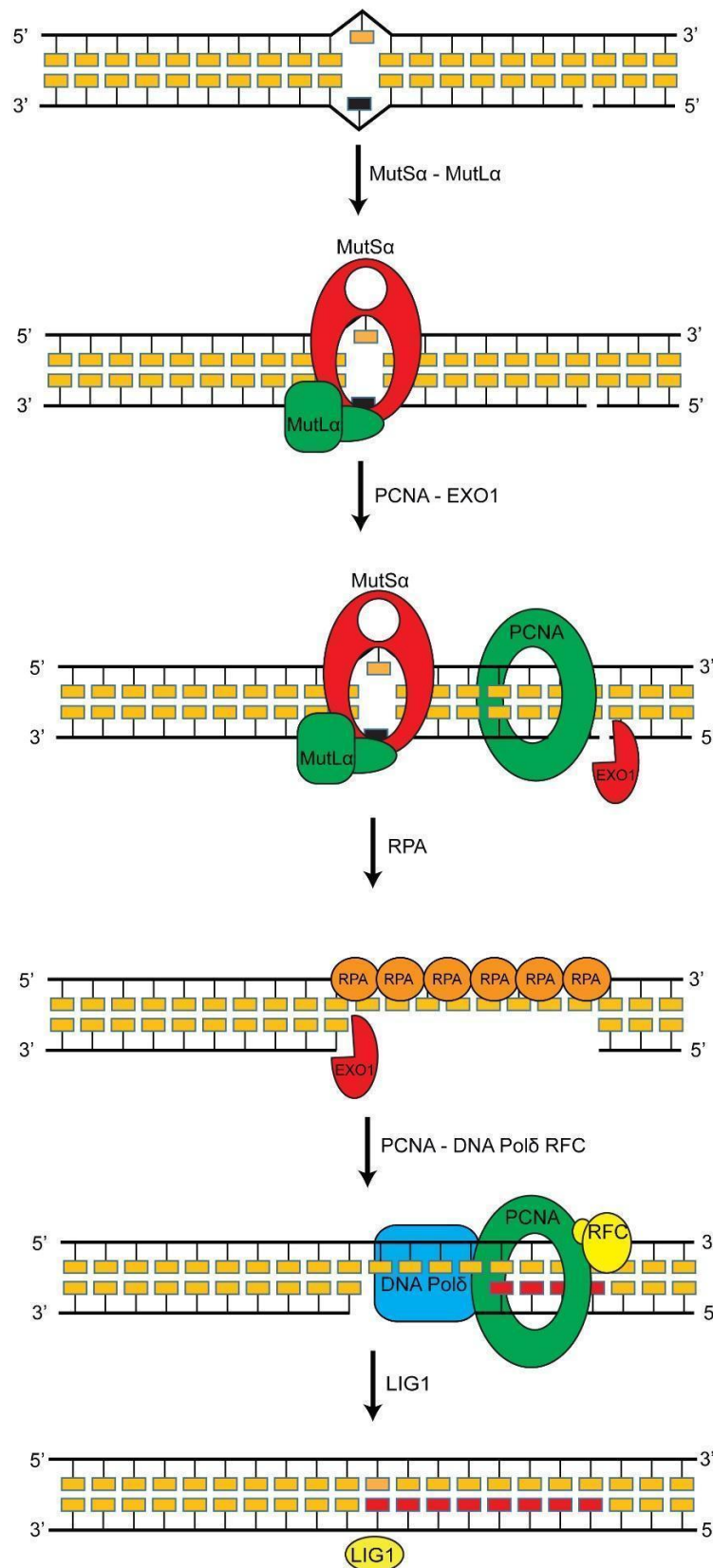


Figure 1.7. Schematic MMR representation. MutSα or MutSβ recognise the mismatches. Activation of MutSα will recruit MutLα protein. MutLα is then activated by the PCNA, which is present in the replication complex, determining a double DNA incision by EXO1. RPA will bind the single stranded DNA and DNA polymerase δ or ε will fill the gapped DNA. Finally, the DNA ligase will ligate the new DNA.

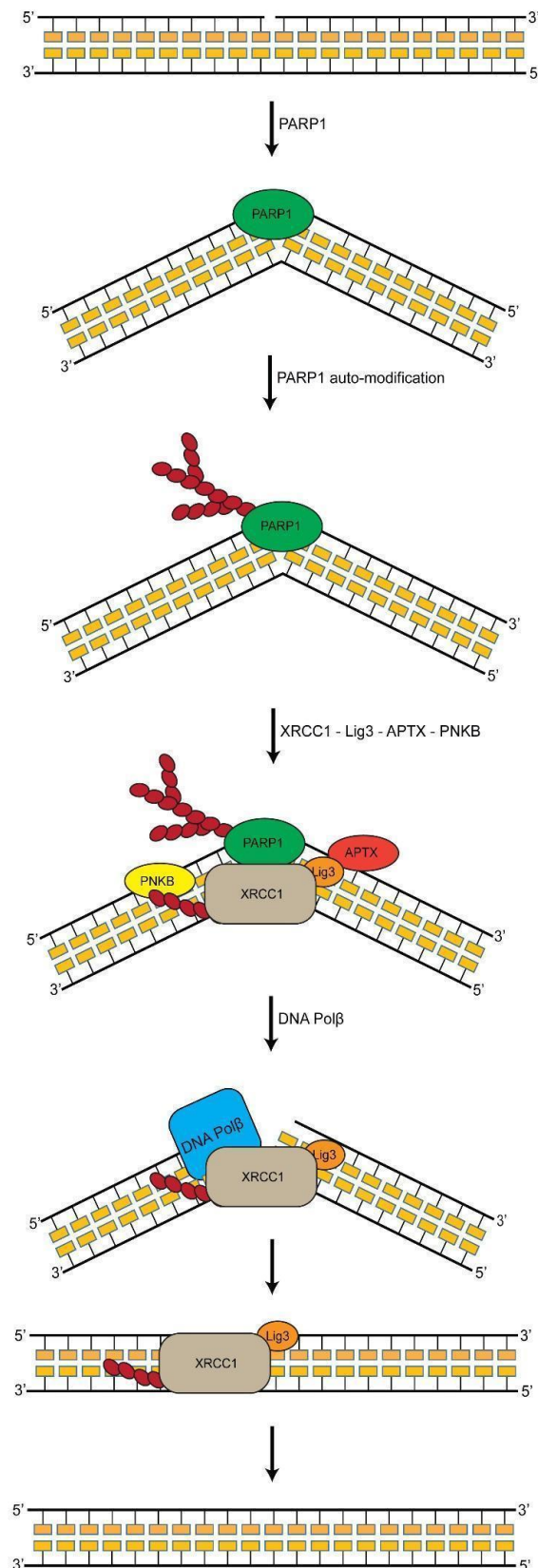


Figure 1.8. Schematic SSB representation. PARP1 interacts with the damaged DNA through its zinc fingers. Activated PARP1 begins to auto-modify itself and trans-modify the recruited factor such as XRCC1. PARP1 auto-modification determines a release of PARP1 from the DNA leaving access to other factors for DNA repair such as polymerases, DNA ligase and APE1.

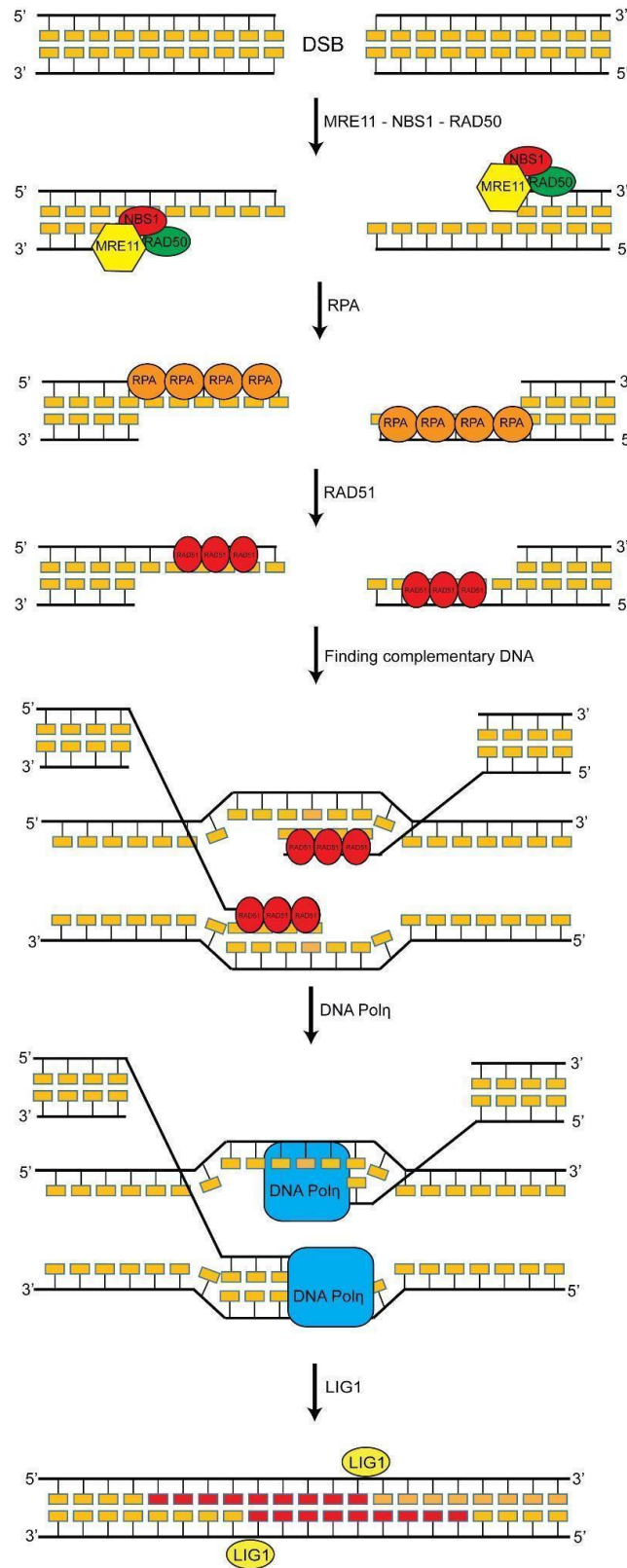


Figure 1.9. Schematic HR representation. MRN complex initiates the resection from the 5' ends of the broken DNA leaving 3' overhangs. RPA binds the 3' single stranded DNA, and RAD51 will replace RPA. This ssDNA-protein complex executes the DNA sequence homology finding. DNA polymerase η will replicate the missing DNA fragment. Finally, DNA ligase I will ligate the newly synthesized DNA.

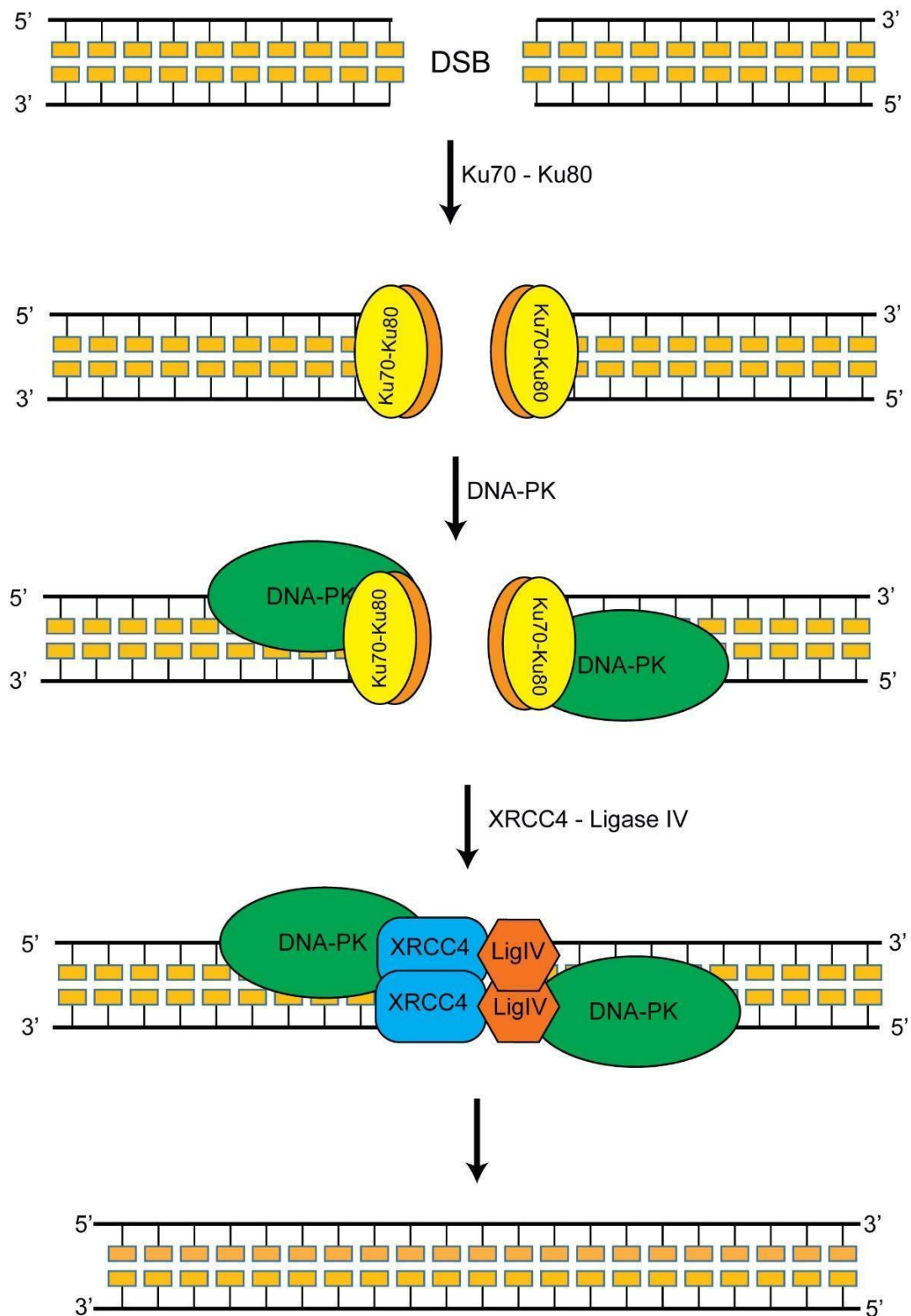


Figure 1.10. Schematic NHEJ representation. Ku70/80 heterodimer binds the end of the blunt DNA and recruits DNA-PK holoenzyme inducing translocation of the heterodimer. Finally, the DNA ligase IV-XRCC4 complex is recruited and ligates the extremity of DNA.

1.3 PAR dependent DNA damage repair

ADP-ribosylation is involved in many biological processes, however, its role in DNA repair is the most characterised and the principal enzyme responsible for this PTM in DDR is PARP1. PARP1 contains six domains divided into three groups; at the N-terminus is the DNA binding domain which consists of three independent zinc fingers named ZnF1, ZnF2 and ZnF3, the central region contains the auto-modification domain called BRCT and the WGR domain, and finally the C-terminus region contains the catalytic domain (Kaufmann *et al.*, 1993; Langelier *et al.*, 2012). PARP1 is activated upon interaction of the zinc fingers with broken DNA (Schreiber *et al.*, 2006; Langelier *et al.*, 2012). The three zinc fingers have different specificities for damaged DNA; the first two zinc fingers seem to be responsible for interaction with single stranded DNA breaks, while the third zinc finger interacts with double strand DNA breaks (Langelier *et al.*, 2012). Structural characterisation of PARP1 has shown that PARP1 activity can be carried out even in the absence of both ZnF2 and BRCT domains which are then considered non-essential for catalysis (Langelier *et al.*, 2012). However, the interaction of the zinc fingers with a DNA termini is absolutely necessary for PARP1 activation (Clark *et al.*, 2012). The BRCT domain, or auto-modification domain, is the main region to be modified by PARP1, although the catalytic domain as well the zinc fingers domains has also been found to be a minor region for auto-modification (Gagné *et al.*, 2015). The catalytic domain is divided into two sub-domains; the proper catalytic domain and the HD domain, or auto-inhibitory domain (Dawicki-McKenna *et al.*, 2015). Interaction with DNA induces a conformational change to the catalytic domain allowing the HD domain to expose the NAD⁺ binding site (Ruf *et al.*, 1998; Amé *et al.*, 2004; Langelier *et al.*, 2012; Eustermann *et al.*,

2015). Upon PARP1 activation, PARP1 begins to auto-modify itself, mainly on the BRCT domain, and modify its other substrates such as; histones, polymerases, topoisomerases, transcriptional factors and other proteins involved in DNA repair (Adami-etz, 1987; Hassa and Hottiger 2008; Kraus, 2008). Unpublished data from the Ahel lab suggests that PARP1 auto-modification in the BRCT domain does not affect PARP1 function and catalytic activity. The majority of DNA damage dependent PARP1 ADP-ribosylation studies have found that the main ADP-ribose acceptor residues are glutamate/aspartate (Ogata *et al.*, 1980; Poirier *et al.*, 1982). However, alternative studies have also suggested lysine as a possible acceptor for ADP-ribose (Messner *et al.*, 2010).

PARP2 and PARP3 are also involved in DNA damage repair but their functions are less well understood. PARP2, as well as PARP1, can catalyse PAR-chain formation, however, just around 10% of the total cellular PAR is produced by PARP2 (Shieh *et al.*, 1998; Schreiber *et al.*, 2002). PARP2 consist of just a WGR domain and a catalytic domain, lacking all of the additional N-terminal regions found in PARP1. PARP2 can interact with single nucleotide gaps (Amé *et al.*, 1999). PARP3, on the other hand, can catalyse only mono-ADP-ribosylation, not chains of PAR. Differently, from the other PARP1 and PARP2, PARP3 seems to be involved in NHEJ and HR of double strand breaks (Wang *et al.*, 2006; Haince *et al.*, 2008).

1.4 Therapeutic roles of PARPs inhibitor

Since PARPs play a role in many biological processes, the modulation of their activity by drugs has been a recurrent topic during the last two decades (Terada *et al.*, 1979; Zaremba *et al.*, 2007; Lord *et al.*, 2017). 3-aminobenzamide was the first identified PARP1 inhibitor drug, tested both *in vivo* and *in vitro*. Later, more specific drugs were developed, based on 3-aminobenzamide (Ben-Hur *et al.*, 1985; Rouleau *et al.*, 2010; Papeo *et al.*, 2013). Due to the similarity of the catalytic domain of all the human PARPs, the identification of drugs which bind specific PARPs without interfering the function of the others has been a considerable challenge. The best characterised PARP1 inhibitor, Olaparib, has recently entered phase III of clinical studies for cancer therapy, despite its ability to modulate also PARP3 activity (De Vos *et al.*, 2012). Modulation of PARP1 by Olaparib seems to increase the efficiency of radiotherapy and chemotherapy for cancer treatment (Ledermann *et al.*, 2012). Olaparib interaction with PARP1 *in vivo* shows an accumulation of DNA damage and increased sensitivity to radiation and alkylating agents, such as platinum analogs which are currently used for chemotherapy (Rouleau *et al.*, 2010; Papeo *et al.*, 2013; Tallis *et al.*, 2013). Moreover, PARP1 inhibitors have been shown to be significantly active for treatment of breast and ovarian cancer without the need of chemotherapy or radiotherapy (Rouleau *et al.*, 2010; De Vos *et al.*, 2012; Papeo *et al.*, 2013). In this case, PARP1 inhibitors rely on a synthetic lethality mechanism which is the process of making a cell inefficient in both homologous recombination and other DNA repair pathways (Rouleau *et al.*, 2010; De Vos *et al.*, 2012; Papeo *et al.*, 2013). For instance, modulation of PARP1 activity leads to a failure in single strand break repair which in turn induces a replication fork collapse, and subsequently DNA double strand break (Sugimura *et al.*, 2008). Tumors lacking BRCA1 and BRCA2 will be

then unable to repair the damage due to the inefficiency of the homologous recombination (De Vos *et al.*, 2012). However, cells which do not lack BRCA1 and BRCA2 can still repair the damage. Due to this difference, PARP1 inhibitors efficacy seems to be selective for cancer cells without having a major effect on non-cancer cells (De Vos *et al.*, 2012). For such reasons, Olaparib is currently used as a drug for tumors lacking BRCA1 and BRCA2. However, synthetic lethality is just one of the possible mechanisms affected by PARPs inhibitor. The anti-tumor effect of PARP1 inhibitors does not rely only on blocking PARP catalytic activity. Indeed depletion of PARP1 has a less toxic effect than PARP1 inhibitor (Murai *et al.*, 2012; Murai *et al.*, 2014). This suggests that PARP1 itself is a mediator of toxicity in the presence of its inhibitor. In this case, PARP1 toxicity probably relies on a trapping mechanism. The inability of PARP1 to auto-modify itself will determine an inability of release from the damaged DNA creating a stable “trapped” PARP1-DNA complex. Such a complex will interfere with the repair of the SSB, and accumulation of this damage will lead to DSBs and cell death (Lindahl, 1993; Helleday, 2011; Kedar *et al.*, 2012).

1.5 Identification of ADP-ribosylation sites

The identification of the sites modified by ADP-ribosylation has been a major challenge for many years, due to technical limitations. The development of techniques which allow for enrichment of mono-ADP-ribosylated peptides, and methods for labelling of ADP-ribosylation sites, prior to mass spectrometry have hugely improved our knowledge of ADP-ribosylation sites. Martello et al., 2016 developed a tool for the identification of endogenous ADP-ribosylation sites from cells/tissue. They isolated proteins from cells/tissue and digested it into peptides using LysC and trypsin. Following digestion, the peptides were treated with PARG, removing all the poly-ADP-ribosylation and enriching mono-ADP-ribosylation. The enriched peptides were then analysed by mass spectrometry for the identification of ADP-ribosylation acceptor residues. Daniels et al., 2014 developed a different tool based on a phosphoproteomic approach and the use of snake venom phosphodiesterase, a pyrophosphatase that cleaves ADPr subunits to phospho(ribose) and 5'-AMP. Purified proteins were treated with the phosphodiesterase, producing a single phospho(ribose) group at the site of modification due to the similarity of phospho(ribosyl) and phosphate groups. Making use of this chemistry, they decided to test existing phosphoproteomic techniques for the enrichment of phospho(ribosyl)ated peptides and finally identified the sites of ADP-ribosylation by mass spectrometry.

1.6 Scope of the thesis

The overall findings of my thesis highlight the important discovery of the mechanism of serine ADP-ribosylation (Ser-ADP-ribosylation) and identified serine as the main residues of ADP-ribosylation in the DDR. The identification of the residues substrate of ADP-ribosylation has been a field of study for many years and is still very active. The discovery of the specific amino acid residue modified by PARP1 will have a big impact in the ADP-ribosylation field since such knowledge is essential to fully understand the mechanism of ADP-ribosylation. Moreover, the further identification of the enzyme responsible for reversing such modification is essential to fully understand ADP-ribosylation regulation and signaling. These findings will also have an impact in the epigenetic field, as histone tails are one of the main substrates of Ser-ADP-ribosylation. A full understanding of Ser-ADP-ribosylation will allow researchers to study the cross-talk of different PTM in the epigenetics.

In the third chapter of my thesis, I will first describe the identification, biochemical and structural characterisation of the Histone Poly(ADP-ribose)ylation Factor 1 (HPF1), a previously unreported protein, followed by the characterisation of the newly reported post-translational modification Ser-ADP-ribosylation, ADP-ribosylation on serine residues, dependent on the PARP1/HPF1 or PARP2/HPF1 complexes. Following on in chapter four, I will show, through a biochemical screening and characterisation, that the newly reported Ser-ADP-ribosylation is a reversible post-translational modification and show that ARH3 is the only human enzyme responsible for Ser-ADP-ribosylation cleavage. Finally, in chapter five, I will report Ser-ADP-ribosylation characterisation in

Drosophila melanogaster; the identification, structural and biochemical characterisation of *Drosophila* Parg as the only *Drosophila* enzyme known so far capable of removal of Ser-ADP-ribosylation.

Chapter 2. Materials and methods

2.1 Standard PCR for cloning

PCR was performed using Phusion® High-Fidelity DNA Polymerase from Thermofisher as suggested from the manufacturer (Cat. No. F-530S) (**Table 2.1**).

PCR Mix				PCR cycles	Temperature	Time
10X Phusion buffer	5 µL		1	Initial Denaturation	98 °C	30s
DMSO	2.5 µL		2	Denaturation	98 °C	5s
dNTP mix	1 µL		3	Annealing	58-60 °C	15s
Primer F (10uM)	2.5 µL		4	Extension	72 °C	30s/kb
Primer R (10uM)	2.5 µL		5	Final Denaturation	72 °C	10m
Template DNA	1 µL (2-5 ng)		6	Hold	4 °C	
Phusion enzyme	0.5 µL					
mQ H2O	35 µL					
				Cycles from 2 to 5 repeated 30 times		

Table 2.1. PCR cloning condition

2.2 Standard mutagenesis

Mutagenesis was performed using QuikChange Lightning Site-Directed Mutagenesis Kit from Agilent as suggested from the manufacturer (Car. No. 210519) (**Table 2.2**).

PCR Mutagenesis Mix				PCR cycles	Temperature	Time
10X Reaction buffer	2 µL		1	Initial Denaturation	95 °C	30s
Quick solution	0.6 µL		2	Denaturation	95 °C	30s
dNTP mix	0.4 µL		3	Annealing	54-58 °C	1m
Primer F (10uM)	0.4 µL		4	Extension	68 °C	1m/kb
Primer R (10uM)	0.4 µL			Hold	4 °C	
Template DNA	1 µL (20 ng)					
Polymerase enzyme	0.4 µL					
mQ H2O	14.8 µL					
				Cycles from 2 to 4 repeated 18 times		

Table 2.2. Mutagenesis condition.

2.3 Cloning, expression and purification of HsHPF1, DmHpf1 and NvHpf1

Cloning

HsHPF1, DmHpf1 and NvHpf1 were cloned into pET28 vector, from synthesised oligo as templates. HsHPF1 and DmHpf1 were cloned into pET28 vector, with an N-Term 6His tag or a C-Term 6His tag and NvHpf1 was cloned into pET28 vector, with a C-Term 6His tag. HsHPF1, DmHpf1 and NvHpf1 were amplified by PCR, using a Phusion polymerase in standard PCR condition (**Table 2.1**), from a synthesised DNA using the primers in table 2.3.

PCR product were purified from the PCR mix and digested 60 minutes at 37°C, in a total volume of 20 µL using 0,5 µL of NEB NcoI restriction enzyme (Cat. No. R0193S) 0,5 µL of NEB XhoI restriction enzyme (Cat. No. R0146S) in NEB 3.1 buffer (Cat. No. B7203S). Digested DNA product and plasmid were loaded in a 1% agarose gel, and gel purified upon separation. DNA concentrations were assessed with nanodrop and 20 ng of digested plasmid were mixed with an excess of 3 or 5 fold of digested insert in water, in a total volume of 5 µL. Mixed DNA were ligate using one volume of Takara mix ligase (Cat. No. 6023) for 1h at 16 °C. 3 µL of ligation mix was used for transformation of 30 µL of *E.coli* DH5α competent cells. Correct insertion was confirmed by sequencing.

Point mutants were obtained by site directed mutagenesis using the Quik-Change Lightning Site-Directed Mutagenesis Kit in standard mutagenesis condition (**Table 2.2**) using the primers in table 2.3.

Primer	Tag	Restriction enzyme	Sequence
HsHPF1 F	N-Term-6His	NcoI	5'-CGTGCGCCATGGGCCATCATCATCATCACGGCGGCATGGTCGGCGGTGGC-3'
HsHPF1 R		XhoI	5'-CCGCTCGAGTCATGCAGCAAGTTGG-3'
HsHPF1 F		NcoI	5'-CGTGCGCCATGGGCATGGTCGGCGGTGGCGGG-3'
HsHPF1 R	C-Term-6His	XhoI	5'-CCGCTCGAGTTAGTGGTGGTGGTGGTGGTGGTCTTTGCAGCAAGTTGGTC-3'
DmHPf1 F	N-Term-6His	NcoI	5'-CGTGCGCCATGGGCCATCATCATCATCATCACGGCGGCATGCCAAAAGAAGATTGC-3'
DmHPf1 R		XhoI	5'-CCGCTCGAGCTACTTAAGTACATCAAAAATG-3'
DmHPf1 F		NcoI	5'-CGTGCGCCATGGGCATGCCAAAAGAAGATTGC-3'
DmHPf1 R	C-Term-6His	XhoI	5'-CCGCTCGAGTTAGTGGTGGTGGTGGTGGTGGTGGTCTTCTTAAGTACATCAAAAATG-3'
NvHPf1 F		NcoI	5'-CGTGCGCCATGGGCATGGCAGAAGTCAGC-3'
NvHPf1 R	C-Term-6His	XhoI	5'-CCGCTCGAGTTAGTGGTGGTGGTGGTGGTGGTGGTCTTCGAGGCTTCTGACAG-3'
Mutagenesis Primers			
HsHPF1 Y238A-R239A F			5'-GATAAAAATGATGTTGGGGCAGCAGAGCTCCCTGAAACAG-3'
HsHPF1 Y238A-R239A R			5'-CTGTTTCAGGGAGCTCTGCTGCCCCAACATCATTTTTATC-3'
HsHPF1 Δ242-346 F			5'-GAGAGCTCCCTTGACTCGAGCACCACCACCAC-3'
HsHPF1 Δ242-346 R			5'-GTGCTCGAGTCAAGGGAGCTCTCGGTACCCAACATC-3'

Table 2.3. Primers for cloning HPF1s in pET28 vector and mutagenesis primers.

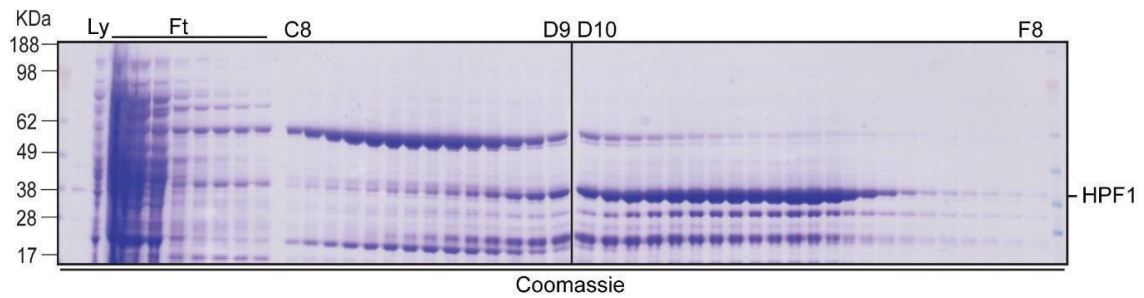
Expression

Rosetta DE3 were transformed with the appropriate plasmid and plated on LB agar in the presence of 30 µg/mL of kanamycin and 30 µg/mL of chloramphenicol. A single colony was picked and grown overnight in LB medium supplemented with the same antibiotics. The pre-culture was added, at 1% final concentration, to LB medium supplemented with antibiotics and cells were grown at 37°C until OD₆₀₀ = 0.6. The culture was then chilled on ice for 1 hour to stop the growth and 1 mM final concentration IPTG was added to induce protein expression. Cells were incubated at 18 °C overnight and collected in the morning by centrifugation at 4000 g for 10 minutes and pellet stored at -20 °C until use.

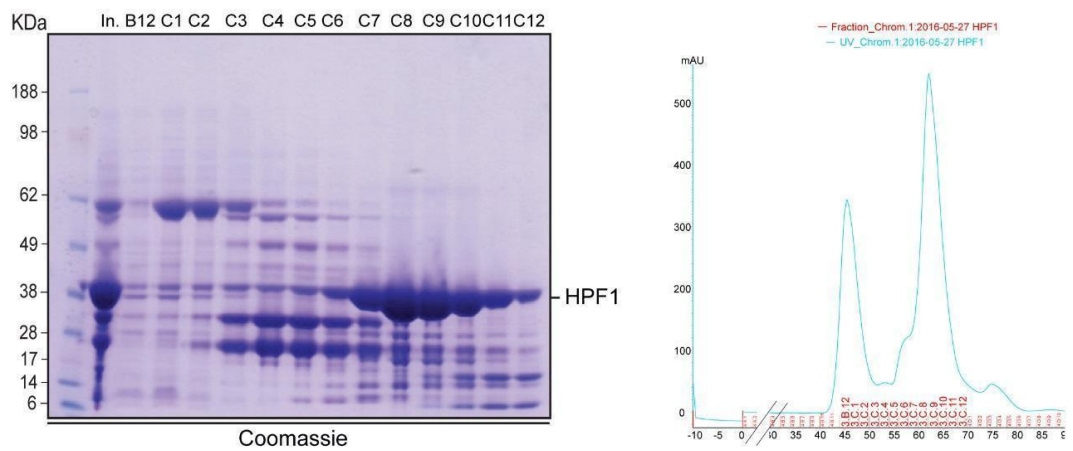
Purification

Pellet from each litre of LB medium was resuspended in 25 mL of 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.5 mM TCEP and EDTA-free tablet protease inhibitor cocktail (Roche, Cat. No. 5056489001) and Benzonase nuclease (Sigma, Cat. No. E1014-25KU), lysis was performed by homogenisation. The lysate was centrifuged for 1 hour at 4 °C at 30,000 g to remove cell debris. Soluble fraction was loaded on a 5 mL HisTrap HP column (GE Healthcare, Cat. No 17-5248-02) pre-equilibrated with 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.5 mM TCEP and 10 mM Imidazole. Protein bound to the column was washed with 10 column volume of equilibration buffer: 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.5 mM TCEP, 20 mM imidazole. Finally protein was eluted with a gradient of imidazole between 20 mM and 500 mM in 15 column volume (**Figure 2.1A**). The eluted protein was concentrated using a column filter with 10 kDa cut-off (Millipore, Cat. No. UFC801024) at 4 °C to a final volume of 2.5 mL. Concentrated protein was loaded onto a Superdex S75 16/60 gel-filtration column (GE Healthcare), with a 2 mL loading loop, pre-equilibrated with 10 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.2 mM TCEP. Eluted protein was concentrated and 10% glycerol was added before snap freezing in liquid nitrogen for storage at -80 °C until use (**Figure 2.1B**). For crystallisation purposes one more step of purification was performed. Eluted protein from gel filtration was loaded onto HiTrap Q Anion exchange HP column (GE Healthcare, Cat. No. 17115401) pre-equilibrated with 10 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.2 mM TCEP. HPF1 does not bind to Q anion column while its contaminant proteins do. HPF1 unbound fraction was concentrated and immediately used for crystallisation (**Figure 2.1C**).

A: 5mL HisTrap HP



B: Superdex S75 16/60 gel-filtration



C: 5mL Q Anion exchange HP

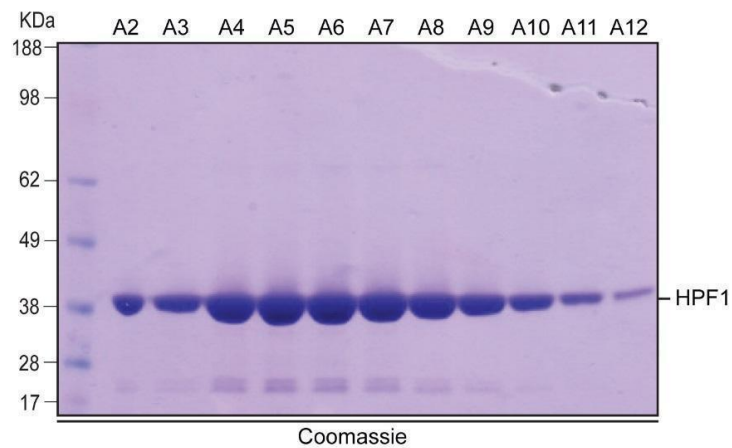


Figure 2.1. A - SDS-PAGE of human HPF1 purification fractions from HisTrap elution. Ly - Lysate, Ft - Flow-through; B - SDS-PAGE and chromatogram of human HPF1 fractions from Superdex S75 16/60 gel-filtration. C - SDS-PAGE of human HPF1 fractions from anion exchange Q flow-through.

2.4 Poly-ADP-ribose binding assay

Poly-ADP-ribose binding assay was performed as described in Ahel *et al.*, 2008 with a slight modification. Briefly, PARP1 was auto-modified in presence of DNA, NAD⁺ and ³²P-NAD⁺ for generation of poly-ADP-ribose by PARP1 auto-modification. Excess of NAD⁺ was removed using a desalting G25 column (GE Healthcare, Cat. No. 27532501). PARP1 was then digested using proteinase K for 30 minutes at room temperature and successfully inactivated at 80 °C for 10 minutes. Recombinant protein was spotted onto nitrocellulose membrane and let dry for a few minutes. The membrane was then blocked with the following buffer: Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% Tween with 5% milk. After removing the blocking buffer, the membrane was incubated with poly-ADP-ribose diluted in Tris-HCl, pH 7.5, 150 mM NaCl, 0.05% Tween with 5% milk. Finally, the membrane was washed in the same buffer and dried prior to analysis by autoradiography.

2.5 Crystallography data collection and processing

2.5.1 Data collection

X-ray diffraction datasets are collected using specialised machinery for X-ray generation and detection of diffracted X-rays, ranging from in-house sources (Skarzynski *et al.*, 2013), to synchrotrons (Owen *et al.*, 2006) to free-electron lasers (Barty *et al.* 2013). Of these, synchrotron radiation, first applied to protein crystals in 1976 (Phillips *et al.*, 1976), has become the most widely used method that accounts for the majority of structures currently deposited in the Protein Data Bank. Protein X-ray diffraction data are typically collected over a range of incident angles by placing the crystal in an X-ray beam and rotating it around an (often arbitrary) axis. Images are collected with varying slicing. The total number of degrees collected, positioning and slicing can be optimised based on crystal symmetry, unit cell dimensions, crystal quality and predicted resistance to radiation damage. Unlike salt crystals, protein crystals possess a high mosaicity due to slight misalignments of its different portions. This mosaicity, moreover, increases due to the use of cryo-protectants and cryo-cooling in liquid nitrogen. In addition, the monochromators currently in use result in a non-perfect monochromation of the X-rays. Such phenomena lead to spots being distributed over a range of degrees.

2.5.2 Data processing

Data processing for X-ray crystallography dataset can be achieved through the use of different programs such as Mosflm (Leslie and Powel, 2007), HKL2000 (Otwinowski and Minor, 1997) and the XDS suite (Kabsch, 2010).

First step of data processing is the data reduction, which can be divided into four different parts: spot finding and auto-indexing, parameter refinement, integration (measurement of spot intensities) and scaling and merging.

2.5.2.1 Spot finding and auto-indexing

The identification of the spots or “indexing” is usually an automated process which gives an estimation of the crystal symmetry and lattice dimensions. However, prior to auto-indexing, a set of spots has to be selected. Such selection can be carried out on contiguous images (Denzo), separated ones (Mosflm), a combination of both (XDS) or through the entire dataset (DIALS) (Gildea *et al.*, 2014). Independently from the method used, the auto-indexing will give as result a primitive triclinic cell. Such results will be compared with all the 44 typical protein lattices through standard matrix transformation. The software will then describe the fit of the primitive cell into the lattices suggesting the best fits.

2.5.2.2 Parameter refinement

The auto-indexing will lead to the identification of the approximate parameters of the unit cell, namely side lengths and angles. If the symmetry is identified correctly, it will lead to more accurate refinement parameters. Two methods are used for parameter refinement. The first is the positional refinement, the optimisation of the fit of

the observed to calculated spot coordinates on the images. The second analyses the relative intensity of partially recorded reflections to optimise the distance of the reciprocal lattice points from the Ewald sphere. Mosflm uses both these methods during cell refinement and during integration; HKL2000 and XDS use the second method just after integration, and DIALS uses just the first method.

2.5.2.3 Integration

Measurement of the spot intensity or “integration” follows the first step of auto-indexing and can be divided into two methods: summation integration and profile fitting (Schreurs *et al.*, 2010). In short, the values attributed to any pixel from a single spot can be summed together giving the intensity of the spot. The limiting factors are: correctly locating the spots (indexing), the identification of all the pixels belonging to the same spot, and the background of the image (noise). For this reason, the background under the spot has to be determined indirectly.

The integration, which allows discriminating the spots from the background, needs to be optimised to increase the signal-to-noise ratio. So far different methods have been developed. Mosflm, for instance, uses the ‘profile tolerance’ parameter, where an octagonal mask of the right dimensions for the spot is applied. XDS and DIALS, on the other hand, identify the spot size by finding the pixels surrounding the spots with the intensity not corresponding only to the background level.

2.5.2.4 Scaling and merging

Due to the crystal’s radiation damage, changes in incident radiation, illuminated volume of the crystal, anisotropic absorption of X-rays by the crystal, and other

physical phenomena, the integrated data are not on a uniform scale, making a scaling process necessary. The process of scaling and merging of the data will give statistic parameters known as R values, which define the agreement between the refined structure and the experimental data. Different softwares have been developed for scaling and merging, such as Aimless (Evans and Murshudov, 2013), Scala (Evans, 2006), Scalepack (Otwinowski and Minor, 1997) and XSCALE (Kabsch, 2010). Merging includes two different aspects, namely merging of the successive images that make up a single spot and merging of symmetry equivalent reflections which should have the same intensity.

2.5.3 Phasing

The last step for structure solution is the identification of the phases (Taylor, 2010). Phases are related to the structure factor amplitudes just through the electron density. Therefore, if we have prior knowledge of the electron density, we can estimate the phase. For protein crystals, there are two different general approaches for solving the phase problem: molecular replacement and experimental phasing.

2.5.3.1 Molecular replacement

Molecular replacement, the most common phasing method, is based on the use of the phases of a structurally similar model that is available, such as a homologous protein or truncated domain (Rossmann and Blow, 1962). In general, a sequence identity of at least 25% may be sufficient for molecular replacement. Traditionally, rotational and translational searches are based on the Patterson functions

(producing maps with peaks at interatomic distance vectors), but alternative search methods are available (Evans and McCoy, 2008).

2.5.3.2 Isomorphous replacement

Heavy atoms, such as gold, mercury and platinum, can be introduced into the protein crystals for solving the phase problem (Groth, 1908). The first application of this method to protein crystallography was carried out by Max Perutz and John Kendrew (Perutz, 1956; Kendrew *et al.*, 1958), who soaked the protein crystals in heavy-atom solutions and solved the phase problem by isomorphous replacement. The amplitudes of a reflection are measured for the native crystal $|FP|$, and for the derivative crystal, $|FPH|$. The difference, $|FH| = |FPH| - |FP|$, can be used to determine the heavy atom's position using Patterson or direct methods. Having localised the heavy atom position, it can be optimised to calculate an accurate $|FH|$ and its phase.

The method used here is single-wavelength anomalous dispersion (SAD). Most commonly, selenium atoms are introduced into the protein by substituting methionine for selenomethionine, and today this method is far more commonly used than heavy atom soaking (Pike *et al.*, 2016). Briefly, this method involves the measurements of the anomalous diffraction differences $\Delta F_{\pm} = |F_{PH(+)}| - |F_{PH(-)}|$. The differences are used to estimate the contribution of the scattering by the heavy atom, such as selenium in the case of selenomethionine-substituted NvHpf1, allowing the use of Patterson methods for identification of the positions of the seleniums (Taylor, 2010).

2.5.4 Phase improvement, model building and refinement

Initially determined phases, especially when obtained experimentally, are not particularly accurate, and a process of optimisation is needed. The first step of this optimisation, known as density modification, relies on prior knowledge of the properties of electron densities in protein crystals (Cowtan, 2010). After density modification, the map is available for model building and refinement, during which the phases are further improved and the map interpreted. Refinement is performed by statistical improvement of the atomic coordinates to better fit the experimental data; it aims to maximise the agreement of the model and the map, improving the map. Such fitting is expressed as the R-work factor, which evaluates the agreement of the amplitudes and phases (expressed in the structure factors) calculated from the model with the experimentally observed amplitudes and phases. However, the refinement method bears the risk of potential model bias since it uses the atomic model and the diffraction data for the calculation of the electron density. The validation of the refinement is then dependent on the contextual calculation of an R-free factor. The R-free factor is calculated exactly as the R-work factor but using just 5-10 % of the experimental reflections that are not used in the refinement, giving a value independent from the refinement itself. Rwork/Rfree are monitored during the refinement to assess the accuracy and iterative improvement of the model. Other parameters used to evaluate model quality include the deviation of bond lengths and angles from ideal values, agreement of dihedral angles with those possible within protein structures (Ramachandran plot), protein B-factor, and the clash score (**Table 3.2**). Images of the final model can be then generated using software such as PyMOL (**Figures 3.2, 3.3, 3.4**).

2.6 HPF1 Crystallisation, data collection, data processing, phasing and model building

Crystallisation trials of Hpf1 homologs were set up at 20°C using commercial screens (Molecular Dimensions) and the sitting-drop vapour-diffusion method. Ten different screens were initially used: Clear Strategy™ Screen I, Clear Strategy™ Screen II, Morpheus™, PACT premier™, Structure Screen 1 + 2, JCSG-plus™, The PGA Screen™, MIDAS™, The Stura Footprint Screen™ and ProPlex. Drops between 250 and 400 nL were set up, consisting of 50% (v/v) of protein solution with a concentration of around 10 mg/mL (see Table 3.1 for details) and 50% (v/v) of reservoir using a Mosquito Crystal robot (TTP Labtech). Initially, NvHpf1 crystals were optimised and diffracted to a final resolution of 2.5 Å in 0.2 M imidazole malate pH 5.5, 15% v/v PEG 600. DmHpf1 crystallised after the first attempt in 0.1 M Tris-HCl pH 8, 2 M ammonium sulphate and diffracted to 2.9 Å. Crystals were soaked in crystallisation solution supplemented with 15% (v/v) glycerol prior to being vitrified by submersion in liquid nitrogen. Selenomethionine-substituted NvHpf1 was purified and crystallised exactly like the native protein. The NvHpf1 construct is 327 amino acids long protein and contains 12 methionines, 10 of which gave anomalous diffraction when substituted for selenomethionine.

Datasets were collected at the Diamond Light Source, beamline DLS/I03 and DLS/I24. The optimal wavelength of 0.9796 Å was used for allowing anomalous diffraction from the selenium present in the selenomethionine-substituted NvHpf1. Structure determination was performed by Dr. Antonio Ariza. Data reduction was performed using Xia2 (Winter *et al.*, 2013). Xia2 is a system for the reduction of macromolecular crystallography data employing existing software. Xia2 can process one or more data set sequences of images at one or multiple wavelengths and convert the

reflections in the images to structure-factor amplitudes. The process is completely automated and does not require any input from the user. Selenomethionine-substituted NvHpf1 crystals showed unit cell dimensions of 88.95 Å, 88.95 Å, 224.13 Å, 90°, 90°, 120° (a, b, c, α , β , γ) and a resolution range of 45.31 Å - 2.80 Å while native NvHpf1 crystals had unit cell dimensions of 89.14 Å, 89.14 Å, 223.26 Å, 90°, 90°, 120° and a resolution range of (Å) 73 Å - 2.4 Å (**Table 3.2**).

To solve the phase problem, molecular replacement (Rossmann and Blow, 1962; Huber, 1965) was not possible due to the lack of any available structure of a protein homologous to Hpf1.

Structure solution was performed using selenomethionine-substituted NvHpf1, via single-wavelength anomalous dispersion (SAD) and achieved using AutoSol (Terwilliger *et al.*, 2009). AutoSol performs automated structure solution using either SAD, multi-wavelength anomalous diffraction (MAD), single isomorphous replacement (SIR) or multiple isomorphous replacement (MIR) data. Different potential solutions of heavy-atom substructures are scored on the basis of numerous criteria using Bayesian statistics or Z-scores. Structure solution for native NvHpf1 and DmHpf1 was performed by molecular replacement using selenomethionine-substituted NvHpf1 as phase model.

Electron density modification was performed with PARROT (Cowtan, 2010). The initial model for native NvHpf1 and DmHpf1 was built using BUCCANEER (Cowtan, 2006). BUCCANEER is a software for automated model building which uses an electron-density likelihood target function for the identification of possible C α positions. BUCCANEER also uses pre-existing methods with different approaches, such as the calculation of a skeleton of ridgelines connecting peaks in the electron density

map and its interpretation in terms of C α positions (Jones *et al.*, 1991), the identification of secondary-structure locations by a six dimensional rotation and translation search considering an “ideal” fragment, and calculation of the electron density (Kleywegt & Jones, 1997).

Finally, refinement was performed in reciprocal space by REFMAC5 (Murshudov *et al.*, 1997) and in real space, with iterative model building, by COOT (Emsley and Cowtan, 2004), excluding 5% of the total reflections for R_{free} calculation.

2.7 Cloning, expression and purification of GST-tagged HPF1 and PARP1

Cloning

GST-tagged variants of PARP1 in pGex4T1 vector were a gift from Dr. John Pascal.

GST-tagged variants of HPF1 in pGex4T1 were cloned from the previously cloned pET28-HPF1. HPF1 was amplified by PCR, using a Phusion polymerase in standard PCR condition (**Table 2.1**), from a synthesised DNA using the primers in table 2.4.

PCR product were purified from the PCR mix and digested 60 minutes at 37°C, in a total volume of 20 µL using 0,5 µL of NEB BamHI restriction enzyme (Cat. No. R0136S) 0,5 µL of NEB XhoI restriction enzyme (Cat. No. R0146S) in NEB 3.1 buffer (Cat. No. B7203S). Digested DNA product and plasmid were loaded in a 1% agarose gel, and gel purified upon separation. DNA concentration were assessed with nanodrop and 20 ng of digested plasmid were mixed with an excess of 3 or 5 fold of digested insert in water, in a total volume of 5 µL. Mixed DNA were ligate using one volume of Takara mix ligase (Cat. No. 6023) for 1h at 16 °C. 3 µL of ligation mix was used for transformation of 30 µL of *E.coli* DH5α competent cells. Correct insertion was confirmed by sequencing.

Primer	Restriction enzyme	Sequence
HsHPF1 F	BamHI	5'-CGTGCGGGATCCGGCATGGTCGGCGGTGGC-3'
HsHPF1 R	XhoI	5'-CCGCTCGAGTCATGCAGCAAGTTGG-3'

Table 2.4. Primers for cloning HPF1 in pGex-4T1 vector.

Expression

Rosetta DE3 cells were transformed with the appropriate plasmid and plated on LB agar in the presence of 100 µg/mL of ampicillin and 30 µg/mL of chloramphenicol. A single colony was picked and grown overnight in LB medium supplemented with the same antibiotics. The pre-culture was added, at 1% final concentration, to LB medium supplemented with antibiotics and cells were grown at 37°C until $OD_{600} = 0.6$. 0.5 mM final concentration IPTG was added for inducing protein expression and cells were incubated at 18°C overnight. Cells were collected in the morning by centrifugation at 4000 g for 10 minutes and pellets stored at -20 °C until use.

Purification

Pellet from each litre of LB medium was resuspended in 25 mL of PBS, 1mM DTT and EDTA-free tablet protease inhibitor cocktail (Roche, Cat. No. 5056489001) and Benzonase nuclease (Sigma, Cat. No. E1014-25KU), lysis was performed by homogenisation. The lysate was centrifuged for 1 hour at 4 °C at 30,000 g to remove cell debris. The supernatant was loaded onto Glutathione resin previously washed twice with water and twice with PBS + 1 mM DTT. 250 µL of 50% resin was used for 25 mL of supernatant. Resin and supernatant were incubated in rotation wheel for 1 h at 4 °C. The resin was poured through a filter column (the flow-through was recovered) and then washed with 10 column volume of PBS and 1 mM DTT. Proteins were eluted by incubation of the resin for 20 minutes at 4 °C with 2 column volume of elution buffer PBS, 1 mM DTT and 10 mM reduced glutathione. The eluted protein was then dialysed overnight at 4°C using an optimal buffer for the follow GST pull-down assay: 150 mM KCl, 12 mM NaCl, 2 mM MgCl₂, 100 µM zinc acetate (if needed), 5 mM DTT and 0.1% Triton X-100. 10% final concentration of glycerol was added to

the pure proteins prior to snap freezing in liquid nitrogen and stored at -80°C until use.

2.7.1 GST pull-down assay

Durapore PVDF 0.22 μm centrifuge tube membranes were moistened with 100 μl TZNK buffer (150 mM KCl, 12 mM NaCl, 2 mM MgCl_2 , 100 μM zinc acetate (if needed), 5 mM DTT and 0.1% Triton X-100). Tubes were then centrifuged for 30 seconds at 4500 g. 55 μl of 50% glutathione-sepharose slurry was loaded in the tube and washed 4 times with 100 μl TZNK buffer. 110 μl of a 5 μM solution of a GST-tagged protein was loaded onto the resin and incubated for 30 minutes at room temperature on a rotating wheel. Proteins were washed twice with 100 μl TZNK/D/T. 110 μl of 5 μM binding partner was then loaded and 10 μl of the mix collected as input. The mix was then incubated for 30 minutes at room temperature. The mix was then centrifuged for 30 seconds at 4500 g and flow-through collected. The resin with the bound protein was washed 6 times with 100 μl TZNK/D/T, once with TZNK buffer and the flow-through was collected. Protein was eluted by adding 50 μl elution buffer 50 mM Tris-HCl pH 8, 200 mM NaCl and 10 mM reduced glutathione and incubated for 30 min at room temperature on a rotating wheel.

2.8 Cloning, expression and purification of HsPARP1 and DmParp1

WT, point mutants and truncations were expressed and purified as previously reported for HsPARP1 (Langelier *et al.*, 2011). Point mutants and truncations were obtained by site directed mutagenesis using the QuikChange Lightning Site-Directed Mutagenesis Kit.

Cloning

HsPARP1 in pET28 with an N-Term 6His tag was a gift from Dr. John Pascal. DmParp1 was cloned into pET28 with an N-Term 6His tag. DmParp1 was amplified by PCR, using a Phusion polymerase in standard PCR condition (**Table 2.1**), from a synthesised DNA using the primers in table 2.5.

PCR product were purified from the PCR mix and digested 60 minutes at 37°C, in a total volume of 20 µL using 0,5 µL of NEB NdeI restriction enzyme (Cat. No. R0111S) 0,5 µL of NEB XhoI restriction enzyme (Cat. No. R0146S) in NEB 3.1 buffer (Cat. No. B7203S). Digested DNA product and plasmid were loaded in a 1% agarose gel, and gel purified upon separation. DNA concentration were assessed with nanodrop and 20 ng of digested plasmid were mixed with an excess of 3 or 5 fold of digested insert in water, in a total volume of 5 µL. Mixed DNA were ligate using one volume of Takara mix ligase (Cat. No. 6023) for 1h at 16 °C. 3 µL of ligation mix was used for transformation of 30 µL of *E.coli* DH5α competent cells. Correct insertion was confirmed by sequencing.

Primer	Restriction enzyme	Sequence
DmParp1 F	NdeI	5'-CGTGC GCATATGATGGATATTGAATTACCTTATCTTGCTG-3'
DmParp1 R	XhoI	5'-GCTCCGCTCGAGTTAATAAGAATACTTGAATTCCATACGAAAC-3'

Table 2.5. Primers for cloning DmParp1 in pET28 vector.

Expression

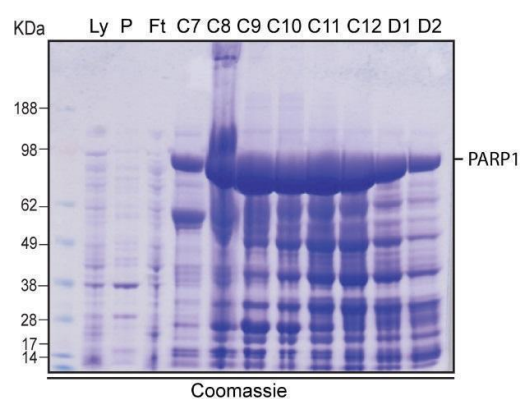
Rosetta DE3 were transformed with the appropriate plasmid and plated on LB agar in the presence of 50 µg/mL of kanamycin and 35 µg/mL of chloramphenicol. A single colony was picked and grown overnight in LB medium supplemented with the same antibiotics and 1% glucose. The pre-culture was added, at 1% final concentration, to LB medium supplemented with antibiotics and cells were grown at 37 °C until $OD_{600} = 0.4-0.5$ before addition of $ZnSO_4$ to the final concentration of 0.1 mM. The culture was subsequently grown at 37°C until $OD_{600} = 0.9-1$. The culture was then chilled in ice for 1 hour to stop the growth. Finally, the addition of IPTG to the final concentration of 0.2 mM was used to induce the expression and the culture was incubated at 16 °C for at least 16 h. The cell pellet was collected by centrifugation of the cell culture for 10 minutes at 4 °C at 4000 g and stored at -20 °C until used.

Purification

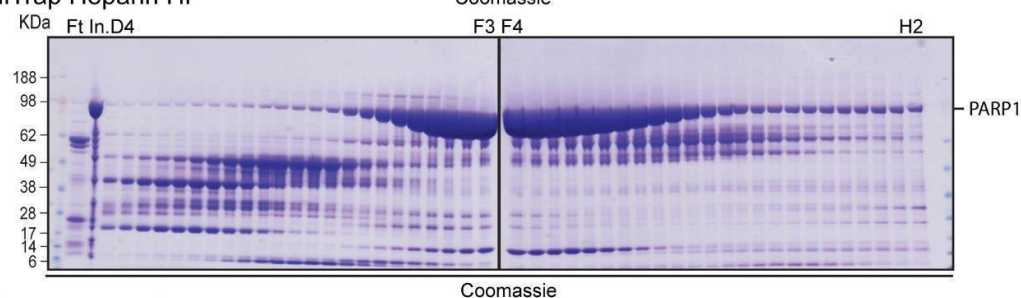
Each litre of pellet was resuspended in 25 mL of resuspension buffer: 25 mM Hepes pH 8.0, 500 mM NaCl, 0.5 mM TCEP and 0.1% NP-40 and EDTA-free tablet protease inhibitor cocktail (Roche, Cat. No. 5056489001), lysis was performed by homogeniser. Cell debris were removed by centrifugation for 1 hour at 4 °C at 30,000 g. The soluble fraction was filtered through 0.22 µm filter syringe and then loaded on a 5 mL HisTrap HP Column (GE Healthcare, Cat. No 17-5248-02) pre-equilibrated with

25 mM Hepes pH 8.0, 500 mM NaCl, 0.5 mM TCEP. The protein bound to the column was first washed with 20 column volume of low salt buffer: 25 mM Hepes pH 8.0, 500 mM NaCl, 0.5 mM TCEP, 20 mM imidazole, followed by 20 column volume of high salt wash buffer: 25 mM Hepes pH 8.0, 1 M NaCl, 0.5 mM TCEP, 20 mM imidazole and again 20 column volume of low salt buffer. Finally, the protein was eluted with a gradient of imidazole between 20 mM and 500 mM in 10 column volume (**Figure 2.2A**). Eluted protein was diluted by half with no-salt buffer: 50 mM Tris pH 7.0, 1 mM EDTA, 0.1 mM TCEP to lower the NaCl concentration to 250 mM. Diluted protein was then loaded onto a 5 mL HiTrap heparin HP column (GE Healthcare, Cat. No. 17040703) pre-equilibrated with 50 mM Tris pH 7.0, 1 mM EDTA, 0.1 mM TCEP and 250 mM NaCl. The bound proteins were eluted with a gradient of NaCl using the increasing concentrations of the buffer 50 mM Tris pH 7.0, 1 mM EDTA, 0.1 mM TCEP, 1 M NaCl (**Figure 2.2B**). The eluted protein was concentrated using a column filter with 10 kDa cut-off (Millipore, Cat. No. UFC801024) at 4 °C to a final volume of 2.5 mL. Concentrated protein was loaded onto a S200 Sephacryl 320 mL size exclusion column (GE Healthcare) pre-equilibrated with 25 mM Hepes pH 8.0, 150 mM NaCl, 1mM EDTA, 0.1mM TCEP, using a 2 mL loading loop (**Figure 2.2C**). Eluted protein from gel filtration was concentrated with column filter with 10 kDa cut-off (Millipore, Cat. No. UFC801024) at 4 °C, snap frozen in liquid nitrogen and saved at -80 °C until used.

A: 5mL HisTrap HP



B: 5 mL HiTrap Heparin HP



C: S200 Sephacryl gel-filtration

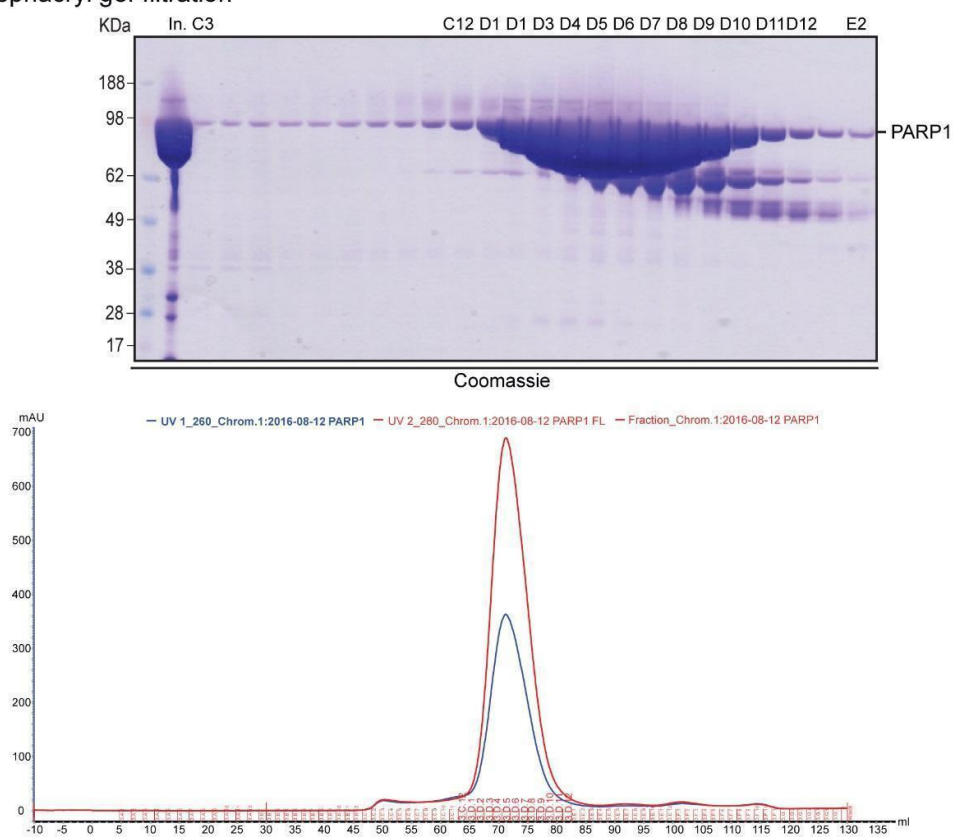


Figure 2.2. **A** - SDS-PAGE of PARP1 purification fractions from HisTrap elution. Ly - Lysate, P - Pellet, Ft Flow-through. Chromatogram not show. **B** - SDS-PAGE of PARP1 fraction from Heparin elution. Chromatogram not show. In. - Input. **C** - SDS-PAGE of PARP1 fraction from S200 Sephacryl gel-filtration.

2.9 PARP1 auto-modification assay

1 μM PARP1 +/- HPF1 was auto-modified in 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM MgCl_2 , DNA and 50 μM NAD^+ . The reaction was performed at room temperature (approximately 20 °C) for 20 minutes and stopped with the addition of 2 μM PARP1 inhibitor Olaparib. Products of the reaction were analysed by SDS-PAGE and immunoblotting with Anti-PAR antibody.

Alternatively, 1 μM PARP1 +/- 1 μM HPF1 were incubated in 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM MgCl_2 , DNA and 50 μM NAD^+ and ^{32}P - NAD^+ . The reaction was performed at room temperature (approximately 20 °C) for 20 minutes and stopped with the addition of 2 μM PARP1 inhibitor Olaparib. The products of the reactions were then analysed by SDS-PAGE and autoradiography.

2.10 *In vitro* glycohydrolase assays

The substrates for the glycohydrolase reaction were ADP-ribosylated using 1 μM PARP1 +/- 1 μM HPF1 in the following buffer: 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM MgCl_2 , DNA and 50 μM NAD^+ and ^{32}P - NAD^+ . The reaction was performed at room temperature (approximately 20 °C) for 20 minutes and stopped with the addition of 2 μM PARP1 inhibitor Olaparib. The product of the reaction was then incubated with the various glycohydrolases for 30 minutes at room temperature and analysed by SDS-PAGE and autoradiography. The concentrations used were 2 μM for recombinant protein substrate, 0.5 mg for histone peptides, and 3 μM for ARH3, DmParg and other glycohydrolase if not stated otherwise.

2.11 Cloning, expression and purification of PARP2

PARP2 was expressed and purified as previously reported (Langelier *et al.*, 2011).

Cloning

PARP2 in pET28 with an N-Term 6His tag was a gift from Dr. John Pascal.

Expression

Rosetta DE3 cells were transformed by heat shock. Cells were recovered for 1 hour in LB media complemented with 10 mM benzamide and plated in LB agar plate supplemented with 30 µg/mL of kanamycin and 30 µg/mL chloramphenicol and 10 mM benzamide.

A single colony was inoculated in an overnight pre-culture supplemented with 30 µg/mL of kanamycin, 30 µg/mL chloramphenicol, 10 mM benzamide and 1% glucose. 1% glucose pre-culture was added, at 1% final concentration, to LB medium supplemented with 30 µg/mL of kanamycin, 30 µg/mL chloramphenicol and 10 mM benzamide. The culture was incubated at 37°C until OD₆₀₀ = 0.8-0.9 and then chilled in ice for 1 hour to stop the growth. Finally, the addition of IPTG to the final concentration of 0.2 mM was used to induce expression and the culture was incubated at 16 °C for at least 16 h. The cell pellet was collected by centrifugation of the cell culture for 10 minutes at 4 °C at 4000 g and stored at -20 °C until used.

Purification

Each litre of pellet was resuspended in 25 mL of resuspension buffer: 25 mM Hepes pH 8.0, 500 mM NaCl, 0.5 mM TCEP, 10 mM benzamide, 0.1% NP-40 and

EDTA-free tablet protease inhibitor cocktail (Roche, Cat. No. 5056489001) prior to cell lysis by homogeniser. Cell debris was removed by centrifugation for 1 hour at 4 °C at 30,000 g and the soluble fraction was filtered through 0.22 µm filter syringe. The filtered fractions were loaded onto a Ni-NTA resin pre equilibrated with 25 mM Hepes pH 8.0, 500 mM NaCl, 0.5 mM TCEP. 1 mL of 50% resin was used for 25 mL of supernatant. The protein bound to the column was first washed with 20 column volume of low salt buffer: 25 mM Hepes pH 8.0, 500 mM NaCl, 0.5 mM TCEP, 20 mM imidazole, followed by 20 column volume of high salt wash buffer: 25 mM Hepes pH 8.0, 1 M NaCl, 0.5 mM TCEP, 20 mM imidazole and again 20 column volume of low salt buffer. Finally, proteins were eluted with a gradient of imidazole between 20 mM and 400 mM in 10 column volume (**Figure 2.3A**). The eluted protein was diluted in half with no salt buffer: 50 mM Tris pH 7.0, 1 mM EDTA, 0.1 mM TCEP to lower the NaCl concentration to 250 mM. The diluted protein was then loaded onto a 5 mL HiTrap heparin HP column (GE Healthcare, Cat. No. 17040703) pre-equilibrated with 50 mM Tris pH 7.0, 1 mM EDTA, 0.1 mM TCEP and 250 mM NaCl. The bound protein was eluted with a gradient of NaCl 50 mM Tris pH 7.0, 1 mM EDTA, 0.1 mM TCEP, 1 M NaCl (**Figure 2.3B**). The eluted protein was concentrated using a column filter with 10 kDa cut-off at (Millipore, Cat. No. UFC801024) at 4 °C to a final volume of 2.5mL. Concentrated protein was loaded onto a S200 Sephacryl 320 mL size exclusion column (GE Healthcare) pre-equilibrated with 25 mM Hepes pH 8.0, 150 mM NaCl, 1mM EDTA, 0.1mM TCEP, using a 2 mL loading loop (**Figure 2.3C**). Eluted protein from gel filtration was concentrated with column filter with 10 kDa cut-off (Millipore, Cat. No. UFC801024) at 4 °C, snap frozen in liquid nitrogen and saved at -80 °C until used.

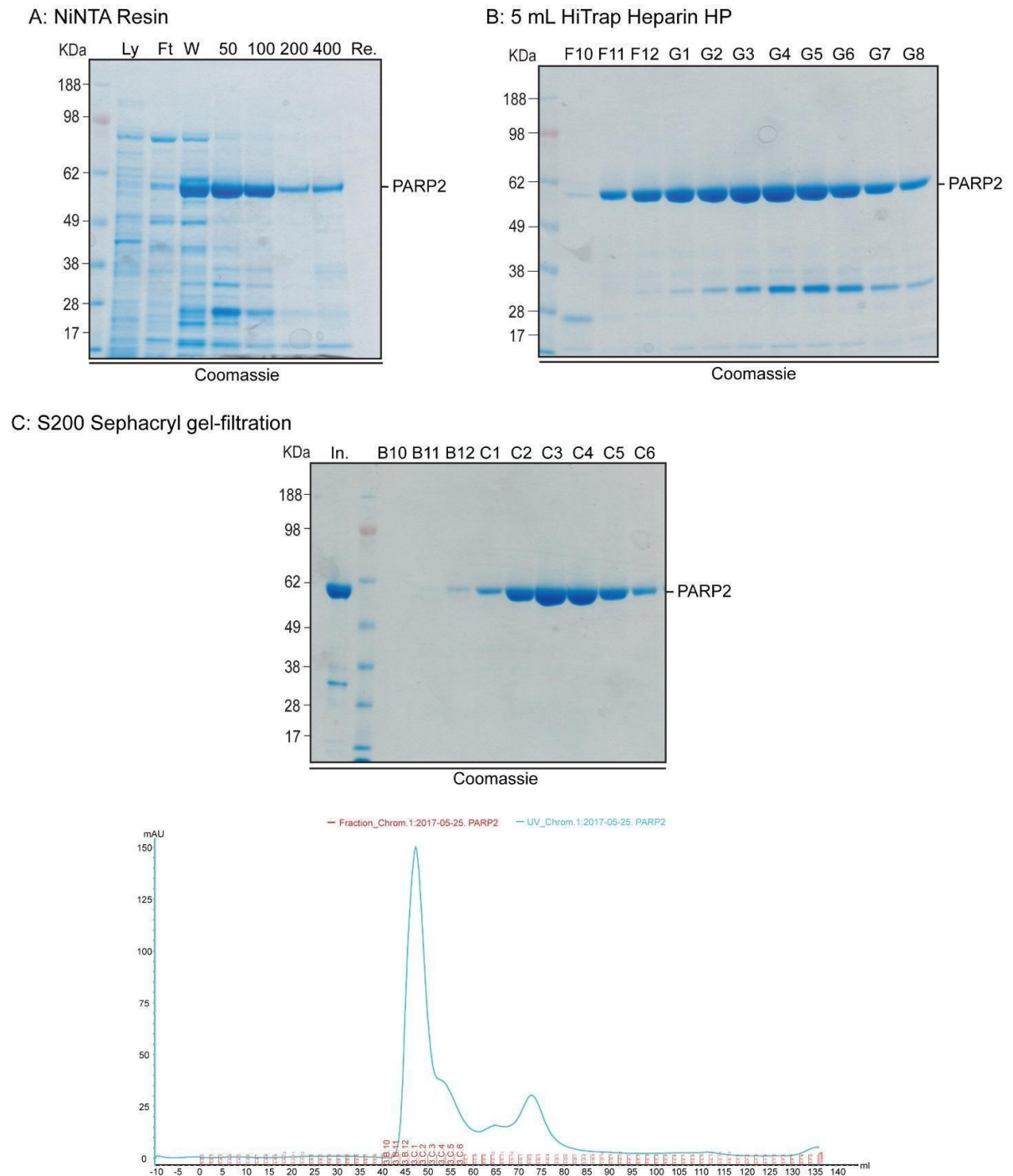


Figure 2.3. **A** - SDS-PAGE of PARP2 purification fractions from Ni-NTA elution. Ly - Lysate, Ft Flow-through, W - Wash with lysis buffer, 50 - Lysis buffer + 50 mM Imidazole, 100 - Lysis buffer + 100 mM Imidazole, 200 - Lysis buffer + 200 mM Imidazole, 400 - Lysis buffer + 400 mM Imidazole, Re. - Resin. **B** - SDS-PAGE of PARP2 fraction from Heparin elution. Chromatogram not shown. **C** - SDS-PAGE of PARP2 fraction from S200 Sephacryl gel-filtration. In. - Input.

2.12 EMSA Electrophoretic mobility shift assay

SsDNA primer 24nt 5'- ACTTGTCTCCAGCTGGGGCTATTT-3' was phosphorylated with T4 PNK in the presence of ^{32}P -ATP for 45 minutes at 37°C. PNK was inactivated at 65°C for 5 minutes. The reaction was passed through a G25 column (GE Healthcare, Cat. No. 27532501) pre-equilibrated with ddH₂O for removing the excess of ATP. Hot labelled ssDNA was mixed with an excess of 1.2X complementary cold ssDNA 5'- AAATAGCCCCAGCTGGAGACAAGT-3'. The mix was incubated at 90 °C for 5 minutes and left to cool down at room temperature allowing the formation of dsDNA. 2.5 nM of labelled DNA were mixed with recombinant HPF1 or recombinant Aprataxin at different concentrations in the following buffer: 50 mM Tris-HCl, pH 8, 30 mM NaCl, 5% glycerol, 1 mM DTT and 0.1 mg/ml BSA. The mix was incubated on ice for 20 minutes and then loaded onto a 4% acrylamide gel and run for 1 hour at 100 V in 0.5X TBE buffer. Gels were then dehydrated and analysed by autoradiography.

2.13 Cloning, expression and purification of ARH1, ARH2 and ARH3

WT forms of ARH1, ARH2 and ARH3 as well as ARH3 mutants were expressed and purified as reported for ARH3 (Kernstock *et al.*, 2006), with a few differences in the protocol.

Cloning

ARH3 expression construct in pDEST17 was a gift from Prof Paul Hergenrother (University of Illinois). ARH1 and ARH2 were cloned from synthesized cDNA into pDEST17 using the gateway cloning technology as recommended from the manufacturer (Invitrogen, Cat. No. 11789-100 and 11791-100) using the primer in table 2.6.

Catalytic mutants of ARH3 were obtained by site directed mutagenesis using the QuikChange Lightning Site-Directed Mutagenesis Kit in standard mutagenesis condition (**Table 2.2**) using the primers in table 2.6.

Primer	AttB	Sequence
ARH1 F	AttB1	5'-GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGGAGAAGTATGTGGCTGCTATG-3'
ARH1 R	AttB2	5'-GGGGACCACTTTGTACAAGAAAGCTGGGTATTAAAGGGAATTACAGTGTCTTCTTTGAC-3'
ARH2 F	AttB1	5'-GGGGACAAGTTTGTACAAAAAGCAGGCTTCATGGTGAGATGCTATGTGGAATC-3'
ARH2 R	AttB2	5'-GGGGACCACTTTGTACAAGAAAGCTGGGTATTACTTCTCTCTGTGGACAGGCGG-3'
Mutagenesis Primers		
ARH3 D77N F		5'-CTTGACTACACAAATgACACAGCCATGGC-3'
ARH3 D77N R		5'-GCCATGGCTGTGTcATTTGTGTAGTACAAG-3'
ARH3 D78N F		5'-CTTGACTACACAgATAACACAGCCATGGC-3'
ARH3 D78N R		5'-GCCATGGCTGTGTTATcTGTGTAGTACAAG-3'
ARH3 D77N-D78N F		5'-CTTGACTACACAAATAACACAGCCATGGC-3'
ARH3 D77N-D78N R		5'-GCCATGGCTGTGTTATTTGTGTAGTACAAG-3'

Table 2.6. Primers for cloning ARH1 and ARH2 in gateway system and primers for ARH3 mutagenesis.

Expression

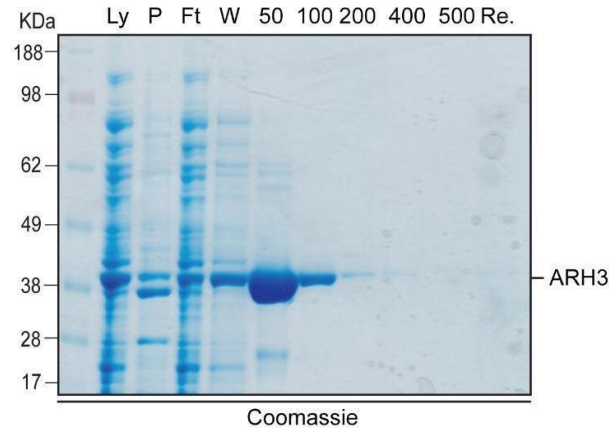
Rosetta DE3 were transformed with the appropriate plasmid and plated on LB agar in the presence of 30 µg/mL of kanamycin and 30 µg/mL of chloramphenicol. A single colony was picked and grown overnight in LB medium supplemented with these antibiotics. The pre-culture was added, at 1% final concentration, to LB medium supplemented with antibiotics and cells were grown at 37°C until OD₆₀₀ = 0.6. 0.5 mM final concentration IPTG was added for inducing protein expression. Cells were incubated at 18°C overnight and collected in the morning by centrifugation at 4,000 g for 10 minutes and the pellet was stored at -20 °C until use.

Purification

Pellet from each litre of LB medium was resuspended in 25 mL of 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10 mM MgCl₂, 1 mM DTT, EDTA-free tablet protease inhibitor cocktail (Roche, Cat. No. 5056489001) and Benzonase nuclease (Sigma, Cat. No. E1014-25KU), lysis was performed by homogenisation. The lysate was centrifuged for 1 hour at 4 °C at 30,000g to remove cell debris. The soluble fraction was loaded on Ni-NTA resin pre-equilibrated with 50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 10 mM MgCl₂, 1 mM DTT and 10 mM imidazole. 1 mL of 50% resin was used for 25 mL of supernatant. Protein bound to the column was washed with 10 column volume with the same buffer. Finally the protein was eluted with a gradient of imidazole between 20 mM and 500 mM in 10 column volume (**Figure 2.4A**). The eluted protein was concentrated using a column filter with 10 kDa cut-off at (Millipore, Cat. No. UFC801024) 4 °C to a final volume of 2.5mL. Concentrated protein was loaded onto a Superdex S75 16/60 gel-filtration column (Ge Healthcare) pre-equilibrated with 10 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.2 mM TCEP. 25 mM Tris-HCl pH 8.0, 150

mM NaCl, 1 mM DTT, using a 2mL loop (**Figure 2.4B**). 10% final concentration of glycerol was added to the pure proteins prior to snap freezing in liquid nitrogen and stored at -80 °C until use.

A: NiNTA Resin



B: Superdex S75 16/60 gel-filtration

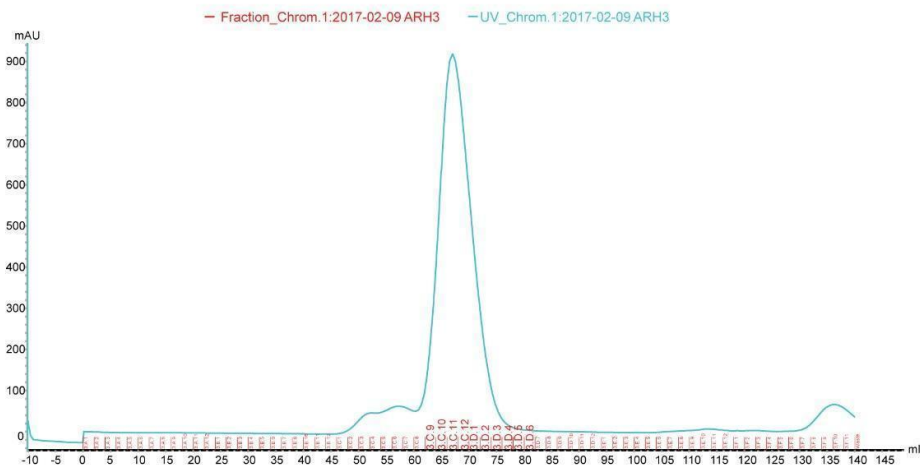
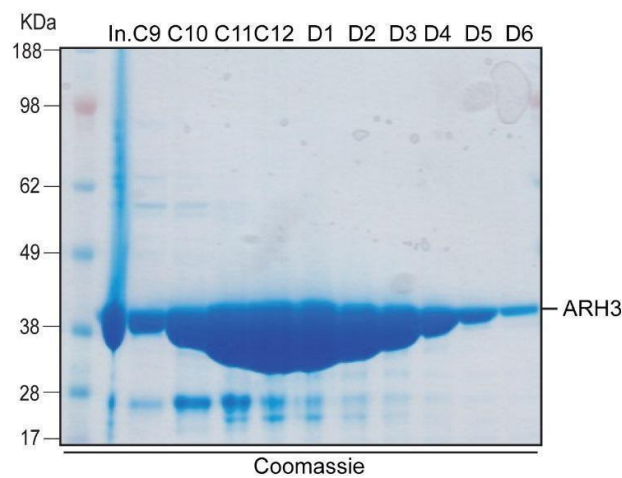


Figure 2.4. A - SDS-PAGE of ARH3 purification fractions from Ni-NTA elution. Ly - Lysate, P - Pellet, Ft Flow-through, W - Wash with lysis buffer, 50 - Lysis buffer + 50 mM Imidazole, 100 - Lysis buffer + 100 mM Imidazole, 200 - Lysis buffer + 200 mM Imidazole, 400 - Lysis buffer + 400 mM Imidazole, 500 - Lysis buffer + 500 mM Imidazole, Re. - Resin. **B** - SDS-PAGE of ARH3 fraction from Superdex S75 16/60 gel-filtration. In. - Input.

2.14 Purification of Ser-ADP-ribosylated histone peptide

Ser-ADP-ribosylated H3 peptide (aa 1-21 MARTKQTARKSTGGKAPRKQL) and Ser-ADP-ribosylated H2B peptide (aa 1-21 MPDPSKSAPAPKKGSKKAVTK) were purified from the mix reaction prior to incubation with the glycohydrolases. After the PARP1/HPF1 reaction, the mix was filtered using a 10 kDa cut-off concentration column (Millipore, Cat. No. UFC501096) in order to remove PARP1/HPF1. NAD⁺ was removed from the mix reaction using a desalting spin column G25 (GE Healthcare, Cat. No. 27532501). Purified peptides were then eluted with 10X buffer to a final concentration of 50 mM Tris-HCl pH 8.0, 100 mM NaCl, 2 mM MgCl₂.

2.15 ARH3 time course

5 µg of modified H3 histone peptide was incubated with ARH3 at a final concentration of 200 nM. An aliquot containing 0.5 µg of H3 histone peptide was taken at regular time intervals. To stop the ARH3 reaction, 1 mM ADPr and LDS sample buffer were added, followed by a 2 minutes incubation at 90 °C. The product of the reaction was analysed by SDS-PAGE and autoradiography.

2.16 Arginine-ADP-ribosylation of K562 and S2 cells extract

Arg-ADP-ribosylation of cell extract was performed as described in (Palazzo et al., 2016). 8-10 cells were collected from a 10 cm dish and washed twice with PBS. Lysis was performed with in the following buffer 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.5% Triton X-100, 0.2 mM DTT, 1 mM Olaparib, 4 mM Pefabloc SC PLUS (Sigma-Aldrich, Cat. No. 11873601001) for 45 minutes at 4 °C. Lysed cells were centrifuged for 30 minutes at 4 °C at 20,000 g. The soluble fraction was diluted five times in the same buffer lacking Triton X-100. Finally, 15 mM MgCl₂ and 1 mCi (37 kBq) of ³²P-NAD⁺ was added. The extracted proteins were incubated with 1 μM recombinant mARTC2.2 for 15 minutes at 30 °C. The Arg-ADP-ribosylated cell extract was then incubated with different glycohydrolases for 45 minutes at room temperature. The product of the reaction was analysed by SDS-PAGE and autoradiography

2.17 Lysine-ADP-ribosylation of H3 histone peptide

Chemical Lys-ADP-ribosylation of Histone H3 peptide was achieved as previously reported (Jankevicius et al., 2013). 2 mg of H3 histone peptide (aa 1-21) was incubated with 1 mM NAD⁺ in the presence of ³²P-NAD⁺ in 25 mM Tris-HCl pH 8.3 at 37 °C for 3 days. Purification of the modified peptide using a ZipTip pipette tips (Millipore, Cat. No. ZTC18M096) was performed in order to remove the excess of NAD⁺. 100% acetonitrile (ACN) was used to pre-wash the tips followed by a 4 times equilibration with 0.1 % trifluoroacetic acid (TFA). The mix reaction containing Lys-ADP-ribosylated histone H3 was eluted with TFA to a final concentration of 0.1% and

loaded into the ZipTip filter. Two washes with 0.1 % TFA followed after the load. Finally, the peptide was eluted with 50 % ACN and 0.1 % TFA. The eluted peptide was dried at 37 °C and resuspended in 25 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM DTT and 2 mM MgCl₂. The purified Lys-ADP-ribosylated histone H3 was then incubated with the different glycohydrolase for 30 minutes at room temperature. The products of the reaction were run on SDS-PAGE and analysed by autoradiography. Chemical digestion with 0.5 M KOH at 37°C for 30 minutes was used as a positive control.

2.18 Recombinant histones and histones peptides

Recombinant Histone tetramer H3/H4 was from NEB (Cat. No. M2509S). Recombinant Histone H1 was purchased from NEB (Cat. No. M2501S), recombinant histone H3 peptide from Sigma (Cat. No. SRP0159-160NMOL) and recombinant H2B from Millipore (14-491). Recombinant Octamer was purified as previously reported (Ahel *et al.*, 2009, Mehrotra *et al.*, 2011). Briefly, single histone (H3, H4, H2A and H2B) were expressed in *E.coli* BL21 untagged from pET vectors. BL21 cells were lysate and inclusion body, including the insoluble histone, separated by centrifugation. A two step purification protocol under denaturing condition, including HPLC and ion-exchange chromatography, allows the inclusion body separation. Unfolded histones were mixed together at equal molar ratio and refolded during dialysis at 4C°. Samples were then loaded on gel filtration chromatography and analysed by SDS-PAGE.

2.19 Cloning, expression and purification of DmParg, DmTarg1, DmTarg2 and DmTarg3

Cloning

DmParg, DmTarg1, DmTarg2 and DmTarg3 were cloned into pET28 with an N-Term 6His tag. DmParg, DmTarg1, DmTarg2 and DmTarg3 were amplified by PCR, using a Phusion polymerase in standard PCR condition (**Table 2.1**), from a synthesised DNA using the primers in table 2.7.

PCR product were purified from the PCR mix and digested 60 minutes at 37°C, in a total volume of 20 µL using 0,5 µL of NEB NdeI restriction enzyme (Cat. No. R0111S) 0,5 µL of NEB NotI restriction enzyme (Cat. No. R0189S) in NEB 3.1 buffer (Cat. No. B7203S). Digested DNA product and plasmid were loaded in a 1% agarose gel, and gel purified upon separation. DNA concentration were assessed with nanodrop and 20 ng of digested plasmid were mixed with an excess of 3 or 5 fold of digested insert in water, in a total volume of 5 µL. Mixed DNA were ligate using one volume of Takara mix ligase (Cat. No. 6023) for 1h at 16 °C. 3 µL of ligation mix was used for transformation of 30 µL of *E.coli* DH5α competent cells. Correct insertion was confirmed by sequencing.

Point mutants and truncations were obtained by site directed mutagenesis using the QuikChange Lightning Site-Directed Mutagenesis Kit in standard mutagenesis condition (**Table 2.2**) using the primers in table 2.7.

Primer	Restriction enzyme	Sequence
DmParg F	NdeI	5'-GTAGGCTACATATGTCTGAAGTCGCCGAT-3'
DmParg R	NotI	5'-GGGGAAAAAGCGCCGCCTAGTCAACAATATCGTCCTTTTCG-3'
DmTarg1 F	NdeI	5'-CTACTAGCTAGCATGTCTCAAAACCTCAGTACAC-3'
DmTarg1 R	NotI	5'-GGGGAAAAAGCGCCGCCTATTTTCCTGTGGAGGCAC-3'
DmTarg2 F	NdeI	5'-GTAGGCTACATATGGCGAACGTGAAGATAGTG-3'
DmTarg2 R	NotI	5'-GGGGAAAAAGCGCCGCCTTAAAGGGAGTCTTTTCGCG-3'
DmTarg3 F	NdeI	5'-GTAGGCTACATATGTCTAGGACTAACTTACAGAGAGATC-3'
DmTarg3 R	NotI	5'-GGGGAAAAAGCGCCGCCTTACTTCATAGCAACATAATTATAAAC-3'
Mutagenesis Primers		
DmParg Δ554-723 F		5'-CTTAAGAAGGTCAGATAAGCGGCCGCACTCGAGC-3'
DmParg Δ554-723 R		5'-GGCCGCTTATCTGACCTTCTTAAGTTCCTTTAATAAAATC-3'
DmParg Δ558-723 F		5'-CGGAGAGGGATAAGCGGCCGCACTCGAGC-3'
DmParg Δ558-723 R		5'-GGCCGCTTATCCCTCTCCGGGCACATCTCTG-3'
DmParg Δ574-723 F		5'-CTGGATTAGGCTAAGCGGCCGCACTCGAGC-3'
DmParg Δ574-723 R		5'-GGCCGCTTAGCCTAATCCAGCTACTCTAGAGCTTCC-3'
DmParg E339A F		5'-GACATGGCTGCGTTCAGGCGGAGATACGCTTTGTTATC-3'
DmParg E339A R		5'-GATAACAAAGCGTATCTCCGCCTGAACGCAGCCATGTC-3'
DmParg E339N F		5'-GACATGGCTGCGTTCAGAAATGAGATACGCTTTGTTATC-3'
DmParg E339N R		5'-GATAACAAAGCGTATCTCATTCTGAACGCAGCCATGTC-3'
DmParg E340A F		5'-CATGGCTGCGTTCAGGAGGCGATACGCTTTGTTATCTG-3'
DmParg E340A R		5'-CAGATAACAAAGCGTATCGCCTCCTGAACGCAGCCATG-3'
DmParg E340N F		5'-CATGGCTGCGTTCAGGAGAATATACGCTTTGTTATCTG-3'
DmParg E340N R		5'-CAGATAACAAAGCGTATATTCTCCTGAACGCAGCCATG-3'
DmParg F231A F		5'-GTCCGAGTACAGCACTGCTCCAGACATCAACTTTAAC-3'
DmParg F231A R		5'-GTTAAAGTTGATGTCTGGAGCAGTGCTGTACTCGGAC-3'
DmParg L333A F		5'-GGGTGGCGGTGTCGCGGGACATGGCTGCGTTC-3'
DmParg L333A R		5'-GAACGCAGCCATGTCCCGCGACACCGCCACCC-3'
DmParg C347A F		5'-GATACGCTTTGTTATCGTCCGGAGCTATTGGTG-3'
DmParg C347A R		5'-CACCAATAGCTCCGAGCGGATAACAAAGCGTATC-3'
DmParg V352A F		5'-CTGTCCGGAGCTATTGGCGGGTAAACTCTTTACG-3'
DmParg V352A R		5'-CGTAAAGAGTTTACCGCCAATAGCTCCGGACAG-3'
DmParg F384A F		5'-GGATATGCCGGAAGCGCCGAGTGGTCCGGCAAC-3'
DmParg F384A R		5'-GTTGCCGGACCACTCGCGCTTCCGGCATATCC-3'
DmParg E423A F		5'-CATCATCAATATCGCGCGGATCTCATGGAAGG-3'
DmParg E423A R		5'-CCTTTCCATGAGATCCGCGCGATATTGATGATG-3'
DmParg E546A F		5'-CTATATGATTTTATTAAAGCGGAACCTAAGAAGGTC-3'
DmParg E546A R		5'-GACCTTCTTAAGTTCGCTTTAATAAAATCATATAG-3'

Table 2.7. Primers for cloning DmParg, DmTarg1, DmTarg2, DmTarg3, and primers for DmParg mutagenesis.

Expression

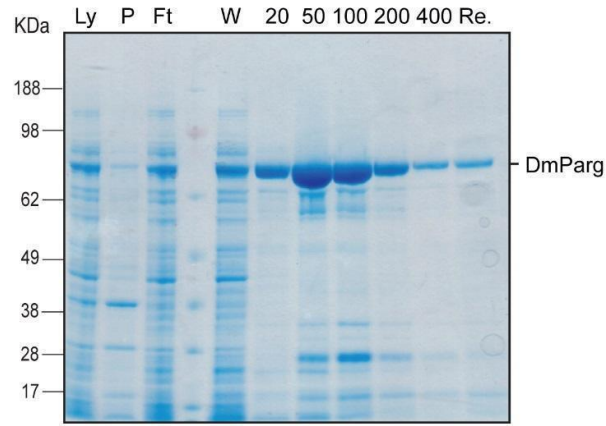
Rosetta DE3 were transformed with the appropriate plasmid and plated on LB agar in the presence of 30 µg/mL of kanamycin and 30 µg/mL of chloramphenicol. A single colony was picked and grown overnight in LB medium supplemented with the same antibiotics. The pre-culture was added, at 1% final concentration, to LB medium supplemented with antibiotics and cells were grown at 37°C until OD₆₀₀ = 0.7. The culture was then chilled on ice for 1 hour to stop the growth and 0.5 mM final concentration IPTG was added for inducing protein expression. Cells were incubated at 18°C overnight and collected in the morning by centrifugation at 4,000 g for 10 minutes. The pellet was stored at -20 °C until use.

Purification

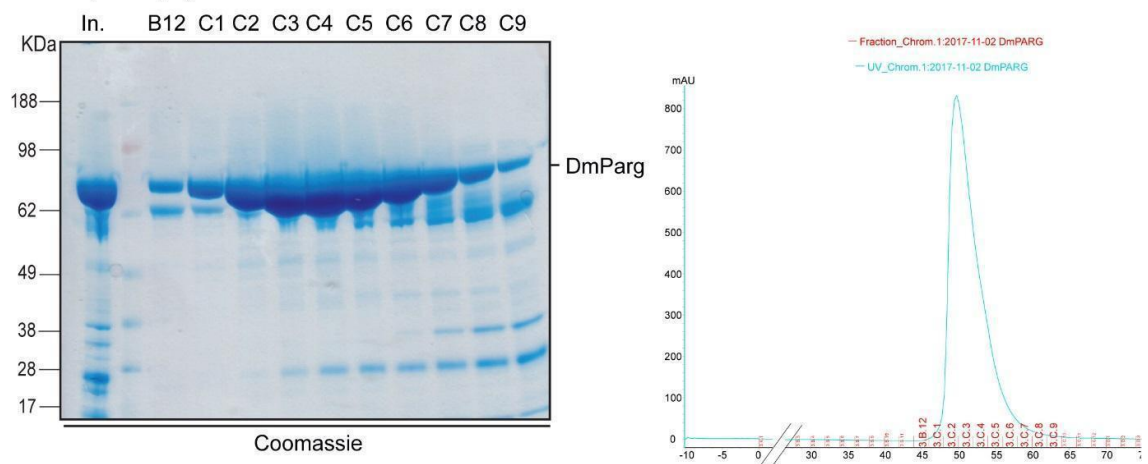
Pellet from each litre of LB medium was resuspended in 25 mL of 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.5 mM TCEP, EDTA-free tablet protease inhibitor cocktail (Roche, Cat. No. 5056489001) and Benzonase nuclease (Sigma, Cat. No. E1014-25KU), lysis was performed by homogenisation. The lysate was centrifuged for 1 hour at 4 °C at 30,000 g to remove cell debris. The soluble fraction was loaded onto a Ni-NTA HisTrap column (GE Healthcare) pre-equilibrated with 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.5 mM TCEP and 10 mM imidazole. 1 mL of 50% resin was used for 25 mL of supernatant. The protein bound to the column was washed with 10 column volume of equilibration buffer: 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.5 mM TCEP, 20 mM imidazole. Finally, the protein was eluted with a gradient of imidazole between 20 mM and 400 mM in 10 column volume (**Figure 2.5A**). Eluted DmTargs were concentrated using a column filter with 10 kDa cut-off (Millipore, Cat. No. UFC801024) at 4 °C to a final volume of 2.5 mL. Concentrated proteins were loaded onto a Superdex S75 16/60 size exclusion column using a 2 mL loading loop (GE

Healthcare), while eluted DmParg was concentrated using a column filter with 10 kDa cut-off (Millipore, Cat. No. UFC801024) at 4 °C to a final volume of 2.5mL and loaded on S200 Sephacryl 320 mL size exclusion column using a 2 mL loading loop (GE Healthcare), in each case pre-equilibrated with 10 mM Tris-HCl pH 8.0, 150 mM NaCl, 0.2 mM TCEP. The eluted protein was concentrated using a column filter with 10 kDa cut-off (Millipore, Cat. No. UFC801024) at 4 °C and 10% glycerol was added before snap freezing in liquid nitrogen and stored at -80 °C until used (**Figure 2.5B**). For DmParg crystallisation, one more step of purification was performed. The protein eluted from gel filtration was diluted in no salt buffer to a final concentration of 10 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.2 mM TCEP and then loaded onto HiTrap Q Anion exchange HP column (GE Healthcare, Cat. No. 17115401) pre-equilibrated with 10 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.2 mM TCEP. The bound fraction was eluted with a gradient of 10 column volume of NaCl of 100 mM to 1 M. The eluted protein was dialysed overnight in 10 mM Tris-HCl pH 8.0, 100 mM NaCl, 0.2 mM TCEP followed by concentration, with a column filter with 10 kDa cut-off (Millipore, Cat. No. UFC801024) at 4 °C and immediate use for crystallisation (**Figure 2.5C**).

A: NiNTA Resin



B: S200 Sephacryl gel-filtration



C: 5mL Q Anion exchange HP

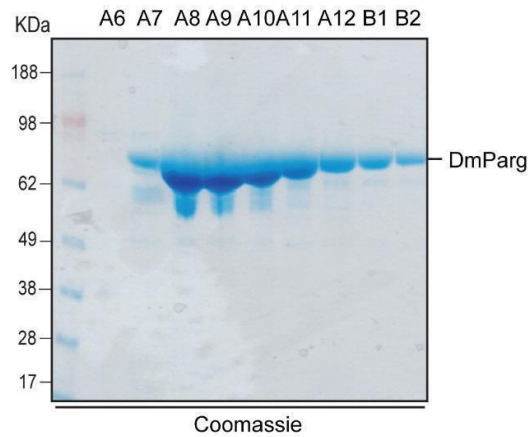


Figure 2.5. **A** - SDS-PAGE of DmParg purification fractions Ni-NTA elution. Ly - Lysate, Ft Flow-through, W - Wash with lysis buffer, 20 - Lysis buffer + 20 mM Imidazole, 50 - Lysis buffer + 50 mM Imidazole, 100 - Lysis buffer + 100 mM Imidazole, 200 - Lysis buffer + 200 mM Imidazole, 400 - Lysis buffer + 400 mM Imidazole, Re. - Resin. **B** - SDS-PAGE of DmParg fraction from S200 Sephacryl gel-filtration. **C** - SDS-PAGE of DmParg fraction from anion exchange Q elution. Chromatogram not show.

2.20 DmParg Δ 554-723 crystallisation and data collection

DmParg crystallisation trials were initially set up at 20°C using the commercial screen as previously described for HPF1. Drops between 250 and 500 nL were set up, consisting of either 1/2 of protein solution and 1/2 of the reservoir or 1/3 of protein solution and 2/3 of the reservoir. Protein concentration was initially 10mg/mL then decreased at 8 mg/mL or 6 mg/mL (see Table 3.1 for details) using a Mosquito Crystal robot (TTP Labtech). Initially, crystals showed no diffraction and were used as micro-seeds for new crystallisation trials. New trials set up at 4°C together using micro-seeding and the Additive Screen (Hamptons) gave crystals that diffracted to a final resolution of 2.4 Å in 19% (v/v) PEG3350, 0.21M Sodium Sulphate, 0.1M BisTris Propane pH7.2. Crystals were soaked in crystallisation solution supplemented with 15% (v/v) glycerol prior to being vitrified by submersion in liquid nitrogen. Datasets were collected at the Diamond Light Source, UK and European Synchrotron Radiation Facility (ESRF), France.

2.21 Proteins accession numbers

Nematostella vectensis Hpf1 XP_001637784, *Drosophila melanogaster* Parp1 NP_001104452, *Drosophila melanogaster* Hpf1 NP_649589, *Drosophila melanogaster* Parg NP_477321, *Tetrahymena thermophila* Parg XP_001013101, *Entamoeba dispar* Parg2 XP_001739172.

Chapter 3. Structural and biochemical characterisation of HPF1 and Ser-ADP-ribosylation

Here I report the structural and biochemical characterisation of a novel protein (C4orf27) which we named histone PARylation factor 1 (HPF1) involved in PARP1/PARP2-dependent DNA damage repair. HPF1 interacts directly with the catalytic domain of either PARP1 or PARP2. This interaction promotes ADP-ribosylation of histone proteins and, at the same time, suppresses auto-modification of PARP1. Moreover, HPF1 induces a novel, previously unreported, PTM named serine-ADP-ribosylation (Ser-ADP-ribosylation) on histone proteins and on auto-modified PARP1. Collectively, my data suggest that HPF1 plays an essential role in PARP1/PARP2-dependent ADP ribosylation and the newly discovered Ser-ADP-ribosylation is an essential PTM which plays an important role in DNA damage repair and other essential pathways.

Figures 3.1, 3.5, 3.6, 3.7, 3.8, 3.9, 3.10 and 3.11 were adapted from Gibbs-Seymour *et al.*, 2016.

Figures 3.12 and 3.13 were adapted from Bonfiglio *et al.*, 2017.

3.1 Identification of HPF1

The Ahel group previously reported the identification of the PBZ domain as a PAR binding domain mainly involved in DNA damage repair (Ahel *et al.*, 2008). PBZ has been then identified, by bioinformatics analysis of genome sequencing data, in a multitude of organisms, including uncharacterised proteins that have not been previously linked with DNA repair. However, we reasoned that some of these proteins could still play a role in ADP-ribosylation-dependent DNA repair, even though such involvement has not been previously reported. Moreover, we reasoned that the human orthologues could have lost or never acquired a PBZ domain, but nonetheless retained the link to ADP-ribosylation and DNA repair that the presence of the PBZ domain in a lower organism hinted at. Bioinformatics screens identified an uncharacterised protein in *Drosophila melanogaster* called CG1218, containing a PBZ domain at the N-terminus and an unknown functional domain, annotated as DUF2228, at the C-terminus (**Figure 3.1A**). Following the discovery of CG1218, the PBZ structure was solved by NMR (Isogai *et al.*, 2010), however no information was obtained about CG1218 biological function. The human homolog has been annotated as C4ORF27, which we later named HPF1 for the reasons explained in section 3.6. Human HPF1 is also an uncharacterised protein, which, however, lacks the PBZ domain and consists solely of the unknown functional domain DUF2228. Given this difference in domain organisation, I aimed to characterise the ability of DUF2228 containing proteins to bind PAR-chain using recombinant *Drosophila* (DmHpf1) and human HPF1. After cloning and establishing a good expression and purification protocols in *Escherichia coli*, recombinant proteins were blotted onto a nitrocellulose membrane, left to dry for a few seconds, and then incubated with the purified ³²P-PAR-chain (**Figure 3.1B**). Dried nitrocellulose membrane were then exposed and analysed by autoradiography,

as expected DmHpf1 can efficiently bind the PAR-chain due to the presence of the PBZ domain while Human HPF1 shows no PAR binding propensity.

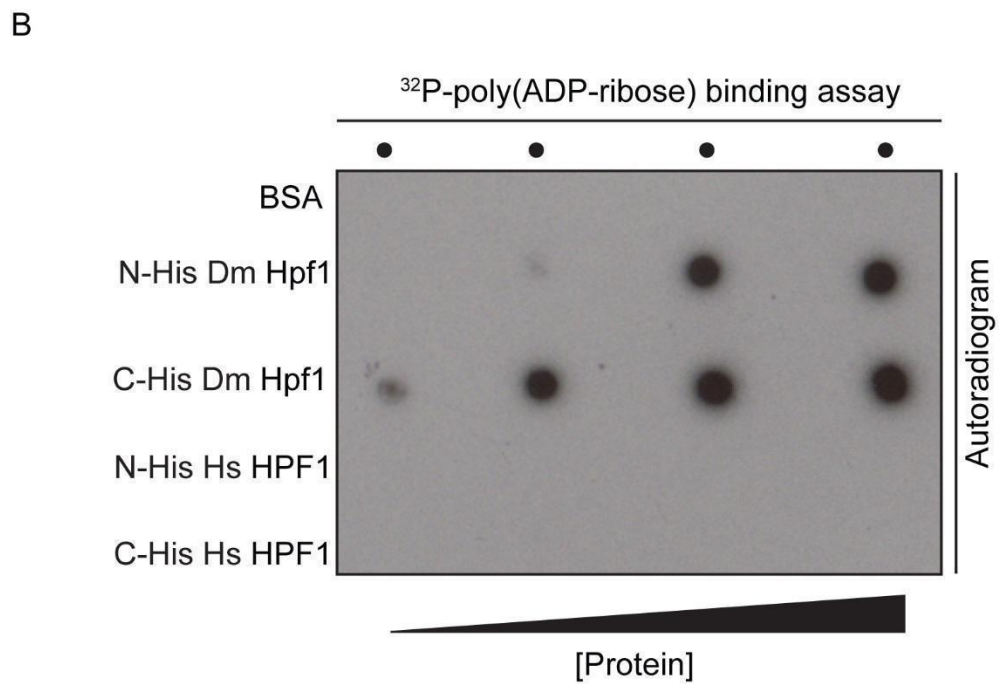
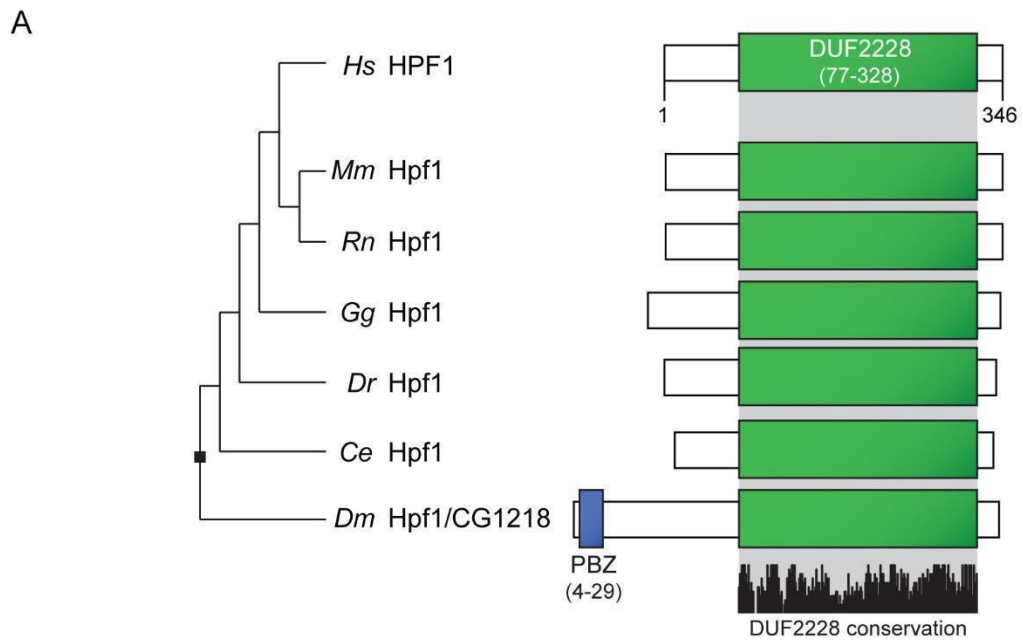


Figure 3.1. **A** - Phylogenetic tree and Schematic representation of HPF1 domain from different organisms generated with Jalview. *Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Gallus gallus*, *Danio rerio*, *Caenorhabditis elegans*, *Drosophila melanogaster*. **B** - Poly-ADP-ribose binding assay.

3.2 HPF1 structural characterisation

Previously, our lab predicted that most if not all proteins containing a PBZ domain are involved in the PARP1-dependent DNA damage response (Ahel *et al.*, 2008). Having identified DmHpf1 and its efficient PAR binding ability, we decided to investigate whether the human HPF1, even though lacking the PBZ domain, played a role in the DNA damage response. Since primary sequence analysis of HPF1 did not yield any conclusions regarding its potential function, I decided to approach the problem by structural characterisation using X-ray crystallography. Previous lab members screened through HPF1 proteins from a number of different species and identified several orthologues that expressed well in *E.coli*. I focused on 2 proteins that purified the best, DmHpf1, *Nematostella vectensis* Hpf1 (NvHpf1) a starlet sea anemone used as model organism for developmental biology and ecology, as well as the human HPF1 homolog which expressed at moderate levels. Approximately 10mg of purified HPF1 could be obtained from an initial 6 litres of culture and a three column purification protocol (HisTrap, gel filtration and anion exchange Q). In order to increase chances of crystallisation, I tried all three different homologs for the structural characterisation, HPF1, DmHpf1 and NvHpf1. Despite the high similarity of the three proteins, the human HPF1 as well as full length DmHpf1 have been reluctant to crystallise (**Table 3.1**). NvHpf1, on the other hand, gave crystals from the first attempt. These initial crystals required several steps of optimisation to improve their diffraction to a final resolution of 2.4 Å (**Table 3.2**). The lack of any structural homologs in the PDB made the phasing by molecular replacement impossible. To solve the phase problem, I opted for an experimental phasing approach using selenomethionine derivative. Selenomethionine-labeled protein expressed and behaved exactly like the na-

tive protein but the crystals diffracted to a final resolution of 2.9 Å. Using the anomalous diffraction from the selenomethionine crystal Dr. Antonio Ariza, a postdoc in the lab, solved the structure of the 2.4 Å native crystal (**Table 3.1**). In order to increase the chance of crystallisation of the DmHpf1, I decided to express a few truncations lacking the PBZ domain, consisting only of the conserved DUF2228 domain. With a single attempt, a truncation of DmHpf1 (DmHpf1 Δ 1-121) crystallized efficiently. Without the need of optimisation, DmHpf1 Δ 1-121 formed crystals that diffracted to a resolution of 2.9 Å and the structure was solved by molecular replacement using the NvHpf1 structure and refined by Dr. Antonio Ariza (**Figure 3.2**). DmHpf1 NvHpf1 shows the same structural organisation with two semi-independent domains linked by an extended loop. The N-terminal domain is organised into four α -helices (α 1- α 2- α 3- α 4), and four β -strands (β 1- β 2- β 3- β 4), while the C-terminus is organised into a five α -helices (α 5- α 6- α 7- α 8- α 9) (**Figure 3.2**). The two Hpf1 structures superimpose almost perfectly with the root-mean-square deviation of 0.94 Å over core 197 C-alpha atoms (**Figure 3.3**). The extended loop (in red) is a highly conserved region among different species and shows two evolutionarily conserved residues, namely Y238 and R239 (human HPF1 residue numbering used), which I later mutated into alanine and identified as essential for HPF1 activity (**Figure 3.4**). NvHpf1 presents two molecules in the asymmetric unit which does not show any major difference between them, while DmHpf1 presents a single molecule in the asymmetric unit. From a preliminary analysis of the structures I was unable to identify any obvious pocket which would suggest the existence of a catalytic site, suggesting the absence of any intrinsic enzymatic activity. Further characterisation was performed by Dr. Antonio Ariza, who checked for secondary structure similarities using DALI and PDBeFold. The NvHpf1 N-terminal domain shows a little structural similarity with a putative glucan synthesis regulator of the SMI1/KNR4-like family from *Bacteroides fragilis* NCTD 9345 (pdb: 3D5P)

(rmsd 2.37 over 89 amino acids aligned and with 8% sequence similarity), which is a SUKH domain. SUKH domains have been shown to mediating protein-protein interactions (Zhang *et al.*, 2011). The NvHpf1 C-terminal domain shows a little structural similarity with the CID-domain of human RPRD1D (rmsd 3.44 over 91 amino acids and with 8% sequence similarity), which acts as a scaffold mediating interaction of different proteins during RNA polymerase II Ser5 dephosphorylation (Ni *et al.*, 2014). These results suggest a role of HPF1 as a scaffold protein that may mediate interactions between two or more different proteins.

Protein name	Expression system	Note	Protein Buffer	mg/mL	drop size a/ drop size b	drop a/drop b	SeMe	Crystal Buffer
NHis-DmHPF1 FL	E. coli Rosetta DE3		25mM Tris pH8, 100mM NaCl, 1mM DTT	12	100+100/100+100	apo/1mM ADPr		
CHis-DmHPF1 FL	E. coli Rosetta DE3		25mM Tris pH8, 100mM NaCl, 1mM DTT	13	100+100/100+100	apo/1mM ADPr		
CHis-NvHPF1 FL	E. coli Rosetta DE3		25mM Tris pH8, 100mM NaCl, 1mM DTT	12.6	100+100/100+100	apo/1mM ADPr		
CHis-NvHPF1 FL	E. coli Rosetta DE3	Phased	25mM Tris pH8, 100mM NaCl, 1mM DTT	12.6	125+125/100+125	apo/seed 25nL	Yes	0.2M Imidazole malate pH 5.5, 15% v/v PEG 600
CHis-NvHPF1 FL	E. coli Rosetta DE3		25mM Tris pH8, 100mM NaCl, 1mM DTT	13.5	125+125/100+125	apo/seed 25nL		
NHis-HsHPF1 FL	Insect cells		25mM Tris pH8, 100mM NaCl, 1mM DTT	12.2	125+125	apo		
NHis-HsHPF1 FL	Insect cells		25mM Tris pH8, 100mM NaCl, 1mM DTT	12.2	125+125/125+125	apo unbound fraction/apo bound fraction		
NHis-HsHPF1 FL	E. coli Rosetta DE3		25mM Tris pH8, 100mM NaCl, 1mM DTT	13.7/18	150+150/150+150	apo 13.7 mg/mL / apo 18 mg/mL		
CHis-HsHPF1 FL	E. coli Rosetta DE3		25mM Tris pH8, 100mM NaCl, 1mM DTT	13.7/18	150+150/150+150	apo 13.7 mg/mL / apo 18 mg/mL		
NHis-HsHPF1 FL	E. coli Rosetta DE3		10mM MES pH5.5, 100mM NaCl, 1mM DTT	8.5	125+125	apo		
NHis-HsHPF1 FL	E. coli Rosetta DE3		10mM ADA pH6, 100mM NaCl, 1mM DTT	11.5	125+125	apo		
NHis-HsHPF1 FL	E. coli Rosetta DE3		10mM AMP SO pH6.3, 100mM NaCl, 1mM DTT	10.7	125+125	apo		
NHis-HsHPF1 FL	E. coli Rosetta DE3		10mM CHES pH8.6, 100mM NaCl, 1mM DTT	10.7	125+125	apo		
NHis-DmHPF1 FL	E. coli Rosetta DE3		10mM Tris pH8, 100mM NaCl, 0.1mM TCEP	12.7	180+180/90+270	apo/apo		0.1M Tris-HCl pH8, 2M ammonium sulphate
NHis-DmHPF1 Δ1-121	E. coli Rosetta DE3	2.5Å	10mM Tris pH8, 100mM NaCl, 0.1mM TCEP	12.7	180+180/90+270	apo/apo		0.2M Imidazole malate pH 5.5, 15% v/v PEG 600
CHis-NvHPF1 FL	E. coli Rosetta DE3	2.4Å	10mM Tris pH8, 100mM NaCl, 0.1mM TCEP	9	200+200/100+300	H3 pep. 1-21 2.1mM/H3 pep. 1-21 2.1mM PARP1 Dloop pep. 2.5mM/PARP1 Dloop pep. 2.5mM		
CHis-NvHPF1 FL	E. coli Rosetta DE3		10mM Tris pH8, 100mM NaCl, 0.1mM TCEP	10	200+200/100+300	HPF1 YR loop peptide 2.5mM/HPF1 YR loop peptide 2.5mM		
PARP1-CAT domain	E. coli Rosetta DE3		10mM Hepes, 150mM NaCl, 0.1mM TCEP	10	200+200/100+300	HPF1 YR loop peptide 2.5mM/HPF1 YR loop peptide 2.5mM		
PARP2-CAT domain	E. coli Rosetta DE3		10mM Hepes, 150mM NaCl, 0.1mM TCEP	10	200+200/100+300	HPF1 YR loop peptide 2.5mM/HPF1 YR loop peptide 2.5mM		
NHis-HsHPF1 FL + NHis-PARP2 FL	E. coli Rosetta DE3		10mM Hepes, 150mM NaCl, 0.1mM TCEP	4.5+7.5	125+125/125+375	apo/apo		
NHis-HsHPF1 FL + NHis-PARP2 FL	E. coli Rosetta DE3		10mM Hepes, 150mM NaCl, 0.1mM TCEP	4.5+7.5	125+125/125+375	62uM 26dsDNA + 100uM ADP/62uM 26dsDNA + 100uM ADP		
NHis-HsHPF1 FL + NHis-PARP2 FL	E. coli Rosetta DE3		10mM Hepes, 150mM NaCl, 0.1mM TCEP	8	125+125/125+375	apo/apo		

Table 3.1. HPF1s crystallisation attempts.

	Se-Met NvHpf1	NvHpf1	DmHpf1Δ1-121
Data Collection			
Synchrotron/beamline	DLS/I03	DLS/I03	DLS/I024
Wavelength (Å)	0.9796 Å	0.9795 Å	0.9796 Å
Data Processing			
Resolution range (Å)	45.31 - 2.801 (2.902 - 2.801)	72.96 - 2.4 (2.486 - 2.4)	52.82 - 2.95 (3.055 - 2.95)
Space group	R 3 :H	R 3 :H	P 65 2 2
Unit cell (Å, °)	88.951 88.951 224.127 90 90 120	89.1419 89.1419 223.261 90 90 120	60.997 60.997 448.924 90 90 120
Total reflections	165989 (13848)	258048 (23233)	234885 (17269)
Unique reflections	16211 (1586)	25881 (2579)	11515 (1121)
Multiplicity	10.2 (8.7)	10.0 (9.0)	20.4 (15.4)
Completeness (%)	99.61 (98.39)	98.45 (95.19)	99.95 (100.00)
Mean I/sigma(I)	16.91 (2.39)	28.37 (1.36)	7.84 (1.50)
Wilson B-factor	70.52	46.85	44.61
R-merge	0.121 (0.7806)	0.3317 (1.411)	0.3322 (0.8484)
R-meas	0.1274 (0.83)	0.3495 (1.498)	0.3402 (0.8775)
R-pim	0.03943 (0.2785)	0.1093 (0.4982)	0.07088 (0.2163)
CC1/2	0.998 (0.788)	0.948 (0.262)	0.994 (0.621)
CC*	0.999 (0.939)	0.986 (0.644)	0.998 (0.875)
Refinement			
Reflections used in refinement	16206 (1585)	25483 (2455)	11510 (1121)
Reflections used for R-free	755 (111)	1305 (130)	562 (51)
R-work	0.2523 (0.3374)	0.2728 (0.3905)	0.3325 (0.4189)
R-free	0.3125 (0.4794)	0.3201 (0.3652)	0.3162 (0.4553)
CC(work)	0.923 (0.728)	0.868 (0.496)	0.884 (0.306)
CC(free)	0.874 (0.434)	0.818 (0.173)	0.884 (0.633)
Number of non-hydrogen atoms	4695	4695	2249
macromolecules	4695	4695	2247
Protein residues	589	589	281
RMSD bonds-length (Å)	0.009	0.014	0.009
RMSD angles (°)	1.37	1.82	1.52
Ramachandran favored (%)	92.4	92.92	80.74
Ramachandran allowed (%)	6.39	5.7	15.19
Ramachandran outliers (%)	1.21	1.38	4.07
Rotamer outliers (%)	6.96	13.15	16.02
Clashscore	2.99	11.62	10.09
Average B-factor	57.44	34.55	44.25
macromolecules	57.44	34.55	44.29

Table 3.2. Data Collection, phasing, and refinement statistics.

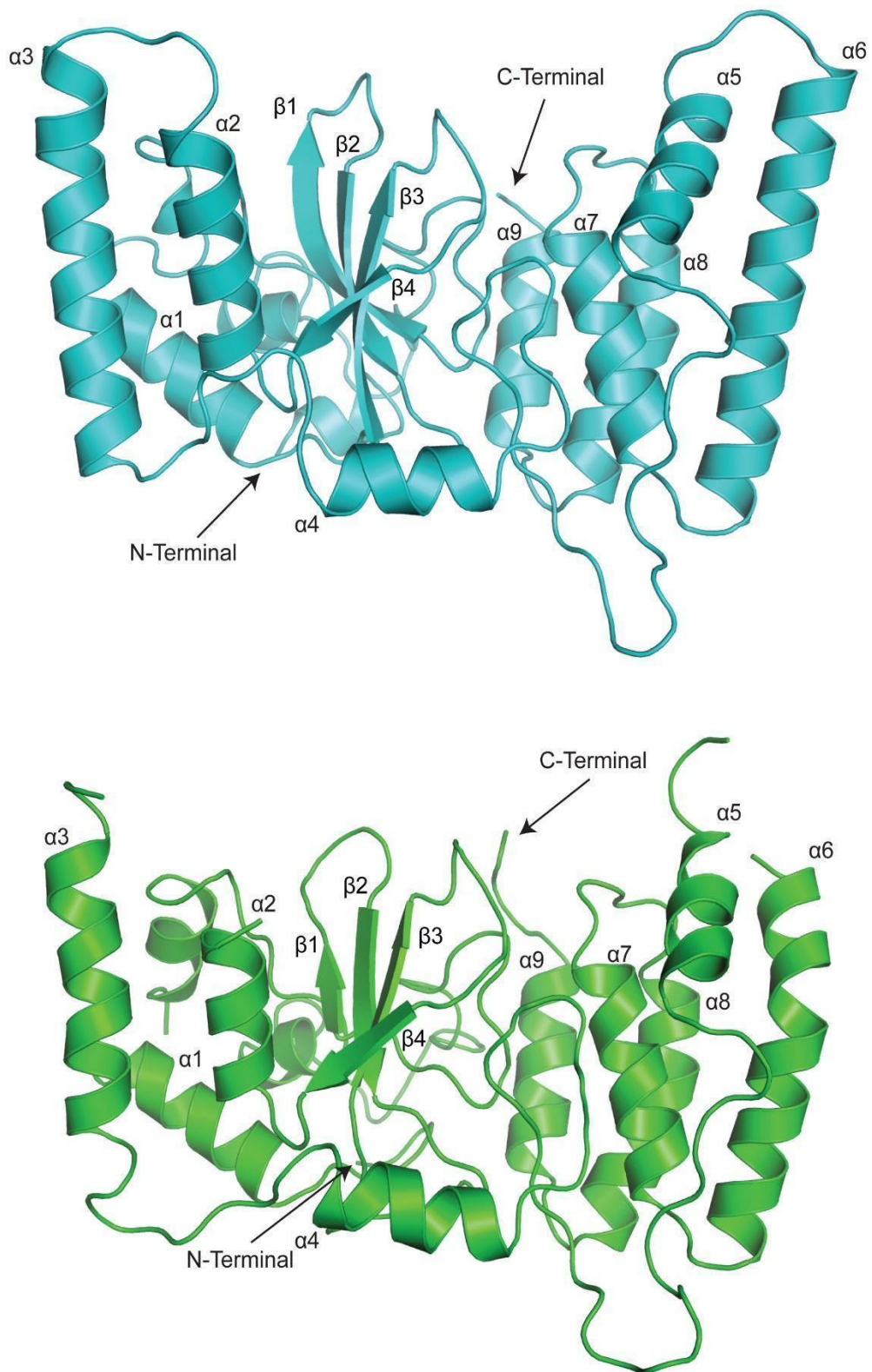


Figure 3.2 NvHpf1 structure at 2.4Å in cyan, DmHpf1Δ1-121 structure at 2.9Å in green.

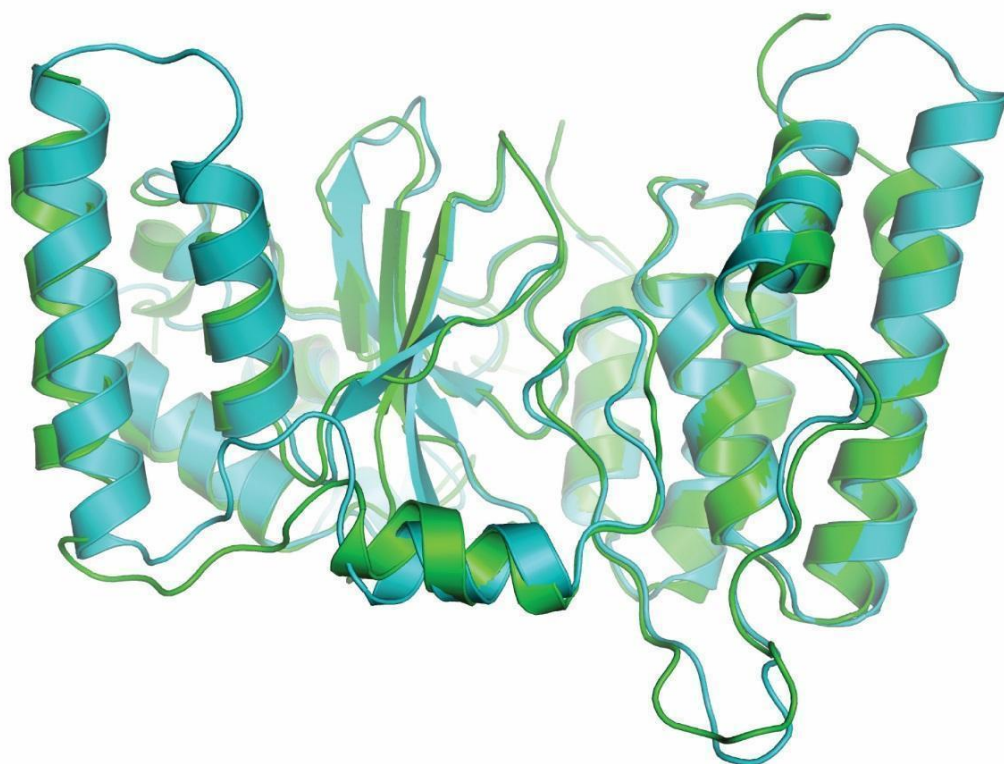


Figure 3.3. NvHpf1 (cyan) and DmHpf1Δ1-121 (green) structures superimposed.

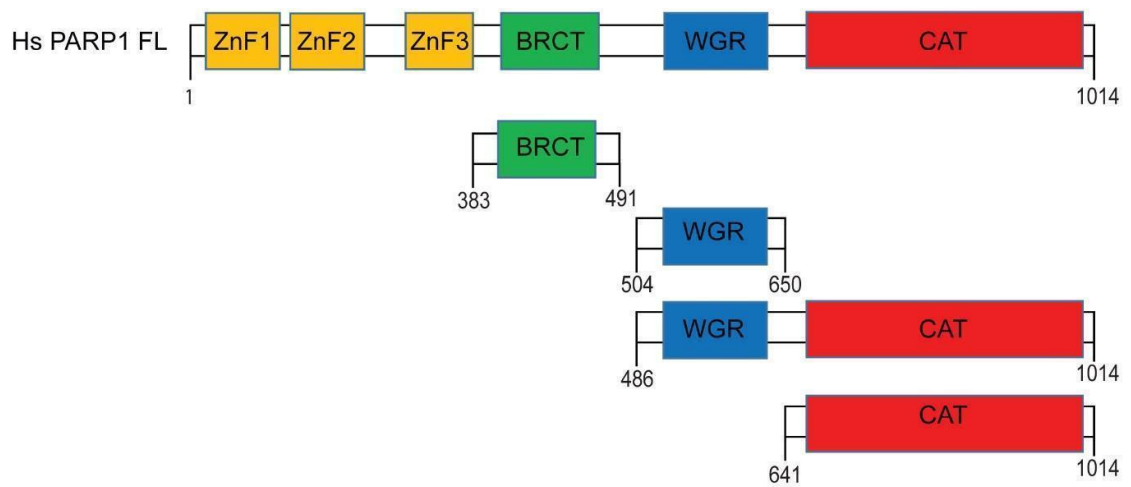
3.3 Functional characterisation of HPF1

While I was working on the structural characterisation, Dr. Gibbs Seymour, a post-doc in the lab, generated human U2-O-S and 293T HPF1^{-/-} cells using CRISPR/Cas9 technology. He transfected these cell lines with YFP-HPF1 and assessed its recruitment at the laser-induced DNA damage. YFP-HPF1 was recruited after a few seconds of laser damage. Moreover, using PARP1 inhibitor or PARP1-KO cells he showed that YFP-HPF1 can localise at the DNA damage site in a PARP1-dependent way, but independently from PARP1 activity, suggesting a recruitment ADP-ribosylation independent. Finally, using mass spectrometry of immunoprecipitated FLAG-HPF1, he identified PARP-1 and the core histones as HPF1 interactors, which were further confirmed by immunoblotting (Gibbs-Seymour *et al.*, 2016). Despite many other interactors, such as BAF180, XRCC1, SMARCA5 as well as PARP2 were also identified, we decided to focus our characterisation on PARP1 and histones as statistically more enriched than any other interactors. Strikingly, the cell based data using the HPF1-KO cells showed that HPF1 is critical for histone ADP-ribosylation (Gibbs-Seymour *et al.*, 2016).

3.4 HPF1 interacts with PARP1 and PARP2 via catalytic domain and with recombinant histones

After the structural characterisation, and in the light of the above mentioned functional characterisation, which show that HPF1 can immunoprecipitate PARP1, PARP2 and core histones and its recruitment on the damaged DNA is PAR independent, I set out to confirm the ability of HPF1 to interact directly with PARP1 and histones. Individual PARP1 domains (**Figure 3.5A**) were cloned into a PGex vector with an N-Term GST tag, expressed in *E.coli* and purified with glutathione resin. GST-PARP1 domain fusions were then used for a GST pull-down assay together with His-HPF1. The assay showed clearly that HPF1 can specifically be pulled-down by the catalytic (CAT) domain of PARP1 (**Figure 3.5B**). Moreover, the same result was obtained using recombinant GST-PARP-2 CAT domain (**Figure 3.6**). This assay, together with the functional characterisation, suggests that HPF1 can directly interact with the catalytic domain of either PARP1 or PARP2. Further pull-down analysis using GST-HPF1 and recombinant histone tetramer (H3/H4) showed the efficient ability of HPF1 to pull down the histone tetramer, confirming the direct interaction (**Figure 3.7**). These assays, together with the structural analysis, suggest that HPF1 could interact with both PARP1 and histones, probably acting as a linker between the two.

A



B

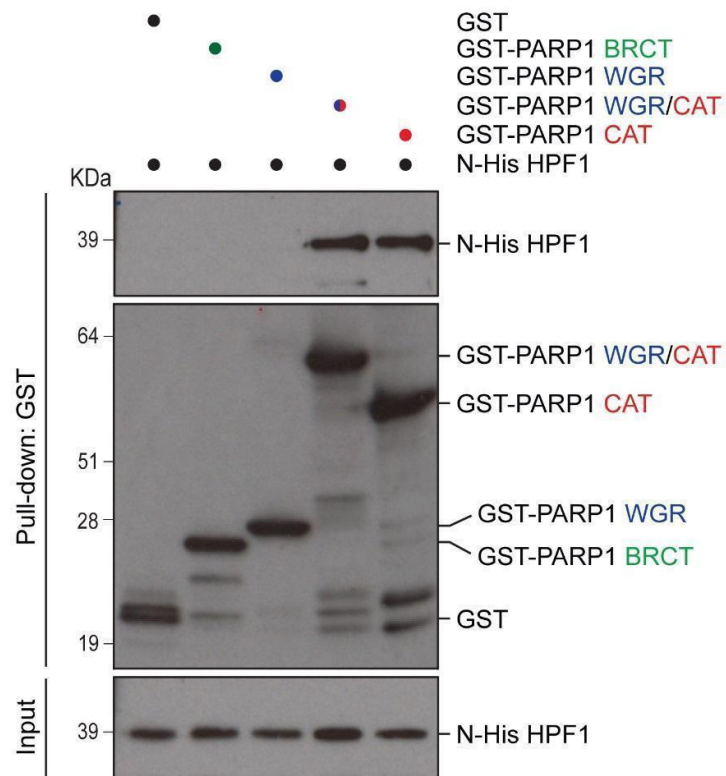


Figure 3.5. A - Schematic representation of human PARP1 and PARP1 truncations. **B** - GST pull-down assay, GST-PARP1 and GST-PARP1 fragments were bound to glutathione resin before addition of human N-His HPF1. Samples were analysed by immunoblotting with GST and His antibody.

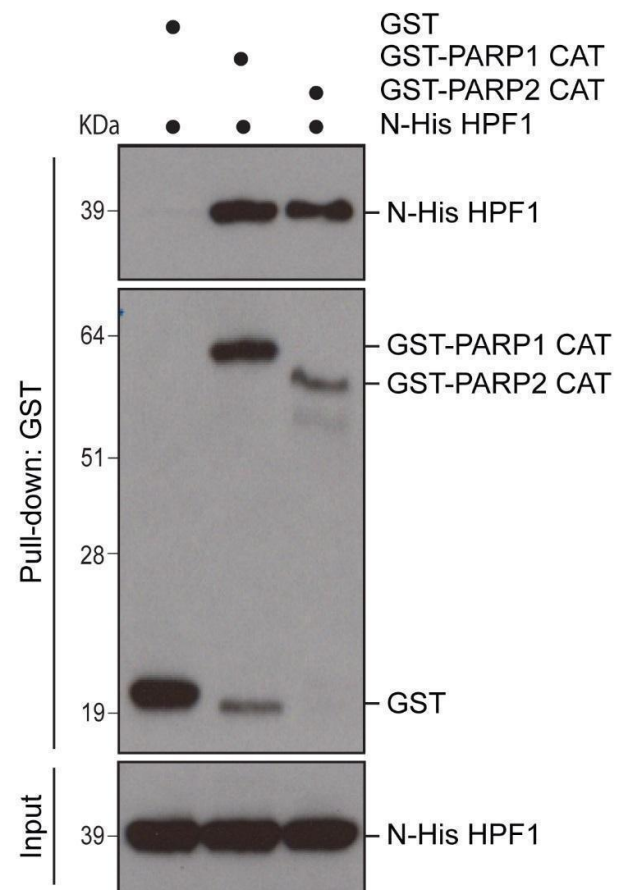


Figure 3.6. GST pull-down assay, GST-PARP1-CAT and GST-PARP2-CAT were bound to glutathione resin before addition of human N-His HPF1. Samples were analysed by immunoblotting with GST and His antibody.

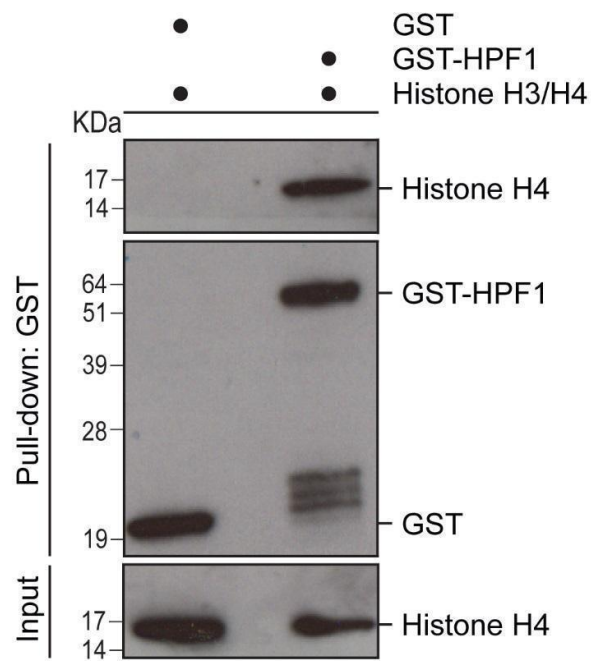


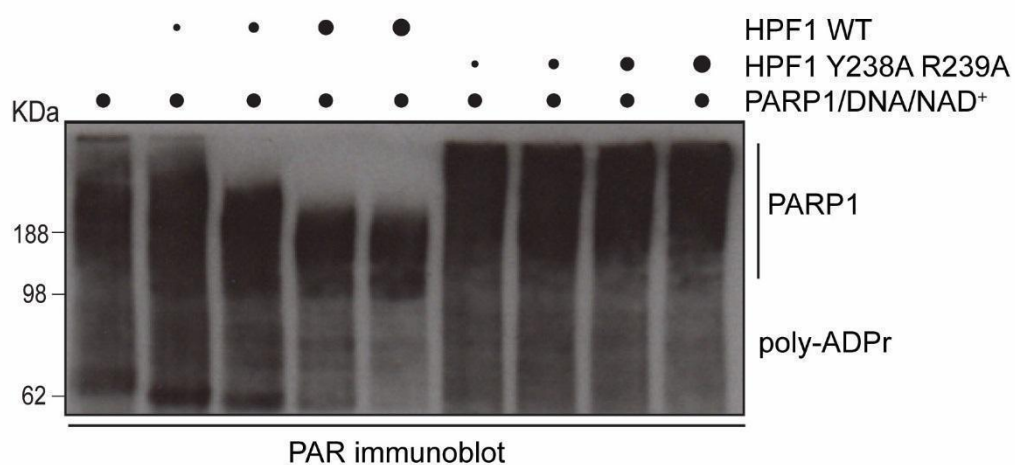
Figure 3.7. GST pull-down assay, GST and GST-HPF1 were bound to glutathione resin before addition of Histone tetramer H3/H4. Samples were analysed by immunoblotting with GST and H4 histone antibody.

3.5 HPF1 suppresses PARP1 auto-modification

Previous data showed that in HPF1-KO cells histone ADP-ribosylation is virtually lost, strongly suggesting that HPF1 may be a critical factor that enables histone PARylation in cells (Gibbs-Seymour *et al.*, 2016). Also, these studies showed that average length of chains on auto-modified PARP1 is increased in HPF1-KO cells. In order to confirm these observations, I aimed to investigate the effect of HPF1 on PARP1 activity *in vitro* using recombinant proteins. I investigated the effect of such modification by PARP1 in the presence or absence of HPF1 and later using various histone substrates (see the next section). Firstly, I investigated the effect of HPF1 on the auto-modification of PARP1. In the presence of activated DNA and NAD⁺, PARP1 auto-modified itself giving a shift in size which can be visualised by SDS-PAGE and immunoblot using anti-PAR antibody (**Figure 3.8A**). This assay clearly showed that in presence of HPF1 the size shift of PARP1 decreases with increasing HPF1 concentration. This result suggests that HPF1 can suppress PARP1 auto-modification. Moreover, the HPF1 mutant Y238A R239A (identified as deficient in PARP1 binding—Gibbs Seymour *et al.*, 2016) showed no ability to suppress PARP1 auto-modification. This assay left the open question whether HPF1 could act as a glycohydrolase removing the PAR-chain similarly to PARG. However, as previously mentioned, the structural characterisation suggested that HPF1 is not an enzyme so its role in reducing PARP1 auto-modification was likely not due to a glycohydrolase activity. In order to confirm this hypothesis, I performed a glycohydrolase assay. In this assay PARP1 was auto-modified in the presence of DNA, NAD⁺, ³²P-NAD⁺ for labelling and in the absence of HPF1. The reaction was stopped with PARP1 inhibitor Olaparib and subsequently HPF1 or the positive control PARG were added (**Figure 3.8B**). As the assay clearly shows, the glycohydrolase PARG can efficiently remove PARP1 auto-

modification, while HPF1 has no effect on it. This suggested that HPF1 is not a glycohydrolase and the decrease of PARP1 auto-modification in the presence of HPF1 is a regulatory mechanism due to an interaction with HPF1 prior to PARP1 auto-modification.

A



B

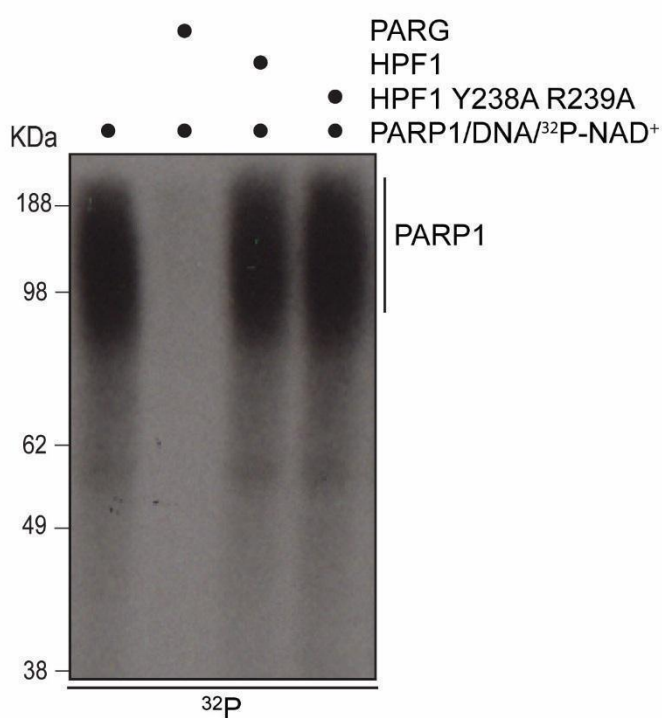


Figure 3.8. A - Increasing amount of recombinant HPF1 was incubated with recombinant PAPR1 in presence of NAD⁺ and DNA. Product of the reactions were analysed by SDS-PAGE and immunoblotting with anti-poly-ADP-ribose antibody. **B** - PARP1 was auto-modified in presence of DNA and ³²P-NAD⁺ to produce PARP1-poly-ADP-ribosylated. The reaction was stopped with addition of PARP1 inhibitor Olaparib before addition of HPF1 and PARG. The reactions were analysed by SDS-PAGE and autoradiography.

3.6 HPF1 promotes PARP1 trans-modification of histones

Next I investigated whether HPF1 has also some effect on the PARP1-dependent ADP-ribosylation of recombinant histones. In this assay, PARP1 was incubated with activated DNA, NAD^+ , ^{32}P - NAD^+ , histone octamer as a substrate and in the presence or absence of HPF1 (**Figure 3.9**). Satisfyingly, the presence of HPF1 in the reaction resulted in a dramatic increase of the PARP1-dependent ADP-ribosylation of the recombinant histone octamer, confirming that HPF1 indeed promotes histone ADP-ribosylation. The presence of histone protein in the reaction determined also a decreased auto-modification of PARP1, likely due to competition with PARP1 as substrate for ADP-ribosylation. APLF, which is not predicted to be a native substrate of PARP1-HPF1, was used as a negative control. The assay showed that HPF1 is indeed unable to promote ADP-ribosylation of APLF, suggesting that HPF1 is a specificity factor for histone ADP-ribosylation. For this reason, we named it Histone Parylation Factor 1 (HPF1). I further show that HPF1 can promote ADP-ribosylation of histone tetramer (**Figure 3.10**) and histone H1 (**Figure 3.11**). Importantly, I could show that HPF1 activity is dependent on the two conserved residues Y238 R239 and that HPF1 Δ 242-346 cannot promote HPF1 activity suggesting that the presence of C-terminal domain is indeed essential. Altogether, my data show that HPF1 suppresses PARP1 auto-modification while promoting histone trans-ADP-ribosylation by PARP1.

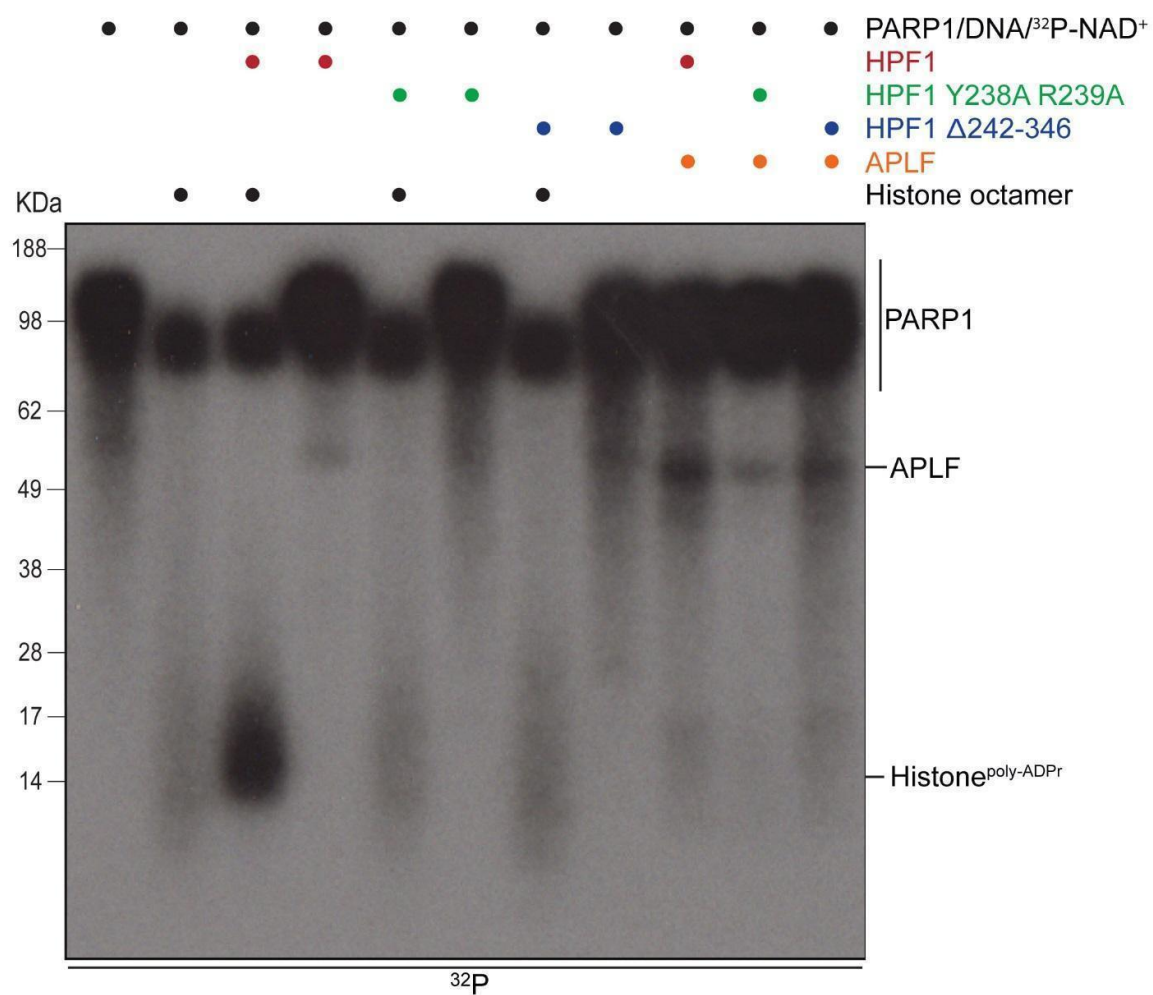


Figure 3.9. PARP1 was incubated in presence of DNA, ^{32}P -NAD $^{+}$ and HPF1. Recombinant histone octamer was used as a substrate for trans-ADP-ribosylation. The reactions were analysed by SDS-PAGE and autoradiography.

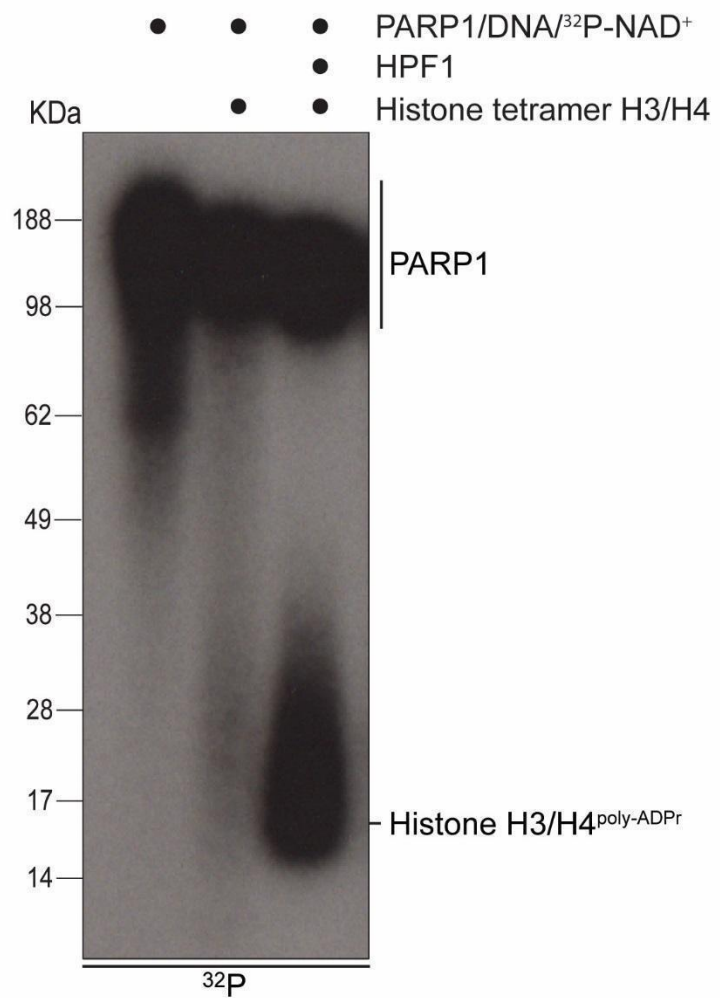


Figure 3.10. PARP1 was incubated in presence of DNA, ³²P-NAD⁺ and HPF1. Recombinant histone tetramer was used as a substrate for trans-ADP-ribosylation. The reactions were analysed by SDS-PAGE and autoradiography.

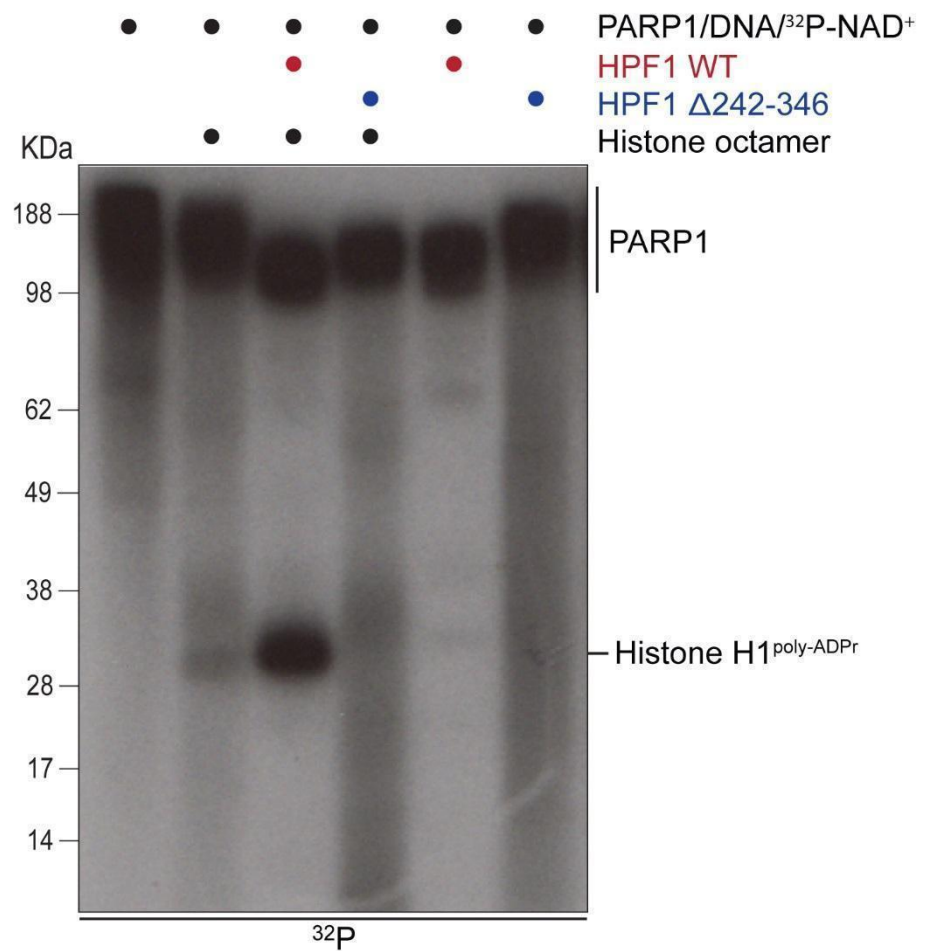


Figure 3.11. PARP1 was incubated in presence of DNA, ³²P-NAD⁺ and HPF1. Recombinant histone H1 was used as a substrate for trans-ADP-ribosylation. The reactions were analysed by SDS-PAGE and autoradiography.

3.7 HPF1 regulates PARP1 and PARP2 activity

Having identified the main role of HPF1 as a factor promoting ADP-ribosylation of histone proteins, I decided to investigate whether HPF1 could regulate also other PARPs involved in DNA repair such as PARP2 and PARP3. I performed a histone PARylation assay with PARP1, PARP2, or PARP3, and the histone tetramer as a substrate. The reaction was carried out in the presence or absence of HPF1 (**Figure 3.12**). The assay showed that HPF1 can promote ADP-ribosylation of histones in a PARP1- and PARP2-dependent manner, but has no effect on PARP3 activity. Sequence analysis of PARP1, PARP2 and PARP3 shows the presence of a conserved loop, named D-loop, in the catalytic domain of PARP1, PARP2 and PARP3. The D-Loop PARP1 and PAR2 are nearly identical while the D-Loop of PARP3 is relatively smaller, suggesting it as a possible specific interaction region for HPF1. However, such hypothesis remains to be confirmed.

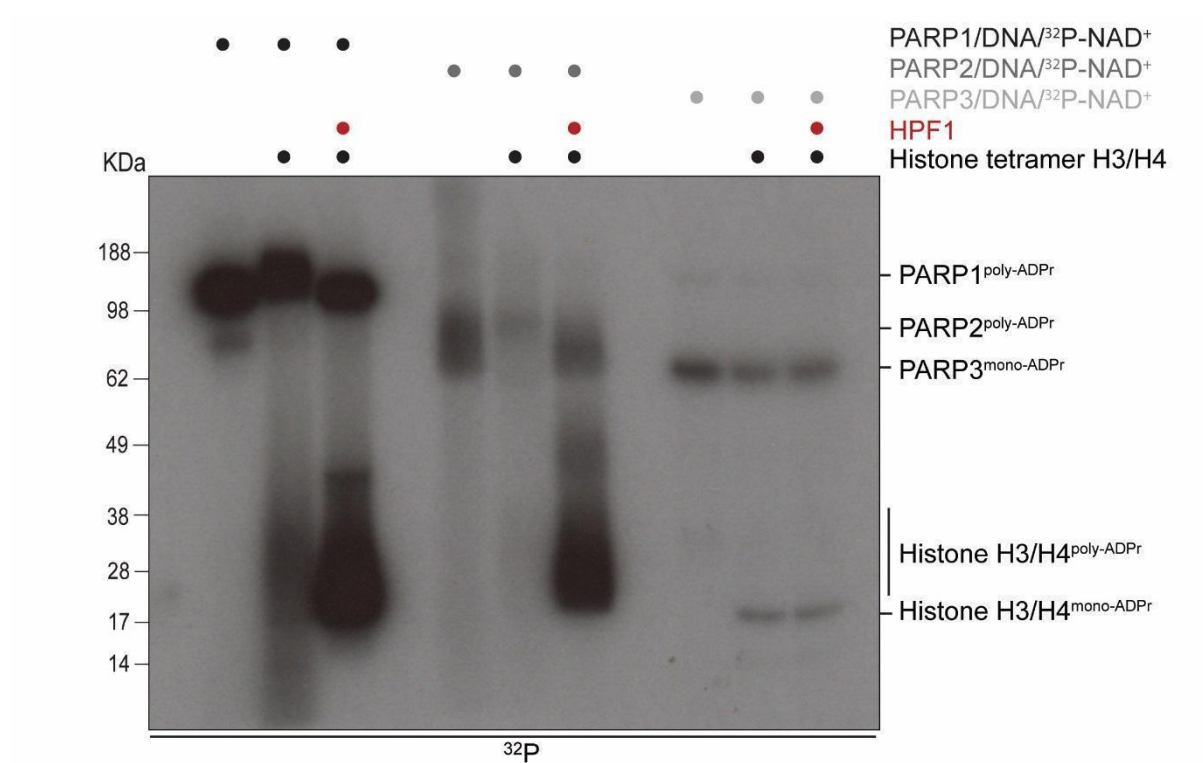


Figure 3.12. PARP1, PARP2 and PARP3 were incubated in presence of DNA, ^{32}P -NAD $^{+}$ and HPF1. Recombinant histone tetramer was used as substrate for trans-ADP-ribosylation. The reactions were analysed by SDS-PAGE and autoradiography.

3.8 Serine ADP-ribosylation specificity

Previously our collaborators, the Ivan Matic group, discovered a novel type of ADP-ribosylation named serine-ADP-ribosylation on histone protein. Implementing a novel proteomic approach they identified serine as a novel substrate for ADP-ribosylation (Leidecker *et al.*, 2016). Having shown that HPF1 affects PARP1 activity, we hypothesised that HPF1 may also steer PARP1 specificity towards Histone residues. To that end, we sent the product of the previous reaction to our collaborator the Matic group (MPI Cologne) who analysed it by mass-spectrometry. To our big surprise, we found that the presence of HPF1 resulted in a switch in the specificity of the PARP1- and PARP2-dependent modification from glutamate/aspartate to serine. Strikingly, we managed to reproduce the exact sites on core histones *in vitro* using just recombinant proteins or histone peptide (Bonfiglio *et al.*, 2017). This raised the question whether Ser-ADP-ribosylation was a specific modification of histones or could be performed on other substrates. I then decided to investigate whether PARP1 auto-modification could also be switched from glutamate/aspartate to serine in a HPF1-dependent way (without the presence of HPF1, PARP1 automodifies on several glutamate residues in this region). Our collaborators identified, by mass spectrometry, 3 serine in the BRCT domain of PARP1 which can be auto-modified. I expressed and purified the auto-modification domain of PARP1 (BRCT, amino acid 374-525) in WT and two mutants (6E/A and 4S/A). I incubated this BRCT WT and mutants with full-length PARP1 in the presence or absence of HPF1 (**Figure 3.13**). The assay showed that the modification of the BRCT domain is HPF1-dependent and the mutation of the serine to alanine abrogated the modification of the BRCT domain. This suggests that the serine present in the BRCT domain are actually the major target of the PARP1 auto-modification in the presence of HPF1. Further, I decided to perform this assay using

just recombinant full-length PARP1 in either WT, 4S/A and 6E/A version in the presence or absence of HPF1 (**Figure 3.14**). The assay showed that the HPF1-dependent auto-modification of the 4S/A mutant was lower compared to the WT, confirming the finding of the previous assay. Differently from the previous assay, however, the auto-modification of PARP1 4S/A does not disappear completely in the presence of HPF1. This effect is likely dependent on the presence of other sites of modification on PARP1, beyond the BRCT domain, which are HPF1 independent.

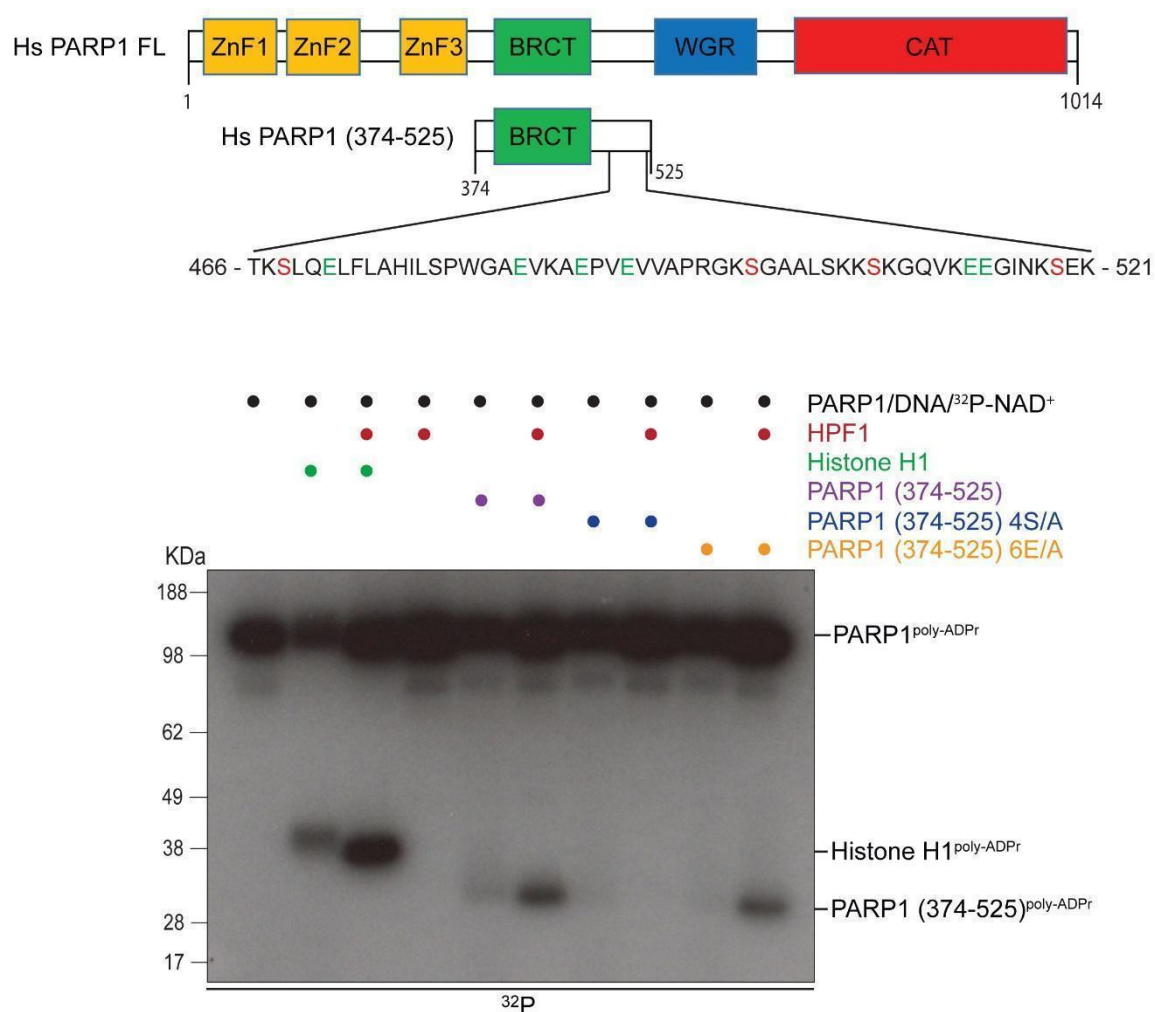


Figure 3.13. Schematic representation of PARP1 with mutated serine highlighted in red and mutated glutamate highlighted in green. PARP1 was incubated in presence of DNA, ³²P-NAD⁺ and HPF1. Recombinant histone H1, BRCT WT, BRCT 4S/A (S468A, S499A, S507A, S519A) and BRCT 6E/A (E471A, E484A, E488A, E491A, E513A, E514A) were used as substrates for trans-ADP-ribosylation. The reactions were analysed by SDS-PAGE and autoradiography. Mutagenesis on the single BRCT domain were performed by Dr. Roko Zaja and Dr. Edward Bartlett.

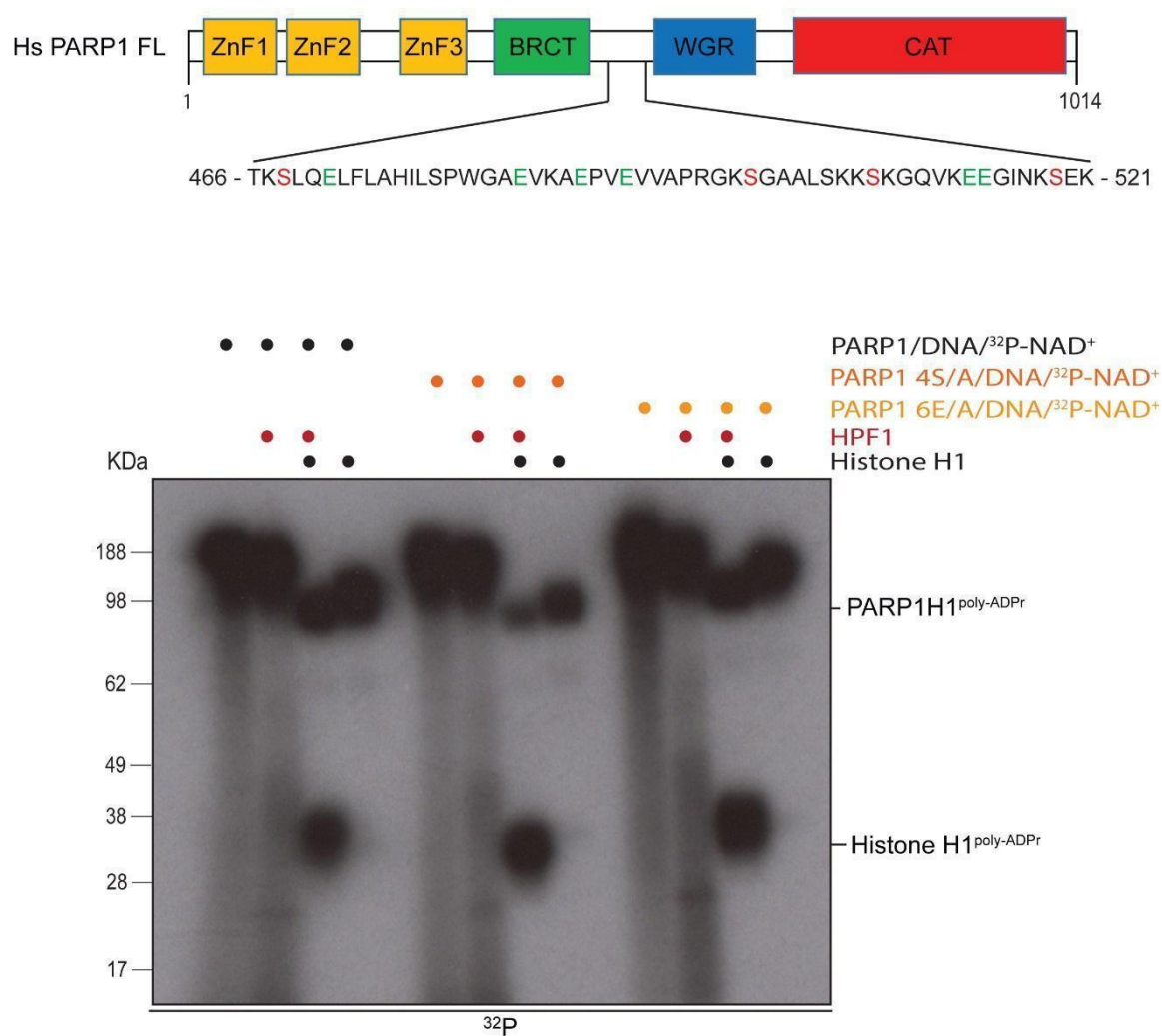


Figure 3.14. Schematic representation of PARP1 with mutated serine highlighted in red and mutated glutamate highlighted in green. PARP1 WT, PARP1 4S/A (S468A, S499A, S507A, S519A) and PARP1 6E/A (E471A, E484A, E488A, E491A, E513A, E514A) were incubated in presence of DNA, ^{32} P-NAD $^{+}$ and HPF1. Recombinant histone H1 was used as substrate for trans-ADP-ribosylation. The reactions were analysed by SDS-PAGE and autoradiography. Mutagenesis in the full-length PARP1 were performed with help of Dr. Edward Bartlett.

3.9 Minimal PARP1 region essential for Ser-ADP-ribosylation

In order to identify the minimal region of PARP1 necessary for Ser-ADP-ribosylation in complex with HPF1, I cloned and expressed several PARP1 truncations and tested their ability to modify histone proteins in the presence of HPF1. In **Figure 3.15** we can see two progressive truncations of PARP1 from the N-terminus, PARP1 Δ ZnF (lacking the three zinc fingers), PARP1 Δ ZnF-BRCT (lacking the three zinc fingers and the BRCT domain). When I tested their ability to modify histone proteins in the presence of HPF1 compared with full-length PARP1, I could clearly see a significant reduction in histone modification in absence of the three ZnF (PARP1 Δ ZnF). These observations suggest that ZnF1, ZnF2 and ZnF3, or at the least some of them, are indeed essential for Ser-AD-ribosylation.

Being PARP1 a DNA dependent enzyme, I decided to investigate whether the DNA is a necessity component even in the presence of HPF1-mediated serine specificity. Dawicki-McKenna *et al.*, 2015 showed that PARP1 Δ HDL lacking the HD domain (an inhibitory part of the catalytic domain) is active even in the absence of DNA. I decided to test PARP1 Δ HDL for Ser-ADP-ribosylation in the presence of HPF1, histone proteins as substrates, and in the presence (**Figure 3.16A**) or absence (**Figure 3.16B**) of DNA. The assay showed clearly that PARP1 Δ HDL can modify histones in the presence of HPF1, confirming its Ser-ADP-ribosylation activity in the presence of DNA. In the absence of DNA, the full length PARP1 completely lost all its activity in terms of auto-modification or modification of histones, while PARP1 Δ HDL still efficiently auto-modified itself and clearly modified histones in the presence of HPF1, although some decrease in activity has been observed. However, it needs to be noted

that PARP1 WT and PARP1 Δ HD were not purified at the same time and the quality of the preparation was different so they are not directly comparable.

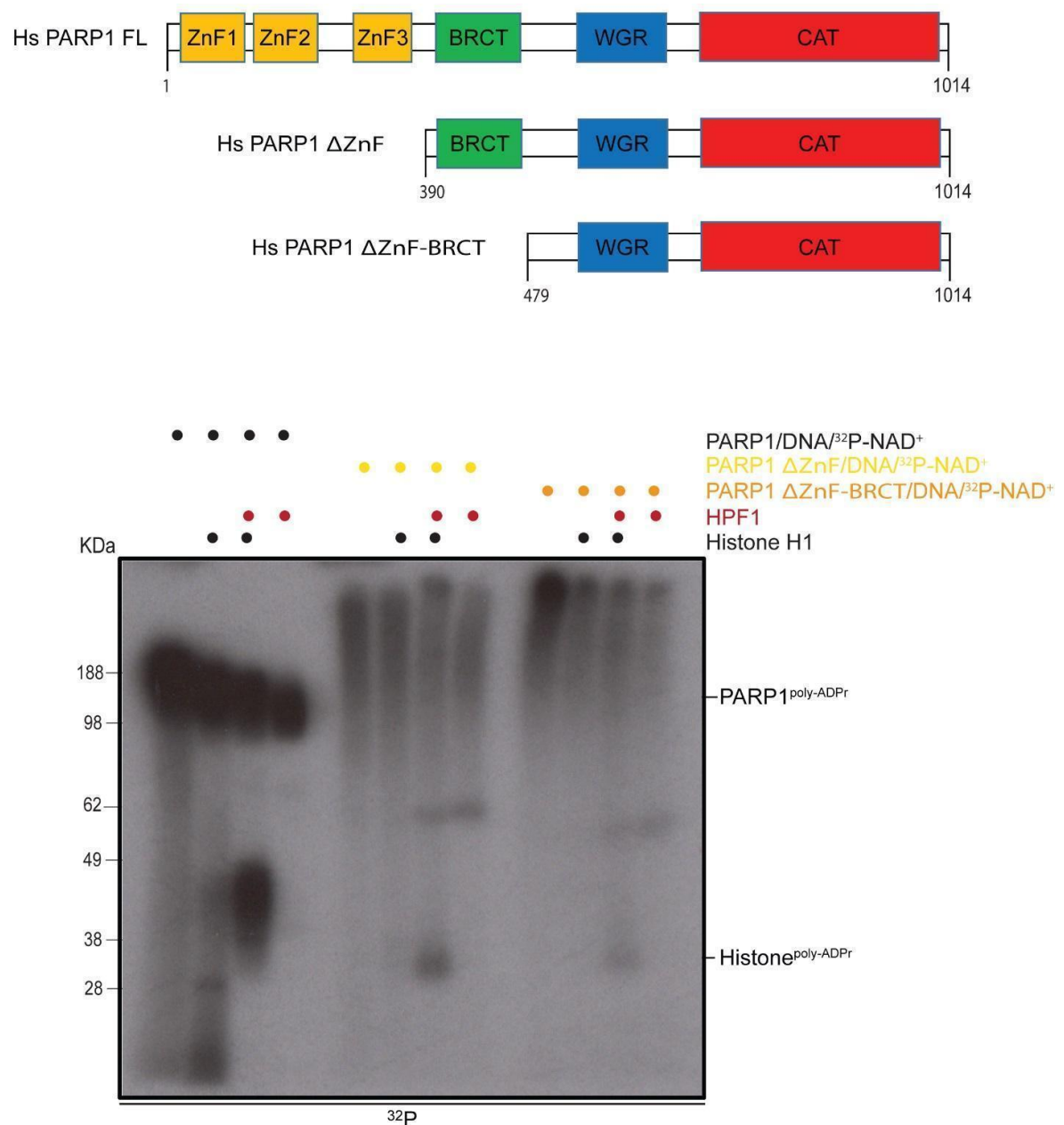


Figure 3.15. PARP1 WT (amino acid 1-1014), PARP1 ΔZnF (amino acid 391-1014) and PARP1 ΔZnF-BRCT (amino acid 480-1014) were incubated in presence of DNA, ³²P-NAD⁺ and N-His HPF1. Recombinant histone H1 was used as a substrate for trans-ADP-ribosylation. The reactions were analysed by SDS-PAGE and autoradiography.

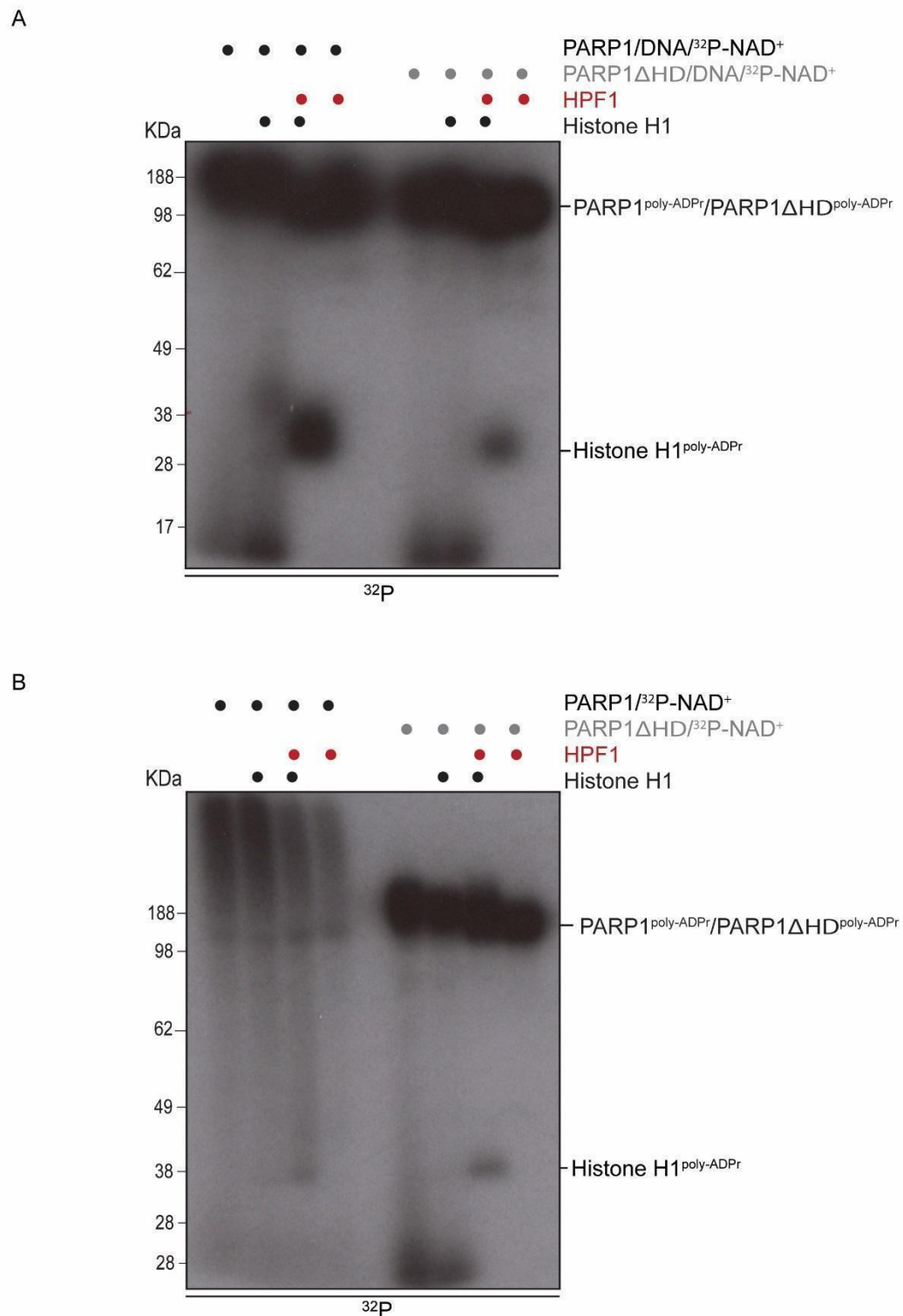


Figure 3.16. A - PARP1 WT, PARP1 Δ HD were incubated in presence ^{32}P -NAD $^{+}$, DNA and HPF1. Recombinant histone H1 was used as substrate for trans-ADP-ribosylation. **B** - Same as previous in absence of DNA. The reactions were analysed by SDS-PAGE and autoradiography. PARP1 Δ HD recombinant protein (Δ 678–787) was a kind gift of Dr. John M. Pascal.

3.10 HPF1 binds ssDNA and dsDNA (EMSA)

In 2009 Ku *et al.*, used a DNA-affinity purification coupled with quantitative proteomics in order to identify protein complex binding to an E-box sequence of HDM2, an oncogene related with cell growth and development in human cells. Ku *et al.*, identified several proteins, among which C4orf27, now known as HPF1. Despite the identification of several proteins, such technique lacks specificity leading to many possible false positive. Hence, HPF1 affinity for DNA remained to be assessed. To confirm HPF1 affinity to DNA, I performed several Electrophoresis Mobility Shift Assays (EMSA) using HMD2 sequence and different HPF1 homologs including human HPF1, DmHpf1 and NvHpf1. As a positive control for the shift, I used a DNA binding protein, namely Aprataxin (Mehrotra *et al.*, 2011). **Figure 3.17** shows the EMSA performed using 24 nucleotides ssDNA and 24 base pairs dsDNA from the DNA sequence of HDM2 E-box. The EMSA showed that both DmHpf1 and NvHpf1 have clear affinity for dsDNA. On the other hand, human HPF1 does not show any obvious affinity for DNA (**Figure 3.17**). This leaves us with the open question whether HPF1 can really bind DNA and in which specific condition or to which specific DNA sequence. Further investigation will be needed to confirm or confute this finding and understand the possible functional implication of such interactions. Different DNA sequences and structures as well as the nucleosome could be used as substrates. Moreover, performing a super-shift with PARP1-DNA/HPF1 and Histones-DNA/HPF1 would give more insight into the formation of such complexes.

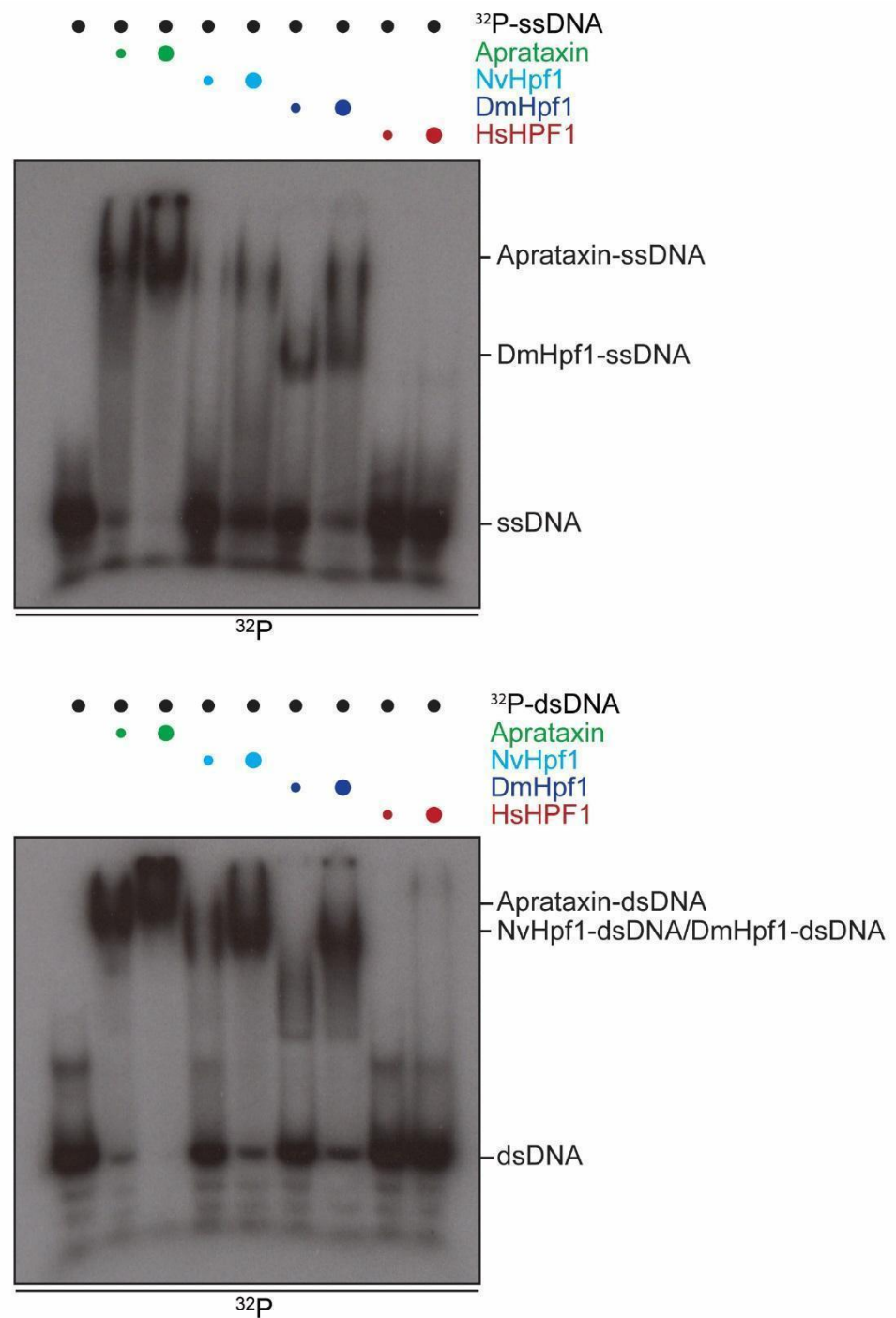


Figure 3.17. EMSA, recombinant Aprataxin, NvHpf1, DmHpf1 and HPF1 were incubated with ^{32}P labelled ssDNA or dsDNA. The reactions were analysed by native gel Electrophoresis and autoradiography.

3.11 Conclusion on HPF1 characterisation

The structure and mechanism of PARP1 ADP-ribosylation has been a heavily explored research topic in recent years (Barkauskaite *et al.*, 2015; Pascal and Ellenberger, 2015). Despite the acquired information, the proper regulation of PARP1 activity has not been deeply understood. In the third chapter of my thesis, a previously uncharacterised protein named HPF1 as a PARP1 and PARP2 interactor protein involved in DNA damage response have been reported and the structures of two different orthologues of Hpf1 from *Drosophila melanogaster* and *Nematostella vectensis* were analysed. This chapter further showed that HPF1 can specifically interact, by *in vitro* pull-down, with the catalytic domain of either PARP1 or PARP2 and recombinant histones. Moreover, it showed that such an interaction acts to repress PARP1 and PARP2 auto-modification and promote histone ADP-ribosylation. Finally, it showed that HPF1 indeed promotes a new type of PTM named Ser-ADP-ribosylation on either recombinant histones or PARP1 itself.

Chapter 4. Ser-ADP-ribosylation reversal by the hydrolase ARH3

Ser-ADP-ribosylation is a recently discovered PTM of proteins involved in genome stability. Previously, our collaborators identified hundreds of proteins as possible targets of Ser-ADP-ribosylation, including both histone and non-histone proteins, making this PTM a widespread modification involved in many biological pathways (Leidecker *et al.*, 2016; Bonfiglio *et al.*, 2017). Here I aimed to investigate the reversal mechanism of Ser-ADP-ribosylation, the novel post-translational modification that we have discovered. Using a biochemical approach, I identified ARH3 as the only human glycohydrolase able to efficiently and completely remove Ser-ADP-ribosylation from histones and other substrates. I performed a biochemical characterisation followed by a set of assays to show the specificity of ARH3 towards Ser-ADP-ribosylation. Finally, I suggested ARH3 as a tool to investigate Ser-ADP-ribosylation without the need of mass spectrometry.

All the figures in this chapter are adapted from Fontana *et al.*, 2017

4.1 Histone Ser-ADP-ribose modification is reversed by the ARH3 hydrolyse

Previous research has continuously reported PARylation as a reversible post-translational modification, and identified several glycohydrolases involved in PAR removal such as PAR glycohydrolase (PARG) which cleave the ribose–ribose bonds between the subunits of the PAR chains (Lin *et al.*, 1997; Slade *et al.*, 2011) or others which are capable of removing the terminal ADP-ribose linked to different amino acid. TARG1, MACROD1 and MACROD2, for instance, remove the terminal ADP-ribose linked to Asp/Glu, while ARH1 that linked to Arg (Feijs *et al.*, 2013; Jankevicius *et al.*, 2013; Rosenthal *et al.*, 2013; Sharifi *et al.*, 2013; Mashimo *et al.*, 2014; Barkauskaite *et al.*, 2015). Since all the previously reported ADP-ribosylation types were found to be reversible, I aimed to investigate whether our newly reported Ser-ADP-ribosylation was also reversible and which enzyme was able to do remove it. I reasoned that one of the previously known glycohydrolase could show such activity, and decided to express and purify all the possible candidate glycohydrolase possessing a DraG-like fold domain or a macrodomain (Glowacki *et al.*, 2002; Oka *et al.*, 2006; Rack *et al.*, 2016), and finally to test their ability to remove Ser-ADP-ribosylation. As an initial substrate, I used the previously reported histone H3 peptide (amino acid 1-21) which does not include any glutamate/aspartate (Bonfiglio *et al.*, 2017). H3 peptide was mono-Ser-ADP-ribosylation by PARP1 in the presence of activated DNA, NAD⁺, ³²P-NAD⁺ and HPF1. After modification, PARP1 activity was stopped with the addition of the inhibitor Olaparib and the mix was incubated with each of the recombinant purified glycohydrolases (**Figure 4.1**). The assay clearly showed that the previously poorly characterised ARH3 was able to remove mono-Ser-ADP-ribosylation from the H3 histone peptide, while all the other enzymes did not show any

discernible activity against this substrate. ARH3 was previously reported as a glycohydrolase able to cleave the ribose–ribose bonds between the subunits of the PAR chains. However, its activity was shown to be less efficient than PARG and the biological role was not completely clear (Mueller-Dieckmann *et al.*, 2006). To further investigate ARH3 activity, I used H1 histone peptide as a substrate of the PARP1/HPF1 complex and a model for poly-Ser-ADP-ribosylation (Bonfiglio *et al.*, 2017). As the assay clearly showed, ARH3 was again able to completely remove poly-Ser-ADP-ribosylation and mono-Ser-ADP-ribosylation from H1 histone, while PARG was just able to remove poly-Ser-ADP-ribosylation leaving a clear band corresponding to H1 mono-Ser-ADP-ribosylation (**Figure 4.2**). These two assays demonstrated clearly for the first time that ARH3 had a completely different function from PARG, despite both being able to cleave the ribose–ribose bonds. In other words, ARH3 is not (or not only) a “backup” for PARG. I further performed a double digestion using PARG/ARH3 showing that the majority of ADP-ribosylation occurs indeed on serine (**Figure 4.3**). In order to characterise ARH3 activity on a more homogeneous substrate, I utilised PARP1 E988Q, a mutant with altered activity allowing just mono-ADP-ribosylation. In this experiment, I modified histone H1 with either the PARP1 WT/HPF1 complex, which produces poly-Ser-ADP-ribosylation on H1, or PARP1 E988Q/HPF1, which generates mono-Ser-ADP-ribosylation on H1. This assay essentially confirmed my previous finding: PARG and ARH3 can both remove poly-Ser-ADP-ribosylation; however, ARH3 alone can remove mono-Ser-ADP-ribosylation (**Figure 4.4**). In order to extend this finding to a non-histone substrate, I investigated the Ser-ADP-ribosylation removal from the auto-modification domain of PARP1. Previously, I reported that the auto-modification domain of PARP1 (amino acid 374-525) could be Ser-ADP-ribosylated by the complex PARP1/HPF1 (Bonfiglio *et al.*, 2017). As in the previous assay, I aimed to investigate the effect of ARH3/PARG on either

poly-Ser-ADP-ribosylation, produced using the complex PARP1 WT/HPF1, and mono-Ser-ADP-ribosylation, produced using the complex PARP1 E988Q/HPF1 (**Figure 4.5**). Consistently, ARH3 was active in removing both poly-Ser-ADP-ribosylation and mono-Ser-ADP-ribosylation from the auto-modification domain of PARP1, while PARG was active just against poly-Ser-ADP-ribosylation.

Overall, these assays establish ARH3 as a very efficient glycohydrolase active on poly-Ser-ADP-ribosylation and mono-Ser-ADP-ribosylation of different substrates and suggest a different biological role of ARH3 compared to PARG.

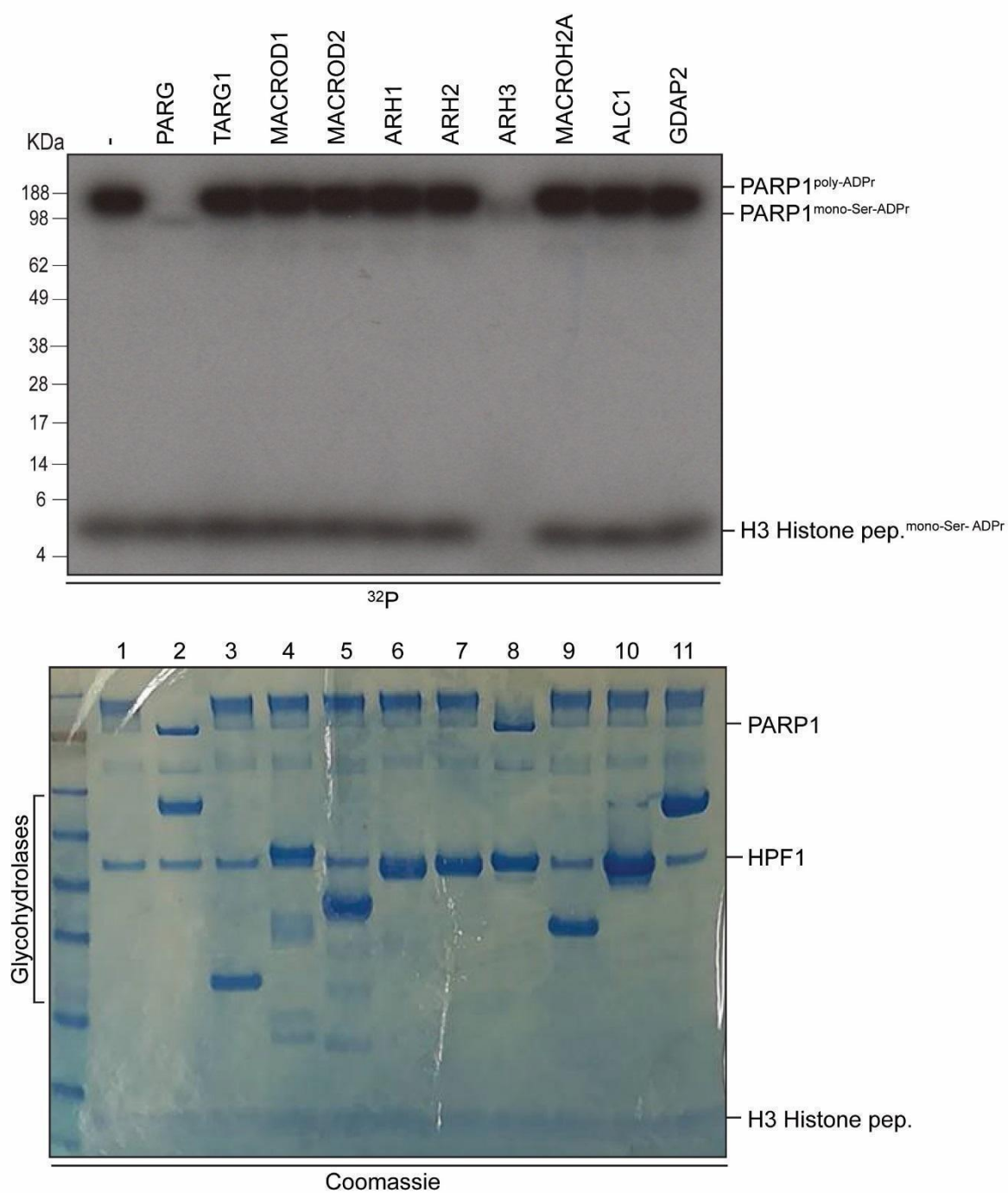


Figure 4.1. Recombinant PARP1 (1 μ M) was incubated with DNA (1 μ M), NAD⁺ 50 μ M), ³²P-NAD⁺ for labelling, HPF1 (1 μ M) for inducing Ser-ADP-ribosylation and H3 peptide (40 μ M) as a substrate of PARP1/HPF1. The reaction was performed for 30 minutes at room temperature and stopped with the addition of PARP1 inhibitor Olaparib (2 μ M). Glycohydrolases (3 μ M) were added and reaction performed for 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography.

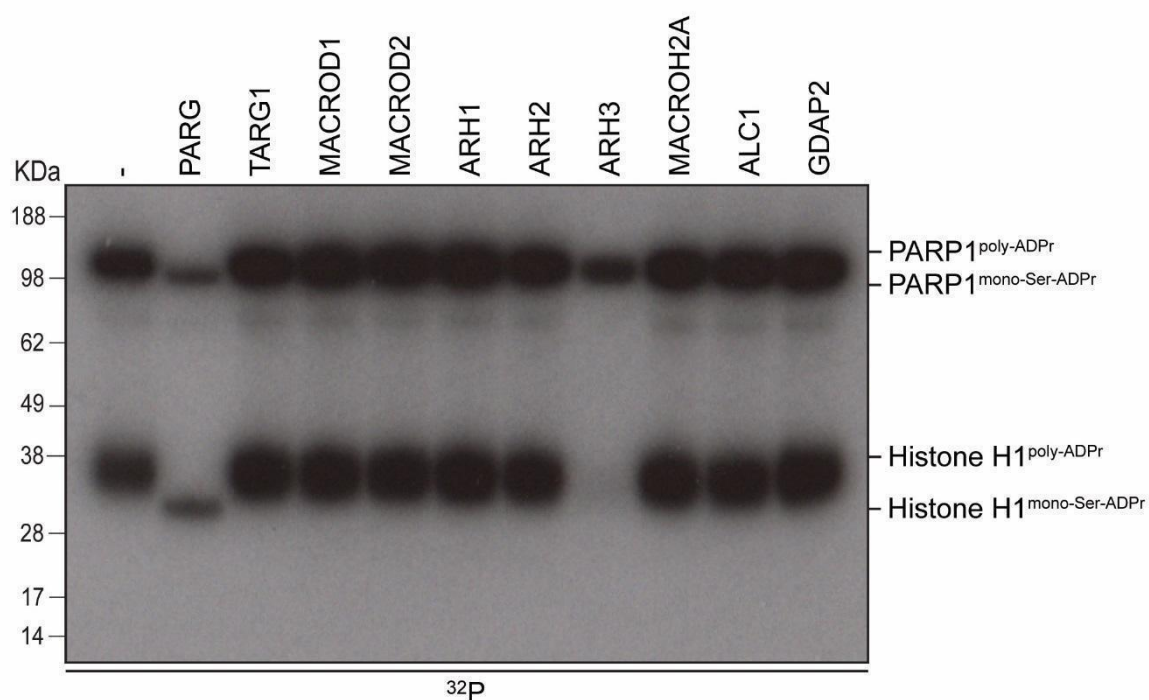


Figure 4.2. Recombinant PARP1 (1 μ M) was incubated with activated DNA (1 μ M), NAD⁺ (50 μ M), ³²P-NAD⁺ for labelling, HPF1 (1 μ M) for inducing Ser-ADP-ribosylation and recombinant Histone H1 (2 μ M) as a substrate of PARP1/HPF1. The reaction was performed for 30 minutes at room temperature and stopped with the addition of PARP1 inhibitor Olaparib (2 μ M). Glycohydrolases (3 μ M) were added and reaction performed for 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography.

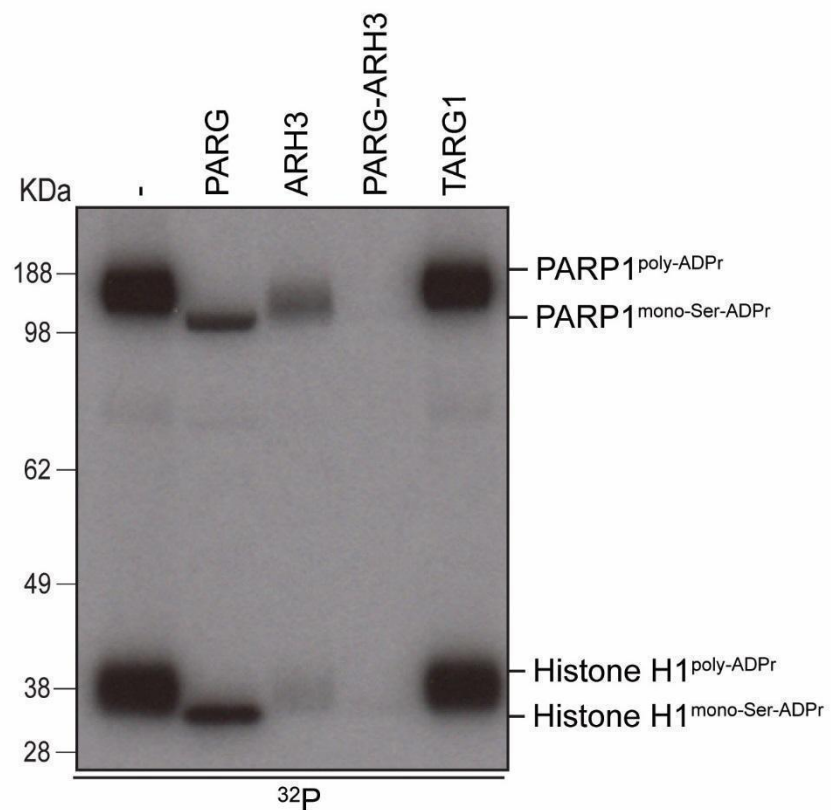


Figure 4.3. Recombinant PARP1 (1 μ M) was incubated with DNA (1 μ M), NAD⁺ (50 μ M), ³²P-NAD⁺ for labelling, HPF1 (1 μ M) for inducing Ser-ADP-ribosylation and recombinant Histone H1 as a substrate of PARP1/HPF1. The reaction was performed for 30 minutes at room temperature and stopped with the addition of PARP1 inhibitor Olaparib (2 μ M). Glycohydrolases (3 μ M) were added and reaction performed for 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography.

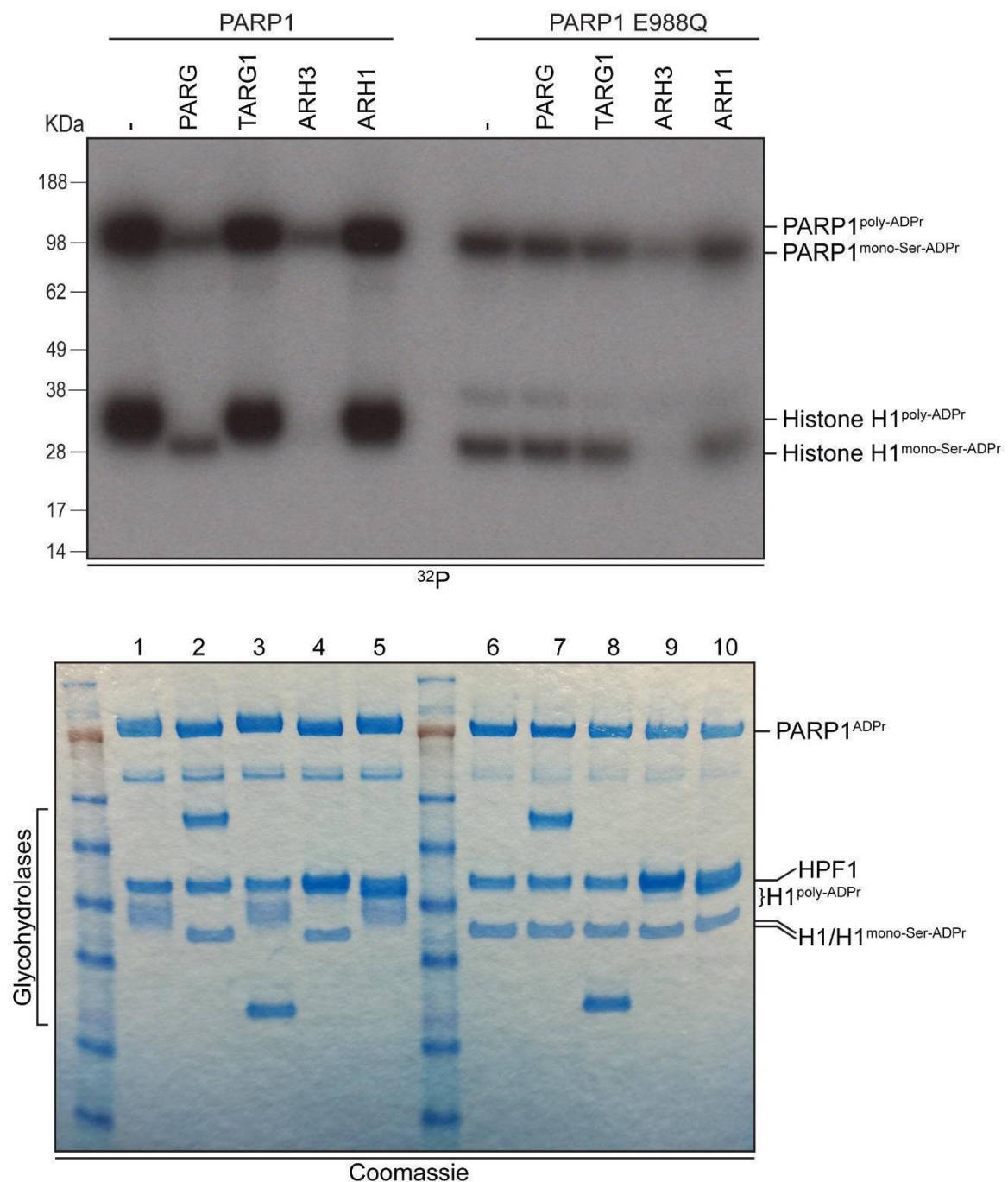


Figure 4.4. Recombinant PARP1 (1 μM) or PARP1 E988Q (1 μM) were incubated with activated DNA (1 μM), NAD $^{+}$, ^{32}P -NAD $^{+}$ for labelling, HPF1 (1 μM) for inducing Ser-ADP-ribosylation and recombinant Histone H1 as a substrate of PARP1/HPF1. The reaction was performed for 30 minutes at room temperature and stopped with the addition of PARP1 inhibitor Olaparib (2 μM). Glycohydrolases (3 μM) were added and reaction performed for 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography.

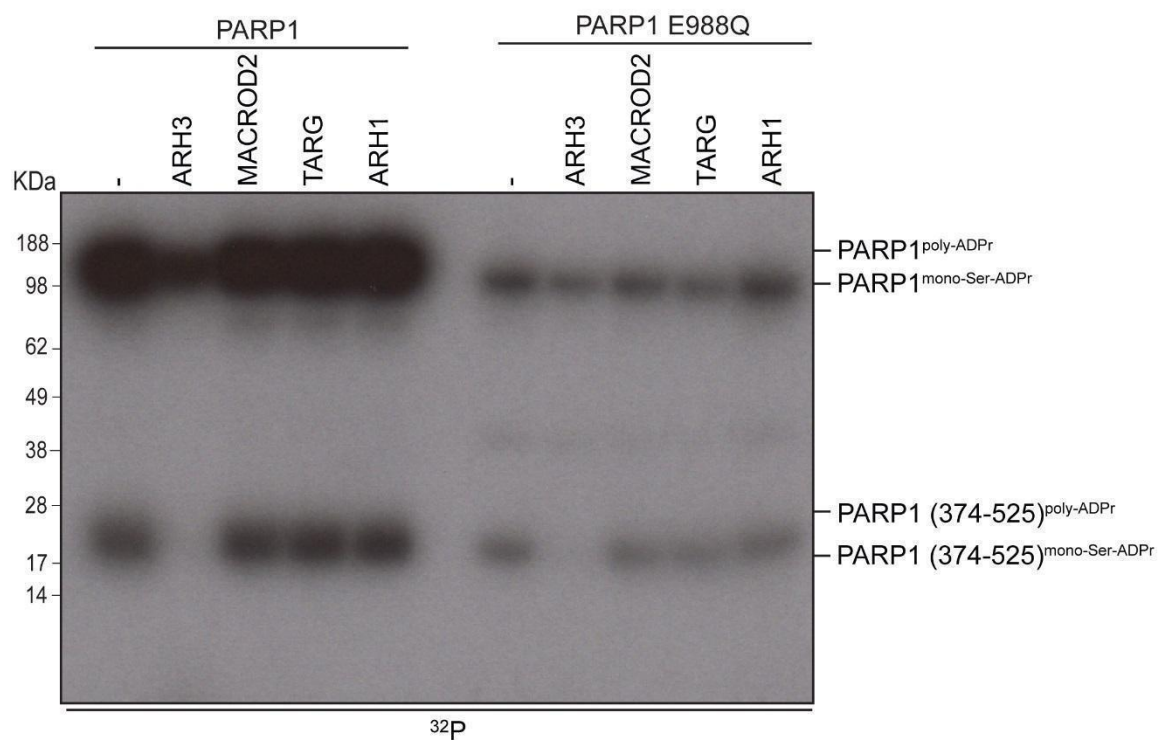


Figure 4.5. Recombinant PARP1 (1 μM) or PARP1 E988Q (1 μM) were incubated with DNA (1 μM), NAD^+ (50 μM), ^{32}P - NAD^+ for labelling, HPF1 (1 μM) for inducing Ser-ADP-ribosylation and PARP1 BRCT domain (aa 374-525, 2 μM) as a substrate of PARP1/HPF1. The reaction was performed for 30 minutes at room temperature and stopped with the addition of PARP1 inhibitor Olaparib (2 μM). Glycohydrolases (3 μM) were added and reaction performed for 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography.

4.2 Characterisation of ARH3 activities

The data showed so far, supports ARH3's hydrolytic activity against poly-Ser-ADP-ribosylation and mono-Ser-ADP-ribosylation. However, the presence in the reaction of other components such as PARP1-HPF1 and NAD⁺ raised the question whether ARH3 itself was enough for removal of poly-Ser-ADP-ribosylation and mono-Ser-ADP-ribosylation or if other components in the reaction play a role together with ARH3. To address this question, I established a protocol for the purification of mono-Ser-ADP-ribosylated H3 histone peptide. After performing an *in vitro* PARylation reaction by PARP1 in the presence of HPF1, the reaction was stopped with addition of the PARP1 inhibitor Olaparib and the peptide recovered by using a spin concentrator column with 10 kDa cut-off which allows the removal of PARP1 and HPF1. Subsequently, the modified peptide was loaded in a G25 desalting column for the removal of excess of NAD⁺. After its modification by PARP1/HPF1 complex and its purification, mono-Ser-ADP-ribosylated H3 histone peptide was incubated with each of the glycohydrolases (**Figure 4.6**). The assay showed clearly that ARH3 itself is sufficient for the removal of mono-Ser-ADP-ribosylation. I also aimed to test this activity using a different substrate. Previously we reported that serine 3 of H2B histone peptide could be Ser-ADP-ribosylated in the presence of PARP1/HPF1 complex (Bonfiglio *et al.*, 2017). As done for the previous assay, the H2B peptide was mono-Ser-ADP-ribosylated by the complex PARP1/HPF1, purified from the mix reaction, and incubated with each glycohydrolase (**Figure 4.7**). As expected, only ARH3 could efficiently remove mono-Ser-ADP-ribosylation from H2B histone peptide.

To further characterise ARH3 activity, I performed site directed mutagenesis of the catalytic amino acids Asp77 and Asp78, either singly or doubly, to alanine. Asp77

and Asp78 were shown to be essential for the coordination of magnesium ions and hydrolysis of poly-ADP-ribose (Oka *et al.*, 2006). Mutation of these two amino acids to alanine resulted in an ARH3 that is catalytically dead against PAR. The same result was achieved with the addition of the chelating agent EDTA (Oka *et al.*, 2006). I aimed to investigate whether the catalytic site, and in particular the two catalytic residues Asp77 and Asp78, were also responsible for the removal of mono-Ser-ADP-ribosylation. ARH3 WT, ARH3 D77N, ARH3 D78N and the double mutant ARH3 D77N D78N were incubated with mono-Ser-ADP-ribosylated H3 (**Figure 4.8A**) and mono-Ser-ADP-ribosylated H2B (**Figure 4.8B**). The assays showed clearly that all the mutants lost their activity toward mono-Ser-ADP-ribosylated histones. Moreover, I incubated ARH3 WT in the presence of chelating agent EDTA, confirming that the magnesium ion is essential for coordination of Asp77 and Asp78 for mono-Ser-ADP-ribosylation hydrolysis. These assays confirm that the already known active site of ARH3 is important for the removal of both poly-ADP-ribosylation and mono-Ser-ADP-ribosylation and suggest a similar mechanism of removal. Finally, I showed that ARH3 activity toward mono-Ser-ADP-ribosylation is, as expected, concentration (**Figure 4.9A**) and time dependent (**Figure 4.9B**).

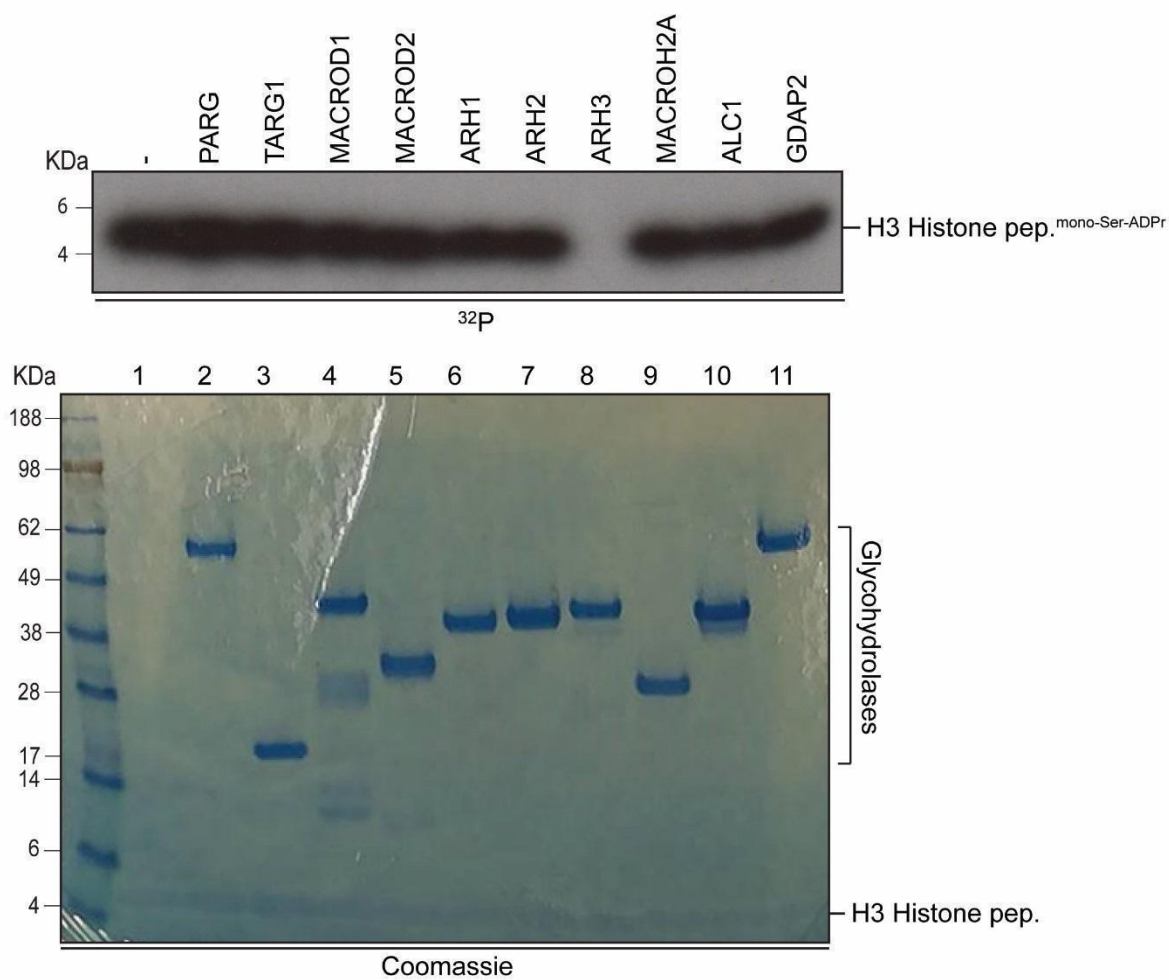


Figure 4.6. Purified H3 histone peptide mono-Ser-ADP-ribosylated (40 μM) was incubated with glycohydrolases (3 μM) 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography.

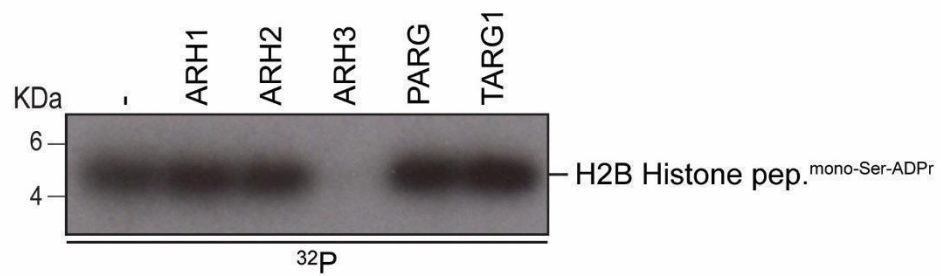


Figure 4.7. Purified H2B histone peptide mono-Ser-ADP-ribosylated (40 μM) was incubated with glycohydrolases (3 μM) 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography.

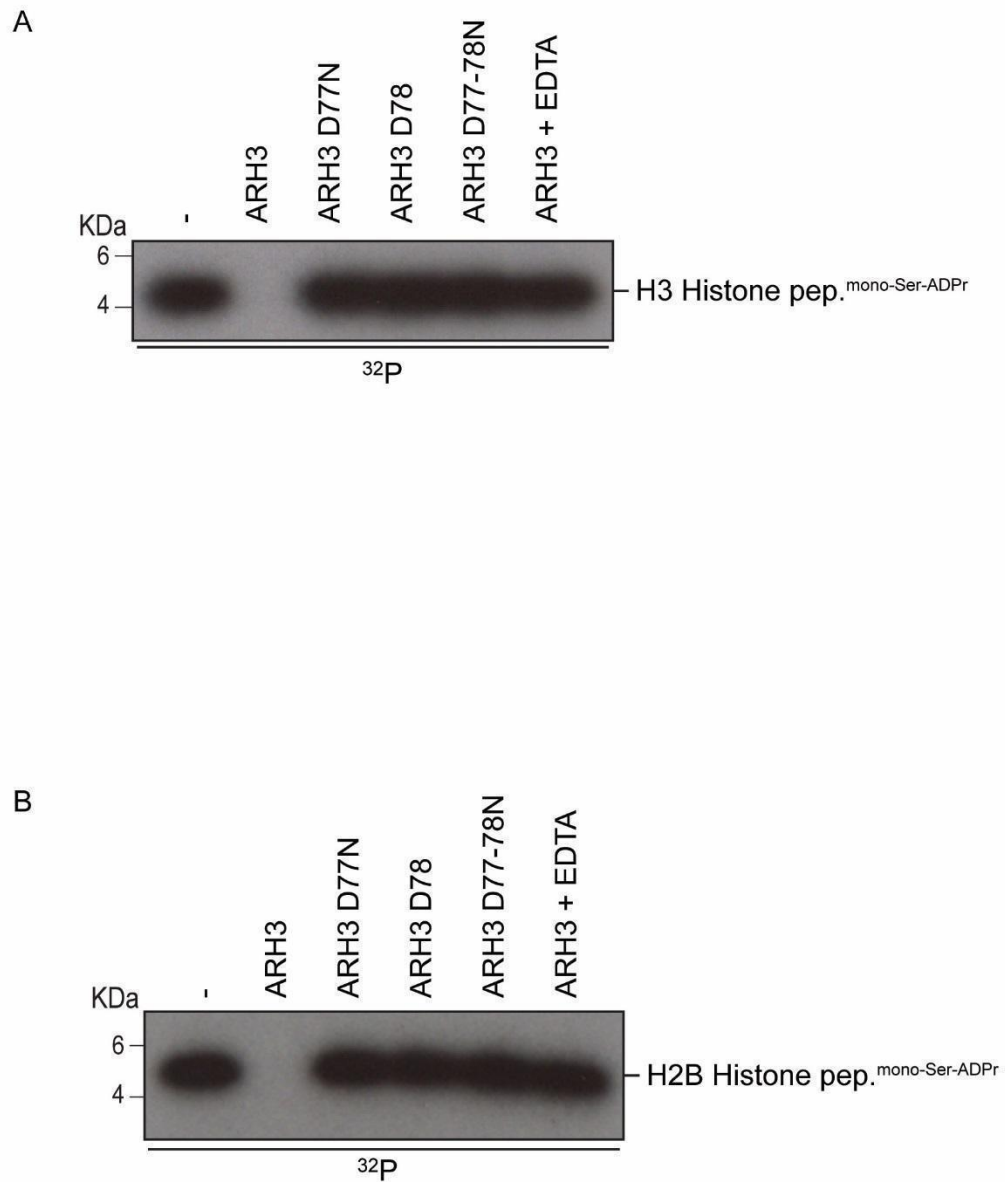
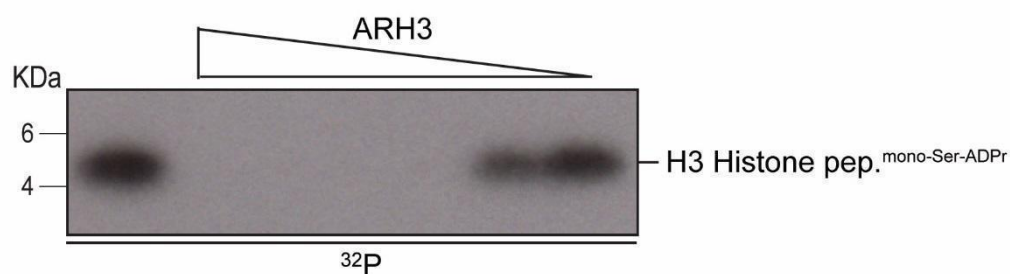


Figure 4.8. A - Purified H3 histone peptide mono-Ser-ADP-ribosylated (40 μM) was incubated with ARH3 WT (3 μM), ARH3 catalytic mutants and ARH3 supplemented with 2mM EDTA. The reactions were performed for 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography. **B** - Performed as in the previous assay, using purified H2B histone peptide mono-Ser-ADP-ribosylated (40 μM) instead.

A



B

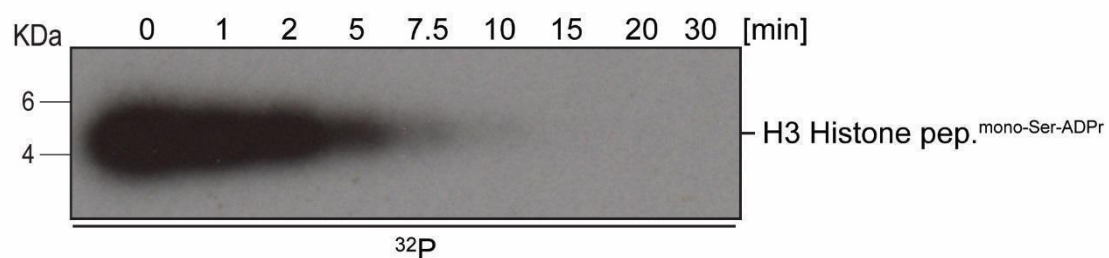


Figure 4.9. **A** - Purified H3 histone peptide mono-Ser-ADP-ribosylated (40 μ M) was incubated with ARH3 from 3 μ M to 50nM. **B** - Purified H3 histone peptide mono-Ser-ADP-ribosylated (5 μ g) was incubated with ARH3 (200nM) and sample (0.5 μ g) taken at the indicated time. The reactions were analysed by SDS-PAGE and autoradiography.

4.3 Specificity of ARH3 activity

Finally, I aimed to assess whether ARH3 is specific for mono-Ser-ADP-ribosylation or shows activity on other mono-ADP-ribosylation types such as those targeting Glu, Arg or Lys. Firstly, I assessed the activity of ARH3 toward mono-Glu-ADP-ribosylation. Prior to the discovery of Ser-ADP-ribosylation, mono-Glu-ADP-ribosylation had always been reported as one of the major ADP-ribosylation occurring in the PARP1-dependent DNA damage repair. PARP1 E988Q was used as a substrate for mono-Glu-ADP-ribosylation, since in the absence of HPF1, it mono-Glu-ADP-ribosylates itself (Sharifi *et al.*, 2013). After auto-modification, PARP1 E988Q was incubated with ARH3, TARG1, MACROD2 and ARH1. TARG1 and MACROD2 are known to be active against mono-Glu-ADP-ribosylation and therefore served as positive controls, while ARH1 is known to remove mono-Arg-ADP-ribosylation (Jankevicius *et al.*, 2013; Oka *et al.*, 2006; Sharifi *et al.*, 2013). The assay in **Figure 4.10A** showed that both MACROD2 and TARG1 were able to remove mono-Glu-ADP-ribosylation, while ARH3 and ARH1 showed no activity against this substrate. With this assay, I showed that ARH3 does not act on mono-Glu-ADP-ribosylation and, therefore, appears specific for mono-Ser-ADP-ribosylation. Secondly, I tested ARH3, ARH2, ARH1 and TARG1 against a mono-Arg-ADP-ribosylation substrate. Mono-Arg-ADP-ribosylation can be produced using the activity of ARTC2.2, a murine enzyme involved in toxic response and previously reported as an Arg-ADP-ribosyltransferase that promotes mono-ADP-ribosylation (Adriouch *et al.*, 2008). Inflammation can be caused by several different factors, such as bacterial infection, and can cause cell damage with release of NAD⁺ which will activate ARTC2.2 (Mendzel *et al.*, 2015). The substrate for mono-Arg-ADP-ribosylation was obtained by incubation of cell extract with ARTC2.2 and ³²P-NAD⁺. **Figure 4.10B** shows the positive control ARH1 as the only enzyme

able to cleave mono-Arg-ADP-ribosylation, while ARH3, ARH2 and TARG1 show no discernible activity. Finally, I generated a chemically modified H3 peptide, which carries ADP-ribose on a Lys residue, as previously reported (Jankevicius *et al.*, 2013). Lys-ADP-ribosylation has been hypothesised for a long time as one of the main ADP-ribosylation types involved in DNA damage response. However, no evidence of a Lys-transferase or a Lys-glycohydrolase has ever been reported either *in vitro* or *in vivo*. Nevertheless, I aimed to assess whether ARH3 could remove mono-Lys-ADP-ribosylation. The incubation of the chemically modified mono-Lys-ADP-ribosylation H3 peptide, showed, that neither ARH3 nor TARG1, have any activity against mono-Lys-ADP-ribosylation (**Figure 4.11**). These assays clearly confirmed that ARH3 is a specific glycohydrolase for the removal of only mono-Ser-ADP-ribosylation with no activity toward glutamate, arginine or lysine mono-ADP-ribosylation.

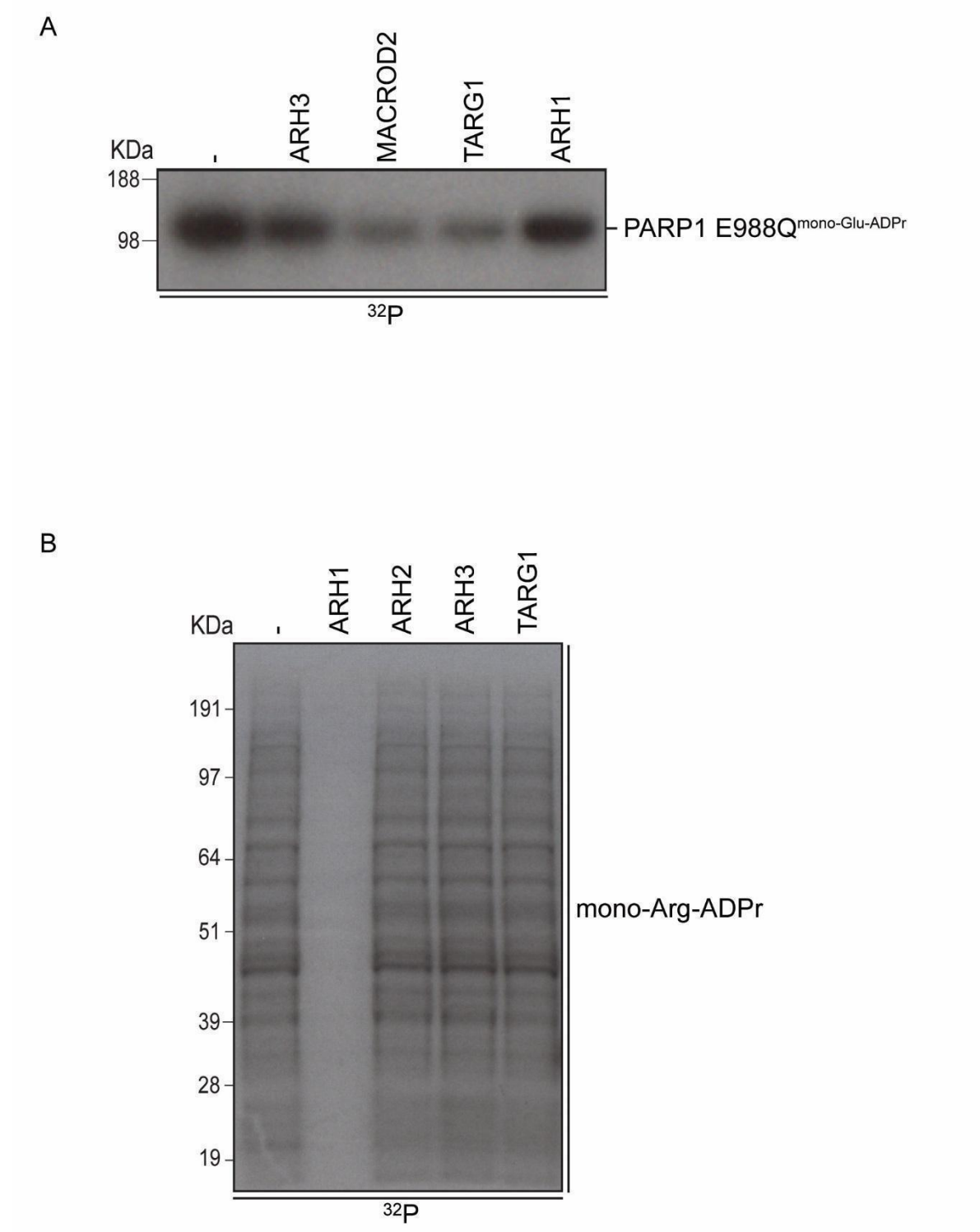


Figure 4.10. A - Auto-modified PARP1 E988Q (1 μ M) was incubated with glycohydrolases (3 μ M) 20 minutes at room temperature. **B** - Mono-Arg-ADP-ribosylated cell extract was incubated with glycohydrolases (3 μ M) for 15 minutes at 30°C. The reactions were analysed by SDS-PAGE and autoradiography.

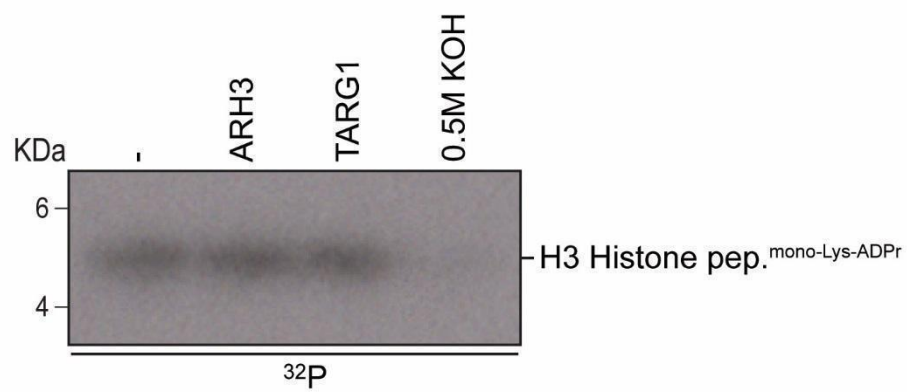


Figure 4.11. Mono-Lys-ADP-ribosylated H3 peptide was incubated (100μM) was incubated with glycohydrolases (3μM) for 30 minutes at room temperature or with 0.5M KOH for 15 minutes at 37°C. The reactions were analysed by SDS-PAGE and autoradiography.

4.4 Conclusion on ARH3 characterisation

Our lab previously reported that Ser-ADP-ribosylation is a common PTM after DNA damage, so a deeper understanding of the enzyme eraser of Ser-ADP-ribosylation will advance our knowledge of ADP-ribosylation regulation in genome stability (Leidecker *et al.*, 2016, Bonfiglio *et al.*, 2017; Fontana *et al.*, 2017; Palazzo *et al.*, 2018). In my project, I have reconstructed *in vitro* Ser-ADP-ribosylation using recombinant proteins and established the minimal components (PARP1/HPF1 or PARP2/HPF1) required for this reaction (Bonfiglio *et al.*, 2017). Our findings led me to investigate whether Ser-ADP-ribosylation is reversible and which enzyme could be responsible for its removal. Here, I demonstrated that ARH3 is the only enzyme among the known human ADP-ribosylation glycohydrolases which shows the ability to remove Ser-ADP-ribosylation. Although, as previously reported, ARH3 can also cleave PAR chains, it is much more efficient against Ser-ADP-ribosylation, suggesting that this latter activity constitutes the main function of ARH3 in the cell. I also showed that the activity of ARH3 against Ser-ADP-ribosylation is dependent on magnesium and the catalytic aspartate residues which were previously found essential for the removal of poly-ADP-ribose. Finally, I showed that ARH3 is specific against Ser-ADP-ribosylation and does not show activity toward glutamate, arginine or lysine. Finally, I propose that ARH3 could be a useful tool for the study and identification of Ser-ADP-ribosylation without the need for mass spectrometry.

In conclusion, I showed that ARH3 can reverse PARP1/HPF1-dependent Ser-ADP-ribosylation on either histone proteins or other substrates. This opens new perspectives on the regulation of ADP-ribosylation and offers new tools for the identification of Ser-ADP-ribosylation.

Chapter 5. Ser-ADP-ribosylation in *Drosophila melanogaster*

ADP-ribosylation is a widespread post-translational modification (PTM), in human cells, involved in many biological processes such as DNA repair, transcription, cell death and chromatin regulation. Recently, our lab identified protein serine residues as a new substrate for ADP-ribosylation involved in DNA repair and dependent on the complex PARP1/HPF1. In addition, I identified ARH3 as the only human enzyme able to remove the terminal ADP-ribose bound to a serine. Despite the increasing understanding of the regulation of ADP-ribosylation over the last few years, little is known about the role of ADP-ribosylation in *Drosophila melanogaster*. Differently from human, which possess 17 PARPs, *Drosophila* possesses a single gene codify Parp1 homolog which however can be spliced in three different isoforms: ParpI (DmParp1), ParpII and ParpE (Tulin *et al.*, 2002). ParpII lacks the auto-modification domain (Kawamura *et al.*, 1998) while ParpE is predicted to lack catalytic domain (Tulin *et al.*, 2002). Their biological functions, however, are not clear yet. DmParp1, as the human homolog, on the other hand has been found playing a role in many biological functions, as regulation of puffing/transcriptional regulation (Tulin *et al.*, 2003), splicing (Ji *et al.*, 2009) and more important genome stability (Ishak *et al.*, 2016). Notably, *Drosophila* possesses a PARP1 (DmParp1) as well as an HPF1 homolog (DmHpf1) but lacks an ARH3 equivalent. Given *Drosophila* as a central model organism of genetics and molecular biology, I aimed to investigate, using a set of biochemical assays, whether Ser-ADP-ribosylation is present in *Drosophila* and dependent on the complex Parp1/Hpf1. The study answers these questions in the affirmative. Moreover, I identified *Drosophila* Parg (DmParg) as the only enzyme among putative ADP-ribosylation glycohydrolases able to reversed the terminal ADP-ribose bound to

serine in *Drosophila* and aimed to structural characterise Parg in order to compare its structure with the human PARG. This comparison was conducted in order to explain why *Drosophila* Parg is able to remove Ser-ADP-ribosylation while human PARG cannot. Finally, I propose that *Drosophila* Parg acts as a combined human PARG/ARH3 and therefore could be a useful tool that allows complete cleavage of both mono- and poly Ser-ADP-ribosylation. Such a tool could be used for testing Ser specificity of PARP-dependent ADP-ribosylation.

5.1 DmParp1/DmHpf1 complex in *Drosophila* can Ser-ADP-ribosylate histone proteins

In the third chapter, I reported that the human complex PARP1/HPF1 can catalyse a new PTM called Ser-ADP-ribosylation. Furthermore, I showed that this modification is reversible, and that removal is catalysed by the human glycohydrolase ARH3 (Fontana *et al.*, 2017). Human HPF1 was initially identified following the discovery of the *Drosophila* orthologue which possesses a PBZ domain. After characterising human Ser-ADP-ribosylation, I wondered whether Ser-ADP-ribosylation is a conserved PTM present in other species, including *Drosophila melanogaster*, and if it can be reversed in these species. *Drosophila* possesses genes coding PARP1 and HPF1 homolog. However, there are no genes resembling ARH3, suggesting the existence of a different mechanism of Ser-ADP-ribosylation removal in this species. Firstly, I aimed to test whether the *Drosophila* DmParp1/DmHpf1 complex can install Ser-ADP-ribosylation on histone proteins. To investigate this, I expressed and purified recombinant DmParp1 and DmHpf1 from *E.coli* cells and compared their activity with the human PARP1/HPF1 complex, using as a substrate the already well-characterised H3 histone peptide (**Figure 5.1**). Despite the low levels of auto-modification activity of DmParp1 compared to PARP1, the assay showed clearly that DmParp1 by itself is not efficient in modifying the histone peptide. However, in the presence of DmHpf1, DmParp1 can modify the peptide as efficiently as the human complex. To confirm that this histone modification is indeed mono-Ser-ADP-ribosylation, I purified the modified histone peptide from the reaction mixture, as already shown in Chapter 4, and treated it with human glycohydrolases TARG1 and ARH3 (**Figure 5.2**). This assay showed that the modification on the histone peptide, introduced by either the human complex or the *Drosophila* complex, cannot be reversed by TARG1 but can

be efficiently reversed by ARH3. As already shown in Chapter 3, ARH3 can only remove mono-Ser-ADP-ribosylation, confirming that both the *Drosophila* and the human complex catalyse Ser-ADP-ribosylation.

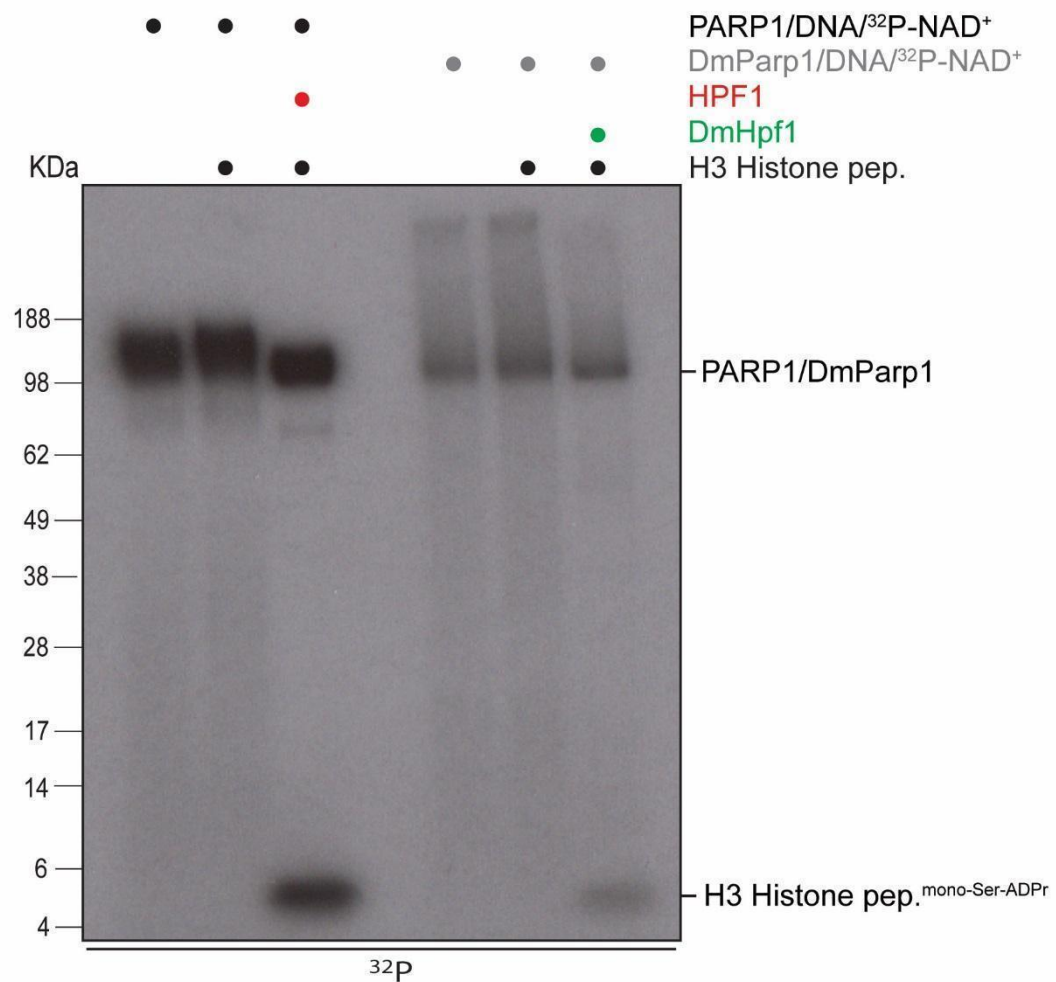


Figure 5.1. Recombinant PARP1 and DmParp1 (1 μM) were incubated in presence of DNA, ^{32}P -NAD $^{+}$ and HPF1 or DmHpf1 (1 μM). H3 histone peptide was used as a substrate for trans-ADP-ribosylation. The reactions were analysed by SDS-PAGE and autoradiography.

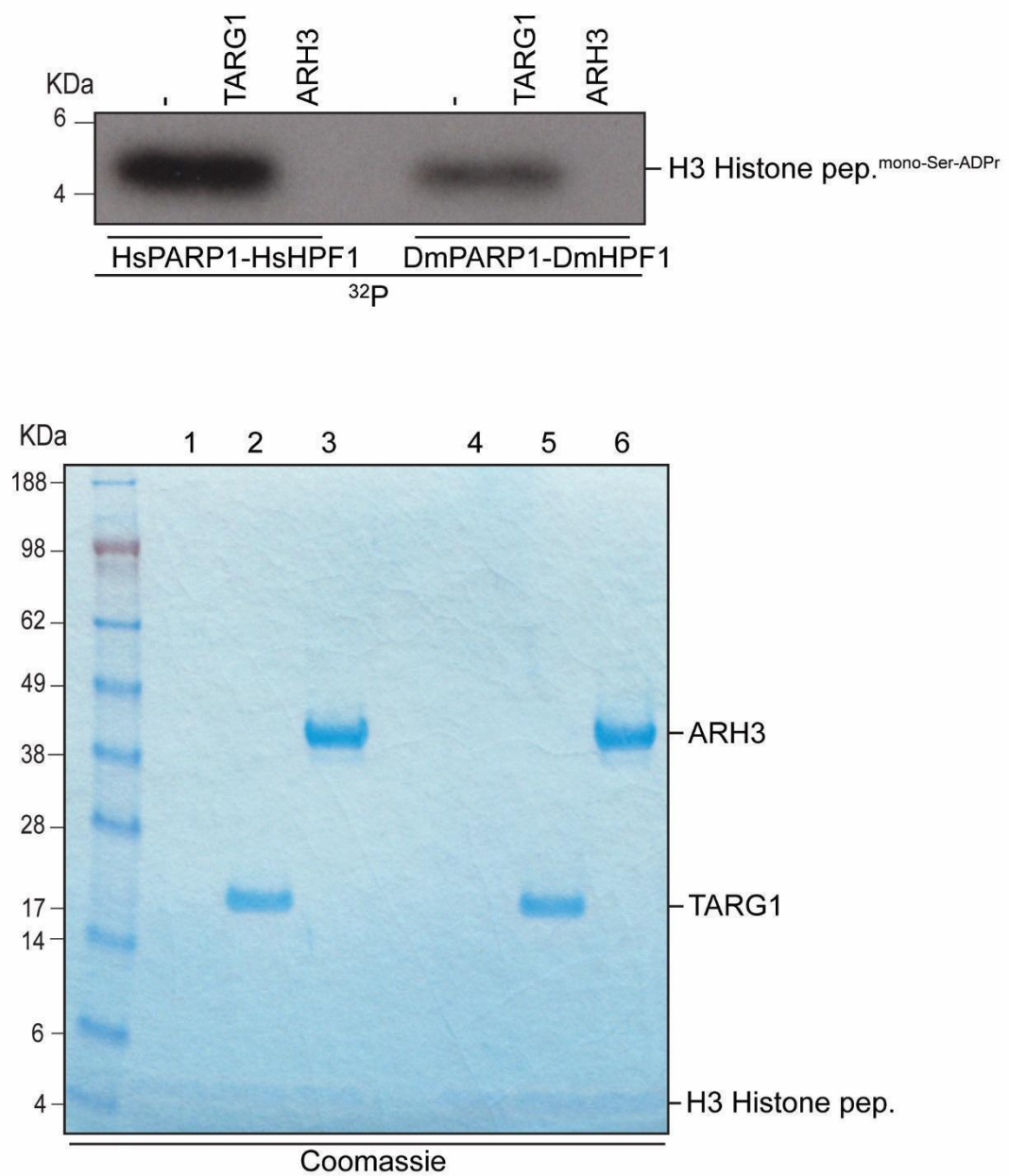


Figure 5.2. Mono-Ser-ADP-ribosylated H3 peptides from the previous reaction (Figure 4.1) were purified and incubated with TARG1 and ARH3. The reactions were analysed by SDS-PAGE and autoradiography.

5.2 DmParg hydrolyses Ser-ADP-ribosylation

After I confirmed the presence of Ser-ADP-ribosylation in *Drosophila*, I wished to identify a potential glycohydrolase which could reverse this PTM. As previously reported, ADP-ribosyl hydrolases usually possess a macrodomain fold (PARG, TARG1, MACROD1 and MACROD2) or a DarG-like fold (ARH1, ARH2 and ARH3) (Glowacki *et al.*, 2002; Oka *et al.*, 2006; Rack *et al.*, 2016). In *Drosophila*, I could only identify proteins belonging to the former class, in particular, DmParg, DmTarg1, DmTarg2 and DmTarg3. I reasoned that one of these hydrolases had to be responsible for removing Ser-ADP-ribosylation, so I cloned and purified all four recombinant proteins from *E.coli* and incubated them with the mono-Ser-ADP-ribosylated H3 peptide (**Figure 5.3**). Surprisingly, while DmTarg1, DmTarg2 and DmTarg3 had no activity against mono-Ser-ADP-ribosylation, DmParg was able to catalyse the removal of this modification from the H3 peptide. Previously, it has been reported the human PARG could cleave poly-ADP-ribosylation but had no activity against mono-ADP-ribosylation (Slade *et al.*, 2011). Moreover, I demonstrated in Chapter 4 that the human form of this enzyme does not remove Ser-ADP-ribosylation either. Therefore, it is striking to see that its *Drosophila* homolog catalyses efficient Ser-ADP-ribosylation removal. The difference in activity between human PARG and *Drosophila* DmParg led me to investigate whether DmParg was still able to cleave poly-ADP-ribose, which is the established function of human PARG. Therefore, I incubated all of the previously purified glycohydrolases with auto-modified PARP1 (**Figure 5.4**). In the absence of HPF1, PARP1 poly-ADP-ribosylates itself on glutamate residues. Whilst DmParg had an additional activity against mono-Ser-ADP-ribosylation, it still retained the canonical activity against poly-ADP-ribosylation. Just like PARG, DmParg effi-

ciently cleaves all of the poly-ADP-ribosylation, leaving a single band which corresponds to the mono-Glu-ADP-ribosylation as previously reported (Slade *et al.*, 2011). An additional implication of this result is that DmParg, even though able to remove mono-Ser-ADP-ribosylation, cannot remove mono-Glu-ADP-ribosylation, suggesting a specific activity toward serine. In order to further characterise DmParg, I incubated the PARP1/HPF1 complex with the histone tetramer H3.1/H4 before treating with a panel of glycohydrolases. The presence of HPF1 in this assay ensures that the modification, whether on PARP1 or histones, is on serine residues, instead of glutamate residues (**Figure 5.5**). I repeatedly observed a distinct double band on the auto-modified PARP1, which suggested that in the presence of HPF1 and the histone tetramer, PARP1 preferentially modifies the histones, resulting in limited auto-modification that consists of some poly- and some mono-Ser-ADP-ribosylated PARP1. If this interpretation is correct, the assay provides a unique opportunity to test separately the activity against poly-ADP-ribosylation and the terminal attachment of mono-Ser-ADP-ribosylation. As expected, PARG could completely trim poly-ADP-ribosylation from both PARP1 and histones, leaving a single band corresponding to mono-Ser-ADP-ribosylation for both substrates. ARH3 was not very efficient in cleaving poly-Ser-ADP-ribosylation attached either to histones or PARP1 but was able to remove mono-Ser-ADP-ribosylation from PARP1. Finally, DmParg was able to completely hydrolyse both poly-Ser-ADP-ribosylation and mono-Ser-ADP-ribosylation from both PARP1 and histones. This assay suggests that the preferred substrate of human ARH3 is indeed mono-Ser-ADP-ribosylation and poly-ADP-ribosylation is only weakly processed by this enzyme. On the contrary, DmParg seems perfectly able to hydrolyse both poly-Ser-ADP-ribosylation and mono-Ser-ADP-ribosylation. Subsequently, I performed a variant of this experiment using the already introduced H3 peptide as a PARP1/HPF1 substrate. Unlike the histone tetramer, the H3 peptide can be just

mono-Ser-ADP-ribosylated (**Figure 5.6**). As in the previous assay, DmParg could completely hydrolyse both poly-ADP-ribosylation and mono-Ser-ADP-ribosylation from auto-modified PARP1 as well as the mono-Ser-ADP-ribosylation from histone H3 peptide. PARG, however, could only trim the poly-ADP-ribosylation of PARP1, leaving a single band corresponding to mono-Ser-ADP-ribosylation PARP1 and showing no activity against the mono-Ser-ADP-ribosylated H3 peptide. In all these assays, the other three *Drosophila* glycohydrolases did not show any activities against any substrates.

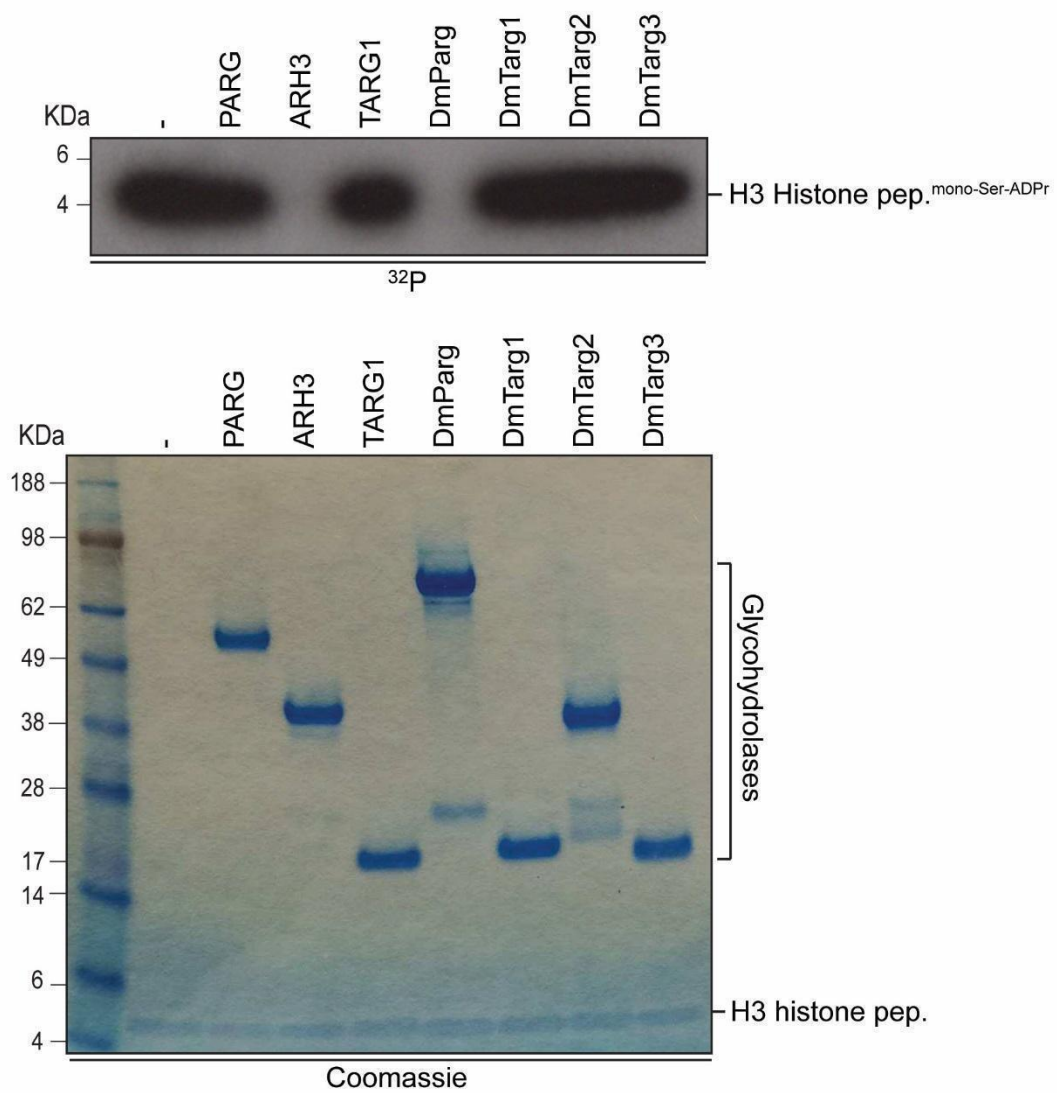


Figure 5.3. Mono-Ser-ADP-ribosylated H3 histone peptide (40 μ M) was incubated with human and *Drosophila* glycohydrolases (3 μ M) for 20 minutes at room temperature. The reactions were analysed by SDP-PAGE and autoradiography.

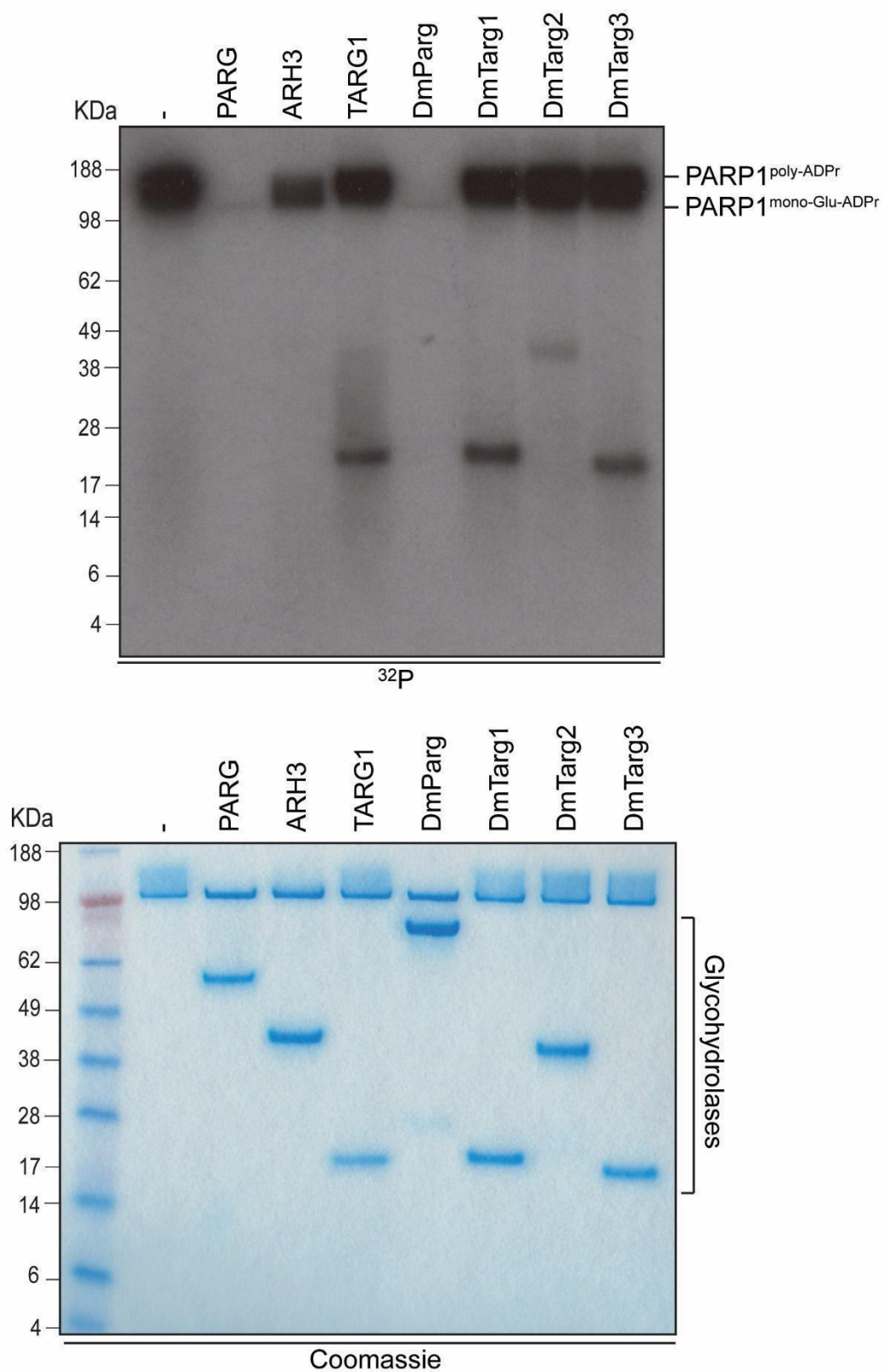


Figure 5.4. Auto-modified PARP1 (1 μ M) was incubated with human and *Drosophila* glycohydrolases (3 μ M) for 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography.

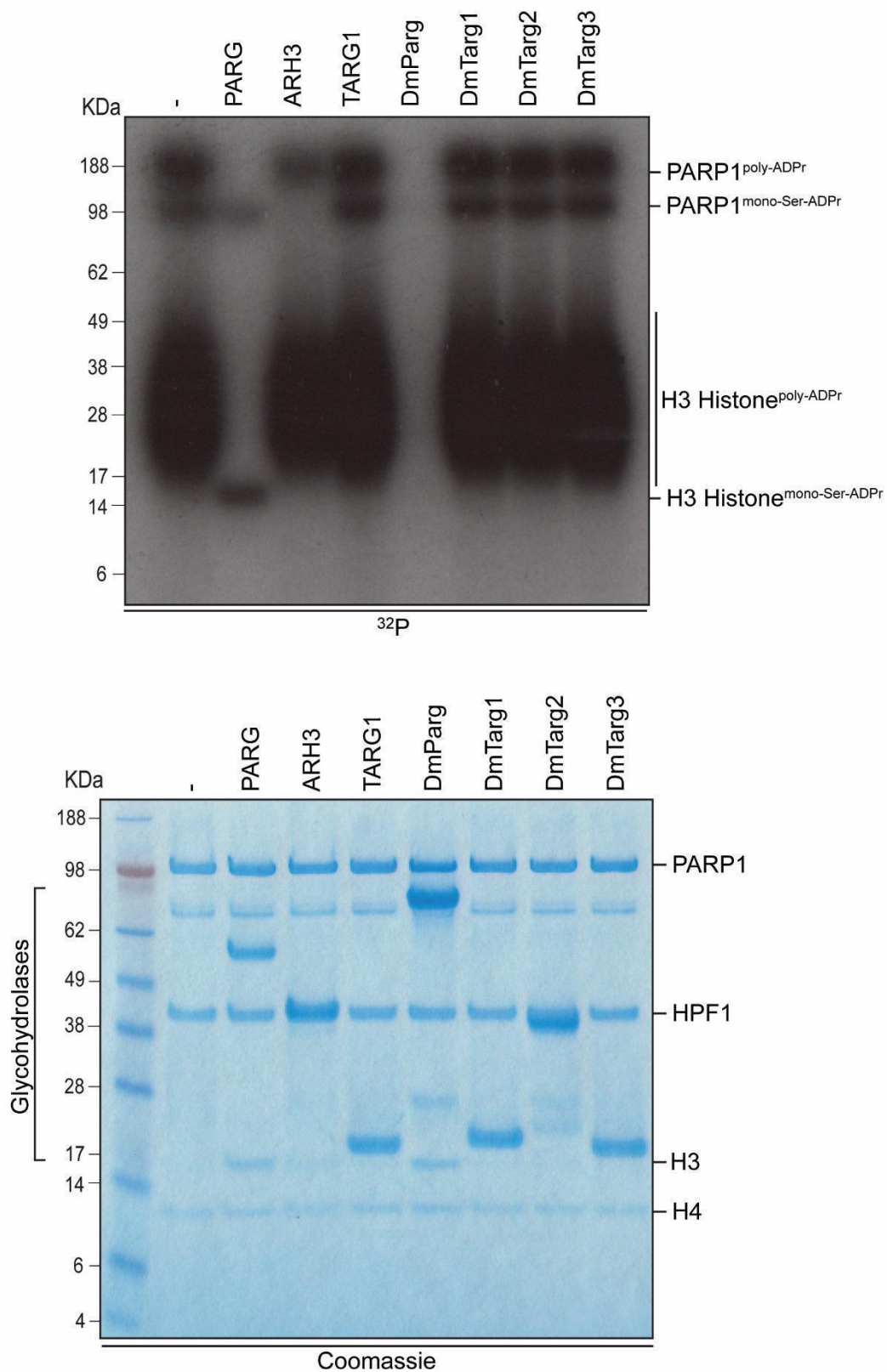


Figure 5.5. Recombinant PARP1 (1 μ M) was incubated with activated DNA (1 μ M), NAD⁺ (50 μ M), ³²P-NAD⁺ for labelling, HPF1 (1 μ M) for inducing Ser-ADP-ribosylation and histone tetramer H3/H4 (2 μ M) as a substrate of PARP1/HPF1. The reactions were incubated with human and *Drosophila* glycohydrolases (3 μ M) for 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography.

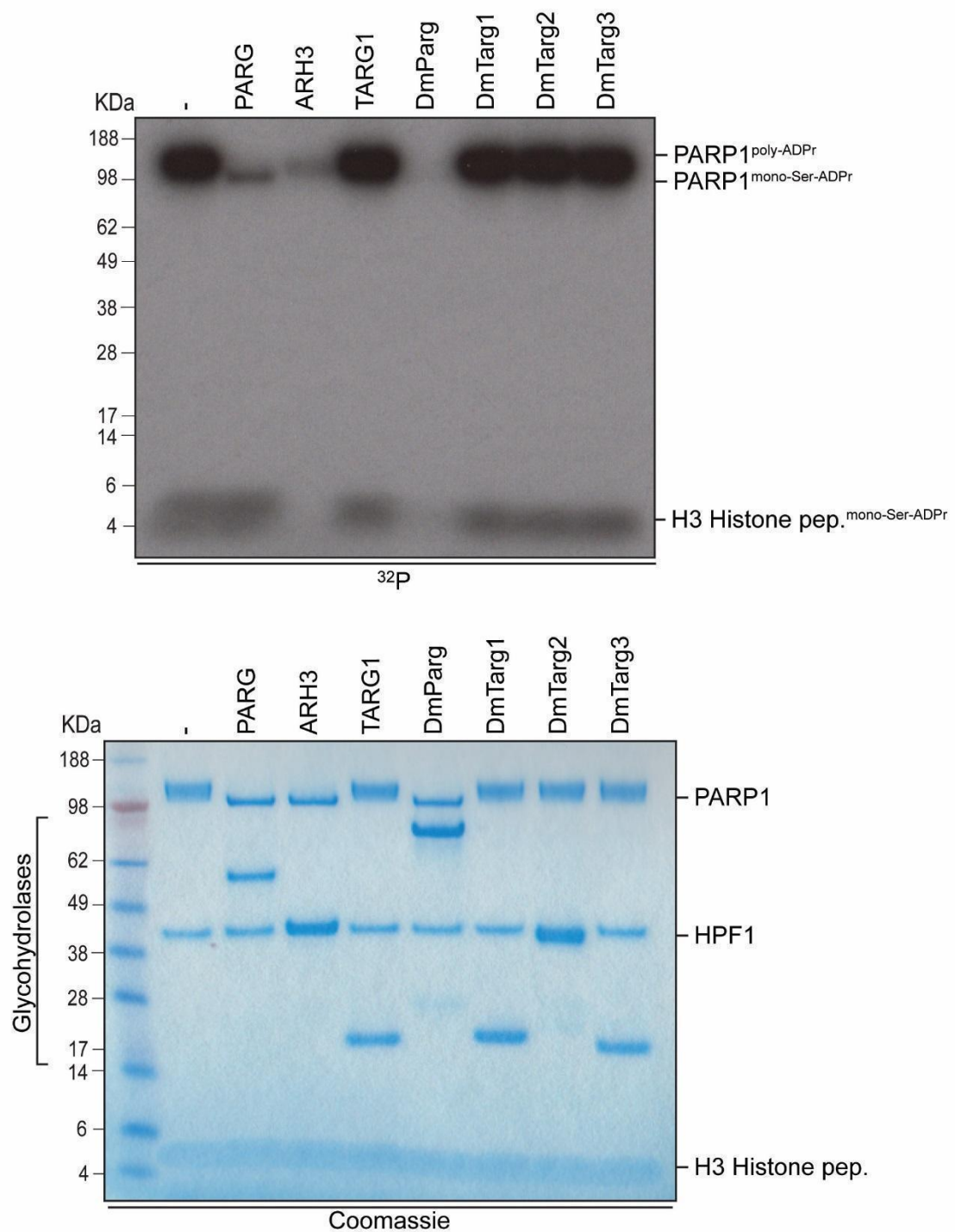


Figure 5.6. Recombinant PARP1 (1 μ M) was incubated with DNA (1 μ M), NAD⁺ (50 μ M), ³²P-NAD⁺ for labelling, HPF1 (1 μ M) for inducing Ser-ADP-ribosylation and histone H3 peptide (40 μ M) as substrate of PARP1/HPF1. The reactions were incubated with human and *Drosophila* glycohydrolases (3 μ M) for 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography.

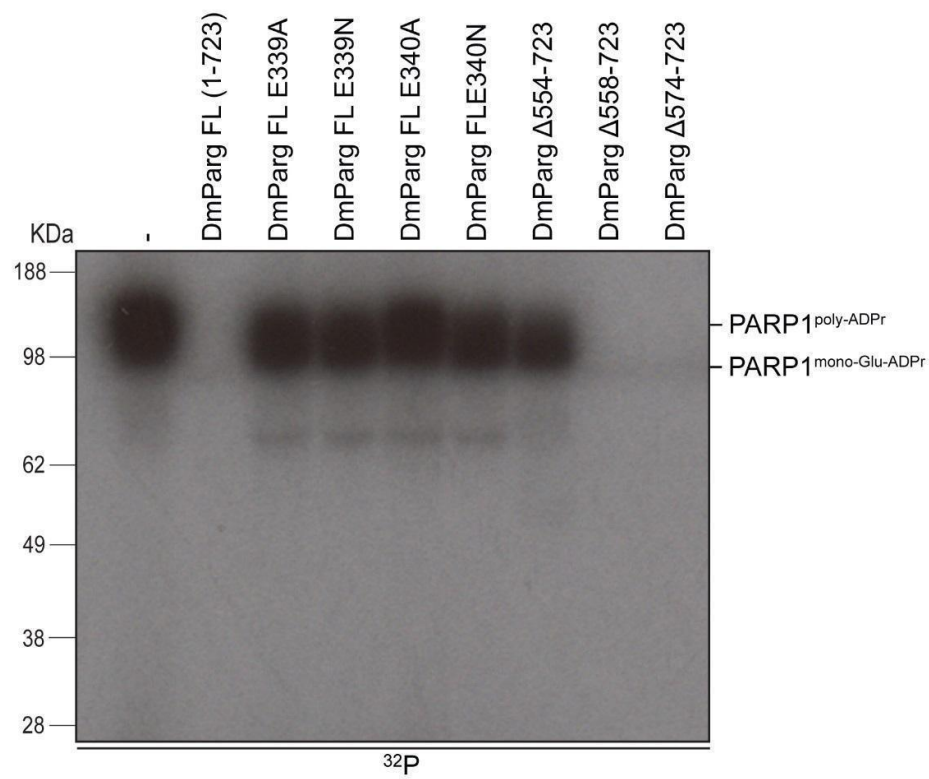
5.3 Characterisation of the serine ADP-ribosylhydrolase activity of DmParg

The previous characterisation of PARG has clearly revealed the catalytic site of the enzyme and the conserved region of the catalytic loop (Slade *et al.*, 2011). The amino acid residues E755 and E756 have been shown to be essential for cleaving poly-ADP-ribosylation. Therefore, I aimed to investigate whether the equivalent residues in DmParg are essential for cleaving mono-Ser-ADP-ribosylation and poly-ADP-ribosylation. I mutated the corresponding glutamates in DmParg (E339 and E340, respectively) to either asparagines or alanines, purified the proteins from *E.coli* and tested their activity. First I incubated WT and mutants with auto-modified PARP1 as a substrate for poly-ADP-ribosylation hydrolysis (**Figure 5.7A**). As expected, the assay showed clearly that all four mutants lost their ability to cleave poly-ADP-ribosylation. I then incubated WT and mutants with mono-Ser-ADP-ribosylation H3 as a substrate for mono-Ser-ADP-ribosylation removal (**Figure 5.7B**). The assay shows that none of the mutants retains the ability to remove the mono-Ser-ADP-ribosylation possessed by the WT enzyme. These two assays show clearly that the same key catalytic residues in DmParg, as in human PARG, are indeed responsible for hydrolysing both poly-ADP-ribosylation and mono-Ser-ADP-ribosylation. Since I found that the catalytic site of DmParg is responsible for hydrolysing both poly-ADP-ribosylation and mono-Ser-ADP-ribosylation, I reasoned that the difference in activity between PARG and DmParg could be located outside the catalytic site in a secondary region that helps to remove the modification from a serine residue but be dispensable for cutting poly-ADP-ribosylation. The idea was that mutating this region would allow separation of function between the two types of hydrolytic activity. In order to identify such po-

tential regions, I performed a series of sequence alignments of several PARG homologs (A representative alignment is reported in **Figure 5.8**). The alignments identified a unique C-terminal region, of about 150 amino acids, that is present in Pargs from all sequenced *Drosophila* species but absent in homolog proteins from any other organism. I reasoned that this region could form a secondary site that contributes uniquely to the removal of mono-Ser-ADP-ribosylation and is dispensable for poly-ADP-ribosylation processing. Therefore, I created a series of truncated DmParg polypeptides lacking various portions of the C-terminal region (DmParg Δ 554-723, DmParg Δ 558-723 and DmParg Δ 574-723) and compared their activity to that of the full-length enzyme DmParg a.a. 1-723. First I incubated the full-length form and truncations with auto-modified PARP1 as a model for poly-ADP-ribosylation (**Figure 5.7A**). Out of the three tested truncations, DmParg Δ 554-723 has lost its activity against poly-ADP-ribosylation, suggesting that DmParg needs a part of the C-terminal extension for catalytic competence, possibly because it contributes to the structural integrity of the enzyme. However, DmParg Δ 558-723 and DmParg Δ 574-723 retained their activity against poly-ADP-ribosylation. I then tested the activity of all the DmParg variants against mono-Ser-ADP-ribosylation H3 (**Figure 5.7B**). I observed the same pattern as in the previous assay, with the longest truncation lacking hydrolytic activity and the two shorter truncations being comparable to WT. This result rules out the C-terminal extension of DmParg as the region that accounts for the unique activity of this enzyme against Ser-ADP-ribosylation. Instead, this extension seems largely dispensable for Ser-ADP-ribosylation, while a part of it is likely required for a proper protein fold. Despite the negative result, these truncations turned out to be important for two reasons. Firstly, they allowed me to identify the minimal essential region of DmParg and, secondly, they turned out to be better material for the further structural characterisation by X-ray crystallography compared to the full-

length protein. However, these assays still left open the question of why DmParg can efficiently hydrolyse both mono-Ser-ADP-ribosylation and poly-ADP-ribosylation while PARG can only cleave poly-ADP-ribosylation.

A



B

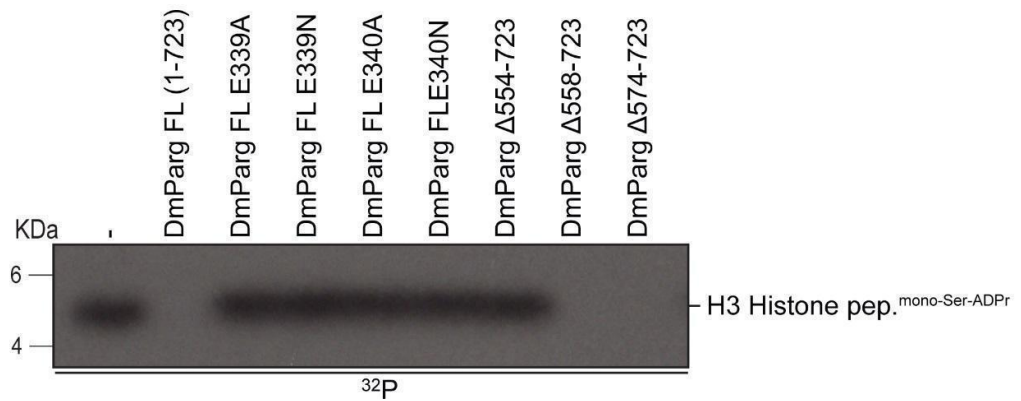


Figure 5.7. A - Recombinant PARP1 (1 μ M) was auto-modified in presence of DNA (1 μ M), NAD⁺ (50 μ M), ³²P-NAD⁺ for labelling. The reactions were incubated with DmParg FL, point mutants and truncations (3 μ M) for 20 minutes at room temperature. **B** - Mono-Ser-ADP- ribosylated H3 histone peptide (40 μ M) was incubated with DmParg FL, point mutants and truncations (3 μ M) for 20 minutes at room temperature. Reactions were analysed by SDP-PAGE and autoradiography.

DmParg	1	-----
TtParg	1	-----
HsPARG	1	MNAGPGCEPCTKRPRWGAATTS PAASDARSFPSRQRRVLPDKDAHVFQVRVPPSPACVPGRAGQHRGSATSLVFKQKTI TS WMDTKGI KT
EdParg2	1	-----
consensus	1	-----
DmParg	1	-----
TtParg	1	-----
HsPARG	91	AESESLDSKENNNTRI ES MMSVQKDNFYQHNVKLE NVS QLS LDKSPT EKSTQYL NQHQT AAMCKWQNEGKHTEQLLESEPTQVTLPVE
EdParg2	1	-----
consensus	91	-----
DmParg	1	-----
TtParg	1	-----
HsPARG	181	QFSNANI DRSPQNDDHS DTDS EENRDNQFLT TTKLANAKQT TE DEQARE AKSHQKCSKS CDPGE DCAS CQQDEI DVPPE S PLS DVGSED
EdParg2	1	-----
consensus	181	-----
DmParg	1	-----
TtParg	1	-----
HsPARG	271	VGTGPKNDNKLTRQESCLGNSPPFEKSEPESPMDVDNS KNS CQDS EADEETS PGFDEQDGS S SQT ANKPSRFQARDADI EFRKRYSTK
EdParg2	1	-----
consensus	271	-----
DmParg	1	-----MSKSPDGGI SEI LLE-----EEPEEN--LANSLD
TtParg	1	-----
HsPARG	361	GGEVRLHFQFEGGESRTGMNDLNAKLPGNI S SLNVECRNSKQHGKKDSKI TDHFMRLLPKAEORRKEQWELKHQRTTERKI PKYVPPHLSPD
EdParg2	1	-----
consensus	361	-----
DmParg	27	DSWRGVSMALHNRPQFELNLPVTAGNLIRVYQOLPI-----RETPIRPIYKSPGKWDSEHVRLPAPESKYPRENPDCSTIIDFR
TtParg	1	-----MIEAEKQENKKI QDYQFPLPQKNSLEWTI QKRTLQDLS-----SGKQKL---
HsPARG	451	KKWLCTPILEMRNMPR--CGI RL LLLRPSANITVIRVDL---LRAGEVPKPPPTHYKDLWNKHFVKMPCSEQNLYTVVDENGERTAGSN
EdParg2	1	-----MENKKPRITSLS
consensus	451	-----
DmParg	110	WEIMTERALLQPIKYCEELQAATFSYNTTYRDQWHF-----RALHQLLDEE---LDESETRVFEEDLLPRIIRLALRLP-----
TtParg	45	-----DSFOSLESTLEILRDSKNQNDERYFNL-----KAWF---EQ---LDKEEQTYFCEQFIPKICOLWLKTK-----
HsPARG	536	WEILQIALLNKFTRPQNKDAIL-KYNVAISKKWDF-----TALIDFWKVF---LEEAEAQHLYQSTLPDMVKKALCEP-----
EdParg2	13	VEKIRTAINKKEFKTCQDVVEFFISQEKERTNKETNKGVGTYGYIALKSEYNNKWSLGLTKRSFEELYQWELAGLKNVMNKLITNNDGTL
consensus	541	-----
DmParg	180	-DEIQSVPLKHHKKNASLSLSQQQLSCLLANAFLCITFPRRNTLKRKSEYSTPDI NFNRLYQSTG-PAVLIEKLKCIWHYFRRVCPTERD
TtParg	103	-KKQLK-NQIPKESKIYEAAPSRIEISYYVSCMFICILKDDQR---KITVKDRLIYKDLVQQTINIRROEKIKKCFYEYLKQALDFSEK
HsPARG	606	-NICTOPIPLKOKMNHSTMSQEQIASLLANAFECTFPRRNA-KMKSEYSSYPDI NFNRLEFGRS-SRKPEKLKTFYCFRRVITE---
EdParg2	103	KNQSTIYIPISSEGISPTVQIKRMEVFGILACSTFNVQWEDRVE-----RTGNISLVYISTNSRVSIERLICTFAYEHVMI EIDSK
consensus	631	-----
DmParg	268	ASNVPITGVVTFVR-RSGLPEHLIDWSQSAAPLGDVPLHWDAGETI EDEGLGLLOVDFANKYLGGGVGHGCVQEEIRFVIPELLVGKIL-
TtParg	186	E---SKEVVIIFORINCQLEEDYENWVDKLKAIKLKNVOLTDDK-LIEDFPGLLOVDFANCDIGGGVGNGLVQEEIRFCVPEIMLVSLLV
HsPARG	689	--KKPTGLVTFTR-QS--LEDFPEWERCEKPLT--RLHVTYEGTIEENGQGLLOVDFANRFVGGGVTSAGLVQEEIRFLINPELITISRL-
EdParg2	184	T---LQEI VTYVERVI---QKTSYDWEHNETPIKTYVFVSGG---LEMSMNVHVDFAKYLQIKHVPVSTQEEILFESVRPECFISLIL
consensus	721	-----
DmParg	356	FTECLRPFEALVMLCAERYSNITGYAGSEMSGNFEDSIPRDSGHRQIAIVAI DALHFAQSHHOYRDLIMERELNKAYIGFVHWMVTP-
TtParg	272	FDSMEANEVIMKGIKOYSDYQGYNSRFVKMGNSKTQKQKRNNPCTIAIDALCFNSSDNOFSVNVSRRELNKSYMGFKQE---D
HsPARG	771	FTEVLHDNECLITIGTEQYSEYTYGAETRYRSRSHEDGSEEDWQHCTEIVAI DALHFRRYLDQFVPEKMRRELNKAYCGFLRPGVSSE
EdParg2	264	FTEVLLDNEAVLIKNTLHISLSEGYQOTKWLGFCKKQIL-----KQONTIAIDSVI---NTCFGKENIDRLNKCVMGFQYA---
consensus	811	-----
DmParg	445	-PPGVATGNWCGGAFGGDSYIKALLQILVCAQLGHPLAY--YIFGNVEFRDDEHFWLFRNDGTIVQQLWSILRSYSRLIKESSKEPR
TtParg	357	QLKTI STGKMGCGAFLGVFDLKFALQWASSRSNKKMI--CTFQDEQTTKQIQOVFDLYKQKNASTI--FLKLVMDVPNSK-----
HsPARG	861	NLSAVATGNWCGGAFGGDARLKALLQILAAAAAENDVY--FTFGDSELMRDLYSWHIFLTERKLLVGDVYKLLRLRYYNEECRNC----
EdParg2	339	-GNKIATGOWGCGSF GNDPLLKMAQYIVLSTLGF EASI HWMNFARY--NELNDYVEICITKGMTIKDVIYILCSFDGDK-----
consensus	901	-----
DmParg	532	ENKASKKKLYDFIKELIKKVRDVPGEAGSAEAGS SRVAGLGEGKSETSAKSSPELNKQPARPQITITQQSTDLLPAQLSQDNSNSSEDOA
TtParg	434	--YMEDYTLLEYLILGK-----EKATSKNS--
HsPARG	944	STPGPDIKLVPFIYHAVESCAETADHSQ-RTGCT-----
EdParg2	416	-----NTFVDYFKQQLQSFHK-----
consensus	991	-----
DmParg	622	LMLSDDEEANAMVEAASLEAKSSYEISNSSTTSKTSSTATKSMGSGGRQLS LLEMLDIHYEKGSAKRPKSPNCSSKAEGSAKSRKEID
TtParg	1081	-----
HsPARG	1081	-----
EdParg2	1081	-----
consensus	1081	-----
DmParg	712	VTDKDEKDDI VD
TtParg	-----	-----
HsPARG	-----	-----
EdParg2	-----	-----
consensus	1171	-----

Figure 5.8. Several Parg sequences have been aligned for the identification of the conserved amino acid, few of them are reported as representative alignment. *Drosophila melanogaster* Parg (DmParg), *Tetrahymena thermophila* Parg (TtParg), *Homo Sapiens* PARG (HsPARG), *Entamoeba dispar* Parg2 (EdParg2). Catalytic glutamates are highlighted in red, additional C-terminal region of DmParg highlighted in green, conserved amino acid in *Drosophila* and *Tetrahymena* are highlighted in yellow.

5.4 Removal of mono-Ser-ADP-ribosylation by Parg is conserved in *Tetrahymena*

The existence of the unique additional C-terminal region in DmParg led me to think that the removal of mono-Ser-ADP-ribosylation by Parg was unique to the *Drosophila* genus. However, after identification of the minimal region of DmParg for the removal of mono-Ser-ADP-ribosylation and in light of the dispensability of the C-terminal domain for this activity, I decided to verify this assumption. Therefore, I aimed to investigate whether the removal of mono-Ser-ADP-ribosylation could be a conserved mechanism across different species. For this purpose, I probed the activity of Parg homologs from two other species: the ciliate *Tetrahymena thermophila* Parg (TtParg) and the amoeba *Entamoeba dispar* Parg2 (EdParg2). Together with DmParg these two proteins represent a wide spectrum of evolutionary diversity. After the purification of the two new Parg homologs from *E.coli*, I tested their activity against the established substrate auto-modified PARP1 (**Figure 5.9A**). The assay showed clearly that WT variants of all tested Pargs are perfectly able to hydrolyse poly-ADP-ribosylation from PARP1, leaving a single band corresponding to mono-Glu-ADP-ribosylation PARP1. The activity was abrogated in a catalytic mutant of TtParg, E256Q. After confirming the ability of TtParg and EdParg to hydrolyse poly-ADP-ribosylation, I aimed to investigate their ability to remove mono-Ser-ADP-ribosylation from the ADP-ribosylated H3 histone peptide (**Figure 5.9B**). Surprisingly, TtParg showed the same activity as DmParg in the removal of mono-Ser-ADP-ribosylation, and this activity was lost in the catalytically dead E256Q mutant. EdParg2, on the other hand, was not able to cleave mono-Ser-ADP-ribosylation, thus classifying along PARG among poly-ADP-ribosylation-specific hydrolases. These assays confirmed that the removal of mono-Ser-ADP-ribosylation by Parg enzyme is not

unique to *Drosophila* but rather a mechanism shared by at least one other organism. Moreover, the identification of TtParg activity will increase our chance to understand this mechanism since high resolution structures of TtParg have been already reported (Barkauskaite *et al.*, 2013).

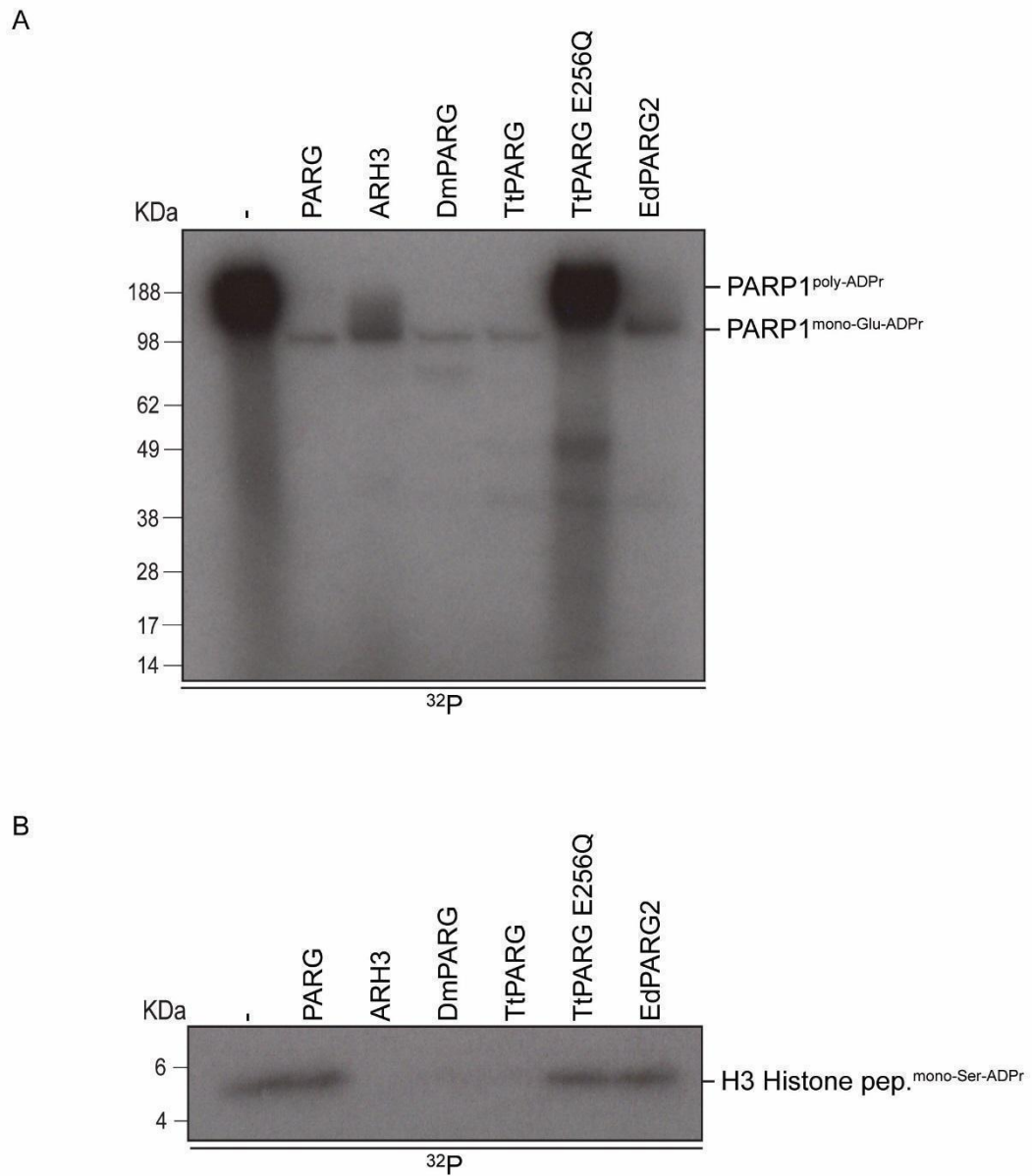


Figure 5.9. A - Recombinant PARP1 (1 μ M) was auto-modified in presence of DNA (1 μ M), NAD⁺ (50 μ M), ³²P-NAD⁺ for labelling. The reactions were incubated with Parg from different species, and ARH3 (3 μ M) for 20 minutes at room temperature. **B** - mono-Ser-ADP-ribosylation H3 histone peptide (40 μ M) was incubated with Parg from different species, and ARH3 (3 μ M) for 20 minutes at room temperature. The reactions were analysed by SDP-PAGE and autoradiography.

5.5 Structural characterisation of DmParg

In order to understand the difference between human PARG and Drosophila Parg, and make a valid model to explain the different activity toward mono-Ser-ADP-ribosylation, I decided to solve the structure of DmParg. DmParg crystallisation attempts are summarised in (**Table 4.1**). Whilst full-length DmParg resisted crystallisation, better results were obtained with DmParg $\Delta 554-723$ (**Figure 5.10**). Initial crystals of DmParg $\Delta 554-723$ were obtained after just a few days at 20°C. Although the crystals were well-formed, with sharp facets, they did not diffract. In order to improve protein organisation in the crystal lattice, I decided to slow down the crystallisation process with two approaches: first by decreasing the temperature from 20°C to 4°C, and second by decreasing the concentration of protein from 10 mg/mL to 8-6 mg/mL. Since protein concentration could have been a limiting factor for crystal nucleation, I used some of the non-diffracting crystals obtained earlier with higher concentrations for microseeding. Different combinations of these approaches led to the formation of crystals that diffracted to a final resolution of a 2.2 Å (**Figure 5.11**). During this phase, I also tried to crystallise DmParg with a few ligands such as ADP-ribose, ADP-HPD and a commercial PARG inhibitor (PDD00017273). Finally, DmParg co-crystallised efficiently in the presence of PDD00017273, yielding the first structure of this inhibitor in complex with any PARG homolog. As of today, I have collected many datasets and obtained phases by molecular replacement using the human PARG, thus generating preliminary models of apo DmParg and DmParg in complex with the inhibitor

Protein name	Note	Temp	mg/mL	drop size a/ drop size b	drop a/drop b	Crystal buffer
DmParg FL	Crystals with no diffraction	20°C	10	125+125/125+375	apo/apo	
DmParg FL		20°C	9	125+125/125+375	600uM ADP/r/600uM ADP-HPD	
DmParg FL		20°C	9	125+125/125+375	600uM ADP/r/600uM ADP-HPD	
DmParg Δ558-723	Crystals with no diffraction	20°C	10	125+125/125+375	apo/apo	
DmParg Δ558-723		20°C	13.6	125+125/125+375	apo/apo	
DmParg Δ558-723		20°C	9	125+125/125+375	600uM ADP/r/600uM ADP-HPD	
DmParg Δ558-723		20°C	9	125+125/125+375	600uM ADP/r/600uM ADP-HPD	
DmParg FL		4°C	8	125+25(seed)+125	microseeding from CS1-G111a/	
DmParg Δ558-723		4°C	8	125+25(seed)+125	microseeding from PACT-D5a/	
DmParg FL	~15Å- ~20Å	4°C	8	125+125/125+375	microseeding from Midas B8a	
DmParg Δ558-723		4°C	8	125+125/125+375	apo/apo	
DmParg Δ558-723	~3.4Å	4°C	8	125+125/125+375	apo/apo	
DmParg NTR50		4°C	8	125+125/125+375	apo/apo	
DmParg Δ558-723		4°C	8	125+125/125+375	1mM ADP/r/1mM ADPr	
DmParg Δ558-723	~2.8Å	4°C	8	125+125/125+375	500uM Inhibitor PDD00017273/ 500uM Inhibitor PDD00017273	Sodium Sulphate 0.2M, BisTris Propane pH6.5 0.1M, PEG3350 20%
DmParg Δ558-723		4°C	8	100+100/100+300	1mM H3 histone peptide/ 1mM H3 histone peptide	
DmParg Δ558-723 E339N		4°C	8	100+100/100+300	1mM Ser-H3 histone/ 1mM Ser-H3 histone	
DmParg Δ558-723 E339N		4°C	8	100+100/100+300	1mM Ser-H2B histone/ 1mM Ser-H2B histone	
DmParg Δ558-723 E339N		4°C	8	100+100/100+300	Branched ADPr/Branched ADPr	
DmParg Δ558-723	~2.7Å	4°C	8	100+25(seed)+125/ 100+15(seed)+375	microseeding	Sodium Sulphate 0.2M, BisTris Propane pH6.5 0.1M, PEG3350 20%
DmParg Δ558-723	~2.6Å	4°C	8	100+25(seed)+125/ 100+15(seed)+375	Microseeding, 500uM Inhibitor PDD00017273/ Microseeding, 500uM Inhibitor PDD00017273	Sodium Sulphate 0.2M, BisTris Propane pH6.5 0.1M, PEG3350 20%
DmParg Δ558-723	~2.4Å	4°C	6	125+125/125+375	microseeding/microseeding	19% v/v PEG3350, 0.21M Sodium Sulphate, 0.1M BisTris Propane pH6.3, 0.1M BisTris Propane pH7.2
DmParg Δ558-723	~2.5Å	4°C	6	125+125/125+375	Microseeding, 500uM Inhibitor PDD00017273/ Microseeding, 500uM Inhibitor PDD00017273	19% v/v PEG3350, 0.21M Sodium Sulphate, 0.1M BisTris Propane pH6.3, 0.1M BisTris Propane pH7.2
DmParg Δ558-723	~2.4Å	4°C	6	125+125/125+375	Microseeding, additive screening/ Microseeding, additive screening	19% v/v PEG3350, 0.21M Sodium Sulphate, 0.1M BisTris Propane pH6.3, 0.1M BisTris Propane pH7.2
DmParg Δ558-723	~2.5Å	4°C	6	125+125/125+375	Microseeding, additive screening, 500uM Inhibitor PDD00017273/ Microseeding, additive screening, 500uM Inhibitor PDD00017273	19% v/v PEG3350, 0.21M Sodium Sulphate, 0.1M BisTris Propane pH6.3, 0.1M BisTris Propane pH7.2

Table 4.1. DmParg crystallisation attempts. Expression system: *E.coli* Rosetta DE3 Protein buffer: 10mM Tris-HCl pH8, 100mM NaCl, 0.2mM TCEP.

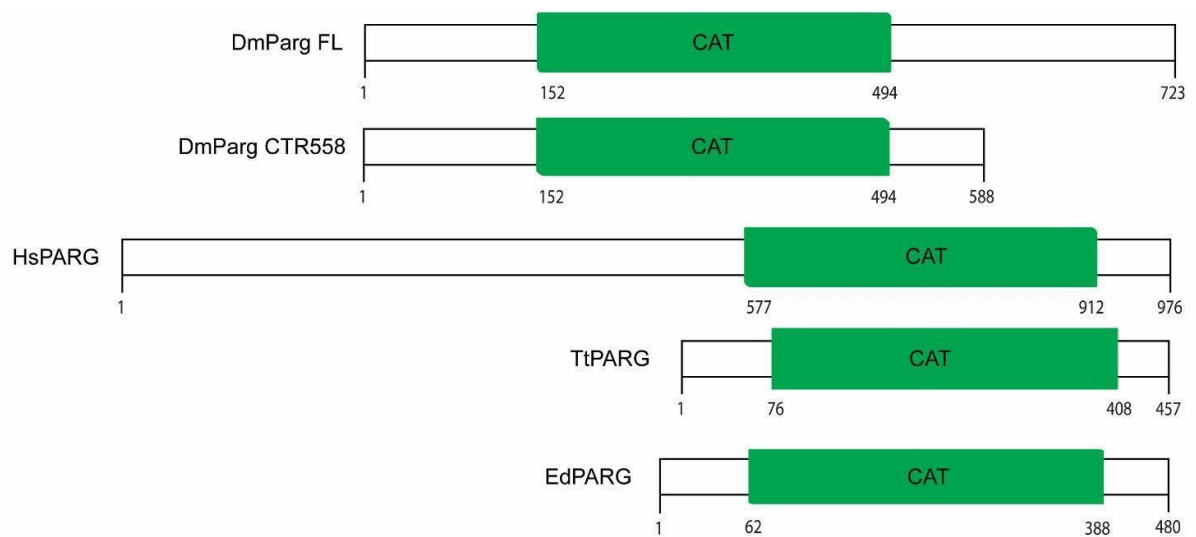


Figure 5.10. Schematic representation of PARG domain, in green the macrodomain or catalytic domain. *Drosophila melanogaster* (DmParg), *Homo Sapiens* PARG (HsPARG), *Tetrahymena thermophila* Parg (TtParg), *Entamoeba dispar* Parg2 (EdParg2).

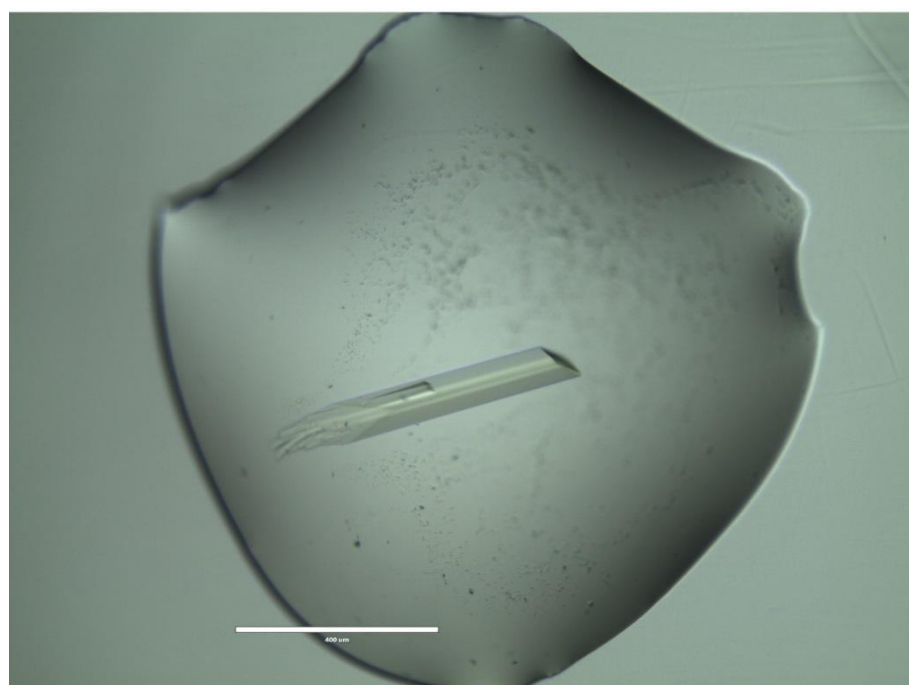
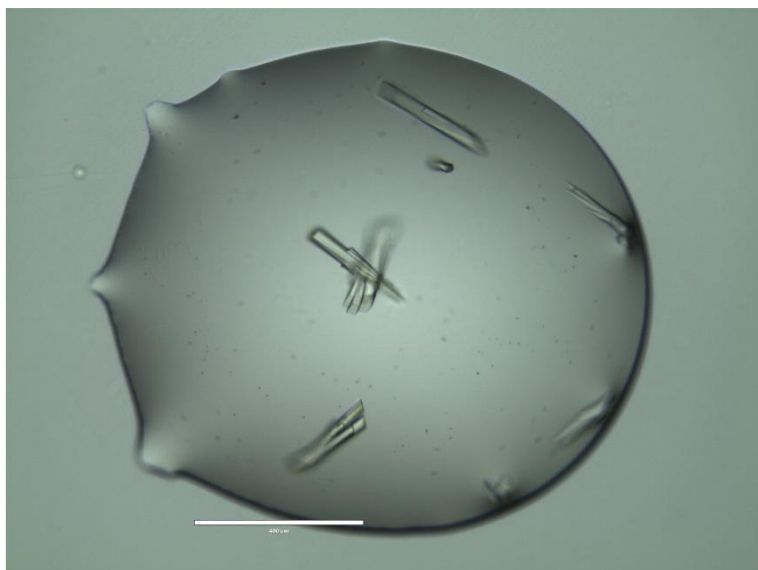


Figure 5.11. Representative DmParg Δ 554-723 apo-crystal grow at 4°C. Protein concentration 6mg/mL. Drop size 250 μ L; 125 μ L protein + 125 μ L crystallisation buffer. Protein buffer: 10mM Tris-HCl pH8, 100mM NaCl, 0.2mM TCEP. Crystallisation buffer: 19% v/v PEG3350, 0.21M Sodium Sulphate, 0.1M BisTris Propane pH 7.2.

5.6 Uncoupling mono-Ser-ADP-ribosylation and poly-ADP-ribosylation glycohydrolase activities

Hanai *et al.*, 2014 reported that knock-out of DmParg in *Drosophila* led to neurodegeneration. This neurodegeneration was of course linked by the authors to the inability of DmParg to remove poly-ADP-ribosylation as this was its only known function. Surprisingly, this phenotype appears unique to *Drosophila* as mutations in Parg homologs in other organisms have never been reported to lead to neurodegeneration. However, in human neurodegeneration has been linked to mutations in TARG1, a hydrolase that removes mono-Glu-ADP-ribosylation (Sharifi *et al.* 2013) and just very recently in ARH3 (Ghosh *et al.*, 2018). In light of the role of terminal-ADP-ribose in determining neurodegeneration, and my finding that DmParg can hydrolyse mono-Ser-ADP-ribosylation in addition to poly-ADP-ribosylation, I postulated that the neurodegeneration in *Drosophila* caused by Parg knock-out might be related to the loss of mono-Ser-ADP-ribosylation removal rather than poly-ADP-ribosylation hydrolysis, linking for the first time mono-Ser-ADP-ribosylation with a *Drosophila* disease. To explore this possibility I aimed to make a DmParg mutant with similar activity to human PARG, i.e. able to hydrolyse poly-ADP-ribosylation and unable to remove mono-Ser-ADP-ribosylation. Such DmParg mutants could be transfected into *Drosophila* DmParg knock-out for testing whether they compensate for the WT protein and rescue the neurodegeneration phenotype. This provided an additional incentive to identify the secondary region(s) that might be necessary for Ser-mono-ADP-ribosylation but not for poly-ADP-ribosylation hydrolysis. Therefore, I reanalysed the alignment of Pargs sequences from several different *Drosophila* species, human PARG, TtParg and EdParg2 and look for amino acids conserved only in *Drosophila* species and TtParg, which could be responsible for cleave mono-Ser-ADP-ribosylation, but not

conserved in PARG or EdParg2 (**Figure 5.8**). I reasoned that conserved amino acid residues that are not directly present in the catalytic site could be the key to explaining the specificity of DmParg and TtParg toward mono-Ser-ADP-ribosylation. From the alignment, I identified 7 such conserved amino acid residues (F231, L333, C347, V352, F384, E423 and E546), 3 of which were located close to the catalytic loop. I then mutated all of these 7 amino acid residues independently into alanine and tested the activity of the mutants against auto-modified PARP1 (**Figure 5.12A**). The assay showed that all the mutants are still active against poly-ADP-ribosylation, leaving the single band corresponding to mono-Glu-ADP-ribosylation. I further tested the 7 mutants against mono-Ser-ADP-ribosylated histone H3 peptide (**Figure 5.12B**). Surprisingly, I found that 4 out of the 7 mutants, DmParg F231A, L333A, C347A and F384A, had difficulty in removing the mono-Ser-ADP-ribosylation from H3 peptide, providing support for the hypothesis that these amino acid residues may selectively affect the removal of mono-Ser-ADP-ribosylation but leaving the PAR hydrolysis intact. To confirm whether these mutants have just overall lower catalytic activities to any substrate or complete separation of function, enzyme kinetics against the mono-Ser-ADP-ribosylation and poly-ADP-ribosylation substrates will be performed and compared to WT protein in the future. Confirming the uncoupling of function, of these mutants, would establish a valuable tool for *in vivo* study in *Drosophila* which could link Ser-ADP-ribosylation and neurodegeneration. Moreover, we could gain a better understanding of how the mechanism of DmParg is linked to its structure.

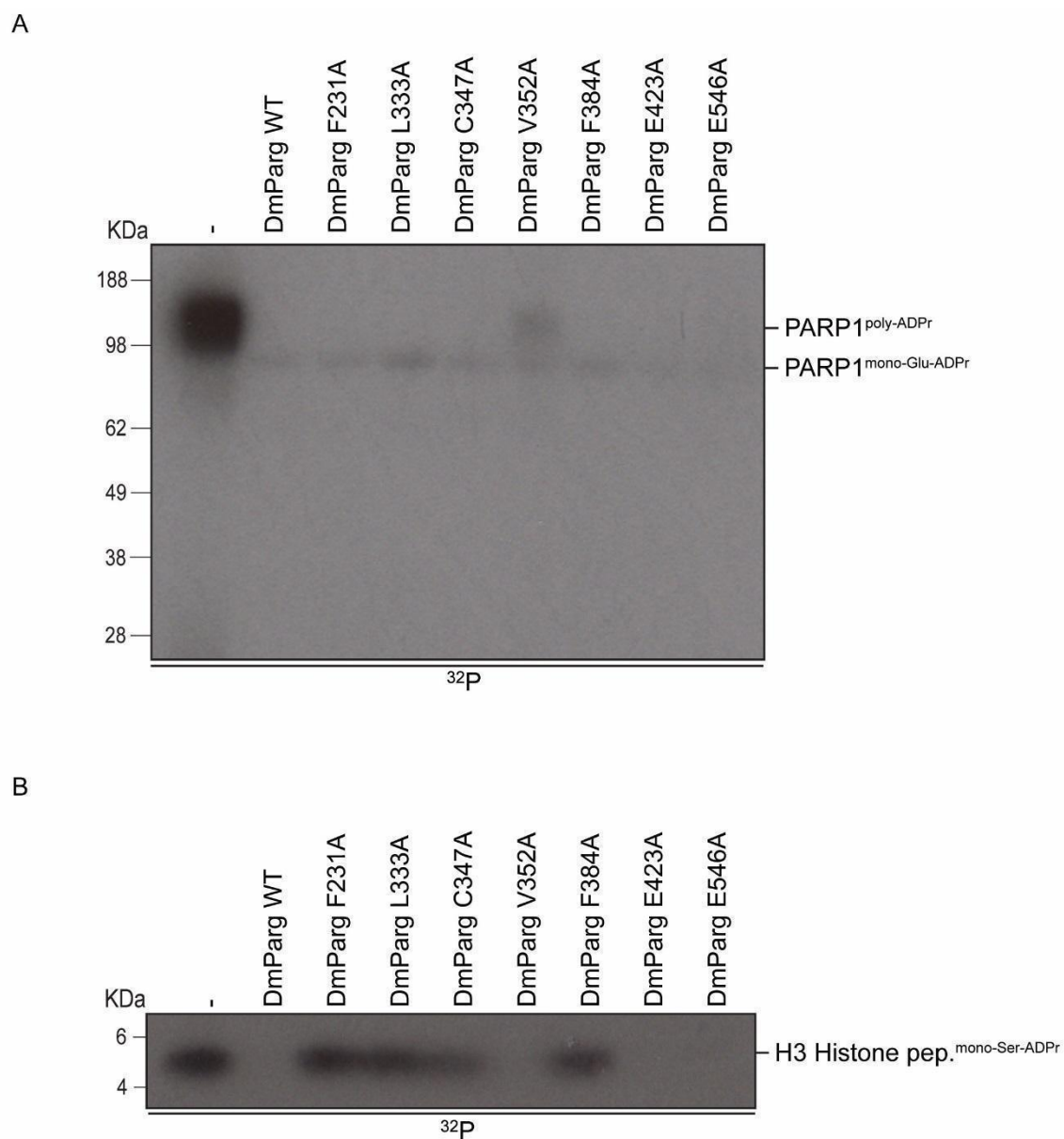


Figure 5.12. A - Recombinant PARP1 (1 μM) was auto-modified in presence of DNA (1 μM), NAD^+ (50 μM), ^{32}P - NAD^+ for labelling. The reactions were incubated with DmParg WT and mutants (3 μM) for 20 minutes at room temperature. **B** - Mono-Ser-ADP-ribosylated H3 histone peptide (40 μM) was incubated with DmParg WT and mutants (3 μM) for 20 minutes at room temperature. The reactions were analysed by SDP-PAGE and autoradiography.

5.7 DmParg specificity assays

I finally asked whether DmParg shows activity against ADP-ribose bound to other amino acids such as glutamate, arginine and lysine, as previously tested for ARH3. First I investigated if DmParg can remove mono-Glu-ADP-ribosylation. As a substrate, I used the auto-modified PARP1 E988Q, which is deficient in poly-ADP-ribosylation activity and therefore auto-modifies itself with mono-Glu-ADP-ribosylation (Sharifi *et al.*, 2013). After PARP1 E988Q auto-modification, I incubated it with a range of glycohydrolases (**Figure 5.13**). Although some decrease of the signal of the auto-modified PARP1 E988Q in the presence of DmParg was observed, it was weaker than for the positive control TARG1, and also weaker than for mono-Ser-ADP-ribosylation hydrolysis under comparable reaction conditions. Therefore, I concluded that DmParg does not have a robust activity against the terminal-ADP-ribose linked to glutamate. I then tested the *Drosophila* glycohydrolase against arginine-linked ADP-ribosylation. As a substrate, I used *Drosophila* S2 extract cell treated with recombinant ARTC2.2 (**Figure 5.14**). The assay showed that the positive control ARH1 was able to efficiently remove the mono-Arg-ADP-ribosylation. None of the other glycohydrolases, including DmParg, showed any discernible activity against the substrate. This assay further suggests that DmParg is specific for mono-Ser-ADP-ribosylation. Finally, I tested each of the glycohydrolases against a chemically generated lysine-linked ADP-ribosylation (Jankevicius *et al.*, 2013) (**Figure 5.15**). Again, none of the tested glycohydrolase exhibited any activities, confirming that DmParg specificity is restricted to mono-Ser-ADP-ribosylation.

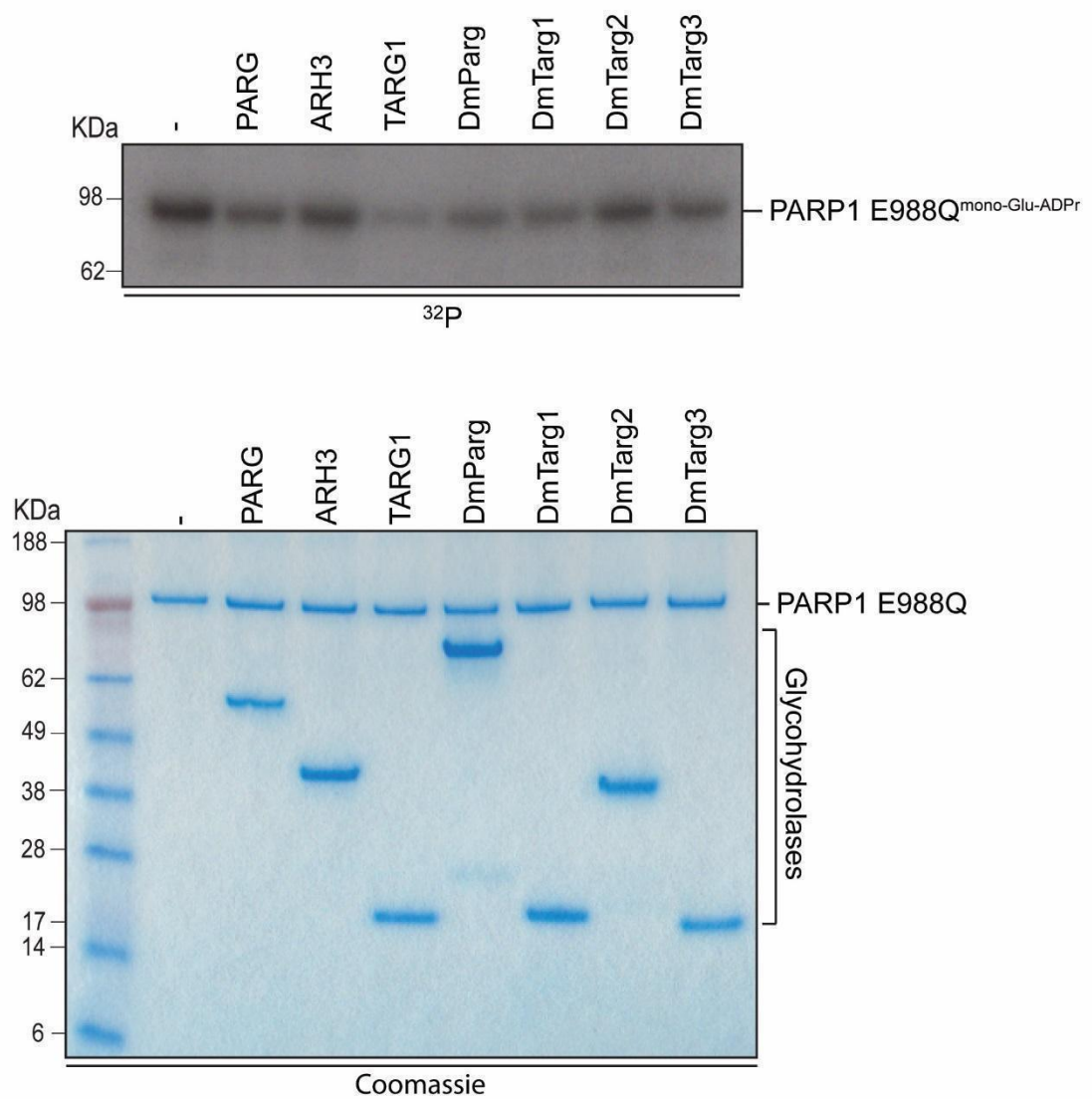


Figure 5.13. Auto-modified PARP1 E988Q (1 μ M) was incubated with human and *Drosophila* glycohydrolases (3 μ M) 20 minutes at room temperature. The reactions were analysed by SDS-PAGE and autoradiography.

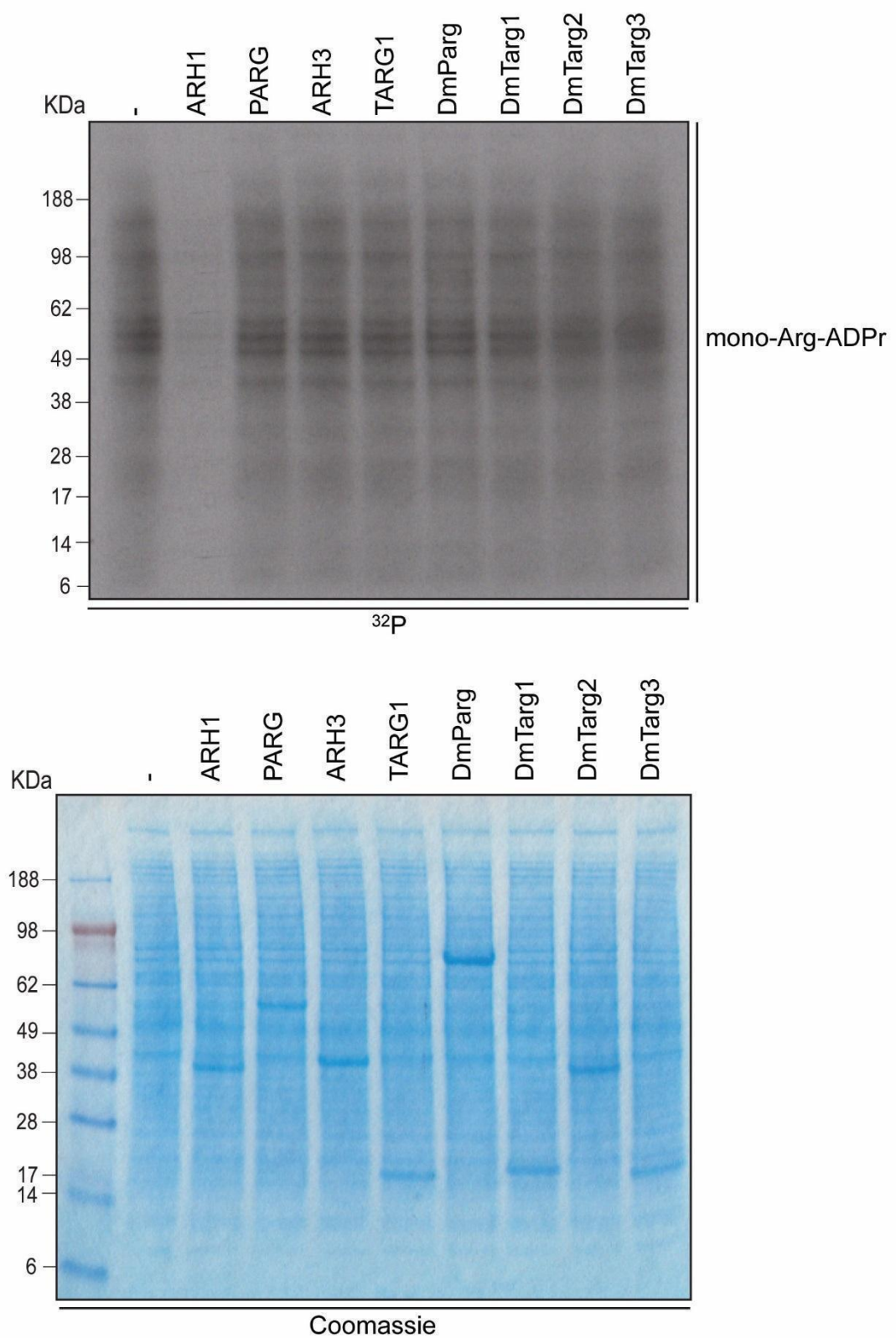


Figure 5.14. Mono-Arg-ADP-ribosylated cell extract was incubated with human and *Drosophila* glycohydrolases (3 μM) for 15 minutes at 30°C. The reactions were analysed by SDP-PAGE and autoradiography.

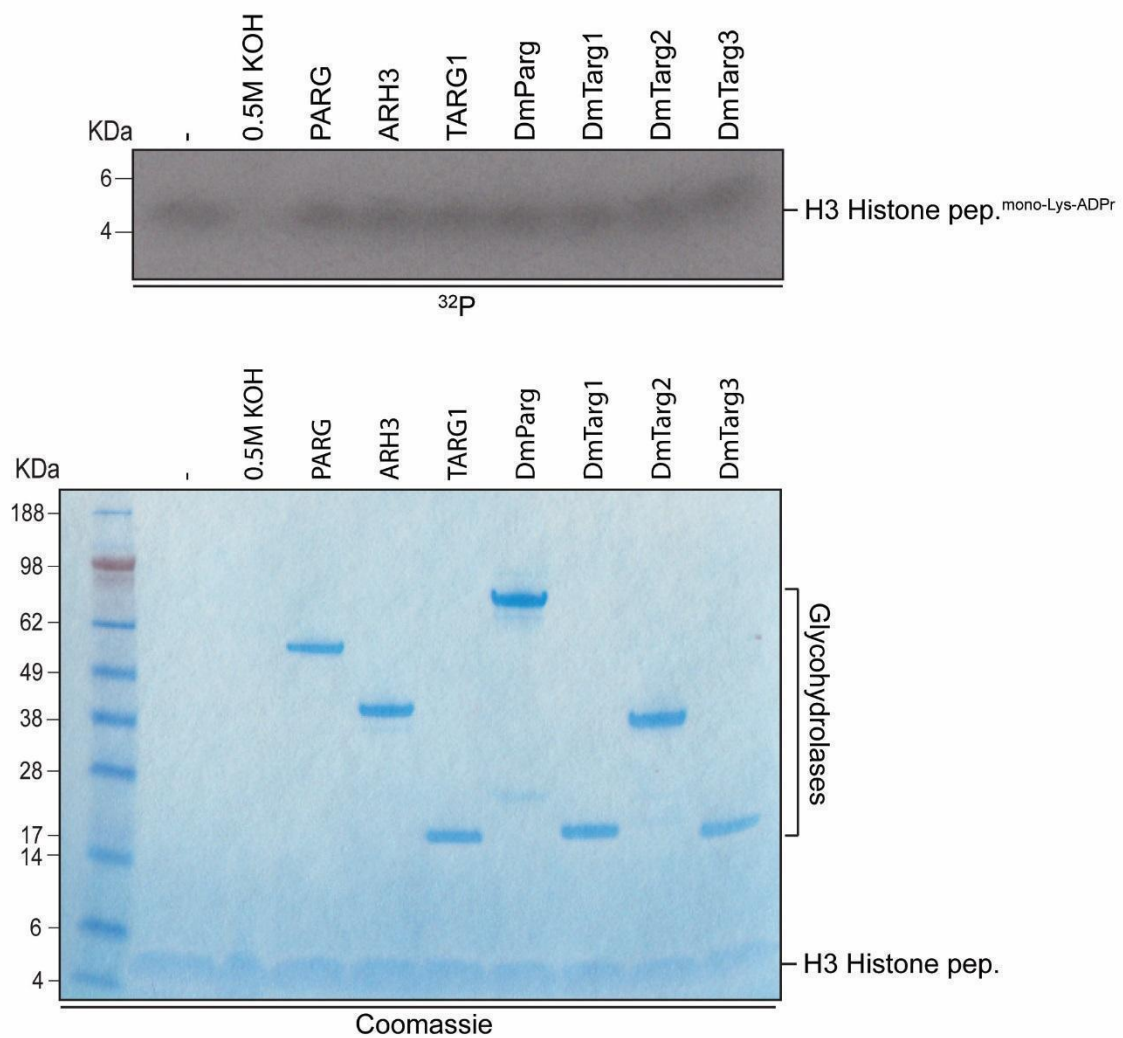


Figure 5.15. Mono-Lys-ADP-ribosylated H3 peptide (100 μ M) was incubated with human and *Drosophila* Glycohydrolases (3 μ M) for 30 minutes at room temperature or with 0.5M KOH for 15 minutes at 37°C. The reactions were analysed by SDP-PAGE and autoradiography.

5.8 Conclusions on Ser-ADP-ribosylation in *Drosophila melanogaster*

Within the last year of my DPhil, I aimed to investigate Ser-ADP-ribosylation in *Drosophila melanogaster*. In this project, I uncovered a new aspect of Ser-ADP-ribosylation regulation which will enhance our understanding of the evolution and function of ADP-ribosylation. Firstly, as described in Chapter 3 of this thesis, I have reconstructed *in vitro* Ser-ADP-ribosylation, this time using the *Drosophila* homologs of PARP1 and HPF1, namely DmParp1 and DmHpf1. I confirmed that the ADP-ribosylation produced in this reaction is indeed bound to a Ser residue and finally identified *Drosophila* DmParg as the *Drosophila* enzyme able to remove mono-Ser-ADP-ribosylation. This was followed by a biochemical characterisation to uncover the mechanism of DmParg. Subsequently, I showed that a Parg homolog from *Tetrahymena thermophila* also possessed the ability to remove Ser-ADP-ribosylation, suggesting a conserved mechanism among different organism. Moreover, I aimed to investigate the difference between PARG and DmParg using a crystallography approach, so far obtaining preliminary data which will be analysed right after the submission of this thesis. Additionally, I aimed to identify the amino acid residues in DmParg which play an essential role in the removal of Ser-ADP-ribosylation but are dispensable for poly-ADP-ribosylation. With this approach, I managed to identify a few promising candidates which will be subjected to further characterisation. Finally, I show that DmParg has specificity for Ser-ADP-ribosylation and does not efficiently remove mono-ADP-ribosylation linked to glutamate, arginine or lysine. To conclude, I propose that DmParg will be a useful tool for future study of ADP-ribosylation as its activity appears to combine those of human PARG and ARH3.

Chapter 6. Discussion

6.1 Aims of the project

The initial aim of this thesis was the identification and characterisation of a novel enzyme/regulator of PARP1 cycle with a particular emphasis of its role in DNA repair. After the initial identification of HPF1 and its structural and biochemical characterisation, the project took a different path with the identification of the Ser-ADP-ribosylation and its important role in DNA repair. The project evolved furtherly with the identification and biochemical characterisation of ARH3 as the human enzyme able to reverse Ser-ADP-ribosylation. Finally, during the last stage of my DPhil, I began the characterisation of Ser-ADP-ribosylation in the model organism *Drosophila melanogaster* which led me to the identification of Parg as the only *Drosophila* enzyme able to reverse Ser-ADP-ribosylation, followed by the structural and biochemical characterisation of this activity. The identification of Ser-ADP-ribosylation has a big impact in the field of ADP-ribosylation as the univocal identification of residues modified by ADP-ribosylation has been a field of study for many years. Despite our still limited knowledge of all the amino acid modified by ADP-ribosylation, the identification of Ser-ADP-ribosylation can be seen as a small piece of the very complexes PTM and epigenetic worlds.

6.2 Identification of HPF1

In Gibbs-Seymour *et al.* 2016, we first show the identification of a novel PARP1 interactor protein involved in DNA damage response which we named HPF1. We show that HPF1 interaction with PARP1 promotes trans-ADP-ribosylation of histones and suppresses PARP1 auto-modification. Functional characterisation shows that HPF1 can be recruited to the DNA damage site in a PARP1-dependent way and the use of PARP1 inhibitor does not affect HPF1 recruitment. Moreover, immunoprecipitation of HPF1 shows that HPF1 can efficiently pull down both PARP1 and histones. Both functional and *in vitro* characterisation data show that, in the absence of HPF1, PARP1 cannot modify histones and hyper-automodifies itself. We suggest that HPF1 is not the only protein able to form a complex with PARP1, as protein quantification from HeLa cells (Hein *et al.*, 2015) shows approximately twenty-fold excess of PARP1 compared to HPF1 (Gibbs-Seymour *et al.* 2015). Indeed, another PARP1 interactor, Timeless, was identified just a few months before HPF1 identification (Xie *et al.*, 2015, Young *et al.*, 2015). Like HPF1, Timeless can be recruited to the DNA damage site in a PARP1-dependent way; however, no evidence has been found of Timeless regulating PARP1 ADP-ribosylation activity (Xie *et al.*, 2015), suggesting different roles for the PARP1/HPF1 and PARP1/Timeless complexes. Both factors, however, interact with the catalytic domain of PARP1. I postulate that HPF1 interacts through its C-terminal domain with the catalytic domain of PARP1/PARP2, as suggested by the functional characterisation, and through its N-terminal domain with the histone substrates, allowing a better orientation of the substrate toward PARP1/PARP2 catalytic site (**Figure 6.1**). Moreover, the interaction with the catalytic site is likely to induce a conformational change in the active site of PARP1/PARP2,

inducing specificity for serine. These models could be verified by structural characterisation of the heterodimer complexes HPF1/ PARP1 and Histone/HPF1, and of the ternary complex Histone/HPF1/PARP1.

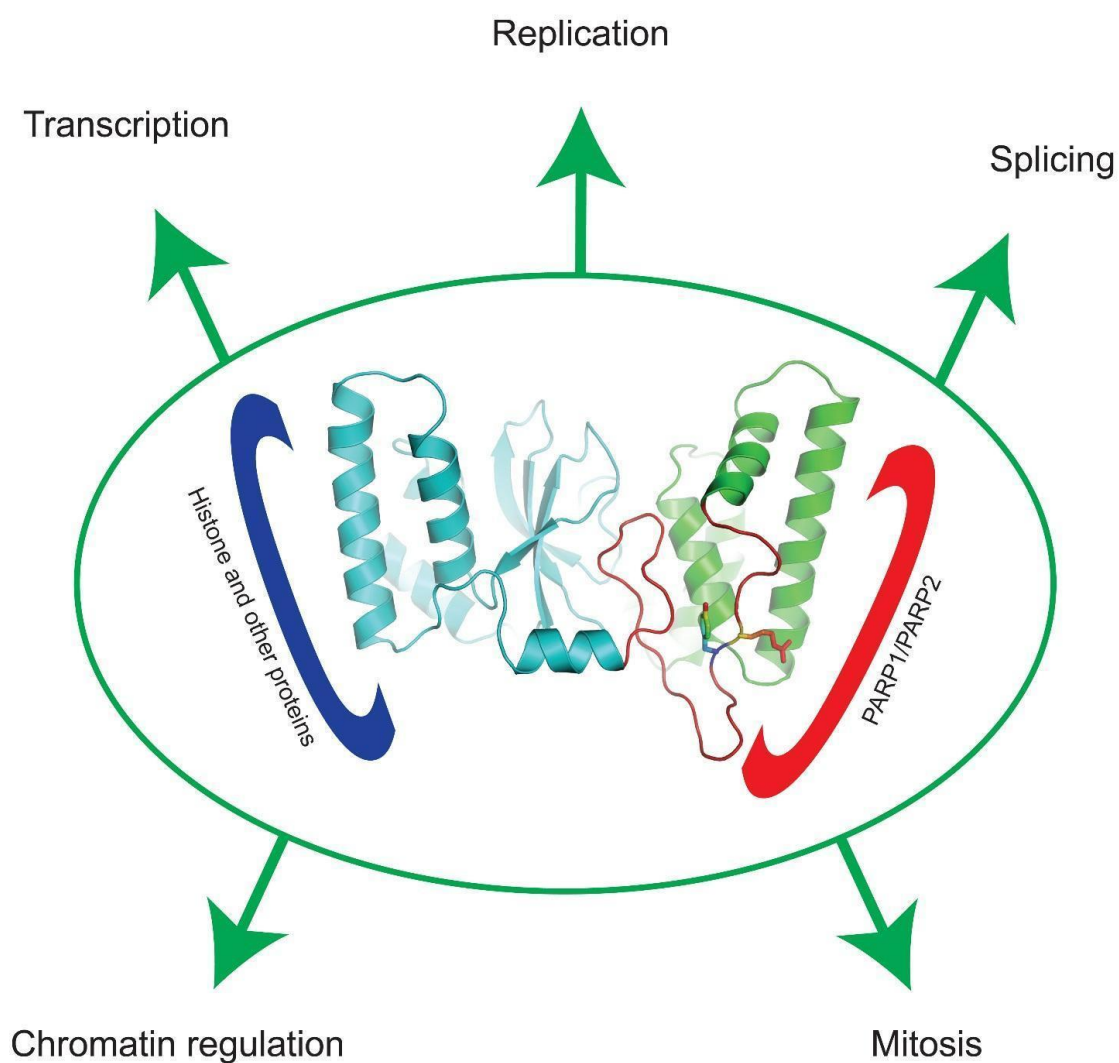


Figure 6.1. NvHpf1 structure showing possible interaction region with PARP1/ PARP2 and histone.

6.3 Ser-ADP-ribosylation HPF1 dependent

Despite the identification of HPF1, the study of PARP1 regulation and activity remains an open field. One of the biggest challenges to overcome is the identification of the amino acid targeted by ADP-ribosylation. In DNA repair, glutamate and aspartate have been considered for a long time as the main ADP-ribosylation acceptors. However, in collaboration with the Matic group, our lab was first to identify a novel form of ADP-ribosylation named Ser-ADP-ribosylation occurring on histone proteins (Leidecker *et al.*, 2016). Following this discovery, we identified HPF1 as the factor that switches PARP1 specificity from glutamate/aspartate to serine. Simultaneously, hundreds of Ser-ADP-ribosylation sites have been identified in human cells using mass spectrometry (Bonfiglio *et al.*, 2017). Despite Ser-ADP-ribosylation HPF1 dependent seems to play an important role in DNA damage response, our collaborator identified several other proteins targeted by this PTM and involved in other essential cellular pathways such as transcription, replication, splicing, mitosis and chromatin regulation. Some of the identified protein include high-mobility group (HMG), chromatin remodelers, histone and DNA methylation systems, ribonucleoprotein particle (HNRNP) proteins, zinc-finger-containing factors and many proteins involved in genome stability (DNA ligase, DNA-PKScs, BLM, XPC, MSH6, Polbeta, FANCI, and Rad54B). The varieties of proteins targeted by Ser-ADP-ribosylation suggest the important role of HPF1 in several physiological pathways (**Figure 6.1**).

Moreover, our collaborator manually reprocessed published proteomic dataset (Martello *et al.*, 2016) looking for information which could suggest serine as a possible substrate for ADP-ribosylation and identified more than 250 residues. More recently, Larsen *et al.*, 2018, implemented a new mass spectrometry protocol and identified 3090 targets for Ser-ADP-ribosylation, hundreds of which overlapped with

phosphorylation sites, suggesting an important cross-talk between the two modifications. However, how many of these targets are dependent from HPF1 remains an open question.

To date, many techniques have been implemented to analyse ADP-ribosylated substrates in order to identify the acceptor site of the modification. After the first identification of Ser-ADP-ribosylation, Bilan *et al.*, 2017 implemented a new mass spectrometry-based technique and confirmed the KS motif as a possible target of ADP-ribosylation; however, these authors suggested both lysine and serine as possible substrates.

Whether the lysine is a real substrate for ADP-ribosylation remains to be determined, as no ARTs or erasers for this modification have been identified to date. Two recent studies have also identified tyrosine as a not abundant but potentially physiologically important alternative acceptor site of PARP1- and HPF1-dependent ADP-ribosylation (Leslie *et al.*, 2018; Bartlett *et al.*, 2018).

Following the identification of HPF1 and the Ser-ADP-ribosylation by the complex HPF1/PARP1, a postdoc in our lab, Dr. Edward Bartlett, investigated in more detail the Ser-ADP-ribosylation site in the histone tails. In his manuscript (Bartlett *et al.*, 2018), using both *in vitro* and functional characterisation, him and his collaborators identified a crosstalk between Ser-ADP-ribosylation and lysine acetylation as mutually exclusive. Acetylation of lysine 9 of histone H3 prevents Ser-ADP-ribosylation from occurring at the adjacent serine 10 of histone H3; this mechanism also works in the opposite direction: Ser-ADP-ribosylation will prevent the acetylation of an adjacent lysine. The same phenomenon was observed for phosphorylation of serine which will prevent the Ser-ADP-ribosylation and *vice versa*, in this case by direct competition for the same hydroxyl group on serine. These findings not only show that Ser-ADP-ribosylation is involved in the crosstalk with other PTMs on histone tails but

add a new piece to the complex picture of epigenetics. These findings partially overlap with a recent report from Tom Muir's laboratory (Liszczyk *et al.*, 2018).

In order to fully understand the mechanism through which HPF1 regulates PARP1 and induce Ser-ADP-ribosylation, in the third chapter of his thesis, I aimed to identify a minimal model for Ser-ADP-ribosylation which would likely increase our chance of co-crystallisation of the complex PARP1-HPF1. Firstly, as I have already solved two apo structures of HPF1, I considered to co-crystallise the full length PARP1/HPF1 or the PARP2/HPF1 complex. I would expect that the co-structure would show the structural change occurring in PARP1 or PARP2 upon HPF1 binding. As the size of the PARP1/HPF1 or the PARP2/HPF1 complex would respectively be 155 kDa and 110 kDa, I decided to identify the minimal region of PARP1 and PARP2 required for Ser-ADP-ribosylation. Previously, using a GST pull-down assay, I showed that the catalytic domain of either PARP1 or PARP2 is enough for the interaction with HPF1. However, whether such a complex is active or not remained an open question. The assay in **Figure 3.14** suggests that the zinc fingers are essential domains for Ser-ADP-ribosylation. Likely this is due to their ability to bind DNA which leads to the activation of PARP1. Despite the fact that HPF1 can interact with the isolated catalytic domain of PARP1, the latter is likely in an inactive conformation and so unable to modify any substrate. Further characterisation will be necessary in order to identify the minimal region of PARP1 for Ser-ADP-ribosylation. For instance, Langelier *et al.*, 2012 showed that ZnF2 and BRCT are not essential for PARP1 activity. Thus testing the HPF1-dependent Ser-ADP-ribosylation activity of a truncated PARP1 variant lacking these two domains, but retaining the other two zinc fingers, would tell us whether ZnF2 and BRCT are important for histone modification. If not, the construct devoid of these two domains could serve as a simpler PARP1 model for structural characterisation.

My results also opened the question of whether the activation by DNA is necessary for Ser-ADP-ribosylation. The presence of DNA would increase the complexity of the complex, decreasing the likelihood of crystallisation. For this reason, I postulated that a DNA-independent PARP1 mutant would be a better model for co-crystallisation with HPF1. In **Figure 3.15**, I investigated the activity of PARP1 Δ H_{HD}, a PARP1 DNA independent, and showed its ability to modify histone protein in the presence of HPF1. Compared to PARP1 WT, PARP1 Δ H_{HD} retains partial Ser-ADP-ribosylation activity even in absence of DNA. If this finding will be confirmed with proper protein preparation, it will greatly simplify the challenge of co-crystallisation of the PARP1/HPF1 complex, as the PARP1 Δ H_{HD}/HPF1 complex is overall smaller and does not need DNA for Ser-ADP-ribosylation.

To sum up, in the third chapter of this thesis I report the structural and biochemical characterisation of the previously uncharacterised DNA damage repair protein HPF1 and identify it as a PARP1 and histone interactor. I also report HPF1 as a specificity factor for PARP1- and PARP2-dependent Ser-ADP-ribosylation of histones and other substrates. Despite the fact that the link connecting ADP-ribose to serine resembles the ADP-ribose-ADP-ribose bond, HPF1 seems to play a role just in the formation of the former but not the latter. However, how HPF1 regulates PARP1/PARP2 auto-modification and induces specificity toward serine modification remains an open question, which could be addressed by further structural characterisation of PARP1/PARP2 in complex with HPF1. In the future, many more questions will need to be addressed on the role of HPF1, for instance: Does HPF1 bind directly to DNA? Does HPF1/PARP1 complex play a role in transcription or homologous recombination? Can inhibition of HPF1 have a therapeutic function? Moreover, the newly identified Ser-ADP-ribosylation opens many questions on the mechanism of its

regulation and also on the implication of other PTMs on serine or other amino acids in the proximity of Ser-ADP-ribosylation.

6.4 Removal of Ser-ADP-ribosylation by ARH3

Following the discovery of Ser-ADP-ribosylation and the *in vitro* reconstruction of the reaction, the focus of my DPhil shifted to the identification of the enzyme erasing Ser-ADP-ribosylation, which is described in Chapter 4 of this thesis. The discovery of Ser-ADP-ribosylation raised the question which enzyme could eventually reverse this novel post-translational modification. Here, I identified ARH3 as the only human enzyme able to efficiently remove Ser-ADP-ribosylation. I showed ARH3 activity is magnesium-dependent and identified two key aspartate residues in the catalytic site. Although ARH3 had been known previously as a glycohydrolase able to cleave PAR-chains and O-acetyl-ADP-ribose, its efficiency has been found inferior to PARG in the removal of long PAR-chains and its real cellular role was not completely understood (Oka *et al.*, 2006; Ono *et al.*, 2006; Mashimo *et al.*, 2013; Fontana *et al.*, 2017). However, ARH3 seems to be more efficient in the cleavage of short PAR-chains compared with PARG (Hatakeyama *et al.*, 1986, Mashimo *et al.*, 2013). ARH3 can be found in different cellular compartments: cytosol, nucleus and mitochondria (Mashimo *et al.*, 2013; Niere *et al.*, 2012). However, ARH3 function always seems to be related to genome stability; for instance, in the cytosol, ARH3 cleaves the short PAR-chain released by the endo-glycosyl activity of PARG, preventing cell death by PARthanatos (Mashimo *et al.*, 2013). Recently, ARH3 mutation has been linked to pediatric neurodegeneration (Ghosh *et al.*, 2018). Moreover, ARH3 knockout MEF

(mouse embryonic fibroblast) cells are more sensitive to hydrogen peroxide treatment compared to WT cells and the presence of PARP1 inhibitor can rescue this phenotype, suggesting a direct link between ARH3 and PARP1 activity in DNA damage (Mashimo *et al.*, 2013). In this chapter, I highlighted the main biochemical activity of ARH3 and gave an overall characterisation of its activity. The identification of ARH3 activity advanced our understanding of the regulation of ADP-ribosylation in DNA damage repair. Finally, although ARH3 is interesting as a research topic and a potential therapeutic target in its own right, due to its specificity in the removal of Ser-ADP-ribosylation and lack of activity toward other ADP-ribosylation types, it can also be used as a simple tool for the study of Ser-ADP-ribosylation of histone proteins in DNA damage response. My *in vitro* findings were eventually confirmed by functional characterisation by a post-doc in our lab, Dr. Luca Palazzo. He used CRISPR/Cas9 technology to generate U2-O-S ARH3^{-/-} cells. After inducing DNA damage with H₂O₂ he confirmed, by PAR immunoblotting, that in the absence of ARH3, ADP-ribosylation could not be removed after 2 hours of induced DNA damage, unlike in the WT control cell line. Finally, using mass spectrometry, our collaborator the Matic group confirmed that Ser-ADP-ribosylation persists in ARH3^{-/-} cell line. Luca Palazzo, moreover, using U2-O-S ARH3^{-/-}, U2-O-S HPF1^{-/-} or U2-O-S PARP1^{-/-} showed that Ser-ADP-ribosylation is actually the most abundant ADP-ribosylation type in DNA damage repair and is dependent on HPF1/PARP1. The absence of ARH3 stabilises it while the absence of HPF1 results in an enrichment of ADP-ribosylation on glutamate/aspartate residues (Palazzo *et al.*, 2018). Months after, my *in vitro* findings were also confirmed by Abplanalp *et al.*, 2017, who performed similar *in vitro* assays and confirmed them using mass spectrometry. They also demonstrated a functional role for ARH3 in chromatin organisation. Despite these recent advances, a more detailed *functional* and *in vitro* characterisation is still lacking. In light of the

involvement of ARH3 in neurodegeneration, detailed understanding of its mechanism holds a promise of allowing novel therapies.

To sum up, I show that Ser-ADP-ribosylation is catalysed by the PARP1/HPF1 complex and can be efficiently reversed by the, previously poorly characterised, ARH3 glycohydrolase. I thus elucidated a new aspect of the ADP-ribosylation cycle and finally propose ARH3 as an easy and useful tool for further investigation of Ser-ADP-ribosylation in DNA repair and other pathways.

6.4 Ser-ADP-ribosylation in *Drosophila melanogaster*

Finally, over the last year of my DPhil, I aimed to investigate Ser-ADP-ribosylation in the model organism *Drosophila melanogaster*. Although *Drosophila* has been a very useful and easy model for the study of ADP-ribosylation compared to mice or other animal models, ADP-ribosylation in *Drosophila* has remained poorly characterised with just a few papers on this topic. PARP1, PARG as well as HPF1 are conserved in *Drosophila*, suggesting the presence of Ser-ADP-ribosylation. However, *Drosophila* does not possess any genes encoding an ARH3 homolog or any other protein with DraG-fold, raising the obvious question whether Ser-ADP-ribosylation can be reversed in this species and, if yes, which enzymes are responsible for this activity. Here I first showed, using a set of biochemical assays, that the *Drosophila* complex DmParp1/DmHpf1 can indeed modify histone proteins on serine residues. Then I went on to successfully show that this modification is indeed reversible and identified *Drosophila* Parg as the only *Drosophila* enzyme able to remove Ser-ADP-ribosylation, and finally performed a structural and biochemical characterisation of its activity. Preliminary analysis of DmParg structure (data not shown) does not suggest

any major differences in the secondary structure compared with human PARG. However, analysis of the catalytic site in human PARG showed that a serine in position 749 (adjacent to the catalytic E755 and E756) leads to the interaction with two lysines from an adjacent α -helix and with the distal ADPr in the catalytic site. Such interactions suggest a high rigidity of the catalytic loop. In contrast, DmParg possesses a correspondent glycine which does not show any interactions suggesting higher flexibility of the catalytic loop. In accordance with the model, TtParg possesses a glycine in the corresponding position. To confirm the model of flexibility of the catalytic loop, mutagenesis of human PARG S749 in glycine and DmParg G334 in serine will be performed.

I suggested *Drosophila* Parg as a useful tool for further investigation of Ser-ADP-ribosylation as it possesses the high efficiency of human PARG in removing PAR-chains as well as the ARH3's ability to remove Ser-ADP-ribosylation. This finding helped to better understand the discrepancy between human PARG and ARH3 activities. Human PARG and ARH3 can both remove PAR-chains while ARH3 alone can remove Ser-ADP-ribosylation. Despite the different structural organisation of these two human enzymes – PARG with a macrodomain and ARH3 with a DraG-fold domain – It was not completely clear why they have different activities as the ADP-ribose/ADP-ribose link is chemically very similar to the ADP-ribose/serine link. Here I also showed that another Parg homolog retained the ability to remove Ser-ADP-ribosylation, suggesting a conserved mechanism that is not unique to *Drosophila* species. I hypothesised that human PARG has lost its ability to remove Ser-ADP-ribosylation after the acquisition of ARH3, which shows a much higher efficiency in removing Ser-ADP-ribosylation compared to *Drosophila* PARG. Moreover, ARH3 activity can be also considered similar to that of other glycohydrolases which remove terminal ADPr, such as TARG1. Indeed, mutations of TARG1 or ARH3 in human are both

associated with severe neurodegeneration. Despite this similarity, it is not clear if the terminal ADPr linked either to glutamate/aspartate or serine is the direct triggering factor of neurodegeneration, for instance, persistent mono-ADP-ribosylation could cause instability on histone preventing other PTM, or if the persistence of mono-ADP-ribosylation determines constant interaction with other factor and genome instability.

In addition, Hanai *et al.*, 2014 reported that knock-out down of DmParg in *Drosophila* leads to neurodegeneration. This neurodegeneration was of course linked by the authors to the inability of DmParg to remove poly-ADP-ribosylation as this was its only known function. Surprisingly, this phenotype appears unique to *Drosophila* as mutations in Parg homologs in other organisms have never been reported to lead to neurodegeneration. However, in human neurodegeneration has been linked to mutations in TARG1, a hydrolase that removes mono-Glu-ADP-ribosylation (Sharifi *et al.* 2013) and just very recently in ARH3 mutation (Ghosh *et al.*, 2018). In light of the role of terminal-ADP-ribose in determining neurodegeneration, and my finding that DmParg can hydrolyse mono-Ser-ADP-ribosylation in addition to poly-ADP-ribosylation, I postulated that the neurodegeneration in *Drosophila* caused by Parg knock-out might be related to the loss of mono-Ser-ADP-ribosylation removal rather than poly-ADP-ribosylation hydrolysis, linking for the first time mono-Ser-ADP-ribosylation with a *Drosophila* disease.

The discovery of DmParg activity has many implications in ADP-ribosylation besides Ser-ADP-ribosylation. For instance, human ARH3 is a glycohydrolase present in different cellular compartments, nucleus cytosol and mitochondria, and is able to cleave short PAR-chains and Ser-ADP-ribosylation. Likely the different activities of ARH3 are dependent on the specific cellular localisation, mainly to cleave Ser-ADP-

ribosylation in the nucleus after DNA damage and to cleave PAR-chains in the cytosol preventing PARthanatos. However, there is not strong evidence of the role of ARH3 in the mitochondria despite being more abundant than in the nucleus. The multiple localisations of ARH3, related to its different functions, makes it difficult to understand its role. Moreover, the presence of PARG with its partial overlapping function with ARH3 makes the picture more complex. In this context, the study of DmParg, as the only identified glycohydrolase in *Drosophila*, would make the study of ADP-ribosylation removal quite simpler. However, so far there are no data showing the cellular localisation of DmParg. Performing some functional experiments to assess DmParg cellular localisation could allow us to understand whether DmParg localise in the mitochondria. Eventually, preventing its internalisation in the mitochondria, which has been challenging for ARH3, and trying to recover the observed phenotype will allow us to understand its function in this particular cellular compartment and eventually ARH3's mitochondrial function.

Finally, the findings highlighted in this thesis with all their possible implications and newly opened questions should suggest how far we are from completely understand the real complexity and regulation of ADP-ribosylation. I foresee outstanding outcome from the study of ADP-ribosylation in the next several years and hopefully, we will better understand how to apply our understanding of ADP-ribosylation for treatment of diseases such as cancer and neurodegeneration.

References

- Abplanalp J, Leutert M, Frugier E, Nowak K, Feurer R, Kato J, Kistemaker HVA, Filippov DV, Moss J, Caffisch A, Hottiger MO. Proteomic analyses identify ARH3 as a serine mono-ADP-ribosylhydrolase. *Nat Commun.* 8(1):2055 (2017).
- Adamietz P. Poly(ADP-ribose) synthase is the major endogenous nonhistone acceptor for poly(ADP-ribose) in alkylated rat hepatoma cells. *Eur J Biochem.* 169(2):365-72 (1987).
- Adams CJ, Pike AC, Maniam S, Sharpe TD, Coutts AS, Knapp S, La Thangue NB, Bullock AN. The p53 cofactor Strap exhibits an unexpected TPR motif and oligonucleotide-binding (OB)-fold structure. *Proc Natl Acad Sci U S A.* 109(10):3778-83 (2010).
- Adriouch S, Bannas P, Schwarz N, Fliegert R, Guse AH, Seman M, Haag F, Koch-Nolte F. ADP-ribosylation at R125 gates the P2X7 ion channel by presenting a covalent ligand to its nucleotide binding site. *FASEB J.* 22(3):861-9 (2008).
- Ahel D, Horejsí Z, Wiechens N, Polo SE, Garcia-Wilson E, Ahel I, Flynn H, Skehel M, West SC, Jackson SP, Owen-Hughes T, Boulton SJ. Poly(ADP-ribose)-dependent regulation of DNA repair by the chromatin remodeling enzyme ALC1. *Science.* 325,1240-3 (2009).
- Ahel I, Ahel D, Matsusaka T, Clark AJ, Pines J, Boulton SJ, West SC. Poly(ADP-ribose)-binding zinc finger motifs in DNA repair/checkpoint proteins. *Nature.* 451(7174):81-5 (2008).
- Ahnesorg P, Smith P, Jackson SP. XLF interacts with the XRCC4-DNA ligase IV complex to promote DNA nonhomologous end-joining. *Cell.* 124(2):301-13 (2006).
- Allen MD, Buckle AM, Cordell SC, Löwe J, Bycroft M. The crystal structure of AF1521 a protein from *Archaeoglobus fulgidus* with homology to the non-histone domain of macroH2A. *J Mol Biol.* 330(3):503-11 (2003).
- Amé JC, Rolli V, Schreiber V, Niedergang C, Apiou F, Decker P, Muller S, Höger T, Ménissier-de Murcia J, de Murcia G. PARP-2, A novel mammalian DNA damage-dependent poly(ADPribose) polymerase. *J Biol Chem.* 274(25):17860-8 (1999).
- Amé JC, Spenlehauer C, de Murcia G. The PARP superfamily. *Bioessays.* 26(8):882-93 (2004).
- Andrabi SA, Umanah GK, Chang C, Stevens DA, Karuppagounder SS, Gagné JP, Poirier GG, Dawson VL, Dawson TM. Poly(ADP-ribose) polymerase-dependent energy depletion occurs through inhibition of glycolysis. *Proc Natl Acad Sci U S A.* 111(28):10209-14 (2014).
- Aravind L. The WWE domain: a common interaction module in protein ubiquitination and ADP ribosylation. *Trends Biochem Sci.* 26,273-5 (2001).
- Barkauskaite E, Brassington A, Tan ES, Warwicker J, Dunstan MS, Banos B, Lafite P, Ahel M, Mitchison TJ, Ahel I, Leys D. Visualization of poly(ADP-ribose) bound to PARG reveals inherent balance between exo- and endo-glycohydrolase activities. *Nat Commun.* 4,2164 (2013).
- Barkauskaite E, Jankevicius G, Ahel I. Structures and Mechanisms of Enzymes Employed in the Synthesis and Degradation of PARP-Dependent Protein ADP-Ribosylation. *Mol Cell.* 58(6):935-46 (2015).
- Bartlett E, Bonfiglio JJ, Prokhorova E, Colby T, Zobel F, Ahel I, Matic I. Interplay of histone marks with serine ADP-ribosylation. *Cell reports.* 24(13):3488-3502 (2018).
- Barty A, Küpper J, Chapman HN. Molecular imaging using X-ray free-electron lasers. *Annu Rev Phys Chem.* 64:415-35 (2013).
- Beard WA, Wilson SH. Structural design of a eukaryotic DNA repair polymerase: DNA polymerase beta. *Mutat Res.* 460(3-4):231-44 (2000).

- Becherel OJ, Jakob B, Cherry AL, Gueven N, Fusser M, Kijas AW, Peng C, Katyal S, McKinnon PJ, Chen J, Epe B, Smerdon SJ, Taucher-Scholz G, Lavin MF. CK2 phosphorylation-dependent interaction between aprataxin and MDC1 in the DNA damage response. *Nucleic Acids Res.* 38(5):1489-503 (2010).
- Ben-Hur E, Chen CC, Elkind MM. Inhibitors of poly(adenosine diphosphoribose) synthetase, examination of metabolic perturbations, and enhancement of radiation response in Chinese hamster cells. *Cancer Res.* 45(5):2123-7 (1985).
- Bilan V, Leutert M, Nanni P, Panse C, Hottiger MO. Combining Higher-Energy Collision Dissociation and Electron-Transfer/Higher-Energy Collision Dissociation Fragmentation in a Product-Dependent Manner Confidently Assigns Proteomewide ADP-Ribose Acceptor Sites. *Anal Chem.* 89(3):1523-1530 (2017).
- Bonfiglio JJ, Fontana P, Zhang Q, Colby T, Gibbs-Seymour I, Atanasov I, Bartlett E, Zaja R, Ahel I, Matic I. Serine ADP-Ribosylation Depends on HPF1. *Mol Cell.* 65(5):932-940 (2017).
- Braun SA, Panzeter PL, Collinge MA, Althaus FR. Endoglycosidic cleavage of branched polymers by poly(ADP-ribose) glycohydrolase. *Eur J Biochem.* 220(2):369-75 (1994).
- Brochu G, Duchaine C, Thibeault L, Lagueux J, Shah GM, Poirier GG. Mode of action of poly(ADP-ribose) glycohydrolase. *Biochim Biophys Acta.* 1219(2):342-50 (1994).
- Bütepage M, Eckerl L, Verheugd P, Lüscher B. Intracellular Mono-ADP-Ribosylation in signaling and disease. *Cells.* 4:569–595 (2015).
- Caldecott KW. XRCC1 and DNA strand break repair. *DNA Repair (Amst).* 2,955-69 (2003).
- Callow MG, Tran H, Phu L, Lau T, Lee J, Sandoval WN, Liu PS, Bheddah S, Tao J, Lill JR, Hongo JA, Davis D, Kirkpatrick DS, Polakis P, Costa M. Ubiquitin ligase RNF146 regulates tankyrase and Axin to promote Wnt signaling. *PLoS One.* 6(7):e22595 (2011).
- Chambon P, Weill JD, Mandel P. Nicotinamide mononucleotide activation of new DNA-dependent polyadenylic acid synthesizing nuclear enzyme. *Biochem Biophys Res Commun.* 11,39-43 (1963).
- Chen D, Vollmar M, Rossi MN, Phillips C, Kraehenbuehl R, Slade D, Mehrotra PV, von Delft F, Crosthwaite SK, Gileadi O, Denu JM, Ahel I. Identification of macrodomain proteins as novel O-acetyl-ADP-ribose deacetylases. *J Biol Chem.* 286(15):13261-71 (2011).
- Ciccia A, Elledge SJ. The DNA damage response: making it safe to play with knives. *Mol Cell.* 40,179-204 (2010).
- Clark NJ, Kramer M, Muthurajan UM, Luger K. Alternative modes of binding of poly(ADP-ribose) polymerase 1 to free DNA and nucleosomes. *J Biol Chem.* 287,32430-9 (2012).
- Cowtan K. Recent developments in classical density modification. *Acta Crystallogr D Biol Crystallogr.* 66(Pt 4):470-8 (2010).
- Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr D Biol Crystallogr.* 62(Pt 9):1002-11 (2006).
- D'Amours D, Desnoyers S, D'Silva I, Poirier GG. Poly(ADP-ribosyl)ation reactions in the regulation of nuclear functions. *Biochem J.* 342(Pt 2),249-68 (1999).
- David KK, Andrabi SA, Dawson TM, Dawson VL. Parthanatos, a messenger of death. *Front Biosci (Landmark Ed).* 14:1116-28 (2009).
- Dawicki-McKenna JM, Langelier MF, DeNizio JE, Riccio AA, Cao CD, Karch KR, McCauley M, Steffen JD, Black BE, Pascal JM. PARP-1 Activation Requires Local Unfolding of an Autoinhibitory Domain. *Mol Cell.* 60(5):755-768 (2015).
- de Murcia G, Jongstra-Bilen J, Ittel ME, Mandel P, Delain E. Poly(ADP-ribose) polymerase auto-modification and interaction with DNA: electron microscopic visualization. *EMBO J.* 2(4):543-8 (1983).

- De Vos M, Schreiber V, Dantzer F. The diverse roles and clinical relevance of PARPs in DNA damage repair: current state of the art. *Biochem Pharmacol.* 84,137- 46 (2012).
- Dunstan MS, Barkauskaite E, Lafite P, Knezevic CE, Brassington A, Ahel M, Hergenrother PJ, Leys D, Ahel I. Structure and mechanism of a canonical poly(ADP-ribose) glycohydrolase. *Nat Commun.* 3,878 (2012).
- Egloff MP, Malet H, Putics A, Heinonen M, Dutartre H, Frangeul A, Gruez A, Campanacci V, Cambillau C, Ziebuhr J, Ahola T, Canard B. Structural and functional basis for ADP-ribose and poly(ADP-ribose) binding by viral macro domains. *J Virol.* 80(17):8493-502 (2006).
- Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr.* 60(Pt 12 Pt 1):2126-32 (2004).
- Evans PR. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* 62, 72–82 (2006).
- Evans PR, McCoy A. An introduction to molecular replacement. *Acta Crystallogr D Biol Crystallogr.* 64(Pt 1): 1–10 (2008).
- Evans PR, Murshudov GN. How good are my data and what is the resolution. *Acta Crystallogr. D Biol. Crystallogr.* 69, 1204–1214 (2013).
- Eustermann S, Brockmann C, Mehrotra PV, Yang JC, Loakes D, West SC, Ahel I, Neuhaus D. Solution structures of the two PBZ domains from human APLF and their interaction with poly(ADP-ribose). *Nat Struct Mol Biol.* 17,241-3 (2010).
- Eustermann S, Wu WF, Langelier MF, Yang JC, Easton LE, Riccio AA, Pascal JM, Neuhaus D. Structural Basis of Detection and Signaling of DNA Single-Strand Breaks by Human PARP-1. *Mol Cell.* 60(5):742-754 (2015).
- Feijs KL, Verheugd P, Luscher B. Expanding functions of intracellular resident mono-ADP-ribosylation in cell physiology. *FEBS J.* 280,3519-29 (2013).
- Fontana P, Bonfiglio JJ, Palazzo L, Bartlett E, Matic I, Ahel I. Serine ADP-ribosylation reversal by the hydrolase ARH3. *Elife.* 6. pii: e28533 (2016).
- Frosina G, Fortini P, Rossi O, Carrozzino F, Raspaglio G, Cox LS, Lane DP, Abbondandolo A, Dogliotti E. Two pathways for base excision repair in mammalian cells. *J Biol Chem.* 271(16):9573-8 (1996).
- Gagné JP, Ethier C, Defoy D, Bourassa S, Langelier MF, Riccio AA, Pascal JM, Moon KM, Foster LJ, Ning Z, Figeys D, Droit A, Poirier GG. Quantitative site-specific ADP-ribosylation profiling of DNA-dependent PARPs. *DNA Repair (Amst).* 30:68-79 (2015).
- Gagné JP, Isabelle M, Lo KS, Bourassa S, Hendzel MJ, Dawson VL, Dawson TM, Poirier GG. Proteome-wide identification of poly(ADP-ribose) binding proteins and poly(ADP-ribose)-associated protein complexes. *Nucleic Acids Res.* 36(22):6959-76 (2008).
- Genschel J, Bazemore LR, Modrich P. Human exonuclease I is required for 5' and 3' mismatch repair. *J Biol Chem.* 277(15):13302-11 (2002).
- Ghosh SG, Becker K, Huang H, Salazar TD, Chai G, Salpietro V, Al-Gazali L, Waisfisz Q, Wang H, Vaux KK, Stanley V, Manole A, Akpulat U, Weiss MM, Efthymiou S, Hanna MG, Minetti C, Striano P, Pisciotta L, De Grandis E, Altmüller J, Nürnberg P, Thiele H, Yis U, Okur TD, Polat AI, Amiri N, Doosti M, Karimani EG, Toosi MB, Haddad G, Karakaya M, Wirth B, van Hagen JM, Wolf NI, Maroofian R, Houlden H, Cirak S, Gleeson JG. Biallelic Mutations in ADPRHL2, Encoding ADP-Ribosylhydrolase 3, Lead to a Degenerative Pediatric Stress-Induced Epileptic Ataxia Syndrome. *Am J Hum Genet.* 103(3):431-439 (2018).
- Gibbs-Seymour I, Fontana P, Rack JGM, Ahel I. HPF1/C4orf27 Is a PARP-1-Interacting Protein that Regulates PARP-1 ADP-Ribosylation Activity. *Mol Cell.* 62(3):432-442 (2016).

- Gibson BA, Kraus WL. New insights into the molecular and cellular functions of poly(ADP-ribose) and PARPs. *Nat Rev Mol Cell Biol.* 13,411-24 (2012).
- Gildea RJ, Waterman DG, Parkhurst JM, Axford D, Sutton G, Stuart DI, Sauter NK, Evans G, Winter G. New methods for indexing multi-lattice diffraction data. *Acta Crystallogr D Biol Crystallogr.* 70(Pt 10):2652-66 (2014).
- Glowacki G, Braren R, Firner K, Nissen M, Kühl M, Reche P, Bazan F, Cetkovic-Cvrlje M, Leiter E, Haag F, Koch-Nolte F. The family of toxin-related ecto-ADP-ribosyltransferases in humans and the mouse. *Protein Sci.* 11(7):1657-70 (2002).
- Groth, P. *Chemische Kristallographie.* Leipzig: Engelmann. Vol. 1, pp. 176–181 (1908).
- Ha HC, Snyder SH. Poly(ADP-ribose) polymerase is a mediator of necrotic cell death by ATP depletion. *Proc Natl Acad Sci U S A.* 96,13978-82 (1999).
- Haince JF, McDonald D, Rodrigue A, Déry U, Masson JY, Hendzel MJ, Poirier GG. PARP1-dependent kinetics of recruitment of MRE11 and NBS1 proteins to multiple DNA damage sites. *J Biol Chem.* 283, 1197-208 (2008).
- Hakmé A, Wong HK, Dantzer F, Schreiber V. The expanding field of poly(ADP-ribosyl)ation reactions. 'Protein Modifications: Beyond the Usual Suspects' Review Series. *EMBO Rep.* 9(11):1094-100 (2008).
- Hassa PO. The molecular "Jekyll and Hyde" duality of PARP1 in cell death and cell survival. *Front Biosci (Landmark Ed).* 14:72-111 (2009).
- Hassa PO, Hottiger MO. The diverse biological roles of mammalian PARPS, a small but powerful family of poly-ADP-ribose polymerases. *Front Biosci.* 13:3046-82 (2008).
- Hatakeyama K, Nemoto Y, Ueda K, Hayaishi O. Purification and characterization of poly(ADP-ribose) glycohydrolase. Different modes of action on large and small poly(ADP-ribose). *J Biol Chem.* 261(32):14902-11 (1986).
- Hein MY, Hubner NC, Poser I, Cox J, Nagaraj N, Toyoda Y, Gak IA, Weisswange I, Mansfeld J, Buchholz F, Hyman AA, Mann M. A human interactome in three quantitative dimensions organized by stoichiometries and abundances. *Cell.* 163(3):712-23 (2015).
- Helleday T. The underlying mechanism for the PARP and BRCA synthetic lethality: clearing up the misunderstandings. *Mol Oncol.* 5(4):387-93 (2011).
- Hombauer H, Campbell CS, Smith CE, Desai A, Kolodner RD. Visualization of eukaryotic DNA mismatch repair reveals distinct recognition and repair intermediates. *Cell.* 147(5):1040-53 (2011).
- Huber R. Die automatisierte Faltmolekülmethode. *Acta Cryst.* 19, 353-356 (1965).
- Ikejima M, Gill DM. Poly(ADP-ribose) degradation by glycohydrolase starts with an endonucleolytic incision. *J Biol Chem.* 15;263(23):11037-40 (1988).
- Isabelle M, Moreel X, Gagné JP, Rouleau M, Ethier C, Gagné P, Hendzel MJ, Poirier GG. Investigation of PARP-1, PARP-2, and PARG interactomes by affinity-purification mass spectrometry. *Proteome Sci.* 8:22 (2010).
- Ishak L, Moretton A, Garreau-Balandier I, Lefebvre M, Alziari S, Lachaume P, Morel F, Farge G, Vernet P, Dubessay P. DNA maintenance following bleomycin-induced strand breaks does not require poly(ADP-ribosyl)ation activation in *Drosophila* S2 cells. *DNA Repair (Amst).* 48:8-16 (2016).
- Isogai S, Kanno S, Ariyoshi M, Tochio H, Ito Y, Yasui A, Shirakawa M. Solution structure of a zinc-finger domain that binds to poly-ADP-ribose. *Genes Cells.* 15(2):101-10 (2010).
- Jankevicius G, Hassler M, Golia B, Rybin V, Zacharias M, Timinszky G, Ladurner AG. A family of macrodomain proteins reverses cellular mono-ADPriboseylation. *Nat Struct Mol Biol.* 20,508-14 (2013).

- Jeggo P, O'Neill P. The Greek Goddess, Artemis, reveals the secrets of her cleavage. *DNA Repair (Amst)*. 1(9):771-7 (2002).
- Ji Y, Thomas C, Tulin N, Lodhi N, Boamah E, Kolenko V, Tulin AV. Charon Mediates Immune Deficiency-Driven PARP-1-Dependent Immune Responses in *Drosophila*. *J Immunol*. 197(6):2382-9 (2016).
- Ji Y, Tulin AV. Poly(ADP-ribosyl)ation of heterogeneous nuclear ribonucleoproteins modulates splicing. *Nucleic Acids Res*. 37(11):3501-13 (2009).
- Jones TA, Zou JY, Cowan SW, Kjeldgaard M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst. A* 47, 110-119 (1991).
- Juarez-Salinas H, Levi V, Jacobson EL, Jacobson MK. Poly(ADP-ribose) has a branched structure in vivo. *J Biol Chem*. 257,607-9 (1982).
- Kabsch W. XDS. *Acta Crystallogr. D Biol. Crystallogr*. 66, 125–132 (2010).
- Kabsch W. Integration, scaling, space-group assignment and post refinement. *Acta Crystallogr. D Biol. Crystallogr*. 66, 133–144 (2010).
- Kadyrov FA, Dzantiev L, Constantin N, Modrich P. Endonucleolytic function of MutL α in human mismatch repair. *Cell*. 126(2):297-308 (2006).
- Kahl S, Nissen M, Girisch R, Duffy T, Leiter EH, Haag F, Koch-Nolte F. Metalloprotease-mediated shedding of enzymatically active mouse ecto-ADP-ribosyltransferase ART2.2 upon T cell activation. *J Immunol*. 165(8):4463-9 (2000).
- Kang HC, Lee YI, Shin JH, Andrabi SA, Chi Z, Gagné JP, Lee Y, Ko HS, Lee BD, Poirier GG, Dawson VL, Dawson TM. Iduna is a poly(ADP-ribose) (PAR)-dependent E3 ubiquitin ligase that regulates DNA damage. *Proc Natl Acad Sci U S A*. 108,14103-8 (2011).
- Karras GI, Kustatscher G, Buhecha HR, Allen MD, Pugieux C, Sait F, Bycroft M, Ladurner AG. The macro domain is an ADP-ribose binding module. *EMBO J*. 24,1911-20 (2005).
- Kass EM, Jasin M. Collaboration and competition between DNA double-strand break repair pathways. *FEBS Lett*. 584,3703-8 (2010).
- Kaufmann SH, Desnoyers S, Ottaviano Y, Davidson NE, Poirier GG. Specific proteolytic cleavage of poly(ADP-ribose) polymerase: an early marker of chemotherapy-induced apoptosis. *Cancer Res*. 53(17):3976-85 (1993).
- Kawamura T, Hanai S, Yokota T, Hayashi T, Poltronieri P, Miwa M, Uchida K. An alternative form of poly(ADP-ribose) polymerase in *Drosophila melanogaster* and its ectopic expression in rat-1 cells. *Biochem Biophys Res Commun*. 251(1):35-40 (1998).
- Kedar PS, Stefanick DF, Horton JK, Wilson SH. Increased PARP-1 association with DNA in alkylation damaged, PARP-inhibited mouse fibroblasts. *Mol Cancer Res*. 10(3):360-8 (2012).
- Kendrew JC, Bodo G, Dintzis HM, Parrish RG, Wyckoff H, Phillips DC. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature*. 181(4610):662-6 (1958).
- Kernstock S, Koch-Nolte F, Mueller-Dieckmann J, Weiss MS, Mueller-Dieckmann C. Cloning, expression, purification, crystallization and preliminary X-ray diffraction analysis of human ARH3, the first eukaryotic protein-ADP-ribosylhydrolase. *Acta Crystallogr Sect F Struct Biol Cryst Commun*. 62(Pt 3):224-7 (2006).
- Kim K, Pedersen LC, Kirby TW, DeRose EF, London RE. Characterization of the APLF FHA-XRCC1 phosphopeptide interaction and its structural and functional implications. *Nucleic Acids Res*. 45(21):12374-12387 (2017).

- Kleine H, Poreba E, Lesniewicz K, Hassa PO, Hottiger MO, Litchfield DW, Shilton BH, Lüscher B. Substrate-assisted catalysis by PARP10 limits its activity to mono-ADP-ribosylation. *Molecular Cell*. 32:57–69 (2008).
- Kleywegt GJ, Jones TA. Template Convolution to Enhance or Detect Structural Features in Macromolecular Electron-Density Maps. *Acta Cryst. D* 53, 179-185 (1997).
- Klungland A, Lindahl T. Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1). *EMBO J*. 16(11):3341-8 (1997).
- Kraus WL. Transcriptional control by PARP-1: chromatin modulation, enhancer-binding, coregulation, and insulation. *Curr Opin Cell Biol*. 20(3):294-302 (2008).
- Krejci L, Altmannova V, Spirek M, Zhao X. Homologous recombination and its regulation. *Nucleic Acids Res*. 40,5795-818 (2012).
- Ku WC, Chiu SK, Chen YJ, Huang HH, Wu WG, Chen YJ. Complementary quantitative proteomics reveals that transcription factor AP-4 mediates E-box-dependent complex formation for transcriptional repression of HDM2. *Mol Cell Proteomics*. (9):2034-50 (2009).
- Lambrech MJ, Brichacek M, Barkauskaite E, Ariza A, Ahel I, Hergenrother PJ. Synthesis of dimeric ADP-ribose and its structure with human poly(ADP-ribose) glycohydrolase. *J Am Chem Soc*. 137(10):3558-64 (2015).
- Langelier MF, Planck JL, Roy S, Pascal JM. Structural basis for DNA damage dependent poly(ADP-ribosyl)ation by human PARP-1. *Science*. 336,728-32 (2012).
- Larsen SC, Hendriks IA, Lyon D, Jensen LJ, Nielsen ML. Systems-wide Analysis of Serine ADP-Ribosylation Reveals Widespread Occurrence and Site-Specific Overlap with Phosphorylation. *Cell Rep*. 24(9):2493-2505 (2018).
- Ledermann J, Harter P, Gourley C, Friedlander M, Vergote I, Rustin G, Scott C, Meier W, Shapira-Frommer R, Safra T, Matei D, Macpherson E, Watkins C, Carmichael J, Matulonis U. Olaparib maintenance therapy in platinum-sensitive relapsed ovarian cancer. *N Engl J Med*. 366(15):1382-92 (2012).
- Leidecker O, Bonfiglio JJ, Colby T, Zhang Q, Atanassov I, Zaja R, Palazzo L, Stockum A, Ahel I, Matic I. Serine is a new target residue for endogenous ADP-ribosylation on histones. *Nat Chem Biol*. (12):998-1000 (2016).
- Leslie AGW, Powell HR. Processing diffraction data with mosflm. *Evolving Methods for Macromolecular Crystallography*. Vol. 245, pp. 41–51 (2007).
- Leslie PDM, Leutert M, Bilan V, Nowak K, Gunasekera K, Ferrari E, Imhof R, Malmström L, Hottiger MO. Comprehensive ADP-ribosylome analysis identifies tyrosine as an ADP-ribose acceptor site. *EMBO Rep*. 19(8).pii:e45310 (2018).
- Li GY, McCulloch RD, Fenton AL, Cheung M, Meng L, Ikura M, Koch CA. Structure and identification of ADP-ribose recognition motifs of APLF and role in the DNA damage response. *Proc Natl Acad Sci U S A*. 107(20):9129-34 (2010).
- Li J, Williams BL, Haire LF, Goldberg M, Wilker E, Durocher D, Yaffe MB, Jackson SP, Smerdon SJ. Structural and functional versatility of the FHA domain in DNA-damage signaling by the tumor suppressor kinase Chk2. *Mol Cell*. 9(5):1045-54 (2002). Li M, Lu LY, Yang CY, Wang S, Yu X. The FHA and BRCT domains recognize ADP-ribosylation during DNA damage response. *Genes Dev*. 27(16):1752–1768 (2013).
- Li N, Chen J. ADP-ribosylation: activation, recognition, and removal. *Mol Cells*. 37(1):9-16 (2014).
- Lieber MR, Lu H, Gu J, Schwarz K. Flexibility in the order of action and in the enzymology of the nuclease, polymerases, and ligase of vertebrate non-homologous DNA end joining: relevance to cancer, aging, and the immune system. *Cell Res*. 18(1):125-33 (2008).

- Lin W, Amé JC, Aboul-Ela N, Jacobson EL, Jacobson MK. Isolation and characterization of the cDNA encoding bovine poly(ADP-ribose) glycohydrolase. *J Biol Chem*. 272(18):11895-901 (1997).
- Lindahl T. An N-glycosidase from *Escherichia coli* that releases free uracil from DNA containing deaminated cytosine residues. *Proc Natl Acad Sci U S A*. 71(9):3649-53 (1974).
- Liszczyk G, Diehl KL, Dann GP, Muir TW. Acetylation blocks DNA damage-induced chromatin ADP-ribosylation. *Nat Chem Biol*. 14(9):837-840 (2018).
- Longley MJ, Pierce AJ, Modrich P. DNA polymerase delta is required for human mismatch repair in vitro. *J Biol Chem*. 272(16):10917-21 (1997).
- Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. *Science*. 355(6330):1152-1158 (2017).
- Loseva O, Jemth AS, Bryant HE, Schüler H, Lehtiö L, Karlberg T, Helleday T. PARP-3 is a mono-ADP-ribosylase that activates PARP-1 in the absence of DNA. *J Biol Chem*. 285(11):8054-60 (2010).
- Malanga M, Pleschke JM, Kleczkowska HE, Althaus FR. Poly(ADP-ribose) binds to specific domains of p53 and alters its DNA binding functions. *J Biol Chem*. 273,11839-43 (1998).
- Malet H, Coutard B, Jamal S, Dutartre H, Papageorgiou N, Neuvonen M, Ahola T, Forrester N, Gould EA, Lafitte D, Ferron F, Lescar J, Gorbalenya AE, de Lamballerie X, Canard B. The crystal structures of Chikungunya and Venezuelan equine encephalitis virus nsP3 macro domains define a conserved adenosine binding pocket. *Virology*. 83(13):6534-45 (2009).
- Mansour WY, Rhein T, Dahm-Daphi, J. The alternative end-joining pathway for repair of DNA double-strand breaks requires PARP1 but is not dependent upon microhomologies. *Nucleic Acids Res*. 38,6065-77 (2010).
- Marteijn JA, Lans H, Vermeulen W, Hoeijmakers JH. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol*. 15,465-81 (2014).
- Martello R, Leutert M, Jungmichel S, Bilan V, Larsen SC, Young C, Hottiger MO, Nielsen ML. Proteome-wide identification of the endogenous ADP-ribosylome of mammalian cells and tissue. *Nat Commun*. 7:12917 (2016).
- Mashimo M, Moss J. Functional Role of ADP-Ribosyl-Acceptor Hydrolase 3 in poly(ADP-Ribose) Polymerase-1 Response to Oxidative Stress. *Curr Protein Pept Sci*. 17(7):633-640 (2016).
- Mashimo M, Kato J, Moss J. ADP-ribosyl-acceptor hydrolase 3 regulates poly (ADP-ribose) degradation and cell death during oxidative stress. *Proc Natl Acad Sci U S A*. 110 (47)18964-18969 (2013).
- Mashimo M, Kato J, Moss J. Structure and function of the ARH family of ADP-ribosyl-acceptor hydrolases. *DNA Repair (Amst)*. 23:88-94 (2014).
- Matsumoto Y, Kim K. Excision of deoxyribose phosphate residues by DNA polymerase beta during DNA repair. *Science*. 269(5224):699-702 (1995).
- Mehrotra PV, Ahel D, Ryan DP, Weston R, Wiechens N, Kraehenbuehl R, Owen-Hughes T, Ahel I. DNA repair factor APLF is a histone chaperone. *Mol Cell*. 41(1):46-55 (2011).
- McIlwraith MJ, Vaisman A, Liu Y, Fanning E, Woodgate R, West SC. Human DNA polymerase eta promotes DNA synthesis from strand invasion intermediates of homologous recombination. *Mol Cell*. 20(5):783-92 (2005).
- Menzel S, Rissiek B, Bannas P, Jakoby T, Miksiewicz M, Schwarz N, Nissen M, Haag F, Tholey A, Koch-Nolte F. Nucleotide-Induced Membrane-Proximal Proteolysis Controls the Substrate Specificity of T Cell Ecto-ADP-Ribosyltransferase ARTC2.2. *J Immunol*. 195(5):2057-66 (2015).

- Messner S, Altmeyer M, Zhao H, Pozivil A, Roschitzki B, Gehrig P, Rutishauser D, Huang D, Caflisch A, Hottiger MO. PARP1 ADP-ribosylates lysine residues of the core histone tails. *Nucleic Acids Res.* 38,6350-62 (2010).
- Miwa M, Saikawa N, Yamaizumi Z, Nishimura S, Sugimura T. Structure of poly(adenosine diphosphate ribose): identification of 2'-[1''-ribosyl-2''-(or 3''-)(1'''-ribosyl)]adenosine-5',5'',5'''-tris(phosphate) as a branch linkage. *Proc Natl Acad Sci U S A.* 76, 595-9 (1979).
- Miwa M, Sugimura T. Splitting of the ribose-ribose linkage of poly(adenosine diphosphate-ribose) by a calf thymus extract. *J Biol Chem.* 246,6362-4 (1971).
- Moss J, Stanley SJ, Nightingale MS, Murtagh JJ Jr, Monaco L, Mishima K, Chen HC, Williamson KC, Tsai SC. Molecular and immunological characterization of ADP-ribosylarginine hydrolases. *J Biol Chem.* 267(15):10481-8 (1992).
- Mu D, Wakasugi M, Hsu DS, Sancar A. Characterization of reaction intermediates of human excision repair nuclease. *J Biol Chem.* 272(46):28971-9 (1997).
- Mueller-Dieckmann C, Kernstock S, Lisurek M, von Kries JP, Haag F, Weiss MS, Koch-Nolte F. The structure of human ADP-ribosylhydrolase 3 (ARH3) provides insights into the reversibility of protein ADP-ribosylation. *Proc Natl Acad Sci U S A.* 103,15026-31 (2006).
- Munnur D, Ahel I. Reversible mono-ADP-ribosylation of DNA breaks. *FEBS J.* 284(23):4002-4016 (2017).
- Murai J, Huang SY, Das BB, Renaud A, Zhang Y, Doroshow JH, Ji J, Takeda S, Pommier Y. Trapping of PARP1 and PARP2 by Clinical PARP Inhibitors. *Cancer Res.* 72(21):5588-99 (2012).
- Murai J, Huang SY, Renaud A, Zhang Y, Ji J, Takeda S, Morris J, Teicher B, Doroshow JH, Pommier Y. Stereospecific PARP trapping by BMN 673 and comparison with olaparib and rucaparib. *Mol Cancer Ther.* 13(2):433-43 (2014).
- Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr D Biol Crystallogr.* 53(Pt 3):240-55 (1997).
- Murzin AG. OB(oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. *EMBO J.* 12(3):861-7 (1993).
- Ni Z, Xu C, Guo X, Hunter GO, Kuznetsova OV, Tempel W, Marcon E, Zhong G, Guo H, Kuo WW, Li J, Young P, Olsen JB, Wan C, Loppnau P, El Bakkouri M, Senisterra GA, He H, Huang H, Sidhu SS, Emili A, Murphy S, Mosley AL, Arrowsmith CH, Min J, Greenblatt JF. RPRD1A and RPRD1B are human RNA polymerase II C-terminal domain scaffolds for Ser5 dephosphorylation. *Nat Struct Mol Biol.* 21(8):686-695 (2014).
- Niere M, Mashimo M, Agledal L, Dölle C, Kasamatsu A, Kato J, Moss J, Ziegler M. ADP-ribosylhydrolase 3 (ARH3), not poly(ADP-ribose) glycohydrolase (PARG) isoforms, is responsible for degradation of mitochondrial matrix-associated poly(ADP-ribose). *J Biol Chem.* 287(20):16088-102 (2012).
- Ogata N, Ueda K, Hayaishi O. ADP-ribosylation of histone H2B. Identification of glutamic acid residue 2 as the modification site. *J Biol Chem.* 255(16):7610-5 (1980).
- Ogata N, Ueda K, Kawaichi M, Hayaishi O. Poly(ADP-ribose) synthetase, a main acceptor of poly(ADP-ribose) in isolated nuclei. *J Biol Chem.* 10;256(9):4135-7 (1981).
- Oka S, Kato J, Moss J. Identification and characterization of a mammalian 39-kDa poly(ADP-ribose) glycohydrolase. *J Biol Chem.* 281(2):705-13 (2006).
- Ono T, Kasamatsu A, Oka S, Moss J. The 39-kDa poly(ADP-ribose) glycohydrolase ARH3 hydrolyzes O-acetyl-ADP-ribose, a product of the Sir2 family of acetyl-histone deacetylases. *Proc Natl Acad Sci U S A.* 103(45):16687-91 (2006).

- Otto H, Reche PA, Bazan F, Dittmar K, Haag F, Koch-Nolte FL. In silico characterization of the family of PARP-like poly(ADPriboseyl)transferases (pARTs). *BMC Genomics*. 6,139 (2005).
- Otwinowski Z, Minor W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol*. 276, 307–326 (1997).
- Owen RL, Juanhuix J, Fuchs M. Current advances in synchrotron radiation instrumentation for MX experiments. *Arch Biochem Biophys*. 602:21-31 (2016).
- Palazzo L, Leidecker O, Prokhorova E, Dauben H, Matic I, Ahel I. Serine is the major residue for ADPriboseylation upon DNA damage. *Elife*. 7. pii: e34334 (2018).
- Palombo F, Gallinari P, Iaccarino I, Lettieri T, Hughes M, D'Arrigo A, Truong O, Hsuan JJ, Jiricny J. GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. *Science*. 268(5219):1912-4 (1995).
- Papeo G, Casale E, Montagnoli A, Cirila A. PARP inhibitors in cancer therapy: an update. *Expert Opin Ther Pat*. 23(4):503-14 (2013).
- Pascal JM, Ellenberger T. The rise and fall of poly(ADPribose): An enzymatic perspective. *DNA Repair (Amst)*. 32:10-6 (2015).
- Pehrson JR, Fried VA. MacroH2A, a core histone containing a large nonhistone region. *Science*. 257,1398-400 (1992).
- Perina D, Mikoč A, Ahel J, Četković H, Žaja R, Ahel I. Distribution of protein poly(ADPriboseylation) systems across all domains of life. *DNA Repair* 23:4–16 (2014).
- Petit C, Sancar A. Nucleotide excision repair: from E. coli to man. *Biochimie*. 81(1-2):15-25 (1999).
- Perutz MF. Isomorphous replacement and phase determination in non-centrosymmetric space groups. *Acta Cryst*. 9, 867-873 (1956).
- Phillips JC, Wlodawer A, Yevitz MM, Hodgson KO. Applications of synchrotron radiation to protein crystallography: preliminary results. *Proc Natl Acad Sci U S A*. 73(1):128-32 (1976). Pike AC, Garman EF, Krojer T, von Delft F, Carpenter EP. An overview of heavy-atom derivatization of protein crystals. *Acta Crystallogr D Struct Biol*. 72(Pt 3): 303–318 (2016).
- Pleschke JM, Kleczkowska HE, Strohm M, Althaus FR. Poly(ADPribose) binds to specific domains in DNA damage checkpoint proteins. *J Biol Chem*. 275(52):40974-80 (2000).
- Poirier GG, de Murcia G, Jongstra-Bilen J, Niedergang C, Mandel P. Poly(ADPriboseylation) of polynucleosomes causes relaxation of chromatin structure. *Proc Natl Acad Sci U S A*. 79,3423-7 (1982).
- Rack JG, Perina D, Ahel I. Macrod domains: Structure, Function, Evolution, and Catalytic Activities. *Annu Rev Biochem*. 85:431-54 (2016).
- Richard DJ, Bolderson E, Cubeddu L, Wadsworth RI, Savage K, Sharma GG, Nicolette ML, Tsvetanov S, McIlwraith MJ, Pandita RK, Takeda S, Hay RT, Gautier J, West SC, Paull TT, Pandita TK, White MF, Khanna KK. Single-stranded DNA-binding protein hSSB1 is critical for genomic stability. *Nature*. 453(7195):677-81 (2008).
- Rosenthal F, Feijs KL, Frugier E, Bonalli M, Forst AH, Imhof R, Winkler HC, Fischer D, Caflisch A, Hassa PO, Lüscher B, Hottiger MO. Macrod domain-containing proteins are new mono-ADPriboseylhydrolases. *Nat Struct Mol Biol*. 20,502-7 (2013).
- Rossmann MG, Blow DM. The detection of sub-units within the crystallographic asymmetric unit. *Acta Cryst*. 15, 24-31 (1962).
- Rouleau M, Patel A, Hendzel MJ, Kaufmann SH, Poirier GG. PARP inhibition: PARP1 and beyond. *Nat Rev Cancer*. 10(4):293-301 (2010).

- Ruf A, Rolli V, de Murcia G, Schulz GE. The mechanism of the elongation and branching reaction of poly(ADP-ribose) polymerase as derived from crystal structures and mutagenesis. *J Mol Biol.* 278,57-65 (1998).
- Sartori AA, Lukas C, Coates J, Mistrik M, Fu S, Bartek J, Baer R, Lukas J, Jackson SP. Human CtIP promotes DNA end resection. *Nature.* 450(7169):509-14 (2007).
- Schreiber V, Amé JC, Dollé P, Schultz I, Rinaldi B, Fraulob V, Ménissier-de Murcia J, de Murcia G. Poly(ADP-ribose) polymerase-2 (PARP-2) is required for efficient base excision DNA repair in association with PARP-1 and XRCC1. *J Biol Chem.* 277(25):23028-36 (2002).
- Schreiber V, Dantzer F, Ame JC, de Murcia G. Poly(ADP-ribose): novel functions for an old molecule. *Nat Rev Mol Cell Biol.* 7,517-28 (2006).
- Schreurs AMM, Xian X, Kroon-Batenburg LMJ. EVAL15: a diffraction data integration method based on ab initio predicted profiles. *J. Appl. Cryst.* 43, 70–82 (2010).
- Sharifi R, Morra R, Appel CD, Tallis M, Chioza B, Jankevicius G, Simpson MA, Matic I, Ozkan E, Golia B, Schellenberg MJ, Weston R, Williams JG, Rossi MN, Galehdari H, Krahn J, Wan A, Trembath RC, Crosby AH, Ahel D, Hay R, Ladurner AG, Timinszky G, Williams RS, Ahel I. Deficiency of terminal ADP-ribose protein glycohydrolase TARG1/C6orf130 in neurodegenerative disease. *EMBO J.* 32,1225-37 (2013).
- Shieh WM, Amé JC, Wilson MV, Wang ZQ, Koh DW, Jacobson MK, Jacobson EL. Poly(ADP-ribose) polymerase null mouse cells synthesize ADP-ribose polymers. *J Biol Chem.* 273(46):30069-72 (1998).
- Skarzynski T. Collecting data in the home laboratory: evolution of X-ray sources, detectors and working practices. *Acta Crystallogr D Biol Crystallogr.* 69(Pt 7):1283-8 (2013).
- Slade D, Dunstan MS, Barkauskaite E, Weston R, Lafite P, Dixon N, Ahel M, Leys D, Ahel I. The structure and catalytic mechanism of a poly(ADP-ribose) glycohydrolase. *Nature.* 477,616-20 (2011).
- Storoni LC, McCoy AJ, Read RJ. Likelihood-enhanced fast rotation functions. *Acta Crystallogr D Biol Crystallogr.* 60(Pt 3):432-8 (2004).
- Sugimura L, Takebayashi S, Taguchi H, Takeda S, Okumura K. PARP-1 ensures regulation of replication fork progression by homologous recombination on damaged DNA. *J Cell Biol.* 183(7):1203-1212 (2008).
- Sugimura T, Miwa M. Poly(ADP-ribose): historical perspective. *Mol Cell Biochem.* 138(1-2):5-12 (1994).
- Szankasi P, Smith GR. A role for exonuclease I from *S. pombe* in mutation avoidance and mismatch correction. *Science.* 267(5201):1166-9 (1995).
- Tallis M, Morra R, Barkauskaite E, Ahel I. Poly(ADP-ribosyl)ation in regulation of chromatin structure and the DNA damage response. *Chromosoma.* 123(1-2):79-90 (2014).
- Tamura M, Nogimori K, Murai S, Yajima M, Ito K, Katada T, Ui M, Ishii S. Subunit structure of islet-activating protein, pertussis toxin, in conformity with the A-B model. *Biochemistry.* 21(22):5516-22 (1982).
- Taylor GL. Introduction to phasing. *Acta Crystallogr D Biol Crystallogr.* 66(Pt 4): 325–338 (2010).
- Terada M, Fujiki H, Marks PA, Sugimura T. Induction of erythroid differentiation of murine erythroleukemia cells by nicotinamide and related compounds. *Proc Natl Acad Sci U S A.* 76(12): 6411–6414 (1979).
- Terwilliger TC, Adams PD, Read RJ, McCoy AJ, Moriarty NW, Grosse-Kunstleve RW, Afonine PV, Zwart PH, Hung LW. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr D Biol Crystallogr.* 65(Pt 6):582-601 (2009).

- Timinszky G, Till S, Hassa PO, Hothorn M, Kustatscher G, Nijmeijer B, Colombelli J, Altmeyer M, Stelzer EH, Scheffzek K, Hottiger MO, Ladurner AG. A macrodomain-containing histone rearranges chromatin upon sensing PARP1 activation. *Nat Struct Mol Biol.* 16,923-9 (2009).
- Tran PT, Simon JA, Liskay RM. Interactions of Exo1p with components of MutLalpha in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A.* 98(17):9760-5 (2001).
- Tulin A, Spradling AC. Chromatin loosening by poly(ADP)-ribose polymerase (PARP) at *Drosophila* puff loci. *Science.* 299(5606):560-562 (2003).
- Tulin A, Stewart D, Spradling AC. The *Drosophila* heterochromatic gene encoding poly(ADP-ribose) polymerase (PARP) is required to modulate chromatin structure during development. *Genes Dev.* 16(16): 2108-2119 (2002).
- Virag L, Robaszkiewicz A, Vargas JM, Javier Oliver F. Poly(ADP-ribose) signaling in cell death. *Mol Aspects Med.* 34(6):1153-67 (2013).
- Vyas S, Matic I, Uchima L, Rood J, Zaja R, Hay RT, Ahel I, Chang P. Family-wide analysis of poly(ADP-ribose) polymerase activity. *Nat Commun.* 5:4426 (2014).
- Wakasugi M, Sancar A. Assembly, subunit composition, and footprint of human DNA repair excision nuclease. *Proc Natl Acad Sci U S A.* 95(12):6669-74 (1998).
- Wakasugi M, Sancar A. Order of assembly of human DNA repair excision nuclease. *J Biol Chem.* 274(26):18759-68 (1999).
- Walker JR, Corpina RA, Goldberg J. Structure of the Ku heterodimer bound to DNA and its implications for double-strand break repair. *Nature.* 412(6847):607-14 (2001).
- Wang M, Wu W, Wu W, Rosidi B, Zhang L, Wang H, Iliakis G. PARP-1 and Ku compete for repair of DNA double strand breaks by distinct NHEJ pathways. *Nucleic Acids Res.* 34(21):6170-82 (2006).
- Wang Y, Dawson VL, Dawson TM. Poly(ADP-ribose) signals to mitochondrial AIF: a key event in parthanatos. *Exp Neurol.* 218(2):193-202 (2009).
- Wang Y, Kim NS, Haince JF, Kang HC, David KK, Andrabi SA, Poirier GG, Dawson VL, Dawson TM. Poly(ADP-ribose) (PAR) binding to apoptosis-inducing factor is critical for PAR polymerase-1-dependent cell death (parthanatos). *Sci Signal.* 4(167):ra20 (2011).
- Wang Z, Michaud GA, Cheng Z, Zhang Y, Hinds TR, Fan E, Cong F, Xu W. Recognition of the iso-ADP-ribose moiety in poly(ADP-ribose) by WWE domains suggests a general mechanism for poly(ADP-ribosylation)-dependent ubiquitination. *Genes Dev.* 26(3):235-40 (2012).
- Wielckens K, Schmidt A, George E, Bredehorst R, Hilz H. DNA fragmentation and NAD depletion. Their relation to the turnover of endogenous mono(ADP-ribosyl) and poly(ADP-ribosyl) proteins. *J Biol Chem.* 10;257(21):12872-7 (1982).
- Winter G, Lobley CM, Prince SM. Decision making in xia2. *Acta Crystallogr D Biol Crystallogr.* 69(Pt 7):1260-73 (2013).
- Wood RD. DNA damage recognition during nucleotide excision repair in mammalian cells. *Biochimie.* 81(1-2):39-44 (1999).
- Wood RD. Nucleotide excision repair in mammalian cells. *J Biol Chem.* 272(38):23465-8 (1997).
- Xie S, Mortusewicz O, Ma HT, Herr P, Poon RY, Helleday T, Qian C. Timeless Interacts with PARP-1 to Promote Homologous Recombination Repair. *Mol Cell.* 60(1):163-76 (2015).
- Yoo S, Dynan WS. Geometry of a complex formed by double strand break repair proteins at a single DNA end: recruitment of DNA-PKcs induces inward translocation of Ku protein. *Nucleic Acids Res.* 27(24):4679-86 (1999).

Young LM, Marzio A, Perez-Duran P, Reid DA, Meredith DN, Roberti D, Star A, Rothenberg E, Ueberheide B, Pagano M. TIMELESS Forms a Complex with PARP1 Distinct from Its Complex with TIPIN and Plays a Role in the DNA Damage Response. *Cell Rep.* 13(3):451-459 (2015).

Zaremba T, Curtin NJ. PARP inhibitor development for systemic cancer targeting. *Anticancer Agents Med Chem.* 7(5):515-23 (2007).

Žaja R, Mikoč A, Barkauskaite E, Ahel I. Molecular Insights into Poly(ADP-ribose) Recognition and Processing. *Biomolecules.* 3(1):1-1 (2013).

Zhang D, Iyer LM, Aravind L. A novel immunity system for bacterial nucleic acid degrading toxins and its recruitment in various eukaryotic and DNA viral systems. *Nucleic Acids Res.* 39(11):4532-52 (2011).

Zhang F, Chen Y, Li M, Yu X. The oligonucleotide/oligosaccharide-binding fold motif is a poly(ADP-ribose)-binding domain that mediates DNA damage response. *Proc Natl Acad Sci U S A.* 111(20):7278-83 (2014).

Zhang Y, Yuan F, Presnell SR, Tian K, Gao Y, Tomkinson AE, Gu L, Li GM. Reconstitution of 5'-directed human mismatch repair in a purified system. *Cell.* 122(5):693-705 (2005).